Revision 2

Unraveling the presence of multiple plagioclase populations and identification of representative two dimensional sections using a statistical and numerical approach

Lilu Cheng[1]; Fidel Costa[1,2]; Roberto Carniel [3]

1, Earth Observatory of Singapore, Nanyang Technological University, 50 Nanyang Av, 639798, Singapore
2, Asian School of the Environment, Nanyang Technological University, 50 Nanyang Av, 639798, Singapore
3, DPIA, Università di Udine, Via delle Scienze 206, 33100 Udine, Friuli, Italy

Abstract:

Many plagioclase phenocrysts from volcanic and plutonic rocks display quite complex chemical and textural zoning patterns. Understanding the zoning patterns and variety of crystal populations holds clues to the processes and time scales that lead to the formation of the igneous rocks. However, in addition to a 'true' natural complexity of the crystal population, the large variety of plagioclase types can be partly artifacts of the use of two-dimensional (2D) petrographic thin sections and random cuts of three-dimensional (3D) plagioclase crystals. Thus, the identification of the true number of plagioclase populations, and the decision of which are 'representative' crystal sections to be used for detailed trace element and isotope analysis is not obvious and tends to be subjective.

Here we approach this problem with a series of numerical simulations and statistical analyses of a variety of plagioclase crystals zoned in 3D. We analyze the effect of increasing complexity of zoning based on 2D chemical maps (e.g. backscattered electron images, BSE).

We first analyze the random sections of single crystals, and then study the effect of mixing of different crystal populations in the samples. By quantifying the similarity of the compositional histogram of about one hundred 2D plagioclase sections it is possible to identify the so-called reference and ideal sections that are representative of the real 3D crystal populations. These section types allow filtering out the random-cut effects and explain more than 90% of the plagioclase compositional data of a given sample. Our method allows the identification of the main crystal populations and representative crystals that can then be used for a more robust interpretation of magmatic processes and time scales.

Keyword: Crystal zoning; Plagioclase; pattern recognition; Modeling; Random cuts;

Introduction

The compositional and textural features of plagioclase from igneous rocks have been investigated for a long time, first using its optical properties and the petrographic microscope (e.g., Hibbard, 1995; Shelley, 1993), and more recently using a combined approach of scanning electron microscope (e.g., Ginibre et al., 2002), electron and ion microprobes (e.g., Blundy and Shimizu, 1991; Singer et al., 1995), rim-to-core analysis (e.g. Bouvet De Maisonneuve et al., 2012; Neill et al., 2015), cathodoluminescence studies (e.g. Higgins et al., 2015), and in situ microdrilling for isotopes (Davidson et al., 2001). Feldspar studies have provided critical clues about the magmatic processes (e.g. magma mixing, assimilation, fractionation, magma ascent, crystal recycling; e.g. Anderson, 1984; Feeley and Dungan, 1996; Landi et al., 2004; Streck, 2008) and their associated time scales (e.g., Costa et al., 2003; Druitt et al., 2012; Stelten et al., 2015; Zellmer et al., 1999). However, a common observation of all of these studies is the extreme variety and the chemical and textural complexity of the plagioclase crystals, including those of open degassing volcanoes such as Stromboli, Etna, Llaima, or Mayon (Bouvet De Maisonneuve et al., 2012; Landi et al., 2004; Nicotra and Viccaro, 2012), e.g. the example from Mayon shown in Fig.1 that we will discuss in detail below. The variety of plagioclase phenocrysts makes it very difficult to establish how many crystal populations are present in the deposit, and to objectively decide whether there are any phenocryst section that can be considered as 'representative' of a population and thus used to derive conclusions about the processes from electron or ion microprobe data (e.g., Singer et al., 1995).

The variety and complexity of plagioclase textures and zoning can reflect the large number of processes and variables that affect its stability (e.g. Yoder et al., 1957), but may also partly be the result that we typically study them using 2D petrographic thin sections that are derived from a set of randomly cut 3D crystals. The 3D to 2D conversion of the crystals can significantly affect their shape (e.g. Higgins, 1994) but also produce an artificial variety of zoning patterns and increase the apparent number of crystal populations (Pearce, 1984; Wallace and Bergantz, 2004). Wallace and Bergantz (2002) proposed a methodology to correlate between different crystals using 1 dimensional (1D) traverses by doing a wavelet analysis of the anorthite (An = 100 x Ca/[Ca+Na]) content. They designed a new crystal phylogeny analysis that showed how different crystals in the same deposits could be related, and shared a larger portion of their history going from core to rim. Later Wallace and Bergantz (2005) recognized the effects of random cuts and proposed a methodology to correct for them, although also acknowledged that loss of information by random cuts missing the inner parts of the crystals is unavoidable.

In this paper, we take a complementary approach by analyzing the 2D compositional features of plagioclase based on BSE images (e.g. An histograms of 2D plagioclase sections; e.g. Cashman and Blundy, 2013). Given the ease with which the BSE images of a large number of crystals can be gathered, it is possible to obtain a large statistical dataset that can be used to characterize the deposits and the crystals. Numerical simulation of 2D crystal sections in thin section may allow understanding the 3D crystals and magma processes in a similar manner to the quantification of crystal size distribution studies (Higgins, 1994; Morgan and Jerram, 2006), olivine and pyroxene zoning patterns and time scales (Pearce, 1984; Shea et al., 2015). However, such analysis for complexly zoned crystals like plagioclase is basically not available.

We first present numerical simulations of 3D plagioclase zoned crystals, which we use to quantify the effect of the 3D-to-2D conversion using compositional maps. We show how it is possible to define some special sections that are fully representative of the 3D crystals by quantifying the similarity between the different 2D sections. We apply these numerical models to a variety of zoning patterns and mixed crystal populations and demonstrate how our approach can retrieve the original true populations and explain > 90% of the compositional data for the samples.

Model set up and strategy for characterizing and comparing compositional zoning

When studying natural samples we do not know a priori how many crystal populations there are in a given sample, and it is difficult to recognize the artifacts of random sampling, in particular if crystals are geometrically and compositionally complex. Our approach is to first construct numerical crystals and perform forward models of their zoning patterns to find measurable variables and statistical procedures to identify the 2D sections that can be confidently taken as representative of the 3D crystal populations. We first study various relationships between populations using simple crystals, which we then make progressively more complex. Later we use these finding to study mixed crystal populations, first in the numerical experiments and then in natural samples.

Strategy: 1D vs. 2D data

The methods proposed in the literature to group crystals, such as wavelet-based correlations (Wallace and Bergantz, 2002) and shared characteristic diagrams (Wallace and Bergantz, 2005) are based on comparison of 1D profiles. However, given the complexity seen in 2D BSE crystal images, it is not straightforward to decide which 1D profile is representative of a given 2D section. Moreover, identification of crystal populations becomes increasingly difficult with increasing degrees of geometric complexity. Thus, we have used another way to classify plagioclase population based on the area frequency compositional distribution of 2D plagioclase sections (Cashman and Blundy, 2013). The advantage is that we have a better overall characterization of the compositional data of the 2D section, but it has the disadvantage that we lose the spatial information, in particular the core-rim relationship. This hampers a detailed understanding of the processes that created the zoning patterns, but this is not the goal of our contribution. We aim at first identifying the crystal populations, and later we can turn to detailed core to rim traverses only for representative sections. Otherwise, given the textural and compositional variability of many natural plagioclase crystals it is impossible to unambiguously choose the crystal sections that are representative to do detailed 1D compositional traverses. A crystal population is made of crystals that have experienced very similar magmatic processes, although in the context of our analysis a crystal population is defined by those crystals that have them same 2D An (or greyscale) distribution (within error; see below). Moreover, the fact that the 2D sections are the result of random cuts already provides a certain degree of information about the core to rim zoning, because for example, the cores can be undersampled, but the rims are not. A large number of compositional information from 2D sections of crystals can be obtained from BSE images calibrated for their gray scale values using a few quantitative analyses and the electron microprobe (Ginibre et al., 2002). It is nowadays possible to produce BSE images of a full thin section using SEM under the same analytical conditions so that all plagioclase crystals can be compared (e.g. Fig. 1).

Numerical 3D simulation of plagioclase crystals and generation of random 2D sections

We first constructed a numerical model of a 3D plagioclase crystal with the geometry and number of faces according to theory (e.g. Deer et al., 1992; Higgins, 1991; Higgins, 2006, Fig. 2) and with five compositional zones using Matlab 2014b software environment (Mathworks, 2014). Note that although plagioclase belongs to the triclinic system (Deer et al., 1992) we used a reference frame of three perpendicular axes that are parallel to the zoning patterns for our numerical simulations. We also performed a numerical simulation of a crystal with an angle of 115° between X and Y axes (close to the theoretical value for anorthite according to Deer et al, 1992) and we later show that this does not affect our results. We varied the ratios between the three dimensions of the crystal (S= shortest dimension, I=intermediate dimension, and L=longest dimension; (Deer et al., 1992; Higgins, 1991; Higgins, 2006) (Fig.2). Previous studies showed that the S:I:L of most plagioclase phenocrysts may range from 1:2.8:4, 1:5:8 to 1:5:5 (e.g. Cheng et al., 2014; Higgins, 1994; Morgan and Jerram, 2006) and we tested an extreme range of shapes from 1:1:2 and 1:2:5 to 1:5:5. In order to simplify the comparisons, we normalized the lengths using an arbitrary "unit". The longest dimension of the plagioclase was fixed to 200 units along the X-axis; then, the intermediate and short dimensions were computed according to the different shapes along the Y and Z axes. For example, if the model uses the shape 1:1:2, then the intermediate and shortest axes (Y and Z) were both set to 100 units (Fig.2).

The plane used to cut the plagioclase can be described with the formula

$$a \cdot x + b \cdot y + c \cdot z + d = 0 \qquad (1)$$

Where *a, b, c* are random values ranging from -1 to 1, and d represents a displacement that takes random values along the longest axis. Different values of *a*, *b*, *c*, *d*, produce different 2D zoning patterns (Fig. 2). For this simple crystal with 5 compositional zones (at 10, 30, 50, 70, and 90 An Mol%) we produced 900 random cuts; 400 random cuts were used for a second example, more complex and more similar to the natural plagioclase (with compositions described in later sections). The number of resulting 2D cuts is lower that the number of random planes because not all planes passed through the crystal (Fig. 3). In addition to the random cuts, we can characterize the compositional zoning of the 3D crystal by three perpendicular 2D sections that pass through the principal-center of the crystal; the one parallel to the X-axis is shown in Figure 2. We call these the ideal sections (**ID;** see caption of Fig 4 for more details). These three sections have the same area compositional histograms but different shapes (Fig. 4).

Characterization of 2D plagioclase zoning patterns and compositional maps

The random cuts produce a large variety of zoning patterns (Fig. 3), ranging from compositionally homogenous sections (e.g. section #4 in Fig. 3) to up to a five-zoned pattern (e.g. section #9 in Fig. 3). To characterize the 2D sections we used the normalized area compositional histograms of each section. This means that for each section we calculated the area with the same composition (i.e. the same An content within a given tolerance, see below) and normalized it to the total area of the crystal section (Fig. 4). We call the random individual sections ***RI*** where I = 1, 2, 3……Z and Z is the total number of sections. To characterize the compositional distribution of all 2D random sections from a 3D crystal we added all the areas of the zones with the same composition from all sections and normalized them to the total area of the sections. We call this distribution the 'all random population', abbreviated as ***RA***. This is an important distribution because it can be used to characterize the various crystal populations in natural samples for which we do not know a priori the zoning of the 3D crystal. In the same way, we can calculate the area compositional distribution for the **ID** sections (Fig. 4). We found that the compositional distributions of the **ID** and all **RA** for the simple 5 zoned crystal are similar but not identical; typically the compositions of the rims are over-represented and those of the crystal centers under-represented in the RA distribution (Fig. 5). This reflects that although all random sections pass through the crystal rims, some sections miss completely the cores (e.g. off-core sections). This gives a similarity of about 75 % between the RA and ID distributions (see below for the precise definition of similarity used) and this difference varies depending on the details of the crystal zoning.

