Terahertz Frequency-Domain Sensing Combined with Quantitative Multivariate Analysis for Pharmaceutical Tablet Inspection

International Journal of Pharmaceutics 632 (2023) 122545

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

journal homepage: www.elsevier.com/locate/ijpharm

Terahertz frequency-domain sensing combined with quantitative multivariate analysis for pharmaceutical tablet inspection

Anis Moradikouchi [a,b,*], Anders Sparén [b], Olof Svensson [b], Staffan Folestad [c], Jan Stake [a], Helena Rodilla [a]

[a] Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
[b] Oral Product Development, Pharmaceutical Technology & Development, Operations, AstraZeneca, Gothenburg, Sweden
[c] Innovation Strategies & External Liaison, Pharmaceutical Technology & Development, Operations, AstraZeneca, Gothenburg, Sweden

ARTICLE INFO

Keywords:
Terahertz spectroscopy
Frequency domain
Non-destructive
Multivariate analysis
Pharmaceutical tablets
Tablet density
API concentration

ABSTRACT

Near infrared (NIR) and Raman spectroscopy combined with multivariate analysis are established techniques for the identification and quantification of chemical properties of pharmaceutical tablets like the concentration of active pharmaceutical ingredients (API). However, these techniques suffer from a high sensitivity to particle size variations and are not ideal for the characterization of physical properties of tablets such as tablet density. In this work, we have explored the feasibility of terahertz frequency-domain spectroscopy, with the advantage of low scattering effects, combined with multivariate analysis to quantify API concentration and tablet density. We studied 33 tablets, consisting of Ibuprofen, Mannitol, and a lubricant with API concentration and filler particle size as the design factors. The terahertz signal was measured in transmission mode across the frequency range 750 GHz to 1.5 THz using a vector network analyzer, frequency extenders, horn antennas, and four off-axis parabolic mirrors. The attenuation spectral data were pre-processed and orthogonal partial least square (OPLS) regression was applied to the spectral data to obtain quantitative prediction models for API concentration and tablet density. The performance of the models was assessed using test sets. While a fair model was obtained for API concentration, a high-quality model was demonstrated for tablet density. The coefficient of determination (R^2) for the calibration set was 0.97 for tablet density and 0.98 for API concentration, while the relative prediction errors for the test set were 0.7% and 6% for tablet density and API concentration models, respectively. In conclusion, terahertz spectroscopy demonstrated to be a complementary technique to Raman and NIR spectroscopy, which enables the characterization of physical properties of tablets like tablet density, and the characterization of API concentration with the advantage of low scattering effects.

1. Introduction

The pharmaceutical industry needs new analytical tools for real-time monitoring and control of their products in the manufacturing area specifically suited for continuous manufacturing. Traditionally, in batch manufacturing, random samples of the finished products are transported to a laboratory to analyze and verify their quality. This process is slow and costly, with a potential risk that the final product should not meet its specifications, leading to batch failure. In this context, there has been a shift from the batch manufacturing process to continuous manufacturing (Plumb, 2005; Ierapetritou et al., 2016), where there is an increasing need for process analytical technologies that can be integrated into the manufacturing line and allow advanced process monitoring and control of the process and product.

Over the years, scientific and technological progress has led to several new techniques for the analysis of chemical and physical properties of drug products. Vibrational spectroscopy includes a series of key technologies that allow non-destructive, non-invasive, and rapid analysis of pharmaceutical materials (Ewing and Kazarian, 2018). NIR and Raman spectroscopy are established reliable techniques in process analytical technology for material analysis based on probing molecular structure and interactions. The use of these techniques combined with multivariate analysis such as partial-least-squares (PLS) and principal component analysis enables qualitative and quantitative analysis of pharmaceutical materials (Laske et al., 2017). NIR spectroscopy in diffuse reflection mode has been extensively used for material quantification (Berntsson et al., 2002) and probing blend uniformity (Lyon

* Corresponding author at: Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden.
E-mail address: anismo@chalmers.se (A. Moradikouchi).

https://doi.org/10.1016/j.ijpharm.2022.122545
Received 11 October 2022; Received in revised form 9 December 2022; Accepted 21 December 2022
Available online 26 December 2022

et al., 2002). However, this technique suffers from high scattering effects and low sample penetration depth (Berntsson et al., 1999), which restricts the extracted information to the surface properties. Moreover, its application for the analysis of low dosage material is limited (Sparén et al., 2015). Raman spectroscopy has been used for synthesis monitoring (Svensson et al., 2000), blend monitoring (Vergote et al., 2004), determination of polymorphs (Aina et al., 2010), and quantification of drugs (Jedvert et al., 1998). Raman spectroscopy offers more chemical specificity than NIR due to probing the fundamental vibrational modes. However, this technique is sensitive to particle size variations due to high scattering effects (Townshend et al., 2012), and if the laser power is too high, long exposure could burn the samples. Low-frequency Raman spectroscopy provides access to the lower frequencies from 300 GHz to 6 THz, but performs poorly when applied to compounds with strong fluorescence (Salim et al., 2020).

Terahertz spectroscopy (Siegel, 2004) is a complementary technique to the spectral information from NIR and Raman spectroscopy for the characterization and identification of materials (Bawuah and Zeitler, 2021). Spectroscopy in the terahertz region, 300 GHz to 10 THz, identifies the low frequency vibrations associated with intermolecular interactions and the morphology of formulated drugs (Claybourn et al., 2007). Moreover, Terahertz spectroscopy enables the characterization of the physical properties of materials like tablet density and porosity (Bawuah et al., 2020; Moradikouchi et al., 2022). Taday (Taday et al., 2004) demonstrated the use of terahertz pulsed spectroscopy for the quantification of paracetamol and aspirin tablets using a PLS calibration model. In a study by Zeitler et al. (2007), combined with principle component analysis, the sensitivity of terahertz pulsed spectroscopy to anhydrous and hydrated pharmaceutical materials was investigated. Hisazumi et al. (2012) demonstrated the applicability of terahertz reflectance spectroscopy with partial least squares regression for the quantification of API in tablets. Yang et al. (2021) used terahertz time-domain spectroscopy (THz-TDS) combined with support vector regression chemometric method to analyze the characteristic spectrum of caffeine in medicine. Peng et al. (2018) proposed THz-TDS combined with support vector regression chemometric method to characterize substances in brain glioma.

