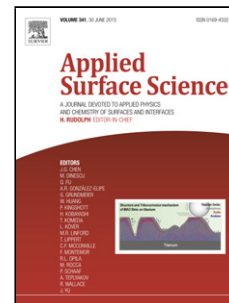


## Accepted Manuscript

Title: Principal Component Analysis of Raman spectra for TiO<sub>2</sub> Nanoparticle Characterization

Authors: Alina Georgiana Ilie, Monica Scarisoareanu, Ion Morjan, Dutu Elena, Maria Badiceanu, Ion Mihailescu



PII: S0169-4332(17)30216-7  
DOI: <http://dx.doi.org/doi:10.1016/j.apsusc.2017.01.193>  
Reference: APSUSC 34988

To appear in: *APSUSC*

Received date: 27-10-2016  
Revised date: 17-1-2017  
Accepted date: 19-1-2017

Please cite this article as: Alina Georgiana Ilie, Monica Scarisoareanu, Ion Morjan, Dutu Elena, Maria Badiceanu, Ion Mihailescu, Principal Component Analysis of Raman spectra for TiO<sub>2</sub> Nanoparticle Characterization, Applied Surface Science <http://dx.doi.org/10.1016/j.apsusc.2017.01.193>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Principal Component Analysis of Raman spectra for TiO<sub>2</sub> Nanoparticle Characterization

Alina Georgiana Ilie<sup>1,2</sup>, Monica Scarisoareanu<sup>1</sup>, Ion Morjan<sup>1</sup>, Dutu Elena<sup>1,2</sup>, Maria Badiceanu<sup>2</sup>, Ion Mihailescu<sup>1</sup>

<sup>1</sup> National Institute for Lasers, Plasma and Radiation Physics, Magurele, Romania

<sup>2</sup> University of Bucharest, Faculty of Physics, Bucharest, Romania

Corresponding author Tel.: +4021-4574489; fax: +4021-4574243 E-mail address:  
[ilie.georgiana@inflpr.ro](mailto:ilie.georgiana@inflpr.ro)

## Highlights

- Principal Component Analysis was applied on TiO<sub>2</sub> nanoparticle Raman spectra.
- A polynomial correlation was found between Sn/Ti percentage and PC-3 scores.
- Anatase crystallite size approximation in good agreement with XRD calculations.
- The method we propose is viable for rapid characterization of TiO<sub>2</sub> nanoparticles.

## Abstract

The Raman spectra of anatase/rutile mixed phases of Sn doped TiO<sub>2</sub> nanoparticles and undoped TiO<sub>2</sub> nanoparticles, synthesised by laser pyrolysis, with nanocrystallite dimensions varying from 8 to 28 nm, was simultaneously processed with a self-written software that applies Principal Component Analysis (PCA) on the measured spectrum to verify the possibility of objective auto-characterization of nanoparticles from their vibrational modes. The photo-excited process of Raman scattering is very sensible to the material characteristics, especially in the case of nanomaterials, where more properties become relevant for the vibrational behaviour. We used PCA, a statistical procedure that performs eigenvalue decomposition of descriptive data covariance, to automatically analyse the sample's measured Raman spectrum, and to interfere the correlation between nanoparticle dimensions, tin and carbon concentration, and their Principal Component values (PCs). This type of application can allow an approximation of the crystallite size, or tin concentration, only by measuring the Raman spectrum of the sample. The study of loadings of the principal components provides information of the way the vibrational modes are affected by the nanoparticle features and the spectral area relevant for the classification.

**Keywords:** Principal Component Analysis (PCA); nanoparticle classification: Raman spectroscopy; TiO<sub>2</sub> nanoparticles; Sn doped TiO<sub>2</sub>;

## Introduction

Raman spectroscopy has become an essential tool for nanoparticle characterization due to the sensible correlation between the vibrational behaviour and material structure, composition and properties, emphasising on the dimensionality dependence and impurity

effects. The quantum size confines the electronic and vibrational wavefunctions. Increase of the specific surface of the material [1], results in shifting, broadening, electron-phonon coupling [2-4], combinations and overtones of the vibrational spectra compared with the bulk material. Such dependencies have been theoretically and experimentally studied [5-7], for TiO<sub>2</sub> nanoparticles, to better understand the structure-property relations. Good predictions were obtained, in case of phonon confinement model applied on the shifting and broadening of the low-frequency mode E<sub>g</sub>, of anatase phase [8]. This semiconducting material has three natural phases, rutile, brookite and anatase which stands out due to its wide band-gap of 3.2 eV, making it suitable for applications in solar energy industry [9], photocatalysis [10-12], chemical sensors [13], electronics and microelectronics [14]. Concerning photocatalysis, the nanosize offers a reduced number of electron-hole recombinations and an increased active surface area that enhances the material performances. Current efforts focus on extending the optical absorption from UV to visible part of the solar spectrum, mainly by doping with carbon, iron, tin or other metals. Doping influences the Raman spectra of the nanomaterial in similar ways the size of the material does [15], depending on the magnitude of its interference in the crystalline structure or on the nanomaterial surface. Doping and size are not the only aspects that produces variation of the Raman spectrum of TiO<sub>2</sub> nanomaterials, but also phase mixtures, phase transitions, structural properties, aspects that make the spectral analysis difficult to approach [16]. Complex areas that use Raman spectroscopy for analysis and interpretation, such as microbiology, genetics, oncology or material science have combined it with Principal Component Analysis (PCA) to discriminate microorganisms at species level [17], differentiate between breast cancer lines [18] or normal/abnormal human sperm cells from Raman spectra of DNA packaging [19], distinguish basal cell carcinoma from its surrounding noncancerous tissue [20] or to determine phase transition pressures in polycrystalline materials [21]. PCA is one of the oldest and best known techniques of multivariate analysis [22] that reduces the dimensionality of a data set. It consists of a large number of interrelated variables, and generates a new set of orthogonal variables, called principal components (PCs). The method concentrates the majority of variations in the Raman spectrum and discharges the noise and redundant data. It is used for exploratory data analysis, making predictive models and cluster analysis. PCA capability to process high quantity of information, taking into account only data that is covariational relevant, makes it a good candidate to analyse Raman spectra of TiO<sub>2</sub> nanoparticles. The determination is performed for size and dopant correlations with the vibrational behaviour, and with reference to principal component loadings, to distinguish which spectral domains are contributing to the selection.

This type of analysis is used in the field of semiconductor materials, but mostly to analyse data from semiconductor manufacturing processes or for improving fault detection and diagnosis on semiconductor manufacturing [23-24]. Although there are many other techniques that can be used in the detailed analysis of nanomaterials, this study presents a viable method of analysis for several nanoparticle characteristics simultaneously (nanocrystallite size and dopant concentration), based only on their Raman spectra. It also allows a more thorough research on the vibrational spectra behaviour, with respect to size and dopant concentration, by analysing the loadings for each PC, which proves relevant for basic understanding of semiconductor nanomaterials. We stress upon that the method offers the possibility to handle hundreds of sample spectra at the same time, its capacity depending only on the computational power available.

In this work we applied PCA on Raman spectra of multiple sets of TiO<sub>2</sub> nanoparticles, covered in turbostratic carbon shells, and/or doped with Sn, with different sizes and proportions of anatase and rutile. A good approximation was determined for anatase nanocrystallite size calculated from XRD measurements, for Sn concentration and turbostratic carbon shell presence determined from EDAX and TEM measurements [25-26]. This method proposes an automated classification of TiO<sub>2</sub> nanoparticles by crystallite size using only Principal Component Analysis on their Raman spectrum that, with at the best of our knowledge, has not been applied in other studies.

## Experimental

### *Materials used*

The TiO<sub>2</sub> nanoparticles analysed in this work consist of three sets of samples synthesized by laser pyrolysis from gaseous precursors [27-29]. The first set, six TSPH samples of TiO<sub>2</sub> nanoparticles, suited for photocatalytic purposes, with Sn and C content, obtained by reactions in the pyrolytic flame of C<sub>2</sub>H<sub>4</sub>, synthetic Air and TiCl<sub>4</sub>, with Ar and SnCl<sub>4</sub> mixture for tin doping, induced by CO<sub>2</sub> laser radiation at 10.60 μm. The nanoparticles were furnace air annealed at 450°C (3 hours) to decrease their carbon content. They were further investigated by Raman spectroscopy, TEM, X-ray diffraction and EDAX to determine the nanoparticle dimensions (7-22 nm for anatase, 8-17 nm for rutile), phase mixtures and Sn content (0-28.5% Sn/Ti), C coating and morphology [30].

The second set, SMT, consists of 6 samples of TiO<sub>2</sub> nanoparticles synthesised by laser pyrolysis from TiCl<sub>4</sub> and C<sub>2</sub>H<sub>4</sub> gas mixtures, with air as oxygen donor, having lower dimensions as XRD analysis reveals (8-14 nm for anatase phase and 7-11 nm for rutile), which are thoroughly studied in a previous article [25]. In Fig.1. XRD diffraction patterns for samples SMT 13, SMT 14 and SMT 15 are exhibited in comparison with the bottom figure that shows the diffractogram of the commercial P25 Degussa sample (about 84% anatase and 16% rutile). Additionally, in Fig.2. a TEM image and grain size histogram of sample SMT14 is presented, showing a loose aggregation of round grains, with a medium grain size of 14.68 nm.

In the third set, TLO, 18 samples of carbon-coated TiO<sub>2</sub> nanoparticles obtained by the same method from TiCl<sub>4</sub>, air and alternatively C<sub>2</sub>H<sub>4</sub> and C<sub>2</sub>H<sub>4</sub>/C<sub>2</sub>H<sub>2</sub> gas mixtures, exhibit higher nanocrystallite dimensions (18-28 nm for anatase phase and 13-47 nm for rutile), as presented in previous articles [26]. Moreover, XRD analysis for samples TLO19, TLO20 and TLO21 (Fig.3) reveals anatase crystallite dimensions of 22 nm, 25 nm and respectively 22 nm. The nanoparticles are round-shaped, and slightly aggregated, as seen in the TEM image of sample TLO16 (Fig.4), where the grain size histogram reports a 21.57 nm medium grain size, in good agreement with the XRD calculated anatase crystallite size (Table 2). The samples are synthesized by the same method but in different geometric systems, gas mixtures, and various experimental conditions to determine an extended range of TiO<sub>2</sub> nanoparticle types, with different morphologies, contents and characteristics that ensure a wide spread of data to be analysed, from specific points of view, invariant to their source.

### *Methods*

The X-ray diffraction measurements were carried out with a PANalytical X'Pert PRO MPD X-ray Diffractometer with Cu X-ray tube providing K $\alpha$  wavelength of 1.5418 Å, a curved graphite monochromator and variable divergence slit placed in the beam pathway to the fixed irradiated area, at room temperature. PANalytical X'Pert High-Score Plus software package was used to analyse the XRD spectra and using Scherrer's formula  $D = k\lambda/\beta\cos\theta$ , the mean crystallite sizes was calculated, using  $k=0.94$  [25,31-33],  $\beta$  the FWHM and  $\theta$  the Bragg angle for the main anatase and rutile peaks, respectively.

Energy-dispersive X-ray Spectroscopy (EDX) was performed inside a Scanning Electron Microscope (SEM), FEI Co., model Inspect S, 0-30 KV accelerating voltage with EDAX Co. SiLi detector. The method of atomic percentage estimation is standardless, we used a percentage of the net intensity for Ti and Sn of the diffuse reflectance spectra.

The Raman Spectroscopy analysis was conducted with a Dispersive Raman microscope JASCO NRS-7200, with confocal optics, under  $\lambda=531.94$  nm laser radiation with 5.7 mW and no attenuator, 1800 l/mm single grating monochromator, 25x1000  $\mu\text{m}$  slit, MPLFLN 20X objective lens, with 1  $\text{cm}^{-1}$  resolution on 4-stage Peltier cooled CCD detector, for 20 second exposure and 6 accumulations per sample, in 5 separated points to assure information homogeneity for the second and the third sets. The spectral window is around 180-920  $\text{cm}^{-1}$  which covers the high frequency Raman spectra of  $\text{TiO}_2$  nanoparticles. For the first set, the spectral window is 180-2700  $\text{cm}^{-1}$ , with 600 l/mm grating and 10 x 1000  $\mu\text{m}$  slit, 1.6 mW laser power, and OD 0.6 attenuator, with same objective lens and resolution. The wider spectral window allows carbon Raman modes visibility. The first set was analysed separately due to its Sn content, unlike the other two sets. Each sample was presented as powder deposited mechanically on Si plates, put onto the spectrometer stage. Before the measurement, the background noise was reduced using the spectrometer software. Multiple sample revealed photoluminescence, feature that led to baseline correction for the Raman spectra, which was carried out with Fityk 0.9.8 Software [34].

#### *Raman spectra processing*

The Raman spectra processing of the  $\text{TiO}_2$  samples was made by 4 self-written Python scripts that prepare the experimental data and apply Principal Component Analysis.

The first program Mediator.py reads the Raman text files provided by the Raman spectrometer and gives a medium Raman spectra of all the input data. It is used to translate the spectra measured in 5 points, on the sample, into only one average Raman spectrum. This partially reduces some of the inhomogeneity errors induced by uneven feature distribution of nanoparticles. The script reads the two columns of the text files, which represent the frequency value in  $\text{cm}^{-1}$  and the intensity of the Raman signal, and assigns a fitting of the intensity to a fixed point on the frequency axis, by an average intensity value for the two nearest neighbour points on the axis, assuring the same frequency points in which the intensity is compared. The next step is to select de common area for all spectra and the normalization of the intensity to the maximum of each spectrum (assigned to  $E_g$  Raman mode of anatase  $\text{TiO}_2$ ), then to compute the mean value of the intensity, of each frequency point, by the number of the spectra initially read. The resulted single mediated Raman spectrum is then written to another text file that will be the input for the next scripts.

