JSCSEN 86(7–8)625–780(2021)

ISSN 1820-7421(Online)

Journal of the Serbian Chemical Society

Electronic version

VOLUME 86

No 7–8

BELGRADE 2021

Available on line at

www.shd.org.rs/JSCS/

The full search of JSCS is available through

DOAJ DIRECTORY OF OPEN ACCESS JOURNALS

www.doaj.org

The **Journal of the Serbian Chemical Society** (formerly Glasnik Hemijskog društva Beograd), one volume (12 issues) per year, publishes articles from the fields of chemistry. The **Journal** is financially supported by the **Ministry of Education, Science and Technological Development of the Republic of Serbia.**

Articles published in the **Journal** are indexed in **Clarivate Analytics products: Science Citation Index-Expanded™** – accessed *via* **Web of Science®** and **Journal Citation Reports®**.

Impact Factor announced 2020: **1.097; 5-year Impact Factor**: **1.023**.

Articles appearing in the **Journal** are also abstracted by: **Scopus, Chemical Abstracts Plus (CAplus℠), Directory of Open Access Journals, Referativnii Zhurnal (VINITI), RSC Analytical Abstracts, EuroPub, Pro Quest** and **Asian Digital Library.**

Publisher:	**Serbian Chemical Society**, Karnegijeva 4/III, P. O. Box 36, 1120 Belgrade 35, Serbia tel./fax: +381–11–3370–467, E-mails: **Society** – shd@shd.org.rs; **Journal** – jscs@shd.org.rs Home Pages: **Society** – http://www.shd.org.rs/, **Journal** – http://www.shd.org.rs/JSCS/
Internet Service:	Contents, Abstracts and full papers (from Vol 64, No. 1, 1999) are available in the electronic form at the Web Site of the **Journal** (http://www.shd.org.rs/JSCS/).
Former Editors:	**Nikola A. Pušin** (1930–1947), **Aleksandar M. Leko** (1948–1954), **Panta S. Tutundžić** (1955**–**1961), **Miloš K. Mladenović** (1962–1964), **Đorđe M. Dimitrijević** (1965–1969), **Aleksandar R. Despić** (1969–1975), **Slobodan V. Ribnikar** (1975–1985), **Dragutin M. Dražić** (1986–2006).
Editor-in-Chief:	BRANISLAV Ž. **NIKOLIĆ**, Serbian Chemical Society (E-mail: jscs-ed@shd.org.rs)
Deputy Editor:	DUŠAN **SLADIĆ**, Faculty of Chemistry, University of Belgrade
Sub editors:	
Organic Chemistry	DEJAN **OPSENICA**, Institute of Chemistry, Technology and Metallurgy, University of Belgrade JÁNOS **CSANÁDI**, Faculty of Science, University of Novi Sad
Biochemistry and Biotechnology	OLGICA **NEDIĆ**, INEP – Institute for the Application of Nuclear Energy, University of Belgrade
Inorganic Chemistry	MILOŠ **ĐURAN**, Serbian Chemical Society
Theoretical Chemistry	IVAN **JURANIĆ**, Serbian Chemical Society
Physical Chemistry	LJILJANA **DAMJANOVIĆ-VASILIĆ**, Faculty of Physical Chemistry, University of Belgrade
Electrochemistry	SNEŽANA **GOJKOVIĆ**, Faculty of Technology and Metallurgy, University of Belgrade
Analytical Chemistry	SLAVICA **RAŽIĆ**, Faculty of Pharmacy, University of Belgrade
Polymers	BRANKO **DUNJIĆ**, Faculty of Technology and Metallurgy, University of Belgrade
Thermodynamics	MIRJANA **KIJEVČANIN**, Faculty of Technology and Metallurgy, University of Belgrade
Chemical Engineering	TATJANA **KALUĐEROVIĆ RADOIČIĆ**, Faculty of Technology and Metallurgy, University of Belgrade
Materials	RADA **PETROVIĆ**, Faculty of Technology and Metallurgy, University of Belgrade
Metallic Materials and Metallurgy	NENAD **RADOVIĆ**, Faculty of Technology and Metallurgy, University of Belgrade
Environmental and Geochemistry	VESNA **ANTIĆ**, Faculty of Agriculture, University of Belgrade
History of and Education in Chemistry	DRAGICA **TRIVIĆ**, Faculty of Chemistry, University of Belgrade
English Language Editors:	LYNNE **KATSIKAS**, Serbian Chemical Society VLATKA **VAJS**, Serbian Chemical Society JASMINA **NIKOLIĆ**, Faculty of Technology and Metallurgy, University of Belgrade
Technical Editors:	VLADIMIR **PANIĆ**, ALEKSANDAR **DEKANSKI**, VUK **FILIPOVIĆ**, Institute of Chemistry, Technology and Metallurgy, University of Belgrade
Journal Manager & Web Master:	ALEKSANDAR **DEKANSKI,** Institute of Chemistry, Technology and Metallurgy, University of Belgrade
Office:	VERA **ĆUŠIĆ**, Serbian Chemical Society

Editorial Board

From abroad: **R. Adžić**, Brookhaven National Laboratory (USA); **A. Casini**, University of Groningen (The Netherlands); **G. Cobb**, Baylor University (USA); **D. Douglas**, University of British Columbia (Canada); **G. Inzelt**, Etvos Lorand University (Hungary); **N. Katsaros**, NCSR "Demokritos", Institute of Physical Chemistry (Greece); **J. Kenny**, University of Perugia (Italy); **Ya. I. Korenman**, Voronezh Academy of Technology (Russian Federation); **M. D. Lechner**, University of Osnabrueck (Germany); **S. Macura**, Mayo Clinic (USA); **M. Spiteller**, INFU, Technical University Dortmund (Germany); **M. Stratakis**, University of Crete (Grecce); **M. Swart**, University de Girona (Cataluna, Spain); **G. Vunjak-Novaković**, Columbia University (USA); **P. Worsfold**, University of Plymouth (UK); **J. Zagal**, Universidad de Santiago de Chile (Chile).

From Serbia: **B. Abramović**, **V. Antić, V. Beškoski**, **J. Csanadi**, **Lj. Damjanović-Vasilić**, **A. Dekanski**, **V. Dondur**, **B. Dunjić**, **M. Đuran, S. Gojković**, **I. Gutman**, **B. Jovančićević**, **I. Juranić**, **L. Katsikas**, **M. Kijevčanin**, **V. Leovac**, **S. Milonjić**, **V.B. Mišković-Stanković, O. Nedić**, **B. Nikolić**, **J. Nikolić, D. Opsenica**, **V. Panić**, **M. Petkovska**, **R. Petrović**, **I. Popović**, **B. Radak**, T. **Kaluđerović Radiočić, N. Radović**, **S. Ražić**, **D. Sladić**, **S. Sovilj**, **S. Šerbanović**, **B. Šolaja**, **Ž. Tešić**, **D. Trivić, V. Vajs**.

Subscription: The annual subscription rate is **150.00 €** including postage (surface mail) and handling. For Society members from abroad rate is **50.00 €**. For the proforma invoice with the instruction for bank payment contact the Society Office (E-mail: shd@shd.org.rs) or see JSCS Web Site: http://www.shd.org.rs/JSCS/, option Subscription.

Godišnja pretplata: Za članove SHD: **2.500,00 RSD**, za penzionere i studente: **1000,00 RSD**, a za ostale: **3.500,00 RSD**:za organizacije i ustanove: **16.000,00 RSD**. Uplate se vrše na tekući račun Društva: **205-13815-62**, poziv na broj **320**, sa naznakom "pretplata za JSCS".

Nota: Radovi čiji su svi autori članovi SHD prioritetno se publikuju.

Odlukom Odbora za hemiju Republičkog fonda za nauku Srbije, br. 66788/1 od 22.11.1990. godine, koja je kasnije potvrđena odlukom Saveta Fonda, časopis je uvršten u kategoriju međunarodnih časopisa (**M-23**). Takođe, aktom Ministarstva za nauku i tehnologiju Republike Srbije, 413-00-247/2000-01 od 15.06.2000. godine, ovaj časopis je proglašen za publikaciju od posebnog interesa za nauku. **Impact Factor** časopisa objavljen 2020. godine iznosi **1,097**, a petogodišnji **Impact Factor 1,023**.

INSTRUCTIONS FOR AUTHORS (2021)

GENERAL

The *Journal of the Serbian Chemical Society* (the *Journal* in further text) is an international journal publishing papers from all fields of chemistry and related disciplines. Twelve issues are published annually. The Editorial Board expects the editors, reviewers, and authors to respect the well-known standard of professional ethics.

Types of Contributions

Original scientific papers	(up to 15 typewritten pages, including Figures, Tables and References) report original research which must not have been previously published.
Short communications	(up to 8 pages) report unpublished preliminary results of sufficient importance to merit rapid publication.
Notes	(up to 5 pages) report unpublished results of short, but complete, original research
Authors' reviews	(up to 40 pages) present an overview of the author's current research with comparison to data of other scientists working in the field
Reviews[a]	(up to 40 pages) present a concise and critical survey of a specific research area. Generally, these are prepared at the invitation of the Editor
Surveys	(about 25 pages) communicate a short review of a specific research area.
Book and Web site reviews	(1 - 2 pages)
Extended abstracts	(about 4 pages) of Lectures given at meetings of the Serbian Chemical Society Divisions
Letters to the Editor	report miscellaneous topics directed directly to the Editor

[a]Generally, Authors' reviews, Reviews and Surveys are prepared at the invitation of the Editor.

Submission of manuscripts

Manuscripts should be submitted using the **OnLine Submission Form**, available on the JSCS Web Site (**http://www.shd-pub.org.rs/index.php/JSCS**). The manuscript must be uploaded as a Word.doc or .rtf file, with tables and figures (including the corresponding captions – above Tables and below Figures), placed within the text to follow the paragraph in which they were mentioned for the first time.

Please note that **Full Names (**First Name, Last Name), **Full Affiliation** and **Country** (from drop down menu) of **ALL OF AUTHORS** (written in accordance with English spelling rules - the first letter capitalized) must be entered in the manuscript Submission Form (Step 3). Manuscript Title, authors' names and affiliations, as well as the Abstract, **WILL APPEAR** in the article listing, as well as in **BIBLIOGRAPHIC DATABASES (WoS, SCOPUS...)**, in the form and in the order entered in the author details

Graphical abstract

Graphical abstract is a one-image file containing the main depiction of the authors work and/or conclusion and must be supplied along with the manuscript. It must enable readers to quickly gain the main message of the paper and to encourage browsing, help readers identify which papers are most relevant to their research interests. Authors must provide an image that clearly represents the research described in the paper. The most relevant figure from the work, which summarizes the content, can also be submitted. The image should be submitted as a separate file in **Online Submission Form - Step 2**.

Specifications: The graphical abstract should have a clear start and end, reading from top to bottom or left to right. Please omit unnecessary distractions as much as possible.

- **Image size:** minimum of 500×800 pixels (W×H) and a minimum resolution of 300 dpi. If a larger image is sent, then please use the same ratio: 16 wide × 9 high. Please note that your image will be scaled proportionally to fit in the available window in TOC; a 150x240 pixel rectangle. Please be sure that the quality of an image cannot be increased by changing the resolution from lower to higher, but only by rescanning or exporting the image with a higher resolution, which can be set in usual "settings" option.
- **Font:** Please use Calibri and Symbol font with a large enough font size, so it is readable even from the image of a smaller size (150 × 240 px) in TOC.
- **File type:** JPG and PNG only.

No additional text, outline or synopsis should be included. Please do not use white space or any heading within the image.

Cover Letter

Manuscripts must be accompanied by a cover letter (strictly uploaded in **Online Submission Step 2**) in which the type of the submitted manuscript and a warranty as given below are given. The Author(s) has(have) to warranty that the manuscript submitted to the *Journal* for review is original, has been written by the stated author(s) and has not been published elsewhere; is currently not being considered for publication by any other journal and will not be submitted for such a review while under review by the *Journal*; the manuscript contains no libellous or other unlawful statements and does not contain any materials that violate any personal or proprietary rights of any other person or entity. All manuscripts will be acknowledged on receipt (by e-mail).

Illustrations

Illustrations (Figs, schemes, photos…) in TIF or EPS format (JPG format is acceptable for colour and greyscale photos, only), must be additionally uploaded (Online Submission Step 2) as a separate file or one archived (.zip, .rar or .arj) file. Figures and/or Schemes should be prepared according to the **Artwork Instructions -** http://www.shd.org.rs/JSCS/jscs-pdf/Artwork_Instructions.pdf!

For any difficulties and questions related to **OnLine Submission Form -** https://www.shd-pub.org.rs/index.php/JSCS/submission/wizard, please refer to **User Guide -** https://openjournal-systems.com/ojs-3-user-guide/ , Chapter **Submitting an Article -** https://openjournalsystems.com/ojs-3-user-guide/submitting-an-article/. If difficulties still persist, please contact JSCS Editorial Office at **JSCS@shd.org.rs**

A manuscript not prepared according to these instructions will be returned for resubmission without being assigned a reference number.

Conflict-of-Interest Statement*: Public trust in the peer review process and the credibility of published articles depend in part on how well a conflict of interest is handled during writing, peer review, and editorial decision making. A conflict of interest exists when an author (or the author's institution), reviewer, or editor has financial or personal relationships that inappropriately influence (bias) his or her actions (such relationships are also known as dual commitments, competing interests, or competing loyalties). These relationships vary from those with negligible potential to those with great potential to influence judgment, and not all relationships represent true conflict of interest. The potential for a conflict of interest can exist whether or not an individual believes that the relationship affects his or her scientific judgment. Financial relationships (such as employment, consultancies, stock ownership, honoraria, paid expert testimony) are the most easily identifiable conflicts of interest and the most likely to undermine the credibility of the journal, the authors, and of science itself. However, conflicts can occur for other reasons, such as personal relationships, academic competition, and intellectual passion.

Informed Consent Statement*: Patients have a right to privacy that should not be infringed without informed consent. Identifying information, including patients' names, initials, or hospital numbers, should not be published in written descriptions, photographs, and pedigrees unless the information is essential for scientific purposes and the patient (or parent or guardian) gives written informed consent for publication. Informed consent for this purpose requires that a patient who is identifiable be shown the manuscript to be published. Authors should identify Individuals who provide writing assistance and disclose the funding source for this assistance. Identifying details should be omitted if they are not essential. Complete anonymity is difficult to achieve, however, and informed consent should be obtained if there is any doubt. For example, masking the eye region in photographs of patients is inadequate protection of anonymity. If identifying characteristics are altered to protect anonymity, such as in genetic pedigrees, authors should provide assurance that alterations do not distort scientific meaning and editors should so note. The requirement for informed consent should be included in the journal's instructions for authors. When informed consent has been obtained it should be indicated in the published article.

Human and Animal Rights Statement* When reporting experiments on human subjects, authors should indicate whether the procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000 (5). If doubt exists whether the research was conducted in accordance with the Helsinki Declaration, the authors must explain the rationale for their approach, and demonstrate that the institutional review body explicitly approved the doubtful aspects of the study. When reporting experiments on animals, authors should be asked to indicate whether the institutional and national guide for the care and use of laboratory animals was followed.

---

*International Committee of Medical Journal Editors ("Uniform Requirements for Manuscripts Submitted to Biomedical Journals"), February 2006

PROCEDURE

All contributions will be peer reviewed and only those deemed worthy and suitable will be accepted for publication. The Editor has the final decision. To facilitate the reviewing process, authors are encouraged to suggest up to three persons competent to review their manuscript. Such suggestions will be taken into consideration but not always accepted. If authors would prefer a specific person not be a reviewer, this should be announced. The Cover Letter must be accompanied by these suggestions. Manuscripts requiring revision should be returned according to the requirement of the Editor, within 60 days upon reception of the reviewing comments by e-mail.

The *Journal* maintains its policy and takes the liberty of correcting the English as well as false content of manuscripts **provisionally accepted** for publication in the first stage of reviewing process. In this second stage of manuscript preparation by JSCS Editorial Office, the author(s) may be required to supply some **additional clarifications and corrections**. This procedure will be executed during copyediting actions, with a demand to author(s) to perform corrections of unclear parts before the manuscript would be published OnLine as **finally accepted manuscript (OLF Section of the JSCS website)**. Please note that the manuscript can receive the status of **final rejection** if the author's corrections would not be satisfactory.