In this paper, we propose terahertz frequency domain spectroscopy (THz-FDS), based on electronic heterodyne techniques (Hübers, 2008), in combination with multivariate analysis for the analysis of API concentration and tablet density of pharmaceutical tablets as a complementary technique to NIR and Raman spectroscopy (Sparén et al., 2015). The high frequency resolution and stability of electronic heterodyne techniques minimizes the propagation of systematic errors when performing reference measurements (Hübers et al., 2011). A quantitative analysis of the tablets was performed including the following steps. First, the complex transmission coefficients in phase and amplitude (S_{21}) (Pozar, 2012) of the tablets were measured and the attenuation spectra were obtained. Then, the spectral data were pre-processed to remove noise and baseline variations. Orthogonal partial least squares (OPLS) regression was applied to the spectral data to establish a prediction model for the tablets. Finally, an independent test set was used to assess the performance of the prediction models for API concentration and tablet density.

2. Methodology

2.1. Tablet preparation

We studied tablets consisting of Ibuprofen as an API, Mannitol as a filler, and Magnesium stearate as a lubricant. A fraction of a full factorial design from Sparén et al. (2015) was used for our experiments. The API, filler, and lubricant were blended in a Turbula 2TF blender (Glen Mills Inc., Switzerland). Tablets were manufactured through direct compaction with a single punch press Korsch EK-0 (Korsch AG, Germany) equipped with flat round 10 mm punches. The design factors were Mannitol particle size, varied at two levels ($d_{50} \sim 91, 450$ μm), and API concentration varied at five levels (~16, 18, 20, 22, and 24 w/w%). Tablets with the applied compaction force of 12 kN were used for the experiments. The Magnesium stearate concentration was constant at ~1 w/w% for all tablets. As shown in Table 1, our sample set included 11 tablet types (A, B, …, M), and for each tablet type there were three replicate samples (A-1, A-2, A-3, B-1, …) to account for the variations in each tablet type, resulting in 33 samples. All tablets were flat-faced, with a nominal weight of 300 mg, a diameter of 10.0 mm, and a thickness, l, in the range of 2.97–3.25 mm, measured using a digital micrometer. The tablet density was calculated from the tablets dimensions and weight.

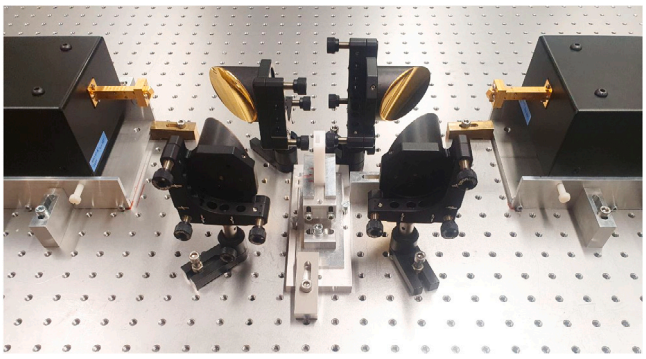

Fig. 1. A photograph of the experimental set-up showing THz S-parameter measurements in transmission mode.

2.2. Measurement set-up

To obtain the attenuation spectra of the tablets, the complex transmission coefficients were measured in frequency domain from 750 GHz to 1.5 THz, with the frequency step of 0.925 GHz. The measurement set-up consisted of a vector network analyzer (VNA) (Keysight PNA-X), frequency extenders WM380 (750 GHz–1.1 THz) and WM250 (1.1–1.5 THz), horn antennas, and four off-axis parabolic mirrors, see Fig. 1. The frequency range was selected based on the presence of the absorption peaks of Mannitol and Ibuprofen in the spectra. During measurements, the intermediate frequency bandwidth was set to 10 Hz in order to minimize the measurement noise level. The signal was focused on the center of the tablets using the parabolic mirrors with a focal length of 76.4 mm. The estimated mid-band beam diameter at the focal point is approximately 1.3 mm at 750 GHz to 1.1 THz and 0.9 mm at 1.1–1.5 THz. Each sample was measured twice by dismounting and mounting the tablets in the sample holder to account for the measurement errors like the white noise of the VNA and tablet displacement in the sample holder. This led to 66 measurements ranging from 750 GHz to 1.5 THz. Before each tablet measurement, the empty sample holder was measured as a reference for relative measurements, which was the ratio between the transmitted signal in the presence of the tablets versus the empty holder. The measurement environment could have been purged with dry nitrogen gas to remove the water vapor and to increase the SNR. In our case, we decided not to use this procedure in order to mimic the production line conditions.

2.3. Data processing

The attenuation coefficient of the samples, $\alpha(f)$, was calculated from the measurement of the scattering parameters using equation (1) (Chen et al., 2017):

$$\alpha(f) = -\frac{1}{l} log \left| \frac{T_s}{T_a} \right|^2, \quad (1)$$

Table 1
Design of experiments for the tablets, a fraction of a full factorial design from Sparén et al. (2015). Design factors are filler particle size at two levels (91, 450 μm) and API concentration at five levels (16, 18, 20, 22, 25 w/w%). There are three replicates for each tablet type.

