

Accelerated discovery of two new structure types in a complex inorganic phase field

C. Collins, M.S. Dyer, M.J. Pitcher, G.F.S. Whitehead, M. Zanella, P. Mandal, J.B. Claridge, G.R. Darling and M.J. Rosseinsky*

Department of Chemistry, University of Liverpool, Liverpool L69 7ZD, UK

New materials are key to advancing technology and the science that underpins it. At present, we do not have efficient approaches to explore the large number of possible elemental compositions for such materials, and of candidate structures at each composition¹. For example, the discovery of new inorganic extended solid structures has relied on deep knowledge of crystal chemistry coupled with extremely time-consuming materials synthesis with systematically varied elemental ratios^{2,3}. Computational methods have been developed to guide synthesis by prediction of new structures at specific compositions⁴⁻⁶ and of new compositions for known crystal structures^{7,8}, with notable successes^{9,10}. The challenge of finding qualitatively new, experimentally realisable compounds, with crystal structures where the unit cell and the atom positions within it differ from known structures, rather than materials related by substitution to known structures, however remains for compositionally complex systems. Many valuable properties arise from substitution into known crystal structures, but materials discovery by this approach alone risks both missing best-in-class performance and attempting design with incomplete knowledge^{8,11}. New structures and compositions are needed to deliver new, potentially important and unexpected properties. Here we report the experimental discovery of two new structure types by computational identification of the region of a complex inorganic composition space that contains them. This is achieved by computing probe structures that capture the chemical and structural diversity of the system and whose energies can be ranked against combinations of currently known materials. Subsequent experimental exploration of the lowest energy regions of the computed phase diagram affords two materials with previously unreported crystal structures featuring unusual structural motifs. This approach will accelerate the systematic discovery of new materials in complex compositional spaces by efficiently guiding synthesis and enhancing the predictive power of the computational tools through expansion of the knowledge base underpinning them.

Calculation can identify compositions where new phases have energies that are competitive with combinations of known materials – such phases are said to be close to the convex hull (the energy surface defined by stable compositions)¹². Although there are infinitely many possible compositions, we can target synthetic effort efficiently by calculating the energies of hypothetical compounds spanning sufficient compositional ranges to locate regions likely to afford new materials. For each

composition, we require a probe structure *i.e.*, a plausible crystal structure for the hypothetical compound. This probe structure must be sufficiently close in energy to the minimum energy achievable at this composition to give a representative stability with respect to competing known phases, but need not be the minimum energy structure itself. Probe structures within a phase diagram can be chosen from a single structural family, if that family has sufficient structural diversity to reflect the attainable long-range and local crystal chemistries at finely spaced compositions, or multiple families can be used to explore the same space.

Computational tools to prioritise selection of synthetic targets in compositionally complex systems must handle the accompanying range of local atomic coordination environments and of extended structural units (hereafter referred to as modules, Extended Data Figure 1 and Methods) that can make up the materials. Oxides are an important, compositionally extensive and structurally varied materials family that exemplifies this challenge. We select the Y-Sr-Ca-Ga-O oxide phase field (the space of all accessible phases defined by a group of elements) because the four metallic elements have diverse charge, size and bonding characteristics that lead to a range of coordination geometries and thus test the scope of the approach. Furthermore, related gallate phase fields contain materials with a range of functional optical and transport properties^{13–16}. There are no quinary oxides containing all four of the metallic elements reported, though there are 88 known ternary (two metals), and 13 quaternary (three metals) oxides¹⁷. The complexity and large unit cell size of likely new structures pose challenges to pure *ab initio* methods^{8,11,18}, stimulating the development of approaches which build chemical knowledge into the possible solutions, such as the Extended Module Materials Assembly (EMMA) method, which spans data mining and *ab initio* approaches and has been successfully applied to the prediction of complex structures¹⁹.

EMMA combines modules and stacking rules identified from the known chemistry of the studied elements to generate new feasible crystal structures. Brute force assembly and structural relaxation of all possible module combinations in search of the lowest energy structure restricts application to narrow compositional ranges with sensible available computing resource. This precludes extensive sampling of phase diagrams that the inter-relationship between structural and compositional degrees of freedom requires for identification of new materials. Here we retain the crystal chemical approach to building up complex structures, but treat the modules as interchangeable components, with Monte-Carlo (MC) driven selection and permutation to rapidly access optimum structures²⁰, rather than exhaustively searching the structure space. Within this new MC-EMMA approach, each MC permutation involves a structural modification selected from the set illustrated in Figure 1a, and calculation of the new energy by relaxation of atomic positions using classical force-fields (FFs). In

addition to changing the order of modules in the stacking sequence, MC-EMMA can vary the length of the sequence, and the choice of modules present in any given sequence. If the energy is reduced, or increased by less than a weighted random number, the new module configuration becomes the basis for further alteration, otherwise the original structure is retained (Extended Data Figure 1). MC-EMMA enables application of the extended module based approach to much larger problems, as exemplified by its ability to identify the lowest energy structure of $\text{YBa}_2\text{Ca}_2\text{Fe}_5\text{O}_{13}$ from 15 modules after 5,242 permutations taking 31 CPU days, 6 orders faster than the 1.778×10^9 possible structures in the original brute force evaluation, which by extension would take about 10^7 CPU days (Extended Data Figure 1b–f). This accelerates structure prediction sufficiently to search complex compositional spaces.

To investigate the potential of probe structure calculation with MC-EMMA to prioritize synthesis, we performed a low-resolution scan of quaternary compositions in the previously experimentally investigated Y-Ba-Ti-O phase field, which contains one reported quaternary compound. Modules were taken from $\text{ABO}_{3.5}$ cubic perovskite-derived structures (Extended Data Figure 4c) as these exhibit a range of local coordination environments found in many structural families. A maximum stacking length of $12 a_p$ (where a_p is the dimension of the cubic perovskite unit) was used to obtain probe structures at 11 quaternary compositions (Extended Data Figure 4a). For each composition we run multiple, independent MC-EMMA calculations from random start points (see Methods), relaxing up to the five lowest energy structures using density functional theory (DFT) to obtain more accurate energies. The lowest energy structure obtained at that composition is then the probe structure. The composition of the only known quaternary oxide in this phase field, $\text{YTi}_2\text{Ba}_3\text{O}_{8.5}$ ²¹ (which is not a cubic perovskite), was successfully identified as the lowest in energy (Extended Data Figure 4b, c).

The MC-EMMA probe structure approach was then applied to the selected, and previously experimentally uninvestigated, Y-Sr-Ca-Ga-O phase field. Higher compositional resolution and larger probe structures were used than in the test example above. Figure 1b shows the perovskite-derived module set selected for probe structure construction: modules with large cations of compositions Y_4 , Y_4O_4 , Ca_4O_4 , Sr_4O_4 , were alternated with Ga_4O_8 and two Ga_4O_4 modules containing the smaller Ga^{3+} cation. Using a fixed 1:1 large cation: small cation ratio and $2a_p$ supercells in the a and b directions, we investigated candidate structures with stacking sequences up to a maximum length of 20 large cation and 20 small cation modules, thereby giving a maximum cell size of $2a_p \times 2a_p \times 20a_p$. The chosen modules and stacking limit allow for 691 different compositions across the phase diagram, which we have sampled with 25 representatives (Extended Data Figure 2a). The probe structures are used to assess stability at a given composition by calculation of their energy relative to the convex hull (Figure

2a), which is constructed using the previously reported compounds in the Y-Sr-Ca-Ga-O phase field (Extended Data Figure 2b). There is a clear energy minimum of only 8 meV/atom above the hull at the composition $\text{Y}_{0.1}\text{Sr}_{0.4}\text{Ca}_{0.5}\text{Ga}_1\text{O}_{2.55}$, with similar low-energy compositions nearby in what we refer to as the low energy region of the phase field, providing a prioritization for experimental synthesis to identify unknown phases.

We explored this low energy region experimentally in 34 subsolidus reactions of the component oxides in air at 1150–1200°C in alumina crucibles (hand ground in acetone and with a single firing of 24 hours). Powder X-ray diffraction (PXRD) allowed identification of the known phases in each sample from the PDF-2 database^{22,23} together with ICSD²⁴ and CAS/Scifinder¹⁷ searches. Application of principal component analysis (PCA, see methods) to the PXRD patterns placed the patterns of 29 samples into three clusters where the majority of Bragg reflections are not accounted for by known materials (Figure 2b, Extended Data Figure 3), and left five unclustered patterns dominated by known compounds. The 29 clustered patterns belonged to samples with compositions around the computed low energy region, demonstrating that this region contains new phases, consistent with the probe structure-based computational predictions that initially guided the syntheses (Figure 2c). Energy Dispersive X-ray Spectroscopy (EDX) analysis of the subsolidus powder sample of the composition with the highest number and relative intensity of such unidentified reflections (nominal; $\text{Y}_{0.015}\text{Sr}_{0.383}\text{Ca}_{0.603}\text{GaO}_{2.508}$, composition 1; Extended Data Figure 3), suggested two distinct phases were present (Figure 2c), with an average EDX measured composition of $\text{Y}_{0.03(1)}\text{Sr}_{0.34(5)}\text{Ca}_{0.64(5)}\text{Ga}_{0.9(2)}\text{O}_x$ from 45 particles. Crystal growth above the solidus at 1300°C for 3 hours followed by cooling at 0.1°C/min to 800°C and then air cooling to room temperature in a Pt crucible afforded two types of crystals which were readily separable based on their distinct habits and colours (Figure 2d). The unit cells determined for each crystal type by single crystal X-ray diffraction (SXRD) revealed that the new structures of these two phases accounted for the majority of PXRD reflections in composition 1 not assignable to known phases.

The structure of phase I $\text{Y}_{0.07(3)}\text{Sr}_{1.01(5)}\text{Ca}_{1.55(5)}\text{Ga}_{2.87(5)}\text{O}_{7.0(2)}$ (determined by EDX, with O content calculated assuming charge neutrality) was solved and refined from the SXRD data with a refined composition of $(\text{Y,Sr})_{1.02(1)}\text{Ca}_{1.48(1)}\text{Ga}_3\text{O}_7$ (Y^{3+} and Sr^{2+} are iso-electronic and therefore indistinguishable by X-ray diffraction). The structure of I (Supplementary Information) is based on the pentagonal **mcm** net of three- and four-connected corner-sharing TO_4 tetrahedra found in $\text{A}_2\text{T}_3\text{O}_7$ melilite²⁵ (Figure 3, Extended Data Figure 5), where layers of larger eight-coordinate A cations are located above and below the centres of the pentagonal rings: $\text{LaSrGa}_3\text{O}_7$ adopts this structure with $\text{T} = \text{Ga}$, $\text{A}_2 = \text{LaSr}$ ²⁶. I retains this GaO_4 connectivity by forming three distinct A cation sites that accommodate the distinct

chemistries of Ca^{2+} , Sr^{2+} and Y^{3+} . This arises from restructuring of the layers to form zig-zag chains of two types of pentagon that alternate along b (Figure 3b). The more regular pentagons accommodate eight-coordinate Sr^{2+} (Figure 3b, 3d, Extended Data Figure 5b). The more distorted pentagons define two distinct sites, Ca1 and Ca2, containing three Ca^{2+} within the same 8 Å intralayer translation containing two Sr^{2+} (Figure 3b), giving I an A cation-rich $\text{A}_{2.5}\text{T}_3\text{O}_7$ composition. The Ca1 site is a centrosymmetric octahedron, described by one terminal and two bridging oxides from each T_3O_7 layer (Figure 3e, Extended Data Figure 5d). The two five coordinate Ca2 sites (Figure 3f) alternate with the octahedral Ca1 site above and below the Ga_3O_7 layer along the distorted pentagon chain (Figure 3g).

The structure of II (Figure 4) was determined to be a 64-fold $4a_p \times 4a_p \times 4a_p$ cation- and anion-defect $\text{A}_{64}\text{B}_{56}\text{O}_{144}$ superstructure of ABO_{3-6} perovskite by combined neutron and X-ray powder diffraction on a ceramic sample of composition 2 (nominal; $\text{Y}_{0.038}\text{Sr}_{0.320}\text{Ca}_{0.848}\text{Ga}_{0.794}\text{O}_{2.416}$; Extended Data Figure 6, Methods and Supplementary Information). One of the two distinct B-site cation layers (Figure 4b) has the defect-free $\text{B}_{16}\text{O}_{32}$ composition, with GaO_6 octahedra (Figure 4e), orientationally disordered GaO_4 tetrahedra (Figure 4f-g) and SrO_7 units (Figure 4h) arranged in a “double checkerboard” pattern: Sr^{2+} and Ga^{3+} sites alternate, and the Ga^{3+} sites themselves alternate between octahedral and tetrahedral. The anion defect $\text{A}_{16}\text{O}_{10}$ layer (Figure 4c, Extended Data Figure 7) is a checkerboard alternation of a rectangle of four six-coordinate Ca^{2+} sites (Figure 4i) surrounding an oxide in a fragment of a perfect AO layer, and a larger anion vacancy-centred rectangle of four site-disordered trigonal bipyramidal CaO_5 / tetrahedral GaO_4 units (Figure 4j-k). Both these coordination environments arise from a highly unusual structural feature of the second B-site cation layer: 25% of the B-site cations are missing, with this site now occupied by O^{2-} anions (Figure 4d, l) that alternate with SrO_7 polyhedra at the centres of squares of GaO_4 tetrahedra. Anion-anion repulsion is minimised by rotation of the neighbouring tetrahedra out of the plane and by the A-site layer anion vacancy locations (Extended Data Figure 7). Four of the A-site neighbour cations are displaced away from this B-site oxide to form the Ca_4O squares, leaving it connected to a tetrahedral cluster of GaO_4 and CaO_5 units (Figure 4e). The resulting CaO_5 geometry is uncommon, although not unknown²⁷: both the new structures found here have five-coordinate Ca^{2+} in different geometries. The refined composition of $(\text{Y},\text{Sr})_{0.903(6)}\text{Ca}_{2.097(6)}\text{Ga}_2\text{O}_6$ is charge-neutral within error and agrees with the EDX measured composition of $\text{Y}_{0.09(3)}\text{Sr}_{0.85(3)}\text{Ca}_{2.12(8)}\text{Ga}_{1.93(7)}\text{O}_{6.0(3)}$.

I and II adopt a range of compositions demonstrated by unit cell size variation (Extended Data Figure 8): phase purity is favoured by incorporation of Y, though both structures can be seen as minor components in multiphase assemblages resulting from Y-free syntheses. Geometry-relaxed DFT calculations on ordered approximant structures to I ($\text{Sr}_2\text{Ca}_3\text{Ga}_6\text{O}_{14}$) and II ($\text{SrCa}_2\text{Ga}_2\text{O}_6$; Extended Data

Figure 9) show that the two new phases are both 5 meV/atom below the previously computed convex hull, and form points on a new computed convex hull in this composition space (Extended Data Figure 3). The lowest energy probe structure is located at 12 meV/atom above this new hull, clearly close enough in energy to the actual phases to identify the correct region for experimental synthesis of new materials by comparing energies at these compositions with energies of mixtures of known phases and suggesting that either the probe structures themselves or even more favourable phases (in this case **I** and **II**) will form. Both **I** and **II** melt congruently despite their compositional complexity, reflecting the energetic stability used to target their synthesis. The lowest energy probe structure composition at which synthesis was attempted but yielded a sample with an unclustered PXRD pattern dominated by previously reported phases, $\text{Y}_{0.1}\text{Sr}_{0.8}\text{Ca}_{0.1}\text{Ga}_1\text{O}_{2.55}$, is 18 meV/atom above the minimum energy probe structure, giving an indication of the energy window within which the new phases were identified.

Using MC-EMMA with a different choice of modules based on the melilite structure allows us to change the ratio of large : small cations to 1:1.2, yet still targets synthesis in the correct region of phase space (Extended Data Figure 4d–f). Although melilite is structurally closer to **I** than perovskite, it has less overall diversity in the coordination environments present in the MC-EMMA generated structures (Extended Data Figure 10). For example, only 3% of the melilite-based structures have six-coordinate gallium, whereas the perovskite affords nearly equal frequency of four-, five- and six-coordinate (Extended Data Figure 10d). The greater structural diversity obtained with perovskite-based modules makes these a better choice to capture the range of possible chemistry occurring across the whole composition space.

