This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2017-6062

Revision 1: Quantitative electron backscatter diffraction data (EBSD) analyses using the dictionary indexing (DI) approach: overcoming indexing difficulties on geological materials

Katharina Marquardt; Bayerisches Geoinstitut, Universität Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Deutschland

Marc De Graef; Materials Science and Engineering Department, Carnegie Institute of Technology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890, USA

Saransh Singh; Materials Science and Engineering Department, Carnegie Institute of Technology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890, USA

Hauke Marquardt; Bayerisches Geoinstitut, Universität Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Deutschland

Anja Rosenthal; Bayerisches Geoinstitut, Universität Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany, AND Laboratoire Magmas et Volcans, Université Blaise Pascal, CNRS IRD-OPGC, Campus Universitaire des Cézeaux, 6 Avenue Blaise Pascal, 63178 Aubière Cedex, France

Sanae Koizuimi; University of Tokyo, Earthquake Research Institute, 1-1-1, Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan

Abstract

Electron backscatter diffraction data (EBSD) yield plentiful information on microstructure and texture of natural as well as experimentally produced mineral and rock samples. For instance, the characterization of microstructures and textures by EBSD allows for the evaluation of phase equilibria. Furthermore, determination of the preferred orientations of crystals using EBSD yields constraints on deformation mechanisms and history of the minerals/rocks. The latter affects bulk rock properties such electrical conductivity and seismic anisotropy. EBSD is also applied to advance our understanding of various phenomena such as seismic wave attenuation in the Earth deep interior that might be caused by the presence of interfacial small degrees of melt fractions, or free fluid phases.

In standard EBSD software solutions, the original EBSD patterns are rarely saved and indexing routines result in many artifacts, such as pseudo-symmetry or unindexed pixels at interfaces that may be misinterpreted as amorphous material, such as a melt.

Here we report the first application of an extension of the dictionary indexing (DI) approach proposed by Chen et al. (2015), an alternative indexing routine, to multiphase geologic materials. The DI method is independent of the EBSD system, and thus of the used detector/software. The DI routine generates simulated EBSD patterns for all possible crystal orientations, taking the sample composition and experimental setups into account. The resulting pattern database is called a dictionary. The experimental Electron Backscattering Pattern (EBSP) images are indexed by comparing them to the dictionary using a dot-product algorithm. We evaluate the new DI method in comparison to standard routines and highlight advantages and disadvantages.

To test and compare the DI's reliability and performance, we apply the routine to two scientifically challenging samples: (1) A nominally anhydrous ('dry') residual eclogite composed of garnet (cubic), clinopyroxene (monoclinic) and an amorphous melt, where the different degrees of hardness of the phases cause surface topology; and (2) a pure forsterite (olivine) polycrystalline sample produced by vacuum sintering (Koizumi et al. 2010). The acquired EBSD patterns are of low quality for the latter as a result of fast data acquisition to reduce the on-line machine time.

We conclude that the new DI method is highly precise and surpasses the performance of previously available methods while being computer time and -memory consuming. We find that the DI method is free of pseudo-symmetry-related problems. Interpolation of data becomes obsolete and high reproducibility is obtained, which minimizes the user impact on the final dataset. The latter is often caused by applying several cleaning steps on EBSD maps with low indexing fraction. Finally, much higher scientific integrity is ensured by the image collections as described above, which requires that all patterns are saved. This in turn allows later re-analyses if required. The DI routine will help to achieve more reliable information on interface properties of geological samples, including amorphous materials, and to improve the accuracy of large-scale Earth mantle process models.

Keywords

EBSD - electron backscatter diffraction, DI - dictionary indexing, EBSD simulation, melt distribution, wetting angles, GBCD - grain boundary character distribution, grain boundary, dunite, olivine, eclogite, Earth's mantle processes

Introduction

Electron backscatter diffraction (EBSD) is a key technique to characterize polycrystalline samples especially when focusing on crystal fabrics (e.g. crystal preferred orientation, CPO; shape preferred orientation, SPO), defect structures such as interfaces, phase-, grain-, and twin boundaries, as well as the distribution of impurities or melt on interfaces (e.g. Faul and Scott 2006; Le Roux et al. 2008; Manthilake et al. 2013; Soustelle et al. 2014).

Manual EBSD techniques have been developed in 1970's, where electron backscatter pattern (EBSP) were collected on electron-sensitive photographic plates (Venables and Harland 1973). Interactive indexing was developed by David Dingley in the late 1980's (e.g. Dingley and Baba-Kishi 1986). Automated EBSD was considerably advanced in the early '90s by Brent Adams and Stuart Wright (Adams et al. 1993) but the first Kikuchi images date back to Kikuchi (Kikuchi 1928). Quantitative texture analysis is routinely performed with millimeter to tens of nanometer spatial resolution. Collected orientation data are used to obtain grain size distributions (Adams et al. 1993), phase fractions, dislocation densities and surface information (Wright and Nowell 2006). Furthermore, the application of EBSD allows for the investigation of deformation mechanisms (e.g. Wheeler et al. 2001; Michibayashi and Mainprice 2004; Warren and Hirth 2006; Prior et al. 2011; Michibayashi et al. 2016) and the identification of new phases (Bandli and Gunter 2012). EBSD has developed into a standard tool for the characterization of fabrics in rocks and materials (e.g. Prior et al. 1999; Schwartz, A.J., Kumar, M., Adams, B.L., Field 2009; Morales et al. 2011). Its application also allows for the derivation of the orientation distribution function (Bunge and Esling 1982), for which 10,000 grains already provide statistically robust information (Maitland and Sitzman 2007;

Wright et al. 2007), while others trust calculate fabric strength from a total number of crystals of only 100 to 150 (e.g. Ismaïl and Mainprice 1998). Many of these types of information can be retrieved from EBSD orientation data using freely available softwares, such as MTEX (Bachmann et al. 2010). MTEX permits highly flexible data processing, scripting for batch processing of large data sets and precisely reproducible data treatment. Likewise, more sophisticated information, such as internal grain deformation and strain measurements using high (angular) resolution EBSD (HR EBSD) (Crawford and Was 1991; Wilkinson et al. 2006; Ram et al. 2016) can be gained. Calculating the geometrically necessary dislocation density has had great impact on deformation quantification in recent years (Cordier et al. 2014). Moreover, grain boundaries are considered a neglected key for deformation in general (Lloyd et al. 1997; Sun et al. 2016). More recently, EBSD has evolved into a robust method that allows for the study of the characteristics of interfaces by enabling the acquisition of maps with high spatial resolution over large areas, including grain boundaries (Saylor et al. 2004b; Dillon and Rohrer 2009; Rohrer 2011; Kelly et al. 2016). Such methodical advances overcome the lack of statistical accuracy (Lloyed et al., 1997).

AltHough-EBSD hardware has evolved quickly, software solutions have remained Hough-transform based (i.e. a routine that allows to identify Kikuchi bands from electron backscatter pattern by orientation and intensity), and have not advanced to meet the requirements for the study of interface properties. In standard software solutions, collected Electron Backscattering Patterns (EBSP) are directly processed and typically only the extracted orientation data are permanently saved, while the original data (the EBSPs) are discarded. We like the reader to note that both Oxford and EDAX systems allow for saving

the EBSPs for example in multiple image formats (jpg, bmp, tiff, etc.). Recently, binary compressed formats (pat and hdf5) have become available. This allows for re-processing of the data using commercial software packages (Wright et al. 2015), but also facilitates a more extensive use of the dictionary indexing (DI) routines (Chen et al. 2015). EBSD users should benefit from these developments, because if the indexing procedure is later found to be wrong or of poor quality, the data can be re-interpreted in contrast to former times (Pinard et al. 2011).

A pixel is also discarded if no indexing solution for the pattern is found by the standard software, even if Kikuchi bands were present in the EBSP. Generally, data processing is still far from optimal, which limits the power of EBSD (Tao and Eades 2005). Furthermore, indexing results and quality vary strongly with the operator's experience, sample surface preparation as well as software settings that are often difficult to access within the various software packages available. For example, the settings of the Hough-transform will strongly influence which bands can be detected in an experimental Kikuchi pattern; these settings include Hough-space resolution, Butterfly-masks, size, minimum peak intensity, minimum distance between peaks, peak symmetry and binned pattern sizes. The indexing result also depends on how many bands are chosen to be included in the indexing routine. However, the optimal value depends, in turn, on the software used in the routine. Different strategies include the use of band orientation, width and intensity. The latter is applied only for the selection of the extent of the use of the hkl's in the database required for indexing. Thus, the information encoded in relative band intensities is discarded (Nolze and Winkelmann 2016). The power of full data acquisition, storage, and the superiority of post-processing for noisy

EBSD patterns, scanning electron microscope (SEM) images and chemical information (Tao and Eades 2005; Payton and Nolze 2013; Wright et al. 2015a,b), as well as the superior accuracy of the DI approach (Ram et al. n.d.) have previously been emphasized.

In this study, we propose an extension of an alternative indexing routine, the DI approach (e.g. Chen et al. 2015), and evaluate the applicability of the method to overcome long-standing difficulties and uncertainties using EBSD on material interface properties related problematics. The here-presented results are obtained from data sets where all EBSPs are saved. We analyzed two challenging samples to test the new DI routine. The first sample, a nominally 'dry' residual eclogite Res2_16 (Rosenthal et al., 2014) contains a melt phase that cannot by indexed by standard EBSD software's. Res2_16 also contains clinopyroxene, a mineral with monoclinic crystal structure, as well as garnet (cubic). The second sample is an uncoated nominally 'dry' forsterite (Mg_2SiO_4) polycrystalline sample with a small average grain size of about 6 µm. In the latter olivine sample, surface charging effects and applied fast acquisition conditions to reduce the on-line machine time result in low pattern quality. For data processing, either the Ametek® EDAX TSL OIM™ Analyses (7) EBSD-software (hereafter: OIM) or the Oxford instruments HKL CHANNEL5 software (hereafter HKL) were employed and compared to the DI approach applied to the same data sets. In both software packages, OIM and HKL, all individual EBSD patterns are Hough-transformed. The Hough-transform replots every pixel of an EBSP by the distance and angle of a vector pointing to this pixel. High intensity peaks are subsequently indexed, using a database calculated from the crystal structure and lattice parameters where the fitting includes angles between bands, which are the angles between lattices planes. For a set of three bands and angles a crystal

orientation can be determined; often several such triplets are used to obtain a more unambiguous indexing.

In contrast, to obtain the standard Euler angle information using the DI method, all experimental EBSPs are compared with a database (dictionary) of simulated EBSPs until a best match is obtained. The EBSPs are simulated using a physics-based forward model (Callahan and De Graef 2013) combined with a uniform sampling of orientation space (Roşca and De Graef 2013; Singh and De Graef 2016). The procedure is explained in detail below.

Samples and Methods

The methods part is divided into three sections. First we introduce the two samples investigated. Then we summarize the specific settings used for EBSP acquisition and indexing using OIM and HKL systems, respectively. Finally, we outline the dictionary indexing (DI) approach.

Sample Materials

Sample 1. The 'dry' (nominally anhydrous) residual bimineralic eclogite, containing sub-/euhedral crystals of garnet up to 70 µm, tabular crystals of clinopyroxene up to ~30-40 µm, and a melt fraction (Res2_16; Rosenthal et al. 2014).

This sample represents a residue, which suffered melt extraction of a 'dry' residual coesite-bearing eclogite at 3 GPa and 1350°C on adiabatic ascent.

Garnet in our sample is a solid solution of mainly pyrope ($Mg_3Al_2Si_3O_{12}$; 52.6 mol.%), almandine ($Fe_3Al_2Si_3O_{12}$; 29.6 mol.%), and grossular ($Ca_3Al_2Si_3O_{12}$; 17.8 mol.%) and contains minor TiO_2 (0.56 wt.%) and Na_2O (0.20 wt.%). Clinopyroxene is a solid solution of jadeite

(NaAlSi$_2$O$_6$; 22.6 mol.%), diopside-hedenbergite (CaMgSi$_2$O$_6$-CaFeSi$_2$O$_6$; 33.0 mol.%), enstatite-ferrosilite (MgMgSi$_2$O$_6$-FeFeSi$_2$O$_6$; 21.8 mol.%), Ca-tschermaks (CaAlAlSiO$_6$; 13.8 mol.%), Ca-eskolaite (Ca$_{0.5}$AlSi$_2$O$_6$; 5.5 mol.%), and Al-buffonite (CaMg$_{0.5}$Ti$_{0.5}$AlSiO$_6$; 3.3 mol.%). The quenched liquid (melt) is of basaltic-andesitic composition.

