




Article

Machine Learning Approach for Determining the Formation of β -Lactam Antibiotic Complexes with Cyclodextrins Using Multispectral Analysis

Mikołaj Mizera ¹, Kornelia Lewandowska ², Andrzej Miklaszewski ³ and Judyta Cielecka-Piontek ^{1,*}

¹ Department of Pharmacognosy, Faculty of Pharmacy, Poznań University of Medical Sciences, Święcickiego 4, 60-781 Poznań, Poland; mikolajmizera@gmail.com

² Institute of Molecular Physics, Polish Academy of Science, ul. Smoluchowskiego 17, 60-179 Poznań, Poland; kornelia.lewandowska@ifmpan.poznan.pl

³ Institute of Materials Science and Engineering, Poznan University of Technology, Pl. M.Sklodowskiej-Curie 5, 60-965 Poznań, Poland; andrzej.miklaszewski@put.poznan.pl

* Correspondence: jpiontek@ump.edu.pl; Tel.: +4861-854-6710

Received: 23 January 2019; Accepted: 17 February 2019; Published: 19 February 2019



Abstract: The problem of determining the formation of complexes of β -lactam antibiotics with cyclodextrins (CDs) and the interactions involved in this process were addressed by machine learning on multispectral images. Complexes of β -lactam antibiotics, including cefuroxime axetil, cefetamet pivoxil, and pivampicillin, as well as CDs, including α CD, β CD, γ CD, hydroxypropyl- α CD, methyl- β CD, hydroxypropyl- β CD, and hydroxypropyl- γ CD, were prepared in all combinations. Thermograms confirming the formation of cyclodextrin complexes were obtained using differential scanning calorimetry. Transmission Fourier-transform infrared (tFTIR) and complementary attenuated total reflectance FTIR (ATR) coupled with machine learning were techniques chosen as a nondestructive alternative. The machine learning algorithm was used to determine the formation of complexes in samples using solely their tFTIR and ATR spectra at the prediction stage. Parameterized method 7 (PM7) was used to support the analysis by molecular modeling of the complexes. The model developed through machine learning properly distinguished samples with formed complexes from noncomplexed samples with a cross-validation accuracy of 90.4%. Analysis of the contribution of spectral bands to the model indicated interactions of ester groups of β -lactam antibiotics with CDs, as well as some interactions of cephem ring in cefetamet pivoxil and penam moiety in pivampicillin. Molecular modeling with PM7 helped to explain experimental results and allowed to propose possible binding modes.

Keywords: cyclodextrin; β -lactam antibiotic; multispectral analysis; machine learning; molecular modeling

1. Introduction

The active pharmaceutical ingredients (APIs) systems with macromolecules such as cyclodextrins (CDs) are a promising way of enhancing physicochemical properties of well-established drugs. The analysis of such systems is the major challenge in the field of quality control due to a problem with the possible different stoichiometric ratio between reagents. An innovative approach for the analysis of the cyclodextrin systems may be based on theoretical methods involving machine learning and molecular modeling. The in silico techniques are particularly valuable when the included substances are APIs, which exhibit many beneficial physicochemical changes when they are in the cavity of a cyclodextrin carrier. Therefore, as a model, APIs in this work as esters of selected derivatives of β -lactam analogs

(prodrugs) was proposed. The benefits of combining prodrugs of β -lactam analogs with cyclodextrins include the possibility of modifying their solubility, increasing absorption, masking the specific taste, and even increasing bactericidal effectiveness [1,2].

Prodrugs are poorly or non-active chemical compounds which are biotransformed *in vivo* to their active form during any stage of absorption, distribution, and metabolism. As defined, the prodrug can be chemical bioprecursor undergoing transformation to an active drug or carrier-linked drug, where the active form is attached to a chemical group such as an ester, amide, or carbamate. Also, another active compound performing a pharmaceutical action can be linked, forming a mutual prodrug which simultaneously releases two active molecules after transformation [3]. Development of prodrugs is a way of addressing issues met during the early stages of API development such as low solubility [4], chemical instability [5], and bad taste [6]; also, changes of pharmacokinetic properties may be achieved, including permeation through the gastrointestinal tract [7], blood–brain penetration [8], modification of therapeutic activity time [9], and pharmacodynamic properties, including toxicity reduction of some antitumor drugs [10] and improvement of therapeutic index [11]. Development of prodrugs by linking with lipophilic moiety may have a negative impact on solubility compared to the parent molecule. As a result, lowered bioavailability creates secondary issues which should be taken into consideration when formulating a drug [12]. Low solubility can be addressed in several ways, including salt formation, development of cocrystals [13], amorphous solid dispersions [14], nanocrystals [15], cryogenic co-grinding [16], and forming complexes with cyclodextrins [17]. Cyclodextrins are useful multifunctional excipients used for the modification of important properties of the parent molecule such as solubility, chemical and physical stability, permeability, and delayed release in the case of pharmaceutical dosage forms [18]. Cyclodextrins include natural cyclodextrins comprising six to eight glucose subunits and modified cyclodextrins substituted with others, including hydroxypropyl, methyl, and sulfobutylether moieties. The specific distribution of functional groups on both sides of linked glucose subunits creates a toroidal shape with a polar surface outside and the lipophilic cavity inside the macromolecule. A number of glucose subunits, as well as chemical modifications, influence the physicochemical properties of the polymer, causing differences in the size of the cavity, polarity, and solubility, thus facilitating the ability to form complexes with specific kinds of guest molecules. Preparation techniques of inclusion complexes can be divided into two main groups: methods carried out in a solid state such as kneading and grinding, and methods performed in media. Grinding of powder API and cyclodextrins were used in the preparation of the β -cyclodextrin complex with acetaminophen, warfarin, indometacin, diazepam, and hydrocortisone acetate [19]. However, the coprecipitation method requires saturated solutions of guest molecules and saturated water solutions of cyclodextrin. The method was applied to prepare, among others, cyclodextrin complexes of silymarin and ketoconazole [20,21].

