Appendix 3

Diffusion model.

$$[^{18}O_i''] = (K \times \sum{}^{18}O - K \times \sum X_O + (K^2 \times (\sum{}^{18}O)^2 - 2K^2 \times \sum{}^{18}O \times \sum X_O +$$

$$K^2 \times (\sum X_O)^2 + 2K \times \sum{}^{18}O + 2K \times \sum X_O + 2K \times \sum X_O + 1)^{\frac{1}{2}} - 1)/(2K) \ \ (A3.1)$$

then:

$$[V_O^{\bullet\bullet}] = -([^{18}O_i''] - \sum{}^{18}O)/(K \times [^{18}O_i'']) \hspace{3cm} (A3.2)$$

and

$$[^{18}O_O^{\times}] = K \times [V_O^{\bullet\bullet}] \times [^{18}O_i''] \hspace{3cm} (A3.3)$$

Fitting procedure. Nonlinear least squares regression (function *lsqnonlin* in MATLAB) was used in order to determine the input parameters necessary to generate the measured profiles, by minimizing the summed squares of the residuals between the modelled $\sum{}^{18}O$ profile and the measured profile converted to ^{18}O per 12 O. Not all parameters were allowed to vary freely – the fitted parameters were the three Ds, K and $\sum X_O$ (initial and interface). The $\sum{}^{18}O$ terms were not fitted, instead $\sum{}^{18}O$ (interface) was determined visually from the SIMS measured profiles, and $\sum{}^{18}O$ (initial) was calculated as the mean of the final (crystal interior) 2-10 points in the measured profile. The general relationship $D^{18}O_O^{\times} < DV_O^{\bullet\bullet} < D^{18}O_i''$ was assumed.

Finding a global minimum in such a system is challenging due to (1) the existence of many local minima and (2) the extremely minor effect of some parameters (mainly K and $D^{18}O_O^{\times}$) on the residuals. For example, for stepped profiles where the step is sharp, changing $\log K$ from 3 to 5 has almost no effect on the profile shape. Therefore, the profiles were fitted using a custom procedure, which generally proceeded as follows. Firstly, an arbitrary value of $\log K$ was chosen, generally between 3 and 5.

With this value fixed, the best fits for the other fitted parameters ($D^{18}O_O^\times$, $DV_O^{\bullet\bullet}$, $D^{18}O_i''$, $\sum X_O$(interface)) were determined using nonlinear least squares regression, with at least 20 random starting guesses for each parameter. Then, $\log K$ was changed (generally reduced) by 0.05, and the best fit parameters from the previous fit were used as initial guesses for the solver, which was run again. From the resulting fit, the sum of the squares of the residuals was calculated. Then K was modified by 0.05 again, and the process repeated, until the model had been run with K values spread over ~3-5 orders of magnitude. The $\log K$ value associated with the lowest residuals was extracted, and the process was repeated with a closer spacing of $\log K$ (0.01), within ±0.2 of the previous best fit $\log K$. The resulting best fit $\log K$ was retained, as were the associated values of $\sum X_O$(interface) and $\log D^{18}O_i''$. Then, to find the best fit values for the final two variables, $D^{18}O_O^\times$ and $DV_O^{\bullet\bullet}$, the already-fitted variables were kept constant, and the diffusion-reaction model was run 10,000 times with random guesses of $DV_O^{\bullet\bullet}$ and $D^{18}O_O^\times$, generally spread over 3-8 orders of magnitude. The constraint $D^{18}O_O^\times < DV_O^{\bullet\bullet} < D^{18}O_i''$ was applied. In some cases, the $DV_O^{\bullet\bullet}$ parameter converged towards one of the other Ds, but the $D^{18}O_O^\times < DV_O^{\bullet\bullet} < D^{18}O_i''$ relationship was always maintained, even where values appear identical to 1d.p. in Table 2. For each of the models, the summed squares of the residuals were calculated. The $DV_O^{\bullet\bullet}$ and $D^{18}O_O^\times$ values associated with the lowest residuals were taken as the best fit. Note that this description does not include a treatment of uncertainties associated with individual fits. These are not provided – further information is provided below (section 'Uncertainties').

