Pharmaceutical Research, Vol. 26, No. 11, November 2009 (© 2009)
DOI: 10.1007/s11095-009-9954-7

Research Paper

A Structural Investigation into the Compaction Behavior of Pharmaceutical Composites Using Powder X-ray Diffraction and Total Scattering Analysis

Michael D. Moore,[1] Alison M. Steinbach,[1] Ira S. Buckner,[1] and Peter L. D. Wildfong[1,2]

Received April 23, 2009; accepted August 10, 2009; published online August 28, 2009

Purpose. To use advanced powder X-ray diffraction (PXRD) to characterize the structure of anhydrous theophylline following compaction, alone, and as part of a binary mixture with either α-lactose monohydrate or microcrystalline cellulose.
Materials and Methods. Compacts formed from (1) pure theophylline and (2) each type of binary mixture were analyzed intact using PXRD. A novel mathematical technique was used to accurately separate multi-component diffraction patterns. The pair distribution function (PDF) of isolated theophylline diffraction data was employed to assess structural differences induced by consolidation and evaluated by principal components analysis (PCA).
Results. Changes induced in PXRD patterns by increasing compaction pressure were amplified by the PDF. Simulated data suggest PDF dampening is attributable to molecular deviations from average crystalline position. Samples compacted at different pressures were identified and differentiated using PCA. Samples compacted at common pressures exhibited similar inter-atomic correlations, where excipient concentration factored in the analyses involving lactose.
Conclusions. Practical real-space structural analysis of PXRD data by PDF was accomplished for intact, compacted crystalline drug with and without excipient. PCA was used to compare multiple PDFs and successfully differentiated pattern changes consistent with compaction-induced disordering of theophylline as a single component and in the presence of another material.

KEY WORDS: composite analysis; pair distribution function; pharmaceutical compaction; PXRD.

INTRODUCTION

The US Food and Drug Administration's Critical Path Initiative to New Medical Products maintains the need for industry to continue investigation and development of analytical methods capable of fundamentally characterizing pharmaceutical systems. The availability of more effective analysis techniques will facilitate product and process understanding, and reduce drug product failures (1). Challenges encountered during drug product development are often the result of unpredicted changes to the physicochemical and mechanical properties of materials used in manufacturing. Characterization during small-scale research should predict performance modifications that may occur at commercial scale. This goal is contingent upon accurate detection and quantification of specific changes in material structure that can be linked with product performance. Some such changes (*i.e.* dissolution rate, friability, *etc.*) can be traced to very subtle structure modifications, where the largest limitation in detection and quantification is inadequate analytical sensitivity, particularly in the context of complex mixtures of materials. At present, research in this area is heavily focused on active pharmaceutical ingredients (API), with substantially fewer reports concerning responses of excipients or composites to processing. As the pharmaceutical industry moves forward, it is imperative that the synergistic functions of both API and excipients be considered as the basis for claiming product understanding at the level of specifically engineered delivery platforms.

Powder X-ray diffraction (PXRD) is universally applied for detecting phase changes and assessing material structure and order. Industrial applications of PXRD rest in the ability to analyze materials from all stages of powder processing, without prior destructive sample preparation. Traditional uses of this technique have focused on analysis of prominent, high-intensity

[1] Duquesne University, Pittsburgh, Pennsylvania, USA.
[2] To whom correspondence should be addressed. (e-mail: wildfongp@duq.edu)

NOTATION

ABBREVIATIONS: MCC, Microcrystalline cellulose; PCA, Principal components analysis; PDF, Pair distribution function; PXRD, Powder X-ray diffraction

SYMBOLS: d_L, Scalar scaling constant (noise); d_R, Scalar scaling constant (reference); f, Atomic scattering factor; $\langle f \rangle$, Average atomic scattering factor; G, Pair distribution function; **I**, Identity matrix; **L**, Noise matrix; ρ, Local number density; ρ_0, Average number density; Q, Scattering vector magnitude; r, Inter-atomic radial distance; **R**, Matrix of reference diffraction pattern(s); S, Structure function; **T**, Scaling matrix; **X**, Matrix of multi-component diffraction patterns; $\mathbf{X}_{fin}$, Matrix of isolated diffraction attributable to the component of interest; δ, Delta function; ∑, Pre-scaling matrix; ', Transpose

Bragg diffraction peaks that are characteristic of a given crystalline phase. In addition to long-range order, the powder diffraction pattern contains information pertaining to intermediate and short-range order. Specifically, diffuse scatter, which occurs between and superimposed on Bragg diffraction peaks, offers additional information pertaining to the short- and intermediate-range structure of materials, when treated appropriately.

In the present work, the compaction behavior of anhydrous theophylline was investigated. Binary composite compacts were also prepared and analyzed, using either a common tablet filler (α-lactose monohydrate) or a common tablet compression aid (microcrystalline cellulose (MCC)) in combination with anhydrous theophylline. PXRD, in combination with total scattering computational methods, was used to assess the effects of powder compaction without reversion of compacts to powders. Furthermore, the use of PCA as an alternative to the traditional weighted agreement factor in comparing PDF transformed diffraction data was explored.

MATERIALS

Binary mixtures comprised of anhydrous theophylline (Lot No. 92577, Knoll AG, Ludwigshafen, Germany) and either crystalline α-lactose monohydrate, referred to as lactose for the remaining paper, (Lot No. 125090020, Acros Organics, Geel, Belgium) or microcrystalline cellulose, referred to as MCC for the remaining paper, (Avicel PH 200, Lot No. M427C, FMC BioPolymer, Mechanicsburgh, PA) were blended and compacted using an automatic single-station Carver Press (Carver, Inc, Model 3887.1SDOA00, Wabash, IN) equipped with a 13 mm flat-faced punch. The approximate median particle size of theophylline was 90 μm. A circumscribed central composite experimental design (2) was used for sample preparation, in which both excipient concentration (0–75% *w/w*) and compaction pressure (67–503 MPa) were varied. Consolidation at each pressure resulted in composite solid fractions ranging from 0.85–0.98. The central point was repeated five times to build precision into the design.