A wide range of analysis techniques was used for the identification of cyclodextrin complexes. The common techniques are FTIR and/or Raman spectroscopy, differential scanning calorimetry (DSC) and thermogravimetric thermal methods, X-ray crystallography, X-ray powder diffraction, and microscopic imaging techniques [22–24]. The DSC technique is the most preferable due to reliability and relative ease of analysis. However, the DSC technique is not specific enough to indicate domains involved in the interaction in guest–host inclusion systems. Moreover, the major drawback of thermal techniques is their destructive nature. Spectral methods, on the other hand, offer more complex measurement results, allowing a deeper insight into the kinds of interactions between API and cyclodextrin. Therefore, they can be used as an alternative method to DSC. Significant benefits of their application include their nondestructive nature and the ability for observations of close-range interactions, making it valuable for analysis of cyclodextrin carriers in inclusion complexes with APIs. However, an interpretation of the multidimensional data coming from the spectral analysis is complex, especially when identification of API domains interacting with cyclodextrins is required. Due to the high dimensionality of spectral data, the analysis may be an objective for a computer-aided approach. Machine learning techniques are applied in the areas where due to complexity or a large amount of data, analysis by a human

may be impossible in the required time. In some areas of analytical sciences, machine learning algorithms helped to significantly improve analysis by speeding up the process or allowing analytical techniques to change to one which is more ecologically, cost-, and time-efficient. In the area of electrochemistry, machine learning techniques allowed scientists to quantify electrolyte salts based on FTIR measurements, meeting the accuracy level of gas chromatography and inductively coupled plasma optical emission spectrometry [25]. In a more general approach, the classifier was proposed to determine the presence of molecular groups in FTIR spectra, resulting in less error-prone techniques comparing to human inspection [26]. Detection of glioma stem cell differentiation in hyperspectral data was handled with principal component analysis (PCA) logistic regression, decision tree, random forest, k-nearest neighbors, and hierarchical cluster analysis, resulting in a high accuracy of analysis [27]. Wavelet-based techniques were applied for feature extraction from FTIR data for classification of plant species [28]. The PCA and support vector machine were used for the discrimination of grape seeds basing on their hyperspectral images [29].

Development of computer-aided analysis algorithms of cyclodextrin complexes has not been reported to the best of our knowledge. The desired method should be able to detect the presence of the complex in a given sample when provided with the spectrum of a complex sample and the spectra of pure ingredients. Assuming a proper spectral dataset of complexed and pure API, machine learning is a relevant technique that is able to learn dependencies directly from spectral data.

As model APIs, the following prodrug β -lactam analogs were proposed: cefuroxime axetil (CA), cefetamet pivoxil (CT), and pivampicillin (PA). The prodrugs compared to acidic forms of selected β -lactam analogs are characterized by better bioavailability, more acceptable taste, and greater chemical stability. As a consequence of changing the acidic nature of hydrophilic β -lactam analogs to lipophilic prodrugs, the solubility decreased significantly. Considering the solubility issue of prodrugs, the use of solubilizers as drug carriers in their formulation such as cyclodextrins is fully justified. In this paper, α -, β -, and γ -cyclodextrins (α CD, β CD, γ CD, respectively) were used along with their synthetic analogs, including hydroxypropyl derivatives of all used natural CDs (HP α CD, HP β CD, HP γ CD) as well as methyl- β CD (m β CD) as carriers.

The aim of this work was to prepare complexes of prodrug β -lactam analogs and CDs and to confirm complex formation based on analysis of spectral and thermal measurements, followed by a development of a machine-learning-based model for the determination of complex formation in samples, based solely on ATR and tFTIR spectra with simultaneous identification of domains involved in the creation of the complex.

2. Results and Discussion

2.1. Determination of Complexes Formation

DSC was used as the reference method to determine the formation of prodrug β -lactam analogs–CDs systems. The DSC thermograms (Supplementary material: Figures S1–S21) of investigated systems were analyzed, and the results were presented in Table 1. For all prodrug β -lactam analogs–CD systems except CA–HP α CD, the complex formation was confirmed. In CA–CDs samples, the main peak of interest was the one corresponding to glass transition of the API around $T_g = 86$ °C. Analysis of thermal curves revealed that all samples except the one with HP α CD expose changes in complexed sample thermograms when compared to the physical mixture. Complexation occurred also in all samples of CT complexes, where changes around $T_g = 163$ °C were considered and in all samples of PA where main changes were noted around the glass transition point $T_g = 138$ °C.

Table 1. Observed changes on the differential scanning calorimetry (DSC) thermograms of complexed samples, pure constituents, and physical mixtures.

| Cyclodextrin | CA | | | CT | | | PA | | |
|----------------|------------------|---------|-----------------------|------------------|---------|------------------------|------------------|---------|----------------------------|
| | Physical Mixture | Complex | Change | Physical Mixture | Complex | Change | Physical Mixture | Complex | Change |
| α CD | 77 °C | - | Intensity ↓ | 71 °C | - | Intensity ↓ | 76 °C | - | Intensity ↓ |
| | 138 °C | - | Intensity ↓ | 126 °C | - | Intensity ↓ | 131 °C | - | Intensity ↓ |
| | | | | 164 °C | - | Intensity ↓ | | | |
| β CD | 114 °C | 116 °C | Peak ↔ Intensity ↓ | 113 °C | - | Intensity ↓ | 120 °C | - | Intensity ↓ |
| | | | | 164 °C | - | Intensity ↓ | | | |
| γ CD | 100 °C | 103 °C | Peak ↔ Intensity ↓ | 86 °C | 100 °C | Peak ↔ Intensity ↑ | 114 °C | 100 °C | Peak ↔ Intensity ↓ |
| | | | | 164 °C | - | Intensity ↓ | | | |
| HP α CD | - | - | No changes | 76 °C | 93 °C | Peak ↔, Intensity ↑ | 84 °C | 84 °C | Intensity ↑ Peak ↔ |
| | | | | 164 °C | - | Intensity ↓ | 138 °C | 140 °C | Intensity ↓ Peak ↔ |
| HP β CD | 83 °C | 87 °C | Peak ↔ Intensity ↓ | 83 °C | 91 °C | Intensity ↓ | 84 °C | 100 °C | Intensity ↑ Intensity ↓ |
| | | | | 164 °C | 164 °C | Intensity ↓ | 138 °C | - | Intensity ↓ Peak ↔ |
| HP γ CD | 91 °C | 85 °C | Peak ↔ Intensity ↓ | 85 °C | 95 °C | Intensity ↑ | 84 °C | 94 °C | Intensity ↓ Peak ↔ |
| | | | | 164 °C | 164 °C | Intensity ↓ | 138 °C | 145 °C | Intensity ↓ Peak ↔ |
| M β CD | 89 °C | 92 °C | Peak ↔ Intensity ↓ | 93 °C | 93 °C | Intensity ↑ | 82 °C | 92 °C | Intensity ↑ Peak ↔ |
| | | | | 165 °C | 165 °C | Intensity ↓ | 138 °C | 145 °C | Intensity ↓ Peak ↔ |

CA: cefuroxime axetil, CT: cefetamet pivoxil, PA: pivampicillin.