Tablet	API particle size d_{50} (μm)	Filler particle size d_{50} (μm)	API concentration (w/w%)	Tablet density (g cm^{-3})
Type A				
A-1	71	91	16.8	0.3061
A-2	71	91	16.8	0.3059
A-3	71	91	16.8	0.3058
Type B				
B-1	71	450	16.0	0.3248
B-2	71	450	16.0	0.3248
B-3	71	450	16.0	0.3246
Type C				
C-1	71	91	18.5	0.3033
C-2	71	91	18.5	0.3031
C-3	71	91	18.5	0.3042
Type D				
D-1	71	450	18.4	0.3206
D-2	71	450	18.4	0.3198
D-3	71	450	18.4	0.3220
Type E				
E-1	71	91	20.6	0.3043
E-2	71	91	20.6	0.3043
E-3	71	91	20.6	0.3042
Type F				
F-1	71	450	19.5	0.3219
F-2	71	450	19.5	0.3209
F-3	71	450	19.5	0.3210
Type G				
G-1	71	91	22.7	0.3014
G-2	71	91	22.7	0.3014
G-3	71	91	22.7	0.3021
Type H				
H-1	71	450	22.9	0.3200
H-2	71	450	22.9	0.3207
H-3	71	450	22.9	0.3191
Type K				
K-1	71	91	24.8	0.3025
K-2	71	91	24.8	0.3014
K-3	71	91	24.8	0.3023
Type L				
L-1	71	450	25.0	0.3180
L-2	71	450	25.0	0.3167
L-3	71	450	25.0	0.3186
Type M				
M-1	95	211	20.1	0.3010
M-2	95	211	20.1	0.3009
M-3	95	211	20.1	0.3003

where l is sample thickness in cm, and T_s and T_a are the measured complex transmission coefficients (S_{21}) of the sample and empty sample holder (air), respectively, presented in supplementary data.

Noise, high-frequency oscillations, and a slanted baseline were present in the measured THz spectra. The oscillations were caused by standing waves in the set-up, and the slanted baseline was caused by scattering effects and frequency-dependent absorption. In order to reduce the effect of such phenomena, the spectra were pre-processed using a combination of smoothing, baseline correction, and normalization filters. It should be noted that in the case of amorphous materials, the baseline contains information about the content and should not be removed. The data pre-processing was performed in SIMCA 17 (Sartorius Stedim Data Analytics, Umeå, Sweden), which is a software for multivariate data analysis.

In the next step, the treated spectra were used to create a model for predicting the API concentration and tablet density, using multivariate analysis. The measured samples were divided into two sets, including a training set and a test set. The training set was used to build the model, and the test set was used to evaluate the robustness of the model for unknown samples. From the three replicates of each tablet type, the first two samples were used in the training set, and the third sample was used in the test set. This led to 44 samples in the training set and 22 samples in the test. The samples were labeled such that the first letter shows tablet type, the second part shows the number of the replicate sample, and the third part shows the number of measurements for each sample.

2.4. Multivariate analysis

Multivariate analysis is an important tool for qualitative and quantitative analysis of multivariate spectral data. We used orthogonal partial least squares (OPLS) regression (Trygg and Wold, 2002) to establish a prediction model for the quantification of API concentration and tablet density. OPLS is a modification of the traditional partial least squares (PLS) regression. OPLS separates the variability in the X spectral data into two parts, one that is predictive of the Y response and another that is orthogonal to Y. In the case of just a single Y variable and with the condition of the same number of components, both PLS and OPLS models fitted to the same data will give identical predictions. In the calibration set, the variances explained for the calibration set (R^2) and the cross-validation (Q^2), the root-mean-square of the calibration (RMSEC) and the cross-validation (RMSECV) were used to assess the model performance. Cross-validation is a model validation method that uses different portions of the calibration set iteratively to train and test the model. In order to further challenge the robustness of the model for the prediction of unknown samples, a test set was applied to the

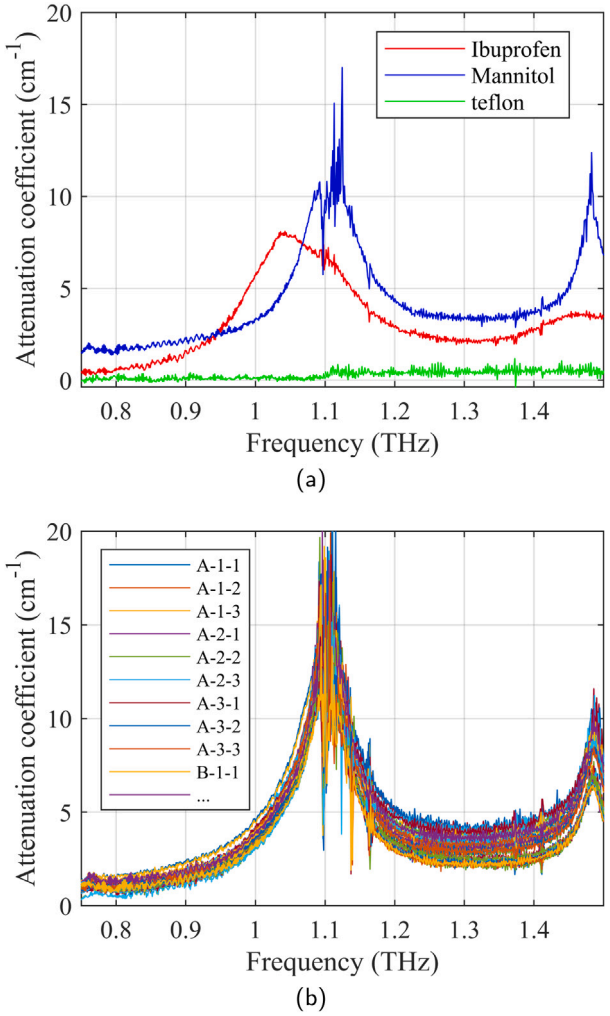

Fig. 2. Spectral data for (a) Ibuprofen, Mannitol, and Teflon as a reference. (b) samples from Table 1.

model. The root-mean-square error of prediction (RMSEP) and mean bias error for the prediction (MBEP) were used to assess how well the test samples fit the calibration model. Relative prediction error (RPE) is the RMSEP normalized with the mean value of the response, measuring the accuracy of the prediction regardless of the nature of the response.