METHODS

Powder X-ray Diffraction (PXRD)

PXRD data were collected using an X'Pert Pro MPD system (PANalytical B.V., Almelo, the Netherlands) equipped with a copper anode (λ=1.5406 Å), an auxiliary elliptical mirror, and X'Celerator™ detector. The operational voltage and amperage were set to 45.0 kV and 40.0 mA, respectively, and diffraction patterns were acquired using an irradiation time of 101.42 s per step and an angular step size of 0.017° 2θ over a range of 2–100° 2θ. Data were collected with the instrument set in transmission geometry, using intact compact samples, sandwiched between two layers of X-ray transparent kapton film and placed on a spinning vertical sample stage (16 rpm).

Pattern Separation

PXRD analysis of binary compacts produced diffraction patterns containing information from both components. To study the effects of powder compaction on a specific constituent, a method for accurately separating diffraction events attributable to each individual component was required. The pattern subtraction technique (3), commonly reported for this purpose, often leaves residual artifact in separated patterns attributable to diffraction by component(s) that are not of interest, or negative peak intensities as a result of overcompensation. An alternative approach for isolating single constituent diffraction patterns from multi-component data has been recently reported in the literature (4).

Generalized Least Squares (GLS) pre-processing has been reported in the near-infrared spectroscopy literature as a multiplicative orthogonalization technique used to make spectra "blind" to interferants while retaining sensitivity to the main analyte (5). Signal pre-processing using this covariance-weighted technique reduces the dimensionality of final models by including prior knowledge. The scaling matrix, $\boldsymbol{T}$, is calculated as follows:

$$\Sigma_L = d_L^2 LL' + I \tag{1}$$

$$\boldsymbol{T} = \mathrm{inv}\left(\Sigma_{\mathrm{L}}^{1/2}\right) \tag{2}$$

where d is a scalar covariance-scaling constant, $\boldsymbol{L}$ is an $m \times n$ noise matrix, $\boldsymbol{L'}$ is the transpose of $\boldsymbol{L}$, and $\boldsymbol{I}$ is an $n \times n$ identity matrix. In the case of PXRD, the matrix $\boldsymbol{L}$ would be comprised of diffraction information from material(s) that are not of interest to a given evaluation. The final scaling of the multi-component diffraction patterns, $\boldsymbol{X}$, is given by

$$\boldsymbol{X}_{FIN} = \boldsymbol{XT} \tag{3}$$

where $\boldsymbol{T}$ multiplicatively suppresses diffraction in $\boldsymbol{X}$ displaying covariance with the noise matrix. A problem is encountered with this particular method when components that are not of interest (noise) diffract at the same angle (2θ) as the component of interest. This issue results in the suppression of diffraction events from the component being isolated. The method was, therefore, modified to account for covariance between the "noise" component(s) and the component of interest by creating a second scaling matrix:

$$\Sigma_R = \boldsymbol{d}_R^2 \boldsymbol{RR}' + \boldsymbol{I} \tag{4}$$

where $\boldsymbol{R}$ is an $m \times n$ matrix of the reference material (*i.e.* a pure component diffraction pattern of the constituent of interest) and the remaining variables are defined above. The scaling matrix, Σ_{R}, was used to suppress angular variables in the noise matrix, which are covariate with the component of interest as follows:

$$\Sigma_{\mathrm{FIN}} = \Sigma_L \bullet \boldsymbol{inv}(\Sigma_R) \tag{5}$$

$$\boldsymbol{T} = \boldsymbol{inv}\left(\Sigma_{\mathrm{FIN}}^{1/2}\right) \tag{6}$$

where $\boldsymbol{T}$ is substituted back into Eq. 3. By compensating for angular variables (2θ) of the noise matrix that are covariate

with those of the constituent of interest, a more accurate separation is obtained. All data manipulations in this study were performed using programs written in-house in the Matlab programming environment (v7.1, MathWorks, Natick, MA).

Pair Distribution Function (PDF)

The PDF is a total scattering method that exploits the Fourier relationship between X-ray diffraction intensity and the real-space arrangement of atoms, given appropriate data treatment. This method has received extensive attention in the inorganic literature (6–9); however, few pharmaceutical applications have been reported (10–12). In the present work, the PDF was calculated for (a) compacts formed from pure components and (b) mathematically isolated diffraction attributable to a single component, detected from PXRD patterns of binary compacts. This provided a sensitive method to assess the structure-specific compaction behavior of theophylline, individually as well as in a consolidated mixture. The PDF, *G(r)*, is defined as

$$G(r) = 4\pi r[\rho(r) - \rho_o] \tag{7}$$

where $\rho(r)$ and ρ_o are the local and average atomic number densities, respectively, and r is the inter-atomic separation distance (13). The PDF calculates the probability of finding atom pairs separated by a distance r, and is obtained by transformation of the reciprocal space structure function, $S(Q)$, according to

$$G(r) = \frac{2}{\pi}\int_0^{Q_{\max}} Q[S(Q) - 1]\sin(Qr)dQ \tag{8}$$

where $S(Q)$ is obtained from a diffraction experiment, and Q is the magnitude of the scattering vector (14). The term Q_{max} is the resolution of the diffraction experiment, which is dependent on the wavelength of radiation used and the maximum diffraction angle (°2θ) of data collection. Corrections were made to the intensity data to obtain a structure function normalized to the total-scattering cross-sectional area, consistent with the assumptions of Egami and Billinge (15).

For a crystalline material, a theoretical PDF may be calculated from a solved crystal structure. The mathematical corrections applied to a theoretical PDF, including compensations for limited Q-resolution and broadening of probability peaks as a result of thermal motion are detailed elsewhere (15,16). Theoretical modeling of the PDF, $G(r)_c$, for a crystalline material is achieved by

$$G(r)_c = \frac{1}{r}\sum_m\sum_n \frac{f_m f_n}{\langle f\rangle^2}\delta(r - r_{m,n}) - 4\pi r\rho_o \tag{9}$$

where f_m and f_n are the atomic form factors for the individual atoms, $\langle f\rangle$ is the mean atomic form factor for the structure, and $r_{m,n}$ is the separation distance between atoms m and n (from the crystal structure), where the delta function is assessed out to a user-defined radial distance $r_{\max}$.

