American Mineralogist, Volume 82, pages 1049–1053, 1997

Ab initio calculations on point defects in forsterite (Mg_2SiO_4) and implications for diffusion and creep

JOHN BRODHOLT

Department of Geological Sciences, University College London, London, U.K.

ABSTRACT

Ab initio calculations on simple ionic vacancies in forsterite show that there are large energetic differences between vacancies on nonequivalent sites. Oxygen defects on the O3 site are between 1 and 3 eV lower in energy than on the O2 and O1 site, respectively. Magnesium defects on M1 sites are 0.8 eV lower in energy than on the M2 sites. These large energy differences mean that there are many orders of magnitude fewer vacancies on the less-favored sites. In the case of magnesium, where diffusion is thought to occur by means of a vacancy-hopping mechanism, the observed anisotropy with respect to crystallographic orientation can be rationalized by the fact that jumps must only be from M1 to M1 sites

INTRODUCTION

Point defects are of considerable interest to geologists because they have an essential role in diffusion, electrical conduction, creep, and seismic attenuation in minerals. For example, diffusion most commonly occurs by a vacancy method, whereby ions jump into neighboring vacancies. At high temperatures electrical conductivity may occur by ionic diffusion, whereas at lower temperatures defects can introduce electronic states in the band gap. Under certain conditions, diffusional creep may be the dominant deformation mechanism, and even when dislocation creep is dominant it may be rate-limited by climb, which again is a diffusion mechanism mediated by vacancies. To understand and be able to make predictions about these processes requires a detailed knowledge of the defect types, their concentrations, and how they affect diffusion.

In this paper we present first-principles quantum mechanical calculations on vacancies in forsterite, Mg_2SiO_4. We find that for both Mg and O large differences in energy exist between vacancies on different nonequivalent sites. This finding agrees with previous pair-potential calculations and means that Mg vacancies in forsterite occur essentially on only the M1 site and that O vacancies only occur on the O3 site.

CALCULATIONS

Calculations were performed using the parallel density functional theory (DFT) code CETEP (Payne et al. 1992). This code uses the generalized gradient approximation (GGA) for the exchange correlation (Perdew et al. 1992) and ab-initio norm-conserving non-local Kleinman-Bylander type pseudopotentials for the valence-core electron interactions (Kleinman and Bylander 1982; Lin et al. 1993). The use of pseudopotentials means that only the valence electrons are handled explicitly, which results in significant savings in computational requirements. Full periodic boundary conditions are employed, and the valence orbitals are represented by a plane-wave expansion. There are no adjustable parameters at this level of theory other than those used in constructing the pseudopotential, and no experimental data are used except for the mass and charge of the electron and Planck's constant. The only significant approximation in these calculations is that of the exchange-correlation functional. A substantial advantage of the plane-wave basis set is that forces and stresses may be easily evaluated using the Hellman-Feynman theorem (Payne et al. 1992) and used to relax both the atomic coordinates and the unit-cell parameters to minimize the total energy (Wentzcovitch et al. 1993). This is especially important with defects because there are significant energy differences (> 1.0 eV) between relaxed and unrelaxed defect structures.

All calculations were performed on one unit cell of 28 atoms containing one charged defect (either Mg^{2+}, O^{2-}, or Si^{4+}). We used a four *k*-point sampling set chosen using the Monkhorst-Pack scheme (Monkhorst and Pack 1976). Plane waves up to an energy of 600 eV were accepted, which gives a basis set of about 10,000 plane waves for each *k* point. We tested for the likely error due to the 600 eV cutoff by comparing energies for the perfect lattice obtained using the 600 eV cutoff with those using 700 and 800 eV cutoffs. The absolute difference between lattice energies obtained using 600 and 700 eV cutoffs is 0.07 eV; between 600 and 800 eV cutoffs it is 0.09 eV. These values suggest that the likely absolute error in any one calculation because of the finite number of plane waves is about 0.1 eV. The error in defect energies is substantially smaller, however, because energy differences converge much faster than absolute energies, and they are negligible in comparison with the size of the defect energies ($\sim$1 eV). We also tested for the effect of