The probe structures used to identify the low-energy region where experimental work was focussed contain five-coordinate calcium sites (Extended Data Figure 10) also present in both the new experimentally isolated structures. Such environments are not present in the modules used to construct the precursors to the probe structures, but emerge from subsequent structural relaxation, demonstrating that the MC-EMMA probe structure approach captures the essential chemistry of a complex compositional space with the precision to guide synthesis towards entirely new structures. In principle, any structure prediction method with sufficient performance and flexibility to generate probe structures with appropriate size and complexity could be used successfully within this approach. The resulting materials open up structural and functional possibilities: **I** is related to melilite, which is one of the best oxide ion conductor families¹⁵, and **II** is a structurally diverse perovskite analogue with demonstrated non-stoichiometry and a range of coordination environments appropriate for substitution by elements from across the periodic table. For example, the observation of the Y-

containing quinary **II** opens up the introduction of lanthanides to explore optical and magnetic properties: substitution of Eu^{3+} for Y^{3+} in **II** introduces the properties necessary for function as a red down-conversion phosphor (Extended Data Figure 3f-g). As the absolute reliability of the predictions increases such complex structures may be computationally accessible directly, but, given the power of modern structure solution tools²⁸, it is the focusing of synthesis into regions where new materials are likely to exist that is the key step demonstrated in Figure 2c. This integrated approach will enable exploration of the vast un-investigated compositional spaces¹ in an efficient manner by targeting those regions containing undiscovered phases and structures.

Figures:

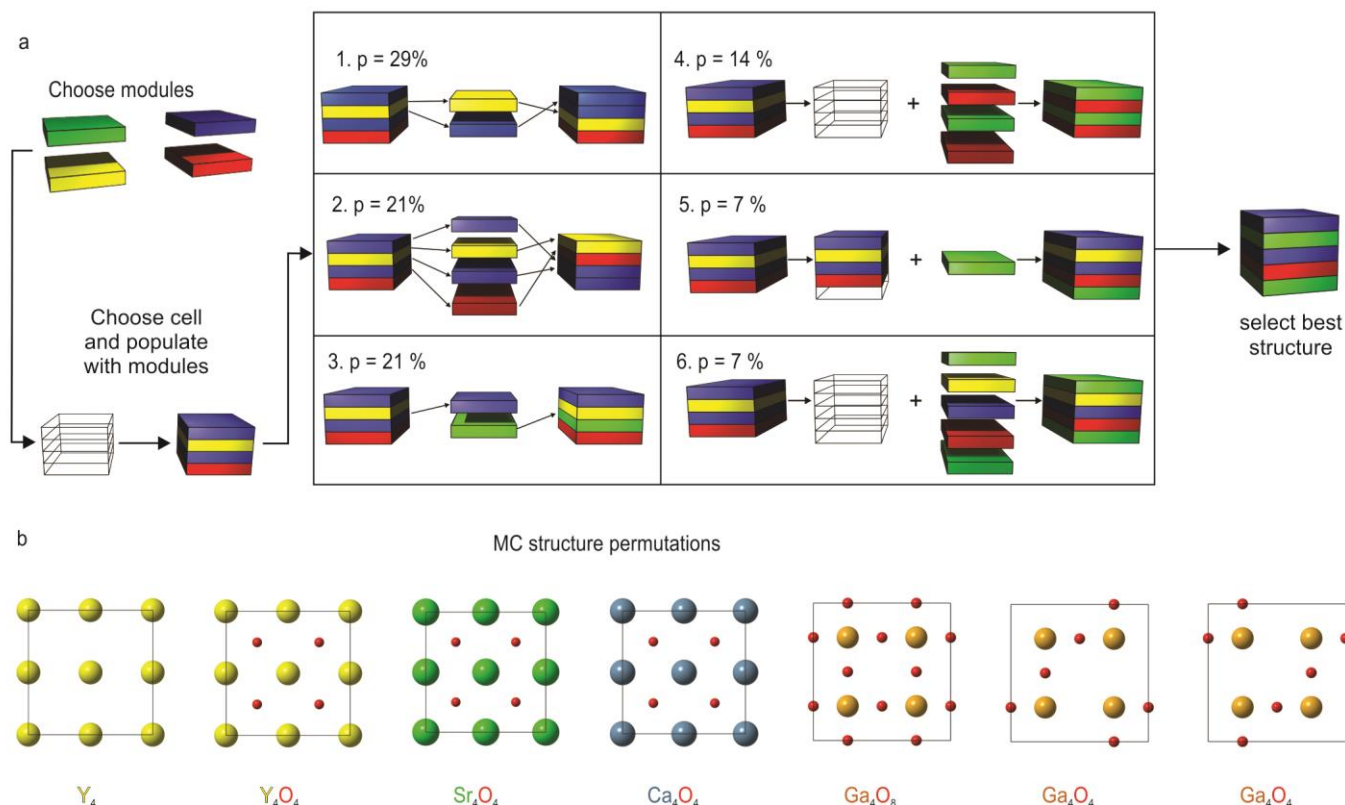


Figure 1. **The MC-EMMA method.** **a** Schematic representation of the MC-EMMA method: The user input is the selection of modules and the composition of the system. The routine then generates a starting structure by selecting an empty starting unit cell, randomly populating it with modules that give the composition under study. Structures are then permuted from the starting structure using a MC routine with six allowed permutations (selected at random but with the (most used) weights, p , indicated (rounded to the nearest whole number, Methods)): **1.** the position of two randomly chosen modules is switched, **2.** the module sequence in the current structure is re-randomised, **3.** one randomly selected module is swapped for a module of a different type, **4.** a new random structure is generated maintaining the current unit cell size, **5.** the size of the current unit cell is altered, adding in new modules if the size is increased (additional modules are chosen at random from the full set but maintaining the overall composition), **6.** a new unit cell is generated and populated with a new set of modules. After each permutation the new structure is relaxed using FFs, with the best structures re-calculated using DFT at the end of the MC routine. The structure with the lowest energy per atom is then selected as the probe structure at that composition. **b** The modules used: Y_4 , Y_4O_4 , Sr_4O_4 and Ca_4O_4 modules contain the larger cations and alternate with the small cation Ga_4O_x modules.

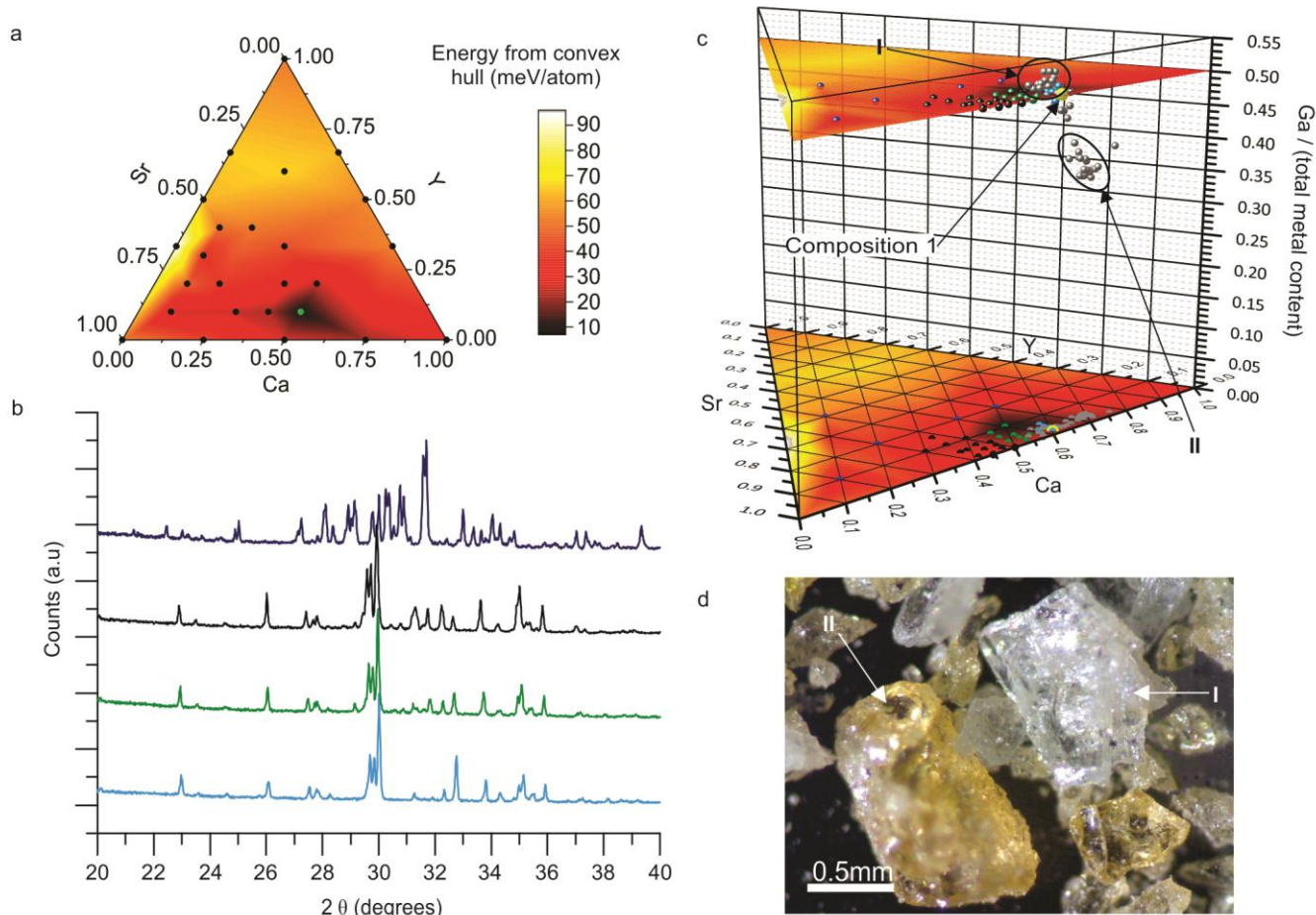


Figure 2. From computation to materials isolation. **a** The energy of the probe structures calculated using MC-EMMA are plotted relative to the convex hull created with previously reported phases (Extended Data Figure 2) on the $\text{Ga} = 0.5$ plane of the Y_2O_3 - SrO - CaO - Ga_2O_3 quaternary diagram, the lowest energy composition is highlighted in green. **b** The PXRD patterns from the 34 initial samples were clustered using principle component analysis. The clustered patterns, shown in black, green, and cyan, are dominated by reflections which do not index to previously reported phases. Composition 1 (cyan PXRD pattern and cluster) has the fewest reflections belonging to previously reported phases. The PXRD patterns which could not be clustered (dark blue) index to a mixture of previously reported phases (Extended Data Figure 3). **c** The nominal compositions of the initial 34 samples (coloured by PXRD pattern cluster as in **b**, composition 1 in yellow); the EDX derived compositions arising from the sample with composition 1 (grey); and the calculated energies represented in **a** shown in the $\text{Ga} = 0.5$ plane. The triangular cross-section of the plot shows the normalised $\text{Y}:\text{Sr}:\text{Ca}$ content, with the height showing the fraction $\text{Ga} / (\text{total metal content})$. The circled groups of observed compositions indicate the presence of two new phases **I** and **II** above and below the $\text{Ga} = 0.5$ plane. All of the data are also projected onto the base of the prism. Samples with PXRD patterns clustered by PCA, along with the EDX determined compositions of phases **I** and **II**, lie within the calculated low energy region shown with dark shading.

d Crystals obtained from the melt sample of composition 1, clearly indicating the presence of two types of crystal based upon colour, with each type found up to $\sim 1 \text{ mm}^3$ in size.

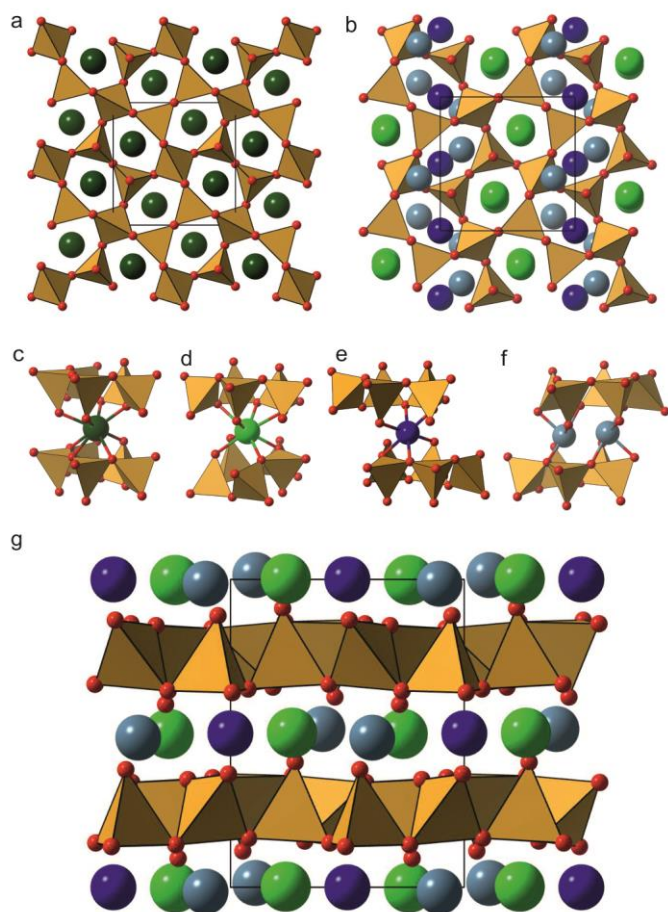


Figure 3. **Structure of Phase I.** **a** $\text{LaSrGa}_3\text{O}_7$ melilite²⁶ viewed perpendicular to the corner-sharing tetrahedral layers; La/Sr(green), Ga (brown) and O (red). **b** A single tetrahedral layer of I and the two adjacent A-site layers. Sr (green), Ca1 (dark blue), Ca2 (light blue), Ga(brown) and O(red). Sr and Ca are located between the chains of more and less regular pentagons respectively. **c** La/Sr site in melilite defined above and below by equivalent pentagons of tetrahedra. **d** Sr site in I - the pentagons are no longer in the same orientation above and below the site. **e** The Ca1 site in I is no longer at the centre of two pentagons, with the co-ordination reduced to six oxygens. **f** Pair of five-coordinate Ca2 sites in I, displaced from the centroids of the neighbouring pentagons. **g** View of I perpendicular to the layers, showing

the alternation between Ca1 and Ca2 on either side of pentagonal rings within the tetrahedral layer. See Extended Data Figure 5 for further details

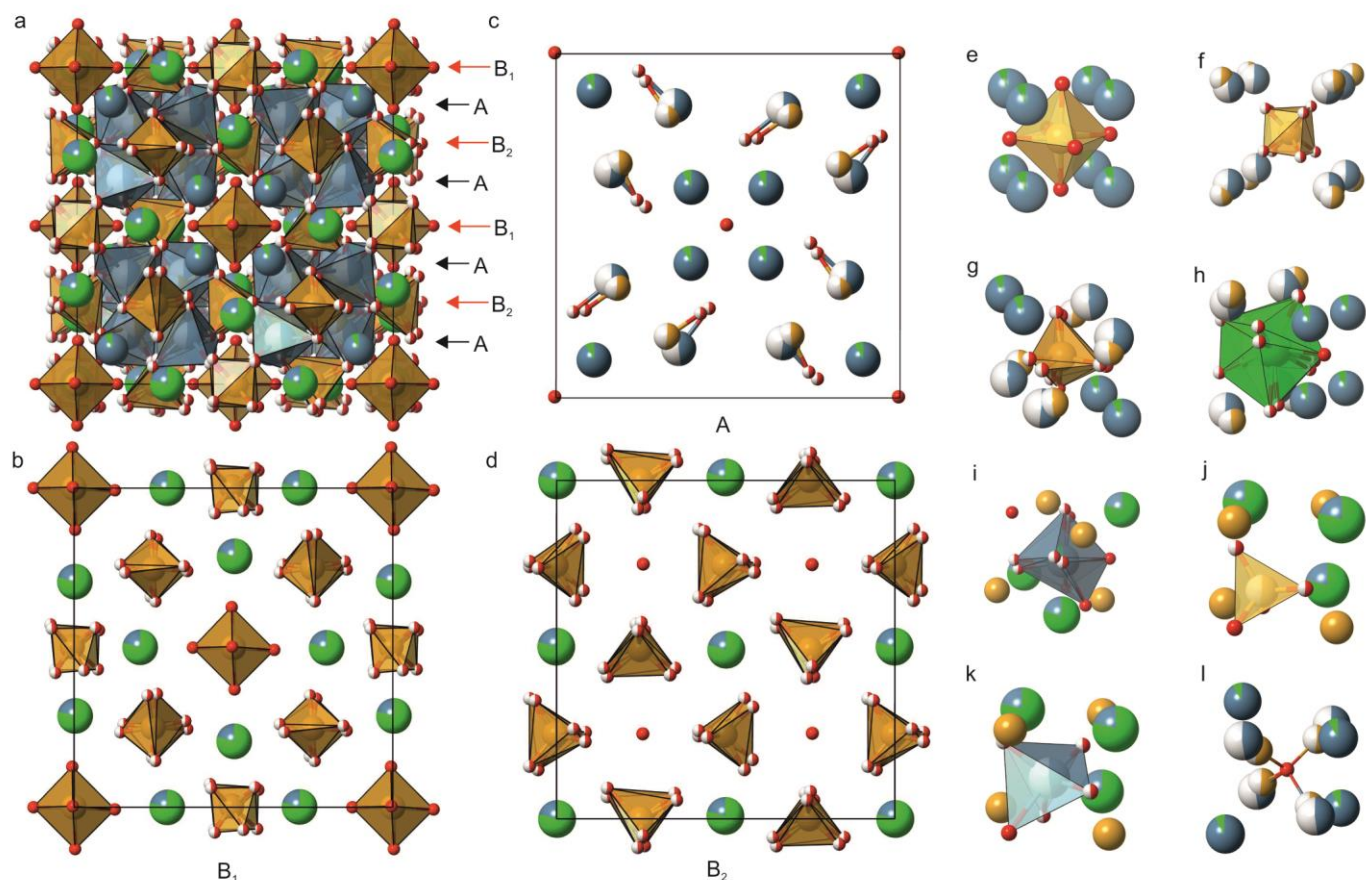


Figure 4: Structure of Phase II. **a** Unit cell of **II** (viewed along the $[100]$ direction). **b** The B_1 non-defective $B_{16}O_{32}$ layer. **c** The anion-deficient $A_{16}O_{10}$ A-site layer. **d** The B_2 cation- and anion-deficient $B_{12}O_{20}$ layer. **b-d** are viewed along $[001]$. **e-k** Cation environments in **II** with nearest neighbour cation sites shown. **e** B_1 layer GaO_6 surrounded by eight Ca sites, the relationship to the $A_{16}O_{10}$ layer is set out in Extended Data Figure 7b. **f** B_1 layer disordered GaO_4 tetrahedra neighbouring eight mixed Ca/Ga sites. **g** B_2 layer disordered GaO_4 tetrahedra neighbouring four Ca and four Ca/Ga sites. **h** B_1 layer SrO_7 environment neighbouring four Ca and four Ca/Ga sites, the site in the B_2 layer is similar. **i** A layer disordered CaO_6 site neighbouring one oxide anion on a B-site vacancy in the B_2 layer, and three Sr and four Ga sites. **j** and **k** respectively represent the GaO_4 and CaO_5 environments disordered on the same site in the A layer, neighbouring one oxide anion on a B-site vacancy, three Sr and four Ga sites. **l** The cation environments around the oxide anion on a B-site vacancy; four CaO_6 sites (**i**) and four mixed Ca/Ga sites (**j-k**). Atoms coloured as follows: Ga (brown), Sr (green), Ca (light blue) and O (red).

