European Journal of Pharmaceutical Sciences 200 (2024) 106836

Contents lists available at ScienceDirect

European Journal of Pharmaceutical Sciences

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

Analysis of the impact of material properties on tabletability by principal component analysis and partial least squares regression

Lena Mareczek [a,b], Carolin Riehl [b], Meike Harms [b], Stephan Reichl [a,*]

[a] Institute of Pharmaceutical Technology and Biopharmaceutics, Technische Universität Braunschweig, Braunschweig 38106, Germany
[b] Department of Orals Development, Merck Healthcare KGaA, Darmstadt 64293, Germany

ARTICLE INFO

Keywords:
Principal component analysis
Partial least squares regression
Multivariate data analysis
Material characterization
Tabletability
Manufacturability

ABSTRACT

Principal component analysis (PCA) and partial least squares regression (PLS) were combined in this study to identify key material descriptors determining tabletability in direct compression and roller compaction. An extensive material library including 119 material descriptors and tablet tensile strengths of 44 powders and roller compacted materials with varying drug loads was generated to systematically elucidate the impact of different material descriptors, raw API and filler properties as well as process route on tabletability. A PCA model was created which highlighted correlations between different powder descriptors and respective characterization methods and, thus, can enable reduction of analyses to save resources to a certain extent. Subsequently, PLS models were established to identify key material attributes for tabletability such as density and particle size but also surface energy, work of cohesion and wall friction, which were for the first time demonstrated by PLS as highly relevant for tabletability in roller compaction and direct compression. Further, PLS based on extensive material characterization enabled the prediction of tabletability of materials unknown to the model. Thus, this study highlighted how PCA and PLS are useful tools to elucidate the correlations between powder and tabletability, which will enable more robust prediction of manufacturability in formulation development.

1. Introduction

Over the last years, chemometrics and multivariate data analysis (MVDA) have gained in importance in the pharmaceutical industry and were given greater consideration in the pharmacopeias (European Directorate for the Quality of Medicines and Health Care, 2023; United States Pharmacopeia, 2016) as well as in multiple guidelines on pharmaceutical development, manufacturing, and control processes (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2020, 2011). Two methods of multivariate data analysis that are commonly used in the pharmaceutical field are the principal component analysis (PCA) and partial least squares regression (PLS), which both reduce dimensionality in large data sets (Ferreira and Tobyn, 2015). In both, PCA and PLS, overarching new independent variables are created by linearly combining correlated descriptors. In PCA, the newly created, independent variables are called principal components (PC). The first principal component describes most of the variance in a data set as it is created from the initial descriptor describing the most variance in the data set and its linearly correlated descriptors. For the second principal component, the initial descriptors are used that explain the next largest variance of the data set without being linearly correlated to the descriptors used for PC 1. This procedure is continued for any additional PC. By then plotting two principal components, the multiple descriptors and therefore multiple dimensions of the data set are reduced to only two dimensions with as little loss in information as possible. In partial least squares regression, a similar approach is applied. Opposed to PCA where the descriptors are analyzed for correlations among themselves, PLS analyzes which initial descriptors (X) explain most variance of one or more responses (Y). Therefore, PCA describes variance in X data whereas PLS directly describes co-variance of descriptor (X) and response (Y) data.

The improved representation of data by dimensionality reduction with PCA and PLS allows for a better evaluation of underlying correlations of samples as well as descriptors and of influences of descriptors on responses. Therefore, PCA and PLS can be useful tools for the pharmaceutical industry as with their application it is possible to better elucidate the relationship of material descriptors and processability, which is a key criteria for a more time and cost efficient formulation development and manufacturing of drug products with less risks (Leane et al., 2015).

* Corresponding author.
E-mail addresses: carolin.riehl@merckgroup.com (C. Riehl), s.reichl@tu-braunschweig.de (S. Reichl).

https://doi.org/10.1016/j.ejps.2024.106836
Received 8 April 2024; Received in revised form 27 May 2024; Accepted 17 June 2024
Available online 18 June 2024

With PCA or PLS, the impact of raw material properties and process parameters on manufacturability for oral solid dosage forms (OSDs) has been evaluated for various pharmaceutical processes and manufacturability aspects.

Peeters et al. applied PCA and PLS to evaluate the impact of initial material properties as well as process settings on granule formation in different stages of the twin-screw wet granulation process and by that identified that for the material attributes the initial particle size, flowability of the powder, water-related powder properties such as water binding capacity and for the process parameters the liquid-to-solid ratio influenced granule size and shape (Peeters et al., 2024). Strong impact for the distribution of binder liquid was highlighted for API solubility as well as the liquid-to-solid ratio. Further, they could demonstrate differences in influencing material attributes in different compartments of the twin-screw granulation. Ryckaert et al. established a T-PLS model to link initial material attributes, drug load and process parameters to twin-screw wet granulation processability and predict suitable formulations for twin-screw wet granulation processing (Ryckaert et al., 2021). They identified dissolution rate, compressibility, water binding capacity, powder density and solubility as key material attributes for granule quality. Further, the T-PLS model demonstrated a first moderate potential to predict formulation and process settings for new APIs in twin-screw wet granulation. The influence of powder properties and process parameters on fill weight and weight variability in capsule filling for dry powder inhalation was determined by Faulhammer et al. using PLS (Faulhammer et al., 2014). Via PLS, a correlation between capsule fill weight with nozzle diameter, dosing chamber length, powder layer depth and powder density was identified, however different extents of these influences were seen in powders with larger particles/good flowability and powders with smaller particles/poor flowability. In a study by Escotet-Espinoza et al. correlations between material properties and phenomenological model coefficients were identified and the effects of initial material properties on processability in direct compression were evaluated by combining PCA and linear regression with the semi-empirical Kuentz-Leuenberger model (Escotet-Espinoza et al., 2018). Subsequently, an empirical model was created that correlated the material attributes to the phenomenological model coefficients to determine a compression design space, which provided a proof of concept for this approach. Dhondt et al. established a T-PLS model for direct compression to correlate tablet quality with initial material properties based on a raw material database, blend design and process parameters, which can be applied in the future to develop data-based preliminary formulations and process parameters for direct compression (Dhondt et al., 2022a). They could show for their data set that the material properties and blend composition had a stronger impact than the process settings on tablet quality in direct compression. Further, loss on drying and powder density descriptors were found to describe the largest variability in their data set and thus could be key factors for tablet quality. Based on the established model, an optimal filler combination was selected for a newly introduced API and tablet quality attributes were predicted, which were in overall good accordance with the experimental data. Thus, it was shown that this multivariate approach could in the future be used as a data based starting point for formulation and process setting selection. However, they also clearly observed limitations of the model such as that only one fixed drug load was used. In-depth information about the influence of API by applying multiple drug loads could provide further insights on processability. Aspects of processability in roller compaction were investigated via multivariate data analysis in studies from Boersen et al., Soh et al. and Souihi et al. (Boersen et al., 2015; Soh et al., 2008; Souihi et al., 2013). In the study by Boersen et al. (Boersen et al., 2015), PLS was applied to identify key process parameters and initial material properties impacting roller compacted ribbons as well as subsequently produced granules and tablets. Further, the study investigated if it is possible to predict the processability in roller compaction of a newly introduced API based on a PLS model established with another API. It was demonstrated that ribbon porosity and granule size in roller compaction were successfully predicted using PLS regression, whereas the prediction of tablet tensile strength was not successful with the study design by Boersen et al. Soh et al. used PCA and PLS as well as a Design of Experiment (DoE) approach to determine key material properties and process parameter for roller compaction processability of different lactoses and microcrystalline celluloses (Soh et al., 2008). The study demonstrated that PLS models that included both process settings as well as initial material properties clearly had higher goodness of fit and predictability compared to models based solely on the process settings, that were not able to predict ribbon and granules properties sufficiently. Strong impact on multiple roller compaction responses was identified for the tablet tensile strength of the initial material, the MCC fraction, tapped density, Kawakita constant, angle of fall as well as span. In a study from Souihi et al., PCA and orthogonal projections to latent structures were applied to gain a deeper understanding on the impact of material attributes of brittle fillers such as mannitol and dicalcium phosphate on the reduction of tablet tensile strength after roller compaction (Souihi et al., 2013). It was demonstrated that brittle materials are less prone to reduction of tablet tensile strength compared to plastically deforming materials and can even display improved tabletability after roller compaction. Further, it was revealed that initial materials specifically designed for direct compression with excellent flow and compaction properties can lead to roller compacted granules with poor flowability compared to the granules manufactured from traditional material not specifically engineered for direct compression.

These studies highlighted the broad applicability of multivariate data analysis for pharmaceutical development and manufacturing. However, for most of these studies only a very limited number of powder characteristics were analyzed as the respective experiments are time and material intensive. For a more holistic understanding of key material attributes for manufacturing processes of oral solid dosage forms, an extensive material library containing a large number of different powder properties is needed as a basis for subsequent evaluation of correlations between material characteristics and processability (Dhondt et al., 2022b; Van Snick et al., 2018). Furthermore, despite the versatile application of PCA and PLS reported in literature for manufacturability aspects of oral dosage forms, the benefits and opportunities arising from applying PCA and PLS for a deeper understanding of the impact and interplay of neat API and excipient properties, drug load and processing route on mechanical tablet properties have not yet been adequately described.

Therefore, this is to our knowledge the first study to apply and combine PCA and PLS to such a large data set with substantial variability by inclusion of 119 material descriptors and 44 materials including neat APIs, fillers and binary blends thereof with varying drug load in two process routes, direct compression and roller compaction, to identify key material descriptors for tablet tensile strength (TTS) prediction as well as the impact of raw material properties of API and excipients as well as the processing route on TTS. With that, the study demonstrates that PCA and PLS can be powerful tools to deepen the understanding on the relationship between different material descriptors and their impact on processability for oral solid dosage forms and thus provides a solid foundation for data-based selection of materials for formulation development in the future.

Firstly, an extensive material library was created including 119 material descriptors as well as the tablet tensile strength at six different compression pressures of 44 powders and roller compacted materials with varying drug load (34 materials to train the model and 10 materials for external model validation). Binary API-filler combinations of two APIs (caffeine, paracetamol) and two fillers (microcrystalline cellulose, mannitol), that have differing mechanical properties, and multiple drug loads per combination were specifically included to systematically analyze the impact of the individual components, API as well as excipients, on the tabletability of formulations. By including initial pre-blend powders as well as respective roller compacted materials, the material

library further enabled to directly compare the influence of the process routes on the mechanical tablet properties.

Secondly, a PCA model was developed to gain a deeper knowledge on the correlations and relations between multiple powder descriptors but also between different characterization methods and to identify the potential for databased reduction in experimental burden with as little loss as possible by PCA.

In the third step, subsequent PLS regression models were established to reveal powder descriptors impacting tablet tensile strength as response for mechanical tablet properties (Berkenkemper et al., 2023) in direct compression and roller compaction. The PLS model was then used to predict the tabletability of 10 materials unknown to the models as external validation. Additionally, the ability to reduce experimental burden by multivariate data analysis with only minor loss in information and predictive ability was evaluated by comparing two PLS models, one using the initial material descriptor set and one using a reduced material descriptor set proposed by the PCA model results as PCA and PLS provide different information and the material descriptors that explain a large variance in the powder and rolled material data (X) do not necessarily have a strong effect on the tablet tensile strength data (Y).

2. Materials and methods

2.1. Materials

Paracetamol (Pcm) was acquired from Fagron Services (Rotterdam, Netherlands). Caffeine (Caf) and spray granulated mannitol Parteck® M200 (Man) were supplied by Merck KGaA (Darmstadt, Germany). Microcrystalline cellulose (MCC, Vivapur® 101) was purchased by JRS Pharma (Rosenberg, Germany) and magnesium stearate was acquired from Peter Greven (Bad Münstereifel, Germany).

For the external validation of the established PLS model, following materials unknown to the model were used. Compound A (CompA) as well as Parteck® M100 (Man100) and a commercially available δ polymorph of D-Mannitol, Parteck® Delta M (DeltaMan), were supplied by Merck KGaA (Darmstadt, Germany).

An overview of the samples used to create the PCA and PLS models is given in Table 1 whereas samples for external validation are described in Table 2.

2.2. Methods

2.2.1. Blend preparation

All powders were sieved through a 1 mm sieve using a Turbosieve BTS 100 (L.B. Bohle, Ennigerloh, Germany) with a speed of 355 rpm.

Binary powder blends of caffeine with Parteck® M200 (20/40/60 % drug load), caffeine with MCC and paracetamol with MCC (20/40/60/80 % drug load) as well as compound A with Parteck® M100 (30/50 % drug load) were produced by blending at 12 rpm for 15 min using a container blender (Servolift GmbH, Offenburg, Germany).

2.2.2. Roller compaction

Roller compaction was performed on a Mini-Pactor® (Gerteis, Rapperswil-Jona, Switzerland) using rolls with a diameter of 25 cm and width of 2.5 cm. A specific compaction force of 3 kN/cm, a gap width of 3 mm and roll speed of 3 rpm were set as process parameters. Exceptions were made for Pcm containing samples, where 1.5 rpm roll speed was used due to paracetamols weak mechanical properties. For the validation materials, the same process parameters were chosen except that Pearlitol® 160C was roller compacted with 9 kN/cm as described in a previous study (Mareczek et al., 2022) to form stable granules. Ribbons were granulated by an oscillating star rotor granulator equipped with a 1.0 mm sieve.

