A microscopic model for a very stable incommensurate modulated mineral: mullite

1993 J. Phys.: Condens. Matter 5 3417

(http://iopscience.iop.org/0953-8984/5/21/004)

J. Phys.: Condens. Matter 5 (1993) 3417–3430. Printed in the UK

A microscopic model for a very stable incommensurate modulated mineral: mullite

S Padlewski†, V Heine† and G D Price‡

† Cavendish Laboratory, Madingley Road, Cambridge CB3 OHE, UK

‡ Department of Geological Sciences, University College of London, Gower Street, London WC1E 6BT, UK

Received 19 February 1992

Abstract. A microscopic model is used to discuss the long-range ordering behaviour of the oxygen vacancies in mullite. This aluminosilicate of composition $Al_2(Al_{2+2x}Si_{2-2x})O_{10-x}(vacancy)_x$ presents an extremely stable incommensurate modulated structure with a wave vector $Q_{Ic}(x)$ which depends on the composition index x. Earlier computer simulations showed that the oxygen vacancies strongly order the Al/Si tetrahedral sites in their local environment. When two vacancies approach too closely, their dressings overlap which results in a repulsive interaction. The system is mapped onto an ANNNI model with frustration and an additional constraint on composition x. The solutions of our model at thermal equilibrium are discussed in q-space as well as in real space in order to calculate $Q_{Ic}(x)$. Reasonable agreement is found with observations for $x < 0.5$. At high concentrations of vacancies it is necessary to renormalize the coupling terms, which then accounts for the doubly modulated structure that occurs in mullite at $x > 0.5$.

1. Introduction

Mullite is a naturally occurring mineral of composition $Al_2(Al_{2+2x}Si_{2-2x})O_{10-x}$ where the composition index x ranges between approximately 0.15 and 0.60 (Cameron 1977). Mullite is structurally closely related to the mineral sillimanite ($x = 0$) and differs essentially by the presence of oxygen vacancies (figure 1) which are formally introduced into the structure via the following substitutional exchange reaction:

$$2Si^{4+} + O^{2-} \longrightarrow 2Al^{3+} + \text{vacancy}. \tag{1.1}$$

Mullite exhibits an incommensurate modulated structure which is extremely stable and persists until the melting point. In this paper we principally discuss the origin of the modulated phase from a microscopic point of view, via a statistical mechanics approach.

In a recent study we have shown with the help of atomistic computer simulations (Padlewski *et al* 1992a, henceforth referred to as I) that a very stable arrangement of Al/Si is found about the O vacancies in mullite (figure 2). Such an Al/Si distribution around the vacancy is extremely strong, being stabilized with respect to disorder by about 1 eV. We called this ordering pattern the dressing of the oxygen vacancy. If two vacancies approach each other too closely, their dressings overlap, and the resultant interaction turns out to be repulsive (Padlewski *et al* 1992b, henceforth referred to II). From our earlier computer modelling four results stand out. Firstly it is the vacancy distribution $\{P_i\}$ that controls the

0953-8984/93/213417+14$07.50

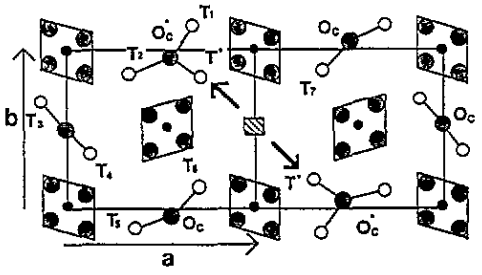

Figure 1. Structure around an O vacancy placed in the centre (shaded). The two adjacent tetrahedrally coordinated T atoms migrate to T* sites near the O_c^* atoms which are then in the middle of three T atoms forming a 3-cluster.

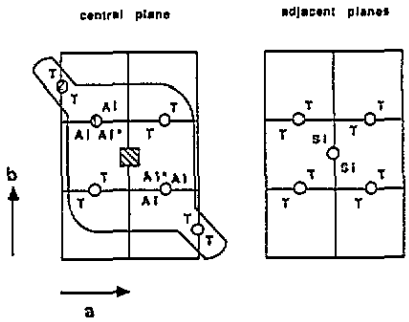

Figure 2. Al/Si ordering around a single O vacancy (shaded). The two Al* atoms would have been linked if the central O_c site had been occupied by an O atom instead of the vacancy.

ordering pattern. Note that in what follows, we use instead of P_i the Ising variables $\{\sigma_i\}$ defined by

$$P_i = \tfrac{1}{2}(1 + \sigma_i) \tag{1.2}$$

where one spin variable is defined on each O_c site i (figure 1). It represents a vacancy for $\sigma_i = +1$ or an O for $\sigma_i = -1$ respectively. Following our earlier formation in II, we can introduce an Ising model for representing the energy of mullite as follows:

$$E(\{\sigma_i\}) = E_0(x) - \frac{1}{2}\frac{2}{N}\sum_i \sum_j \frac{C_{ij}}{4}\sigma_i\sigma_j \tag{1.3}$$

where $E(\{\sigma_i\})$ is the energy per (abc) unit cell with respect to the distribution of vacancies $\{\sigma_i\}$. The cell of size (abc) is defined to be the unit cell of the disordered structure containing one formula unit of mullite as given above. The term $E_0(x)$ has no importance in the context of the present study. We will consider it in detail in a further work (Padlewski *et al* 1993) in order to discuss the stability field of mullite. Note that the composition variation is controlled by the exchange reaction (1.1) which fixes the concentration of vacancies. In particular the set $\{\sigma_i\}$ is required to fulfill the constraint

$$\frac{1}{N}\sum_i \sigma_i = \bar{\sigma} = x - 1 \tag{1.4}$$

where x is the composition index. In equation (1.3) the sum runs over the N O_c sites of the crystal ($N/2$ cells with two O_c sites per unit cell) and one needs to introduce the factor $2/N$ for normalization.

