HAL
open science

DL_MULTI - A molecular dynamics program to use distributed multipole electrostatic models to simulate the dynamics of organic crystals

Maurice Leslie

To cite this version:

Maurice Leslie. DL_MULTI - A molecular dynamics program to use distributed multipole electrostatic models to simulate the dynamics of organic crystals. Molecular Physics, 2008, 106 (12-13), pp.1567-1578. 10.1080/00268970802175308 . hal-00513210

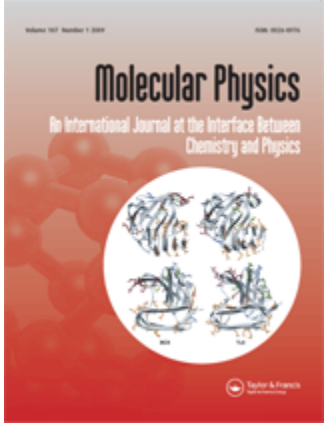
HAL Id: hal-00513210

https://hal.science/hal-00513210v1

Submitted on 1 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



DL_MULTI - A molecular dynamics program to use distributed multipole electrostatic models to simulate the dynamics of organic crystals

Journal:	*Molecular Physics*
Manuscript ID:	TMPH-2008-0086.R1
Manuscript Type:	Full Paper
Date Submitted by the Author:	29-Apr-2008
Complete List of Authors:	Leslie, Maurice; STFC Daresbury Laboratory
Keywords:	Ewald sum, multipoles, molecular dynamics, polymorphism
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
dlmultivec.tex	

scholarONE™
Manuscript Central

DL_MULTI — A molecular dynamics program to use distributed multipole electrostatic models to simulate the dynamics of organic crystals

M Leslie

STFC Daresbury Laboratory, Daresbury, Warrington WA4 4AD

Abstract

DL_MULTI has been developed to extend the Molecular Dynamics simulation program DL_POLY [1] to model rigid molecules whose intermolecular interactions include a distributed multipole model for the electrostatic interactions. The adaptations use anisotropic atom-atom potentials, corresponding to atomic multipoles up to hexadecapole. The lattice sums of these multipoles are evaluated using the Ewald method, using a technique using Stone's S functions [2] which describes the multipoles in a molecule fixed reference frame. An algorithm for determining suitable cutoffs is described and errors in the direct space part of the Ewald summation discussed. Thus DL_MULTI provides a general purpose MD program for studying polar rigid organic molecules in their liquid and crystalline states with a realistic intermolecular potential suitable for studying polymorphism. Example applications to uracil and 5-azauracil show that, with the new summation method, a realistic electrostatic model can be used without excessive computer time being used.

For Peer Review Only

Keywords

Ewald sum; multipoles; molecular dynamics; polymorphism

1 Introduction

Most molecular and macromolecular simulations with long range electrostatic interactions are performed in periodic boundary conditions, and there are many techniques for summing periodic charge distributions. Methods based on the Ewald sum are widespread and have been reviewed by Toukmaji [3]. All rely on dividing the slowly converging lattice sums into two rapidly converging sums, one in direct space and one in reciprocal space. Particle-Mesh Ewald (PME) [4] uses Fast Fourier Transforms (FFT) to accelerate the reciprocal space sum, but uses a monopole expansion of the electrostatic charge density. The Fast Multipole Method (FMM) was originally developed for a single non-periodic cell [5]. Schmidt and Lee [6] extended the method to periodic cells, and have also combined the method with FFT techniques [7]. The FMM expands the charge density of subcells of the simulation cell, containing many molecules in molecular systems. In contrast, Smith [8], describes an Ewald simulation in which the charge density of a single molecule is expanded as distributed multipoles up to quadrupole. Toukmaji [9] has combined PME methods with electrostatic dipole interactions on individual molecules. Unlike the FMM expansions, Smith uses Cartesian multipoles to expand the charge density. Challacombe et al. [10] compare a number of these multipole methods. The present work uses a distributed multipole analysis (DMA) in spherical harmonics to expand the charge density of an isolated molecule. A unique feature which sets it apart from all other earlier work is the definition of the spherical harmonics in an axis system fixed with respect to the molecular axes. This removes the need to rotate the spherical harmonics to the laboratory reference frame after each time step, a procedure which requires the use of Wigner matrices. Stone [2] has used this technique to describe the interaction between two isolated molecules, and derives S functions to describe the interaction. This paper will derive an Ewald expansion of the distributed multipole analysis in terms of the S functions.

For Peer Review Only

Our primary interest is in the solid state and in particular polymorphism, although the program has application to other systems. The pharmaceutical, pigment and energetic materials industries produce polar organic molecules, usually in a crystalline form. Hence, polymorphism, the ability of many of their products to crystallise in different crystal structures is a major concern of these industries [11]. Polymorphs have different physical properties, and so an unexpected change to the crystalline form during processing or storage is a disaster, but the controlled use of a metastable polymorphic form can be advantageous. A specific polymorphic form has to be used in a drug formulation, and so the regulatory and patent issues involved in polymorphism have led to considerable activity in devising methods of finding all the polymorphs of a given molecule, either experimentally or by computational means.

Computational methods of predicting which crystal structures will be observed for a given molecule from the chemical diagram have evolved over the last decade [12]. There are still many issues to be solved before polymorphism can be predicted [12]. Methods that have had some success in the international blind tests of crystal structure prediction [13, 14] have been based on searching for the global minimum in the static lattice energy. However, most such searches find far more distinct crystal structures within the energy range of possible polymorphism than known polymorphs [15]. Some of these hypothetical energetically feasible crystal structures may indeed correspond to as yet undiscovered polymorphs, indeed, such predicted structures have been used to characterise polymorphs discovered [16] after the predictions were published [17]. However, most of the energetically feasible unobserved structures are probably not observed because the simple static lattice energy search does not reflect the complex interplay of kinetic and thermodynamic factors involved in the nucleation and growth of crystals under different conditions of solvent, temperature and pressure. The phenomena of concomitant crystallisation [18], where more than one polymorph crystallises in the same experiment, or disappearing polymorphs [19], where it proves extremely difficult to produce new samples of a well characterised polymorph after a more stable form has been produced, warn that understanding polymorphism sufficiently to produce a reliable predictive model is a major scientific challenge.

For Peer Review Only

One key ingredient to understanding polymorphism will be a proper understanding of the effects of temperature on the dynamics of organic crystal structures. The prediction of the relative stability of different structures as a function of temperature requires free energies rather than lattice energies, and a careful consideration of the most appropriate approximations [20]. Harmonic approximation estimates of the relative free energies of different known polymorphs [21] show that entropy differences were small compared with the lattice energy differences. Similarly, harmonic approximation estimates of the intermolecular zero-point energy and entropy at the melting point of pyridine [22] produced a different energetic ordering than the lattice energy for the many structures predicted to be a few kJ/mol more stable than the known polymorphs. However, the use of the harmonic approximation for the motions within such soft organic crystals is questionable, and the motions for which there is significant anharmonicity will vary considerably between different crystal structures. For example, the modes of hydrogen bonded chains of carboxylic acids will differ considerably from the modes of crystal structures based on hydrogen bonded dimers, yet acetic acid [23], formic acid and tetrolic acid [24] all have many energetically feasible crystal structures based on both motifs, and indeed a dimer and chain based polymorph are known for tetrolic acid. Relatively few comparisons [25] have been made of the phonon frequencies of molecular solids as calculated from molecular dynamics simulations [26] and the use of various harmonic and other approximations. Clearly, a deeper insight into the effect of temperature on the motions within molecular crystals is required before their relative thermodynamic stability can be properly assessed.

The kinetics as well as the thermodynamics of phase transitions between different polymorphs of organic crystals are also important for polymorph prediction. Constant pressure molecular dynamics methods [27], which allow the shape of the simulation cell to change, have been used to study a few simple transformations in molecular solids. This method could be used more widely to study known polymorphic phase transitions. Such studies will show how to diagnose when the transformation between a hypothetical low energy structure and a known structure is so facile that the hypothetical structure would never be observed. Obviously closely related crystal structures are

For Peer Review Only

already eliminated from crystal structure prediction searches, but when there is a relationship, e.g. the same hydrogen bonded sheet, stacked in different translationally related ways, then determining whether two polymorphs will be observed (as in the case of indigo [28]) or not, depends on the kinetics.

Thus, there is a clear need for a program to be able to perform a wide range of molecular dynamics simulations on organic crystal structures, which can be readily used for a wide range of organic molecules, in the low symmetry space groups which they generally adopt. The major barrier to providing a program that can simulate organic crystal structures to the required accuracy is that it must be capable of using model intermolecular potentials that are sufficiently realistic to reproduce the crystal structures. For polar and hydrogen bonded molecules, the potential has to correctly model the balance between the hydrogen bonding interactions, the π—π interactions involved between the aromatic rings and the van der Waals interactions of the hydrocarbon groups. High accuracy in the intermolecular potential is required in crystal structure prediction, as the lattice energy searches rely on the relative accuracy of the lattice energies of the thousands to millions of hypothetical structures considered. This requirement is proving a major incentive for the development of more accurate methods of modelling intra and intermolecular potentials for organic molecules [12, 29, 30]. It is clear that accurate models for the electrostatic forces are required for all organic molecules (with the possible exception of the saturated hydrocarbons) as this contribution is very dependent on the relative orientation of the molecules in the different crystal structures. The use of distributed multipole electrostatic models [31, 2], where the charge distribution of each molecule is represented by sets of atomic multipoles derived from the ab initio charge density of the molecule, has improved [32] the range of molecules where the lattice energy minimum reproduces the experimental crystal structure to reasonable accuracy. Since the distributed multipole model automatically represents the anisotropic electrostatic forces arising from lone pairs and π electron density etc, such models predict the directionality of hydrogen bonding and other intermolecular interactions [33, 34]. Thus a general method of modelling the intermolecular forces between organic molecules, that is sufficiently realistic for simulating the organic solid state of a range of organic

For Peer Review Only

molecules, will require the use of anisotropic atom-atom multipolar electrostatic models.