Reviews of the PDF, both experimental and theoretical, are found elsewhere (13,15). All intensity corrections and PDF calculations were performed using software developed in-house in the Matlab programming environment based on published equations. Theoretical PDF patterns for anhydrous theophylline were calculated using the solved Cambridge Structural Database (CSD) crystal structure, (refcode: *BAPLOT01*).

Principal Components Analysis (PCA)

PCA is a multivariate, bilinear decomposition technique used to identify the orthogonal basis set, which explains the maximum amount of variance in a data matrix, $\boldsymbol{X}$, where $\boldsymbol{X}$ is n observations of an m-dimensional vector. PCA decomposes data into loadings and scores, where loadings represent the weight of each variable for a given principal component (PC), and scores represent the weight of the PC for each sample. Vectors comprising a data matrix, $\boldsymbol{X}$, possessing comparable score values for a given PC are similar; dissimilar score values indicate differences between vector features. Reviews of principal components analysis may be found elsewhere in the literature (17,18). PCA models were calculated in the Matlab programming environment using the PLS_Toolbox (v3.0, Eigenvector Research, Manson, WA).

Traditional comparison of PDF transforms have involved calculation of a weighted agreement factor, R_{wp}, between the properly scaled sample PDF and a reference, albeit a theoretically calculated or empirically derived pattern (15). In this study, PCA was used to assess variations among PDF patterns as an alternative to the weighted agreement factor commonly employed for this purpose. This method may be preferable to the R_{wp} calculation when there are a large number of PDFs to compare, as in this study. PCA is also applicable when a reference pattern is not available or comparison to a reference pattern is inappropriate for a given analysis.

RESULTS

Pure Theophylline Compaction

In Fig. 1, the PXRD patterns (average of 3 samples) for theophylline compacted at 67 MPa and 503 MPa are shown as blue and red solid lines, respectively. The original patterns were normalized and were corrected for absorption, Laue monotonic scattering, Compton scattering, *etc.* (15) to remove physical artifacts (*i.e.,* sample thickness and solid fraction). The red and blue dashed lines in Fig. 1 respectively represent one standard deviation (n=3) for diffraction from theophylline compacts prepared at 67 MPa and 503 MPa. The Bragg peak positions of samples compacted at higher pressures occur at the same diffraction angle (°2θ) relative to the sample compacted at 67 MPa, indicating low probability of residual uniform strain. Moreover, additional PXRD peaks, which would be indicative of a polymorphic conversion, are not observed. One noticeable difference between the patterns, however, is increasing diffuse scatter over the 23–30 °2θ range for samples compacted at higher pressures.

Fig. 2 (solid lines) shows the average PDF pattern (n=3) for each of the corresponding samples in Fig. 1. The dashed lines represent one standard deviation as calculated from the three replicates. At lower radial distances in the PDF (r<30 Å), peak positions and intensities are highly correlated,

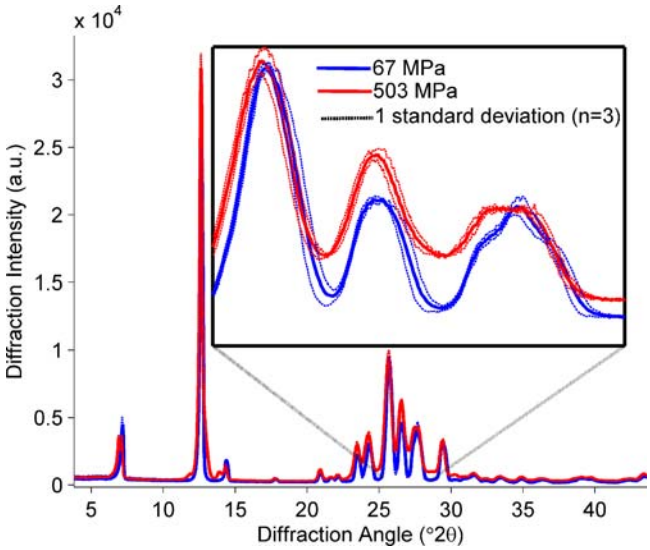

Fig. 1. Mean PXRD patterns (average of n=3) for anhydrous theophylline compacted at 67 MPa (*blue solid line*) and 503 MPa (*red solid line*). The *red* and *blue dashed lines* correspond to 1 standard deviation from the mean of samples compacted at 67 MPa and 503 MPa, respectively.

irrespective of compaction pressure. In contrast, the Fig. 2 inset, which focuses on r>40 Å, shows a dampening of the probability peaks, as well as peak shifts, which become significant with increased compaction pressure.

In an attempt to arrive at a possible materials-based explanation for the alterations observed in the PDF of anhydrous theophylline upon compaction at high pressures, simulations based on reverse Monte Carlo (6,19) refinement were performed. Briefly, as unit cell translation and expansion were calculated from the crystal structure solution, spatial permutations were applied to randomly selected atomic coordinates. These permutations included linear deviations from periodicity, as well as alterations to axial orientation (*e.g.,* the molecule may be rotated 120° with respect to the *c*-axis). The fully translated model, calculated to a defined r_{max}, therefore contained a certain percentage of atomic coordinates that varied from their original positions defined by the average structure. PDF patterns for computationally altered structures were subsequently calculated, and the data were mathematically compared with experimentally derived PDF patterns using an agreement factor.

When simulating atomic displacements from average structure, it is important to define the proper constraints to yield physically-meaningful simulations. One of the constraints incorporated in the models calculated for this study was the rigid body assumption; *i.e*, the molecule was held rigid such that permutations to atomic coordinates could not result in changes to intra-molecular bond distances, angles, or conformations. An additional constraint prevented atomic overlap of coordinates resulting from simulated displacements. The final constraint allowed only atomic coordinates residing at r values comprising long-range order (r>40 Å) to be randomly selected for permutation, as this region defines the marked differences between the PDFs presented in Fig. 2.

The experimentally derived PDF pattern for theophylline compacted at 67 MPa (blue line) and the PDF pattern calculated from the crystallographic structure of theophylline (black circles) are shown in Fig. 3a. Generally speaking, the difference plot indicates pattern similarity, especially at longer r values. This result is expected as the compaction pressure is insufficient to induce large deviations from the average crystallographic structure.

