Phys Chem Minerals (2004) 31: 12–21
DOI 10.1007/s00269-002-0268-0

ORIGINAL PAPER

M. S. Craig · M. C. Warren · M. T. Dove
J. D. Gale · D. Sanchez-Portal · P. Ordejon
J. M. Soler · E. Artacho

Simulations of minerals using density-functional theory based on atomic orbitals for linear scaling

Received: 5 July 2001 / Accepted: 16 May 2002

Abstract The use of quantum mechanics methods within the formalism of density functional theory requires a method to represent the electron wave functions. We compare the use of strictly localized basis functions based on atomic orbitals with the use of plane waves for the study of mineral properties and behaviour. Strictly localized functions enable the computational resources to scale linearly with the size of the system, whereas plane-wave methods scale more as the cube power of the system size, and for this reason the use of localized functions will be preferred for studies of large sizes. We present test results obtained from studies of cation ordering in spinel, garnet and amphibole phases, the high-pressure displacive phase transition in cristobalite, and the intercalation of organic molecules into pyrophyllite. We conclude that the use of localized basis sets provides a useful route forward for quantum mechanical studies of large-scale mineral problems.

M. S. Craig · M. C. Warren · M. T. Dove (✉) · E. Artacho
Department of Earth Sciences,
University of Cambridge, UK
e-mail: martin@esc.cam.ac.uk

J. D. Gale
Department of Chemistry,
Imperial College, London, UK

D. Sanchez-Portal
Department of Physics,
University of Illinois,
Urbana IL 61801, USA

P. Ordejon
Institut de Ciencia de Materials de Barcelona – CSIC,
08193 Bellaterra, Spain

J. M. Soler
Department of Physics,
Lyman Laboratory,
Harvard University,
Cambridge, MA 02138, USA

E. Artacho
Departamento de Fisica de la Materia Condensada
and Instituto Nicolas Cabrera,
Universidad Autonoma,
28049 Madrid, Spain

Keywords Linear scaling · SIESTA · Density-functional Theory · Atomic Orbitals

Introduction

In recent years, simulation studies of minerals using quantum mechanics have become increasingly important in helping to understand the behaviour and properties of minerals. Many of the studies in this area have used density-functional theory (DFT) (Payne et al. 1992), which allows exchange and correlation effects to be incorporated directly and efficiently, although in an approximate way. These studies sometimes also use pseudopotential methods to represent the inner (core) electrons in order to reduce the computational demands of the calculations. By contrast, Hartree–Fock methods incorporate exchange interactions exactly, but correlation effects are either neglected or incorporated as an a posteriori correction (Chartier et al. 1999). In applications to studies of minerals, density-functional theory calculations have typically represented the electron wave functions by superposition of plane waves, allowing the amplitude of each wave to be adjusted in the calculation in order to minimize the energy (the variational principle of DFT of Hohenberg and Kohn 1964). Plane waves are a natural choice for a system with periodicity, but they suffer from the fact that the time taken for a calculation scales with the cube of the number of atoms. Electronic structure calculations are demanding of computer power, and the scope of such calculations is always limited by the availability of computing resources. Developments in hardware (such as the availability of large parallel supercomputers) and developments in algorithms enable larger calculations to be performed. However, neither types of development can grow at a sustained rate to allow significant increases in N^3. Instead, it is necessary to look again at the basic methodology, in order to seek for methods that enable computational demands to scale linearly with sample size. The most promising methods are those that

represent the electron wave functions by localized functions rather than by plane waves (Ordejon 1998; Goedecker 1999).

Attaining linear scaling is possible by making explicit use of the near-sightedness principle (Kohn 1996), by which the effect of a local perturbation becomes negligibly small in a region sufficiently far away. The simplest way to relate this locality with linear scaling is thinking about studying a large system piecewise, the size of the (overlapping) pieces being determined by the range of the interactions. The computational effort then becomes proportional to the number of pieces and scales, therefore linearly with the system size. This argument gives a simple image for understanding linear scaling, but most methods nowadays use more elaborate ways of exploiting locality (Ordejon 1998). The main point in any case is to use strictly localized wave functions at every level of the calculations. Paradoxically, the long-range electrostatic interactions do not pose particular problems: the corresponding term in the energy and in the one-electron potential can be treated efficiently in linear-scaling fashion (White et al. 1994; Ordejon et al. 1996; Goedecker 1999).

Effective linear scaling is obtained when the size of the system becomes larger than the locality range. This gives a crossing size below which conventional methodologies are of advantage. This crossing point depends on system characteristics (most importantly whether the system is metal or insulator) and on differences among the particular implementations. It is clear, however, that the importance of these techniques will only grow with time as more powerful computers allow the study of larger and more complex systems.

Strictly localized numerical atomic orbitals have already proven able to provide efficient basis sets, well adapted to linear-scaling techniques (Sankey and Niklewski 1989; Artacho et al. 1999; Kenny et al. 2000; Junquera et al. 2001). The present study offers a comparison of the representation of the electron wave functions by plane waves and these localized functions. The comparison is centred on some mainstream problems in mineral physics, namely (1) cationic substitution and ordering, using a spinel and an amphibole as representative examples, (2) displacive phase transitions, studying cristobalite and (3) surface adsorption and intercalation, particularly for small aromatic molecules in/on pyrophyllite. The implementation of the localized atomic basis sets used in this study is in the program SIESTA, which has been described by Ordejon et al. (1996) and Artacho et al. (1999).

We have several aims in presenting this comparison of the two approaches: (1) to compare them in terms of accuracy, exploring levels of accuracy for different atomic basis functions, (2) to compare also their performance in terms of computational demands, and (3) to contrast methods in different types of applications. This evaluation of the performance of the linear-scaling method, as implemented within the SIESTA program, will allow the consideration of more complex systems, like surfaces and other non-periodic systems, for which the standard plane-wave techniques are not as applicable. In addition, the atomic-orbital method offers controlled accuracy, allowing less accurate but quicker calculations.

