TECHNICAL ARTICLE

Interlaboratory study on the quantification of calcium phosphate phases by Rietveld refinement

Nicola Döbelin[a)]
RMS Foundation, Bischmattstrasse 12, 2544 Bettlach, Switzerland

(Received 25 November 2014; accepted 17 April 2015)

An interlaboratory study (ILS, round robin) was conducted to assess the accuracy and precision of the phase quantification of calcium phosphate (CaP) bioceramics by X-ray diffraction (XRD) and Rietveld refinement. For that purpose, a mixture of hydroxyapatite and β-tricalcium phosphate, two CaP phases commonly used in synthetic bone graft substitutes, was prepared and sent to 12 laboratories for XRD analysis. Results from 26 different instruments were received and evaluated statistically according to ASTM E691 – 13. The statistical analysis revealed that the reproducibility standard deviation of phase quantities was approximately two times greater than the repeatability standard deviation, which is obtained by repeating the analysis on a single instrument configuration multiple times. The 95% reproducibility limit for phase quantities was $R = \pm 1.67$ wt%. The study also demonstrated that several participants overinterpreted their data in an attempt to refine crystallite sizes of the minor phase. © *2015 International Centre for Diffraction Data.*
[doi:10.1017/S088571561500038X]

Key words: phase quantification, Rietveld refinement, calcium phosphate, interlaboratory study, repeatability, reproducibility

I. INTRODUCTION

Synthetic calcium phosphate ceramics (CaP) are widely used for the treatment of large bone defects because of their chemical similarity to the mineral part of bone (Döbelin *et al.*, 2010). Besides the various compositions of apatite [hydroxyapatite (HA), carbonate apatite, chloroapatite, and fluorapatite], which is the crystalline phase of bone mineral, a large number of other CaP phases exist in the $CaO–P_2O_5$ system (Welch and Gutt, 1961; Elliott, 1994). The chemical variation among these CaP phases is mostly found in the Ca:P molar ratio and the presence or absence of H^+, OH^-, or H_2O groups. However, the physico-chemical properties of the different phases such as thermodynamic (Chow, 2001; Raynaud *et al.*, 2002a) and kinetic (Dorozhkin, 2002) characteristics, or the reactivity with H_2O (Dorozhkin, 2011), can differ vastly. Some of these properties are of particular interest for medical application of CaP materials, for example as a substitute for bone grafts (Bohner *et al.*, 2012), because they determine the biocompatibility and resorption characteristics of a CaP device after implantation in the human body. Despite the minute chemical variations, major differences in crystal structure topologies are usually evident, which in turn means that CaP phases can be relatively easily identified by their crystal structure. For that reason powder X-ray diffraction (XRD) is one of the most common techniques to identify and quantify CaP phases or mixtures thereof.

Crystalline phases are identified from XRD patterns by comparing the angle of diffraction peaks with reference positions stored in a database. However, much more information on the crystal structures and phase quantities can be obtained by analysing the pattern with Rietveld refinement (Rietveld, 1969; Dinnebier and Billinge, 2008; Madsen and Scarlett, 2008), a process iteratively matching the calculated diffraction pattern of a set of crystal structure models to a measured pattern. This technique allows obtaining precise phase quantities, unit-cell dimensions, and crystallite sizes, and sometimes also atomic coordinates, ionic substitutions, stress and strain, and thermal vibration parameters. Rietveld refinement marks the current state-of-the-art of powder XRD data processing. However, as many complex analytical techniques, it requires skilled personnel, well-calibrated equipment, and very careful sample preparation. In practical use, the operators processing XRD data with Rietveld refinement are faced with two major problems: (1) Owing to the iterative data processing, a common problem is to determine the quality of a refinement and to recognize whether a refinement has converged or allows further optimization. If a refinement is considered complete even though more information could be extracted from the data, then the dataset is underprocessed. (2) On the other hand, refining more parameters than the data quality permits may lead to arbitrary results or unstable refinements. In that case, the data are overprocessed. The progress and convergence of a Rietveld refinement employing a strategy with a reasonable number of free parameters on datasets obtained from highly crystalline phases can usually be followed by goodness-of-fit parameters and the straightening difference curve. However, in case of problematic samples with a large number of phases, severe peak broadening, texture, an amorphous signal in the background, or fluorescence, even experienced users may struggle to determine a robust refinement strategy and thus find themselves in doubt of the validity of the results.

While experienced users often learned to spot and avoid data over- and underprocessing, less experienced users

a)Author to whom correspondence should be addressed. Electronic mail: nicola.doebelin@rms-foundation.ch

would greatly benefit from a complementary technique to verify their refinement results. However, several other techniques may provide similar results, but the differences are important enough to render direct comparisons impossible. Solid-state vibrational [Fourier transform infrared spectroscopy (FTIR), Raman] and absorption spectroscopies (NMR) provide short-range order information, but no long-range structural order. Phase quantifications in the presence of X-ray amorphous or nanocrystalline phases can therefore differ vastly from powder XRD results. Electron backscatter diffraction/selected area electron diffraction (EBSD/SAED) and single-crystal XRD do not interact with large sample volumes and cannot be used to quantify phases in multi-phase powder samples. Neutron diffraction is often not accessible for small institutions or independent laboratories. And last but not least none of the chemical analyses [X-ray fluorescence (XRF), energy-dispersive X-ray spectroscopy (EDX), inductive coupled plasma (ICP), AAS, and others] provide structural information and can be used to identify polymorphs.

XRD has become one of the standard tools in the field of CaP biomaterials research because of its ready availability and its potential to discriminate CaP phases even in complex mixtures. The majority of publications in this field of research feature phase identifications from peak positions. However, in the past decade the number of publications using Rietveld refinement for quantitative phase analysis has been constantly increasing. Several groups specifically addressed the accuracy of Rietveld refinement for CaP phase analysis and compared it to reference mixtures and conventional quantification methods. A significant step towards higher accuracy in HA and β-tricalcium phosphate (β-TCP) phase quantification in biphasic samples was already achieved by the use of a digital diffractometer instead of a powder camera, as shown by Toth *et al.* (1991), who improved the quantification based on reference intensity ratio (RIR) calibration curves from a 20 wt% range to a 3 wt% range by a change of instrument. Keller (1995) determined the HA and β-TCP phase content of commercially used binary mixtures by RIR quantification (single-peak analysis) and Rietveld refinement. The calibration curve for RIR quantification was established beforehand for the two strongest peaks HA 211 and β-TCP 20–10. HA phase quantities determined by Rietveld refinement were systematically higher by 2 ± 1 wt% than the results determined by RIR quantification. The author pointed out that single-peak-based analyses are inherently sensitive to insufficiently ground samples and textured sample surfaces, whereas the full-pattern-based approach of Rietveld refinement reduces the weighting of individual peak errors and allows for modelling of texture. In a similar study Raynaud *et al.* (2001) prepared reference mixtures to establish a calibration for phase quantification by RIRs for biphasic mixtures in the systems β-TCP–HA and HA–CaO. They determined a maximum absolute uncertainty for β-TCP and HA phase quantities of 3.5 wt% for mixtures of 50:50 wt%. Uncertainties for other mixtures were generally lower, reaching 0.5 wt% for nominal β-TCP and HA fractions of 5 wt%, respectively. Rietveld refinement was performed for comparison, and a systematic overinterpretation of β-TCP up to 6 wt% was observed in several compositions. Better agreement between reference mixtures and Rietveld refinements was observed by Reid and Hendry (2006). In a series of reference samples containing three CaP phases (α-TCP, β-TCP, and HA) a maximum difference between nominal and refined phase quantities of 1.4 wt% was determined.