Recall from Fig. 2 that the PDF of the high pressure sample shows dampening and shifts in peak position relative to the lower pressure sample at r>40 Å. An acceptable fit between the PDF of the sample compacted at high pressure and the theoretical PDF will require alterations to the original calculated PDF to accommodate the probability peak aberrations. Fig. 3b shows the experimentally derived PDF for theophylline compacted at 503 MPa (red line), a refined PDF from the Monte Carlo simulation (black circles), and the difference between the two PDFs (green line, shown below Fig. 3b). The simulated PDF shows the same probability peak shifts and dampening as the experimental PDF of the sample compacted at a high pressure. The goodness of fit between the calculated PDF and the experimental PDF can be seen in the difference plot (green) shown below Fig. 3b. The simulated PDF was modeled by pseudo-randomly selecting molecules in the defined spherical volume (radius = 60 Å) and inducing a deviation from the average crystallographic structure. The number of molecular permutations was converted to a percentage of total molecules probed in the simulation. Therefore, the simulated PDF in Fig. 3b (black circles) differs from the calculated PDF in Fig. 3a (black circles) as a result of 0.928% of the molecules deviating from the defined average structure.

PDF of Mathematically Isolated Theophylline

Fig. 4a shows the PXRD pattern of a sample (blue circles) comprised of 25% *w*/*w* theophylline and 75% *w*/*w* MCC, compacted at 118 MPa. Diffraction data attributable

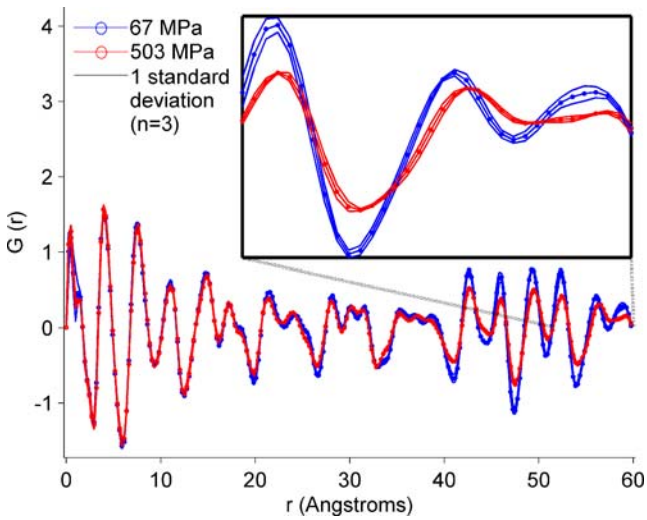

Fig. 2. Mean PDF patterns (average of n=3) of theophylline compacted at 67 MPa (*blue solid line*) and 503 MPa (*red solid line*). The *red* and *blue dashed lines* correspond to 1 standard deviation from the mean of samples compacted at 67 MPa and 503 MPa, respectively.

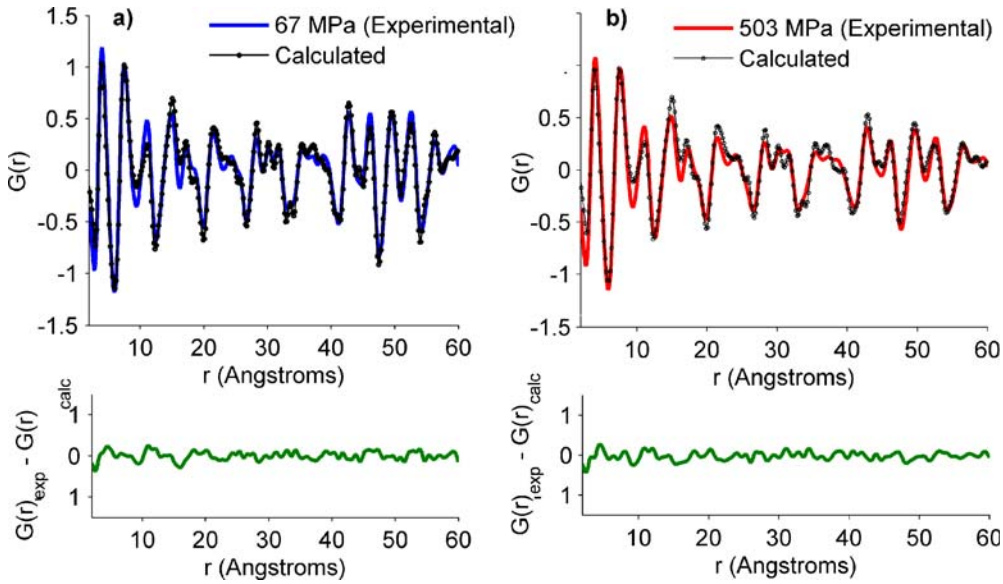

Fig. 3. **a** Experimental PDF pattern for theophylline compacted at 67 MPa (*black*), calculated PDF pattern from the crystal structure (*blue circles*), and the difference plot (*green*, *below*); **b** Experimental PDF pattern for theophylline compacted at 503 MPa (*red*), a simulated PDF pattern from reverse Monte Carlo refinement (*blue circles*), and the difference plot (*green*, *below*).

only to theophylline (red) are superimposed over the diffraction pattern of the aforementioned sample. Most of the diffuse scatter observable in the diffraction pattern of the original binary sample (blue circles) is absent when theophylline diffraction was isolated (red data). In Fig. 4b, the PDF transform of the mathematically isolated theophylline (red) is shown superimposed with the calculated theoretical PDF pattern for theophylline (black circles).