To conclude this introduction, it is worth making a few comments on the wider context of simulation methods. Empirical models have been used to tackle a number of problems concerned with mineral properties, such as cation ordering, with considerable success (Bosenick et al. 2001; Warren et al. 2001). However, there are cases where quantum mechanics are certainly necessary, for example when investigating effects associated with changing chemical bonds, or where pressure changes bond lengths beyond the limits over which the empirical model was fitted. Certainly, in these cases, quantum-mechanical methods are expected to give much more accurate calculations of energies. Unfortunately, there are many problems that are simply too large for the use of plane-wave implementation of DFT; some of these are used as examples in this paper. It is hoped that by using linear-scaling methods with localized basis functions it will be possible for quantum-mechanics methods to have a greater impact in studies of mineral behaviour. In this paper we use examples from a number of areas of potential interest to demonstrate the power of linear-scaling methods.

Methods

Both methods compared in this work belong to the family of first-principles methods based on density-functional theory (DFT). The essential ingredients and approximations common to both techniques are outlined in the following. Firstly, the adiabatic approximation is assumed to decouple electronic from nuclear degrees of freedom, i.e. the quantum-mechanical ground-state energy of the electrons for given nuclear positions defines the potential energy for the motions of the nuclei. By virtue of this, not only electronic properties are accessible to these methods, but also the properties associated to preferred positions and dynamics of the nuclei, namely, structural, vibrational and elastic properties, for example. Furthermore, if the dynamics of the nuclei are solved for their classical time evolution (called molecular-dynamics simulations), these methods offer thermodynamic information as well.

The basic goal is thus finding the quantum-mechanical ground-state energy of the electrons in the system for given nuclear positions. To deal with the many-electron problem, the density-functional theorem is employed (Hohenberg and Kohn 1964) which asserts that the energy of the ground state of a many-body system can in principle be obtained as a functional of its particle density. The local density approximation (Kohn and Sham 1965) provides a practical (albeit approximate) implementation of that functional. It proposes an alternative system of formally non-interacting particles under the influence of a local potential, which depends on the electron density at the point where it is evaluated, and which includes the electron–electron interaction in an effective way. The essentially non-local effects of electron exchange and correlation are described by an effective local potential obtained by fitting (Perdew and Zunger 1981) to the result of an independent, very accurate calculation for the full many-body problem of the free electron gas (Ceperley and Alder 1980). The original many-particle problem is thus approximated by a one-particle problem that has to be solved self-consistently (the potential depends on the density which, in turn, depends on the solution).

This scheme proved successful beyond expectations, considering the inherent complexity of the many-electron problem. Many implementations of it have been used for studying a large variety of systems (see Payne et al. 1992), showing accuracy of around 1% in structural properties of many materials, 5–10% in binding energies, and below 5% in elastic and vibrational properties. The generalized-gradient approximated (GGA) functionals proposed by several authors represent a step towards the exact DFT functional, by considering that the exchange-correlation potential at a point depends not only on the electronic density but also on its gradient and (in some cases) on higher derivatives. The scheme of Perdew et al. (1996) (PBE) is used below.

The inclusion of core electrons in the calculation increases the computational load since (1) the number of electrons is higher, and (2) their wave functions vary very rapidly. They are, however, very insensitive to variations in the chemical environment. Therefore, core electrons are not explicitly considered in the calculations but are replaced by non-local norm-conserving pseudopotentials, also calculated from first principles. In this work the pseudopotentials within the linear-scaling method were obtained according to the improved Troullier and Martins scheme (1991), whereas the ones within the PW method were the ultrasoft pseudopotentials of Vanderbilt (1990).

When some core electrons are shallow in energy and overlapping in space with the valence electrons, better accuracy can be obtained by including a part of the core in the particle density for the computation of the (non-linear) exchange-correlation potential, the so-called partial-core exchange-correlation correction (Louie et al. 1982). Additionally, scalar relativistic effects of the core electrons can also be included in the calculations (Bachelet and Schluter 1982).

After all these considerations, the problem to be solved is a one-particle quantum-mechanical problem with an effective potential. This problem is solved by expanding the one-particle wave functions in a finite basis and diagonalizing the Hamiltonian matrix that results. It is in these two steps (building the H matrix and solving it) that the methods compared here diverge. The PW method is a more mature and quite efficient one, but cube-scaling. The basis set it uses is made of plane waves up to a certain kinetic-energy cutoff (Payne et al. 1992). This gives a basis with the virtue of its systematic improvability (higher cutoff gives a better basis) but quite inefficient (many plane waves needed per electron) and made of delocalized functions. The linear-scaling method, on the other hand, is based on strictly localized, numerical atomic orbitals (NAOs, see below).

In the SIESTA program, the building of the Hamiltonian and overlap matrices is performed using a combination of two techniques (Ordejon et al. 1996; Sanchez-Portal et al. 1997). On one hand, the overlap matrix and some terms of H including the kinetic energy (the so-called two-centre integrals) allow a very efficient numerical integration in one single variable and as a function of just the distance between every two atoms. The remaining terms are calculated by replacing the three-dimensional integrals by summations over a finite grid, a discretization of 3-D space. One parameter controls the precision of these integrations, namely, the fineness of the grid, usually expressed as an energy cutoff, an index borrowed from PWs, the square root of which gives the inverse of that fineness.

The linear-scaling solution of the eigenvalue problem posed by the Kohn–Sham Hamiltonian requires abandoning the standard diagonalization procedure. In fact, for the ground-state properties sought in these calculations there is no need of knowing all the eigenvalues separately, it is just the sum of the occupied ones that is of relevance to the total energy, i.e. the trace of the Hamiltonian in occupied space, $Tr_{\mathrm{occ}}\{H\}$. Since this trace is invariant under any change of basis of occupied space, a localized basis in the line of Wannier functions can be sought instead. They are obtained by minimizing $Tr_{\mathrm{occ}}\{H\}$ under the constraint that the wave functions vary within localized regions of space. Finally, it is necessary to avoid the localized-wave-function orthogonalization (quadratic scaling), and this is accomplished by minimizing a slightly modified functional instead of $Tr_{\mathrm{occ}}\{H\}$, which shares the same minimum with it (Ordejon 1998).