The experiment ran in an end-loaded Boyd-England half-inch (1.27 cm) piston-cylinder apparatus at the Research School of Earth Sciences (RSES), Australian National University (ANU) for 146 hours. Experimental and analytical procedures are outlined in Rosenthal et al. (2014).

Sample 2. The forsterite sample was prepared using the vacuum sintering method of Koizumi et al. (2010). The polycrystalline sample was synthesized from mineral powders with particle sizes of < 100 nm. The source powders of high purity colloidal SiO$_2$ and Mg(OH)$_2$ were mixed in a ball mill in highly pure ethanol solvent for about a day. This is followed by a calcination step in an alumina tube furnace at 960°C under the flow of oxygen to remove the decomposition products of H$_2$O and CO$_2$ efficiently. One of the most important points to obtain these highly dense compacts is to use fully reacted nm-particle as reaction are associated with volume changes that can disturb the densification. The calcined powders were compacted to the desired shape and pressed at 200 MPa for 10 minutes. Sintering was performed at vacuum of about 1* 10-3 Pa at temperature of 1380°C for 10 min. The final dense polycrystalline sample contains less than 2 vol% of enstatite to buffer the silica activity. The average grain size is about 6 µm.

EBPS acquisition and standard analysis using commercial EBSD software packages

EBSD of Sample 1. The EBSD patterns of the garnet-clinopyroxene-melt sample were

198 acquired using a Zeiss Leo Gemini 1530 Schottky FE-SEM operated at an accelerating voltage
199 of 20 keV and a probe current of about 2.0 -2.5 nA (60 mm aperture and high current
200 option).
201 The specimen was placed at 15.2 mm working distance, which corresponds to a sample-
202 scintillator distance fitted by the DI algorithm of 14.828 mm. The scans were run at
203 0.3µm*0.3µm step size. The microscope is equipped with a NordlysS camera from Oxford
204 Instruments with a CCD resolution of 1344*1024*12bit. Acquisition was set to 0.152 seconds
205 per pattern and good pattern quality was obtained, while having a carbon coating of 4nm
206 thickness. Patterns were captured with 4*4 binning on the camera, with window averaging of
207 8, an averaged background for background subtraction, maximum camera gain, and no
208 image processing besides automated contrast enhancement (details are given in Oxford
209 Channel 5 User Manual). After binning, the pattern was further compressed by 50% into
210 "*.jpg" format. These further compressed images were analyzed using the DI routine.
211 Automated indexing routine and analyses were performed using HKL.
212 Garnet was indexed using the cubic symmetry Ia3d and lattice parameters of a=11.459 Å,
213 which may vary up to 11.87 Å related to substitution. Clinopyroxene was indexed using the
214 crystal parameters of diopside ($CaMgSi_2O_6$) in the monoclinic symmetry and lattice
215 parameters included in the HKL database, $a = 9.746\,\dot{A}, b = 8.990\,\dot{A},\ c = 5.251\dot{A}$ and
216 $\alpha = 90°, \beta = 105.629°,\ \gamma = 90°$. For the automated SEM-EBSD measurements, the
217 reliability of the indexing procedure is critical (Prior et al. 1999). The reference file consisted
218 of the 8 strongest reflections for garnet, the 15 strongest reflections for clinopyroxene and
219 no reference for the melt phase. The reflections were calculated from the scattering factors

of the atoms using the kinematical diffraction theory. Even tHough-multiple scattering occurs in the interaction volume of the electron beam with the bulk sample (dynamic diffraction), the kinematical intensities of the reflections are approximately correct. In more than 56% of the data points a good solution was found, the other 44% were not indexed. The criteria for rejection of a solution were (1) if less than 5 reflection bands could be identified, where three non-coaxial bands are considered sufficient to yield a unique solution; (2) if the mean angular deviation of the calculated reflection bands in Hough-space was larger than 1.3°. One source of problems arising from the indexing algorithm is a low indexing success rate across boundaries. This is caused by the overlapping diffraction patterns from two crystals at grain boundaries and results in a low confidence index and ultimately no solution. To obtain maps with 100% indexed data points, post processing would be required to assign grain boundary data points to orientations that are identical to the orientations of the majority of its eight neighbors, and thus one or the other grain. However, such cleaning of data is not recommended in the presence of a melt phase, as the system has no way to distinguish between different types of unindexed patterns. Thus, no post processing was performed. Nevertheless, it might be possible to discriminate the phases based on the Band Contrast map, where band contrast is lower in an amorphous region; band-free patterns can be compared to a pattern resulting from the overlap of two (band-rich) patterns.

EBSD of sample 2. The EBSD patterns of the forsterite sample were acquired on a FEI Quanta SEM at Carnegie Mellon University (Pittsburgh, PA, USA) using a probe current of 2 nA, 30 kV acceleration voltage, and a step size of 0.2µm*0.2 µm. The microscope is equipped with a DigiView 5 camera with 1395*1040*12 bit resolution and the EDAX/TSL OIM DC software

(version 7.0; EDAX/ TSL, Draper, UT, USA). Patterns were captured with 4*4 binning on the camera. A maximum camera gain, background removal but no image processing were applied. The recorded patterns have a dimensionality of 61*61 pixels in .jpg format.

For enstatite we used orthorhombic symmetry Pbca and lattice parameters of $a = 18.241\ \dot{A}, b = 8.83\dot{A}\ c = 5.185\ \dot{A}$. For forsterite, we assumed orthorhombic symmetry Pbnm with the lattice parameters $a = 4.761\ \dot{A}, b = 10.225\ \dot{A}\ c = 5.994\ \dot{A}$.

The general parameters were set to a binned pattern size of 120 pixels with a theta step size of 1°, i.e. a Rho fraction of 80%. The Max Peak Count was set to 11, the minimum to 5. These parameters were kept constant for all indexing runs.

To obtain crystal orientation maps, data were processed using OIM. First, the Kikuchi bands utilized for orientation derivation were detected using the classical two-dimensional Hough-transform-based algorithm. They were applied to images, which were subject to the following image processing: background subtraction, median smoothing, dynamic background division, and normalization of intensity histogram. We used a classic Hough-convolution mask of 9*9. The minimum and maximum Hough-peak magnitudes were set to 5 and 19 respectively. Peak symmetry was set to 0.6, and 7 bands were chosen to index forsterite, 25 for enstatite.

Dictionary indexing (DI) approach

The DI approach uses a set of pre-computed EBSD patterns (EBSPs), known as “the dictionary,” to find a best matching pattern for each experimental EBSP (Park et al. 2013; Chen et al. 2015). A schematic working procedure of the dictionary indexing process is

illustrated in figure 1. To obtain simulated EBSPs as similar to the experimental EBSPs as possible, a physics-based Generalized Forward Projector (GFP) (Callahan and De Graef 2013) of the EBSD signal is employed. This EBSP generation process is combined with a geometrical model of the detector system (including the pattern center coordinates, the detector pixel size, and the distance between detector and sample), and a uniform sampling of the orientation space *SO(3)*; the resulting electron backscatter diffraction pattern is called the master pattern, depicted in figure 2. This master pattern can be thought of as the back-scattered electron (BSE) yield on a spherical surface for a hypothetical spherical sample with an electron source at its center. *SO(3)* is the set of special orthonormal 3x3 matrices that represent 3D orientations. The sampling approach is based on an equal-volume mapping between a cubical grid, the "cubochoric space," and the northern hemisphere of unit quaternions, which is isomorphic with *SO(3)*. The approach thus creates a uniform sample of orientations that provide the pattern labels in the dictionary (Roşca et al. 2014). The spacing of the sampling grid corresponds to an average angular step size; selecting a finer grid reduces the angular step size in orientation space. The details of the sampling approach, as well as applications to constant misorientation sampling, are described elsewhere (Singh and De Graef 2016). The uniform orientation sample is obtained by creating a uniform cubical grid, mapping all the grid points into Rodrigues space (Morawiec and Field 1996; Morawiec 2010; Roşca et al. 2014), and keeping only those points that fall inside the fundamental zone (FZ) for the relevant crystal symmetry (one of the 3D rotational point groups).

For cubic symmetry (rotational point group **432**) and *N=100* sampling points (an average angular step size of *1.4*°) along the cubochoric semi-edge, a total of N_d*=333,227* unique

sampling points are found (e.g for garnet). For the monoclinic rotational point group **2** instead, the number of unique sampling points (orientations) at the same angular step size increases to N_d*=4,000,424,* (e.g. for clinopyroxene). Similar considerations yield 2,000,037 sampling points for forsterite and enstatite.

The required input parameters for the creation of a pattern dictionary are: (1) the crystal structure; (2) the microscope acceleration voltage and the sample tilt angle; (3) the detector geometry parameters, including the detector tilt angle, the pattern center coordinates, and the distance between the illuminated region on the sample and the scintillator screen; and (4) a uniform orientation sampling grid with a step size of about *1.4°* (i.e., *N=100*). Of these input parameters only (1) is a true operator choice, the remainders are fixed as instrument parameters or extracted in a fitting routine from one EBSP of the map center (Singh et al. n.d.).

The dictionary patterns are then computed for the pattern size and geometric conditions appropriate to the experimental EBSPs. In this study, we use the dot product between patterns as the similarity metric. Because of the large number of dictionary patterns needed for lower symmetry crystal structures (i.e., orthorhombic, monoclinic, and triclinic), it was found to be more efficient to compute these dictionary patterns in real time during the indexing run instead of pre-computing them and storing them on disk.

The master patterns for the crystal structures used in this study are shown in figure 2. On the top row of figure 2, the master patterns for forsterite, garnet and enstatite are shown as stereographic projections. In all projections, the crystallographic a-axis points horizontally towards the right, and the reciprocal c*-axis is normal to the plane of the projection. In the

bottom row of figure 2, two projections are shown for clinopyroxene: the northern hemisphere projection on the left and the southern hemisphere on the right. These master patterns are sampled using bi-linear interpolation during the computation of the dictionary EBSD patterns.

Pattern matching is carried out using a dot-product approach. As an example, consider the dot-product between two 4-D vectors with components ***v*** *= (23,18,5,31)* and ***w*** *= (24,16,6,30):* the dot-product can be defined as:

$$\boldsymbol{v}.\boldsymbol{w} = |\boldsymbol{v}||\boldsymbol{w}|\cos(\theta_{\boldsymbol{vw}}),$$

where vertical bars indicate the vector norm, and $\theta_{\boldsymbol{vw}}$ is the angle between the two vectors. This relation can be rewritten by introducing normalized vectors,

$$\cos(\theta_{\boldsymbol{vw}}) = \hat{\boldsymbol{v}}.\hat{\boldsymbol{w}},$$

where the hat indicates a unit length vector. With $\hat{\boldsymbol{v}}$ = (0.5363,0.4197,0.1166,0.7223) and $\hat{\boldsymbol{w}}$ = (0.5708,0.3805,0.1427,0.7135), we obtain $\hat{\boldsymbol{v}}.\hat{\boldsymbol{w}}$ =0.9983, so that $\theta_{\boldsymbol{vw}}$ = 3.38°. Using a third vector **u** = (18,21,11,25), it is easy to show that $\theta_{\boldsymbol{vu}}$ = 13.34° and $\theta_{\boldsymbol{wu}}$ = 14.27°. From the magnitude of the angle we can see that the vectors **v** and **w** are similar, whereas the pairs **v-u** and **w-u** are dissimilar, assuming we define similarity in terms of an angular threshold of 5°. If we instead employed a threshold of 15°, all three vectors would be considered to be similar. In the DI approach, the same type of analysis is applied to the experimental and simulated EBSPs. The angular threshold of the DI routine is the same as the angular resolution of the simulations. Each EBSP, with number of pixels in x and y direction N_x by N_y, is reorganized as a column vector of N_xN_y elements containing the grayscale values of the EBSP and normalized to unit length. This reformatting is performed for all N_e experimental

patterns and N_d dictionary patterns. Representing the experimental patterns by the vectors $\widehat{\boldsymbol{v}}_j$ with $1 \leq j \leq N_e$ and the dictionary patterns by $\widehat{\boldsymbol{w}}_k$ with $1 \leq k \leq N_d$, the dictionary indexing approach then involves the computation of all dot products $\widehat{\boldsymbol{v}}_j . \widehat{\boldsymbol{w}}_k$ and, for each j, ranking of the N_d products from largest to smallest. For a given experimental pattern *j*, the *k*-value for which the dot-product is largest (i.e., the smallest angle between the two vectors) represents the dictionary pattern that is most similar to this experimental pattern *j*. Since each dictionary pattern corresponds to an orientation from the uniform sample of orientations, determination of the highest dot-product (or the smallest angle θ_{jk}) for a given *j* is equivalent to indexing the experimental pattern *j*. The fit quality (CI in OIM), in terms of the **highest dot product value** drops after the best match and decreases drastically when a mismatch of more than few degree is reached. These disorientations obtained from different dot-product fits are evaluated in figure 3. We match one experimental EBSP with all simulated EBSPs and plot them ordered according to decreasing dot-product. The best match has the minimal disorientation that is maximally the resolution of the simulated dictionary, then the second best match is plotted and the disorientation between simulated pattern one and two is the disorientation displayed on the abscissa. Furthermore the average dot product values are displayed directly as a 2D map in figure 4a, and figure 5f.