The second part of the experimental studies was the spectral analysis of the prodrug β -lactam analogs–CDs complexes obtained in relation to the changes of the positions of changes of the bands and vibration intensity observed for the APIs and cyclodextrins themselves. Using the theoretical approach, the theoretical spectra of API and cyclodextrin were determined, and, on the basis of the high consistency, the characteristic band was assigned to all analyzed compounds. Spectra of all samples in the dataset are presented in the supplementary material (Figures S22–S42). The spectra were cropped to range 400–2000 cm^{-1} and subjected to analysis with the machine learning approach.

2.2. Molecular Modeling

A molecular modeling approach was used in order to estimate the thermodynamics of complexation of the investigated prodrug β -lactam analogs–CDs systems. The optimized geometries of APIs along with the molecular electrostatic potentials (MEPs) calculated were presented in Figure 1. The optimized geometries and electrostatic potential maps for all investigated cyclodextrins are collected in the supplementary material (Figures S43–S49). Optimized geometries of prodrug β -lactam analogs and CDs were set as initial geometries for docking using molecular dynamics. Preparation of initial complexes conformations was done by placing the β -lactam analog structure in 3D space facing one of the molecular domains towards one of two sides of cyclodextrins, resulting in a diverse set of conformations. The molecular domains considered as possible for inclusion in the docking study were identified in regard to the nonpolar (green) regions visualized on the MEPs, for CA: 1-acetoxyethyl and furyl groups, for CT: 2,2-dimethylpropanoyloxymethyl and methoxyiminoacetyl groups, and for PA: 2,2-dimethylpropanoyloxymethyl and phenol groups. The polar regions included for CA: [(Aminocarbonyl)oxy]methyl group, for CT: 2-amino-1,3-thiazol-4-yl group, and they were considered as possible domains susceptible to the association on the surface of CDs. The geometries of the conformations were optimized, revealing possible stable conformations according to the heat of formation calculated.

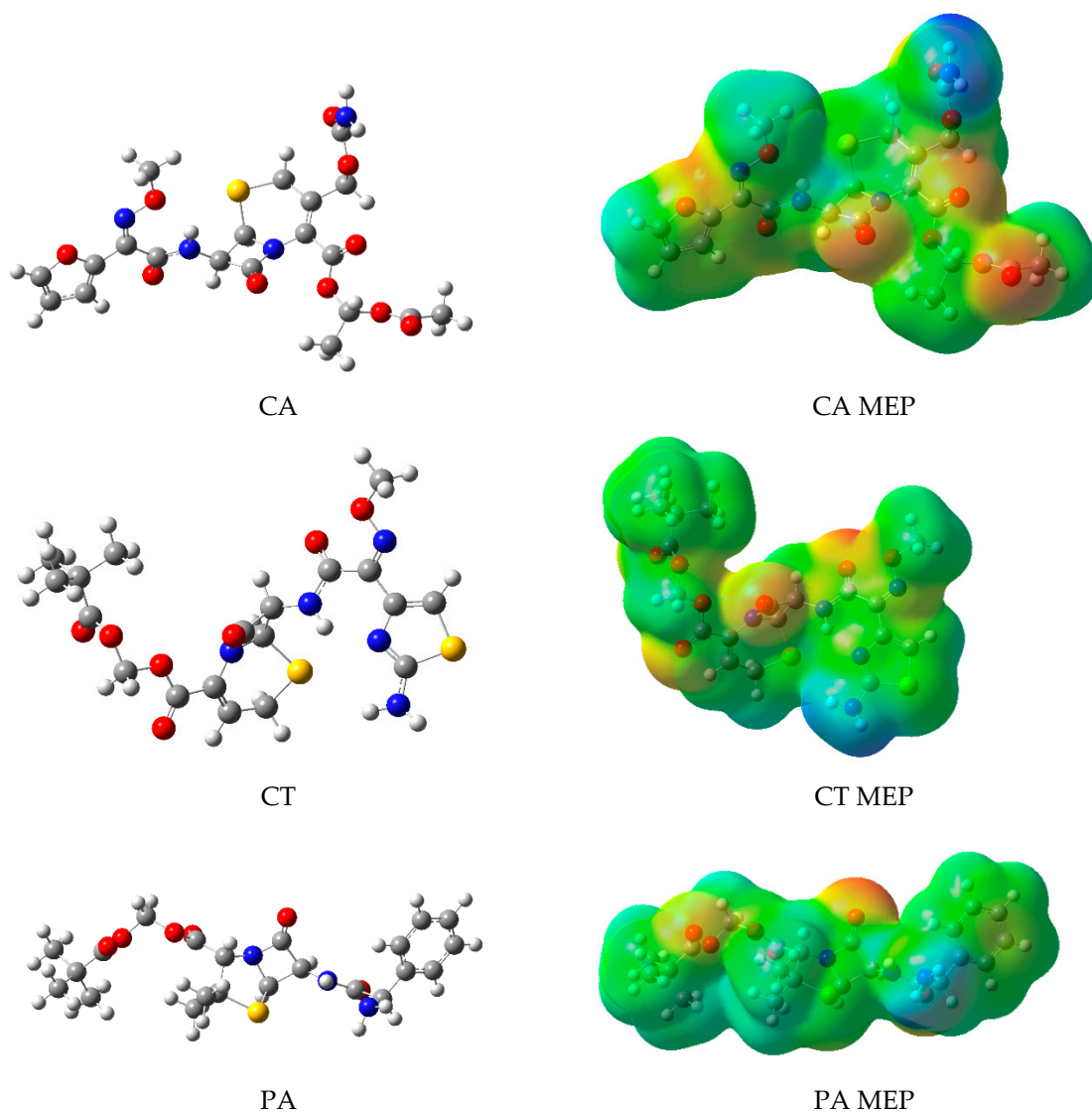


Figure 1. Optimized structure of prodrug β -lactam analogs and their maps of electrostatic potentials (MEPs).