When finally accepted manuscript is ready for printing, the corresponding author will receive a request for proof reading, which should be performed within 2 days. Failure to do so will be taken as the authors agree with any alteration which may have occurred during the preparation of the manuscript for printing.

Accepted manuscripts of active members of the Serbian Chemical Society (all authors) have publishing priority.

MANUSCRIPT PRESENTATION

Manuscripts should be typed in English (either standard British or American English, but consistent throughout) with 1.5 spacing (12 points Times New Roman; Greek letters in the character font Symbol) in A4 format leaving 2.5 cm for margins. For Regional specific, non-standard characters that may appear in the text, save documents with Embed fonts Word option: *Save as -> (Tools) -> Save Options… -> Embed fonts in the text.*

The authors are requested to seek the assistance of competent English language expert, if necessary, to ensure their English is of a reasonable standard. The Serbian Chemical Society can provide this service in advance of submission of the manuscript. If this service is required, please contact the office of the Society by e-mail (jscs-info@shd.org.rs).

Tables, figures and/or schemes must be embedded in the main text of the manuscript and should follow the paragraph in which they are mentioned for the first time. **Tables** must be prepared with the aid of the **WORD table function**, without vertical lines. The minimum size of the font in the tables should be **10 pt**. Table columns must not be formatted using multiple spaces. Table rows must not be formatted using any returns (enter key; ↵ key) and are **limited to 12 cm width**. Tables should not be incorporated as graphical objects. **Footnotes to Tables** should follow them and are to be indicated consequently (in a single line) in superscript letters and separated by semi-column.

Table caption must be placed above corresponding Table, while **Captions of the Illustrations** (Figs. Schemes...) must follow the corresponding item. **The captions, either for Tables or Illustrations**, should make the items comprehensible without reading of the main text (but clearly referenced in), must follow numerical order (Roman for Tables, Arabic for Illustrations), and should not be provided on separate sheets or as separate files.

High resolution Illustrations (named as Fig. 1, Fig. 2… and/or Scheme 1, Scheme 2…) in **TIF or EPS format** (JPG format is acceptable for photos, only) **must be additionally uploaded as a separate files or one archived (.zip, .rar) file.**

Illustrations should be prepared according to the ARTWORK INSTRUCTIONS - http://www.shd.org.rs/JSCS/jscs-pdf/Artwork_Instructions.pdf. **!**

All pages of the manuscript must be numbered continuously.

DESIGNATION OF PHYSICAL QUANTITIES AND UNITS

IUPAC recommendations for the naming of compounds should be followed. SI units, or other permissible units, should be employed. The designation of physical quantities must be in italic throughout the text (including figures, tables and equations), whereas the units and indexes (except for indexes having the meaning of physical quantities) are in upright letters. They should be in Times New Roman font. In graphs and tables, a slash should be used to separate the designation of a physical quantity from the unit

(example: p / kPa, j / mA cm^{-2}, t / °C, T_0 / K, τ / h, ln (j / mA cm^{-2})…). Designations such as: p (kPa), t [min]…, are not acceptable. However, if the full name of a physical quantity is unavoidable, it should be given in upright letters and separated from the unit by a comma (example: Pressure, kPa; Temperature, K; Current density, mA cm^{-2}…). Please do not use the axes of graphs for additional explanations; these should be mentioned in the figure captions and/or the manuscript (example: "pressure at the inlet of the system, kPa" should be avoided). The axis name should follow the direction of the axis (the name of y-axis should be rotated by 90°). Top and right axes should be avoided in diagrams, unless they are absolutely necessary.

Latin words, as well as the names of species, should be in *italic*, as for example: *i.e.*, *e.g.*, *in vivo*, *ibid*, *Calendula officinalis* L., *etc*. The branching of organic compound should also be indicated in *italic*, for example, *n*-butanol, *tert*-butanol, *etc*.

Decimal numbers must have decimal points and not commas in the text (except in the Serbian abstract), tables and axis labels in graphical presentations of results. Thousands are separated, if at all, by a comma and not a point.

Mathematical and chemical equations should be given in separate lines and must be numbered, Arabic numbers, consecutively in parenthesis at the end of the line. All equations should be embedded in the text Complex equations (fractions, integrals, matrix…) should be prepared with the aid of the **Microsoft Equation 3.0** (or higher) or **MathType** (Do not use them to create simple equations and labels). **Using the Insert -> Equation option, integrated in MS Office 2010 and MS Office 2013, as well as insertion of equation objects within paragraph text IS NOT ALLOWED.**

ARTICLE STRUCTURE

- TITLE PAGE;
- MAIN TEXT – including Tables and Illustrations with corresponding captions;
- SUPPLEMENTARY MATERIAL (optional**)**

Title page

- **Title** in bold letters, should be clear and concise, preferably 12 words or less. The use of non-standard abbreviations, symbols and formulae is discouraged.
- AUTHORS' NAMES in capital letters with the full first name, initials of further names separated by a space and surname. Commas should separate the author's names except for the last two names when 'and' is to be used. In multi-affiliation manuscripts, the author's affiliation should be indicated by an Arabic number placed in superscript after the name and before the affiliation. Use * to denote the corresponding author(s).
- *Affiliations* should be written in italic. The e-mail address of the corresponding author should be given after the affiliation(s).
- *Abstract:* A one-paragraph abstract written of 150 – 200 words in an impersonal form indicating the aims of the work, the main results and conclusions should be given and clearly set off from the text. Domestic authors should also submit, on a separate page, an Abstract - Izvod, the author's name(s) and affiliation(s) in Serbian (Cyrillic letters). (Домаћи аутори морају доставити Извод (укључујући имена аутора и афилијацију) на српском језику, исписане ћирилицом, иза Захвалнице, а пре списка референци.) For authors outside Serbia, the Editorial Board will provide a Serbian translation of their English abstract.
- *Keywords:* Up to 6 keywords should be given. Do not use words appearing in the manuscript title
- RUNNING TITLE: A one line (maximum five words) short title in capital letters should be provided.

Main text – should have the form:

- **INTRODUCTION,**
- **EXPERIMENTAL (RESULTS AND DISCUSSION),**
- **RESULTS AND DISCUSSION (EXPERIMENTAL),**
- **CONCLUSIONS,**
- **NOMENCLATURE (optional) and**
- ***Acknowledgements*: If any.**
- REFERENCES (Citation of recent papers published in chemistry journals that highlight the significance of work to the general readership is encouraged.)

The sections should be arranged in a sequence generally accepted for publication in the respective fields. They subtitles should be in capital letters, centred and NOT numbered.

- The INTRODUCTION should include the aim of the research and a concise description of background information and related studies directly connected to the paper.
- The EXPERIMENTAL section should give the purity and source of all employed materials, as well as details of the instruments used. The employed methods should be described in sufficient detail to enable experienced persons to repeat them. Standard procedures should be referenced and only modifications described in detail. On no account should results be included in the experimental section.

Chemistry

Detailed information about instruments and general experimental techniques should be given in all necessary details. If special treatment for solvents or chemical purification were applied that must be emphasized.

Example: Melting points were determined on a Boetius PMHK or a Mel-Temp apparatus and were not corrected. Optical rotations were measured on a Rudolph Research Analytical automatic polarimeter, Autopol IV in dichloromethane (DCM) or methanol (MeOH) as solvent. IR spectra were recorded on a Perkin-Elmer spectrophotometer FT-IR 1725X. ^{1}H and ^{13}C NMR spectra were recorded on a Varian Gemini-200 spectrometer (at 200 and 50 MHz, respectively), and on a Bruker Ultrashield Advance III spectrometer (at 500 and 125 MHz, respectively) employing indicated solvents (*vide infra*) using TMS as the internal standard. Chemical shifts are expressed in ppm (δ / ppm) values and coupling constants in Hz (J / Hz). ESI-MS spectra were recorded on Agilent Technologies 6210 Time-Of-Flight LC-MS instrument in positive ion mode with CH_3CN/H_2O 1/1 with 0.2 % HCOOH as the carrying solvent solution. Samples were dissolved in CH_3CN or MeOH (HPLC grade purity). The selected values were as follows: capillary voltage = 4 kV, gas temperature = 350 °C, drying gas flow 12 L min^{-1}, nebulizer pressure = 310 kPa, fragmentator voltage = 70 V.The elemental analysis was performed on the Vario EL III- C,H,N,S/O Elemental Analyzer (Elementar Analysensysteme GmbH, Hanau-Germany). Thin-layer chromatography (TLC) was performed on precoated Merck silica gel 60 F254 and RP-18 F254 plates. Column chromatography was performed on Lobar LichroPrep Si 60 (40-63 μm), RP-18 (40-63 μm) columns coupled to a Waters RI 401 detector, and on Biotage SP1 system with UV detector and FLASH 12+, FLASH 25+ or FLASH 40+ columns pre packed with KP-SIL [40-63 μm, pore diameter 6 nm (60 Å)], KP-C18-HS (40-63 μm, pore diameter 9 nm (90 Å) or KP-NH [40-63 μm, pore diameter 10 nm (100 Å)] as adsorbent. Compounds were analyzed for purity (HPLC) using a Waters 1525 HPLC dual pump system equipped with an Alltech, Select degasser system, and dual λ 2487 UV-VIS detector. For data processing, Empower software was used (methods A and B). Methods C and D: Agylent Technologies 1260 Liquid Chromatograph equipped with Quat Pump (G1311B), Injector (G1329B) 1260 ALS, TCC 1260 (G1316A) and Detector 1260 DAD VL+ (G1315C). For data processing, LC OpenLab CDS ChemStation software was used. For details, see Supporting Information.

1. Synthesis experiments

Each paragraph describing a synthesis experiment should begin with the name of the product and any structure number assigned to the compound in the Results and Discussions section. Thereafter, the compound should be identified by its structure number. Use of standard abbreviations or unambiguous molecular formulas for reagents and solvents, and of structure numbers rather than chemical names to identify starting materials and intermediates, is encouraged.

When a new or improved synthetic method is described, the yields reported in key experimental examples, and yields used for comparison with existing methods, should represent amounts of isolated and purified products, rather than chromategraphically or spectroscopically determined yields. Reactant quantities should be reported in weight and molar units and for product yields should be reported in weight units; percentage yields should only be reported for materials of demonstrated purity. When chromatography is used for product purification, both the support and solvent should be identified.

2. Microwave experiments

Reports of syntheses conducted in microwave reactors must clearly indicate whether sealed or open reaction vessels were used and must document the manufacturer and model of the reactor, the method of monitoring the reaction mixture temperature, and the temperature-time profile. Reporting a wattage rating or power setting is not an acceptable alternative to providing temperature data. Manuscripts describing work done with domestic (kitchen) microwave ovens will not be accepted except for studies where the unit is used for heating reaction mixtures at atmospheric pressure.

3. Compound characterization

The Journal upholds a high standard for compound characterization to ensure that substances being added to the chemical literature have been correctly identified and can be synthesized in known yield and purity by the reported preparation and isolation methods. For **all new** compounds, evidence adequate to establish both **identity** and **degree of purity** (homogeneity) must be provided.

Identity *- Melting point.* All homogeneous solid products (*e.g.* not mixtures of isomers) should be characterized by melting or decomposition points. The colors and morphologies of the products should also be noted.

Specific rotations. Specific rotations based on the equation $[\alpha]^t; D = (100\ \alpha) / (l\ c)$ should be reported as unitless numbers as in the following example: $[\alpha]^{20}$; D = -25.4 (c 1.93, $CHCl_3$), where c / g mL^{-1} is concentration and l / dm is path length. The units of the specific rotation, (deg mL) / (g dm), are implicit and are not included with the reported value.

Spectra/*Spectral Data.* Important IR adsorptions should be given.

For all new diamagnetic substances, NMR data should be reported (^1H, ^{13}C, and relevant heteronuclei). ^1H NMR chemical shifts should be given with two digits after the decimal point. Include the number of protons represented by the signal, signal multiplicity, and coupling constants as needed (*J* italicized, reported with up to one digit after the decimal). The number of bonds through which the coupling is operative, ^xJ, may be specified by the author if known with a high degree of certainty. ^{13}C NMR signal shifts should be rounded to the nearest 0.01 ppm unless greater precision is needed to distinguish closely spaced signals. Field strength should be noted for each spectrum, not as a comment in the general experimental section. Hydrogen multiplicity (C, CH, CH_2, CH_3) information obtained from routine DEPT spectra should be included. If detailed signal assignments are made, the type of NOESY or COSY methods used to establish atom connectivity and spatial relationships should be identified in the Supporting Information. Copies of spectra should also be included where structure assignments of complex molecules depend heavily on NMR interpretation. Numbering system used for assignments of signals should be given in the Supporting Information with corresponding general structural formula of named derivative.

HPLC/LCMS can be substituted for biochemistry papers where the main focus is not on compound synthesis.

HRMS/*elemental analysis.* To support the molecular formula assignment, HRMS data accurate within 5 ppm, or combustion elemental analysis [carbon and hydrogen (and nitrogen, if present)] data accurate within 0.5 %, should be reported for new compounds. HRMS data should be given in format as is usually given for combustion analysis: calculated mass for given formula following with observed mass: (+)ESI-HRMS *m*/*z*: [molecular formula + H]$^+$ calculated mass, observed mass. Example: (+)ESI-HRMS *m*/*z*: calculated for [$C_{13}H_8BrCl_2N + H^+$] 327.92899, observed 327.92792.

NOTE: in certain cases, a crystal structure may be an acceptable substitute for HRMS/elemental analysis.

Biomacromolecules. The structures of biomacromolecules may be established by providing evidence about sequence and mass. Sequences may be inferred from the experimental order of amino acid, saccharide, or nucleotide coupling, from known sequences of templates in enzyme-mediated syntheses, or through standard sequencing techniques. Typically, a sequence will be accompanied by MS data that establish the molecular weight.

Example: Product was isolated upon column chromatography [dry flash (SiO_2, eluent EA, EA/MeOH gradient 95/5 → 9/1, EA/MeOH/NH_3 gradient 18/0.5/0.5 → 9/1/1, and flash chromatography (Biotage SP1, RP column, eluent MeOH/H_2O gradient 75/25 → 95/5, N-H column, eluent EA/Hex gradient 6/3 → EA). was obtained after flash column chromatography (Biotage SP NH column, eluent hexane/EA 4:6 → 2:6). Yield 968.4 mg (95 %). Colorless foam softens at 96-101 °C. $[\alpha]^{20}$; D = +0,163 (c =2.0×10^{-3} g/mL, CH_2Cl_2). IR (ATR): 3376w, 2949m, 2868w, 2802w, 1731s, 1611w, 1581s, 1528m, 1452m, 1374s, 1331w, 1246s, 1171m, 1063w, 1023m, 965w, 940w, 881w, 850w, 807w, cm^{-1}. ^1H NMR (500 MHz, $CDCl_3$, δ): 8.46 (*d*, 1H, J = 5.4, H-2'), 7.89 (*s*, 1H, J = 2.0, H-8'), 7.71 (*d*, 1H, J = 8.9, H-5'), 7.30 (*dd*, 1H, J_1 = 8.8, J_2 = 2.1, H-6'), 6.33 (*d*, 1H, J = 5.4, H-3'), 6.07 (*s*, HN-Boc, exchangeable with D_2O), 5.06 (*s*, 1H, H-12), 4.92-4.88 (*m*, 1H, H-7), 4.42 (*bs*, H-3), 3.45 (*s*,CH_3-N), 3.33 (*bs*, H-9'), 3.05-2.95 (*m*, 2H, H-11'), 2.70-2.43 (*m*, 2H, H-24) and HN, exchangeable with D_2O), 2.07 (*s*, CH_3COO), 2.04 (*s*, CH_3COO), 1.42 (*s*, 9H, $(CH_3)_3$C-N(Boc)), 0.88 (*s*, 3H, CH_3-10), 0.79 (*d*, 3H,J = 6.6, CH_3-20), 0.68 (*s*, 3H, CH_3-13). ^{13}C NMR (125 MHz, $CDCl_3$, δ): 170.34, 170.27, 151.80, 149.92, 148.87, 134.77, 128.36, 125.11, 121.43, 117.29, 99.98, 75.41, 70.82, 50.43, 49.66, 47.60, 47.33, 44.97, 43.30, 41.83, 41.48, 37.65, 36.35, 35.44, 34.89,

34.19, 33.23, 31.24, 28.79, 28.35, 27.25, 26.45, 25.45, 22.74, 22.63, 21.57, 21.31, 17.85, 12.15. (+)ESI-HRMS (*m/z*): calculated for $[C_{45}H_{67}ClN_4O_6 + H]^+$ 795.48219, observed 795.48185. Combustion analysis for $C_{45}H_{67}ClN_4O_6$: Calculated. C 67.94, H 8.49, N 7.04; found C 67.72, H 8.63, N 6.75. HPLC purity: method A: RT 1.994, area 99.12 %; method C: RT 9.936, area 98.20 %.

