King Saud University

Arabian Journal of Chemistry

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

Development of machine learning model and analysis study of drug solubility in supercritical solvent for green technology development

Evgeniy V. Kostyrin [a], Vadim V. Ponkratov [b], Ahmed Salah Al-Shati [c],*

[a] *Bauman Moscow State Technical University, Moscow, Russian Federation*
[b] *Financial University under the Government of the Russian Federation, Moscow, Russian Federation*
[c] *Department of Chemical Engineering and Petroleum Industries, Al-Mustaqbal University College, 51001 Hillah, Babylon, Iraq*

Abstract   The current research is focused on development of machine learning model for estimation of pharmaceutical solubility in supercritical CO_2 as the green solvent. The main aim is to assess the suitability of supercritical processing for preparation of nanomedicine. Oxaprozin was taken as model drug for the solubility measurements, and its solubility was determined at different operational conditions by variation of temperature and pressure of the process. Artificial Neural Network (ANN) model was implemented for simulation of the drug solubility, and the best model was obtained with R^2 greater than 0.99 for the training and validation as well. The tested model was then exploited to understand the process, and it turned out that both pressure and temperature had major and considerable influence on the solubility of Oxaprozin in supercritical carbon dioxide as solvent. However, the effect of pressure was shown to be more significant on the solubility compared to the effect of pressure, which was attributed to the effect of pressure on the density of the supercritical solvent. The developed ANN model was indicated to be robust in estimating the values of drug solubility in wide range of conditions which can save time and cost of the measurements.

1. Introduction

Manufacturing of nanomedicine has been of great importance for pharmaceutical industry given that these nano drugs can provide better solubility compared to other micron sized drugs (Khoshmaram, 2021). Due to their enhanced surface area and subsequently free energy, the drug nanoparticles possess higher solubility in aqueous media which can resolve the issue of poor water solubility for majority of drugs specifically BCS Class II of drugs according to Biopharmaceutical Classification system (Saeed, 2021; Agostinho, 2021; Elworthy, 2008). Apart from nanonization method, other techniques have been used in this area to enhance the solubility of drug substances such as

* Corresponding author.
 E-mail address: ahmed.salah@mustaqbal-college.edu.iq (A. Salah Al-Shati).

ELSEVIER | **Production and hosting by Elsevier**

co-crystallization, amorphous solid dispersion, salt formation, and so on (Okamoto, 2022; Jennotte, 2022; de Assis, 2022; Butreddy, 2022; Queiroz, 2022; Morissette, 2004; Mahapatra, 2010; Ma, 2020).

The method of drug nanonization can be performed by various unit operations among which supercritical processing is efficient owing to its superior characteristics in pharmaceutical processing such as green technology, easy, and low-cost processing. The method can be categorized as top-down approach unlike the bottom-up approach where nanoparticles are synthesized in the solution phase and controlled to prepare the nanosized drug particles. The method of supercritical processing utilizes a solvent, mainly CO_2, and the drug is dissolved in the solvent for further processing (Zabihi, 2021; Zabihi, 2021; Pishnamazi, 2020; Bagheri et al., 2018). For utilizing the process for different drug molecules, the first step is to determine the solubility of the drug in the supercritical solvent, thereby the process can be realized to be feasible for the manufacturing the nanomedicine.

For determination of drug solubility in supercritical solvents such as CO_2, the simple method of gravimetric is usually employed which is based on weighing the drug powder before and after contacting with the solvent. This method is easy and facile which has attracted much attention in measuring solubility of drugs in different solvents (Keshavarz, 2019; de Souza, 2018). The method can be employed for measuring the solubility of drugs in supercritical solvents as well in which the treatment is done on the drug with the solvent and the weight difference is measured accordingly (Pishnamazi, 2021; Zabihi, 2020; Cao, 2021; Hartono et al., 2001).

