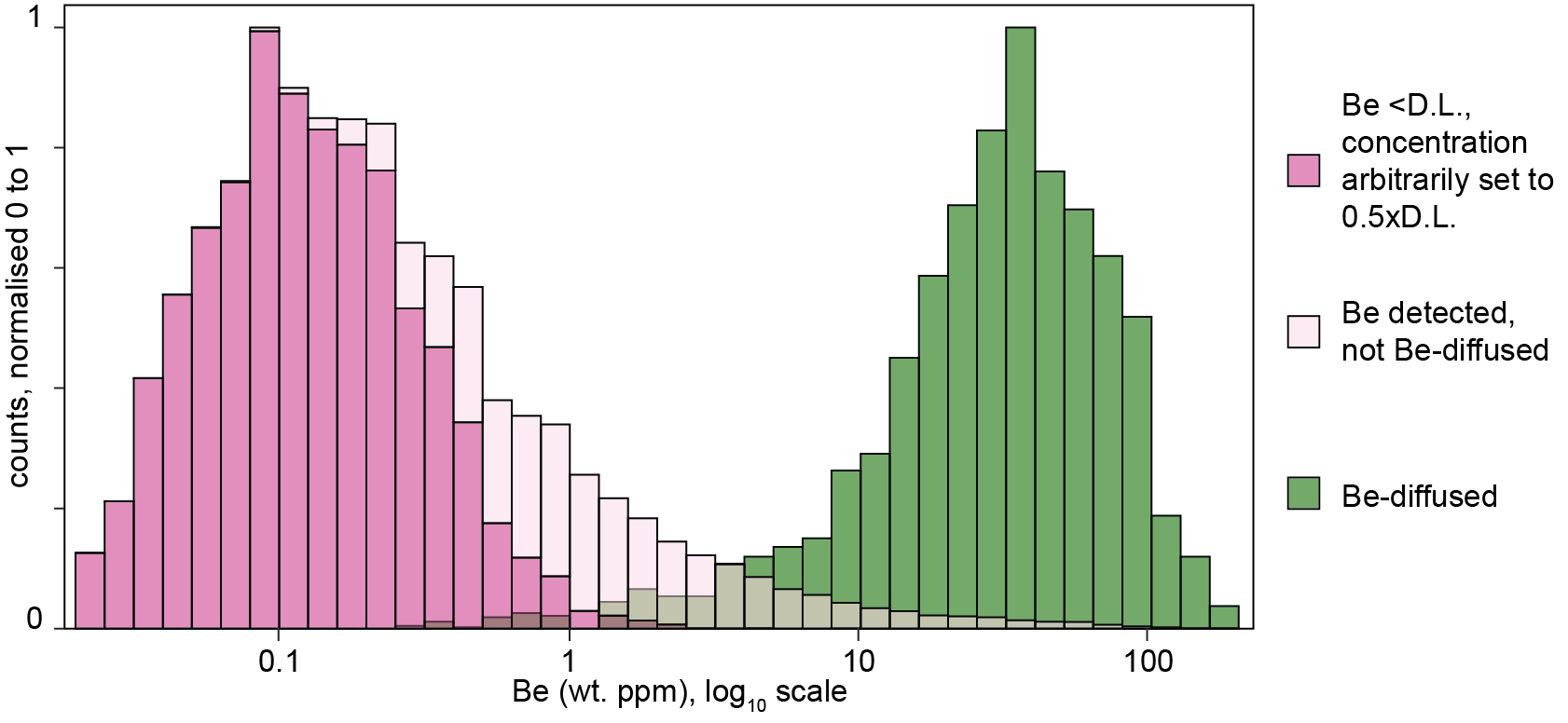
**Supplementary material**

**Beryllium and hydrogen incorporation in corundum**

*Be concentrations in Be-diffused and natural corundum crystals*

In the main text it was stated that the normal Be concentrations in natural corundum were generally below 2 at. ppm, and the range for Be concentration in Be-diffused crystals is generally a few at. ppm to 10s at. ppm. These values are drawn from an evaluation of 100,000 LA-ICP-MS analyses of corundum conducted at laboratories of the Gemological Institute of America (GIA) drawn from a larger database. Histograms showing the distribution of Be concentrations in corundum crystals determined as Be diffused, and determined as not Be-diffused, are presented in the following figure. Data where Be was below the detection limit of analysis is also included in the not-Be-diffused population for the purpose of determining percentiles – in these data, the Be concentration is assigned an arbitrary value of 0.5 x detection limit. 90% of not-Be-diffused crystals have Be concentrations < 1.6 at. ppm, 95% have Be < 3.2 at. ppm. For the Be-diffused samples, the 5th and 95th percentiles are 4.5 and 133 at. ppm, respectively.