Purity **-** Evidence for documenting compound purity should include one or more of the following:

a) Well-resolved high field 1D ^{1}H NMR spectrum showing at most only trace peaks not attributable to the assigned structure and a standard 1D proton-decoupled ^{13}C NMR spectrum. Copies of the spectra should be included as figures in the Supporting Information.
b) Quantitative gas chromatographic analytical data for distilled or vacuum-transferred samples, or quantitative HPLC analytical data for materials isolated by column chromatography or separation from a solid support. HPLC analyses should be performed in two diverse systems. The stationary phase, solvents (HPLC), detector type, and percentage of total chromatogram integration should be reported; a copy of the chromatograms may be included as a figure in the Supporting Information.
c) Electrophoretic analytical data obtained under conditions that permit observing impurities present at the 5 % level.

HRMS data may be used to support a molecular formula assignment **but cannot be used as a criterion of purity**.

4. Biological Data

Quantitative biological data are required for all tested compounds. Biological test methods must be referenced or described in sufficient detail to permit the experiments to be repeated by others. Detailed descriptions of biological methods should be placed in the experimental section. Standard compounds or established drugs should be tested in the same system for comparison. Data may be presented as numerical expressions or in graphical form; biological data for extensive series of compounds should be presented in tabular form. Tables consisting primarily of negative data will not usually be accepted; however, for purposes of documentation they may be submitted as supporting information. Active compounds obtained from combinatorial syntheses should be resynthesized and retested to verify that the biology conforms to the initial observation.

Statistical limits (statistical significance) for the biological data are usually required. If statistical limits cannot be provided, the number of determinations and some indication of the variability and reliability of the results should be given. References to statistical methods of calculation should be included. Doses and concentrations should be expressed as molar quantities (*e.g.*, mol/kg, μmol/kg, M, mM). The routes of administration of test compounds and vehicles used should be indicated, and any salt forms used (hydrochlorides, sulfates, *etc.*) should be noted. The physical state of the compound dosed (crystalline, amorphous; solution, suspension) and the formulation for dosing (micronized, jet-milled, nanoparticles) should be indicated. For those compounds found to be inactive, the highest concentration (*in vitro*) or dose level (*in vivo*) tested should be indicated.

- The RESULTS AND DISCUSSION should include concisely presented results and there significance discussed and compared to relevant literature data. The results and discussion may be combined or kept separate.
- The inclusion of a CONCLUSION section, which briefly summarizes the principal conclusions, is recommended.
- NOMENCLATURE is optional but, if the authors wish, a list of employed symbols may be included.
- REFERENCES should be numbered sequentially as they appear in the text. Please note that any reference numbers appearing in the Illustrations and/or Tables and corresponding captions must follow the numbering sequence of the paragraph in which they appear for the first time. When cited, the reference number should be superscripted in Font 12, following any punctuation mark. In the reference list, they should be in normal position followed by a full stop. Reference entry must not be formatted using Carriage returns (enter key; ↵ key) or multiple space key. The formatting of references to published work should follow the *Journal'*s style as follows:

Journals[a]:	A. B. Surname1, C. D. Surname2, *J. Serb. Chem. Soc.* **Vol** (Year) first page Number (https://doi.org/doi)[b]
Books:	A. B. Surname1, C. D. Surname2, *Name of Book*, Publisher, City, Year, pp. 100-101 (https://doi.org/doi)[b]
Compilations:	A. B. Surname1, C. D. Surname2, in *Name of Compilation*, A. Editor1, C. Editor2, Ed(s)., Publisher, City, Year, p. 100 (https://doi.org/doi)[b]
Proceedings:	A. B. Surname1, C. D. Surname2, in *Proceedings of Name of the Conference or Symposium*, (Year), Place of the Conference, Country, *Title of the Proceeding*, Publisher, City, Year, p. or Abstract No. 100
Patents:	A. B. Inventor1, C. D. Inventor2, (Holder), Country Code and patent number (registration year)
Chemical Abstracts:	A. B. Surname1, C. D. Surname2, *Chem. Abstr*. CA 234 567a; For non-readily available literature, the Chemical Abstracts reference should be given in square brackets: [C.A. 139/2003 357348t] after the reference
Standards:	EN ISO 250: *Name of the Standard* (Year)
Websites:	Title of the website, URL in full (date accessed)

[a] When citing Journals, the International Library Journal abbreviation is required. Please consult, *e.g.*, https://images.webofknowledge.com/WOK46/help/WOS/A_abrvjt.html

[b] doi should be replaced by doi number of the Article, for example: http://dx.doi.org/10.2298/JSC161212085B (as active link). If doi do not exist, provide the link to the online version of the publication.

Only the last entry in the reference list should end with a full stop.

The names of all authors should be given in the list of references; the abbreviation *et al.* may only be used in the text. The original journal title is to be retained in the case of publications published in any language other than English (please denote the language in parenthesis after the reference). Titles of publications in non-Latin alphabets should be transliterated. Russian references are to be transliterated using the following transcriptions:

ж→zh, х→kh, ц→ts, ч→ch, ш→sh, щ→shch, ы→y, ю→yu, я→ya, э→e, й→i, ь→’.

Supplementary material

Authors are encouraged to presentthe information and results non-essential to the understanding of their paper as SUPPLEMENTARY MATERIAL (can be uploaded inStep 4 of Online Submission). This material may include as a rule, but is not limited to, the presentation of analytical and spectral data demonstrating the identity and purity of synthesized compounds, tables containing raw data on which calculations were based, series of figures where one example would remain in the main text, etc. The Editorial Board retain the right to assign such information and results to the Supplementary material when deemed fit. Supplementary material does not appear in printed form but can be downloaded from the web site of the JSCS.

Mathematical and chemical equations should be given in separate lines and must be numbered, Arabic numbers, consecutively in parenthesis at the end of the line. All equations should be embedded in the text Complex equations (fractions, integrals, matrix…) should be prepared with the aid of the Microsoft Equation 3.0 (or higher) or MathType (Do not use them to create simple equations and labels). Using the Insert -> Equation option, integrated in MS Office 2010 and MS Office 2013, as well as insertion of equation objects within paragraph text IS NOT ALLOWED.

Deposition of crystallographic data

Prior to submission, the crystallographic data included in a manuscript presenting such data should be deposited at the appropriate database. Crystallographic data associated with organic and metal-organic structures should be deposited at the Cambridge Crystallographic Data Centre (CCDC) by e-mail to **deposit@ccdc.cam.ac.uk**

Crystallographic data associated with inorganic structures should be deposited with the Fachinformationszentrum Karlsruhe (FIZ) by e-mail to **crysdata@fiz-karlsruhe.de**. A deposition number will then be provided, which should be added to the reference section of the manuscript.

For detailed instructions please visit the JSCS website:
https://www.shd-pub.org.rs/index.php/JSCS/Instructions

ARTWORK INSTRUCTIONS

JSCS accepts only **TIFF** or **EPS** formats, as well as **JPEG** format (only for colour and greyscale photographs) for electronic artwork and graphic files. **MS files** (Word, PowerPoint, Excel, Visio) **NOT acceptable**. Generally, scanned instrument data sheets should be avoided. Authors are responsible for the quality of their submitted artwork. Every single Figure or Scheme, as well as any part of the Figure (A, B, C…) should be prepared according to following instructions (every part of the figure, A, B, C…, must be submitted as an independent single graphic file):

TIFF

Virtually all common artwork and graphic creation software is capable of saving files in TIFF format. This 'option' can normally be found under 'the 'Save As...' or 'Export...' commands in the 'File' menu.

TIFF (Tagged Image File Format) is the recommended file format for bitmap, greyscale and colour images.

- Colour images should be in the RGB mode
- When supplying TIFF files, please ensure that the files are supplied at the correct resolution:
    1. Line artwork: minimum of 1000 dpi
    2. RGB image: minimum of 300 dpi
    3. Greyscale image: minimum of 300 dpi
    4. Combination artwork (line/greyscale/RGB): minimum of 500 dpi
- Images should be tightly cropped, without frame and any caption.
- If applicable please re-label artwork with a font supported by JSCS (Arial, Helvetica, Times, Symbol) and ensure it is of an appropriate font size.
- Save an image in TIFF format with LZW compression applied.
- It is recommended to remove Alpha channels before submitting TIFF files.
- It is recommended to flatten layers before submitting TIFF files.

Please be sure that quality of an image cannot be increased by changing the resolution from lower to higher, but only by rescanning or exporting the image with higher resolution, which can be set in usual "settings" facilities.

EPS

Virtually all common artwork creation software, such as Canvas, ChemDraw, CorelDraw, SigmaPlot, Origin Lab…, are capable of saving files in EPS format. This 'option' can normally be found under the 'Save As...' or 'Export...' commands in the 'File' menu.

For vector graphics, EPS (Encapsulated PostScript) files are the preferred format as long as they are provided in accordance with the following conditions:

- when they contain bitmap images, the bitmaps should be of good resolution (see instructions for TIFF files)
- when colour is involved, it should be encoded as RGB
- an 8-bit preview/header at a resolution of 72 dpi should always be included
- embed fonts should always included and only the following fonts should be used in artwork: Arial, Helvetica, Times, Symbol
- the vertical space between the parts of an illustration should be limited to the bare necessity for visual clarity
- no data should be present outside the actual illustration area
- line weights should range from 0.35 pt to 1.5 pt
- when using layers, they should be reduced to one layer before saving the image (Flatten Artwork)

JPEG

Virtually all common artwork and graphic creation software is capable of saving files in JPEG format. This 'option' can normally be found under 'the 'Save As...' or 'Export...' commands in the 'File' menu.

JPEG (Joint Photographic Experts Group) is the acceptable file format **only for colour and greyscale photographs**. JPEG can be created with respect to photo quality (low, medium, high; from 1 to 10), ensuring file sizes are kept to a minimum to aid easy file transfer. Images should have a minimum resolution of 300 dpi. Image width: minimum 3.0 cm; maximum 12.0 cm.

Please be sure that quality of an image cannot be increased by changing the resolution from lower to higher, but only by rescanning or exporting the image with higher resolution, which can be set in usual "settings" facilities.

SIZING OF ARTWORK

- JSCS aspires to have a uniform look for all artwork contained in a single article. Hence, it is important to be aware of the style of the journal.
- Figures should be submitted in black and white or, if required, colour (charged). If coloured figures or photographs are required, this must be stated in the cover letter and arrangements made for payment through the office of the Serbian Chemical Society.
- As a general rule, the lettering on an artwork should have a finished, printed size of 11 pt for normal text and no smaller than 7 pt for subscript and superscript characters. Smaller lettering will yield a text that is barely legible. This is a rule-of-thumb rather than a strict rule. There are instances where other factors in the artwork, (for example, tints and shadings) dictate a finished size of perhaps 10 pt. Lines should be of at least 1 pt thickness.
- When deciding on the size of a line art graphic, in addition to the lettering, there are several other factors to address. These all have a bearing on the reproducibility/readability of the final artwork. Tints and shadings have to be printable at the finished size. All relevant detail in the illustration, the graph symbols (squares, triangles, circles, *etc*.) and a key to the diagram (to explain the explanation of the graph symbols used) must be discernible.
- The sizing of halftones (photographs, micrographs,...) normally causes more problems than line art. It is sometimes difficult to know what an author is trying to emphasize on a photograph, so you can help us by identifying the important parts of the image, perhaps by highlighting the relevant areas on a photocopy. The best advice that can be given to graphics suppliers is not to over-reduce halftones. Attention should also be paid to magnification factors or scale bars on the artwork and they should be compared with the details inside. If a set of artwork contains more than one halftone, again please ensure that there is consistency in size between similar diagrams.

General sizing of illustrations which can be used for the Journal of the Serbian Chemical Society:

- Minimum fig. size: 30 mm width
- Small fig. size - 60 mm width
- Large fig. size - 90 mm width
- Maximum fig. size - 120 mm width

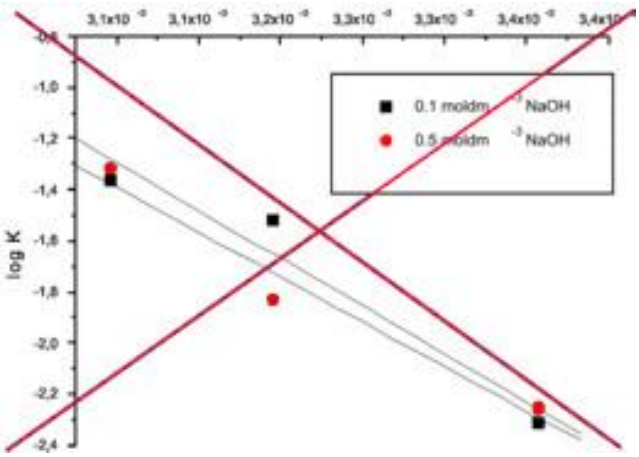
Pixel requirements (width) per print size and resolution for bitmap images:

	Image width	A	B	C
Minimal size	30 mm	**354**	**591**	**1181**
Small size	60 mm	**709**	**1181**	**2362**
Large size	90 mm	**1063**	**1772**	**3543**
Maximal size	120 mm	**1417**	**2362**	**4724**

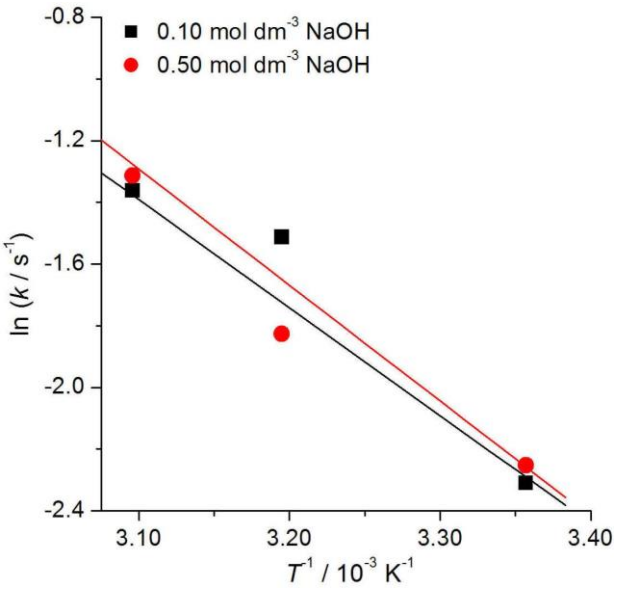
A: 300 dpi > RGB or Greyscale image
B: 500 dpi > Combination artwork (line/greyscale/RGB)
C: 1000 dpi> Line artwork

The designation of physical quantities and graphs formatting

The designation of physical quantities on figures must be in italic, whereas the units are in upright letters. They should be in Times New Roman font. In graphs a slash should be used to separate the designation of a physical quantity from the unit (example: p / kPa, t °C, T_0 / K, τ / h, ln (j / mA cm^{-2})…). Designations such as: p (kPa), t [min]…, are not acceptable. However, if the full name of a physical quantity is unavoiable, it should be given in upright letters and separated from the unit by a comma (example: Pressure, kPa, Temperature, K…). Please do not use the axes of graphs for additional explanations; these should be mentioned in the figure captions and/or the manuscript (example: "pressure at the inlet of the system, kPa" should be avoided). The axis name should follow the direction of the axis (the name of y-axis should be rotated by 90°). Top and right axes should be avoided in diagrams, unless they are absolutely necessary. Decimal numbers must have decimal points and not commas in the axis labels in graphical presentations of results. Thousands are separated, if at all, by a comma and not a point.