Heats of formation were calculated for each ingredient of complexes as well as for the complexes. The results of the final optimization step were used to calculate changes in standard enthalpy of complex creation according to the equation:

$$\Delta_r H^\ominus = \Delta_{tw} H_{API \cdot CD}^\ominus - (\Delta_{tw} H_{API}^\ominus + \Delta_{tw} H_{CD}^\ominus)$$

where: $\Delta_r H^\ominus$ is the standard enthalpy of the complexation reaction, $\Delta_{tw} H_{API \cdot CD}^\ominus$ is the standard enthalpy of the complex creation, $\Delta_{tw} H_{API}^\ominus$ is the standard enthalpy of the API creation, and $\Delta_{tw} H_{CD}^\ominus$ is the standard enthalpy of cyclodextrin creation. The values were summarized in Table 2. Molecular modeling studies revealed binding modes of β -lactam analogs and CDs in complexes assuming a water environment. Binding modes of the investigated system were presented in the supplementary material S50–S52. Significantly positive complexation energy indicating that complexation is unlikely was observed for CA–HP α CD. For all other cases, the results suggest that in a simulation environment of the isolated system, the complexation would occur.

Table 2. Binding enthalpies of complexes conformations acquired according to machine learning results (experimentally favored) in relation to conformations favored in a simulated water environment.

| | Cefuroxime axetil | | Cefetamet pivoxil | | Pivamipicillin | |
|-------------|--------------------|------------------------|--------------------|------------------------|--------------------|------------------------|
| | Simulation Favored | Experimentally Favored | Simulation Favored | Experimentally Favored | Simulation Favored | Experimentally Favored |
| α | −157.31 | −134.83 | −239.54 | −235.75 | −197.79 | −149.85 |
| β | −238.22 | −145.17 | −279.69 | −131.33 | −263.39 | −195.71 |
| γ | −255.57 | −196.71 | −291.79 | −178.73 | −243.19 | −125.94 |
| HP α | 486.92 | 558.60 | −250.64 | −182.03 | −235.03 | −180.66 |
| M β | −256.15 | −168.48 | −185.90 | −144.22 | −236.37 | −173.36 |
| HP β | −904.60 | 352.80 | −268.65 | −259.74 | −193.11 | −223.94 |
| HP γ | −352.92 | −230.65 | −282.78 | −263.54 | −246.71 | −293.35 |

2.3. Machine Learning

DSC is the method of choice for investigating cyclodextrin complexes as it provides reliable information about the complex formation in the sample by a simple comparison of several major thermal peaks on thermograms of the complexed system in a physical mixture. On the other hand, an application of a more specific spectral method such as tFTIR/ATR is limited due to complex, interpretively difficult dependencies between changes in spectra of isolated constituents, physical mixtures, and complexed samples. The motivation for the use of a machine learning technique lies in the application of tFTIR/ATR spectra as a specific fingerprint of the sample, carrying information about bands involved in complex formation. The presented algorithm utilized a dataset of known data constituted by spectra of samples divided to one of two classes—complexed and noncomplexed—depending on whether the DSC confirmed complex formation in the sample or not. Automated protocol of search for the most useful spectral bands ensured that only the significant vibrations would be taken into consideration for machine learning. The accuracy of the classification model in one-out testing was 90.1%, and the exact results are presented in Table 3.

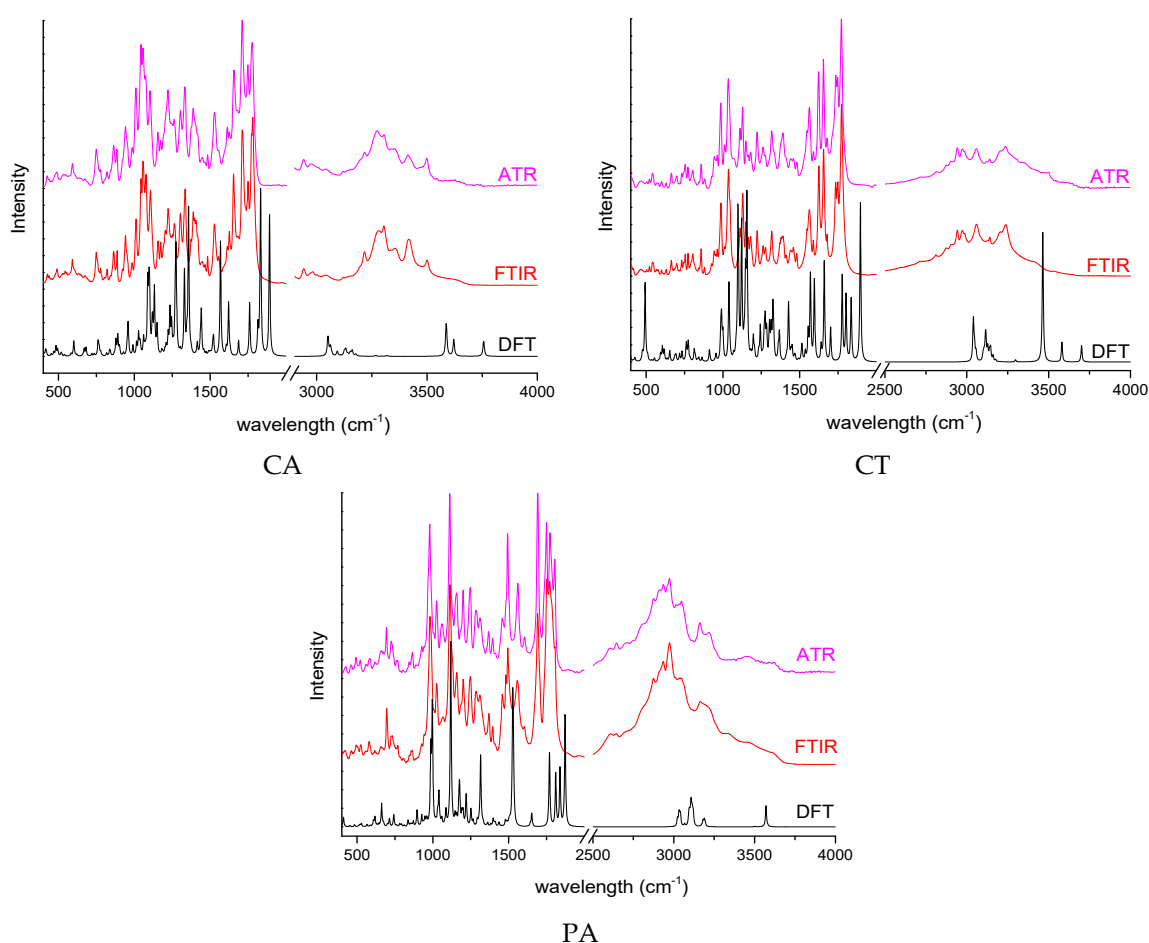
Table 3. Testing predictions of the model, where True—complexed sample, False—physical mixture/noncomplexed sample.