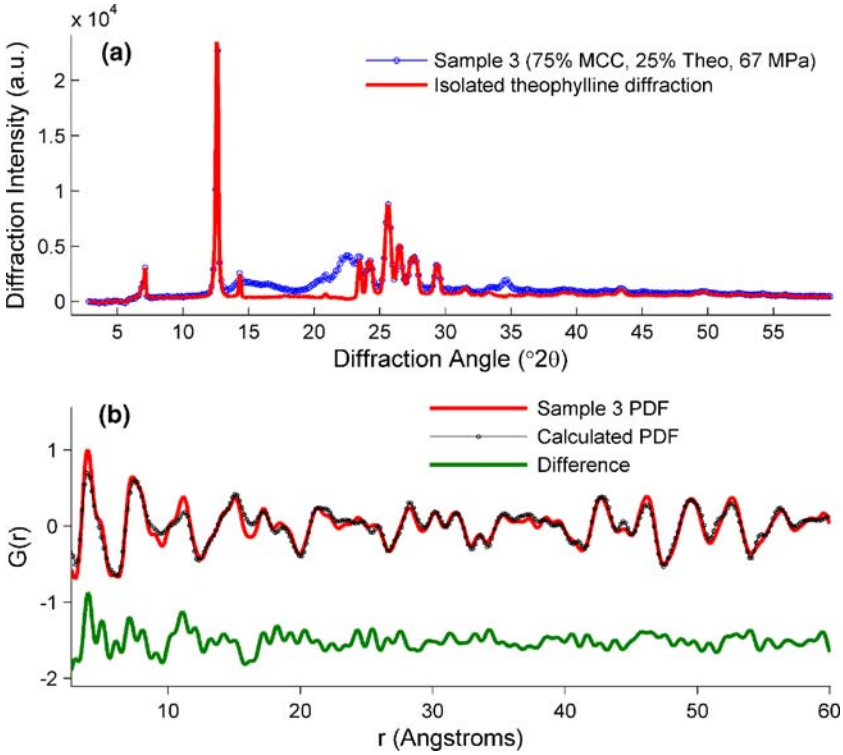

Fig. 4. **a** PXRD pattern (*zoomed*) of a compact containing 75% MCC and 25% theophylline (sample3) compressed at 67 MPa (*blue circles*) and the isolated diffraction attributable to theophylline (*red*); **b** The calculated pair distrubtion function of anhydrous theophylline (crystal structure ref: BAPLOT01) (*black circles*), the pair distribution function of isolated theophylline from sample 3 (*red*), and the difference between the calculated and experimental (*green*).

Fig. 5a shows the diffraction pattern of a sample (blue circles) containing 25% *w*/*w* theophylline and 75% *w*/*w* lactose, compacted at 118 MPa. Mathematically isolated theophylline diffraction (red) is superimposed over the total diffraction data from this sample. Many of the Bragg diffraction peaks observed in the original sample pattern (blue circles), which are strictly attributable to lactose, are absent from the isolated theophylline diffraction pattern. In Fig. 5b, the PDF transform of the isolated theophylline (red) is shown with the calculated PDF pattern for theophylline (black circles).

PCA of PDF Transformed Data

As introduced, PCA is a bilinear decomposition method, where principle components (PCs) are calculated in a manner to explain the maximum amount of variance such that PC_n and PC_{n+1} are orthogonal. Similar PDF patterns will have similar scores for a given PC, while dissimilar PDF patterns will have drastically different score values for the same PC. Examples employing PCA score cluster analysis to ordinary PXRD patterns are found in the pharmaceutical literature (20,21); however, the authors are unaware of the existence of previous applications of PCA cluster analysis to PDF data.

Three different PCA models were calculated, where each varied in the number/type of samples included. The first analysis incorporated PDF patterns for pure theophylline and theophylline mathematically isolated from binary mixtures with MCC. The resulting scores plot is shown in Fig. 6. First, in the dimension of the first PC, where 36.28% of the overall variance is explained, the score values for pure theophylline samples compacted at low pressures are drastically different than those for samples compacted at high pressures. The variance explained by the first PC, therefore, is weighed differently in PDF patterns for samples compacted at low pressures relative to samples compacted at high pressures. Second, PDF patterns for theophylline data mathematically isolated from binary samples containing MCC compacted at low pressures had similar score values for the first PC relative to those calculated for pure theophylline compacted at low pressures. This indicates that PDF patterns of samples compacted at low pressures were similar in the dimension of the first PC, regardless of MCC concentration. Additionally, all samples containing MCC compacted at higher pressures had score values similar to pure theophylline compacted at high pressures in the dimension of the first PC. Overall, it was observed that scores of PDF data from samples compacted at low pressures formed one general cluster, while those from samples compacted at higher pressures formed another cluster.

The second PCA performed incorporated PDF data obtained from theophylline mathematically isolated from binary mixtures containing lactose, as well as pure theophylline compacts. The PCA scores plot for these data is shown in

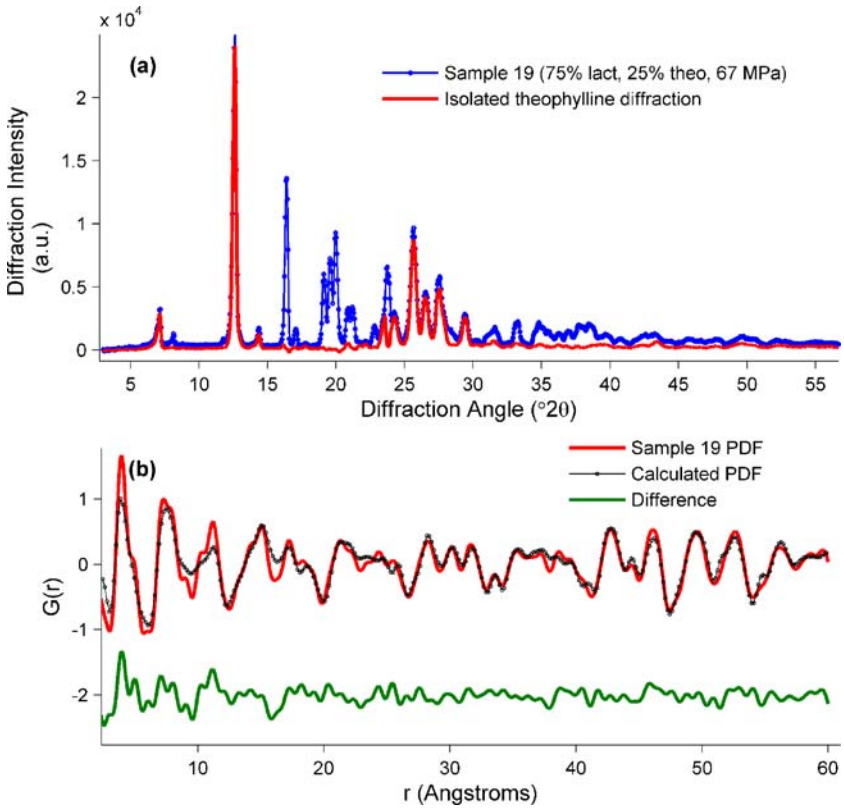

Fig. 5. a PXRD pattern of a compact containing 75% lactose and 25% theophylline (sample 19) compressed at 67 MPa (*blue circles*) and the isolated diffraction attributable to theophylline (*red*); **b** The calculated PDF pattern for anhydrous theophylline (crystal structure ref: *BAPLOT01*) (*black circles*), the PDF pattern for isolated theophylline from sample 19 (*red*), and the difference between the calculated and experimental (*green*).