INCORRECT



CORRECT

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS

J. Serb. Chem. Soc. Vol. 86, No. 7–8 (2021)

CONTENTS*

Published by the Serbian Chemical Society
Karnegijeva 4/III, P.O. Box 36, 11120 Belgrade, Serbia
Printed by the Faculty of Technology and Metallurgy
Karnegijeva 4, P.O. Box 35-03, 11120 Belgrade, Serbia

* For colored figures in this issue please see electronic version at the Journal Home Page: http://www.shd.org.rs/JSCS/

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 86 (7–8) 625–637 (2021)
JSCS–5449
Original scientific paper

Quantitative structure–activity relationship modelling of influenza M2 ion channels inhibitors

IVANKA G. STANKOVA[1]*, RADOSLAV L. CHAYROV[1], MICHAELA SCHMIDTKE[2], DANCHO L. DANALEV[3]**, LIUDMILA N. OGNICHENKO[4], ANATOLY G. ARTEMENKO[4], VALERY A. SHAPKIN[5] and VICTOR E. KUZ'MIN[4]

[1]Department of Chemistry, South-West University "NeofitRilski", Blagoevgrad, 2700, Bulgaria, [2]Friedrich Schiller University, Department of Virology and Antiviral Therapy, Jena, 207745, Germany, [3]University of Chemical Technology and Metallurgy, Biotechnology Department, 1756 Sofia, 8 blvd. Kliment Ohridski, Bulgaria, [4]A.V. Bogatsky Physico-Chemical Institute of Ukrainian National Academy of Sciences, Department of Molecular Structure and Chemoinformatics, 86, Lustdorfskaya doroga, Odessa, 65080, Ukraine and [5]Department of Department of Theoretical Foundations of Chemistry, Odessa National Polytechnic University, 1, Shevchenko ave., Odessa 65044, Ukraine

(Received 9 May 2020, revised 15 March, accepted 5 May 2021)

Abstract: A series of adamantane derivatives (rimantadine and amantadine) incorporating amino-acid residues are investigated by simplex representation of molecular structure (SiRMS) approach in order to found correlation between chemical structures of investigated compounds and obtained data for antiviral activity and cytotoxicity. The obtained data from QSAR analysis show that adamantane derivatives containing amino acids with short aliphatic non-polar residues in the lateral chain will have good antiviral activity against the tested virus A/H3N2, strain Hong Kong/68 with low cytotoxicity. QSAR experiments and *in vitro* data also show good correlation and reveal that modified adamantine derivatives including guanidated in the lateral chain amino acid and *β*-amino acids as substituents show low to none activity.

Keywords: QSAR study; molecular simplex; adamantine derivatives; amantadine; rimantadine; amino acids.

INTRODUCTION

Adamantane derivatives have been used successfully for the prevention and treatment of influenza A virus infection for more than 30 years.[1,2] The aminoadamantanes (amantadine and rimantadine) block M2 proton channel and thus stop virus replication. However, they are no longer effective because of widespread drug resistance. S31N is the predominant and amantadine-resistant M2 mutant, present in almost all of the circulating influenza A strains as well as in the pan-

*,** Corresponding authors. E-mail: (*)ivastankova@abv.bg; (**)ddanalev@uctm.edu
https://doi.org/10.2298/JSC200509036S

demic 2009 H1N1 and the highly pathogenic H5N1 flu strains.[3] Structural and biochemical studies of the S31N mutant showed that replacing ^{31}Ser, which is located in the helix–helix interface, with the bulkier Asn results insubstantially weaker helix–helix packing. Since the pocket is composed of residues from two adjacent TM helices, the stability and physical properties of the pocket depend on the dynamics and conformation of helical packing.[4] The M2 proton channel from influenza A virus, a prototype for a class of viral ion channels known as viroporins, conducts protons along a chain of water molecules and ionizable side-chains, including ^{37}His. Drugs inhibit proton conduction by binding to an aqueous cavity adjacent to M2's proton-selective filter, thereby blocking access of proton to the filter, and altering the energetic landscape of the channel and the energetics of proton-binding to ^{37}His. According to Gaiday *et al.* studied cage compounds inhibit the M2 ion channel by binding to the ^{37}His residue. The adamantane cage fits into a pocket formed by ^{41}Trp residue, while the hydrogen bond is formed between hydrogen atom of ammonium nitrogen and the nitrogen of histidine residue.[6] One of the possible approaches to restore the antiviral properties of this class of compounds is to incorporate more than one functional group in their molecule. Amino acids and peptides are promising alternative for such kind of modification because of their multi functionality.

Herein, we report the QSAR analysis of some adamantanes modified with different substituents and the role of specific modifications on the biological activity.

EXPERIMENTAL

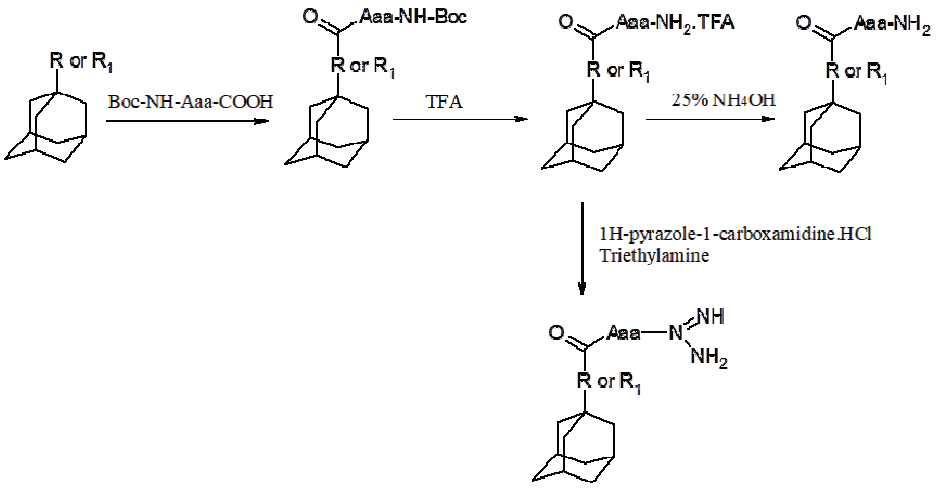
Chemical synthesis

Compounds aimed to this study were previously synthesized according to Scheme 1.

The synthetic protocol of target compounds was as follow: 3 mmol of 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate (TBTU) was dissolved in 15 mL CH_2Cl_2. Further, the corresponding Boc-$^{\alpha}$N-protected amino acid (3 mmol) and DIPEA (3.1 mmol) were added to a solution of TBTU. The obtained mixture was stirred at room temperature for 30 min and 3 mmol of rimantadine or amantadine and 3 mmol *N*,*N*-dimethylaminopyridine (DMAP) were also added. This mixture was stirred at room temperature for another 3 h, and then evaporated to dryness. After evaporation the obtained oil was purified by flash chromatography in system ethyl acetate:*n*-hexan (50:50 volume ratio).

1 eq. of Boc-$^{\alpha}$N-Aaa-amantadine or Boc-$^{\alpha}$N-Aaa-rimantadine was dissolved in 10-fold excess of trifuoroacetic acid (TFA) at 0 °C. The reaction mixture was stirred until fully deprotection of Boc-group under chromatographic control in systems chloroform-methanol (95:5 volume ratio). The excess of TFA evaporated and the obtained oil was dissolved in 10 mL methanol. Further 25 % ammonium hydroxide was added until the pH reached around 8. The solvent was evaporated under vacuum. Tha obtained crystals were dissolved in ethyl acetate and washed with 3×25 mL water. The organic layers were combined, dried on anhydrous Na_2SO_4 and solvent removed under vacuum. The yield of all compounds as well as ^{1}H- and ^{13}C-NMR analysis confirming their structures are given in Chayrov *et al.*[17]

For guanidated analogues synthesis the following procedure was used.



Scheme 1. Synthesis of amino acid derivatives of adamantine; R = $–NH_2$ is amantadine (Am), R = CH_3CHNH_2 is rimantadine (Rim), Aaa = amino acid; for amantadine Aaa = Ala; Phe; Phe(4-F); Val; for rimantadine Aaa = Ala; Gly; Ile; Leu; Phe; D-Phe(4-F); L-Phe(4-F); Val; beta-Ala; Tyr.

0.67 mmol of Boc-deprotected Aaa-rimantadine or Aaa-amantadine was dissolved in 2 mL of acetonitrile. Further to the obtained solution 1.00 mmol of 1*H*-pyrazole-1-carboxamidine hydrochloride and 2.0 mmol triethylamine were added. The reaction was run for 48 h at room temperature. At the end of the reaction time the solvent was evaporated under vacuum. The obtained crude oil was dissolved in 25 mL chloroform and washed several times with 5 % $NaHSO_4$ (pH 3). The organic layers were combined, dried with anhydrous Na_2SO_4 and solvent was evaporated under vacuum. The residue was crystalized in methanol/diethyl ether.

Biological studies

The results obtained by the realized biological tests[17] are presented in Table I.

TABLE I. Data of biological tests of studied compounds; *CC* is cytotoxic concentration; *HNTC* is high-nontoxic concentration

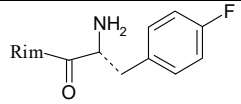
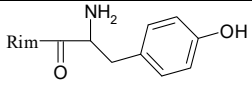
No.	Structural formula of compound (abbreviation)	CC_{50} / μM MDCK cells		*HNTC* / μM MDCK cells		IC_{50} / μM Against A/Hong Kong/68	log (IC_{50} / μM)
		Average	Class numb.	Average	Class numb.	Average	–
1	 (D-Phe(4-F)-Rim)	19.91	**1**	6.23	**1**	100.00	**2.00**
2	 (L-Tyr-Rim)	> 100	**0**	9.97	**1**	0.52	**–0.28**

TABLE I. Continued

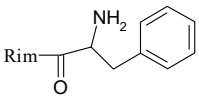
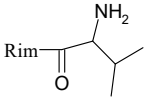
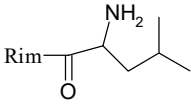
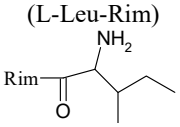
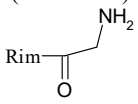
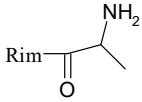
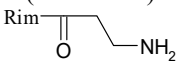
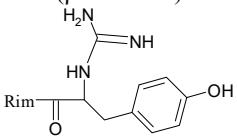
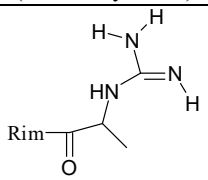
No.	Structural formula of compound (abbreviation)	CC_{50} / μM MDCK cells		*HNTC* / μM MDCK cells		IC_{50} / μM Against A/Hong Kong/68	log (IC_{50} / μM)
		Average	Class numb.	Average	Class numb.	Average	–
3	 (L-Phe-Rim)	21.93	**1**	4.75	**1**	100.00	**2.00**
4	 (L-Val-Rim)	70.33	**1**	10.20	**1**	100.00	**2.00**
5	 (L-Leu-Rim)	> 100	**0**	11.66	**1**	0.75	**–0.12**
6	 (L-Ile-Rim)	58.35	**1**	5.33	**1**	100.00	**2.00**
7	 (L-Gly-Rim)	> 100	**0**	13.89	**1**	0.11	**–0.96**
8	 (L-Ala-Rim)	> 100	**0**	14.65	**1**	1.53	**0.18**
9	 (β-Ala-Rim)	> 100	**0**	20.78	**1**	15.72	**1.20**
10	 (Gua-L-Tyr-Rim)	> 100	**0**	36.69	**1**	8.19	**0.91**
11	 (Gua-L-Ala-Rim)	> 100	**0**	> 100	**0**	41.90	**1.62**

TABLE I. Continued

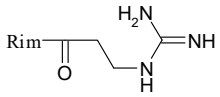
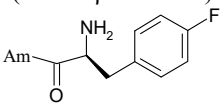
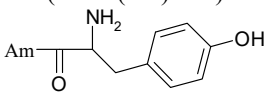
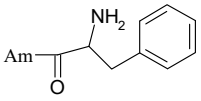
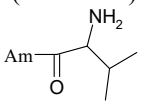
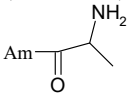
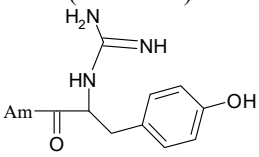
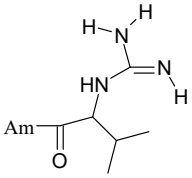
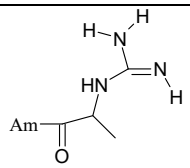
No.	Structural formula of compound (abbreviation)	CC_{50} / μM MDCK cells		*HNTC* / μM MDCK cells		IC_{50} / μM Against A/Hong Kong/68	log (IC_{50} / μM) –
		Average	Class numb.	Average	Class numb.	Average	
12	 (Gua-L-β-Ala-Rim)	> 100	**0**	> 100	**0**	100.00	**2.00**
13	 (L-Phe(4-F)-Am)	83.61	**1**	15.35	**1**	0.81	**–0.09**
14	 (L-Tyr-Am)	> 100	**0**	> 100	**0**	3.93	**0.59**
15	 (L-Phe-Am)	89.40	**1**	15.10	**1**	0.75	**–0.12**
16	 (L-Val-Am)	> 100	**0**	> 100	**0**	1.32	**0.12**
17	 (L-Ala-Am)	> 100	**0**	> 100	**0**	1.41	**0.15**
18	 (Gua-L-Tyr-Am)	> 100	**0**	> 100	**0**	59.71	**1.78**
19	 (Gua-L-Val-Am)	> 100	**0**	> 100	**0**	17.13	**1.23**

TABLE I. Continued

No.	Structural formula of compound (abbreviation)	CC_{50} / μM		*HNTC* / μM		IC_{50} / μM	log (IC_{50} / μM)
		MDCK cells		MDCK cells		Against A/Hong Kong/68	
		Average	Class numb.	Average	Class numb.	Average	–
20	 (Gua-L-Ala-Am)	> 100	**0**	> 100	**0**	18.37	**1.26**
21	Am-H	> 100	**0**	n/a	**0**	0.39	**–0.41**
22	Rim-H	–	**–**	–	**–**	0.06	**–1.25**

QSAR calculations

In the present study the simplex representation of molecular structure (SiRMS) approach[7,8] was used for calculation of structural descriptors for all investigated compounds. The simplex approach is based on isolating and counting the number of molecular fragments (pairs, triples, quadruples of atoms).

Any molecule in the framework of SiRMS can be represented as a system of different specific fragments (simplexes) of fixed composition and topology. Various atomic characteristics can be used for the vertex differentiation in the simplex, such as the uniqueness of the atom (atom nature or a more detailed type), electronegativity, partial charge, lipophilicity, electronic polarizability (refraction), H-bond donor/acceptor potential, van der Waals interactions, *etc.* For atomic characteristics having real values (electronegativity, lipophilicity *etc*.) at the preliminary stage the range of values is divided into a certain number of groups G (G is a tuning parameter of models and can vary, as a rule from 3 to 7; see the example showing how using vertexes differentiation by atomic charges in Supplementary materials).

In addition, electronegativity, refraction, molecular weight and octan-1-ol–water partition coefficient (Log *P*) are calculated as integral characteristics that describe the whole molecule.

The calculation of descriptors was carried out at the 2D level of molecular structure representation. The 2D-QSAR models are the most popular in structure-property studies.[7,8] In this case only molecular topology is taken into account, i.e. all information is extracted from the structural formula.

The relationships between the calculated molecular descriptors and the studied properties of investigated molecules were established by methods of partial least squares (PLS)[9] and random forest (RF).[10]

The removal of highly correlated and constant descriptors, the trend vector method,[11] genetic algorithm (GA)[12] and the automatic variable selection (AVS) strategy[7] have all been used to select the descriptors in PLS. The removal of highly correlated descriptors is not necessary for PLS analysis since descriptors are reduced to a series of uncorrelated latent variables. This procedure frequently helps to obtain more adequate models. During this procedure one descriptor from each pair having a pair correlation coefficient r satisfying $|r|>0.90$ is eliminated. It was previously discovered[13] that descriptors involved in the best Trend Vector models (several decades of models with approximately identical quality) form a good subset for their subsequent usage in PLS. The noise elimination can be one of the more probable

explanations of the success of the trend vector procedure. GA is used as a tool for the selection of adequate PLS models. We used the small set of following algorithm parameters as mutation rate = 0.3, crossingover rate = 0.7, type of crossingover is double. Descriptors, from the best model obtained by a preliminary AVS procedure, are normally used as the starting "population". In this work we run GA only once. The GA is definitely not a tool for the elucidation of the global maximum or minimum, and very often subsequent AVS procedures and different enumerative techniques allow one to increase the quality of the PLS models obtained.