Hence, we have extended the general purpose MD simulation package DL_POLY [1] to simulate rigid organic models whose intermolecular interactions are described by a distributed multipole electrostatic model. The resulting program DL_MULTI allows Molecular Dynamics calculations to be performed with the same models that are used in simulating molecular clusters and surfaces in the program ORIENT [35] and organic crystal structures and properties by static lattice energy minimisation in the program DMAREL [36, 37]. The extension of DL_POLY to use anisotropic atom-atom interactions, of the form dictated by the multipole expansion of the electrostatic energy, with their associated non-central forces and torques [2, 36, 38] essentially follows ORIENT and DMAREL.

DL_MULTI has already been used to study liquid water [39] at constant pressure and temperature, using high rank atomic multipoles, and the liquid structure and self-diffusion coefficients show significant sensitivity to the higher rank multipoles. The results of this paper show that, with the new Ewald summation method, DL_MULTI is capable of simulating organic crystal structures at constant pressure and temperature. Future applications will use this simulation tool to analyse the molecular motions and phase transitions within organic crystals, as a tool to understand polymorphism.

The performance of DL_MULTI, and particularly the problem with cutoffs and the Ewald summation, are illustrated by example simulations on uracil and 5-azauracil. These were chosen as similar rigid hydrogen bonding molecules, with a variety of hydrogen bonding donors and acceptors, which adopt very different crystal structures. Uracil adopts a crystal structure with hydrogen bonding in all three directions. Although the observed structure was predicted [40] to be the global minimum in the lattice energy, there are other hypothetical structures involving different hydrogen bonds within the energy range of possible polymorphism. In contrast, 5-azauracil adopts a crystal structure of hydrogen bonded sheets, which was correctly predicted [41] as the global minimum in the lattice energy without knowledge of the experimental structure. The competitive local minima were all based on the same hydrogen bonded sheet motif.

For Peer Review Only

2 Force Field

Anisotropic atom-atom potentials have previously been used in MD simulations, but with specific programming for the individual system [42]. In some cases, the electrostatic model has included atomic dipoles. Other studies have emphasised the use of atom-atom anisotropy in the short range interactions, such as the study of molecular chlorine [43, 44], bromine and iodine [45], or the anisotropic carbon only potential model for butane [42]. The program DL_MULTI provides the facility to use an accurate distributed multipole electrostatic model, which in conjunction with an isotropic repulsion-dispersion potential, should provide sufficiently realistic potentials for the simulation of a range of C/H/N/O rigid organic molecules. It could be easily extended to use anisotropic repulsion models, using the methodology within DMAREL [46] when needed to model atoms such as Cl which have a non-spherical repulsion in organic molecules.

The ab initio charge distribution of the rigid molecular model has to be calculated relative to a set of molecule-fixed axes. For uracil and 5-azauracil, the molecular structures and charge distribution were obtained by optimisation of the MP2(fc) 6-31G** wavefunction, using the program Gaussian [47]. This charge density is then subjected to a distributed multipole analysis (DMA) [31] performed using the program GDMA [48] to represent the charge distribution by a set of multipoles $\tilde{Q}^A_{l_1\kappa_1}$ up to hexadecapole ($l_1 \leq 4$), referred to the molecule fixed axes, on each atomic site A. The electrostatic interaction energy of a periodic lattice of molecules can then be evaluated, up to total multipole moment for the interaction of $l = 4$ ($l = l_1 + l_2$). Although a limited amount of work has been carried out with higher values of total l, the contribution to the energy from terms of $l = 4$ is found to be small for the organic systems which are usually studied. The program DL_MULTI uses molecular fixed axes defined by the centre of mass and principal inertial axes. The input multipoles, which are defined with respect to a user defined axis system, are transformed, using Wigner matrices, to the molecular principal moment of inertia reference frame. Since this choice of axes still reflects any molecular symmetry, there will be a number of zero multipole moment components according to the molecular symmetry.

The repulsion-dispersion potential is assumed to be of an isotropic atom-atom form and the

For Peer Review Only

existing capabilities of DL_POLY are used. For these examples, the repulsion-dispersion form used was:

$$E_{rep-dis} = \frac{1}{2}\sum_{A}\sum_{B}\sum_{N}\left(\left(\bar{A}_A\bar{A}_B\right)^{\frac{1}{2}}\exp\left(-\frac{1}{2}\left(\bar{B}_A+\bar{B}_B\right)r_{ABN}\right) - \frac{\left(\bar{C}_A\bar{C}_B\right)^{\frac{1}{2}}}{r^6_{ABN}}\right) \quad (1)$$

where atom A is in molecule 1 and atom B is in molecule 2 and A and B must be in different molecules if $N = 0$. $\bar{A}$ $\bar{B}$ and $\bar{C}$ are the short range potential parameters. This interaction between a pair must be summed over all sites A within a unit cell and over all sites B in the entire lattice, translated by direct space lattice vectors N, to give the energy per unit cell. The parameters used for C, N, O and hydrogens bonded to carbon (HC) were taken from the empirically fitted parameters of Williams [49, 50], and the parameters for the polar hydrogen atoms (HN) from the extension of this parameters set to hydrogen bonding molecules [32].

3 Description of the Ewald method

The equation for the interaction between two spherical harmonic point multipoles is first expanded as an infinite series of irregular spherical harmonics equation (3.3.3) of [2]. This equation is derived from an expansion of $1/r_{AB}$.

$$E_{AB} = \frac{1}{4\pi\epsilon_0}\sum_{l_1,l_2}\sum_{m_1m_2m}(-1)^{l_2}\begin{pmatrix} l_1+l_2 \\ l_1 \end{pmatrix} \times \hat{Q}^A_{l_1m_1}\hat{Q}^B_{l_2m_2}I_{l_1+l_2,m}(r)\begin{pmatrix} l_1 & l_2 & l_1+l_2 \\ m_1 & m_2 & m \end{pmatrix} \quad (2)$$

For Peer Review Only

where

$$\begin{pmatrix} l_1+l_2 \\ l_1 \end{pmatrix} = \left(\frac{(2l_1+2l_2+1)!}{(2l_1)!(2l_2)!}\right)^{\frac{1}{2}}$$

and $I_{l_1+l_2,m}(r)$ is an irregular spherical harmonic, $\hat{Q}^A_{l_1m_1}$ is a multipole moment operator, which is a regular spherical harmonic of the charge distribution about the centre. $\begin{pmatrix} l_1 & l_2 & l_1+l_2 \\ m_1 & m_2 & m \end{pmatrix}$ is a Wigner $3j$ symbol [52]. The energy is summed over all atoms and over all unit cells as for the repulsion dispersion energy (equation (1)) except that the interactions within a molecule for $N=0$ are included and subtracted out later on. This is necessary in order to make an atom in a molecule interact with a periodic array of other atoms so that an Ewald transformation can be carried out.

$$\begin{aligned} E_{coul} &= \frac{1}{2}\sum_A\sum_B\sum_N \frac{1}{4\pi\epsilon_0}\sum_{l_1,l_2}\sum_{m_1m_2m}(-1)^{l_2}\begin{pmatrix} l_1+l_2 \\ l_1 \end{pmatrix} \\ &\times \hat{Q}^A_{l_1m_1}\hat{Q}^B_{l_2m_2}I_{l_1+l_2,m}(r_{ABN})\begin{pmatrix} l_1 & l_2 & l_1+l_2 \\ m_1 & m_2 & m \end{pmatrix} \end{aligned} \tag{3}$$

The irregular spherical harmonics can be replaced by an expression involving a sum of terms which are l^{th} order differentials of $1/r$. Using the notation of Tough and Stone equation (2) of [53]

$$I_{lm}(\underset{\sim}{r}) = (-1)^l\left[2^l/(2l)!\right]^{1/2}\sum_{\alpha_1\ldots\alpha_l}\nabla_{\alpha_l}\ldots\nabla_{\alpha_1}(r^{-1})\,\langle\alpha_1\ldots\alpha_l|12\ldots l;m\rangle \tag{4}$$

where ∇_α is the Cartesian component of the gradient operator ∇ and $\langle\alpha_1\ldots\alpha_l|12\ldots l;m\rangle$ is the Cartesian-spherical transformation coefficient [53]. When the irregular spherical harmonics are summed over a lattice we can replace the lattice sum of (r^{-1}) by the Ewald function to give the energy of interaction between a site multipole and a lattice of multipoles at the other sites. The desired expression for the Ewald function denoted $A(\underset{\sim}{r})$ is taken from Saunders et al. equation (44) of [54]

$$A(\underset{\sim}{r}) = -\pi/(\gamma V) + B(\gamma,\underset{\sim}{r}) + C(\gamma,\underset{\sim}{r}) \tag{5}$$

$$B(\gamma,\underset{\sim}{r}) = {\sum_N}' \frac{1-\text{erf}(\gamma^{1/2}|\underset{\sim}{r}-\underset{\sim}{r}_N|)}{|\underset{\sim}{r}-\underset{\sim}{r}_N|} \tag{6}$$