Comparing between 2D sections using the similarity of compositional maps

To quantify the difference or similarity between two compositional distributions such as the ***RA*** and ***ID***, we used the technique of histogram intersection, which is used in other fields, e.g. in color index (Swain and Ballard, 1991), and also in problems of computer vision (e.g. Kumar et al., 2012). Given a pair of histograms, in our case RA and ID, each containing j=1...n bins of An content, the mismatch between the histograms can be defined as the absolute difference (abs) between the two:

$$Mismatch = \sum_{j=1}^{n} abs(ID_j - RA_j) \qquad (2)$$

and the normalized mismatch as:

$$Normalized\ Mismatch\_ID * RA = \frac{\sum_{j=1}^{n} abs(ID_j - RA_j)}{\sum_{j=1}^{n} \max(ID_j, RA_j)} \times 100 \qquad (3)$$

where the max is the largest frequency of each bin in the two histograms. From this, we defined the similarity between two compositional histograms as (Fig. 6):

$$Similarity\ ID\_RA\ (S\_ID * RA) = 100 - Normalized\ Mismatch\_ID * RA \qquad (4)$$

For example, the similarity between the RA and the ID distributions (*S_ID*RA*) for the simple crystal is about 75 % (Fig. 6). The similarity depends on the number of bins, which in our case depends in turn on the An range we choose for each bin. Many authors advise that for real data sets histograms based on 5-20 bins usually suffice, noted by Scott (1979). Calibration of greyscale of BSE images of plagioclase typically produces An values with a precision of 1 to 2 mol% (Ginibre et al., 2002). We tested bin sizes of 1, 2, 3 mol% An and found that the 3 mol% (the number of bins is more than 30) is quite similar to using 1 mol% (see Supplementary **Appendix 1**), so we used a bin size of 3 mol% An unless otherwise noted. We will use the similarity between different sections or groups of sections in this manner for the rest of the manuscript (e.g. Fig. 6). The uncertainty of the similarity was calculated to be up to 1 % by assuming a relative analytical uncertainty of 0.5 % on the An content in this study.

Reference and ideal sections

Using the method we described above, we can calculate the similarity between any section types and/or between groups of sections. This is the first step to be able to identify ideal or representative 2D sections and thus different crystal populations in natural samples. For example, the compositional similarity between the ID and RA distributions can be calculated to be about 75 % (Fig. 6). The simplest approach to determine the number of crystal populations and representative sections would be to use directly the ID and RA distribution, but this is not possible for several reasons. The RA distribution actually does not correspond to any individual crystal section, so although it characterizes the overall plagioclase population we still need to identify representative sections. Moreover, although we know which the ID sections in our numerical experiments are, this is not the case in natural samples. These problems appear when we want to filter out the random cuts for a single crystal population, but they become even more apparent when there are multiple crystal populations rather than a single one. Thus, we have designed a strategy where we first try to filter out the effects of random cuts using proxies to characterize the RA distribution and the number of populations, and then we can identify the best 2D sections using more strict criteria and the ID sections.

The first step is to calculate the similarity distribution between the overall random population (***RA***) and each random section (R1, R2, R3 ...RZ) (named ***S_RA*R1-RZ***; Fig. 7). We find that it shows one main peak about 70-80% similarity that includes 25% of the sections (a total of 180 sections out of 745), and another at 30 % similarity that includes about 14% of the sections. Moreover, more than 50% of the random sections have a similarity (S_RA*R1-RZ) ≥ 70 % with that of the overall population, and thus they could be considered as a reasonably representative sample of the 3D crystal. A much smaller number of sections (about 5 %) have a similarity ≥ 90% with the overall random population and we call them the reference sections (***REF***). Thus, we can use the RA distribution to identify individual reference sections from the overall population. However, this method works if there is only a single 3D crystal population, but not if there are more, since then the similarity values are much lower and it is not possible to identify the REF sections although they must exist in data sample (see detail in the population mixing sections).

Another way of identifying the reference sections is by calculating the similarity between a given individual section and all the rest of the random sections (e.g., not using the RA directly; S_R1*R1-RZ; Fig. 8). The similarity distribution pattern of each individual section and the rest is quite variable, but it is apparent that reference sections are more abundant for similarities ≥ 70 % (Fig. 8a). This finding is akin to the effect of random cuts of crystal shapes, where some characteristic and important shapes, areas, and dimensions are more frequent because of the geometrical symmetry of the object (e.g., Higgins, 1994). In our case, it is apparent that even if the cuts were done at random, the compositional histograms of sections parallel or perpendicular to the main geometry axes have the same compositional distribution and thus are more frequent to begin with than any other section (Fig. 8b). See Supplementary **Appendix 2** for more details about the method involved in identification of these sections. We also discuss later how this method can be improved by using thresholds of similarities.

Another important group of sections that need to be characterized are the IDs, and we have calculated the similarity distribution between ID and each individual section (RI) [here called ***S_ID*R1-RZ***; Fig. 7]. The distribution has a peak at about 20% similarity, which means that most random sections are quite different from the ID, which reflects the fact that the random 2D cut effect significantly changes the compositional maps, with many sections that are off-center. However, there are more sections with a similarity (S_ID*R1-RZ; Fig. 7) ≥ 90 % to the ID than to the overall random distribution (S_RA*R1-RZ; Fig. 7). Thus, although the random cut effects are important, there are still many 2D sections that are very similar (S_ID*R1_RZ $\geq$ 90 %) to the ideal sections. These random sections record the most complete information and we will also call them ideal sections because they are $\geq$ 90% similar to the theoretical ID sections.

The reference sections (***REF***) are different from the ideal sections in that they typically do not record the inner parts of crystals. However, since we do not know a priori how many crystal populations there are in a given natural dataset, we have to use the method of calculating the similarity between each random section (Fig. 8). It is apparent that the most frequent sections for similarities $\geq$ 90% are the ideal ones (see Supplementary **Appendix 2** for more details). Later we quantify better these relationships using threshold values; here we wanted to illustrate that it is possible to identify the ID and REF by comparing each individual section to the rest in a systematic and statistical manner. In typical studies of plagioclase zoning, petrologists tend to use sections with shapes close to the most classical shapes of plagioclase, i.e. rectangular or square. If we do a subsample of the rectangular and square sections (sides >=5) from our random samples we find that about 36% of those have a similarity with the ID of 90 % or higher, which means that, in the case of a single crystal population, the likelihood of choosing an ID section if the shape is rectangular is much higher.

So far, we have used a crystal with three perpendicular axes, but plagioclase belongs to the triclinic system and thus one axis at an angle that deviates significantly from 90^0 from the rest (Deer et al. 1992). To test whether this has effect on our method we have also done a simulation with a crystal that has different angles between the axes. The Fig.9 shows the wire-structure of a zoned crystal with an angle of 115 ^{o} between two axes that is close to triclinic system of anorthite plagioclase according to Deer et al. 1992. In total, we produce 900 cuts to get 714 2D patterns. We still could find that it shows one main peak about 70-80% similarity that includes 25% of the sections, and another at 35 % similarity that includes about 20% of the sections. The ID distribution has a peak at about 20% similarity, and there are more sections with a similarity ≥ 90 % (Fig.9). The two distributions are similarity to that of three perpendicular axes model. Thus, the approach described above for the identification of the different ID and REF sections is still valid. This is because the effects of random cuts are much larger than the small deviations deriving from the “incorrect” use of three perpendicular axes.

Models with complex crystal zoning patterns

Natural plagioclase crystals are however often more complex that the five-zone crystal simulation (Fig. 2) we have used so far (e.g. Ginibre et al. 2002). Thus, we also generated three other types of 3D crystals (Fig. 10): Type A with patterns of An values that generally increase from rim to core with oscillatory zoning and fine variations of less than 5% An, Type B with large and abrupt An changes between core and rim, and Type C which combines the zoning characteristics of the previous two types. Calculation of the similarity properties for the 3 types of crystals shows that they can be treated in the same manner as the simpler crystal, although in detail the similarity thresholds for the identification of the reference and ideal sections are somewhat different (Fig. 10). The similarity histogram of RA and each individual random section has main peaks at 70-90% similarity for crystals of Type A, at 70-80% similarity for those of Type B, and at 80-90% similarity for those of Type C, whereas the similarities of ID and each random section have also high frequency for all crystal zoning types at similarity ≥90 % (Fig. 10).

To further refine our approach to find the reference and ideal sections we need to do an additional step that is critical when we are dealing with multiple crystal populations (e.g. for cases of magma mingling). We first calculate the similarity between each random section, and then we calculate the similarity between the section that has the largest number of section pairs with a minimum similarity threshold, starting from 50 up to 100%, with that of the RA distribution (please see Supplementary **Appendix 2** for a step by step description of this approach). We determine where this similarity is at a maximum (Fig. 11). By doing this we find that for example, for Type A crystals we have a similarity ≥ 90 % between some random sections and the RA for a similarity threshold between 70-80 % (Fig. 11a). This means that these random sections can be taken as representatives of the Reference sections and thus used to characterize one crystal population. Note that this threshold is also higher than the mean threshold, which provides additional constraints to identify the appropriate threshold value. Different zoning types give slightly different thresholds, but at about 80% threshold, the sections are all ≥90 % similar to the RA and thus can be used to identify the reference section. Similar relations can be used to identify the ideal sections (Fig. 11b) but these require higher thresholds at 90-100%. Finally, we also did a Monte Carlo simulation where we tested up to 1200 3D crystals with different An zoning patterns. We found that our inferences of similarity thresholds at 70-80%, and 90-100% are robust and able to identify the Reference and Ideal sections, respectively, independent of the type of zoning (Fig. 12).

Multiple crystal populations

The next level of complexity is to be able to identify the reference and ideal sections for mixed crystal populations. We designed a numerical experiment were we first generated 50-50 mixtures of two crystal populations involving three different scenarios (Fig. 13): (1) two crystal populations with different An core compositions but the same rim compositions, as representative of mafic-silicic magma interactions (e.g. Feeley and Dungan, 1996); the similarity between these two populations is relatively low, about 15 % (Fig. 13 a,d,g). (2) two populations with a similar An compositional range but different zoning patterns; they have a higher similarity of about 30% (Fig. 13 b,e,h). (3) two crystal populations where one of the two started to grow at a later time, and thus it shares the same more recent history but is missing the earlier one (Fig. 13 c,f,i). In this last scenario, the two populations have the highest similarity of about 60%. The more similar the two crystal populations are before the mixing, the harder it will be to identify and separate them.

We find that the similarity histograms for the three mixing scenarios (Fig. 13 g,h,i) are quite different from those of single crystal populations only affected by the random cut effect (e.g. Fig. 10). Most notably the maximum of similarity tends to reach much lower values (Fig. 13 g,h,i). In these examples it is apparent that the similarity information of the 2D crystal sections and the RA reflects the combination of both the cut effect and the mixing of the two populations. In effect, the overall RA that we might measure in a natural sample is a "weighted average" that is the result of mixing the RA of each individual population (Fig. 13 g,h,i).

