Supplementary Materials for

Computationally-assisted Identification of Functional Inorganic Materials

Matthew S. Dyer, Christopher Collins, Darren Hodgeman, Philip A. Chater, Antoine Demont, Simon Romani, Ruth Sayers, Michael F. Thomas, John B. Claridge, George R. Darling, and Matthew J. Rosseinsky

correspondence to: m.j.rosseinsky@liverpool.ac.uk

This PDF file includes:

Materials and Methods

Calculations

Classical force-field calculations were conducted using the General Utility Lattice Programme (GULP) (*33*), where the potential is split into two components; long range Coulombic interactions and a short range mostly repulsive component of the potential using Buckingham potentials with the form:

$$\varphi_{ij}(r_{ij}) = A_{ij}\exp\left(-\frac{r_{ij}}{\rho_{ij}}\right) - \frac{C_{ij}}{r_{ij}^6}$$

Polarizability is included in the model using the shell model (*34*) incorporated into GULP. The first iteration of the force-field was created using solely literature values to find the best combination to describe the $YBa_2Ca_2Fe_5O_{13}$ ($10a_p$) experimental structure (*35-39*). The force-field was then altered where necessary to recreate the experimental lattice parameters to within 5% and to reproduce the energetics calculated using DFT in the original report of the material (*26*). The resulting force-field parameters (A (eV) / ρ_{ij} (Å) / C_{ij} (eVÅ^{-6})) are 22764.3 / 0.149 / 42 for O-O, 20717.5 / 0.24203 / 0 for Y-O, 4818 / 0.3067/ 0 for Ba-O, 1244.5 / 0.3299 / 0 for Fe-O / 2272.741 / 0.2986 / 0 for Ca-O and 6276 / 0.22074/ 0 for Cu-O. Shells were attached to Ba, Ca and O cores with spring constants of 34.05, 34.05 and 42 eVÅ^{-2} and charges of 1.831, 1.281, −2.24 q_e respectively. Geometry relaxations were performed with a maximum of 1000 steps. Some initial structures failed to converge, and were rejected from consideration as final candidate structures. We observed that the rejected structures were generally very high in energy or contained unphysical geometries.

Density Functional Theory (DFT) calculations were performed using the Vienna Ab-intio Simulation Package (VASP) (*40*) using the PBE exchange correlation functional (*41*) and treating core-electrons using the projector augmented wave method (*42*). The plane-wave cut-off was set to 450 eV, forces on atoms were relaxed to less than 0.01 eV/Å and unit cell parameters relaxed. A Γ-centered *k*-point grid was used for each calculation, where the number of *k*-points in each direction was set to the lowest number to satisfy the condition:

Real-space lattice vector (Å) × *k-points on lattice vector* ≥ *20*

For AO layer metal atoms (Y, Ba and Ca) the first sub-valent *s* orbital was treated as valence, for BO_2 layer metal atoms (Fe and Cu) the first sub-valent *p* orbital was treated as valence. To correct for the known failings of standard functionals to treat localised *d*-electrons, we used the PBE+U method in the rotationally invariant form which uses only a single effective parameter, $U_{eff} = U - J$ (*43*). We used a U_{eff} parameter of 4.0 eV for both Cu and Fe as suggested in previous studies (*44*). Structures were set-up with G-type anti-ferromagnetic ordering, the magnetic structure of related observed phases.

The DFT calculations were performed to obtain accurate structures and relative energies of formation. Accordingly the DFT setup was first tested against experimental enthalpies of formation for $LaMO_3$ perovskites (*45*) and the computational setup optimized to complete the calculation as quickly as possible whilst still achieving the same result in terms of stabilities of the perovskites relative to their binary oxides. For calculations of electronic and magnetic properties not calculated in this paper (e.g.

accurate band gap energies) a more accurate method than PBE+U would probably be needed for the treatment of highly correlated electrons.

The creation of all possible permutations with 16 AO_{1-x} modules and 16 BO_{2-x} modules is computationally infeasible. Simply doubling the modules used for $Y_2Ba_2Ca_4Fe_8O_{21}$ results in over 10^8 total permutations. Instead structures with a $16a_p$ repeat were constructed by building all possible permutations in $8a_p$ with $2Ba_4O_4$ + $4Ca_4O_4$ + Y_4 + Y_4O_4 + $6Fe_4O_8$ + $2Fe_4O_4$ modules, and expanding to $16a_p$ using inversion symmetry to introduce body centring, which resulted in the 47,040 structures relaxed using force-fields.

For complex systems such as the ones studied here, it is very unlikely that a classical force-field can be used to completely reliably rank the energies of all the structures. We trust that the force-field is able to distinguish between structures which are grossly unstable, and those which are relatively stable. At this point we turn to the more reliable DFT to rank the lowest energy structures. In general, the cut off chosen for structures calculated using DFT will depend both on the reliability of the force-field, and the energy spectrum of structures.

In the case of the $10a_p$ material $YBa_2Ca_2Fe_5O_{13}$, the force-field calculations provided three structures with very similar energies, which were separated in energy from the fourth most stable structure. We therefore chose these three low energy structures for further DFT calculations, and also took the fourth lowest lying structure as an additional check. In this case, the stability of the lowest energy structure ranked by force-fields was found to be overestimated, and much reduced using DFT. The other three structures, however, were ranked in the same order with both methods.

For the calculations on the $8a_p$ material, a single structure was found to be much more stable than the others using force-fields, with an energy gap to the second structure much larger than the energy gaps found in the $10a_p$ case (0.42 eV/FU). Although it was likely that such a large energy difference could be trusted, we calculated the first 20 structures using DFT, since in this case the final result was unknown. The ranking of the 20 structures was similar in both cases, but with some change in order between structures at higher energies. The first and second most stable structures were the same using both force-fields and DFT.

Site Disorder

In many oxide crystal structures with two different cations or more, crystallographic sites are occupied by a random distribution of cations, rather than being solely occupied by one species. This is difficult to model in calculations using periodic boundary conditions, when structure and energetics are important the traditional method is to use the supercell approach. Here a single crystallographic unit cell is repeated to form a larger supercell and the cations are distributed over the periodic repeats of crystallographic sites. The supercell approach was combined with finite temperature Monte Carlo sampling by Zacate and Grimes (*28*) to calculate the expected fractional occupancy of cation sites within the Ca_2FeAlO_5 Brownmillerite structure at high temperature.

In the spirit of this work, we performed simulations starting from the ordered low energy $16a_p$ structure resulting from our EMMA calculations, repeated into a 3×3×1 supercell with 18 cations in each $AO_{1-\delta}$ layer. One random swap of a pair of Y, Ba and

Ca ions was made at each step, and the resulting structure relaxed using force-fields. If the new structure had a lower energy than the previous structure, it was kept in the next swapping step, otherwise it was accepted with a probability of $\exp(-\Delta E/kT)$, where ΔE is the energy difference in eV, and the temperature T was 1160 K. Four separate simulations were performed in parallel, each with 1100 swapping steps in total. The simulations were stopped once the final 100 steps produced no structures lower in energy.

The resulting lowest energy structure was 0.29 eV/FU lower in energy than the initial, fully ordered structure. Mixing occurred between the A1 and A3 sites (Table S1), such that their compositions became A1: $Y_{0.39}Ca_{0.61}$, A3: $Y_{0.31}Ca_{0.69}$. This is compared to the experimentally refined compositions of A1: $Y_{0.63}Ca_{0.37}$ and A3: $Ca_{0.77}Y_{0.13}Ba_{0.10}$. The other three sites kept their initial compositions, of A2: Ba, A4: Ca and A5: Y, in contrast to the experimentally refined structure which found mixed site occupancies on all A-sites.

Although the site swapping procedure does produce a structure which was predicted to be more stable at 0 K, and which contains some of the site mixing seen experimentally, there is clearly considerable difference still between the occupancies produced in calculations and that refined in experiment. Some of these differences could be improved by using a larger supercell, or by integrating occupancies at high temperature during the Monte Carlo simulation to reproduce the occupancies at synthesis temperatures, assuming that these are then kinetically trapped in upon cooling. In addition, more accurate relative stabilities would be expected if DFT calculations were used at each step rather than force-field calculations. Further complications arise from the non-stoichiometric composition of the experimentally obtained pure material, which results in more Y and Ba and less Ca than in the stoichiometric case. All of these issues could be addressed, but at significantly increased computational cost. In fact, even the relatively simple site-swapping procedure described above is prohibitively expensive to use for all of the structures relaxed during more EMMA calculations. As available computational resources are increased in the future, it will be possible to use such methods to improve the theoretical treatment of complex oxide structures, and so improve the ability to rank structures generated during EMMA calculations.

Synthesis

Polycrystalline samples were prepared via a direct solid state reaction of Y_2O_3 (99.999%), $BaCO_3$ (99.95%), $CaCO_3$ (99.995%), Fe_2O_3 (99.945%) and CuO (99.995%), all sourced from Alfa Aesar, mixed and ground by hand in the desired cationic ratios and annealed at 1200°C alumina crucibles lined with platinum foil under ambient air atmosphere. Several cycles of regrinding and firing were performed to ensure phase homogeneity and to complete the reaction process. Phase identification and purity was examined by powder X-ray diffraction collected on a Panalytical diffractometer using Co $K\alpha_1$ radiation in Bragg Brentano geometry.