Biphasic materials containing stoichiometric HA and β-TCP can either be mixed from pure components, or prepared by thermal treatment of Ca-deficient HA (CDHA, $Ca_{10-x}(HPO_4)_x(PO_4)_{6-x}(OH)_{2-x}$). Mixing two components gives full control over the bulk phase composition, but it bears the risk of inhomogeneous blending leading to a residual risk for phase fluctuations among individually sampled specimens. On the other hand, thermal treatment of CDHA is a simple but effective way to obtain materials with a homogeneous HA and β-TCP phase distribution, though the material usually reveals its precise phase composition only after the calcination because of the extraordinary sensitivity to variations of the bulk molar Ca:P ratio. The process of CDHA decomposition above a critical temperature has been described many times in the past (Riboud, 1973; Ishikawa *et al.*, 1993; Raynaud *et al.*, 2001; 2002a; 2002b; Wilson *et al.*, 2005; Nilen and Richter, 2008; Tõnsuaadu *et al.*, 2012). The onset of the separation was reported to occur at 700 °C (Nilen and Richter, 2008), 800 (Jackson *et al.*, 2004), 850 (Tõnsuaadu *et al.*, 2012), and 900 °C (Wilson *et al.*, 2005). According to Tõnsuaadu *et al.* (2012), dehydroxylation of HA begins at 850–900 °C, the precise temperature depending on the humidity of the atmosphere. Gopal and Calvo (1972) presented evidence that the naturally occurring mineral whitlockite transforms into the structure of β-TCP upon heating at or above 900 °C. The β-TCP phase formed by decomposition of CDHA at 900 °C or above is therefore expected to exhibit space group $R3c$ of the ideal β-TCP structure for β-$Ca_3(PO_4)_2$, both in absence (Dickens *et al.*, 1974) and presence of Mg contamination (Schroeder *et al.*, 1977).

As Rietveld refinement becomes more widely accepted and used by research groups with less expertise in XRD in general and Rietveld refinement in particular than the studies described above, the question of how accurate and precise refinement results of such groups are arises. Interlaboratory studies (ILS, round robin) are a common strategy to validate results from analytical procedures lacking a complementary method for verification. In the present study, we report the results of an ILS for phase quantification of CaP powders by powder XRD and Rietveld refinement, in which 12 laboratories from six European countries participated and contributed datasets from 26 different instruments. The ILS and statistical data evaluation were conducted strictly according to ASTM E691 – 13 (2013).

II. REFERENCE MATERIAL

A. Description

The reference sample contained two high-temperature phases commonly used in bone graft substitutes: HA [$Ca_5(PO_4)_3OH$] and β-TCP [β-$Ca_3(PO_4)_2$] (Table I). HA impurities in β-TCP samples and *vice versa* are realistic and common scenarios, which are caused by minor mismatches of the molar Ca:P ratio, and therefore particularly addressed in international standards specifying the phase purity of CaP bone graft substitutes (ASTM F1185 – 03, 2003; ASTM F1088 – 04, 2004; ISO 13779-3, 2008; ASTM F2024 – 10, 2010; ISO 13175-3, 2012). Based on published information on the thermal stability of CDHA and its decomposition into stoichiometric HA and β-TCP at elevated temperatures, it was decided to prepare a reference material by thermal treatment of a CDHA precursor at 900 °C. This procedure was seen as a compromise between

https://doi.org/10.1017/S088571561500038X

TABLE I. Crystallographic information for the HA and β-TCP structure models disclosed to the participants at the beginning of the study.

	Hydroxyapatite	β-Tricalcium phosphate
Abbreviation	HA	β-TCP
Chemical formula	$Ca_5(PO_4)_3OH$	$Ca_3(PO_4)_2$
Reference	PDF# 01-074-0565 (Sudarsanan and Young, 1969; ICDD, 2013)	PDF# 04-008-8714 (Dickens *et al.*, 1974; ICDD, 2013)
Space group (number)	$P6_3/m$ (176)	$R3c$ (161)
a (Å)	9.424 (4)	10.439 (1)
c (Å)	6.879 (4)	37.375 (6)
Structural density (g cm^{-3})	3.15	3.07
Z	2	21

inducing separation into stoichiometric phases, recrystallization, annealing, and a moderate amount of grain growth, without increasing the risk for HA dehydroxylation and excessive grain growth. Larger grains and sintering of agglomerates would have made intense milling a necessity for XRD sample preparation and thus made the samples prone to the formation of amorphous fractions, or microabsorption effects biasing the result of phase quantifications. On the other hand, the as-obtained reference material was essentially free of preferred orientation, microabsorption effects, micro strain, amorphous content, and nanocrystalline particles.

B. Preparation and characterization

Approximately 200 g of a commercial Ph. Eur. grade CDHA powder (Calipharm T 118 BP, Item No. 12714-330, Batch B00217A, Brenntag Schweizerhall AG, Switzerland) were annealed by thermal treatment for 12 h at 900 °C in a sinter furnace (CDF 15/1B, Carbolite, Hope Valley, England), which initiated decomposition into stoichiometric β-TCP and HA. Since the two primary phases were not mixed from pure materials but formed by thermal decomposition of a homogeneous single-phase precursor, a precise true value for the phase composition was not known at the time of powder preparation. However, the preparation route had the advantage of very homogeneous phase distribution, as the two phases formed *in situ* homogeneously throughout the bulk of the material, and particle separation or incomplete homogenization could be expected to be negligible.

Scanning electron microscopy (SEM) images showed agglomerates of several micrometres in diameter (Figure 1). Despite considerable fusion caused by the thermal treatment, it was still evident that the agglomerates were formed by sintered primary particles of less than 1 μm in diameter. The agglomerates were of irregular shape and their primary particles were randomly oriented. The micrographs suggested that no tendency for preferred orientation was to be expected from this material. The crystallite size, which can only be smaller or equal to the primary particle size, was well below the critical size of approximately 5 μm that is required for non-spotty diffraction patterns (McCusker *et al.*, 1999). In this size range microabsorption effects are largely irrelevant, particularly

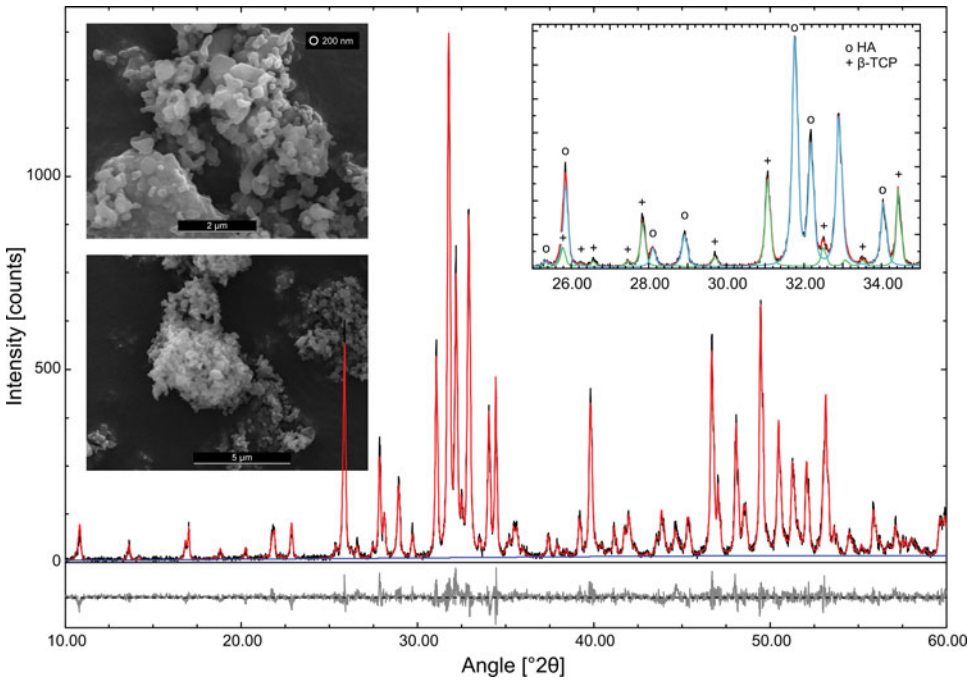

Figure 1. (Color online) Example refinement showing the overall complexity of the diffraction pattern. The enlarged section from 25° to 35° 2θ shows the contributions of β-TCP and HA in the range of strongest intensity, featuring the 100% β-TCP (0 2 10) peak at 31.02° 2θ and the 100% HA (2 1 1) peak at 31.77° 2θ. SEM image inlays show the particle and agglomerate morphology of the reference powder.