The fitting routine above was developed for the profiles from the dry experiments, where the position of the interface can be easily identified. For the wet experiments, the modelling is more complex, because the exact position of the interface is not

known – the depth profiles include a contribution from an overgrowth. Effectively, by choosing to place the interface at a different position in the profile, the geometry of the portion attributed to diffusion can be stepped, as with the dry experiments, or can be close to an error function form. Using the $^{18}O/(^{18}O+^{16}O)$ profiles themselves to determine the position of the interface would require some circular reasoning, thus the measured OH/O profiles (see Electronic Appendix 2 and Table S2, electronic supplement) were used. These should be mostly independent of $^{18}O/(^{18}O+^{16}O)$, and should record the position of the interface between a nominally dry starting material and an overgrowth formed in a wet experiment. The issue is that H diffusivity is expected to be much higher than O (e.g., Reynes et al. 2018 and references therein), but, interestingly, the OH/O profiles consistently show a much sharper step than the $^{18}O/(^{18}O+^{16}O)$ profiles. A potential explanation is that some H in the overgrowth is incorporated as molecular H_2O or some other immobile crystal defect. Nevertheless, H diffusion in spessartine and grossular garnet is related to oxidation of Fe and Mn (Reynes et al. 2018), hence it is not possible to rule out that the different chemical composition of YAG might also affect H diffusivity. Regardless of the exact mechanism, the sharp step in OH/O provides an independent constraint on the position of the effective diffusion interface. Therefore, to model the profiles from the wet experiments, the interface was placed at two positions – one representing the mid-point of the sharp drop in OH/O, and one at the point where the overgrowth value starts to decrease from the rimward value (Fig. A3-1). Then, the same fitting routine described above was used to fit the $^{18}O/(^{18}O+^{16}O)$ profiles, and for each wet experiment two sets of fit parameters are given (Table 2).
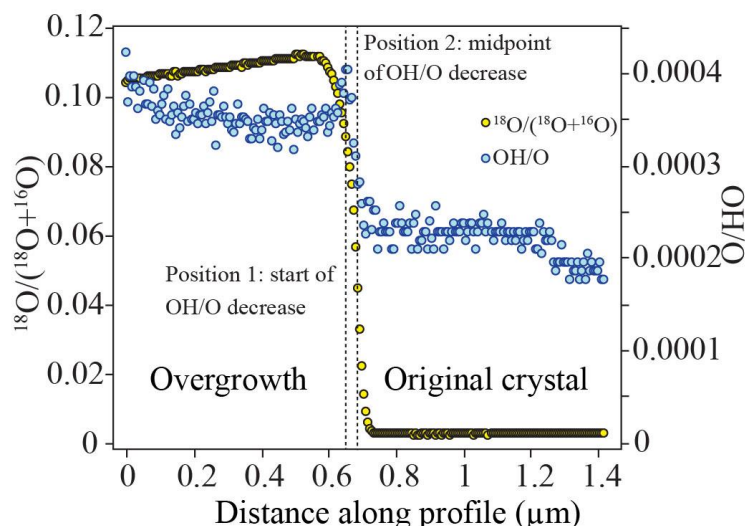
Figure A3-1. Locating the edge of the original crystal using the OH/O signal. The original crystal edge was placed at two positions, marked 1 and 2, as described in the text.

The method described above was developed to fit the profiles with the stepped shapes, but could fit equally well the profiles showing the close to error function forms, but with extended tails (Fig. 7c). The profiles that showed error function forms were fitted to Equation 1.

Uncertainties

We do not provide uncertainties on model fit parameters, but instead provide an estimated range of uncertainties. The following describes how this estimate was derived, and the associated caveats.

