STRUCTURAL SCIENCE
CRYSTAL ENGINEERING
MATERIALS

Supporting information for article:

The modulated average structure of mullite

Volume 71 (2015)

Johannes Birkenstock, Václav Petříček, Bjoern Pedersen, Hartmut Schneider and Reinhard X. Fischer

S1. Tables for section "3.2 Average structure analyses – neutron and X-ray data"

Table S1 Final structure parameters for the average structure refinements. The three rows supply the values for the refinements based on X-ray data, on *neutron data* and on **combined X-ray and neutron data**, respectively.

Atom	x y z	U_{11}	U_{22}	U_{33}	U_{12}	U_{equiv}	Principal axes
Al1	0 0 0	0.0075(3)	0.0066(3)	0.0052(3)	0.0008(1)	0.0065(2)	0.0052 0.0062 0.0080
	0 0 0	*0.0061(4)*	*0.0058(4)*	*0.0020(5)*	*0.0009(3)*	*0.0047(3)*	*0.0020 0.0051 0.0069*
	0 0 0	**0.0065(3)**	**0.0061(2)**	**0.0034(3)**	**0.00089(12)**	**0.00534(14)**	**0.0034 0.0054 0.0072**
T	0.14889(6) 0.34026(6) ½	0.0063(3)	0.0080(3)	0.0068(3)	-0.0005(1)	0.0070(2)	0.0062 0.0068 0.0081
	0.14901(9) 0.34009(13) ½	*0.0044(4)*	*0.0080(4)*	*0.0032(5)*	*-0.0013(20*	*0.0052(3)*	*0.0032 0.0040 0.0084*
	0.14897(5) 0.34018(5) ½	**0.0051(2)**	**0.0074(2)**	**0.0051(2)**	**-0.00062(11)**	**0.00587(14)**	**0.0049 0.0051 0.0076**
T*	0.2624(2) 0.2055(3) ½	0.0059(7)	0.0064(8)	0.0049(7)	0.0012(6)	0.0058(4)	0.0049 0.0049 0.0074
	0.2618(4) 0.2069(5) ½	*0.0006(14)*	*0.007(1)*	*0.004(2)*	*0.0003(10)*	*0.0038(9)*	*0.0006 0.0038 0.0067*
	0.2621(2) 0.2056(2) ½	**0.0050(7)**	**0.0069(7)**	**0.0040(6)**	**0.0007(5)**	**0.0053(4)**	**0.0040 0.0047 0.0071**
O1	0.3582(1) 0.4227(1) ½	0.0139(5)	0.0191(5)	0.0063(4)	-0.0081(3)	0.0131(3)	0.0063 0.0080 0.0250
	0.35840(6) 0.42242(9) ½	*0.0119(4)*	*0.0191(4)*	*0.0022(4)*	*-0.0082(2)*	*0.0111(2)*	*0.0022 0.0066 0.0245*
	0.35843(5) 0.42246(7) ½	**0.0127(3)**	**0.0193(3)**	**0.0034(3)**	**-0.00836(14)**	**0.01184(15)**	**0.0034 0.0071 0.0250**
O2	0.1277(1) 0.2186(1) 0	0.0153(4)	0.0134(5)	0.0109(4)	-0.0064(3)	0.0132(3)	0.0079 0.0109 0.0209
	0.12740(6) 0.21847(7) 0	*0.0126(3)*	*0.0132(3)*	*0.0088(4)*	*-0.0066(1)*	*0.0116(2)*	*0.0063 0.0088 0.0195*
	0.12738(5) 0.21852(6) 0	**0.0140(3)**	**0.0133(3)**	**0.0098(3)**	**-0.00687(13)**	**0.01238(15)**	**0.0068 0.0098 0.0205**
O3	½ 0 ½	0.009(3)	0.020(4)	0.029(2)	0.002(2)	0.019(2)	0.0086 0.0200 0.0058
	½ 0 ½	*0.018(2)*	*0.017(2)*	*0.023(1)*	*-0.0055(14)*	*0.020(1)*	*0.0123 0.0231 0.0233*
	½ 0 ½	**0.0167(16)**	**0.0183(17)**	**0.0238(11)**	**-0.0054(11)**	**0.0196(9)**	**0.0120 0.0230 0.0238**
O4	0.4517(9) 0.0480(9) ½	0.006(3)	0.009(3)	0.012(2)	0.0003(19)	0.0088(16)	0.0058 0.0091 0.0115
	0.4500(6) 0.0522(5) ½	*0.011(2)*	*0.004(2)*	*0.008(1)*	*0.0005(10)*	*0.0076(9)*	*0.0038 0.0078 0.0112*
	0.4493(6) 0.0515(4) ½	**0.0108(14)**	**0.0058(12)**	**0.0100(9)**	**-0.0008(8)**	**0.0089(7)**	**0.0058 0.0100 0.0110**

Comments on the three series of refinements with different sets of constraints:

In the first refinement with fixed Si/Al distribution and Si only on the T position (table S2) the refined compositions and the molar fractions of Al_2O_3 were in perfect agreement with the expected composition of 2:1 mullite.