TABLE II. Results of the trace element analysis. All values are given in ppm. Values below 2.0 ppm were omitted.

Element	Concentration	Element	Concentration	Element	Concentration
Ca	214 500	Si	913.0	Mn	62.4
P	456 700	K	15.2	Fe	180.9
B	36.2	Mo	4.6	Zr	2.1
Na	392.5	Ti	24.3	Ni	3.6
Mg	1353.8	V	3.7	Zn	5.7
Al	292.8	Cr	9.9	Sr	267.0

among phases with similar linear absorption coefficients. A moderate amount of diffraction peak broadening because of small crystallite size was expected, though.

The particle size distribution measured by laser diffraction (Beckman Coulter LS 13 320, Krefeld, Germany) after ultrasonication in ethanol confirmed the results observed on SEM images. The mean particle size by number was 204 nm and the median was 181 nm. A low number (but large volume) of particles between 1 and 20 μm was observed. Based on SEM images, this signal was interpreted as agglomerates not completely disintegrated by ultrasonication. These chunks of randomly oriented primary particles were unlikely to cause spottiness, orientation effects, or phase inhomogeneities in XRD patterns because of the random arrangement of primary particles within the agglomerates.

Trace elements were identified and quantified by ICP–mass spectroscopy (ICP–MS; Perkin Elmer ELAN 6000, USA) on supernatant samples diluted 1:100 in 2% HNO_3 solution (HNO_3, 70% purified by distillation, No 225711, Sigma Aldrich, Switzerland), calibrated using a global standard (No IV-ICPMS-71A, Inorganic ventures, USA). The main elements Ca and P were quantified in a second series using 1:10 000 dilution and standards for Ca (No 19051, Sigma Aldrich, Switzerland) and P (No 38338, Sigma Aldrich, Switzerland, 38338). Results of the analysis are shown in Table II. The strongest contamination was found to be Mg (1353.8 ppm), which is an impurity commonly observed in CaP phases. The compatibility of Mg with the HA lattice is relatively low (Neuman and Mulryan, 1971), while at the same time it is high with the β-TCP lattice (Schroeder *et al.*, 1977; Enderle *et al.*, 2005). The Mg impurity could therefore be expected to mainly accumulate in the β-TCP phase and substitute up to 2.05% of all Ca atoms. According to Enderle *et al.* (2005) this amount of substitution was expected to lead to shrinkage of the β-TCP unit cell by 0.22% (*a*-axis) and 0.11 % (*c*-axis), respectively. The bulk molar Ca:P ratio calculated from the elemental analysis was 1.657 (±0.061). Calculating an expected phase ratio and propagating the standard deviation to phase quantities resulted in a standard deviation of 34 wt% for β-TCP, and 41 wt% for HA phase content. Chemical quantification of Ca and P was therefore not precise enough to verify the phase content determined by XRD because of the extraordinary sensitivity of the β-TCP:HA phase ratio for the bulk molar Ca:P ratio.

III. PROTOCOL AND PARTICIPANTS

To obtain comparable results, the participants were instructed (i) not to apply any mechanical treatment such as milling or sieving to the reference powder prior to data collection, and (ii) not to use any sample preparation or instrument setting specifically optimized for the ILS. Statistical results were supposed to represent standard measurement conditions rather than best-case scenarios. Five datasets from independently prepared samples were requested. The main phases β-TCP and HA were disclosed prior to data collection, and structural models for both phases (Sudarsanan and Young, 1969; Dickens *et al.*, 1974) were suggested, but not required to use. This information was provided for the participants' convenience and to allow those less familiar with the CaP phase system to focus on the refinement procedure rather than on phase identification. But all participants were free and encouraged to use whatever structure model they preferred, and to employ whatever refinement strategy they considered appropriate. They were asked to report values for phase quantities, unit-cell dimensions, and crystallite sizes determined by Rietveld refinement from all five datasets. Anonymous codes were assigned to each set of results, using capital letters *A* to *L* for the laboratory, and Roman numerals I to IV for instruments, as many laboratories provided datasets from more than one instrument.

IV. STATISTICAL ANALYSIS

The statistical evaluation of all refined parameters was conducted according to ASTM E691 – 13 (2013). In a first step, average values and standard deviations were calculated for each configuration from the five individual reported values. In accordance with ASTM E691 – 13, these results were called the "cell average" $\bar{x}$, and "cell standard deviation" s, with the term "cell" being equivalent to "instrument" in this study. The average calculated from all cell averages $\bar{x}$ was called the "average of cell averages" $\bar{\bar{x}}$, and the standard deviation of all $\bar{x}$ scattered around was called the "standard deviation of cell averages" $s_{\bar{x}}$. The latter were calculated from laboratory instrument datasets only. The datasets submitted by laboratories *K* and *L* using synchrotron radiation were ignored to obtain precision statements relevant for laboratory instruments. Calculations of repeatability standard deviation s_r and reproducibility standard deviation s_R, as well as 95% repeatability and reproducibility limits r and R were performed according to ASTM E691 – 13 (2013) and ASTM E177 – 13 (2013). In addition to the statistical parameters specified in ASTM E691 – 13, we also present the average standard deviation $\bar{s}$, which is the mean value of all cell standard deviations s.

To interpret the cell averages $\bar{x}$ and cell standard deviations s for each instrument configuration in the context of all contributed datasets, two consistency values were calculated: (1) the between-laboratory data consistency h; and (2) the

https://doi.org/10.1017/S088571561500038X

within-laboratory data consistency k. Equations for both values can be found in ASTM E691 – 13. h consistency values express the ratio of the cell deviation $\bar{x} - \bar{\bar{x}}$ to the standard deviation of the cell averages $s_{\bar{x}}$. The meaning of a consistent h value is to be read as "a cell average $\bar{x}$ does not deviate from $\bar{\bar{x}}$ more than expected". Limits for consistency at the 0.5% significance level are tabulated in ASTM E691 – 13. k consistency values express the cell standard deviation s relative to the repeatability standard deviation s_r. The meaning of a consistent k value is to be read as "the data in one cell does not scatter more than expected". All equations and tabulated limits for k consistency at the 0.5% significance level are given in ASTM E691 – 13.

V. RESULTS

An example refinement of a dataset measured on a PANalytical CubiX^3 diffractometer (PANalytical, The Netherlands) using graphite-monochromated Cu$K\alpha$ radiation, and refined with BGMN (Bergmann *et al.*, 1998), is shown in Figure 1. The enlarged section shows the individual contributions of β-TCP and HA. The strong peaks of both phases were relatively well separated and only moderate peak broadening due to small crystallites was observed. Hence no excessive peak overlap was encountered.

Not all participants provided all requested parameters: (1) Laboratory *K* did not report refined unit-cell dimensions. Owing to the variable wavelength of synchrotron radiation, this participant refined the wavelength instead. (2) Laboratories *B* and *K* did not report crystallite sizes, because they refined the instrument resolution function instead. (3) Several laboratories only provided isotropic crystallite sizes for HA, whereas the majority of participants refined anisotropic HA crystallite sizes.

Some reported cell averages $\bar{x}$ or cell standard deviations s of the various refined parameters were clearly different from the majority of the reported values. However, none of these presumable outliers could be related to clerical, sampling, or procedural errors. Therefore, all values were used for the statistical calculations, and none were flagged as outliers. This decision was in accordance with the procedure laid out in ASTM E691 – 13 (2013). Repeatability standard deviation s_r, reproducibility standard deviation s_R, as well as 95% repeatability and reproducibility limits r and R are shown in Table III for all refined parameters.