The next script `Open&Process.py` reads the mediated spectra from all the samples we want to analyse, applies the same fitting procedure as the one before, to ensure identical frequency points, the X vector, and constructs a matrix Y, with all the Raman intensities of the samples, and their names, the N vector, and then cuts the data matrix and vector to the common spectral window dimensions. The Y matrix, X and N vectors are written into a text file that is read by the next script, which is optional.

Depending of the matrix dimensions, to ensure a rapid and easy processing of the data by a normal computer, sometimes it is necessary to reduce the input data by excluding points that are informationally redundant. The `Reduce.py` script reads the pre-processed data and homogeneously deletes points in order to reduce the dimensions of the matrix, as the user ordered, and writes the new text file. Reducing points under 1 cm<sup>-1</sup> value doesn't affect the information due to the measurement resolution. After this step, each iteration must be verified with the previous one to ensure no relevant data loss. Each script has a test that compares the initial spectra to the final one, to verify that the spectrum has not been distorted.

The `PCA.py` script is the main program that applies the Principal Component Analysis onto the data matrix using numpy and scipy libraries and gives the final results. First the data is read from the text file and the Y matrix is transformed into the covariance matrix M:

$M = \sum_{i=1}^L (\underline{Y}_i - \underline{\mu}_Y) (\underline{Y}_i - \underline{\mu}_Y)^T$  where  $\underline{\mu}_Y$  is the mean value for each data vector  $\underline{\mu}_Y = \frac{1}{n} \sum_{i=1}^n \underline{Y}_i$  and  $n$  the vector length.

Using `linalg.py` library we compute the sorted eigen values  $\lambda_i$  and eigen vectors  $\Lambda_i$  of the covariance matrix, and leave the user to choose the information percentage that he wants to conserve, and as a consequence to that, the number of eigen values and vectors that the program will keep. The number of eigen values kept, will become the number of principal components (PC) that will be analysed. The information percentage is calculated by the formula:  $\epsilon = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^n \lambda_i} \times 100\%$  [35] where  $m$  is the number of eigen values kept, and  $n$  the initial number of values, representing in fact the  $M_{(n,n)}$  square matrix dimension. The script plots, for each PC kept, the contribution to the information percentage conserved. After this choice is made, the subset of the eigenvectors is saved, as new vector basis, in a matrix  $K_{(m,n)}$ , and the Karhunen-Loève transformation [36-37] is applied:

$KLM = \text{abs}(K_{(m,n)}) * Y^T$  where the KLM matrix contains the values for each principal component PC- i, for every Raman spectrum sample.



The program then plots the loadings of each principal component PC-i, as the corresponding eigen vector  $\Lambda_i$  versus the frequency vector X, and the scores for 2D PCA and, if it's necessary, 3D PCA. The fitting of the PC values with experimental measurements was accomplished with Curve Expert Professional 2.4.0.

The scripts presented above have been used to process separately the 6 TSPH Raman spectra within spectral window 250-1000  $\text{cm}^{-1}$ , and then the other 24 SMT and TLO Raman spectra of  $\text{TiO}_2$  nanoparticles within spectral window of 250-840  $\text{cm}^{-1}$ .

## Results and Discussion

Anatase phase of titanium dioxide has six Raman active modes  $A_{1g}+2B_{1g}+3E_g$ , determined by group analysis  $D_{4h}(I4_1/amd)$  [38], and identified in the first-order Raman spectrum of single crystal  $\text{TiO}_2$  by Ohsaka [39] at 144  $\text{cm}^{-1}$  ( $E_g$ ), 197  $\text{cm}^{-1}$  ( $E_g$ ), 399  $\text{cm}^{-1}$  ( $B_{1g}$ ), 516  $\text{cm}^{-1}$  ( $A_{1g}+B_{1g}$ ) and 639  $\text{cm}^{-1}$  ( $E_g$ ). Rutile  $D_{4h}(P4_2/mnm)$  single crystal Raman modes were measured by Porto et al. [40] at 143  $\text{cm}^{-1}$  ( $B_{1g}$ ), 447  $\text{cm}^{-1}$  ( $E_g$ ), 612  $\text{cm}^{-1}$  ( $A_{1g}$ ), 826  $\text{cm}^{-1}$  ( $B_{2g}$ ), consistent with the four Raman active modes theoretically determined from group theory  $A_{1g}+B_{1g}+B_{2g}+E_g$ . The nanophase of  $\text{TiO}_2$  spectrum is similar with the one measured in bulk anatase/rutile, with lower intensities and broader linewidth. The finite size of the nanoparticles makes surface modes become predominant, due to the increase in surface/volume ratio [41]. Both shifts and broadening of Raman modes are attributed to effects of the decreasing particle size on the force constants and vibrational amplitudes of the nearest neighbour bonds, phonon confinement and structural defects. It is possible to theoretically or experimentally determine the size of nanoparticles from a measurement of the Raman bands [42-44]. The strongest mode of  $\text{TiO}_2$  anatase nanoparticles, is the low frequency phonon band 144  $\text{cm}^{-1}$  ( $E_g$ ) of anatase, which has been widely studied in order to determine temperature effect in  $\text{TiO}_2$  nanoparticles [45, 41] and nanoparticle dimension. Combining the phonon confinement model with size distributions and theoretical phonon dispersions, the  $E_g$  mode intensity and frequency, as a function of particle size, was calculated in good agreement with experimental data [46-47]. This type of methods involve deconvolution of the Raman spectra for each sample and takes into account only the size effects on the Raman modes, while PCA allows processing of multiple samples simultaneously, without any demanding computational requirements, and decomposes the Raman spectra variations into linear independent behaviours. In our study we focused on the high-frequency Raman modes of  $\text{TiO}_2$  nanoparticles (Fig.5b), Fig. 13) to determine the nanocrystallite dimensions, Sn concentration and carbon content presence.

*TiO<sub>2</sub> nanoparticles doped with Sn.*

The measured Raman spectrum of the first set of samples, TSPH, is shown in Fig.5b) where the anatase and rutile TiO<sub>2</sub> high frequency modes are given. For TSPH01 and TSPH4, the D and G modes of turbostratic carbon indicate a consistent quantity of carbon that accompanies the TiO<sub>2</sub> nanophase. The Fig.5a) represents the spectral range that is taken into consideration by the PCA script, eliminating the carbon modes. The six spectra are transformed into a matrix and analysed. More than 99% of information is concentrated in only three principal components, as it is presented in Fig. 6. The respective loadings for each principal component is plotted in Fig.7, and symbolise vectors of a new basis that confines the information contained in the variations of the measured spectrum, and separates the contribution of spectral domains into linear independent values. The values of the first and second PC are plotted in Fig.8 where a segregation of two types of samples is obvious. This segregation is consistent with carbon content seen in the extended Raman spectrum, but that certain spectral range isn't processed by the scripts. Examining the loadings for PC-1 and PC-2 we observe that the spectral domains with great impact are the ones correlated with the Raman modes of anatase/rutile TiO<sub>2</sub>. Respectively, for PC-1 the domains are sharp, indicating a quantification of the intensity modes, as for PC-2 loadings, the broadened domains show a quantification in shifting and broadening of modes, but also of intensification of the A<sub>1g</sub>+B<sub>1g</sub> mode of anatase. The first component contains 96% of the relevant information in the measured spectra, giving us an overall image of the average spectrum of the TiO<sub>2</sub> nanoparticles. The PC-1 values measure a global variation that is consistent with the presence of carbon shell on the nanoparticles, indicating surface interaction between TiO<sub>2</sub> and carbon, which affects the vibrational behaviour of the TiO<sub>2</sub> nanoparticle in such way it becomes an independent characteristic of the Raman spectra variation.

The plot of PC-1 versus PC-3 is presented in Fig. 9, where we obtain an almost linear placement of points representing the samples, in the order of their Sn/Ti ratio, reported by the EDAX measurements (see Table 1). The loadings for the first and third principal components have a common point of interest, originated at 518 cm<sup>-1</sup> frequency, a vibrational mode of anatase. The correlation between Sn content and this mode can be attributed either to Sn insertion into the crystalline structure of TiO<sub>2</sub>, that should translate into shifting and broadening of the vibrational modes, or an external contribution of the metallic Sn or tin oxides to the Raman spectra of TiO<sub>2</sub>. In other studies, the variation of this mode has been proven to indicate a phase transition of TiO<sub>2</sub> from anatase to rutile [48] and that Sn presence in TiO<sub>2</sub> nanoparticle

synthesis enhances anatase-rutile transitions [49]. The fitting of PC-1 and PC-3 values with Sn/Ti concentration revealed that the first PC has a very poor contribution, and the best results, were obtained by a 3-rd order polynomial fitting of the values of PC-3 and tin concentration (Fig. 10), with  $R = 0.998$  correlation coefficient and  $\varepsilon = 0.9$  [% ] standard error:

$$\text{Sn [\%]} = -186.579 + 1789.53 \cdot \text{PC-3} - 5360.196 \cdot \text{PC-3}^2 + 5358.284 \cdot \text{PC-3}^3$$

This 3-rd order polynomial model is useful in relations where an inflection point appears, as the case in study, where tin concentration influence on the Raman spectra becomes steeper for values around 10-15% Sn/Ti content.

The low contribution of the first principal component suggests that the Sn effect on  $\text{TiO}_2$  nanoparticles, regarding their vibrational modes, can only be attributed to phase transition enhancement. A good correlation between tin concentration and values of a principal component of spectral data give rise to the possibility of approximating Sn content, in  $\text{TiO}_2$  nanoparticle samples, only by measuring and processing their Raman spectra, with respect to a data base of samples with well predetermined Sn content, and fitting the data.

Another linear manifestation of the PC values is exhibited in the PC-2 versus PC-3 plot, in Fig. 11, where the sample placements are correlated with the anatase nanocrystallite dimensions, calculated from XRD measurements (Table 1). The PC-2 and PC-3 also have in common the  $518 \text{ cm}^{-1}$  domain, which suggests an influence of the Sn content on the anatase crystallite dimension, which is to be expected if tin favours rutile phase growth. I. Rangel-Vázquez justified the diminishing of  $\text{TiO}_2$  anatase crystallite size by the difference between the ionic radius of  $\text{Ti}^{4+}$  and the one of  $\text{Sn}^{4+}$ , which introduces a perturbation in the crystal structure of anatase in a way that inhibits the growth of the crystallites [50]. As a fact, a high value of PC-3, which is directly correlated with high concentration of Sn, will correspond to a small anatase crystallite dimension. But the phase transition supported by Sn isn't enough to evaluate the dimensions of anatase. The second PC gives a quantification of the Raman modes variation in width and position, that linearly fits with the PC-3 value and nanocrystallite dimension, with  $R = 0.996$  correlation coefficient and  $\varepsilon = 0.26 \text{ nm}$  standard error:  $D_A(\text{nm}) = 19.9634 + 10.316 \cdot \text{PC-2} - 67.882 \cdot \text{PC-3}$ . Similar to the previous case, this linear dependence of the PCs with  $\text{TiO}_2$  anatase nanocrystalline dimensions makes this technique suitable for crystallite dimension approximation, as highlighted in Fig. 12. One utilises the vibrational spectrum of the  $\text{TiO}_2$  nanoparticle and a data base of  $\text{TiO}_2$  nanoparticles, previously characterised to determine their anatase crystallite dimension, ending by linearly fitting the relevant PCs. To verify that this

behaviour is consistent with other TiO<sub>2</sub> samples, we applied the same procedures on the next two sets of Raman spectra, simultaneously analysing 24 samples, with variable carbon content.

The rutile phase isn't discussed in this case because it has a lower percentage in the samples (see Table 1 and Table 2), and does not distinguish its contribution to the analysis.

### *Undoped TiO<sub>2</sub> samples*

The Raman spectra of SMT and TLO samples (Fig. 13) was processed the same way as the first set. The common spectral window was analysed for all the 24 spectra, focusing on the high-frequency modes of the TiO<sub>2</sub> nanoparticles, where the anatase and shy rutile modes can be discriminated. 96.5% of the information calculated from the eigen values is present in the first PC, while the second PC contains 1.5% of information, leaving the rest with contributions below 1%. This indicates that the majority of variations in the Raman spectra has been concentrated into the first two PCs. Comparing the loadings for PC-1 and PC-2 (Fig. 14), the points with high contribution to the PC value correspond to anatase vibrational modes, with peaks from 389-395 cm<sup>-1</sup> for B<sub>1g</sub>, 510-517 cm<sup>-1</sup> for A<sub>1g</sub>+B<sub>1g</sub> and 631-640 cm<sup>-1</sup> for E<sub>g</sub>, accompanied by the modes broadening. This characteristics are linear independent behaviours of the variation of vibrational modes, and the weights of each sample in this new basis describe a linear dependence between them, that is correlated with the crystallite dimension of anatase TiO<sub>2</sub> nanoparticles (Fig. 15). Fitting the PC-1, PC-2 and crystallite dimensions, calculated from XRD measurements (Table 2), we obtain very good results using linear approximation with R= 0.937 correlation coefficient and  $\varepsilon = 1.86$  nm standard error, and linear logarithmic approximation with R= 0.941 correlation coefficient and  $\varepsilon = 0.92$  nm standard error:  $D_A[\text{nm}] = 30.501 - 7.181 \cdot \ln(\text{PC-1}) - 4.245 \cdot \ln(\text{PC-2})$ . Comparing the results using Akaike information criterion [51] and F-test [52], the linear logarithmic fit offers the best approximation of the data, but this function is unlikely to be the one governing this behaviour, and may be a good approximation for this set of data only. We choose to use the linear fit, as in the previous application, allowing us to determine easily the TiO<sub>2</sub> nanocrystallite size by processing the Raman spectrum:

$D_A(\text{nm}) = A + B \cdot \text{PC-i} + C \cdot \text{PC-j}$ . Depending on the data set, the order of the PCs correlated with anatase dimensions may vary. In this case  $A=3.050$ ,  $B=-7.181$  for  $i=2$ ,  $C=-4.245$  for  $j=3$ .