Online Content: Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

References

- 1 Jansen, M. A concept for synthesis planning in solid-state chemistry. *Angew Chem Int Ed* **41**, 3747-3766 doi:10.1002/1521-3773(20021018)41:20<3746::AID-ANIE3746>3.0.CO;2-2 (2002).
- 2 Siegrist, T. & Vanderah, T. A. Combining magnets and dielectrics: Crystal chemistry in the BaO-Fe₂O₃-TiO₂ system. *Eur J Inorg Chem*, 1483-1501 doi:10.1002/ejic.200390192 (2003).
- 3 Vanderah, T. A., Loezos, J. M. & Roth, R. S. Magnetic dielectric oxides: Subsolidus phase relations in the BaO:Fe₂O₃:TiO₂ system. *J Solid State Chem* **121**, 38-50, doi:10.1006/jssc.1996.0006 (1996).
- 4 Woodley, S. M. & Catlow, R. Crystal structure prediction from first principles. *Nat Mater* **7**, 937-946, doi:10.1038/nmat2321 (2008).
- 5 Lyakhov, A. O., Oganov, A. R., Stokes, H. T. & Zhu, Q. New developments in evolutionary structure prediction algorithm USPEX. *Comput Phys Commun* **184**, 1172-1182, doi:10.1016/j.cpc.2012.12.009 (2013).
- 6 Wang, Y. C. *et al.* Materials discovery via CALYPSO methodology. *J Phys: Condens Matter* **27**, 203203 doi:10.1088/0953-8984/27/20/203203 (2015).
- 7 Hautier, G., Fischer, C., Ehrlacher, V., Jain, A. & Ceder, G. Data Mined Ionic Substitutions for the Discovery of New Compounds. *Inorg Chem* **50**, 656-663, doi:10.1021/ic102031h (2011).
- 8 Schön, J. C. How can Databases assist with the Prediction of Chemical Compounds? *Z Anorg Allg Chem* **640**, 2717-2726, doi:10.1002/zaac.201400374 (2014).
- 9 Gautier, R. *et al.* Prediction and accelerated laboratory discovery of previously unknown 18-electron ABX compounds. *Nat Chem* **7**, 308-316, doi:10.1038/Nchem.2207 (2015).
- 10 Chen, H. L., Hautier, G. & Ceder, G. Synthesis, Computed Stability, and Crystal Structure of a New Family of Inorganic Compounds: Carbonophosphates. *J Am Chem Soc* **134**, 19619-19627, doi:10.1021/ja3040834 (2012).
- 11 Cerqueira, T. F. T. *et al.* Identification of Novel Cu, Ag, and Au Ternary Oxides from Global Structural Prediction. *Chem Mater* **27**, 4562-4573, doi:10.1021/acs.chemmater.5b00716 (2015).
- 12 Morris, A. J., Grey, C. P., Needs, R. J. & Pickard, C. J. Energetics of hydrogen/lithium complexes in silicon analyzed using the Maxwell construction. *Phys Rev B* **84**, 224106, doi:10.1103/PhysRevB.84.224106 (2011).
- 13 Ishihara, T., Matsuda, H. & Takita, Y. Doped LaGaO₃ perovskite type oxide as a new oxide ionic conductor. *J Am Chem Soc* **116**, 3801-3803, doi:10.1021/ja00088a016 (1994).
- 14 Feng, M. & Goodenough, J. B. A superior oxide-ion electrolyte. *Eur J Solid State Inorg Chem* **31**, 663-672 (1994).
- 15 Kuang, X. *et al.* Interstitial oxide ion conductivity in the layered tetrahedral network melilite structure. *Nat Mater* **7**, 498-504, doi:10.1038/nmat2201 (2008).
- 16 Antic-Fidancev, E., Hölsä, J., Lastusaari, M. & Lupei, A. Dopant-host relationships in rare-earth oxides and garnets doped with trivalent rare-earth ions. *Phys Rev B* **64**, 195108, doi:10.1103/PhysRevB.64.195108 (2001).
- 17 Scifinder (Chemical Abstracts Service, Columbus, OH, 2015).
- 18 Zhu, Q., Oganov, A. R., Glass, C. W. & Stokes, H. T. Constrained evolutionary algorithm for structure prediction of molecular crystals: methodology and applications. *Acta Crystallogr, Sect B* **68**, 215-226, doi:10.1107/S0108768112017466 (2012).
- 19 Dyer, M. S. *et al.* Computationally Assisted Identification of Functional Inorganic Materials. *Science* **340**, 847-852, doi:10.1126/science.1226558 (2013).
- 20 Wales, D. J. & Scheraga, H. A. Review: Chemistry - Global optimization of clusters, crystals, and biomolecules. *Science* **285**, 1368-1372, doi:10.1126/science.285.5432.1368 (1999).
- 21 Kuang, X., Jing, X., Loong, C.-K., Lachowski, E. E., Skakle, J. M. S. & West, A. R. A new hexagonal 12-layer perovskite-related structure: Ba₆R₂Ti₄O₁₇ (R = Nd and Y). *Chem Mater* **14**, 4359-4363, doi:10.1021/cm020374m (2002).
- 22 X'Pert HighScore Plus. 2.2a ed (PANalytical B. V A, The Netherlands, 2006).
- 23 Powder Diffraction File PDF-2 (International Center for Diffraction Data P, U. S. A., 2007).
- 24 ICSD Web (FIZ Karlsruhe, Germany, 2015).

- 25 O' Keeffe, M. & Hyde, B. G. Plane Nets in Crystal-Chemistry. *Philos Trans R Soc, A* **295**, 553-618, doi:10.1098/rsta.1980.0150 (1980).
- 26 Skakle, J. M. S. & Herd, R. Crystal chemistry of (RE,A)₂M₃O₇ compounds (RE=Y, lanthanide; A=Ba, Sr, Ca; M=Al, Ga). *Powder Diffr* **14**, 195-202, doi:10.1017/S0885715600010526 (1999).
- 27 Deiseroth, H. J. & Müller-Buschbaum, H. Über Erdalkalimetall-Oxogallate. I. Zur Kenntnis von CaGa₄O₇. *Z Anorg Allg Chem* **382**, 149-156, doi:10.1002/zaac.19713820206 (1971).
- 28 Cheetham, A. K. & Goodwin, A. L. Crystallography with powders comment. *Nat Mater* **13**, 760-762, doi:10.1038/nmat4044 (2014).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Extended materials: are available in the online version of the paper.

Acknowledgements: We thank EPSRC for funding under EP/N004884. MJR is a Royal Society Research Professor. We thank the Materials Chemistry Consortium (EPSRC: EP/L000202) and the PRACE-3IP project (FP7 RI-312763) for provision of computer time on the ARCHER UK National Supercomputing Service, and the N8 consortium (EPSRC: EP/K000225/1) and the University of Liverpool for access to local high performance computing resources. We thank ISIS and Diamond Light Source for provision of beamtime and Chiu Tang, Claire Murray, and Dominic Fortes for assistance on the I11 and HRPD instruments. We acknowledge use of Hartree Centre resources in this work. The STFC Hartree Centre is a research collaboratory in association with IBM providing High Performance Computing platforms funded by the UK's investment in e-Infrastructure. We thank Professor R. Seshadri and Dr. T. A. Strom (Materials Research Laboratory and Department of Materials, UC Santa Barbara) for measurement of the phosphor quantum yield. Authors are grateful to Karl Dawson, from the Centre for Materials and Structures at the University of Liverpool, for the sample preparation with the Focussed Ion Beam (FIB).

Author contributions: The MC-EMMA approach was defined by CC, MSD, MJR, JBC and GRD. The code was written by CC and MSD. CC performed all synthetic work and powder diffraction characterisation in the initial identification of **I** and **II**, with subsequent substitution synthesis and characterisation by PM. GFSW solved the single crystal structure, MJP and JBC advised CC on the Rietveld refinement. MZ performed electron microscopy and measured excitation and emission spectra. JBC supervised the crystallographic work. MJR directed the project. CC, MSD and MJR wrote the first draft. All authors contributed to development of the manuscript and to discussion as the project developed.

Author information: Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to rossein@liverpool.ac.uk.

Methods

The MC-EMMA code

The description of complex solid state structures is enabled by considering them as a combination of structural units such as blocks, rods or layers derived as fragments of archetypal structures, which we have called modules¹⁹. MC-EMMA assembles crystal structures from modules (layered structural units extended in two dimensions) identified in known crystal structure motifs containing the elements of interest, using stacking rules appropriate to the generic structure type (*e.g.*, perovskite) chosen for construction (exemplified in Extended Data Figure 1b-g). Monte Carlo-driven selection is used to evolve structures, rather than exhaustively building and testing all possible structures (Figure 1a, Extended Data Figure 1a). At a given composition, MC-EMMA enumerates all possible module sets and stacking lengths (up to a maximum, user-specified length of $2N_{\max}$ modules, a value that will largely be determined by availability of computing resources), giving structures with the correct ratio of elements (for example for $\text{Ca}_2\text{Ga}_2\text{O}_5$ with $N_{\max} = 2$, using the modules shown in Figure 1b, the two possible module sets would be; $[2 \times \text{Ca}_4\text{O}_4 + \text{Ga}_4\text{O}_8 + \text{Ga}_4\text{O}_4$ (option 1)] and $[2 \times \text{Ca}_4\text{O}_4 + \text{Ga}_4\text{O}_8 + \text{Ga}_4\text{O}_4$ (option 2)]). The minimum possible stacking length is chosen and populated with a random choice of modules with the correct composition, to give an initial structure for the MC search.

For a given composition, N_{\max} separate MC-EMMA runs are performed in parallel, starting from randomly generated initial structures. The geometries of these are optimized and the energies calculated using classical force-fields (FFs, see below). The starting structures (*i.e.*, before geometry optimization) are then modified using a weighted random choice of one of the permutations shown in Figure 1a. The most common set of weights used in the calculations (used in two thirds of cases) is shown in Figure 1a. These weights are tuneable by the user and the best choice is likely system dependent, although we have not yet attempted a systematic optimisation. Generally we give larger weights to the module swapping permutations 1-3 which have a smaller effect on the structure. The more significant structural permutations 4-6 are necessary to hop out of local energy minima and to grow the structure. We increase the weight of these when there is poor convergence of an MC-EMMA run. For example, using perovskite probe structures for Y-Sr-Ca-Ga-O we adjusted the weights for 18 compositions to 1: 17 %, 2: 25 %, 3: 25 %, 4: 17 %, 5: 8 % and 6: 8 %, increasing the number of more significant structural permutations so that we explored larger regions of the configuration space than obtained with a higher frequency of single module swaps (permutation 1).

If a generated structure is the same as one previously calculated, it is discarded and a different structure is generated. If not, the geometry of the new candidate structure is optimized by relaxation of all atomic positions and its energy calculated using FFs. If the relaxed energy of the new candidate structure is lower than that of the previous structure, then the candidate structure is accepted and becomes the basis for further permutations. Candidate structures in which the energy is higher are accepted with a probability of $\exp(-\Delta E/\Theta)$, where ΔE is the energy difference in eV/atom, and Θ a reduced temperature in eV. Structures in which the geometry optimization did not converge are automatically rejected.

The MC reduced temperature parameter was set to be equal to 0.1 eV – however, if a MC run became unstable, the temperature parameter was reduced to 0.01 eV.

Structural space was explored in this way until the number of consecutive permutations rejected was equal to the square of the maximum stacking length, N_{max}^2 , as long as the number of permutations exceeded $2N_{max}^2$. The overall MC-EMMA process is illustrated in Extended Data Figure 1a. Two break points exist in the process. The break after the permutation step ends the calculation if upon attempting structural permutations no new structures can be found, and the break after the maximum number of rejections has been exceeded.

Construction of probe structures with MC-EMMA

The N_{max} lowest energy structures produced at a composition by MC-EMMA were compared, and unique structures identified. The geometries of up to five structures ranked as lowest in energy by FFs were then optimised using more accurate DFT calculations, and the structures were re-ranked according to their DFT energies. If less than five unique structures were produced by the MC-EMMA runs, then all structures were recalculated with DFT. The structure with the lowest DFT energy was then used as the probe structure to assess the likelihood of a new compound existing at this composition (see “Comparison to convex hull”).

Test identification of the known quaternary oxide composition in the Y-Ba-Ti-O phase field with probe structure calculation

We have tested the probe structure approach by generating probe structures for 11 quaternary compositions across the experimentally explored Y-Ba-Ti-O phase field, to see if the composition of the only known quaternary, $YTi_2Ba_3O_{8.5}^{21}$, emerges as a low-energy composition. We used cubic perovskite-derived modules with the in-plane cell dimensions of $2a_p \times 2a_p$ (where a_p is the dimension of the standard cubic perovskite unit, Extended Data Figure 4a), and used one module set ($2 \times Ti_4O_4$, Ti_4O_8 , Y_4 , Y_4O_4 , Y_4O_8 , Ba_4O_4 , $2 \times Y_2Ba_2O_8$). We did not impose an alternation of module types by cation size, so there is no restriction on which cation modules can be adjacent to each other, allowing us to freely choose the ratios of the cations. The maximum stacking length, $2N_{max}$, was chosen to be 24, (maximum cell length of $\sim 12a_p$), giving a minimum of 576 structures at each composition with energies computed by FFs. Up to five of the lowest energy structures were re-ranked using DFT at each composition. Energies were calculated at $T = 1498$ K (see “Comparison to convex hull” below). The energies of the 11 quaternary compositions are plotted in Extended Data Figure 4b-c, the lowest of these (42 meV/atom above the hull) corresponds to the known quaternary composition.

Perovskite probe structures in the Y-Sr-Ca-Ga-O phase field

In the detailed survey of the experimentally unexplored Y-Sr-Ca-Ga-O phase field, perovskite-related modules and stacking rules (Figure 1b) were chosen as this offers a broad range of local structural chemistry (for example, the oxygen vacancy fragments allow coordination numbers of less than six for gallium) in the resulting computed probe structures (example probe structures are shown in Extended Data 10). Other modules and stacking rules could be followed, as exemplified by melilite modules in the next section. Four modules with larger cations (Y_4 , Y_4O_4 , Sr_4O_4 and Ca_4O_4) and three modules with smaller cations ($2 \times Ga_4O_4$ and Ga_4O_8) were chosen (Figure 1b), based upon the known structural motifs found in perovskites. Each module had in-plane cell dimensions of $2a_p \times 2a_p$ to allow for structural flexibility within the module. A stacking rule of alternating larger cation modules and smaller cation modules was imposed. N_{max} was set to 20 larger cation and 20 smaller cation modules, resulting in a maximum cell length of $20a_p$. Probe structures were compared to the convex hull constructed from the previously reported phases in the phase field, with $T = 1498$ K (see “Comparison to convex hull below”). The relative energy of the probe structures with respect to the convex hull was then calculated and is plotted in Figure 2a.