A. Phase quantities

All participants refined two phases, β-TCP and HA, and normalized the sum to 100 wt%. No other phases were identified. Cell standard deviations s ranged from 0.05 to 0.83 wt% with a mean value of 0.22 and a median at 0.17 wt% (Figure 2). A *t*-test showed that many configurations reported phase quantities with a significant ($p < 0.05$) or highly significant ($p < 0.01$) systematic offset of $\bar{x}$ from $\bar{\bar{x}}$. The 95% reproducibility limit R was 1.67 wt% (Table III). Between-laboratory consistency h of all instrument configurations was below the critical value $h_{\text{critical}} = \pm 2.60$. Only configuration *A*:IV approached, but did not exceed, the limit. However, the excessively high cell standard deviation s observed with configurations *B*:I and *I*:I resulted in within-laboratory consistency k exceeding the critical value $k_{\text{critical}} = 1.88$ (Figure 3).

B. Unit-cell dimensions

Refined unit-cell parameters for HA and β-TCP are shown in Figure 4. Interestingly, the standard deviation of cell averages $s_{\bar{x}}$ of the β-TCP *c*-axis was 1.01×10^{-3} nm and hence clearly greater than $s_{\bar{x}}$ of all other axes, which ranged from 1.9×10^{-4} to 2.7×10^{-4} nm (Table III). Again, many configurations reported mean values $\bar{x}$ significantly ($p < 0.05$) or highly significantly ($p < 0.01$) different from the global average $\bar{\bar{x}}$. The origin of the excessive offset obtained from configurations *D*:I and *F*:I could not be identified. In case of configuration *K*:I; however, the unit-cell dimensions were not refined because of correlations with the refined wavelength of the synchrotron radiation.

Configuration *D*:I exceeded the critical value for between-laboratory consistency for all unit-cell dimensions (Figure 5), whereas *F*:I unit-cell dimensions were consistent. Within-laboratory consistency was exceeded by a small margin by configurations *C*:I, *C*:II, *I*:II, and *J*:II, and by a large margin by the synchrotron configuration *K*:I for reasons mentioned above.

C. Crystallite sizes

Three configurations out of 26 did not report refined crystallite sizes. Out of the other 23 configurations, 14 refined anisotropic and 9 refined isotropic HA crystallite sizes. For β-TCP only isotropic crystallite sizes were reported. Mean refined values are shown in Figure 6. The most prominent

TABLE III. Precision statistics of all refined parameters.

Parameter	$\bar{\bar{x}}$	$\bar{s}$	$s_{\bar{x}}$	s_r	s_R	r	R
Weight fraction β-TCP (wt%)	28.03	0.22	0.54	0.29	0.60	0.81	1.67
Weight fraction HA (wt%)	71.97	0.22	0.54	0.29	0.60	0.81	1.67
Unit-cell length β-TCP *a* (Å)	10.4099	0.000 45	0.0027	0.0005	0.0027	0.0015	0.0077
Unit-cell length β-TCP *c* (Å)	37.3561	0.002 10	0.0102	0.0024	0.0104	0.0066	0.0292
Unit-cell length HA *a* (Å)	9.4221	0.000 37	0.0026	0.0005	0.0026	0.0013	0.0073
Unit-cell length HA *c* (Å)	6.8845	0.000 28	0.0019	0.0003	0.0019	0.0009	0.0054
Crystallite size β-TCP (nm)	212.87	17.39	110.30	31.04	113.74	86.91	318.48
Crystallite size HA ⟨001⟩ (nm)	106.99	4.78	16.67	6.69	17.72	18.74	49.60
Crystallite size HA ⟨100⟩ (nm)	92.69	3.25	14.38	4.54	14.95	12.72	41.85

$\bar{\bar{x}}$, is the average of cell averages; $\bar{s}$, the average cell standard deviation; $s_{\bar{x}}$, the standard deviation of cell averages; s_r, the repeatability standard deviation; s_R, the reproducibility standard deviation; r, the 95% repeatability limit; and R, the 95% reproducibility limit.

https://doi.org/10.1017/S088571561500038X

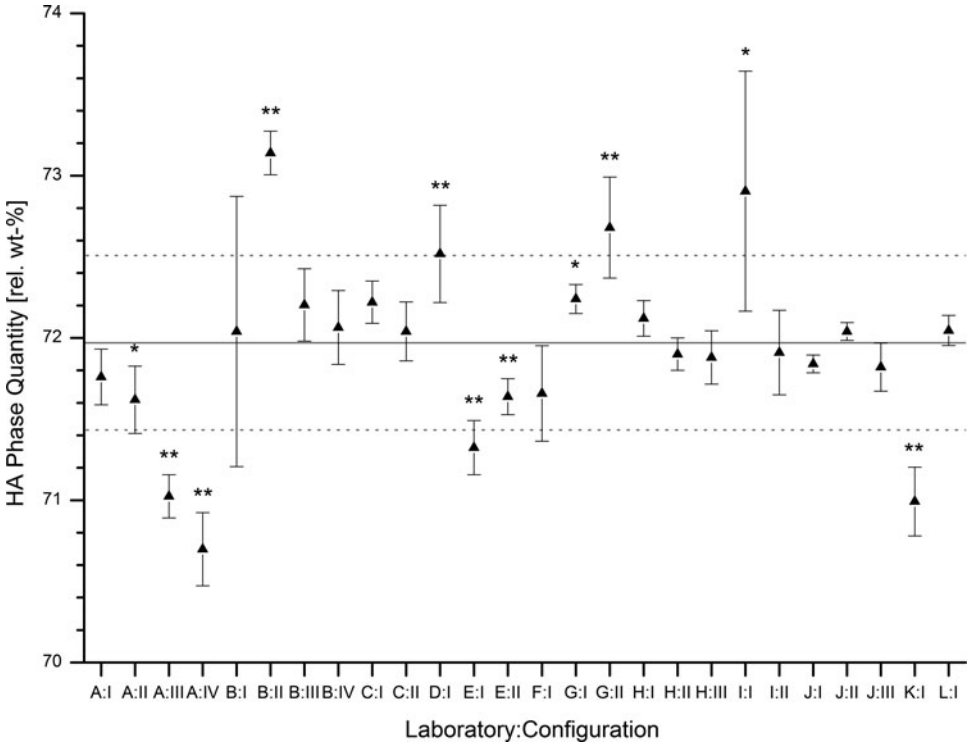

Figure 2. Cell average values $\bar{x}$ and standard deviations s of refined HA phase quantities. Horizontal lines mark the average of cell averages $\bar{\bar{x}}$ (solid line) and standard deviations of cell averages $s_{\bar{x}}$ (dotted lines). Cell average values significantly different from the average of cell averages are labelled with *($p < 0.05$) or ** ($p < 0.01$).

feature of this illustration is the excessively high scattering of refined β-TCP sizes reported by laboratories *B, D*, and *G*. The standard deviation of cell averages $s_{\bar{x}}$ was one order of magnitude greater than $s_{\bar{x}}$ of HA crystallite sizes in directions ⟨100⟩ and ⟨001⟩.

h consistency values for refined crystallite sizes expose two inconsistent results (Figure 7): Laboratory *B*:IV reported inconsistent β-TCP sizes, and laboratory *G*:I reported inconsistent sizes for HA in direction ⟨100⟩. All other laboratories reported crystallite sizes consistent with the average values.