The second main point arising from our earlier work (I and II) is that the vacancies interact via the way they organize/order the Al/Si atoms around themselves, i.e. the 'dressing'. Most important in this are the T* atoms (figure 1) which are released by the vacancy and have to attach themselves to an O atom at an adjacent O_c called O_c^*.

Thirdly we considered the coupling terms defined in figure 3 and determined them with help of computer modelling (II). Some of the C_{ij} are strongly repulsive (see table 1) and they effectively determine the vacancy ordering pattern, i.e. the incommensurate structure. They are C_{110}, C_{001}, C_{220} and to a lesser extent C_{111}. In the ab-plane the strong repulsion

Table 1. Numerical value in eV obtained for the coupling between the vacancies. The four first terms result from the overlapping effect and are all repulsive as expected.

C_{110}	C_{220}	C_{001}	C_{111}	C_{200}	C_{400}
−1.02	−0.36	−0.72	−0.17	+0.20	+0.04
C_{020}	C_{040}	C_{310}	C_{002}	C_{130}	
+0.22	+0.05	−0.02	+0.19	+0.08	

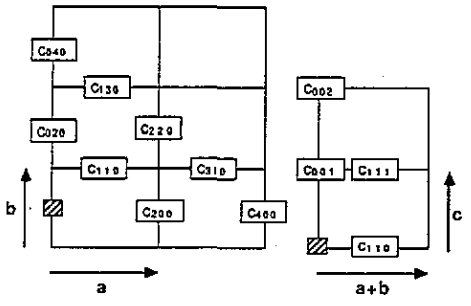

Figure 3. Field of interactions around a central vacancy (shaded) considered in our modelling. The interactions denoted by bold characters result from the overlapping effect.

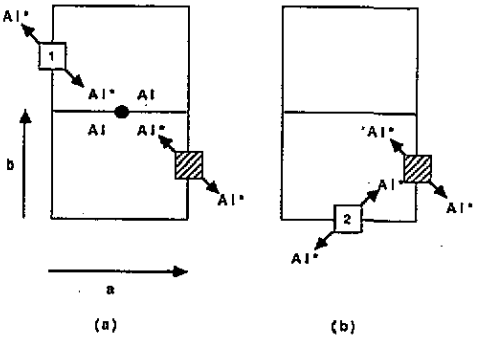

Figure 4. Energetically highly unsatisfactory situations arising when two vacancies approach too closely. In (*a*) a cluster with four ions is formed around one O. This situation is unfavourable because of the excess of charge. In (*b*) one of the four Al* atoms has no O_c^* atom to attach itself to.

arises because when the vacancies approach very closely; one vacancy interferes with the T* atom of the other vacancy and its attachment to its O_c^* atom. Two such situations are illustrated in figure 4. In one case the O_c^* atom is not there because that is the position of the other vacancy (figure 4(*b*)). In the other case the O_c^* atom already has an Al* atom from the other vacancy, and cannot accept a second Al* (figure 4(*a*)).

Finally our fourth point arising from our earlier work (II) is that we gave a simple heuristic demonstration that the C_{ij} of table 1 might be expected to lead to an incommensurate (IC) structure. However there were two defects which will be addressed in the present work:

(a) We show that by taking a modulated cosine distribution of σ_i the energy (1.3) turned out to be a minimum for an IC wave vector q_{Ic}. Such an approach does not mention temperature, entropy or free energy: it is purely heuristic. However, as is discussed in section 3, this analysis is valid at the transition temperature.

(b) This simple approach does not give any variation of $q_{Ic}(x)$ with the x which is observed in mullite (Cameron 1977) and which relates to the behaviour of the model at lower temperature (section 3 and 4).

In this paper, we discuss at thermal equilibrium the solutions of the energy model (1.3) constrained by (1.4). This is different from the usual Ising model because of the additional constraint (1.4). Two distinct methods can be used in order to address this problem:

(a) one can use the Lagrange multiplicator technique. In this one switches on a virtual field which is coupled to the average 'magnetization', i.e. the vacancy concentration. At thermal equilibrium, the field needs to be tuned in order to consider only the solution with the appropriate composition, i.e. magnetization. Note that formal studies on the Ising model

with competition, as in our model, show a sophisticated response to an applied field. In particular, it is known that the devil's staircase picture can also be produced as a function of applied field (Yokoi *et al* 1981, Yeomans 1988). This approach is however quite tedious because of the additional input parameter (external field).

(b) We can also search for solutions, at thermal equilibrium, using trial forms which already fulfill the constraint (1.4) as long as these forms are general enough. This approach removes the difficulty caused by the constraint without introducing any other parameter.

For the following we shall adopt the latter approach.

In an extensive experimental investigation of mullite, Cameron (1977) reported the dependence of the wave vector on composition. He found in particular that the wavelength of modulation, which occurs along the a-direction, increases with the composition index x.

In section 2 we derive from the energetic model (1.3) the expression for the free energy within a mean-field framework. We determine the transition temperature $T_c(x)$ (figure 5) as well as the wave vector of modulation q_{Ic} at T_c which is fixed by the maximum of the Fourier transform of the coupling terms C_{ij} (figure 6). The results are the same as in our earlier work II which we now see to be applicable at T_c. As already remarked above, this wave vector does not vary with composition so that we need to consider lower temperatures.

