Structure and oxide ion conductivity: local order, defect interactions and grain boundary effects in acceptor-doped ceria

## **Supporting Information**

### S1 - Experimental details: sample preparation, XRD and Raman spectroscopy

**Synthesis:** For the sample preparation,  $Ce(NO_3)_3 \cdot 6H_2O$ ,  $Yb(NO_3)_3 \cdot 5H_2O$ ,  $Er(NO_3)_3 \cdot 6H_2O$  and  $Sm(NO_3)_3 \cdot 6H_2O$  (Sigma-Aldrich,  $\geq 99.9$  %), and citric acid (Sigma-Aldrich, 99.5%) were used. Typically, the total metal cation concentration was about 0.02M, and the citric acid concentration about 0.04M. The pH was adjusted by adding a concentrated ammonia solution dropwise until pH = 6. CeO<sub>2</sub> from Acros Organics (99.9%) was used for the undoped ceria samples. Sm<sub>2</sub>O<sub>3</sub>, Yb<sub>2</sub>O<sub>3</sub> and Er<sub>2</sub>O<sub>3</sub> (Sigma-Aldrich, 99.9%) were used as standards for Raman and XAFS experiments. All chemicals were used as supplied without further purification. The sintered pellets used for the impedance spectroscopy experiments had diameter and thickness on the order of 5 and 1 mm, respectively.

**XRD:** Diffractometer alignment and peak profile parameters were refined using corundum and LaB<sub>6</sub> standards. Rietveld refinement was carried out with pseudo-Voigt functions for profile fitting (GSAS function #3).

**Raman:** The Jobin Yvon Type V 010 LabRam Raman single grating spectrometer was equipped with a double super razor edge filter and a Peltier cooled CCD camera, and the spectra were acquired in quasi-backscattering geometry. The resolution of the spectrometer (600 lines/mm grating) is 3 wavenumbers (cm<sup>-1</sup>), laser power is less than 1 mW, and focused on a 20 (10) μm spot through a 20x (50x) microscope objective on the surface of the sample.

#### S2 - XAS: experimental details and data analysis

For the X-ray absorption experiments, the sintered pellets were crushed manually, and the powders were mixed with polyvinylpyrrolidone and pressed to form self-standing pellets. The Sm L<sub>2</sub>-edge fluoresence data were corrected for sample self-absorption using Viper (the Sm L<sub>2</sub>-edge was used instead of the L<sub>3</sub> for the lower interference from the intense Ce fluorescence lines). Theoretical EXAFS signals were generated by embedding the absorber atoms in a spherical CeO<sub>2</sub> cluster with radius about 8 Å having a cubic fluorite structure. The mixing of cerium and dopant cations in the second shell was accounted for by appropriately averaging the corresponding scattering paths. In any case, given that all the cations under consideration are lanthanides, the differences between the scattering amplitude and phases would be negligible. For instance, the effect of neglecting the Er/Ce random substitution in the second shell around  $Er^{3+}$  in EDC10M is one-two orders of magnitude smaller than the fitting uncertainty for the different parameters.

The important MS paths were preliminarily checked by using disorder factors calculated with the correlated Debye model implemented in Feff8.4. In general, multiple scattering effects were found to be most important for the Ce edge: they were accounted for by considering the triangular Ce-O-O (two vertices of the CeO<sub>8</sub> cube) and colinear Ce-O-Ce-O (diagonal of the CeO<sub>8</sub> cube) paths. A colinear path Sm-O-Sm-O was also included in the model to fit Sm-edge data. For the analysis of Er and Yb edges, multiple scattering paths were not included.

The overall amplitude factor,  $S_0^2$ , takes into account all multielectron effects that reduce the EXAFS signal amplitude. It was determined for each edge by fitting a standard compound (i.e.  $CeO_2$ ,  $Sm_2O_3$ ,  $Yb_2O_3$ ,  $Er_2O_3$ ), and fixed throughout the following analyses: its values are  $0.69\pm0.004$ ,  $0.93\pm0.005$ ,  $0.93\pm0.005$ , and  $0.75\pm0.004$  for Ce, Er, Yb and Sm, respectively. The respective relative uncertainty was added to that of N. All the EXAFS fittings were performed in R-space (approx. 1.0 - 4.5 Å) after k<sup>2</sup>-weighing of the raw data. The fit residual (R-factor) was around 5-7% for all samples. Experimental EXAFS data in k-space are shown in figures SI.1 and SI.2. According to the Nyquist criterion, the number of independent points was about 2 times higher than the number of fitted parameters. The correlations between two different structural and conductivity parameters are reported using the R<sup>2</sup> coefficient for linear regression.