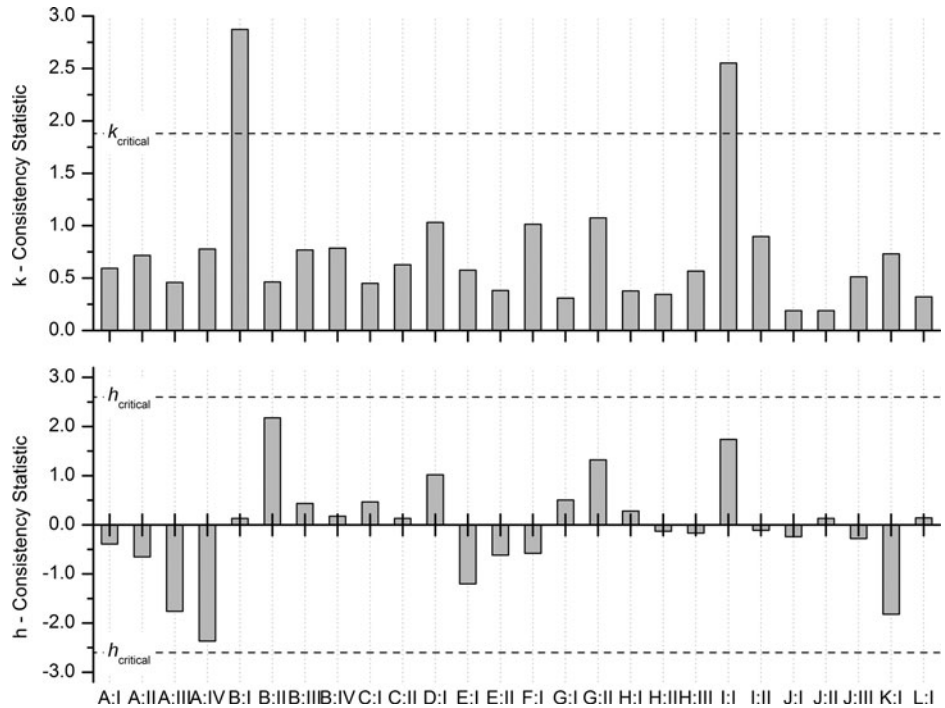

Figure 3. Within-laboratory consistency *k* and between-laboratory consistency *h* statistics for refined HA phase quantity. Owing to normalization of the phase quantities, β-TCP consistency *k* was identical and *h* was inverse to the corresponding HA consistency parameters.

https://doi.org/10.1017/S088571561500038X

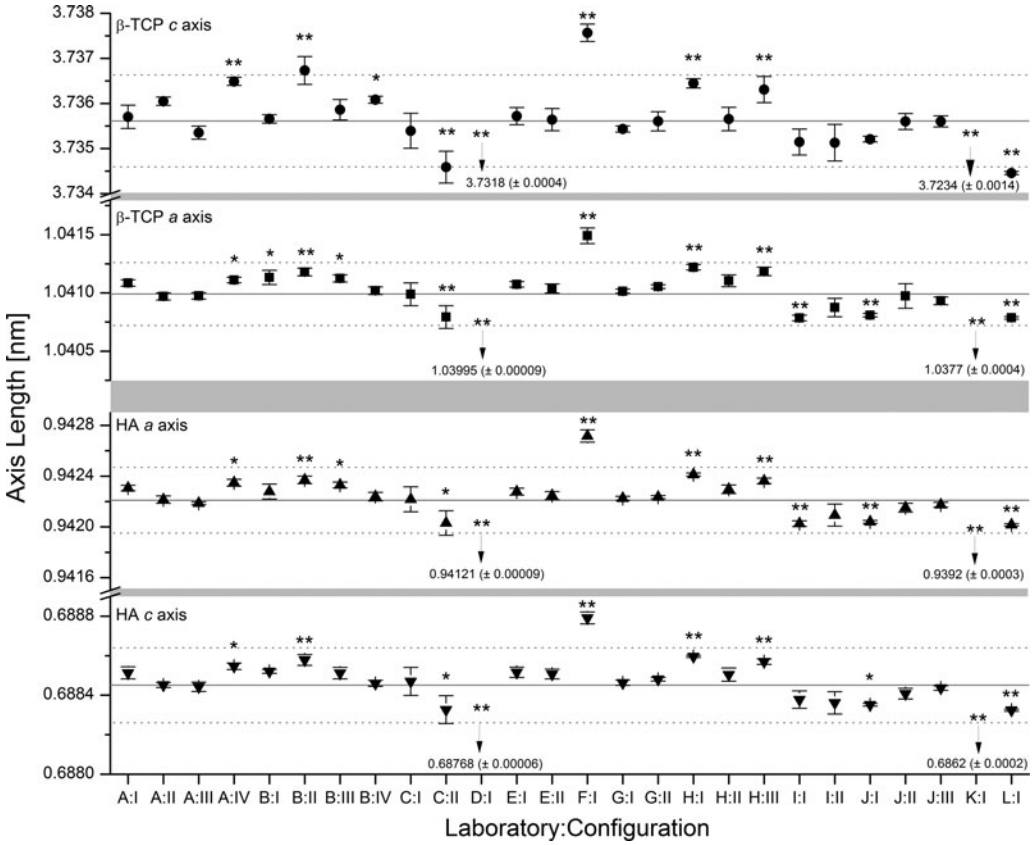

Figure 4. Cell average values $\bar{x}$ and standard deviations s of refined unit-cell dimensions. Horizontal lines mark the average of cell averages $\bar{\bar{x}}$ (solid line) and standard deviations of cell averages $s_{\bar{x}}$ (dotted lines). Cell average values significantly different from the average of cell averages are labelled with *($p < 0.05$) or **($p < 0.01$).

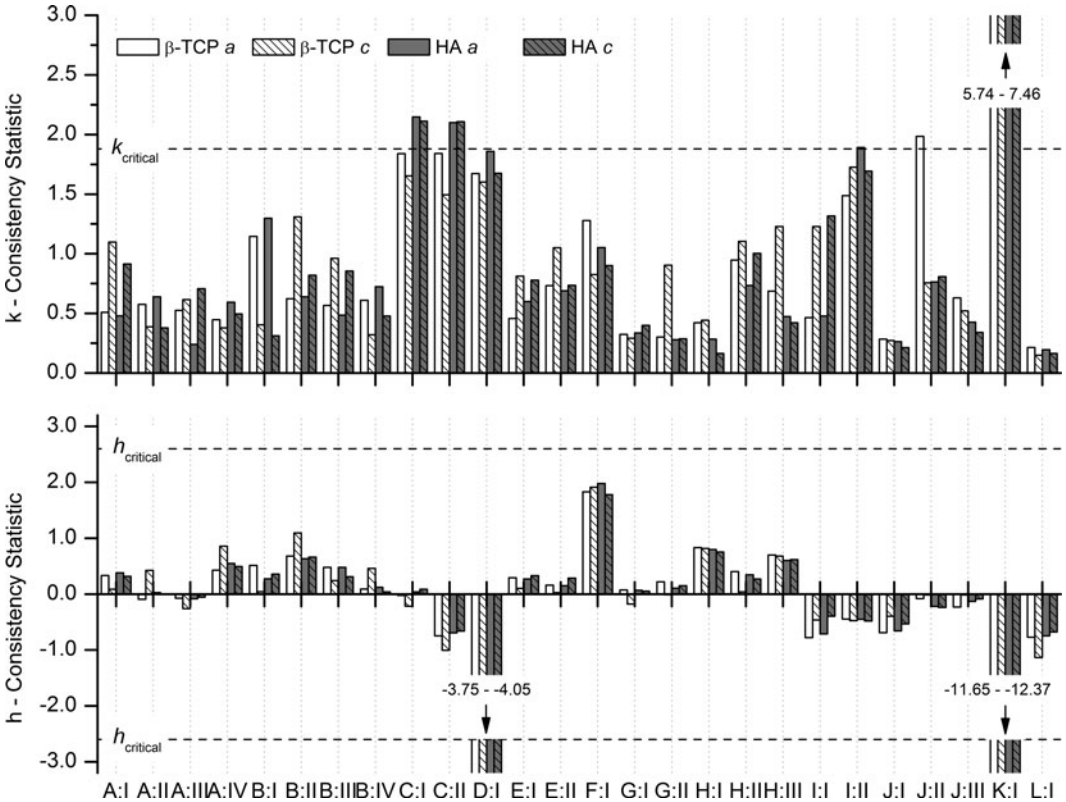

Figure 5. Within-laboratory consistency k and between-laboratory consistency h statistics for refined unit-cell dimensions.