According to the QSAR/QSPR OECD principles,[14] applicability domain (AD) of developed models was estimated. A ellipsoid model of structural space,[7] Williams plot[15] and tree AD approach[16] were used for AD estimation.

Within the framework of SiRMS approach it is possible to define the relative influence of the different physical and chemical factors on the character of the molecules interaction with the biological target.[7,8] For this purpose it is necessary to sum and compare absolute values of PLS regression coefficients of descriptors for all groups used for atom differentiation.

The clear interpretation is one of the advantages of SiRMS approach.[8] On the basis of developed QSAR models the influence of each atom over a particular property can be calculated. The contribution of each atom in the molecule can be defined as the ratio of the sum of PLS regression coefficients for all simplexes containing this atom to the number of atoms in the simplex. The atomic contribution depends on the number of simplexes that include this atom. The number of simplexes is not constant. It varies in different molecules depending on other constituents. Thus, this contribution is non-additive. The analysis of such information allows selecting different fragments which have negative or positive influence on a considered property.

The descriptors calculation and data analysis were performed using HiTQSAR software,[7] which was developed in the department of molecular structure and chemoinformatics of the A. V. Bogatsky Physico-chemical institute of National Academy of Sciences of Ukraine.

RESULTS AND DISCUSSION

During the first step of synthetic scheme Boc-$^{\alpha}$N-protected amino acid are bonded to rimantadine or amantadine. Next step is reaction of Boc-group deprotection with treatment with 10 fold excess of TFA. Finally target compounds are obtained as a free bases on their amino functions by treatment with 25 % ammonium hydroxide till pH around 8.

Thus using structures and results from biological investigations of target rimantadine and amantadine analogues QSAR analysis was carried out and the effects of the substituent on the *in vitro* cytotoxicity (CC_{50}), *HNTC* and antiviral activity (against human influenza virus A/H3N2 strain/Hong Kong/68) were also investigated.[17]

In this study we used a dataset consisting of 22 derivatives (13 are rimantadine derivatives and 9 – amantadine derivatives, Table I).

In addition, the derivatives Tyr-rimantadine, Ala-rimantadine, *β*-ala-Rimantadine, and Ala-amantadine, Val-amantadine, which exhibit the lowest cytotoxicity were guanidated using 1*H*-pyrazole-1-carboxamidine reagent (Scheme 1).[18]

All compounds were tested against human influenza virus A/H3N2 strain Hong Kong/68 and their activity is described by Chayrov *et al.*[17] The antiviral

activity against influenza virus strain A/H3N2 strain Hong Kong/68 *in vitro* of new analogs of amantadine and rimantadine conjugated with amino acids reveal that the highest antiviral activity combined with low cytotoxicity was demonstrated by the rimantadine derivative conjugated with the simplest in structure glycine. Moreover, glycyl-rimantadine presented a high stability profile after incubation in human plasma for 24 h. Interestingly, the analogues of amantadine with the amino acids L-phenylalanine and L-(4-F)-phenylalanine exhibited high activity although lower than those of the amantadine in the same concentration. In addition, amantadine and rimantadine analogues conjugated with guanidine showed low toxicity but they also exhibited low activity. The various aliphatic and aromatic amino acids as substitutes in the rimantadine molecule have no significant effect on antiviral activity. The use of β-alanine reduces antiviral activity. Guanidation of the rimantadine and amantadine analogues do not increase antiviral activity unlike guanidated oseltamivir analogues.

Using SiRMS approach for calculation of structural descriptors a total of about 2000 different structural characteristics were calculated for the investigated compounds.

The ranges of dividing the atom properties into intervals are, as already mentioned above, tuning characteristics and the following schemes were used in the calculation of simplex descriptors: electronegativity: A < 2.19 ≤ B < 2.5 ≤ C < 3 ≤ ≤ D, refraction: A < 1.5 ≤ B < 3 ≤ C < 8 ≤ D, atomic charge: A < –0.16 ≤ B < < –0.10 ≤ C < –0.04 ≤ D < 0.01 ≤ E < 0.07 ≤ F < 0.13 ≤ G, lipophilicity: A < –1.51 ≤ B < –0.96 ≤ C < –0.42 ≤ D < 0.13 ≤ E < 0.68 ≤ F < 1.23 ≤ G, VDW attraction: A < 50 ≤ B < 100 ≤ C < 250 ≤ D < 400 ≤ E < 650 ≤ F < 2000 ≤ G, VDW repulsion: A < 20.000 ≤ B < 32.000 ≤ C < 50.000 ≤ D < 100.000.

All atoms corresponding to the simplex vertices were also divided into three groups: A – acceptors of potential H-bond, D – donors of potential H-bond and I – indifferent ones.

On the first stage of this study, 2D QSAR model reflecting the structural influence of investigated compounds on their antiviral activity against human influenza (virus A/H3N2 strain Hong Kong/68) was developed. This model (Model A1) was used for interpretation, *i.e.*, for estimation of influence of different structural factors on investigated activity (on value of log IC_{50}), but this model was not intended for prediction of activity. Model A1 was built using the PLS method with two latent variables (based on 9 descriptors in the final) and with the following statistical characteristics: determination coefficient for the training set R^2 = = 0.88, coefficient of determination for cross-validation (leave-one-out) Q^2 = 0.77, standard error S(ws) = 0.35, S(cv) = 0.52 for training set and cross-validation, respectively. A "randomization" procedure (Y-Scrambling) was used to confirm the "non-randomness" of the developed QSAR model.[7] The statistical characteristics obtained using the Y-Scrambling procedure were lower in indices than in

the final model: R^2 (Y-scr) = 0.26 ± 0.03, Q^2 (Y-scr) = 0.10 ± 0.03. Thus, the non--randomness of the established relationship between the structure of the investigated compounds and their antiviral activity can be stated (see Supplementary material to this paper).

The analysis of obtained Model A1 testifies that, electrostatic (atomic charges) factors and lipophilicity have the largest influence on changing of antiviral activity (41 and 37 %, respectively). The relative influence of atom nature is 22 %.

As already mentioned above, on the basis of developed QSAR models the contribution of each atom in the molecule can be calculated. The analysis of such information allows selecting different fragments which have negative or positive influence on a considered property.

The sequence of mean relative influence of various substituents (R) for amantadine and rimantadine derivatives on antiviral activity (log IC_{50}) is shown on Fig. 1.

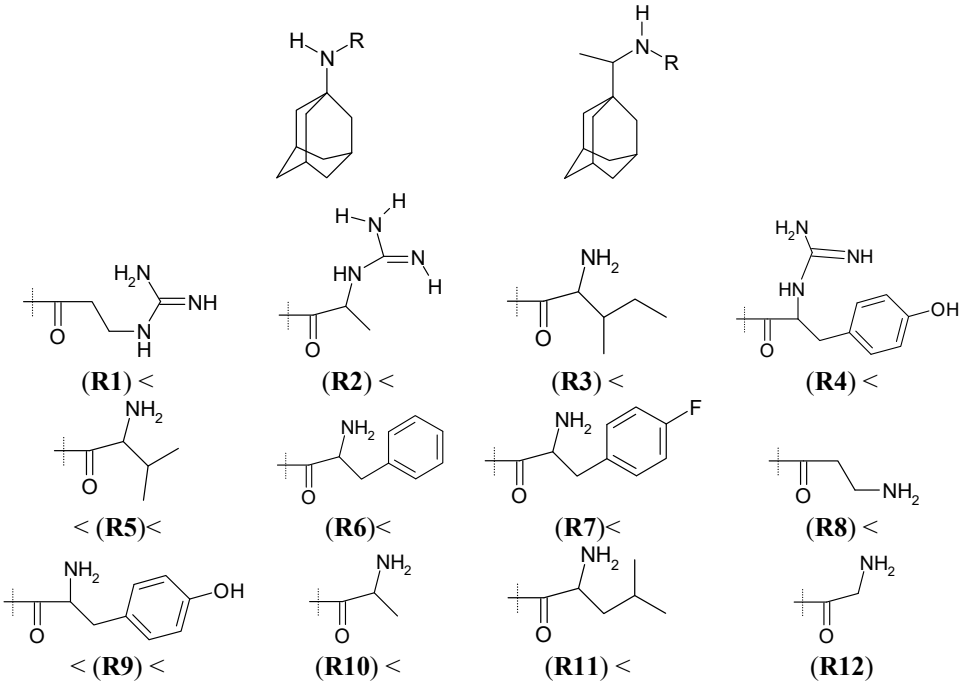


Fig. 1. Averaged relative influence of substituents (R) for amantadine and rimantadine derivatives on antiinfluenza activity (log IC_{50}) by the Model A1.

In addition, the influence of identical substituents for amantadine and rimantadine derivatives was estimated separately:

a) for amantadine: R2 < R4 < R5 < R6 <R7 < R9 < R10;

b) for rimantadine:R2 < R5 < R6 <R7 < R4 < R10 < R9.

In general, the sequence of changes in the relative influence of the substituents is the same for derivatives of amantadine and rimantadine, a noticable difference can only be noted for hydroxyl-containing substituents R4 and R9. Averaged relative influence of Rim- and Am-centers on antiinfluenza activity by the Model A1 equals 0.75 and –0.19, respectively.

RF model was developed for estimation of the predictive ability of the QSAR model. RF models were constructed according to the described original RF algorithm.[10] RF is an ensemble of single decision trees.[16] Each tree has been grown as: a) A bootstrap sample, which will be a training set for the current tree, is produced from the whole training set of N compounds. Compounds which are not in the current tree training set are placed in an out-of-bag (OOB) set (OOB set size is ~N/3); b) the best split by classification and regression tree (CART) algorithm[10] among the m randomly selected descriptors from whole set of M ones in each node is chosen. The value of m is just one tuning parameter for which RF models are sensitive; c) each tree is grown to the largest possible extent (there is no pruning).

RF model (Model A2) was built using the descriptor's set from Model A1. Obtained model (trees count = 50, randomly selected descriptors count = 3) has following statistical characteristics: determination coefficient for the training set $R^2_{(ws)}$ = 0.95, coefficient of determination for out-of-bag (OOB) set $R^2_{(OOB)}$ = = 0.65, $RMSE_{(ws)}$ = 0.26, $RMSE_{(oob)}$ = 0.62. The characteristics of the OOB just reflect the predictive ability of the QSAR model, which can be considered satisfactory. The greater accuracy of prediction is observed for compounds with medium activity (see Supplementary material).

Further in this study, QSAR models reflecting the structural influence of the investigated compounds on their cytotoxicity (CC_{50}) and *HNTC* were developed. The cytotoxicity data (CC_{50}) show (Table I) that among the 21 compounds, only 6 exhibit cytotoxicity (CC_{50} values are less than 100). For *HNTC* data were found that among the 21 compounds, only 12 exhibit activity (property values are less than 100).

Thus, the original activity values were coded as follows: 1-active (property < < 100) and 0-inactive (property > 100). In these cases RF method was used for decision of classification tasks.

For the first case since the dataset is unbalanced, *i.e.*, the count of active and inactive molecules is significantly different, a special procedure for balance was used. The count of inactive molecules was constant (15 molecules) and the count of active ones was duplicated. In the first series, 6 active molecules (all of active and inactive – 21 molecules) were used for developed Model B1 **(**trees count = 100, randomly selected descriptors count = 5); in the second series the count of active molecules was increased twofold, *i.e.*, 12 active molecules (a total of 27

molecules) were used for developed Model B2 (trees count = 150, randomly selected descriptors count = 13). The resulting QSAR models for the training set showed an unmistakable classification. The predictive ability of the QSAR models was evaluated using the "out-of-bag" (OOB) procedure. The quality of the classification models was assessed according to the following statistical characteristics: Matthew's correlation coefficient (*MCC*), specify (*SPC*), accuracy (*ACC*) and sensitivity (*SEN*):

$$MCC = \frac{(TP \cdot TN + FP \cdot FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

$$SPC = \frac{TN}{(TN + FP)}$$

$$ACC = \frac{(TP + TN)}{(TP + TN + FP + FN)}; \; SEN = \frac{TP}{(TP + FN)}$$

where *TP* – true positive, *TN* – true negative, *FP* – false positive, *FN* – false negative.

The obtained data are presented in Table II.

TABLE II. Statistical parameters for classification models for CC_{50} (Models B1 and B2) and HNTC (Model B3) for "out-of-bag" set

Model	Set	*MCC*	*ACC*	*SPC*	*SEN*
B1	6 + 15 = 21	0.77	0.90	0.88	1
B2	12 + 15 = 27	0.86	0.93	0.93	0.92
B3	12 + 9 = 21	0.81	0.90	0.89	0.91

As it can be seen from Table II, the balancing of models leads to a significant quality improvement. Model B2 could be considered as the most appropriate.

For *HNTC* data the duplication procedure was not used, since the dataset is more balanced (active molecules 12 and inactive 9). Model B3 (trees count = = 150, randomly selected descriptors count = 20) was also built for *HNTC* dataset, statistical parameters are shown in Table II.

As can be seen from this table, the statistical characteristics of models B1––B3 are quite acceptable.

It should be noted that all molecules are in domain applicability (DA) of Models A1, A2 and B1–B3.

CONCLUSION

The realized QSAR studies have found good compliance between theoretical predictions and in vitro revealed activity and cytotoxicity. Compounds 2, 5 and 7 have both good antiviral activity against tested virus A/H3N2, strain Hong Kong/68 combined with low cytotoxicity which is in a perfect correlation with QSAR pre-

diction studies. All guanidated derivatives have lower activity according to both QSAR investigation and in vitro assay. Based on the obtained results we can conclude that more bulky side chain residues are not tolerated by the mechanism of interaction with M2 ion channels of influenza virus. Unfortunately, the same conclusion is revealed for structures creating less conformational freedom including *β*-amino acids as moiety in the lateral chain of amantadine and rimantadine.

SUPPLEMENTARY MATERIAL

Additional data are available electronically at the pages of journal website: https://www.shd-pub.org.rs/index.php/JSCS/index, or from the corresponding author on request.

И З В О Д

МОДЕЛОВАЊЕ КВАНТИТАТИВНОГ ОДНОСА СТРУКТУРА–АКТИВНОСТ ИНХИБИТОРА М2 ЈОНСКОГ КАНАЛА ВИРУСА ГРИПА

IVANKA G. STANKOVA[1], RADOSLAV L. CHAYROV[1], MICHAELA SCHMIDTKE[2], DANCHO L. DANALEV[3], LIUDMILA N. OGNICHENKO[4], ANATOLY G. ARTEMENKO[4], VALERY A. SHAPKIN[5] и VICTOR E. KUZ'MIN[4]

[1]Department of Chemistry, South-West University "NeofitRilski", Blagoevgrad, 2700, Bulgaria, [2]Friedrich Schiller University, Department of Virology and Antiviral Therapy, Jena, 207745, Germany, [3]University of Chemical Technology and Metallurgy, Bio-technology Department, 1756 Sofia, 8 blvd. Kliment Ohridski, Bulgaria, [4]A.V. Bogatsky Physico-Chemical Institute of Ukrainian National Academy of Sciences, Department of Molecular Structure and Chemoinformatics, 86, Lustdorfskaya doroga, Odessa, 65080, Ukraine и [4]Department of Department of Theoretical Foundations of Chemistry, Odessa National Polytechnic University, 1, Shevchenko ave., Odessa 65044, Ukraine

Серија нових деривата адамантана који садрже фрагменте аминокиселина испитана је користећи метод симплекс представљања молекулских структура (енг. simplex representation of molecular structure, SiRMS) са циљем да се утврди однос између хемијске структуре испитиваних једињења и резултата антивиралне и цитотоксичне активности. Добијени резултати QSAR анализе показују да ће деривати адамантана који садрже аминокиселине кратких неполарних алифатичних група у бочном низу имати добре антивиралне активности према A/H3N2 сoју Hong Kong/68 вируса као и малу цитотоксичност. QSAR експерименти и резултати in vitro активности показују добру корелацију и указују да деривати који садрже гуанидо фрагмент у бочном низу и *β*-аминокиселину као супституент, имају малу или не показују активност.