The photoelectron wavevector is defined as a function of energy as:

$$k = \sqrt{\frac{2m_e}{\hbar^2} (E - E_0)}$$
(SI.1)

where  $E_0$  is the "real" edge position. In practice, this is not known *a priori*, so it is refined during the fit by introducing an edge shift,  $\Delta E_0 = E'_0 - E_0$ , where  $E'_0$  is the edge position used in Feff to generate the theoretical EXAFS signals. The wavevector then becomes:

$$k = \sqrt{\frac{2m_e}{\hbar^2} \left(E - E_0' + \Delta E_0\right)} \tag{SI.2}$$

The experimental error on each point in R-space was calculated by using the high-frequency part of the spectrum between 15 and 25 Å, and is between 0.01 and 0.03 for all data. Confidence intervals on the fitted results were then evaluated by taking into account the correlation between parameters with contour plots of the likelihood function,  $\exp(-\chi^2)$ . Example contour plots are shown in Figure SI.3.

Coordination numbers for the second and third shells were fixed as follows: the second shell coordination number was always fixed at 12, and the substitution between Ce<sup>4+</sup> and the dopants on the cation site was accounted for by mixing the corresponding amplitudes and phases according to the nominal stoichiometry. In any case, since the aliovalent cations considered here are all lanthanides, the difference between these cations as scatterers is negligible. The third shell coordination number was reduced in order to account for an average oxygen defectivity, since local effects average out at such a distance from the absorber. Therefore, for 10%, 20% and 30% doping level, the coordination number of the third shell was fixed at 23, 23 and 22, respectively.



**Figure SI.1** – Experimental EXAFS data. On the left: data on the Ce L<sub>3</sub>-edge. From top to bottom: EDC10M, EDC20M, EDC30M, YbDC10M, YbDC20M, YbDC30M, SDC10M, SDC20M, SDC30M. On the right: data on the Er L<sub>3</sub>-edge. From top to bottom: EDC10M, EDC10N, EDC20M, EDC20M, EDC20N, EDC30N.



**Figure SI.2** – Experimental EXAFS data. On the left: data on the Sm L<sub>2</sub>-edge. From top to bottom: SDC10M, SDC10N, SDC20M, SDC20N, SDC30M, SDC30M, SDC30N. On the right: data on the Yb L<sub>3</sub>-edge. From top to bottom: YbDC10M, YbDC10N, YbDC20M, YbDC30M, YbDC30N.



**Figure SI.3** – Contour plots of the likelihood function  $exp(-\chi^2)$  for two pairs of correlated parameters for sample YbDC10M on the Yb L<sub>3</sub>-edge. Upper plot: R and  $\Delta E_0$ ; lower plot: N and  $\sigma^2$ . Different colors enclose regions with values = 0.0-0.2, 0.2-0.4, 0.4-0.6, 0.6-0.8 and 0.8-1.0, respectively. In particular, the ellipse of the 68.3% confidence level lies in the second smallest region (depicted in violet).

The XANES peaks A, B and C was identified through minima in the second derivative. The peak intensity was determined with both: a) peak intensity at the maximum; b) deconvolution of Lorenzian lineshapes and integration. It is found that both methods give comparable results, although the latter is relatively less stable due to the higher number of refined parameters. For peak intensity normalization, both the A/C and A/(B+C) ratios were calculated: again, both methods give comparable results, although the latter suffers from increased propagation of uncertainty.

