## **Supporting information**

## Polymorphic phase study on Nitrogen-doped TiO<sub>2</sub> nanoparticles: Effect on oxygen site occupancy, dye sensitized solar cells efficiency and hydrogen production

Emerson C. Kohlrausch,<sup>a</sup> Maximiliano J. M. Zapata,<sup>b</sup> Renato V. Gonçalves,<sup>c</sup> Sherdil Khan,<sup>b</sup> Mauricio de O. Vaz,<sup>b</sup> Jairton Dupont,<sup>d</sup> Sérgio R. Teixeira,<sup>b</sup> Marcos J. L. Santos<sup>a\*</sup>

In order to calculate the percentage of anatase and rutile phase in the samples, in addition to the concentration of defects, we have obtained XRD spectra of the samples and by Rietveld Refinement using Anatase (JCPDS # 84-1286; CIF # 20-2243) and Rutile (JCPDS # 76-0649; CIF # 34-372) the diffraction profile was built.

The fitting was performed according to the following sequence:

Scale factor, zero point of detector, background (using linear interpolation), lattice parameters, atomic positions, overall Debye-Waller factor anisotropic, peak shape and asymmetry parameters, atom occupancies, finally microstructural parameters: size (using harmonics spherical corrections) and strain effects.

In a mixture of N crystalline phases the weight fraction W<sub>j</sub> of phase is given by:

$$W_j = \frac{S_j Z_j M_j V_j / t_j}{\sum S_i Z_i M_i V_i / t_i}$$

where  $S_j$  is the scale factor of phase j,  $Z_j$  is the number of formula units per unit cell for phase j,  $M_j$  is the mass of the formula unit,  $V_j$  is the unit cell volume,  $t_j$  is the Brindley particle absorption contrast factor for phase j.

The crystallographic structure factor  $F_i$  is calculated in FULLPROF by using formula:

$$F_{h} = \sum_{j=1}^{n} O_{j} f_{j}(h) exp(-B_{j}(h)/4) \sum_{s=1}^{m} T_{js}(h) exp2pii(h^{T}S_{s}rj + h^{T}t_{s})$$

where *n* is the number of atoms in the asymmetric unit, *m* is the number of the reduced set of symmetry operators.  $O_j$  is the occupation factor,  $f_j(h)$  is the scattering length,  $B_j$  is the isotropic temperature parameter,  $r_j$  is the position vector of atom *j*.