The linear-scaling solution just described is obviously more efficient than the cube-scaling diagonalization for large systems. For small systems, however, the diagonalization is competitive and more convenient. The critical size separating these two regimes depends on the kind of system. Large-gap insulators present highly localized electrons and allow efficient linear scaling, thereby reducing the critical size. As a general rule of thumb for the SIESTA method for non-metallic systems, the critical sizes are to be found around 100 atoms. Many of the calculations presented below have been performed on smaller systems and are therefore solved with diagonalization. This is of no relevance for the purpose of this work of comparing the behaviour of the different basis sets. It is important to stress, however, that within SIESTA, the building of the Kohn–Sham Hamiltonian is always done in a linear-scaling fashion.

For given atomic positions, the self-consistent solution of the electronic problem gives the electronic total energy. Using the Hellmann–Feynman theorem derivatives of the energy are also obtained, most importantly the forces on the atoms and the stress tensor. This allows the relaxation of the system including atomic positions and/or lattice parameters, with the possible additional constraints of externally applied pressure, hydrostatic, uniaxial etc.

Both SIESTA and codes based on the plane-wave schemes, such as the CASTEP code used in this study, have been implemented for parallel computers. The results presented below were obtained using two such machines, namely a Cray T3E (up to 64 nodes in a single job) and an SGI Origin 2000 (up to 16 nodes in a single job).

Atomic-orbital basis sets

There are three main aspects that determine the NAO basis sets: (1) the number of orbitals, (2) their extent (cutoff radii) and (3) their radial shape. The shape is defined by solving the isolated atomic problem within DFT and for the same pseudopotential as used in the calculations. This procedure gives reasonable shapes well adapted to the pseudopotentials. To control their extent, the isolated atom is put into a spherically symmetric confining potential that forces the orbitals to become strictly zero beyond a certain cutoff radius. A reasonably balanced definition of all the cutoff radii for different orbitals and different atomic species is obtained by defining one single parameter, the energy shift ΔE. This is the raise in energy that the orbital suffers when confined (Artacho et al. 1999). In the studies presented in this paper, different values of the energy shifts were used for different systems.

The number of orbitals per atom in a basis set is defined by the number of angular-momentum channels and the number of different orbitals in a channel. Starting by the simplest, the minimal basis incorporates only those atomic-orbital shells needed for the free atom, with one single orbital per shell (also called single-zeta or SZ). For example, one s orbital and one p shell (one p_x, p_y and p_z) for C, O or Si, one s orbital and a d shell for Fe. The basis can then be doubled (DZ), introducing two different orbitals for an s channel, or tripled (TZ), using three different orbitals. It is possible to use mode complex basis sets, and some of these were investigated in this study.

To add angular flexibility to the basis, it is customary to include polarization orbitals, additional basis functions of higher angular momentum. The procedure to obtain the shape of the polarization orbitals as well as higher zeta ones is explained by Artacho et al. (1999). Typically, we included one polarization orbital with a double zeta basis, hence DZP. For TZP we used two polarization orbitals rather than just one.

In the cationic substitution problem (see below), the energy is calculated for the exchange of different cations between different positions. Since the orbitals are associated to the atoms, the exchange involves not only the atoms but also the basis, and therefore, the energy obtained has an undesired added term due to this, the so-called basis set superposition error (BSSE). It can be corrected by introducing the basis set of both cations simultaneously, so that both positions have the same orbitals independent of which atom sits on them. In addition to correcting the deficiency, the change in energy obtained by introducing BSSE will gauge the quality of the basis. In practice we found that the BSSE was not significant in our example problems.

Calculations using the localized orbital method of the SIESTA code were compared with those of conventional plane wave methods (mostly using the CASTEP code) for several mineral systems of current interest. In each case the main intention was to compare the predictions of the methods rather than to make an in-depth scientific study of the system. We also investigate the convergence of results with the size of the basis set used for the SIESTA calculations, since this is less straightforward than in plane-wave codes where we have a single parameter (the plane-wave cutoff) to determine the basis size.

In the following, results are presented for the comparison between both methods. The problems have been selected to be challenging and delicate in the comparison, with small energy differences and subtle rearrangements of atoms, representative of questions of interest to the mineral physics community.

Results and discussion

Selection of test systems

In this section we report a number of test calculations. As we noted in the introduction, these have been chosen to be representative of a wide range of problems that are currently being studied by various methods, and we are focusing on properties that are not easy to calculate. The range of examples includes cation ordering and solid solutions (spinel, garnets and amphiboles), mineral interactions with organic molecules (benzene in pyrophyllite) and pressure-induced displacive phase transitions (cristobalite). In each case we are focusing on the need to calculate small changes in energies, in addition to the calculations of basic structures. These are not easy to measure experimentally, and quantum-mechanics methods will have a significant role to play in providing information about energetics associated with transition processes. Our starting point, Mg/Al ordering in spinel, is useful because we have a benchmark against our own previously published studies performed with plane-wave methods (Warren et al. 2000a, b) – it was the computational demands of this earlier study that indicated than an alternative method for large systems is necessary. We have used this example to carry out a detailed test of the accuracies of different basis sets. The example of cation ordering in amphibole is a case where plane-wave methods could not have been performed with existing resources. This is also true for the example of cation ordering in the garnet solid solution, and in this case the quantum-mechanical calculations performed using siesta have been used to support the results of an earlier extensive study using empirical potentials (Bosenick et al. 2000). The study of the pressure-induced displacive phase transition in cristobalite was motivated by recent experimental results (Dove et al. 2000), and in this case we were able to compare the SIESTA results with plane wave ab initio calculations. The example of organic molecules interacting with pyrophyllite surfaces is an example of the type of study one would like to perform in order to tackle the issue of the binding of pollutant organic molecules to the surfaces of soil minerals (Craig and Dove 2002). In all cases bar one (one set of calculations on spinel) we always allow both the atomic positions and lattice parameters to relax in the calculations.