3. Results and discussions

One tablet consisting of pure ibuprofen and one tablet of pure Mannitol were measured to identify the absorption peaks of each tablet ingredient individually across the frequency range. Fig. 2a shows the attenuation spectra of pure Ibuprofen, pure Mannitol, and Teflon. Ibuprofen showed two broad peaks centered around 1.05 THz and 1.47 THz. Mannitol showed two peaks centered around 1.10 THz and 1.48 THz, in agreement with Mannitol type β (Allard et al., 2011). The peaks occurring at around 1.10 THz were not completely resolved due to the lower signal-to-noise ratio (SNR) at the beginning of the frequency band of the extenders, the low SNR at absorption peaks, and water absorption lines. Teflon was just used for reference and as expected, did not show any absorption peaks. Fig. 2b shows the attenuation spectra of all measured tablets, and like Mannitol, the peaks at around 1.10 THz were not completely resolved (see supplementary material). In the THz spectra of the samples, the peak seen at around 1.1 THz and below 1.5 THz are the superposition of the Ibuprofen and Mannitol peaks that heavily overlap. Therefore, simple univariate calibration method would not be selective for this formulation.

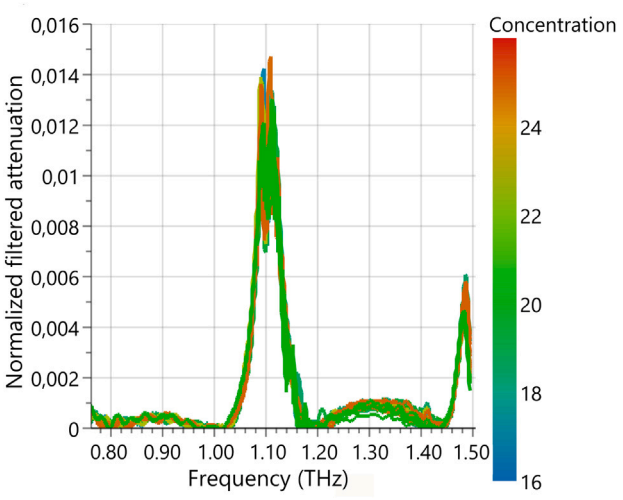

Fig. 3. Pre-processed spectral data for API concentration, after noise effect reduction, baseline correction, and normalization.

3.1. OPLS model for API concentration

To model the API concentration, first the attenuation data were pre-processed, and then an OPLS analysis was applied to the data to obtain a prediction model. For the pre-processing, a Savitzky–Golay smoothing filter (Savitzky and Golay, 1964) was used to reduce the effect of noise and oscillations. The filter was calculated from moving quadratic sub-models with 31 data points long, and edge effects were excluded. Then, an asymmetric least squares algorithm (Peng et al., 2010) was used for each spectrum to correct baseline variations. Finally, each spectral value was normalized with respect to the sum of the signal intensity over the frequency range. The treated attenuation spectra of the tablets are shown in Fig. 3, where different API concentrations are marked with different colors.

In the next step, OPLS regression was applied to the scaled and centered training set to establish a prediction model for API concentration, having the frequency points as the X independent variables and API concentration as the Y response variable. For the characterization of API, the parts of the spectrum including the absorption peaks have the dominant role in modeling. A first data evaluation revealed differences in the THz spectra of the tablets measured on a second day under higher humidity conditions, which complicated the modeling. After excluding those measurements affected by humidity variation (tablets B and H), we succeeded in developing a good model with R^2 and Q^2 values of 0.98 and 0.92, respectively. The training set consisted of 36 samples and the test set of 18 samples.

An OPLS score plot in Fig. 4 shows the clusters of the training samples based on their similarity for modeling the API concentration. The variations across the horizontal axes show the predictive component for the response, and the variations across the vertical axes show the orthogonal components, which are unrelated to the response and express the variations within clusters. In this plot, the training samples are clustered across the horizontal direction corresponding to the five API levels. Tablets are ordered from left to right by increasing API concentration. Moreover, the variations between replicate samples for each tablet type (tablets labeled with the same first digit but different second digit) and the repeatability of the measurements (different third digit in the label) can be observed in this plot.

Fig. 5a shows the observed versus predicted values for API concentration based on the OPLS regression. The R^2 was 0.98, and the RMSEC and RMSECV were 0.40 and 0.77, respectively. Five clusters of samples are observed corresponding to the five levels of API concentration. To challenge the robustness of the OPLS model, a test set was used.

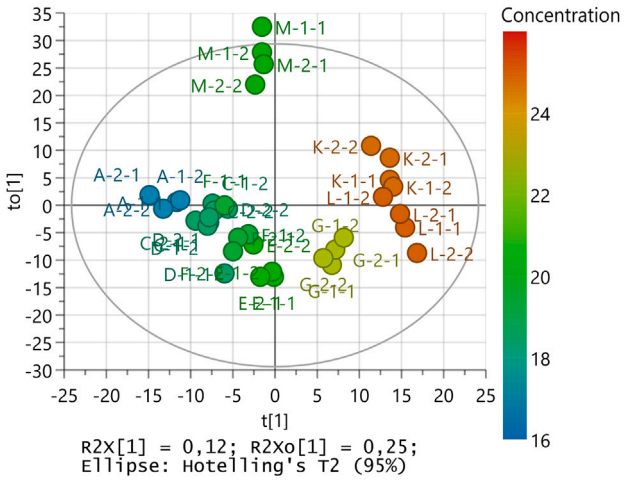

Fig. 4. Score plot for API concentration. Tablets are clustered with respect to the API concentration across the horizontal axes. Tablets are labeled with three digits that represent type, replicate, and measurement number, respectively.