(Примљено 9. маја 2020, ревидирано 15. марта, прихваћено 6. маја 2021)

REFERENCES

1. R. L. Tominack, F. G. Hayden, *Infect. Dis. Clin. North. Am.* **1** (1987) 459.
2. R. B. Belshe, M. H. Smith, C. B. Hall, R. Betts, A. J. Hay, *J. Virol.* **62** (1988) 1508 (https://jvi.asm.org/content/62/5/1508)
3. J. Wang, C. Ma, J. Wang, H. Jo, B. Canturk, G. Fiorin, W. F. DeGrado, *J. Med. Chem.* **56** (2013) 2804 (https://doi.org/10.1021/jm301538e)
4. R. M. Pielak, J. J. Chou, *Biochem. Biophys. Res. Commun.* **401** (2010) 58 (https://doi.org/10.1016/j.bbrc.2010.09.008)
5. J. Wang, J. X. Qiu, C. Soto, W. F. DeGrado, *Curr. Opin. Struct. Biol.* **21** (2011) 68 (https://doi.org/10.1016/j.sbi.2010.12.002)

6. A. V. Gaiday, I. A. Levandovskiy, K. G. Byler, T. E. Shubina, in *Proceedings of International Conference on Computational Science*, 2008, Berlin, Germany, pp. 360–368 (https://doi.org/10.1007/978-3-540-69387-1_40)
7. V. E. Kuz'min, A. G. Artemenko, E. N. Muratov, *J. Computer-Aided Molec. Des.* **22** (2008) 403 (https://doi.org/10.1007/s10822-008-9211-x)
8. V. E. Kuz'min, A. G. Artemenko, E. N. Muratov, P. G. Polischuk, L. N. Ognichenko, A. V. Liahovsky, A. I. Hromov, E. V. Varlamova, *Recent Advances in QSAR Studies: Methods and Applications*,Springer, Dordrecht, 2010, p.127 (ISBN 978-1-4020-9783-6)
9. S. Rännar, F. Lindgren, P. Geladi, S. Wold, *J. Chemometrics* **8** (1994) 111 (https://doi.org/10.1002/cem.1180080204)
10. L. Breiman, *Machine Learning* **45** (2001) 5 (https://doi.org/10.1023/A:1010933404324)
11. R. E. Carhart, D. H. Smith, R. Venkataraghavan, *J. Chem. Inform. Comp. Sci.* **25** (1985) 64 (https://doi.org/10.1021/ci00046a002)
12. K. Hasegawa, Y. Miyashita, K. Funatsu, *J. Chem. Inform. Comp. Sci.* **37** (1997) 306 (https://doi.org/10.1021/ci960047x)
13. V. E. Kuz'min, A. G. Artemenko, P. G. Polischuk, E. N. Muratov, A. I. Hromov, A. V. Liahovskiy, S. A. Andronati, S. Y. Makan, *J. Mol. Model* **11** (2005) 457 (http://doi.org/10.1007/s00894-005-0237-x)
14. *OECD (2014) Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models, OECD Series on Testing and Assessment*, OECD Publishing, Paris, 2014, p.154 (http://doi.org/10.1787/9789264085442-en)
15. M. Meloun, J. Militku, M. Hill, *Analyst* **127** (2002) 433 (http://doi.org/10.1039/B110779H)
16. P. G. Polischuk, E. N. Muratov, A. G. Artemenko, O. G. Kolumbin, N. N. Muratov, V. E. Kuz'min, *J. Chem. Inf. Mod.* **49** (2009) 2481 (http://doi.org/10.1021/ci900203n)
17. R. Chayrov, N. A. Parisis, M. V. Chatziathanasiadou, E. Vrontaki, K. Moschovou, G. Melagraki, H. Sbirkova-Dimitrova, B. Shivachev, M. Schmidtke, Y. Mitrev, M. Sticha, T. Mavromoustakos, A. G. Tzakos, I. Stankova, *Molecules* **25** (2020) 3989 (https://doi.org/10.3390/molecules25173989)
18. A. Chintakrindi, Ch. D'souza, M. Kanyalkar, *Mini Rev. Med. Chem.* **12** (2012) 1273 (https://doi.org/10.2174/138955712802761997).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 86 (7–8) S120–S124 (2021) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Quantitative structure–activity relationship modelling of influenza M2 ion channels inhibitors

IVANKA G. STANKOVA[1]*, RADOSLAV L. CHAYROV[1], MICHAELA SCHMIDTKE[2], DANCHO L. DANALEV[3]**, LIUDMILA N. OGNICHENKO[4], ANATOLY G. ARTEMENKO[4], VALERY A. SHAPKIN[5] and VICTOR E. KUZ'MIN[4]

[1]Department of Chemistry, South-West University "NeofitRilski", Blagoevgrad, 2700, Bulgaria, [2]Friedrich Schiller University, Department of Virology and Antiviral Therapy, Jena, 207745, Germany, [3]University of Chemical Technology and Metallurgy, Biotechnology Department, 1756 Sofia, 8 blvd. Kliment Ohridski, Bulgaria, [4]A.V. Bogatsky Physico-Chemical Institute of Ukrainian National Academy of Sciences, Department of Molecular Structure and Chemoinformatics, 86, Lustdorfskaya doroga, Odessa, 65080, Ukraine and [5]Department of Department of Theoretical Foundations of Chemistry, Odessa National Polytechnic University, 1, Shevchenko ave., Odessa 65044, Ukraine

J. Serb. Chem. Soc. 86 (7–8) (2021) 625–637

Fig. S-1 depicts the example showing how, using vertexes differentiation by atomic charges, the simplexes descriptors are generated at 2D level.

The contribution of each of nine SiRMS descriptors used in the Model A1 is represented in Table S-I.

The distribution of prediction errors (ΔY = mod($Y_{predicted} - Y_{observed}$)) demonstrates that a greater accuracy of prediction is observed for compounds with medium (Fig. S-3).

*,** Corresponding authors. E-mail: (*)ivastankova@abv.bg; (**)ddanalev@uctm.edu

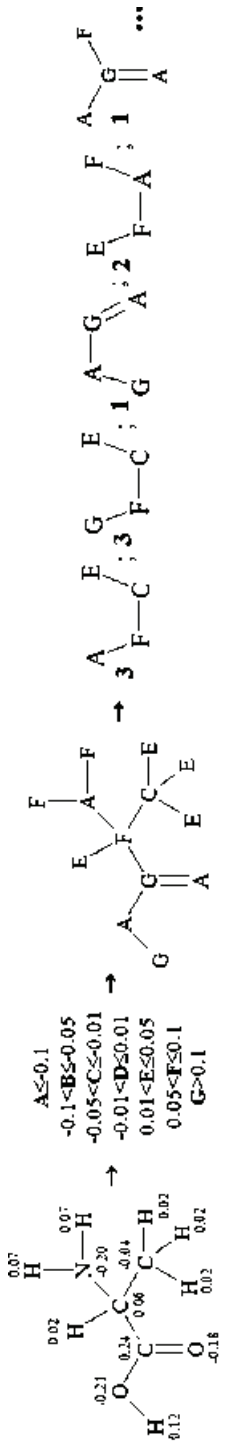
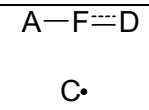
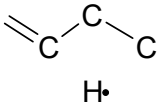
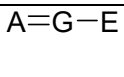
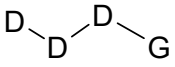
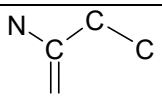
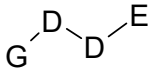
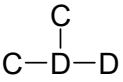
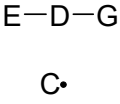


Fig. S-1. Depiction of the development of simplex descriptors at 2D level.

TABLE S-1. Structural parameters and their VIP scores

No.	Type differentiation vertexes simplexes	Structural fragment	VIP scores
1	Atom charge		0.78
2	Type		0.45
3	Atom charge		0.44
4	lipophilicity		0.41
5	type		0.41
6	lipophilicity		0.39
7	lipophilicity		0.33
8	Atom charge		0.30
9	lipophilicity		0.23

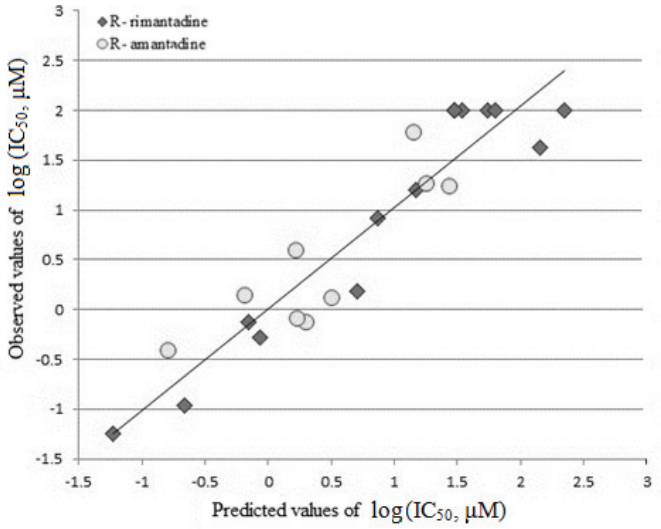


Fig. S-2. Observed *vs.* predicted diagram of antiviral activity (log IC_{50}) for 22 molecules (Model A1).

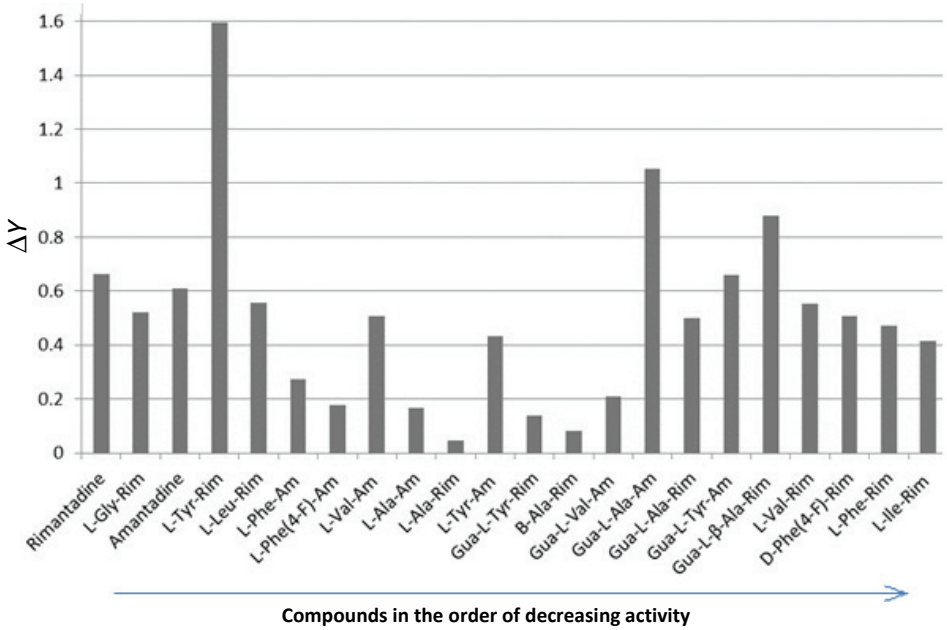


Fig. S-3. The distribution of activity prediction errors (Δ*Y*) by Model A2.

TABLE S-II. Predicted classes of activity by models B1 – B3 for "out-of-bag" set

Compound name	CC_{50} class			*HNTC* class	
	Observed	Predicted (Model B1)	Predicted (Model B2)	Observed	Predicted (Model B3)
L-Val-Rim	1	0	1	1	1
L-Phe-Am	1	1	1	1	1
L-Phe(4-F)-Am	1	0	1	1	0
L-Val-Am	0	0	1	0	0
Gua-L-Val-Am	0	0	0	0	0
L-Tyr-Am	0	0	0	0	1
Gua-L-Tyr-Am	0	0	0	0	0
L-Gly-Rim	0	0	0	1	1
L-Ile-Rim	1	1	1	1	1
L-Leu-Rim	0	0	0	1	1
L-Phe-Rim	1	1	1	1	1
β-Ala-Rim	0	0	0	1	1
D-Phe(4-F)-Rim	1	1	1	1	1
Amantadine	0	0	0	0	0
Gua-L-β-Ala-Rim	0	0	0	0	0
L-Tyr-Rim	0	0	0	1	1
Gua-L-Tyr-Rim	0	0	0	1	1
L-Ala-Rim	0	0	0	1	1
Gua-L-Ala-Rim	0	0	0	0	0
L-Ala-Am	0	0	0	0	0
Gua-L-Ala-Am	0	0	0	0	0

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. *86* (7–8) 639–649 (2021)
JSCS–5450
Original scientific paper

Synthesis and biological profiling of novel isocoumarin derivatives and related compounds

MILENA R. SIMIĆ[1]*#, SLAVICA ERIĆ[2], IVAN BORIĆ[3], ANNAMARIA LUBELSKA[4], GNIEWOMIR LATACZ[4], KATARZYNA KIEC-KONONOWICZ[4], SANDRA VOJNOVIĆ[5], JASMINA NIKODINOVIĆ-RUNIĆ[5] and VLADIMIR M. SAVIC[1#]

[1]University of Belgrade, Faculty of Pharmacy, Department of Organic Chemistry, Vojvode Stepe 450, 11221 Belgrade, Serbia; [2]University of Belgrade, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, Vojvode Stepe 450, 11221 Belgrade, Serbia, [3]IQVIA, Clinical Department, Rudnička 2, 11000 Belgrade, Serbia, [4]Jagiellonian University, Medical College, Department of Technology and Biotechnology of Drugs, Medyczna 9, 30-688 Kraków, Poland and [5]University of Belgrade, Institute for Molecular Genetics and Genetic Engineering, Laboratory for Microbial Molecular Genetics and Ecology, Vojvode Stepe 444a, 11042 Belgrade, Serbia

(Received 1 December 2020, revised 20 March, accepted 23 March 2021)

Abstract: In the continuation of our study of substituted isocoumarins a series of novel 3-azolyl isocoumarin and their thio derivatives, including some related lactone compounds was prepared and biologically profiled against *C. albicans* showing moderate activity with *MIC* values in range of 4–60 μg mL^{-1}, in general. The additional characterisation of selected compounds was carried out by exploring their activity on CYP3A4 and CYP2D6 enzymes, while experiments on mutagenicity were performed by AMES test. The representative isocoumarins **3b**, **4a** and **4b** showed lower inhibitory activity on CYP enzymes, when compared to the reference inhibitors, ketoconazole and quinidine. Compound **4a** showed a higher mutagenic potential than the other two compounds. Further characterization included cytotoxicity profiling against normal MRC5 cells.

Keywords: antifungal compounds; isocoumarins; *Candida albicans*; CYP enzymes; AMES.

INTRODUCTION

Isocoumarin derivatives have been intensively studied in recent decades.[1–3] In synthetic chemistry, they have been widely explored as starting compounds for the preparation of other heterocyclic derivatives.[4,5] Furthermore, these compounds have been subject of detailed research in medicinal chemistry due to their promising general biological potential.[6–8] Isocoumarin skeleton can be con-

* Corresponding author. E-mail: milena@pharmacy.bg.ac.rs
# Serbian Chemical Society member.
https://doi.org/10.2298/JSC201201025S

sidered as a privileged structure, as it is frequently encountered in naturally occurring compounds and also in plethora of synthetic derivatives exhibiting whole array of physiological properties. Natural isocoumarins may or may not have C(3) double bond but are rarely unsubstituted. They possess various substituents diversely positioned around the core, although most frequent substitution pattern, represented by structures in Fig. 1, is characterized by the oxygenated aromatic ring and functionality at C(3).

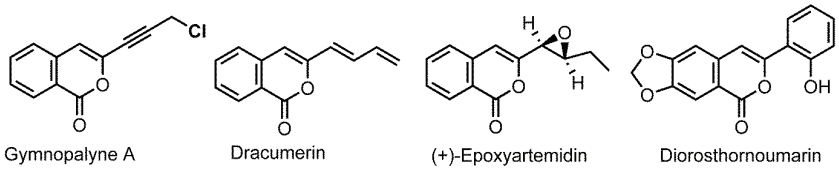


Fig. 1. Structures of some natural isocoumarins.

The C(3)-derived isocoumarins, either natural or synthetic, are known to have broad range of biological activities, such as anti-inflammatory, antimalarial, antimicrobial, antifungal, cytotoxic, antiangiogenic.[6–8] We have disclosed earlier a novel class of these derivatives possessing a heterocyclic ring at C(3).[9] Those compounds showed excellent antifungal properties, in some instances even better than voriconazole, which was used as a standard. Some derivatives demonstrated activity against azole resistant strains such as *Candida krusei* and *Candida parapsilosis*. Brief exploration of the mechanism of action suggested that the selected compounds did not cause any degradation or any interactions with *Candida albicans* DNA, implying that the origin of the biological profile of studied isocoumarins is not related to structural modifications of DNA molecules.