Cation ordering in spinel

Spinel, $MgAl_2O_4$, contains both tetrahedral and octahedral sites, normally occupied by Mg and Al cations respectively. The structure is shown in Fig. 1. On heating, there is progressive exchange of the Mg and Al sites, which has been studied recently using a combination of ab initio plane-wave methods and Monte Carlo simulations (Warren et al. 2000a, b). For this reason, it is a

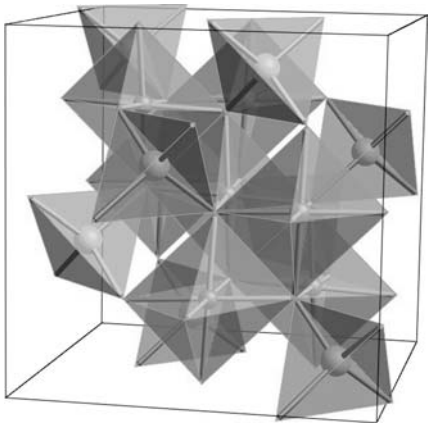

Fig. 1 Structure of spinel, showing the tetrahedral and octahedral cation sites

useful test case for calculations of cation-ordering energies with SIESTA. We have therefore performed a series of calculations comparing the energy of the ordered structure with that of an inverse structure in which all the tetrahedral sites are occupied by Al cations, and half the octahedral sites by Mg cations. A set of simulations with different localized basis sets, and with a fully converged plane-wave basis (using a modification of SIESTA rather than CASTEP in order to ensure consistent use of equivalent pseudopotentials) was performed, within the LDA. Only the Γ point was sampled in the electronic band structure, with the intention of systematically testing the basis set rather than producing a more accurate, but slower, calculation.

Geometry optimization simulations of the normal ordered structure using SIESTA gives an initial indication of the effect of basis sets – the results are presented in Table 1. These may be compared to the experimental volume of 133.0 $Å^3$. It is clear from the results that the polarized basis sets all give reasonable volumes, much better than those given by the basis sets without polarization. In this regard, it should be noted that SZP gives a better result than DZ, with both having the same number of orbitals.

To study these effects in more detail, single-point energies (i.e. with no structure relaxation) were calculated for both a normal and an ordered inverse structure with the different basis sets. The results are presented in Fig. 2. The absolute energy of a single structure will obey the variational principle and so can give insight into general factors affecting basis set efficiency. Firstly, we found that the number of different orbitals used had a much greater effect on the total energy than the energy shift, which determines the size of an individual orbital. Secondly, adding polarization orbitals to the cations, i.e. *d* on Al or *p* on Mg, was more effective in lowering the energy than doubling the number of atomic orbitals with the double-zeta method, i.e. SZP is a better choice than DZ. This effect was also found in the geometry optimization described above. Finally, when constructing a polarization orbital for Al, the atomic *d* orbitals were more efficient than polarization of the *p* orbitals.

From a DZP set, it was also found that doubling the polarization orbitals had a great effect on converging ΔE. Even if no *d* orbitals were used for Mg, increasing the maximum angular momentum component considered to $\ell = 3$, also improved the result. With an energy shift of 0.2 eV and a basis set of DZ with double polarization, a value for ΔE of 0.2 eV (20%) away from the plane-wave value was obtained. When using a TZP basis set, the value of ΔE was reduced to 0.08 eV (6%) away from the plane-wave value.

The sensitivity of ΔE to the choice of basis functions is due to the fundamental changes in bonding between the two structures. Mg and Al ions are being exchanged between tetrahedral and octahedral sites, so good representation of the bonding in both coordinations is needed for both ions. The addition of ghost cations to correct for BSSE, which allows both sets of orbitals to be present in both sites, did not dramatically improve the result.

This example highlights the differing sensitivity to a basis set of total energy, optimized geometry and small energy differences. The information required from such calculations must thus be considered when choosing basis sets.

High-pressure phase transition in cristobalite

These calculations were motivated by the recent solution of the high-pressure phase of cristobalite (Dove et al. 2000), whose crystal structure is shown in Fig. 3. The calculations were intended to test the relative stability of the new monoclinic structure compared to the low-temperature tetragonal structure. Calculations were performed at pressures of 0 and 2 GPa, with the experiments predicting the transition to be in this range. The simulation cell contained 24 atoms (i.e. two unit cells of the tetragonal structure), and a *k*-point sampling grid of $1 \times 2 \times 2$ was used. For the SIESTA calculations, convergence was obtained with a TZP basis set and an energy shift of 0.03 eV. These calculations were compared with plane-wave results calculated with

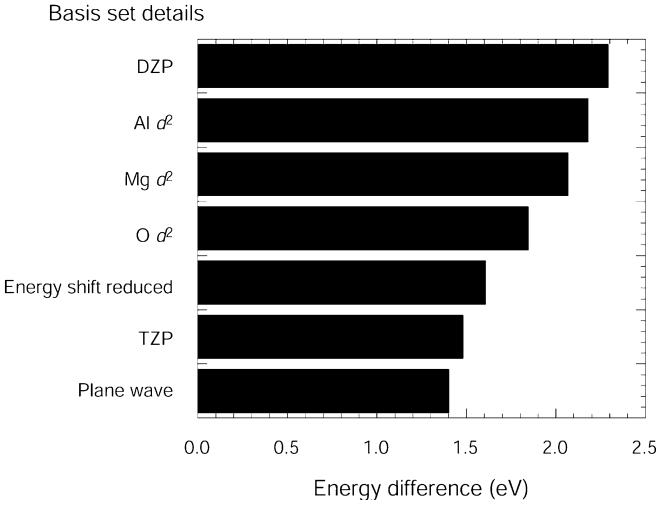

Fig. 2 Convergence of energy difference between normal and inverse phases of spinel, ΔE, with basis set and energy shift. A value from equivalent plane-wave calculations acts as the fully converged standard. The initial DZP basis set consists of s^2p^3 orbitals on Mg and $s^2p^6d^5$ on Al and O. The improvements added for each calculation are shown *against each point*, with d^2 representing a doubling of the polarization (*d*-orbital) set