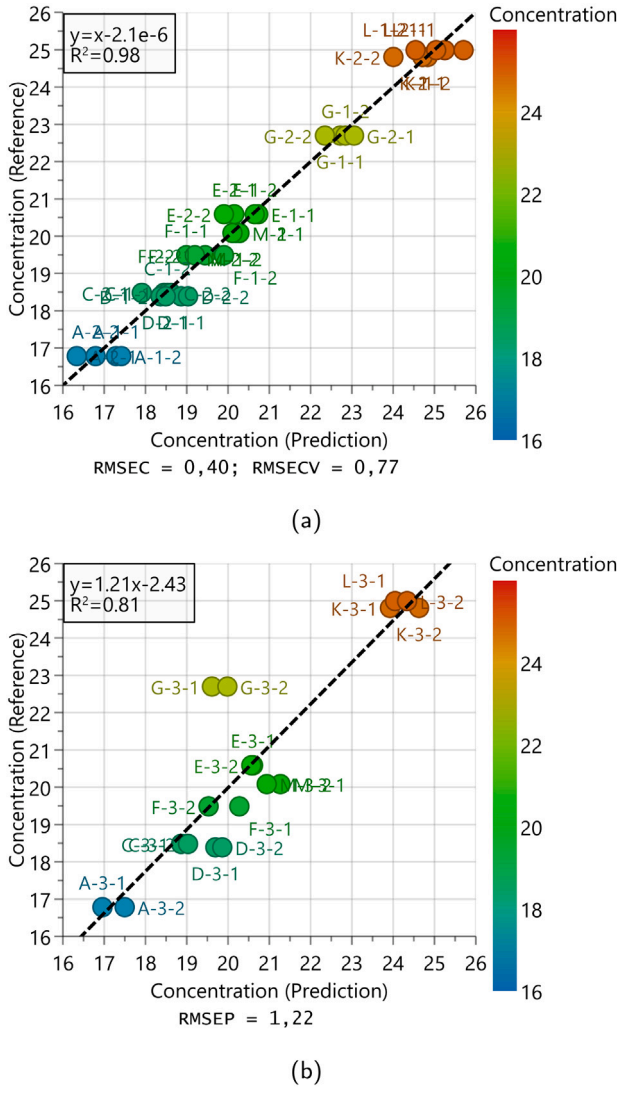

Fig. 5. Prediction plot for API concentration (a) calibration set (b) test set. Tablets are labeled with three digits that represent type, replicate, and measurement number, respectively.

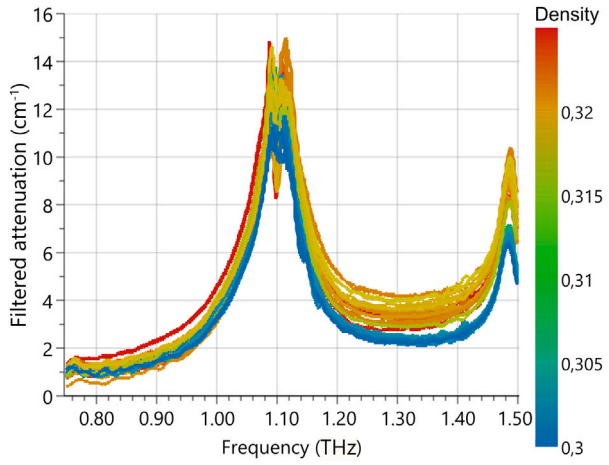

Fig. 6. Pre-processed spectral data for the quantification of tablet density after applying Savitzky–Golay smoothing filter for removing the high frequency noise.

Fig. 5b shows the comparison between the reference and prediction results for the test samples with five clusters corresponding to the API concentration. For Sample G-3, we obtained predicted values about 3% lower than the expected API level. One possible reason for this could be that the API may not have been evenly distributed across the tablet, while the THz beam illuminated mainly the center of the samples, which could lead to sub-sampling. The terahertz spectrum for this sample indicated a lower level of API than expected (data not shown), which supports this hypothesis. The relative prediction error and RMSEP were 6% and 1.22, respectively, which were less good than the results in paper (Sparén et al., 2015) because Raman and NIR techniques spectra have more selective bands and contain more chemical information than THz spectroscopy. It must be noticed that in Sparén et al. (2015), a larger number of tablets were included in the model. The quantification of the API concentration could probably be further improved by measuring in a controlled environment and completely resolving the absorption peaks. The latter can be achieved by purging the measurement system to remove the water lines and/or averaging the measurements over time, with the cost of increasing the complexity of the system and the measurement time.

3.2. OPLS model for tablet density

To model the tablet density, the spectral data were treated to reduce the noise and oscillations using the Savitzky–Golay smoothing filter. In this case, the baseline correction and data normalization were not conducted because of the correlation between the baseline and tablet density through the particle size and porosity (Moradikouchi et al., 2022). Fig. 6 shows the treated spectra after applying the smoothing filter, where different tablet densities were assigned with a different color. As can be observed, the spectral data were clustered into two groups, corresponding to the tablets with different initial particle sizes. The attenuation for tablets with smaller initial particle sizes was lower than those with larger particles. At the end of the band, tablets with larger particle size showed higher attenuation due to scattering effects (Shen et al., 2008; Wu et al., 2008).

Next, OPLS regression was applied to the 44 training samples, having the spectral data centered, to establish a prediction model for the tablet density. For the characterization of the tablet density based on the attenuation, higher frequencies play the dominant role due to the correlation between tablet density and particle size. The OPLS score plot in Fig. 7 shows two clusters of training samples across the horizontal direction with respect to the particle size, expressing the correlation between tablet density and the variations in the particle

Table 2
A summary of the OPLS models and their performance for the quantification of API concentration and tablet density.

