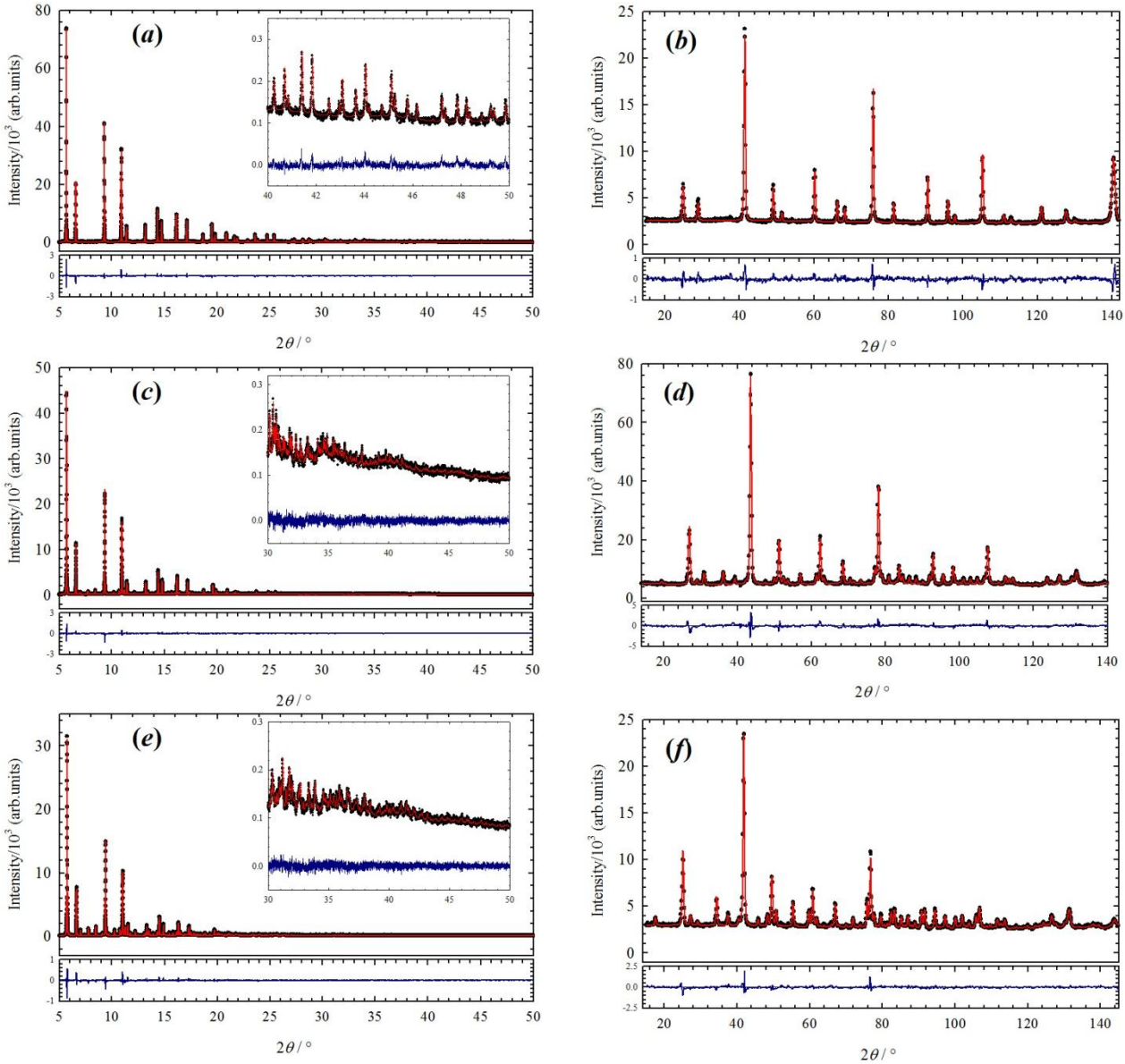

Fig SI1. *Combined* XRPD (*a*-*c*-*e*) and NPD (*b*-*d*-*f*) Rietveld refinements for *x*=0.250 (*a*,*b*), *x*=0.500 (*c*,*d*) and *x*=0.750 (*e*,*f*).