2.2.3. Particle size distribution

The CamSizer X2® (Retsch GmbH, Haan, Germany) was utilized to

Table 1
Overview of training sample set for PCA and PLS.

Abbreviation	Process stage	API	Drug load	Filler	Filler load
0 %CafMan-P	Powder	/	0 %	Parteck® M200	100 %
20 %CafMan-P	Powder	Caffeine	20 %	Parteck® M200	80 %
40 %CafMan-P	Powder	Caffeine	40 %	Parteck® M200	60 %
60 %CafMan-P	Powder	Caffeine	60 %	Parteck® M200	40 %
100 % CafMan-P	Powder	Caffeine	100 %	/	0 %
0 %CafMan-G	Roller compacted material	/	0 %	Parteck® M200	100 %
20 %CafMan-G	Roller compacted material	Caffeine	20 %	Parteck® M200	80 %
40 %CafMan-G	Roller compacted material	Caffeine	40 %	Parteck® M200	60 %
60 %CafMan-G	Roller compacted material	Caffeine	60 %	Parteck® M200	40 %
100 % CafMan-G	Roller compacted material	Caffeine	100 %	/	0 %
0 %CafMCC-P	Powder	/	0 %	Vivapur® 101	100 %
20 %CafMCC-P	Powder	Caffeine	20 %	Vivapur® 101	80 %
40 %CafMCC-P	Powder	Caffeine	40 %	Vivapur® 101	60 %
60 %CafMCC-P	Powder	Caffeine	60 %	Vivapur® 101	40 %
80 %CafMCC-P	Powder	Caffeine	80 %	Vivapur® 101	20 %
100 % CafMCC-P	Powder	Caffeine	100 %	/	0 %
0 %CafMCC-G	Roller compacted material	/	0 %	Vivapur® 101	100 %
20 %CafMCC-G	Roller compacted material	Caffeine	20 %	Vivapur® 101	80 %
40 %CafMCC-G	Roller compacted material	Caffeine	40 %	Vivapur® 101	60 %
60 %CafMCC-G	Roller compacted material	Caffeine	60 %	Vivapur® 101	40 %
80 %CafMCC-G	Roller compacted material	Caffeine	80 %	Vivapur® 101	20 %
100 % CafMCC-G	Roller compacted material	Caffeine	100 %	/	0 %
0 %PcmMCC-P	Powder	/	0 %	Vivapur® 101	100 %
20 % PcmMCC-P	Powder	Paracetamol	20 %	Vivapur® 101	80 %
40 % PcmMCC-P	Powder	Paracetamol	40 %	Vivapur® 101	60 %
60 % PcmMCC-P	Powder	Paracetamol	60 %	Vivapur® 101	40 %
80 % PcmMCC-P	Powder	Paracetamol	80 %	Vivapur® 101	20 %
100 % PcmMCC-P	Powder	Paracetamol	100 %	/	0 %
0 %PcmMCC-G	Roller compacted material	/	0 %	Vivapur® 101	100 %
20 % PcmMCC-G	Roller compacted material	Paracetamol	20 %	Vivapur® 101	80 %

(continued on next page)

Table 1 (*continued*)

Abbreviation	Process stage	API	Drug load	Filler	Filler load
40 % PcmMCC-G	Roller compacted material	Paracetamol	40 %	Vivapur® 101	60 %
60 % PcmMCC-G	Roller compacted material	Paracetamol	60 %	Vivapur® 101	40 %
80 % PcmMCC-G	Roller compacted material	Paracetamol	80 %	Vivapur® 101	20 %
100 % PcmMCC-G	Roller compacted material	Paracetamol	100 %	/	0 %

determine particle size distribution (PSD) of powders and roller compacted materials by dynamic image analysis using the X-Jet module and a dispersing air pressure of 25 kPa. Measurement was conducted in triplicate and samples were divided with the sample splitter RT 12.5 (Retsch GmbH, Haan, Germany).

2.2.4. Density

Pycnometric density (PycD) measurement of the powders and roller compacted materials was performed with nitrogen via gas displacement technique using an Ultrapyc 1200e (Quantachrome Instruments, Boynton Beach, USA). A medium test cell was filled to at least 80 % with the sample and mass was weighed. The arithmetic mean of $n = 3 \pm S.D.$ was determined.

The bulk (PF-0) and tapped (PF-500) packing fraction were determined with a Granupack™ (GranuTools, Awans, France) by applying 500 taps and a tapping frequency of 1 Hz. Bulk (ρ_B) and tapped density (ρ_T) were calculated as m/V_0 and m/V_t. To achieve better comparability between different materials, the bulk and tapped packing fraction were calculated by dividing the bulk and tapped density through the respective pycnometric density as PF-0 = ρ_B/PycD and PF-500 = ρ_T/PycD. Hausner ratio was calculated as HR = ρ_T/ρ_B and Carr index was calculated as Carr = $100*(1 - \rho_B/\rho_T)$. Powders and roller compacted materials were measured in triplicate.

2.2.5. Powder rheometer

With the dynamic flow measurement of the FT4 powder rheometer (Freeman Technology, Tewkesbury, UK), the resistance of a sample against flow induced by a blade moving through the powder bed is determined. With the combined "stability and variable flow rate" method, multiple descriptors of the powders and roller compacted materials such as the basic flow energy (BFE), specific energy (SE), stability index (SI), flow rate index (FRI) and conditioned bulk density (CBD) were measured according to established protocols (Freeman, 2007) with a 48 mm diameter steel blade and 50 mm diameter glass vessels. BFE was normalized to sample mass and normalized BFE (nBFE) was used in the following PCA to avoid doubled input from bulk density effects.

The powder rheometer was furthermore utilized to determine the powder compressibility as *v/v*% reduction under pressure as well as the air pressure drop as measurement against air flow through the powder bed. The combined measurement of powder compressibility and air pressure drop was performed in a glass vessel with a diameter of 50 mm at applied compression forces of 1 and 15 kPa, with each load being held for 60 s. For pressure drop determination, an air velocity of 2 mm/s was maintained.

All measurements were performed in triplicate.

2.2.6. Ring shear tester

The flowability as flow function coefficient (Ffc), the cohesion (Coh), angle of internal friction (AIF) and effective angle of internal friction (AIFe) were determined by ring shear tester RST-XS (Dietmar Schulze Schüttgutmesstechnik, Wolfenbüttel, Germany) in triplicate according to Jenike (Jenike, 1964). A normal preshear stress of 9000 Pa and consolidation stresses of 1800, 4500 and 7200 Pa were applied. Evaluation was performed with the RSV 95 software (Dietmar Schulze Schüttgutmesstechnik, Wolfenbüttel, Germany) using straight line segments for regression.

Wall friction angle (WFA) was determined in triplicate with a maximum wall normal stress of 19 000 Pa and minimum wall normal stress of 950 Pa. 10 wall normal stresses linearly distributed between the maximum and minimum were set for the measurement and 4 repetitions were done for each wall normal stress within one measurement. As trends for the different materials were comparable at all used wall normal stresses, wall friction angles at 7000 Pa (WFA-7) and 19 000 Pa (WFA-19) were used exemplary in the following PCA and PLS.

2.2.7. Specific surface area & surface energy

Specific surface area (SSA) and surface energy of the powders and roller compacted materials were measured using inverse gas chromatography with the iGC-SEA (Surface Measurement Systems Ltd., Alperton, UK). Data was analyzed with the SEA Analysis software (Surface Measurement Systems Ltd., Alperton, UK). Silanized glass columns (3 mm inner diameter) were filled with a sample mass that corresponded to a total surface area of approximately 0.5 m^2 and stoppered using silanized glass wool at both ends. Dead volume was determined by methane injections. Retention times of probe molecules and methane were determined using a flame ionization detector. Samples were conditioned for 60 min at measurement settings of 30 °C, 0 % relative humidity and 10 cm^3/min nitrogen carrier gas flow.

The SSA was determined according to Brunauer-Emmett-Teller (Brunauer et al., 1938; Thielmann et al., 2007) from the isotherm of physical adsorption of octane molecules onto a solid's surface in the pressure range (p/p0) from 0.05 to 0.35.

Surface energy measurement was performed with the same sample columns and measurement settings as described for SSA.

According to Fowkes, the total surface energy (γ_t) is directly correlated to the work of cohesion (WoC) and can be divided into a dispersive and polar contribution (Fowkes, 1964). The dispersive surface energy (γ_d) was determined with a series of alkanes, heptane, octane and nonane, according to the Dorris and Gray approach (Dorris and Gray, 1980).

Table 2
Overview of test sample set for external validation.

Abbreviation	Process stage	API	Drug load	Filler	Filler load
0 %CmpAMan100-P	Powder	Compound A	0 %	Parteck® M100	100 %
30 %CmpAMan100-P	Powder	Compound A	30 %	Parteck® M100	70 %
50 %CmpAMan100-P	Powder	Compound A	50 %	Parteck® M100	50 %
100 %CmpAMan100-P	Powder	Compound A	100 %	Parteck® M100	0 %
0 %CmpAMan100-G	Roller compacted material	Compound A	0 %	Parteck® M100	100 %
30 %CmpAMan100-G	Roller compacted material	Compound A	30 %	Parteck® M100	70 %
50 %CmpAMan100-G	Roller compacted material	Compound A	50 %	Parteck® M100	50 %
100 %CmpAMan100-G	Roller compacted material	Compound A	100 %	Parteck® M100	0 %
DeltaMan-P	Powder	/	/	Parteck® Delta M	100 %
DeltaMan-G	Roller compacted material	/	/	Parteck® Delta M	100 %

Polar surface energy (γ_{ab}) is determined with a Lewis acidic probe (chloroform) and a Lewis base probe (ethyl acetate) based on the polarization approach described by Dong et al. (Dong et al., 1989) and the Della Volpe scale (Della Volpe and Siboni, 2000, 1997). Surface energy was analyzed for three different surface coverages (0.02 n/n_m, 0.06 n/n_m, 0.2 n/n_m). Surface polarity γ_{pol} was calculated as the share of γ_{ab} in γ_t.

KaKb describes the Lewis acidity-basicity ratio of a material surface and was determined with the Polarization method and the modified Gutmann approach by Papirer (Gamble et al., 2012; Papirer et al., 1988). The same samples, measurement settings, surface coverages and alkane series used for surface energy determination were applied for KaKb and chloroform, ethyl acetate, ethanol and acetonitrile were used as polar probes.

Measurements were performed with $n = 1$ as the relative standard deviation of surface energy measurements with the iGC-SEA was determined in other studies to be below 4 % (Reutenauer, 2002).

2.2.8. Tablet compression

Tablet compression of the powders and roller compacted materials was performed using a Styl'One Evolution compaction simulator (Medelpharm, Beynost, France) with round, flat faced punches with a diameter of 11.28 mm and a default compression profile without pre-compression at 20 % compression speed. Maximum punch velocity was 30 mm/s. 25 tablets per powder blend or roller compacted material were produced at compression pressures of 50, 100, 150, 200, 300 and 400 MPa. A Quantos dosing system QB1 (Mettler Toledo, Columbus OH, USA) was utilized to weigh a sample mass of 400 mg per tablet with a deviation of <1 %. The weighted powders and roller compacted materials were then filled into the tablet press die manually. Only external lubrication of the punches and die with magnesium stearate was used for tableting in order to minimize the impact on the mechanical properties of the materials.

Weight, thickness, diameter and breaking force of the tablets ($n = 10$) were determined using a MultiCheck VI (Erweka GmbH, Langen, Germany) with a constant tablet breaking speed of 2.3 mm/s.

Diametral tablet tensile strength (TTS) was calculated according to Fell and Newton (Fell and Newton, 1970):

$$\textit{Tablet tensile strength } [MPa] = \frac{2 * F}{\pi * d * t}$$

with F as the tablet breaking force (N), d as tablet diameter (mm) and t as tablet thickness (mm).

No experimental TTS values could be acquired neither for 0 %CafMan powder and respective granules at 300 and 400 MPa compression pressure nor 20 %CafMan powder and granules at 400 MPa compression pressure because of very high tablet ejection forces due to extensive sticking of mannitol to the die. For these four samples, TTS values at the respective compression pressures were determined by linear regression with at least 4 data points.