Melilite probe structures in the Y-Sr-Ca-Ga-O phase field

The probe structure approach does not depend upon the use of a 1:1 ratio of larger cations to smaller cations (as demonstrated in the Y-Ba-Ti-O example above) or perovskite-derived modules. We explored the Y-Sr-Ca-Ga-O phase field using modules derived from the structure of melilites, choosing a 1:1.2 ratio of larger cations (Y, Sr, Ca) to smaller Ga cations (this is the ratio found in I). The modules with smaller cations were based upon the tetrahedral net of melilite^{25,26}, with increasing levels of interstitial oxygen in the centre of the pentagonal rings, to allow for the varied charge from the larger cation modules (Extended Data Figure 4d). The modules containing larger cations have 4 cations in sites occupied in $LaSrGa_3O_7$, augmented with a fifth cation at the module origin to give the ratio of 1:1.2 large to small cations (compared to 1:1.5 in $LaSrGa_3O_7$). With varying decorations of the different larger cations and different translational offsets, we generated 156 large cation modules (Extended Data Figure 4e). The same stacking rule used for the perovskite derived modules was imposed, with larger cation modules and smaller cation modules alternating in the stacking direction. These modules are 4 times larger in a and b than in the perovskite case, so we reduced N_{max} to 4. A total of 21 probe structures were constructed using MC-EMMA as described above, with the exception of increasing the minimum number of structures sampled to 576 (rather than $2N_{max}^2$) and running 12 parallel MC-EMMA runs rather than N_{max} . Energies were compared to the convex hull at 0 K (see “Comparison to convex hull” below). The lowest energy quinary probe structure in the calculated phase field (Extended Data Figure 4f) has the composition $Y_{0.125}Sr_{1.1875}Ca_{1.1875}Ga_3O_{7.0625}$, and is 32 meV/atom above the convex hull. It has a similar Y : Sr : Ca ratio to that in the low energy region around $Y_{0.1}Sr_{0.4}Ca_{0.5}GaO_{2.55}$ found using perovskite probe structures, but with a different ratio of larger to smaller cations.

Comparison to convex hull

The stability of the probe structure was calculated relative to the convex hull constructed from compounds reported on ICSD with crystal structures from which ordered DFT models could be assembled. For the Y-Sr-Ca-Ga-O phase field: Y_2O_3 ²⁹, Ga_2O_3 ³⁰, CaO ³¹, SrO ³¹, $\text{Ca}_2\text{Ga}_2\text{O}_5$ ³², $\text{Ca}_5\text{Ga}_6\text{O}_{14}$ ³³, CaGa_2O_4 ³⁴, CaGa_4O_7 ,²⁷ $\text{Sr}_{10}\text{Ga}_6\text{O}_{19}$ ³⁵, $\text{Sr}_3\text{Ga}_2\text{O}_6$ ³⁵, $\text{Sr}_3\text{Ga}_4\text{O}_9$ ³⁶, $\text{Sr}_4\text{Ga}_2\text{O}_7$ ³⁷, SrGa_2O_4 ³⁸, $\text{Y}_3\text{Ga}_5\text{O}_{12}$ ³⁹ and YGaO_3 ⁴⁰. For the Y-Ba-Ti-O phase field the reported reference compounds used were: $\text{Ba}_2\text{Ti}_9\text{O}_{20}$ ⁴¹, Ba_2TiO_4 ⁴², $\text{Ba}_4\text{Ti}_{13}\text{O}_{30}$ ⁴³, $\text{Ba}_6\text{Ti}_{17}\text{O}_{40}$ ⁴⁴, BaO ⁴⁵, BaTi_2O_5 ⁴⁶, BaTi_4O_9 ⁴⁷, $\text{BaTi}_5\text{O}_{11}$ ⁴⁸, BaTiO_3 ⁴⁹, TiO_2 ⁵⁰, Y_2BaO_4 ⁵¹, Y_2O_3 ²⁹, $\text{Y}_2\text{Ti}_2\text{O}_7$ ⁵², Y_2TiO_5 ⁵³, $\text{Y}_4\text{Ba}_3\text{O}_9$ ⁵⁴. DFT energies were calculated for each of these reference energies and then the convex hull constructed using the pymatgen package⁵⁵.

In the original EMMA work¹⁹, the $16a_p$ material identified is metastable at room temperature and stabilized by entropy at the synthesis temperature. To allow for this possibility, the energy of the probe structure was adjusted for finite temperature effects by including the contribution from configurational entropy at temperature, T . The configurational entropy, TS_{tot} was calculated from:

$$S_{tot} = \sum_j S_j$$

Where j indicates the partially occupied site and S_j is the configurational entropy for a given site defined by:

$$S_j = -k_B \left[\frac{n}{Z} \right] \times \sum_i f_i \ln(f_i)$$

Where k_B , n , Z and f_i are the Boltzmann constant, number of atoms in the same site per cell, the number of formula units per unit cell and the fractional occupancy of the site by species i respectively. Species were determined to be on equivalent sites based upon the number of each type of next-neighbour cation polyhedra.

Calculations on ordered approximant gallate structures

Once the structures of the two new gallate phases **I** and **II** had been determined, representative model structures were constructed so that their stabilities could be compared with previously reported compounds. The model structures were built with Sr, Ca, Ga and O, and without Y. This is reasonable since both phases are observed in low phase fractions in Y-free experimental samples (Extended Data Figure 8). The model for **I** had a composition of $\text{Sr}_2\text{Ca}_3\text{Ga}_6\text{O}_{14}$, with each A-site given the composition of the majority species for that site in the crystal structure. The model for **II** was constructed in the primitive $P1$ cell of the $F1$ pseudocubic cell, with a composition of $\text{SrCa}_2\text{Ga}_2\text{O}_6$. Two clusters around B-site oxide ions are present within the cell: one cluster consisting of four GaO_4 tetrahedra, and one consisting of four CaO_5 trigonal bipyramids (Extended Data Figure 9). The model structures for both phases were structurally optimized using DFT, and their DFT energies used directly to assess their energies relative to the convex hull at 0 K (no configurational entropy contributions were included).

Classical Force-Field Calculations

Force-field calculations were performed using GULP⁵⁶. Buckingham short-range potentials were used between cations and anions, and between two oxide ions, with a radial cutoff of 12 Å. The long-range electrostatic energy was calculated with atomic charges split between harmonically coupled cores and shells to model polarization⁵⁷. All force-field parameters were obtained from the literature^{58–64}. Cell parameters and atomic positions were optimized until the norm of the gradient was lower than 0.001. For the Y-Sr-Ca-Ga-O calculations, we can see from the top two panels in Extended Data Figure 10d that the distribution of Ga coordination environments is almost the same for FF-relaxed and DFT-relaxed structures, indicating that the force-fields are adequate for use in the first-pass MC-driven search for the probe structures.

Density Functional Theory Calculations

DFT calculations were performed using the Vienna ab-initio simulation package VASP package⁶⁵ using the PBE functional⁶⁶ and PAW potentials⁶⁷. The unit cell and atomic positions of each structure were optimised until forces were less than 0.01 eV/Å, using a plane wave cut-off energy of 520 eV. *K*-point grids were established to fulfil minimum requirements of the following equation along each of the three crystallographic axes:

$$k_i \times r_i \geq 30$$

Where r_i is the real lattice vector i (in Å) and k_i is the number of k -points in the same direction.

Synthesis of materials within the low-energy region

Samples for the initial evaluation grid were synthesised from the stoichiometric quantities of their pre-dried binary oxides/carbonates and hand ground (Y₂O₃ (99.99%, Alfa Aesar), SrCO₃ (99.99%, Alfa Aesar), CaCO₃ (99.95%, Alfa Aesar), Ga₂O₃ (99.99%, Alfa Aesar) and Eu₂O₃ (99.99%, Alfa Aesar)). For the initial 34 compositions studied (including composition 1 referred to in the main text), samples were heated in Al₂O₃ crucibles for 24 hours at 1150 or 1200°C with a typical heating rate of 5 °C/min (all compositions were initially trialled at 1200°C, some samples were found to have melted and were therefore re-synthesized at 1150°C). For the last seven compositions synthesised to improve phase purity from composition 1 to compositions 2 and 3 (see “Refinement of the structures of phases I and II” for compositions 2 and 3) and the resulting NPD sample, the powders were also subjected to a pre-firing step of 120 hours at 1000°C before regrinding and firing at the final temperature until the reaction was observed to have completed by PXRD. Firings were performed in 24 hour steps and the reaction found to be complete in either 48 hours or 72 hours.

Crystal growth

For the melt of composition 1 the starting materials were prepared as described above, targeting 4 g of sample, heated in a Pt crucible to 1300°C (approximately 100°C above the melting point) for 3 hours in a box furnace. The sample was then cooled to 800°C at 0.1°C/min, then the power supply to the furnace was turned off, and the sample was allowed to cool to room temperature. Crystals were removed from the crucible by hand using a spatula, and two types of crystal were

distinguished according to their two distinct colours (as shown in figure 2d). Crystals for X-ray diffraction studies were then separated by hand under a microscope based upon crystal colour. For both phases, crystals were found to be up to $\sim 1 \text{ mm}^3$ in size.

Characterisation methods

Single crystal X-ray diffraction (SXRD) data were collected on a Rigaku MicroMaxTM-007 HF with a molybdenum rotating anode microfocus source and a Saturn 724+ detector. Data integration and reduction was performed by the CrysAlisPro system software⁶⁸ with structure solutions and refinements performed using SHELX-2013^{69–71}.

Powder X-ray diffraction (PXRD) data were collected on the samples used in phase exploration with a Bruker D8 Advance diffractometer using a monochromatic $\text{CuK}\alpha_1$ source in transmission geometry using foil sample holders. For the structure refinement of **II**, synchrotron PXRD data were collected at Diamond Light Source using the I11 beamline with a wavelength of $0.826281(10) \text{ \AA}$ with a step size of 0.004° (d space range $18.94\text{--}0.48 \text{ \AA}$). Time of flight (TOF) NPD data were collected at the ISIS spallation source on the HRPD beamline, with data collected on the 168° and 90° detector banks (d -spacing range $3.73\text{--}0.68 \text{ \AA}$) used in subsequent Rietveld refinements performed using TOPAS Academic⁷² (*F1* structure) and Jana 2006⁷³ (*F432* structure) software packages. PXRD patterns for Eu-substituted **II** were collected on a Phillips PANalytical diffractometer (using a monochromated $\text{CoK}\alpha_1$ source) in Bragg-Brentano geometry.

Indexing of PXRD patterns and Principal Component Analysis (PCA)⁷⁴ was performed using Panalytical Highscore Plus software²². PCA attempts to decompose the entire set of diffraction patterns into linear combinations of basis vectors. In the dataset of the PXRD patterns of the 34 initial samples, 85% of the variance is described using the first three components (used as axes in Extended Data Figure 3a). This results in 29 of the diffraction patterns being clustered into three groups according to their similarity, and a fourth group containing the five patterns which could not be clustered.

Transmission Electron Microscopy (TEM) Energy Dispersive X-ray Spectroscopy (EDX) was performed on a JEOL 2000FX equipped with an EDAX detector. Specimens were prepared by dispersing 10 mg of finely ground sample in 3 ml of ethanol, errors on reported compositions are derived from the standard deviation of the measured points. A few drops of the suspension were deposited on the holey carbon film of a copper TEM grid. Spectra were collected for more than 10 minutes for each particle in order to maximize the signal to noise ratio. Quantification was performed with Genesis software provided by EDAX. Selected Area Electron Diffraction (SAED) and Convergent Beam Electron Diffraction (CBED) were performed on an aberration-corrected Jeol 2100 Transmission Electron Microscope. Samples were prepared following the same procedure used for EDX and loaded on a double tilt holder with an angle range of $\pm 20\text{--}25^\circ$.

Samples of **II** substituted with Eu, with nominal composition $\text{Eu}_{0.016}\text{Sr}_{0.174}\text{Ca}_{0.426}\text{Ga}_{0.384}\text{O}_{1.2}$, were tested as down-conversion phosphors⁷⁵. Room temperature photoluminescence spectra were collected with a Shimadzu RF-5301PC

spectrofluorophotometer, depositing the powder sample on a glass slide. Quantum yields at room temperature were measured using a Horiba Fluoromax-4 Fluorometer with a 150 W xenon arc lamp for excitation and a 150 mm Spectralon™-coated Quantum-F integrating sphere attachment. The samples were directly mounted as powders on a cavity in an anodized aluminium plate, and were then covered with quartz glass for measurement. The quantum yield was calculated based on the method of de Mello et al⁷⁶. Results are presented in section “Eu³⁺ substitution into II to Produce a Down-Conversion Phosphor”, and Extended Data Figure 3f, g.

Refinement of the structures of phases I and II

Refinement of the structure of I proceeded smoothly from single crystal X-ray diffraction (SXRD). Although neither neutron nor X-ray diffraction can distinguish Y from Sr, assignment of the non-Ca density on the Ca sites to Y rather than Sr, and assuming that Sr solely occupies the eight co-ordinate site at a rate that matches the EDX content, gives an A_{2.5} site composition Y_{0.06}Sr_{0.96}Ca_{1.48} that compares well with the EDX ratio of the 2.5 A cations of Y_{0.07}Sr_{0.96}Ca_{1.47}. The charge of the Y³⁺ can be balanced by the introduction of interstitial oxygen, as demonstrated for A₂Ga₃O₇ melilites¹⁵.

SXRD revealed that phase II was based on a 64-fold $4a_p \times 4a_p \times 4a_p$ expansion of a primitive perovskite unit cell of edge a_p , but the associated twinning coupled with possible lowering of metric symmetry undetectable with the single crystal diffractometers used prevented structure refinement from these data. Instead, refinement required a powder sample with a higher fraction of II than present in composition 1 (26.49(8)%).

Samples of 1.2 g at the EDX-defined composition for II (Y_{0.09(3)}Sr_{0.85(3)}Ca_{2.12(8)}Ga_{1.93(7)}O_{6.0(3)}) plus 6 nearby compositions were hand ground under acetone, sintered at 1000 °C for 5 days, then re-ground and re-fired at 1200°C in 24 hour steps until the reaction had completed as observed by PXRD (reaction times ranging between 48 and 72 hours). A high purity sample was synthesised, allowing for the combined Rietveld refinement of II from synchrotron powder X-ray diffraction and high resolution neutron powder diffraction (composition 2; Y_{0.038}Sr_{0.320}Ca_{0.848}Ga_{0.794}O_{2.416}) containing 87.59(7) wt% of II together with I. A sample with a 97.0(1) wt% phase fraction of II by Rietveld refinement of PXRD data was later obtained at composition 3, Y_{0.038}Sr_{0.360}Ca_{0.848}Ga_{0.754}O_{2.396}.

The enhanced Q-resolution of the powder diffractometers revealed a triclinic distortion ($\Delta d/d \approx 3 \times 10^{-4}$) and permitted successful refinement in space group *F1* despite the extensive pseudosymmetry (Extended Data Figure 6). Convergent Beam Electron Diffraction (CBED) confirmed the absence of point symmetry and thus the space group assignment (Extended Data Figure 6e–g).

The triclinic symmetry of II was confirmed via SAED/CBED measurements, on a crystallite on the [101] zone axes (SAED is shown in Extended Data Figure 6d, with theoretical diffraction spots in red). The CBED measurement at high camera length (Extended Data Figures 6e–g), exhibits no symmetry on the diffraction spots, indicating that the material must have symmetry belonging to space group No.1, with absences in the X-ray powder diffraction confirming it as the face-centred

setting of the space group; *F1*. The Rietveld refinement was performed in the standard primitive setting of the *F1* unit cell, *P1* with the unit cell parameters; $a = 10.9317(7) \text{ \AA}$, $b = 10.93359(5) \text{ \AA}$, $c = 10.93023(8) \text{ \AA}$, $\alpha = 59.9980(5)^\circ$, $\beta = 59.9781(4)^\circ$, $\gamma = 59.9672(5)^\circ$. The structure was transformed back into the *F1* setting after completion. The two settings of the space group are related by the following matrix (F centred to Primitive):

$$\begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{pmatrix}$$

The refinement used a starting model from the initial SXRD solution. As the refinement progressed, groups of atoms were observed to come together such that a series of restraints could be put in place for the final refinement, increasing precision without decreasing the quality of the fit. In the final refinement, the following restraints were imposed upon the model:

The partial occupancies of all of the mixed (Sr/Ca)O₇ cation sites within the cell were constrained to be identical and sum to 1, and a single thermal parameter was refined across all such sites. This was also the case for the (Ca/Sr)O₆ cation sites. Each of the purely Ga cation sites were constrained to have a fixed occupancy of 1 and the same thermal displacement parameter.