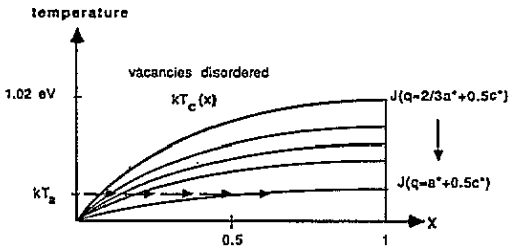

Figure 5. Plots of critical temperature in the mean-field theory versus the composition index x. A schematic contour plot is given of T_c for various assumed modulation wave vectors. The wave vector of modulation is studied on the isotherm $T_a = 1500\,°C$ which is assumed to correspond to the annealing temperature on formation.

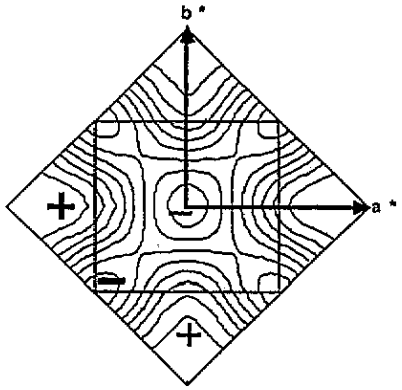

Figure 6. Contour plot of $C(q)$, the Fourier transform of the coupling term, in the (a^*, b^*) plane of reciprocal space at $Q = 0.5c^*$. The figure presents the first and second Brillouin zones because there are two O_c sites per unit cell. The maximum occurs at $q_{Ic} \simeq 0.5c^* + 0.69a^*$ which corresponds to the wave vector of modulation at T_c.

At this level another question needs to be considered: what $Q_{Ic}(T)$ would we expect the samples, measured at room temperature, to correspond to? We do not expect the vacancies to be very mobile except at high temperature because of the large energy $\simeq 1$ eV (I) involved in the 'dressing' because vacancy motion requires the movement of four T^* atoms (two before, two others after) as well as one O_c atom. Indeed from the study of synthesized mullites we know that there is little change of the structure below about the melting point 1850 °C or lower formation temperature by solid state transformation (McConnell 1991). We therefore assume that the observed structures correspond to equilibrium at some effective annealing temperature T_a at which the vacancies are still just mobile enough. There seems

no reason to suppose that T_a depends much on composition x and we therefore choose T_a in our calculations to be constant for all x ($T_a = 1500\,^{\circ}\mathrm{C}$ in our study).

In section 3, we employ a reciprocal space analysis with no lock-on effects to investigate the solution at T_a but with $T_a/T_c(x)$ not too small, applicable to small x. In this region the modulated structure requires higher harmonics (squaring up) for an improved description. The asymmetry which results from the constrained magnetization (1.4) allows in principle the presence of even harmonics in the Fourier analysis, so we include them in a first approach. We also consider the situation with no even harmonics. In both cases the wavelength of modulation increases with composition in agreement with observation.

In section 4, we use a real space analysis with lock-on effects to investigate the solutions at T_a but with smaller $T_a/T_c(x)$, applicable at higher composition. The findings of section 3 enable us to consider an effective one-dimensional system which is very similar to an ANNNI model (Yeomans 1988, Selke 1988). The solution has an important degree of squaring up in this limit. We find that the wavelength of modulation versus composition behaves according to a devil's staircase picture (Smith and Yeomans 1983, Selke 1988, Yeomans 1988). We also discuss in section 4 why we expect that the devil's staircase can hardly be observed.

Both approaches mentioned above (reciprocal and real space analysis) are in reasonable agreement with observation at low concentration. However from the experimental work of Nakajima and Ribbe (1981), it appears that for $x > 0.5$ the wave vector of modulation presents a doubly modulated structure along the directions $\boldsymbol{a}$ and $\boldsymbol{c}$. This so-called twinning (in the reciprocal space) does not occur with our energy model (1.3) where the modulated structure remains along the direction $\boldsymbol{a}$, even at very high composition. This inconsistency is discussed in section 5. We actually find that beyond $x = 0.5$, the coupling terms need to be renormalized because of the presence of a vacancy 'intruder' beside the central vacancy. The corrections are estimated numerically with the use of computer simulation similar to those in II. We find that the vacancy intruder modifies strongly the coupling terms along the direction $\boldsymbol{c}$ which then give the double modulation. Finally, the conclusions are summarized in section 6.

2. The transition temperature and the modulated structure

In this section, we first write down in a mean field framework the free energy expression for the Ising model (1.3). This is used to determine the transition temperature $T_c(x)$ which turns out to depend strongly on composition. We also determine the wave vector of modulation $q_{Ic}(T_c)$ at the transition temperature.

The Bragg–Williams approximation is known to overestimate T_c by a factor of about 1.5 but the phase diagram of the ANNNI model and the variation of the wave vector with respect to temperature is given quite well (Bak 1982, Selke 1988, Yeomans 1988). In this approximation the free energy per (abc) unit cell is given by

$$F_{MF}(T) = E_0(x) - \frac{1}{2}\frac{2}{N}\sum_i \sum_j \frac{C_{ij}}{4}\langle\sigma_i\rangle\langle\sigma_j\rangle + kT\frac{2}{N}\sum_i \left[\frac{1+\langle\sigma_i\rangle}{2}\log\left(\frac{1+\langle\sigma_i\rangle}{2}\right) + \frac{1-\langle\sigma_i\rangle}{2}\ln\left(\frac{1-\langle\sigma_i\rangle}{2}\right)\right]. \quad (2.1)$$

This F_{MF} is still in the form of a Landau free energy functional and to obtain the structure at thermal equilibrium we need to minimize the above expression with respect to the set