0003–004X/97/1112–1049$05.00

TABLE 1. Summary of cell parameters and bond lengths (Å)

	Calculated	Data*
a	10.15	10.19
b	5.96	5.978
c	4.71	4.753
Si-O1	1.58	1.614
Si-O2	1.62	1.654
Si-O3	1.60	1.633
M1-O1	2.08	2.081
M1-O2	2.06	2.067
M1-O3	2.13	2.131
M2-O1	2.17	2.177
M2-O2	2.05	2.045
M2-O3a	2.07	2.065
M2-O3b	2.20	2.210

* Data from Kudoh and Takéuchi (1985).

k-point sampling by comparing four *k*-point results with those obtained using a nine and twelve Monkhorst-Pack *k*-point set. These tests were performed on both the relaxed lattice and a lattice containing an O1 defect. For both the relaxed lattice and the lattice with an O1 defect, the energy differences between four and nine *k* points and four and twelve *k* points were 0.04 and 0.045 eV, respectively. Again this means that the error in defect energies because of incomplete *k*-point sampling is negligible.

CORRECTING INTERACTIONS BETWEEN DEFECTS

The energy calculated for one unit cell with a charged defect is actually the energy for a periodic array of unit cells each containing a defect. This is not the energy we want. The energy of interest is that of an infinite lattice containing only one defect. Although we could get around this by increasing the size of the periodically repeating unit until defects are essentially isolated from each other, the heavy computational requirements of these calculations makes this impossible. It can be shown, however, that we can correct for the dominant term by taking into account the energy of interaction between the defects (Leslie and Gillan 1985). For a cubic unit cell this term is:

$$E_{\text{corr}} = \tfrac{1}{2}\alpha q^2/\epsilon_0 L$$

where α is the Madelung constant for the cell, L is the cell length, ϵ_0 is the dielectric constant for the perfect crystal, and q is the net charge of the defect. This term is in fact the energy of a periodically repeating array of charges surrounded by a neutralizing background in a dielectric medium of the crystal rather than a vacuum. For forsterite the unit cell is orthorhombic so the Madelung constant is not readily available. This energy correction can be calculated simply, however, since it is equivalent to the Ewald energy for an array of charges in a vacuum (where $\epsilon_0 = 1.0$) with the appropriate box dimensions and dividing by ϵ_0 for forsterite. For the zero pressure cell parameters used here the correction is then:

$$E_{\text{corr}}(\text{eV}) = 2.524q^2/\epsilon_0$$

where $q = 2$ for O and Mg ion defects and $q = 4$ for Si. A value of 6.2 was used for the dielectric constant of forsterite (Weast and Astle 1981).

TABLE 2. Comparison of cell parameters (Å) calculated using different combinations of GGA, LDA, and *k* points

	LDA 1*k*pt	GGA 1*k*pt	LDA 4*k*pt	GGA 4*k*pt	Data*
a	9.988	10.05	10.08	10.15	10.19
b	6.07	6.07	5.99	5.96	5.978
c	4.64	4.65	4.69	4.71	4.753

* Data from Kudoh and Takéuchi (1985).