The indexing accuracy in commercial OIM systems is described using the confidence index (CI), which is defined in terms of the number of Kikuchi bands that can be reproduced by different matching solutions. In the OIM analysis software, for example, one orientation solution is derived from three intersecting Kikuchi bands. Typically more than three bands are found and all possible combinations of band triplets receive a vote. The orientation with

352 the most votes is chosen as the solution. The CI is then defined as the difference in votes
353 received by the highest and second highest-ranking solutions (*V1* and *V2* respectively)
354 divided by the number of total possible votes, V_{ideal}:

$$CI = \frac{(V_1 - V_2)}{V_{ideal}}$$

355 Low CI data are typical for very poor pattern quality, e.g. containing many scratches or grain
356 boundaries where patterns of two adjacent grains overlap. Also, a value of CI = 0 does not
357 necessarily indicate incorrect indexing (technical note from EBSD). Finally, contrary to the
358 highest dot product value in the DI approach, the CI is not an absolute measure as it depends
359 on how many bands were chosen by the operator to be included in the indexing routine.
360 In HKL, the mean angular deviation (MAD) is used to express the reliability of indexing, where
361 an average angular misfit between the detected and indexed Kikuchi bands is given. The
362 highest dot product value of the DI approach is thus more similar to MAD then to CI.
363 It should be noted that in the DI approach there is no analysis of any of the features present
364 in the EBSPs, such as the background intensity profile and the Kikuchi bands. In contrast to
365 the more commonly used Hough-space analysis, the DI routine weights all pattern pixels
366 equally and no portion of the pattern is discarded. In the bottom row of figure 4e we display
367 representative experimental electron backscatter patterns chosen arbitrarily from the set of
368 saved patterns.
369 In some cases, it is found to be useful to process the EBSD patterns to enhance contrast. Any
370 processing steps were performed on both the experimental patterns and the simulated
371 dictionary patterns to avoid the introduction of artifacts in one set that would complicate the
372 matching process. The processing carried out on all of the patterns used for this paper

consists of a single step using an adaptive histogram equalization (AHE) filter (Pizer 1987), which spreads the intensities in the pattern histogram to cover the available range of [0,255] more evenly. As a result, the pattern contrast is significantly enhanced, as apparent when comparing the patterns inserted in figure 4b. The red curve represents the frequency of specific gray values before applying the AHE filter; the blue curve shows the filtered data. The AHE filter is then followed by pattern normalization to generate vectors of unit length, which are subsequently used in the dot-product computation.

In addition to the dot-product computations between experimental and simulated patterns (referred to as "pattern indexing" hereafter), there are several other dot-product based operations that can provide valuable information about the sample as well as the quality of the indexing results. We will describe some of them in the following and compare the derived information with similar information provided by standard software solutions. Examples of each of these maps will be presented below (see "Results").

Average Dot Product (ADP) Map. Consider an experimental region of interest (ROI) with M_x by M_y sampling points and a sampling step size D; we will label a sampling point by its row and column indices, *r* and *c*, respectively, in the ROI. After application of the AHE filter and pattern normalization, one can generate an average dot-product (ADP) map by computing the average dot-product value, $\alpha_{r,c}$, for each pattern with its four nearest neighbor patterns as follows:

$$\alpha_{r,c} = \frac{1}{4}(\hat{\boldsymbol{v}}_{r,c}.\hat{\boldsymbol{v}}_{r+1,c} + \hat{\boldsymbol{v}}_{r,c}.\hat{\boldsymbol{v}}_{r-1,c} + \hat{\boldsymbol{v}}_{r,c}.\hat{\boldsymbol{v}}_{r,c+1} + \hat{\boldsymbol{v}}_{r,c}.\hat{\boldsymbol{v}}_{r,c-1}).$$

For patterns at the ROI edges and corners, the number of contributing dot products will be equal to 3 and 2, respectively. A 2-D map of the $a_{r,c}$ values then displays how similar each

sampled pixel location is to its nearest neighbors. For pixels on or near a grain or phase boundary, for instance, the average dot-product will in general be lower than for pixels inside a grain. This results from the fact that the orientation of neighboring grains is different, which leads to changes in the EBSP upon crossing the boundary (figure 4b, c, d, figure 5h, figure 6d, figure 7d). The ADP is therefore related to indexing quality and can be calculated for every pixel from the stored EBSPs. A measure similar to the average dot product in conventional Hough-based indexing is the kernel average misorientation map (KAM), where each pixel has a value equal to the average disorientation of this pixel with respect to its neighbors. The ADP-map is compared to the KAM-map in figure 5 b, e, h. This will highlight grain and phase boundaries, i.e. similar to KAM maps obtained in OIM analyses. However, grain boundaries appear black in the ADP map which highlights similarity, and white in the map KAM which highlights dissimilar orientations (figure 5 b,e,h).

Orientation Similarity Map. Since the DI approach returns a ranked list of dot products between an individual experimental pattern and all of the dictionary patterns, one can store not only the best match (highest dot-product) value, as in the highest DP-map (figure 5f), but also a series of near-matches. For each sampling point, we consider the list of *M* near-matches (*M* is typically 20 or 30), and compare this list to the corresponding lists of the four nearest neighbors of the sampling point. If we write the ordered list of near-matches for a given point (*r,c*) as a set $S_{r,c}$, then we can define the near-match similarity index, $\eta_{r,c}$, as the average value of the cardinalities (#) of the intersections with the neighboring sets:

$$\eta_{r,c} = \frac{1}{4}\left(\#(S_{r,c} \cap S_{r-1,c}) + \#(S_{r,c} \cap S_{r+1,c}) + \#(S_{r,c} \cap S_{r,c-1}) + \#(S_{r,c} \cap S_{r,c+1})\right).$$

Plotting this index as a function of the sampling point location then results in the orientation

similarity (OS) map, as displayed in figure 5g and 5i, where it compares to the CI map displayed in figure 5c. Near a grain or phase boundary, neighboring sets will be more different than in the grain interior, so that the OS map shows grain and phase boundaries too.

The Image Quality (IQ) parameter used in standard software solutions essentially refers to how easily the Kikuchi bands can be detected by the Hough-transform approach.
In the OIM software package, IQ is defined as the average height of the detected peaks in the Hough-transforms multiplied by 5, this is plotted in figure 5a. The quality of diffraction patterns is dependent on the strain in the diffraction volume of material, the phase of the material, the presence of impurities, camera parameters, vacuum, and coating conditions among other factors. Thus IQ is not an absolute measure of the quality of a pattern (adapted from the EDAX manual of OIM TSL version 7.2).
Since the DI approach does not focus on detecting individual bands, an alternative definition of the image (or pattern) quality is needed. In the present work, we chose to employ the *"pattern sharpness" Q parameter* defined by Lassen (Krieger Lassen et al. 1994). The pattern sharpness is given by:

$$Q = 1 - \frac{J}{J_{res} w_{tot}},$$

where

$$J = \sum_{h=-\frac{N}{2}}^{\frac{N}{2}} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} w(h,k)|\boldsymbol{q}|^2;$$

$$J_{res} = \frac{1}{N^2} \sum_{h=-\frac{N}{2}}^{\frac{N}{2}} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} |\boldsymbol{q}|^2;$$

$$w_{tot} = \sum_{h=-N/2}^{N/2} \sum_{k=-N/2}^{N/2} w(h,k).$$

The function *w(h,k)* is the power spectrum of the experimental EBSD pattern, and the vectors **q** are the frequency vectors with components *(h,k)*. The sharper the Kikuchi bands, the higher the high-frequency content of the power spectrum and, therefore, the closer *Q* will be to unity. The sharpness parameter makes it possible to compare EBSD patterns from different grains or phases on a uniform scale that does not depend on the number of Kikuchi bands in the pattern. An example Q map for forsterite is displayed in figure 5d; the values range from 0.40 (black) to 0.48 (white).

Note that the only operator choice using the DI method is the choice of used phases, which also includes the atom positions of the asymmetric unit and the crystal class. Using a representative EBSP the imaging parameters are refined. The latter yield the pattern center coordinates and the sample-scintillator distance. The resulting pattern orientations can then be analyzed using the conventional techniques in commercial software such as OIM or HKL; or using the freely available matlab tool box MTex (Bachmann et al. 2010).

Results

Sample 1. Figure 3 displays the average dot product for the top fifty matches for each sampling point of the partially molten residual eclogite averaged over the complete data set

as a function of the disorientation angle between the top match and the lower matches for both the garnet (red) and clinopyroxene (blue) phases. The orientation for which the highest dot product occurs is used as the orientation assigned to the experimental pattern. The next nearest matches all have lower dot products and correspond to a gradually increasing disorientation with respect to the best match pattern. For disorientations larger than about 10°, the dot product values level out and oscillate around a background value.

The histograms in Figure 4b correspond to the raw EBSD pattern (top inset and nearly Gaussian curve), and the EBSD pattern after adaptive histogram equalization (bottom inset and nearly flat curve). The equalization filter enhances the contrast in the patterns, and is applied to both experimental and dictionary patterns before computation of the dot products. Figure 4a & c illustrate the average dot product (ADP) map before and after equalization. Individual grains in Figure 4a indicate contrast differences as a result of electron channeling in different crystal orientations. The brightest regions correspond to the melt phase. In figure 4c instead, the main contrast differences are produced by the different phases. That is, the order of decreasing 'brightness' in Fig. 4c is clinopyroxene > garnet > melt. In Figure 4d an area of a garnet grain next to a melt pool is encircled. The slight variations in the gray level are caused by the amorphous melt pattern superimposing a weaker pattern of a garnet. This visualizes the third dimension captured by the interaction volume of the electron beam within the sample interior.

A raw EBSD phase map collected using the HKL system is displayed in figure 6a. In the presence of melt at the interfaces, data cleanup is not recommended. Figure 6e shows the phase map derived using the four individual highest dot product maps displayed in figure 6

f,g,h,i for clinopyroxene, garnet, melt pools and garnet-clinopyroxene phase boundaries, respectively. The joint histogram in figure 6b shows the dot product values for the best fit for clinopyroxene and garnet; there are 250,000 points in this scatter plot. The two dominant clusters correspond to high dot products obtained using the garnet dictionary and the clinopyroxene dictionary, the smaller cluster to the lower left corresponds to EBSD patterns that have a poor match against both dictionaries; those points are assigned to the melt phase. Points above the diagonal of the joint histogram have a larger best match with respect to garnet and are hence assigned to the garnet phase (red) in the phase map (figure 6e). The remaining points have a larger best match for clinopyroxe and are represented in blue in the phase map. The curves along the horizontal and vertical axes in the joint histogram represent the projections of the histogram onto the clinopyroxene and garnet axes. Note that the clinopyroxene and garnet point clouds consist of overlapping blobs that correspond to the individual grains in the microstructure. Overlap patterns obtained near grain boundaries typically correspond to points in between the three major clusters of the joint histogram. The color map in Figure 6e is obtained by merging f-i together with different color values. The polar representation of the joint histogram in figure 6c provides a better separation between the clinopyroxene and garnet phases and the melt phase, and was used to obtain the binary maps in panels 6f tHough-6i. The lines limiting the fields between the three phases were chosen at the local point density minimum, and for the clinopyroxene garnet mixed patterns attributed a width of 0.5 rad of the polar angle.