SI-1 *Rietveld refinement strategy*

XRPD and NPD patterns were analyzed via the Rietveld method [1] as implemented in the GSAS software suite [2] which features the graphical interface EXPGUI [3]. The background was fitted by Chebyshev polynomials. Line profile was fitted using a modified pseudo-Voigt function [4] accounting for asymmetry correction [5]. XRPD data were corrected for anomalous scattering considering *f'* and *f''* values taken from [6] and for absorption through the Lobanov empirical formula [7]. In the final refinement cycles, different strategies were followed for fluorite and C-type samples. For the former, the cell parameter and the *mean square displacement* were allowed to vary along with the background and the line profile parameters. For the latter, atom positions and *anisotropic displacement parameters* were also allowed to vary. For the samples investigated by means of both probes (see Table 2 in main text), *combined* X-ray and neutron refinements were performed: the profile parameters as well as the background coefficients were refined separately for both patterns, while the same structural model, in terms of displacement parameters, cell parameter and in C-type samples atom positions, was applied to both datasets. A NIST Si standard was used to calibrate wavelength and instrumental profile of both diffractometers.

[1] Rietveld, H. M. *J. Appl. Crystallogr.* **1969**, *2*, 65-71

[2] Larson A. C.; Von Dreele, R. B. *General Structural Analysis System (GSAS)*; Los Alamos National Laboratory Report LAUR **2004** 86-748.

[3] Toby, B. H. *J. Appl. Crystallogr,*. **2001**, *34*, 210-214.

[4] Thompson, P.; Cox, D. E.; Hastings, J. B. *J. Appl. Cryst.* **1987**, *20*, 79-83.

[5] Finger, L. W.; Cox, D. E.; ,Jephcoat; A. P. *J. Appl. Cryst.* **1994**, *27*, 892-900.

[6] Brennan S.; Cowan; P. L.. *Rev. Sci. Instrum.* **1992**, *63*, 850.

[7] Lobanov, N. N.; Alte da Veiga, L. *6th European Powder Diffraction Conference*, **1998**, Abstract P12-16