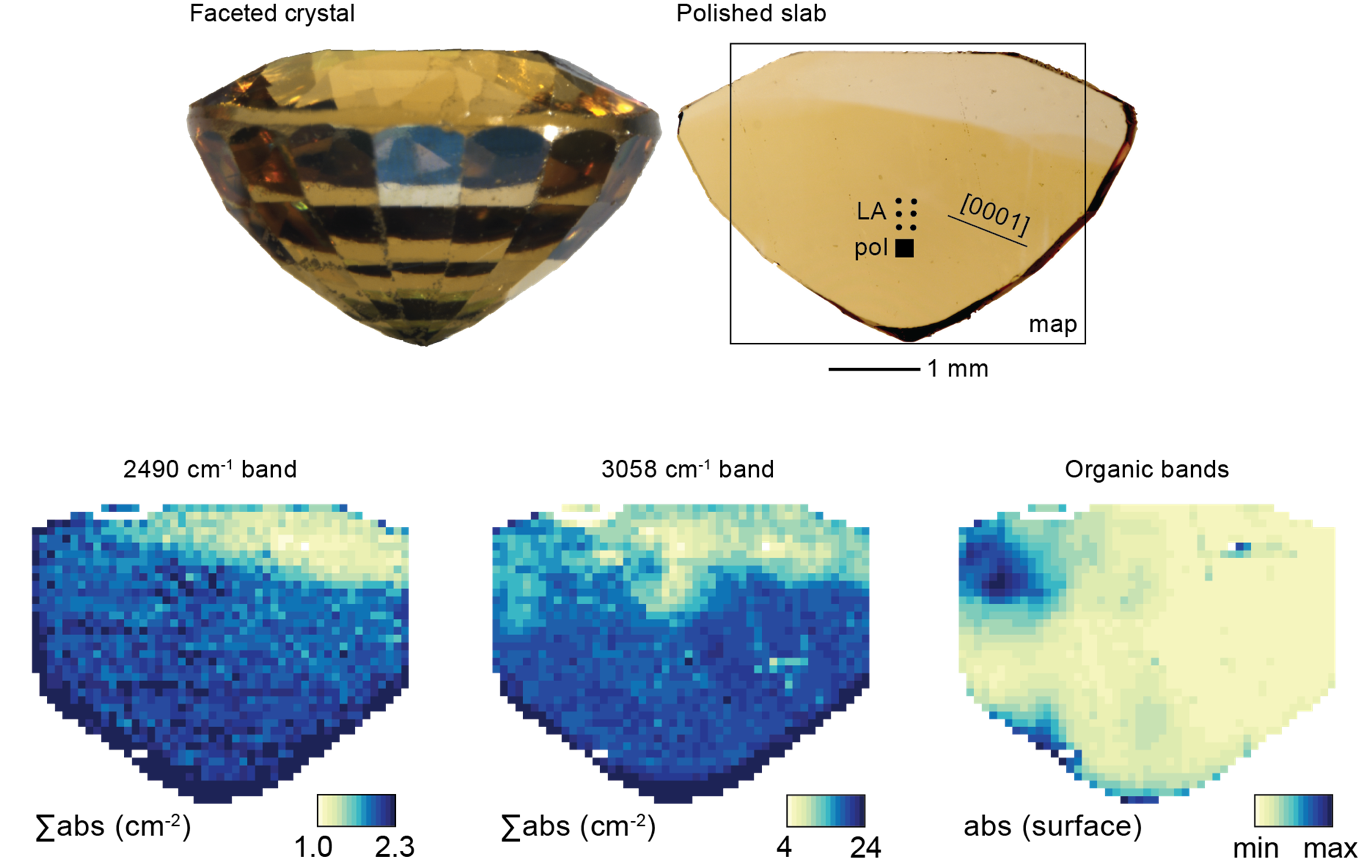
*Spectra and metadata from the 1000 crystal dataset*

These are provided in the accompanying online dataset. In each file, the top rows contain metadata, including path length, Be concentration ranges and colour, and the remaining rows contain the spectral data (wavenumber and absorbance). In many cases, spectra are affected by considerable noise in the >3500 cm-1 region from water.