To identify the two crystal populations we calculated the similarities between the different sections and we focused only on those with a similarity threshold around 80 % (as suggested by Fig. 11). We then calculated the similarity of these sections with the overall RA for each of the scenarios and found low similarities ranging from 46 to 77 % (Fig. 14 a,c,e). This means that there has to be other reference sections in the population that would explain the full dataset. Thus, we removed from the pool the sections with a given similarity threshold (e.g. >80 %) to the first reference section, and chose another section with a threshold around 80% and calculated whether the similarity with the overall population increased by mixing them in different proportions. We keep doing this using least squares minimization between the RA distributions and that of mixing different sections in different proportions. Once we found that the overall similarity increased to >90 % (Fig. 14 a,c,e) we stopped, and we considered that we were able to explain >90% of the overall compositional distribution of the plagioclase.

For the three scenarios above we were able to obtain similarities of the two reference sections and the RA of each scenario between 92 and 95 % (Fig. 14 b,d,f). This approach worked for the reference sections as well as for the ideal sections. We also tested the effect of varying proportions of mixing ratios between two crystal populations (e.g., 10-90, 70-30) using scenario two (Fig. 15). We found that using the same approach of reference/ideal sections and least squares minimization we are still able to recover the two populations. However, when the proportion of one population gets close to 90% or more it becomes more difficult to identify any other, until it becomes eventually undetectable. An example of application of our approach to mixed crystal populations from a lava of Mayon volcanoes can be found in Supplementary Appendix 3.

How many crystals are needed to represent a given sample

A related important question when studying natural crystals is how many random sections are needed to characterize the population(s) and build a reliable RA distribution. We used the Type C plagioclase (see Section 6) as the sample test. We made 400 random cuts and calculated their RA (RA_{400}). We then compared their similarity to an increasing amount of random sections, randomly extracting 10 of them (RA_{10}) until 300 (RA_{300}). Because of the large variability of the RA when we sample a small number of sections at random (e.g. RA_{10}; Fig. 16a), we calculated their RA multiple times, and thus their similarity also varies (Fig. 14b). We found that within about 60-100 random picks of random sections we characterize the full variability (Fig. 16b). Moreover, we find that the similarity between the distributions obtained from 80 (RA_{80}) to 100 (RA_{100}) random sections with that of the 400 sections (RA_{400}) is consistently higher than 95%, and thus we conservatively suggest that 100 random 2D sections are enough to represent RA of the entire crystal population.

Implications for identification of plagioclase populations and representative crystal sections

The large variety of compositional and textural features of many plagioclase crystals from igneous rocks can be party explained by the effects of 2D random cuts of 3D plagioclase crystals. Using the compositional histogram of 2D sections of many plagioclase crystals and statistical analyses based on the compositional similarity between the different sections it is possible to account for the effects of the random cuts and separate them from the effects of the presence of different crystal populations. With our method, we can identify the different crystal populations that are at least 90 % similar and the proportions of the different populations. The identification of the different populations by using reference and ideal sections removes the problem of a subjective choice of sections and allows studing the crystal sections that are more representative of the samples. These can be further studied using detailed electron microprobe, ion microprobe and/or microdrilling of isotopes, which should lead to a much more robust understanding of magmatic processes based on plagioclase crystal records.

Acknowledgements

We acknowledge discussion with D. Ruth on this topic, and C. Newhall for the 1947 sample of Mayon. G. Fabbro, J. Herrin, D. Ruth, Sri Budhi, W. Li, D. Schonwalder, C. Widijiwayanti, L. Zeng, T. Flaherty, O. Bergal, and D. Nurfani are thanked for their participation in the expert elicitation exercise on the number of plagioclase types. Reviews by G. Bergantz, P. Izbekov, M. Higgins, and editorial handling and reviews of J. Hammer significantly improved the presentation of ideas in this manuscript and greatly appreciated. This research is supported by the National Research Foundation of Singapore and the

Singapore Ministry of Education under the Research Centres of Excellence initiative, under the "Crystal pattern" project.

References

Anderson, A.T. (1984) Probable relations between plagioclase zoning and magma dynamics, Fuego Volcano, Guatemala. American Mineralogist, 69(7-8), 660-676.

Blundy, J.D., and Shimizu, N. (1991) Trace element evidence for plagioclase recycling in calc-alkaline magmas. Earth and Planetary Science Letters, 102(2), 178-197.

Bouvet De Maisonneuve, C., Dungan, M.A., Bachmann, O., and Burgisser, A. (2012) Petrological Insights into Shifts in Eruptive Styles at Volcán Llaima (Chile). Journal of Petrology.

Cashman, K., and Blundy, J. (2013) Petrological cannibalism: the chemical and textural consequences of incremental magma body growth. Contributions to Mineralogy and Petrology, 166(3), 703-729.

Cheng, L.-L., Yang, Z.-F., Zeng, L., Wang, Y., and Luo, Z.-H. (2014) Giant plagioclase growth during storage of basaltic magma in Emeishan Large Igneous Province, SW China. Contributions to Mineralogy and Petrology, 167(2), 1-20.

Costa, F., Chakraborty, S., and Dohmen, R. (2003) Diffusion coupling between trace and major elements and a model for calculation of magma residence times using plagioclase. Geochimica et Cosmochimica Acta, 67(12), 2189-2200.

Davidson, J., Tepley III, F., Palacz, Z., and Meffan-Main, S. (2001) Magma recharge, contamination and residence times revealed by in situ laser ablation isotopic analysis of feldspar in volcanic rocks. Earth and Planetary Science Letters, 184(2), 427-442.

Deer, W.A., Howie, R.A., and Zussman, J. (1992) An introduction to the rock-forming minerals. Longman London.

Druitt, T.H., Costa, F., Deloule, E., Dungan, M., and Scaillet, B. (2012) Decadal to monthly timescales of magma transfer and reservoir growth at a caldera volcano. Nature, 482(7383), 77-80.

Feeley, T.C., and Dungan, M.A. (1996) Compositional and Dynamic Controls on Mafic—Silicic Magma Interactions at Continental Arc Volcanoes: Evidence from Cordón El Guadal, Tatara-San Pedro Complex, Chile. Journal of Petrology, 37(6), 1547-1577.

Ginibre, C., Kronz, A., and WoÈrner, G. (2002) High-resolution quantitative imaging of plagioclase composition using accumulated backscattered electron images: new constraints on oscillatory zoning. Contributions to Mineralogy and Petrology, 142(4), 436-448.

Hibbard, M. (1995) Petrography to petrogenesis.Prentice Hall, New Jersey.

Higgins, M.D. (1991) The origin of laminated and massive anorthosite, Sept Iles layered intrusion, Quebec, Canada. Contributions to Mineralogy and Petrology, 106(3), 340-354.

Higgins, M.D. (1994) Determination of crystal morphology and size from bulk measurements on thin sections: numerical modelling. American Mineralogist, 79, 113-119.

Higgins, M.D., 2006. Quantitative Textural Measurements in Igneous and Metamorphic Petrology. Cambridge: Cambridge University Press. Cambridge, UK.

Higgins, M.D., Voos, S. and Vander Auwera, J., 2015. Magmatic processes under Quizapu volcano, Chile, identified from geochemical and textural studies. Contributions to Mineralogy and Petrology, 170(5): 51.

Kumar, N., Belhumeur, P.N., Biswas, A., Jacobs, D.W., Kress, W.J., Lopez, I.C., and Soares, J.V. (2012) Leafsnap: A computer vision system for automatic plant species identification. Computer Vision–ECCV 2012, p. 502-516. Springer.

Landi, P., Métrich, N., Bertagnini, A., and Rosi, M. (2004) Dynamics of magma mixing and degassing recorded in plagioclase at Stromboli (Aeolian Archipelago, Italy). Contributions to Mineralogy and Petrology, 147(2), 213-227.

Mathworks. (2014) MATLAB version 2014b. Natick, Massachusetts: The MathWorks Inc.

Morgan, D.J., and Jerram, D.A. (2006) On estimating crystal shape for crystal size distribution analysis. Journal of Volcanology and Geothermal Research, 154(1–2), 1-7.

Neill, O.K., Larsen, J.F. Izbekov, P.E. and Nye, C.J., 2015. Pre-eruptive magma mixing and crystal transfer revealed by phenocryst and microlite compositions in basaltic andesite from the 2008 eruption of Kasatochi Island volcano. American Mineralogist, 100(4): 722-737.

Nicotra, E., and Viccaro, M. (2012) Unusual magma storage conditions at Mt. Etna (Southern Italy) as evidenced by plagioclase megacryst-bearing lavas: implications for the plumbing system geometry and summit caldera collapse. Bulletin of Volcanology, 74(4), 795-815.

Pearce, T.H. (1984) The analysis of zoning in magmatic crystals with emphasis on olivine. Contributions to Mineralogy and Petrology, 86(2), 149-154.

Scott, D.W., 1979. On optimal and data-based histograms. Biometrika, 66(3): 605-610.

Shea, T., Costa, F., Krimer, D., and Hammer, J.E. (2015) Accuracy of timescales retrieved from diffusion modeling in olivine: A 3D perspective. American Mineralogist, 100(10), 2026-2042.

Shelley, D. (1993) Igneous and metamorphic rocks under the mcroscope: Classification, textures, microstructures and mineral preferred orientations. Chapman & Hall, New York, 445p.

Singer, B.S., Dungan, M.A., and Layne, G.D. (1995) Textures and Sr, Ba, Mg, Fe, K, and Ti compositional profiles in volcanic plagioclase: clues to the dynamics of calc-alkaline magma chambers. American Mineralogist, 80(7-8), 776-798.

Stelten, M.E., Cooper, K.M., Vazquez, J.A., Calvert, A.T., and Glessner, J.J.G. (2015) Mechanisms and Timescales of Generating Eruptible Rhyolitic Magmas at Yellowstone Caldera from Zircon and Sanidine Geochronology and Geochemistry. Journal of Petrology.

Streck, M.J. (2008) Mineral Textures and Zoning as Evidence for Open System Processes. Reviews in Mineralogy and Geochemistry, 69(1), 595-622.

Swain, M.J., and Ballard, D.H. (1991) Color indexing. International journal of computer vision, 7(1), 11-32.

Wallace, G.S., and Bergantz, G.W. (2002) Wavelet-based correlation (WBC) of zoned crystal populations and magma mixing. Earth and Planetary Science Letters, 202(1), 133-145.

Wallace, G.S., and Bergantz, G.W. (2004) Constraints on mingling of crystal populations from off-center zoning profiles: A statistical approach. American Mineralogist, 89(1), 64-73.

Wallace, G.S., and Bergantz, G.W. (2005) Reconciling heterogeneity in crystal zoning data: an application of shared characteristic diagrams at Chaos Crags, Lassen Volcanic Center, California. Contributions to Mineralogy and Petrology, 149(1), 98-112.

Yoder, H., Stewart, D., and Smith, J. (1957) Ternary feldspars. Carnegie Institution of Washington Year Book, 56, 206-214.

Zellmer, G.F., Blake, S., Vance, D., Hawkesworth, C., and Turner, S. (1999) Plagioclase residence times at two island arc volcanoes (Kameni Islands, Santorini, and Soufriere, St. Vincent) determined by Sr diffusion systematics. Contributions to Mineralogy and Petrology, 136(4), 345-357.

Figure captions

Figure 1. Composite BSE image of a thin section from a basaltic andesite lava of an eruption of Mayon (Philippines). Note the large variety of zoning and textures of the plagioclase phenocrysts. Similarly, complex plagioclase crystals are found in many other igneous rocks. The BSE image of the thin section was acquired using a JEOL JSM-7800F Scanning Electron Microscope (SEM) of the Asian School of the Environment (Nanyang Technological University). In total, we collected 660 individual BSE images that were collated to build a whole thin section image, with a resolution of 8089*5563 pixels after the SEM run 12 hours. Image mosaic was obtained using the same contrast and brightness when we were taking BSE images using the Aztec software from Oxford. The thin section contains about 750 plagioclase phenocrysts (crystals larger than 250 μm in the shortest dimension).