All of the mixed CaO₅/GaO₄ sites were constrained to have the same thermal parameters and the total occupancy of each site was constrained to 1. On the assumption that the site was either a GaO₄ tetrahedron, or a CaO₅ trigonal bipyramid, the occupancy of the split oxygen positions surrounding the mixed Ca/Ga sites was fixed to match that of the cation species to which they were bound, for example an oxygen site 1.8 Å from a mixed Ca/Ga site would be assigned as bound to the Ga component and its occupancy fixed to match. Bond restraints were then applied to the tetrahedral oxygens and the equatorial component of the trigonal bi-pyramids along with the axial bond to the B-site oxygen sites. Bond restraints were set to be equal to the sum of the two ionic radii. The oxide anions on B-site defects and those in the GaO₆ octahedron had a fixed occupancy of 1. Finally all of the oxygen thermal parameters in the structure of **II** were constrained to be identical and refined.

During initial powder refinements of **II**, the occupancies of the mixed Ca/Ga sites were found to divide into two clusters per asymmetric unit, such that all four Ca/Ga sites surrounding an oxide ion B-site defect had the same set of partial occupancies within error. The asymmetric unit contains two B-site defects, and the Ca/Ga ratios surrounding one cluster were the inverse of those on the other. Following this observation, constraints were applied so that each group of four sites has the same occupancy parameter and that the second group has the inverse occupancy of the first, thus under the constraint that the total occupancy of each site was one, all mixed Ca/Ga sites within the asymmetric unit could be refined using a single occupancy parameter. In the final structure, this parameter refined such that one cluster contained 52(2) % Ga and 48(2) % Ca, and the other 48(2) % Ga and 52(2) % Ca.

To understand the disorder observed in **II**, the crystal structure can be described as the superposition of many local orderings, based upon the ordering of Ca/Ga sites around each of the B-site defect oxide anions (Extended Data Figure 9e). The constraints used in the refinement can be understood in terms of the superposition of the particular orderings shown in Extended Data Figure 9c, d. These two models are constructed by choosing one B-site defect oxide anion within the asymmetric unit to be fully surrounded by GaO₄ tetrahedra or CaO₅ trigonal bipyramids and vice versa for the other B-site defect oxide anions (Extended Data Figure 9f, g). The choice of oxygen sites to include within each model was then made based upon bond distances to the cation sites as chosen in each model. One of these two models (Extended Data Figure 9c) was then used as the basis for the DFT model of **II** (Extended Data Figure 9b).

The BVS of the five co-ordinate Ca sites in **II** has a mean value of +2.0, with a BVS range between +1.7 and +2.3, confirming that the sites were refined to have chemically sensible environments. The largest deviation in calculated BVS values from those expected for the refined ions was found to be ± 0.4 (for the 7 co-ordinate Sr11 site) across all of the cation environments. This, and the location of the calculated structure of **II** below the convex hull, supports the refined structure.

Due to the extreme pseudo-symmetry and disorder found in **II**, bond lengths for a large number of the cation sites could only be determined to two significant figures. To illustrate that the precision on bond distances is a result of the pseudo-symmetry a cubic approximation of **II** was constructed. The initial cubic indexation of **II** (from single crystal) yielded the space group $Fm\bar{3}m$, however, this model does not allow for the observed tetrahedral ordering observed in the *F1* solution. In order to retain the same tetrahedral ordering observed in **II**, the symmetry was reduced to *F432*. Refinement in this space group for the same data ranges as for the *F1* solution yields a $\chi^2 = 5.81$ (vs. 3.55 for the triclinic solution), however the precision on bond distances is increased to a minimum of four significant figures. Powder CIF for both the *F1* and *F432* solutions are available as Supplementary Information. The CBED data demonstrate that the symmetry is *F1*, the *F432* refinement is provided to demonstrate that the bond length precision in *F1* is due to the high pseudo-symmetry.

Eu³⁺ substitution into **II to Produce a Down-Conversion Phosphor**

The identification of Y-containing quinary gallates with the probe structure approach suggests the substitution of open-shell lanthanide cations to explore optical and magnetic properties. Eu was substituted for Y in **II**, with phase pure samples synthesised at the nominal composition Eu_{0.032}Sr_{0.348}Ca_{0.852}Ga_{0.768}O_{2.4} (composition 4), using Pt crucibles and the same procedure described for composition 2. Phase purity was confirmed by Le Bail refinement of the PXRD pattern using Jana 2006⁷³, starting from the unit cell and space group symmetry of **II** (Extended Data Figure 3f) and EDX measurement yielding the composition: Eu_{0.056(6)}Sr_{0.32(4)}Ca_{0.88(2)}Ga_{0.73(2)}O_x from 10 particles (inset in Extended Data Figure 3f). Attempts to substitute higher concentrations of Eu into **II** resulted in the presence of impurity phases.

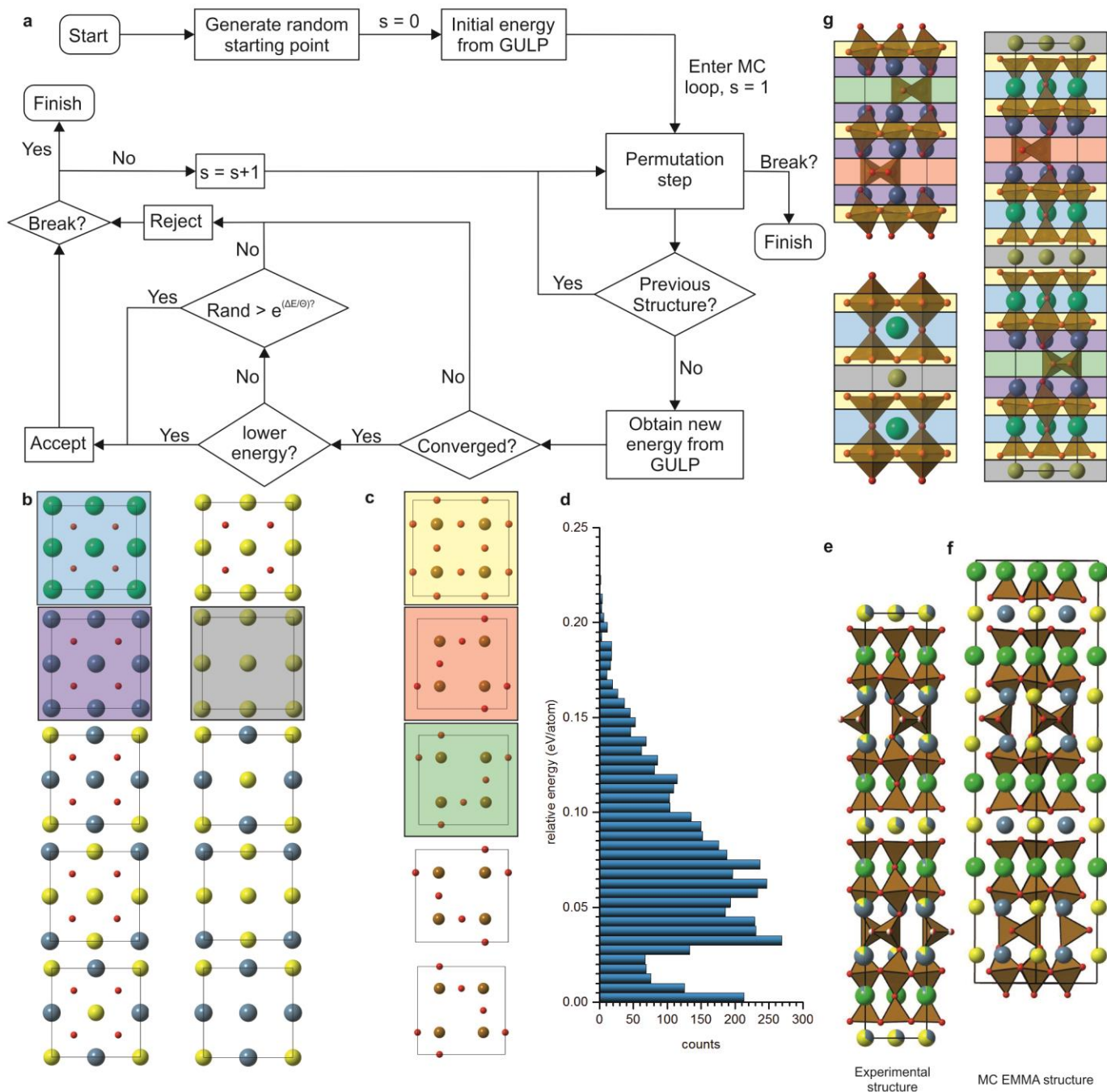
Excitation spectroscopy shows that composition 4 can be excited at a wavelength of 460 nm, which can be readily produced by conventional blue LEDs⁷⁵. Excitation at this wavelength produced an emission peak in the red region of the

visible spectrum at 610 nm (Extended Data Figure 3g). Quantum yields were measured on three different samples of composition 4, giving values of 3 %, 3 % and 10 %. These results show that **II** is a new phosphor parent structure.

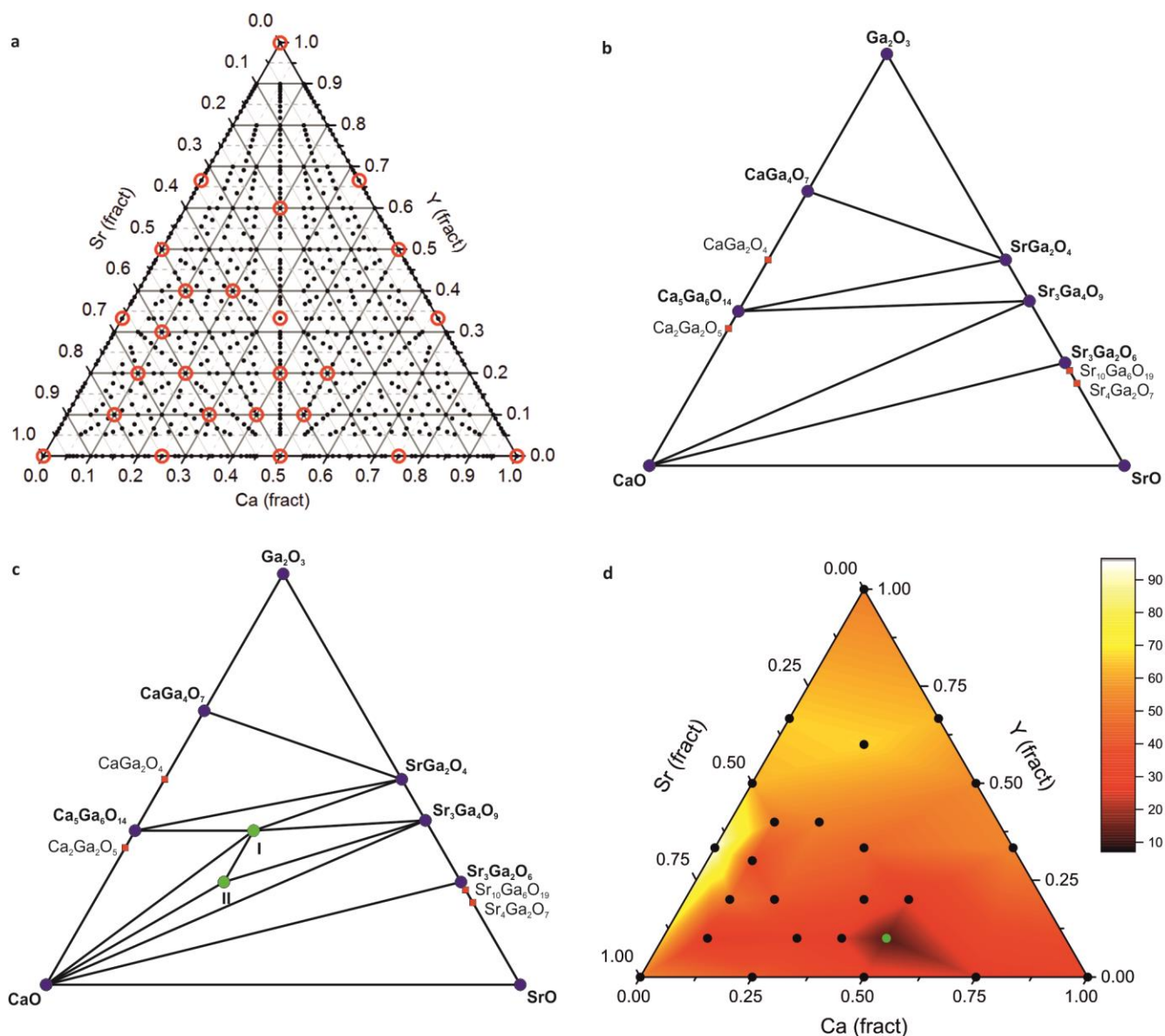
Code availability: The MC EMMA method was implemented in python and is dependent on the Atomic Simulation Environment (ase, <https://wiki.fysik.dtu.dk/ase>). A copy of the code used for this study is available at <http://pcwww.liverpool.ac.uk/~msd30/software/MC-EMMA.html>.

Data availability: The datasets generated during and/or analysed during the current study are available from the University of Liverpool, and can be found at: <http://datacat.liverpool.ac.uk/id/eprint/82>