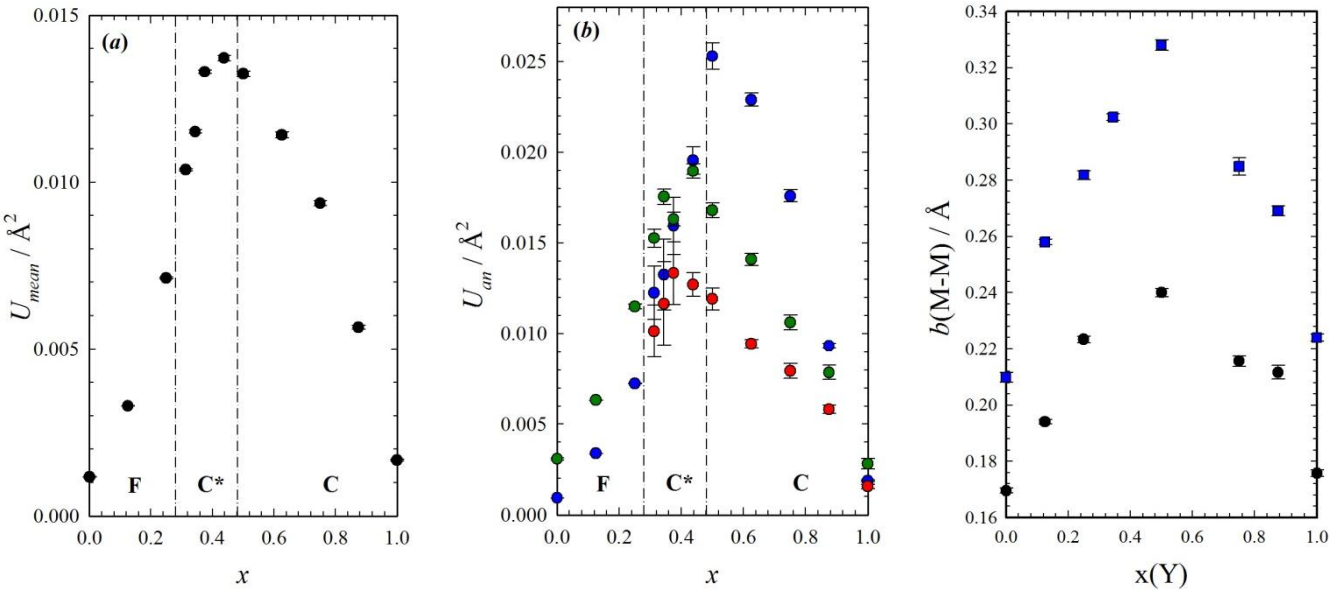

Fig. SI2. (*a*) U_{mean} and (*b*) U_{an} as determined from Rietveld refinements. Colors stand for different sites, according to Fig. 1 of the manuscript: blue indicate M1 site in C-type and M site in fluorite, red M2 site in C-type and green O sites. (*c*) PDF peak MM_1 width as determined from XRPD (black circles) and NPD (blue squares).

SI-2: *PDF data reduction*

XRPD real space data were processed using the software PDFGetX2: data were corrected for background, sample self-absorption, multiple and Compton scattering. Empty kapton capillary and air were measured to properly subtract background contribution.

Neutron real space data were processed by regrouping repeated scans of the D4c detectors over the available range of scattering angles in the reciprocal space. For each sample, empty container and sample environment contributions were subtracted from the raw data, taking into account attenuation effects [8] as well as the incoherent-scattering contribution. Standard multiple-scattering [8] and Placzek [9] corrections were applied.

CeO_2 was used to determine PDF instrumental parameters for both XRPD and NPD datasets.

[8] Howe, M. A.; McGreevy, R.L.; Zetterstrom, P. NFL Studsvik internal report **1996**.

[9] Yarnell, J. L.; Katz, M. J.; Wenzel, R. G.; Koenig, S. H. *Phys. Rev*. **1973**, *A7*, 2130-2144.

SI-3: *Displacement parameters*

The overall Debye -Waller is determined in a Rietveld refinement by constraining all ions to the same *msd*, here denoted as U_{mean}. The U_{mean} values for the CeO_2-Y_2O_3 system are plotted in Fig. SI2 (*a*). The largest disorder is observed in the C* region, at x~0.4. U_{mean} does not provide, though, any information about the spatial distribution of disorder. In C-type samples, this information can be extracted by assuming a different *msd* for each of the crystallographic sites. This strategy, however, led to unphysical results: negative O *msd* were found when only the XRPD dataset was available. This problem was overcome by considering the atomic *anisotropic displacement parameters* for cationic sites and isotropic *msd* for oxygen ions, i.e. the lightest atoms, which less contribute to the X-ray scattered intensity. Whenever also NPD datasets were available, *ADP* were considered even for the O1 site. The quadratic sum of the *U*11, *U*22, *U*33 components of the anisotropic tensor is here defined as U_{an} and is reported in Fig. SI2 (*b*). When only XRPD are available, isotropic O *msd* are considered. From Fig. SI2 (*b*) it can be inferred that the most disordered site is M1 in C-type. Moreover, the largest disorder on the M1 site occurs at doping amounts larger than M2 and O sites.

The peak width as derived from PDF analysis for the MM_1 pair is shown in Fig. SI2 (*c*) for NPD (blue squares) and XRPD (black circles) data. The compositional evolution resembles the one of *msd* parameters. The larger width determined on NPD data are in accordance with i) the lower *r*-resolution (lower Q_{max}) of D4c data and ii) the occurrence of O-O pairs at nearly the same distances.