RESULTS

Table 1 summarizes the zero pressure cell parameters and bond lengths for the relaxed structure with no defects. At zero pressure, the calculated cell parameters are $a = 10.15$, $b = 5.96$, and $c = 4.71$ Å. (Note that in this paper all directions and cell parameters are relative to space group *Pnma.* In *Pnma* $a > b > c$, whereas in *Pbnm* $b > c > a$.) These cell parameters compare well with the experimental values of $a = 10.19$, $b = 5.978$, and $c = 4.753$ Å (Fugino et al. 1981). The metal-oxygen bond lengths are also in extremely good agreement with the experimental data, whereas Si-O bond lengths are somewhat smaller than experimental values by 0.03 Å. The order of increasing bond length is correct, however, with Si-O1 < Si-O3 < Si-O2. There is a small improvement in cell parameters and bond lengths over those calculated using one *k* point and the local density approximation (LDA) given in Brodholt et al. (1996). This improvement is due to a combination of increased *k*-point sampling and using GGA instead of LDA. The two effects can be separated from each other by comparing LDA and GGA one *k*-point (Γ point) results with four *k*-point results. These are summarized in Table 2. GGA expands the cell in the *a* direction in comparison to LDA, makes no difference in the *c* direction, but interestingly the *b* cell edge is reduced slightly in the four *k*-point case. Even with GGA and four *k* points, however, the cell parameters remain slightly smaller than the experimental values; this contrasts with recent work on metals and semiconductors that shows LDA tends to produce lattice constants that are too small and GGA tends to produce lattice constants that are too large (Filippi et al. 1994; Khein et al. 1995; Winkler et al. 1995). A greater effect on cell parameters comes, however, from better *k*-point sampling (Table 2). As mentioned previously, increasing the *k*-point sampling to nine or twelve points does not change the cell parameters further.

Figure 1 shows the structure of forsterite and indicates the positions of the three nonequivalent O sites (O1, O2, and O3) and the two nonequivalent metal sites (M1 and M2). Table 3 summarizes the calculated energy of the fully relaxed structure with no defect together with the relaxed lattice energy of the unit cell containing one charged defect for each atom type and each nonequivalent site. Note that, unlike pair-potential calculations, the en-

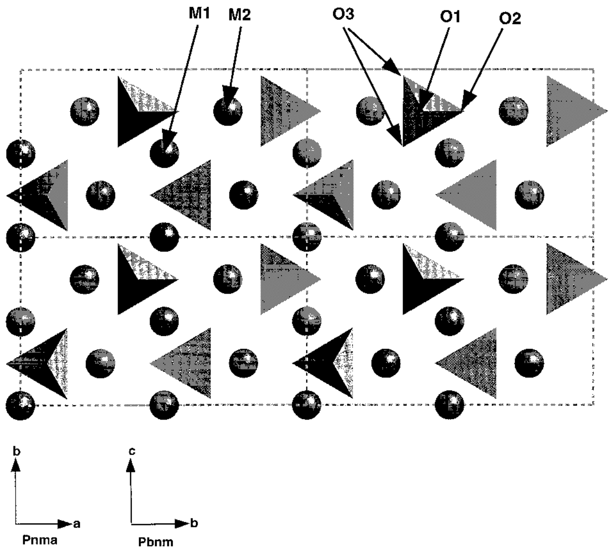

FIGURE 1. Four unit cells of forsterite showing the two different cation sites and the three different O sites. Also shown are the coordinates for *Pnma* and *Pbnm* space groups. In this paper we use *Pnma.*

ergy difference between the unit cell with and without a defect is not equal to the vacancy energy. Because the vacancy energy is the energy required to remove the ion to infinity, the energy of the isolated ion at infinity must be accounted for, which, unlike in pair-potential calculations, is not zero. In fact the vacancy energy is not important, because ions are not removed to infinity but create a formula unit of either the same or another phase. Nevertheless we can calculate the full Schottky defect energy and compare it with that calculated using pair potentials. The full Schottky defect energy is the energy needed to remove one formula unit of forsterite and add it to the surface. This is 30.25 eV, which compares with 34.1 eV calculated with empirical pair potentials (Wright and Catlow 1994; Jaoul et al. 1995). (To calculate the Schottky defect energy from these pair-potential calculations, the energy of a formula unit of forsterite is needed. Using the same interatomic potentials, this is 212.5 eV pfu.)