We follow the method of Anvi (1976), described in Press et al. (2007), where uncertainties are determined using the constant chi-square method. Such a method has been previously used for estimating uncertainties in diffusion models (e.g., Van Orman et al. 2009). After finding the best fit parameters (K, three Ds, interface and background concentrations), as above, a large number of simulations are run with

parameters varied randomly around those associated with the best fit. For each, the

chi-squared statistic (χ^2) is determined:

$$\chi^2 = \sum_{i=0}^{N-1}\left(\frac{y_i - y(x_i|a_0...a_{M-1})}{\sigma_i}\right)^2 \tag{A3.4}$$

where N is the number of data points, M is the number of adjustable parameters, a,

and σ is the standard deviation associated with each measurement point (eq. 15.1.6

from Press et al. 2007). In this case, y would be $^{18}O/(^{18}O+^{16}O)$, and x is the distance

along the profile. From Equation (A3.4), this calculation requires an uncertainty to be

assigned to each point along the measured profile. In Tables S1 to S3 (electronic

supplement), the formal uncertainties are presented as 2SE, which are inappropriate

for this method. As an alternative, if we make the assumption that the model does

indeed fit the data well and we assume that σ is constant for all points, this can be

recalculated using:

$$\sigma^2 = \sum_{i=0}^{N-1}[y_i - y(x_i)]^2/(N-M) \tag{A3.5}$$

(eq. 15.1.7 from Press et al. 2007) which assumes that the reduced χ^2 value is equal

to 1, ie. $\chi^2 = N\text{-}M$. Whilst this allows the uncertainties on fit parameters to be

estimated, it precludes providing any independent assessment of the goodness of fit.

With these uncertainty estimates, a large number (10^4-10^6) of curves are simulated,

randomly varying the six parameters of interest ($D^{18}O_O^{\times}$, $DV_O^{\bullet\bullet}$, $D^{18}O_i''$, $\sum X_O(initial)$,

K, $\sum^{18}O(interface)$) around their best-fit values. For each model, χ^2 is determined and

tabulated along with the input parameters. Then, the range of values of the input

parameters that give a $\Delta\chi^2$ (minimum χ^2 plus some value) of less than some fixed

value is determined. For 95% confidence intervals, $\Delta\chi^2 = 4$, if uncertainties are

reported separately (rather than simultaneously).

The figures below (Figs. A3-2 and A3-3) show this method applied to two

measured curves, one with a large number of points (YHPD4c1, N=240 points), and

one with a smaller number (YHPD3ns2, N=26 points). The data describing the curves can be found in Table S3 (electronic supplement).