2.2.9. Principal component analysis and partial least squares regression

Principal component analysis was applied to deepen the understanding of how different material descriptors of powders and roller compacted materials, and respective methods to determine them, correlate to each other. Partial least squares regression was utilized to elucidate which material descriptors correlate to the tablet tensile strength (TTS) and therefore are potential key attributes for mechanical tablet properties. Principal component analysis, partial least squares regression and the respective projections and predictions were performed with the Unscambler X 10.5.1. software (Camo software inc., Magnolia, USA). Before PCA and PLS, the data were preprocessed. Descriptors and samples for which more than 10 % of data were missing were excluded from the data set. Non-linear data such as the powder compressibility, pressure drop, flow function coefficient, cohesion, angle of internal friction, effective angle of internal friction and wall friction angle were logarithmized. To identify possible outliers, the Hotelling's T^2 plot was constructed and no outliers in the training sample set for PCA and PLS were identified using the 99 % confidence threshold. The total variance in X or Y data explained by the different numbers of principal components in a PCA model or factors in a PLS model is calculated as 100 * (initial variance – residual variance) / (initial variance). The residual variance is calculated as sum of squares of the X- or Y-residuals divided by the number of degrees of freedom.

2.2.9.1. PSD PCA and iGC surface energy PCA. Particle size distribution and surface energy measurement generated many more data points compared to the other powder characterization methods. In order to avoid overrepresentation of these two methods in the descriptor dataset for multivariate data analysis, a PCA was executed in advance for each surface energy and PSD data to reduce the number of data points without loss of information. Therefore, the PCA described in this study includes the principal components of these preliminary PCAs of PSD and iGC instead of directly using the respective raw data. For the PSD-PCA, PSD data was normalized and mean-centered and for the iGC-PCA, surface energy data was mean centered and unit variance scaled. For both PCAs the NIPALS (Nonlinear Iterative Partial Least Square) algorithm was applied, which is a commonly used iterative approximation method to calculate the principal components (Wold, 1974). A maximum of 100 iterations was used. Detailed information on the PSD-PCA and iGC-PCA can be found in the supplementary material A1 and A2.

2.2.9.2. PCA of material descriptors. Due to the preliminary PCA models established for the particle size distribution data and surface energy data the number of powder and material descriptors and therefore dimensions for the PCA of the material descriptors could be reduced from 119 to 29. PCA was then performed with these 29 descriptors acquired from material characterization of the powders and roller compacted materials (listed in Table 3) and the 34 materials from the training sample set described in Table 1. All descriptors including the principal components from PSD-PCA and iGC-PCA were mean centered and unit variance scaled before the analysis. For the algorithm, singular value decomposition (SVD) was used.

In addition, a PCA of the TTS of the training sample set was performed to identify the correlation between the tensile strengths at different compression pressures. The TTS values were mean centered and unit variance scaled. The NIPALS algorithm with a maximum of 100 iterations and a 10-fold cross validation with randomized segmentation was applied.

Further information can be found in the supplementary material A3 and A4.

2.2.9.3. PLS of material descriptors and tablet tensile strength. PLS regression models were created with the 34 materials from the training sample set to identify the correlations of material descriptors of powders and roller compacted materials with their tablet tensile strength. Therefore, the 29 powder and roller compacted material descriptors listed in Table 3 were used as input parameter and the TTS at 200 MPa compression pressure as result parameter. To evaluate the ability to reduce experimental burden by PCA, a second PLS model was created with a smaller input data set by exclusion of material descriptors based on the PCA described in 2.2.9.2 (excluded descriptors are marked with * in Table 3) and compared to the PLS model established with the initial 29 descriptors. The input parameters as well as TTS at 200 MPa as result parameter were mean centered and unit variance scaled for the partial least squares regression. The NIPALS algorithm with a maximum of 100 iterations was applied. Details of the PLS model with the 29 initial input parameters are described in supplementary material A5 and details of the PLS model with the reduced input parameter set are described in

Table 3
Overview of material descriptors of powders and roller compacted materials for PCA and as PLS input data (X). Further information on the preliminary PCAs of PSD and surface energy data can be found in supplementary material A1 and A2. Descriptors excluded as indicated by PCA described under 3.1 for the second PLS regression model described under 3.4 are marked with *.

Abbreviation	Descriptor	Pretreatment before mean centering and unit variance scaling
nBFE	Basic flow energy	BFE normalized to sample mass
SI	Stability index	
FRI	Flow rate index	
SE	Specific energy	
CBD*	Conditioned bulk density	
PF-0, PF-500	Bulk packing fraction, tapped packing fraction	
HR, Carr*	Hausner ratio, Carr index	
PycD	Gas pycnometric density	
CPS-1, CPS-15*	Powder compressibility at 1 and 15 kPa compression pressure	Logarithmized
PD-1, PD-15*	Pressure drop at 1 and 15 kPa compression pressure	Logarithmized
Ffc*	Flow function coefficient	Logarithmized
Coh*	Cohesion	Logarithmized
AIF*, AIFe	Angle of internal friction, effective angle of internal friction	Logarithmized
WFA-7, WFA-19*	Wall friction at 7 and 19 kPa pressure	Logarithmized
SSA	Specific surface area	
iGC-PC1	iGC-principal component 1: positively impacted by work of cohesion (WoC), total surface energy (γ_t) and polar surface energy (γ_{ab})	PCA of mean-centered and unit variance scaled surface energy data
iGC-PC2	iGC-principal component 2: positively impacted by dispersive surface energy (γ_d) negatively impacted by polarity (γ_{pol})	PCA of mean-centered and unit variance scaled surface energy data
iGC-PC3	iGC-principal component 3: negatively impacted by the acidity/basicity ratio (KaKb)	PCA of mean-centered and unit variance scaled surface energy data
iGC-PC4	iGC-principal component 4: negatively impacted by the total surface energy d90 value (γ_t d90)	PCA of mean-centered and unit variance scaled surface energy data
iGC-PC5	iGC-principal component 5: weak positive impact by KaKb at 6 % coverage (KaKb6) and weak negative impact by total surface energy and work of cohesion at 20 % surface coverage (γ_t20, Woc20)	PCA of mean-centered and unit variance scaled surface energy data
PSD-PC1	PSD-principal component 1: positively impacted by particle sizes 40–100 μm, negatively impacted by 200–400μm	PCA of normalized and mean-centered particle size distribution data
PSD-PC2	PSD-principal component 2: positively impacted by particle sizes 400–800 μm, negatively impacted by particle sizes 100–200 μm	PCA of normalized and mean-centered particle size distribution data
PSD-PC3	PSD-principal component 3: positively impacted by fine fraction (particle sizes $\leq$ 20 μm)	PCA of normalized and mean-centered particle size distribution data

supplementary material A7.

2.2.9.4. Validation of PCA and PLS model. For internal validation of the created PCA and PLS regression models, a 10-fold cross validation with randomized segmentation for each of the models was performed.

For external validation of the PLS models, TTS at 200 MPa compression pressure of new materials unknown to the model, the test data set (Table 2), was predicted by the PLS model based on the new material's powder descriptors and subsequently compared to the experimentally determined TTS data of the materials. Predictions were performed based on only the first two factors of the PLS model as indicated by the root mean square error with the same algorithm and data pretreatment that was applied for the creation of the PLS model.

The deviation estimates the reliability of the predicted Y values and is calculated from sample leverage, the overall model error and residual variance, which gives the mean squared residuals corrected for the respective degrees of freedom:

$$Y\ Deviation = \sqrt{ResYValVar\left(\frac{ResXValSamp_{pred}}{ResXValTot} + Hi + \frac{1}{I_{Cal}}\right)\left(1 - \frac{a+1}{I_{Cal}}\right)}$$

ResYValVar is the residual variance per Y using the validation samples, $ResXValSamp_{pred}$ is the residual X variance per sample using the validation samples, ResXValTot is the total residual X variance using the validation samples, a is the number of principal components or factors, H is the leverage of the samples and i and I gives number of samples.

To predict the tablet tensile strength of the test sample set, the scores of the test sample set for the preliminary PSD-PCA and iGC-PCA is needed as input parameter. Therefore, the test samples were projected onto the respective training sample map generated by the PSD-PCA and iGC-PCA model without affecting the existing structure of the preliminary PCAs. For each PCA model, the same data pretreatment and algorithms that were applied to the training samples were also applied to the test data for the projections. Further information can be found in the supplementary material A6.

3. Results and discussion

3.1. Principal component analysis (PCA) of material descriptors of powders and roller compacted materials

PCA and PLS are common methods to reduce dimensionality in a data matrix with as little loss of information as possible. These reduced dimensions can be well graphically represented and underlying correlations between the initial descriptors (PCA) as well as descriptors and responses (PLS) become much more evident, which enables the explanation of variability and detection of interactions that cannot be captured by bivariate analysis. Therefore, PCA was applied to the 34 materials (Table 1), which included powders and roller compacted materials of neat APIs, fillers and their respective binary blends, as well as their descriptors obtained for this study (Table 3) to elucidate how different powder descriptors are related, how the API, excipient and drug load impact blend properties and how different process routes transfer into the material properties.

By applying preliminary PCA to the surface energy and particle size distribution data to avoid overrepresentation of these two methods in the descriptor set to reduce the data points with as little loss in information as possible, the surface energy data was reduced from 30 to 5 and the particle size distribution data was reduced from 68 to 3 descriptors and therefore dimensions. In the preliminary PCA regarding the particle size distribution data, 3 principal components explained over 97 % of the variance in the PSD data with PSD-PC 1 being strongly influenced by particles with particle sizes between 40 and 100 μm as well as particle sizes between 200 and 400 μm, PSD-PC 2 was strongly influenced by particles with particle sizes between 400 and 800 μm as well as 100–200 μm, and PSD-PC 3 was mainly influenced by particles with particle size $\leq$ 20 μm, thus the fine fraction. Further details such as the correlation loadings plots of this preliminary PSD-PCA can be found in supplementary material A1. In the preliminary PCA regarding the surface energy data, over 97 % of the variability in the surface energy data was explained with 5 principal components. Clusters and strong correlations were found between the same surface energy descriptors at different surface coverage percentages as well as the respective surface energy distribution descriptors. iGC-PC 1 was strongly influenced by the descriptors of work of cohesion, polar surface energy descriptors and descriptors of the total surface energy whereas iGC-PC 2 was mainly

impacted by the multiple descriptors of dispersive surface energy and polarity. iGC-PC 3 was mostly impacted by descriptors of the acidity-basicity ratio, iGC-PC 4 was strongly influenced by the total surface energy d90 value and iGC-PC 5 had multiple impact factors such as the acidity-basicity ratio at 6 % surface coverage, the total surface energy and work of cohesion at 20 % surface coverage. Further details of the preliminary iGC-PCA can be found in supplementary material A2. Thus, by including the 3 principal components from the preliminary PSD-PCA instead of the initial 68 particle size descriptors and the 5 principal components from the preliminary iGC-PCA instead of the initial 50 surface energy descriptors, the number of PSD and surface energy descriptors in the subsequent PCA with the other material descriptors was extensively reduced with minimal loss in PSD and surface energy information but without risking overrepresentation of these methods in the PCA.

By applying preliminary PCA to the iGC and PSD data, the initial 119 material descriptors were combined to the 29 descriptors shown in Table 3. PCA of the resulting 29 powder descriptors (Table 3) reduced them into 6 independent principal components. Hence, overall 119 material descriptors and therefore dimensions were reduced to only 6 dimensions with as little loss in the description of data variance as possible as these 6 principal components combined described 89 % of the variance in the data (Fig. 1).

The principal components can be displayed in two-dimensional plots of the correlation loadings to visualize the relations of the different material descriptors (Fig. 2). The ellipse in the plots indicates 100 % explained variance. The importance of individual descriptors for a principal component is indicated by their position in the plot as descriptors closer to the edges of the plot have higher impact on the PC whereas variables within the center are less relevant for the PCs. Descriptors close to each other in the loadings plot are positively correlated whereas descriptors in opposite directions have a negative correlation. The first percentage in the brackets after a PC's label describes how much X data variance is explained by this PC with the complete trainings data set, the second percentage describes the predicted variance of X data by this PC from cross validation. Principal components 1 and 2 explain and predict most of the variance in the data set and with increasing PC number the variance explained by the PC decreases.