#### S3 - Impedance spectroscopy

Representative impedance Nyquist plots acquired from the microcrystalline samples are shown in Fig. SI.4. The spectra consist of three different contributions arising from the bulk (grain interior) at high frequencies, grain boundary (GB) at intermediate frequencies, and electrode at low frequencies. Each contribution is modelled by an equivalent circuit consisting of a resistance R in parallel with a constant phase element Q, whose capacitance is defined by

$$C = (R^{1-n}Q^n)^{1/n}$$
 (SI.3)

where n is a fitting parameter. In those cases in which the bulk signal is incomplete or not detectable (e.g. because the frequency corresponding to the characteristic time constant exceeds the maximum frequency accessible by the spectrometer), the equivalent circuit is composed of one resistance R (bulk) in series with two RQ circuits (GB and electrode).

The impedance spectra of all microcrystalline samples exhibit 3 distinct contributions, which in the light of the capacitance values can be safely assigned (from high to low frequencies) to the bulk ( $C_{\infty}$  = 7 pF, from which  $\varepsilon_{\infty}$  = 25 results), the grain boundary ( $C_{gb}$  = 1 nF) and the electrode contribution, respectively.



Figure SI.4 – Impedance spectra at 400°C for SDCM and EDCM samples. On the left: a) SDC10M; b) SDC20M; c) SDC30M. On the right: d) EDC10M; e) EDC20M; f) EDC30M.



Figure SI.5 – Impedance spectra at 400°C for YbDCN and EDCN samples. On the left: a) YbDC10N; b) YbDC20N; c) YbDC30N. On the right: d) EDC10N; e) EDC20N; f) EDC30N.

Fig. SI.5 displays the impedance spectra collected at 400°C from the YbDCN and EDCN samples. Generally, the spectra were modelled using one resistance  $R_{\infty}$  in series with a  $R_{gb}Q_{gb}$  circuit. Due to the largely depressed spectra of the 30 at.% doped samples, only the total resistance was could be considered for these compositions. The  $C_{gb}$  values are comprised between 0.1 nF and 10 pF, while the bulk values are about 6 pF ( $\epsilon_{\infty}$  = 25).

# S4 - Results of the XRD analysis

Figure SI.6 depicts the XRD pattern with the corresponding Rietveld refinement obtained for the YbDC10N sample.



Figure SI.6 – Rietveld refinement of the XRD pattern of YbDC10N.

Table SI.1 - Results from the Rietveld refinements on XRD data: lattice parameter (a), calculated cation-cation distance (M-M) and calculated cation-anion distance (M-O). Uncertainty is about 10-4 Å on the lattice parameter, and distances are truncated to 0.01 Å for comparison with EXAFS results.

	a (Å)	M-0 (Å)	M-M (Å)	M-O (Å)
		first shell	second shell	third shell
CeO2	5.4105	2.34	3.83	4.49
EDC10N	5.4085	2.34	3.82	4.48
EDC10M	5.4051	2.34	3.82	4.48
EDC20N	5.4048	2.34	3.82	4.48
EDC20M	5.3999	2.34	3.82	4.48
EDC30N	5.3992	2.34	3.82	4.48
EDC30M	5.3896	2.33	3.81	4.47
YbDC10N	5.4053	2.34	3.82	4.48
YbDC10M	5.3993	2.34	3.82	4.48
YbDC20N	5.3966	2.34	3.82	4.47
YbDC20M	5.3875	2.33	3.81	4.47
YbDC30N	5.3879	2.33	3.81	4.47
YbDC30M	5.3762	2.33	3.80	4.46
SDC10N	5.4241	2.35	3.84	4.50
SDC10M	5.4234	2.35	3.83	4.50
SDC20N	5.4334	2.35	3.84	4.51
SDC20M	5.4322	2.35	3.84	4.50
SDC30N	5.4443	2.36	3.85	4.51
SDC30M	5.4421	2.36	3.85	4.51

#### S5 - Calculation of interatomic distances in the ion-packing model

(after Yashima et al., J. Am Ceram. Soc. 1992, 75, 1541)

1

The interatomic distances are calculated using tabulated values for the ionic radii of cations in 6-fol, 7-fold and 8fold coordination and oxygen ions in 4-fold coordination.