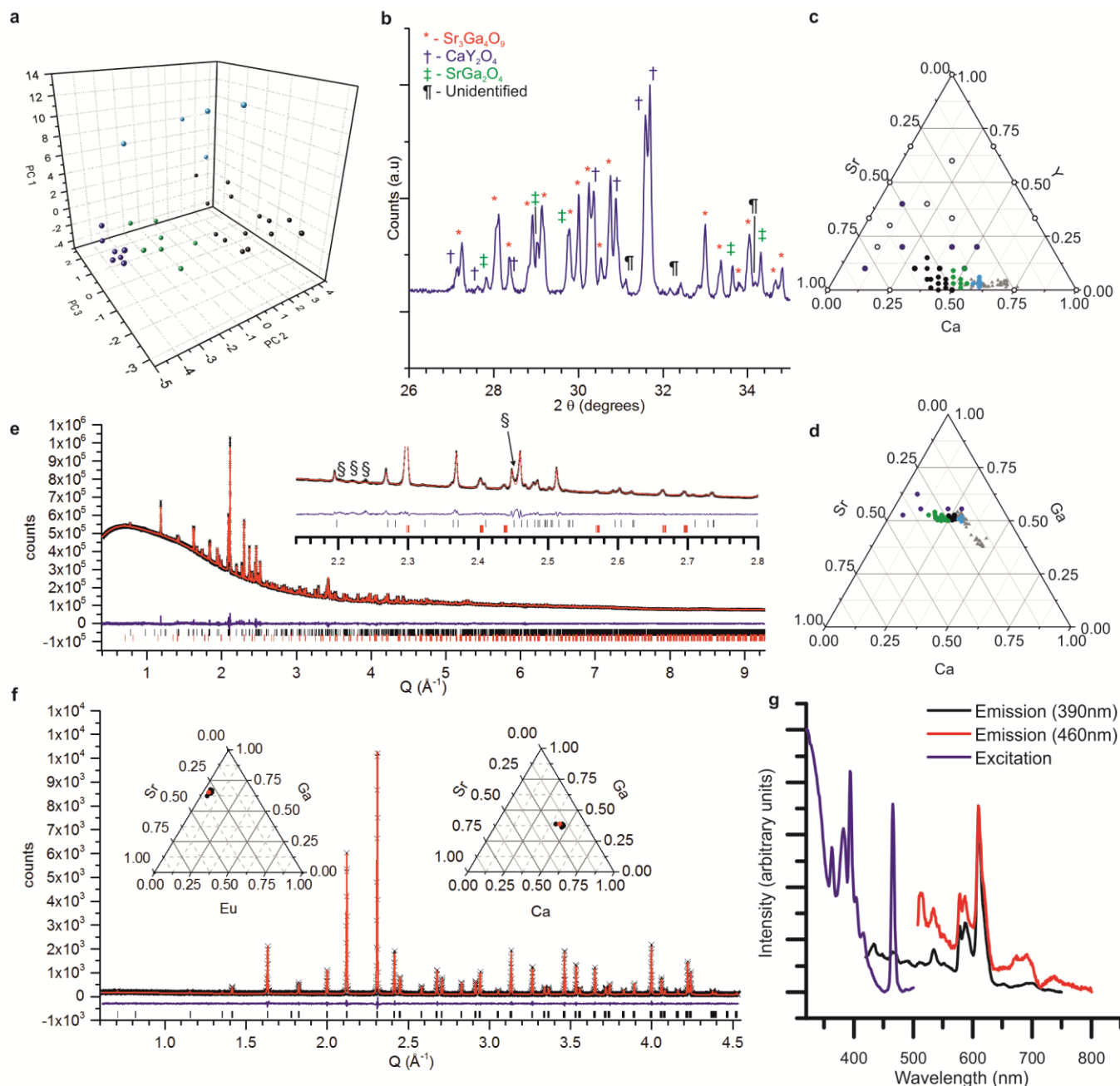
Extended Data captions



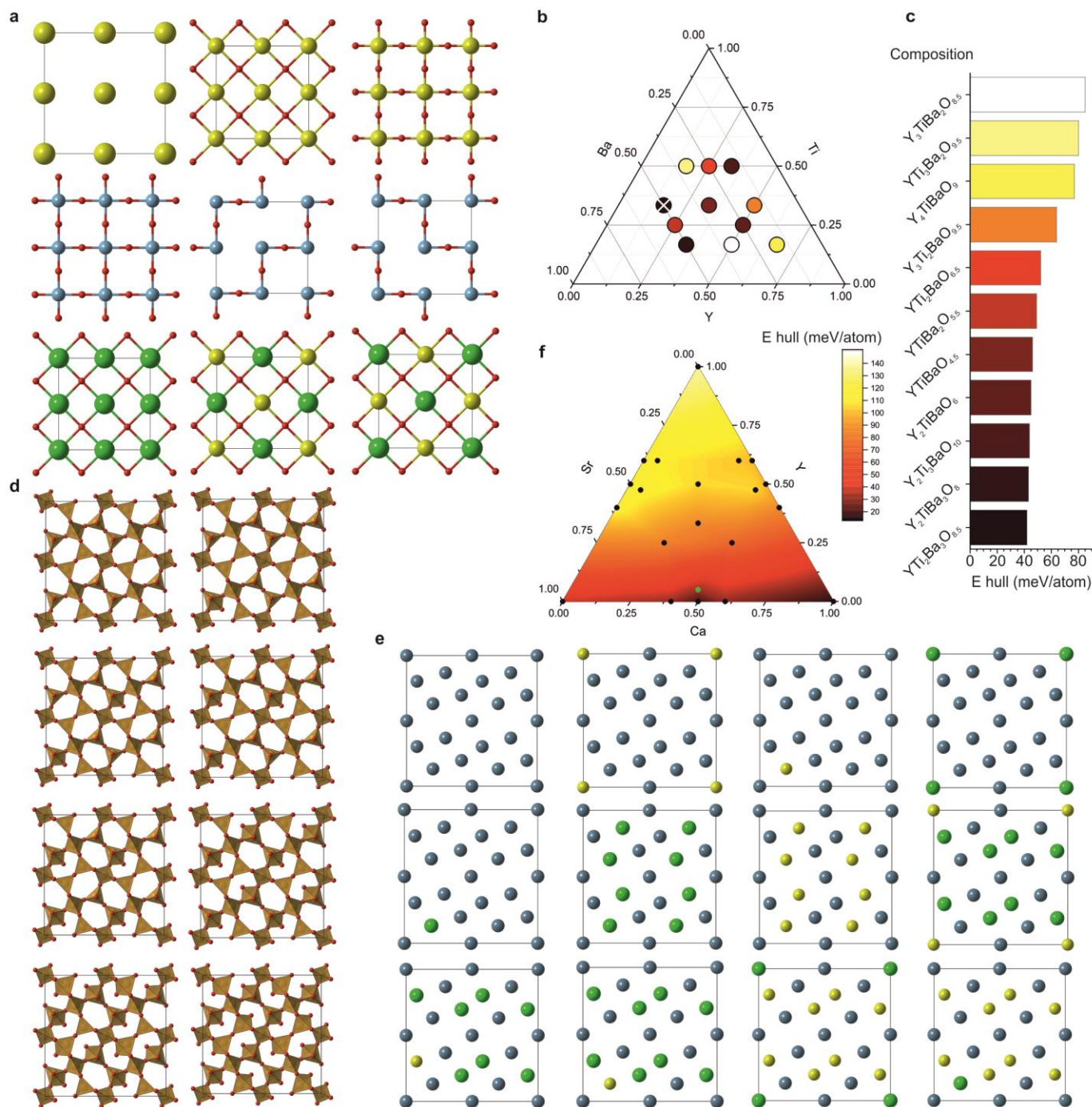
Extended Data Figure 1: **a** The flowchart of the MC-EMMA method. **b** Large cation modules used in testing of the MC-EMMA method on the structure of YBa₂Ca₂Fe₅O₁₃; left, top-bottom; Ba₄O₄, Ca₄O₄, Ca₃YO₄, CaY₃O₄, Ca₂Y₂O₄ and right, top-bottom; Y₄O₄, Y₄Ca₂Y₂, CaY₃ and Ca₃Y. **c** Small cation modules used in MC-EMMA testing top - bottom; FeO₂ and 4 rotational variants of FeO. **d** energies of all converged structures using FF from YBa₂Ca₂Fe₅O₁₃ in eV/atom relative to the lowest energy structure. **e** the experimental structure of YBa₂Ca₂Fe₅O₁₃. **f** The lowest energy structure found by MC-EMMA, viewed along the [110] direction. The correct stacking sequence is observed, with mixed Y/Ca modules, capturing the main large cation site disorder in the experimental structure. This test of the MC-EMMA method was performed using FFs only, with FF parameters previously used with EMMA on the same system¹⁹. **g** the structures of Ca₂Fe₂O₅⁷⁷, YBa₂Fe₃O₈⁷⁸ and YBa₂Ca₂Fe₅O₁₃⁷⁹ (ideal) overlaid with coloured blocks indicated the MC-EMMA modular description using modules from **b** with the corresponding colour.



Extended Data Figure 2: **a** Compositions for perovskite-derived probe structures in the Y-Sr-Ca-Ga-O phase field: 25 compositions highlighted with red circles are selected from the 691 possible with the modules and cell size limits in the MC-EMMA search. Given the low experimental Y content and absence of Y in the DFT model structures of phases I and II, we show computed convex hulls in the reduced SrO-CaO-Ga₂O₃ phase diagram (*i.e.* the $y = 0$ plane of the Y₂O₃-SrO-CaO-Ga₂O₃ phase diagram shown in Extended Data 3). **b** The convex hull calculated using previously reported phases. Blue circles are phases calculated to be on the hull, red squares are phases with DFT energies above the hull. **c** The computed convex hull including the two new phases discovered in this study, phase I (Sr₂Ca₃Ga₆O₁₄) and phase II (SrCa₂Ga₂O₆), as green circles on the hull. **d** Energy of probe structures relative to the new convex hull shown in **c** which includes the newly discovered phases I and II. The energies are plotted on the Ga = 0.5 slice of the Y₂O₃-SrO-CaO-Ga₂O₃ phase diagram with black points showing the compositions of the probe structures. The DFT calculated energies per formula unit of the experimentally reported phases are Ca₂Ga₂O₅: -56.278 eV, Ca₅Ga₆O₁₄: -156.837 eV, CaGa₂O₄: -43.540 eV, CaGa₄O₇: -73.995 eV, Sr₁₀Ga₆O₁₉: -214.803 eV, Sr₃Ga₂O₆: -67.690 eV, Sr₃Ga₄O₉: -99.074 eV, Sr₄Ga₂O₇: -79.737 eV, SrGa₂O₄: -43.254 eV, Y₃Ga₅O₁₂: -145.076 eV, YGaO₃: -38.055 eV, Y₂O₃: -45.772 eV, Ga₂O₃: -30.178 eV, CaO: -12.806 eV, SrO: -12.055 eV, Sr₂Ca₃Ga₆O₁₄: -156.143 eV, SrCa₂Ga₂O₆: -68.999 eV.

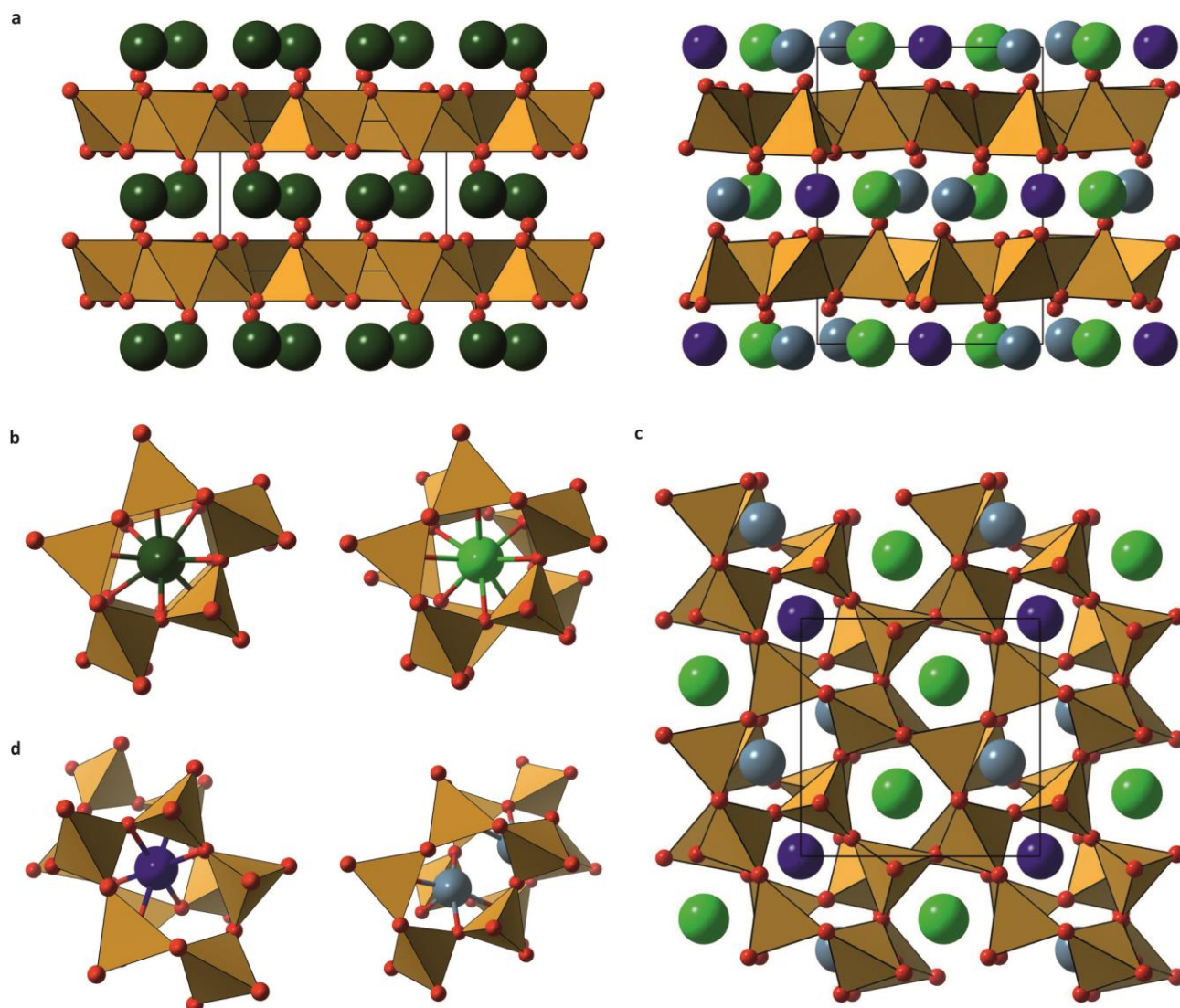


Extended Data Figure 3: **a** Principal component analysis results of the PXRd patterns of the initial 34 compositions sampled, with the analysis performed in Highscore Plus²²; five samples do not cluster (dark blue), with the remainder grouped into three clusters (black, green and cyan), with composition 1 (referenced in main text and Figure 2) falling into the cyan cluster. **b** representative PXRd pattern from the non-clustered samples represented in **a**, with symbols indicating known phases in the pattern^{36,38,80}. **c** $\text{Ga} = 0.5$ and **d** $\text{Y} = 0.0$ slices of $\text{Y}_2\text{O}_3\text{-SrO-CaO-Ga}_2\text{O}_3$ quaternary diagram, with nominal compositions coloured as in **a**, calculated compositions in black open circles and EDX data from composition 1 in grey. **e** Synchrotron PXRd data (PSD, $\lambda = 0.825781 \text{ \AA}$) of composition 1, using the final refined models for each new phase. Reflections from the new phases I (black, 73.4(1) wt%) and II (red, 26.6(1) wt%) are marked. The inset indicates the location of the most intense reflections not indexed to either of the new phases with §. **f** Le Bail fit to PXRd data of $\text{Eu}_{0.032}\text{Sr}_{0.348}\text{Ca}_{0.852}\text{Ga}_{0.768}\text{O}_{2.4}$ (composition 4) inset: EDX data, black: observed points, red: nominal composition. **g** Excitation (blue, detection wavelength 611 nm) and photoluminescence emission spectra for composition 4, excited at 390 nm (black) and 460 nm (red).

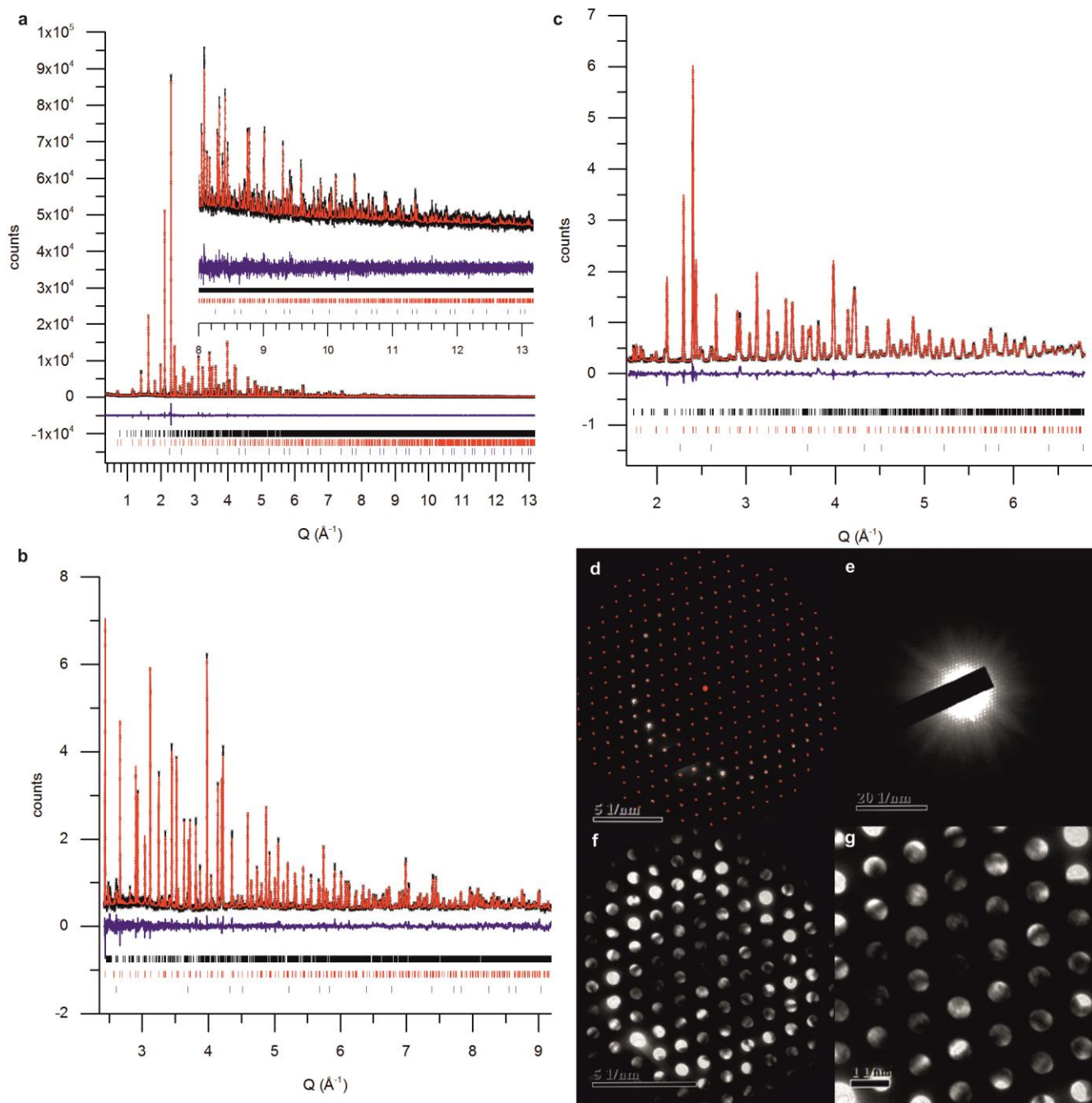


Extended Data Figure 4: **a** The nine modules used for the Y-Ba-Ti-O diagram (Y= yellow, Ti = cyan, Ba = green and O = red). **b** Compositions of probe structures in the Y-Ba-Ti-O phase field, filled colours indicate the energy from the convex hull (E hull), as shown in **c**, the point highlighted by the white cross is the lowest energy composition, $\text{YTi}_2\text{Ba}_3\text{O}_{8.5}$, which is also the only experimentally known quaternary **d** Eight smaller cation modules used for melilite probe structures, with composition $\text{Ga}_{24}\text{O}_{56+\delta}$, $\delta = 0, 1, 2, 3, 4, 8, 12, 16$ (top row to bottom row). **e** 13 decorations of the larger cation module used for melilite probe structures, containing up to three elements: A (blue) A' (yellow) A'' (green). The module set is constructed using the combinations $[\text{A}=\text{Ca}, \text{A}'=\text{Y}, \text{A}''=\text{Sr}]$, $[\text{A}=\text{Sr}, \text{A}'=\text{Y}, \text{A}''=\text{Ca}]$ and $[\text{A}=\text{Y}, \text{A}'=\text{Ca}, \text{A}''=\text{Sr}]$ (giving 39 combinations). For each module in **d** and **e**, three variants were created by the translations $[\text{+}0.5, 0, 0]$ and $[0, \text{+}0.5, 0]$ and by rotation of 90° about the c axis (resulting in 32 small cation modules and 156 large cation modules). **f** The energy