The most interesting result, though, is that very large differences in energy exist between different sites for the same atomic species. The O vacancy on the 03 site is 1.07 eV lower in energy than on the O2 site and 2.75 eV lower than on the O1 site. Similarly an Mg vacancy on the M1 site is 0.81 eV lower in energy than on the M2 site. These differences in energy are very large and important because defect concentrations are exponentially inversely dependent on defect energy. Assuming a similar pre-exponential factor, the concentration ratio of two similar ions on two different sites is:

$$\frac{n_2}{n_1} = \exp(-\Delta E/kT)$$

TABLE 3. Summary of lattice energies for the undefected lattice (RL) and with the six different ionic vacancies

	E (eV)	E_{corr} (eV)
RL	−7547.40	
O1	−7125.86	−7124.54
O2	−7127.54	−7126.22
O3	−7128.61	−7127.29
Si	−7402.76	−7397.48
M1	−7505.35	−7504.03
M2	−7504.54	−7503.22

Notes: E_{corr} is the calculated lattice energy plus the correction for interactions between defects. The O3 vacancy is energetically much more favorable than either the O1 or O2. Similarly an Mg^{2+} vacancy on the M1 site is 0.8 eV lower in energy than on the M2 site.

where ΔE is the energy difference for the defect formation and n is the number of vacancies. Taking the two Mg sites with $\Delta E = 0.81$ eV, the ratio of vacancy concentrations on the more favorable to the less favorable site is 10^{17} at room temperature and only decreases to 10^2 at the melting temperature (2163 K). This ratio is even greater for vacancies on different O sites. Essentially, ionic vacancies exist only on one of the different possible sites.

These results are in qualitative agreement with those calculated using simple empirical pair potentials (Jaoul et al. 1995). In that study, a defect on the M1 site was about 1.9 eV more likely than on the M2 site, and an O3 defect was 3.6 and 1.2 eV lower in energy than O1 and O2 defects, respectively. Although they found a greater difference between the two metal sites than in the present study, the result is essentially the same; there are at least two orders of magnitude more defects on the M1 than on the M2 site, and even greater differences between vacancy concentrations on the O sites. Although this does not prove the results, the fact that similar results are obtained from a completely independent source greatly strengthens the conclusions.

IMPLICATIONS FOR DIFFUSION

Diffusion is controlled to a large part by the height of the energy barrier over which a migrating ion must cross. Although it is certainly possible to calculate these at the same level of theory as for the defects above, the computational cost of finding the minimum energy path is prohibitive unless the exact path is well constrained by symmetry. For this reason we have not attempted this at present. Nevertheless, it is possible to make some heuristic observations based on these results that relate to diffusion in forsterite.

The observation that there are effectively no vacancies on the M2 site means that, assuming a vacancy mechanism, diffusing Mg^{2+} ions can only jump from M1 to M1 sites. Figure 2 depicts the shortest straight-line paths that a diffusing Mg ion may take, but with the restriction that jumps must be from M1 to M1 sites only. For M1 ions diffusing in the [100] direction, the shortest paths would be successive jumps of 3.6 and 3.2 Å, if jumps to M2