In the Sn doped samples, the carbon presence was correlated with the first PC, as an independent characteristic, but in this case the carbon isn't highlighted in any of the PCs, only the anatase dimension being obviously correlated with PC values. We must take into account

that the first set was thermally treated, leaving only the turbostratic shells on the surface of the nanoparticle, and eliminating additional carbon resulted from the pyrolysis synthesis, while the second and third set of samples was analysed in their raw form. This aspect, combined with the feature dispersion of samples, made the carbon content a characteristic that was not separated from other nanoparticle properties, or not a relevant information from a covariational point of view.

Figure 16 shows the dimensions calculated by both functions, in comparison with the dimensions calculated from XRD measurements. The higher errors are presented in samples with significant rutile concentrations, indicating that the analysis quantifies modes broadening due to convolution of anatase and rutile modes as broadening caused by crystallite dimensions. Also, samples with smaller dimensions give wider Raman modes that overlap for rutile/anatase phases, introducing more uncertainty regarding the dimension/phase contribution. Taking into account that these samples are a mixture of two phases grown together as a  $\text{TiO}_2$  nanoparticle, in a one-step synthesis, accompanied by turbostratic carbon, this method, of calculating the nanocrystallite dimension of anatase  $\text{TiO}_2$  from its Raman modes, has proven results in good agreement with the experimental measurements. This analysis can be carried out onto many spectra simultaneously, and given its linear behaviour, it requires solving a basic fitting problem to determine rapidly the anatase  $\text{TiO}_2$  nanocrystallite dimension.

## Conclusions

Principal Component Analysis was applied on the high-frequency Raman spectra of three sets of  $\text{TiO}_2$  nanoparticle samples synthesized by laser pyrolysis with anatase/rutile mixed phases and crystallite size in range of 8-28 nm. The principal component values calculated revealed a good correlation between vibrational modes and Sn concentration, carbon content and nanocrystallite size. Using basic fitting of the PCs values in the first set of six samples, we calculated Sn concentration with  $\varepsilon = 0.9$  [%] error, and attributed this dependency to the variations in  $519\text{ cm}^{-1}$  mode of anatase, as revealed by the loading of the third PC. The crystallite size of anatase  $\text{TiO}_2$  was determined by linearly fitting the values for second and third PC with  $\varepsilon = 0.92$  nm error, compared with the XRD calculated crystallite dimensions.

The same method was applied on the remaining two sets of 24 samples, linearly fitting the values for the first and second PC, and determined the crystallite size of anatase  $\text{TiO}_2$  with  $\varepsilon = 1.86$  nm standard error, compared with the XRD calculated crystallite dimensions. The

method we propose is viable for rapid determination of specific characteristics for TiO<sub>2</sub> nanoparticles, such as their nanocrystallite size. It simultaneously processes all the Raman spectra with respect to a vibrational spectral database required, of specific TiO<sub>2</sub> nanoparticles, depending on the property we want to evaluate.

### **Acknowledgements**

This work has been financed by the National Authority for Research and Innovation in the frame of Nucleus programme- contract 4N/2016 and the project CETAL 8PM/I 26.11.2008.

## References

- [1] G. Gouadec, P. Colomban. Raman Spectroscopy of Nanomaterials: How Spectra Relate to Disorder, Particle Size and Mechanical Properties, *Progress in Crystal Growth and Characterization of Materials* 53 (2007) 1-56.
- [2] Y. Iida, M. Furukawa, T. Aoki, T. Sakain, Raman spectrum of ultrafine anatase powders derived from hydrolysis of alkoxide, *Applied Spectroscopy* 52 (1998) 673-678.
- [3] D. Bersani, P.P. Lottici, X.Z. Dingn, Phonon confinement effects in the Raman scattering by TiO<sub>2</sub> nanocrystals, *Applied Physics Letters* 72 (1998) 73-75.
- [4] W.F. Zhang, Y.L. He, M.S. Zhang, Z. Yin, Q. Chen, Raman scattering study on anatase TiO<sub>2</sub> nanocrystals, *Journal of Physics D: Applied Physics* 33 (2000) 912-916.
- [5] M. Ivanda, S. Music, M. Gotic, A. Turkovic, A.M. Tonejc, O. Gamulin, The effects of crystal size on the Raman spectra of nanophase TiO<sub>2</sub>, *Journal of Molecular Structure* 480–481 (1999) 641–644.
- [6] K. A. Farrell, Synthesis Effects on Grain Size and Phase Content in the Anatase-Rutile TiO<sub>2</sub> System, *Materials Science and Engineering*, Faculty of the Worcester Polytechnic Institute, 2001.
- [7] S. K. Gupta, R. Desai, P. K. Jha, S. Sahoo, D. Kirin, Titanium dioxide synthesized using titanium chloride: size effect study using Raman spectroscopy and photoluminescence, *Journal Raman Spectroscopy* 41 (2010) 350–355.
- [8] K.R. Zhu, M.S. Zhang, Q. Chena, Z. Yina, Size and phonon-confinement effects on low-frequency Raman mode of anatase TiO<sub>2</sub> nanocrystal, *Physics Letters A* 340 (2005) 220–227.
- [9] C. Colmenares, M.A. Aramendia, A. Marinas, J.M. Marinas, F.J. Urbano, Syn-thesis, characterization and photocatalytic activity of different metal-doped titania systems, *Applied Catalysis A: General* 306 (2006) 120–127.
- [10] A. L. Linsebigler, G. Lu, and J. T. Yates Jr., Photocatalysis on TiO<sub>2</sub> surfaces: principles, mechanisms, and selected results, *Chemical Reviews* 95 (1995) 735–758.
- [11] H.C. Lee, H. Jung, J.-M. Oh and J. H. Choy, A new nanohybrid photocatalyst between anatase (TiO<sub>2</sub>) and layered titanate, *Bulletin of the Korean Chemical Society*, 23 (2002) 477-480.

- [12] E. Figgemeier , W. Kylberg , E. Constable, M. Scarisoreanu, R. Alexandrescu, I. Morjan, I. Soare, R. Birjega, E. Popovici, C. Fleaca, L. Gavrilă-Florescu, G. Prodan, Titanium dioxide nanoparticles prepared by laser pyrolysis: Synthesis and photocatalytic properties, *Applied Surface Science* 254 (2007) 1037–1041.
- [13] A. P. Caricato, A. Luches, R. Rella, Nanoparticle Thin Films for Gas Sensors Prepared by Matrix Assisted Pulsed Laser Evaporation, *Sensors* 9 (2009) 2682-2696.
- [14] D. Rathee, S.K. Arya, M. Kumar, Analysis of TiO<sub>2</sub> for microelectronic applications: effect of deposition methods on their electrical properties, *Frontiers of Optoelectronics in China* 4 (2011) 349-358.
- [15] M. Scepănovi, S. Askračić, M. Grujić-Brojin, A. Golubović, Z. Dohčević-Mitrović, A. Kremenović, Z.V. Popović, Low-Frequency Raman Spectroscopy of Pure and La-Doped TiO<sub>2</sub> Nanopowders Synthesized by Sol-Gel Method, *Acta Physica Polonica A* 116 (2009) 99-102.
- [16] D. Wolfgang, *Laser Spectroscopy*, Springer-Verlag Berlin Heidelberg, 2008.
- [17] P. Mobili, A. Londero, G. De Antoni, A. Gomez-Zavaglia, Multivariate analysis of Raman spectra applied to microbiology: Discrimination of microorganisms at the species level, *Revista Mexicana de Física* 56 (2010) 378-385.
- [18] A. C. S. Talari, C. A. Evans, I. Holen, R. E. Coleman, Raman Spectroscopic analysis differentiates between breast cancer cell line, *Journal of Raman Spectroscopy* 46 (2015) 421–427.
- [19] T. Huser, C.A. Orme, C.W. Hollars, M. H. Corzett, R. Balhorn, Raman spectroscopy of DNA packaging in individual human sperm cell distinguishes normal from abnormal cells, *Journal of Biophotonics* 2 (2009) 322-332.
- [20] A. Nijssen, T.C. Bakker Schut, F. Heule, P.J. Caspers, D.P. Hayes, M.H. Neumann, G.J. Puppels, Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy, *Journal of Investigative Dermatology* 119 (2002) 64-9.
- [21] R.Y. Sato-Berru , E.V. Mejia-Uriarte , C. Frausto-Reyes , M. Villagran-Muniz , H. Murrieta S, J.M. Saniger, Application of principal component analysis and Raman spectroscopy in the analysis of polycrystalline BaTiO<sub>3</sub> at high pressure, *Spectrochimica Acta Part A* 66 (2007) 557–560.



- [22] I.T. Jolliffe, *Principal Component Analysis*, Second Edition, Springer Series in Statistics, 2002.
- [23] A. Thieullen, M. Ouladsine, J. Pinaton, Application of Principal Components Analysis to Improve Fault Detection and Diagnosis on Semiconductor Manufacturing Equipment, Control Conference (ECC), 2013.
- [24] G.A. Cherry, S.J. Qin, Multiblock principal component analysis based on a combined index for semiconductor fault detection and diagnosis, *IEEE Transactions on Semiconductor Manufacturing*, 19 (2006) 159-172.
- [25] M. Scarisoreanu, I. Morjan, R. Alexandrescu, R. Birjega, I. Voicu, C. Fleaca, E. Popovici, I. Soare, L. Gavrilă-Florescu, O. Cretu, G. Prodan, V. Ciupina, E. Figgemeier, Effects of some synthesis parameters on the structure of titania nanoparticles obtained by laser pyrolysis, *Applied Surface Science* 253 (2007) 7908–7911.
- [26] M. Scarisoreanu, R. Alexandrescu, I. Morjan, R. Birjega, C. Luculescu, E. Popovici, E. Dutu, E. Vasile, V. Danciu, N. Herlin-Boime, Structural evolution and optical properties of C-coated TiO<sub>2</sub> nanoparticles prepared by laser pyrolysis, *Applied Surface Science* 278 (2013) 295–300.
- [27] R. Alexandrescu, I. Morjan, M. Scarisoreanu, R. Birjega, E. Popovici, I. Soare, L. Gavrilă-Florescu, I. Voicu, I. Sandu, F. Dumitrache, G. Prodan, E. Vasile, E. Figgemeier, Structural investigations on TiO<sub>2</sub> and Fe-doped TiO<sub>2</sub> nanoparticles synthesized by laser pyrolysis, *Thin Solid Films* 515 (2007) 8438–8445.
- [28] R. Alexandrescu, I. Morjan, M. Scarisoreanu, R. Birjega, C. Fleaca, I. Soare, L. Gavrilă, V. Ciupina, W. Kylberg, E. Figgemeier, Development of the IR laser pyrolysis for the synthesis of iron-doped TiO<sub>2</sub> nanoparticles: Structural properties and photoactivity, *Infrared Physics & Technology* 53 (2010) 94–102.
- [29] R. Alexandrescu, M. Scarisoreanu, I. Morjan, R. Birjega, C. Fleaca, C. Luculescu, I. Soare, O. Cretu, C.C. Negrila, N. Lazarescu, V. Ciupina, Preparation and characterization of nitrogen-doped TiO<sub>2</sub> nanoparticles by the laser pyrolysis of N<sub>2</sub>O-containing gas mixtures, *Applied Surface Science* 255 (2009) 5373–5377.
- [30] M. Scarisoreanu, C. Fleaca, I. Morjan, A.-M. Niculescu, C. Luculescu, E. Dutu, A. Ilie, I. Morjan, L. Gavrilă Florescu, E. Vasile, C. I. Fort, High photoactive TiO<sub>2</sub>/SnO<sub>2</sub> nanocomposites

prepared by laser pyrolysis, *Applied Surface Science* (2016), <http://dx.doi.org/10.1016/j.apsusc.2016.12.122> .

[31] E. Valatka, Z. Kulėšius, TiO<sub>2</sub>-mediated photoelectrochemical decoloration of methylene blue in the presence of peroxodisulfate, *Journal of Applied Electrochemistry*, 37 (2007) 415-420.

[32] G. Mariappan, P. Vijayan, C. Suresh, K. Shanthi, Titania Nanoparticles Modified with Nitrogen: Enhanced Visible-light Photocatalytic Activity, *J. Environ. Nanotechnol*, 3 (2014) 67-72.

[33] A. V. Agafonov, A. V. Vinogradov, Sol–gel synthesis, preparation and characterization of photoactive TiO<sub>2</sub> with ultrasound treatment, *Journal of Sol-Gel Science and Technology*, 49 (2009) 180-185.

[34] M. Wojdyr, Fityk: a general-purpose peak fitting program, *Journal of Applied Crystallography* 43 (2010) 1126-1128.

[35] L.I. Smith, A Tutorial on Principal Components Analysis, *Information Fusion* 51:52 (2002).

[36] S. S. Ghoman, Z. Wang, P.C. Chen, R.K. Kapania, A POD-based Reduced Order Design Scheme for Shape Optimization of Air Vehicles, 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Honolulu, Hawaii, 2012.

[37] S. Sapatnekar, Overcoming variations in nanometer-scale technologies, *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 1 (2011) 5–18.

[38] W. Ma, Z. Lu, M. Zhang, Investigation of structural transformation in nanophase titanium dioxide by Raman spectroscopy, *Applied Physics A* 66 (1998) 621–627.

[39] T. Ohsaka, Temperature Dependence of the Raman Spectrum in Anatase TiO<sub>2</sub>, *Journal of the Physical Society of Japan* 48 (1980) 1661-1668.

[40] S. P. S. Porto, P. A. Fleury, T. C. Damen, Raman Spectra of TiO<sub>2</sub>, MgF<sub>2</sub>, ZnF<sub>2</sub>, FeF<sub>2</sub>, and MnF<sub>2</sub>, *Physical Review* 154 (1967) 522-526.