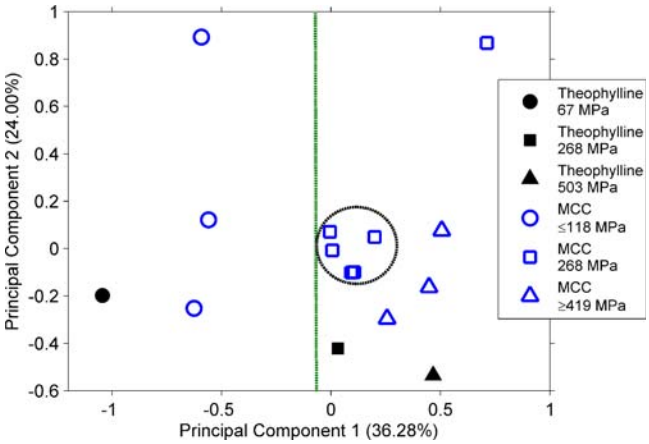

Fig. 6. PCA scores plot corresponding to the first and second principal components for analysis of PDF patterns from pure theophylline samples (*black*, *closed symbols*) and PDF patterns from theophylline mathematically isolated from binary compacts containing MCC (*blue*, *open symbols*). *Black dashed-line circles* surround the replicate samples.

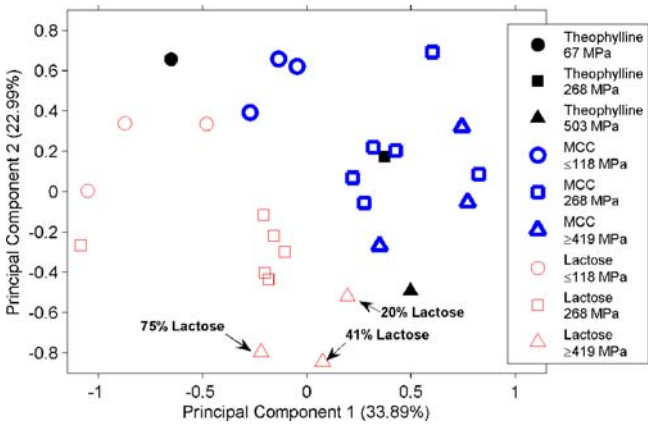

Fig. 8. PCA scores plot corresponding to the first and second principal components for analysis of all samples, including: PDF patterns for pure theophylline samples (*black*, *closed symbols*), PDF patterns for theophylline mathematically isolated from samples containing MCC (*blue*, *bold open symbols*), and PDF patterns for theophylline mathematically isolated from samples containing lactose (*red*, *un-emboldened open symbols*).

Fig. 7, where the dimension of the first PC explained 46.88% of the overall variance. Score values for PDF patterns from pure theophylline samples compacted at low pressures were considerably different relative to PDF patterns from pure theophylline compacted at high pressures. Samples compacted at low pressures tended to cluster together, regardless of lactose content; as did samples compacted at high pressures. A particularly interesting sample was the compact containing 75% *w*/*w* lactose compacted at 268 MPa, which clustered with the samples compacted at low pressures. These data indicate that the PDF pattern resulting from the isolated theophylline diffraction data of this sample is more similar to PDF data for samples compacted at low pressures than samples compacted at high pressures.

The final PCA performed incorporated data from PDF patterns obtained from all samples in the circumscribed central composite experimental design, resulting in the scores plot shown in Fig. 8. In Figs. 6 and 7, scores of samples

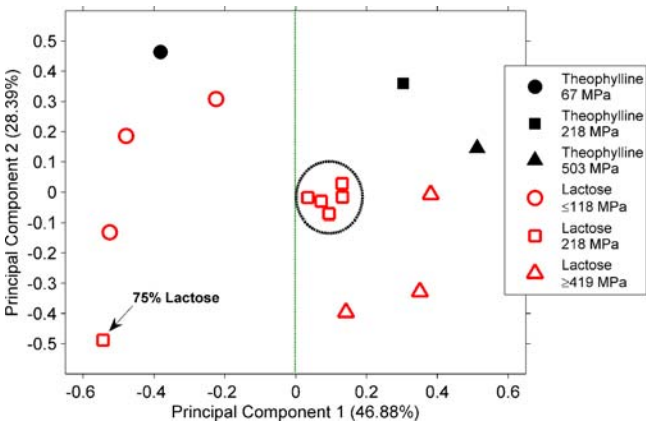

Fig. 7. PCA scores plot corresponding to the first and second principal components for analysis of PDF patterns for pure theophylline samples (*black*, *closed symbols*) and PDF patterns for theophylline mathematically isolated from binary compacts containing lactose (*red*, *open symbols*). *Black, dashed line circles* surround the replicate samples.

compacted at low pressures formed one cluster while samples compacted at high pressures formed a separate cluster in the dimension of the first PC. In Fig. 8, delineation between score values is not as obvious and merits further explanation. Beginning with the pure theophylline compacts (filled symbols), the score values in the dimension of the first PC, where 33.89% of the overall variance is explained, are drastically different for samples compacted at low pressures *versus* those compacted at high pressures. The scores in the dimension of the first PC of PDF data from samples containing MCC (blue, unfilled symbols) again tended to cluster with pure theophylline samples compacted at similar pressures, irrespective of MCC concentration. Contrary to the previous analyses of PDF data for samples containing lactose, a clear-cut boundary delineating clusters of similarly compacted samples in the dimension of the first PC is not observed. Interestingly, almost all samples containing lactose (red, unfilled symbols) resulted in PDF patterns having score values (in the first PC) that were most similar to all other samples compacted at low pressures. The score values for all lactose samples compacted at the intermediate pressure (268 MPa) clustered around samples compacted at low pressures. Further, a lactose-containing sample compacted at 419 MPa clustered in the same area.