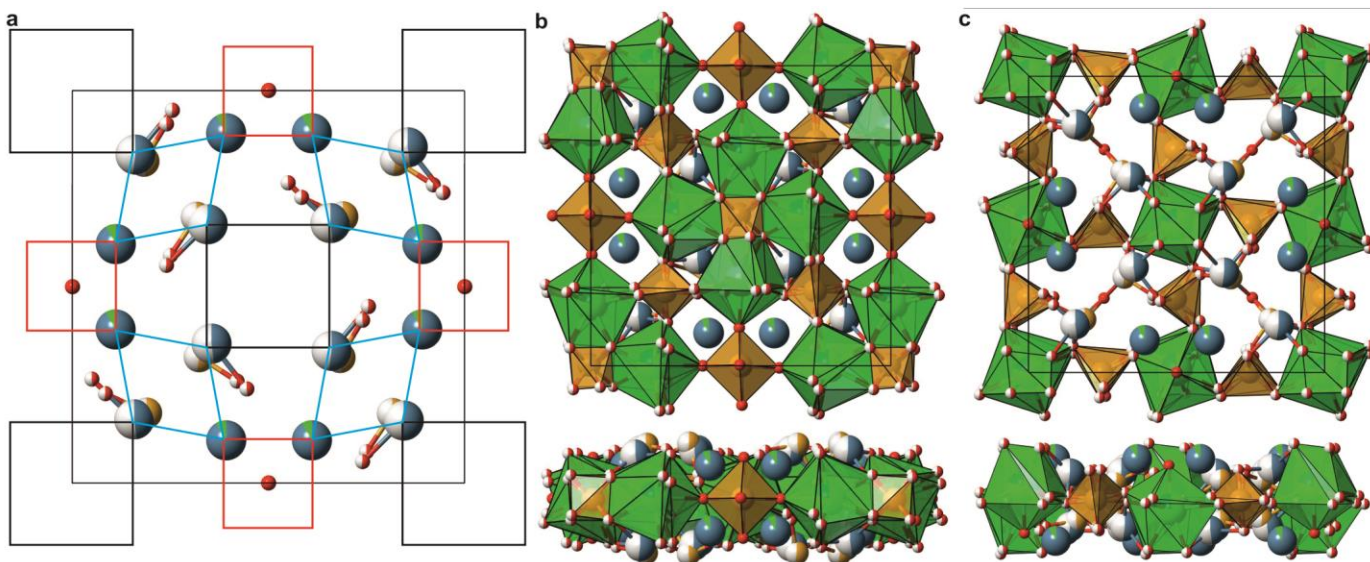
of the probe structures calculated using MC-EMMA with melilite modules are plotted relative to the convex hull created with previously reported phases (Extended data Figure 2), the lowest energy quinary is highlighted in green.



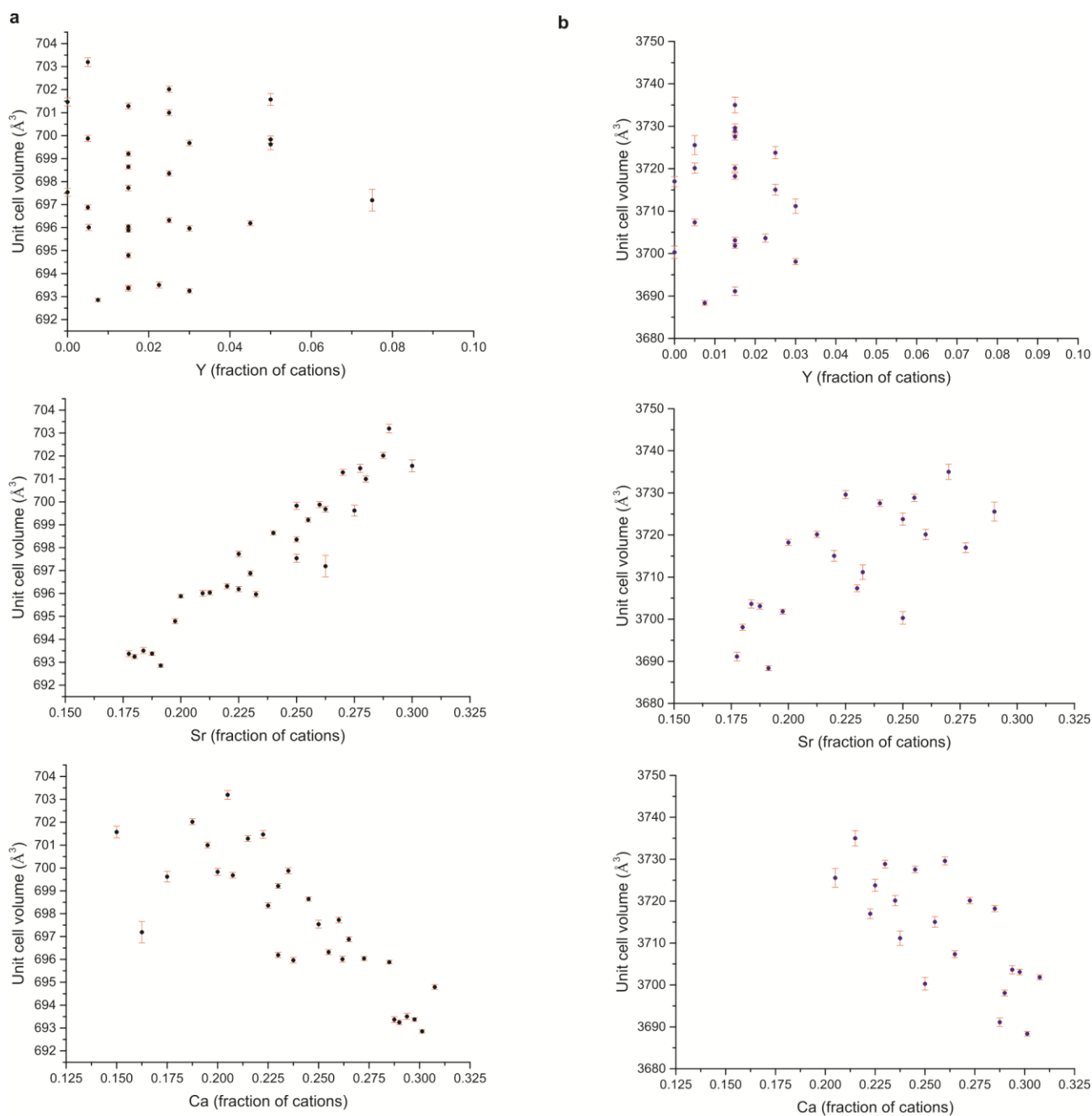
Extended Data Figure 5: **a** [100] view of LaSrGa₃O₇ melilite^{15,26} (left) and **I** (right). The difference between the Ca,Sr cation occupancy on either side of the Ga₃O₇ layers in **I** doubles the *c* parameter compared to melilite (Figure 3g). **b** [001] view of the Sr²⁺ sites in melilite (left) and **I** (right), showing the superposition in **I** of three- and four-connected tetrahedra in the pentagonal rings that coordinate Sr²⁺ from the two neighbouring Ga₃O₇ layers, in contrast to the identical ring orientations in melilite. **c** Two Ga₃O₇ layers in **I** viewed along [001] with one layer of cations between them. **d** [001] view of the octahedral Ca1 site (dark blue, left) and five-coordinate Ca2 site (light blue, right) in **I** where the displacement of the two pentagons contrasts with the near-superposition of the coordinating rings in the strontium-containing chains. The contacts from the Ca1 site to the bridging oxides displace the two pentagons above and below it relative to each other. Both Ca2 sites also coordinate to a bridging oxide from each ring. The Ca2 sites lie over the shared edge of two rings in one layer and are displaced towards two of the tetrahedra away from the ring centroid in the second layer. Mixed La/Sr (dark green), Sr (light green), Ca1 (dark blue), Ca2 (light blue), Ga(brown) and O(red).



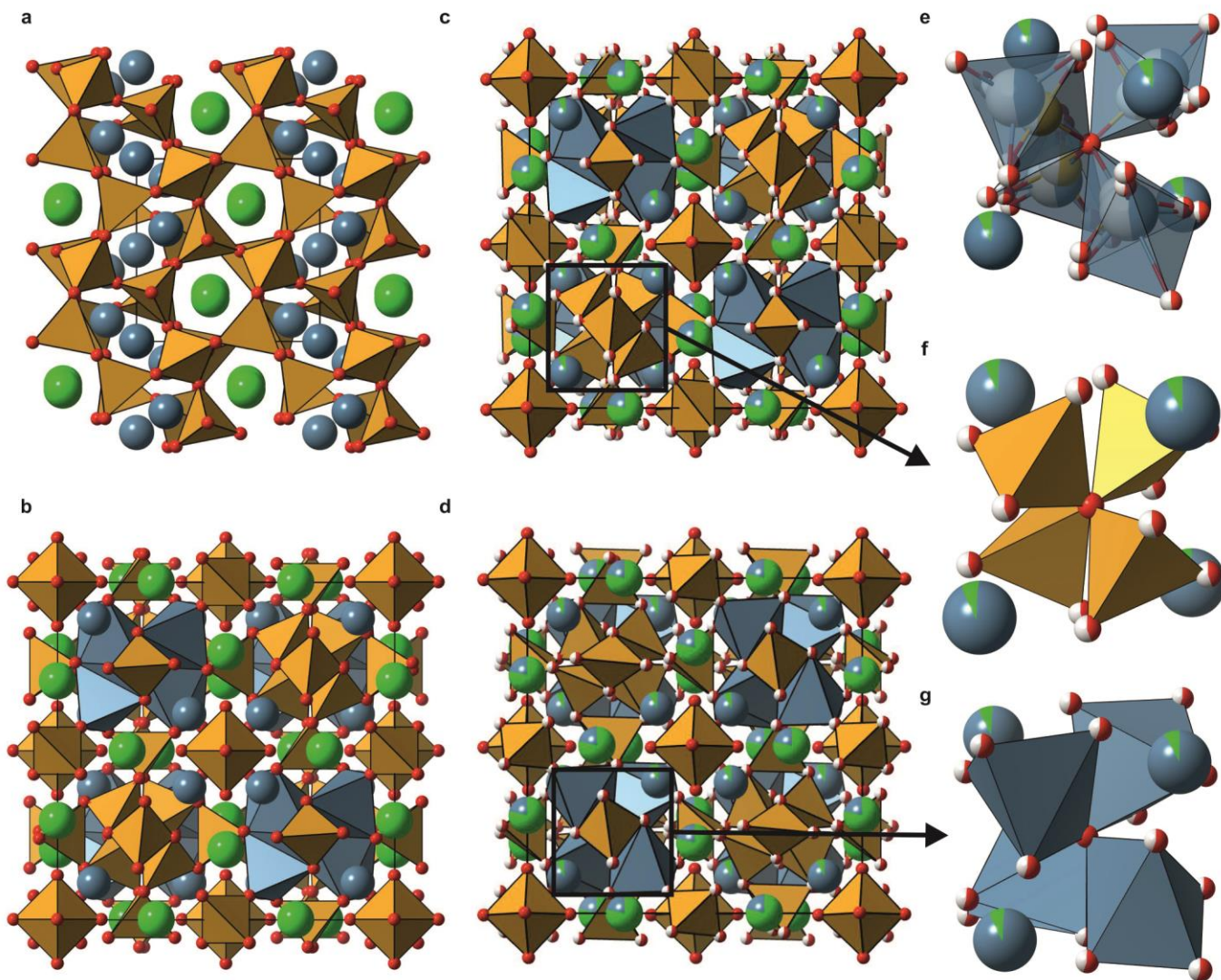
Extended Data Figure 6: Room temperature Rietveld refinement of sample with nominal composition 2 ($\text{Y}_{0.038}\text{Sr}_{0.320}\text{Ca}_{0.848}\text{Ga}_{0.794}\text{O}_{2.416}$) containing 87.59(7) wt% phase II using **a** synchrotron X-ray powder diffraction ($\chi^2 = 3.84$) and neutron powder diffraction data from **b** 168° bank, $\chi^2 = 1.59$ and **c** 90° bank, ($\chi^2 = 4.9$) with a combined $\chi^2 = 3.55$ with 418 parameters. Space group $F1$. $a = 15.4538(1)$ Å, $b = 15.4573(1)$ Å, $c = 15.4621(1)$ Å, $\alpha = 89.9990(7)^\circ$, $\beta = 89.9669(7)^\circ$, $\gamma = 89.9847(7)^\circ$. **d** SAED and **e–g** CBED pattern at increasing camera length of II along [101]. The absence of symmetry confirms the space group assignment.



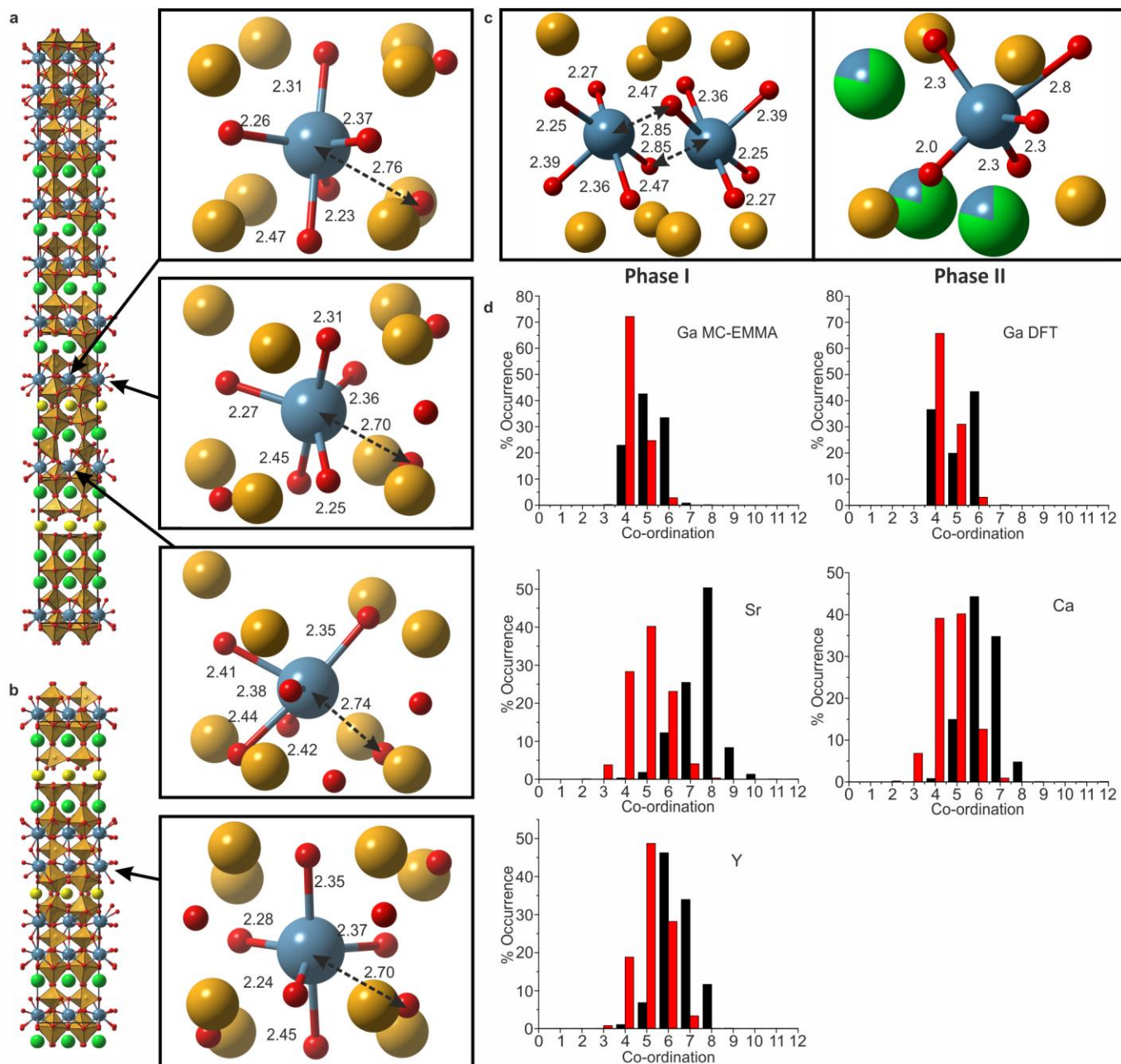
Extended Data Figure 7: **a** Structural units within the anion-deficient $A_{16}O_{10}$ layer of **II**. The smaller Ca_4O rectangles (red) have dimensions of $3.32 \text{ \AA} \times 3.42 \text{ \AA}$, while the larger anion vacancy centred $(Ca/Ga)_4$ rectangles (black) are $4.76 \text{ \AA} \times 4.60 \text{ \AA}$. The light blue rhombus with 2 Ca and 2 Ca/Ga vertices is also marked. **b** top: Relation between polyhedra in the B_1 layer and the anion vacancies in the $A_{16}O_{10}$ layer. The oxide at the centre of Ca_4O is located above the octahedral Ga, while the tetrahedral Ga is formed by rotation of the four coordinating oxides out of the B_1 plane coupled with its location below the anion vacancies at the centre of the $(Ca/Ga)_4$ rectangle. The SrO_7 unit is formed from one ordered oxygen from an octahedron in the B_1 plane, two oxygens from rotationally disordered GaO_4 tetrahedra, two oxygens from tetrahedra within the B_1 plane and two oxygens from the mixed Ca/Ga environment in the A layers; one above and one below the B_1 plane. **b** bottom: perpendicular view of B_1 layer with adjacent A layers. **c** Relation between the $A_{16}O_{10}$ and B_2 layer structures. The oxide occupying the B site vacancy in the B_2 layer has two of the four A site cation neighbours (which form the rhombus shown in **a**) displaced away from it to form the Ca_4O square, with the two Ca/Ga sites displacing below the $A_{16}O_{10}$ plane to coordinate to the oxide. Rotation of the B_2 layer oxides out of the plane at the four neighbouring GaO_4 tetrahedra reduces anion-anion repulsion with the oxide ion on the B-site.