Fig. 2a displays the correlation loadings plot of PC 1 and PC 2, which together explained already 57 % and predicted 43 % and thus the majority of variance in the data set. High relevance for PC 1, which explained 30 % data variance but predicted only 7 %, and a strong positive correlation is demonstrated for Hausner ratio (HR), the Carr index (Carr), powder compressibility at 15 kPa (CPS-15), the angle of internal friction (AIF) and effective angle of internal friction (AIFe), the specific energy (SE), the cohesion (Coh) and PSD-PC 1. Furthermore, a high relevance for PC 1 and a negative correlation to the above mentioned descriptors was shown for the flow function coefficient (Ffc). Materials with a high flowability display high values of Ffc. High values of Hausner ratio or Carr index for a material are indicators of low flowability (European Directorate for the Quality of Medicines and Health Care, 2022). The PCA model identifying this known negative correlation between Ffc and Hausner ratio and Carr index indicated that the model is capable to indicate those correlations between powder descriptors correctly. Carr and HR are both ratios of a material volume or density without and with application of mechanical force which also applies for the powder compressibility at 15 kPa measured by FT4 powder rheometer. Therefore, even though the force is applied by tapping for HR and Carr and by compression punch with defined force for powder compressibility, the CPS-15 is positively correlated to the Carr and HR as all 3 descriptors contain comparable information which is another confirmation for the model. The angle of internal friction and the effective angle of internal friction are measures of the internal resistance per unit area against flow. AIF indicates internal friction at the onset of flow whereas AIFe is the descriptor of internal friction at steady state flow, but both describe cohesive and frictional forces between adjacent particles which explains the positive correlation in PC 1. High AIF and AIFe values are a sign of high cohesive and friction forces which can lead to low flowability and high volumes of entrained air in the powder bulk, which can still be densified more by applying force than already less cohesive, efficiently packed powders. Therefore, it is reasonable, that the PCA demonstrated a positive correlation between AIF and AIFe with HR, Carr and CPS-15 and a negative correlation to Ffc. The specific energy SE measured by FT4 powder rheometer describes the energy needed for a powder to flow in an unconfined, low stress environment. High specific energy values often indicate high cohesive forces between the particles of a material, which explains its positive correlation to the cohesion in the powder bulk measured by RST and the AIF and AIFe and negative correlation to the Ffc. For PSD-PC 1, mainly the smaller PSD values between 40 and 100 μm were relevant with a positive impact as shown in the supplementary material A1. High fractions of these smaller particles in the material positively correlated with the descriptors of high cohesion and negatively correlated with Ffc in the PCA model as small particles often lead to higher surface area compared to larger particles resulting in more interparticle interactions. Further, PSD-PC 1 also showed that larger particles in the size range of 200–400 μm had reduced cohesive forces. Therefore, it can be concluded that the largest variance in the data set of powders and roller compacted materials, represented by PC 1, was mainly explained by descriptors of flow and cohesive forces within the materials. Descriptors from different methods expected to correlate (positively or negatively) from prior knowledge were identified correctly by the model.

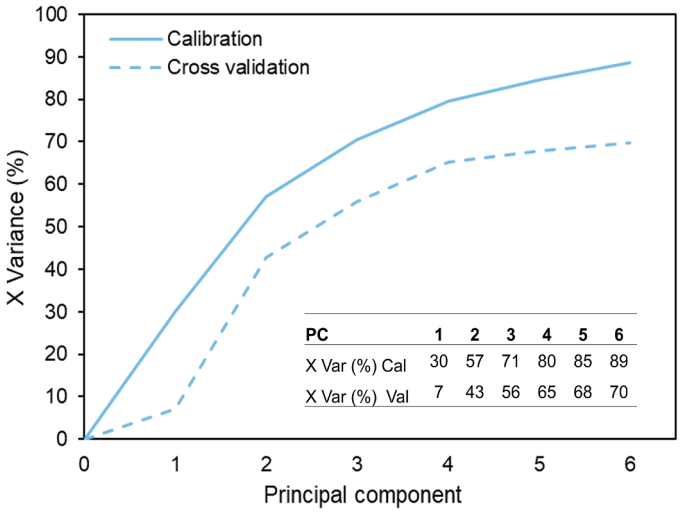

Fig. 1. Cumulative explained X-variance for the principal components of the PCA model.

Relevant material descriptors for the second PC, which still explained 27 % and predicted 36 % of data variance, were the bulk and tapped packing fraction, the conditioned bulk density, the pressure drop and PSD-PC 3, that are positively correlated to each other. A negative correlation of importance for these descriptors was found with the wall friction angles (WFA-7, WFA-19) as well as iGC-PC 1.

IGC-PC 1 was mainly described by the total and polar surface energy as well as work of cohesion and is therefore a descriptor for interparticle interaction on particle surfaces. The wall friction angle is an indicator of adhesive forces as it describes friction between the particles of a material with a wall material and can therefore be considered another descriptor for interactions of particles, which explains the positive correlation between the iGC-PC 1 and the wall friction angles. Wall friction angles at both different wall normal stresses are highly positively correlated, which confirmed again that the different measurement settings in this case result in the same information. High adhesive and cohesive forces can lead to low densities, especially bulk densities, in the materials as material with high interparticle interactions are less efficiently packed

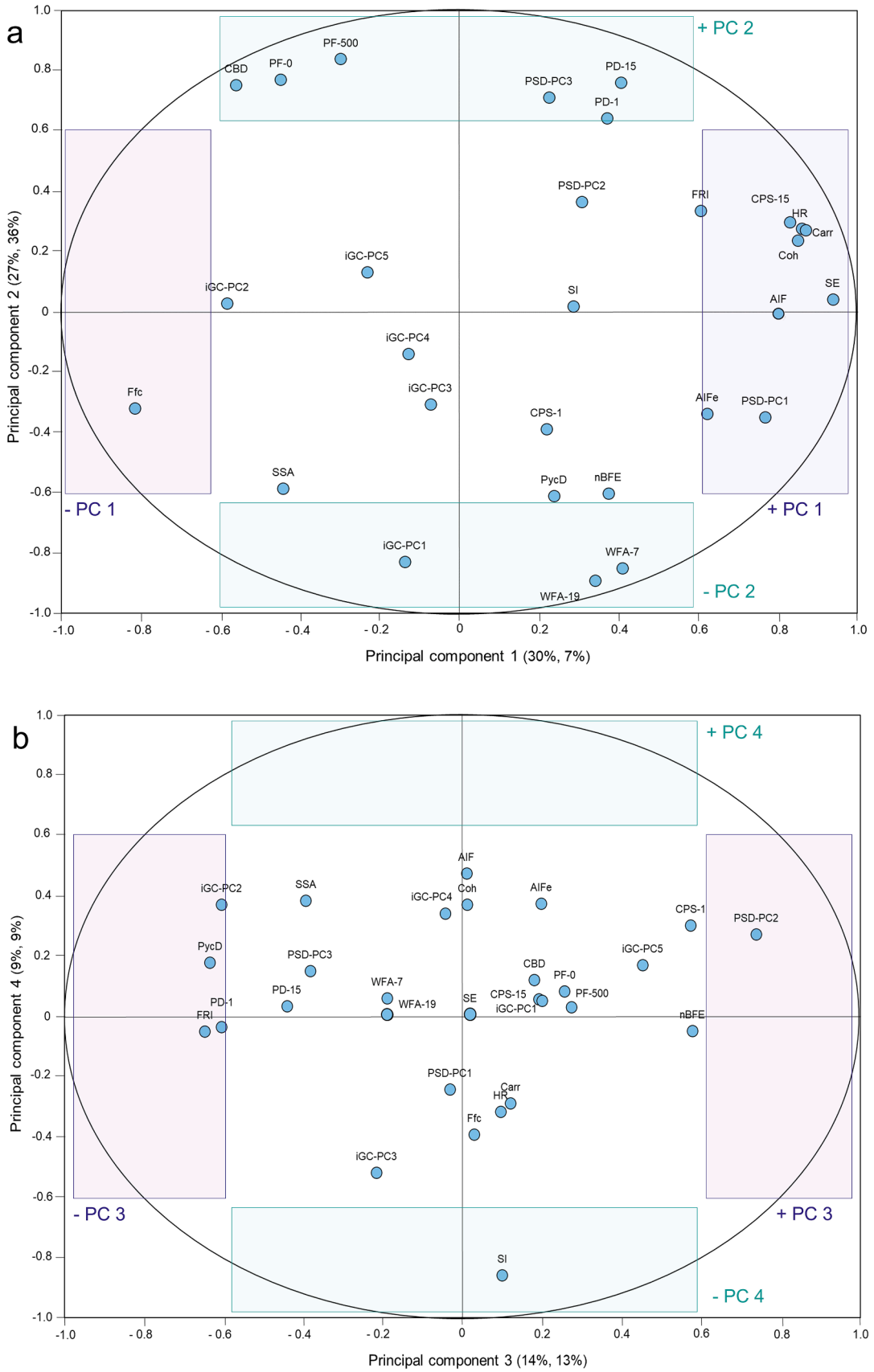

Fig. 2. Correlation loadings plots of PC 1 and PC 2 (a) as well as PC 3 and PC 4 (b) from principal component analysis. The first percentage value in the brackets after the PC title describes the variance in the data explained by the PC whereas the second percentage is the variance predicted by the PC in cross-validation. Areas of strong positive (+) and negative (-) impact for the PCs are highlighted in violet and green.

and have high volumes of air entrapped in the bulk which is demonstrated by the negative correlation of wall friction angles and iGC-PC 1 with bulk and tapped packing fraction as well as the conditioned bulk density. The strong positive correlation between the packing fractions determined with the Granupack™ and the CBD measured via the FT4 powder rheometer demonstrated that these different measurement methods lead to the same density information. The air pressure drop at two different compression pressures displayed a strong positive correlation indicating the different measurement settings here again lead to similar information.

PSD-PC 3, which contains the fine fraction ($\leq$ 20 μm) as main contributor with positive impact, showed a positive correlation to the pressure drop (PD-1, PD-15) as well as the density descriptors CBD, PF-0 and PF-500.

A high fine fraction is often related to higher interparticle interaction which is in alignment with a high air pressure drop, as air cannot flow easily through the powder bed. Despite potentially leading to higher cohesion in a powder by increased surface area, fine fraction could be displaying a positive correlation to the density descriptors as the very small particles below 20 μm fill in gaps between the larger particles in a powder bed leading to higher density. Negative correlations with PSD-PC 3 were shown for the wall friction angles (WFA-7, WFA-19) and iGC-PC 1, which was mainly impacted by the total surface energy and work of cohesion. As iGC-PC 1 and WFA also indicate high interparticle interaction potential, one could have expected a positive correlation, however the opposite was shown by PCA. One hypothesis is that surface energy and wall friction angle describe more the adhesive forces in a material whereas air pressure drop in a powder is more a descriptor of cohesive forces within the powder bulk. Further, a high pressure drop could also occur due to irregular particle shape, mechanical interlocking or surface roughness, and is therefore not only dependent on cohesiveness of a powder.

Fig. 2b shows the correlations loadings plot of PC 3 and PC 4, which together explained 23 % and predicted 22 % of the variance in the data set. For PC 3, which explained 14 % and predicted 13 % of data variance, the PSD-PC 2 was of highest importance which was mainly described by the positive impact of particle sizes between 400 and 800 μm and negative impact of particle sizes between 100 and 200 μm. Further negative impact on PC 3 was demonstrated by the flow rate index, pycnometric density and iGC-PC 2. The flow rate index describes if a material is sensitive to changes in the flow rate during processing. High values of FRI are common for more cohesive powders as they are more sensitive to changes in flow rate than non-cohesive or granular materials due to, for example, high air contents in the cohesive materials. A larger particle size (high values at 400 to 800 μm and low values at 100 to 200 μm) and thus a high PSD-PC 2 often leads to less interparticle interaction due to a lower surface area and therefore negatively correlates to the FRI. The iGC-PC 2 was positively impacted by the dispersive surface energy and negatively impacted by the surface polarity of a material. High values of dispersive surface energy could lead to overall more interparticle interaction and therefore also to higher FRI explaining the positive correlation between iGC-PC 2 and FRI.

Principal component 4 explained and predicted 9 % of the variance in the data set and was mainly described by the stability index with a negative impact. SI is an indicator of robustness of a material during processing and the stress applied by the FT4 powder rheometer's rotational forces. Instability in a material could be a result of deaeration of cohesive particles or electrostatic charge but also segregation due to particle size, attrition or deagglomeration, which describes the change in physical size and particle shape through mechanical stress. Therefore, roller compacted granules, for example, often tend to be less robust compared to initial powders.

Principal component 5 and 6 will not be discussed in detail here as PC 5 and 6 together only explained 9 % and predicted 5 % of the variance in the data set. Further information on PC 5 and 6 can be found in the supplementary material A4.

Score plots show the position of the analyzed samples in the two-dimensional representation of the PCs and are useful to determine differences and similarities among samples. Samples at similar positions within the area spanned by the two PCs under consideration are similar in terms of the descriptors subsumed in the PCs. Samples that lie far away from each other in the plot differ in their properties. As such, the combination with the correlation loadings plot enables to identify which descriptors are responsible for differences between the samples by the position of the sample in the score plot of the two respective components. Thus, this PCA model included all powders and roller compacted materials of Table 1 to identify how drug load, different excipients, different APIs as well as different processing of powders and roller compacted material impacted the descriptors of the respective material.

In principal component 1, which was as mentioned above mainly impacted by flowability descriptors, the samples could be distinguished by the excipients they contain (Fig. 3a). Samples containing Parteck® M200 were all located more on the left side of the score plot whereas samples containing MCC were more on the right side. Parteck® M200 had a very high flowability compared to MCC and therefore also a higher ffc value. The ffc value was located on the left side of the correlation loadings plot and thus the Parteck® M200 containing samples were in accordance with that also occurring on the left side of the score plot of PC 1 and PC 2. MCC containing samples displayed lower flowability than Parteck® M200 containing samples and had higher values of HR and Carr index and were therefore located on the right of the score plot as HR and Carr are on the correlation loadings plot. Thus, PC 1 illustrated that the samples could be distinguished mainly based on their flowability due to the excipients they contain even at high drug loads of 60 or 80 % which highlighted the influence the excipient had on the flowability of the blends.