The uncertainties on fitted parameters are considerably larger in the latter, as expected. These are extreme cases, thus we can estimate that a reasonable range of uncertainties (2σ) on the three Ds is $\log_{10}D^{18}O_O^{\times}$: 0.03-1; $\log_{10}DV_O^{\bullet\bullet}$: 0.02-0.7; $\log_{10}D^{18}O_i''$: 0.01-0.05. The reason that uncertainties for each fit are not presented in Table 2 in the main text is that, in order to compute these uncertainties, we are forced to 1) assume *a priori* that the model fits the data well, which is not fully clear given certain observations from the data and discussion (e.g. that the stepped shapes are not observed following 1-atm experiments; the following discussion re. ^{16}O) and 2) determine σ for each point, assuming that 1) is correct.
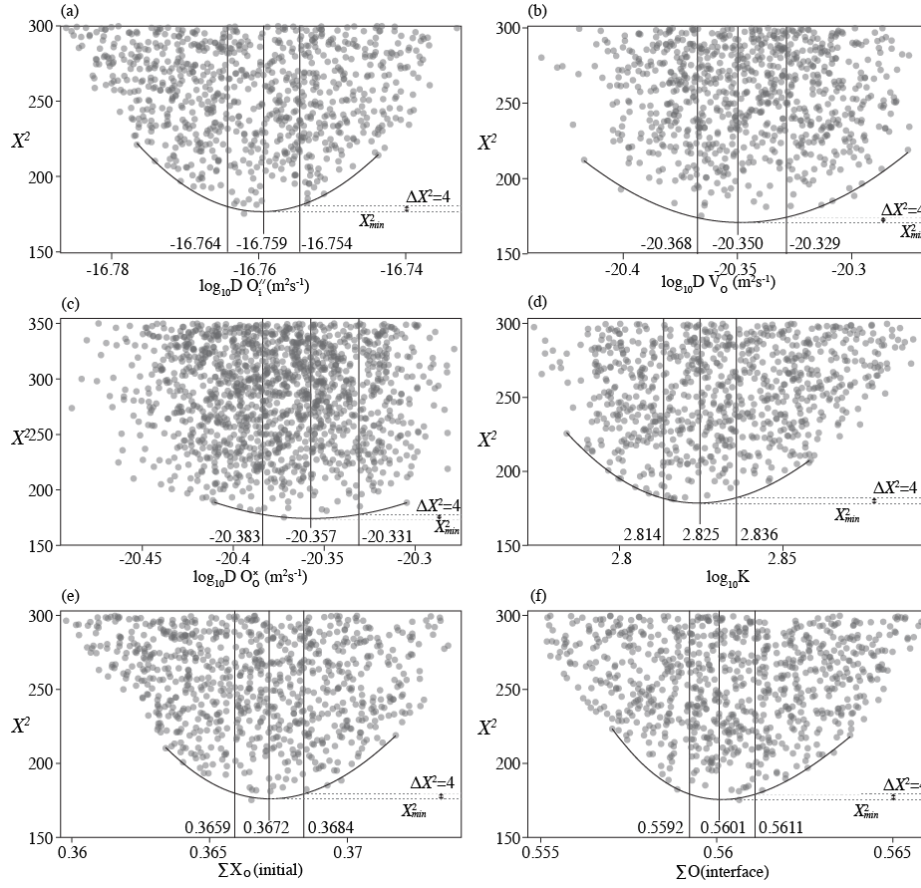
Figure A3-2. The result of $>10^6$ diffusion-reaction models with input parameters ($D^{18}O_O^{\times}$, $DV_O^{\bullet\bullet}$, $D^{18}O_i^{''}$, $\Sigma X_O(initial)$, K, $\Sigma^{18}O(interface)$) varying randomly around the best fit values for experiment YHPD4c1. For each model, the chi-squared statistic was determined, as described above. The best fit value is associated with the minimum chi-squared χ^2_{min}, and the 95% confidence interval is associated with $\chi^2_{min} + 4$. The uncertainties associated with each point were estimated using Equation (A3.5) prior to modelling, these appear to have been slight overestimates given $\chi^2_{min} \approx 170$ and N=240.
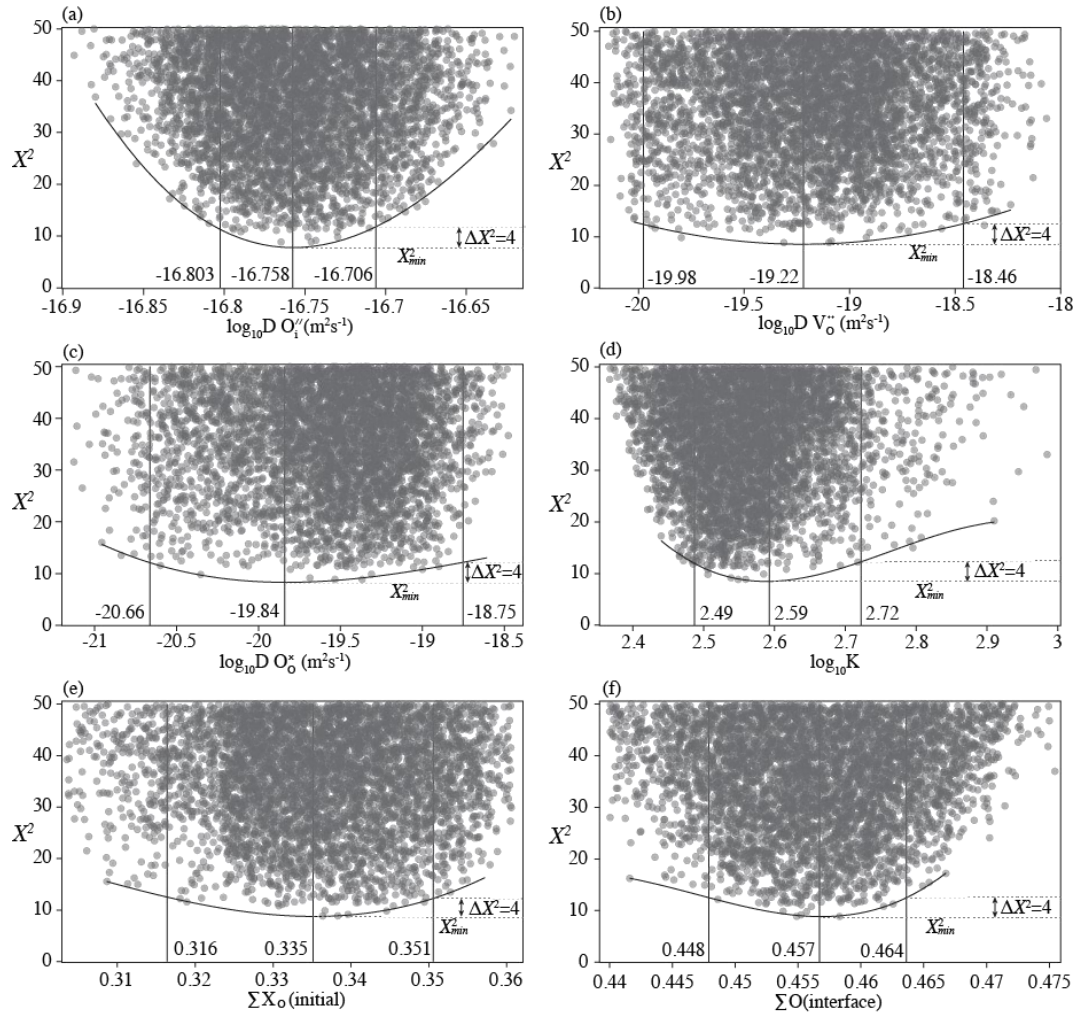
Figure A3-3. As Figure A3-2, for experiment YHPD3ns2, with considerably fewer data points.