Model	Calibration set	Test set	R^2	Q^2	RMSEC	RMSEP (test set)	MBEP (test set)	RPE (test set)
API concentration	36 samples	18 samples	0.98	0.92	0.40	1.22	−0.06	6.0%
Tablet density	44 samples	22 samples	0.97	0.94	0.002	0.002	−0.0002	0.7%

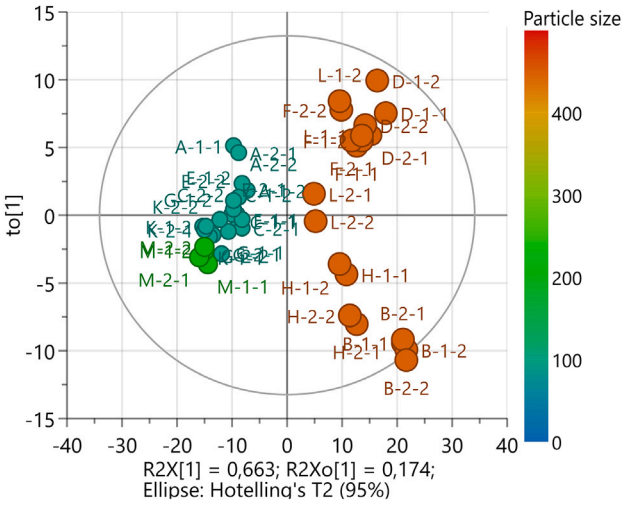

Fig. 7. Score plot for tablet density. Tablets are clustered across the horizontal axes with respect to particle size. Tablets are labeled with three digits that represent type, replicate, and measurement number, respectively. Small and large particle sizes are shown with small and large circles, respectively.

size. Tablets with larger particles are clustered on the right side marked with larger circles, and tablets with smaller particle sizes are on the left side marked with smaller circles. The vertical direction expresses the variation within the clusters due to API concentration.

Fig. 8a shows the reference versus prediction values for tablet density. The RMSEC and RMSECV for the obtained OPLS model were both 0.002, and the R^2 was 0.97. These values show the high accuracy and the goodness of the regression model. It should be noticed that for the quantification of tablet density, samples measured under higher humidity condition (B and H) were not excluded, indicating that humidity variation in the lab environment was not as critical as for quantifying the API concentration.

The robustness of the obtained OPLS model was further assessed by applying the 22 test samples to the model. Fig. 8b presents the comparison between the reference and prediction results for the test samples. The prediction of tablet density on the test samples gave a relative prediction error of 0.7%, and RMSEP and MBEP of 0.002 and 0.0002, respectively. These results show that the OPLS model for tablet density is significantly accurate and precise and performs better than the OPLS model for API concentration. The summary of the two models' performance is given in Table 2. Moreover, in both Figs. 8 a and b, we see two clusters of tablets corresponding to the small and large particle sizes marked with smaller and larger circles respectively, showing that initial particle size impacted the tablet density. Besides that, the results show that within each cluster of tablets, as the API concentration increased, the tablet density decreased being in line with the results from our previous study (Moradikouchi et al., 2022). Also, it can be observed that initial particle size has a higher impact than API concentration on the tablet density.

4. Conclusion

In this paper, we explored the THz-FDS technique together with multivariate data analysis for the characterization of pharmaceutical

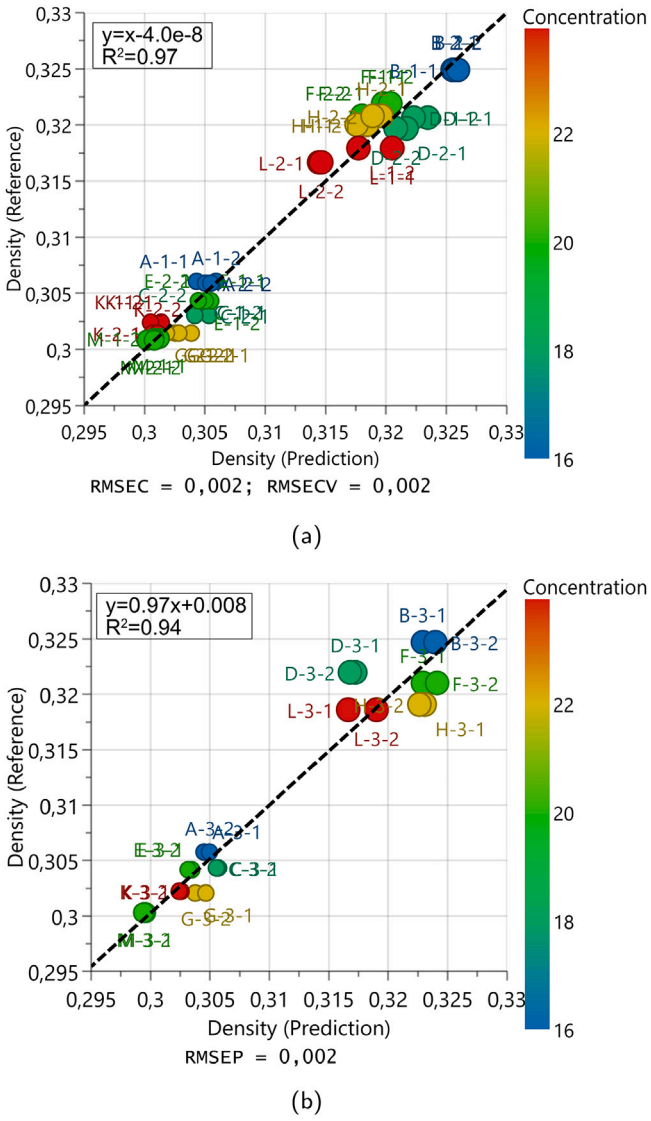

Fig. 8. Prediction plot for tablet density (a) calibration set (b) test set. Tablets are labeled with three digits that represent the type, replicate, and measurement number, respectively. Small and large particle sizes are shown with small and large circles, respectively.

tablets. A quantitative OPLS method was proposed to model the API concentration and tablet density based on terahertz transmission measurements at a frequency range of 750 GHz to 1.5 THz. The OPLS model for the API concentration achieved an R^2 value of 0.98 and a relative prediction error of 6% for tablets with a 2% difference in the level of the API concentration. The results showed that THz-FDS is less precise than NIR and Raman spectroscopy for the quantification of API concentration, but it brings the advantage of being less affected by particle size variations, which could prevent systematic bias errors. To optimize the determination of API content, the THz measurements should preferably be conducted in a controlled low humidity environment. The OPLS model for the tablet density achieved an R^2 value of 0.97 and an

excellent relative prediction error of 0.7%. This highlights the potential of terahertz spectroscopy for characterizing the physical properties of tablets, which is a challenge for other non-destructive techniques. In summary, THz-FDS in combination with multivariate analysis showed to be a promising process analytical tool that complements NIR and Raman techniques for the characterization of the physical properties of tablets and API concentration with the advantage of low scattering effects.