Figure 2. Wire-structure representation of zoned 3D plagioclase crystal, with different colors representing different compositions and the three perpendicular geometric axes. Two-dimensional (2D) cuts of the crystal in different orientations and locations as occurs during the making of petrographic thin sections will produce different 2D images. Three planes go through the center and are perpendicular to one of the main geometric axes (e.g. section 1). These have the most complete zoning information of the 3D crystal and are called ideal sections (ID). Many others will occur at random locations and miss part of the information (section 2). Please note that the different apparent shapes of section #2 in the 3D and 2D views are just due to the angle of view, they have actually the same shape.

Figure 3. Example of the variety of 2D zoning patterns and shapes produced from 100 random cuts through the 3D crystal shown in Figure 2. Blank means the planes do not go

through the crystal. 2D sections with black underline are the ideal sections, those with the red dots underlines are the reference sections (see details in the text). Based on these patterns, we can calculate compositional maps distribution of each section and that of all the sections.

Figure 4. Examples of histograms of compositional distributions of 2D sections akin to those that would be obtained from BSE images that were calibrated for the anorthite content (An %). (a) Distribution of a random section (R1); (b) Distribution of one ideal section (the three ideal sections have the same distribution, ID); (c) composite distribution obtained from adding the histograms of all random sections (RA) shown in Figure 3. Note that the ID = Ideal sections are perpendicular to one of the geometric axes of the crystal. This is the most representative of the 3D crystal. It is known in the models but can't be measured and is therefore unknown in natural samples. RI (1...Z) = Individual random section. It can be measured and thus it is known in natural samples. RA = All random sections 'combined'. It can be calculated from natural samples. REF = Reference sections. Individual random sections with a similarity higher than 90% with the RA. They can be determined from natural samples.

Figure 5. Comparison of histograms obtained from adding the histograms of all the random sections (RA) with that of the ideal sections (ID) for crystals with different shapes. Note how the RA distribution underestimates the core compositions and overestimates the rims because the random cut effect might miss the cores but not the rims. The range of crystal shapes we have explored do not significantly affect the relation between the RA and ID distribution. The numbers in the upper right corner of each graph are the different aspect values of the crystals.

Figure 6. Example of calculation of similarity between ID and RA distributions of An (using 1-2-2 shape and 3 mol% as the An for each bin). Note that there is a similarity of 74 % (or a

mismatch of 26 %) between the two distributions, implying that using all the random sections as a whole we can have a first order idea (e.g. 74 % similarity) of the ID distribution for a single crystal population. See details in text about how to calculate the similarity. The exact relations between the ID and RA distributions depends on the style of zoning (see text for other examples).

Figure 7. Similarity of the ID and RA distributions with each individual random section. Note that how the similarity distribution with the RA has two main peaks although it is only one crystal, and that of the ID has mainly one peak. The sections with a similarity ≥ 90% with the RA are called reference sections (REF), and those with a similarity ≥ 90% with the ID are called ideal sections. Vertical dotted line marks the 90% similarity.

Figure 8. Similarity plots between each random section and the rest and identification of the reference and ideal sections. (a) It can be seen that there is a large range of similarity values and distributions but (b) the sections with highest frequency and similarity between 70% and 100% correspond with the reference section, and (c) those with the highest frequency and similarity (≥ 90%) correspond to the ideal sections.

Figure 9. (a) Wire-structure of a zoned crystal with an angle of 115 º between two axes and thus close to triclinic system of anothite plagioclase (Deer et al., 1992). (b) Similarity of the ID and RA distributions with each individual random section. Note that the relations between the different sections are the same for this crystal and those that have three perpendicular axes and which we have used for the rest of simulations. This means that our approach is also valid for triclinic crystals.

Figure 10. Three different types (A, B, C) of complexly zoned 3D plagioclase models (a through c), 1D traverses through the ideal sections (d through f), frequency histograms of An for their corresponding RA sections (g through i), and similarity calculations between the random sections and the ID and RA (j through l). Note that in the similarity histogram for all three types of plagioclase (j through l) between the RA with all individual section which has a peak similarity ≥ 70-80%, and the similarity histogram of ID with all individual section has a lot of sections with similarity ≥ 90%. Vertical dashed red line is the mean of the similarity of the S_RA*R1-RZ. See text for more discussion.

Figure 11. Relationship between the similarity threshold and the similarity between the RA and the ID sections for the three types of plagioclase zoning. The figures (a, c, e) show that it is possible to identify the REF sections by choosing a similarity threshold of 80% for the three kinds of zoned crystals. In a similar manner the b, d, f show that the ID sections can be determined by choosing > 90% of similarity threshold for most zoned crystals. See text and Suppl. Appendix 2 for more details and discussion.

Figure 12. Monte Carlo experiments (repeated 1200 times) that show that the method to find the reference and ideal sections works for a large variety of crystal zoning types to obtain a similarity better than 90%.

Figure 13. Compositional and similarity relationships produced by mixing two different crystal populations at 50-50. The similarity between the two populations increases from 15 % to 59 %. The more similar the two populations are, the more difficult it is to uniquely identify and distinguish them. Panels (a), (b), and (c), are the 3D plagioclase we used; Panels (d), (e),

(f) are 1D profiles through the center of the crystals; Blue line in (g), (h), (i) correspond to the RA of population 1; Green line correspond to the RA of population 2, and red bar correspond to the RA of two mixed populations. Note that here we used lines and dots rather than bars in the histograms for a better visualization of the fits. Yellow bar in (j), (k) and (i) are the distribution of similarities; red dotted line is the mean of the similarity.

Figure 14. Examples of fits of compositional histogram distributions for the three mixing scenarios (1, 2, 3) shown in Fig. 13. Note how the distribution of the known RA (line with dots and light green, red, and blue colors extracted from Fig. 13) can be reproduced to a high similarity (S_RA*REF1REF2 >90 %) when we use two references sections as opposed to only one section (S_RA*REF1) [left hand side panels (a), (c), (e)]. We can also reproduce very well the ideal sections [Right hand panels (b), (d), (f)]. Note that some lines may be hidden behind others. See text for more discussion. Note that here we used lines and dots rather than bars in the histograms for a better visualization of the fits.

Figure 15. Example of varying mixing proportion of the two populations from scenario 2 [90-10, panels (a),(c),(e); 70-30, panels (b), (d), (f)]. Please compare with the 50-50 proportion shown in Fig. 14. It is still possible to find the reference and ideal sections of the two populations, although when one population becomes less than about 10% it becomes increasingly difficult. See text for more discussion. RA1, RA2 and RA are the RA of population 1, population 2 and two mixed populations respectively; REF1, REF2 is the REF of population 1, population 2 respectively; REF1REF2 are equal to 0.9*REF1+0.1*REF2 or 0.7*REF1+0.3*REF2; ID of P1, ID of P2 are the ID of population 1 and population2; ID1 and ID2 are the almost ideal sections of population 1 and population 2. Note that here we used lines and dots rather than bars in the histograms for a better visualization of the fits.

Figure 16. Simulation aimed at calculating how many crystals we need to determine properly the correct distribution of RA. (a) Relationship between number of crystals and the similarity of RA of these number of crystals and RA of 375 crystals; (b) assumsing 10 crystals are used to calculate the RA, how many times should be used to determine the real distribution of similarities between RA of 10 crystals and 375 crystals. See text for more discussion.