For Peer Review Only

$$
\begin{aligned}
C(\gamma, \underset{\sim}{r}) &= \frac{4\pi}{V} \sum_n{}' |\underset{\sim}{K}_n|^{-2} \exp\left(-\frac{|\underset{\sim}{K}_n|^2}{4\gamma} + i\underset{\sim}{K}_n \cdot \mathbf{r}\right) \\
&= \frac{8\pi}{V} \sum_n{}'' |\underset{\sim}{K}_n|^{-2} \exp\left(\frac{-|\underset{\sim}{K}_n|^2}{4\gamma}\right) \cos(\underset{\sim}{K}_n \cdot \underset{\sim}{r}) \\
&= \frac{8\pi}{V} \sum_n{}'' G(K_n) \cos(\underset{\sim}{K}_n \cdot \underset{\sim}{r}) \qquad (7)
\end{aligned}
$$

where $\underset{\sim}{r}_N$ and $\underset{\sim}{K}_n$ are lattice vectors in direct and reciprocal space respectively and γ is a positive constant which determines the convergence of the direct and reciprocal space sums. The term $-\pi/(\gamma V)$ cancels out for electroneutral cells if the same screening parameter γ is used for all terms. $\sum_N'$ in equation (6) implies that the sum extends over all direct space lattice vectors except that the $|\underset{\sim}{r} - \underset{\sim}{r}_N|^{-1}$ term is omitted if $\underset{\sim}{r} - \underset{\sim}{r}_N = 0$. $\sum_n'$ in equation (7) indicates that the reciprocal lattice vector $\underset{\sim}{K}_n = 0$ is omitted and $\sum_n''$ takes advantage of the symmetry of reciprocal space and implies that the sum is over a hemisphere of reciprocal space omitting $\underset{\sim}{K}_n = 0$. As will be shown later, the $\underset{\sim}{K}_n = 0$ term is either going to diverge ($l < 2$), be indeterminate or have a finite limit ($l = 2$) or equal zero ($l > 2$). The justification of omitting it is that its value depends on the nature of the macroscopic surface of the crystal being simulated. To evaluate the Ewald sum of the lattice energy, $A(\underset{\sim}{r})$ from equation (5) is substituted for the lattice sum of (r^{-1}) in equation (3) using (4) to expand the irregular spherical harmonic.

We will consider the term in reciprocal space first. The Cartesian differential gradient operator will only operate on $\underset{\sim}{r}$ in equation (7). Now

$$
\nabla_\alpha \cos(\underset{\sim}{K}_n \cdot \underset{\sim}{r}) = K_{n\alpha} \frac{\partial}{\partial(\underset{\sim}{K}_n \cdot \underset{\sim}{r})} \cos(\underset{\sim}{K}_n \cdot \underset{\sim}{r}) \qquad (8)
$$

Hence

$$
\begin{aligned}
\nabla_{\alpha_l} \ldots \nabla_{\alpha_1} \cos(\underset{\sim}{K}_n \cdot \underset{\sim}{r}) &= K_{n\alpha 1} \ldots K_{n\alpha l} \frac{\partial^l}{\partial(\underset{\sim}{K}_n \cdot \underset{\sim}{r})^l} \cos(\underset{\sim}{K}_n \cdot \underset{\sim}{r}) \\
&= K_{n\alpha 1} \ldots K_{n\alpha l} \cos^{(l)}(\underset{\sim}{K}_n \cdot \underset{\sim}{r}) \qquad (9)
\end{aligned}
$$

For Peer Review Only

where the notation $\cos^{(l)}(\underset{\sim}{K}_n \cdot \underset{\sim}{r})$ is used for the lth differential of $\cos(\underset{\sim}{K}_n \cdot \underset{\sim}{r})$. Substituting into equation (4) summed over lattice vectors gives

$$\begin{aligned}\left(\sum_N I_{lm}(\underset{\sim}{r})\right)_{rl} &= (-1)^l \left[2^l/(2l)!\right]^{1/2} \frac{8\pi}{V} \sum_n'' G(K_n) \cos^{(l)}(\underset{\sim}{K}_n \cdot \underset{\sim}{r}) \\ &\times \sum_{\alpha_1 \ldots \alpha_l} K_{n\alpha 1} \ldots K_{n\alpha l} \langle \alpha_1 \ldots \alpha_l | 12 \ldots l; m \rangle \end{aligned} \tag{10}$$

Tough and Stone equation (1) of [53] give a corresponding expression for the regular spherical harmonics of r.

$$R_{lm}(\underset{\sim}{r}) = \left[(2l)!/2^l\right]^{1/2} (l!)^{-1} \sum_{\alpha_1 \ldots \alpha_l} r_{\alpha_l} \ldots r_{\alpha_1} \langle \alpha_1 \ldots \alpha_l | 12 \ldots l; m \rangle \tag{11}$$

This is of the same form as equation (10) with the distance r replaced with the reciprocal lattice vector $\underset{\sim}{K}_n$. This is the key transformation which allows the evaluation of the reciprocal space terms. Making this substitution in (10) gives

$$\begin{aligned}\left(\sum_N I_{lm}(\underset{\sim}{r})\right)_{rl} &= (-1)^l \left[2^l/(2l)!\right]^{1/2} \frac{8\pi}{V} \sum_n'' G(K_n) \cos^{(l)}(\underset{\sim}{K}_n \cdot \underset{\sim}{r}) \\ &\times \left[2^l/(2l)!\right]^{1/2} l! R_{lm}(\underset{\sim}{K}_n) \\ &= \frac{(-1)^l}{(2l-1)!!} \frac{8\pi}{V} \sum_n'' G(K_n) \cos^{(l)}(\underset{\sim}{K}_n \cdot \underset{\sim}{r}) R_{lm}(\underset{\sim}{K}_n) \end{aligned} \tag{12}$$

where

$$1/(2l-1)!! = 2^l l!/(2l)! = 1/1.3.5 \ldots (2l-1) \tag{13}$$

Substituting the lattice summed irregular solid harmonics into equation (3) gives (From now on the subscript coul will be dropped).

$$\begin{aligned} E_{rl} &= \frac{1}{2} \sum_A \sum_B \frac{1}{4\pi\epsilon_0} \sum_{l_1,l_2} \sum_{m_1 m_2 m} (-1)^{l_2} \begin{pmatrix} l_1 + l_2 \\ l_1 \end{pmatrix} \\ &\times \hat{Q}^A_{l_1 m_1} \hat{Q}^B_{l_2 m_2} \begin{pmatrix} l_1 & l_2 & l_1 + l_2 \\ m_1 & m_2 & m \end{pmatrix} \frac{(-1)^l}{(2l-1)!!} \frac{8\pi}{V} \end{aligned}$$

For Peer Review Only

$$\times \quad \sum_n'' G(K_n)\cos^{(l)}(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_{AB}) R_{lm}(\underset{\sim}{K}_n) \tag{14}$$

This equation is still not in a convenient form as the multipole moment operators $\hat{Q}^A_{l_1m_1}$ and $\hat{Q}^B_{l_2m_2}$ are still defined in the laboratory reference frame and we need them in the molecule-fixed reference frame. However, the transformation used by Stone [51] may be used. The transformation of the reference frames is given by

$$\hat{Q}^A_{l_1m_1} = \sum_k \tilde{Q}^A_{l_1k_1} D^l_{mk}(\Omega_A)^* \tag{15}$$

where $D^l_{mk}(\Omega_A)$ is a Wigner rotation matrix [52] and Ω_A describes the orientation of molecule A. The transformation leads to an expression which contains Stone's S functions defined by equation (3.3.7) of [2]

$$\begin{aligned} \bar{S}^{k_1k_2}_{l_1l_2j}(\Omega_A,\Omega_B,\hat{K}_n) &= i^{l_1-l_2-j}\begin{pmatrix} l_1 & l_2 & j \\ 0 & 0 & 0 \end{pmatrix}^{-1} \sum_{m_1m_2m} \begin{pmatrix} l_1 & l_2 & j \\ m_1 & m_2 & m \end{pmatrix} \\ &\times \quad D^{l_1}_{m_1k_1}(\Omega_A)^* D^{l_2}_{m_2k_2}(\Omega_B)^* C_{jm}(\theta,\phi) \end{aligned} \tag{16}$$

where $C_{jm}(\theta,\phi)$ is the modified spherical harmonic $\sqrt{4\pi/(2l+1)}Y_{jm}$, and

$$R_{lm}(\underset{\sim}{K}_n) = K_n^l C_{lm}(\theta,\phi), \qquad I_{lm}(K_n) = K_n^{-l-1} C_{lm}(\theta,\phi) \tag{17}$$