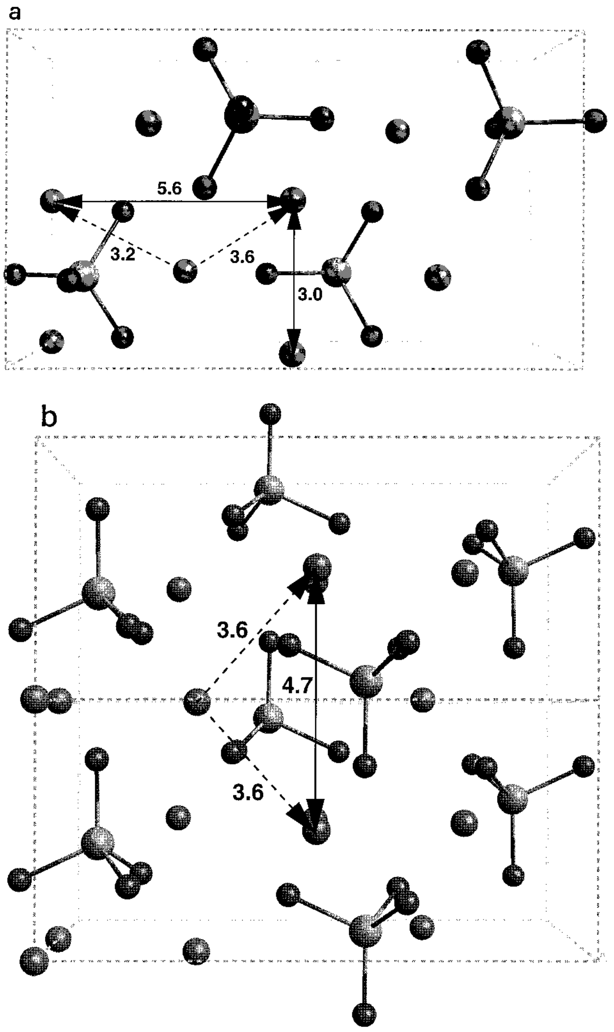

Figure 2. (**a**) View of the forsterite structure looking down [001]. The solid lines show the shortest straight-line paths and distances (in Å) that jumping cations must take to diffuse in the [100] and [010] directions under the restriction that jumps must be from M1-to-M1 sites only. The dashed lines show that shorter jumps exist in the [100] direction if M1-to-M2 jumps were allowed. Because effectively no M2 vacancies exist, cations migrating in the [100] direction must take the longer jumps of ~5.6 Å rather than the shorter 3.2 and 3.6 Å jumps. Cations diffusing in the [010] direction do not need to pass into M2 sites and are therefore unaffected by the absence of vacancies on the M2 site. (**b**) Same as **a** except looking down [010]. As with cations diffusing in the [100] direction, the absence of M2 vacancies changes the shortest diffusing jump lengths from 3.6 to 4.7 Å.

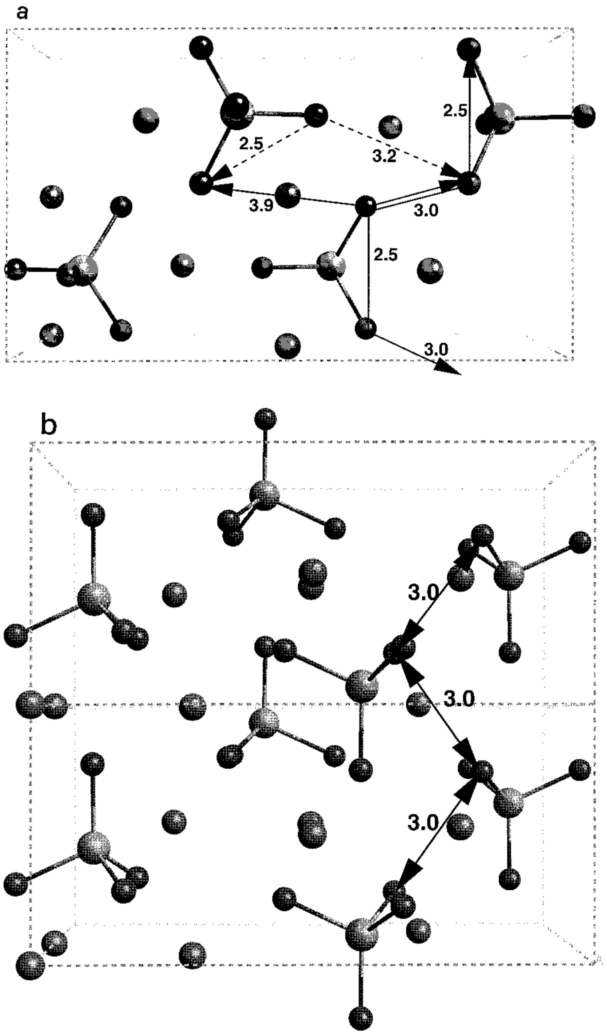

Figure 3. (**a**) Same as Figure 2a but showing jumps for a diffusing O ion (assuming a vacancy method). The absence of vacancies on the O1 and O2 sites only affects diffusion in the [100] direction. The solid lines show the shortest jumps avoiding O1 and O2 sites, whereas the dashed lines show that somewhat shorter jumps are possible if O3-to-O2 jumps occur. (**b**) For an O atom diffusing in the [001] direction the shortest paths are from O3 to O3 and so the absence of vacancies on O1 and O2 sites is probably not important for this direction.