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-2017-5929CCBYNCND

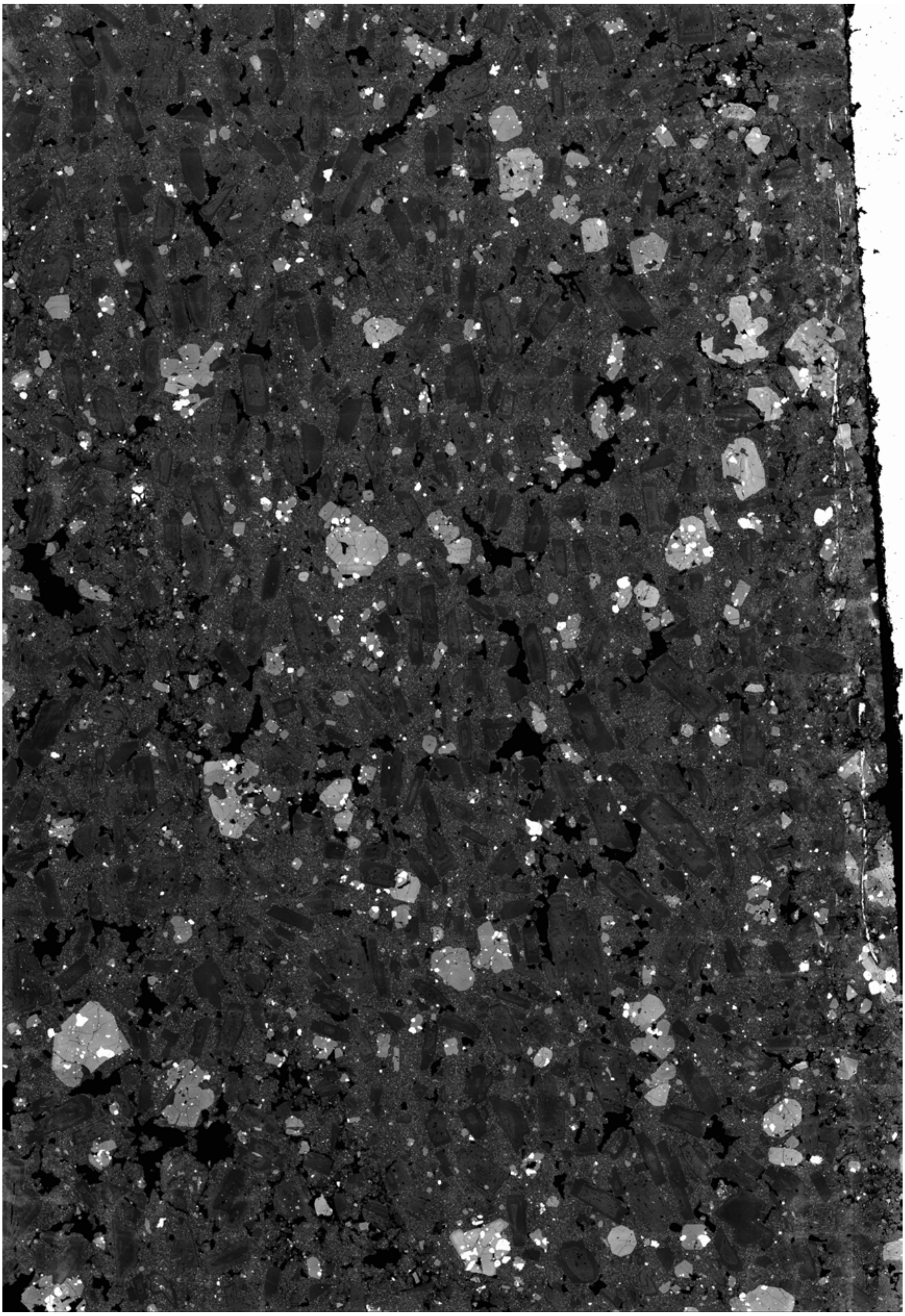

Figure 1

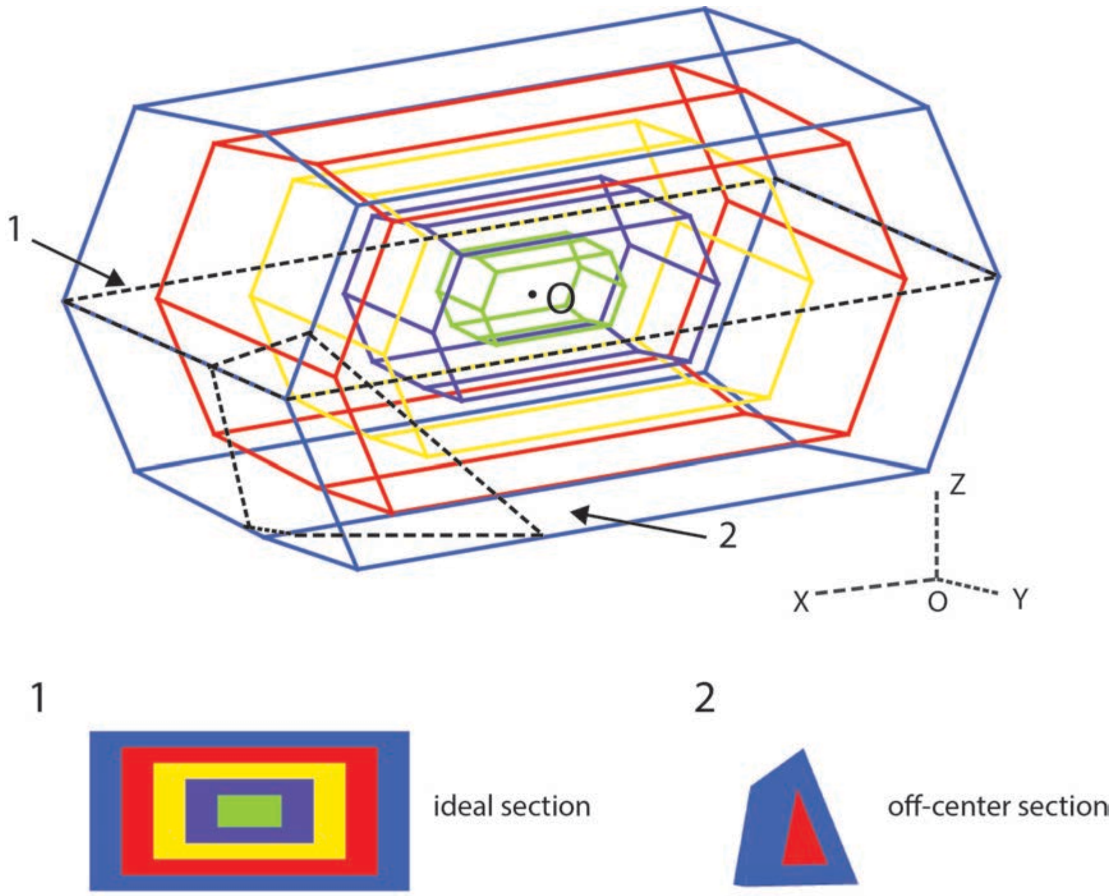

Figure 2

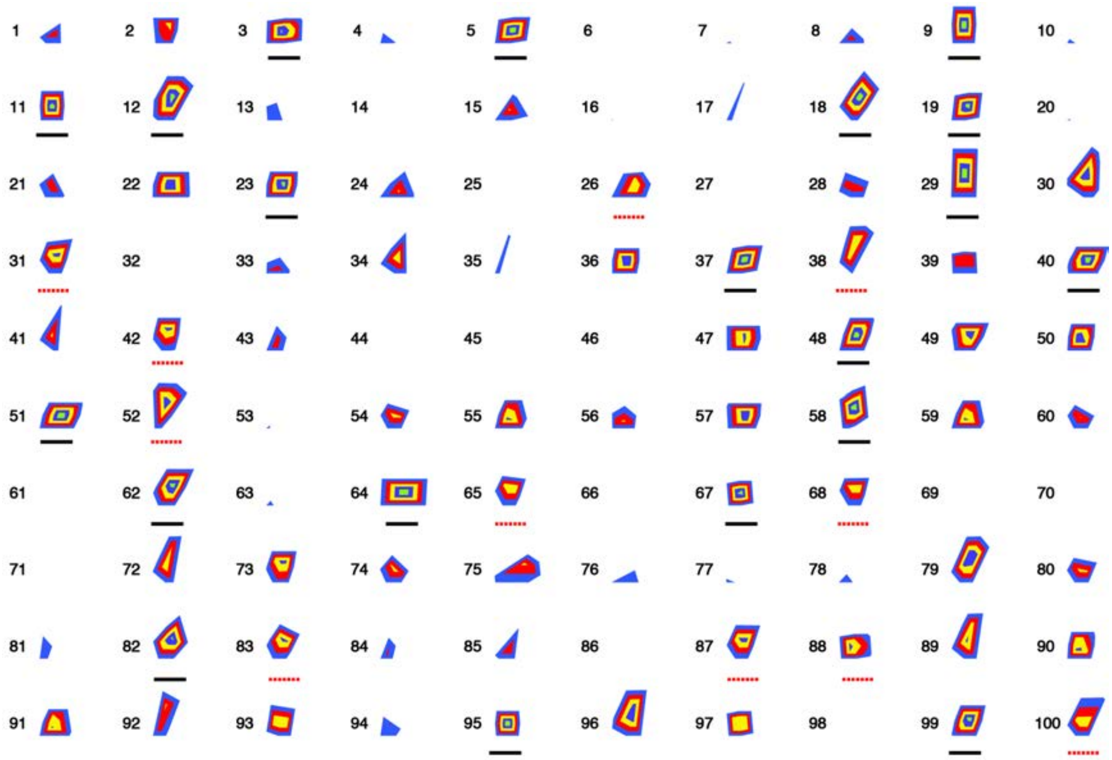
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Figure 3

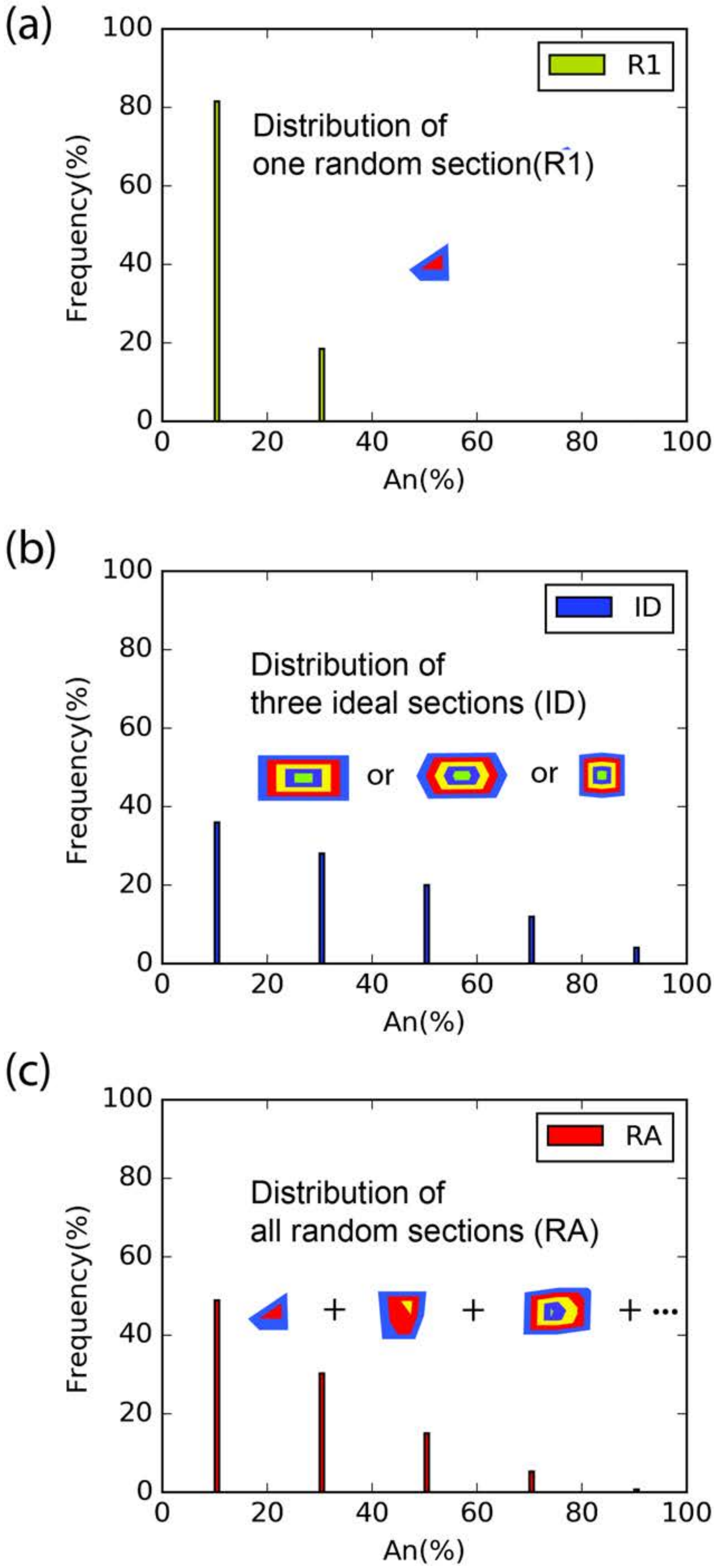
(a)
R1
Distribution of
one random section(R1)
Frequency(%)
An(%)
0
20
40
60
80
100
(b)
ID
Distribution of
three ideal sections (ID)
or
or
Frequency(%)
An(%)
(c)
RA
Distribution of
all random sections (RA)
+
+
+ ···
Frequency(%)
An(%)