Table 1 Optimized volumes of spinel calculated with different basis sets, compared with experiment. Volumes are in $Å^3$. The experimental (*Exp*) value is from Sawada (1995)

SZ	DZ	SZP	DZP	Exp
126.9	127.1	132.5	133.1	132.3

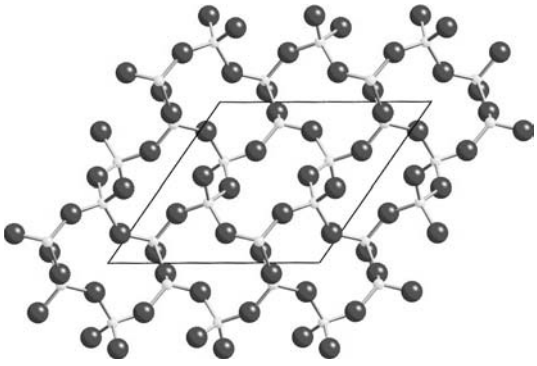

Fig. 3 Crystal structure of the high-pressure phases of cristobalite. (Dove et al. 2000)

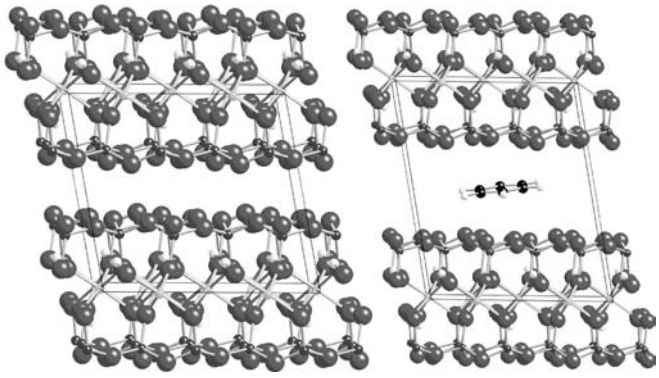

Fig. 5 Structure of pyrophyllite (*left*), and with intercalated benzene (*right*)

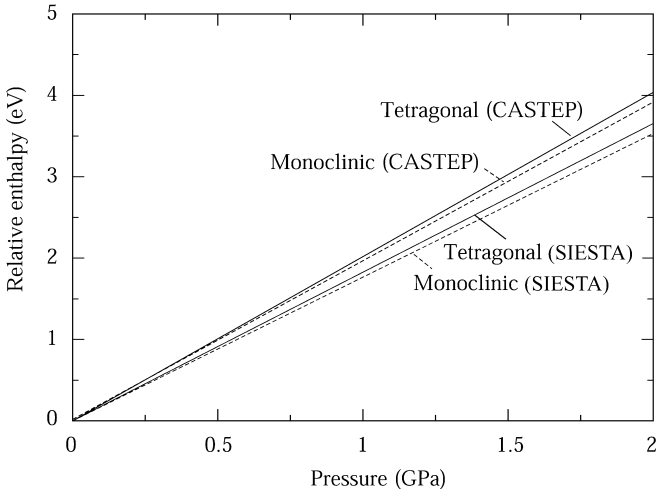

Fig. 4 Enthalpies of the two phases of cristobalite calculated at pressures of 0 GPa and 2 GPa using SIESTA and CASTEP (*lines* are drawn to connect the two points in each calculation)

CASTEP using ultrasoft pseudopotentials and a plane-wave cutoff of 380 eV.

Relative enthalpies at both pressures are shown in Fig. 4. This highlights the fact that both calculations predict the existence of the phase transition, although both results predict a transition pressure closer to 0 GPa than the experimental pressure (0.5–1.5 GPa, with considerable hysteresis). The enthalpy differences between the two phases are similar in both types of calculations.

The main differences are seen in the crystal structures. The lattice parameters and Si–O bond lengths are given in Table 2. The main difference is the *a* lattice parameter, which is the lattice parameter that is most sensitive to pressure, leading to relatively large differences in volumes. These can be accounted for by noting that the volume of the high-pressure phase of cristobalite is particularly sensitive to the degree of buckling of the network of SiO_4 tetrahedra, which is itself a low-energy process and hence easily changed. Differences in the two types of calculation appear to correspond to an effective pressure offset. This may be due to the use of different pseudopotentials, as plane-wave calculations using an alternative set also gave large differences in the *a* lattice parameter. The length of the Si–O bond is consistent across both structures and pressures. Again, the difference between plane waves and SIESTA can probably be attributed to pseudopotentials, as bond lengths can be very characteristic of these. The experimental bond length is 1.603 Å.

Crystal structure of pyrophyllite

Pyrophyllite, $AlSi_2O_5(OH)$, is an example of a simple layered clay mineral, which we have used as part of our ongoing project studying adsorption properties of such surfaces. Its structure is shown in Fig. 5. It is also known to be a difficult type of material to model using empirical potentials, as the aluminosilicate layers are held together by weak van der Waals forces. We compare the structure predicted by CASTEP and SIESTA. GGA was used for these calculations, as it is known to be superior to LDA for problems with surfaces where there is a large gradient in electron density. Two *k* points were used to sample the band structure (using a $2 \times 1 \times 1$ grid in reciprocal space), and a DZP basis with an energy shift 0.14 eV was used.

The calculated structures (Table 3) are similar to each other, and compare well with the experiment. The agreement includes lattice parameters, bond lengths and the interlayer spacing.