Sample 2. The EBSD raw data of sample 2 were indexed with the OIM software (upper row in figure 7), and using the DI routine (lower row in figure 7). Representative electron backscatter patterns are displayed in figure 7 in the bottom row. The results are visualized using the inverse pole figure (IPF) color scheme of OIM. The OIM data reveal more random orientation pixels mainly along grain boundaries relative to other areas of the sample. In contrast, the DI method returns grains that are all well delineated, and only a few grain boundary points are indexing randomly (figure 7a versus c).

To obtain grain boundary trace/segment information, and thus the orientation of the adjacent crystals, their disorientation and the best trace of orientation of the grain boundary plane intersecting with the sample surface, the data sets must be free of any randomly indexed pixels. This is however not a requirement if the user is only interested in simply obtaining misorientation data, as there only the Euler angles of the adjacent grains are necessary. Thus, the OIM data must undergo a cleanup procedure as described above (see "Methods"). The raw data and cleaned data are presented in figures 7a,b respectively. The spotted grains in figure 7a are indicative of the indexing ambiguities inherent to the OIM analysis. In figure 7b, grain boundary segments, highlighted in white, can subsequently be analyzed for twinning and disorientations between grains. To extract the same information using the DI method, the grain boundaries are delineated using the ADP-map, shown in figure 7d.

The noisy appearance of the phase map obtained using OIM in Figure 7a,b is a result of incorrectly indexed (or un-indexed) patterns. However, using the DI approach in Figure 7c,d, as there is always a largest dot product, no un-indexed points are observed. The need for a

post-indexing cleanup step is therefore removed. To evaluate the reliability of the DI routine, and the conventional Hough-space based indexing, we compare the extracted disorientation distributions with the expected random disorientation distribution calculated for orthorhombic symmetry (Fig. 8). The disorientation distribution obtained from the DI map agrees perfectly well with the calculated random distribution, considering the limited number of grains analyzed. In contrast, the indexing routine using the OIM routine results in a high fraction of 60°[100] pseudo-symmetry related neighboring sampling points (compare Fig. 8b and Fig. 8c).

Discussion

This study investigates the potential of the DI approach, which allows for the indexing of every EBSD pattern with an absolute value of confidence. If patterns from one phase receive low dot product values during the first DI run, dictionaries for additional phases can be calculated until all EBSD patterns can be assigned to a phase. The fact that there is always a best match means that no patterns go unindexed, not even the overlap patterns near grain boundaries for which the list of near matches contains the two different orientations on each side of the grain boundary. The potential of developing unsupervised decision trees to differentiate different pattern types, such as shifted background, as a result of a pore, noise background, overlapping pattern at grain boundaries or grain interiors has been shown in Chen et al. (2015) and is thus not repeated here. Note however, that unambiguous indexing at interfaces or in relation with pseudo-symmetry is a clear advantage of the DI approach, and this superiority becomes evident when evaluating the maps of figure 7 by displaying

their disorientation between grains and between OIM and DI maps displayed in figure 8. The DI approach uses information from every pixel of an EBSP with equal weight; this is in strong contrast to Hough-transform based methods, where only some peaks in Hough-space are used for indexing. The Hough-transform based methods therefore, in principle, disregard most of the information contained in weak Kikuchi bands or information contained in the local absence of Kikuchi-bands. Consequently we identify the approach of using all information of an EBSP equally as the source of success of the DI indexing technique.

The usage of all EBSP pixels equally is in our opinion also the cause for successful treatment of overlapping patterns, for example at grain boundaries. Recently, Wright et al. (2015) reported very powerful indexing results using neighbor pattern averaging with a re-indexing (NPAR) post processing routine, which results in large indexing improvement. Nevertheless, Wright et al. (2015) conclude: ”The two methods produce comparable results except for the presence of more mis-indexed points at the grain boundaries in the NPAR data than in the dictionary data”. We think that additionally it should be noted that NPAR is intrinsically smoothing local information over several pixels at the expense of a decreased spatial resolution, which is not the case for the DI approach. The strength of the DI approach in comparison is therefore that it retains local information, such as small in-grain misorientations, twin-lamellae or grain boundaries and interfaces with high accuracy and precision, even at high noise levels. The DI routine results in maps in which all patterns are indexed and attributed with an absolute error of indexing. Despite the fact that the resulting maps of the DI approach have no data “holes”, this does not necessarily mean that all patterns have been positively identified, for example in the case where the actual phase was

562 not included in the matching procedure. The calculation of absolute errors is an important
563 innovation as it immediately indicates the necessity for reanalysis.

564 The DI approach leads to several improvements over the currently available standard
565 software solutions: (i) As the original data are saved, patterns and results can be reviewed
566 and revised at any time; (ii) indexing will always return a highest dot product, which allows
567 for an easy assessment on indexing quality and whether or not additional phases must be
568 considered. This is because the dot product value is normalized to the interval [-1,1], so that
569 the highest dot product also represents the confidence index. (iii) Absolute errors on the data
570 interpretation are ensured as the mean least square error on every indexing result is
571 obtained. (iv) Assumptions regarding the presence of pseudo-symmetries are obsolete for
572 the systems investigated (Figure 7). (v) No data interpolations in the form of clean-up
573 procedures with intrinsic assumptions are required. While this is uniquely great for the DI
574 approach, recent developments show how pure data interpolation can be overcome by using
575 EBSD pattern post-processing in OIM, where, in case of low pattern quality, the EBSPs at a
576 given point are averaged with the patterns of the neighboring points (Wright et al. 2015). For
577 the latter is must be noted that small variations in orientation are smeared out over several
578 pixels. (vi) As the entire original EBSD patterns are routinely saved, additional information
579 can be extracted. For example gray level distributions in the electron backscatter pattern are
580 influenced by the chemical composition of the sample. At high backscattering angles the
581 intensity of the Kikuchi lines increases with increasing mean atomic number. Therefore, maps
582 similar to BSE images can be calculated, by considering only those rows of an experimental
583 EBSP sensitive to chemical composition (in most setups the upper rows of an EBSP). Similar

584 approaches have also been included in the OIM software recently, where gray values of
585 various predefined regions of an EBSP are summed and displayed as a grayscale intensity
586 map (Wright et al. 2015).
587 The drawbacks of the DI method as compared to standard software solutions are: (i) The
588 relatively long dictionary computing time, which can amount to several days for low
589 symmetry crystal structures, depending on the number of experimental patterns to be
590 indexed. This limitation can be reduced by accelerating these computations on computer
591 clusters, by means of GPUs (graphical processing units), high-performance computers, or
592 shared memory platforms. . The current implementation of the DI approach uses both the
593 GPU platform using OpenCL as well as the shared memory architecture using OpenMP to
594 speed up the computations. An additional simplification which gives equivalent results
595 involves calculating the binned pattern directly using the correct values of scaled pattern
596 centers, sometimes referred to as x-star and y-star but assuming that the ``apparent"
597 physical detector pixel size is binning*δ, where δ is the true physical detector size. This
598 typically reduces the pattern computation time by an order of magnitude for 4x or 8x binned
599 patterns. It has been shown in Ram et al. (n.d.) that the accuracy of the indexing remains
600 unaffected even for an EBSPs which are 25x25 pixels. Use of highly binned patterns together
601 with using the apparent physical detector pixel size gives us another avenue for improving
602 the indexing rate of the DI approach without compromising with the quality of the final
603 results. (ii) The need to save all patterns for subsequent indexing; while disk space is
604 relatively inexpensive, managing large numbers of pattern files can become problematic;
605 furthermore, the DI method is memory intensive, especially as the commercial indexing

606 systems are not designed to save the patterns in compressed image formats and to transfer
607 the data quickly to memory. Nevertheless, this situation is quickly improving with the
608 incorporation of new data formats (e.g. .pat and .hdf5). Final orientation data can be further
609 analyzed using the commercial software such as OIM or HKL, or using the freely available
610 matlab tool box MTex (Bachmann et al. 2010), which allow for the calculation of
611 geometrically necessary dislocations, misorientations or specific twin-boundaries. Note that
612 the concept of determining GND is still debated and Zisman (2016) suggested an alternative
613 approach which would indeed allow determining GND from our data, especially with a DI-
614 calculated for higher angular precision.
615 We point out that, while collecting EBSPs with short acquisition times and small data sizes is
616 ideal for the comparison of different indexing routines – these parameters are neither the
617 optimal acquisition parameters for Hough-based indexing nor for the DI approach.
618 Finally, we summarize the discussion analogous to Tao and Eades (2005), who discuss the
619 problems attributed to Hough-transform based EBSD processing and mapping and comment:
620 ""Saving all of the data" seems a cure for everybody and everything" at least in EBSD-based
621 research.

622 **Implications**

623 **The importance of high quality EBSD indexing for the analysis of interface characteristics of**
624 **polycrystalline and multiphase rock/material samples**: Rocks consist of crystal grains
625 separated by phase and grain boundaries, generally interfaces. Their presence impacts the
626 bulk rock properties such as diffusivity, electrical conductivity, deformation (sliding), or
627 reactivity. Furthermore, the presence and distribution of melt fractions in the interfacial

network of rocks/materials strongly impact the bulk rock physical and chemical properties (McKenzie 1989; Watson and Lupulescu 1993; Faul 1997; Schäfer and Foley 2002; Garapic et al. 2013). The wettability of specific interfaces and thus, as a function of interface energy may control melt extraction and percolation (Bagdassarov et al. 2000; Faul 2000). Here, we demonstrate that spatially high resolved EBSD data obtained by the DI routine help to minimize the individual bias of the extent of wettability by a quantitative evaluation of melt distribution. Acquired EBSD data are also suited to determine whether or not melt fractions are mobile, able to segregate and to be extracted from its host rock at upper mantle conditions, provided high-spatial resolution EBSD data are available.

Grain boundary properties and structure relations are best revealed when examining grain boundary plane orientation distributions (GBPD) and thus, internal surfaces (Saylor et al. 2003; Rohrer et al. 2004; Papillon et al. 2009; Marquardt et al. 2015). Coincident site lattice (CSL) grain boundaries and grain misorientations instead are of negligible importance in comparison (Randle 2002; Rohrer 2007). The grain boundary plane distribution is obtained by using a stereological approach based on 2-D EBSD data (Saylor et al. 2004a). The characterization of grain boundaries using EBSD data is however ultimately limited by the indexing quality and clean up procedures of the EBSD data (Rohrer et al. 2004) as well as the character and conditions of the grain boundaries (i.e. such as wet or dry). Pseudo-symmetric indexing as required for forsterite may introduce grain boundary planes into a data set, or even remove those if artificially cleaned. As quantified in this study, the pseudo-symmetric relation of 60° misorientations around the common 100 direction is caused in the case for olivine by the nearly hexagonal closed packed oxygen sub-lattice (Poirier 1975).

650 This study demonstrates that correct indexing at interfaces is crucial for interface distribution
651 studies and that indexing melt using EBSD will help to understand melt distribution in
652 partially molten rocks. Similarly, the investigation of grain boundary characteristics requires
653 high accuracy (and preferably precision) of indexing.
654 The DI approach overcomes long standing difficulties to index patterns of low quality,
655 including patterns affected by sample surface charging, low indexing rates, as well as
656 difficulties arising from pseudo-symmetric indexing. Therefore, we are convinced that future
657 EBSD-based work will greatly benefit from the DI approach. Examples for such are in
658 particular (i) the characterization of melt distribution in partially molten rocks, (ii) grain
659 boundary character distribution where minerals are affected by mis-indexing problems (e.g.
660 caused by pseudo-symmetry), and (iii) indexing of low-quality EBSD patterns (e.g. caused by
661 insufficiently polished surfaces of minerals/rocks/ materials).
662 The direct characterization of the melt distribution by the DI method would help to better
663 understand recent experimental studies on pre-melting at low-angle grain boundaries
664 (Levine et al. 2016).