SI-4: *Superstructure peaks broadening*

The main problem in performing an accurate refinement of the whole powder diffraction pattern of C*samples is that none of the profile functions available in the GSAS-EXPGUI package successfully model both the main and the superstructure peaks using a single C-type phase. The presence of C-type nanodomains in a fluorite matrix could in principle be implemented in a Rietveld refinement by considering the coexistence of both phases, characterized by a different line profile. However, we did not success for two main reasons:

- The size affects ALL the peaks of the corresponding phase, while in our experimental patterns ONLY the superstructure ones are further broadened. The Rietveld refinement performed employing a C-type phase with a small particle size adapted to fit the broadened superstructure peaks together with a fluorite phase accounting for sharp structure peaks is shown in Fig. SI3. This model produces calculated structure peaks with very large tails, typical of a bimodal distribution, evidently

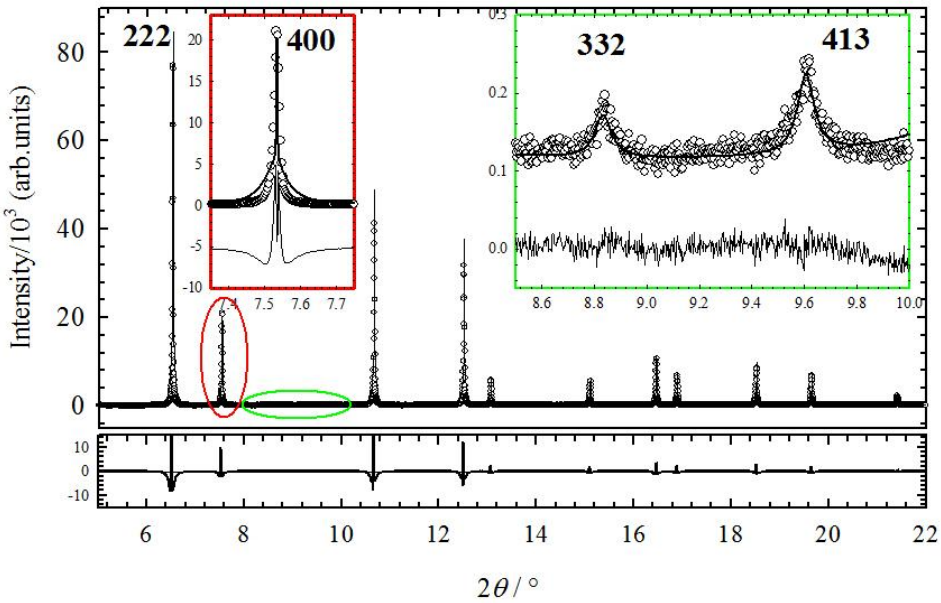

Fig. SI3. Fit of the XRPD pattern of x=0.344 performed with Rietveld refinement considering a biphasic system composed of a fluorite phase accounting for the sharp structure peaks and a C-type phase accounting for the nanodomains. The size related parameters of the C-type phase was selected to get a good fit of 332 and 413 superstructure peaks, magnified in the green inset. The fit of the 400 main reflection is shown in the red panel. Note that the scale in the insets is the same as in the main figure.

inconsistent with the experimental profile (see red panel in Fig. SI3).

- The above modeling neglects any form of interactions between the C-type phase and the fluorite matrix. In fact, we propose that these interaction effects are the source of the peak broadening.

The failure of the above refinement procedures led us to look for a different modeling approach, which is the one reported in the main text.

SI-5: *PDF modeling*

The determination of the extent of disorder via PDF is somewhat arbitrary: in principle it can be considered as the *r*-region in which it is possible to resolve differences in PDF peaks, in terms of shape or position, by refining experimental data against a reference structural model. A bad fit is an indication of disorder. Nevertheless, the extent of disorder is hard to define especially when G(*r*) curves reflect its presence in terms of a large peak broadening. In order to determine whether the long range structural model is apt to describe the G(*r*) curves throughout different *r*-ranges two long range models, defined as *average*, were considered.