Due to the difficulty of measuring the drug solubility in wide range of temperature and pressure, computational techniques can be used for estimation of drug solubility in all range of pressure and temperature which can be further used for design and optimization of nanomedicine manufacturing. Usually, thermodynamic models are employed for estimating drug solubility in solvents (Sodeifian, 2018; Sodeifian, 2020; Sodeifian et al., 2020), however recently machine learning based models have been successfully employed in estimation of drug solubility in supercritical carbon dioxide (Chinh Nguyen, 2022; Tianhao, 2022). The models of machine learning are shown to be accurate in estimation of drug solubility in supercritical CO_2, provided that measured data are available as function of temperature and pressure. The models can be utilized to find the relationship between solubility and pressure/temperature so that the cost and time of measurements can be saved.

In machine learning models, a set of data are used to train the network, and once the network has been trained, it can be used for prediction of process. However, the testing needs to be performed to check the validity of the model prior to use for process description and prediction. Selection of the algorithm is of fundamental importance and depends on the type of data and amount of data points for training and validation. Artificial neural networks (ANN) have attracted more attention in description of complex processes and can be used to estimate the solubility of drugs in supercritical carbon dioxide.

The main aim of the current study is to utilize the Artificial neural networks (ANN) method as machine learning technique for simulation of drug solubility in supercritical CO_2. We selected Oxaprozin as the model drug in this work and developed ANN model for the first time in prediction of drug solubility in supercritical CO_2. A bunch of solubility data for Oxaprozin was collected from resources and used for training and validation of the ANN model. Finally, the trained ANN model was used to assess the effect of parameters on the drug solubility.

2. Materials and methods

The data of Oxaprozin was collected from resources (Khoshmaram, 2021) and used for model development in this study. There are 32 datasets describing the solubility values of the drug as function of temperature and pressure (T & P). The

molecular structure of Oxaprozin is illustrated in Fig. 1. The set of data was obtained following a design of experiment work, with variations of temperature between 308 and 338 K, and pressure between 120 and 400 bar in the process, as reported by (Khoshmaram, 2021). The data was used to
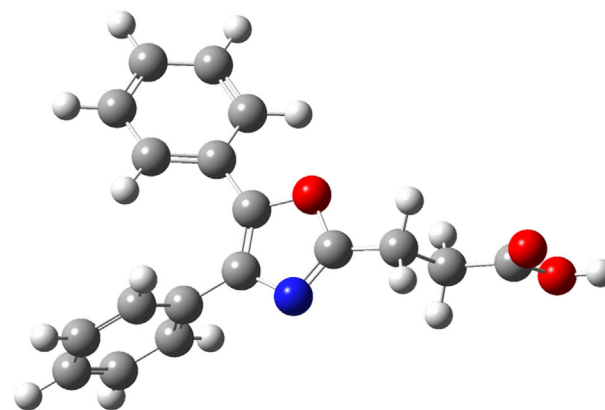

Fig. 1 Chemical structure of Oxaprozin drug studied in this work.

Table 1 Observed and predicted values of Oxaprozin solubility in supercritical solvent at different conditions using ANN.