665 **Acknowledgement**

666 The CMU portion of this work is supported by the Office of Naval Research on contract no.
667 N00014-12-1-0075 and the computational facilities of the Materials Characterization Facility
668 under grant no. MCF-677785. The BGI portion of this work was partially supported by the
669 German Research Foundation (DFG) with grant no MA 6278/3-1 to KM. AR thanks the EU and
670 the Laboratoire Magmas et Volcans for the award of a Marie-Curie International Incoming
671 Fellowship (302637) and an Auvergne Fellowship, respectively. We also acknowledge the

672 Earthquake Research Institute's cooperative research program, which allowed for the sample
673 synthesis at ERI in Tokyo. We a obliged to Farangis Ram for her assistance to rescan, thus re-
674 index existing EBSPs using the OIM software and Tom Nuhfer for running the electron
675 microscopy facility at CMU. Last but not least we are grateful for the detailed comments of
676 Pierre Vonlanthen and one anonymous reviewer. The comments significantly improved the
677 conciseness of the manuscript and allowed for a more facetted discussion.

678

679 **References**

680 Adams, B.L., Wright, S.I., and Kunze, K. (1993) Orientation imaging: The emergence of a new
681 microscopy. Metallurgical Transactions A, 24, 819–831.
682 Bachmann, F., Hielscher, R., and Schaeben, H. (2010) Texture Analysis with MTEX – Free and
683 Open Source Software Toolbox. Solid State Phenomena, 160, 63–68.
684 Bagdassarov, N., Laporte, D., and Thompson, A.B. (2000) Physics and Chemistry of Partially
685 Molten Rocks, 271 p. Kluwer Academic Publishers.
686 Bandli, B.R., and Gunter, M.E. (2012) Electron backscatter diffraction from unpolished
687 particulate specimens: Examples of particle identification and application to inhalable
688 mineral particulate identification. American Mineralogist, 97, 1269–1273.
689 Bunge, H., and Esling, C. (1982) Reproducible Textrues. Textrures and Microstructures, 5, 87–
690 94.
691 Callahan, P.G., and De Graef, M. (2013) Dynamical Electron Backscatter Diffraction Patterns.
692 Part I: Pattern Simulations. Microscopy and Microanalysis, 19, 1255–1265.
693 Chen, Y.H., Park, S.U., Wei, D., Newstadt, G., Jackson, M.A., Simmons, J.P., De Graef, M., and

694 Hero, A.O. (2015) A Dictionary Approach to Electron Backscatter Diffraction Indexing *.
695 Microscopy and microanalysis, 21, 739–752.
696 Cordier, P., Demouchy, S., Beausir, B., Taupin, V., Barou, F., and Fressengeas, C. (2014)
697 Disclinations provide the missing mechanism for deforming olivine-rich rocks in the
698 mantle. Nature, 507, 51–56.
699 Crawford, D.C., and Was, G.S. (1991) Grain boundary character distributions in Ni-16Cr-9Fe
700 using selected area channeling patterns: methodology and results. Journal of Electron
701 Microscopy Technique, 19, 345–60.
702 Dillon, S.J., and Rohrer, G.S. (2009) Mechanism for the development of anisotropic grain
703 boundary character distributions during normal grain growth. Acta Materialia, 57, 1–7.
704 Dingley, D.J., and Baba-Kishi, K. (1986) Use of electron backscatter diffraction patterns for
705 determination of crystal symmetry elements. Scanning Electron Microscopy II, 383–391.
706 Faul, U.H. (1997) Permeability of partially molten upper mantle rocks from experiments and
707 percolation theory. Journal of Geophysical Research, 102, 10299.
708 Faul, U.H. (2000) Chapter 3 Constraints on the Melt Distribution in Anisotropic Polycrystalline
709 Aggregates Undergoing Grain Growth. In N. Bagdassarov, D. Laporte, and A.B.
710 Thompson, Eds., Physics and Chemistry of Partially Molten Rocks Vol. Chapter 3, pp. 67–
711 92. Kluwer Academic Publishers, Dordrecht, 2000.
712 Faul, U.H., and Scott, D. (2006) Grain growth in partially molten olivine aggregates.
713 Contributions to Mineralogy and Petrology, 151, 101–111.
714 Garapic, G., Faul, U.H., and Brisson, E. (2013) High-resolution imaging of the melt distribution
715 in partially molten upper mantle rocks: evidence for wetted two- grain boundaries.

716 Geochemistry, Geophysics, Geosystems, 14, 1–11.
717 Ismaïl, W. Ben, and Mainprice, D. (1998) An olivine fabric database: An overview of upper
718 mantle fabrics and seismic anisotropy. Tectonophysics, 296, 145–157.
719 Kelly, M.N., Glowinski, K., Nuhfer, N.T., and Rohrer, G.S. (2016) The five parameter grain
720 boundary character distribution of α-Ti determined from three-dimensional orientation
721 data. Acta Materialia, 111, 22–30.
722 Kikuchi, S. (1928) Diffraction of cathode rays by mica. Proceedings of the Imperial Academy,
723 4, 271–274.
724 Koizumi, S., Hiraga, T., Tachibana, C., Tasaka, M., Miyazaki, T., Kobayashi, T., Takamasa, A.,
725 Ohashi, N., and Sano, S. (2010) Synthesis of highly dense and fine-grained aggregates of
726 mantle composites by vacuum sintering of nano-sized mineral powders. Physics and
727 Chemistry of Minerals, 37, 505–518.
728 Krieger Lassen, N.C., Juul Jensen, D., and Conradsen, K. (1994) On the statistical analysis of
729 orientation data. Acta Crystallographica Section A, 50, 741–748.
730 Le Roux, V., Tommasi, A., and Vauchez, A. (2008) Feedback between melt percolation and
731 deformation in an exhumed lithosphere-asthenosphere boundary. Earth and Planetary
732 Science Letters, 274, 401–413.
733 Levine, J.S.F., Mosher, S., and Rahl, J.M. (2016) The role of subgrain boundaries in partial
734 melting. Journal of Structural Geology, 89, 181–196.
735 Lloyd, G.E., Farmer, A.B., and Mainprice, D. (1997) Misorientation analysis and the formation
736 and orientation of subgrain and grain boundaries. Tectonophysics, 279, 55–78.
737 Maitland, T., and Sitzman, S. (2007) Electron Backscatter Diffraction (EBSD) Technique and

738 Materials Characterization Examples. In Scanning Microscopy for Nanotechnology pp.
739 41–75.
740 Manthilake, M.A.G.M., Miyajima, N., Heidelbach, F., Soustelle, V., and Frost, D.J. (2013) The
741 effect of aluminum and water on the development of deformation fabrics of
742 orthopyroxene. Contributions to Mineralogy and Petrology, 165, 495–505.
743 Marquardt, K., Rohrer, G.S., Morales, L., Rybacki, E., Marquardt, H., and Lin, B. (2015) The
744 most frequent interfaces in olivine aggregates: the GBCD and its importance for grain
745 boundary related processes. Contributions to Mineralogy and Petrology, 170, 40.
746 McKenzie, D. (1989) Some remarks on the movement of small melt fractions in the mantle.
747 Earth and Planetary Science Letters, 95, 53–72.
748 Michibayashi, K., and Mainprice, D. (2004) The Role of Pre-existing Mechanical Anisotropy on
749 Shear Zone Development within Oceanic Mantle Lithosphere: an Example from the
750 Oman Ophiolite. J. Petrology, 45, 405–414.
751 Michibayashi, K., Mainprice, D., Fujii, A., Uehara, S., Shinkai, Y., Kondo, Y., Ohara, Y., Ishii, T.,
752 Fryer, P., Bloomer, S.H., and others (2016) Natural olivine crystal-fabrics in the western
753 Pacific convergence region: A new method to identify fabric type. Earth and Planetary
754 Science Letters, 443.
755 Morales, L.F.G., Boudier, F., and Nicolas, A. (2011) Microstructures and crystallographic
756 preferred orientation of anorthosites from Oman ophiolite and the dynamics of melt
757 lenses. Tectonics, 30.
758 Morawiec, A. (2010) Orientations and Rotations Computations in Crystallographic Textures,
759 200 p. Springer Berlin/ Heidelberg.

760 Morawiec, A., and Field, D.P. (1996) Rodrigues parameterization for orientation and
761 misorientation distributions. Philosophical Magazine A, 73, 1113–1130.
762 Nolze, G., and Winkelmann, A. (2016) Crystallometric and projective properties of
763 backscattered Kikuchi diffraction patterns. Journal of Applied Crystallography.
764 Papillon, F., Rohrer, G.S., and Wynblatt, P. (2009) Effect of Segregating Impurities on the
765 Grain-Boundary Character Distribution of Magnesium Oxide. Journal of the American
766 Ceramic Society, 92, 3044–3051.
767 Park, S.U., Wei, D., Graef, M. De, Shah, M., Simmons, J., and Hero, A.O.. (2013) EBSD IMAGE
768 SEGMENTATION USING A PHYSICS-BASED FORWARD MODEL. In Image Processing (ICIP),
769 2013 20th IEEE International Conference on pp. 3780–3784.
770 Pinard, P.T., Lagacé, M., Hovington, P., and Thibault, D. (2011) An Open-Source Engine for the
771 Processing of Electron Backscatter Patterns : EBSD-Image. Micorscopy and
772 Microanalyses, 1–12.
773 Pizer, S. (1987) Adaptive Histogram Equalization and Its Variations.pdf. Computer Vision,
774 Graphics, and image processing.
775 Poirier, J.P. (1975) On the slip systems of olivine. Journal of Geophysical Research, 80, 4059–
776 4061.
777 Prior, D.J., Boyle, A.P., Brenker, F., Cheadle, M.C., Austin, D., Lopez, G., Peruzzo, L., Potts, G.J.,
778 Reddy, S., Spiess, R., and others (1999) The application of electron backscatter
779 diffraction and orientation contrast imaging in the SEM to textural problems in rocks.
780 American Mineralogist, 84, 1741–1759.
781 Prior, D.J., Rutter, E.H., and Tatham, D.J. (2011) Deformation Mechanisms, Rheology and

Tectonics: Microstructures, Mechanics and Anisotropy, 354 p.

Ram, F., Li, Z., Zaefferer, S., Hafez Haghighat, S.M., Zhu, Z., Raabe, D., and Reed, R.C. (2016) On the origin of creep dislocations in a Ni-base, single-crystal superalloy: an ECCI, EBSD, and dislocation dynamics-based study. Acta Materialia, 109, 151–161.

Ram, F., Singh, S., and De Graef, M. (n.d.) Error Analysis of Crystal Orientations Obtained by the Dictionary Approach to EBSD Indexing. Ultramicroscopy.

Randle, V. (2002) The coincidence site lattice and the " sigma enigma ." Materials Characterization, 47, 411–416.

Rohrer, G.S. (2007) The distribution of grain boundary planes in polycrystals. JOM: The Journal of The Minerals, Metals & Materials Society (TMS), 59, 38–42.

Rohrer, G.S. (2011) Measuring and Interpreting the Structure of Grain-Boundary Networks. Journal of the American Ceramic Society, 94, 633–646.

Rohrer, G.S., Saylor, D.M., El Dasher, B., Adams, B.L., Rollett, A.D., and Wynblatt, P. (2004) The distribution of internal interfaces in polycrystals. Zeitschrift für Metallkunde, 95, 197–214.

Roşca, D., and De Graef, M. (2013) Area-preserving projections from hexagonal and triangular domains to the sphere and applications to electron back-scatter diffraction pattern simulations. Modelling and Simulation in Materials Science and Engineering, 21, 55021.