In the second constrained refinement the average fractions of Si and Al on both the T and the T* positions were refined while the overall composition was kept ideal for 2:1 mullite (table S3). All three refinements gave a weak indication that a small portion of the Si might also occupy the T* position to 8(5) % (combined data), to 11(7) % (neutron data) or to 23(15) % (X-ray data),

respectively. The values agree within one standard uncertainty, yet they are also zero within two standard uncertainties. It is clear that the value from the X-ray data alone is not significant due the similar scattering powers of Si^{4+} and Al^{3+}. But unfortunately the values from the neutron data and from the combined data suffer from similarly large relative uncertainties of more than 60 %. Accordingly, a significantly clear picture on the true Si/T* occupancy could not be derived from these data. However, in a recent ^{27}Al and ^{29}Si MAS-NMR study of mullite, King (2014) suggested that a minor part of Si also enters the triclusers which is consistent with our findings. It should be noted, though, that the R_{obs} values are almost identical to those of the refinements on composition only (table S2). Accordingly, from these data a model with Si on the T position only can hardly be distinguished from a model with Si distributed over the T and the T* positions.

In the third refinement (table S4, supplement) we performed a simultaneous constrained refinement on the overall composition and the distribution of Si and Al over the T* and the T site (table S4) at a time. For all three data sets the SOF factor of O4 is nicely reproducible within three standard uncertainties and accordingly the overall composition of 2:1 with x = 0.4 is undoubted. With regard to the Si distribution over T and T* sites the X-ray data only gave a physically meaningless negative value for Si on T*. This is clearly due to the very similar X-ray scattering powers of Si^{4+} and Al^{3+}. For the neutron data 20(13) % of the Si was found on T* sites, for the combined data it was 14((7) %. Again these values weakly indicate a partial occupation of the T* by Si but again these values are also zero within two standard uncertainties. The R_{obs} values are again almost identical to the previous ones.

We conclude that the refinement of SOF parameters relevant for the composition is equivalently valid for X-ray and neutron data sets while a conclusive result for the refinement of SOF parameters on the Si distribution is not really feasible with none of our data although the small positive values for Si on the T* position in case of the neutron data are supported by the ^{29}Al MAS-NMR study of mullite (King, 2014). Accordingly we suggest that wqe have final results for the neutron data and for the combined data in table S4.

The rather insignificant results the Si distribution even in case of the neutron data is somewhat surprising. We think that this may be understood as follows:

For steric reasons the T* position can be occupied at maximum to 50% as the two directly neighboring T* positions are too close to each other to be occupied simultaneously (see fig. 1). Of these available sites, e.g., from the neutron data in table S3 only 39.4(4) % are occupied by the sum of (Al,Si) atoms, with the standard uncertainty derived from the occupation factor of O4. At maximum (30±20) % in T* are Si atoms, i.e., at maximum the overall Si/T* occupation of the available sites amounts to 12±8% only, with a relative one sigma uncertainty of 67%! This means that the small contribution to the diffracted intensities induced by the replacement of Al ↔ Si is weighted by a factor of 0.12 and accordingly very small. This may explain the lack of sensitivity even for the

neutron data. As a result, within 1.5 standard uncertainties we cannot distinguish between the case of 60 % Si/T* atoms on the available sites and that of 0%. This is strongly supported by the fact that the R_{obs} values are almost exactly identical for all three refinement series, including the case where a zero occupation of Si/T* was assumed (table S2).

Excluding the unreasonable refinement based on X-ray data in table S4 the compositional values only vary within one standard uncertainty between 0.097(4) and 0.1004(9) and accordingly produced a relative one sigma uncertainty of only 4 to 1%. This translates to compositional values of 0.388(16) to 0.402(4) in terms of the x-value, of 1.94(8):1 to 2.008(18):1 in terms of the Al_2O_3:SiO_2 ratio and of $Al_{4.78(3)}Si_{1.22(3)}O_{9.612(16)}$ to $Al_{4.803(7)}Si_{1.197(7)}O_{9.5984(4)}$ in terms of overall composition per average unit cell. This clearly indicates that the data are sensitive to the overall composition and that these values are not affected by the large uncertainties in case of refined Si/T* occupations in mullite.