sites were allowed (shown by dashed line in Figure 2a). The restriction of M1-to-M1 sites though increases the jump distance to 5.6 Å. The other option is by an intermediate interstitial site, but this also effectively adds to the jump distance because the interstitial ion is at a higher energy than in the original M1 site. Similarly, in the [001] direction the absence of M2 vacancies increases jump distances from successive jumps of 3.6 Å to one of 4.7 Å. For an M1 cation diffusing in the [010] direction, the shortest jumps are in fact from M1 to M1, avoiding M2 sites altogether. Experimentally, Mg diffusion in forsterite is found to be quite anisotropic, with tracer diffusion studies in almost pure forsterite showing $D_b \sim 4D_c \sim 6D_a$ (Chakraborty et al. 1994). The order of decreasing jump distance is exactly the same as the order of increasing diffusion rates. Under the reasonable assumption that smaller jumps are more likely to succeed than larger ones, the anisotropy of Mg diffusion in forsterite can be simply related to the jump distance. If M2 vacancies were allowed then one would expect a different order of diffusion anisotropy with diffusion in the [100] direction being similar or slightly slower than in [001], which is in disagreement with the experimental data.

Because O vacancies exist effectively only on the O3 site, we should be able to look at the restrictions imposed by only O3-to-O3 hoping for O diffusion too. Experimentally, however, O diffusion is found to be independent of oxygen fugacity (Jaoul et al. 1980), which means that the mechanism for O diffusion in forsterite is likely to be by an interstitial mechanism rather than by vacancy hopping (Gérard and Jaoul 1989). This makes an analysis of O jumps based on available vacancies less useful than for the cations. Nevertheless, for completeness, if we do look at jump distances, with the caveat that jumps are only from O3 to O3 sites, the longest jump distance is about 3.9 Å in the [100] direction vs. jumps of 2.5 and 3.0 Å in the [010] direction and 3.0 Å in the [001] direction (Fig. 3). The [100] direction is the only direction effected by the absence of O1 and O2 vacancies. If O2 vacancies were allowed, the shortest jumps in the [100] direction would be 2.5 and 3.2 Å instead of 3.9 Å (dashed line in Fig. 3). However, this is a much smaller increase in jump distance than occurs in the cation sites where the absence of M2 vacancies increases jump distances from 3.6 to 5.6 Å. An alternative mechanism for O diffusion may be by means of a cooperative motion of two ions whereby both the O3 and O2 ions move one site together. The intermediate state could involve two ions out of place with a lower energy than one interstitial with two vacancies (R. Wentzcovitch, personal communication).

It should be remembered that diffusion rates are controlled by the height and shape of the energy barrier as well as the jump geometry, so a true analysis of diffusion is much more complicated than the simple arguments made here. Nevertheless, we feel that the calculations on defect energetics given here are robust and at least qualitatively correct, and that a mechanism for cation diffusion must take into consideration the relative absence of M2 vacancies. In addition, this conclusion should also be considered for diffusion of other atomic species, such as H, that may be associated with cation defects.

ACKNOWLEDGMENTS

I gratefully acknowledge the Royal Society for receipt of a University Research Fellowship. I am also grateful to Renata Wentzcovitch and an anonymous reviewer for their constructive criticisms and suggestions. Thanks are also given to Atul Patel, Keith Refson, David Price, Dave Dobson, and Nicola Richmond for discussion and support.

REFERENCES CITED

Brodholt, J.P., Patel, A., and Refson, K. (1996) An ab initio study of the compressional behavior of forsterite. American Mineralogist, 81, 257–260.