| No.	Temperature (K)	Pressure (bar)	Observed	Predicted
1	3.08E + 02	1.20E + 02	8.19E−05	9.08574E-05
2	3.08E + 02	1.60E + 02	1.58E−04	0.000142509
3	3.08E + 02	2.00E + 02	2.24E−04	0.000208631
4	3.08E + 02	2.40E + 02	2.80E−04	0.000279781
5	3.08E + 02	2.80E + 02	3.44E−04	0.000347313
6	3.08E + 02	3.20E + 02	4.06E−04	0.000407448
7	3.08E + 02	3.60E + 02	4.73E−04	0.000465522
8	3.08E + 02	4.00E + 02	5.33E−04	0.000537732
9	3.18E + 02	1.20E + 02	7.55E−05	8.79459E−05
10	3.18E + 02	1.60E + 02	1.41E−04	0.000162564
11	3.18E + 02	2.00E + 02	2.45E−04	0.000256985
12	3.18E + 02	2.40E + 02	3.56E−04	0.000361461
13	3.18E + 02	2.80E + 02	4.64E−04	0.000466362
14	3.18E + 02	3.20E + 02	5.58E−04	0.000565694
15	3.18E + 02	3.60E + 02	6.60E−04	0.000661188
16	3.18E + 02	4.00E + 02	7.66E−04	0.000764756
17	3.28E + 02	1.20E + 02	5.34E−05	4.01401E−05
18	3.28E + 02	1.60E + 02	1.28E−04	0.000146205
19	3.28E + 02	2.00E + 02	3.02E−04	0.00027729
20	3.28E + 02	2.40E + 02	4.14E−04	0.000422876
21	3.28E + 02	2.80E + 02	5.82E−04	0.000572095
22	3.28E + 02	3.20E + 02	7.87E−04	0.000716947
23	3.28E + 02	3.60E + 02	8.51E−04	0.000856172
24	3.28E + 02	4.00E + 02	1.03E−03	0.000997934
25	3.38E + 02	1.20E + 02	3.31E−05	0.000042786
26	3.38E + 02	1.60E + 02	9.09E−05	0.000105806
27	3.38E + 02	2.00E + 02	2.98E−04	0.000284763
28	3.38E + 02	2.40E + 02	4.81E−04	0.000482399
29	3.38E + 02	2.80E + 02	6.77E−04	0.000686093
30	3.38E + 02	3.20E + 02	8.89E−04	0.000885361
31	3.38E + 02	3.60E + 02	1.08E−03	0.001075718
32	3.38E + 02	4.00E + 02	1.24E−03	0.001261726

Diagram
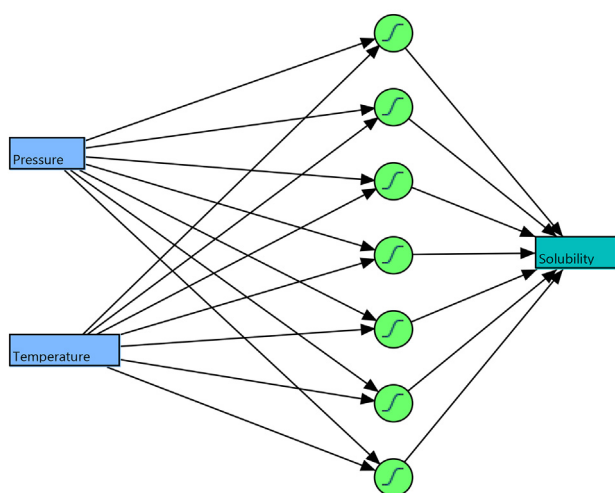
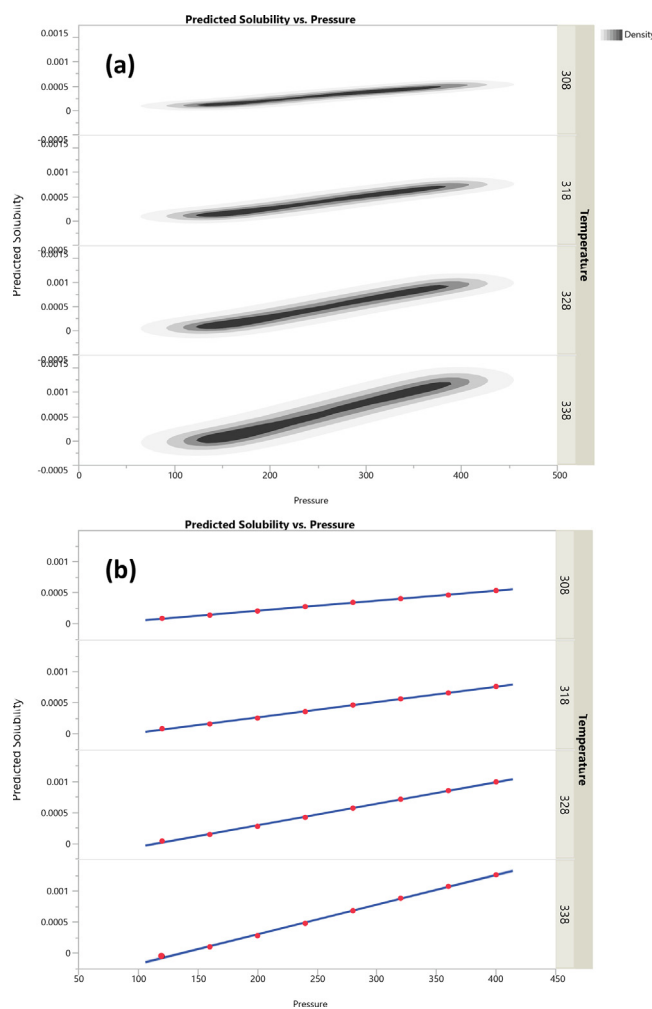**Fig. 2** Design of the model employed in this work for simulation of Oxaprozin solubility in supercritical solvent.