[41] H.C. Choi, Y.M. Jung, S. Bin Kim, Size effects in the Raman spectra of TiO<sub>2</sub> nanoparticles, *Vibrational Spectroscopy* 37 (2005) 33–38.

- [42] A. Turkovic, M. Ivanda, S. Popovic, A.M. Tonejc, M. Gotic, P. Dubcek, S. Music, Comparative Raman, XRD, HREM and SAXS studies of grain sizes in nanophase  $\text{TiO}_2$ , *Journal of Molecular Structure* 410 (1997) 271-273.
- [43] M. Gotic, M. Ivanda, S. Popovic, S. Music, A. Sekulic, A. Turkovic, K. Furic, Raman investigation of nanosized  $\text{TiO}_2$ , *Journal of Raman spectroscopy* 28 (1997) 555-558.
- [44] S. Music, M. Gotic, M. Ivanda, S. Popovic, A. Turkovic, R. Trojko, A. Sekulic, K. Furic, Chemical and micro structural properties of  $\text{TiO}_2$  synthesized by sol-gel procedure, *Materials Science and Engineering: B* 47 (1997) 33-40.
- [45] K.R. Zhu, M.S. Zhang, Q. Chena, Z. Yina, Size and phonon-confinement effects on low-frequency Raman mode of anatase  $\text{TiO}_2$  nanocrystal, *Physics Letters A* 340 (2005) 220–227.
- [46] M. Ivandaa, S. Music, M. Gotic, A. Turkovic, A.M. Tonejc, O. Gamulin, The effects of crystal size on the Raman spectra of nanophase  $\text{TiO}_2$ , *Journal of Molecular Structure* 480–481 (1999) 641–644.
- [47] S. K. Gupta, R. Desai, P. K. Jha, S. P. Sahoo, D. Kirin, Titanium dioxide synthesized using titanium chloride: Size effect study using Raman and Photoluminescence, *Journal of Raman Spectroscopy* 41 (2010) 350-355.
- [48] C.D. Terwilliger, Excess thermodynamic properties and size-dependent segregation phenomena in ultrafine-grained titanium dioxide, Massachusetts Institute of Technology, Dept. of Materials Science and Engineering, 1993.
- [49] Z.M. Shi, L. Yan, L.N. Jin, X.M. Lu, G. Zhao, The phase transformation behaviors of  $\text{Sn}^{2+}$ -doped Titania gels, *Journal of Non-Crystalline Solids* 353 (2007) 2171–2178.
- [50] I. Rangel-Vázquez, G. Del Angel, V. Bertin, F. González, A. Vázquez-Zavala, A. Arrieta, J.M. Padilla, A. Barrera, E. Ramos-Ramirez, Synthesis and characterization of Sn doped  $\text{TiO}_2$  photocatalysts: Effect of Sn concentration on the textural properties and on the photocatalytic degradation of 2,4-dichlorophenoxyacetic acid, *Journal of Alloys and Compounds* 643 (2015) S144–S149.
- [51] K.P. Burnham, D.R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.), Springer-Verlag, New York, 2002.
- [52] R.G. Lomax, *Statistical Concepts: A Second Course*, Routledge Taylor & Francis Group, New York, 2012.

## Figure captions

Fig.1. XRD analysis of SMT13-15 samples, confronted with a commercial sample P25 Degussa of nano-TiO<sub>2</sub>, emphasizing the main peaks for anatase (A) and rutile (R) titania.

Fig.2. TEM image of TiO<sub>2</sub> sample showing nanopowder morphology and particle diameter distribution of SMT14, in which the maximum value of the Log Normal fitting function is also displayed.

Fig.3. XRD diffraction patterns for TLO19-21 samples, revealing a higher amount of TiO<sub>2</sub> rutile phase in sample TLO19 (50% Rutile). For comparison, the bottom of the figure shows the diffractogram of the commercial P25 Degussa sample.

Fig.4. General TEM image of sample TLO16 and its grain size histogram, displaying spherical-shaped TiO<sub>2</sub> nanoparticles with 21.57 nm medium grain dimension.

Fig.5. Measured Raman spectra of TSPH samples a) and the spectral window processed for TSPH samples b). The high-frequency vibrational modes for anatase and rutile TiO<sub>2</sub> are visible in the selected range.

Fig.6. Logarithm of eigenvalue based information vs. PC number for TSPH samples.

Fig.7. Loadings for the three PCs used in TSPH analysis, showing the critical spectral points that determine the scores for each of the six Sn-doped TiO<sub>2</sub> samples.

Fig.8. Two-dimensional PCA-scores for first and second principal component that highlights a segregation between TSPH samples with carbon shells visible in their Raman spectra.

Fig.9. PCA values for first and third PC, showing a good correlation with tin concentration in TSPH samples.

Fig.10. Third degree polynomial fitting of the PC-3 scores with Sn/Ti percentage.

Fig.11. Second vs. third PC scores of TSPH samples indicating a good correlation between PCs values and nanocrystallite dimension.

Fig.12. Comparison between anatase crystallite sizes calculated from XRD measurements and PCA linear fittings of PC-2 and PC-3 values.

Fig.13. Raman spectra of SMT and TLO samples, in the spectral range processed by the PCA scripts, for different size TiO<sub>2</sub> anatase nanocrystallites: TLO4 - 28 nm, TLO20 - 25 nm, TLO9

– 18.7 nm, SMT13 -13 nm, SMT10 -8.1 nm. The selected area includes all the high-frequency Raman modes for anatase and rutile  $\text{TiO}_2$  as in the first analysed set.

Fig.14. Loadings corresponding to the first and second principal components, for the 24 analysed  $\text{TiO}_2$  un-doped samples, highlighting the spectral points which affect mostly the PCs scores.

Fig.15. Two-dimensional PCA-values for the first and second principal component, containing ~98% of data information, that exhibit a good correlation between PCs scores and anatase nanocrystallite size.

Fig.16. Comparison between  $\text{TiO}_2$  anatase nanocrystallite sizes calculated from XRD measurements and linear or linear logarithmic fitting of the PCs scores, for each sample. The SMT samples have smaller nanoparticle size, under 15 nm, while TLO samples grow up to 28 nm.