Roşca, D., Morawiec, a, and De Graef, M. (2014) A new method of constructing a grid in the space of 3D rotations and its applications to texture analysis. Modelling and Simulation in Materials Science and Engineering, 22, 75013.

804 Rosenthal, A., Yaxley, G.M., Green, D.H., Hermann, J., Kovács, I., and Spandler, C. (2014)
805 Continuous eclogite melting and variable refertilisation in upwelling heterogeneous
806 mantle. Scientific reports, 4, 6099.
807 Saylor, D.M., Morawiec, A., and Rohrer, G.S. (2003) The relative free energies of grain
808 boundaries in magnesia as a function of five macroscopic parameters. Acta Materialia,
809 51, 3675–3686.
810 Saylor, D.M., El Dasher, B.S., Rollett, A.D., and Rohrer, G.S. (2004a) Distribution of grain
811 boundaries in aluminum as a function of five macroscopic parameters. Acta Materialia,
812 52, 3649–3655.
813 Saylor, D.M., El Dasher, B., Pang, Y., Miller, H.M., Wynblatt, P., Rollett, A.D., and Rohrer, G.S.
814 (2004b) Habits of Grains in Dense Polycrystalline Solids. Journal of the Physical Society
815 of Japan, 26, 724–726.
816 Schäfer, F.N., and Foley, S.F. (2002) The effect of crystal orientation on the wetting behaviour
817 of silicate melts on the surfaces of spinel peridotite minerals. Contributions to
818 Mineralogy and Petrology, 143, 254–261.
819 Schwartz, A.J., Kumar, M., Adams, B.L., Field, D.P. (2009) Electron Backscatter Diffraction in
820 Materials Science, 2nd ed. (D.P. Schwartz, A.J., Kumar, M., Adams, B.L. & Field, Ed.).
821 Singh, S., and De Graef, M. (2016) Orientation Sampling for Dictionary-Based Diffraction
822 Pattern Indexing Methods, 1–22.
823 Singh, S., Ram, F., and De Graef, M. (n.d.) Accurate Determination of EBSD Detector
824 Parameters and Sample Misalignment Using Dynamically Simulated Patterns.
825 Microscopy & Microanalysis.

826 Soustelle, V., Walte, N.P., Manthilake, M. a. G.M., and Frost, D.J. (2014) Melt migration and
827 melt-rock reactions in the deforming Earth's upper mantle: Experiments at high
828 pressure and temperature. Geology, 42, 83–86.
829 Sun, X.-Y., Cordier, P., Taupin, V., Fressengeas, C., and Jahn, S. (2016) Continuous description
830 of a grain boundary in forsterite from atomic scale simulations: the role of disclinations.
831 Philosophical Magazine, 96, 1757–1772.
832 Tao, X., and Eades, A. (2005) Errors, artifacts, and improvements in EBSD processing and
833 mapping. Microscopy and Microanalysis, 11, 79–87.
834 Venables, J.A., and Harland, C.J. (1973) Electron back-scattering patterns—A new technique
835 for obtaining crystallographic information in the scanning electron microscope.
836 Philosophical Magazine, 27, 1193–1200.
837 Warren, J.M., and Hirth, G. (2006) Grain size sensitive deformation mechanisms in naturally
838 deformed peridotites. Earth and Planetary Science Letters, 248, 423–435.
839 Watson, E.B., and Lupulescu, A. (1993) Aqueous fluid connectivity and chemical transport in
840 clinopyroxene-rich rocks. Earth and Planetary Science Letters, 117, 279–294.
841 Wheeler, J., Prior, D.J., Jiang, Z., Spiess, R., and Trimby, P.W. (2001) The petrological
842 significance of misorientations between grains. Contributions to Mineralogy and
843 Petrology, 141, 109–124.
844 Wilkinson, A.J., Meaden, G., and Dingley, D.J. (2006) High-resolution elastic strain
845 measurement from electron backscatter diffraction patterns: New levels of sensitivity.
846 Ultramicroscopy, 106, 307–313.
847 Wright, S.I., and Nowell, M.M. (2006) EBSD Image Quality Mapping. Microscopy and

Microanalysis, 12, 72–84.

Wright, S.I., Nowell, M.M., and Bingert, J.F. (2007) A Comparison of Textures Measured Using X-Ray and Electron Backscatter Diffraction. Metallurgical and Materials Transactions A, 38, 1845–1855.

Wright, S.I., Nowell, M.M., Lindeman, S.P., Camus, P.P., De Graef, M., and Jackson, M.A. (2015) Introduction and comparison of new EBSD post-processing methodologies. Ultramicroscopy, 159, 81–94.

Zisman, A. (2016) Choice of scalar measure for crystal curvature to image dislocation substructure in terms of discrete orientation data. Journal of the Mechanical Behavior of Materials, 25, 15–22.

Figure captions

Figure 1. Schematic of the dictionary indexing process. A Generalized Forward Projector (GFP) model for EBSD (light blue box) is combined with a realistic detector and noise model (yellow box) and a uniform sampling of orientation space (orange) to generate a pattern dictionary (dark blue box), which consists of parts of the master pattern depicted in figure 2. The experimental patterns (pink box) are then compared with the dictionary patterns using a pattern matching engine (green box), in this case a simple dot product comparison. The result is a set of indexed EBSD patterns (bottom green-filled box) which can then be analyzed similar to standard EBSD data, e.g. using commercial software or MTex.

Figure 2. EBSD master patterns for (top row) forsterite, garnet, and enstatite, and (bottom row) clinopyroxene; all patterns are represented as stereographic projections, with the crystallographic a-axis oriented horizontally towards the right and the reciprocal c*-axis normal to the projection plane. The projections for clinopyroxene show both the northern (left) and southern (right) hemispheres; for the other structures, the two hemispheres have identical projections.

Figure 3. Graph showing the dot product values obtained by multiplying an experimental EBSP with the dictionary pattern versus the disorientation between the experimental EBSP and the simulated EBSP. The indexing routine in the dictionary approach is based on a ranking of similarity between experimental EBSP and simulated EBSP using computed dot

products $\hat{\boldsymbol{v}}_j . \hat{\boldsymbol{w}}_k$, where the experimental patterns are given by the vectors $\hat{\boldsymbol{v}}_j$ with $1 \leq j \leq N_e$ and the dictionary patterns by $\hat{\boldsymbol{w}}_k$ with $1 \leq k \leq N_d$. The largest dot product value represents the highest similarity between experiment and simulation; plotted is the ranking from large to small of each j's top 15 products or more. In green this is displayed for low quality EBSP of forsterite; in blue for high quality EBSP of clinopyroxene and in red for garnet. The assumption for the disorientation scale is that the dictionary finds the correct misorientation.

Figure 4. Illustration of image processing prior to nearest neighbor similarity map calculation. (a) Average dot-product (ADP) map for the raw background subtracted patterns, and b) shows a histogram (red) for a background-subtracted EBSP outlined in red, along with the EBSP and histogram after adaptive histogram equalization in blue and outlined in blue. (d) ADP map after applying adaptive histogram equalization before calculating the average dot-products. (d) Magnified region from (c), with an interesting region where backscatter information of crystal and melt overlap. The lowest row, e) displays exemplary raw EBSPs.

Figure 5. Comparison of common measures of OIM and the DI-approach. In the first row some measures of OIM are displayed. Pattern quality, judged by how easy Kikuchi bands can be detected, a) OIM IQ-map; disorientation between indexed pattern b) OIM KAM-map; and the confidence index, c) OIM CI-map. In the second row comparable measures used in the DI-approach. d) EBSP image quality expressed through the Q parameter, e) inverse average dot product ADP-map, f) highest dot product, g) KAM map obtained by using the average

orientation for the 20 highest dot products (20HDP), with a misorientation maximum of 10°. h) ADP-map. i) orientation similarity map, OS-map.

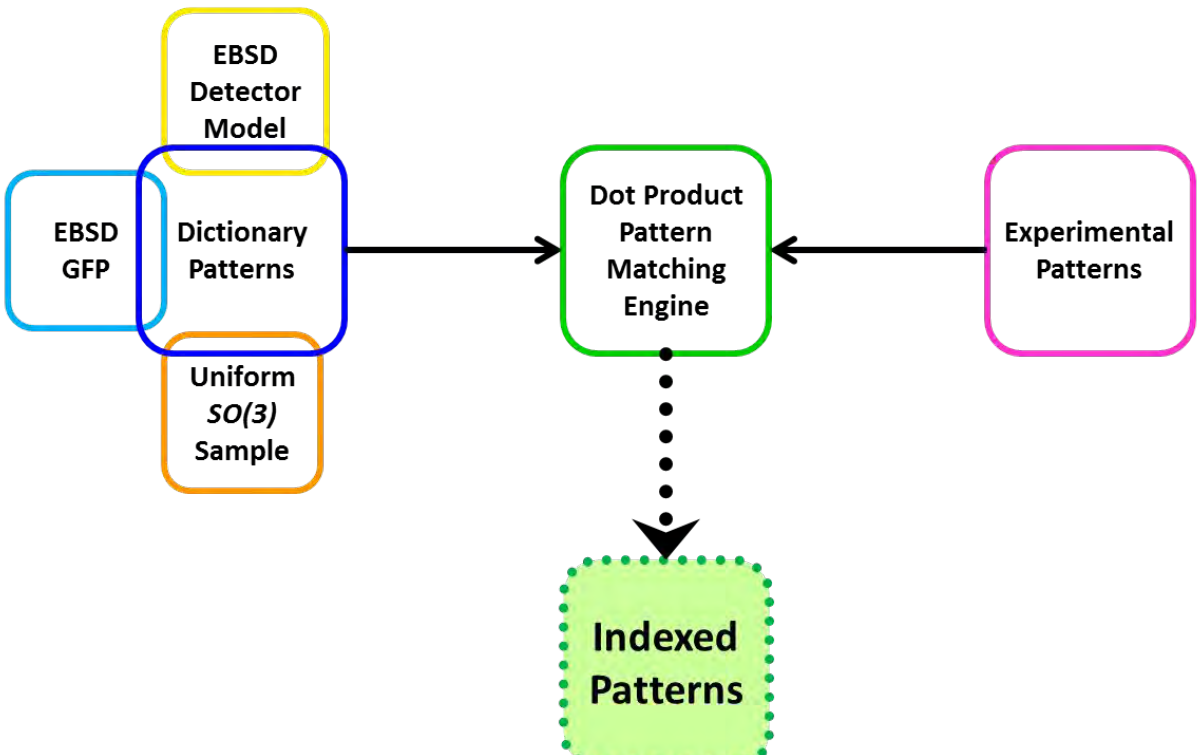
Figure 6. Phase identification using different indexing routines. a) indexing using Hough-transform based indexing in HKL, the map was acquired next to the map displayed in d,e,f,g,h,I, using the same settings as for the later. 44% of pixels stay without a solution. Post processing is largely prevented by the presence of the melt phase. b) Indexing by means of DI. The dot product values of garnet and clinopyroxene are plotted versus each other. c) Plot of the polar distance versus the polar angle determined using the DI approach. The three phases can be clearly separated. d) average dot product map obtained from DI indexing. Grain, phase and twin boundaries are nicely delineated. e) Phase identification in a 2D map representation using the differentiation obtained from b and c respectively. Blue: cpx; Red: Garnet, Yellow: Amorphous (melt) phase and white: wetted Grt-Cpx phase boundaries. f-i) Dictionary indexing confidence index map (DI CI). Light colors correspond to high dot-products, dark are low (i.e., poor or no pattern matches); the dot product for experimental images vs. the cpx dictionary is displayed in f. g) Garnet DI CI, h) Melt pools DI CI, i) Grt-Cpx-phase boundaries.