| | CA | | | | CT | | | | PA | | | |
|----------------|-----------|------------|-------------------|------------|-----------|------------|-------------------|------------|-----------|------------|-------------------|------------|
| | Complexed | | Physical mixtures | | Complexed | | Physical mixtures | | Complexed | | Physical mixtures | |
| | Predicted | True value | Predicted | True value | Predicted | True value | Predicted | True value | Predicted | True value | Predicted | True value |
| α CD | True | True | True | False | True | True | False | False | True | True | False | False |
| β CD | True | True | False | False | True | True | False | False | True | True | False | False |
| γ CD | True | True | False | False | True | True | False | False | True | True | False | False |
| HP α CD | False | False | False | False | True | True | False | False | True | True | False | False |
| HP β CD | False | True | False | False | True | True | False | False | True | True | False | False |
| HP γ CD | True | True | False | False | True | True | True | False | True | True | False | False |
| M β CD | True | True | False | False | False | True | False | False | True | True | False | False |

Analysis of the importance of bands being used as input data for the classification model allowed the specification of spectral regions, which, according to the acquired model, are responsible for changes caused by the formation of the complex. The bands selected by the algorithm (Table 4) were matched with the bands on the FTIR/ATR spectra of pure β -lactam analogs, supported by theoretical calculations (Figure 2). The binding modes involving β -lactam analog domains indicated by the algorithm were simulated and compared against the most optimal ones acquired by docking. The binding enthalpies in relation to enthalpies of energetically favored binding modes in a water environment were reported in Table 2. Structures of binding modes derived from machine learning results are collected in the supplementary material (Figures S53–S55).

Table 4. The spectral bands contributing to model, most important features first.

| API | Peak [cm^{-1}] | Theoretical | Description |
|-----|---------------------------|-------------|---|
| CA | 1772 | 1832 | C=O s in (acetyloxy)ethyl group |
| | | | C=O s in ((aminocarbonyl)oxy)methyl group |
| | 1264 | 1274 | C–O s in (acetyloxy)ethyl group C–H w in ((aminocarbonyl)oxy)methyl group |
| CT | 1774 | 1838 | C=O s in pivoxil group |
| | 1458 | 1479 | CH ₂ sc in pivoxil group |
| | 1277 | 1275 | C–C s between the thiazol ring and cephem group C–N–H b-ip between the thiazol ring and cephem group |
| PA | 1774 | 1838 | C=O s in pivoxil group |
| | 1458 | 1479 | CH ₂ sc in pivoxil group |
| | 1371 | 1362 | C–H b in β -lactam ring |
| | 1283 | 1293 | C–C s in pivoxil group |

**Figure 2.** Experimental and theoretical spectra of CA, CT, and PA.

Analysis showed that CA binds mostly with the axetil group, where interactions of carbonyl groups in the (acetyloxy)ethyl group and ((aminocarbonyl)oxy)methyl group are observed. The different mode of binding is possibly observed for the CA–HP β CD system. Molecular modeling studies showed that among all CA–CDs confirmed complexes, only CA–HP β CD showed positive enthalpy of binding when docked with the axetil domain included into the HP β CD lipophilic cavity. The machine learning model predicted a false negative for this example, which may suggest that the considered system has a unique binding mode in the whole dataset and cannot be predicted based on other examples. The most probable binding mode for CT according to machine learning analysis

involves interactions of the pivoxil group and some interaction among the thiazole ring and cephem group with the cyclodextrins. The interactions of thiazole and cephem groups may be observed in Figure S54. The considered groups interact with polar, hydroxypropyl-rich moieties placed on the outer side of cyclodextrin. The algorithm predicted one false negative and one false positive for the CT systems. A false positive for a physical mixture of CT-HP γ CD may indicate that spontaneous complexation occurred during storing, which is observed in systems of large, lipophilic CD such as HP γ CD. The false negative for a system of m β CD-CT may indicate that a different binding mode than in all other samples is present in the experimental sample. PA complexed and noncomplexed systems were predicted entirely correctly. According to the analysis of the most important bands involved in the prediction of complexation, the pivoxil group is mainly involved in the creation of the complex. The contribution of bands related to the C-H group in the β -lactam ring is also a significant variable for the classification model. Molecular modeling showed that deep inclusion may also involve a β -lactam ring in complexes with synthetic derivatives such as hydroxypropyl-substituted CD (Figure S55D–G).

2.4. Discussion

The presented study demonstrates the potential application of machine learning algorithm in the area of CD inclusion complexes by elucidating binding modes of CA, CT, and PA in complexes with CDs. The identified molecular domains of CA indicate the axetil group as the main interacting domain for all CDs systems except HP β CD. According to results published by Shah et al., the axetil group may actually be involved in the interaction with macromolecules after complexation with HP β CD [30]. The reason for the divergence in these results may be due to differences in the preparation method and underlines the necessity of human supervision over computer-aided analysis. An approach for the fully computer-based elucidation of the binding mode of the CA-HP β CD complex was investigated in the study by Sapte et al [31]. The binding modes were simulated using the force-field-based method MMFF94, revealing that the most energetically favored mode involves the inclusion of CA's furan ring. Application of the force-field method is a rough estimate of the outcome of a physical process of cyclodextrin complex preparation. The source of divergence may lay in the complexity of the inclusion phenomena, which is dependent on various environmental conditions unfeasible to estimate on the level of accuracy of static force-field simulation. The combined outcomes of discussed and presented papers allow us to draw the conclusion that theoretical simulation of the isolated system in certain binding modes does not cover conformational diversity related to the thermodynamics of complex inclusion in the real sample and is greatly affected by preparation method and intermolecular interactions. The simulation approach, to meet the requirements of a useful analytical technique, must be supported by an analysis of experimental results, whether it be conducted by human labor or by a machine learning algorithm.