Extended Data Figure 8: Observed unit cell volumes for **a**, phase I and **b**, phase II over the composition range in the initial synthetic study.



Extended Data Figure 9: **a** The fully ordered model of **I** used for DFT calculations. The combination of the three distinct layers in **II** with their disordered occupancies produces several near-equivalent three-dimensional connectivities. **b** The connectivity used in the DFT calculations is shown here. **c–d** The two ordered models of **II** used to understand the disorder present in the Rietveld refined structure, with different choices of cation clusters (shown in **f** and **g** below) around the B-site oxygen anions highlighted. **e** The disordered cation cluster identified within **II**, surrounding each of the B-site oxygen anions, which can be understood in terms of **f** the cluster when all four cations in the cluster are set to GaO₄ with the associated oxygen positions and **g** the cluster when all four cations are set to CaO₅ with the associated oxygen positions.



Extended Data Figure 10: Five coordinate calcium sites in **a** the most stable, and **b** the fourth most stable probe structures in the low-energy region of the phase diagram. Ca–O distances of less than 2.5 Å are shown with bonds, the next largest Ca–O distance is shown with a dashed arrow. All distances are given in Ångström. **c** Ca–O environments in the computed probe structures are compared with the experimentally refined Ca–O environments in phase I and phase II. **d** Frequency of occurrence of co-ordination numbers for each of the cation species in the Y-Sr-Ca-Ga-O phase field in the DFT-relaxed probe structures (shown for Ga, from $n = 170$ calculated structures) and in the FF converged structures generated using MC-EMMA ($n = 194,912$) using perovskite-derived modules (black) and melilite-derived modules (red). Bonds were counted as being in the co-ordination sphere if the M–O interaction contributed $> 5\%$ of the nominal charge when calculating the central cations bond valence sum of the central cation.

References

- 29 Faucher, M. & Pannetier, J. Refinement of the Y_2O_3 Structure at 77 K. *Acta Crystallogr, Sect B* **36**, 3209-3211, doi:10.1107/S0567740880011351 (1980).
- 30 Geller, S. Crystal Structure of Beta- Ga_2O_3 . *J Chem Phys* **33**, 676-684, doi:10.1063/1.1731237 (1960).
- 31 Ghebouli, M. A., Ghebouli, B., Bouhemadou, A., Fatmi, M. & Bouamama, K. Structural, electronic, optical and thermodynamic properties of $Sr_xCa_{1-x}O$, $Ba_xSr_{1-x}O$ and $Ba_xCa_{1-x}O$ alloys. *J Alloy Compd* **509**, 1440-1447, doi:10.1016/j.jallcom.2010.11.097 (2011).
- 32 Kahlenberg, V. & Shaw, C. S. J. $Ca_2Ga_2O_5$: a new high pressure oxogallate. *Z Kristallogr* **216**, 206-209, doi:10.1524/zkri.216.4.206.23257 (2001).
- 33 Bilyi, A. I. et al. Crystal-Structure of $Ca_5Ga_6O_{14}$. *Kristallografiya* **31**, 1217-1219 (1986).
- 34 Deiseroth, H. J. & Müller-Buschbaum, H. Alkaline-Earth Oxogallates .5. Crystal-Structure of Monoclinic $CaGa_2O_4$. *Z Anorg Allg Chem* **402**, 201-205, doi:10.1002/zaac.19734020211 (1973).
- 35 Kahlenberg, V. The crystal structures of the strontium gallates $Sr_{10}Ga_6O_{19}$ and $Sr_3Ga_2O_6$. *J Solid State Chem* **160**, 421-429, doi:10.1006/jssc.2001.9259 (2001).
- 36 Kahlenberg, V. & Parise, J. B. $Sr_3Ga_4O_9$ - a strontium gallate with a new type of tetrahedral layer structure. *Z Kristallogr* **216**, 210-214, doi:10.1524/zkri.216.4.210.23254 (2001).
- 37 Kahlenberg, V., Lazic, B. & Krivovichev, S. V. Tetrastrontium-digalliumoxide ($Sr_4Ga_2O_7$) - synthesis and crystal structure of a mixed anion strontium gallate related to perovskite. *J Solid State Chem* **178**, 1429-1439, doi:10.1016/j.jssc.2005.02.009 (2005).
- 38 Kahlenberg, V., Fischer, R. X. & Shaw, C. S. J. Polymorphism of strontium monogallate: The framework structures of beta- $SrGa_2O_4$ and ABW-type gamma- $SrGa_2O_4$. *J Solid State Chem* **153**, 294-300, doi:10.1006/jssc.2000.8768 (2000).
- 39 Nakatsuka, A., Yoshiasa, A. & Takeno, S. Site Preference of Cations and Structural Variation in $Y_3Fe_{5-x}Ga_xO_{12}$ ($0 \leq x \leq 5$) Solid-Solutions with Garnet Structure. *Acta Crystallogr, Sect B* **51**, 737-745, doi:10.1107/S0108768194014813 (1995).
- 40 Geller, S., Jeffries, J. B. & Curlander, P. J. Crystal-Structure of a New High-Temperature Modification of $YGaO_3$. *Acta Crystallogr, Sect B* **31**, 2770-2774, doi:10.1107/S0567740875008862 (1975).
- 41 Fallon, G. D. & Gatehouse, B. M. The crystal-structure of $Ba_2Ti_9O_{20}$ - A hollandite related compound. *J Solid State Chem* **49**, 59-64, doi:10.1016/0022-4596(83)90216-5 (1983).
- 42 Wu, K. K. & Brown, I. D. crystal-structure of beta-barium orthotitanate, beta- Ba_2TiO_4 , and bond strength bond length curve of Ti-O. *Acta Crystallogr, Sect B* **29**, 2009-2012, doi:10.1107/s0567740873005959 (1973).
- 43 Tillmanns, E. Crystal structure of tetrabarium tridecatitinate. *Inorg Nucl Chem Lett* **7**, 1169-1171, doi:10.1016/0020-1650(71)80059-4 (1971).
- 44 Hofmeister, W., Tillmanns, E. & Baur, W. H. Refinement of barium tetratitinate, $BaTi_4O_9$, and hexabarium 17-titinate, $Ba_6Ti_{17}O_{40}$. *Acta Crystallogr, Sect C* **40**, 1510-1512, doi:10.1107/s0108270184008544 (1984).
- 45 Liu, L. G. Dense modification of BaO and its crystal structure. *J. Appl. Phys.* **42**, 3702-&, doi:10.1063/1.1659673 (1971).
- 46 Tillmanns, E. Refinement of barium dititanate. *Acta Crystallogr, Sect B* **30**, 2894-2896, doi:10.1107/S0567740874008405 (1974).
- 47 Hofmeister, W., Tillmanns, E. & Baur, W. H. Refinement of barium tetratitinate, $BaTi_4O_9$, and hexabarium 17-titinate, $Ba_6Ti_{17}O_{40}$. *Acta Crystallogr, Sect C* **40**, 1510-1512, doi:10.1107/s0108270184008544 (1984).
- 48 Tillmanns, E. Die Kristallstruktur von $BaTi_5O_{11}$. *Acta Crystallogr, Sect B* **25**, 1444-1452, doi:10.1107/S0567740869004195 (1969).
- 49 Buttner, R. H. & Maslen, E. N. Structural parameters and electron difference density in $BaTiO_3$. *Acta Crystallogr, Sect B* **48**, 764-769, doi:10.1107/S010876819200510X (1992).
- 50 Sabine, T. M. & Howard, C. J. Determination of the oxygen x parameter in rutile by neutron powder methods. *Acta Crystallogr, Sect B* **38**, 701-702, doi:10.1107/S0567740882003896 (1982).
- 51 Kobayashi, H. et al. Order-Disorder Transition of BaM_2O_4 Bodies (M: La, Nd, Sm, Gd, Ho or Y) Synthesized by Sintering of $BaCO_3$ - M_2O_3 Mixtures. *J Ceram Soc Jpn* **102**, 583-586, doi:10.2109/jcersj.102.583 (1994).
- 52 Chtoun, E., Hanebali, L., Garnier, P. & Kiat, J. M. X-Rays and neutrons rietveld analysis of the solid solutions $(1-x)A_2Ti_2O_7$ - $xMgTiO_3$ (A = Y or Eu). *Eur J Solid State Inorg Chem* **34**, 553-561 (1997).

- 53 Mumme, W. G. & Wadsley, A. D. The structure of orthorhombic Y_2TiO_5 , an example of mixed seven- and fivefold coordination. *Acta Crystallogr, Sect B* **24**, 1327-1333, doi:10.1107/S0567740868004243 (1968).
- 54 Szymanik, B., Buckley, R. G., Trodahl, H. J. & Davis, R. L. Structure and decomposition of ceramic $\text{Ba}_3\text{Y}_4\text{O}_9$. *Solid State Ionics* **109**, 223-228, doi:10.1016/S0167-2738(98)00087-3 (1998).
- 55 Ong, S. P. *et al.* Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comp Mater Sci* **68**, 314-319, doi:10.1016/j.commatsci.2012.10.028 (2013).
- 56 Gale, J. D. & Rohl, A. L. The General Utility Lattice Program (GULP). *Mol Simulat* **29**, 291-341, doi:10.1080/0892702031000104887 (2003).
- 57 Dick, B. G. & Overhauser, A. W. Theory of the Dielectric Constants of Alkali Halide Crystals. *Phys Rev* **112**, 90-103, doi:10.1103/PhysRev.112.90 (1958).
- 58 Ruiz-Trejo, E., Islam, M. S. & Kilner, J. A. Atomistic simulation of defects and ion migration in LaYO_3 . *Solid State Ionics* **123**, 121-129, doi:10.1016/S0167-2738(99)00092-2 (1999).
- 59 Bush, T. S., Gale, J. D., Catlow, C. R. A. & Battle, P. D. Self-Consistent Interatomic Potentials for the Simulation of Binary and Ternary Oxides. *J Mater Chem* **4**, 831-837, doi:10.1039/jm9940400831 (1994).
- 60 Schröder, K. P., Sauer, J., Leslie, M., Catlow, C. R. A. & Thomas, J. M. Bridging Hydroxyl-Groups in Zeolitic Catalysts - a Computer-Simulation of Their Structure, Vibrational Properties and Acidity in Protonated Faujasites (H-Y Zeolites). *Chem Phys Lett* **188**, 320-325, doi:10.1016/0009-2614(92)90030-Q (1992).
- 61 Maglia, F. *et al.* Incorporation of trivalent cations in synthetic garnets $\text{A}_3\text{B}_5\text{O}_{12}$ (A = Y, Lu-La, B=Al, Fe, Ga). *J Phys Chem B* **110**, 6561-6568, doi:10.1021/jp055713o (2006).
- 62 Baetzold, R. C. Atomistic simulation of ionic and electronic defects in $\text{YBa}_2\text{Cu}_3\text{O}_7$. *Phys Rev B* **38**, 11304–11312, doi:10.1103/PhysRevB.38.11304 (1988).
- 63 Woodley, S. M., Battle, P. D., Gale J. D. & Catlow, C. R. A. The prediction of inorganic crystal structures using a genetic algorithm and energy minimisation, *Phys Chem Chem Phys* **1**, 2535-2542, doi:10.1039/a901227c (1999).
- 64 van Beest, B. W. H., Kramer, G. J. & van Santen, R. A. Force-fields for silicas and aluminophosphates based upon abinitio calculations, *Phys Rev Lett* **64**, 1955–1958, doi:10.1103/PhysRevLett.64.1955 (1990)]
- 65 Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys Rev B* **54**, 11169-11186, doi:10.1103/PhysRevB.54.11169 (1996).
- 66 Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys Rev Lett* **77**, 3865-3868, doi:10.1103/PhysRevLett.77.3865 (1996).
- 67 Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys Rev B* **59**, 1758-1775, doi:10.1103/PhysRevB.59.1758 (1999).
- 68 CrysAlis PRO (Agilent Technologies UK Ltd, UK, 2011).
- 69 Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr, Sect A* **64**, 112-122, doi:10.1107/S0108767307043930 (2008).
- 70 Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Crystallogr, Sect C* **71**, 3-8, doi:10.1107/S2053229614024218 (2015).
- 71 Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2: a complete structure solution, refinement and analysis program. *J Appl Crystallogr* **42**, 339-341, doi:10.1107/S0021889808042726 (2009).
- 72 TOPAS Academic: General Profile and Structure Analysis Software for Powder Diffraction v. 5 (Bruker AXS, Karlsruhe, Germany, 2007).
- 73 Petricek, V., Dusek, M. & Palatinus, L. Crystallographic Computing System JANA2006: General features. *Z Kristallogr* **229**, 345-352, doi:10.1515/zkri-2014-1737 (2014).
- 74 Barr, G., Dong, W. & Gilmore, C. J. High-throughput powder diffraction. II. Applications of clustering methods and multivariate data analysis. *J Appl Crystallogr* **37**, 243-252, doi:10.1107/S0021889804000391 (2004).
- 75 McKittrick, J. & Shea-Rohwer, L. E. Review: Down conversion materials for solid-state lighting. *J Am Ceram Soc*, **97**, 1327–1352, doi:10.1111/jace.12943 (2014).
- 76 de Mello, J. C., Wittmann, H. F. & Friend, R. H., An improved experimental determination of external photoluminescence quantum efficiency. *Adv Mater* **9** 230–232, doi:10.1002/adma.19970090308 (1997).
- 77 Bertaut, E. F., Blum, P. & Sagnieres, A. Structure of $\text{Fe}_2\text{O}_3\cdot\text{CaO}$ and brownmillerite. *Acta Crystallogr, Sect B* **12**, 149–159 doi:10.1107/S0365110X59000433 (1959).

- 78 Huang, Q., Karen, P., Karen, V. L., Kjekshus, A., Lynn, J. W., Mighell, A. D., Rosov, N. & Santoro, A. Neutron-powder-diffraction study of the nuclear and magnetic structures of yttrium barium iron oxide ($\text{YBa}_2\text{Fe}_3\text{O}_8$) at room temperature. *Phys Rev B* **45**, 9611–9619 doi:10.1103/PhysRevB.45.9611 (1992).
- 79 Demont, A., Dyer, M. S., Sayers, R., Thomas, M. F., Tsiamtouri, M., Niu, H. N., Darling, G. R., Daoud-Aladine, A., Claridge, J. B. & Rosseinsky, M. J. Stabilization of a Complex Perovskite Superstructure under Ambient Conditions: Influence of Cation Composition and Ordering, and Evaluation as an SOFC Cathode. *Chem Mater* **22**, 6598–6615, doi:10.1021/cm102475n (2010).
- 80 Barry, T. L. & Roy, R. New Rare Earth-Alkaline Earth Oxide Compounds - Predicted Compound Formation and New Families Found. *J Inorg Nucl Chem* **29**, 1243–1248, doi:10.1016/0022-1902(67)80365-8 (1967).