Chakraborty, S., Farver, J.R., Yund, R.A., and Rubie, D.C. (1994) Mg tracer diffusion in synthetic forsterite and San Carlos olivine as a function of P, T and $_{O2}$. Physics and Chemistry of Minerals, 21, 489–500.

Filippi, C., Singh, D.J., and Umrigar, C.J. (1994) All-electron local-density and generalized-gradient calculations of the structural properties of semiconductors. Physical Review, B50, 14,947–14,951.

Fugino, K., Sasaki, S., Takéuchi, Y., and Sadanaga, R. (1981) X-ray determination of electron distributions in forsterite, fayerlite and tephroite. Acta Crystallographica, B37, 513–518.

Gérard, O. and Jaoul, O. (1989) Oxygen diffusion in olivine. Journal of Geophysical Research, 94, 4119–4128.

Jaoul, O., Froidevaux, C., Durham, W.B., and Michaut, M. (1980) Oxygen self-diffusion in forsterite: implications for the high-temperature creep mechanism. Earth and Planetary Science Letters, 47, 391–397.

Jaoul, O., Bertran-Alvarez, Y., Liebermann, R.C., and Price, G.D. (1995) Fe-Mg interdiffusion in olivine up to 9 GPa at T = 600–900 °C; experimental data and comparison with defect calculations. Physics of Earth and Planetary Interiors, 89, 199–218.

Khein, A., Singh, D.J., and Umrigar, C.J. (1995) All-electron study of gradient corrections to the local-density functional in metallic systems. Physical Review, B51, 4105–4109.

Kleinman, L. and Bylander, D.M. (1982) Efficacious form for model pseudopotentials. Physical Review Letters, 48, 1425–1428.

Kudoh, Y. and Takéuchi, Y. (1985) The crystal structure of forsterite Mg_2SiO_4 under higher pressure up to 149 kb. Zeitschrift für Kristallographie, 171, 291–302.

Leslie, M. and Gillan, M.J. (1985) The energy and elastic dipole tensor of defects in ionic crystals calculated by the supercell method. Journal of Physical Chemistry, 18, 973–982.

Lin, J.S., Qteish, A., Payne, M.C., and Heine, V. (1993) Optimized and transferable nonlocal seperable ab initio pseudopotentials. Physical Review, B47, 4174–4180.

Monkhorst, H.J. and Pack, J.D. (1976) Special points for brillouin-zone integrations. Physical Review, B13, 5188–5192.

Payne, M.C., Teter, M.P., Allan, D.C., Arias, T., and Joannopoulos, J.D. (1992) Iterative minimization techniques for ab-initio total energy calculations: Molecular dynamics and conjugate gradients. Review of Modern Physics, 64, 1045–1097.

Perdew, J.P., Chevary, J.A., Vosko, S.H., Jackson, K.A., Pedersen, M.R., Singh, D.J., and Fiolhais, C. (1992) Atoms, molecules solids and surfaces: Applications of the generalized gradient approximation for exchange and correlation. Physical Review, B46, 6671–6687.

Weast, R.C. and Astle, M.J., Eds. (1981) CRC Handbook of Chemistry and Physics, CRC Press, Inc., Boca Raton, Florida.

Wright, K. and Catlow, C.R.A. (1994) A computer-simulation study of (OH) deffects in olivine. Physics and Chemistry of Minerals, 20, 515–518.

Wentzcovitch, R.M., Martins, J.M., and Price, G.D. (1993) Ab-initio molecular dynamics with variable cell shape: Application to $MgSiO_3$. Physical Review Letters, 70, 3947–3050.

Winkler, B., Milman, V., Hennion, B., Payne, M.C., Lee, M.-H., and Lin, J.S. (1995) Ab initio total energy study of brucite, diaspore and hypothetical hydrous wadsleyite. Physics and Chemistry of Minerals, 22, 461–467.

MANUSCRIPT RECEIVED OCTOBER 4, 1996
MANUSCRIPT ACCEPTED JULY 15, 1997