The samples were furthermore distinguishable by their drug load on PC 2 (Fig. 3b). PC 2 was mainly impacted by descriptors of density and indicators of interparticle interaction. Packing fraction, CBD and PSD-PC 3 were all located in the upper area of the loadings plot of PC 1 and PC 2. As drug load of the samples increased from negative to positive values of PC 2, samples with high drug and the neat APIs were also located in the upper part of the score plot. This showed that with increasing drug load the densities of the samples as well as the fine fraction increased. Further experiments will show if this observation is specific only for the here chosen APIs or more generally valid. Samples with low drug load or neat excipients were set on the bottom half of the score plot which is in accordance with high values of wall friction and iGC-PC 1. This indicated that samples with low drug load and neat excipients displayed higher adhesive forces, higher work of cohesion and higher total surface energy compared to samples with high drug load. Therefore, it was clearly visualized on PC 2 that the different drug loads of the blends differed mainly in their densities as well as their adhesive and cohesive forces.

While the excipients' impact on data variability was covered by PC 1 (highest variance in dataset explained), variability introduced by the APIs and the impact of API on the blends were well shown in PC 3 as paracetamol containing samples were located right and caffeine containing samples were located left on the score plot of PC 3 and PC 4 (Fig. 3c). PC 3 was mainly impacted by the PSD-PC 2, which was displayed on the right side of the correlation loadings plot and was positively described by larger particle sizes between 400 and 800 μm and less on 100–200 μm. Paracetamol containing samples were therefore also located more to the right in the score plot as they demonstrated higher shares of particles with a size of 400–800 μm and lower on 100–200 μm and with that, consequently, higher values for PSD-PC 2. PC 3 was further described by a negative impact of iGC-PC 2 and with weaker impact by the pycnometric density and flow rate index, which were all located on the left side of the correlations loadings plot of PC 3 and PC 4. Hence, the score plot also showed that caffeine containing samples on the left side of the plot had higher values of iGC-PC 2 than paracetamol containing samples which indicated that the caffeine

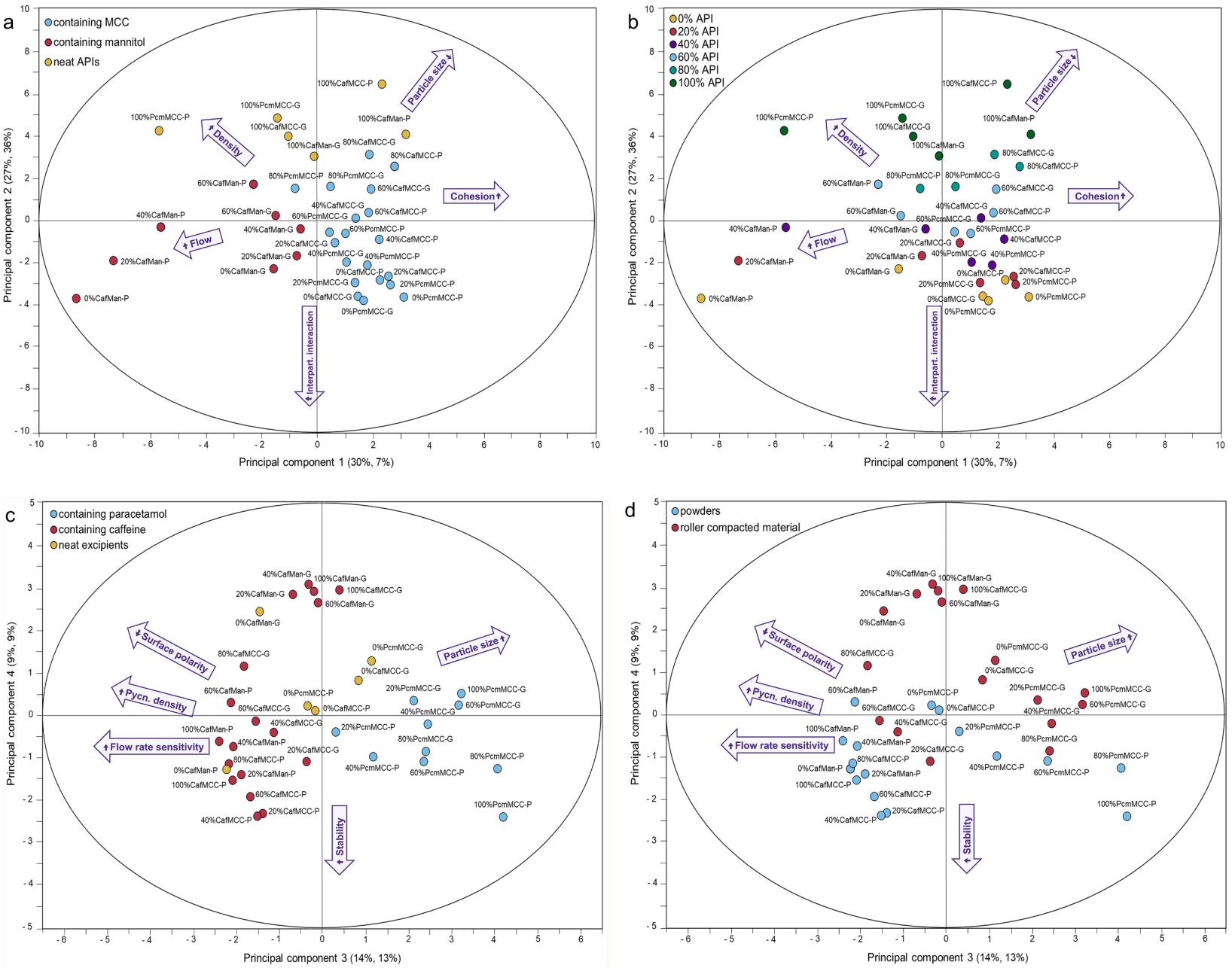

Fig. 3. Score plots from PCA for differentiation of samples according to excipients on PC 1/PC 2 (a), drug load on PC 1/PC 2 (b), APIs on PC 3/PC 4 (c) and processing stage on PC 3/PC 4 (d) The PCs were selected based on their ability to best discriminate the four distinguishing features.

samples had a higher dispersive energy and lower surface polarity than the paracetamol samples. Caffeine containing samples further had higher pycnometric density and were more sensitive to changes in the flow rate during processing than paracetamol containing samples irrespective of the excipients used. PC 3 therefore illustrated, that the material descriptors of the two APIs caffeine and paracetamol mainly differed in particle size but also pycnometric density, surface polarity and stability against flow rate changes, which was also reflected in the blends.

Furthermore, a trend was visible on PC 4 between the two different process states of the samples, the initial powders and the respective roller compacted materials (Fig. 3d). PC 4 was mainly impacted by the stability index SI, which is located on the bottom area of the correlation loadings plot of PC 3 and PC 4. The initial powders are also located rather in the bottom area of the score plot of PC 3 and PC 4 as they have higher SI values closer to 1 compared to the roller compacted materials. Even though the differentiation between the powders and roller compacted materials is not as clear as previously described for the APIs, the excipients or the drug loads, PC 4 visualized that initial powders are more stable when forced to flow compared to the roller compacted materials, which can be explained by segregation, attrition, or deagglomeration of the granular particles. The score plots demonstrated that the APIs, excipients and drug loads used could be well distinguished by different powder descriptors whereas differentiation between initial powders and respective roller compacted materials was only partly visible in the SI. Therefore, analysis of the loadings and scores of the PCA indicated that the initial material properties prevailed over the process route of direct compression and roller compaction in the present dataset.

Principal component analysis of the material descriptors clearly showed that the largest variability in the powder and roller compacted material data set was described by flow descriptors of cohesive forces. Correlations of different methods and descriptors were revealed by PCA, such as the strong correlation highlighted between HR, Carr, Coh and Ffc as well as AIF and AIFe as they give similar information. Further, a strong positive correlation was found by PCA in CBD measured with FT4 powder rheometer and the PF-0 and PF-500 from Granupack measurements. PD-1 and PD-15, CPS-1 and CPS-15 as well as WFA-7 and WFA-19 displayed high positive correlation and provided similar information, thus illustrating that reduction of multiple settings within one characterization method would be possible with almost no loss in information. Therefore, it was demonstrated that PCA allows for a data-based reduction of multiple settings within one characterization method and thus, time resources. However, although strong correlations between descriptors from different methods was demonstrated, each of the powder characterization methods described here often leads to multiple different material descriptors and none of the methods could have been excluded completely without loss in information, thus the reduction of analyses by PCA and therefore material resources and consequently also development costs is only possible to a certain extent.

By systematically changing API and excipient in the powder systems and including different drug loads, the PCA was able to highlight how the different materials impacted the blend properties. Excipients strongly influenced the flow descriptors of the blends whereas APIs drove the differentiation of the blends by their particle size, surface polarity and dispersive surface energy. The different drug loads differed mostly in their densities as well as the cohesive-adhesive forces. Initial powders and roller compacted materials showed only differentiation in the PCA in the stability index, however the process differentiation was not as clear as differentiation between the materials highlighting again the importance of initial material properties for formulation development.

3.2. Partial least squares regression (PLS) for tabletability

Partial least squares regression was applied to better elucidate the impact of material descriptors on tablet tensile strength in direct compression and roller compaction. For creation of the PLS model all 34 materials in Table 1, thus neat APIs, fillers and binary blends thereof with varying drug load in two different process routes, roller compaction and direct compression, were included to identify how drug load, different excipients, different APIs as well as different processing of powders and roller compacted material impacted the tablet tensile strength of the respective material.

All 29 powder and roller compacted material descriptors in Table 3 were included as input parameters (X) in the PLS model to minimize the risk of excluding a descriptor that is potentially critical for mechanical tablet properties of materials yet unknown to the PLS model.

A PCA of the TTS values at different applied compression pressures showed their strong positive linear correlation to each other which was also described by Berkenkemper et al. (Berkenkemper et al., 2023). Thus, the TTS at one compression pressure, the TTS at 200 MPa compression pressure, was used as response (Y) in the PLS model. Detailed information on the PCA of TTS at different compression pressures can be found in the supplementary material A3.

As in PCA, the first newly created overarching independent variable in the PLS model, called factor 1, describes the most variance of the response (Y) data explained by the input (X) data and factor 2 describes the second most variance in the response (Wold et al., 2001). Fig. 4a displays the cumulative explained variance of the TTS data at 200 MPa compression pressure for the number of factors included in the model. It was concluded that already with the first factor most of the variance in the tensile strength data was explained (73 %) and predicted with cross validation (69 %). The root mean square error (RMSE) is displayed in Fig. 4b and has the unit of the response parameter, therefore MPa. RMSE was overall relatively large but decreased with more factors included in the model. Thus, although the first two factors explain together already 84 % of the TTS data variance, 5 factors were included to create the model with a low RMSE (0.45 MPa for the calibration and 0.85 MPa for the cross validation).

The input data, thus the 29 powder and roller compacted material descriptors described in Table 3, was plotted in the two-dimensional space of the new independent variables (factors) in a correlation loadings plot. 100 % explained variance of the response, the TTS data, were again indicated by the ellipse in the plot to reflect variable importance. The first percentage in the brackets after a factor's label describes how much variance in the response (Y) data is explained by this factor with the complete trainings data set, the second percentage describes the predicted variance of Y by this factor from cross validation. Input parameters close to the response are positively correlated whereas diagonally opposed input parameter are negatively correlated to the response.

In the correlation loadings plot of factor 1 and 2 (Fig. 5) it became evident, as already shown by the explained variance plot, that factor 1 describes the majority of the variance in the TTS data as TTS at 200 MPa compression pressure has a high value of 0.85 for factor 1 and is located on the right side of the plot but has a much smaller value of 0.34 for factor 2 indicating already less relevance of factor 2 for the variance in the response data.

A strong impact on factor 1 and strong positive correlation to TTS was seen for the wall friction angles WFA-7 and WFA-19, the effective angle of internal friction, the mass-normalized basic flow energy as well as iGC-PC 1, which was mainly described by the positive impact of total and polar surface energy (γ_t, γ_{ab}) and work of cohesion WoC. Thus, high values of these powder and roller compacted material descriptors indicate high tablet tensile strength. High internal friction and basic flow energy are indicators for high cohesive forces in a powder bulk whereas high wall friction angles hint at high adhesive forces in a material. High γ_t and WoC values also suggest high potential for interparticle interactions. The positive correlation between TTS and WFA-7, WFA-19, AIFe, nBFE, γ_t, γ_{ab} and WoC can be explained by these interparticle interactions due to adhesive and cohesive forces, which could lead to more interparticle bonding during the tableting process and consequently to harder tablets and higher tablet tensile strength values.