train and test the ANN model for simulation of the solubility data. The detailed description of the process and solubility measurements are reported elsewhere (Khoshmaram, 2021).

3. Modeling and simulation

In this study the Artificial Neural Network (ANN) method was used for simulation of Oxaprozin solubility in supercritical solvent at different pressure and temperature as listed in Table 1. Generally, ANNs are well suited for learning and classifying data, as they are inspired by human brains and nerves. The ANN model consists of a set of layers: input, hidden, and output layers with neurons (nodes) and weights (connections) which are distributed parallel to each other (Pelalak, 2020; Pelalak, 2021). Here, a three-layer, an input layer with two neurons (pressure (bar) and temperature (K)), an output layer with one neuron (Oxaprozin solubility), and one hidden layer with seven nodes are used in which the topology of this network is presented in Fig. 2. The experimental data were divided to training and testing data. The basis of ANN model is according to data training in which a meaningful relation is generated between input and output data according to the
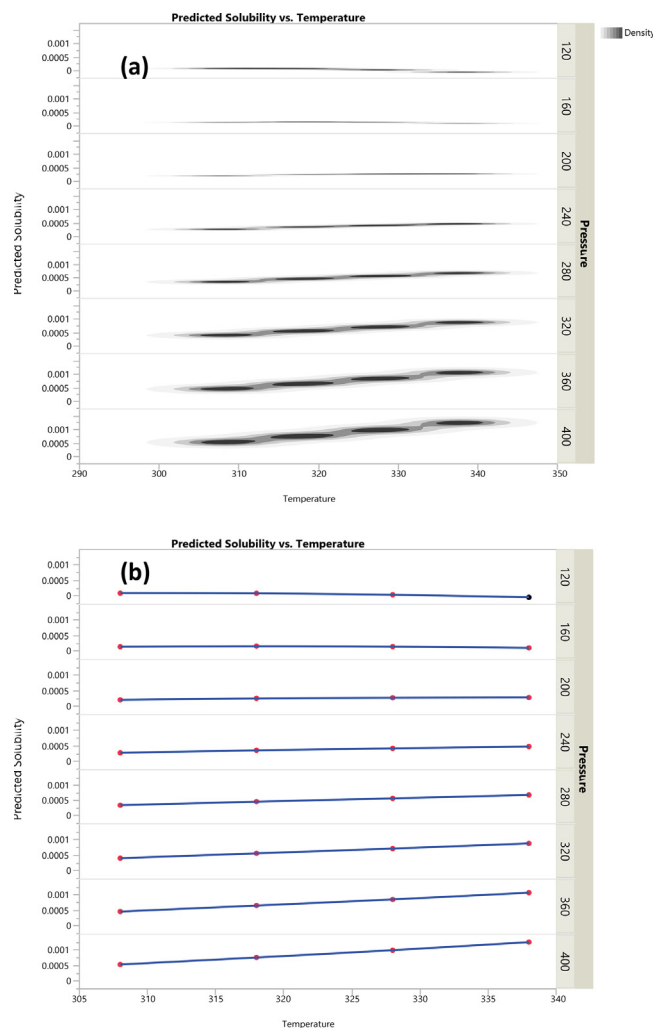

Fig. 3 Scatterplot matrix of predicted solubility of Oxaprozin drug in CO_2 vs pressure.