$\{\langle\sigma_i\rangle\}$. The solution $\{\langle\sigma_i\rangle\}$ also needs to fulfill the constraint (1.4). For the following, we find it convenient to expand the 'magnetization' per site around the average which is fixed by (1.4):

$$\langle\sigma_i\rangle = \bar{\sigma} + \Delta_i \tag{2.2}$$

where the terms Δ_i represent the deviation of the vacancy concentration from the mean, and hence the ordering pattern. Note that the variable Δ_i is confined between $-x$ (no vacancy) and $2 - x$ (occupied by a vacancy) where x is the composition index. The equation (1.4) becomes the constraint

$$\sum_i \Delta_i = 0. \tag{2.3}$$

Just below the transition temperature, the amplitude of the modulated structure around the disorder is small ($\Delta_i \ll 1$) and hence we can insert the form (2.2) into (2.1) and expand the entropic term for the free energy around the average magnetization. One can reasonably consider the lowest-order terms (up to the fourth order in our analysis). The mean-field free energy then takes the general form

$$F_{\mathrm{MF}}(T) = \text{constant} - \frac{1}{2}\frac{2}{N}\sum_{ij}\frac{C_{ij}}{4}\Delta_i\Delta_j + \frac{2}{N}\sum_i\sum_{k\geqslant 2} C_k(\bar{\sigma}, T)\Delta_i^k \tag{2.4a}$$

where the first three terms $C_k(\sigma, T)$ are given by

$$\begin{aligned} C_2(\bar{\sigma}, T) &= \tfrac{1}{2}kT/(1-\bar{\sigma}^2) \\ C_3(\bar{\sigma}, T) &= \tfrac{1}{3}kT\bar{\sigma}/(1-\bar{\sigma}^2)^2 \\ C_4(\bar{\sigma}, T) &= \tfrac{1}{4}kT(1+3\bar{\sigma}^2)/3(1-\bar{\sigma}^2)^3. \end{aligned} \tag{2.4b}$$

Just below the transition temperature, the modulated structure is described by a pure sinusoidal wave:

$$\Delta_i = 2A\cos(\boldsymbol{q}\cdot\boldsymbol{r}_i + \epsilon) \tag{2.5}$$

where ϵ is a phase, A the amplitude of the modulation and $\boldsymbol{r}_i$ the position of an O_c site i. By inserting (2.5) into (2.4a), one obtains an expression of the form

$$F_{\mathrm{MF}}(T) = \text{constant} + 2[2C_2(\bar{\sigma}, T) - C(q)/4]A^2 + \text{higher order} \tag{2.6a}$$

with

$$C(q) = \sum_j C_{ij}\cos[\boldsymbol{q}\cdot(\boldsymbol{r}_i - \boldsymbol{r}_j)]. \tag{2.6b}$$

Thus as T is lowered, the wave vector actually appearing at the temperature transition will be the $\boldsymbol{q}_{\mathrm{Ic}}$ which maximizes $C(\boldsymbol{q})$. In fact we find that the maximum of $C(\boldsymbol{q})$ occurs within the $(\boldsymbol{a}^*, \boldsymbol{b}^*)$ plane of the reciprocal space lying at height $Q = 0.5\boldsymbol{c}^*$ (figure 6).

$$\boldsymbol{q}_{\mathrm{Ic}} \simeq 0.5\boldsymbol{c}^* + 0.69\boldsymbol{a}^* \tag{2.7a}$$

and

$$C(q_{\mathrm{Ic}}) \simeq 4.08 \text{ eV}. \tag{2.7b}$$

As q_{Ic} lies outside the Brillouin zone, the observed wave vector is

$$Q_{\mathrm{Ic}} = a^* - q_{\mathrm{Ic}} = 0.5c^* + q_a a^* \tag{2.8a}$$

with

$$q_a \simeq 0.31. \tag{2.8b}$$

Note that in figure 6 we included the second Brillouin zone for the reason mentioned above. The form (2.8*a*) is in agreement with observations (Angel and Prewitt 1986, 1987, Cameron 1977) and this result has already been discussed at length in II. The transition temperature can be deduced from (2.6*a*):

$$2C_2(\sigma, T_c) - C(q)/4 = 0$$

i.e.

$$kT_c = (2x - x^2)C(q_{\mathrm{Ic}})/4. \tag{2.9}$$

This has the usual form of solution theory, proportional to a $c(1-c)$ where c is the concentration $x/2$ of vacancies. Note that the transition temperature increases with the composition x (figure 5) so that if we fix the temperature T_a, the effective temperature $T_a/T_c(x)$ felt by the sample decreases as we increase x. Note also that over the whole range of composition where mullite exists, $0.17 < x < 0.59$, the transition temperature remains well above the melting point $T_m \simeq 1850\,^{\circ}\mathrm{C}$ (Holland and Carpenter 1986), even if we multiply by a factor $\simeq 0.67$ to correct for the mean-field approximation, so that the disordered structure is never observed. This result is also in agreement with observations (Cameron 1977).

3. High-temperature approximation in reciprocal space

So far so good, but it is also well known that the wave vector of mullite is composition dependent (Cameron 1977). At $T_c(x)$ the critical wave vector is determined by the maximization of $C(q)$ and so does not depend on the composition x, i.e. on average 'magnetization'. This inconsistency can be resolved if we consider the squaring up of the modulation at lower temperature. This applies in particular if one increases the composition x (figure 5) which reduces the ratio $T_a/T_c(x)$ as already remarked.