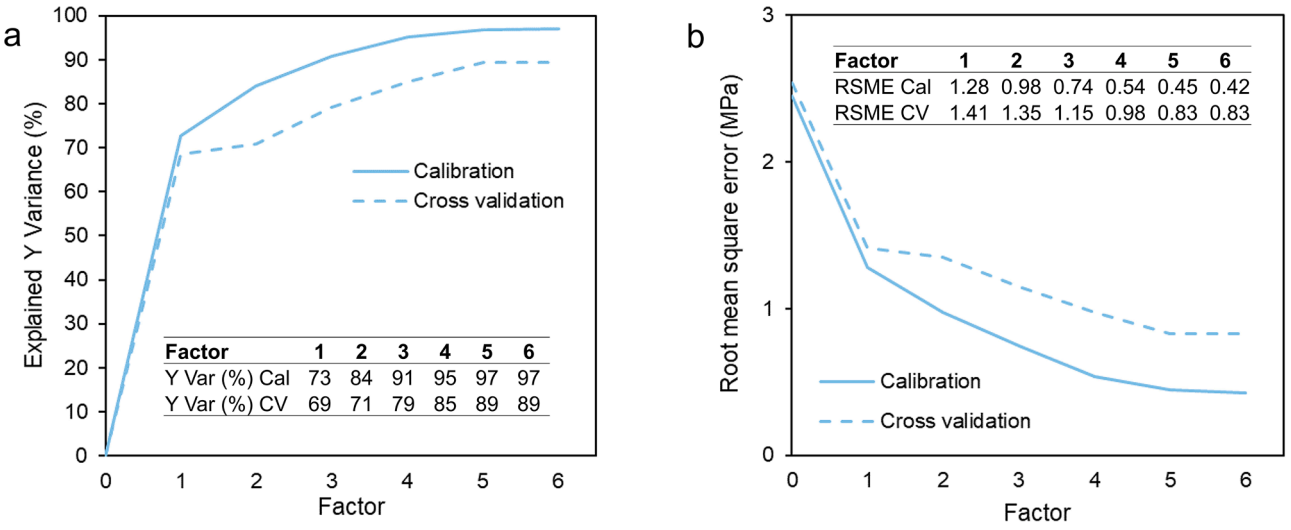

Fig. 4. Cumulative explained Y-variance (a) and root mean square error (RMSE) (b) for the factors of the PLS model.

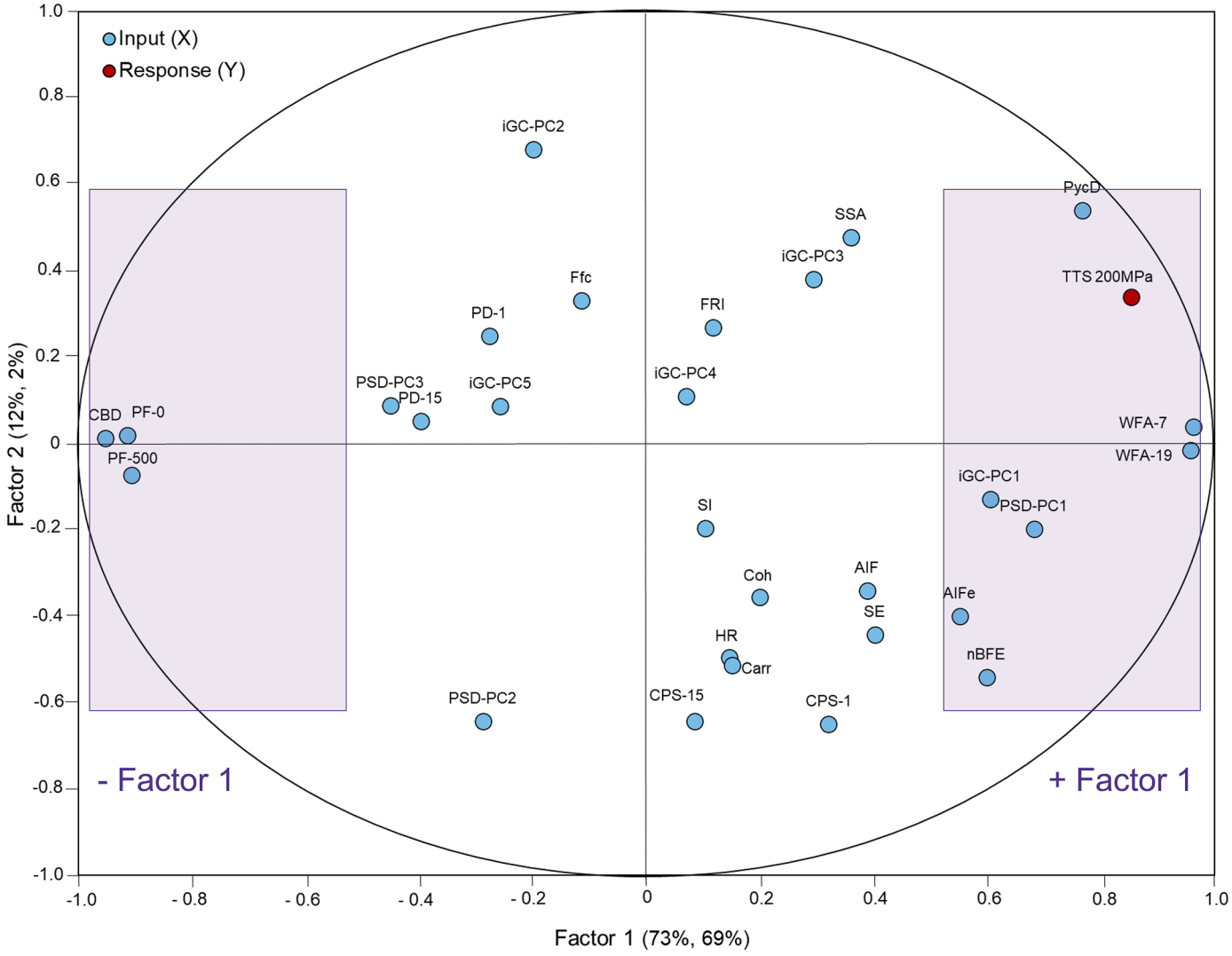

Fig. 5. Correlation loadings plot of factor 1 and 2 of the PLS regression model. Areas of strong positive (+) and negative (-) impact for factor 1 are highlighted in violet.

The PLS model further displayed a positive correlation between the TTS and PSD-PC1, which was positively impacted by particle sizes 40–100 μm, as smaller particles can create more bonds per cross-sectional area during tablet compression compared to larger particles and thus lead to stronger tablets.

Furthermore, a strong correlation of TTS with density descriptors was shown by the PLS model as TTS displayed a positive correlation with the pycnometric density as well as a negative correlation to the density descriptors PF-0, PF-500 and CBD, which is supported by the findings of the multivariate analysis of different MCCs by Thoorens et al. (Thoorens et al., 2015). Higher TTS values in low bulk and tapped density materials could be attributed to multiple reasons such as cohesive properties of the material leading to entrapped air in the bulk which is released under compression or large size distribution that could impact particle packing.

By including a large set of 119 powder and roller compacted material descriptors, it was possible to show systematically for the first time that a relatively small group of key descriptors is sufficient to obtain information about the mechanical tablet properties of a pharmaceutical powder in roller compaction and direct compression. The identified key descriptors for tabletability in direct compression and roller compaction included descriptors already discussed in literature, such as descriptors of density or particle size (Soh et al., 2008; Thoorens et al., 2015) but also included the surface energy, work of cohesion and wall friction angle which were identified for the first time as key material properties for tabletability in both roller compaction and direct compression. Further it was demonstrated that multiple other material descriptors, such as descriptors of flow, do not impact tablet tensile strength, which opposes the findings of Souihi et al. (Souihi et al., 2013), where a correlation of the flowability of mannitol, determined with the FT4 powder rheometer by compressibility and permeability, with the final blend tablet tensile strength was discussed. This could be likely attributed to the fact that compressibility and permeability measured by the FT4 powder rheometer are impacted by multiple factors including interparticle interactions as it was shown by the correlations in the PCA in this study and discussed under 3.1. principal component 1 and 2 where permeability displayed as pressure drop was correlated to descriptors of adhesive forces such as the wall friction and compressibility was correlated to descriptors of cohesive forces such as the effective angle of internal friction. As descriptors of adhesive and cohesive forces displayed a correlation to TTS in the PLS of this study it can be reconstructed that a potential correlation of permeability and compressibility was identified in the study by Souihi et al. However, it clearly demonstrates the importance of systematic and extensive material characterization as it was shown in the PLS with 119 material descriptors and 34 materials including neat APIs, fillers and their binary mixtures with varying drug load in the here described study that not flowability as described by Souihi et al. but the underlying interparticle interactions such as cohesive and adhesive forces impact the tablet tensile strength. Furthermore, in the PCA, the largest variance in the data of the powder and roller compacted material descriptors was described by indicators of flowability as shown in PC 1. However, in PLS regression, the largest variance in the response data, the tablet tensile strength, was described by descriptors of density, particle size and interparticle interaction in factor 1. Thus, it was highlighted, that the different information that PCA and PLS analysis deliver can be synergistically combined to obtain a deeper understanding on key attributes for manufacturability of oral solid dosage forms.

Comparable to PCA, the score plot of the PLS factors shows the position and therefore differences and similarities of the analyzed samples in the two-dimensional representation of the factors. In combination with the correlation loadings plot it is again possible to identify which input parameters are responsible for the identified differences with regard to their impact on tablet tensile strength.

In the plot of factor 1 and 2, the samples can be distinguished on factor 1 by their drug load as drug load increased from right to left (Fig. 6a). In the corresponding correlation loadings plot, the response tablet tensile strength was mainly explained by factor 1 and was positioned on the positive, right side of the plot with positively correlating powder and roller compacted material descriptors such as the pycnometric density, WFA-7 and WFA-19, AIFe, nBFE, iGC-PC 1, which was mainly described by the positive impact of total and polar surface energy (γ_t, γ_{ab}) and work of cohesion WoC as well as the PSD-PC1, which was positively impacted by particle sizes 40–100 μm whereas the negatively correlating input parameters PF-0,PF-500 and CBD were located on the left side of the plot. As the neat excipients MCC and Parteck® M200, so 0 % drug load, and the lower drug loads were positioned to the right and neat APIs and higher drug loads were on the left, the score plot indicated that neat excipients had highest tablet tensile strengths and with increasing drug load the TTS decreased as well as WFA, AIFe, nBFE, pycnometric density, surface energy and WoC as well as the share of particles with 40–100 μm particle size. Conversely, neat APIs and samples with high drug loads have higher values for PF-0, PF-500 and CBD compared to samples with lower drug load, which is described in detail in the score section of the PCA under 3.1.

As discussed under 3.1., the impact of processing route on the materials was mainly distinguished in the PCA due to their different stabilities when exposed to rotational stress. This stability index SI did not demonstrate an impact on the first two factors of the PLS and no correlation to TTS was shown, thus no differentiation between the two process routes direct compression and roller compaction, was visible on the first two factors in the PLS regression (Fig. 6b). This highlights that the initial material descriptors have such a high impact on mechanical tablet properties that at least in terms of direct compression and roller compaction and the process parameters utilized in this study, a change in process route does not automatically lead to a change in tablet processability regarding the mechanical properties.

Thus, not only were key material attributes for tablet tensile strength in direct compression and roller compaction identified from a large set of 119 material descriptors, but it was also identified how drug load, different excipients, different APIs as well as different processing of powders and roller compacted material impacted the tablet tensile strength of the respective material by including 34 neat APIs, fillers and binary blends thereof with varying drug load in two different process routes, roller compaction and direct compression, in the creation of the PLS model.