*Preparation and analyses of a Be-diffused corundum crystal*

As described in the text, one Be-diffused crystal was selected from the research collection of the GIA for analysis. The following figure shows the initial faceted crystal and the slab cut and polished from the centre. The position at which polarised spectra were acquired is shown (“pol”), as well as the [0001] direction. The LA-ICP-MS spot locations are also shown (“LA”).

Also shown are the mapped intensities of the two unresolved bands (2490 and 3062 cm-1) mapped by FTIR spectroscopy, as described in the main text, and the organic bands. The intensities of the two unresolved bands are corrected to 1 cm thickness and to total absorbance using polarization factors presented in the main text. The absorbance of both unresolved bands corresponds to the crystal colour, with the stronger yellow colour associated with higher absorbance. The organic bands are patchy and not correlated with the unresolved 2490 and 3062 cm-1 bands.



Laser ablation inductively coupled plasma mass spectrometry analyses were conducted at the laboratory of the GIA in New York, USA, using a ESI 213 nm laser coupled to a Thermo iCAP-Q quadrupole mass spectrometer. The laser conditions were: 55 µm diameter circular spot, 20 Hz, 10-11 Jcm-2. The counted isotopes and time-per-sweep are presented in the following table. The data were processed using an in-house spreadsheet. The internal standard was Al, and the external standards were in-house trace element-doped corundum crystals and NIST 610 and 612 glasses, with analyte concentrations for the Be diffused crystal in the following table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Element | Isotope(s) | Time per sweep (s) | Concentration (at. ppm) | Concentration (wt. ppm) |
| Be | 9 | 0.01 | 29.9 (0.6) | 13.2 (0.3) |
| Mg | 25 | 0.01 | 10.4 (0.9) | 4.6 (0.4) |
| Al | 27 | 0.005 | 103.2 (26.9) | 45.7 (11.9) |
| Ti | 47 | 0.01 | 17.7 (1.2) | 7.8 (0.5) |
| V | 51 | 0.05 | 4.6 (0.2) | 2.1 (0.1) |
| Cr | 53 | 0.05 | 3.5 (0.4) | 1.5 (0.2) |
| Mn | 55 | 0.01 | 1.8 (0.3) | 0.8 (0.1) |
| Fe | 57 | 0.05 | 14800 (290) | 6550 (130) |
| Zn | 66 | 0.01 | 1.8 (1) | 0.8 (0.4) |
| Ga | 71 | 0.01 | 454 (1.7) | 201 (0.8) |
| Nb | 93 | 0.01 | 0.08 (0.05) | 0.03 (0.02) |
| Ta | 179 | 0.01 | 0.11 (0.03) | 0.05 (0.01) |