Figure 1

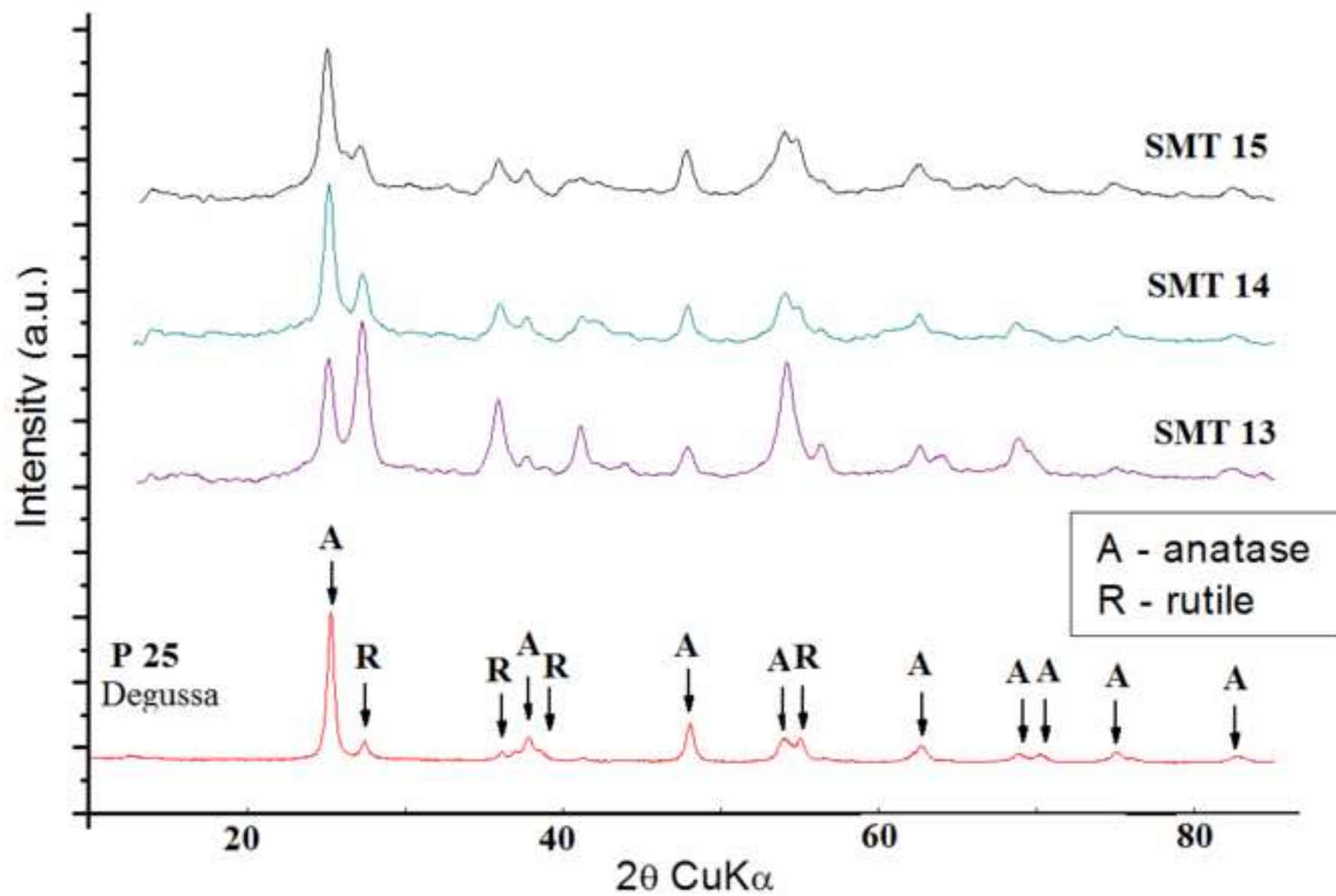


Figure 2

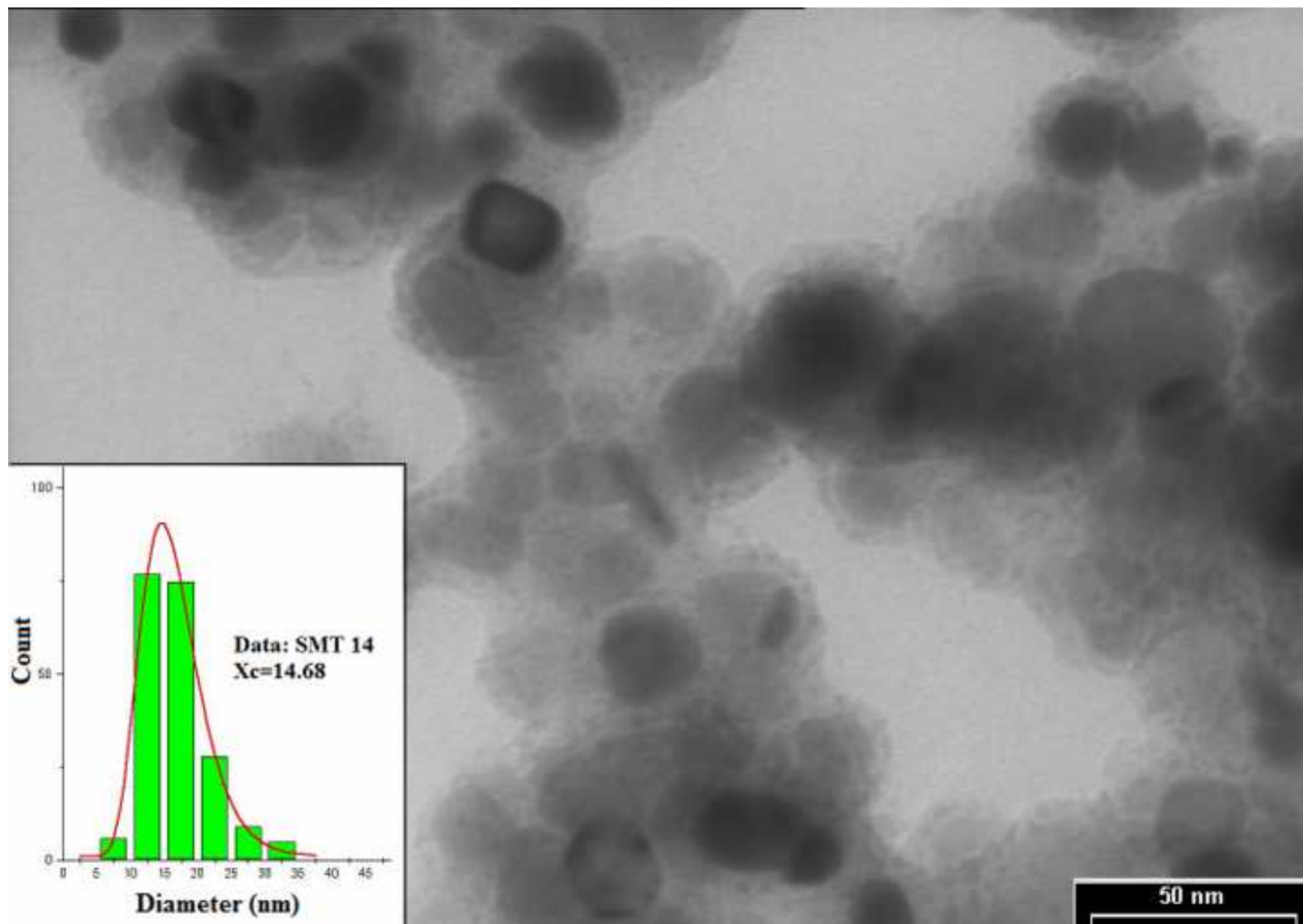


Figure 3

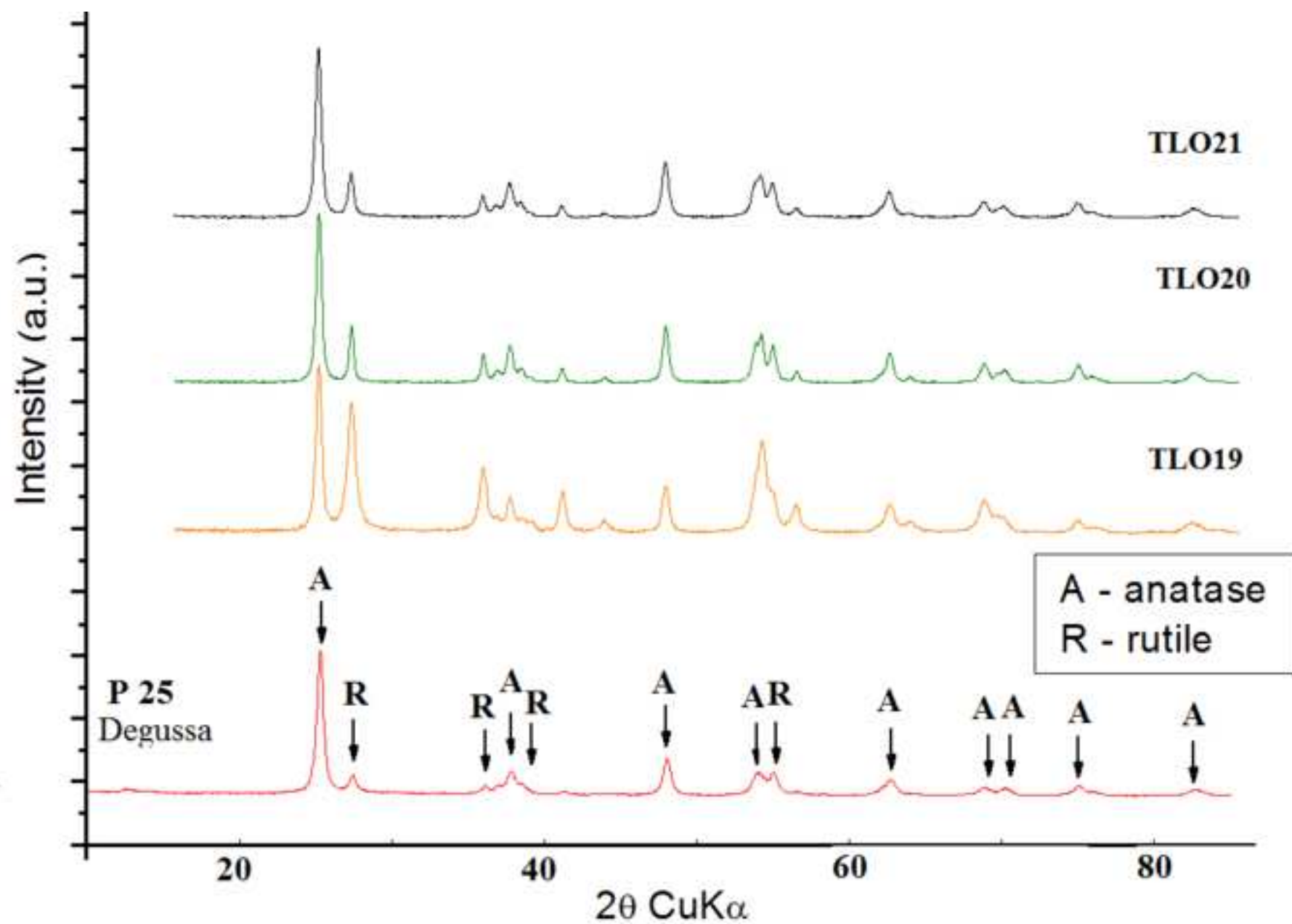




Figure 4

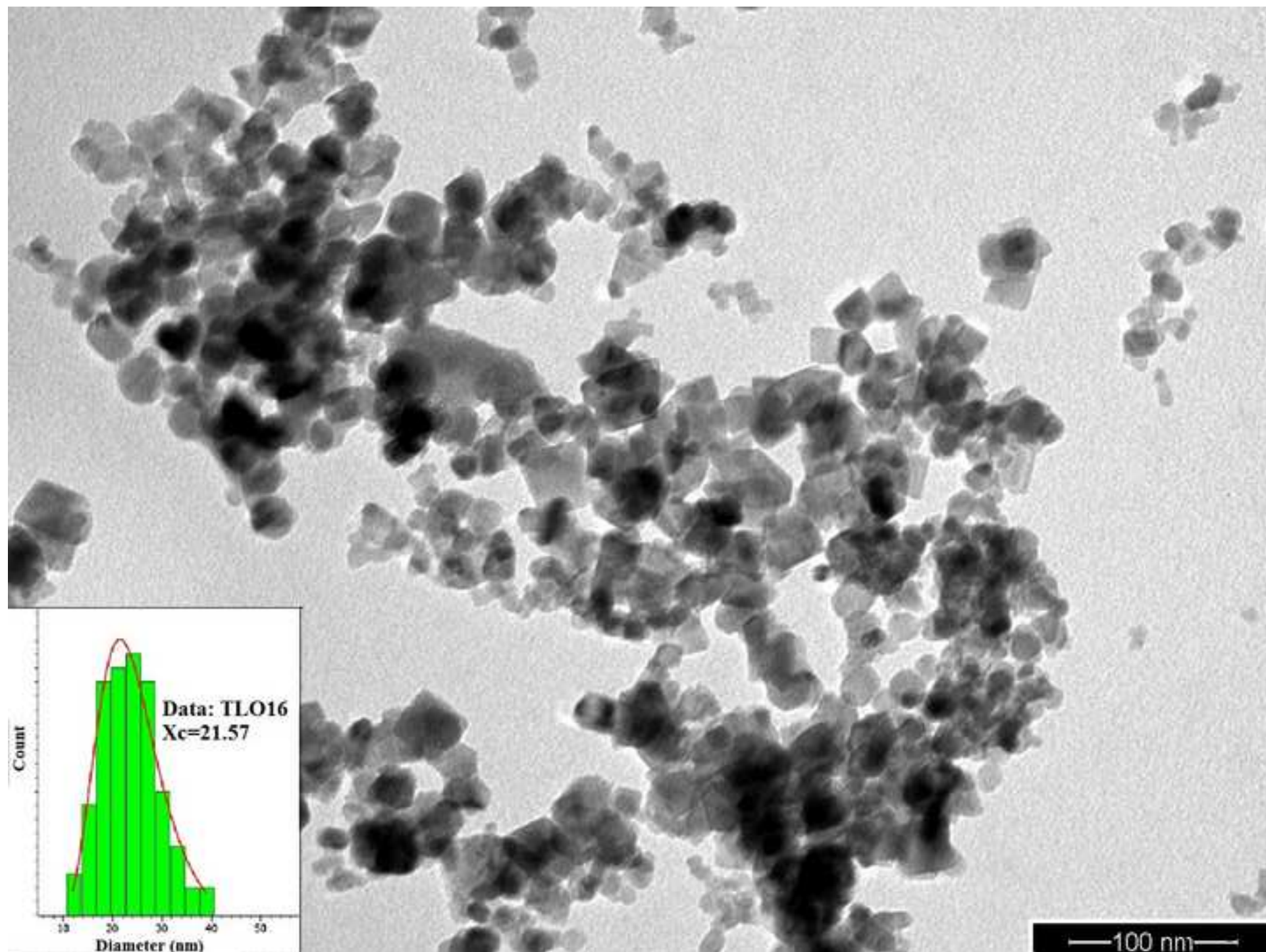


Figure 5

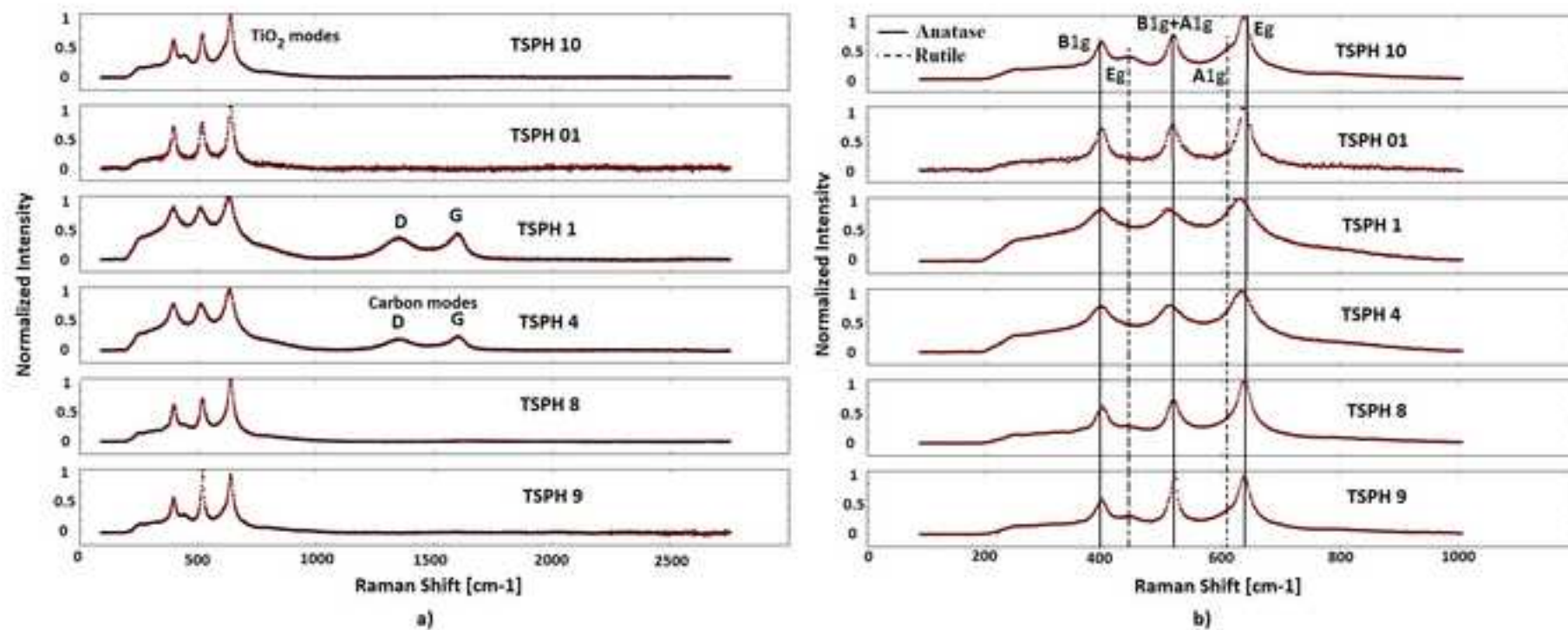


Figure 6

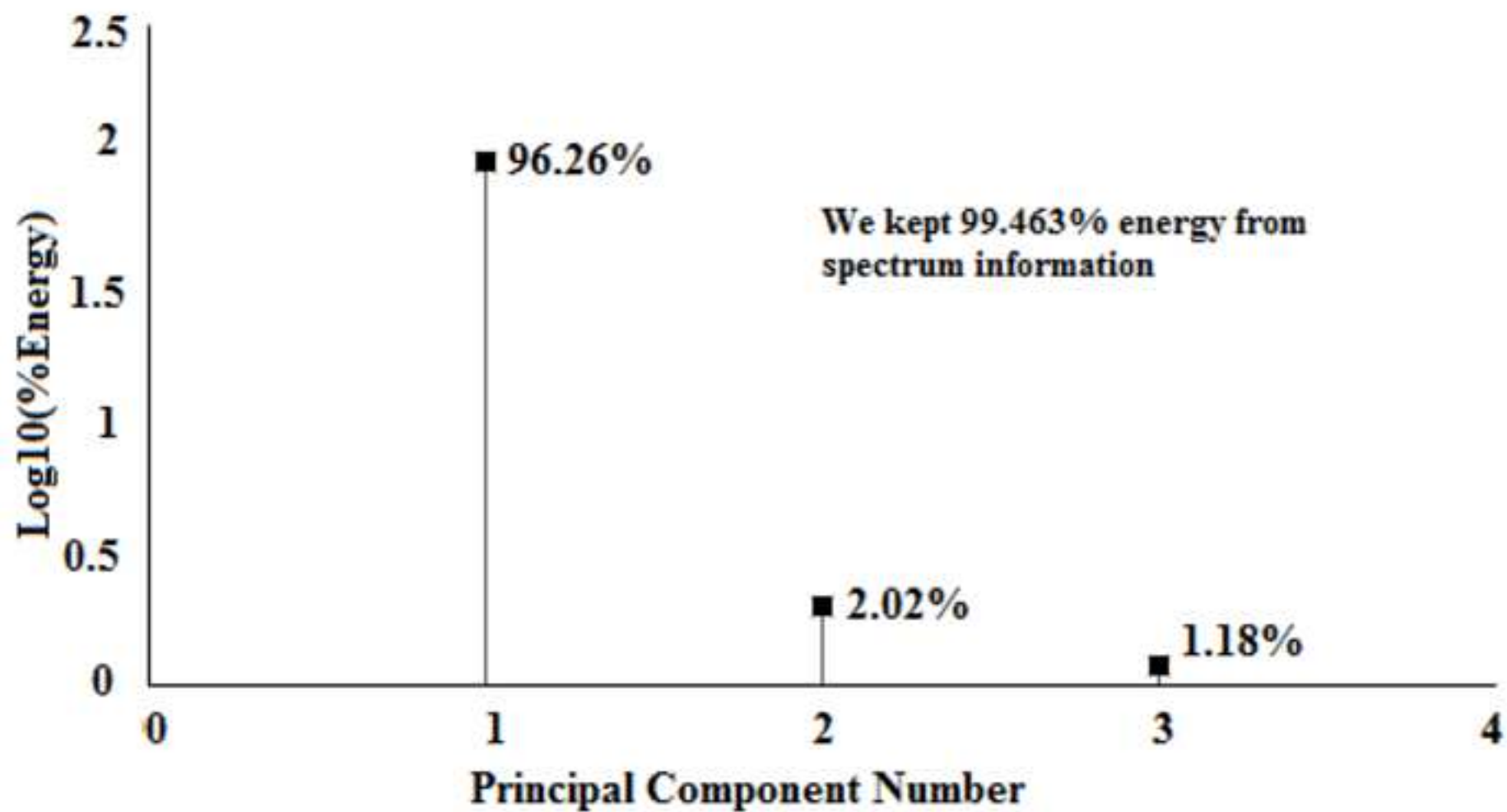


Figure 7

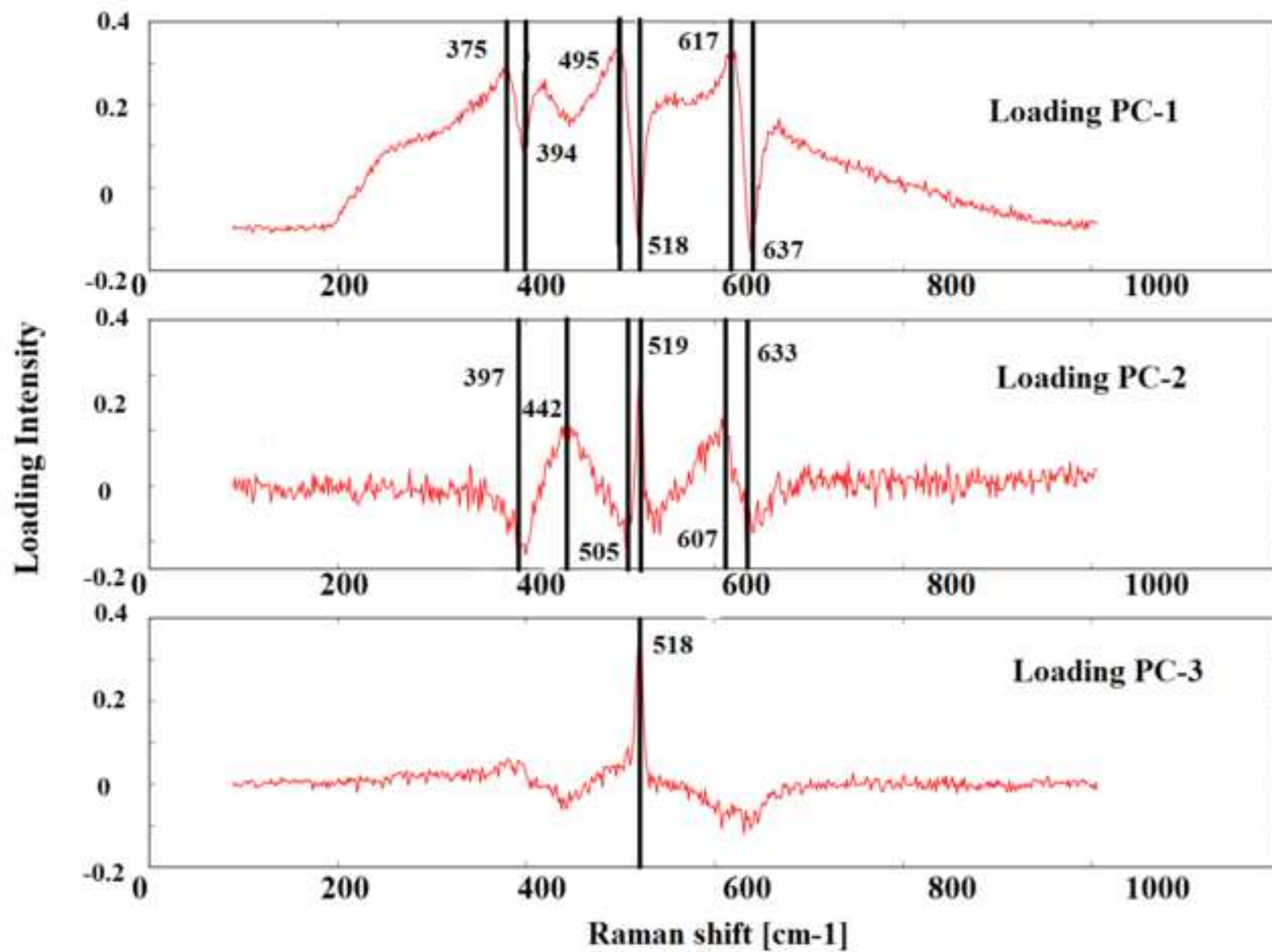


Figure 8

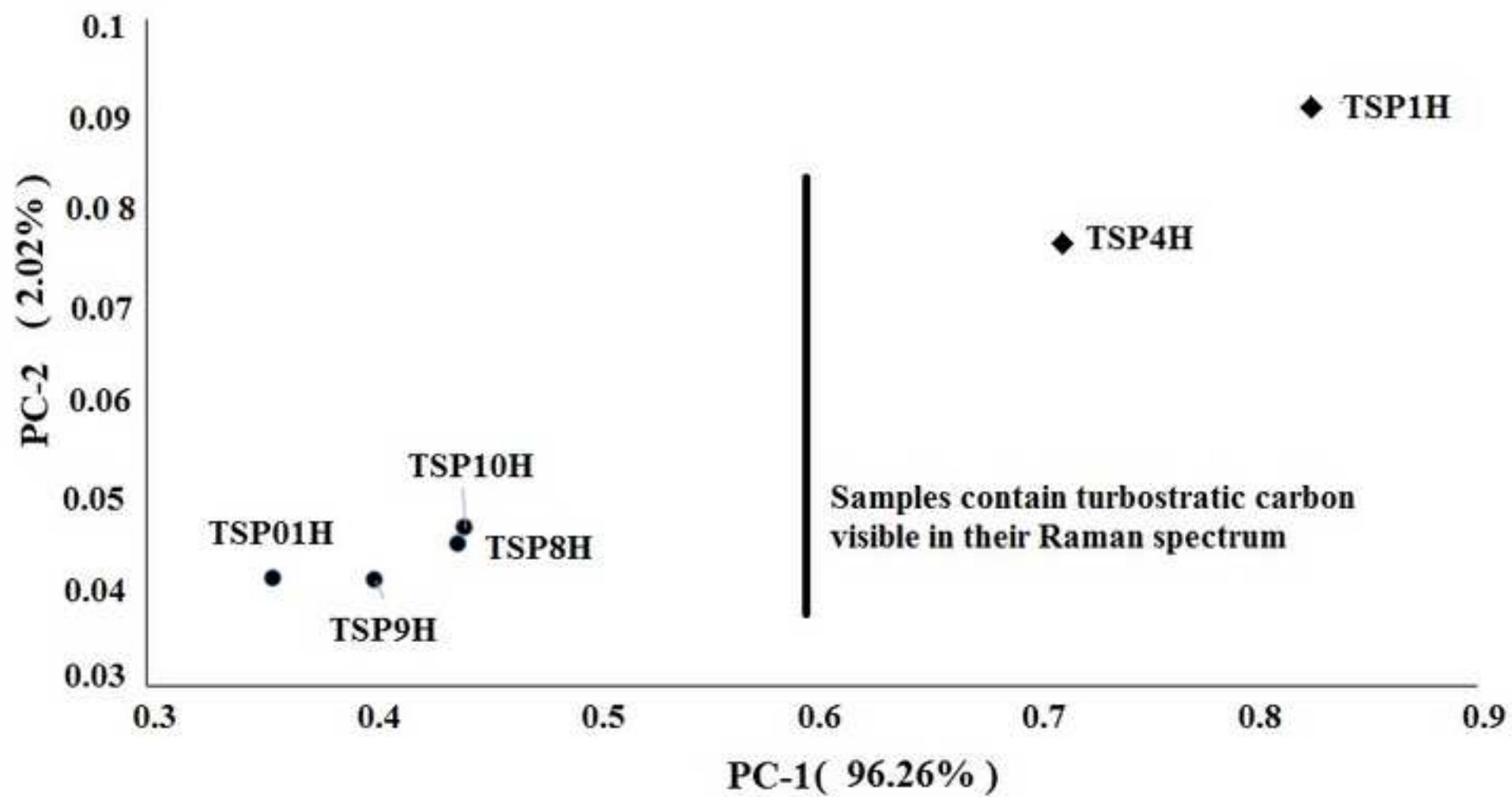


Figure 9

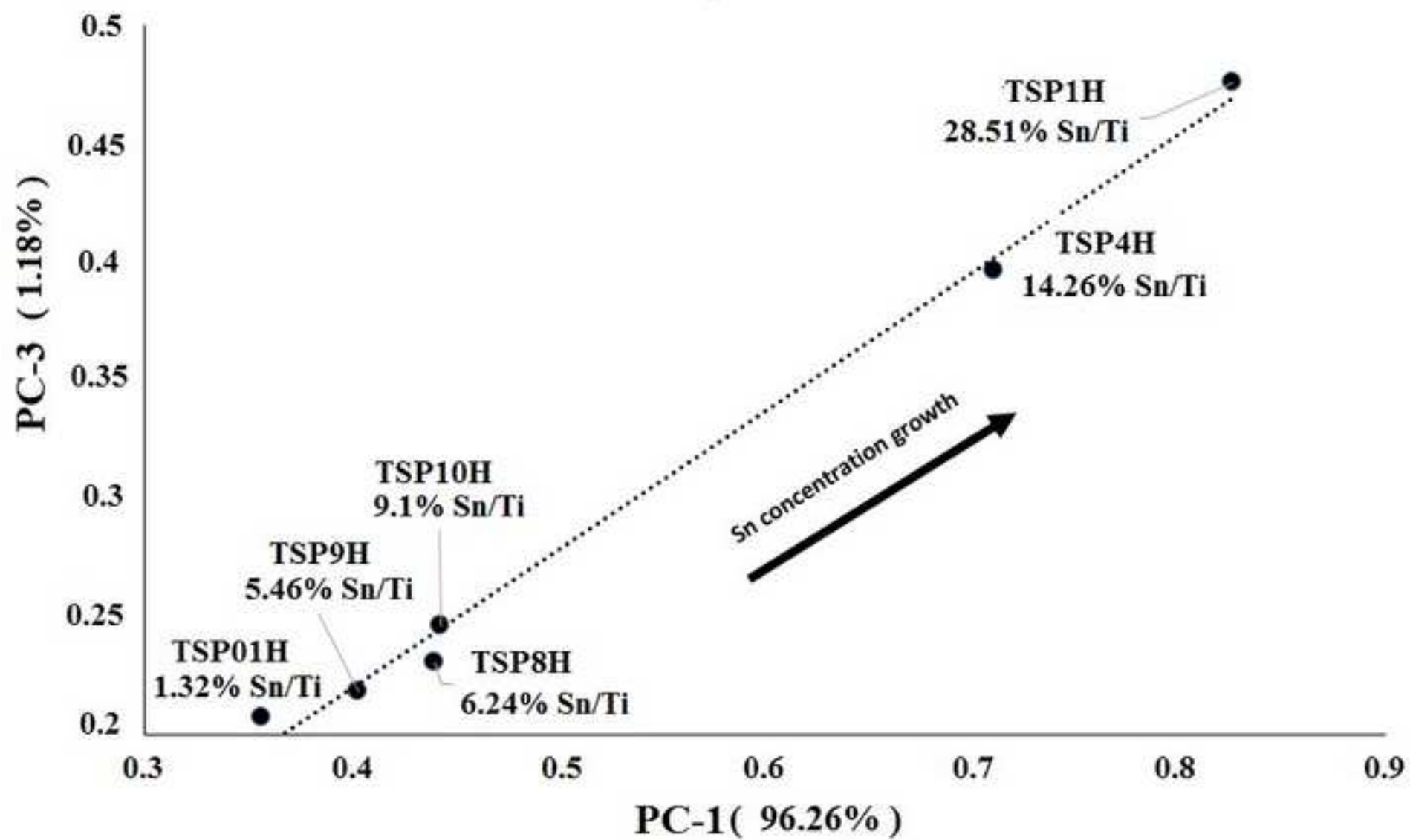




Figure 10

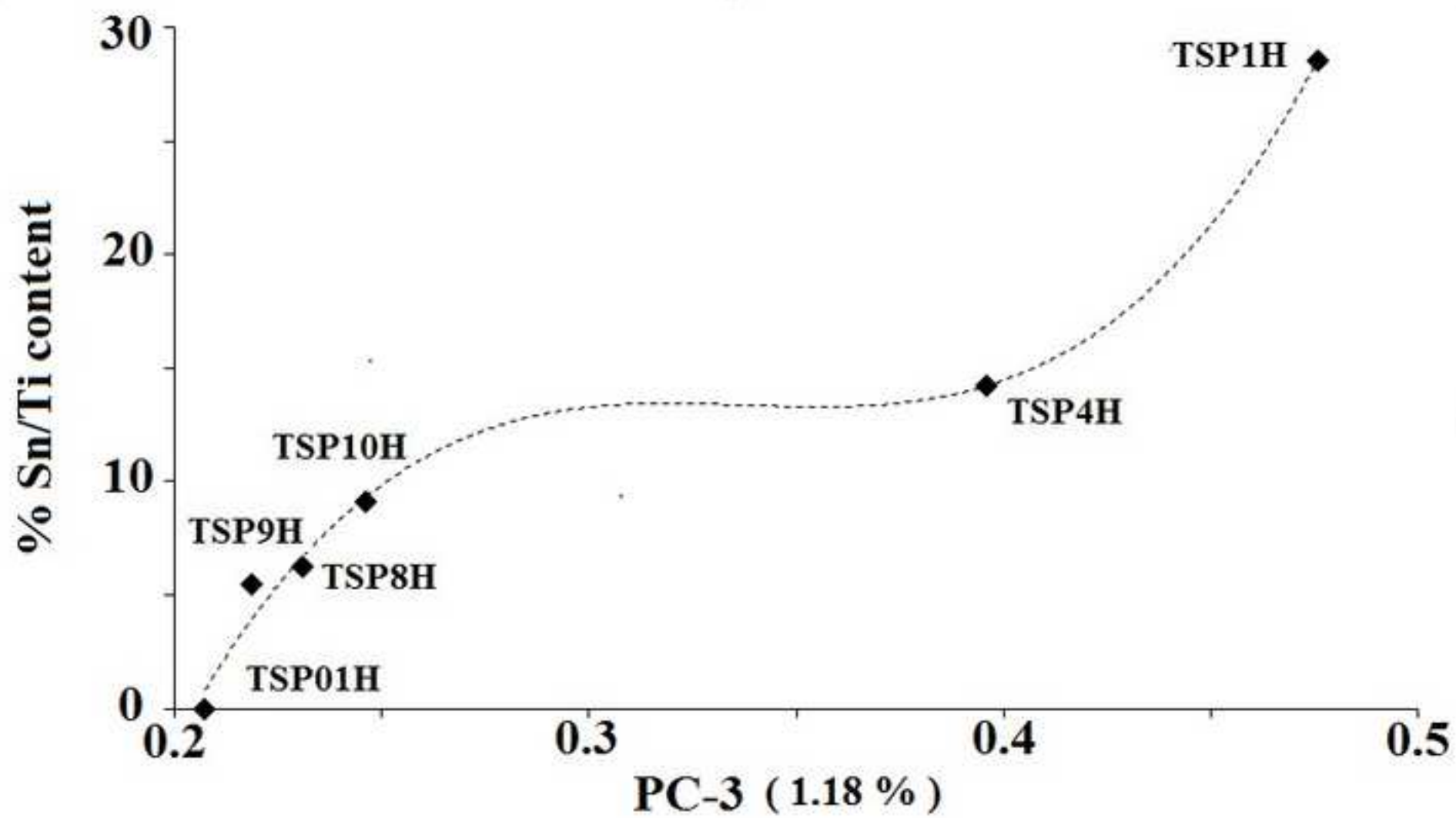


Figure 11

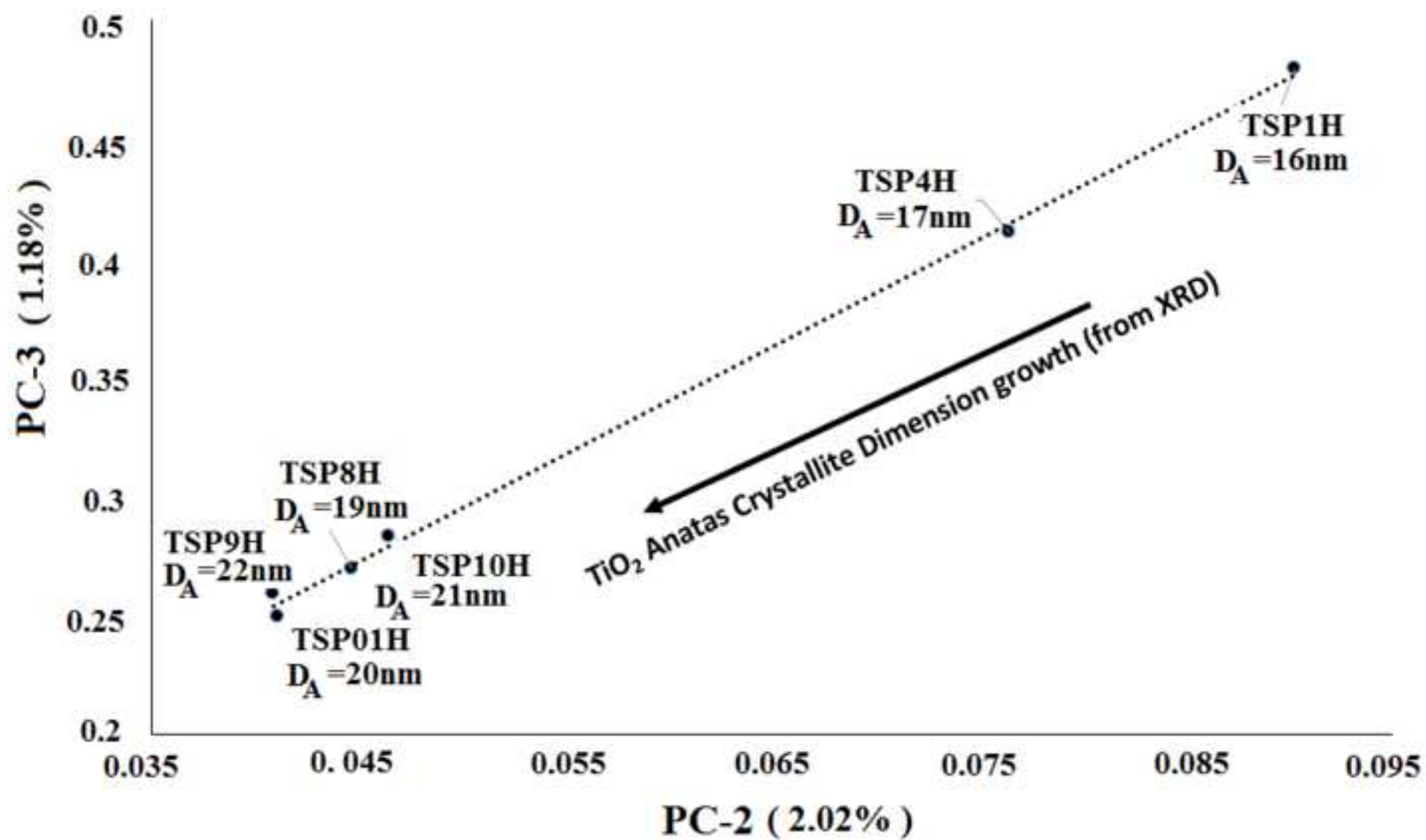




Figure 12

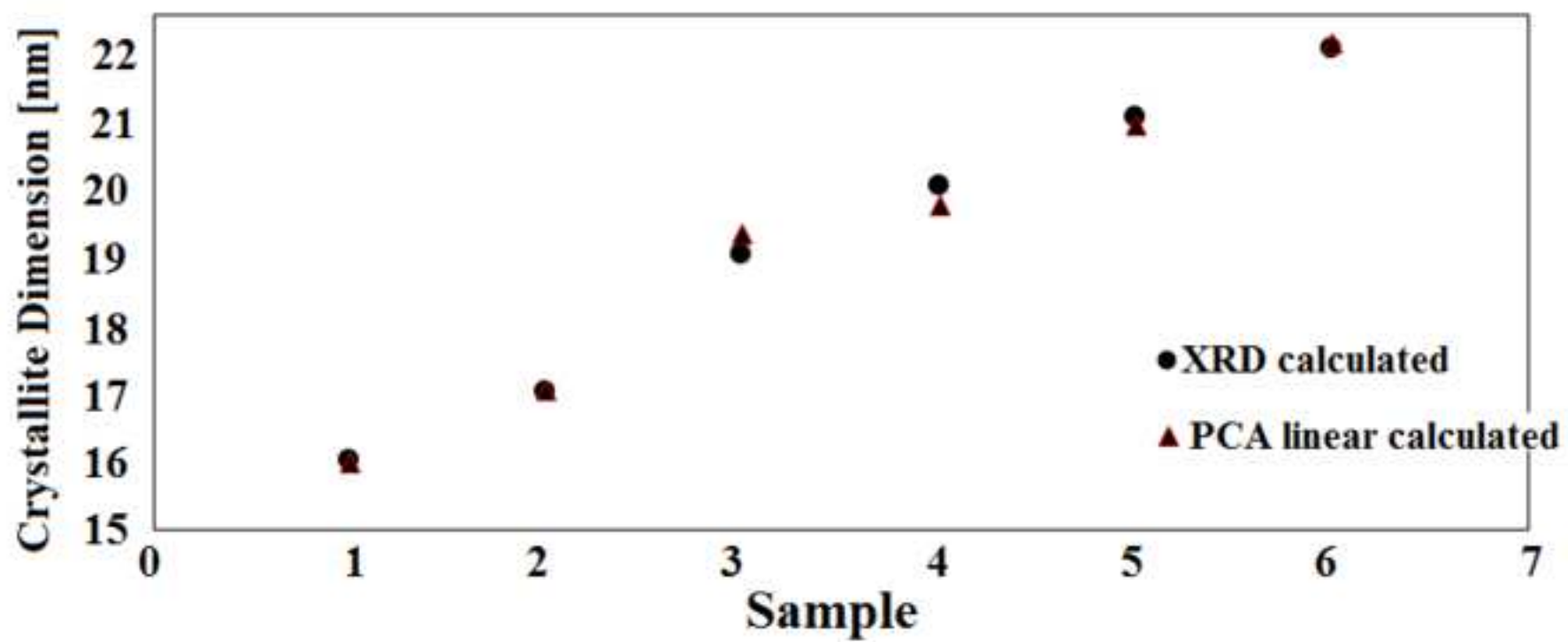


Figure 13

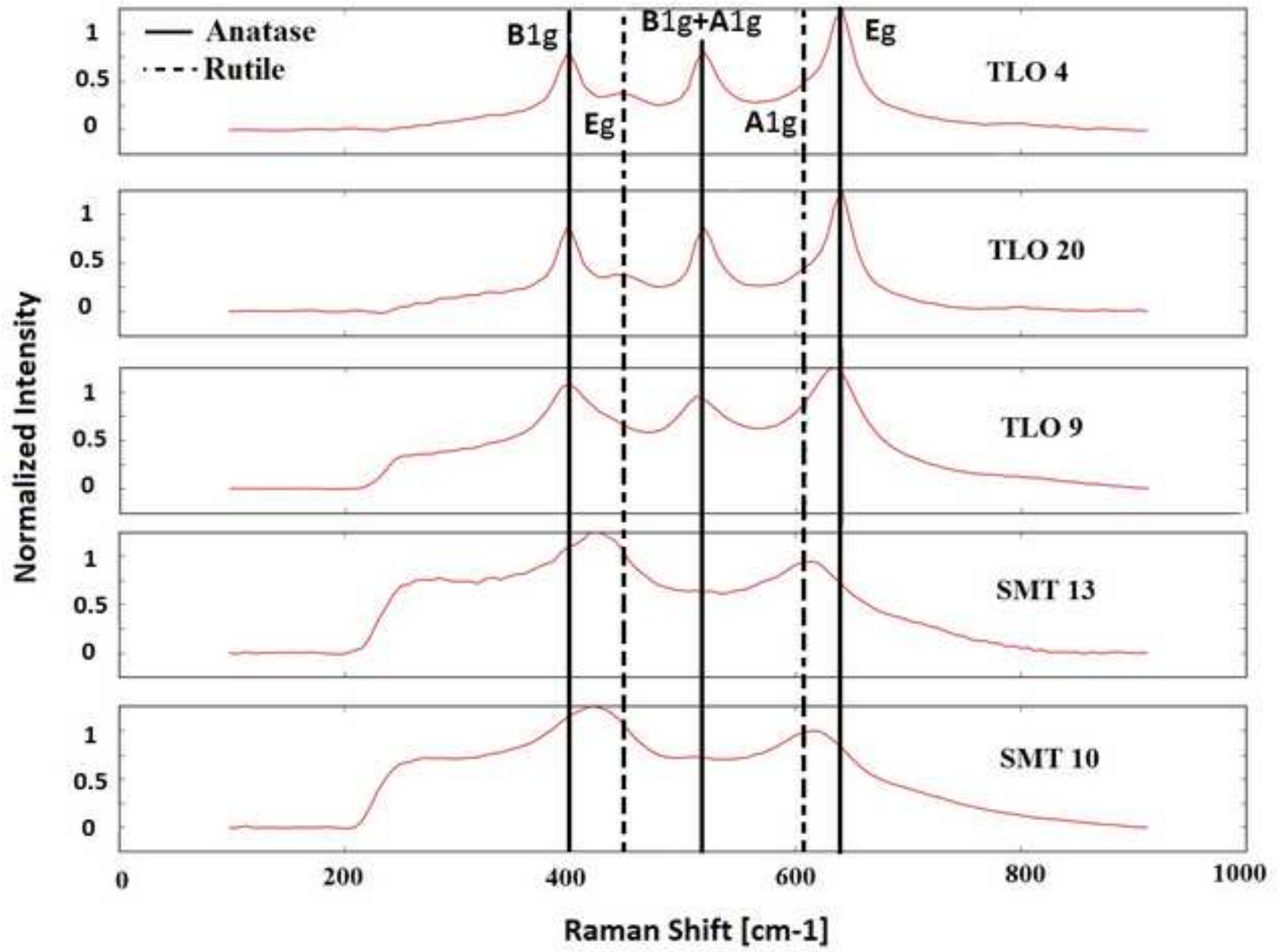


Figure 14

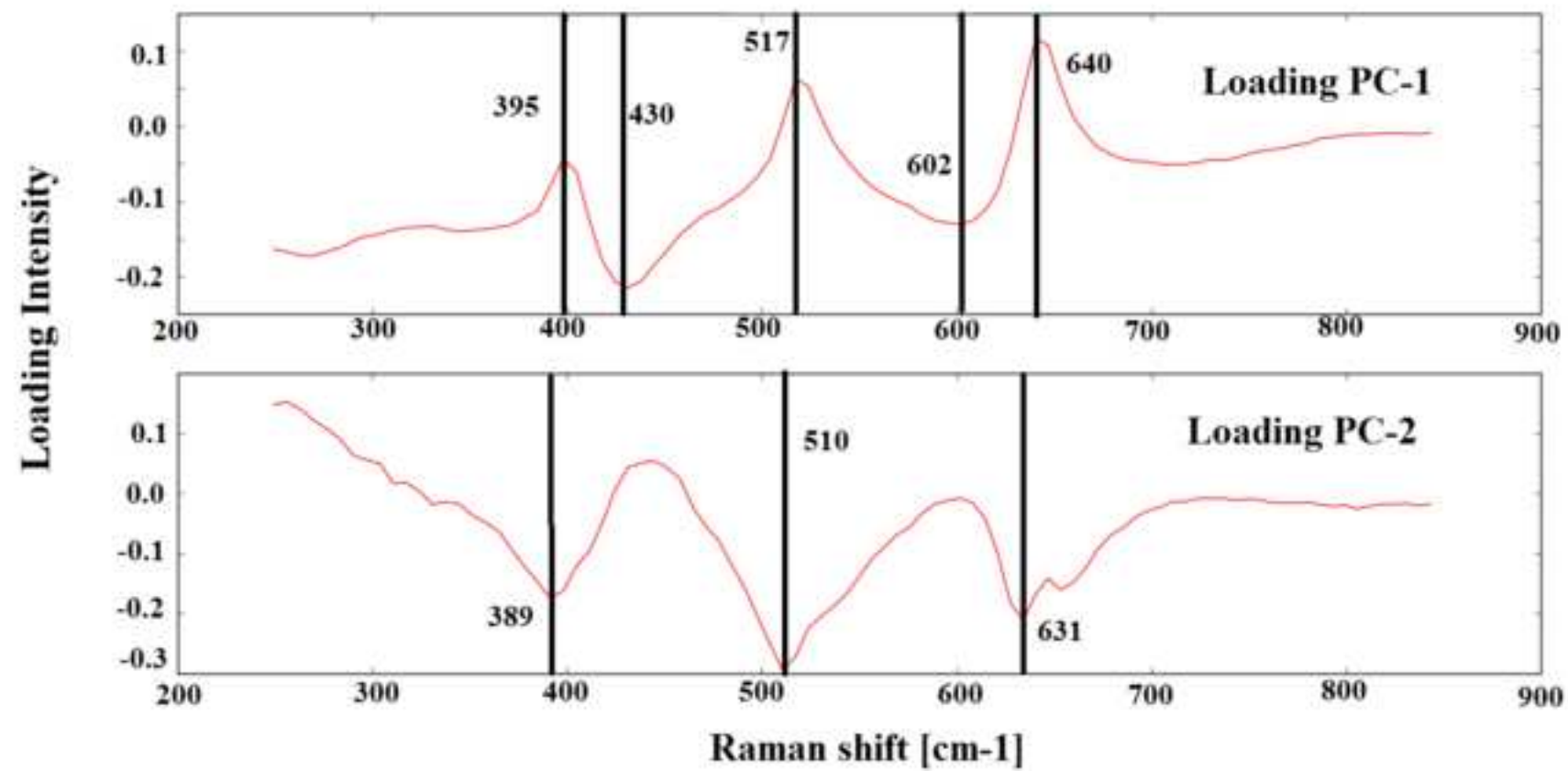


Figure 15

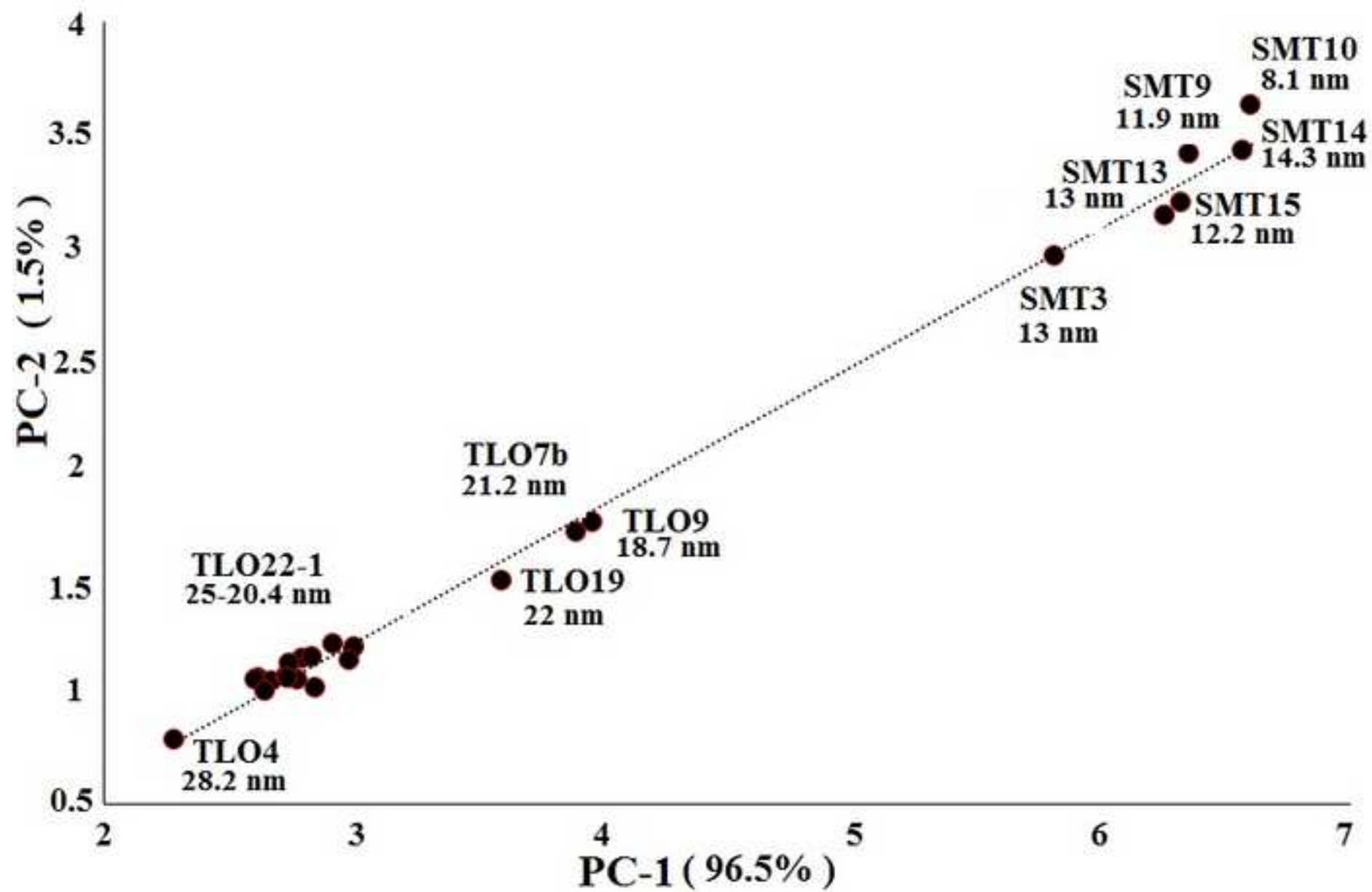
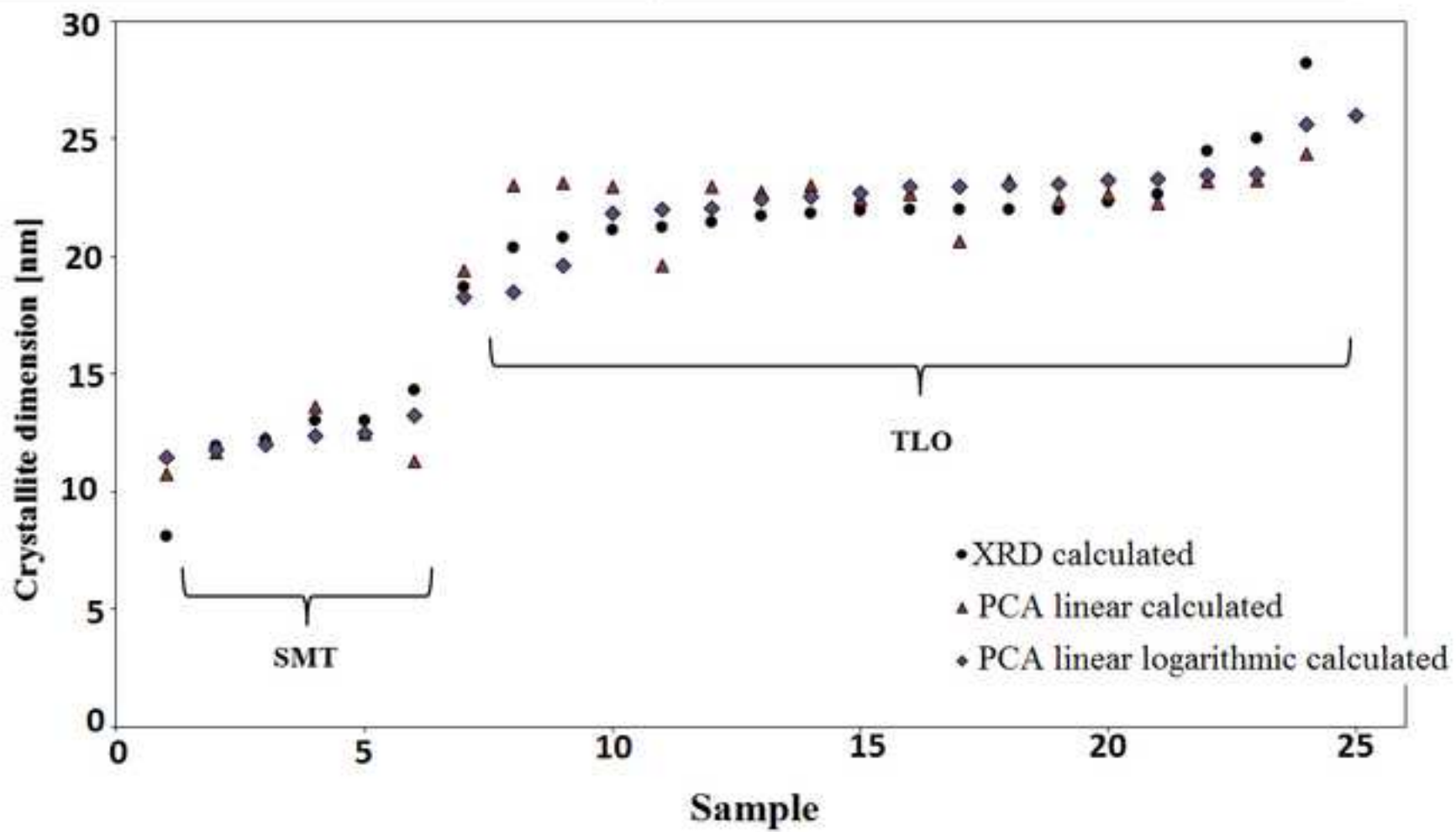


Figure 16



**Table 1.** TSPH Sample characteristics revealed by XRD and EDS measurements

| <b>Sample</b>                   | TSP01H | TSP1H | TSP4H | TSP8H | TSP9H | TSP10H |
|---------------------------------|--------|-------|-------|-------|-------|--------|
| <b>Sn/Ti [%]</b>                | 1.32   | 28.51 | 14.26 | 6.24  | 5.46  | 9.10   |