Table S2 Results and compositions derived from refinements of the average 2:1 mullite structure with constrained chemical composition and fixed Si/Al distribution. Refinements were performed on X-ray data, on neutron data and on combined X-ray and neutron data, respectively.
SOF = site occupation factor = $\frac{multiplicity\ of\ site}{multiplicity\ of\ general\ position} occupation\ factor.$

Quantity	Constraint	Start value	X-ray data	neutron data	combined data
SOF(O4)	free variable (= $^{x}/_{4}$)	0.1	0.1004(9)	0.0976(14)	0.1003(8)
SOF(O3)	= 1/4– 3/2·SOF(O4)	0.1	0.0994(14)	0.1037(27)	0.0996(12)
SOF(Al3/T*)	= SOF(O4)	0.1	0.1004(9)	0.0976(14)	0.1003(8)
SOF(Si2/T)	= 1/4–SOF(O4)	0.15	0.1496(9)	0.1524 (18)	0.1497(8)
Derived composition		$Al_{4.8}Si_{1.2}O_{9.6}$	$Al_{4.803(8)}Si_{1.197(8)}O_{9.598(13)}$	$Al_{4.780(15)}Si_{1.220(15)}O_{9.61(3)}$	$Al_{4.793(7)}Si_{1207(7)}O_{9.603(12)}$
Al_2O_3 content [mol%]		66.7	66.7(4)	66.2(8)	66.5(4)
x		0.4	0.40(1)	0.39(3)	0.40(1)
R_{obs} [%]	–	–	2.14	4.96	4.82 (X-ray 2.12, n 5.04)

Table S3 Results of refinements with constrained Si/Al distribution on T and T* sites and fixed chemical composition of $Al_{4.8}Si_{1.2}O_{9.6}$. Refinements were performed on X-ray data, on neutron data and on combined X-ray and neutron data, respectively. SOF = see table S2.

Quantity	Constraint	Start value	X-ray data	neutron data	combined data
SOF(Al2/T)	1/4 + SOF(Si/T*)	¼	0.284(22)	0.267(11)	0.262(8)
SOF(Si2/T)	0.15–SOF(Si/T*)	0.15	0.116(22)	0.133(11)	0.138(8)
SOF(Al3/T*)	0.1[1] - SOF(Si/T*)	0.1	0.066(22)	0.083(11)	0.088(8)
SOF(Si3/T*)	free variable	0	0.034(22)	0.017(11)	0.012(8)
R_{obs} [%]	–	–	2.13	4.97	4.82 (X-ray 2.12, n 5.04)

[1] The reference value 0.1 is taken from table S2 and represents the ideal composition of 2:1 with SOF(O4) = 0.1. The sum of cations on T* must be identical to this value (see text).

Table S4 Results of refinements with constrained Si/Al distribution on T and T* sites and constrained chemical composition. Refinements were performed on X-ray data, on neutron data and on combined X-ray and neutron data, respectively. SOF = see tableS 2. The results taken from X-ray data were physically meaningless due to unreasonable values in the Si,Al distribution (*italic characters*). For the neutron and the combined refinement there is no clear indication whether Si shares the T* position or not.

Quantity	Constraint	Start value	X-ray data	neutron data	combined data
SOF(O4)	free variable	0.1	0.092(3)	0.097(4)	0.0987(11)
SOF(O3)	= 1/4 – 3/2·SOF(O4)	0.1	0.111(4)	0.104(6)	0.1019(17)
SOF(Al2/T)	= 1/4 + SOF(Si/T*)	¼	*0.45(6)*	0.28(2)	0.271(11)
SOF(Si2/T)	= 1/4 – SOF(O4) – SOF(Si/T*)	0.15	***-0.05(6)***	0.12(6)	0.130(11)
SOF(Al3/T*)	= SOF(O4) – SOF(Si/T*)	0.1	*-0.11(6)*	0.07(2)	0.078(11)
SOF(Si3/T*)	free variable	0	*0.20(6)*	0.03(2)	0.021(11)
R_{obs} [%]	–	–	2.15	4.97	4.82 (X-ray 2.12, n 5.04)

S2. Tables related to section "The modulated structure – first q-vector"

Table S5 Interatomic distances in [Å] calculated for modulated 2:1 mullite and the average structure. Notes: [1)] The distance from T to O4 is different for the two TO_4 tetrahedra in a tricluster! [2)] Standard uncertainties from error propagation. [3)] Values for the average structure. Values for the TO_4 tetrahedra in tricluster were not explicitly listed. [4)] min and max values refer to the minima and maxima in the top diagram of Fig. 5 with corresponding average values.