3. Materials and Methods

3.1. Materials

Active pharmaceutical ingredients—cefuroxime axetil (CA), cefetamet pivoxil (CT), and pivampicillin (PA)—were synthesized in the Institute of Biotechnology and Antibiotics. Cyclodextrins, including α CD, β CD, γ CD, HP α CD, HP β CD, HP γ CD, and m β CD, were purchased from Sigma Aldrich Chemie. Complexes of API and cyclodextrins (1:1) were prepared by coprecipitation using saturated methanol solutions of APIs and saturated water solutions of CDs. The solutions of each binary system were shaken at 60 RPM in a 4.0 ml flask till complete evaporation of solvents and kept in a desiccator under a controlled humidity environment.

3.2. Instrumentation

As a base method for the identification of prodrug β -lactam analogs-CDs systems, DSC was used. The thermal experiment was carried out with DSC apparatus TA Instruments DSC Q20, with

the heating speed at 10 °C/min in the range 30–180 °C. The tFTIR spectra of complexes as well as pure substances constituting the given complex, along with their physical mixtures, were recorded with FTIR Bruker IFS 66v/S with a DTGS detector (Bruker, Billerica, MA, USA). The samples were put in KBr pills in 1:100 ratio by applying 8 metric tonnes of pressure in a hydraulic press. The ATR FTIR spectra were obtained with ATR BRUKER VERTEX 70 (Bruker, Billerica, MA, USA) apparatus with DLaTGS detector directly on powder. The vibrational infrared spectra were recorded in the range between 400 and 4000 cm^{-1} .

3.3. Computation

The theoretical investigation consisted of several stages which allowed the acquisition of the optimal geometry of isolated molecules and the ingredients of the complexes. Initial conformation of the α -, β -, and γ -cyclodextrins were acquired from their crystal structure [32–34]. Initial conformations of the remaining cyclodextrins, as well as the APIs, were obtained with the application of a conformational search using a genetic algorithm implemented in Open Babel library [35]. Conformation search was followed by DFT B3LYP 6-311G(3df,3dp) geometry optimization of the most promising conformers of prodrug β -lactam analogs and DFT B3LYP 6-31G(d,p) for CDs using the Gaussian09 package [36]. Molecular docking and the relaxation part of the docked system through molecular dynamics were applied in order to establish stable conformations of complexes. The prodrug β -lactam analogs were divided into three subdomains branching out from penam or cephem substructures. The domains were placed facing two possible sides of cyclodextrins and optimized using MMFF94. Relaxation of manually docked structures was done in ChemBio3D 13.0 using molecular dynamics simulation lasting 100 ps with 2 fs timestep in a water environment to determine the most stable conformation. The dielectric constant of water 80.2 was used during the simulation. The length of the simulation was determined by an observation of two scenarios: for the unstable pose, a drift of the guest molecule away from CD, or for the stable pose, a drop in the root-mean-square deviation of distances between geometric centroids of the guest molecule and CD below 0.75 Å. For all simulated systems, the scenarios happened within 100 ps of simulation. The most energetically favored conformations of systems using MMFF94 were optimized with parametrized method 7 (PM7) implemented in MOPAC [37]. As a result, thermodynamically stable conformations were acquired and set as the geometry for calculation of FTIR spectra using PM7 as well.

An automated machine learning approach was applied for the selection of a classification model, allowing the optimizations of hyperparameters in regard to an accuracy metric. The goal of this approach was to maximize the accuracy of a classification of tFTIR/ATR spectra of given samples to one of two classes indicating that the sample contains cyclodextrin complex or is a physical mixture. The code in Python for data manipulation and processing was developed and is available upon request. The machine learning code was based on TPOT [38] and scikit-learn [39] libraries. The TPOT allowed us to select the best-suited algorithm from the ones available in scikit-learn and to optimize parameters of the model implementing the selected algorithm. The algorithm selected by TPOT was an ensemble of randomized decision trees, as implemented in the ExtraTreesClassifier class of scikit-learn. The default parameters of the model were optimal. Input data were multispectral tFTIR/ATR images in the range 400–2000 cm^{-1} of all samples, including complexed systems as well as physical mixtures for a total of 42 samples. The model was tested with a leave-one-out, cross-validation (LOOCV) approach. The LOOCV involved iterative testing of the model, where during each iteration, one of the samples was used for testing and the rest for training. The protocol involves an iteration through all of the 42 samples in a dataset dividing them into 41 training samples in each iteration and a test sample indicated by an iterator. The protocol used each sample as a test sample once throughout the whole testing procedure. The test score is then computed as a mean value from partial scores, and the model is randomly initialized on the beginning of each test iteration, preventing any data leaks from a training set to a test set.

4. Conclusions

Cyclodextrin complexes were successfully prepared using the coprecipitation method in 20 of 21 systems of prodrug β -lactam analogs. The DSC as the method of choice for the confirmation of complex identification was used. Both melting point temperature shifts, as well as changes in the intensity of endothermic peaks on the DSC curves, were considered as indicators for complex formation. The only system which did not expose any significant changes in the DSC curve of the complexed sample in relation to the physical mixture was CA-HP α CD; thus, it was considered as noncomplexed. μ FTIR and ATR spectra were subjected to automatic analysis for the detection of an inclusion complex using a machine learning approach. The algorithm built according to the automatic machine learning paradigm was able to successfully predict cyclodextrin complex formation using only spectral data as inputs. Although learning of the model required DSC results as calibration data, only spectral results were used during testing, thereby calculating predictions reported in this study. Analysis of the importance of a given band for a trained classifier allowed us to explain the modes of binding which are observed on average for all investigated complexes of the given prodrug β -lactam analogs. Application of automated machine learning allowed us to achieve notably accurate results of more than 90% correctly classified spectra of prodrug β -lactam analogs-CDs systems to complexed/noncomplexed classes. The molecular modeling allowed us to confirm conclusions deduced from the machine-learning algorithm outcome. However, the most energetically favored binding modes predicted for isolated docked and optimized poses were not entirely in agreement with results acquired in the experiment. That suggests that environmental thermodynamic effects play a crucial role in the formation of particular binding modes in the experimental condition. The functionalities of computer algorithms based on the machine learning approach enable a better insight from the multispectral data gathered in experiments. A possibility to switch from a destructive thermal into a nondestructive, information-rich spectral method without additional human labor is a major milestone for the analysis of β -lactam analogs-CDs systems.