Figure 7. Comparison of OIM indexing (upper row) and data fitting using the dictionary approach (lower row). In both examples the EBSP were indexed with forsterite being the only phase. Representative EBSPs are depicted in the third row. a) OIM indexing, raw data, on the left an enlargement of the data is shown. b) Cleaned OIM data set, highlighting the grain boundaries in white. c) Same data set as in a, but indexed using the dictionary approach,

927 note the absence of randomly -indexed pixels in the data set and the consistency of indexing
928 across grain boundaries and in the crystal interior; examples are encircled in the enlargement
929 of a and c. d) Grain boundaries highlighted (dark) using the dictionary data. Displayed is the
930 average dot product map (ADP) that shows average dot product value for each pixel the of
931 the EBSD pattern with its four nearest neighbors. In the third row (e) the color legend and
932 typical EBSD patterns are displayed. The blow-ups of the panels a,b,c, and d are on the left
933 and right hand side of the respective pannels.
934
935 **Figure 8:** Disorientation distribution plots resulting from the different analyses approaches,
936 a) OIM raw indexing (red) and OIM after pseudo-symmetry correction (blue); b) dictionary
937 approach raw data (green), no second graph is displayed as no corrections apply to the
938 dictionary; c) discrepancy in indexing between DI and OIM displayed as disorientation
939 between the respectively indexed pixels. The disorientation between two pixels indexed
940 using the DI or Hough-transform based methods is minimal if the same or similar results are
941 obtained. Different indexing results of same pixels are related to pseudo-symmetric relations
942 as well as other issues such as the assignment of grain boundary pixels to one or the other
943 grain.
944

This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2017-6062

Figure 1



This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2017-6062

Figure 2

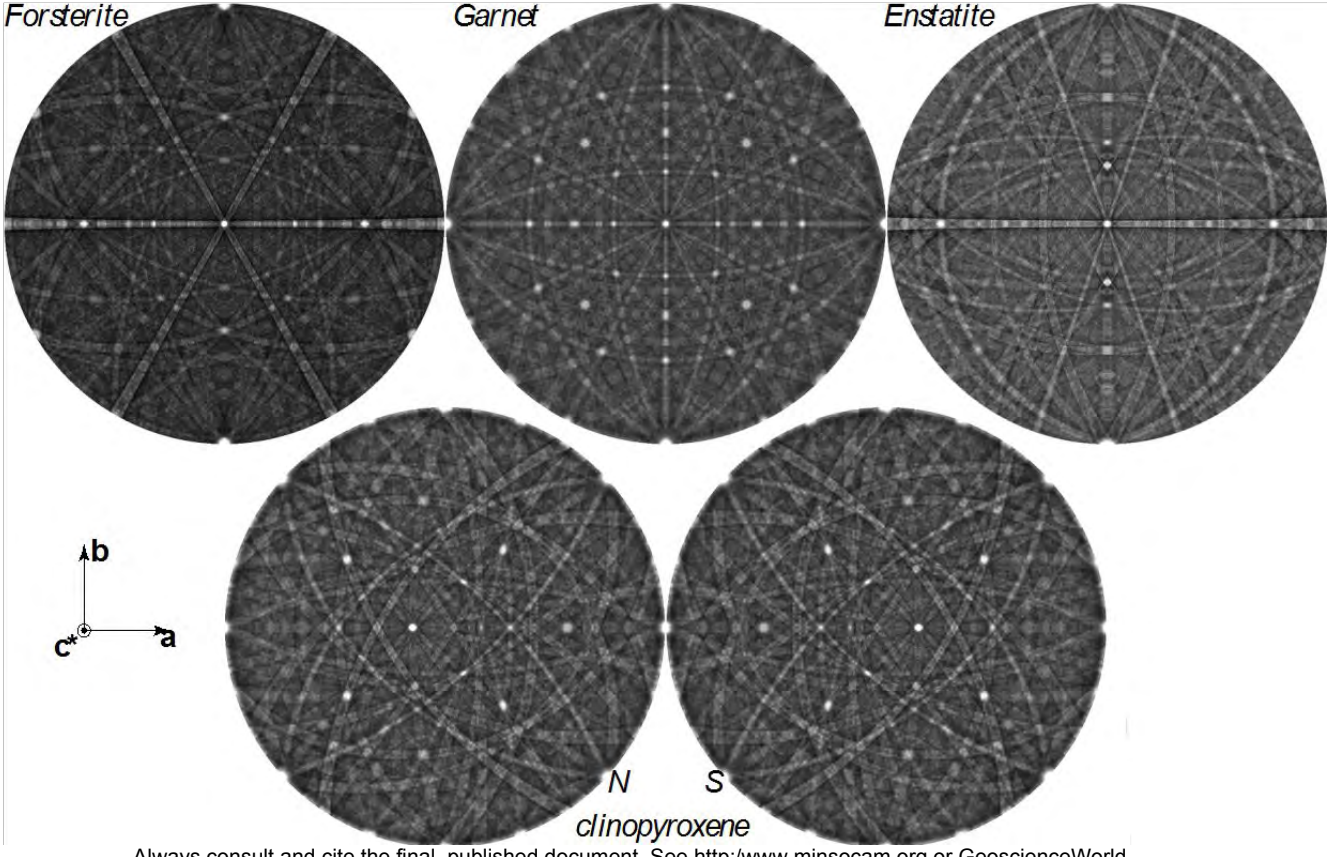
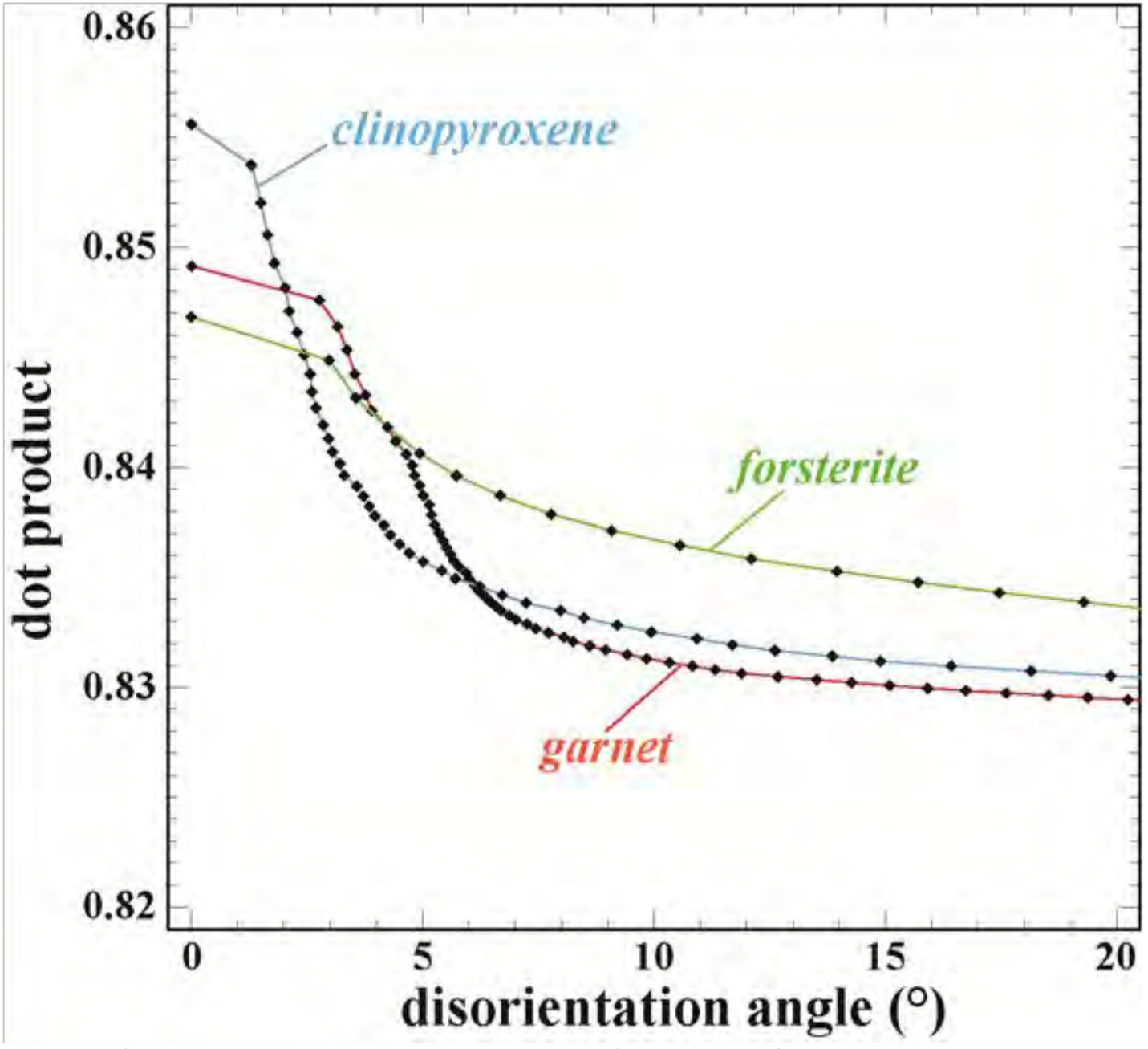
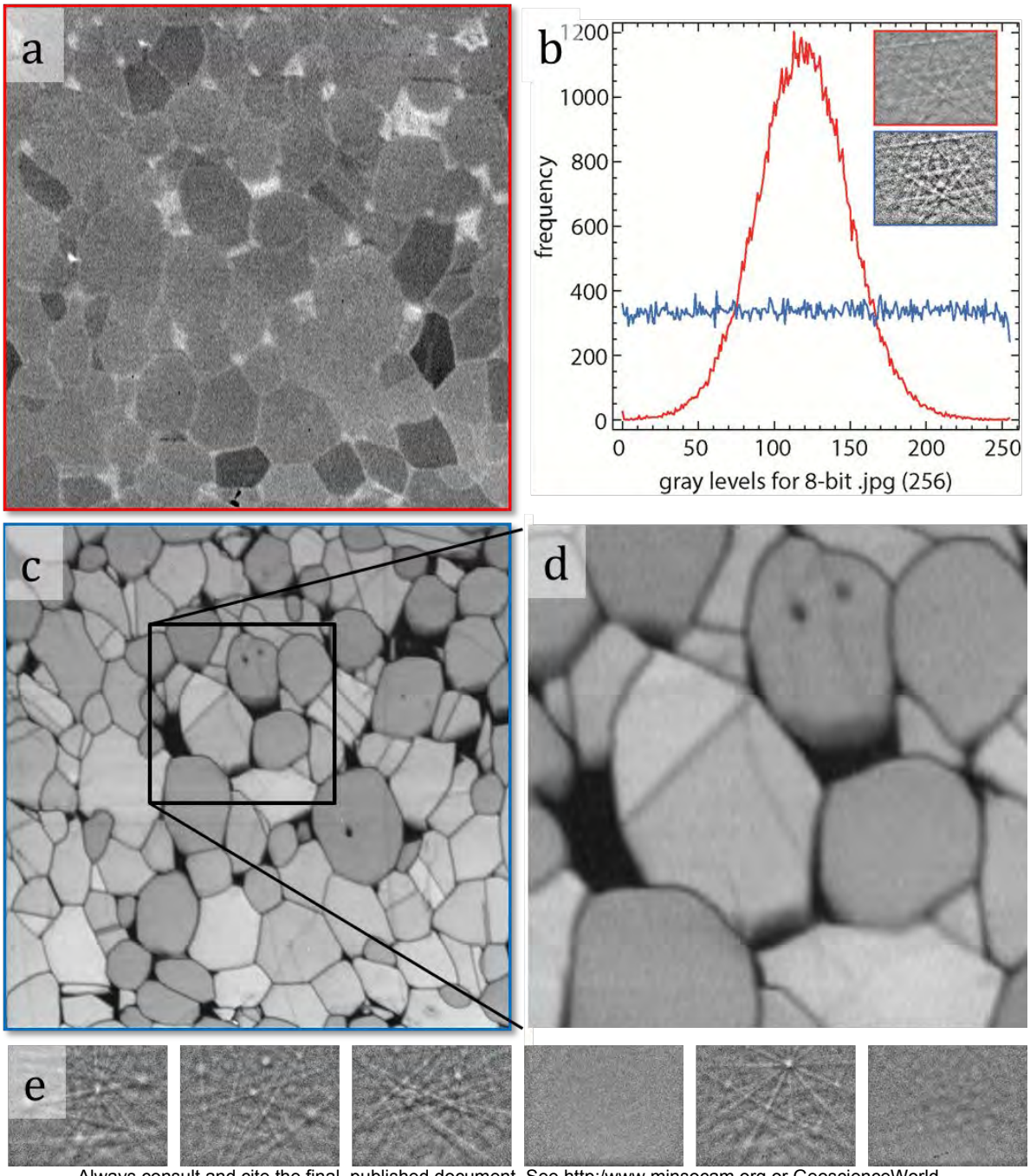


