Data in Brief 41 (2022) 107991

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

UHPLC-HRMS data from non-targeted screening for biotransformation products of cytostatic drug imatinib

Žiga Tkalec[a,b], Noelia Negreira[c], Miren López de Alda[c,*], Damiá Barceló[c], Tina Kosjek[a,b,*]

[a] Jožef Stefan Institute, Department of Environmental Sciences, Ljubljana, Slovenia
[b] Jožef Stefan International Postgraduate School, Ljubljana, Slovenia
[c] Water, Environmental and Food Chemistry Unit, Department of Environmental Chemistry, Institute of Environmental Assessment and Water Research, Spanish National Research Council, Barcelona, Spain

ARTICLE INFO

Article history:
Received 22 December 2021
Revised 10 February 2022
Accepted 18 February 2022
Available online 23 February 2022

Keywords:
Biodegradation
Wastewater
Anticancer
Drug
Transformation product

ABSTRACT

Imatinib is a selective tyrosine kinase inhibitor used to treat chronic myeloid leukemia. It enters the environment by excretion from the body through urine and feces and is transferred with wastewater to a wastewater treatment plant. There, it can be degraded by activated sludge, forming a number of biotransformation products. Presence of imatinib and its potential transformation products in the environment can impose a high risk to aquatic organisms and human health, therefore it is important to obtain knowledge of its environmental fate. The data presented here is a result of a simulated biodegradation of imatinib at two levels of activated sludge using a batch biotransformation setup, with and without carbon source. The data was acquired with UHPLC-HRMS/MS and processed by MzMine2.36 [1]. The dataset presents a table of $[M+H]^+$ features with retention times and corresponding MS/MS data. With development of new data mining tools this data can be used to identify new transformation products of imatinib and with it fully understand its environmental fate and the risk associated with its presence in the environment.

* Corresponding authors.
E-mail addresses: mlaqam@cid.csic.es (M. López de Alda), tina.kosjek@ijs.si (T. Kosjek).

https://doi.org/10.1016/j.dib.2022.107991

Specifications Table

Subject	*Analytical chemistry*
Specific subject area	*Non-targeted analysis of simulated environmental samples*
Type of data	Raw and table
How the data were acquired	*Data was acquired using HPLC coupled to a Thermo Scientific Q-Exactive triple quadrupole-Orbitrap high-resolution mass spectrometer (San Jose, CA, USA) operating in positive mode ESI(+). Chromatographic separation was achieved on a Hibar HR column (50 × 2.1 mm, 2 µm, Merck).*
Data format	*Abundance table in .csv, mass spectrometry data in .mgf and .raw files of the acquired data*
Description of data collection	*Mass spectrometry data was collected in the full scan range of 70-1000 m/z with mass resolution of 70,000 full width at half maximum (FWHM) at 200 m/z. MS/MS data was collected using data dependant acquisition fragmenting the 10 most abundant ions with resolution of 17,500 FWHM at 200 m/z. This data was pre-processed using MzMine 2.36.*
Data source location	• *Institution: Jožef Stefan Institute, Department of Environmental Sciences* • *City/Town/Region: Ljubljana* • *Country: Slovenia*
Data accessibility	Mendeley data: doi:10.17632/bby22v4bdr.1
Related research article	***For a published article:*** [1] Ž. Tkalec, N. Negreira, M. López de Alda, D. Barceló, T. Kosjek, A novel workflow utilizing open-source software tools in the environmental fate studies: The example of imatinib biotransformation, Sci. Total Environ. 797 (2021) 149063. https://doi.org/10.1016/j.scitotenv.2021.149063.

Value of the Data

- The data describes the biotransformation of the cytostatic drug imatinib. It consists of degradation kinetics at two levels of activated sludge in carbon rich and carbon poor media. Non-targeted screening data contains structural information about the biotransformation products formed from imatinib during microbial degradation.
- This data is of value for people researching the environmental fate of organic contaminants as well as data scientists interested in data mining and developing new identification tools.
- The data can be further mined by ever-developing identification and data mining tools and produce even more valuable insights on biotransformation and lead to broader understanding of processes during wastewater treatment. Further identification of other biotransformation products can help enhance the current knowledge on the risk associated with the introduction of the cytostatic imatinib into the environment.

1. Data Description

The dataset contains MzMine 2.32 processed non-targeted data and consists of:

- raw LC-MS data files obtained with Q-Orbitrap mass spectrometer
- quantitative data file “Imatinib_biotransformation_quant.csv” contains abundance and retention time data for each detected ion for which MS/MS data was acquired. The data table includes row ID, feature m/z and retention time and abundances in each sample.
- The file Imatinib_biotransformation_MSMS.mgf contains merged MS/MS data for each ion presented in the first table. The dataset is not filtered to allow processing for future use.

2. Experimental Design, Materials and Methods

2.1. Experimental design

Imatinib was spiked into an artificial wastewater matrix at the level of 1 mg/L. Its biodegradation was studied in two media, i.e. the nutrient-mineral medium (IMA-A, IMA-B, IMA-2A, IMA-2B) which simulated the composition of actual wastewater [2], and in the mineral medium (IMA-C, IMA-D, IMA-2C, IMA-2D), which was devoid of other carbon sources except for the spiked compound imatinib. To account for possible abiotic degradation and sorption the biomass was inhibited *via* the addition of formaldehyde (IMA-2B, IMA-2D). To follow possible differences in biotransformation kinetics each biodegradation setup was studied at two levels of activated sludge, e.g. high – 50 mL (IMA-A, IMA-B) and low – 10 mL (IMA-2A, IMA-2B, IMA-2C, IMA-2D) per 400 mL of total batch volume. Spontaneous degradation of the parent compound was controlled by adding imatinib into the distilled water (batch IMA-2E), while concurrent matrix biodegradation was followed in a non-spiked batch recorded as IMA-E. Sampling was performed in 24-h intervals on eight consecutive days, which is reflected by the last number of the datafile name.