In order to explore further the structure–activity relationship (SAR) for this class of compounds and perform wider biological characterisation, a small series of additional isocoumarin derivatives and related compounds were synthesised and tested for antifungal activity on *C. albicans.* Further biological profiling was carried out by exploring cytotoxicity on healthy human MRC5 cells, inhibitory potential on cytochrome P450 (CYP) enzymes, and mutagenicity using AMES assay.

EXPERIMENTAL

General

The NMR spectra were recorded on Bruker Ascend 400 (400 MHz) and Bruker Avance III (500 MHz) spectrometers. Chemical shifts are given in parts per million (δ) downfield from tetramethylsilane as the internal standard. Deuterochloroform was used as a solvent, unless otherwise stated. Mass spectral data were recorded using Agilent Techologies 6520 Q-TOF spectrometer coupled with Agilent 1200 HPLC or Agilent Technologies 5975C MS coupled with Agilent Technologies 6890N GC. IR spectra were recorded on IR Termo Scientific Nicolet iS10 (4950) spectrometer. Melting points were determined using Boetius PHMK 05 apparatus without correction. Flash chromatography was carried out using silica gel 60

(230–400 mesh) while thin layer chromatography was carried out using alumina plates with 0.25 mm silica layer (Kieselgel 60 F_{254}, Merck). Compounds were seen by staining with potassium permanganate solution and Dragendorff reagent. The solvents were purified by distillation before use.

Analytical and spectral data, as well as additional experimental details are given in the Supplementary material to this paper.

General procedure for synthesis of thioisocoumarins

Thioisocoumarins were synthesised following literature procedure.[10] To a solution of isocoumarin (0.12 mmol) in dry toluene (2 mL) Lawesson's reagent (48.5 mg, 0.12 mmol) was added under nitrogen atmosphere and the mixture was refluxed for 12 h. After the reaction was complete, as indicated by TLC, the mixture was cooled to room temperature and solvent was removed under reduced pressure. The residue was purified by flash chromatography in order to get the clean the product.

General procedure for synthesis of azolyl-methylisocoumarins

Azolyl-methylisocoumarins were synthesised from 3-bromomethylisocoumarin and corresponding azoles following literature procedure.[11,12]

General procedure for synthesis of 3-azolyl-phthalides

Starting compound, 3-bromophthalide (**7**) was prepared from phthalide and NBS.[13] 3-Azolyl-phthalides **8a**–**c** were synthesised from 3-bromophthalide and corresponding azoles following the procedure from literature.[14]

Biology

Antifungal activity. Standard disc diffusion assay was done for the preliminary screen using 0.25 mg of test compounds per disc. Briefly, late stationary phase cells of *C. albicans* ATCC 10231 were spread on potato dextrose agar plates (HiMedia Laboratories, Mumbai, India) and sterile cellulose discs (6 mm diameter) were applied to the surface (HiMedia Laboratories, Mumbai, India). Plates were incubated at 37 °C and zones of inhibition were measured. As a control, voriconazole (0.25 and 0.025 mg/disc) was also included in disc diffusion assay. Broth microdilution assays were carried out in RPMI medium (Sigma Aldrich, Germany) according to the standard NCCLS M27-A2 with the highest concentration of 0.5 mg mL^{-1} applied.[15,16]

Cytotoxicity. Cytotoxicity of the compounds was assessed against human lung fibroblast MRC5 cell line obtained from the American Type Culture Collection (ATCC) using MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) reduction assay.[17] The highest concentration of the compounds used was 0.25 mg mL^{-1}. The MTT assay was performed after 48 h treatment with compounds three times in three replicates and the IC_{50} values (concentrations at which 50 % cell growth inhibition occurred) were calculated in comparison to control (untreated cells) that were arbitrarily set to 100 %.

In vitro *inhibition of CYP3A4 and CYP2D6 enzymes*. The study was performed using commercial tests CYP3A4 P450-Glo™ and CYP2D6 P450-Glo™ provided by Promega (Madison, WI, USA). These assays use the luminescent measurement of the potential inhibition effect on the conversion of the beetle D-luciferin derivative into D-luciferin by respective CYP isoform.[18] The tests were performed in white polystyrene, flat-bottom Nunc™ MicroWell™ 96-well microplates (Thermo Scientific, Waltham, MA, USA). The bioluminescence signal was measured with the microplate reader Perkin Elmer (Waltham, MA, USA) in luminescence mode. Ketoconazole (KE) and quinidine (QD) used as the positive controls were obtained

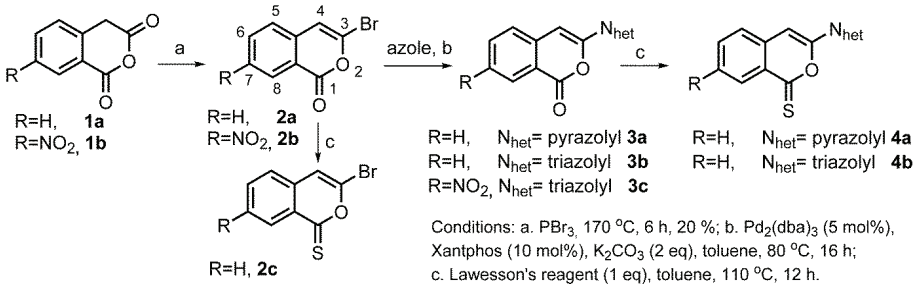
from Sigma–Aldrich. The isocoumarin derivatives were tested in triplicate in concentration range 0.01–25 μM, whereas reference inhibitors 0.0001–0.1 (QD) and 0.01–10 μM (KE), respectively.

AMES test. Ames microplate fluctuation protocol (MPF) assay was performed with *Salmonella typhimurium* TA100 strain, enabling the detection of base substitution mutations.[19] The bacterial strain, as well as exposure and indicator medium, were purchased from Xenometrix AG (Allschwil, Switzerland). *S. typhimurium* TA100 bacteria were exposed to 2 concentrations (1 and 10 μM) of a test agent, the reference was mutagen nonyl-4-hydroxyquinoline-*N*-oxide (4-NQNO, 0.5 μM) and a negative control (1 % DMSO) for 90 min in the medium containing sufficient histidine to support approximately two cell divisions. After 90 min, the exposure cultures were diluted in pH indicator medium lacking histidine, and aliquoted into 48 wells of a 384-well plate. Within two days, cells which have undergone the reversion to histidine prototrophy (revertants) reduced the pH of the indicator medium. The colour change (from purple to yellow) was detected visually by the microplate reader EnSpire Perkin Elmer (Waltham, MA, USA) at 420 nm. The number of wells containing revertants were counted for each dose and compared to the medium control baseline, which is an average of mean revertants from negative control (1 vol.% DMSO) plus the standard deviation. The two-fold increase in number of positive wells, over the medium control baseline indicated the mutagenicity borderline.

RESULTS AND DISCUSSION

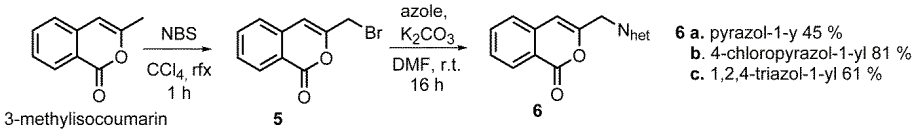
Chemistry

Various synthetic methods for the preparation of isocoumarin derivatives have been reported in the literature.[20–24] Scheme 1 shows our general approach to the synthesis of these compounds.[9] We employed the general Pd-catalysed processes for the functionalisation of the C(3) position of isocoumarin core.[25–27] The common intermediates, which provided access to several classes of compounds, were bromo derivatives **2a/b** (Scheme 1), synthesised from homophthalic anhydride.[28,29] The sulphur-containing derivatives such as **2c** and **4a/b** were prepared by the reaction of appropriate starting material and Lawesson's reagent under typical conditions (Scheme 1). The heterocyclic derivatives were synthesized *via* well-established Pd-promoted aminations, combining bromoisocoumarin scaffold and a distinctive antifungal pharmacophores such as the various azole rings (Scheme 1).[30–33]



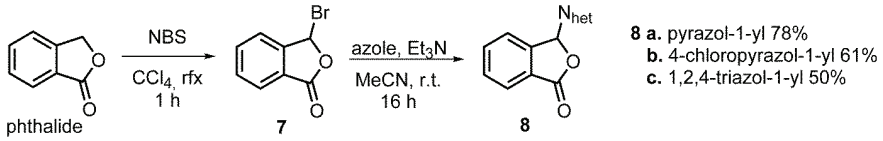
Scheme 1. Synthesis of 3-substituted isocoumarins.

For further structure–activity studies (SAR), several compounds possessing the methylene linker between the isocoumarin scaffold and the azole moiety were prepared using bromination/nucleophilic substitution sequence following the literature procedure (Scheme 2).[12]



Scheme 2. Synthesis of 3-azolylmethylisocoumarins.

With the intention to explore more chemical space defined by the above mentioned isocoumarins, an additional related class of compounds possessing phthalide moiety, was also prepared. Some phthalides previously demonstrated to have antifungal activity against plant pathogens.[34] Therefore we chose for a small series of azole containing phthalides in order to supplement SAR information (Scheme 3). [14,35]



Scheme 3. Synthesis of 3-azolylphthalides.

Biology

Antifungal activity and cytotoxicity. The antifungal and cytotoxic properties of isocoumarins from our current study and three compounds from our previous work for comparison are presented in Table I. The preliminary biological characterisation was carried out against *C. albicans* using disc diffusion assay with 250 μg of compound per disc.[15,16] For the determination of minimum inhibitory concentration (*MIC*) of the active compounds, in comparison with the clinically used voriconazole, the broth dilution method was applied. *MIC* was not determined for the compounds that showed poor activity against *C. albicans* in the disc diffusion assay. Together with the antifungal properties the cytotoxicity was also determined using human MRC5 cells and 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay,[17] which is also presented in Table I. The difference between the activities of nitro derivatives **2b** and **3c** (entries 2 and 6, Table I) and the corresponding compounds without this capability **2a** and **3b** (entries 1 and 5, Table I) is notable. The effect of thiocarbonyl group in place of carbonyl on the biological profile was also briefly explored. This structural alteration was expected to be important if the lacton moiety was involved in the shaping of biological character of this class of compounds. Indeed, the derivat-

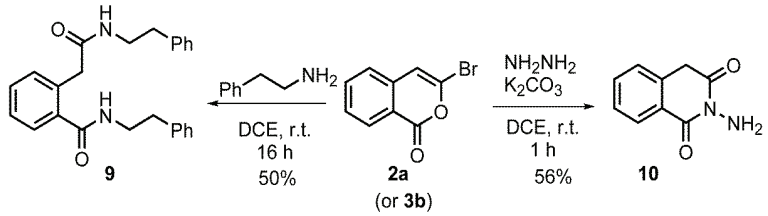
ives of this type **2c**, **4a**, **4b** (entries 3, 7, 8, Table I) showed good activity, in particular **2c**, suggesting that this structural features should be considered in designing novel isocoumarin antifungals. It should be also noted that azolyl-thio-isocoumarins, although showing comparable antifungal activity against *C. albicans*, seem to have slightly higher toxicity against normal cells when compared to the parent *O*-analogues. Unfortunately, none of derivatives possessing more flexible azole moiety such as **6a**–**c** (entries 9–11, Table I) showed any activity against *C. albicans.* Further explored class, 3-azolyl phthalides, compounds **8a**–**c**, have also shown that they were completely inactive (entries 12–14, Table I) in this study. Analysing the structural properties of the active compounds, in particular comparing the isocoumarin and phthalide derivatives, but also isocoumarins with and without directly attached azoles, it seems that the activity might be related to the electrophilicity of the lactone carbonyl.

Table I. Antifungal and cytotoxic properties of isocoumarins and derivatives; –: *MIC* was not determined; *: compounds synthesised in our previous work

Entry	Compound	*MIC* / μg mL^{-1} (*C. albicans*)	IC_{50} / μg mL^{-1} (MRC5)
1	**2a***	15.6	20
2	**2b**	60.0	100
3	**2c**	4.0	7
4	**3a***	31.2	30
5	**3b***	20	40
6	**3c**	>500	70
7	**4a**	20	5
8	**4b**	20	15
9	**6a**	-	-
10	**6b**	-	-
11	**6c**	-	-
12	**8a**	>400	>100
13	**8b**	>400	>100
14	**8c**	>400	>100
15	**9**	>500	20
16	**10**	40	90
17	Voriconazole	7.8	40

Therefore, the susceptibilities of active compounds **2a** and **3b** towards nucleophilic species[36] such as hydrazine and 2-phenylethylamine (Scheme 4) were studied. Both compounds underwent ring opening reaction under mild conditions at room temperature. Although the SAR suggested this possibility it was not a pleasing outcome, as it suggests possible mode of action of these compounds, because it would make them promiscuous towards various biological targets and might also support the compounds toxicity against MRC5 cell lines. However, the differences in *C. albicans*/MRC5 activities for some of the tested compounds suggest the involvement of more specific mechanism of action as well. The ring

opening products were also tested for antifungal activity against *C. albicans* and while the compound **9** did not show any antifungal activity, the compound **10** was moderately active (entries 15 and 16, Table I).



Scheme 4. Ring opening of bioactive isocoumarins.

Interaction with CYP3A4 and CYP2D6. The profiling of novel biologically active compounds against CYP enzymes is very important due to the fact that these enzymes have in metabolism of xenobiotics. The most active compounds surround the azole pharmacophore, a potential ligand for Fe^{3+}, containing the cytochrome P450 family of enzymes, prompted the study in this direction. Majority of drugs are metabolised by only few of CYP enzymes, amongst which CYP3A4 and CYP2D6 are perhaps the most important. Studying *in vitro* properties of the synthesised novel compounds as the inhibitors of these two isoforms in particular is a good basis for predicting their interactions *in vivo*.[37,38] Two compounds from our current (**4a**, **b)**, and one (**3b)** from our previous work, were selected for this investigation because they showed the same antifungal activity on *C. albicans*, but also various cytotoxicity. The inhibition potential of compounds was established using the luminescent assay in comparison to ketoconazole for CYP3A4 and quinidine for CYP2D6. The results outlined in Fig. 2 suggest that all three compounds, **3b** and **4a** and **b**, have significantly lower inhibition capacity towards CYP3A4 than ketoconazole. While ketoconazole at 0.1 μM concentration has strong inhibitory effect, the tested compounds, in particular **4b**, are significantly less potent. The calculated IC_{50} values supported this observation (IC_{50}: **4a** > 25 μM, **4b**, 16.8 μM, **3b**, 16.4 μM, ketoconazole, 0.14 μM).

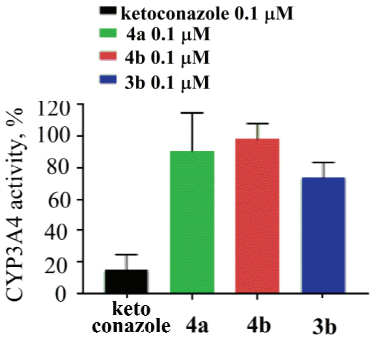


Fig. 2. Inhibition of CYP3A4 enzyme by selected compounds and standard inhibitor.

Further study of our compounds related to CYP2D6 resulted in similar outcome, Fig. 3. The comparison of inhibition effect of **4a** and **b** and **3b** with quinidine revealed very mild influence of our compounds on CYP2D6 activity. While quinidine at 0.1 μM concentration reduced activity of CYP2D6 for >80 %, all other tested compounds showed almost negligible effect. IC_{50} values found for all tested compounds **4a** and **b** and **3b** were >25 μM, whereas for quinidine were 0.01 μM).

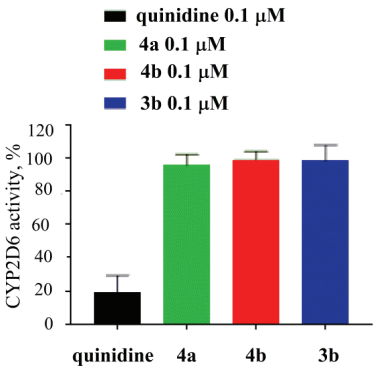


Fig. 3. Inhibition of CYP2D6 enzyme by selected compounds and standard inhibitor.