3.3. External validation of PLS model

Beyond cross validation, the PLS model and the correlations of material descriptors and TTS derived were validated by applying the model to unknown samples to predict their TTS at 200 MPa compression pressure with the powder descriptors shown to be relevant by the model. Powders and respective roller compacted materials of neat compound A, Parteck® M100 and blends of compound A with Parteck® M100 with 30 % and 50 % drug load as well as neat Parteck® Delta M were used as test set. Details on the test sample compositions are shown in Table 2. The TTS of the test sample set was predicted using the first two factors of the established PLS model as suggested by the lowest RMSE of prediction for two factors with 0.88 MPa (Fig. 7a). The Hotelling's T^2 scores in Fig. 7b illustrate that none of the test samples were classified as outliers by the model but the neat compound A powder and roller compacted material demonstrated high Hotelling's T^2 scores as compound A is a micronized material and therefore stresses the model with a particle size distribution outside of the range that has been included in the model. In Fig. 8, the reference and the TTS values at 200 MPa compression pressure predicted by the PLS model are displayed. The model was able to predict tabletability trends in the test sample set correctly as tabletability decreased with increasing drug load in the Compound A-Parteck® M100 blends in the powders as well as roller compacted materials. Largest deviations of predicted and reference TTS were found in the neat compound A powder as well as the powder blends with compound A as

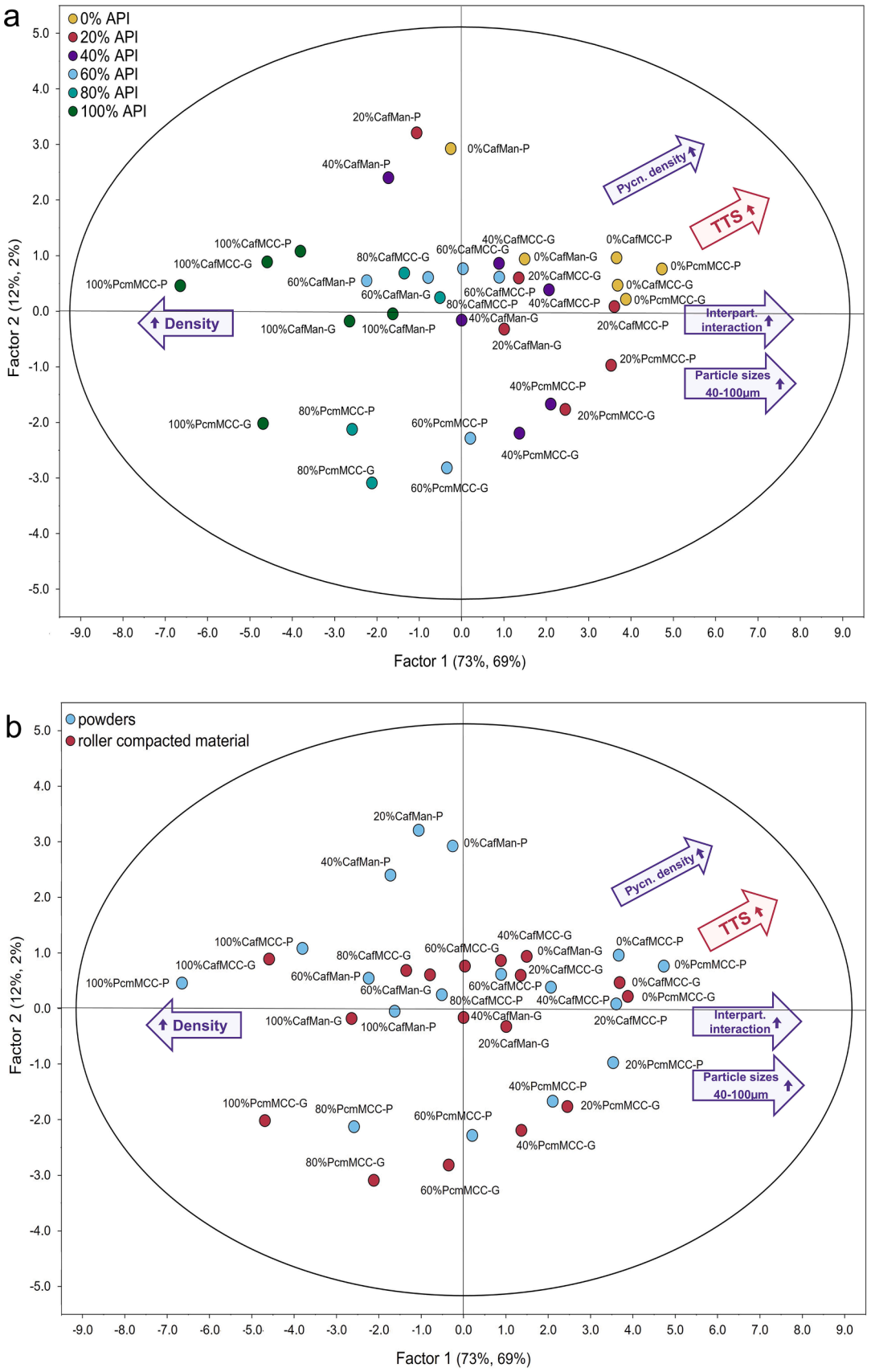

Fig. 6. Score plots from PLS for differentiation of samples on factor 1 and 2 according to drug load (a) and processing stage (b).

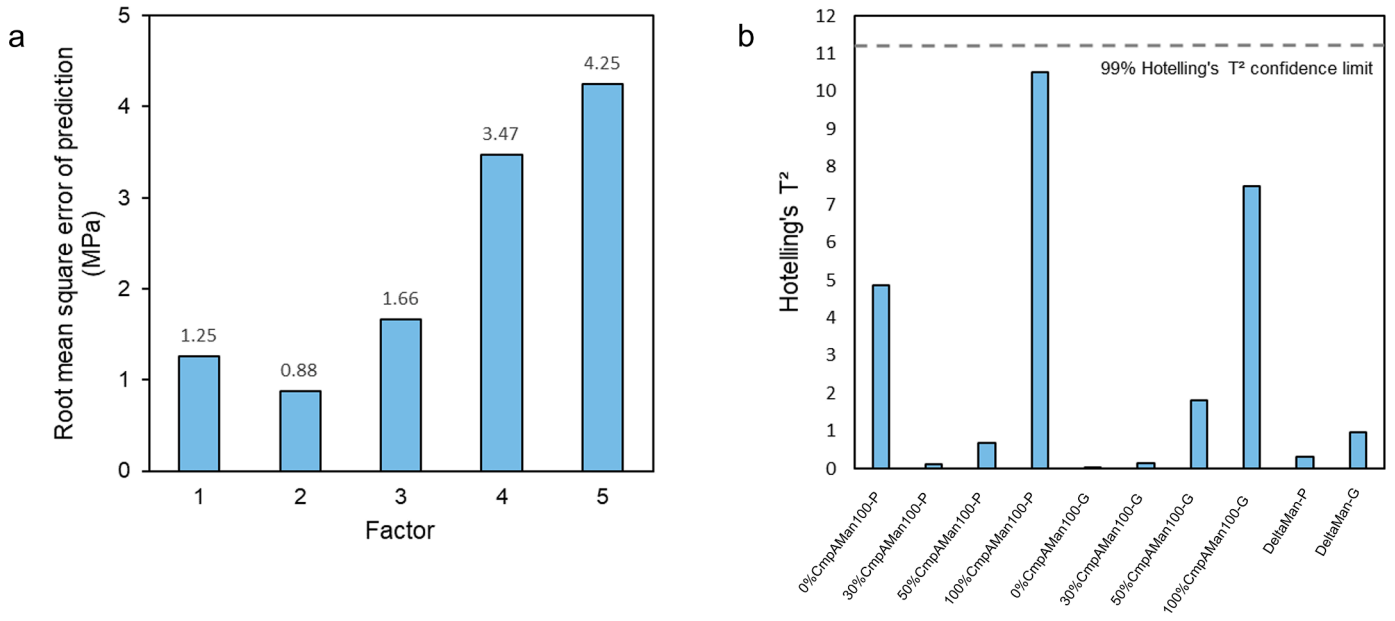

Fig. 7. Hotelling's T^2 plot for the test samples to externally validate the PLS model using 2 factors.

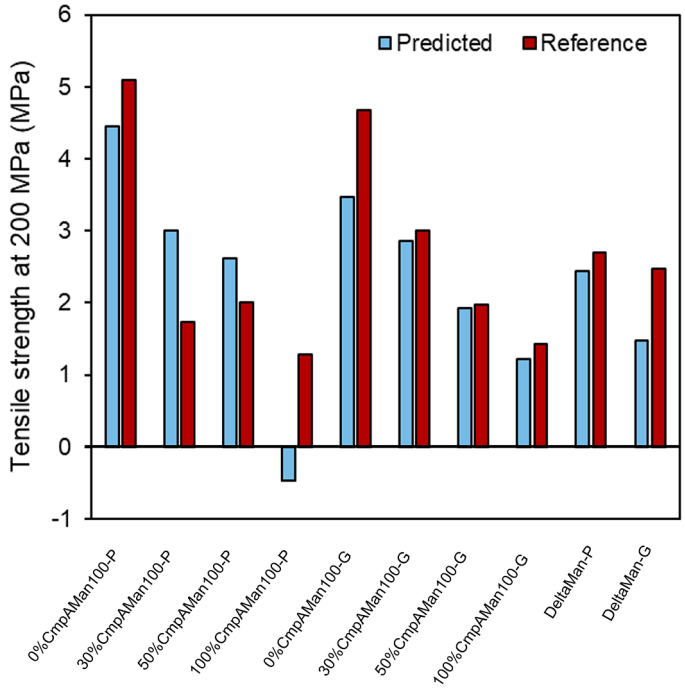

Fig. 8. Predicted TTS at 200 MPa compression pressure of test sample set using the established PLS model with 2 factors and reference TTS (experimentally obtained).

compound A is a micronized material with a particle size range yet unknown to the PLS model. As the roller compacted materials have particle sizes that are in the range of PSDs included in the model, the TTS prediction of roller compacted compound A and compound A blends is more accurate. Therefore, the external validation illustrated that the here established PLS model was able to identify important powder descriptors impacting the TTS and was able to correctly predict tabletability trends in unknown materials in direct compression as well as roller compaction. Opposed to the study of Boersen et al. (Boersen et al., 2015), where no direct correlation between the tabletability of the final tablets and the physical attributes of the initial materials was found and tabletability prediction of an unknown material was not feasible, key material attributes of the initial materials for tablet tensile strength were identified in the here described study via PCA and PLS. By using an extensive data set of 119 material descriptors and 44 materials including neat APIs and fillers as well as their binary blends with varying drug load in two process routes, direct compression and roller compaction for creation and validation of the model, correlations between powder properties and mechanical tablet properties were clearly demonstrated and the tabletability trends of unknown materials to the model were successfully predicted. However, also clear limitations of the model were identified by including materials stressing the model such as the micronized compound A. Further details on the external validation can be found in supplementary material A6.

Thus, the multivariate data analysis approach in this study provides a solid first basis to predict TTS of pharmaceutical powders and roller compacted materials by systematic characterization of a large set of material descriptors. However, it is important to note that the model cannot claim having included the entirety of all powder descriptors possibly impacting TTS. Additionally, the systematic selection of samples resulted in only a limited number of different materials and thus limited general transferability to all materials.

3.4. Comparison of established PLS model to a PLS model with input data reduced via PCA

As PCA is applied to determine which powder and roller compacted material properties are correlated and describe the largest variance in the powder and roller compacted materials data (X) whereas PLS reveals correlations of powder and roller compacted material properties with tablet tensile strength and which powder properties describe the largest variance in the tablet tensile strength data (Y), PCA and PLS provide different information, and the material descriptors that explain a large variance in the powder and rolled material data (X) do not necessarily strongly impact the tabletability (Y) as demonstrated in 3.1 and 3.2. Therefore, by combing the information from PCA and PLS and comparing two PLS models, one using the initial material descriptor set as described in 3.2 and a second PLS model using a material descriptor set proposed by the identified correlations in the PCA model in 3.1, the ability to reduce experimental burden by multivariate data analysis with only minor loss in information and predictive ability was evaluated.

The second PLS model was created in accordance with the PLS model

described in 3.2., but with a smaller set of powder and roller compacted material descriptors as input parameters. The input parameters were chosen based on the results of the PCA model discussed under 3.1. To obtain an objective selection of descriptors, availability and simplicity were preferred compared to quality of the characterization technique (e.g. ρb over CBD, HR over Ffc), which led to the reduced material descriptor set (excluded descriptors marked with * in Table 3). This descriptor set was used as input (X) for the here discussed PLS model shown in detail in supplementary material A7. As response for the PLS model, again the tablet tensile strength at 200 MPa compression pressure was used in accordance with the PLS model described under 3.2. The first two factors of the PLS model with reduced set of input descriptors explained already 85 % and predicted 72 % of the variance of the TTS data, with factor 1 explaining (72 %) and predicting (67 %) the majority of the variance in the TTS data. The model was created with 5 factors to minimize the root mean square error which had a value of 0.43 MPa for the calibration data set and 0.85 MPa for the cross validation of the PLS model with reduced input descriptors. Therefore, it was shown that the PLS model created with a reduced set of input parameters as indicated by the PCA under 3.1. had comparable ability to explain and predict the variance in the tablet tensile strength data with a comparable RMSE as the PLS model created with all initial 29 powder and material descriptors discussed in 3.2. Further, the key material attributes for tabletability in roller compaction and direct compression which were identified by second PLS model were in accordance with the results of the PLS model described under 3.2 (see supplementary material A7). For the external validation of the second PLS model, it was seen that the root mean square error of prediction for the first two factors was higher with 1.50 MPa than the RMSE of prediction of the external validation of the PLS model with all 29 material descriptors. Further, as shown in supplementary material A7, stronger deviations from predicted to reference TTS were seen especially for neat compound A powder and roller compacted material compared to the PLS model with all 29 material descriptors as the micronization of compound A here has stronger impact on the model as less other material descriptors are included in the model compared to the initial PLS model under 3.2. Therefore, the comparison of the PLS model with the initial 29 powder and roller compacted material descriptors and a PLS model created with a reduced set of input material descriptors as indicated by the PCA under 3.1 demonstrated that data based reduction of material descriptors as well as characterization methods and therefore reduction of experimental burden is generally possible by principal component analysis, which could be especially useful for formulation development where often only limited amounts of material are available. However, it has to be noted, that although PCA highlighted strong correlations between different descriptors as well as methods, in this study no method could have been excluded completely as discussed under 3.1 and that the reduction of experimental burden is possible only to a limited extent without loss in information. Thus, the systematic evaluation of potential reduction of experimental burden by PCA with a direct comparison of two PLS models with all initial descriptors as well as exclusion of descriptors via PCA analysis highlighted that although descriptors can be reduced to a certain extent by PCA, still a complex and large set on material characterization methods and descriptors is needed for a deeper understanding of the impact of material properties on processability. This was further underlined by the fact, that the prediction of tablet tensile strength of unknown materials was much more precise when all 29 material descriptors were included.