Figure 4

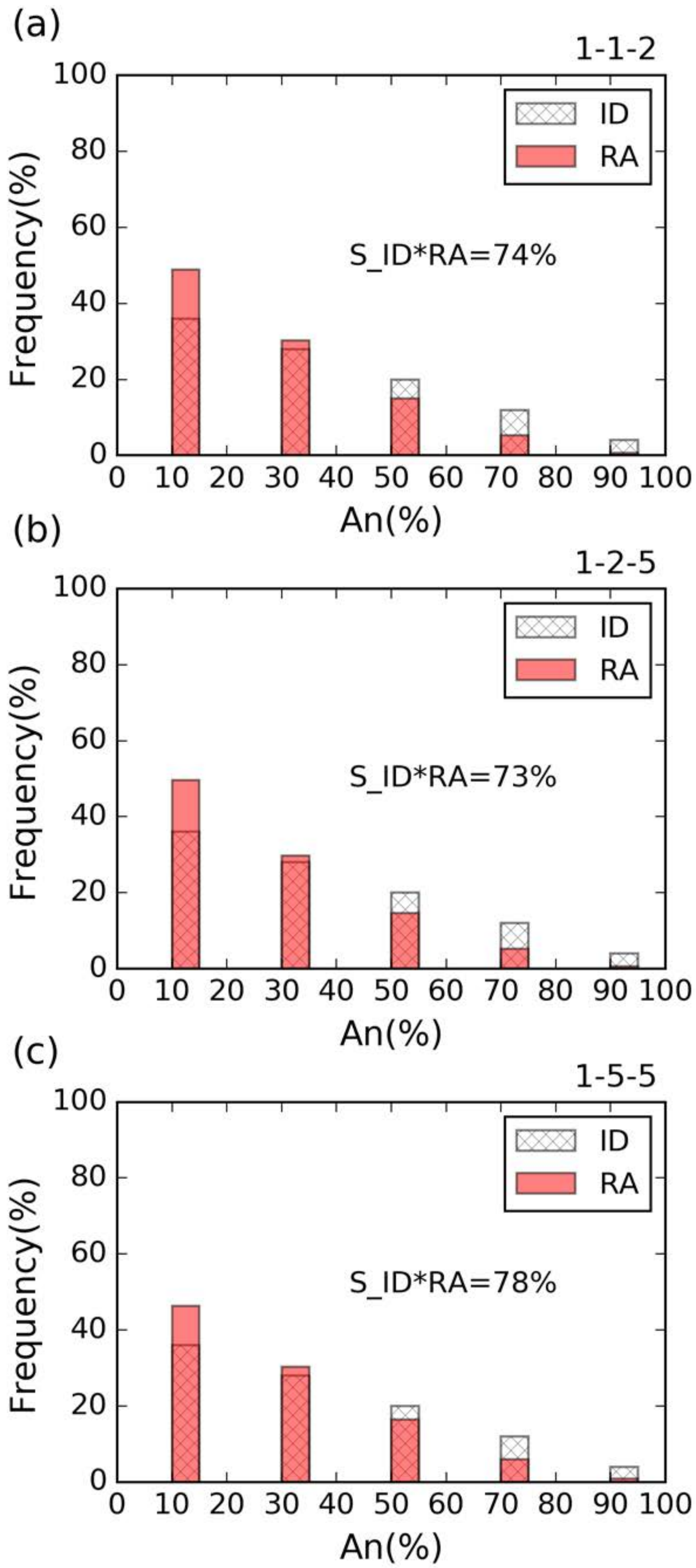

Figure 5

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-2017-5929CCBYNCND

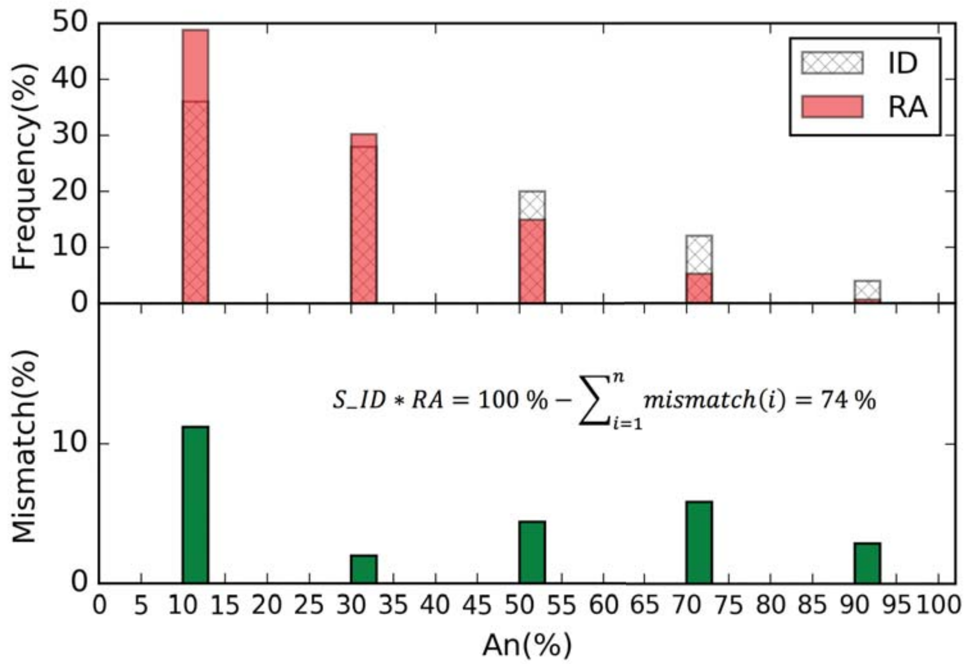

Figure 6

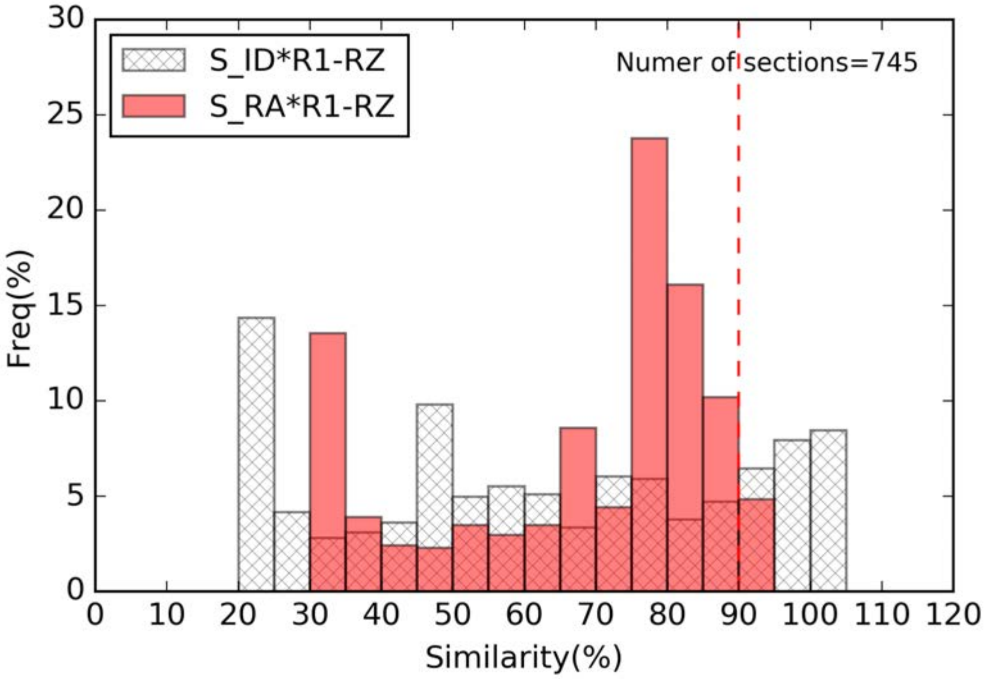
S_ID*R1-RZ
S_RA*R1-RZ
Numer of sections=745
Freq(%)
Similarity(%)
0
5
10
15
20
25
30
0
10
20
30
40
50
60
70
80
90
100
110
120

Figure 7

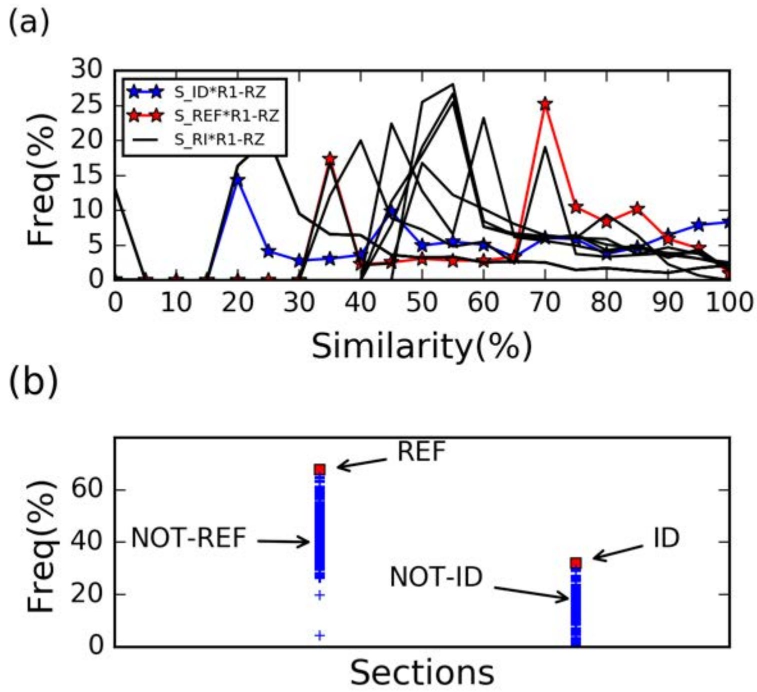
(a)
S_ID*R1-RZ
S_REF*R1-RZ
S_RI*R1-RZ
Freq(%)
Similarity(%)
(b)
REF
NOT-REF
ID
NOT-ID
Freq(%)
Sections