Supplementary Materials: The supplementary material is provided in the "Supplementary material.docx" file and contains detailed results of spectra and thermal measurements (Figures S1–S42). The conformations of complexes simulated with theoretical methods were also presented (Figures S43–S55).

Author Contributions: Conceptualization, M.M. and J.C.-P.; Data curation, M.M.; Formal analysis, M.M. and J.C.-P.; Funding acquisition, M.M. and J.C.-P.; Investigation, M.M., K.L., A.M., and J.C.-P.; Methodology, M.M. and J.C.-P.; Project administration, M.M. and J.C.-P.; Resources, M.M., K.L., A.M., and J.C.-P.; Software, M.M.; Supervision, J.C.-P.; Validation, M.M.; Visualization, M.M.; Writing—original draft, M.M. and J.C.-P.; Writing—review & editing, M.M. and J.C.-P.

Funding: The scientific work was funded from the budget resources for science in the years 2015–2018 as a research project within the program "Diamond Grant". This research was supported in part by PL-Grid Infrastructure.

Acknowledgments: We would like to thank Professor Irena Oszczapowicz from the Institute of Biotechnology and Antibiotics for delivering samples of β -lactam analogs used in this scientific work.

Conflicts of Interest: All authors declare no conflict of interest.

References

1. Paczkowska, M.; Mizera, M.; Szymanowska-Powałowska, D.; Lewandowska, K.; Błaszczak, W.; Gościańska, J.; Cielecka-Piontek, J. β -Cyclodextrin complexation as an effective drug delivery system for meropenem. *Eur. J. Pharm. Biopharm.* **2016**, *99*, 24–34. [[CrossRef](#)]
2. Paczkowska, M.; Szymanowska-Powałowska, D.; Mizera, M.; Siąkowska, D.; Błaszczak, W.; Piotrowska-Kempisty, H.; Cielecka-Piontek, J. Cyclodextrins as multifunctional excipients: Influence of inclusion into β -cyclodextrin on physicochemical and biological properties of tebipenem pivoxil. *PLoS ONE* **2019**, *14*, e0210694. [[CrossRef](#)]
3. Aljuffali, I.A.; Lin, C.F.; Chen, C.H.; Fang, J.Y. The codrug approach for facilitating drug delivery and bioactivity. *Expert Opin. Drug Deliv.* **2016**, *13*, 1311–1325. [[CrossRef](#)]

4. Jornada, D.H.; dos Santos Fernandes, G.F.; Chiba, D.E.; de Melo, T.R.F.; dos Santos, J.L.; Chung, M.C. The prodrug approach: A successful tool for improving drug solubility. *Molecules* **2015**, *21*, 42. [[CrossRef](#)]
5. Gupta, D.; Gupta, S.V.; Lee, K.D.; Amidon, G.L. Chemical and enzymatic stability of amino acid prodrugs containing methoxy, ethoxy and propylene glycol linkers. *Mol. Pharm.* **2009**, *6*, 1604–1611. [[CrossRef](#)]
6. Karaman, R. Prodrugs for Masking the Bitter Taste of Drugs. In *Application of Nanotechnology in Drug Delivery*; InTech: London, UK, 2014; pp. 399–445.
7. Liu, Y.; Zhao, D.; Sun, M.; Wei, W.; Wang, Y.; Zhou, J.; Kan, Q. Covalently mucoadhesive amphiphilic prodrug of 5-fluorouracil for enhanced permeation and improved oral absorption. *Drug Deliv. Transl. Res.* **2018**, *8*, 645–656. [[CrossRef](#)]
8. Placzek, A.T.; Ferrara, S.J.; Hartley, M.D.; Sanford-Crane, H.S.; Meinig, J.M.; Scanlan, T.S. Sobetirome prodrug esters with enhanced blood–brain barrier permeability. *Bioorgan. Med. Chem.* **2016**, *24*, 5842–5854. [[CrossRef](#)]
9. Leifer, F.; Omiatek, D.; Malinin, V.; Ong, J.; Li, Z.; Klecha, P.; Perkins, W. Prolonged activity of inhaled treprostinil prodrug nanoparticles in a rat model of pulmonary arterial hypertension. *Eur. Respir. J.* **2014**, *44*, P2356.
10. Chu, K.S.; Finniss, M.C.; Schorzman, A.N.; Kuijter, J.L.; Luft, J.C.; Bowerman, C.J.; DeSimone, J.M. Particle replication in nonwetting templates nanoparticles with tumor selective alkyl silyl ether docetaxel prodrug reduces toxicity. *Nano Lett.* **2014**, *14*, 1472–1476. [[CrossRef](#)]
11. Li, S.L.; Hou, Y.; Hu, Y.; Yu, J.; Wei, W.; Lu, H. Phosphatase-triggered cell-selective release of a Pt (IV)-backboned prodrug-like polymer for an improved therapeutic index. *Biomater. Sci.* **2017**, *5*, 1558–1566. [[CrossRef](#)]
12. Beaumont, K.; Webster, R.; Gardner, I.; Dack, K. Design of ester prodrugs to enhance oral absorption of poorly permeable compounds: Challenges to the discovery scientist. *Curr. Drug Metab.* **2003**, *4*, 461–485. [[CrossRef](#)]
13. Soares, F.L.; Carneiro, R.L. Green synthesis of ibuprofen–nicotinamide cocrystals and in-line evaluation by Raman spectroscopy. *Cryst. Growth Design* **2013**, *13*, 1510–1517. [[CrossRef](#)]
14. Trasi, N.S.; Taylor, L.S. Dissolution performance of binary amorphous drug combinations—Impact of a second drug on the maximum achievable supersaturation. *Int. J. Pharm.* **2015**, *496*, 282–290. [[CrossRef](#)]
15. Ige, P.P.; Baria, R.K.; Gattani, S.G. Fabrication of fenofibrate nanocrystals by probe sonication method for enhancement of dissolution rate and oral bioavailability. *Colloids Surf. B* **2013**, *108*, 366–373. [[CrossRef](#)]
16. Penkina, A.; Semjonov, K.; Hakola, M.; Vuorinen, S.; Repo, T.; Yliruusi, J.; Heinämäki, J. Towards improved solubility of poorly water-soluble drugs: Cryogenic co-grinding of piroxicam with carrier polymers. *Drug Dev. Ind. Pharm.* **2015**, *42*, 378–388. [[CrossRef](#)]
17. Cappello, B.; Di Maio, C.L.; Iervolino, M.; Miro, A. Improvement of solubility and stability of valsartan by hydroxypropyl-beta-cyclodextrin. *J. Inclusion. Phenom. Macrocycl. Chem.* **2006**, *54*, 289–294. [[CrossRef](#)]
18. Jambhekar, S.S.; Breen, P. Cyclodextrins in pharmaceutical formulations I: Structure and physicochemical properties, formation of complexes, and types of complex. *Drug Discov. Today* **2016**, *21*, 356–362. [[CrossRef](#)]
19. Miller, L.A.; Carrier, R.L.; Ahmed, I. Practical considerations in development of solid dosage forms that contain cyclodextrin. *J. Pharm. Sci.* **2007**, *96*, 1691–1707. [[CrossRef](#)]
20. Ghosh, A.; Biswas, S.; Ghosh, T. Preparation and evaluation of silymarin β -cyclodextrin molecular inclusion complexes. *J. Young Pharm.* **2011**, *3*, 205–210. [[CrossRef](#)]
21. Marzouk, M.A.; Kassem, A.A.; Samy, A.M.; Amer, R.I. Comparative evaluation of ketoconazole- β -cyclodextrin systems prepared by coprecipitation and kneading. *Drug Discov. Ther.* **2010**, *4*, 380–387.
22. Mohan, P.K.; Sreelakshmi, G.; Muraleedharan, C.V.; Joseph, R. Water soluble complexes of curcumin with cyclodextrins: Characterization by FT-Raman spectroscopy. *Vib. Spectr.* **2012**, *62*, 77–84. [[CrossRef](#)]
23. Aigner, Z.; Hassan, H.B.; Berkesi, O.; Kata, M.; Erős, I. Thermoanalytical, FTIR and X-ray studies of gemfibrozil-cyclodextrin complexes. *J. Therm. Anal. Calorimetry* **2005**, *81*, 267–272. [[CrossRef](#)]
24. Mura, P. Analytical techniques for characterization of cyclodextrin complexes in the solid state: A review. *J. Pharm. Biomed. Anal.* **2015**, *113*, 226–238. [[CrossRef](#)]
25. Ellis, L.D.; Buteau, S.; Hames, S.G.; Thompson, L.M.; Hall, D.S.; Dahn, J.R. A New Method for Determining the Concentration of Electrolyte Components in Lithium-Ion Cells, Using Fourier Transform Infrared Spectroscopy and Machine Learning. *J. Electrochem. Soc.* **2018**, *165*, A256–A262. [[CrossRef](#)]