https://doi.org/10.1017/S088571561500038X

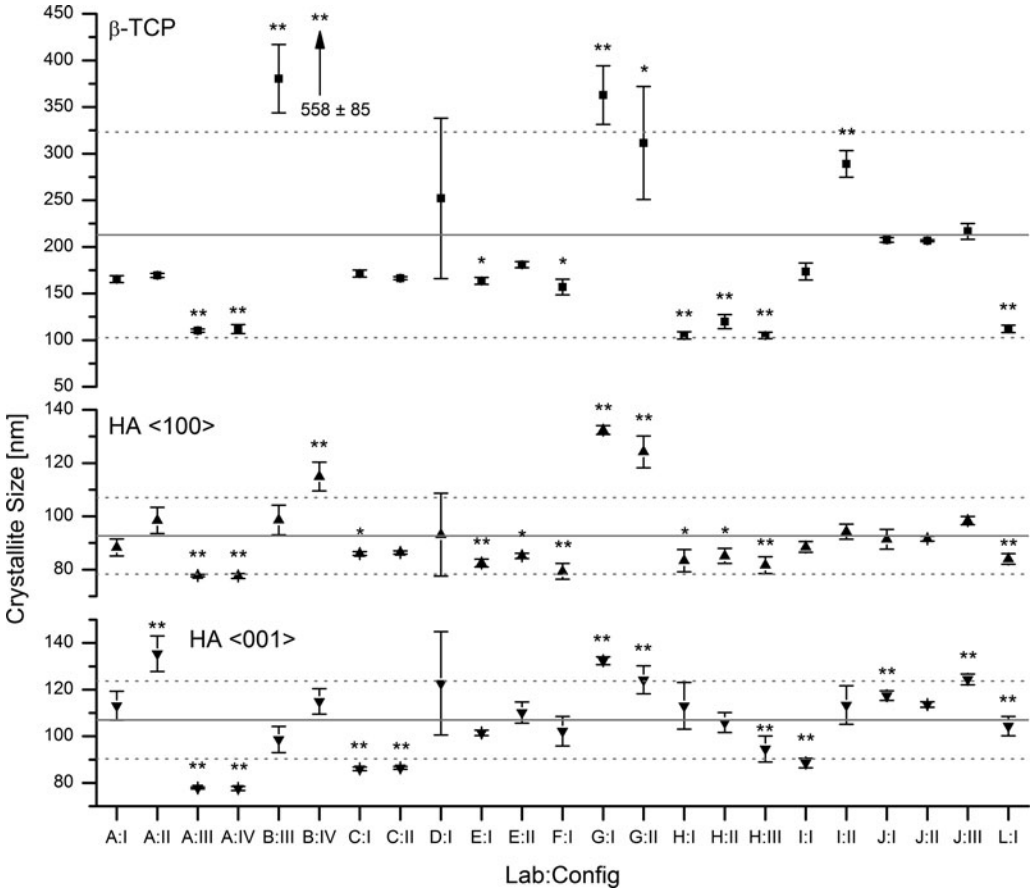

Figure 6. Refined crystallite sizes for β-TCP (isotropic, top), HA ⟨100⟩ (middle), and HA ⟨001⟩ (bottom). Horizontal lines mark the average of cell averages $\bar{\bar{x}}$ and standard deviations of cell averages $s_{\bar{x}}$ calculated from laboratory instruments *A* to *J*. No results were reported from configurations *B*:I, *B*:II, and *K*:I. Cell average values significantly different from the average of cell averages are labelled with *($p < 0.05$) or **($p < 0.01$).

Within-laboratory consistency values k exceeded the critical limit for the following configurations: *B*:IV, *D*:I, and *G*:II. All other configurations reported consistent k values for refined crystallite sizes.

VI. DISCUSSION

Most national and international standards specifying XRD and Rietveld refinement for the characterization of CaP biomaterials aim at identification and quantification of the crystalline phases. The precision estimate of refined phase quantities is therefore seen as the most relevant outcome of this study. The cell averages of the phase quantities $\bar{x}$ (Figure 2) reached a repeatability standard deviation $s_r =$ 0.29 wt%. However, the same figure also illustrates that 11 out of 26 datasets differed significantly from the average of cell averages $\bar{\bar{x}}$. It has to be assumed that these datasets are systematically biased and increasing the number of repetitions would not suffice to eliminate the offset from the true phase quantities. Phase quantifications by XRD and Rietveld refinement clearly exhibit a substantial risk for systematic bias of data measured and processed under repeatability conditions. When measured and processed under reproducibility conditions, the standard deviation of phase quantities $s_R = 0.60$ wt % was 2.07 times greater than the repeatability standard deviation s_r. This ratio between repeatability and reproducibility conditions was remarkably similar to the results published by Stutzman (2005), who conducted an ILS on phase quantification of hydraulic cements. Even though Stutzman's reference sample was a more complex mixture of nine phases and the standard deviations were generally higher, the ratio between reproducibility and repeatability standard deviation was 1.84 and thus close to the value observed in this study.

Refined unit-cell dimensions $\bar{\bar{x}}$ (Table III) of the HA phase showed less than 0.1 % deviation from the reference structure (Table I). On the other hand, the refined dimensions of β-TCP were 0.24% (*a*-axis) and 0.13% (*c*-axis) shorter than the reference values. This correlates well with the shrinkage expected for 2.05% Mg substitution (Enderle *et al.*, 2005) and confirms the previously mentioned assumption that Mg contamination (Table II) preferentially accumulates in the β-TCP phase. The shrinkage also explains a moderate shift of β-TCP diffraction peaks towards higher 2θ values (Figure 1). A statistical evaluation of refined unit-cell parameters exposed that the standard deviation of cell averages $s_{\bar{x}}$ was a constant fraction of the refined averages of cell averages $\bar{\bar{x}}$ (Table III). In other words, long cell dimensions resulted in greater standard deviations than short dimensions. The ratio $s_{\bar{x}}/\bar{\bar{x}}$ for all four refined axes was in the range from 0.026 to 0.028%. Lower absolute precision for larger unit-cell dimensions is not surprising, considering that large d values generate diffraction peaks at low 2θ angles. Datasets measured at constant sampling intervals in 2θ feature lower angular resolution at large d values because of the non-linear relation of d and 2θ. Low-angle resolution is further

https://doi.org/10.1017/S088571561500038X

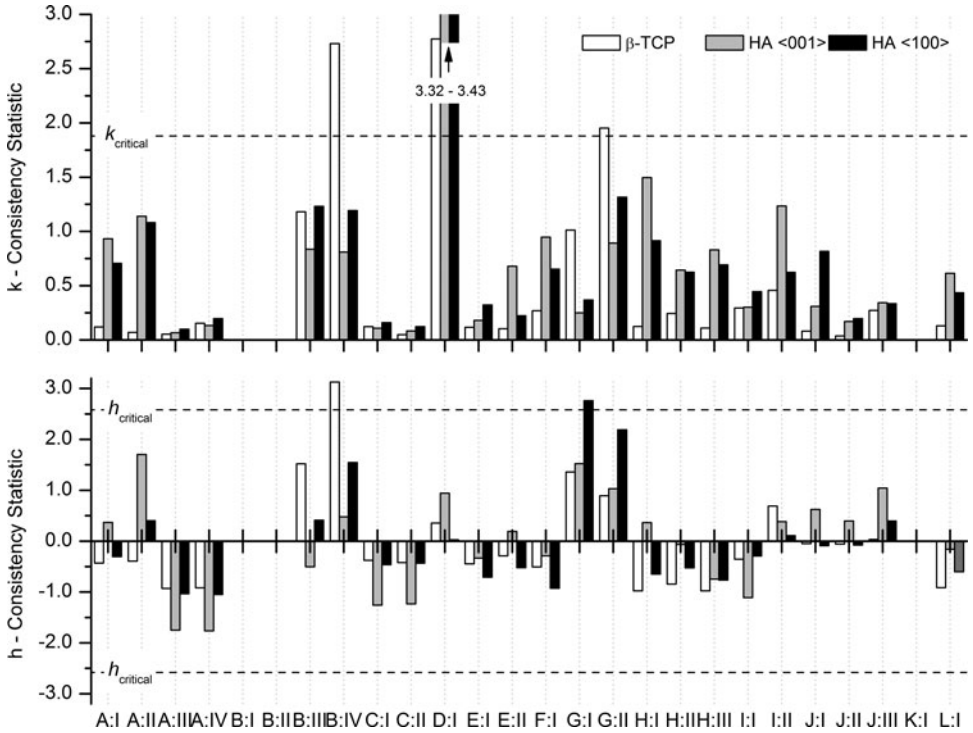

Figure 7. Within-laboratory consistency k and between-laboratory consistency h statistics for refined crystallite sizes.

impaired by increased peak asymmetry and full width at half maximum (FWHM) towards low 2θ angles observed with most instruments.