Figure 8

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-2017-5929CCBYNCND

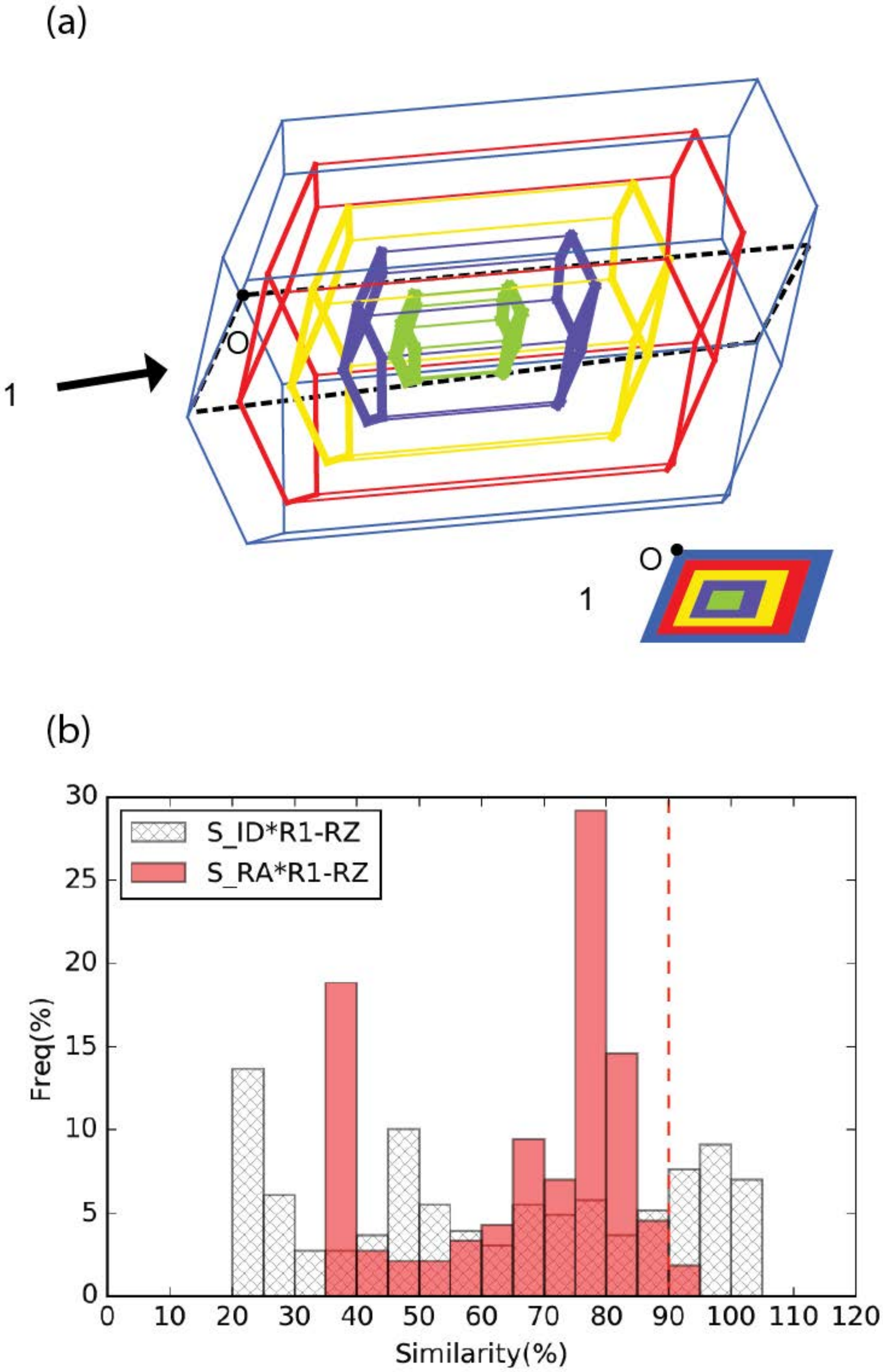

Figure 9

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-2017-5929CCBYNCND

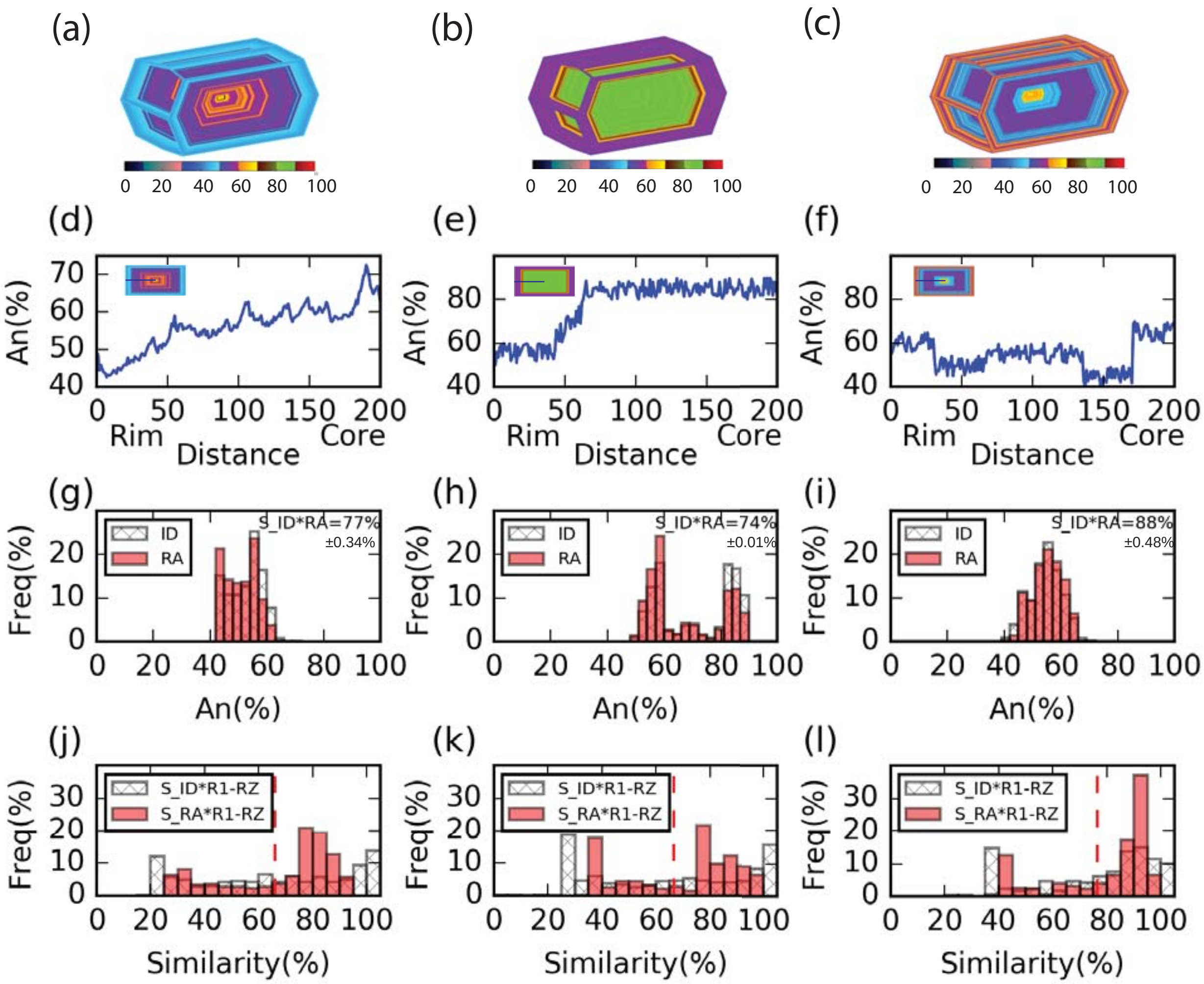

Figure 10

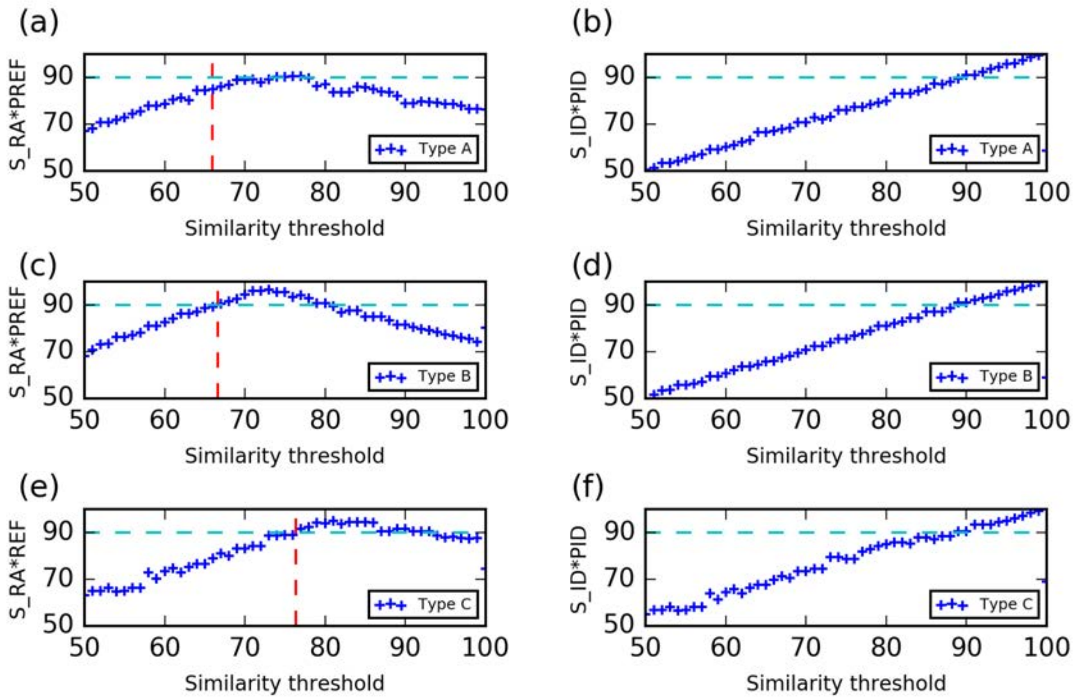
(a)
(b)
(c)
(d)
(e)
(f)
S_RA*PREF
S_ID*PID
S_RA*REF
Type A
Type B
Type C
Similarity threshold
50
60
70
80
90
100

Figure 11

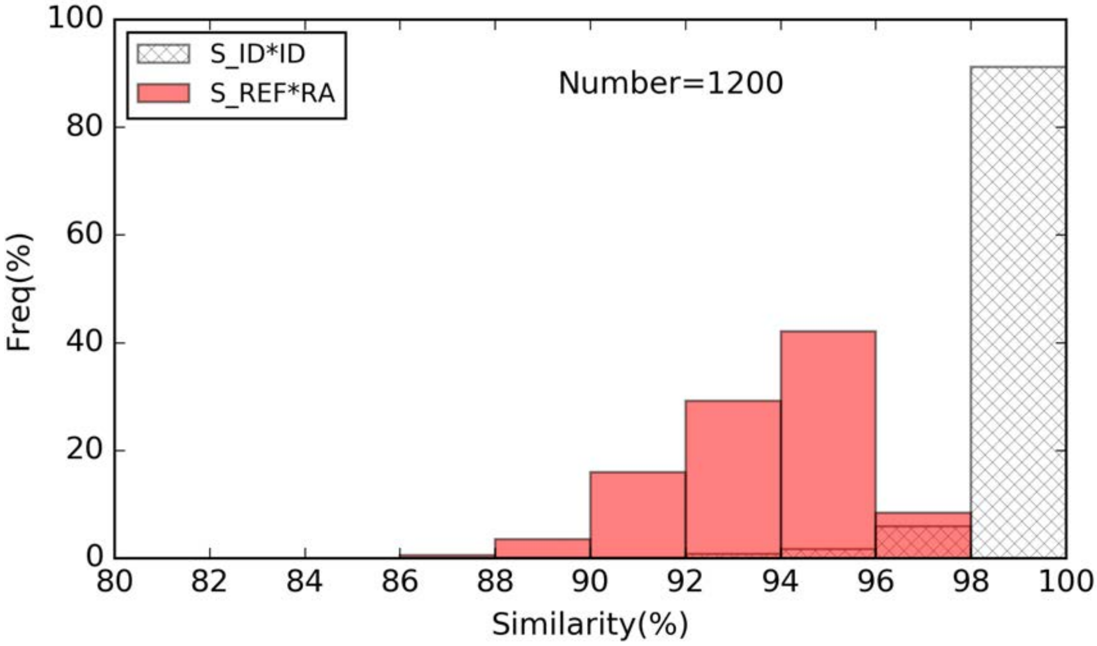
S_ID*ID
S_REF*RA
Number=1200
Freq(%)
Similarity(%)
0
20
40
60
80
100
80
82
84
86
88
90
92
94
96
98
100

Figure 12

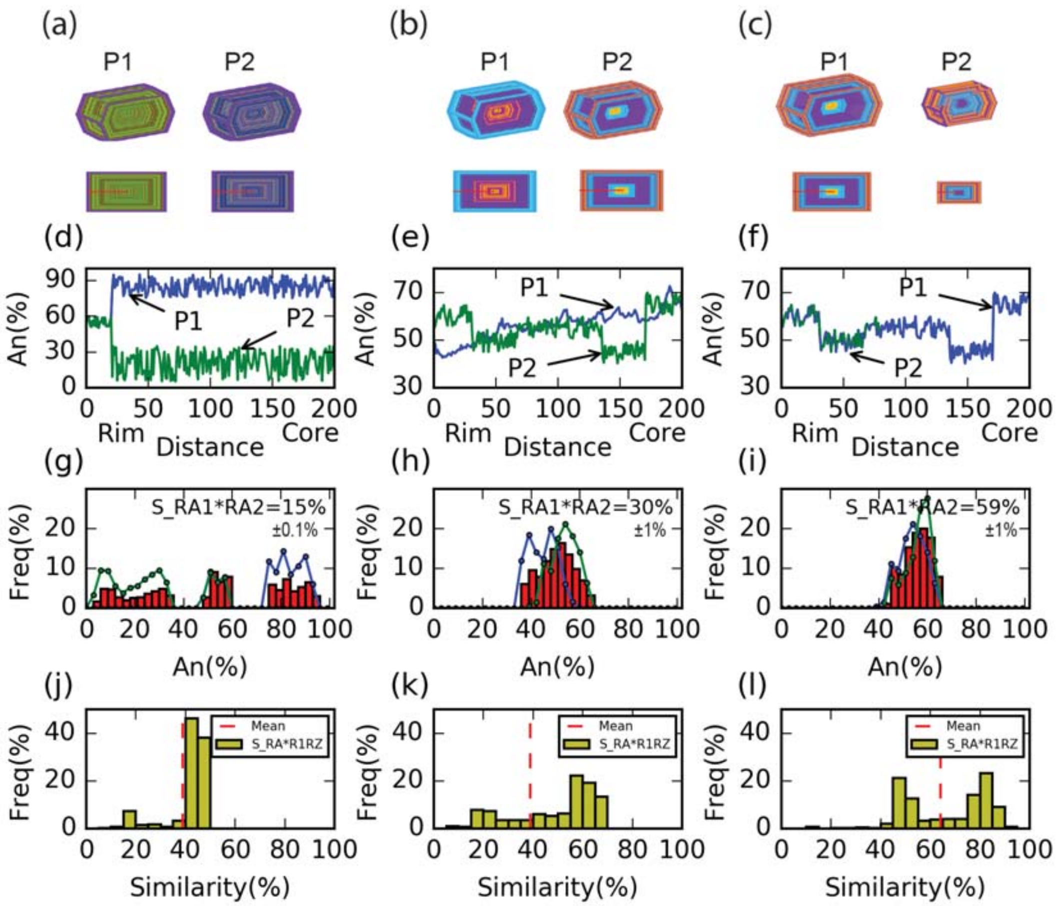
(a)
P1
P2
(b)
P1
P2
(c)
P1
P2
(d)
(e)
(f)
An(%)
P1
P2
Rim
Distance
Core
(g)
(h)
(i)
S_RA1*RA2=15%
±0.1%
S_RA1*RA2=30%
±1%
S_RA1*RA2=59%
±1%
Freq(%)
An(%)
(j)
(k)
(l)
Mean
S_RA*R1RZ
Similarity(%)