A *fluorite* model was applied varying the cell parameter, an overall scale factor, one metal and one oxygen *msd* parameter. When dealing with first neighbor pairs, a parameter accounting for correlated motion was considered, too. Five parameters in all. A *C-type* model was applied varying the cell parameter, an overall scale factor, two metal and one oxygen *msd*, the *x*(M2) coordinate and a parameter accounting for correlated motion of nearest neighbor atom pairs. Seven parameters in all. The four O atomic coordinates (three for O1 and one for O2) were kept fixed to the values determined via Rietveld analysis. It should be noted that better fits are produced by including *ADP* parameters at the expense of larger parameter correlations. The presence of two metal ions is taken into account by randomly populating cation positions with Ce^{4+} and Y^{3+} ions in due proportions. As mentioned in main text, refinements were performed in 20 Å wide steps up to 400 Å. A Nyquist grid was employed to reduce oversampling in the Fourier Transform process. When moving to shorter interatomic distances, a custom grid of 0.01 Å was utilized to better appreciate details of PDF refinements.

SI-6: *the peak C in pure CeO_2*

By looking at Fig. 6 of the main text, the reader could argue that a second M-M peak in fluorite appears even in pure ceria. This would imply that peak MM_2 is not induced by doping. Actually, the contribution at ~4.1 Å in pure ceria is due to a termination ripple arising from the combination of the Q_{max} truncation effect and very low *msd*. By increasing the *msd*, as occurs in doped samples, this effect reduces. This ripple is indeed correctly modeled employing an undistorted CeO_2 phase [**12**]. Conversely, in doped samples the observed peak is expected to bear some structural information. Anyhow, we cannot exclude the position of the peak MM_2 to be somehow affected by the presence of spurious oscillations, especially for x=0.125.

The incorrect normalization of the S(*Q*) function at high *Q* values, due to the presence of Bragg peaks, rebounds on the quality of the PDF pattern in the direct space, producing some ripples on a wide *r*-range. This cannot though be considered a remarkable response of disorder, since the ripples are actually artifacts. According to the very low *msd* determined by Rietveld refinements, CeO_2 can be considered as an ordered material from local to average scale.

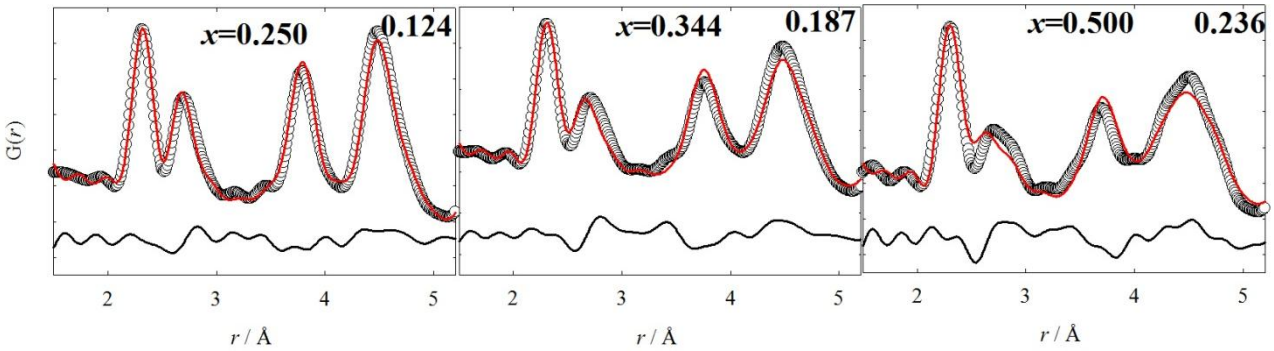

Fig. SI4. PDF refinements performed in the 1.5 < *r* < 5.2 Å range on NPD data on *x*=0.250 (*left*), *x*=0.344 (*middle*) and *x*=0.500 (*right*), employing the *biphasic* model with cell parameters and *x*(M2) fixed as the values determined from XRPD *biphasic* model. Empty circles: experimental data; red solid lines: calculated profile; difference curves shown below each plot. Fit residuals R_w are given in the top right corner of each panel.

SI-7: ***biphasic model strategy***

The following parameters were refined applying the *biphasic* model against XRPD datasets: a different cell parameter for fluorite and C-type phases, one overall *msd* for cations and another one for O ions, one overall parameter accounting for correlated motion and the *x* coordinate of M2 site. Also the O atomic coordinates and the occupation factor of the O2 site were refined against NPD data.