DISCUSSION

Pure Theophylline Compaction

The PXRD patterns of pure theophylline compacted at different pressures exhibit an increase in diffuse scatter over the 23–30 °2θ range. With a median particle size of 90 μm, it is anticipated that the effects on pattern features and further mathematical manipulations will be minimal, as peak broadening and increased diffuse scatter tend to occur for particle sizes in the nanometer scale. Replicate samples (n=3) of pure theophylline compacted at 67 and 503 MPa were analyzed to estimate contributions of sample variability to the PXRD patterns. At both compaction pressures, the standard deviation intervals (Fig. 1, dashed lines) indicate that the increase

in diffuse scatter is not attributable to random sample variations, as the magnitude of the standard error is not sufficient to explain the scattering deviations between the PXRD patterns of the samples compacted at two different pressures (Fig. 1, solid lines).

The previously discussed pattern distortion can be associated with many different physical and structural anomalies (13,22); however, a discussion of each is beyond the scope of this manuscript. Rather, it is proposed that the introduction of disorder into a crystalline material may result in the alterations observed in the average diffraction patterns (solid lines) in Fig. 1. The PDF method is relatively sensitive to instances where materials exhibit long-range order; however, significant structural distortions may also be present that would not be representative of the average crystallographic structure. These "crystallographically challenged materials" result in dampened PDF features as a function of increasing radial distances relative to their pure crystalline counterparts (6). It may, therefore, be suggested that the features observed as a function of increasing compaction pressure in Fig. 2 are the result of increasing deviations from the average crystallographic structure of anhydrous theophylline. Replicate samples at each compaction pressure were carried through the PDF transformation to assess the effect of sample variability on the resulting PDF patterns. As shown by the standard deviation intervals (Fig. 2, dashed lines), the probability peak dampening and shifting can not be adequately explained by sample-to-sample variability, as the magnitude is not sufficient to justify the pattern distortions. This is further supported by data in Fig. 3. Key features in the experimental PDF and the pattern generated by simulating distortions to 0.928% of the molecules from their defined average position (Fig. 3b) agree well with one another. This agreement further supports the observation that compaction-related structural distortions resulted in the types of probability peak changes observed in Fig. 2.

It is worthwhile to note that the PDF is a one-dimensional representation of a three-dimensional structure; therefore, multiple simulated models may result in the same one-dimensional PDF representation. Though the simulations in this study were properly constrained to prevent physically meaningless solutions, other random structure permutations may give an acceptable or better answer than the one yielded in Fig. 3b. The proposed model is presented only to show that the PDF differences associated with compaction are consistent with compaction-induced disorder.

PDF of Mathematically Isolated Theophylline

The significant disorder of MCC presents a potential problem with respect to separation of diffuse scatter specifically attributable to MCC from the diffuse scatter from theophylline introduced as a result of compaction. The green difference plot in Fig. 4b shows that theophylline diffraction data mathematically isolated from the theophylline:MCC mixtures agrees well with the theoretical PDF of theophylline. Given the low compaction pressure used to prepare the samples for Fig. 4b, significant changes to the anhydrous theophylline structure were not anticipated, and accurate mathematical isolation of theophylline diffraction was, therefore, expected.

Different issues with isolation of scattering from theophylline:lactose compacts occurs owing to the crystallinity of the excipient. Specifically, isolation of theophylline diffraction from regions in the PXRD pattern where both components have 2θ peak overlap may result in multiple anomalies with the isolated pattern, including inaccurate peak intensity values, peak shape distortions, and/or the disappearance of theophylline peaks altogether. The green difference plot in Fig. 5b shoes that PDF-transformed theophylline diffraction data mathematically isolated from the theophylline:lactose mixtures agrees well with the theoretical PDF of theophylline. Again, considering the low pressure used to prepare compacts for Fig. 5b, significant structural changes to anhydrous theophylline were not anticipated.

PCA of PDF Transformed Data

The circumscribed central composite design of experiments used in the present work directly built sample variability into each model through replication of the center point ($n=5$), which corresponded to a compaction pressure of 268 MPa and an excipient concentration of 48% *w*/*w*. The variance in the PDF attributable to sample variation is therefore susceptible to explanation by PCA. In the dimension of the first PC, however, the replicate samples are observed to cluster together (dashed circles in Figs. 6 and 7), thereby precluding sample variability from acting as the sole contributor to the data segregation observed in each plot.

The data-segregation correlated to compaction pressure observed in Fig. 6 occurred regardless of the concentration of MCC present. As described above, long-range atomic correlation dampening ($r \geq 40$ Å) was concluded to be most likely attributable to structural differences arising from compaction. Ultimately, these data strongly support that the variance explained by the first PC is correlated to the probability peak dampening observed in PDF data (Fig. 2).

The same data segregation observed in the first analysis (Fig. 6) was also observed in the second PCA model (Fig. 7), which suggests that clustering corresponds with the same probability peak dampening explained above. Recall that the point representing the binary sample containing 75% *w*/*w* lactose compacted at 268 MPa clustered with the samples compacted at low pressures. Interpretation of this behavior suggests that the presence of this much lactose may provide some degree of "protection" against the theophylline structural changes interpretable from the PDF; however, this warrants further investigation.