K_n is the magnitude of $\underset{\sim}{K}_n$. The angles θ and ϕ define the direction of the reciprocal lattice vector $\underset{\sim}{K}_n$ in this case rather than the direction vector between two sites. Using this transformation in equation (14) leads to an expression for the reciprocal lattice part of the energy

$$\begin{aligned} E_{rl} &= \frac{1}{2}\sum_A\sum_B \frac{1}{4\pi\epsilon_0}\sum_{l_1,l_2}\sum_{k_1k_2}\begin{pmatrix} l_1+l_2 \\ l_1 \end{pmatrix} \\ &\times \quad \tilde{Q}^A_{l_1k_1}\tilde{Q}^B_{l_2k_2}\frac{(-1)^l}{(2l-1)!!}\frac{8\pi}{V}\sum_n'' G(K_n)\cos^{(l)}(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_{AB}) K_n^l \\ &\times \quad \bar{S}^{k_1k_2}_{l_1l_2l}(\Omega_A,\Omega_B,\hat{K}_n) \end{aligned} \tag{18}$$

For Peer Review Only

Following Stone equation (12) of [51] we can remove the complex arithmetic by defining entirely real multipoles and S functions. Let

$$\tilde{Q}_{lk} = \sum_{\kappa} \tilde{Q}_{l\kappa} X_{\kappa k} \tag{19}$$

and

$$\bar{S}^{\kappa_1\kappa_2}_{l_1 l_2 l} = \sum_{k_1 k_2} X_{\kappa_1 k_1} X_{\kappa_2 k_2} \bar{S}^{k_1 k_2}_{l_1 l_2 l} \tag{20}$$

Equation (18) then becomes, with κ replacing k throughout,

$$\begin{aligned} E_{rl} &= \frac{1}{2} \sum_A \sum_B \frac{1}{4\pi\epsilon_0} \sum_{l_1,l_2} \sum_{\kappa_1\kappa_2} \begin{pmatrix} l_1 + l_2 \\ l_1 \end{pmatrix} \\ &\times \tilde{Q}^A_{l_1\kappa_1} \tilde{Q}^B_{l_2\kappa_2} \frac{(-1)^l}{(2l-1)!!} \frac{8\pi}{V} \sum_n{}'' G(K_n) \cos^{(l)}(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_{AB}) K_n^l \\ &\times \bar{S}^{\kappa_1\kappa_2}_{l_1 l_2 l}(\Omega_A, \Omega_B, \hat{K}_n) \end{aligned} \tag{21}$$

Although we could use the equation in this form, a triple sum is involved, twice over the particles in the cell and once over the reciprocal lattice vectors. In Ewald sums of charges the expression will factorize into the product of two sums over particles in the cell. We have investigated the factorisation of equation (20) for the terms in the multipole expansion up to $l_1 + l_2 = 4$ as given by Price et al. [55] and by Stone [2] Appendix F. Such a factorisation is indeed possible for all of these interactions. The factorisation of $\cos^{(l)}(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_{AB})$ is straightforward.

$$\cos^{(l)}(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_{AB}) = \sum_{\sigma=1}^{2} s_\sigma \mathrm{cs}_\sigma(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_A) \mathrm{cs}_\sigma(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_B) \qquad s_\sigma = \pm 1 \tag{22}$$

where cs represents a cosine or sine function as appropriate. The factorisation of Stone's S functions will be written

$$\begin{aligned} \sum_{\kappa_1\kappa_2} \begin{pmatrix} l_1 + l_2 \\ l_1 \end{pmatrix} \bar{S}^{\kappa_1\kappa_2}_{l_1 l_2 l}(\Omega_A, \Omega_B, \hat{K}_n) \tilde{Q}^A_{l_1\kappa_1} \tilde{Q}^B_{l_2\kappa_2} &= \\ \sum_{\tau} U_{\tau l_1 l_2} S_{F l_1 \tau}(\Omega_A, \hat{K}_n) S_{F l_2 \tau}(\Omega_B, \hat{K}_n) \end{aligned} \tag{23}$$

For Peer Review Only

where $U_{\tau l_1 l_2}$ is a constant. Explicit forms of the S_F terms in this expansion are given in Appendix A. The substitution of these two factorisations into equation (20) leads to our final expression for the reciprocal lattice contribution to the energy.

$$\begin{aligned} E_{rl} &= \frac{1}{2}\frac{1}{4\pi\epsilon_0}\sum_{l_1,l_2}\frac{(-1)^l}{(2l-1)!!}\frac{8\pi}{V}\sum_n{}'' G(K_n)\sum_{\sigma=1}^{2}\sum_{\tau} U_{\tau l_1 l_2} s_\sigma \\ &\times \sum_A \mathrm{cs}_\sigma(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_A) S_{Fl_1\tau}(\Omega_A, \hat{K}_n) K_n^{l_1} \\ &\times \sum_B \mathrm{cs}_\sigma(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_B) S_{Fl_2\tau}(\Omega_B, \hat{K}_n) K_n^{l_2} \end{aligned} \tag{24}$$

We can simplify this expression further for the rigid molecule case. Here, a large number of atoms will have the same local axes. The sum over all atoms in the unit cell A may be replaced by a double sum over molecules M and atoms within a molecule M_A. The functions S_F as given in Appendix A can be written

$$S^A_{Fl_1\tau}(\Omega_A, \hat{K}_n) = \sum_\kappa s^A_{Fl_1\tau\kappa_1}(\Omega_A, \hat{K}_n)\tilde{Q}^A_{l_1\kappa_1} \tag{25}$$

Where the term $s^A_{Fl_1\tau\kappa_1}(\Omega_A, \hat{K}_n)$ depends only on molecule orientation. Hence

$$\begin{aligned} &\sum_A \mathrm{cs}_\sigma(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_A) S_{Fl_1\tau}(\Omega_A, \hat{K}_n) K_n^{l_1} = \\ &\sum_M \sum_\kappa \sum_{A_M} \left\{ \mathrm{cs}_\sigma(\underset{\sim}{K}_n \cdot \underset{\sim}{r}_A) \tilde{Q}^A_{l_1\kappa_1} \right\} s^A_{Fl_1\tau\kappa_1}(\Omega_A, \hat{K}_n) K_n^{l_1} \end{aligned} \tag{26}$$

We will now consider the direct space term arising from the substitution of equation (5) into equation (4). We start from the lattice sum of irregular spherical harmonics equivalent to equation (10).

$$\begin{aligned} (\sum_N I_{lm}(\underset{\sim}{r}))_{dl} &= (-1)^l \left[2^l/(2l)!\right]^{1/2} \sum_{\alpha_1\ldots\alpha_l} \nabla_{\alpha_l}\ldots\nabla_{\alpha_1} \sum_N{}' \\ &\times \frac{1-\mathrm{erf}(\gamma^{1/2}|\underset{\sim}{r}-\underset{\sim}{r}_N|)}{|\underset{\sim}{r}-\underset{\sim}{r}_N|} \langle \alpha_1\ldots\alpha_l | 12\ldots l; m\rangle \end{aligned} \tag{27}$$

For Peer Review Only

It is convenient to redefine the error function in terms of the function F_0 (Equation (51c)) of [54].

$$F_m(p) = \int_0^1 \exp(-pu^2)u^{2m}\mathrm{d}u \tag{28}$$

$$F_0(p) = 2^{-1}(\pi/p)^{1/2}\mathrm{erf}(p^{1/2}) \tag{29}$$

We will also make use of the operator $(s^{(-1)}\hat{s})^n$ which is defined

$$(s^{(-1)}\hat{s})^n = \left(\frac{1}{s}\frac{\mathrm{d}}{\mathrm{d}s}\right)^n \tag{30}$$

and note the following relations.

$$\frac{\mathrm{d}}{\mathrm{d}x} = \nabla_x = x(s^{-1}\hat{s}) \tag{31}$$

$$(s^{(-1)}\hat{s})^n s^{-1} = (-1)^n(2n-1)!!s^{-2n-1} \tag{32}$$

$$(s^{(-1)}\hat{s})^n F_0(\alpha s^2) = (-2\alpha)^n F_n(\alpha s^2) \tag{33}$$

Equation (27) then becomes

$$\begin{aligned}
(\sum_N I_{lm}(\underset{\sim}{r}))_{dl} &= (-1)^l \left[2^l/(2l)!\right]^{1/2} \sum_N{}' \sum_{\alpha_1\ldots\alpha_l} \left(\underset{\sim}{r} - \underset{\sim}{r}_N\right)_{\alpha_1} \ldots \left(\underset{\sim}{r} - \underset{\sim}{r}_N\right)_{\alpha_l} \\
&\times \langle\alpha_1\ldots\alpha_l|12\ldots l;m\rangle \left(|\underset{\sim}{r} - \underset{\sim}{r}_N|^{(-1)}\widehat{|\underset{\sim}{r} - \underset{\sim}{r}_N|}\right)^l \\
&\times \left[\frac{1}{|\underset{\sim}{r} - \underset{\sim}{r}_N|} - 2\left(\frac{\gamma}{\pi}\right)^{1/2} F_0\left(\gamma|\underset{\sim}{r} - \underset{\sim}{r}_N|^2\right)\right]
\end{aligned} \tag{34}$$