During synthesis of **1** and **2**, samples were heated at 5°C per minute to 1200°C and cooled at approximately 10°C per minute. Two cycles of heating and regrinding were performed with heating times of 12 and 72 hours. Synthesis using this heating regime at the composition of **1** ($Y_{2.24}Ba_{2.28}Ca_{3.48}Fe_{7.44}Cu_{0.56}O_{21}$) reproducibly produces >80% **1** on an approximately 5 g scale, alongside $10a_p$ and Y_2O_3 impurities. Phase pure **1** was accessible at the same composition by rapidly quenching the sample from 1200°C to room temperature by removing the sample from the furnace and placing it into a metal

block (Fig. S10). Phase pure samples were only possible on a smaller scale of a few grams. Structural refinements were performed on a sample produced on a 5 g scale containing >80% **1** alongside minor impurities. Physical properties measurements were performed on samples containing >95% **1** synthesised by quenching on a smaller scale.

Energy-dispersive X-ray spectroscopy (EDX) and selected area electron diffraction (SAED) were performed using a JEOL 2000FX TEM on the $Y_{0.9}Ba_{1.7}Ca_{2.4}Fe_{4.5}Cu_{0.5}O_{13}$ sample synthesised to investigate Cu substitution into $YBa_2Ca_2Fe_5O_{13}$ revealed a small number of crystallites with an EDX-derived composition of $Y_{2.1(2)}Ba_{1.8(2)}Ca_{4.1(2)}Fe_{7.4(2)}Cu_{0.6(2)}$ and a longest unit cell repeat of $8a_p$. Using a ternary phase diagram representation of the A-site cation composition, the composition required to give a pure phase was targeted by synthesising a range of compositions around the EDX composition (Fig. 4F). After optimisation of the A-site composition to produce **1**, EDX gave a cation composition of $Y_{2.07(15)}Ba_{2.49(20)}Ca_{3.44(5)}Fe_{7.63(18)}Cu_{0.37(7)}$ which is close to the targeted composition of $Y_{2.24}Ba_{2.28}Ca_{3.48}Fe_{7.44}Cu_{0.56}$.

Diffraction

Time-of-flight neutron diffraction (ND) data were collected from the backscattering (2θ = 168°, *d*-spacing range 2.6 to 0.65 Å), 2θ = 90° (*d*-spacing range 4.0 to 0.85 Å) and 2θ = 30° (*d*-spacing range 2.6 to 10 Å) banks on the HRPD instrument at the ISIS facility, Rutherford Appleton Laboratories. Synchrotron X-ray diffraction (SXRD) data were obtained on Beamline I11 at Diamond Light Source, UK, at a wavelength of 0.826145(3) Å over a 2θ range of 2 to 150° and a step size of 0.003°.

Powder diffraction data were analyzed in a combined X-ray and neutron Rietveld refinement using the computer program GSAS (*46*). The starting model for Rietveld refinement of **1** was derived from the lowest energy $16a_p$ $Y_2Ba_2Ca_4Fe_8O_{21}$ structure from EMMA. The *P*1 structure from EMMA was transformed into the experimentally (SAED) determined symmetry of *Imma* using FindSym (*47*) to determine the unit cell origin and merge atoms within a tolerance of 0.25 Å, giving an observed symmetry of $P2_1/c$, and then Vesta (*48*) to impose *Imma* symmetry and remove duplicate atoms. The resulting *Imma* model contains disordering of oxygen positions around the tetrahedral Fe site (site B3) and one of the octahedral Fe sites (site B4). This transformation to give the starting model for Rietveld refinement of **1** is represented in Fig. S5.

The ND time-of-flight diffractometer zero offset was refined for the 168° bank, all three constants were refined for the 90° and 30° banks 2 and 3, and the zero point was refined for the SXRD data giving a total of 8 refined diffractometer parameters. A total of 42 background parameters were refined; 4 for each ND histogram; 24 for SXRD high 2θ; 6 for SXRD low 2θ. Lattice parameters for **1** were refined separately for ND and SXRD data. Bragg reflections matching $10a_p$ and Y_2O_3 impurities were also observed in the sample. Structures for $10a_p$ $Ba_{1.6}Ca_{2.3}Y_{1.1}Fe_5O_{13}$ (*26*) (including the magnetic structure) and Y_2O_3 (*49*) were included in the Rietveld refinement and determined to be present at 15.0(2) weight% and 3.67(14) weight%, respectively, alongside the dominant $16a_p$ phase, **1**. Only the lattice parameters, scale factors and profile of the impurity phases were refined.

In order to improve the peak profile fitting of the broad, asymmetric low order Bragg reflections the SXRD data-set was split into two histograms covering low 2θ and high 2θ ranges with a *d*-spacing range of 0.90 to 6.76 Å and 6.76 to 36 Å, respectively.

Anisotropic strain broadening (*50*) was refined for the $16a_p$ phase in the high 2θ SXRD data-set and the highest resolution (168°) ND data-set to account for slight differences in peak shape for reflections corresponding to the long stacking axis. A total of 39 profile parameters were used in the refinement for all histograms and phases; 11 for ND 168°; 6 for ND 90°; 3 for ND 30°; 15 for SXRD high 2θ; 4 for SXRD low 2θ.

The relative occupancy of Y, Ba and Ca for each A-site were refined while constraints were applied to force a total occupancy of unity on each A-site. A soft chemical composition restraint was implemented to drive the refinement towards the experimental A-site composition of $Y_{2.24}Ba_{2.28}Ca_{3.48}$. A single isotropic displacement parameter was refined for each atom type; A-site (Y/Ba/Ca), B-site (Fe/Cu) and O. The atomic coordinates of each atom was refined within the symmetry allowed by each Wyckoff position. During the refinement, six of the disordered O positions from the EMMA structure moved close to a higher symmetry Wyckoff position. These atoms were moved onto the higher symmetry sites to reduce the number of refined atomic coordinates.

Possible B-site magnetic structures were determined by co-representational analysis using the computer program SARAh (*51*) and were screened against the observed neutron scattering data. The best fit was observed for a G-type antiferromagnetic structure (propagation vector k = 0,0,0) with moments (anti)parallel to the *a*-axis. The magnitude of the magnetic moment along the *a*-axis on each B site was refined in magnetic space group *Pmn´a*; sites B2 and B3 (see Fig. S5) refined to a moment consistent with occupation by Fe^{3+} (3.89(8) μ_B and 3.97(6) μ_B, respectively), while sites B1 and B4 refined to a significantly lower moment (2.89(10) μ_B and 2.48(11) μ_B, respectively). As Cu and Fe are practically indistinguishable by X-ray and neutron atomic scattering, the Cu position was inferred on sites B1 and B4 from the refined magnetic moment, as Cu^{2+} has a significantly lower magnetic moment than Fe^{3+}. The experimental Cu composition was shared equally between the two identified Cu positions (sites B1 and B4) to give two sites with occupancies of $Fe_{0.86}Cu_{0.14}$ which were not refined.

The result of the joint Rietveld refinement of **1**, together with their starting values from EMMA, are shown in Table S1 and a Rietveld plot is shown in Fig. S13. CCDC 887926 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif, or by emailing data_request@ccdc.cam.ac.uk, or by contacting The Cambridge Crystallographic Data Centre, 12, Union Road, Cambridge CB2 1EZ, UK; fax: +44 1223 336033.

Mössbauer spectroscopy

Mössbauer data were obtained at 300 K using a conventional constant acceleration Mössbauer spectrometer incorporating a ≈25 mCi source of ^{57}Co in a Rh matrix (Fig. S7). The observed data were fitted to Fe^{3+} in octahedral, square pyramidal and tetrahedral environments in a 2:1:1 ratio consistent with the calculated and observed crystal structure of **1**.

Titrations to determine oxygen content

Iodometric titrations were performed to determine the oxygen content in compound **1**. Approximately 50 mg of material was dissolved in 60 ml of 1 M hydrochloric acid with an excess of potassium iodide, in an argon atmosphere. Upon complete dissolution,

the solution was titrated with 0.1 M sodium thiosulphate. As sodium thiosulphate was added, the solution turned from orange to clear; starch solution was added just before the end point, where the solution turns from dark purple to clear. Prior to titrations of compound **1**, the sodium thiosulphate solution was standardised by titrating it against potassium iodate, which was dissolved in distilled water with potassium iodide and 1 M sulphuric acid. Titrations were repeated three times to ensure reproducibility.

Assuming reduction to Fe^{2+} and Cu^{+} species, and using the nominal composition for the cations, the oxygen content was calculated giving a full composition of $Y_{2.24}Ba_{2.28}Ca_{3.48}Fe_{7.44}Cu_{0.56}O_{20.92(2)}$. This corresponds to an average Fe/Cu charge state of 2.95+ and is in good agreement with the fact that attempts to refine oxygen content away from full site occupancy during Rietveld refinement were unsuccessful and the observation of only Fe^{3+} environments in the Mössbauer data. Formal charge states of Fe^{3+} (consistent with Mössbauer spectra) and Cu^{2+} would yield in an oxygen content of $O_{20.84}$, notably close to the observed value.