Almost all samples containing lactose (Fig. 8, red, unfilled symbols) had PDF patterns with score values (in the dimension of the first PC) that were most similar to all other samples compacted at low pressures. The score values for all lactose samples compacted at 268 MPa clustered around samples compacted at low pressures, as did a lactose-containing sample compacted at 419 MPa. The clustering of this particular sample with others compacted at low pressures may not seem intuitive given that the remaining two lactose samples (containing less % *w*/*w* lactose) compacted at a high pressure did not. When sample composition is taken into consideration, however, a trend in the dimension of the first PC is observable. At a compaction pressure of 419 MPa, as the amount of lactose is decreased (Fig. 8, red triangles), the

sample score values became increasingly similar to the rest of the samples compacted at high pressure. Overall, every sample containing lactose (with the exception of the two compacted at 419 MPa having low % *w*/*w* concentrations) resulted in PDF patterns similar to all samples compacted at low pressures. These PDF data, therefore, did not exhibit the probability peak dampening displayed by PDF-transformed data of (a) pure theophylline samples compacted at high pressures and (b) PDF-transformed data of mathematically isolated theophylline from samples containing MCC compacted at high pressures. Hence, the samples containing >50% *w*/*w* lactose did not demonstrate any detectable structural modifications as a function of compaction pressure in the range investigated.

CONCLUSION

In this study, the effects of powder consolidation on solid structure were examined. Simulated data suggest that the compaction-related changes observed in real-space analyses (PDF) of theophylline are primarily attributable to molecular deviations from average crystalline order. The analysis of pharmaceutically relevant, intact consolidated systems was performed using PXRD, which afforded structure-sensitive capabilities without the introduction of artifacts arising from traditional sample preparation (*e.g.* subsequent grinding to revert to powder samples). The application of the PDF transform to an individual constituent, in the presence of an excipient, was successfully accomplished as a result of a novel, accurate mathematical pattern separation technique. Ultimately, this permitted practical real-space structural analysis of a single component in the presence of another material, thereby enabling the investigation of compound consolidation effects (*i.e.* the effect a material has on the compaction behavior of another). The use of PCA as an alternative to the weighted agreement factor for comparing multiple PDF patterns was successfully demonstrated. Relative to binary compacts formed with microcrystalline cellulose, structural changes to theophylline in the presence of α-lactose monohydrate were not observed (except in the two samples compacted with the least amount of lactose). This suggests the inclusion of lactose provides protection to theophylline from structural changes that would otherwise occur at high pressures. Further exploration of this particular conclusion, however, deserves additional attention and will be the focus of future experiments.

REFERENCES

1. U.S. Food and Drug Administration (FDA). Challenge and opportunity on the critical path to new medical products. http://www.fda.gov/oc/initiatives/criticalpath/.
2. Gabrielsson J, Lindberg N-O, Lundstedt T. Multivariate methods in pharmaceutical applications. J Chemom. 2002;16:141–60.
3. Phadnis NV, Cavatur RK, Suryanarayanan R. Identification of drugs in pharmaceutical dosage forms by X-ray powder diffractometry. J Pharm Biomed Anal. 1997;15:929–43.
4. Moore MD, Cogdill RP, Short SM, Hair CR, Wildfong PLD. The use of net analyte signal orthogonalization in the separation of multi-component diffraction patterns obtained from X-ray powder diffraction of intact compacts. J Pharm Biomed Anal. 2008;47:238–47.
5. Martens H, Hoy M, Wise BM, Bro R, Brockhoff PB. Prewhitening of data by covariance-weighted pre-processing. J Chemom. 2003;17:153–65.
6. Billinge SJL, Kanatzidis MG. Beyond crystallography: the study of disorder, nanocrystallinity and crystallographically challenged materials with pair distribution function. Chem Commun. 2004;7:749–60.
7. Jeong IK, Mohiuddin-Jacobs F, Petkov V, Billinge SJL. Local structure of In Ga As semiconductor alloys by high-energy synchrotron x-ray diffraction. Phys Rev B. 2001;63:1–9.
8. Petkov V. Atomic-scale structure of glasses using high-energy x-ray diffraction. J Am Ceram Soc. 2005;88:2528–31.
9. Petkov V, Qadir D, Shastri SD. Rapid structure determination of disordered materials: study of GeSe2 glass. Solid State Commun. 2004;129:239–43.
10. Bates S, Kelly RC, Ivanisevic I, Schields P, Zografi G, Newman AW. Assessment of defects and amorphous structure produced in raffinose pentahydrate upon dehydration. J Pharm Sci. 2007;96:1418–33.
11. Sheth AR, Bates S, Muller FX, Grant DJW. Local structure in amorphous phases of piroxicam from powder x-ray diffractometry. Cryst Growth Des. 2005;5:571–8.
12. Newman A, Engers D, Bates S, Ivanisevic I, Kelly RC, Zografi G. Characterization of amorphous API:Polymer mixtures using X-ray powder diffraction. J Pharm Sci. 2008;97:4840–56.
13. Warren BE. X-ray diffraction. New York: Dover; 1990.
14. Peterson PF, Bozin ES, Proffen T, Billinge SJL. Improved measures of quality for the atomic pair distribution function. J Appl Crystallogr. 2003;36:53–64.
15. Egamiand T, Billinge SJL. Underneath the Bragg Peaks. Structural analysis of complex materials. Oxford: Elsevier; 2003.
16. Jeong I-K, Proffen T, Mohiuddin-Jacobs F, Billinge SJL. Measuring correlated atomic motion using X-ray diffraction. J Phys Chem A. 1999;103:921–4.
17. Haalandand DM, Thomas EV. Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. Analytical Chemistry. 1988;60:1193–202.
18. Wold S, Esbensen K, Geladi P. Principal component analysis. Chemom Intell Lab Syst. 1987;2:37–52.
19. Toby BH, Egami T, Jorgensen JD, Subramanian MA. Observation of a local structural change at T_c for $Tl_2Ba_2CaCu_2O_8$ by pulsed neutron diffraction. Phys Rev Lett. 1990;64:2414–7.
20. Jorgensen AC, Miroshnyk I, Karjalainen M, Jouppila K, Siiria S, Antikainen O, *et al*. Multivariate data analysis as a fast tool in evaluation of solid state phenomena. J Pharm Sci. 2006;95:906–16.
21. Norrman M, Stahl K, Schluckebier G, Al-Karadaghi S. Characterization of insulin microcrystals using powder diffraction and multivariate data analysis. J Appl Crystallogr. 2006;39:391–400.
22. Masadeh AS, Bozin ES, Farrow CL, Paglia G, Juhas P, Karkamkar A, *et al*. Quantitative size-dependent structure and strain determination of CdSe nanoparticles using atomic pair distribution function analysis. Condens Matter. 2007;76:1–10.