Benzene intercalated into pyrophyllite

Clay minerals such as pyrophyllite can have the ability to intercalate small molecules between the aluminosilicate layers, accompanied by an expansion of the interlayer region to accommodate the molecule. Pillared clays give an example of this behaviour. We tested the effect of inserting a benzene molecule in between the layers of

Table 2 Comparison of calculations of the two phases of cristobalite using both SIESTA and CASTEP. The structure of phase II is from Dove et al. (2001), and experimental data for lattice parameters are after Palmer and Finger (1994)

	SIESTA	CASTEP	Exp
Low cristobalite 0 Gpa			
a (Å)	4.790	4.990	4.972
c (Å)	6.579	6.960	6.922
Si–O (Å)	1.626	1.589	1.603
Vol (Å^3)	301.961	346.296	342.160
Cristobalite II (2 GPa)			
a (Å)	8.386	8.453	8.457
b (Å)	4.516	4.675	4.706
c (Å)	8.737	9.252	9.312
β (°)	124.638	125.387	125.19
Si–O (Å)	1.626	1.590	–
Vol (Å^3)	272.187	298.100	302

Table 3 Comparison of calculations of pyrophyllite structure. Experimental data (*Exp*) from Wardle and Brindley 1972

	SIESTA	CASTEP	Exp
a (Å)	5.17	5.12	5.16
b (Å)	9.00	8.91	8.96
c (Å)	9.25	9.52	9.35
α (°)	91.09	90.64	91.03
β (°)	100.81	100.56	100.37
γ (°)	89.89	89.67	89.75
Si–O (tet–tet) (Å)	1.65	1.59	1.63
Al–O (Å)	1.92	1.85	1.96
O–H (Å)	0.95	0.96	–
Interlayer spacing (Å)	2.30	2.80	2.50

pyrophyllite, with full cell relaxation to enable the layers to be pushed apart. The structure is shown in Fig. 5, and the resultant geometries predicted by SIESTA and CASTEP are shown in Table 4. Both show good agreement for the cell parameters and interlayer separation. These are about 3 Å larger than pure pyrophyllite. Calculation of an isolated benzene molecule also allows the calculation of the energy change in inserting a benzene molecule into the structure. This was calculated to be 0.74 eV for CASTEP and 0.36 eV for SIESTA, i.e. energetically unfavourable in both cases. These two results are in reasonably good agreement, given that the energy difference being measured is quite small. Moreover, in contrast to the comparison of energy differences in spinel, different pseudopotential sets were used in the two sets of calculations. This was necessary in principle because the CASTEP allows the use of ultrasoft pseudopotentials in order to improve the speed of the calculation, whereas ultrasoft pseudopotentials are not built into SIESTA.

Cation ordering in garnet solid solutions

This is another example of a study of cation ordering, in this case Ca–Mg ordering over equivalent sites in pyrope and grossular garnets. The purpose of this calculation was to confirm a striking result obtained using empirical calculations, which predicted an unusually significant third-nearest-neighbour interaction (Bosenick et al. 2000). The unit cell contained 160 atoms, which would make the calculation computationally challenging using conventional plane wave methods. The SIESTA calculations used a DZP basis and an energy shift of 0.27 eV.

Table 4 Comparison of the structures of pyrophyllite with intercalated benzene

	CASTEP	SIESTA
a (Å)	10.249	10.370
b (Å)	8.915	8.992
c (Å)	12.508	12.603
Layer separation (Å)	5.8 (1)	5.6 (1)

Table 5 Comparison of garnet structures calculated by SIESTA and obtained by experiment (*Exp*) (Armbruster et al. 1992 for pyrope; Ganguly et al. 1993 for grossular)

	Pyrope		Grossular	
	SIESTA	Exp	SIESTA	Exp
a (Å)	11.359	11.452	11.728	11.848
Si–O (Å)	1.656	1.634	1.694	1.646
Al–O (Å)	1.877	1.886	1.891	1.926
(Ca,Mg)–O (Å)	2.228	2.269	2.348	2.405

The experimental lattice parameters of pyrope and grossular are 11.456 and 11.846 Å, respectively. The corresponding values calculated by SIESTA were 11.359 and 11.728 Å, respectively. The calculated values are 1% lower in both cases. Other structural details are compared with experiment in Table 5.

We produced configurations of 2 Mg cations in a cubic cell of grossular and of 2 Ca cations in a cubic cell of pyrope. One set of configurations had the two dilute cations in nearest-neighbour sites (distance 3.4 Å) and the other set had the two cations in third-neighbour sites (distance 6.4 Å) – these two distances are shown in Fig. 6. For both pyrope and grossular we calculated the energy associated with replacing one nearest-neighbour pair with one third-neighbour pair as 0.13 eV. This is close to the value of 0.11 eV calculated using empirical pair potentials (Bosenick et al. 2000), and confirms the detailed analysis carried out with the pair potentials.

Cation ordering in the amphibole glaucophane

SIESTA was used to calculate the energies of ordering of Mg and Al cations over the octahedral sites in the amphibole glaucophane, $Na_2(Mg_3Al_2)Si_8O_{22}(OH)_2$. The structure is shown in Fig. 7. There are three distinct octahedral sites, labelled M1 M2 and M3 (see Fig. 7). Experimental evidence points to ordering of Al cations in M2 and Mg cations in M1 and M3. We have

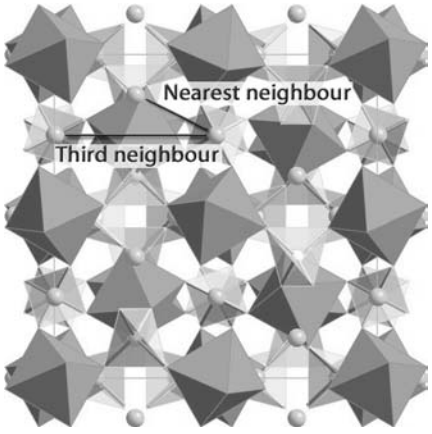

Fig. 6 Structure of garnet, showing octahedral and tetrahedral sites and the dodecahedral Mg/Ca cation sites. The *nearest-* and *third-neighbour* distances are indicated