So as we cool down the system below T_c, the modulated structure is not perfectly described by a pure sinusoidal wave but by a rather more complicated shape which requires higher harmonics. In general, we can represent the modulation by a Fourier series:

$$\Delta_j = \sum_{n \neq 0} A(n) \exp[-\mathrm{i}(nq \cdot r_j + \epsilon(n))] \tag{3.1}$$

where q is the wave vector, the term $A(n)$ is the amplitude of the nth harmonic and $\epsilon(n)$ its phase. They need to fulfil the property

$$\epsilon(-n) = -\epsilon(n) \tag{3.2a}$$

and

$$A(-n) = A(n) \tag{3.2b}$$

in order to produce a real modulation. We do not include the term $n = 0$ in (3.1) because we have already fixed the average magnetization in (2.2). Another important remark is in order: the expansion of the free energy (2.4*a*) contains odd-order terms so that we formally need to consider even harmonics in (3.1). By inserting (3.1) into (2.4*a*) which is taken up to $k = 4$, one obtains the following form:

$$F_{\mathrm{MF}}(T) = \text{constant} + \frac{2}{N}\sum_{n>0}\left(2C_2(\bar{\sigma}, T) - \frac{C(nq)}{4}\right)A(n)^2$$
$$+ C_3(\bar{\sigma}, T)\frac{2}{N}\sum_{n_1+n_2+n_3=0}\left(\prod_{i=1}^{3}A(n_i)\right)\cos\left(\sum_{i=1}^{3}\epsilon(n_i)\right)$$
$$+ C_4(\bar{\sigma}, T)\frac{2}{N}\sum_{n_1+\ldots+n_4=0}\left(\prod_{i=1}^{4}A(n_i)\right)\cos\left(\sum_{i=1}^{4}\epsilon(n_i)\right) \tag{3.3}$$

where we have purposely excluded lock-on terms (Heine 1989). In the above expression, the quartic entropic term causes the harmonics to interact with each other. The strength of the coupling depends on the composition and also on the temperature through $C_3(\sigma, T)$ and $C_4(\sigma, T)$ as seen in (2.4*b*). By minimizing the free energy (3.3) with respect to the $\{A(n), \epsilon(n)\}$ and q we find that the wave vector $q_{\mathrm{lc}}(x)$ depends on composition through the σ in (2.4*b*). For instance, we have considered the first three harmonics $n = 1$, 2 and 3 in the free energy form (3.3). As we only consider an incommensurately modulated structure, the free energy is left invariant by a global shift of the modulation so that we can set $\epsilon(1) = 0$ (Parlinski 1985). The cubic term (3.3) couples for example the first three harmonics via the term

$$12C_3(\sigma, T)A(1)A(2)A(3)\cos[\epsilon(2) - \epsilon(3)]. \tag{3.4a}$$

The quartic term couples them via the term

$$24C_4(\sigma, T)A(1)A(2)^2A(3)\cos[2\epsilon(2) - \epsilon(3)] \tag{3.4b}$$

and so on.

The modulated structure is found by minimizing the free energy (3.3) with respect to eight independent variables, i.e. $F_{MF}(q, \{A(n)\}, \epsilon(n)\}, n = 1, 2, 3)$. This can be performed with a standard Newton–Raphson procedure. As discussed in section 1, we assume that the experimental samples correspond to equilibrium at some annealing temperature $T_a = 1500\,^{\circ}\mathrm{C}$ for all x, below which temperature the vacancies are supposed to be effectively immobile.

In figure 7 we compare the resulting calculated wavelength of modulation at T_a with the experimental data of Cameron (1977). The wave vector q_{lc} continues to have the form (2.8*a*) as expected. We also carried out the same calculation but with only the odd harmonics included, $n = 1$, 3 and 5 which give the dashed line in figure 7. In both cases the

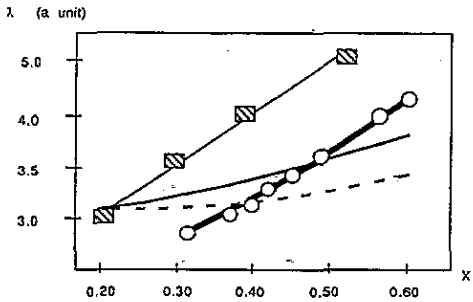

Figure 7. Comparison of wavelength between experiment (circles) and the findings of our model for various x at $T_a = 1500\,°C$. Full curve, the wavelength obtained by considering the three first harmonics $n = 1$, 2 and 3 in the modulation; broken curve, considering only the odd harmonics $n = 1$, 3 and 5; squares, commensurate wavelengths from the real space treatment. In all cases, the wavelength increases with composition in agreement with experiment.

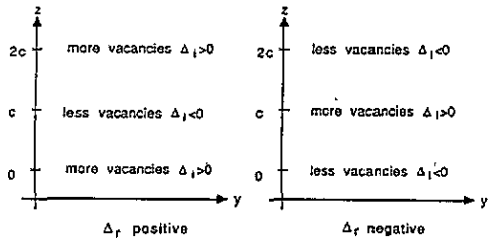

Figure 8. Two types of bc atomic layer with partial ordering of the vacancies. These layers interact with each other with repulsive coupling terms so that a modulated structure emerges.

wavelength is of about the right magnitude and increases with composition, in qualitative agreement with experiment.

An important remark is in order: the even harmonics when Q_{IC} has the form (2.8), i.e. with modulation along the a-direction, would give some segregation of the vacancies with different average vacancy concentrations on different bc-planes, which in principle should show up in diffraction. However, there is no experimental evidence for $2Q_{Ic}$ satellites in the diffraction pattern (Angel and Prewitt 1986). From our calculation, the amplitude of $A(2)$ does not exceed 5% of $A(1)$ over the whole range of composition, so it would be practically impossible to observe this satellite (its intensity would be 400 times smaller!). We conclude that $A(2)$ and some segregation of the vacancies on different bc-planes is consistent with the experimental findings and improves agreement with experiment (figure 7).