Fig. 4 Scatterplot matrix of predicted solubility of Oxaprozin drug in CO_2 vs temperature.

weights and biases in the designed network. After training step, the validation of model is evaluated with a portion of data according to the holdback method. The calculations of ANN for prediction of Oxaprozin drug solubility were achieved in *JMP* software. The validity of the provided ANN network was studied according to the linear regression between the experimental data and the predicted data by model. Also, the two and three dimensional diagrams predicting the drug solubility value (output) as a function of pressure and temperature of process (inputs) were obtained to analyze the process and understand the influence of inputs on the output parameter.

4. Results and discussion

The best dataset was obtained for the validation based on the validation method employed in the designed network. The observed and predicted data versus temperature and pressure are listed in Table 1. It is seen that 32 data points have been used in the model among which 33 % was used for the testing the model. It is clearly observed that the trained model is well capable of predicting the drug solubility in supercritical CO_2 at various temperature and pressure, while R^2 values of greater than 0.99 are obtained for both the training and validation.

The fitted and validated model was then used in this study in order to assess the effects of temperature and pressure on the variations of Oxaprozin solubility in the solvent. The results are useful to evaluate the solubility of Oxaprozin at different conditions so that one can save time and costs associated with measurements. The results of simulations are indicated in Figs. 3–6. Furthermore, the fitting accuracy and the residual of fitting are illustrated in Figs. 7 and 8, respectively. It is clearly observed that both temperature and pressure have influence on the drug solubility, and the values of
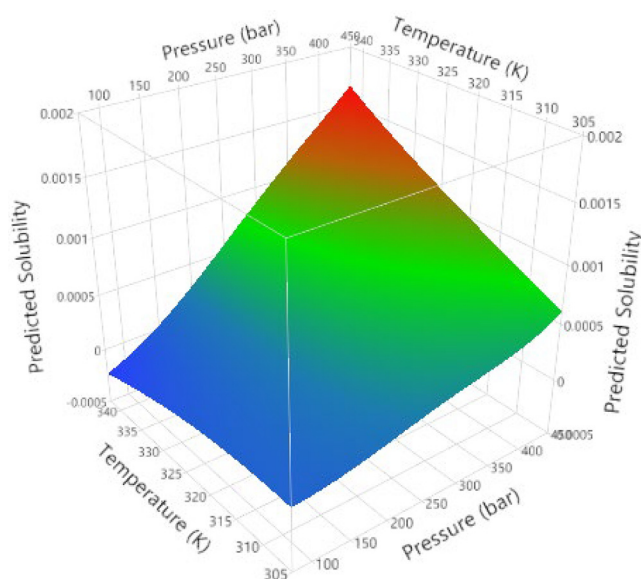

Fig. 6 3D surface plot of predicted solubility of Oxaprozin drug in CO_2.