Figure 13

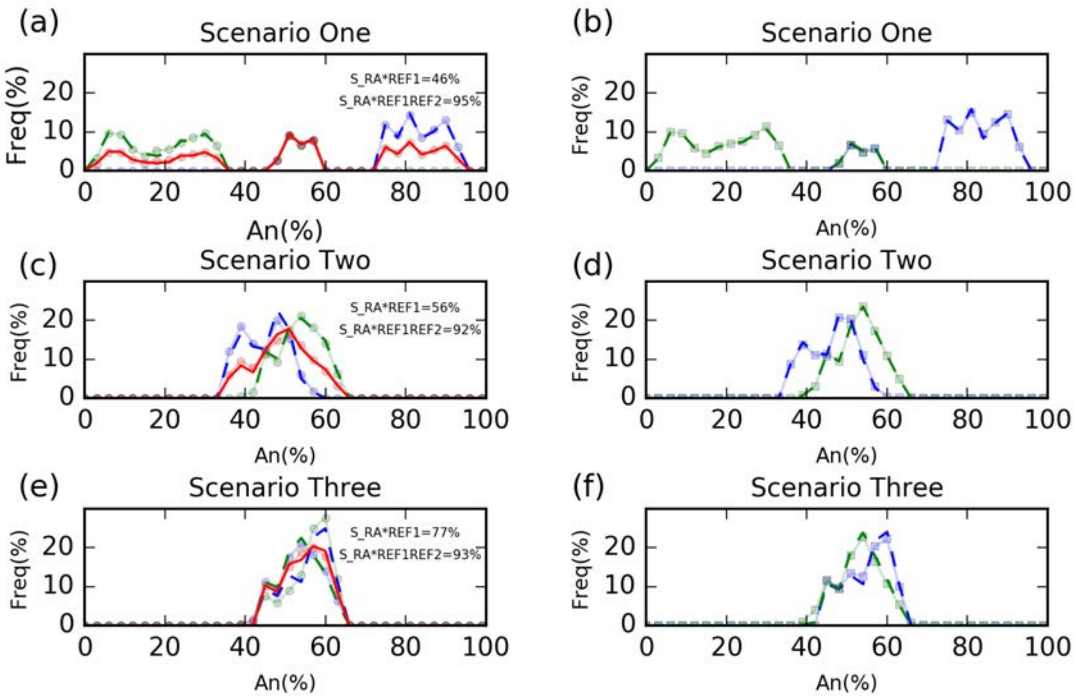
(a)
Scenario One
S_RA*REF1=46%
S_RA*REF1REF2=95%
Freq(%)
An(%)
(b)
Scenario One
Freq(%)
An(%)
(c)
Scenario Two
S_RA*REF1=56%
S_RA*REF1REF2=92%
Freq(%)
An(%)
(d)
Scenario Two
Freq(%)
An(%)
(e)
Scenario Three
S_RA*REF1=77%
S_RA*REF1REF2=93%
Freq(%)
An(%)
(f)
Scenario Three
Freq(%)
An(%)

Figure 14

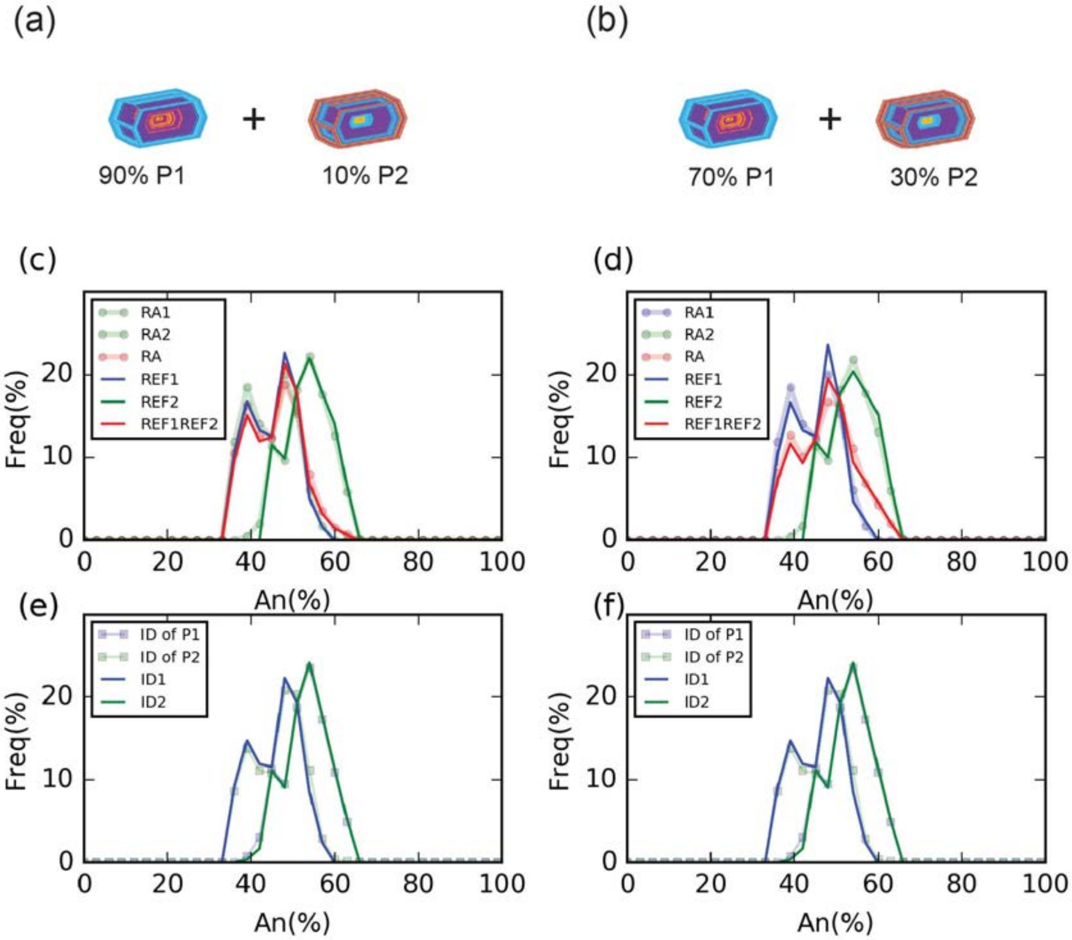
(a)
(b)
90% P1
10% P2
70% P1
30% P2
(c)
(d)
RA1
RA2
RA
REF1
REF2
REF1REF2
Freq(%)
An(%)
(e)
(f)
ID of P1
ID of P2
ID1
ID2

Figure 15

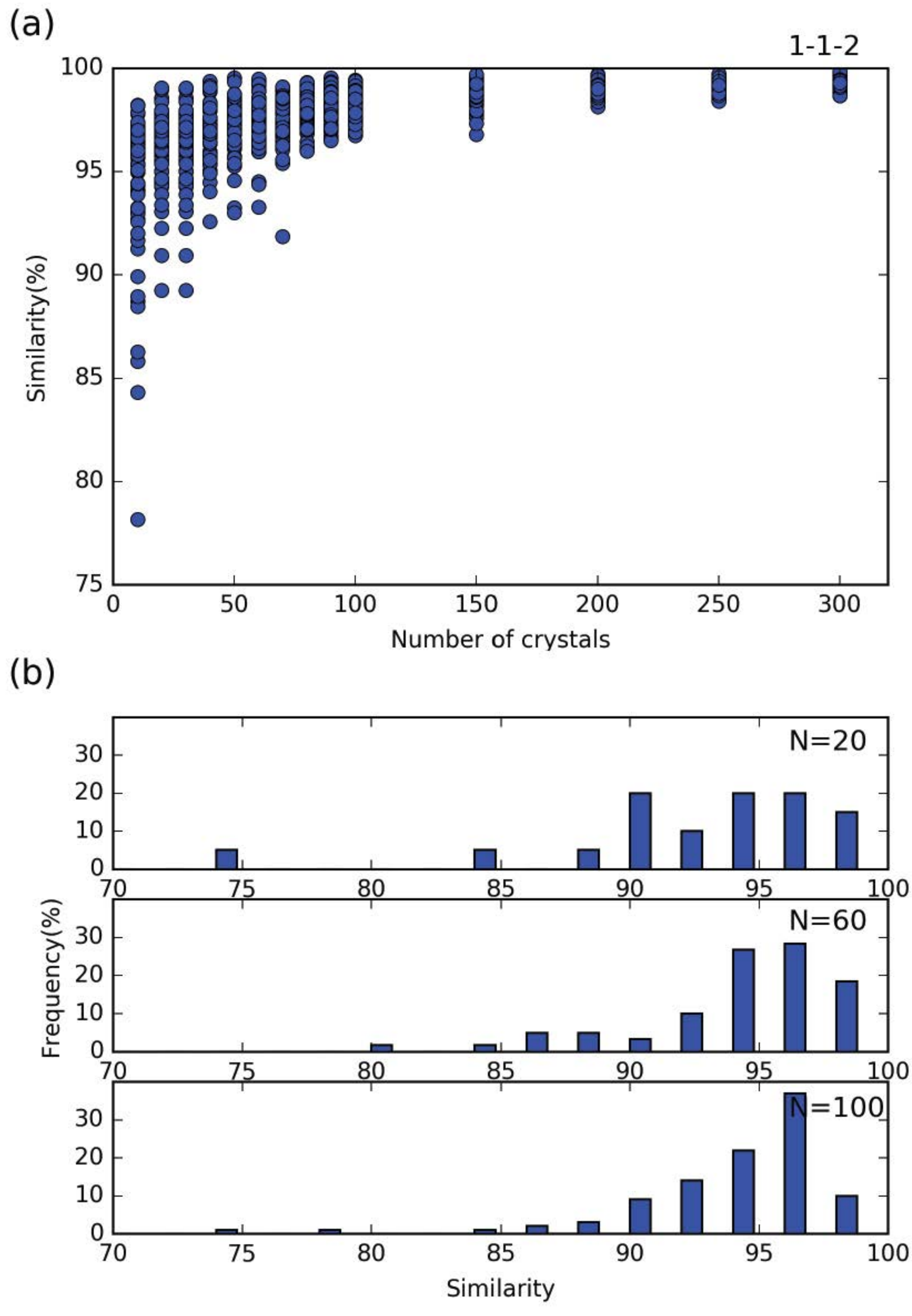
(a)
1-1-2
Similarity(%)
Number of crystals
(b)
N=20
N=60
N=100
Frequency(%)
Similarity

Figure 16