Figure 3



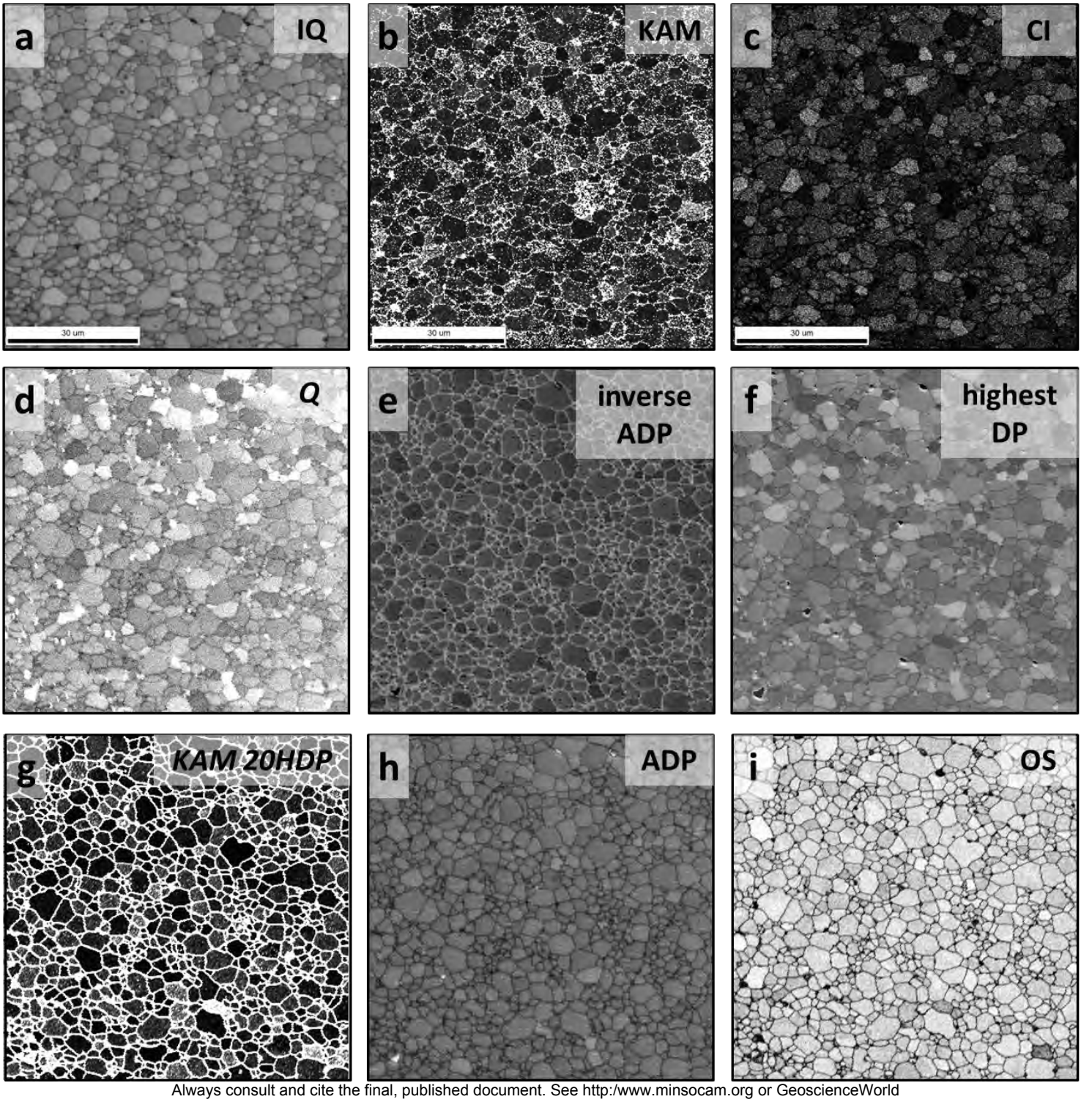
This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2017-6062

Figure 4



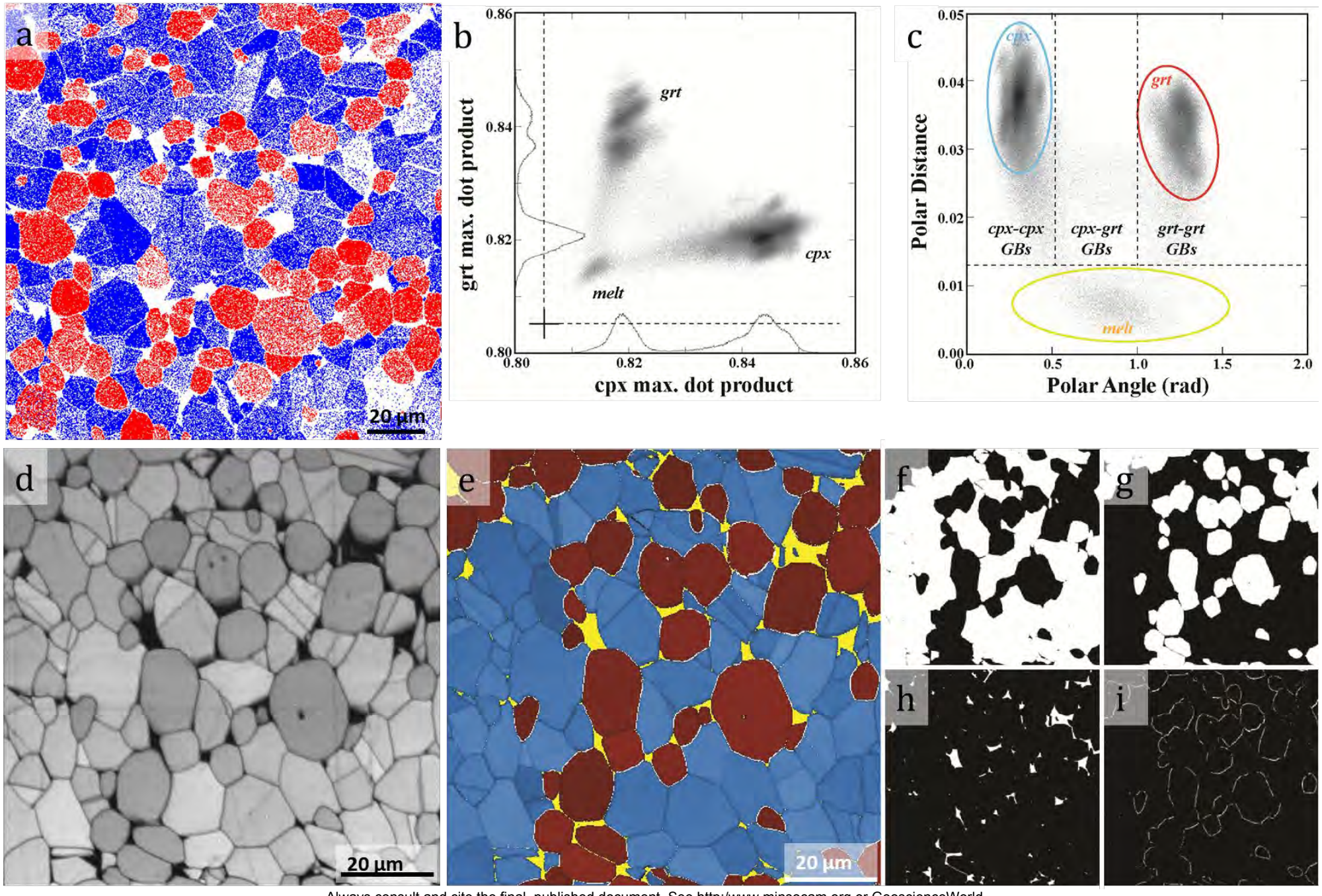
This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2017-6062

Figure 5



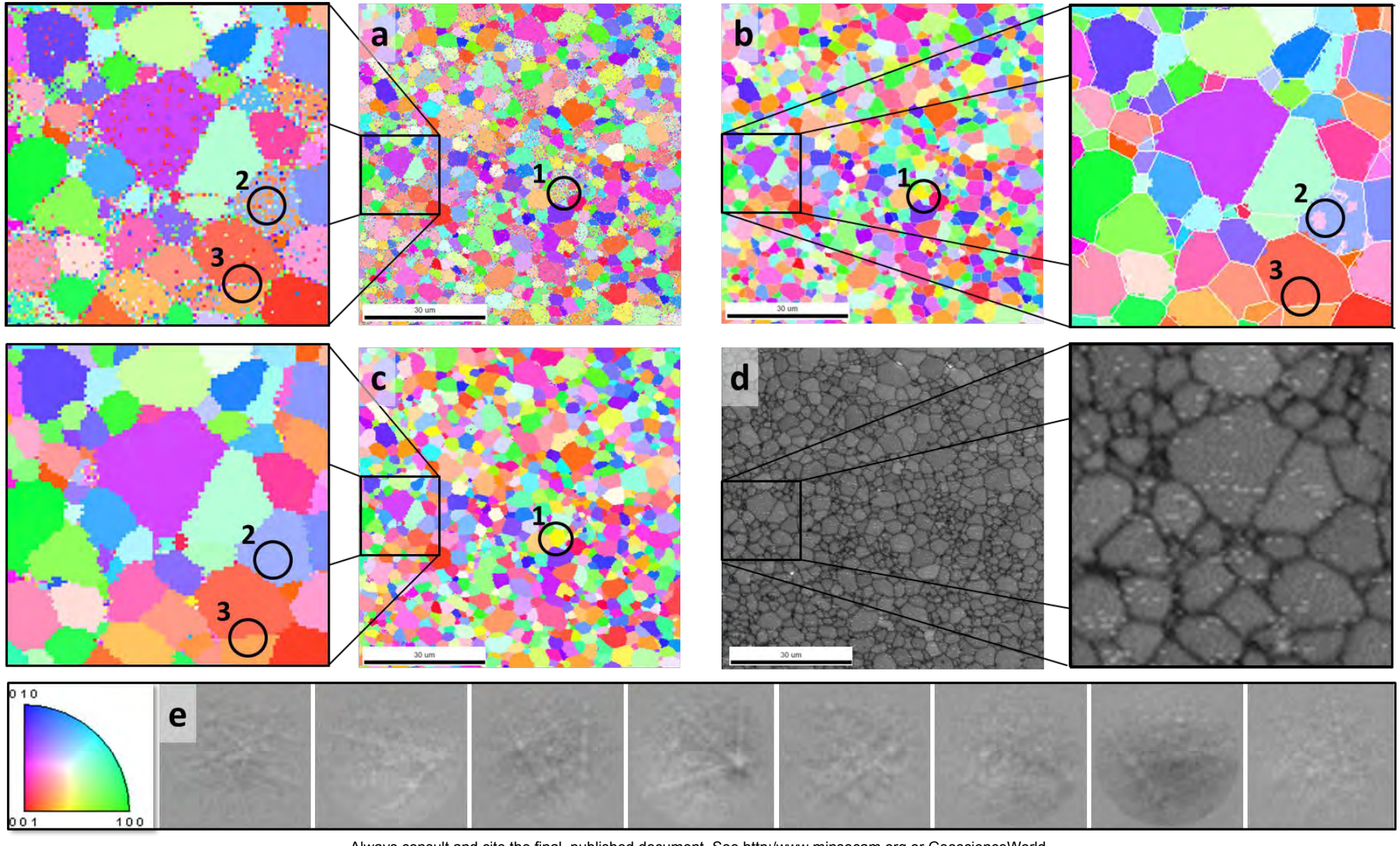
This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2017-6062

Figure 6



This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2017-6062

Figure 7



This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
(DOI will not work until issue is live.) DOI: http://dx.doi.org/10.2138/am-2017-6062

Figure 8