Crystallite sizes were clearly the most critical parameters refined in this study, as numerous factors contribute to systematic errors of the results: a very precise mathematical description of the instrument resolution function is required to accurately deconvolute the crystallite size broadening effect from the instrument broadening. For software based on the fundamental parameters approach, the fundamental parameters must precisely describe the optical elements of the instrument, including all misalignments. At the level of refinement strategy, surface roughness, sample transparency, and micro strain may be misinterpreted as crystallite size-related peak broadening to a certain degree. And last but not least the complexity of the diffraction pattern and the phase abundance determine whether the peak shape of a specific phase is defined well enough to allow anisotropic, isotropic, or no refinement of the crystallite size. Refining crystallite sizes despite insufficient data quality may lead to unstable refinements and random results.

The striking difference in data scattering between β-TCP and HA crystallite sizes (Table III, Figure 6) raised the question whether β-TCP crystallite sizes submitted by laboratories *B, D*, and *G* should be regarded as outliers and consequently be excluded from the precision estimate, as ASTM E 691–13 (2013) recommends for erroneous data. However, careful examination of the raw and refined data did not expose any clerical, sampling, or procedural errors. The inaccurate results seemed rather to be related to raw data quality (signal-to-noise ratio, peak resolution, and step size) resulting in more peak overlap and less defined β-TCP peak shapes, and to the refinement strategy. As the choice of the refinement strategy is an integral part of Rietveld refinement, excluding the datasets in question from the statistical evaluation would have excessively sanitized the precision of crystallite sizes in an unjustified manner. Hence none of the reported crystallite size data were eliminated from the precision estimate.

Precipitated HA crystals often grow in anisotropic shapes relative to the crystallographic *c*-axis (Neira *et al.*, 2009), which results in different amounts of peak broadening for direction ⟨001⟩ compared with all other crystallographic directions. Relative phase quantities are calculated from all phases' scale factors, unit-cell volumes, and the total masses of the unit-cell contents after convergence of the refinement (Dinnebier and Billinge, 2008). These parameters are predominantly sensitive to accurate fits of peak positions and integrated intensities, but less so to the shape of peaks. One might thus be tempted to assume that poor fits of anisotropic HA peaks widths would not bias phase quantifications, provided the integrated intensities were refined accurately. But this would ignore the fact that anisotropic crystal shapes often introduce preferred orientation effects which can be reduced, but not entirely eliminated, by careful sample preparation. Quite often refinement of anisotropic peak broadening for directions ⟨100⟩ and ⟨001⟩ as well as preferred orientation is therefore necessary to obtain accurate fits of HA 00*l* peaks, hence accurate calculations of integrated peak intensities, and ultimately accurate phase quantities. However, in our reference material moderate heat treatment and sintering effects during the preparation resulted in almost spherical HA particles. This explains why anisotropic refinements of HA crystallite sizes resulted in similar dimensions for directions ⟨100⟩ and ⟨001⟩ and why some participants could refine the size isotropically without obtaining substantially worse fits. Several mean values $\bar{x}$ were significantly different from the global mean value $\bar{\bar{x}}$ (Figure 6). In case of anisotropic refinement, the differences were not random but correlated ($R^2 = 0.77$). In other words, in most datasets with a significant bias both

https://doi.org/10.1017/S088571561500038X

dimensions were either smaller or greater than the global average $\bar{\bar{x}}$. This systematic offset was likely to be caused by one of the calibration issues discussed above.

The reproducibility of β-TCP crystallite size was one order of magnitude greater than reproducibilities reached for HA. This was mainly related to the fact that five configurations reported excessive standard deviations. All other datasets reached standard deviations comparable to HA crystallite sizes (Figure 6). Examination of the diffraction contributions of β-TCP and HA did not reveal any inherent sample-related factors impairing the approximation of the β-TCP peak shape. The strongest peaks were not significantly overlapped, neither by neighbouring β-TCP peaks, nor by those of the HA phase (Figure 1), and all instrument alignment and calibration related effects would have affected both phases to the same extent. However, the lower phase content of β-TCP, which was the primary reason for the lower intensity of the strongest peaks compared with HA, resulted in a lower signal-to-noise ratio and counting noise covering the tails of the peaks. Even with maximum peak intensities far stronger than the counting noise, separate refinement of crystallite size and micro strain may become impossible if the peak tails, which manifest the Lorentzian characteristics of the broadening effect to a large degree, are dominated by noise. Whether or not refinement of micro strain played a role in unreliable crystallite sizes was not known, as no further details on the refinement strategy were reported by the participants. Those participants who reported inconsistent β-TCP crystallite sizes were therefore advised to revise their refinement strategy and/or improve their counting statistics.

VII. CONCLUSION

The ILS on quantification of HA and β-TCP bioceramics by powder XRD and Rietveld refinement, conducted according to ASTM E691–13, provided detailed statistical data on the repeatability and reproducibility of the results reported by a number of independent laboratories. The group of participants comprised operators of various levels of experience using a mix of different instruments and software, and employing their established data collection and evaluation techniques. The study thus provided results not obtained under best-case conditions, but under realistic conditions taking into account variable skills, time constraints, and optimizations of instrument configurations.

The precision and accuracy of refined phase quantities generally depends on the context, namely on the number of phases in the sample, these phases' pattern complexity, particle and crystallite size and shape, chemical composition, and many more parameters. Nevertheless were our results consistent with previous studies in that the multilaboratory reproducibility parameters of phase quantifications were roughly two times greater than the within-laboratory repeatability parameters. Based on the ratio between repeated and reproduced phase quantifications, laboratories may prefer to multiply their standard deviation with factor 2 to report reproducibility standard deviations s_R rather than repeatability standard deviations. In a conservative approach, the reproducibility standard deviations may further be multiplied by factor 2.8 (with reference to ASTM E691 – 13 (2013) and/or ASTM E177 – 13 (2013)) to obtain 95% reproducibility intervals R. In our study, the latter approach resulted in $R = \pm 1.67$ wt%.

Our ILS also revealed that the data quality of several configurations was insufficient to refine crystallite sizes of the minor phase β-TCP. The problem manifested itself in excessively high standard deviations, which could be easily spotted without comparing the results with other more robust refinements. Laboratories experiencing inconsistent crystallite sizes from repeated data collections and refinements are therefore advised to revise their sample preparation, data collection settings, and refinement strategy, until refined crystallite sizes are consistent.

The statistical information presented in this study may be considered for reporting and validation by laboratories determining phase quantities of HA and β-TCP ceramics by powder XRD, as suggested by several national and international standards. Laboratories working in other chemical systems may also benefit from the relation between within-laboratory repeatabilities and multilaboratory reproducibilities in general.

ACKNOWLEDGEMENTS

The author acknowledges the support of all anonymous contributors to this ILS. Without their efforts and data contributions this study would not have been realized. Benjamin Andreatta of RMS Foundation, Switzerland, is acknowledged for his support with ICP-MS analyses.

ASTM E177 – 13 (**2013**). "Standard practice for use of the terms precision and bias in ASTM test methods," in Annual Book of ASTM Standards (ASTM International, West Conshohocken, PA), Vol. 14.02.

ASTM E691 – 13 (**2013**). "Standard practice for conducting an interlaboratory study to determine the precision of a test method," in Annual Book of ASTM Standards (ASTM International, West Conshohocken, PA), Vol. 14.02.

ASTM F1088 – 4 (**2004**). "Standard specification for beta-tricalcium phosphate for surgical implantation," in Annual Book of ASTM Standards (ASTM International, West Conshohocken, PA), Vol. 13.01.

ASTM F1185 – 03 (**2003**). "Standard specification for composition of hydroxyapatite for surgical implants," in Annual Book of ASTM Standards (ASTM International, West Conshohocken, PA), Vol. 13.01.