Electron microscopy

Electron diffraction (ED) and energy-dispersive spectroscopy (EDS) analyses were carried out with a JEOL 2000FX electron microscope equipped with an EDAX analyzer. High-angle annular dark field scanning transmission electron microscopy (HAADF-STEM) was performed with a JEOL 2100FCs Field Emission Transmission Electron Microscope with CEOS spherical aberration probe corrector, operated at 200kV. The high-angle annular dark field (HAADF) images were acquired using inner and outer collection semi angles of 74 and 196 mrad, respectively. Environmental noise sources were removed by using an annular filter in the FFT of the raw image, which improves the image qualitatively, while flattening the intensity response across the image field. The beam convergence semi angle was 14 mrad and spherical aberration (C_s) was controlled to within +/− 10 microns.

Simulated HAADF-STEM images were produced using the computer program QSTEM (*52*) using the same instrumental parameters as those used in the experiment. In order to produce a representative image a $10a \times b \times 10c$ super-cell of the $16a_p$ unit cell was constructed from the structure determined by Rietveld refinement. Within each A-site layer, individual A-sites were randomly assigned an atom type (Y, Ba or Ca) which were consistent with the experimentally determined atomic composition for that layer.

Preliminary structural characterizations were performed on the $16a_p$ material by selected area electron diffraction (SAED) (Fig. S3). A set of subcell reflections characteristic of a cubic perovskite structure was observed. Additionally, sharp superstructure reflections show the existence of large inter-reticular distances along one of the directions, in agreement with low angle peaks observed by PXRD. Reconstruction of the reciprocal space was carried out by rotation around this long axis, yielding a large perovskite superstructure with unit cell dimensions $a_p\sqrt{2} \times 16a_p \times a_p\sqrt{2}$, $\alpha = \beta = \gamma = 90°$ (where a_p is the cubic perovskite unit cell parameter), while the diffraction spots are controlled by the condition *hkl*: $h + k + l = 2n$, *hk*0: $h,k = 2n$ consistent with space groups *Imma* and *I2mb*. In parallel EDX was performed on 18 crystals of this phase (the presence of the $16a_p$ repeat was checked for each crystal prior to elemental analysis) allowing determination of an average cationic composition. Given the obvious presence of Cu Kα radiation on all the EDX spectra (on a Nickel TEM grid), a copper content was

taken into account despite a large relative error due to weakness of the peak intensity. Normalizing the EDX results to eight ABO_3 perovskite units yields the cation composition $Y_{2.07(15)}Ba_{2.49(20)}Ca_{3.44(5)}Fe_{7.63(18)}Cu_{0.37(7)}$.

Solid oxide fuel cell cathode processing and measurement

DC conductivity

To achieve samples with a high relative density, powder of **1** was ballmilled in isopropanol for 12 hours, uniaxially pressed into pellets and isostatically pressed (Autoclave Engineers Cold Isostatic Press) under a pressure of 200 MPa followed by sintering in air at 1100°C for 48 hours. The relative density was greater than 98 %, measured using an Archimedean balance. The circular pellets were cut into bars for electrical conductivity measurements using the standard DC four-probe method (Keithley 2182 Nanovoltmeter and Keithley 220 Current Source). Gold wires were attached to the bars (held in position with a small amount of gold paste) to make the I-V probes with four-in-a-line contact geometry. The measured DC conductivity is shown as a function of temperature in Fig. S9.

AC impedance

$Ce_{0.9}Gd_{0.1}O_{2-\delta}$ electrolyte substrates were obtained from commercial powder (Fuel Cell Materials) by uniaxial pressing into disks (3 tonnes), and sintered at 1400°C for 5 hours. The relative density of the electrolyte substrates (measured using an Archimedean balance) was 98-99 %.

The cathode inks **1** were prepared by ballmilling the powder with an organic binder (Fuel Cell Materials) for 18 hours, in the weight ratio 67:33 powder to binder. 2 g of powder was milled using a Fritsch pulverisette planetary mill at 350 rpm for 24 hours. 16 5 mm Zirconia balls were used as the grinding media in zirconia pots, along with 8 ml of isopropanol. The electrode ink was applied to both surfaces of the electrolyte substrate by screen printing. The cell was dried at 100°C for 1 hour between the application of each layer followed by a final heating in air at 950°C for 1 hour to achieve good adherence of the electrodes to the electrolyte surface. From SEM measurements, the thickness of the deposited cathode after the final heating was 20 μm. The contacts for the electrical measurement were gold wire and gauze fixed in position with gold paste, which were attached to the cell by heating to 800°C for 1 hour.

AC impedance measurements were recorded over the frequency range 1 MHz to 0.01 Hz using a Solartron 1260 FRA with a modulation potential of 10 mV, over the temperature range of 773-1073 K in static air. The symmetrical cell was held for 90 minutes at each temperature to allow thermal equilibration and measurements were made using ZPlot v.2.9b (Scribner Associates) every 50°C. The area specific resistance (ASR) of the cathode was calculated by normalizing the measured resistance for the electrode area and dividing by two to take into account the symmetry of the cell.

Thermal stability

Thermal stability tests of **1** were performed in air on the as-synthesised powder and on powder which was ballmilled in isopropanol for 20 hours. Samples were annealed at 750°C and 950°C in air for 5 hours and subsequent phase stability was characterized by

PXRD (Fig. S8). There was no visible sign of degradation of the $16a_p$ phase, **1**, after the thermal treatment.

To study chemical compatibility with the electrolytes $Ce_{0.9}Gd_{0.1}O_{2-\delta}$ (GDC, Fuel Cell Materials), $Ce_{0.8}Sm_{0.2}O_{2-\delta}$ (SDC, Fuel Cell Materials) and $La_{0.9}Sr_{0.1}Ga_{0.8}Mg_{0.2}O_{3-\delta}$ (LSGM, PI-KEM LTD), milled powder of **1** and electrolyte were mixed in a weight ratio 1:1 by hand grinding, pressed into pellets and calcined at 950°C for 5 hours. The pellets were then ground to powders for characterization by PXRD.

Supplementary Text

Stability of $16a_p$ phases

We calculated the energy of the predicted $16a_p$ structure with composition $Y_2Ba_2Ca_4Fe_8O_{21}$ using DFT, and with some mixing of the A-sites. To capture some of the observed A-site mixing, the oxygen deficient A-site layers (A1) were constructed with 50% Y and 50% Ca, with the remaining Y distributed within the majority Ca layers neighboring the Fe T_d sites (A3). The calculated energy was then compared with the energy of known binary and ternary and quaternary oxides. Although $Y_2Ba_2Ca_4Fe_8O_{21}$ was more stable than the binary oxides, $Y_2O_3 + 2BaO + 4CaO + 4Fe_2O_3$ by 2.35 eV/FU, it was unstable relative to the ternary oxides, $2YFeO_3 + 2Ca_2Fe_2O_5 + Ba_2Fe_2O_5$ by 0.13 eV/FU and relative to $YFeO_3 + Ca_2Fe_2O_5 + YBa_2Ca_2Fe_5O_{13}$ by 0.17 eV/FU. Thus DFT calculations at 0 K predict that a structure with the $16a_p$ repeat should not form at this composition. Although this seems at odds with the experimental isolation of **1** and **2**, these may well be metastable phases at low (i.e. lower than synthesis) temperature. The appearance of $10a_p$ in slowly cooled samples of $16a_p$ (Fig S8) shows that $10a_p$ is more stable than $16a_p$ at temperatures below the synthesis temperature of 1473 K. The formation of pure $16a_p$ only via quenching indicates that $16a_p$ is indeed metastable. This is in agreement with the $16a_p$ being unstable in the 0 K DFT calculations, and with temperature-dependence of the phase equilibria, suggesting a key controlling role for the entropy in determining phase stability in this realistic complex oxide synthesis example. Indeed, a calculation of mixing entropy based on the experimentally refined fractional compositions of the A-sites in $Y_{1.1}Ba_{1.6}Ca_{2.3}Fe_5O_{13}$ (2.26×10^{-4} eV/FU/K) and **1** (4.45×10^{-4} eV/FU/K) gives an entropy difference of 2.19×10^{-4} eV/FU/K, corresponding to an energetic contribution of 0.32 eV at the synthesis temperature of 1473 K favoring the observed phase **1**. This is more than enough energy to compensate for the calculated enthalpy difference in DFT of 0.17 eV/FU, and to explain the ability to synthesize **1** and **2** by quenching from 1473 K. The mixing entropy, S_{mix}, was calculated using the equation $S_{mix} = -k_B \Sigma N_a x_{ia} \ln x_{ia}$, where k_B is the Boltzmann constant, N_a is the number of times site a appears in the full unit cell, and x_{ia} is the refined fractional occupancy of species i at the site a. Contributions are summed over all species and A sites.