4. Conclusion

In this study, principal component analysis and partial least squares regression were applied to a large data set with substantial variability by inclusion of 44 materials with varying drug loads in two process routes, direct compression and roller compaction, as well as 119 material descriptors to identify key material attributes for mechanical tablet properties in roller compaction and direct compression. By that, PCA and PLS proved to be powerful tools to elucidate the impact of powder descriptors, process route, drug load as well as raw API and excipient properties on mechanical tablet properties and could enable a more databased selection of materials for formulation development in the future. PCA of an extensive number of material descriptors enabled a deeper knowledge of relations between powder descriptors and revealed many correlations and co-dependencies of the different material descriptors as well as their characterization methods. To systematically evaluate if the identified correlations of material descriptors by PCA could enable the reduction of analyses and thus cost in formulation development, two subsequent PLS models were created, one with the initial set of material descriptors and one with less powder and roller compacted material descriptors as indicated by the PCA. By that, it was demonstrated that the data based reduction of time and material resources can be enabled by multivariate data analysis to a certain extent. However, for a substantial understanding of the impact of powder characteristics on tabletability and precise prediction of mechanical tablet properties complex and extensive material characterization methods are needed.

The PLS regression model created with 34 materials and 119 material descriptors revealed underlying correlations between powder and mechanical tablet properties represented by the tablets' tensile strengths and was able to systematically identify key powder descriptors, which were indicators of density such as the packing fraction as well as particle size descriptors but also descriptors of adhesive and cohesive forces such as the wall friction angle, total surface energy and work of cohesion, which were identified as key material attributes for tabletability in roller compaction and direct compression for the first time by PLS. With this model it was achieved to predict the tabletability trends in 10 materials unknown to the model successfully.

By including initial powders as well as the respective roller compacted materials, the PLS model further indicated that the initial material descriptors often outweigh the different process routes of direct compression and roller compaction regarding their effect on tabletability. On the one hand, the intentional use of only a limited number of different materials in varying API-excipient combinations as well as different drug loads allowed for a more detailed assessment of the influence of the respective material on blend properties. On the other hand, the small number of different materials also presented clear limitations of the PLS model as a broader variation in materials is needed to improve and strengthen the model's robustness and general validity. Although this study does not claim to have addressed the entirety of all relevant material descriptors for mechanical tablet properties, by being able to explain 73 % of the data variance with the established PLS model and to predict tabletability trends of materials unknown to the model, an initial basis for tabletability prediction for direct compression and roller compaction was created, which can be built upon in the future.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Lena Mareczek: Conceptualization, Investigation, Formal analysis, Writing – original draft. **Carolin Riehl:** Conceptualization, Supervision, Writing – review & editing. **Meike Harms:** Software, Writing – review & editing. **Stephan Reichl:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Data availability

Further information on the research data can be found in the supplementary material

Acknowledgements

The authors gratefully acknowledge Axel Becker, Anke Marx and Dirk Wandschneider for providing analytical support in conducting the inverse gas chromatography measurements as well as the scientific discussion on multivariate data analysis.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejps.2024.106836.

References

Berkenkemper, S., Klinken, S., Kleinebudde, P., 2023. Multivariate data analysis to evaluate commonly used compression descriptors. Int. J. Pharm. 637, 122890 https://doi.org/10.1016/j.ijpharm.2023.122890.

Boersen, N., Carvajal, M.T., Morris, K.R., Peck, G.E., Boersen, N., Carvajal, M.T., Morris, K.R., Peck, G.E., 2015. The influence of API concentration on the roller compaction process: modeling and prediction of the post compacted ribbon, granule and tablet properties using multivariate data analysis. Drug Dev. Ind. Pharm. 41, 1470–1478. https://doi.org/10.3109/03639045.2014.958754.

Brunauer, S., Emmett, P.H., Teller, E., 1938. Adsorption of Gases in Multimolecular Layers. J. Am. Chem. Soc. 60, 309–319. https://doi.org/10.1021/ja01269a023.

Della Volpe, C., Siboni, S., 2000. Acid – base surface free energies of solids and the definition of scales in the Good – van Oss – Chaudhury theory. J. Adhes. Sci. Technol. 14, 235–272. https://doi.org/10.1163/156856100742546.

Della Volpe, C., Siboni, S., 1997. Some reflections on acid-base solid surface free energy theories. J. Colloid Interface Sci. 195, 121–136. https://doi.org/10.1006/jcis.1997.5124.

Dhondt, J., Bertels, J., Kumar, A., Hauwermeiren, D.Van, Ryckaert, A., Van Snick, B., Klingeleers, D., Vervaet, C., De Beer, T., 2022a. A multivariate formulation and process development platform for direct compression. Int. J. Pharm. 623, 121962 https://doi.org/10.1016/j.ijpharm.2022.121962.

Dhondt, J., Eeckhout, Y., Bertels, J., Kumar, A., Van Snick, B., Klingeleers, D., Vervaet, C., De Beer, T., 2022b. A multivariate methodology for material sparing characterization and blend design in drug product development. Int. J. Pharm., 121801 https://doi.org/10.1016/j.ijpharm.2022.121801.

Dong, S., Brendlé, M., Donnet, J.B., 1989. Study of solid surface polarity by inverse gas chromatography at infinite dilution. Chromatographia 28, 469–472. https://doi.org/10.1007/BF02261062.

Dorris, G.M., Gray, D.G., 1980. Adsorption of n-alkanes at zero surface coverage on cellulose paper and wood fibers. J. Colloid Interface Sci. 77, 353–362. https://doi.org/10.1016/0021-9797(80)90304-5.

Escotet-Espinoza, M.S., Vadodaria, S., Ravendra, S., Muzzio, F.J., Ierapetritou, M.G., 2018. Modeling the effects of material properties on tablet compaction: a building block for controlling both batch and continuous pharmaceutical manufacturing processes. Int. J. Pharm. 543, 274–287. https://doi.org/10.1016/j.ijpharm.2018.03.036.

European Directorate for the Quality of Medicines & Health Care, 2023. 5.21. Chemometric Methods. Eur. Pharmacop. 11.1.

European Directorate for the Quality of Medicines & Health Care, 2022. 2.9.36. Powder Flow. Eur. Pharmacopoeia 11.0.

Faulhammer, E., Fink, M., Llusa, M., Lawrence, S.M., Biserni, S., Calzolari, V., Khinast, J. G., 2014. Low-dose capsule filling of inhalation products : critical material attributes and process parameters. Int. J. Pharm. 473, 617–626. https://doi.org/10.1016/j.ijpharm.2014.07.050.

Fell, J.T., Newton, J.M., 1970. Determination of tablet strength by the diametral-compression test. J. Pharm. Sci. 59, 688–691. https://doi.org/10.1002/jps.2600590523.

Ferreira, A.P., Tobyn, M., 2015. Multivariate analysis in the pharmaceutical industry : enabling process understanding and improvement in the PAT and QbD era Multivariate analysis in the pharmaceutical industry : enabling process understanding and improvement in the PAT and QbD era. Pharm. Dev. Technol. 20, 513–527. https://doi.org/10.3109/10837450.2014.898656.

Fowkes, F.M., 1964. Attractive Forces at Interfaces. Ind. Eng. Chem. 56, 40–52. https://doi.org/10.1021/ie50660a008.

Freeman, R., 2007. Measuring the flow properties of consolidated, conditioned and aerated powders — A comparative study using a powder rheometer and a rotational shear cell. Powder Technol 174, 25–33. https://doi.org/10.1016/j.powtec.2006.10.016.

Gamble, J.F., Leane, M., Olusanmi, D., Tobyn, M., Supuk, E., 2012. Surface energy analysis as a tool to probe the surface energy characteristics of micronized materials — A comparison with inverse gas chromatography. Int. J. Pharm. 422, 238–244. https://doi.org/10.1016/j.ijpharm.2011.11.002.

International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2020. ICH guideline Q12 on technical and regulatory considerations for pharmaceutical product lifecycle management - Step 5.

International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2011. ICH Quality Implementation Working Group Points to Consider (R2) ICH-Endorsed Guide for ICH Q8/Q9/Q10 Implementation.

Jenike, A.W., 1964. Storage and flow of solids. Bull. No.123 Utah Eng. Exp. Stn. 53, 1–198. https://doi.org/10.2172/5240257.

Leane, M., Pitt, K., Reynolds, G., Anwar, J., Charlton, S., Crean, A., Creekmore, R., Davies, C., DeBeer, T., De-Matas, M., Djemai, A., Douroumis, D., Gaisford, S., Gamble, J., Stone, E.H., Kavanagh, A., Khimyak, Y., Kleinebudde, P., Moreton, C., Paudel, A., Storey, R., Toschkoff, G., Vyas, K., 2015. A proposal for a drug product Manufacturing Classification System (MCS) for oral solid dosage forms. Pharm. Dev. Technol. 20, 12–21. https://doi.org/10.3109/10837450.2014.954728.

Mareczek, L., Riehl, C., Harms, M., Reichl, S., 2022. Understanding the multidimensional effects of polymorphism, particle size and processing for D-mannitol powders. Pharmaceutics 14, 2128. https://doi.org/10.3390/pharmaceutics14102128.

Papirer, E., Balard, H., Vidal, A., 1988. Inverse gas chromatography: a valuable method for the surface characterization of fillers for polymers (glass fibres and silicas). Eur. Polyme J. 24, 783–790. https://doi.org/10.1016/0014-3057(88)90015-8.

Peeters, M., Barrera Jiménez, A.A., Matsunami, K., Ghijs, M., dos Santos Schultz, E., Roudgar, M., Vigh, T., Stauffer, F., Nopens, I., De Beer, T., 2024. Analysis of the effect of formulation properties and process parameters on granule formation in twin-screw wet granulation. Int. J. Pharm. 650, 123671 https://doi.org/10.1016/j.ijpharm.2023.123671.

Reutenauer, S., 2002. Reproducibility of the dispersive component of surface energy measured by inverse gas chromatography - part i, surface measurement systems ltd. Technical Note 801.

Ryckaert, A., Van Hauwermeiren, D., Dhondt, J., De Man, A., Funke, A., Djuric, D., Vervaet, C., Nopens, I., De Beer, T., 2021. TPLS as predictive platform for twin-screw wet granulation process and formulation development. Int. J. Pharm. 605, 120785 https://doi.org/10.1016/j.ijpharm.2021.120785.

Soh, J.L.P., Wang, F., Boersen, N., Pinal, R., Garnet, E., Carvajal, M.T., Cheney, J., Valthorsson, H., Pazdan, J., Soh, J.L.P., Wang, F., Boersen, N., Pinal, R., Peck, E., Carvajal, M.T., Cheney, J., Valthorsson, H., Utility, J.P., Soh, J.L.P., Wang, F., Boersen, N., Pinal, R., Peck, G.E., Carvajal, M.T., 2008. Utility of Multivariate Analysis in Modeling the Effects of Raw Material Properties and Operating Parameters on Granule and Ribbon Properties Prepared in Roller Compaction Utility of Multivariate Analysis in Modeling the Effects of Raw Material Properties. Drug Dev. Ind. Pharm. 34, 1022–1035. https://doi.org/10.1080/03639040801925990.

Souihi, N., Dumarey, M., Wikström, H., Tajarobi, P., Fransson, M., Svensson, O., Josefson, M., Trygg, J., 2013. A quality by design approach to investigate the effect of mannitol and dicalcium phosphate qualities on roll compaction. Int. J. Pharm. 447, 47–61. https://doi.org/10.1016/j.ijpharm.2013.02.036.

Thielmann, F., Burnett, D.J., Heng, J.Y.Y., 2007. Determination of the Surface Energy Distributions of Different Processed Lactose Determination of the Surface Energy Distributions. Drug Dev. Ind. Pharm. 33, 1240–1253. https://doi.org/10.1080/03639040701378035.

Thoorens, G., Krier, F., Rozet, E., Carlin, B., Evrard, B., 2015. Understanding the impact of microcrystalline cellulose physicochemical properties on tabletability. Int. J. Pharm. 490, 47–54. https://doi.org/10.1016/j.ijpharm.2015.05.026.

United States Pharmacopeia, 2016. <1039>Chemometrics. United States Pharmacop. Natl. Formul. https://doi.org/10.31003/USPNF_M2345_02_01.

Van Snick, B., Dhondt, J., Pandelaere, K., Bertels, J., Mertens, R., Klingeleers, D., Di Pretoro, G., Remon, J.P., Vervaet, C., De Beer, T., Vanhoorne, V., 2018. A multivariate raw material property database to facilitate drug product development and enable in-silico design of pharmaceutical dry powder processes. Int. J. Pharm. 549, 415–435. https://doi.org/10.1016/j.ijpharm.2018.08.014.

Wold, H., 1974. Causal flows with latent variables: partings of the ways in the light of NIPALS modelling. Eur. Econ. Rev. 5, 67–86.

Wold, S., Sjostrom, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58, 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1.