Type	Atom		average[4)]	min[4)]	max[4)]	Average structure	Angel & Prewitt (1986)[3)]
Octahedra	Al1	–O1 (4x)	1.8977(7)	1.8901(7)	1.9054(7)	1.8934(3)	1.8936(5)
		–O2 (2x)	1.9416(9)	1.9415(9)	1.9416(9)	1.9331(4)	1.9366(9)
		mean	1.9123(3)	1.9072(3)	1.9174(3)	1.90663(14)	1.9079(3)
Dicluster T_2O_7	T	–O1	1.7109(12)	1.7023(12)	1.7192(12)	1.7073(5)	1.7102(8)
		–O2 (2x)	1.7289(8)	1.7184(8)	1.7393(8)	1.7255(3)	1.7273(5)
		–O3	1.670(3)	1.646(6)	1.693(6)	1.6656(4)	1.6676(2)
		mean	1.7097(3)	1.6963(3)	1.7227(3)	1.70598(13)	1.70810(17)
Tricluster $T_2T^*O_{10}$	T	–O1	1.7109(12)	1.7023(12)	1.7192(12)	1.7073(5)	
		–O2 (2x)	1.7289(8)	1.7184(8	1.7393(8)	1.7255(3)	
		–O4	1.733(11)	1.691(12)	1.774(12)	1.726(4)	
		mean	1.7254(3)	1.7075(3)	1.7430(3)	1.72108(13)	
	T[1)]	–O1	1.7109(12)	1.7023(12)	1.7192(12)	1.7073(5)	
		–O2 (2x)	1.7289(8)	1.7184(8	1.7393(8)	1.7255(3)	
		–O4	1.755(13)	1.689(15)	1.821(15)	1.782(3)	
		mean	1.7309(4)	1.7070(4)	1.7547(4)	1.73508(12)	
	T*	–O1	1.821(3)	1.807(3)	1.836(3)	1.8146(17)	1.8166(11)
		–O2 (2x)	1.7739(16)	1.7611(16)	1.7865(16)	1.7699(9)	1.7727(7)
		–O4	1.895(12)	1.848(11)	1.941(11)	1.846(4)	1.8522(41)
		mean	1.8160(7)	1.7943(7)	1.8375(7)	1.8036(7)	1.8036(7)

S3. Hypothetical superstructures compatible with the chemical composition of 2:1 mullite

To argue why mullites form disordered and weakly modulated structures is not straight forward. We tackle this question from the opposite end: Is it possible to construct superstructures with ordered arrangements of the basic building units? And if yes, what hinders the mullite structure to adopt such a superstructure? To start with we consider the peculiarities of the structural motifs found in mullites first. It is a basic crystal chemical principle that similar building units in a crystal tend to locate in similar environments which is usually nicely compatible with a long-range ordered structure. In

mullites the rule of similar environments is broken at several levels of complexity, at three of which mullite differs from the closely related sillimanite:

1.) Al is located in both tetrahedra and octahedra in the same crystal structure. However, this is also true for the sillimanite structure, which is perfectly long-range ordered.