ASTM F2024 – 10 (**2010**). "Standard practice for X-ray diffraction determination of phase content of plasma-sprayed hydroxyapatite coatings," in Annual Book of ASTM Standards (ASTM International, West Conshohocken, PA), Vol 13.01.

Bergmann, J., Friedel, P. and Kleeberg, R. (**1998**). "BGMN – a new fundamental parameters based Rietveld program for laboratory X-ray sources, it's use in quantitative analysis and structure investigations," Commission of Powder Diffraction, International Union of Crystallography, CPD Newslett. **20**, 5–8.

Bohner, M., Galea, L. and Doebelin, N. (**2012**). "Calcium phosphate bone graft substitutes: failures and hopes," J. Eur. Ceram. Soc. **32**, 2663–2671.

Chow, L. C. (**2001**). "Solubility of calcium phosphates," in *Octacalcium Phosphate*, edited by L. C. Chow and E. D. Eanes (Karger, Basel), Vol. **18**, pp. 94–111.

Dickens, B., Schroeder, L. W. and Brown, W. E. (**1974**). "Crystallographic studies on the role of Mg as a stabilizing impurity in β-$Ca_3(PO4)_2$ I. The crystal structure of pure β-$Ca_3(PO4)_2$," J. Solid State Chem. **10**, 232–248.

Dinnebier, R. E. and Billinge, S. J. L. (Eds.) (**2008**). *Powder Diffraction: Theory and Practice* (Royal Society of Chemistry, Cambridge).

Döbelin, N., Luginbühl, R. and Bohner, M. (**2010**). "Synthetic calcium phosphate ceramics for treatment of bone fractures," Chimia **64**, 723–729.

Dorozhkin, S. V. (**2002**). "A review on the dissolution models of calcium apatites," Prog. Cryst. Growth Charact. Mater. **44**, 45–61.

Dorozhkin, S. V. (**2011**). "Self-setting calcium orthophosphate formulations: cements, concretes, pastes and putties," Int. J. Mater. Chem. **1**, 1–48.

Elliott, J. C. (**1994**). *Structure and Chemistry of the Apatites and other Calcium Orthophosphates* (Elsevier, Amsterdam).

Enderle, R., Gotz-Neunhoeffer, F., Gobbels, M., Muller, F. A. and Greil, P. (**2005**). "Influence of magnesium doping on the phase transformation temperature of beta-TCP ceramics examined by Rietveld refinement," Biomaterials **26**, 3379–3384.

Gopal, R. and Calvo, C. (**1972**). "Structural relationship of Whitlockite and $\beta Ca_3(PO_4)_2$," Nat. Phys. Sci. **237**, 30–32.

ICDD (**2013**). PDF-4+2013 (Database), edited by Dr. Soorya Kabekkodu, (International Centre for Diffraction Data, Newtown Square, PA, USA).

Ishikawa, K., Ducheyne, P. and Radin, S. (**1993**). "Determination of the Ca/P ratio in calcium-deficient hydroxyapatite using X-ray-diffraction analysis," J. Mater. Sci.-Mater. Med. **4**, 165–168.

ISO 13175-3 (**2012**). *Implants for surgery - Calcium phosphates - Part 3: Hydroxyapatite and beta-tricalcium phosphate bone substitutes* (ISO, Geneva, Switzerland).

ISO 13779-3 (**2008**). *Implants for surgery - Calcium phosphates - Part 3: Chemical analysis and characterization of crystallinity and phase purity* (ISO, Geneva, Switzerland).

Jackson, L. E., Barralet, J. E. and Wright, A. J. (**2004**). "Rietveld analysis in sintering studies of Ca-deficient hydroxyapatite," Key Eng. Mater. **254–256**, 297–300.

Keller, L. (**1995**). "X-ray powder diffraction patterns of calcium phosphates analyzed by the Rietveld method," J. Biomed. Mater. Res. **29**, 1403–1413.

Madsen, I. C. and Scarlett, N. V. Y. (**2008**). "Quantitative phase analysis" in *Powder Diffraction: Theory and Practice*, edited by R. E. Dinnebier and S. J. L. Billinge (Royal Society of Chemistry, Cambridge), pp. 298–331.

McCusker, L. B., Von Dreele, R. B., Cox, D. E., Louer, D. and Scardi, P. (**1999**). "Rietveld refinement guidelines," J. Appl. Crystallogr. **32**, 36–50.

Neira, I. S., Kolen'ko, Y. V., Lebedev, O. I., Van Tendeloo, G., Gupta, H. S., Guitián, F. and Yoshimura, M. (**2009**). "An effective morphology control of hydroxyapatite crystals via hydrothermal synthesis," Cryst. Growth. Des. **9**, 466–474.

Neuman, W. F. and Mulryan, B. J. (**1971**). "Synthetic hydroxyapatite crystals IV. Magnesium incorporation," Calcif. Tissue Res. **7**, 133–138.

Nilen, R. W. N. and Richter, P. W. (**2008**). "The thermal stability of hydroxyapatite in biphasic calcium phosphate ceramics," J. Mater. Sci.: Mater. Med. **19**, 1693–1702.

Raynaud, S., Champion, E., Bernache-Assolant, D. and Laval, J.-P. (**2001**). "Determination of calcium/phosphorus atomic ratio of calcium phosphate apatites using X-ray diffractometry," J. Am. Ceram. Soc. **84**, 359–366.

Raynaud, S., Champion, E., Bernache-Assollant, D. and Thomas, P. (**2002a**). "Calcium phosphate apatites with variable Ca/P atomic ratio I. Synthesis, characterisation and thermal stability of powders," Biomaterials **23**, 1065–1072.

Raynaud, S., Champion, E. and Bernache-Assollant, D. (**2002b**). "Calcium phosphate apatites with variable Ca/P atomic ratio II. Calcination and sintering," Biomaterials **23**, 1073–1080.

Reid, J. W. and Hendry, J. A. (**2006**). "Rapid, accurate phase quantification of multiphase calcium phosphate materials using Rietveld refinement," J. Appl. Crystallogr. **39**, 536–543.

Riboud, P. V. (**1973**). "Composition et stabilité des phases a structure d'apatite dans le système CaO-P_2O_5-oxyde de Fer-H_2O a haute temperature," Ann. Chim. **8**, 381–390.

Rietveld, H. M. (**1969**). "A profile refinement method for nuclear and magnetic structures," J. Appl. Crystallogr. **2**, 65–71.

Schroeder, L. W., Dickens, B. and Brown, W. E. (**1977**). "Crystallographic studies of the role of Mg as a stabilizing impurity in β-$Ca_3(PO_4)_2$ II. Refinement of Mg-containing β-$Ca_3(PO_4)_2$," J. Solid State Chem. **22**, 253–262.

Stutzman, P. (**2005**). "Powder diffraction analysis of hydraulic cements: ASTM Rietveld round-robin results on precision," Powder Diffr. **20**, 97–100.

Sudarsanan, K. and Young, R. A. (**1969**). "Significant precision in crystal structure details: Holly springs hydroxyapatite," Acta Crystallogr. **B25**, 1534–1543.

Tõnsuaadu, K., Gross, K. A., Plusuma, L. and Veiderma, M. (**2012**). "A review on the thermal stability of calcium apatites," J. Therm. Anal. Calorim. **110**, 647–659.

Toth, J. M., Hirthe, W. M., Hubbart, W. G., Brantley, W. A. and Lynch, K. L. (**1991**). "Determination of the ratio of HA/TCP mixtures by x-ray diffraction," J. Appl. Biomater. **2**, 37–40.

Welch, J. H. and Gutt, W. (**1961**). "High-temperature studies of the system calcium oxide-phosphorus pentoxide," J. Chem. Soc. **874**, 4442–4444.

Wilson, R. M., Elliott, J. C., Dowker, S. E. P. and Rodriguez-Lorenzo, L. M. (**2005**). "Rietveld refinements and spectroscopic studies of the structure of Ca-deficient apatite," Biomaterials **26**, 1317–1327.