solubility change with variation of these two factors. However, it is observed that the effect of pressure on the Oxaprozin solubility if more significant than the effect of temperature. This observation could be attributed to the effect of pressure on the density of the solvent and consequently enhance of solvation power of the supercritical solvent with increasing the pressure. Also, it is observed that the highest value of Oxaprozin solubility is observed at the highest pressure and temperature, i.e., 400 bar and 338 K.
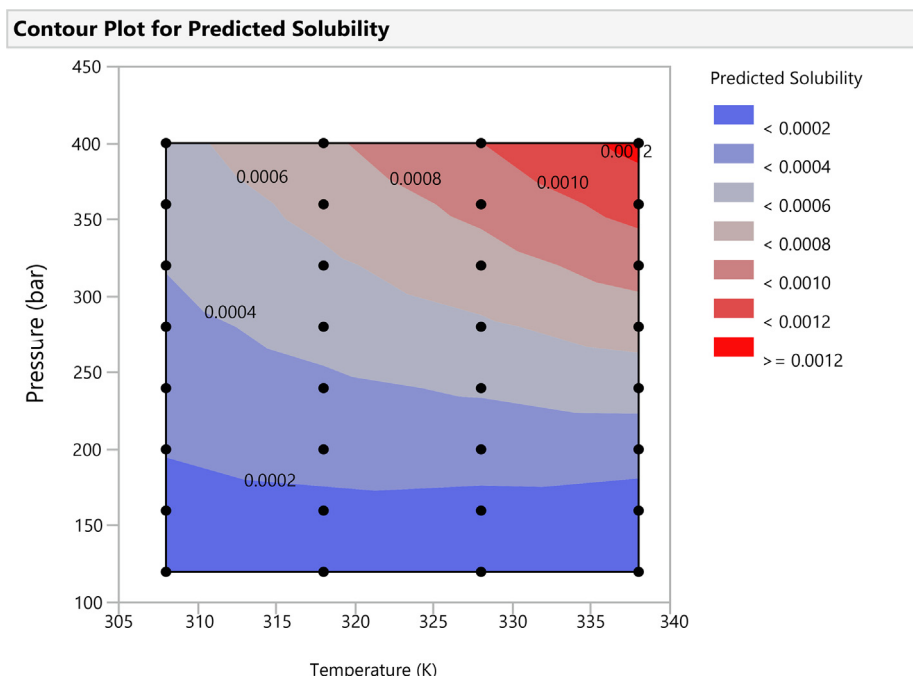
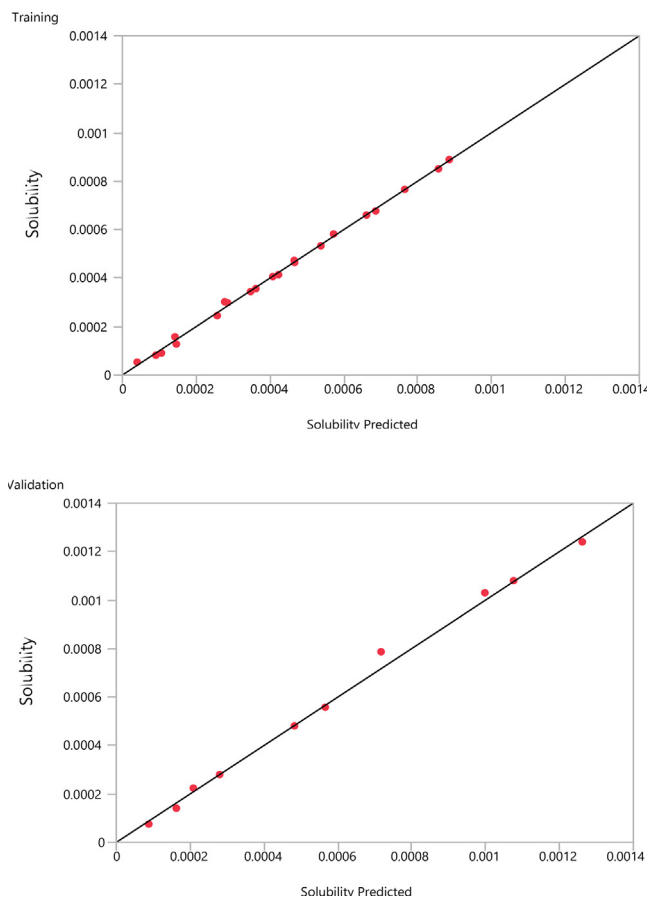

Fig. 5 Surface map of Oxaprozin drug solubility in CO_2.