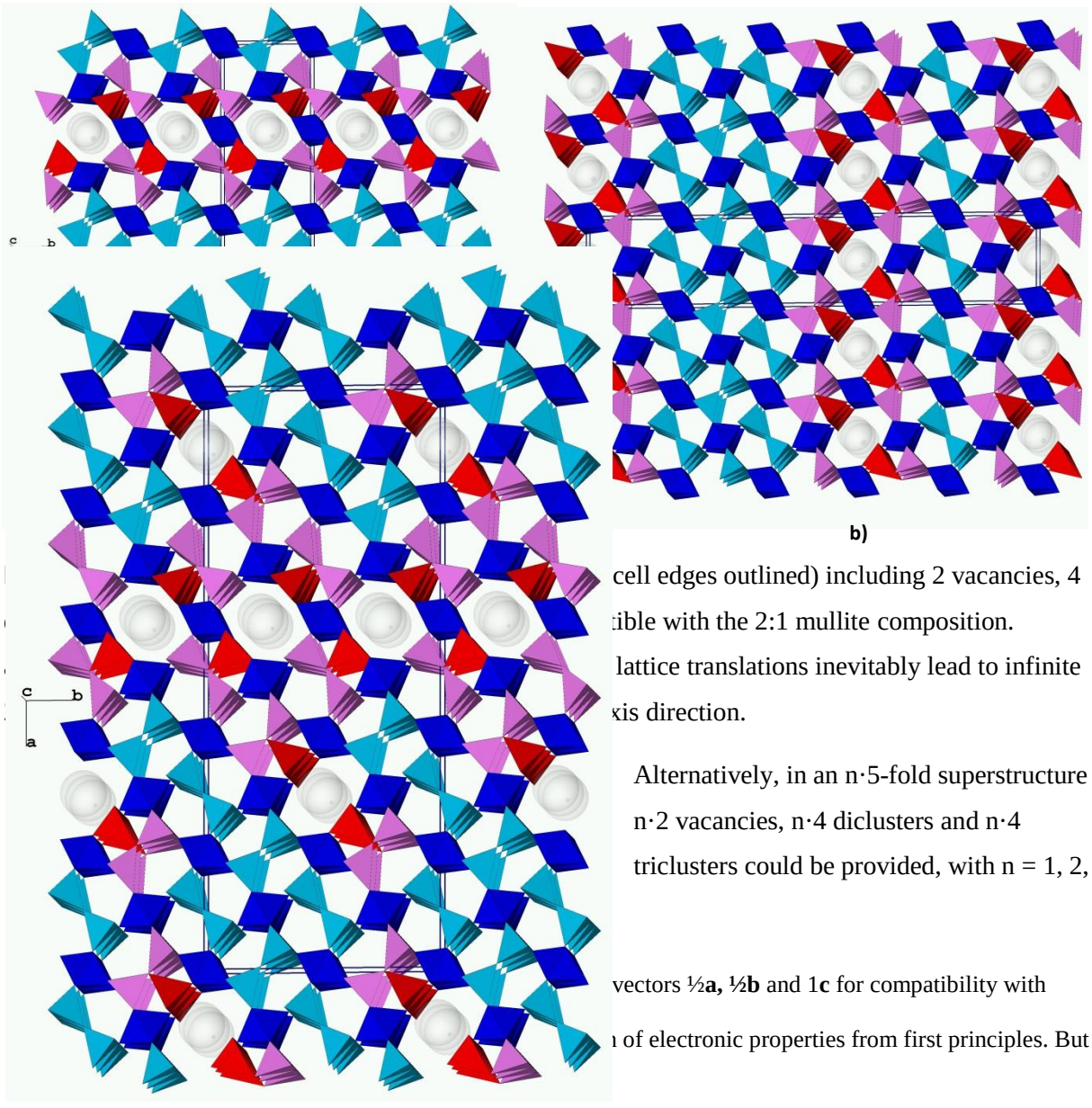
2.) Si and Al share the tetrahedral TO_4 sites and possibly the T^*O_4 sites to some extent without ordering. However, this is often observed for phases generated at high temperature due to the entropy term that promotes mutual replacement and disorder for ions such as Si^{4+} and Al^{3+} behaving crystal-chemically similarly in similar environments. From this point of view it is rather surprising that sillimanite is so well ordered with respect to Si and Al.

3.) Al is found in two different types of distorted tetrahedra, TO_4 and T^*O_4, having a significantly different geometric distortion. This is clearly not ideal from energy considerations.

4.) From the two primary building units TO_4 and T^*O_4, two different types of composite building units are formed, the T_2O_7 diclusters (formed by two TO_4 tetrahedra) and the $T^*T_2O_{10}$ triclusters (formed by one T^*O_4 and two TO_4 tetrahedra). It seems energetically unfavorable that diclusters and triclusters coexist and on average occupy identical cavities enclosed in between the octahedral units.

The need to build a rather complex structure from simple, rather rigid primary building units is rationalized as follows for 2:1 mullite ($Al_{4.8}Si_{1.2}O_{9.6}$). Two formula units of Al per unit cell are located in the AlO_6 octahedra forming the edge-sharing chains parallel to the **c**-axis while 2.8 Al and 1.2 Si share the tetrahedral sites. In sillimanite there are no O3-vacancies and no T* sites, in 2:1 mullite – on average – there are 0.4 vacancies, 0.8 diclusters T_2O_7 and 0.8 triclusters $T^*T_2O_{10}$ per unit cell, each of them in similar environments midst between the octahedral chains. As these units are very dissimilar, in neighboring unit cells within an **a**-**b**-plane there can be a strongly varying electron and mass density in the spaces between the octahedral units.