Searching a phase diagram

To demonstrate that EMMA could be used to investigate phase diagrams, we have searched the (Y, Ba, Ca, Fe^{3+}, O) phase diagram for layered phases with $7a_p$ long axis repeats using force-fields within the EMMA method as described for $Y_2Ba_2Ca_4Fe_8O_{21}$ and at the compositions shown in Fig. S11. We chose the three lowest energy candidate structures for each of the compositions $YBa_2Ca_4Fe_7O_{18}$, $YBa_4Ca_2Fe_7O_{18}$ and $YBa_3Ca_3Fe_7O_{18}$ which had the best stability calculated using force-fields relative to the

ternary oxides $YFeO_3$ + $x$$Ca_2Fe_2O_5$ + $y$$Ba_2Fe_2O_5$, and recalculated the energies using DFT. The lowest energy structure for each composition, shown in Fig. S11, were structurally similar, with two tetrahedral Fe layers with opposing orientations and one square pyramid Fe layer all separated by octahedral Fe layers. Using DFT, the material with composition $YBa_2Ca_4Fe_7O_{18}$ was stable relative to the ternary oxides by 0.02 eV/FU, but was unstable relative to $Ca_2Fe_2O_5$ and $YBa_2Ca_2Fe_5O_{15}$ by 0.02 eV/FU. $YBa_4Ca_2Fe_7O_{18}$ and $YBa_3Ca_3Fe_7O_{18}$ were both found to be unstable relative to ternary oxides and $YBa_2Ca_2Fe_5O_{15}$ when using DFT. Attempts to synthesize a material with a $7a_p$ repeat failed at all compositions, and only phases with $10a_p$ and $16a_p$ repeats are isolable. In this case we must assume that entropic effects are not sufficient to stabilize the $7a_p$ phase at high temperatures relative to competing phases. A more thorough examination of the phase diagram would be possible by searching for phases with all possible repeats (taking into account the fact that multiples of smaller cells occur in larger ones). For example, all phases with repeats up to $12a_p$ could be found by searching for $7a_p$, $9a_p$, $10a_p$ (includes $2a_p$ and $5a_p$), $11a_p$ and $12a_p$ (includes $3a_p$, $4a_p$ and $6a_p$). The ability to accurately predict isolable phases would rely on a better treatment of finite temperature effects than simple 0 K DFT calculations, however, extending the methodology used for simpler systems (*53*).

Extracting Information from Ranked Structures

A large database of different structures and their properties results from the construction and ranking of structures using EMMA. It is possible to query this database and obtain general information which can both aid in the understanding of the materials and structural families studied and can allow the EMMA structural search to be refined.

As an example, we looked for correlations between the cell volume, atomic coordination numbers and structural stability in the 9,863 fully relaxed structures obtained from EMMA calculations on $Y_2Ba_2Ca_4Fe_8O_{21}$ with a $16a_p$ repeat. The resulting plots are shown in Fig. S12. One might expect that the densest structures would be the most stable for materials dominated by ionic interactions, however the energy versus volume plot in Fig. S12A shows that the structures with the lowest volume are not those with the lowest energy. Further information is gained by looking at the ratio between the average Ca and Y coordination numbers in each structure. We find that, in general, the structures with the lowest volumes have Ca in low coordination environments and that those with lowest energies have Y in low coordination environments. This reflects the known preference for Y^{3+} ions to be 8-coordinate, even though this has a cost of lowered density compared to 8-coordinate Ca^{2+} ions. It is interesting to consider how this difference might be altered by the application of external pressure, under which conditions the denser structures involving 8-coordinate Ca^{2+} ions might be preferred. The result also supports our choice of starting with modules which have O vacancies in Y rather than Ca layers for our search for stable structures at ambient pressure.

For each structure, and for each atom within, we plotted the structure's energy against the coordination number of that atom. For the purposes of this analysis, the coordination number is defined as the number of counter ions within a distance of 145% of the sum of the ionic radii. This gives us a picture of which coordination environments are preferred for each species in low energy structures, as shown in Fig. S12B-F. We find that low energy structures tend to have Y in 8- and/or 12-coordinate sites, Ba in 12-

coordinate sites, Ca in 10-coordinate sites and O in 6-coordinate sites. There appears to be little preference for Fe between 4-, 5- and 6-cordinate sites. These results are perhaps not surprising, given the expected coordination environments for these ions, but do give a statistical backing to what chemical intuition suggests. The different preferences for the A-site cations and the flexibility of the Fe^{3+} ion are probably the determining factors in the creation of a complex structure with so many different coordination environments for the cations for this choice of elements.

We chose to investigate relationships between cell volume, coordination environment and stability, however, there are in principle many other correlations that could be investigated between structural parameters (e.g. cell size and shape, coordination environments, ion-ion distances) and calculated properties (e.g. stability, band gap, magnetic moment), depending on the particular choices made during the EMMA calculations. The wealth of information generated by EMMA calculations is, in itself, a benefit of the EMMA method.

Source code

An implementation of the EMMA method in python is available online at http://pcwww.liv.ac.uk/~msd30/software/EMMA.html and depends on the Atomic Simulation Environment (ase) https://wiki.fysik.dtu.dk/ase/.

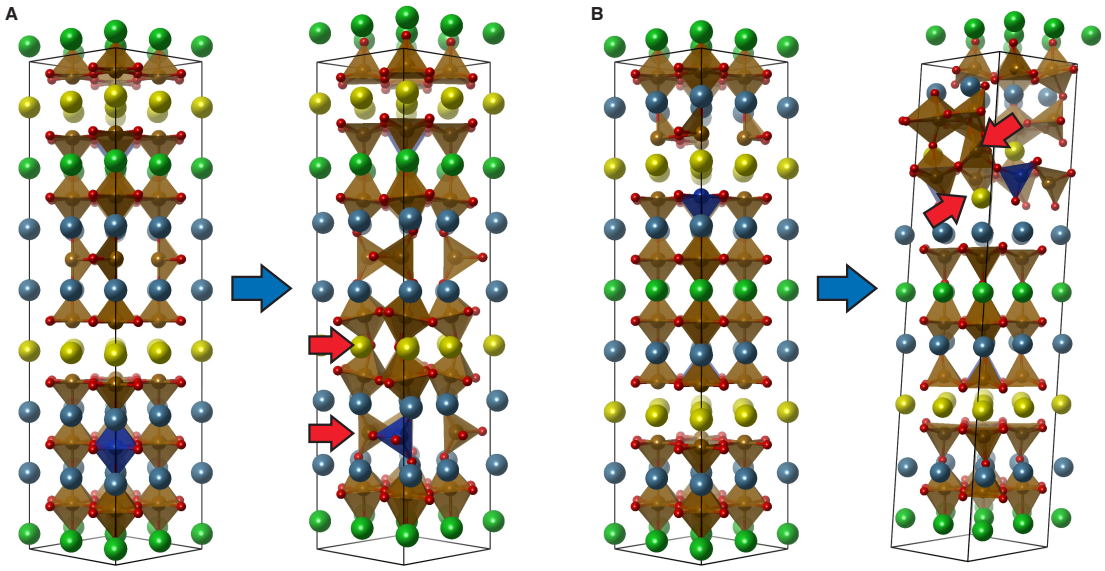

Fig. S1.

The initial and final geometries of two candidate structures with composition $Y_2Ba_2Ca_4Fe_{7.5}Cu_{0.5}O_{21}$ are shown. The structure in A) is the most stable resulting structure and shows oxygen migration over a considerable distance moving oxygen vacancies into different layers than in the initial structure (red arrows). B) shows a structure much higher in energy which relaxes to a clearly non-orthorhombic final structure in which Y and Fe atoms have swapped layers (red arrows).

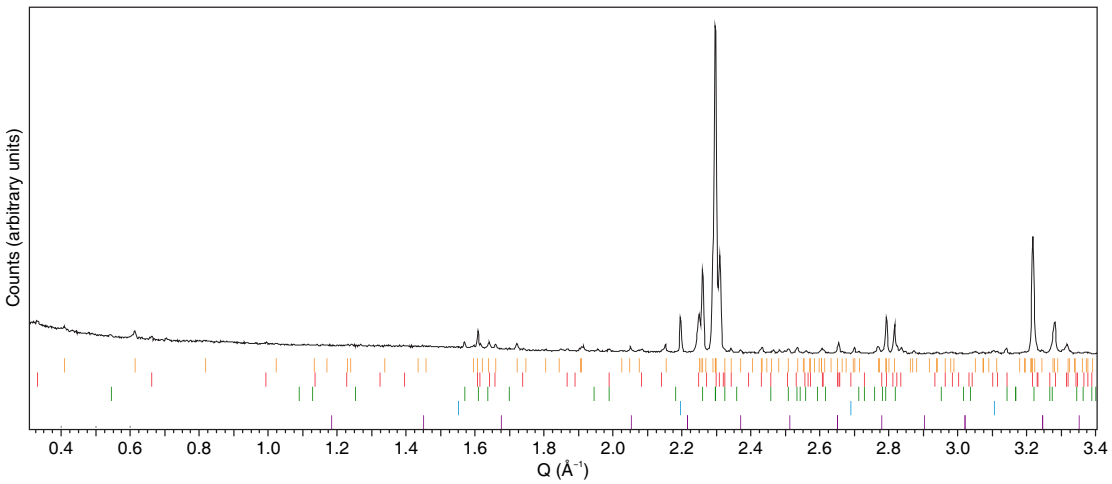

Fig. S2.