4. Low-temperature approximation in real space

Finally, in order to carry the analysis to still lower reduced temperature (higher concentration), lock-on of the modulation to the periodicity of the crystal needs to be considered. For large value of x the ratio $T_a/T_c(x)$ is small and hence the amplitude becomes large, of the order of unity. In this limit higher-order terms in the free energy expansion (3.3) are required. The Umklapp terms in (3.3) also need to be included, because in the squared-up regime they are known to lower the free energy at commensurate wave vectors (Heine 1989, Parlinski 1985). This problem is best addressed in real space.

This approach leads us again to minimize the free energy (2.1) with respect to $\{\Delta_i\}$ under the constraint (2.3). However this problem can be simplified if we include the findings of sections 2 and 3. In particular, according to the Q_{Ic} in (2.8), one clearly finds that the vacancies tend to lie along lines in the b-direction, spaced a distance $2c$ apart. Hence each bc atomic layer can appear in two different ways according to the phase of modulation, as shown in figure 8. We can describe both types of layer by writing

$$\Delta_i \text{ on layer } r = \Delta_r \cos(\tfrac{1}{2}c^* \cdot r_i) \tag{4.1}$$

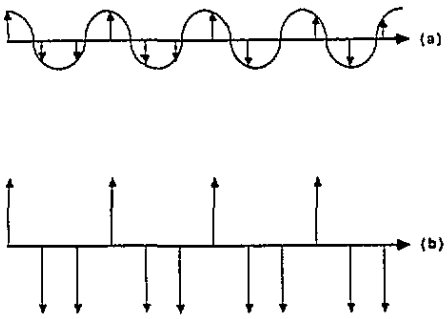

Figure 9. Squaring-up effect on a modulation. In (*a*) the energy of the ordering pattern described by the IC sine wave is left invariant by a shifting of the phase. In (*b*) the squaring up of the modulation locks the phase and the wavelength onto the lattice.

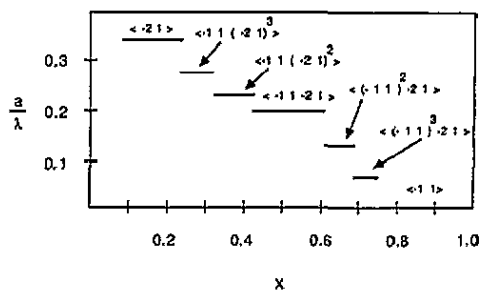

Figure 10. Wavelength of modulation versus composition x on the isotherm $T_a = 1500\,^{\circ}\mathrm{C}$ in the real space treatment. Only commensurate wavelengths can be considered. Note that the wavelength increases with the composition. The solution $\langle -21 \rangle$ gives a wavelength $\lambda = 3$ (in units of a) and $\langle -11 \rangle$ has $\lambda = \infty$. The intermediate solutions have intermediate wavelength.

with Δ_r being positive or negative respectively for the two types of layer. Note that this definition implicitly removes the segregation of vacancies between *bc*-planes because the constraint (2.3) is always fulfilled on each layers. Hence we can consider an effective one-dimensional free energy which is obtained by inserting (4.1) into (2.1). The free energy per (*abc*) unit cell takes the form:

$$F_{MF}(T) = \text{constant} - \frac{2N_p}{N}\sum_r (J_0\Delta_r\Delta_r + J_1\Delta_r\Delta_{r+1/2} + J_2\Delta_r\Delta_{r+1}) + \frac{kT}{2}\frac{2N_p}{N} \times \sum_r [(x \pm \Delta_r)\log(x \pm \Delta_r) + (2 - x \pm \Delta_r)\log(2 - x \pm \Delta_r)]. \tag{4.2}$$

The index r runs only along the direction $\boldsymbol{a}$ and N_p is the total number of O_c sites within a *bc*-plane. The coupling terms are respectively given by

$$J_0 = (C_{002} + C_{020} + C_{040} - C_{001})/4 = +0.29 \text{ eV}$$
$$J_1 = (2C_{110} + 2C_{130} - 4C_{111})/4 = -0.30 \text{ eV} \tag{4.3}$$
$$J_2 = (2C_{220} + C_{200})/4 = -0.13 \text{ eV}.$$

Note that the expression of the energy in (4.2) presents a similar structure to an ANNNI model. In this the IC structure results from the competition which occurs between J_1 and J_2 because in (4.3) J_2 is antiferromagnetic and $|J_1| < 4|J_2|$ (Yeoman 1988, Selke 1988). However, the physical origin and structure of the entropy terms are quite different from the usual ANNNI model (Villain and Gordon 1980, Smith and Yeomans 1983, Yeomans 1988). The difference arises because of the constraint on composition (1.4).

In what follows we perform a series of calculations for commensurate structures with q_a in the range of interest. We calculate $F_c(T_a, x)$ for each of these phases and we isolate the one that has the lowest free energy. These structure are obtained by squaring up the Δ_r of a sinusoidal modulation (figure 9). The phase of the modulation locks on to the periodicity of the crystal (Heine 1989) and our variational calculation of $F_c(T_a, x)$ allows it to find its best phase to minimize the free energy. We defer until later a discussion of how this relates to the real situation where $Q_{IC}(x)$ appears to vary continuously or nearly continuously with x (Cameron 1977).

Thus the general procedure that we followed to determine the solutions $\{\Delta_i\}$ at different compositions can be stated as follows:

(a) We selected a set of trial solutions, e.g. $\langle -11\rangle$, $\langle -21\rangle$, $\langle -21-11\rangle$ and where the sequences $\langle -n, m \ldots\rangle$ represent n consecutive $\Delta_r < 0$, m consecutive $\Delta_r > 0$ and so on. We also constrained the wavelength to be not larger than $6a$.

(b) We minimized the free energy $F_c(T_a, x)$ of each of these configurations for various compositions with periodic boundary conditions.