To provide space for vacancies, diclusters and triclusters in integral units, a hypothetical ordered 2:1 superstructure of mullite must at least be five-fold to provide integral numbers of sites for 2 vacancies (5·0.4), 4 diclusters (5·0.8) and 4 triclusters (5·0.8). A five-fold superstructure could only form in a linear arrangement of average mullite cells in direction of the **a**- or the **b**-axis. A 5-fold supercell with **c'** = 5·$\mathbf{c}_{\mathbf{mullite}}$ is excluded because this would inevitably induce a regular pattern of vacancies along the vectors[2] [220] and [2-20] which cannot exist as shown in the supplement D)[3]. Examples of hypothetical 5**a**- and 5**b**-supercells are shown in fig. S1. However, the vacancies, the diclusters, and the triclusters inevitably line up two-dimensionally in planes normal to the elongated supercell axis, forming parallel sheets with extremely high concentrations of either one of these composite building units. A layered structure would result where the electron and mass densities would vary strongly normal to these planes. The linking octahedra would face completely different environments in different directions normal to the octahedral chains which seems crystal-chemically unfavorable.



cell edges outlined) including 2 vacancies, 4
ible with the 2:1 mullite composition.
lattice translations inevitably lead to infinite
xis direction.

Alternatively, in an n·5-fold superstructure n·2 vacancies, n·4 diclusters and n·4 triclusters could be provided, with n = 1, 2,

vectors ½**a,** ½**b** and 1**c** for compatibility with
of electronic properties from first principles. But

Figure S2 Hypothetical 5**a**·2**b**·1**c** superstructure with 4 vacancies, 8 di- and 8 triclusters per unit cell (edges weakly). Symbols as defined in fig. 1.

3, ….. To avoid lining-ups in at least one of the two planar directions the structure would have to adopt at least a 5·2 (n = 2) superstructure. For example, in the 5**a**·2**b**·1**c** superstructure in fig. S2 the vacancies are still partially lined up in planes (here parallel to **b**-**c**) similar to a five-fold supercell but partially only lined up linearly parallel to the **c**-axis. In fig. 6 (main article) two examples of chemically compatible 5x5x1 supercells are displayed but these still suffer from straight line-ups of the vacancies (and accordingly the triclusters) parallel to the **c**-axis. To avoid any straight lining-up of vacancies and thereby provide a smoother distribution of density the size of the supercell would have to be increased and adopt at least a 5·2·2 superstructure (n = 4). In such a structure one could think of an arrangement where we have no simple straight lining-up of the structural units extending throughout the whole crystal but rather a regularly kinked line-up of these along the **c**-direction. This would clearly be more favorable than the planar or linear line-ups. But still this would result in planar arrangements – of thicker sheets now – with different densities with respect to these structural units. So even with such a 20-fold supercell the structure would still remain rather unfavorable.

The same arguments would hold for any other simple superstructure with small n: Any ordered arrangement of vacancies, diclusters and triclusters would produce kinked linear, straight linear or even planar elements of strongly different mass and electron densities running throughout the structure and accordingly any such arrangement must be rather unfavorable. It can also be shown that a construction of the mullite structure starting from one of the preferred pair-correlation vectors would lead to similar problems based on energy considerations (see section D). Furthermore is can be shown that some pairs of vacancies are not feasible within the mullite structure while they were included in the model provided by Freimann & Rahman (2001, see section E).

S4. Hypothetical construction of the mullite structure starting from preferred vacancy correlation vectors

The preferred paired-vacancy correlation vector <401> (after Freimann and Rahman) is used in fig. S3 to construct a different type of hypothetical supercell. If arranged in an ordered (3**a**·2**c**)-superstructure this vector would induce the same number of <201> correlated pairs of vacancies alternating with <401>. <201> is again one of the preferred correlation vectors. An even higher number of <020> vectors is also induced. A perfectly ordered 3**a**·2**c**-superstructure would automatically produce diffraction patterns with sharp satellites at **q** = ($^1/_3$ 0 ½) without diffuse scattering. This **q**-vector is quite similar to that observed for the first set of satellites where $\mathbf{q_1}$ = (~0.31 0 ½). If the order along this modulation wave is less perfect such that the alternation is sometimes not kept perfectly – e.g. by interweaving a sequence like <401> <201> <401> <401> <201> – the average repeat could be stretched from 3**a**·2**c** to 3.19**a**·2**c** which is the modulation wavelength associated with $\mathbf{q_1}$! Thus it may be concluded that locally distributed 3**a**·2**c** supercells may be frequent motifs of the structure and contribute to the observed satellites. However, the better the order, the worse the result in terms of energetic considerations: This type of cell clearly would suffer again from a very regular lining up of the vacancies and of the composite tetrahedral units.