An X-ray powder diffraction pattern of the material resulting from synthesis at the DFT studied composition $Y_2Ba_2Ca_4Fe_{7.5}Cu_{0.5}O_{21}$. The tick marks represent: the apparently $8a_p$ structure as the majority phase (≈57.2 %, orange), a structure with $10a_p$ repeat (≈26.4 %, red), one with a $3a_p$ repeat (≈13.9 %, green), a cubic perovskite (2.2%, blue) and Y_2O_3 (0.3 %, purple). Estimated weight fractions were determined for each phase present by Rietveld refinement.

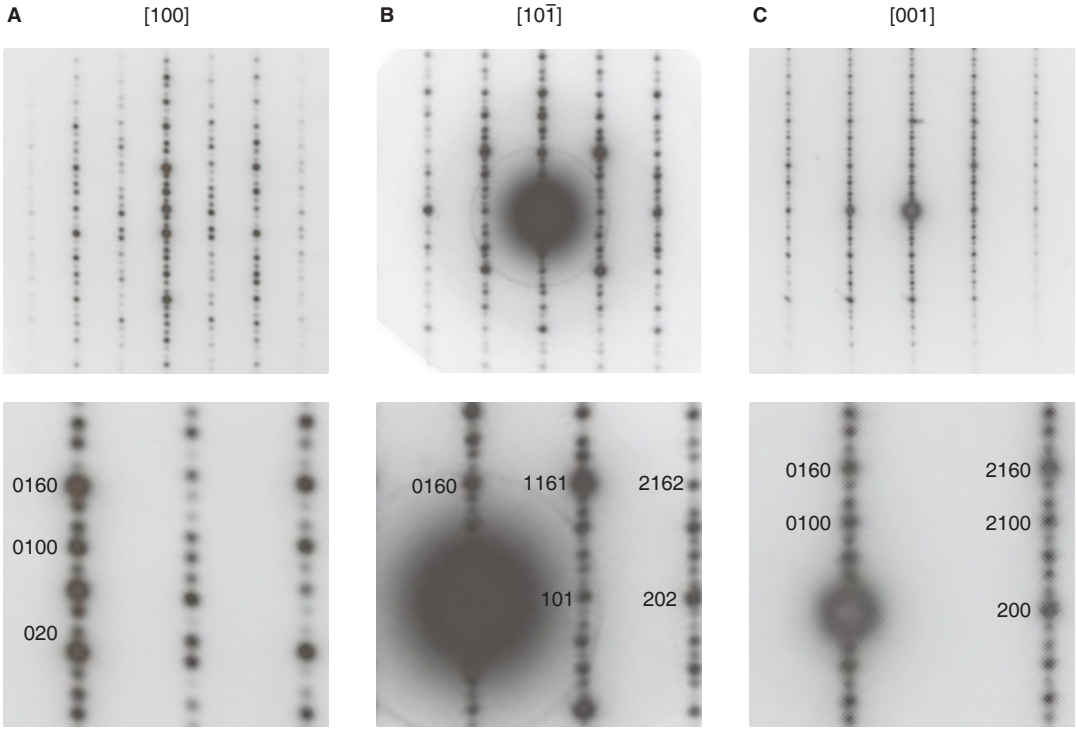

Fig. S3.

A) [100], B) [$10\bar{1}$], and C) [001] experimental SAED patterns (above) with their respective selected portions (below) showing indexing of the $16a_p$ unit cell.

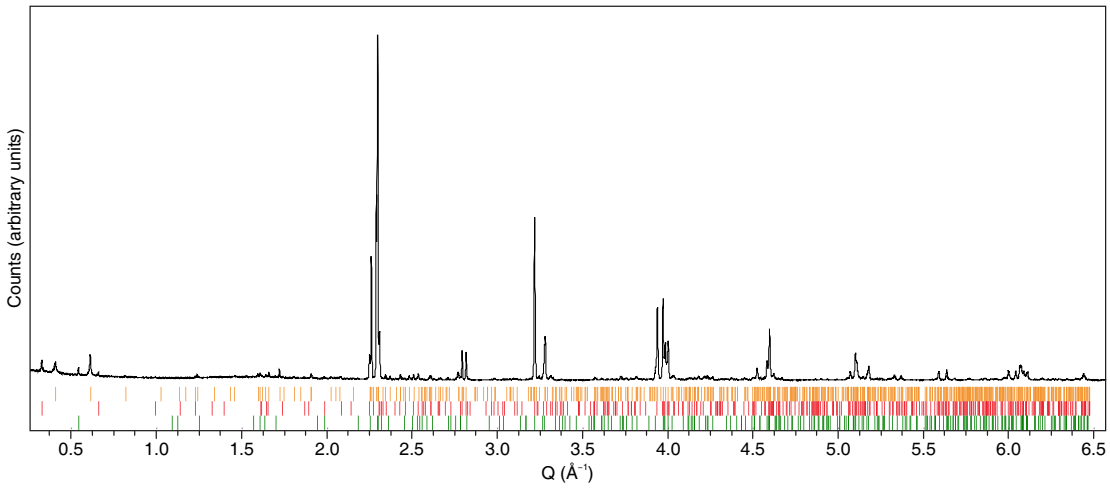

Fig. S4.

A synchrotron X-ray powder diffraction pattern of the Cu free compound **2** with composition $Y_{1.95}Ba_{2.1}Ca_{3.95}Fe_8O_{21}$. The tick marks represent: the $16a_p$ phase, **2**, as the majority phase (≈80 %, orange), a phase with a $10a_p$ repeat (≈10 %, red) and one with a $3a_p$ repeat (≈10 %, green). Estimated weight fractions were determined for each phase present by Rietveld refinement.

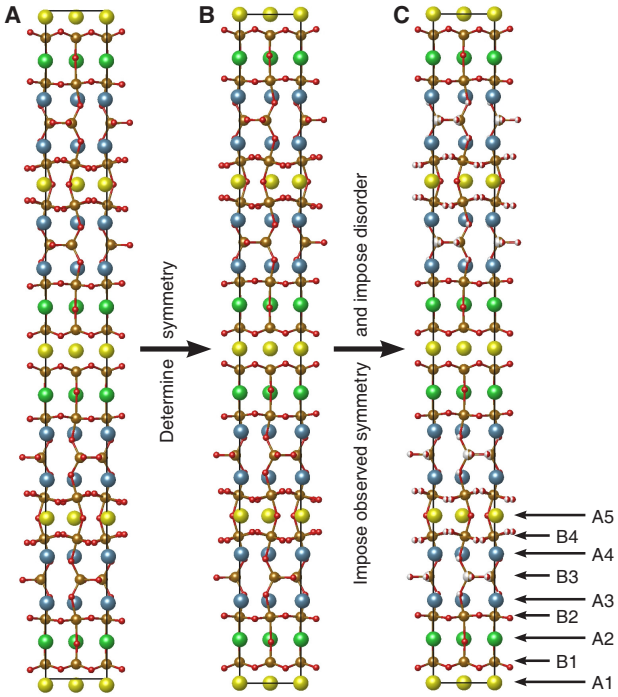

Fig. S5.

Generation of the symmetrical 16a_p model for Rietveld refinement. Atoms are colored as follows: Blue Ca, green Ba, yellow Y, brown Fe and red O. The unit cell is shown in black. A) *P*1 structure from EMMA. B) $P2_1/c$ model after symmetry interrogation with FindSym. C) *Imma* model used for Rietveld refinement, including disordered oxygen positions. Disordered, half occupied O and Fe sites are shown as half white. A-site and B-site labels used throughout the text are shown in the bottom right hand corner.