Fig. 7 Training and validation data computed for solubility of Oxaprozin drug in CO_2.
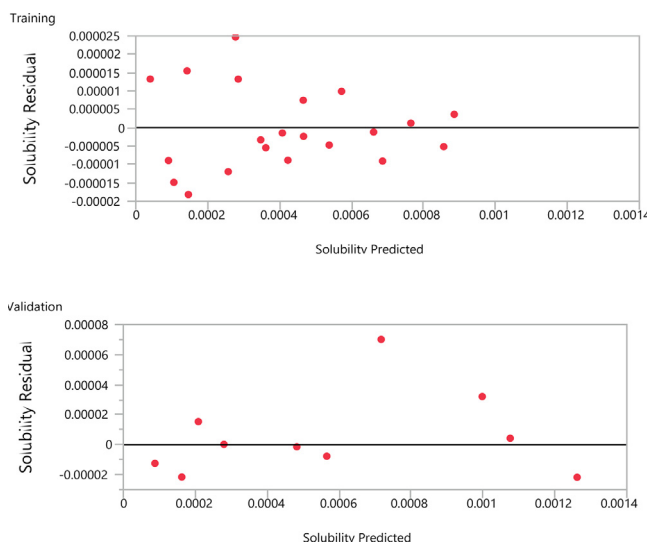

Fig. 8 The residual of fitting for data computed for solubility of Oxaprozin in CO_2.

5. Conclusions

We developed a machine learning model based on ANN (Artificial Neural Network) for estimation and analysis of drug solubility in supercritical carbon dioxide as solvent. The neural model consisted of one hidden layer and non-linear activation functions, while two inputs including temperature and pressure were considered for the model. The network was trained using the experimental data collected from literature to predict the value of solubility as the sole output. The training and validation steps turned out excellent results with R^2 of fitting more than 0.99 for both the training and validation steps. The tested model was then utilized to assess the effect of pressure and temperature on the solubility of Oxaprozin, and it was revealed that pressure had more significant influence on the Oxaprozin solubility in supercritical carbon dioxide. The model developed in this work indicated to be accurate and robust in prediction of drug solubility in supercritical solvents and can be used for development of supercritical processing for advanced pharmaceutical manufacturing of solid-dosage oral formulations.

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.

Acknowledgement

References

Agostinho, D.A.S. et al, 2021. New non-toxic biocompatible dianionic ionic liquids that enhance the solubility of oral drugs from BCS class II. J. Ionic Liquids 1 (1), 100003.

Bagheri, H., Ali Mansoori, G., Hashemipour, H., 2018. A novel approach to predict drugs solubility in supercritical solvents for RESS process using various cubic EoS-mixing rule. J. Mol. Liq. 261, 174–188.

Butreddy, A. et al, 2022. Hot-melt extruded hydroxypropyl methylcellulose acetate succinate based amorphous solid dispersions: Impact of polymeric combinations on supersaturation kinetics and dissolution performance. Int. J. Pharm. 615, 121471.

Cao, Y. et al, 2021. Neural simulation and experimental investigation of Chloroquine solubility in supercritical solvent. J. Mol. Liq. 333, 115942.

Chinh Nguyen, H. et al, 2022. Computational prediction of drug solubility in supercritical carbon dioxide: Thermodynamic and artificial intelligence modeling. J. Mol. Liq. 354, 118888.

de Assis, J.M.C. et al, 2022. Hot-melt extrudability of amorphous solid dispersions of flubendazole-copovidone: An exploratory study of the effect of drug loading and the balance of adjuvants on extrudability and dissolution. Int. J. Pharm. 614, 121456.

de Souza, B. et al, 2018. Solubility Measurement and Thermodynamic Modeling of N-(4-Methylphenyl-Z-3-chloro-2-(phenylthio)propenamide in 12 Pure Solvents at Temperatures Ranging from 278.15 to 318.15 K. J. Chem. Eng. Data 63 (5), 1419–1428.

Elworthy, T.R. et al, 2008. Orally bioavailable prodrugs of a BCS class 2 molecule, an inhibitor of HIV-1 reverse transcriptase. Bioorg. Med. Chem. Lett. 18 (24), 6344–6347.

Hartono, R., Mansoori, G.A., Suwono, A., 2001. Prediction of solubility of biomolecules in supercritical solvents. Chem. Eng. Sci. 56 (24), 6949–6958.

Jennotte, O. et al, 2022. Development of amorphous solid dispersions of cannabidiol: Influence of the carrier, the hot-melt extrusion parameters and the use of a crystallization inhibitor. J. Drug Delivery Sci. Technol. 71, 103372.