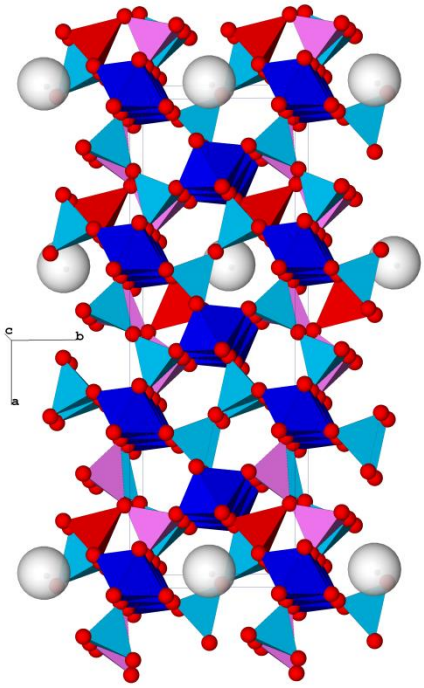


Figure S3 A hypothetical 3**a**·2**c** commensurate superstructure with two O3 vacancies (large light spheres) per supercell separated by the frequent correlation vector <401> (vectors given with base $^{\mathbf{a}}/_2$, $^{\mathbf{b}}/_2$ and **c**). Note that the vectors <201> and <020> are generated in equal amounts automatically due to the periodic arrangement in the superstructure. Light blue = regular T_2O_7 dicluster, light green & red = tricluster with two distorted TO_4-tetrahedra (green) and one T^*O_4 (red).

There is a second drawback to this model: the hypothetical (3**a**·2**c**)-supercell with two vacancies would only provide for 0.3;$^{-}$ vacancies per average unit cell instead of 0.4, and for 1.16;$^{-}$ diclusters and 0.5 triclusters instead of 0.8 each. The stretching of the average repeat distance to 3.19**a** by additional <401> correlations would even decrease the number of vacancies. Accordingly, if such motifs should be responsible for the $\mathbf{q_1}$ satellites they must be interwoven with other motifs of higher vacancy concentration in order to yield the correct average composition. This can be accomplished both in the direction of the modulation and normal to it. Both would introduce additional disorder.

Thus, the real structure obviously needs to incorporate additional paired-vacancy correlation vectors and the according motifs of vacancy-dicluster-tricluster-arrangements. This is what was demonstrated by Freimann and Rahman (2001): The correct composition of 2:1 is realizable by at least one but possibly several distinct linear combinations of the more frequent paired-vacancy correlation vectors. The chemical composition is one hard constraint in deriving possible linear combinations. On the other hand, on average any valid linear combination must be compatible with the observed modulations in the structural arrangement, which sets another constraint on the possible linear combinations and their relative frequencies. An idea of that was presented above for the $\mathbf{q_1}$ satellites. The set of satellites with $\mathbf{q_2}$ = (0 0.4021(5) 0.1834(2)) corresponds to a modulation wave running along 2.49**a** + 5.45**c**. Correspondingly $\mathbf{q_3}$ = (0 0.4009(5) -0.1834(2)) provides for a modulation along 2.49**a** - 5.45**c**. Again no single supercell or correlation vector can be found which could give rise to these superperiods as only integer integral numbers of the structural units can exist. However, a combination of various <0vw> vectors could contribute to such modulations on average.

S5. Possible and impossible local arrangements of the structural building units

Freimann & Rahman (2001) listed 15 preferred paired-vacancy correlation vectors <lmn> in the following order of prominence: <310>, <111>, <022>, <201>, <330>, <401>, <131>, <130>, <042>, <113>, <060>, <600>, <312>, <222>, <040>, with probabilities of occurrence running from 30.3 to 22.5 % (the probability that a given vacancy is involved in the respective vacancy pair). As noted above the vectors are given with respect to the new basis vectors $^{\mathbf{a}}/_2$, $^{\mathbf{b}}/_2$ and **c**, i.e., as vectors [l·$^{\mathbf{a}}/_2$, m·$^{\mathbf{b}}/_2$, n·**c**]) and the brackets "<>" indicate that a set of "equivalent ones" is included which usually means sets of *symmetrically* equivalent directions in a lattice. With orthorhombic lattice symmetry *mmm* a set <uvw> would include the eight vectors [uvw], [u;¯vw], [uv;¯w], [uvw;¯], [u;¯v;¯w], [u;¯vw;¯], [uv;¯w;¯] and [u;¯v;¯w;¯], which reduces to a set of four or two vectors if one or two of the coefficients are zero, respectively.