CRediT authorship contribution statement

Anis Moradikouchi: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Anders Sparén:** Conceptualization, Methodology, Formal analysis, Writing review & editing, Supervision, Resources, Project administration. **Olof Svensson:** Formal analysis, Methodology, Writing – review & editing. **Staffan Folestad:** Conceptualization, Writing – review & editing, Supervision, Resources, Project administration. **Jan Stake:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition, Resources. **Helena Rodilla:** Conceptualization, Methodology, Writing – review & editing, Supervision, Resources, Project administration.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jan Stake reports financial support was provided by Swedish Foundation for Strategic Research.

Data availability

Data will be made available on request.

Acknowledgments

The measurements were carried out at the Kollberg Laboratory, at Chalmers University of Technology, Gothenburg, Sweden.

Funding

This research was funded by a grant from the Swedish foundation for strategic research (SSF), ID 17–0011, and from AstraZeneca, Gothenburg, Sweden.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijpharm.2022.122545.

References

Aina, A., Hargreaves, M.D., Matousek, P., Burley, J.C., 2010. Transmission Raman spectroscopy as a tool for quantifying polymorphic content of pharmaceutical formulations. Analyst 135 (9), http://dx.doi.org/10.1039/c0an00352b.

Allard, J.-F., Cornet, A., Debacq, C., Meurens, M., Houde, D., Morris, D., 2011. Improved detection sensitivity of D-mannitol crystalline phase content using differential spectral phase shift terahertz spectroscopy measurements. Opt. Express 19 (5), 4644. http://dx.doi.org/10.1364/oe.19.004644.

Bawuah, P., Markl, D., Farrell, D., Evans, M., Portieri, A., Anderson, A., Goodwin, D., Lucas, R., Zeitler, J.A., 2020. Terahertz-based porosity measurement of pharmaceutical tablets: A tutorial. J. Infrared Millim. Terahertz Waves 41 (4), 450–469. http://dx.doi.org/10.1007/s10762-019-00659-0.

Bawuah, P., Zeitler, J.A., 2021. Advances in terahertz time-domain spectroscopy of pharmaceutical solids: A review. TrAC - Trends Anal. Chem. 139, 116272. http://dx.doi.org/10.1016/j.trac.2021.116272.

Berntsson, O., Burger, T., Folestad, S., Danielsson, L.G., Kuhn, J., Fricke, J., 1999. Effective sample size in diffuse reflectance near-IR spectrometry. Anal. Chem. 71 (3), http://dx.doi.org/10.1021/ac980652u.

Berntsson, O., Danielsson, L.G., Lagerholm, B., Folestad, S., 2002. Quantitative in-line monitoring of powder blending by near infrared reflection spectroscopy. Powder Technol. 123 (2–3), http://dx.doi.org/10.1016/S0032-5910(01)00456-9.

Chen, W., Peng, Y., Jiang, X., Zhao, J., Zhao, H., Zhu, Y., 2017. Isomers identification of 2-hydroxyglutarate acid disodium salt (2HG) by terahertz time-domain spectroscopy. Sci. Rep. 7 (1), http://dx.doi.org/10.1038/s41598-017-11527-z.

Claybourn, M., Yang, H., Gradinarsky, L., Johansson, J., Folestad, S., 2007. Terahertz spectroscopy for pharmaceutical applications. Handb. Vib. Spectrosc. 1–14. http://dx.doi.org/10.1002/9780470027325.s8914.

Ewing, A.V., Kazarian, S.G., 2018. Recent advances in the applications of vibrational spectroscopic imaging and mapping to pharmaceutical formulations. Spectrochimica Acta - Part A: Mol. Biomol. Spectrosc. 197, http://dx.doi.org/10.1016/j.saa.2017.12.055.

Hisazumi, J., Watanabe, T., Suzuki, T., Wakiyama, N., Terada, K., 2012. Using terahertz reflectance spectroscopy to quantify drug substance in tablets. Chem. Pharmaceut. Bull. 60 (12), 1487–1493. http://dx.doi.org/10.1248/cpb.c12-00524.

Hübers, H.W., 2008. Terahertz heterodyne receivers. IEEE J. Sel. Top. Quantum Electron. 14 (2), 378–391. http://dx.doi.org/10.1109/JSTQE.2007.913964.

Hübers, H.W., Kimmitt, M.F., Hiromoto, N., Bründermann, E., 2011. Terahertz spectroscopy: System and sensitivity considerations. IEEE Trans. Terahertz Sci. Technol. 1 (1), 321–331. http://dx.doi.org/10.1109/TTHZ.2011.2159877.

Ierapetritou, M., Muzzio, F., Reklaitis, G., 2016. Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. AIChE J. 62 (6), http://dx.doi.org/10.1002/aic.15210.

Jedvert, I., Josefson, M., Langkilde, F., 1998. Quantification of an active substance in a tablet by NIR and Raman spectroscopy. J. Near Infrared Spectrosc. 6 (1–4), http://dx.doi.org/10.1255/jnirs.148.

Laske, S., Paudel, A., Scheibelhofer, O., Sacher, S., Hoermann, T., Khinast, J., Kelly, A., Rantannen, J., Korhonen, O., Stauffer, F., De Leersnyder, F., De Beer, T., Mantanus, J., Chavez, P.F., Thoorens, B., Ghiotti, P., Schubert, M., Tajarobi, P., Haeffler, G., Lakio, S., Fransson, M., Sparen, A., Abrahmsen-Alami, S., Folestad, S., Funke, A., Backx, I., Kavsek, B., Kjell, F., Michaelis, M., Page, T., Palmer, J., Schaepman, A., Sekulic, S., Hammond, S., Braun, B., Colegrove, B., 2017. A review of PAT strategies in secondary solid oral dosage manufacturing of small molecules. J. Pharmaceut. Sci. 106 (3), http://dx.doi.org/10.1016/j.xphs.2016.11.011.