(c) Finally we isolated at a given composition the phase that minimized the free energy.

The results are presented in figure 10. Note that the wavelength of modulation increases with the composition. With T_a near T_c at low concentration (figure 5), we find the phase $\langle -21\rangle$ which corresponds to this wave vector $Q_C = 0.5c^* + 1/3a^*$ in the series of (2.8a) (the subscript 'C' stands for 'commensurate'). At high concentration the system is represented by the phase $\langle -11\rangle$, with corresponding wave vector $Q_c = 0.5c^* + 0a^*$. Such an ordering is also called the i-alumina structure. In principle, if we had considered all possible wavelengths, the plateaus in figure 10 would be narrower and would be joined by an infinite sequence of intermediate phases, a 'devil's staircase' (Selke 1988, Yeomans 1988). In any case these structures contain locally the two symmetries discussed by McConnell and Heine (1985).

Returning to the comparison with the observed continuous $Q_{IC}(x)$, we also represented in our earlier figure 7 the commensurate value of Q_c at the centre of each stability range of figure 10. In reality equilibrium might never actually be achieved to such a high degree, i.e. actually following the width of the plateaus on the devil's staircase. To obtain the modulation Q_{Ic} one only needs equilibrium over about three atomic planes, but to obtain full equilibrium for a commensurate structure one would need to achieve equilibrium over a range of few repeat distances, i.e. a few times six planes for $\langle(-21)^2 - 11\rangle$. Thus it is sensible and realistic to join up the squares on figure 7. Perhaps some of the scatter in experimental data around the smooth curve can be explained by a slight tendency towards lock-on.

5. Double incommensurate ('twinning') structure

Detailed experimental investigations show that in fact the structure contains a second modulation along the c-direction beyond the composition $x \simeq 0.5$ (Nakajima and Ribbe 1981) of the form

$$Q_{Ic}(x) = q_a(x)a^* + [0.5 - q_c(x)]c^*. \tag{5.1}$$

This result is not predicted by our model and one can understand the reason for this from table 1: the interactions $C_{001} < 0$ and $C_{002} > 0$ accommodate very well the antiferro-ordering along the c-direction and hence have no reason to give an IC ordering along this direction.

So far we have assumed that the coupling terms are independent of the composition. In reality, a renormalization is needed in the region of high vacancy concentration. An inspection of the structure (e.g. figure 11) shows that for $x > 0.5$, mullite needs to accommodate more than one vacancy per four O_c sites. Consider the vacancy on the bc-plane $X = 0$ in figure 11, and the nearby sites 2, 3, 5 and 6 on the next plane at $X = a/2$. When $x = 0.5$, the mean vacancy concentration is $P = x/2 > 0.25$. This means that at least one of the four sites mentioned must on average be occupied by a vacancy

'intruder'. All are unfavourable sites in the sense of figure 4 as discussed in II. The one least unfavourable, and hence most likely to be occupied, is site 5. It would result in an (Si, Si, Al*) 3-cluster, on the plane $X = 0$. Taking account of the high Al concentration for large x, i.e. the reaction exchange (1.1), we think it would tend to turn itself into the (Si, Al, Al*) 3-cluster which is also energetically more favourable because of the lower total cation charge around the O_c^* ion as discussed in I. This in turn will affect the values of the coupling terms C_{001} and C_{002}. In particular the positive sign (energetically favourable) of C_{002} depended as we saw in II on the pair of Si atoms in between two sites a separation $2c$ apart.

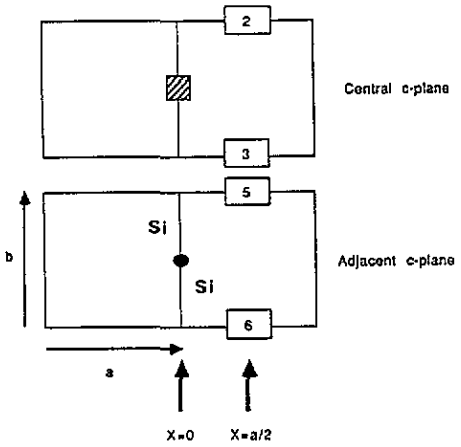

Figure 11. Sites accessible by an intruder vacancy beside the central vacancy (shaded) at high concentration. Using table 1 we find that it is located preferentially at $X = a/2$ on site 5 or 6.

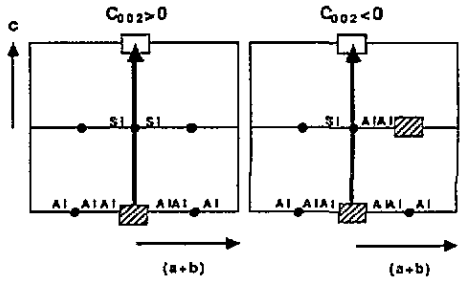

Figure 12. Effect of the vacancy intruder on the interaction C_{002}. The likely (Si, Si) bond (left) which initially mediated a ferro-type interaction is destroyed. The (Al, Si Al*) 3-cluster produced (right), turns C_{002} into a repulsive interaction. The intruder progressively appears at high vacancy concentration beyond $x > 0.5$.