Although some instrumental aberrations might slightly affect PDF peak positions, as a first attempt we tried to constrain atom positions to be exactly the same for XRPD and NPD datasets. The best way to do that would be to perform a combined simultaneous refinement of the same structural model against XRPD and NPD G(*r*) curves. This is in principle possible with the PDFGUI software, but we didn't manage to get reliable results. A good fit of XRPD was obtained at the expense of the fit of NPD data. In particular, *msd*(O) was found to be overestimated in XRPD data. To overcome this problem, the refinements were first performed against XRPD PDFs, the results were then applied to NPD data and so on in a recursive way. The first fit against XRPD PDF is done to accurately determine the cations related parameters. All XRPD PDF peaks involve indeed M ions. Due to their low weight in XRPD PDF, the position of O ions is constrained to the ones of pure Y_2O_3.

The cell parameters and the *x*(M2) coordinate as obtained from the XRPD PDF fit were kept fixed in the NPD data refinement, in order to retain the same interatomic distances involving M ions, whereas the O atomic coordinates and the occupation factor *f*(O2) of the O2 site in C-type were refined. This refinement strategy produced *f*(O2) and *C frac* values close to those reported in the main text. However, the fit of NPD G(*r*) curves is poor. PDF refinements for the same samples discussed in the main text are reported in Fig. SI4. An accurate perusal of NPD refinements reveals, though, that especially when the dopant amount is low, all peak positions are correctly described but the second one, corresponding to O-O pairs. This is particularly true of *x*=0.344 and *x*=0.500. As obtained by direct analysis, doping produces both the shrinking of M-O and the lengthening of O-O pairs. XRPD detects only the first of these effects, whilst NPD detects both 1st M-O and O-O pairs. It is then impossible to fit the OO peak by constraining the cell parameters to the values determined from XRPD.

The *biphasic* model approach would then require two different cell parameters for describing the cationic and anionic substructures. This approach was actually applied and discussed in the main text and is supported also by the large oxygen and cation *msd* obtained against XRPD and NPD G(*r*) curves, respectively.

SI-8: the *biphasic model at the mesoscopic scale*

The same estimated sizes of C-type domains are obtained by applying the average C-type or the *biphasic* model. The lack of structural coherence is proven by the *biphasic* model through the vanishing of both *x*(M2) and *C frac* parameters. In this case, *x*(M2) is observed to be much larger than by involving a single C-type phase. The *x*(M2) value obtained employing a single C-type phase turns then out to be an average between a fluorite configuration, where *x*(M2)=0, and a C-type configuration, where *x*(M2) is negative. Owing to the large correlations between *x*(M2) and phase fraction, as well as between the cell parameters of the two phases, we decided to refer to the refinements with a single C-type phase.

Table SI.1 Results of Rietveld refinements on (*a*) fluorite, (*b*) C* and (*c*) C-type samples. The probe employed, that is X-rays (X), neutrons (N) or both (X+N), is indicated for each sample. U_{an} stands for either the quadratic sum of *U*11, *U*22, *U*33 when *ADP* are considered, or the *msd* value.

(*a*) x	0 X+N	0.125 X+N	0.250 X+N
phase	F	F	F
a	5.406813()	5.405629(6)	5.403096(9)
U(M)	0.00090(1)	0.00339(1)	0.00724(2)
U(O)	0.00318(6)	0.00653(7)	0.01149(12)
$R(F^2)$	0.0302-0.0386	0.0397-0.0476	0.0466-0.0524
R_{wp}	0.1014-0.0453	0.0489-0.0651	0.0511-0.0535