Substituting (11) and using (32), (33) and (13) in (34) gives

$$\begin{aligned}
(\sum_N I_{lm}(\underset{\sim}{r}))_{dl} &= \frac{(-1)^l}{(2l-1)!!} \sum_N{}' R_{lm}(\underset{\sim}{r} - \underset{\sim}{r}_N) \left\{(-1)^l(2l-1)!!|\underset{\sim}{r} - \underset{\sim}{r}_N|^{-2l-1}\right. \\
&- \left. 2\left(\frac{\gamma}{\pi}\right)^{1/2}(-2\gamma)^l F_l\left(\gamma|\underset{\sim}{r} - \underset{\sim}{r}_N|^2\right)\right\} \\
&= \sum_N{}' R_{lm}(\underset{\sim}{r} - \underset{\sim}{r}_N) \left\{|\underset{\sim}{r} - \underset{\sim}{r}_N|^{-2l-1}\right. \\
&- \left. \frac{1}{(2l-1)!!} 2\left(\frac{\gamma}{\pi}\right)^{1/2}(2\gamma)^l F_l\left(\gamma|\underset{\sim}{r} - \underset{\sim}{r}_N|^2\right)\right\}
\end{aligned} \tag{35}$$

For Peer Review Only

Substituting into the equation for the direct space part of the lattice energy (3).

$$
\begin{aligned}
E_{dl} &= \frac{1}{2}\sum_A\sum_B \frac{1}{4\pi\epsilon_0}\sum_{l_1,l_2}\sum_{m_1m_2m}(-1)^{l_2}\begin{pmatrix} l_1+l_2 \\ l_1 \end{pmatrix} \\
&\times \hat{Q}^A_{l_1m_1}\hat{Q}^B_{l_2m_2}\sum_N{}' R_{lm}(\underset{\sim}{r}-\underset{\sim}{r}_N)\Bigg\{|\underset{\sim}{r}-\underset{\sim}{r}_N|^{-2l-1} \\
&- \frac{1}{(2l-1)!!}2\left(\frac{\gamma}{\pi}\right)^{1/2}(2\gamma)^l F_l\left(\gamma|\underset{\sim}{r}-\underset{\sim}{r}_N|^2\right)\Bigg\}\begin{pmatrix} l_1 & l_2 & l_1+l_2 \\ m_1 & m_2 & m \end{pmatrix}
\end{aligned} \tag{36}
$$

Now the equation for the S functions (16) is used in (36) where the S function now depends on the direction of the intersite vector $\underset{\sim}{r}-\underset{\sim}{r}_N$. This gives the final expression for the direct space part of the lattice energy.

$$
\begin{aligned}
E_{dl} &= \frac{1}{2}\sum_A\sum_B \frac{1}{4\pi\epsilon_0}\sum_{l_1,l_2}\sum_{k_1k_2}\begin{pmatrix} l_1+l_2 \\ l_1 \end{pmatrix} \\
&\times \tilde{Q}^A_{l_1k_1}\tilde{Q}^B_{l_2k_2}\sum_N{}' \bar{S}^{k_1k_2}_{l_1l_2l}(\Omega_A,\Omega_B,\widehat{\underset{\sim}{r}-\underset{\sim}{r}_N}) \\
&\Bigg\{|\underset{\sim}{r}-\underset{\sim}{r}_N|^{-l-1} - \frac{|\underset{\sim}{r}-\underset{\sim}{r}_N|^l}{(2l-1)!!}2\left(\frac{\gamma}{\pi}\right)^{1/2}(2\gamma)^l F_l\left(\gamma|\underset{\sim}{r}-\underset{\sim}{r}_N|^2\right)\Bigg\}
\end{aligned} \tag{37}
$$

We now need to consider a number of additional terms. Firstly, the Ewald method introduces the interaction of a particle with itself, $(A=B, N=0)$, which must be explicitly subtracted out (a self-interaction energy). This has been considered for the Cartesian tensor formulation by Smith [8], where there is an interaction for all pole orders. In contrast for the spherical harmonic formalism there is only a term for the charge-charge energy. Referring to [8] equation (48) the terms that come into the self interaction energy are those in the direct space sum for which the G functions defined in his equations (38) - (42) have a term which does not depend on $\underset{\sim}{r}_{jg}$. For the spherical multipole formalism all other terms for all higher multipoles depend on the interatomic separation $\underset{\sim}{r}_{AB}$, so there will be no equivalent self interaction term. The term for the charge-charge can be calculated in the usual way by expanding the error function as a power series [8]. Secondly, where

For Peer Review Only

A and B are atoms in the same molecule the explicit form of the interaction energy between them will need to be subtracted, for all multipole orders. This is because it was included to allow the Ewald transformation to be carried out. A third additional term arises from the reciprocal space terms which differ from the Cartesian case. From equation (21) it is clear that as $G(K_n)$ varies as $|\underset{\sim}{K}_n|^{-2}$, overall the terms in E_{rl} vary as $|\underset{\sim}{K}_n|^{l-2}$. For total multipole moment $l < 2$ the $\underset{\sim}{K}_n = 0$ terms will diverge. The justification for omitting the $\underset{\sim}{K}_n = 0$ terms is that surface charges will rearrange to eliminate any macroscopic field across the system. There has been recent discussion in the literature about this and its application to polar organic molecules [56, 57, 58]. For $l > 2$ the $\underset{\sim}{K}_n = 0$ term is zero. The situation for total multipole moment $l = 2$ is more interesting, corresponding to dipole dipole and quadrupole charge interactions. For example, equation (24) for dipole dipole interactions becomes

$$E_{rl} = \frac{1}{4\pi\epsilon_0}\frac{2\pi}{3V}\sum_{\tau} U_{\tau 11}(-1) \times \sum_{A} S_{F1\tau}(\Omega_A, 0) \times \sum_{B} S_{F1\tau}(\Omega_B, 0) \tag{38}$$

(For $\underset{\sim}{K}_n = 0$ the extra factor of 2 introduced in equation (7) must be left out). The terms involving S_{F11} are indeterminate and have a limit which depends on the direction from which zero is approached. However the second dipole dipole term has a finite limit

$$\sum_{\tau=2}^{4}\sum_{A} S_{F1\tau}(\Omega_A, 0) = \mu \tag{39}$$

where μ is the total dipole moment of the cell due to the dipoles (not including any dipole moment of the cell due to the point charges). The present version of DL_MULTI allows this term to be included or excluded at the choice of the user. In any event the magnitude of the term is small. Including the $\underset{\sim}{K}_n = 0$ term will give an identical sum for dipole dipole interactions as is obtained using the Cartesian formalism from [8]. It is interesting to note that the individual terms in the direct and reciprocal space parts for the dipole dipole interaction are however completely different. As has been noted above, for dipole dipole interactions there is no self interaction term in the spherical harmonic case. The Cartesian direct space sum also contains a first order term in equation (39) of [8] which is absent from the spherical harmonic formalism. The reciprocal space

For Peer Review Only

term in contrast has an additional term in the spherical harmonic case. This has a term with a finite limit for $\underset{\sim}{K}_n = 0$ which needs to be included. If this term is omitted, DL_MULTI gives an energy of zero for a unit cell with one particle consisting of a lattice of dipoles all pointing in the same direction, in agreement with [59].

Calculation of the forces and torques for the lattice summed interactions from these equations is straightforward and the formulae will not be given here. The torques are derived by calculating the derivatives of the energy with respect to the quaternion parameters used to define the orientation of the molecules, and calculating the torques from these using the equations of quaternion algebra. This will directly calculate the torque about the centre of mass of the molecule without needing to calculate any site-site torques. DL_MULTI also calculates the virial to compute the pressure, and the strain derivatives of the energy for constant pressure MD simulations.

4 Cutoffs

In practice all summations need to be truncated at some finite cutoff. Drawing on the experience of Steinbach and Brooks [60] it was found necessary to pay particular attention to these. These authors note that, in a system with fixed multipoles, there can be a drift in the total energy of the system during a molecular dynamics simulation. Normally, convergent long range forces, which decay as the inverse fourth or greater power of the interatomic separation would be evaluated by direct summation to a given cutoff distance. The tiny error introduced by two atoms moving so that their separation goes over the cutoff distance cancels when the atoms move to being within the cutoff distance again if the potential only depends on the atomic separation. However, when the atoms interact by an anisotropic potential, the changes in the relative orientation whilst the atoms are separated by more than the cutoff distance, prevent this cancellation being exact. It is important to note that this energy change is independent of the integration time step and is not due to errors in the integration method. The energy drift will be exacerbated in the examples we have been studying because it is in the solid state and the motions are periodic. Using the Ewald method for higher order multipoles, where the sums are absolutely convergent, although not

For Peer Review Only

strictly necessary, is desirable in order to reduce the energy drift. Even then it was found necessary to use large cutoffs for the higher multipole terms in direct space.

The following method is used to calculate the cutoffs in direct and reciprocal space. User supplied accuracy values ϵ_l^d and ϵ_l^r are provided for both direct and reciprocal space, for each total pole order l. Normally these will be set to be all equal, however in some cases it may be necessary to increase the accuracy for higher order poles in direct space to improve energy stability. It is also important to define the accuracy carefully. Usually we are interested in the cutoff needed so that the final term in the summation is less than a user defined proportion of a characteristic interaction energy of the whole system, rather than less than the total energy for this pole order summed to infinity. This means that the cutoffs for high order poles can be reduced if there are terms giving large energies from the monopoles. A characteristic energy of the system E_{char} is defined as the largest interaction energy for any two pole orders separated by the mean interatomic separation in the system. From equation (37)

$$\begin{aligned}\epsilon_l^d E_{char} &= \frac{1}{4\pi\epsilon_0} \times \left(\tilde{Q}^A \tilde{Q}^B\right)_{max} \\ &\quad \left\{ r_{cut,l}^{-l-1} - \frac{r_{cut,l}^l}{(2l-1)!!} 2 \left(\frac{\gamma}{\pi}\right)^{1/2} (2\gamma)^l F_l\left(\gamma r_{cut,l}^2\right) \right\}\end{aligned} \tag{40}$$

where $\begin{pmatrix} l_1 + l_2 \\ l_1 \end{pmatrix} \bar{S}_{l_1 l_2 l}^{k_1 k_2}(\Omega_A, \Omega_B, \widehat{\underset{\sim}{r} - \underset{\sim}{r}_N})$ from equation (37) is taken to be of order unity.