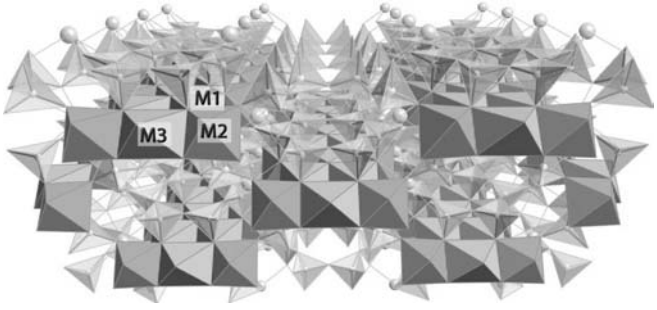

Fig. 7 Structure of glaucophane, showing octahedral and tetrahedral sites. The three octahedral sites involved in Mg/Al ordering are indicated (M_1, M_2, M_3)

performed calculations of the ordering energies using empirical pair potentials, obtaining energies of forming different neighbours (exchange interactions) and energies of cations on specific sites (chemical potential). The empirical models suggested a lowest-energy state with M1 and M2 sites having half occupancy of Al and Mg, in contrast to the experimental result. This state was calculated to have an energy lower than the experimentally ordered state by 0.08 eV per formula unit.

Nine configurations were analyzed using SIESTA. Each configuration contains 328 atoms, which is too large for plane-wave codes to calculate in a reasonable time. The experimentally ordered structure was found to be the lowest-energy in these calculations. The lowest-energy state from the empirical models was also found to be low compared with other configurations, but was higher by 0.06 eV for formula unit than the experimentally ordered structure. From these calculations it was possible to extract chemical potentials for Al cations on each of the three octahedral sites, taking into account the neighbouring exchange energies obtained from the empirical calculations. In fact, the exchange interactions favour ordering of Al on both M1 and M2 sites by 0.44 eV per formula unit, but the chemical potentials favour ordering of Al on M2 sites over M1 sites by 0.48 eV per site. There is thus a fine balance between the ordering preferred by the exchange interactions and the ordering preferred by the chemical potentials. SIESTA appears to have succeeded in properly reproducing this balance as judged by experiment.

Table 6 Comparison of glaucophane structure calculated by SIESTA and obtained by experiment (*Exp*) (Papike and Clark 1968)

	SIESTA	Exp
a (Å)	9.647	9.541
b (Å)	17.815	17.740
c (Å)	5.304	5.295
β (°)	105.8	103.6
Si–O (Å)	1.649	1.621
Al–O (Å)	1.935	1.944
Mg–O (Å)	2.025	2.092
Na–O (Å)	2.581	2.498
O–H (Å)	0.985	–

The structure of the ordered configuration is compared with experiment in Table 6, again showing a reasonable level of agreement.

Computational resources

In practice it is very difficult to make a simple comparison of the computational resources required to run a plane-wave code such as CASTEP with those of SIESTA. For example, SIESTA can be reasonably used at various levels of convergence from very fast minimal basis set calculations, to highly converged calculations, comparable with converged plane-wave calculations. Furthermore, the scaling of resources required with system size is, as described, different for plane-wave codes and those using localized basis functions, so plane waves are more efficient for smaller systems, localized basis functions for larger systems. Also, we have typically run these simulations on different machines, and using different numbers of processors. Finally, it is important to note that while CASTEP is a mature and highly optimized code, SIESTA is still rather new, and comparisons of computer time may not accurately reflect the relative efficiencies of the underlying methods.

Tests of scaling with system size confirmed the result that a plane-wave code such as CASTEP scales approximately with the cube of the system size (with the square if volume is added containing no new atoms). When used in diagonalization (non-order-N) mode, SIESTA scales approximately with the square of the system size. Addition of vacuum does not have any significant effect on the resources required, which is of great importance in modelling surfaces and molecules. Linear scaling can also be obtained, as described previously, though we have not used it in this work as it is less efficient for systems of the size of that we have considered.

A direct comparison between converged calculations can be attempted for the calculation of benzene intercalated in pyrophyllite. This was run with CASTEP and SIESTA on a Cray T3E using eight nodes. The CPU time per minimization step was about three times larger for CASTEP than SIESTA (35741s to 10218s). This system (92 atoms) is therefore at a point where a converged calculation using the two methods is currently of comparable speed. Memory requirements were also similar for these calculations (139 Mb node^{-1} for CASTEP to 150 Mb node^{-1} for SIESTA), and are expected to scale similarly to time.

Conclusions

In this paper we have reported a number of tests of the SIESTA approach to ab initio simulations of minerals. These tests have been chosen to be reasonably challenging, particularly with respect to the calculation of small energy changes associated with displacive phase transitions and cation-ordering processes. Where practical, the results have been compared with plane-wave calculations, but for the larger systems only the SIESTA approach can give results in a reasonable time scale.

As far as crystal structures are concerned, SIESTA is reasonably accurate, and compares well with plane-wave calculations. As a rough guide, our results show that DZP basis sets are capable of giving good results for structures. In the examples, SIESTA also performs reasonably well in calculation of energy differences in the cases where we can compare with plane-wave results. The example of spinel has given an idea of the need to be careful with choice of basis sets. DZP basis sets give acceptable results for energy differences, and TZP gives good energy differences. The example of cristobalite has shown that SIESTA is able to capture small energy differences associated with displacive phase transitions, where there is a balance between changes in bonding energy and volume energy at high pressure. In cases where we have not been able to compare with plane-wave calculations, we have been encouraged by agreement with experimental data (glaucophane) or with empirical calculations (garnet). However, we have also performed a study of Al–Si ordering in a pyroxene, and do not find the same level of agreement between SIESTA, plane waves and empirical potential calculations, for reasons that are not clear to us.

We have shown that SIESTA is able to be used to tackle computationally demanding systems that are currently out of the reach of plane-wave codes. Often it is useful to perform calculations for a range of conditions, such as pressure, composition or degree of order. Such parametric studies require relatively rapid turnaround of results, and the examples presented in this paper have shown that SIESTA is able to produce results of suitable quality in a reasonable time scale for mineralogical research. We will present results for simulations of cations and molecules on mineral surfaces obtained using SIESTA in later publications.