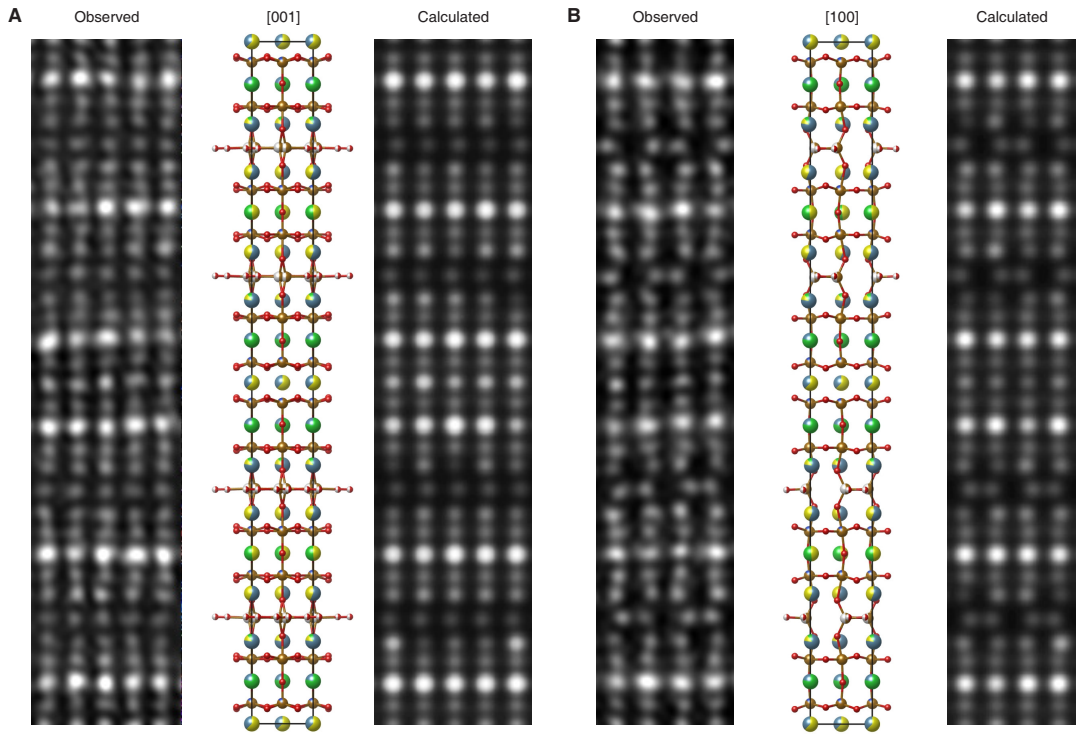

Fig. S6.

Observed and calculated HAADF-STEM images of **1**, as viewed down the A) [001] and B) [100] axes. The refined crystal structure is shown for comparison.

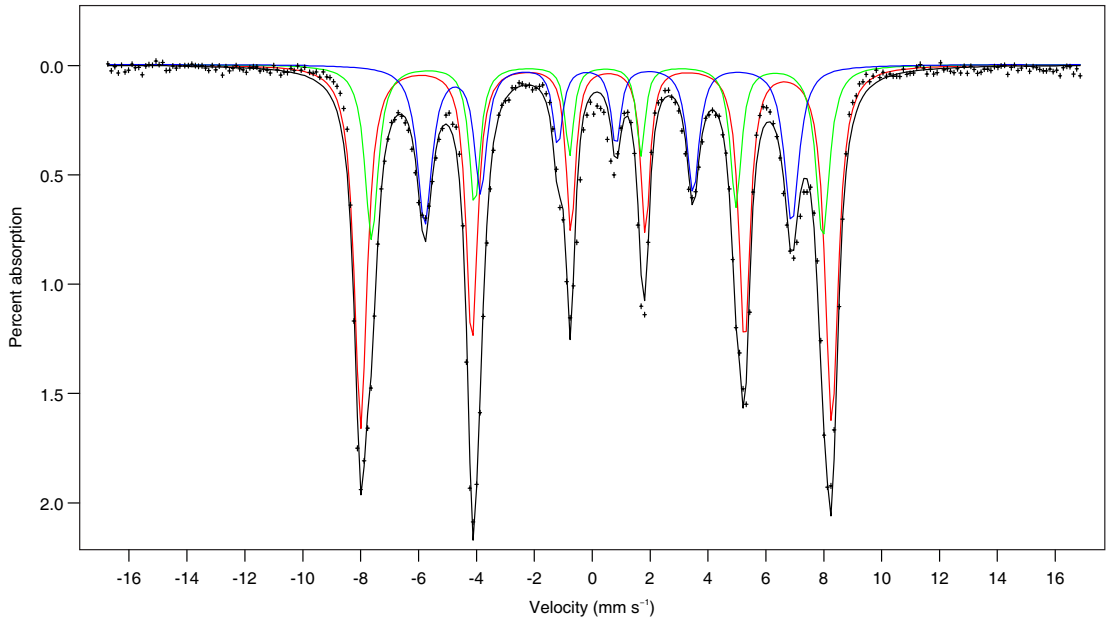

Fig. S7.

^{57}Fe Mossbauer spectrum of **1**, $Y_{2.32}Ba_{2.16}Ca_{3.52}Fe_{7.44}Cu_{0.56}O_{21}$, showing the observed data (black crosses) and calculated fit (black line). Individual components of the calculated fit corresponding to octahedral (red line), square pyramidal (green line) and tetrahedral (blue line) Fe^{3+} are shown. $\chi^2 = 9.78$ for 15 variables.

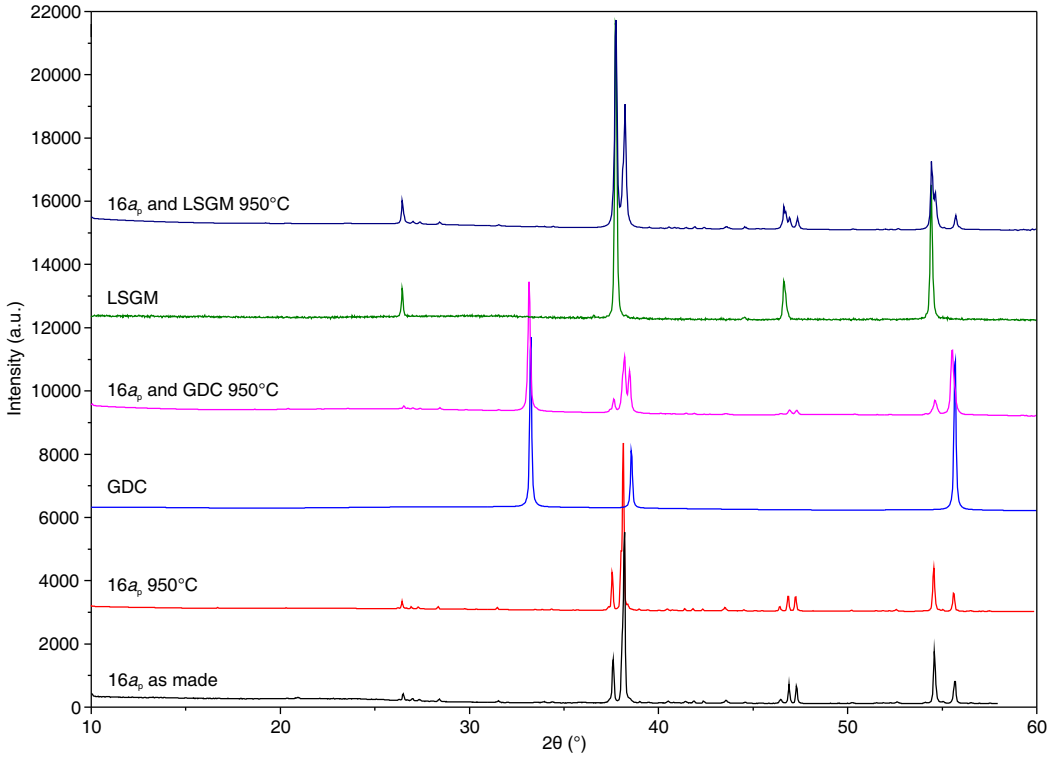

Fig. S8.

X-ray diffraction patterns showing the stability of the synthesised 16a_p material, **1**. The patterns show that the material is both thermally stable and compatible with the two common electrolyte materials, $Ge_{0.9}Gd_{0.1}O_{2-\delta}$ (GDC) and $La_{0.9}Sr_{0.1}Ga_{0.8}Mg_{0.2}O_{3-\delta}$ (LSGM).

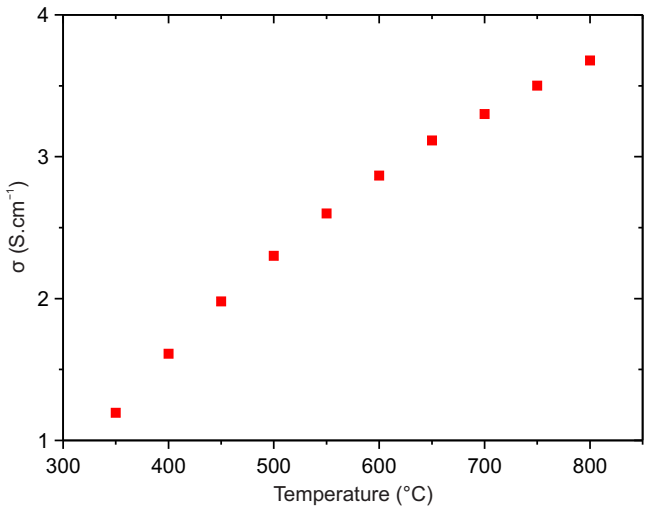

Fig. S9.

The DC conductivity of $Y_{2.24}Ba_{2.28}Ca_{3.48}Fe_{7.44}Cu_{0.56}O_{21}$ (**1**) collected upon cooling from 800°C to 500°C.