Lyon, R.C., Lester, D.S., Lewis, E.N., Lee, E., Yu, L.X., Jefferson, E.H., Hussain, A.S., 2002. Near-infrared spectral imaging for quality assurance of pharmaceutical products: Analysis of tablets to assess powder blend homogeneity. AAPS PharmSciTech 3 (3), http://dx.doi.org/10.1007/BF02830615.

Moradikouchi, A., Sparén, A., Folestad, S., Stake, J., Rodilla, H., 2022. Terahertz frequency domain sensing for fast porosity measurement of pharmaceutical tablets. Int. J. Pharm. 618, http://dx.doi.org/10.1016/j.ijpharm.2022.121579.

Peng, J., Peng, S., Jiang, A., Wei, J., Li, C., Tan, J., 2010. Asymmetric least squares for multiple spectra baseline correction. Anal. Chim. Acta 683 (1), http://dx.doi.org/10.1016/j.aca.2010.08.033.

Peng, Y., Shi, C., Xu, M., Kou, T., Wu, X., Song, B., Ma, H., Guo, S., Liu, L., Zhu, Y., 2018. Qualitative and quantitative identification of components in mixture by terahertz spectroscopy. IEEE Trans. Terahertz Sci. Technol. 8 (6), 696–701. http://dx.doi.org/10.1109/TTHZ.2018.2867816.

Plumb, K., 2005. Continuous processing in the pharmaceutical industry: Changing the mind set. Chem. Eng. Res. Des. 83 (6 A), http://dx.doi.org/10.1205/cherd.04359.

Pozar, D.M., 2012. Microwave engineering. In: John Wiley &Sons, Inc, 4th Ed. pp. 178–188.

Salim, M., Fraser-Miller, S.J., Rzinš, K.B., Sutton, J.J., Ramirez, G., Clulow, A.J., Hawley, A., Beilles, S., Gordon, K.C., Boyd, B.J., 2020. Low-frequency raman scattering spectroscopy as an accessible approach to understand drug solubilization in milk-based formulations during digestion. Mol. Pharmaceut. 17 (3), 885–899. http://dx.doi.org/10.1021/acs.molpharmaceut.9b01149.

Savitzky, A., Golay, M.J., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36 (8), http://dx.doi.org/10.1021/ac60214a047.

Shen, Y.C., Taday, P.F., Pepper, M., 2008. Elimination of scattering effects in spectral measurement of granulated materials using terahertz pulsed spectroscopy. Appl. Phys. Lett. 92 (5), 1–4. http://dx.doi.org/10.1063/1.2840719.

Siegel, P.H., 2004. Terahertz technology in biology and medicine. IEEE Trans. Microwave Theory Techniques 52 (10), 2438–2447. http://dx.doi.org/10.1109/TMTT.2004.835916.

Sparén, A., Hartman, M., Fransson, M., Johansson, J., Svensson, O., 2015. Matrix effects in quantitative assessment of pharmaceutical tablets using transmission Raman and near-infrared (NIR) spectroscopy. Appl. Spectrosc. 69 (5), 580–589. http://dx.doi.org/10.1366/14-07645.

Svensson, O., Josefson, M., Langkilde, F.W., 2000. The synthesis of metoprolol monitored using Raman spectroscopy and chemometrics. Eur. J. Pharmaceut. Sci. 11 (2), http://dx.doi.org/10.1016/S0928-0987(00)00094-4.

Taday, P.F., Van Der Weide, D., Wood, K., Chamberlain, M., Roskos, H., Phillips, C., Newnham, D., Towrie, M., Appelquist, I., 2004. Applications of terahertz spectroscopy to pharmaceutical sciences. Phil. Trans. R. Soc. A 362 (1815), 351–364. http://dx.doi.org/10.1098/rsta.2003.1321.

Townshend, N., Nordon, A., Littlejohn, D., Andrews, J., Dallin, P., 2012. Effect of particle properties of powders on the generation and transmission of Raman scattering. Anal. Chem. 84 (11), 4665–4670. http://dx.doi.org/10.1021/ac203446g.

Trygg, J., Wold, S., 2002. Orthogonal projections to latent structures (O-PLS). J. Chemometr. 16 (3), http://dx.doi.org/10.1002/cem.695.

Vergote, G.J., De Beer, T.R., Vervaet, C., Remon, J.P., Baeyens, W.R., Diericx, N., Verpoort, F., 2004. In-line monitoring of a pharmaceutical blending process using FT-Raman spectroscopy. Eur. J. Pharmaceut. Sci. 21 (4), http://dx.doi.org/10.1016/j.ejps.2003.11.005.

Wu, H., Heilweil, E.J., Hussain, A.S., Khan, M.A., 2008. Process analytical technology (PAT): Quantification approaches in terahertz spectroscopy for pharmaceutical application. J. Pharm. Sci. 97 (2), 970–984. http://dx.doi.org/10.1002/jps.21004.

Yang, Q., Wu, L., Shi, C., Wu, X., Chen, X., Wu, W., Yang, H., Wang, Z., Zeng, L., Peng, Y., 2021. Qualitative and quantitative analysis of Caffeine in medicines by terahertz spectroscopy using machine learning method. IEEE Access 9, 140008–140021. http://dx.doi.org/10.1109/ACCESS.2021.3116980.

Zeitler, J.A., Kogermann, K., Rantanen, J., Rades, T., Taday, P.F., Pepper, M., Aaltonen, J., Strachan, C.J., 2007. Drug hydrate systems and dehydration processes studied by terahertz pulsed spectroscopy. Int. J. Pharm. 334 (1–2), 78–84. http://dx.doi.org/10.1016/j.ijpharm.2006.10.027.