(*b*) x	0.3125 X	0. 34375 X	0. 375 X	0.4375 X
phase	C*	C*	C*	C*
a	10.79171(5)	10.7853(6)	10.77719(2)	10.77380(5)
xM2	-0.00267(12)	-0.00449(9)	-0.00623(5)	-0.01159(3)
x(O1)	0.3697(10)	0.3716(16)	0.3773(8)	0.3785(3)
y(O1)	0.1264(11)	0.1286(18)	0.1353(7)	0.1406(3)
z(O1)	0.377(2)	0.377(3)	0.3781(9)	0.3785(5)
x(O2)	0.379(3)	0.379(3)	0.3762(15)	0.3759(11)
U11(M1)	0.0123(15)	0.013(2)	0.0159(16)	0.0195(8)
U12(M1)	0.0076(9)	0.0097(10)	0.0112(6)	0.0127(5)
U11(M2)	0.0127(6)	0.0131(1)	0.0070(6)	0.0072(2)
U22(M2)	0.0057(5)	0.00880(10)	0.0176(12)	0.0078(3)
U33(M2)	0.012(2)	0.013(2)	0.015(2)	0.0231(8)
U23(M2)	-0.0029(8)	-0.0045(7)	-0.0035(5)	-0.0082(3)
Uan(M1)	0.0122(15)	0.013(2)	0.0159(16)	0.0195(8)
Uan(M2)	0.0101(14)	0.0117(23)	0.0133(17)	0.0127(7)
U(O)	0.0153(5)	0.0175(4)	0.0163(4)	0.0190(4)
$R(F^2)$	0.0745	0.0896	0.0436	0.0410
R_{wp}	0.0795	0.0863	0.0652	0.0703

(*c*) x	0.500 X+N	0.625 X	0.750 X+N	0.875 X+N	1 X+N
phase	C	C	C	C	C
a	10.76044(9)	10.72268(9)	10.69187(4)	10.65293(7)	10.601027(9)
xM2	-0.01501(5)	-0.01986(2)	-0.02405(2)	-0.02830(2)	-0.03241(2)
x(O1)	0.3807(3)	0.3835(2)	0.3860(2)	0.3884(1)	0.3908(1)
y(O1)	0.1413(2)	0.1462(3)	0.1472(2)	0.1502(1)	0.1518(1)
z(O1)	0.3798(3)	0.3797(3)	0.3809(2)	0.3799(2)	0.3803(1)
x(O2)	0.3816(6)	0.3830(7)	0.3821(8)	0.3844(11)	-
U11(M1)	0.0253(9)	0.0229(4)	0.0176(3)	0.00932(14)	0.00189(6)
U12(M1)	0.0169(5)	0.0164(3)	0.0136(3)	0.0062(2)	0.00048(13)
U11(M2)	0.0070(2)	0.00547(11)	0.00579(12)	0.00468(11)	0.00141(7)
U22(M2)	0.0076(3)	0.0096(2)	0.00781(16)	0.00570(15)	0.00171(10)
U33(M2)	0.0211(10)	0.0133(4)	0.0102(3)	0.0071(2)	0.00158(12)
U23(M2)	-0.0096(3)	-0.0090(3)	-0.0061(2)	-0.0031(2)	-0.00021(9)
U11(O1)	0.0257(10)	-	0.0142(8)	0.0107(7)	0.0032(5)
U22(O1)	0.0134(10)	-	0.0115(7)	0.0078(6)	0.0026(5)
U33(O1)	0.0113(7)	-	0.0061(6)	0.0050(5)	0.0027(4)
U12(O1)	-0.0108(7)	-	-0.00719(5)	-0.0033(4)	-0.00004(32)
U13(O1)	-0.0006(7)	-	-0.0073(6)	-0.0004(4)	-0.0009(3)
U23(O1)	-0.0047(6)	-	-0.00033(6)	-0.0005(6)	0.0005(4)
Uan(M1)	0.0253(9)	0.0229(4)	0.0176(3)	0.00932(14)	0.00189(9)
Uan(M2)	0.0119(5)	0.00943(3)	0.0080(2)	0.0058(2)	0.00157(5)
Uan(O)	0.0168(9)	0.0141(8)	0.0106(7)	0.0079(6)	0.00282(9)
$R(F^2)$	0.0521-0.0729	0.0334	0.0386-0.0528	0.0228-0.0472	0.0237-0.0334