Considering ^{16}O

A potential issue with our model is that we consider ^{18}O in the framework one would generally use to describe diffusion of a trace element into a lattice that is initially free of that element. A more complete/correct model would also consider ^{16}O (assuming that ^{17}O is negligible). Whilst we do not incorporate ^{16}O into the main results of this study, below is a description of one possible way in which it could be included into the model, as well as some of the issues that are encountered when doing so, and the potential implications.

If we start with Equation (4) from the main text:

$$^{18}O_i'' + V_O^{\bullet\bullet} = {}^{18}O_O^{\times} \tag{A3.6}$$

a similar reaction is written for ^{16}O:

$$^{16}O_i'' + V_O^{\bullet\bullet} = {}^{16}O_O^{\times} \tag{A3.7}$$

For these two reactions, we have five unknowns ($^{18}O_i''$, $^{16}O_i''$, $^{18}O_O^{\times}$, $^{16}O_O^{\times}$ and $V_O^{\bullet\bullet}$).

Therefore, five equations are needed. The most obvious choices are the two Ks, which

we call K1 and K2 for Equations A3.8 and A3.9.

$$K1 = \frac{[^{18}O_O^{\times}]}{[^{18}O_i''][V_O^{\bullet\bullet}]} \tag{A3.8}$$

$$K2 = \frac{[^{16}O_O^{\times}]}{[^{16}O_i''][V_O^{\bullet\bullet}]} \tag{A3.9}$$

along with the constraint that the number of O sites does not change, and is equal to

12:

$$\sum X_O = 12 = [V_O^{\bullet\bullet}] + [^{18}O_O^{\times}] + [^{16}O_O^{\times}] \tag{A3.10}$$

and relationships describing the concentration of ^{18}O and ^{16}O:

$$\sum{}^{18}O = [^{18}O_i''] + [^{18}O_O^{\times}] \tag{A3.11}$$

$$\sum{}^{16}O = [^{16}O_i''] + [^{16}O_O^{\times}] \tag{A3.12}$$

In this case the solution is found numerically – firstly the concentration of $^{16}O_i''$ is

found as the root of A3.13, where z=$[^{16}O_i'']$:

$$K2^2 \times z^3 - K1 \times K2 \times z^3 + K1 \times K2 \times \sum{}^{18}O \times z^2 + 2 \times K1 \times K2 \times \sum{}^{16}O \times z^2 +$$

$$K2^2 \times \sum X_O \times z^2 - K1 \times K2 \times \sum X_O \times z^2 + K2 \times z^2 - K2^2 \times \sum{}^{16}O \times z^2 -$$

$$K1 \times z^2 + K1 \times K2 \times \sum{}^{16}O \times \sum X_O \times z - K1 \times K2 \times \sum{}^{16}O \times \sum{}^{18}O \times z +$$

$$2 \times K1 \times \sum{}^{16}O \times z - K1 \times K2 \times \sum{}^{16}O^2 \times z - K2 \times \sum{}^{16}O \times z - K1 \times \sum{}^{16}O^2$$

$$\tag{A3.13}$$

then, the values of the other unknowns are found:

$$[V_O^{\bullet\bullet}] = -\frac{[^{16}O_i''] - \sum{}^{16}O}{K2 \times [^{16}O_i'']} \tag{A3.14}$$

$$[^{18}O_i''] = \frac{\sum^{18}O}{K1 \times [V_O^{\bullet\bullet}] + 1} \tag{A3.15}$$

$$[^{16}O_O''] = K2 \times [^{16}O_i''] \times [V_O^{\bullet\bullet}] \tag{A3.16}$$

$$[^{18}O_O''] = K1 \times [^{18}O_i''] \times [V_O^{\bullet\bullet}] \tag{A3.17}$$

If K1=K2, then there is an analytical solution for $[^{16}O_i'']$:

$$[^{16}O_i''] = (\sum^{16}O \times (K1^2 \times \sum^{16}O^2 + 2 \times K1^2 \times \sum^{16}O \times \sum^{18}O - 2 \times K1^2 \times \sum^{16}O \times$$

$$\sum X_O + K1^2 \times \sum^{18}O^2 - 2 \times K1^2 \times \sum^{18}O \times \sum X_O + K1^2 \times \sum X_O^2 + 2 \times K1 \times$$

$$\sum^{16}O + 2 \times K1 \times \sum^{18}O + 2 \times K1 \times \sum X_O + 1)^{\frac{1}{2}} - \sum^{16}O + K1 \times \sum^{16}O^2 +$$

$$K1 \times \sum^{16}O \times \sum^{18}O - K1 \times \sum^{16}O \times \sum X_O)/2 \times (K1 \times \sum^{16}O + K1 \times \sum^{18}O))$$

$$\tag{A3.18}$$

Eq. A3.13 may have multiple roots within a reasonable range of $^{16}O_i''$ concentrations,
but only one is associated with positive concentrations of all five unknowns.

With the system defined, a diffusion-reaction model is constructed. In this case, the

root of Equation A3.13 is found at every grid point independently after every

diffusion step, which makes it somewhat more computationally intensive than the

relatively simple model described above, where only ^{18}O is considered (which has an

analytical solution).

When this is modelled, the outcome is profiles with forms broadly describable as

error functions for ^{18}O (Fig. A3-4). Stepped shapes for ^{18}O could not be generated

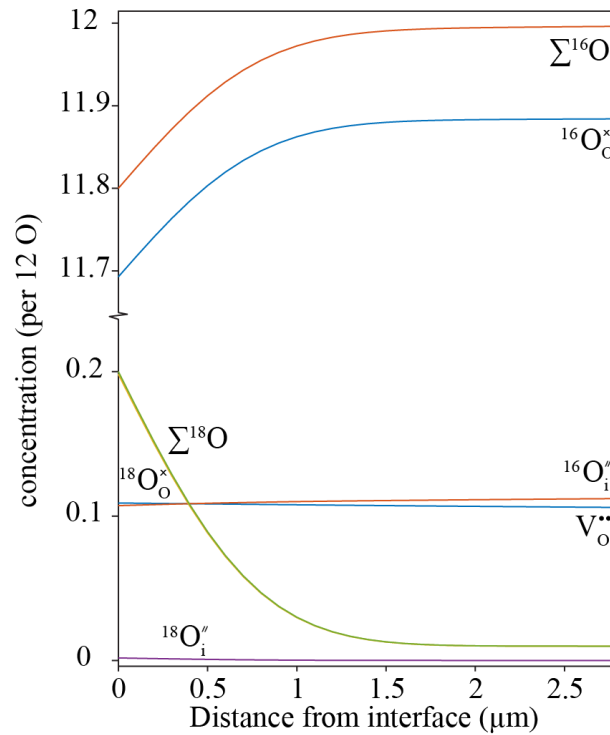with any reasonable initial or boundary conditions.