Looking closer at the paired-vacancy vectors in fig. 6 it is clear that there are some vectors [uv0] present in the **a**-**b**-layers which were not listed among the 15 preferred ones. E.g., in fig. 6a) this applies to the dashed vectors ±[200], ±[020], ±[220], ±[22;¯0], ±[400] and ±[420]. These vectors are also shown by the authors in their fig. 7 but their frequencies were found to be smaller than those of the above ones and thus were not included in the truncated list of preferred vectors.

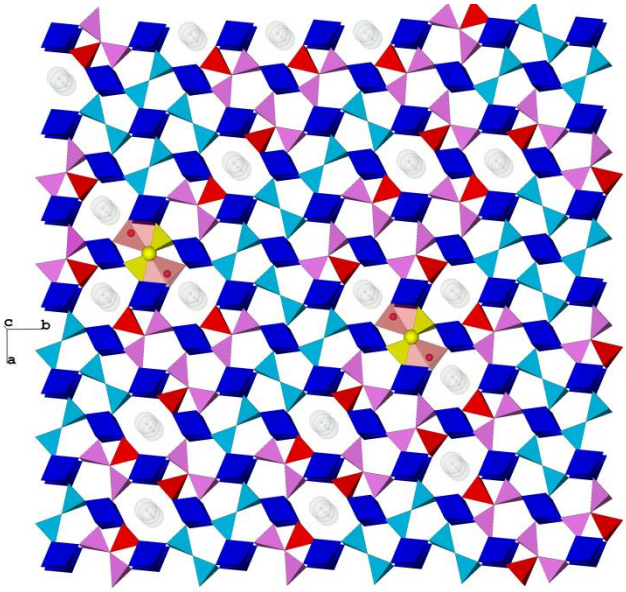


Figure S4 Structure plot of a square section (subcells in a clipped square with the cell coordinates (given as row,column) 7,9 to 13,15) of fig.7a in Freimann and Rahman (2001) with two implicitly four-coordinated O3 atoms (yellow spheres) which are hardly feasible in the structure.

Special constraints apply to the vectors ±[220] and ±[22;¯0]: They should be part of the set <220> and therefore they should be equivalent with respect to the lattice symmetry but one pair of them, ±[220], can only be realized with two paired vacancies in 0½½ (in terms of the respective average cell) while the other one, ±[22;¯0], can only exist with vacancies in the ½0½ positions, respectively. If a ±[220] vector would form with vacancies in ½0½ positions then a four-coordinated O3 atom would result between the two related vacancies with two T^*O_4 and two TO_4 tetrahedra in a $T^*_2T_2O_{13}$ group. This would be rather improbable 1) because of the unfavorable fourfold coordination of the O3 atom, 2) because the two T*-O3 distances would be even much larger (> 2 Å) than the already enlarged value of T*-O4 and 3) because a fourfold junction between four tetrahedra is energetically rather unfavorable. The same arguments are valid for ±[22;¯0] in the case of two paired vacancies in 0½½. We expect that this reasoning holds for compositions with $x \leq 0.67$ where the structure is not "over-saturated" with triclusters. For $x > 0.67$ Fischer, et al. (1994) argued that four-coordinated oxygen may no longer be avoided because at this critical x-value all tetrahedra are involved in triclusters so that diclusters no longer exist.

However, in fig. 7 (top) of Freimann and Rahman (2001) all of the four displayed ±[220] correlation vectors are included for pairs of ½0½ vacancies. This is shown here in fig. S4: a section of 6x6 subcells of their schematic pattern containing two of these pairs is transferred into a polyhedral structure drawing. It shows that two O3 atoms inevitably have to be four-coordinated in the middle of such vacancy-pairs. I.e., the model from which the diffuse scattering was calculated was in part questionable. At the same time none of the allowed ±[220] correlations were applied to paired 0½½ vacancies. For ½0½ vacancies no allowed ±[220] and also no forbidden [22;¯0] correlations are shown. As their calculated diffuse scattering patterns strongly resembled the observed patterns we still think that the overall model of Freimann and Rahman was predominantly correct for the most part but it should be noted that some details were not defined in a crystal-chemically sound manner. On top of

the principles of Freimann and Rahman used for real structure constructions this further constraint must be obeyed, avoiding four-coordinated oxygen in mullites with $x \leq 0.67$.