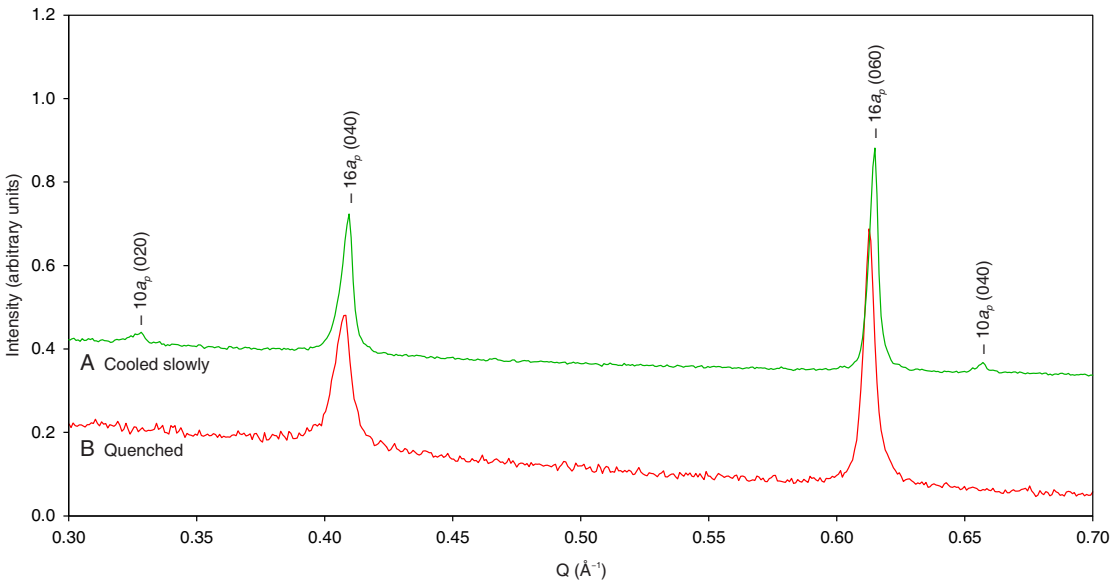

Fig. S10.

Powder synchrotron X-ray diffraction patterns of $Y_{2.24}Ba_{2.28}Ca_{3.48}Fe_{7.44}Cu_{0.56}O_{21}$ samples produced by A) cooling at approximately 10°C per minute and B) quenched rapidly to room temperature. Bragg reflections due to the $16a_p$ phase, **1** and the $10a_p$ phase are indicated.

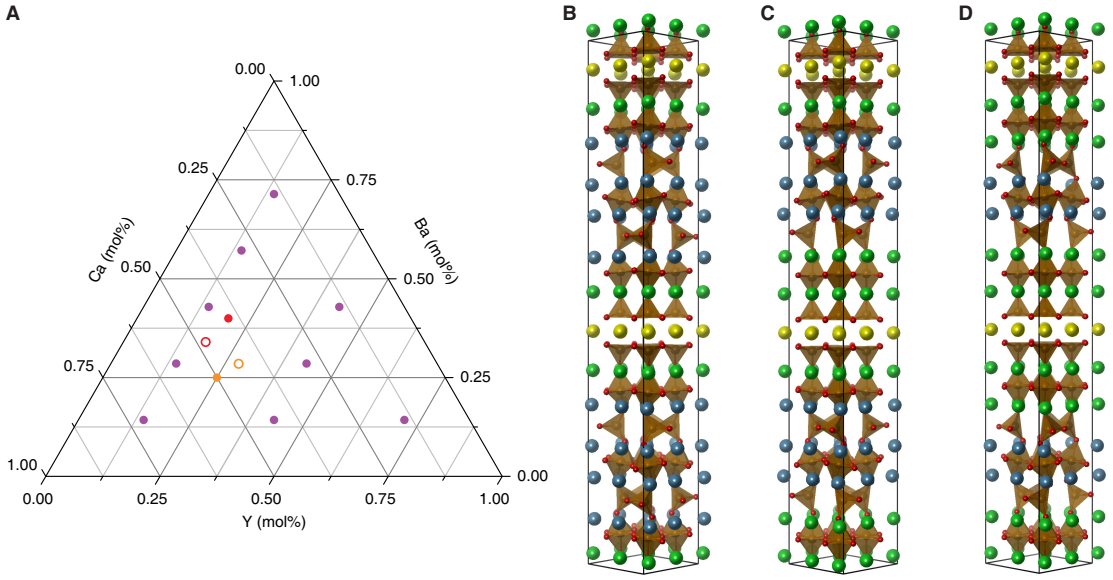

Fig. S11.

A) $7a_p$ compositions studied using EMMA and experimental solid state synthesis (purple filled circles). $10a_p$ (red) and $16a_p$ (orange) compositions studied in calculations (filled circles) and experiments (open circles) are shown for reference. B-D) The three lowest energy structures calculated using DFT for a $7a_p$ repeat. The compositions and formation energy relative to the ternary oxides $YFeO_3$ + $x$$Ba_2Fe_2O_5$ + $y$$Ca_2Fe_2O_5$ for each structure are B) $YBa_2Ca_4Fe_7O_{18}$: −0.02 eV/FU, C) $YBa_3Ca_3Fe_7O_{18}$: +0.19 eV/FU, and D) $YBa_4Ca_2Fe_7O_{18}$: +0.36 eV/FU. The DFT calculations were performed in doubled ($14a_p$) unit cells to allow for G-type antiferromagnetic ordering of the Fe spins. EMMA calculations using force-fields were carried out at the compositions $Y_5BaCaFe_7O_{20}$, $Y_3BaCa_3Fe_7O_{19}$, $Y_3Ba_2Ca_2Fe_7O_{19}$, $Y_3Ba_3CaFe_7O_{19}$, $YBaCa_5Fe_7O_{18}$, $YBa_2Ca_4Fe_7O_{18}$, $YBa_3Ca_3Fe_7O_{18}$, $YBa_4Ca_2Fe_7O_{18}$ and $YBa_5CaFe_7O_{18}$. At each composition, more than one set of starting modules was chosen, allowing the oxygen vacancies to reside either in Y or FeO layers. The lowest energy structures for $YBa_2Ca_4Fe_7O_{18}$, $YBa_3Ca_3Fe_7O_{18}$ and $YBa_4Ca_2Fe_7O_{18}$ had the lowest formation energies calculated relative to the ternary oxides using force-fields, and the three lowest energy structures at each of these compositions was then optimized using DFT. The structures shown in this figure are those with the lowest energy calculated using DFT at each composition.

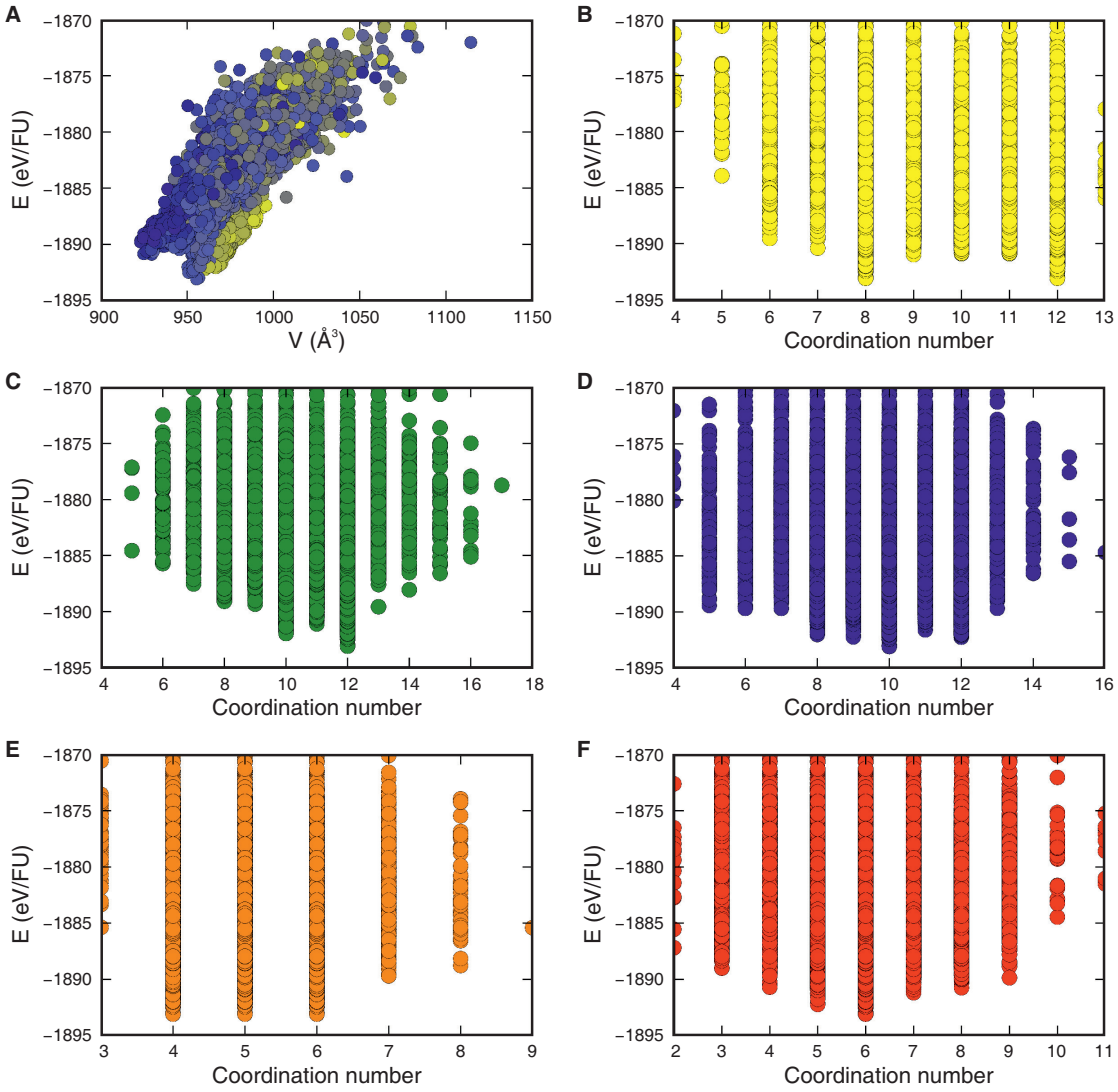

Fig. S12.