| <b>Sn/Ti [%]<br/>Polynomial</b> | 1.80   | 28.52 | 14.25 | 6.82  | 4.44  | 9.08   |
| <b>A/R [%]</b>                  | 100    | 77.32 | 82.71 | 82.54 | 69.87 | 65.41  |
| <b>D<sub>A</sub> [nm]</b>       | 20     | 16    | 17    | 19    | 22    | 21     |
| <b>DA linear<br/>PCA [nm]</b>   | 19.71  | 15.97 | 17.02 | 19.31 | 22.09 | 20.89  |

**Table 2.** SMT and TLO sample characteristics revealed by XRD and calculated using PCA analysis and linear, or linear logarithmic, fitting of PCs scores.

| Sample | A/R [%] | D <sub>A</sub> [nm] | D <sub>A</sub> linear logarithmic PCA [nm] | D <sub>A</sub> linear PCA [nm] |
|--------|---------|---------------------|--|--------------------------------|
| SMT3   | 74      | 13                  | 13.27                                      | 13.62                          |
| SMT9   | 72      | 11.9                | 12.02                                      | 11.66                          |
| SMT10  | 69      | 8.1                 | 11.48                                      | 10.74                          |
| SMT13  | 38      | 13                  | 12.49                                      | 12.49                          |