Acknowledgements We are pleased to acknowledge financial support from NERC and EPSRC. Calculations were performed on the computers of the CSAR Service (Manchester, UK) and of the Cambridge High Performance Computer Facility (Cambridge, UK).

References

Armbruster T, Geiger CA, Lager GA (1992) Single-crystal X-ray structure study of synthetic pyrope almandine garnets at 100 and 293 K. Am Mineral 77: 512–521

Artacho E, Sanchez-Portal D, Ordejon P, Garcia A, Soler JM (1999) Linear-scaling ab-initio calculations for large and complex systems. Phys Stat Sol (B) 215: 809–817

Bachelet GB, Schluter M (1982) Relativistic norm-conserving pseudopotentials. Phys Rev (B) 25: 2103–2108

Bosenick A, Dove MT, Geiger CA (2000) Simulation studies of pyrope–grossular solid solutions. Phys Chem Miner 27: 398–418

Bosenick A, Dove MT, Myers ER, Palin EJ, Sainz-Diaz CI, Guiton B, Warren MC, Craig MS, Redfern SAT (2001) Computational methods for the study of energies of cation distributions: applications to cation-ordering phase transitions and solid solutions. Mineral Mag 65: 193–219

Ceperley DM, Alder BJ (1980) Ground state of the electron gas by a stochastic method. Phys Rev Lett 45: 566–569

Chartier A, D'Arco P, Dovesi R, Saunders VR (1999) Ab initio Hartree-Fock investigation of the structural, electronic, and magnetic properties of Mn_3O_4. Phys Rev (B) 60: 14042–14048

Craig MS, Dove MT (2002) Ab initio simulations of adsorption of organic molecules on pyrophyllite surfaces. Surface Science (in press)

Dove MT, Craig MS, Keen DA, Marshall WG, Redfern SAT, Trachenko KO, Tucker MG (2000) Crystal structure of the high-pressure monoclinic phase-II of cristobalite, SiO_2. Mineral Mag 64: 569–576

Ganguly J, Cheng W, O'Neill HStC (1993) Syntheses, volume, and structural changes of garnets in the pyrope-grossular join: implications for stability and mixing properties. Am Mineral 78: 583–593

Goedecker S (1999) Linear scaling electronic structure methods. Rev Mod Phys 71: 1085–1123

Hohenberg P, Kohn W (1964) Inhomogeneous electron gas. Phys Rev (B) 136: 868–871

Junquera J, Paz O, Sanchez-Portal D, Artacho E (2001) Numerical atomic orbitals for linear-scaling calculations. Phys Rev B64: Art. No. 235111

Kenny SD, Horsfield AP, Fujitani H (2000) Transferable atomic-type orbital basis sets for solids. Phys Rev (B) 62: 4899–4905

Kohn W (1996) Density functional and density matrix method scaling linearly with the number of atoms. Phys Rev Lett 76: 3168–3171

Kohn W, Sham LJ (1965) Self-consistent equations including exchange and correlation effects. Phys Rev 136: A1133–A1138

Louie SG, Froyen S, Cohen ML (1982) Nonlinear ionic pseudopotentials in spin-density-functional calculations. Phys Rev (B) 26: 1738–1742

Ordejon P (1998) Order-N tight-binding methods for electronic structure and molecular dynamics. Comp Mater Sci 12: 157–191

Ordejon P, Artacho E, Soler JM (1996) Self-consistent Order-N density-functional calculations for very large systems. Phys Rev (B) 53: R10441–R10444

Palmer DC, Finger LW (1994) Pressure-induced phase transition in cristobalite: an X-ray powder diffraction study to 4.4 GPa. Am Mineral 79: 1–8

Papike JJ, Clark JR (1968) The crystal structure and cation distribution of glaucophane. Am Mineral 53: 1156–1173

Payne M, Teter M, Allan D, Arias T, Joannopoulos JD (1992) Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. Rev Mod Phys 64: 1045–1097

Perdew JP, Zunger A (1981) Self-interaction correction to density-functional approximations for many-electron systems. Phys Rev (B) 23: 5048–5079

Perdew JP, Burke K, Ernzerhof M (1996) Generalized gradient approximation made simple. Phys Rev Lett 77: 3865–3868

Sanchez-Portal D, Ordejon P, Artacho E, Soler JM (1997) Density-functional method for very large systems with LCAO basis sets. Int J Quantum Chem 65: 453–461

Sankey OF, Niklewski DJ (1989) Ab initio multicenter tight-binding model for molecular-dynamics simulations and other applications in covalent systems. Phys Rev (B) 40: 3979–3995

Sawada H (1995) An electron density residual study of magnesium aluminium oxide spinel. Mater Res Bull 30: 341–345

Troullier N, Martins JL (1991) Efficient pseudopotentials for plane-wave calculations. Phys Rev (B) 43: 1993–2006

Vanderbilt D (1990) Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. Phys Rev (B) 41: 7892–7895

Wardle R, Brindley GW (1972) The crystal structures of pyrophyllite, 1Tc, and of its dehydroxylate. Am Mineral 57: 732–750

Warren MC, Dove MT, Redfern SAT (2000a) Ab initio simulations of cation ordering in oxides: application to spinel. J Phys: Condens Matt 12: L43–L48

Warren MC, Dove MT, Redfern SAT (2000b) Disordering of $MgAl_2O_4$ spinel from first principles. Mineral Mag 64: 311–317

Warren MC, Dove MT, Myers ER, Bosenick A, Palin EJ, Sainz-Diaz CI, Guiton B, Redfern SAT (2001) Monte Carlo methods for the study of cation ordering in minerals. Mineral Mag 65: 221–248

White C, Johnson B, Gill P, Head-Gordon M (1994) The continuous fast-multipole method. Chem Phys Lett 230: 8–16