A user supplied value for $r_{cut,0}$ is used in this equation with ϵ_0^d to calculate a value for γ. The same value of γ is then used for all pole orders. Using this value of γ and ϵ_l^d for $l > 0$ values of $r_{cut,l}$ are calculated. A check is made that these values are all less than $r_{cut,0}$. This will usually be the case as the expansion of the charge distribution as a DMA in spherical harmonics gives values of $\left(\tilde{Q}^A \tilde{Q}^B\right)_{max}$ which decrease with increasing l. If the value of one of the $r_{cut,l}$ is greater than $r_{cut,0}$ the value of γ is reset and the process repeated until a $r_{cut,l}$ less than the user supplied value is obtained.

For Peer Review Only

Turning now to reciprocal space, from (21)

$$\epsilon_l^r E_{char} = \frac{1}{2}\frac{1}{4\pi\epsilon_0}\left(\tilde{Q}^A\tilde{Q}^B\right)_{max}\frac{(-1)^l}{(2l-1)!!}\frac{8\pi}{V} K_{n,cut,l}^{-2}\exp\left(\frac{-K_{n,cut,l}^2}{4\gamma}\right)K_{n,cut,l}^l \quad (41)$$

where $\begin{pmatrix} l_1+l_2 \\ l_1 \end{pmatrix}\bar{S}_{l_1 l_2 l}^{\kappa_1\kappa_2}(\Omega_A,\Omega_B,\hat{K}_n)$ from equation (21) is taken to be of order unity.

These equations are solved for each total pole order l to determine $K_{n,cut,l}$. Overall $\epsilon_l^r E_{char}$ depends on $K_{n,cut,l}^{l-2}$. This means that for $l < 2$ the equations will always have a solution for $K_{n,cut,l}$. However for $l \geq 2$ there may not be a solution. This is because the terms in the energy for $l > 2$ have a maximum for a certain K_n, and the maximum may be less than the required accuracy. For $l = 2$ there is a finite limit for $K_n = 0$ which may be less than the desired accuracy. In these cases the reciprocal space sum for these pole orders is omitted entirely.

5 Results for example structures 5-azauracil and uracil

As examples to illustrate the method, simulations were performed on 5-azauracil and uracil. These examples are presented in this paper as an illustration of the capabilities of DL_MULTI and detailed analysis of the simulations will not be presented here. Table (1) compares with experiment the relaxed static lattice structures from DMAREL [40, 41] and the results from a DL_MULTI simulation.

This was carried out using the NST constant temperature, constant isotropic stress ensemble [1] using a Berendsen thermostat and barostat to equilibrate the sample. Other simulation details for both simulations were time step 3 fs, temperature 310 K, equilibration steps 2500, simulation steps 2500. For 5-azauracil a cutoff of 13.5 Åwas used and for uracil 9.0 Å. The test cases had, for uracil, 96 molecules (1152 atoms) in a 2 x 2 x 6 supercell, for 5-azauracil, 192 molecules (2112 atoms) in a 4 x 2 x 3 supercell. The final lattice parameters were averaged over the 2500 steps after equilibration. For uracil, there is a slight expansion of the lattice in the c direction compared with

For Peer Review Only

the DMAREL simulation and the cell angle β changes to be much closer to the experimental value. For 5-azauracil there is a more isotropic expansion of the lattice compared with DMAREL. Lattice expansion is expected in the DL_MULTI simulations because of thermal effects. Also, the isotropic short range parameters were fitted to reproduce room temperature crystal structures using static lattice simulations, so dynamic simulations might be expected to give slightly too high values for the cell parameters.

As a further test of the cutoff criteria, additional NVE simulations [1] were carried out using the relaxed structure of 5-azauracil as a starting point. The NST simulations had used values in equations (40), (41) $\epsilon_l^d = 0.0000001$ for $l > 0$ and ϵ_l^r and $\epsilon_0^d = 0.0001$, NVE simulations were carried out using these values for the ϵ and also using lower precision for ϵ_l^d for $l > 0$. Values of 0.00001 and 0.0001 were used. The results for the total cell energy are presented in figure (1).

As can be seen, when the thermostat is switched off after 2500 time steps there is an energy drift for the two lower precision calculations, but the high precision simulation is stable.

These simulations are both small in terms of system size and simulation time, but they do illustrate how DL_MULTI performs. The simulations were carried out using 4 processors of an IBM SP2 parallel processor. The 5-azauracil simulation takes 3 s per time step on this machine. Preliminary calculations on larger systems suggest that scaling to a larger number of processors on a parallel machine will be limited by the global summations needed for the reciprocal space sums. However, the number of reciprocal lattice points needed to achieve a given accuracy decreases with system size, so large simulations on multiprocessor machines will be feasible using DL_MULTI.

6 Acknowledgements

Stephen Spence, Graeme Day, Aileen Gray and Sarah (Sally) L. Price are thanked for their assistance in developing and testing DL_MULTI whilst funded by an EPSRC ROPA grant. Dr. W. Smith also gave valuable advice.

For Peer Review Only

Appendix A

In this appendix we derive the factorizations of the explicit expressions for the S functions as defined by equation (23). The expressions for the functions S_F are given first, followed by the equations for the interactions. $\underset{\sim}{\alpha}_w$ refers to one of the three molecule-fixed axes. First we define β and η which will be used in the following expressions. (jkl are the indices 1,2,3 used cyclically).

$$\beta_w = \underset{\sim}{K}_n \cdot \underset{\sim}{x}_1, \underset{\sim}{y}_1, \underset{\sim}{z}_1$$

The vector $\underset{\sim}{\eta}$ depends on the outer vector product of two molecule-fixed axes.

$$\underset{j=1,3}{\eta_{1j}} = z_j z_j \qquad \underset{j=1,3}{\eta_{1j+3}} = \sqrt{2} z_k z_l$$

$$\underset{j=1,3}{\eta_{2j}} = x_j x_j - y_j y_j \qquad \underset{j=1,3}{\eta_{2j+3}} = \sqrt{2}(x_k x_l - y_k y_l)$$

$$\underset{j=1,3}{\eta_{3j}} = 2 y_j z_j \qquad \underset{j=1,3}{\eta_{3j+3}} = \sqrt{2}(y_k z_l + y_l z_k)$$

$$\underset{j=1,3}{\eta_{4j}} = 2 z_j x_j \qquad \underset{j=1,3}{\eta_{4j+3}} = \sqrt{2}(z_k x_l + z_l x_k)$$

$$\underset{j=1,3}{\eta_{5j}} = 2 x_j y_j \qquad \underset{j=1,3}{\eta_{5j+3}} = \sqrt{2}(x_k y_l + x_l y_k)$$

Charge

$$S_{F01} = \tilde{Q}_{00}$$

Dipole

$$S_{F11} K_n = \sqrt{3} \sum_{w=1}^{3} \tilde{Q}_{1w} \beta_w$$

$$\underset{\sim}{S}_{F1,j+1} K_n = K_n \sum_{w=1}^{3} \tilde{Q}_{1w} \underset{\sim}{\alpha}_w$$

For Peer Review Only

Quadrupole

$$S_{F21}K_n^2 = -\sqrt{3}/2K_n^2\tilde{Q}_{21}$$

$$S_{F22}K_n^2 = 5\sqrt{3}/2\beta_z^2\tilde{Q}_{21} + 5\left(\beta_x\beta_z\tilde{Q}_{22} + \beta_y\beta_z\tilde{Q}_{23}\right) + 2.5\left(2\beta_x\beta_y\tilde{Q}_{25} + \left(\beta_x^2 - \beta_y^2\right)\tilde{Q}_{24}\right)$$

$$\begin{aligned}\underset{\sim}{S}_{F2,j+2}K_n^2 &= \sqrt{3}K_n \times \left\{\left[\sqrt{3}\beta_z\tilde{Q}_{21} + \beta_x\tilde{Q}_{22} + \beta_y\tilde{Q}_{23}\right]\underset{\sim}{z}\right. \\ &+ \left.\left[\beta_z\tilde{Q}_{22} + \beta_x\tilde{Q}_{24} + \beta_y\tilde{Q}_{25}\right]\underset{\sim}{x} + \left[\beta_z\tilde{Q}_{23} + \beta_x\tilde{Q}_{25} - \beta_y\tilde{Q}_{24}\right]\underset{\sim}{y}\right\}\end{aligned}$$

$$\underset{\sim}{S}_{F2,j+5}K_n^2 = K_n^2 \times \left\{\sqrt{\frac{3}{2}}\tilde{Q}_{21}\underset{\sim}{\eta}_1 + \frac{1}{\sqrt{2}}\left(\tilde{Q}_{24}\underset{\sim}{\eta}_2 + \tilde{Q}_{23}\underset{\sim}{\eta}_3 + \tilde{Q}_{22}\underset{\sim}{\eta}_4 + \tilde{Q}_{25}\underset{\sim}{\eta}_5\right)\right\}$$