2.2. Chemical analysis

Chromatographic separation was performed on a Waters Acquity (Waters, Milford, MA, USA) UHPLC system using a reversed phase Hibar HR (50 × 2.1 mm, 2 µm, Merck) column. The separation was achieved using water (A) and methanol (B) as mobile phases at a flow of 0.3 mL/min and at 25°C with an injection volume of 10 µL. The elution gradient was 0-1 min, 5% B; 3 min, 20% B; 6 min, 80% B; 7 min 100% B; 10-12 min, 5% B. Mass spectrometry data was acquired on a Thermo Scientific Q-Exactive quadrupole-Orbitrap (San Jose, CA, USA) mass spectrometer operating in positive electrospray ionisation mode (ESI(+)). Data was acquired using data dependant acquisition with the 5 most abundant ions being fragmented by higher-energy collisional dissociation (HCD) with normalized collision energy at 40 % and activation time of 120 ms. With that, full scan data was acquired with resolution of 70,000 FWHM at *m/z* 200 and data dependant acquisition of 17,500 FWHM at *m/z* 200. Spray voltage was 3000 V, sheath gas flow was set at 40 PSI, auxiliary gas at 10 PSI with capillary and vaporizer temperatures at 350°C and 400°C, respectively. Acquisition was controlled by Xcalibur 2.2 (Thermo Fischer Scientific).

2.3. Data pre-processing

Raw data was imported to MzMine 2.36 [3], a software widely used for processing of mass spectrometry non-targeted data (Sailwal et al., 2020; Sarpe and Schriemer, 2017). Mass detection was done with Exact mass module using 1.0E06 as noise threshold. Chromatograms were built with minimum time span of 0.02 min and minimum height 1.0E06. Mass error tolerance was 0.005 *m/z*. Chromatograms were deconvoluted using Wavelets (ADAP) module with signal-to-noise threshold of 10 and intensity window SN as signal-to-noise estimator. Minimum feature height was set to 1.0E06. Area threshold was set at 100 with peak duration range and retention time wavelet range of 0.02-0.90 min. Isotopic peaks were grouped with 0.005 *m/z* of mass tolerance and 0.1 min of retention time tolerance. Maximum allowed charge was 1. Ions were aligned using RANSAC alignment module with 0.005 *m/z* of mass tolerance and 0.2 min of retention time tolerance. Tolerance of 0.1 min was allowed after correction. Number of points was set to 20 % with threshold value of 3. After alignment, ions were filtered for duplicates with New average filter mode and 0.005 *m/z* of mass tolerance and 0.2 min of retention time tolerance.

CRediT Author Statement

Žiga Tkalec: Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing; **Noelia Negreira:** Investigation, Formal analysis, Visualization, Writing – review & editing; **Miren López de Alda:** Project administration, Conceptualization, Validation, Writing – review& editing, Supervision, Funding acquisition; **Damiá Barceló:** writing – review & editing; **Tina Kosjek:** Conceptualization, Investigation, Validation, Writing – review & editing, Supervision.

Ethics Statement

The authors declare that the manuscript meets all the rules and conditions described in the "Ethics in publishing" section standards (https://www.elsevier.com/journals/data-in-brief/2352-3409/guide-for-authors). This work did not include any investigations involving animal experimentation or human participants.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Imatinib biotransformation (Original data) (Mendeley Data)

Acknowledgments

This work has received funding from the European Commission through the projects CytoThreat (265264) and SOLUTIONS (603437), the Slovenian Research Agency - ARRS program P1-0143: Cycling of substances in the environment, mass balances, modelling of environmental processes and risk assessment the Spanish Ministry of Science and Innovation (Project CEX2018-000794-S), and the Generalitat de Catalunya (Consolidated Research Group 2017 SGR 01404-Water and Soil Quality Unit). Žiga Tkalec acknowledges the Slovenian Research Agency – ARRS for the Young researcher grant. The work reflects only the authors' views. The Community in not liable for any use that may be made of the information contained therein.

The authors thank Mrs. Vesna Mislej from Central Wastewater Treatment Plant in Ljubljana and Mrs. Marjetka Levstek from Wastewater treatment plant Domžale-Kamnik for providing wastewater samples.

References

[1] Ž. Tkalec, N. Negreira, M. López de Alda, D. Barceló, T. Kosjek, A novel workflow utilizing open-source software tools in the environmental fate studies: the example of imatinib biotransformation, Sci. Total Environ. 797 (2021) 149063, doi:10.1016/j.scitotenv.2021.149063.

[2] T. Kosjek, N. Negreira, M.L. de Alda, D. Barceló, Aerobic activated sludge transformation of methotrexate: Identification of biotransformation products, Chemosphere 119 (2015) S42–S50, doi:10.1016/j.chemosphere.2014.04.081.

[3] T. Pluskal, S. Castillo, A. Villar-Briones, M. Orešič, MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, BMC Bioinformatics 11 (2010) 1–11, doi:10.1186/1471-2105-11-395.