Figure A3-4. An example of a model incorporating ^{16}O, as described by the equations above. In this model, $K1=K2=1000$; $\Sigma^{16}O$ (initial)$=11.99$; $\Sigma^{18}O$ (initial)$=0.01$; $\Sigma^{16}O$ (interface)$=11.8$; $\Sigma^{18}O$ (interface)$=0.2$; $\Sigma X_O=12$; $D^{16}O_O^\times=D^{18}O_O^\times=10^{-19}m^2s^{-1}$; $D^{16}O_i''=D^{18}O_i''=10^{-17}m^2s^{-1}$; $DV_O^{\bullet\bullet}=10^{-17}m^2s^{-1}$; total time $=10^6$ s.

Therefore, simply expanding the model presented in the text by adding ^{16}O does not enable the observed stepped profile shapes to be recreated. This supports the assertion, as stated in the main text, that the model used to fit the profiles is a simplification. There must be some physical process occurring in these experiments that is not incorporated into the model (e.g., the coupled flux of another element, or certain defects, that might be shown in the CL data), and of which we cannot provide a satisfactory description. However, as long as the basic assumption, that there are at least two diffusion mechanisms for O, and the ability for O to move between the associated substitution mechanisms, is correct, then our diffusion coefficients should remain valid.

Implications.

The bulk concentration (C) as a function of time (t) is given by:

$$C(t) = C_0 + (C_1 - C_0)\left(1 - \frac{6}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\exp\left(\frac{-Dn^2\pi^2 t}{R^2}\right)\right) \qquad (A3.19)$$

Whereas C at any radial position is given by:

$$C(r,t) = C_0 + (C_1 - C_0)\left(1 + \frac{2R}{\pi r}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\sin\left(\frac{n\pi r}{R}\right)\exp\left(\frac{-Dn^2\pi^2 t}{R^2}\right)\right) \qquad (A3.20)$$

It is also informative to consider the composition at the crystal core, i.e. $r\to 0$, as a function of time:

$$C(r\to 0,t) = C_0 + (C_1 - C_0)\left(1 + 2\sum_{n=1}^{\infty}(-1)^n\exp\left(\frac{-Dn^2\pi^2 t}{R^2}\right)\right) \qquad (A3.21)$$

Additional fits of complexly shaped profiles using the reaction-diffusion model. In this section, all fits for complexly shaped profiles using the reaction-diffusion model are reported. The name of each profile indicates (1) the experimental charge (e.g., YHPD-1), including information on garnet composition (Y = YAG, P = pyrope), pressure (HP = piston cylinder apparatus experiments) and presence/absence of water during the experiment (D = dry, W = wet); (2) the analytical method used to measure the profile (i.e., c = CAMECA IMS-1280, ns = NanoSIMS); (3) the number ID of the measured profile as reported in the supplementary Tables S1 and S3. For wet experiments, 2 sets of parameters are reported for each profile indicated by the letters 'a' and 'b', depending on the chosen location of the interface.

YHPD-3_ns1

YHPW-1_c1a

YHPW-1_c1b

YHPW-1_c2b

YHPW-2_c1a

YHPW-2_c1b

YHPW-2_c2a

YHPW-2_c2b

YHPW-2_c3a

YHPW-2_c3b

YHPW-3_c1a

YHPW-3_c1b

PHPW-1_c1a

PHPW-1_c1b

PHPW-1_c2a

PHPW-1_c2b

PHPW-1_c3a

PHPW-1_c3b