Octopole

$$\begin{aligned}S_{F31}K_n^3 &= \frac{35}{2\sqrt{3}}\beta_z^3\tilde{Q}_{31} + \frac{\sqrt{2}.35}{4}\beta_z^2\left[\beta_x\tilde{Q}_{32} + \beta_y\tilde{Q}_{33}\right] \\ &+ \frac{\sqrt{5}.7}{2}\beta_z\left[\left(\beta_x^2 - \beta_y^2\right)\tilde{Q}_{34} + 2\beta_x\beta_y\tilde{Q}_{35}\right] \\ &+ \frac{\sqrt{10}.7}{4.\sqrt{3}}\left[\beta_x\left(\beta_x^2 - 3\beta_y^2\right)\tilde{Q}_{36} - \beta_y\left(\beta_y^2 - 3\beta_x^2\right)\tilde{Q}_{37}\right]\end{aligned}$$

$$S_{F32}K_n^3 = -\frac{15}{2\sqrt{3}}K_n^2\beta_z\tilde{Q}_{31} - \frac{\sqrt{2}.5}{4}K_n^2\left(\beta_x\tilde{Q}_{32} + \beta_y\tilde{Q}_{33}\right)$$

$$\begin{aligned}\underset{\sim}{S}_{F3,j+2}K_n^3 &= K_n \times \\ &\left\{\left[\frac{3}{2}\left(5\beta_z^2 - K_n^2\right)\tilde{Q}_{31} + \frac{\sqrt{6}.5}{2}\left(\beta_x\beta_z\tilde{Q}_{32} + \beta_y\beta_z\tilde{Q}_{33}\right) + \right.\right. \\ &\left.\left.\frac{\sqrt{15}}{2}\left(\left(\beta_x^2 - \beta_y^2\right)\tilde{Q}_{34} + 2\beta_x\beta_y\tilde{Q}_{35}\right)\right]\underset{\sim}{z}\right. \\ &+ \left[\frac{\sqrt{6}}{4}\left(5\beta_z^2 - K_n^2\right)\tilde{Q}_{32} + \sqrt{15}\left(\beta_x\beta_z\tilde{Q}_{34} + \beta_y\beta_z\tilde{Q}_{35}\right) + \right. \\ &\left.\frac{\sqrt{10}.3}{4}\left(\left(\beta_x^2 - \beta_y^2\right)\tilde{Q}_{36} + 2\beta_x\beta_y\tilde{Q}_{37}\right)\right]\underset{\sim}{x}\end{aligned}$$

For Peer Review Only

$$+ \quad \left[\frac{\sqrt{6}}{4}\left(5\beta_z^2 - K_n^2\right)\tilde{Q}_{33} + \sqrt{15}\left(\beta_x\beta_z\tilde{Q}_{35} - \beta_y\beta_z\tilde{Q}_{34}\right) + \frac{\sqrt{10}.3}{4}\left(-2\beta_x\beta_y\tilde{Q}_{36} + \left(\beta_x^2 - \beta_y^2\right)\tilde{Q}_{37}\right)\right] \underset{\sim}{y}\Bigg\}$$

Hexadecapole

$$\begin{aligned}
S_{F41}K_n^4 &= \frac{1}{8}\left[\left(7\beta_z^2 - 3K_n^2\right)\left(5\beta_z^2 - K_n^2\right) - 8\beta_z^2K_n^2\right]\tilde{Q}_{41} \\
&+ \frac{\sqrt{10}}{4}\left[\left(7\beta_z^2 - 3K_n^2\right)\left(\beta_x\beta_z\tilde{Q}_{42} + \beta_y\beta_z\tilde{Q}_{43}\right)\right] \\
&+ \frac{\sqrt{5}}{4}\left[\left(7\beta_z^2 - K_n^2\right)\left(\beta_x^2 - \beta_y^2\right)\tilde{Q}_{44} + 2\beta_x\beta_y\tilde{Q}_{45}\right] \\
&+ \frac{\sqrt{70}}{4}\left[\beta_x\beta_z\left(\beta_x^2 - 3\beta_y^2\right)\tilde{Q}_{46} - \beta_y\beta_z\left(\beta_y^2 - 3\beta_x^2\right)\tilde{Q}_{47}\right] \\
&+ \frac{\sqrt{35}}{8}\left\{\left[\left(\beta_x^2 - 3\beta_y^2\right)\beta_x^2 + \left(\beta_y^2 - 3\beta_x^2\right)\beta_y^2\right]\tilde{Q}_{48} + 2\left(\beta_x^2 - \beta_y^2\right)2\beta_x\beta_y\tilde{Q}_{49}\right\}
\end{aligned}$$

The equations for the factorization of S (equation (23)) are given by

Charge-Charge $\quad S_{F01}^A S_{F01}^B$

Charge-Dipole $\quad -1/\sqrt{3}S_{F01}^A S_{F11}^B$

Dipole-Charge $\quad +1/\sqrt{3}S_{F11}^A S_{F01}^B$

Dipole-Dipole $\quad -S_{F11}^A S_{F11}^B + \sum_{j=1}^{3} S_{F1j+1}^A S_{F1j+1}^B$

Charge-Quadrupole $\quad S_{F01}^A\left((1/\sqrt{3})S_{F21}^B + (\sqrt{3}/5)S_{F22}^B\right)$

Quadrupole-Charge $\quad S_{F01}^B\left((1/\sqrt{3})S_{F21}^A + (\sqrt{3}/5)S_{F22}^A\right)$

Dipole-Quadrupole $\quad -S_{F11}^A\left(S_{F21}^B + S_{F22}^B\right) - \sum_{j=1}^{3} S_{F1j+1}^A S_{F2j+2}^B$

Quadrupole-Dipole $\quad +S_{F11}^B\left(S_{F21}^A - S_{F22}^A\right) - \sum_{j=1}^{3} S_{F1j+1}^B S_{F2j+2}^A$

For Peer Review Only

Quadrupole-Quadrupole $S^A_{F21}S^B_{F21} + S^A_{F21}S^B_{F22} + S^A_{F22}S^B_{F21} + (7/5)S^A_{F22}S^B_{F22} - (5/3)\sum_{j=1}^{3} S^A_{F2,j+2}S^B_{F2,j+2} + \sum_{j=1}^{6} S^A_{F2,j+5}S^B_{F2,j+5}$

Charge-Octopole $+S^A_{F01}\left((\sqrt{3}/7)S^B_{F31} + (\sqrt{3}/5)S^B_{F32}\right)$

Octopole-Charge $-S^B_{F01}\left((\sqrt{3}/7)S^A_{F31} + (\sqrt{3}/5)S^A_{F32}\right)$

Dipole-Octopole $-S^A_{F11}\left(S^B_{F31} + S^B_{F32}\right) + \sum_{j=1}^{3} S^A_{F1,j+1}S^B_{F3,j+2}$

Octopole-Dipole $-S^B_{F11}\left(S^A_{F31} + S^A_{F32}\right) + \sum_{j=1}^{3} S^B_{F1,j+1}S^A_{F3,j+2}$

Charge-Hexadecapole $S^A_{F01}S^B_{F41}$

Hexadecapole-Charge $S^B_{F01}S^A_{F41}$

References

[1] W. Smith and T.R. Forester, Journal of Molecular Graphics **14**, 136 (1996).

[2] A.J. Stone, *The Theory of Intermolecular Forces*, (Clarendon Press, Oxford, 1996).

[3] A. Toukmaji and J. Board, Computer Physics Communications, **95**, 73 (1995).

[4] T. Darden, D. York and L. Pedersen, J. Chem. Phys., **98**, 10089 (1993).

[5] L. Greengard and V. Rokhlin, J Comput. Phys., **73**, 325 (1987).

[6] K. Schmidt and M. Lee, J. Stat. Phys., **63**, 1223 (1991).

[7] K. Schmidt and M. Lee, J. Stat. Phys., **89**, 411 (1997).

[8] W. Smith, CCP5 Newsletter Number 46 (October 1998), 18 This article can be downloaded from ftp//ftp.dl.ac.uk/ccp5.newsletters/46/ps/smith.ps.gz.

For Peer Review Only

[9] A. Toukmaji, C. Sagui, J. Board, and T. Darden, J. Chem. Phys., **113**, 10913 (2000).

[10] M. Challacombe, C. White, and M. Head-Gordon, J. Chem. Phys., **107**, 10131 (1997).

[11] J. Bernstein, *Polymorphism in Molecular Crystals*, (Clarendon Press, Oxford, 2002).

[12] A. Gavezzotti, CrystEngComm, **4**, 343 (2002).

[13] J.P.M. Lommerse et al., Acta Crystallog. Section B - Structural Science, **56**, 697 (2000).

[14] W.D.S. Motherwell et al., Acta Crystallog. Section B - Structural Science, **58**, 647 (2002).

[15] T. Beyer, T. Lewis and S.L. Price, CrystEngComm, **3**, 183 (2001).

[16] M.L. Peterson et al., J. Am. Chem. Soc., **124**, 10958 (2002).

[17] T. Beyer,G.M. Day and S.L. Price, J. Am. Chem. Soc., **123**, 5086 (2001).