A) A plot of the energy of each fully relaxed structure against its volume. The dots are colored according to the ratio of the average Ca coordination number to the average Y coordination number, with dark blue representing structures in which Ca is coordinated to fewer O atoms than Y, and yellow the converse. B-F) Plots of the energy of structures against the coordination number of each atom within that structure, for Y, Ba, Ca, Fe and O respectively. For the purposes of this analysis, the coordination number is defined as the number of counter ions within a distance of 145% of the sum of the ionic radii.

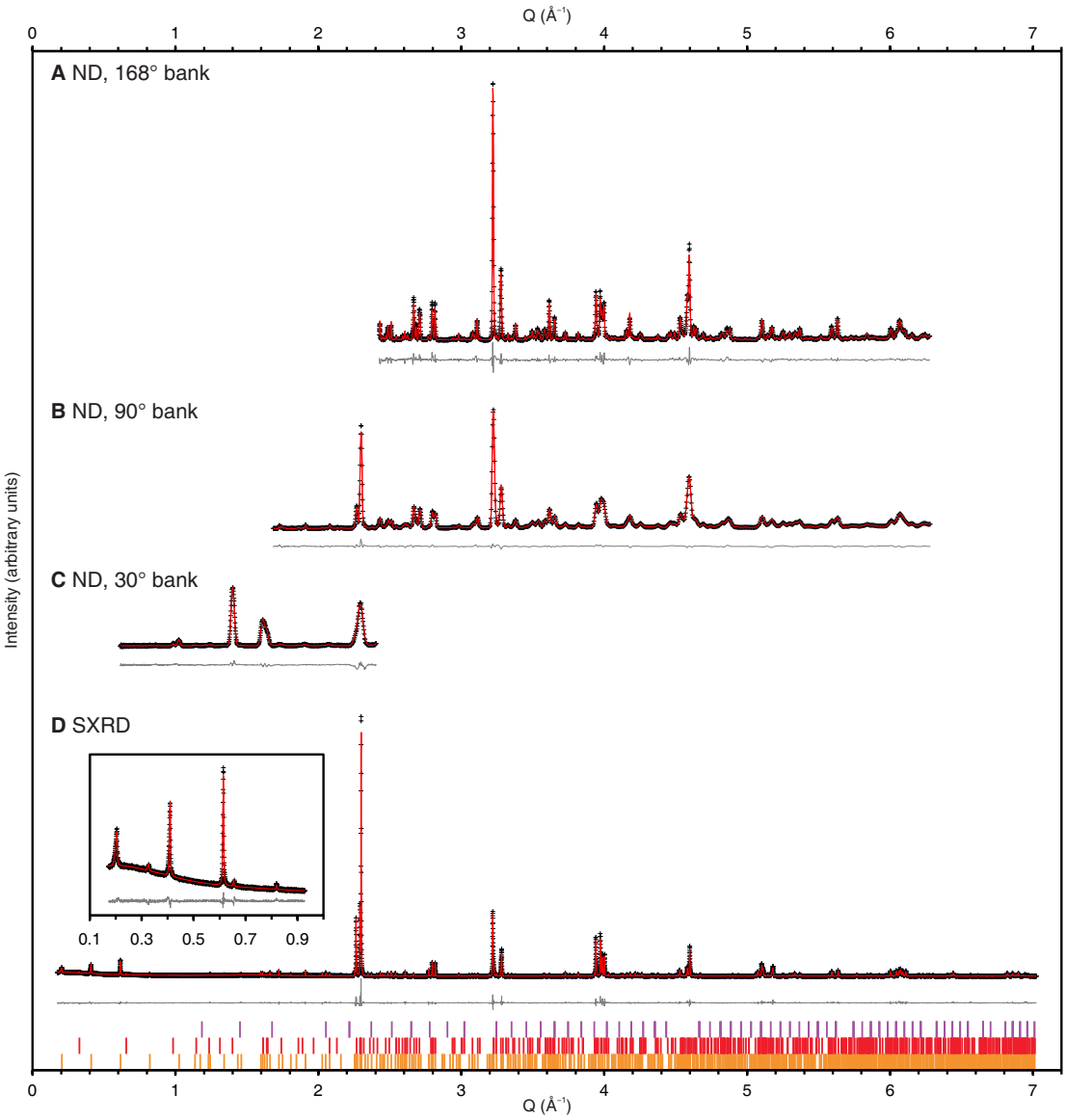

Fig. S13.

Final Rietveld plots for powder diffraction data from A) ND 168° bank, B) ND 90° bank, C) ND 30° bank and D) SXRD data with inset showing enlarged low 2θ data. Observed data are shown as black crosses, calculated data as a red line and difference as a grey line. Peak positions for 16a_p **1** (orange), 10a_p (red) and Y_2O_3 (purple) are indicated. χ^2 = 10.87 for 71 variables. R_{wp} / R_{exp}; ND 168° 0.0799 / 0.0174; ND 90° 0.0468 / 0.0063; ND 30° 0.0815 / 0.0463; SXRD high 2θ 0.100 / 0.0409; SXRD low 2θ 0.0328 / 0.0301.

Table S1.

Refined crystallographic data for **1** and starting values from EMMA.

	Parameter	Rietveld	EMMA
	a (Å)	5.484964(25)	5.56182
	b (Å)	61.25901(32)	61.55240
	c (Å)	5.550477(25)	5.57481
Site	Parameter	Rietveld	EMMA
A1	Position: 4a	0,0,0	0,0,0
	Occupancy	$Y_{0.627(4)}Ca_{0.373(4)}$	Y
A2	Position: 8h	0,0.062392(28),0.9977(7)	0,0.06555,0.99690
	Occupancy	$Ba_{0.787(2)}Ca_{0.213(2)}$	Ba
A3	Position: 8h	0,0.12077(5),0.9756(8)	0,0.12337,0.97632
	Occupancy	$Ca_{0.768(9)}Y_{0.132(7)}Ba_{0.100(5)}$	Ca
A4	Position: 8h	0,0.19120(5),0.9770(7)	0,0.19270,0.97272
	Occupancy	$Ca_{0.566(1)}Y_{0.434(1)}$	Ca
A5	Position: 4e	0,0.25,0.9868(9)	0,0.25,0.97795
	Occupancy	$Y_{0.488(2)}Ba_{0.512(2)}$	Y
B1	Position: 8h	0,0.030041(31),0.5008(6)	0,0.03128,0.49803
	Occupancy	$Fe_{0.86}Cu_{0.14}$	Fe
B2	Position: 8h	0,0.09470(4),0.4891(8)	0,0.09962,0.49929
	Occupancy	Fe	Fe
B3	Position: 16j	0.9450(4),0.15521(5),0.43229(32)	0.94244,0.15895,0.43195
	Occupancy	Fe	Fe
B4	Position: 8h	0,0.217745(29),0.4867(6)	0,0.21917,0.48128
	Occupancy	$Fe_{0.86}Cu_{0.14}$	Fe
O1	Position: 8g	0.25,0.02180(8),0.25	0.25,0.02234,0.25
O2	Position: 8g	0.75,0.02445(7),0.75	0.75,0.02324,0.75
O3	Position: 8h	0.5,0.06107(8),0.034(4)	0.5,0.06157,0.01199
O4	Position: 8g	0.75,0.09480(7),0.75	0.75,0.09665,0.75
O5	Position: 8h / 16j	0.5,0.12861(7),0.9347(9)	0.52821,0.13255,0.93059
O6	Position: 4e / 8i	0.5,0.25,0.0500(14)	0.53640,0.25,0.11429
O7	Position: 8g / 16j	0.25,0.28288(8),0.75	0.29486,0.22583,0.70720
O8	Position: 8g / 16j	0.25,0.28929(9),0.25	0.20219,0.21332,0.19796
O9	Position: 8h / 16j	0.5,0.31872(5),0.9354(8)	0.53161,0.31444,0.92061
O10	Position: 16j	0.9024(7),0.34338(11),0.1161(7)	0.90360,0.34174,0.12207
O11	Position: 8g	0.75,0.40037(8),0.25	0.75,0.39896,0.25

Space group *Imma*. Refined composition $Y_{2.25(1)}Ba_{2.29(1)}Ca_{3.47(2)}Fe_{7.44}Cu_{0.56}O_{21}$, including soft chemical composition restraint.