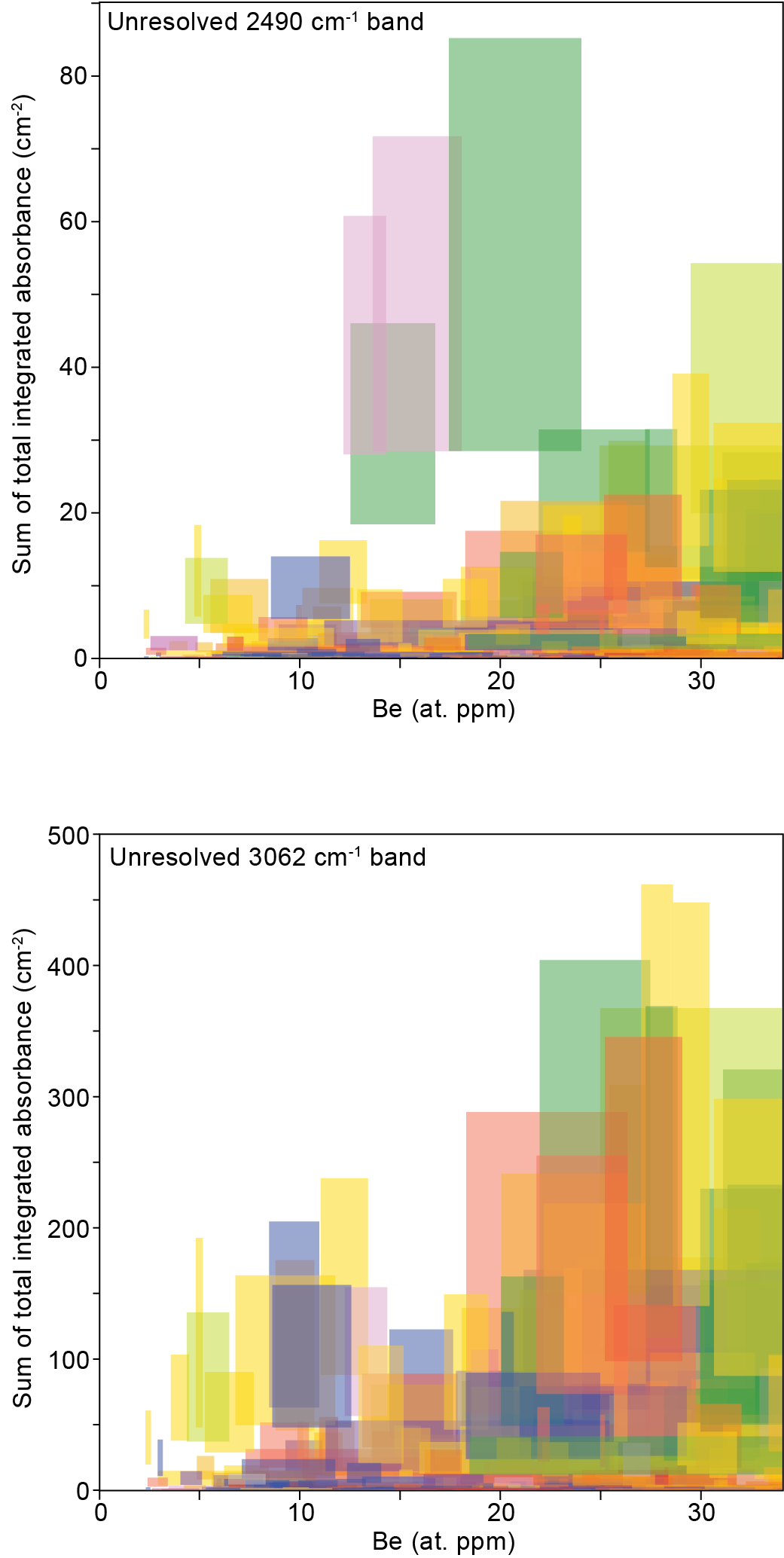
The conversion from wt. ppm (as measured) to at. ppm is made using the formula at. ppm of element= ((molar mass Al2O3/5)/(molar mass element)) x wt. ppm element, i.e., assumes infinite dilution, so will be inaccurate for Fe. Values represent the mean and standard deviation of six analyses. Other elements that were below detection limits: B, Na, Si, K, Ca, Sc, Co, Ni, Cu, Y, Zr, Mo, Ru, Rh, Pd, Ag, Cd, Sn, La, Ce, Nd, Hf, W, Ir, Pt, Au, Pb, Bi, Th, U.

*Polarised Spectra*

The raw polarised spectra are provided in the accompanying dataset. The nominal polarizer angle and corrected angle are presented, with the correction made using Al-O overtones. 90 degrees corresponds to E||[0001].

*Relationship between Be concentration and both unresolved bands*

In the main text, a figure is presented showing the concentration of Be versus the summed integrated area of both unresolved bands (2490 and 3062 cm-1). It was stated that the relationship between the Be concentration and each of the unresolved bands was similar to the relationship between Be concentration and their sum. This is shown in the following figures, which both show an unoccupied region in the top left.



*Crystallographic Information Files (cif)*

Defect structures are provided in cif format in the accompanying dataset.