Following the same type of computer modelling as in II, we have shown that the conversion of the (Si, Si) pair of figure 12 to the 3-cluster (Si, Al, Al*) does indeed lead to a dramatic reduction of C_{002}. With such an 'intruder' vacancy on site 5, we find that $C_{001} = -0.80$ eV and $C_{002} = -0.25$ eV. Note that the positive sign of C_{002} is lost. Instead we have both couplings negative, which results in the same type of competition as discussed in section 4 in connection with the $q_a a^*$ modulation. By the same argument the competition of C_{001} and C_{002} now leads to a modulation with an extra q_c in the wave vector of the form (5.1). If we use the findings of II we can estimate q_c at the transition temperature

$$q_c \simeq 0.25 \tag{5.2}$$

which should apply for a very high vacancy concentration. In reality q_c increases from 0 at $x \simeq 0.5$ to 0.3 at $x \simeq 0.6$ (Nakajima and Ribbe 1981). Clearly for such a range of x the effective C_{002} will be some linear interpolation between the value +0.19 eV in table 1 and the value −0.25 eV found here, which will give a lower value of q_c. To give a proper theory is beyond the capabilities of our mean-field theory but we regard our results to be in qualitative and even order of magnitude agreement with experiment.

6. Summary and discussion

We have used a microscopic model which has successfully enabled us to discuss the modulated structure of mullite. Our energy model considers the O vacancy as the central object. An isolated vacancy is surrounded by an ideal 'dressing' cloud of Al/Si ordering. If two vacancies approached too closely to one another, their 'dressings' overlap and result in repulsive interactions. Using the field of interaction calculated recently in II, we have built up an Ising model with similar features to those in an ANNNI model with frustration. We estimated the solutions within the mean-field framework corresponding to thermal equilibrium at some annealing temperature $T_a = 1500\,°C$ on formation. Two approaches were used.

(a) We employed an analysis in the reciprocal space including and excluding the even harmonics. This approximation is expected to be valid only for T_a near T_c and hence at low x. Taking $x = 0.3$, the calculated wavelength λ is $\sim 3.2a$ and the observed on $2.8a$. We regard this as satisfactory agreement. We also find gratifying that the calculated wavelength increases with x qualitatively as observed, although the method in reciprocal space is less valid at the larger x.

(b) At lower effective temperature $T_a/T_c(x)$, i.e. at higher concentration x, the lock-on on the lattice is best addressed by following a real space analysis. The behaviour of $Q_c(x)$ follows a devil's staircase picture which is in reasonable agreement with observation if we ignore the discontinuities between the steps for the reasons discussed in section 4. The results agree with the reciprocal space method at low x, so that the calculated wavelength is again a bit high (figure 7), but it now has the right slope in its variation with x.

We also discussed the structure of mullite at very high concentration. Experimentally another modulation occurs along the c-direction beyond $x \simeq 0.5$. This result was not predicted by our original model. However, if we include the vacancy intruder which occurs at such composition, we find that the coupling terms C_{001} and C_{002} along the c-direction turn out to compete with each other in the sense of giving an additional incommensurate modulation in the c-direction, as observed.

We have seen the role of entropy in all the calculations. In our mean-field approximation, the wave vector q_{Ic} (T_c) of modulation at $T_c(x)$ is always the same. The observed variation of $q_{Ic}(x)$ with composition x arises in our calculations from the variation of the ratio $T_a/T_c(x)$ with x, via the entropy terms in (2.1) and subsequent equations. It is the residual entropy of disorder of the vacancies along the lines of possible sites in figure 8. We shall show in subsequent articles (Padlewski *et al* 1993, 1994) that the residual entropy also plays a key role in establishing the stability field of mullite in the SiO_2/Al_2O_3 phase diagram, indeed that it is a stable material at all. Another manifestation of the residual entropy is the existence of strong diffuse scattering. This has been discussed by Butler *et al* (1993). They adopted a similar model to our own with interactions between vacancies equivalent to our C_{ij}, and evaluated the degrees of short- and long-range order by a Monte Carlo simulation. Our C_{ij} gave results in broad agreement with the observed diffuse scattering, but some moderate change in their values was required to fit the observation completely.

Acknowledgments

We wish to thank Professor J D C McConnell and Dr R J Angel for useful and helpful discussion. SP would also like to thank the French Ministry of Research and Industry for providing a grant.

References

Angel R J and Prewitt C T 1986 *Am. Mineral.* **71** 1472–82
—— 1987 *Acta Crystallogr.* B **43** 116–26
Bak P 1982 *Rep. Prog. Phys.* **45** 587–629
Bertram U C, Heine V, Jones J L and Price G D 1990 *Phys. Chem. Minerals* **17** 326–33
Butler B D, Welberry T R and Withers R L 1993 *Phys. Chem. Minerals* to be published
Cameron W E 1977 *Am. Mineral.* **62** 747–55
Chakraborty K G and Tucker J W 1986 *Physica* A **137** 111
Holland T J B and Carpenter M A 1986 *Nature* **320** 151
Heine V 1989 *Phase Transitions* **15** 311–23
McConnell J D C 1991 private communication
McConnell J D C and Heine V 1985 *Phys. Rev.* B **31** 6140–2
Nakajima Y and Ribbe P H 1981 *Am. Mineral.* **66** 142–7
Padlewski S, Heine V and Price G D 1992a *Phys. Chem. Minerals* **18** 373–8
—— 1992b *Phys. Chem. Minerals* **19** 19
—— 1993 *Phys. Chem. Minerals* in preparation
Padlewski S, McConnell J D C, Heine V and Price G D 1994 *Am. Mineral.* in preparation
Parlinski K 1985 *J. Phys. C: Solid State Phys.* **18** 5667–82
Selke W 1988 *Phys. Rep.* **170** 214–64
Smith J and Yeomans J 1983 *J. Phys. C: Solid State Phys.* **16** 5305–20
Villain J and Gordon M B 1980 *J. Phys. C: Solid State Phys.* **13** 3117–33
Yeomans J 1988 *Solid State Physics* vol 41 (New York: Academic) pp 151–200
Yoikoi C S O, Coutinho-Filho M D and Salinas S P 1981 *Phys. Rev.* B **24** 4047–61