| SMT14  | 68      | 14.3                | 11.76                                      | 11.30                          |
| SMT15  | 74      | 12.2                | 12.34                                      | 12.25                          |
| TLO1   | 91.3    | 24.5                | 23.34                                      | 23.17                          |
| TLO2   | 90.3    | 22.3                | 22.54                                      | 22.68                          |
| TLO4   | 67.4    | 28.2                | 25.61                                      | 24.37                          |
| TLO6   | 91.8    | 21.1                | 22.96                                      | 22.95                          |
| TLO7b  | 88.8    | 21.2                | 18.46                                      | 19.62                          |
| TLO8   | 89.8    | 21.5                | 22.93                                      | 22.93                          |
| TLO9   | 92.3    | 18.7                | 18.23                                      | 19.40                          |
| TLO10  | 90.8    | 20.8                | 23.24                                      | 23.11                          |
| TLO11  | 89.2    | 20.4                | 23.04                                      | 22.99                          |
| TLO12  | 94.4    | 21.7                | 22.74                                      | 22.81                          |
| TLO14  | 90.2    | 22                  | 22.42                                      | 22.61                          |
| TLO15  | 88.1    | 22.7                | 21.83                                      | 22.24                          |
| TLO16  | 85.9    | 21.9                | 22.10                                      | 22.42                          |
| TLO17  | 89.4    | 21.8                | 23.03                                      | 22.99                          |
| TLO19  | 50.4    | 22                  | 19.61                                      | 20.60                          |
| TLO20  | 70.8    | 25                  | 23.42                                      | 23.22                          |
| TLO21  | 75.7    | 22                  | 23.49                                      | 23.26                          |
| TLO22  | 90.6    | 22                  | 22.00                                      | 22.34                          |