Keshavarz, L. et al, 2019. Influence of Impurities on the Solubility, Nucleation, Crystallization, and Compressibility of Paracetamol. Cryst. Growth Des. 19 (7), 4193–4201.

Khoshmaram, A. et al, 2021. Supercritical process for preparation of nanomedicine: Oxaprozin case study. Chem. Eng. Technol. 44 (2), 208–212.

Ma, Y. et al, 2020. Recent Progress in Continuous Crystallization of Pharmaceutical Products: Precise Preparation and Control. Org. Process Res. Dev. 24 (10), 1785–1801.

Mahapatra, A.K. et al, 2010. Model pharmaceutical co-crystallization: Guest-directed assembly of caffeine and aromatic tri-hydroxy and dicarboxylic acids into different heteromolecular hydrogen bonding networks in solid state. J. Mol. Struct. 963 (1), 63–70.

Morissette, S.L. et al, 2004. High-throughput crystallization: polymorphs, salts, co-crystals and solvates of pharmaceutical solids. Adv. Drug Deliv. Rev. 56 (3), 275–300.

Okamoto, T. et al, 2022. Comparison of improvements of aqueous dissolution of structurally analogous hydrophobic drugs by amorphous solid dispersion. Colloids Surf. A 632, 127744.

Pelalak, R. et al, 2020. Degradation of sulfonamide antibiotics using ozone-based advanced oxidation process: Experimental, modeling, transformation mechanism and DFT study. Sci. Total Environ. 734, 139446.

Pelalak, R. et al, 2021. Extraction of ingredients from tea leaves using oxidative enzymatic reaction and optimization of extraction conditions. Sci. Rep. 11 (1), 1–19.

Pishnamazi, M. et al, 2020. Using static method to measure tolmetin solubility at different pressures and temperatures in supercritical carbon dioxide. Sci. Rep. 10 (1).

Pishnamazi, M. et al, 2021. Chloroquine (antimalaria medication with anti SARS-CoV activity) solubility in supercritical carbon dioxide. J. Mol. Liq. 322, 114539.

Queiroz, M.B. et al, 2022. Co-crystallized sucrose-soluble fiber matrix: Physicochemical and structural characterization. LWT 154, 112685.

Saeed, A.M. et al, 2021. Comparative Bioavailability of Two Formulations of Biopharmaceutical Classification System (BCS) Class IV Drugs: A Case Study of Lopinavir/Ritonavir. J. Pharm. Sci. 110 (12), 3963–3968.

Sodeifian, G. et al, 2018. A comprehensive comparison among four different approaches for predicting the solubility of pharmaceutical solid compounds in supercritical carbon dioxide. Korean J. Chem. Eng. 35 (10), 2097–2116.

Sodeifian, G. et al, 2020. Experimental and thermodynamic analyses of supercritical CO2-Solubility of minoxidil as an antihypertensive drug. Fluid Phase Equilib. 522, 112745.

Sodeifian, G., Sajadian, S.A., Derakhsheshpour, R., 2020. Experimental measurement and thermodynamic modeling of Lansoprazole solubility in supercritical carbon dioxide: Application of SAFT-VR EoS. Fluid Phase Equilib., 507

Tianhao, Z. et al, 2022. Prediction of busulfan solubility in supercritical CO2 using tree-based and neural network-based methods. J. Mol. Liq. 351, 118630.

Zabihi, S. et al, 2020. Experimental Solubility Measurements of Fenoprofen in Supercritical Carbon Dioxide. J. Chem. Eng. Data 65 (4), 1425–1434.

Zabihi, S. et al, 2021. Thermodynamic study on solubility of brain tumor drug in supercritical solvent: Temozolomide case study. J. Mol. Liq. 321.

Zabihi, S. et al, 2021. Tenoxicam (Mobiflex) Solubility in Carbon Dioxide under Supercritical Conditions. J. Chem. Eng. Data 66 (2), 990–998.