[18] J. Bernstein R.J. Davey and J.O. Henck, Angewandte Chemie-International Edition, **38**, 3441 (1999).

[19] J.D. Dunitz and J. Bernstein, Accounts Chem. Res., **28**, 193 (1995).

[20] B. van Eijck, J. Comput. Chem., **22**, 816 (2001).

[21] A. Gavezzotti and G. Filippini, J. Am. Chem. Soc., **117**, 12299 (1995).

[22] A.T. Anghel,G.M. Day and S.L. Price, CrystEngComm, **4**, 348 (2002).

[23] W.T.M. Mooij, B.P. van Eijck, S.L. Price, P. Verwer and J. Kroon, J. Comput. Chem., **19**, 459 (1998).

[24] T. Beyer and S.L. Price, J. Phys. Chem. B, **104**, 2647 (2000).

[25] R. Righini, Physica, **131B**, 234 (1985).

[26] R.M. Lynden-Bell, R.W. Impey and M.L. Klein, Chem. Phys., **109**, 25 (1986).

[27] M.L. Klein, Ann. Rev. Phys. Chem., **36**, 525 (1985).

For Peer Review Only

[28] S.L. Price and T. Beyer, Trans. Amer. Cryst. Assocn, **33**, 23 (1998).

[29] B. van Eijck, W. Mooij, and J. Kroon, J. Comput. Chem., **22**, 805 (2001).

[30] M. Tremayne et al., 2002, J. Amer. Chem. Soc., **126** 7071 (2004).

[31] A.J. Stone and M. Alderton, Mol. Phys., **56**, 1047 (1985).

[32] D.S. Coombes, S.L. Price, D.J. Willock and M. Leslie, Journal of Physical Chemistry, **100**, 7352 (1996).

[33] A.D. Buckingham, P.W. Fowler, and A.J. Stone, International Reviews in Physical Chemistry, **5**, 107 (1986).

[34] S.L. Price, J. Chem. Soc. Farad. Trans., **92**, 2997 (1996).

[35] A. Stone, A. Dullweber, M.P. Hodges, P.L.A. Popelier and D.J. Wales, ORIENT. ORIENT can be downloaded from http://www-stone.ch.cam.ac.uk/programs.html.

[36] D.J. Willock, S.L. Price, M. Leslie, and C.R.A. Catlow, J. Comput. Chem., **16**, 628 (1995).

[37] S.L. Price, D.J. Willock, M. Leslie, and G.M. Day, DMAREL This manual can be downloaded from http//www.ucl.ac.uk/ ucca17p/dmarelmanual/dmarel.html. (2001).

[38] P.L.A. Popelier and A.J. Stone, Mol. Phys., **82**, 411 (1994).

[39] S.Y. Liem, P.L.A. Popelier,and M. Leslie, Int. J. Quantum Chemistry, **99** 685 (2004).

[40] S.L. Price and K.S. Wibley, Journal of Physical Chemistry A, **101**, 2198 (1997).

[41] B.S. Potter, R.A. Palmer,R. Withnall, B.Z. Chowdhry and S.L. Price, Journal of Molecular Structure, **486**, 349 (1999).

[42] P.M. Rodger, A.J. Stone, and D.J. Tildesley, Molecular Simulation, **8**, 145 (1992).

[43] P.M. Rodger, A.J. Stone, and D.J. Tildesley, Mol. Phys., **63**, 173 (1988).

For Peer Review Only

[44] P.M. Rodger, A.J. Stone, and D.J. Tildesley, Journal of the Chemical Society-Faraday Transactions II, **83**, 1689 (1987).

[45] P.M. Rodger, A.J. Stone, and D.J. Tildesley, Chemical Physics Letters, **145**, 365 (1988).

[46] J.B.O. Mitchell, S.L. Price, M. Leslie,D. Buttar, and R.J. Roberts, Journal of Physical Chemistry A, **105**, 9961 (2001).

[47] M.J. Frisch, J.A. Pople, et al., GAUSSIAN. (Gaussian Inc.), Pitsburgh. (1998)

[48] B.S. Potter,R.A. Palmer, R. Withnall, B.Z. Chowdhry, and S.L. Price, GDMA. GDMA Can Be Downloaded from http//www-stone.ch.cam.ac.uk/programs.html. (1999).

[49] D.E. Williams and S.R. Cox, Acta Crystallog. Section B - Structural Science, **40**, 404 (1984).

[50] S.R. Cox L.Y. Hsu and D.E. Williams, Acta Crystallog. Section A - Foundations of Crystallography, **37**, 293 (1981).

[51] A.J. Stone, in *Theoretical Models of Chemical Bonding, Part 4*, edited by Maksić, Z., (Springer-Verlag, 1991).

[52] D.M. Brink and G.R. Satchler it Angular Momentum, (Oxford University Press, London, 1968).

[53] R.J.A. Tough and A.J. Stone, J. Phys. A, **10**, 1261 (1977).

[54] V.R. Saunders C. Freyria R. Dovesi L. Salasco and C. Roetti, Mol. Phys., **77**, 629 (1992).

[55] S.L. Price, A.J. Stone, and M. Alderton, Mol. Phys., **52**, 987 (1984).

[56] B. van Eijck, and J. Kroon, J. Phys. Chem. B, **101**, 1096 (1997).

[57] B. van Eijck and J. Kroon, J. Phys. Chem. B, **104**, 8089 (2000).

[58] W. Wedemeyer et al., J. Phys. Chem. B, **104**, 8090 (2000).

[59] S. de Leeuw, J. Perram and E. Smith, Proc. R. Soc. Lond. A, **373**, 27 (1980).

For Peer Review Only

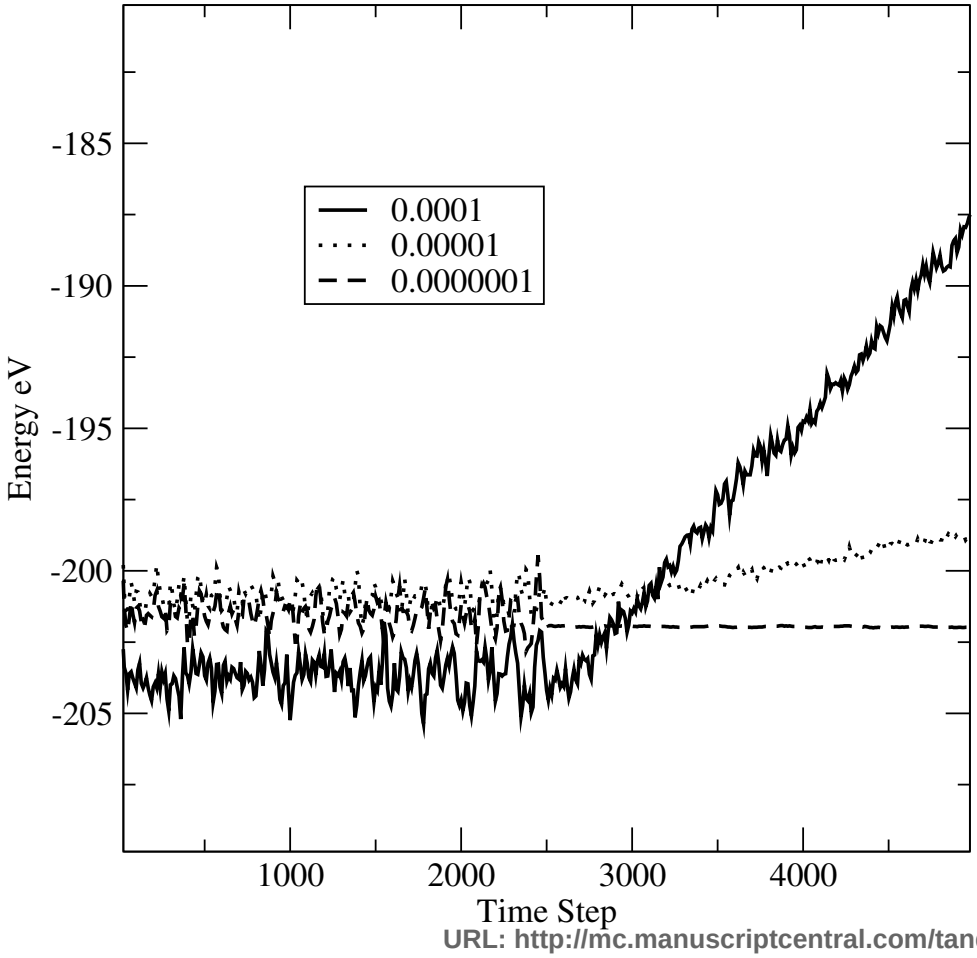
[60] P.J. Steinbach and B.R. Brooks, J. Comp. Chem., **15**, 667 (1994).

For Peer Review Only

Table 1: Experimental and theoretical cell constants for uracil and 5-azauracil

		Experiment	Simulation	
			DMAREL	DL_MULTI
uracil	a/ Å	11.938	12.17	12.11
	b/ Å	12.376	12.66	12.65
	c/ Å	3.655	3.70	3.81
	β/ deg	120.9	124.0	121.8
5-azauracil	a/ Å	6.5135	6.814	6.927
	b/ Å	13.5217	13.781	13.98
	c/ Å	9.5824	9.197	9.356

For Peer Review Only


0.0001
0.00001
0.0000001
-185
-190
-195
-200
-205
Energy eV
1000
2000
3000
4000
Time Step