26. Nalla, R.; Pinge, R.; Narwaria, M.; Chaudhury, B. Priority based functional group identification of organic molecules using machine learning. In *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*; ACM: New York, NY, USA, 2018; pp. 201–209.
27. Toplak, M.; Birarda, G.; Read, S.; Sandt, C.; Rosendahl, S.M.; Vaccari, L.; Borondics, F. Infrared Orange: Connecting Hyperspectral Data with Machine Learning. *Synchrotron Radiat. News* **2017**, *30*, 40–45. [[CrossRef](#)]
28. Kharyuk, P.; Nazarenko, D.; Oseledets, I. Comparative study of Discrete Wavelet Transforms and Wavelet Tensor Train decomposition to feature extraction of FTIR data of medicinal plants. *arXiv*, 2018; arXiv:1807.07099.
29. Zhao, Y.; Zhang, C.; Zhu, S.; Gao, P.; Feng, L.; He, Y. Non-destructive and rapid variety discrimination and visualization of single grape seed using near-infrared hyperspectral imaging technique and multivariate analysis. *Molecules* **2018**, *23*, 1352. [[CrossRef](#)]
30. Shah, M.; Pore, Y.; Dhawale, S.; Burade, K.; Kuchekar, B. Physicochemical characterization of spray dried ternary micro-complexes of cefuroxime axetil with hydroxypropyl- β -cyclodextrin. *J. Inclusion. Phenom. Macrocycl. Chem.* **2013**, *76*, 391–401. [[CrossRef](#)]
31. Sapte, S.; Pore, Y. Inclusion complexes of cefuroxime axetil with β -cyclodextrin: Physicochemical characterization, molecular modeling and effect of L-Arginine on complexation. *J. Pharm. Anal.* **2016**, *6*, 300–306. [[CrossRef](#)]
32. Puliti, R.; Mattia, C.A.; Paduano, L. Crystal structure of a new α -cyclodextrin hydrate form. Molecular geometry and packing features: Disordered solvent contribution. *Carbohydr. Res.* **1998**, *310*, 1–8. [[CrossRef](#)]
33. Ramos, A.I.; Braga, T.M.; Silva, P.; Fernandes, J.A.; Ribeiro-Claro, P.; Lopes, M.D.F.S.; Braga, S.S. Chloramphenicol cyclodextrin inclusion compounds: Co-dissolution and mechanochemical preparations and antibacterial action. *CrystEngComm* **2013**, *15*, 2822–2834. [[CrossRef](#)]
34. Harata, K. Crystal structure of γ -cyclodextrin at room temperature. *Chem. Lett.* **1984**, *13*, 641–644. [[CrossRef](#)]
35. O'Boyle, N.M.; Banck, M.; James, C.A.; Morley, C.; Vandermeersch, T.; Hutchison, G.R. Open Babel: An open chemical toolbox. *J. Cheminform.* **2011**, *3*, 33. [[CrossRef](#)]
36. Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. *Gaussian 09, Revision E.01*; Gaussian, Inc.: Wallingford, CT, USA, 2009.
37. Stewart Computational Chemistry. *Stewart, Stewart Computational Chemistry*; Stewart Computational Chemistry: Colorado Springs, CO, USA, 2007.
38. Olson, R.S.; Urbanowicz, R.J.; Andrews, P.C.; Lavender, N.A.; Moore, J.H. Automating biomedical data science through tree-based pipeline optimization. *Appl. Evol. Comput.* **2016**, 123–137. [[CrossRef](#)]
39. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. *Scikit-Learn: Machine Learning in Python*; JMLR 12; JMLR, Inc. and Microtome Publishing: USA, 2011; pp. 2825–2830.

Sample Availability: Samples of the CA, CT and PA complexes with CDs are available from the authors.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).