R are distances, r<sub>X,p</sub> is the ionic radius of species X in p-fold coordination, x is the dopant fraction, and a<sub>X,p</sub> is a coefficient relative to the X species in p-fold coordination, which is a function of p, X, and of the model. The a<sub>Ce,p</sub> and a<sub>M,p</sub> coefficients for the different dopant-vacancy clusters are given in Yashima, M.; Ishizawa, N.; Yoshimura, M. J. Am Ceram. Soc. 1992, 75, 1541.

The Ce-O distance is given by:

$$R_{Ce-O} = r_{O,4} + \frac{r_{Ce,8} + x \cdot \left(\sum_{p} a_{Ce,p} r_{Ce,p}\right)}{1 - x}$$
(SI.4)

The M-O distance is given by:

$$R_{M-O} = r_{O,4} + \sum_{p} a_{M,p} r_{M,p}$$
(SI.5)

The Ce-cation distance is given by:

$$R_{Ce-cat} = \sqrt{\frac{2}{3}} \left[ R_{Ce-O} + r_{Ce,8} + r_{O,4} + x \cdot \left( \sum_{p} a_{Ce,p} r_{Ce,p} + \sum_{p} a_{M,p} r_{M,p} \right) \right]$$
(SI.6)

The M-cation distance is given by:

$$R_{M-cat} = \sqrt{\frac{2}{3}} \left[ R_{M-O} + r_{Ce,8} + r_{O,4} + x \cdot \left( \sum_{p} a_{Ce,p} r_{Ce,p} + \sum_{p} a_{M,p} r_{M,p} \right) \right]$$
(SI.7)

The cation-cation distance is a weighted average (Vegard's law):

$$R_{cat-cat} = x \cdot R_{M-cat} + (1-x) \cdot R_{Ce-cat}$$
(SI.8)

## **S6 - Statistical lattice simulations**

The occurrence of configurations around Ce and M corresponding to different dopant-vacancy associations were calculated with statistical lattice simulations. The purpose of the simulations is to calculate the average coordination numbers M-O, Ce-O (shown in figure SI.7), M-M, M-Ce and Ce-Ce and to correlate these with the occurrence of the basic association clusters (i.e. Ce4V, Ce3MV, Ce2M2V, CeM3V or M4V), by averaging over a large enough portion of a crystallite with different degrees of clustering. In a typical simulation, a CeO<sub>2</sub> lattice box comprising about 20000 atoms is generated and the dopants and vacancies are placed by varying two parameters: a) M-M attraction (expressed as the ratio between M-M pairs and total cation-cation pairs); b) M-V attraction (as the ratio between M-V pairs and total cation-anion pairs). The coordination numbers are then counted around each cation and averaged. In comparison to isolated models A-E, this results in a realistic assortment of M-V pairs and M-V-M clusters that are expected to occur randomly in heavily doped regimes.



**Figure SI.7** – Coordination numbers around the dopant and Ce for 10% (black), 20% (red) and 30% doping (blue). Dots are statistical lattice simulations under different M-M and M-V attraction rules, lines are linear fits. The lines converge at the point where pure CeO<sub>2</sub> and M<sub>2</sub>O<sub>3</sub> are totally separated (i.e. N(Ce-O) = 8, N(M-O) = 6).

The fraction of M-M or M-V pairs over the total cation-cation or cation-anion pairs is used as a quantitative measure of the M-M or M-V clustering. In the simulations the defects are placed either randomly or by following the association models A to E. The non-random models represent extreme situations of M-V attraction and M-M ordering, and it is extremely unlikely that only one single motif is present in a statistical ensemble of particles.



**Figure SI.8** – Relative abundance of the different cation configurations around an oxygen vacancy, if all atoms are placed randomly and no preferential interaction occurs.

As is shown in figure SI.8, by placing the dopants and vacancies randomly in the lattice the following results are obtained: a) at 10% doping, the random configuration corresponds to a mixture of model A (70%) and model B (30%). The occurrence of dopant pairs is expected not to be significant; b) at 20% doping, the random configuration is composed of model B and model A (both around 40%), and model C (15%). The majority of the vacancies reside in the vicinity of at least one dopant. About half of the dopant atoms are expected to be paired, and the other half isolated; c) at 30% doping, the random configuration is composed of model D (15%). About 7 dopants out of 10 are clustered in groups of two or three, and the majority of the vacancies reside near them.