Mutagenicity (AMES test). Further profiling of the previously selected compounds **4a** and **b** and **3b** involved the determination of their mutagenic potential by the AMES mutagenicity assay with *Salmonella typhimurium* TA100 strain, which allows for the detection of mutagens that cause base-pair substitution and nonyl-4-hydroxyquinoline-*N*-oxide (NQNO), as a standard compound. This test developed few decades ago shows bacterial mutagenicity, and, consequently, may indicate carcinogenic properties of studied compounds. The profiling results summarised in Fig. 4 suggest that compounds **3b** and **4b** do not possess mutagenic properties either at 1 or 10 μM concentration (red dotted line shows muta-

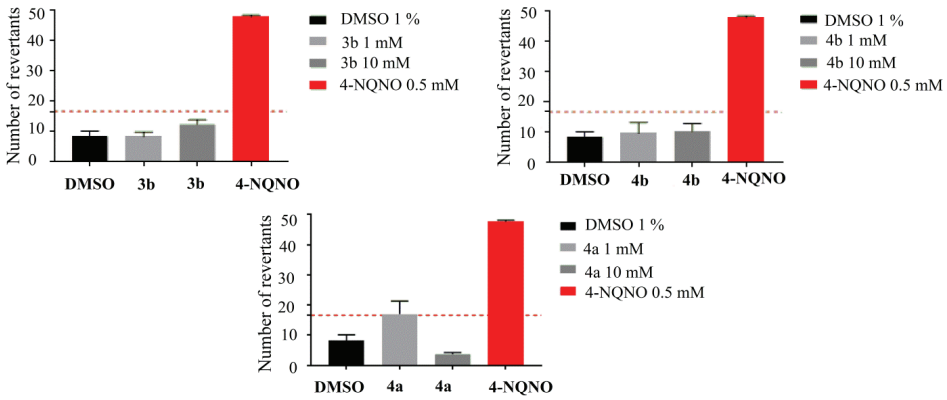


Fig. 4. Evaluation of mutagenic potential of selected compounds by AMES test. DMSO (1 vol.% in growth medium) was used as negative control.

genicity borderline).

Contrary to these, compound **4a** at 1 μM concentration showed mutagenicity, while at 10 μM the number of revertants observed were below the control level of DMSO and thus may indicate the toxic effect of this compound on *S. typhimurium* strain.

CONCLUSION

As an addition to our previous research, a series of azolyl-isocoumarins and related compounds were synthesised and tested against *C. albicans*. Novel azolyl-thioisocoumarin derivatives showed moderate antifungal activity, while azolyl-isocoumarins with the methylene linkers, andazolylphthalides, were inactive in this assay. Further profiling showed that the isocoumarin derivatives **4a** and **b** and **3b** demonstrated lower inhibitory activity on CYP3A4 and CYP2D6 enzymes when compared to ketoconazole and quinidine as reference inhibitors. The same compounds were tested by AMES assay and only **4a** showed activity at the border of mutagenicity. The presented results will be useful in further modification of isocoumarins in order to search for more efficient antifungal derivatives of this class of compounds.

SUPPLEMENTARY MATERIAL

Additional data are available electronically at the pages of journal website:.https://www.shd-pub.org.rs/index.php/JSCS/index, or from the corresponding author on request.

Acknowledgements. Financial support from the Ministry of Education, Science and Technological Development of the Republic of Serbia (451-03-68/2020-14/200161 and 451-03-68/2020-14/200042) is greatly appreciated. We also thank the Medical College, Jagiellonian University, Krakow, for support.

ИЗВОД

СИНТЕЗА И БИОЛОШКО ПРОФИЛИСАЊЕ НОВИХ ИЗОКУМАРИНСКИХ ДЕРИВАТА И СЛИЧНИХ ЈЕДИЊЕЊА

МИЛЕНА Р. СИМИЋ[1], СЛАВИЦА ЕРИЋ[2], ИВАН БОРИЋ[3], ANNAMARIA LUBELSKA[4], GNIEWOMIR LATACZ[4], KATARZYNA KIEC-KONONOWICZ[4], САНДРА ВОЈНОВИЋ[5], ЈАСМИНА НИКОДИНОВИЋ-РУНИЋ[5] и ВЛАДИМИР М. САВИЋ[1]

[1]Универзитет у Београду, Фармацеутски факултет, Катедра за органску хемију, Војводе Степе 450, 11221 Београд, [2]Универзитет у Београду, Фармацеутски факултет, Катедра за фармацеутску хемију, Војводе Степе 450, 11221 Београд, [3]IQVIA, Clinical Department, Рудничка 2, 11000 Београд, [4]Jagiellonian University, Medical College, Department of Technology and Biotechnology of Drugs, Medyczna 9, 30-688 Kraków, Poland и [5]Универзитет у Београду, Институт за молекуларну генетику и генетичко инжењерство, Лабораторија за молекуларну генетику и екологију микроорганизама, Војводе Степе 444а, 11042 Београд

Синтетисана је серија нових 3-азолил-изокумарина и сличних лактонских деривата и евалуирана је њихова антифунгална активност на *Candida albicans*, где су показали умерену активност (*MIC* 4–60 μg mL^{-1}). Испитана је и интеракција одабраних изокумаринских деривата са хуманим CYP3A4 и CYP2D6 ензимима помоћу луминисцентног теста, док им је мутагени потенцијал одређен AMES тестом. Испитивани изокумарини **3b**, **4a** и **4b** не показују значајну интеракцију са наведеним CYP ензимима у поређењу са

референтним инхибиторима. Једињење **4a** показује већи мутагени потенцијал у односу на друга два. Додатна биолошка карактеризација је укључила одређивање цитотоксичности према нормалним MRC5 ћелијама.

(Примљено 1. децембра 2020, ревидирано 20. марта, прихваћено 23. марта 2021)

REFERENCES

1. S. Pal, V. Chatare, M. Pal, *Curr. Org. Chem*. **15** (2011) 782 (https://doi.org/10.2174/138527211794518970)
2. L. Pochet, R. Frederick, B. Masereel, *Curr. Pharm. Des*. **10** (2004) 3781 (https://doi.org/10.2174/1381612043382684)
3. A. Saeed, *Eur. J. Med. Chem*.**116** (2016) 290 (https://doi.org/10.1016/j.ejmech.2016.03.025)
4. S. Roy, S. Roy, B. Neuenswander, D. Hill, R. C. Larock, *J. Comb. Chem*. **11** (2009) 1128 (https://dx.doi.org/10.1021%2Fcc9001197)
5. P. Manivel, K. Prabakaran,Y. Suneel, S. M. Ghouse, P. M. Vivek, E. Ubba, I. Pugazhenthi, Fazlur-Rahman Nawaz Khan, *Res. Chem. Intermed*. **41** (2015) 2081 (https://doi.org/10.1007/s11164-013-1333-7)
6. Z. Xiao, S.Chen, R. Cai, S. Lin, K. Hong, Z. She, *Beilstein J. Org. Chem*. **12** (2016) 2077 (https://doi.org/10.3762/bjoc.12.196)
7. J. S. Kumar, B. Thirupataiah, R. Medishetti, A. Ray, S. D. Bele, K. A. Hossain, G. S. Reddy, R. K. Edwin, A. Joseph, N. Kumar, G. G. Shenoy, C. M. Rao, M. Pal, *Eur. J. Med. Chem.* **201** (2020) 112335 (https://doi.org/10.1016/j.ejmech.2020.112335)
8. H. Hussain, I. R. Green, *Expert. Opin. Ther. Pat*. **27** (2017) 1267 (https://doi.org/10.1080/13543776.2017.1344220)
9. M. Simic, N. Paunovic, I. Boric, J. Randjelovic, S. Vojnovic, J. Nikodinovic-Runic, M. Pekmezovic, V. Savic, *Bioorg. Med. Chem. Lett*. **26** (2016) 235 (https://doi.org/10.1016/j.bmcl.2015.08.086)
10. P. Manivel, S. M. Roopan, D. P. Kumar, F. N. Khan, *Phosphorus Sulfur Silicon Relat. Elem*. **184** (2009) 2576 (https://doi.org/10.1080/10426500802529507)
11. L. Liu, J. Hu, X-C. Wang, M-J. Zhong, X-Y. Liu, S-D. Yang, Y-M. Liang, *Tetrahedron* **68** (2012) 5391 (https://www.sciencedirect.com/science/article/pii/S0040402012006746?via%3Dihub)
12. M. Biagetti, A. Capelli, A. Accetta, L. Carzaniga, U.S. Pat. Appl. Publ. (2015), US 20150166549 A1 20150618
13. I. A. Koten, R. J. Sauter, *Org. Synth*. **42** (1962) 26 (http://orgsyn.org/Result.aspx)
14. H. Stark, M. Krause, A. Rouleau, M. Garbarg, J-C. Schwartz, W. Schunack, *Bioorg. Med. Chem*. **9** (2001) 191 (https://doi.org/10.1016/S0968-0896(00)00237-6)
15. *Reference Method for Broth Dilution Antifungal Susceptibility Testing of Yeasts; Approved Standard – Second edition for yeasts (2002), NCCLS document M27-A2*, National Committee for Clinical Laboratory Standards, Wayne, PA, 2002
16. M. Balouiri, M. Sadiki, S. K. Ibnsouda, *J. Pharm. Anal*. **6** (2016) 71 (https://doi.org/10.1016/j.jpha.2015.11.005)
17. M. B. Hansen, S. E. Nielsen, K. Berg, *J Immunol Methods* **119** (1989) 203 (https://doi.org/10.1016/0022-1759(89)90397-9)
18. K. Kamiński, A. Rapacz, J. J. Łuszczki, G. Latacz, J. Obniska, K. Kieć-Kononowicz, B. Filipek, *Bioorg. Med. Chem*. **23** (2015) 2548 (https://doi.org/10.1016/j.bmc.2015.03.038)
19. https://www.aniara.com/mm5/PDFs/Literature/Xenometrix_AmesII-Technical-Doc.pdf (accessed 1 November, 2019)

20. M. P. Drapeau, L. J. Gooßen, *Chem. Eur. J.* **22** (2016) 18654 (https://doi.org/10.1002/chem.201603263)
21. D. A. Loginov, V. E. Konoplev, *J. Organomet. Chem*. **867** (2018) 14 (https://doi.org/10.1016/j.jorganchem.2017.11.013)
22. K. Suman, K. Prabhakara Rao, V. Anuradha, M. V. Basaveswara Rao, M. Pal, *Mini Rev. Med. Chem.* **18** (2018) 1064 (http://dx.doi.org/10.2174/1389557518666180117093706)
23. N. Panda, P. Mishra, I. Mattan, *J. Org. Chem*. **81** (2016) 1047 (https://doi.org/10.1021/acs.joc.5b02602)
24. A. P. Molotkov, M. A. Arsenov, D. A. Kapustin, D. V. Muratov, N. E. Shepel, Y. V. Fedorov, A. F. Smolyakov, E. I. Knyazeva, D. A. Lypenko, A. V. Dmitriev, A. E. Aleksandrov, E. I. Maltsev, D. A. Loginov, *ChemPlusChem* **85** (2020) 334 (https://doi.org/10.1002/cplu.202000048)
25. Y. S. Kumar, C. Dasaradhan, K. Prabakaran, P. Manivel, F-R. N. Khan, E. D. Jeong, E. H. Chung, *Tetrahedron Lett*. **56** (2015) 941 (https://doi.org/10.1016/j.tetlet.2014.12.114)
26. B. H. Yang, S. L. Buchwald, *J. Organomet. Chem*. **576**, (1999) 125 https://doi.org/10.1016/S0022-328X(98)01054-7
27. J. F. Hartwig, in *Modern Arene Chemistry*, D. Astruc (Ed.), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2002, p.107 (https://doi.org/10.1002/3527601767.ch4)
28. T. X. Nguyen, M. Abdelmalak, C. Marchand, K. Agama, Y. Pommier, M. Cushman, *J. Med. Chem*. **58** (2015) 3188 (https://doi.org/10.1021/acs.jmedchem.5b00136)
29. H. Duddeck, M. Kaiser, *Spectrochim. Acta, A* (1985) 913 (https://doi.org/10.1016/0584-8539(85)80224-5)
30. Y. J. Jingjun, S. Buchwald, *J. Am. Chem. Soc.* **124** (2002) 6043 (https://doi.org/10.1021/ja012610k)
31. X. Che, C. Sheng, W. Wang, C. Yongbing, X. Yulan, J. Haitao, D. Guoqiang, M. Zhenyuan, Y. Jianzhong, Z. Wannian, *Eur. J. Med. Chem*. **44** (2009) 4218 (https://doi.org/10.1016/j.ejmech.2009.05.018)
32. S. Sandhu, H. Shukla, R. Aharwal, S. Kumar S, S. Shukla, *Nat. Prod. J*. **4** (2014) 140 (https://doi.org/10.2174/2210315504021410091006 32)
33. D. A. Erlanson, S. W. Fesik, R.E. Hubbard, W. Jahnke, H. Jhoti, *Nat. Rev. Drug Discov*. **15** (2016) 605 (https://doi.org/10.1038/nrd.2016.109)
34. X-L. Yang, S. Zhang, Q-B. Hu, D-Q. Luo, Y. Zhang, *J. Antibiot. (Tokyo)* **64** (2011) 723 (https://doi.org/10.1038/ja.2011.82)
35. Y-Y. Xu, A-R. Qian, X-F. Cao, C-Y. Ling, Y-B. Cao, R-L. Wang, Y-S. Li, Y-S. Yang, *Chinese Chem. Lett.* **27** (2016) 703 (https://doi.org/10.1016/j.cclet.2016.01.040)
36. J. W. Harper, J. C. Powers, *Biochemistry* **24** (1985) 7200 (https://doi.org/10.1021/bi00346a028)
37. T. Niwa, Y. Imagawa, H. Yamazaki, *Curr. Drug Metab.* **15** (2014) 651 https://doi.org/10.2174/1389200215666141125121511
38. T. Saarikoski, T. I. Saari, N. M. Hagelberg, J. T. Backman, P. J. Neuvonen, M. Scheinin, K. T. Olkkola, K. Laine, *Eur. J. Clin. Pharmacol*. **71** (2015) 321 (https://doi.org/10.1007/s00228-014-1799-2).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 86 (7–8) S125–S130 (2021) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Synthesis and biological profiling of novel isocoumarin derivatives and related compounds

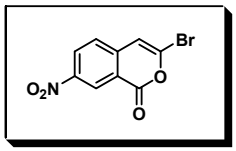
MILENA R. SIMIĆ[1]*, SLAVICA ERIĆ[2], IVAN BORIĆ[3], ANNAMARIA LUBELSKA[4], GNIEWOMIR LATACZ[4], KATARZYNA KIEC-KONONOWICZ[4], SANDRA VOJNOVIĆ[5], JASMINA NIKODINOVIĆ-RUNIĆ[5] and VLADIMIR M. SAVIC[1]

[1]University of Belgrade, Faculty of Pharmacy, Department of Organic Chemistry, Vojvode Stepe 450, 11221 Belgrade, Serbia; [2]University of Belgrade, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, Vojvode Stepe 450, 11221 Belgrade, Serbia, [3]IQVIA, Clinical Department, Rudnička 2, 11000 Belgrade, Serbia, [4]Jagiellonian University, Medical College, Department of Technology and Biotechnology of Drugs, Medyczna 9, 30-688 Kraków, Poland and [5]University of Belgrade, Institute for Molecular Genetics and Genetic Engineering, Laboratory for Microbial Molecular Genetics and Ecology, Vojvode Stepe 444a, 11042 Belgrade, Serbia

J. Serb. Chem. Soc. 85 (7–8) (2021) 639–649

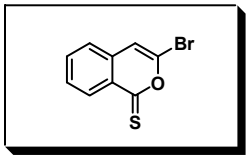
SYNTHETIC PROCEDURES AND CHARACTERIZATION DATA OF SYNTHESISED COMPOUNDS

*3-Bromo-7-nitro-1*H*-isochromen-1-one (**2b**)*



Compound as synthesised from 4-nitro homophthalic anhydride following the modified literature procedure.[1,2] Flash chromatography (SiO_2, 7:3 v/v petroleum ether-diethyl ether, Rf = 0.30) afforded **2b** (20 %) as a pale yellow amorphous solid, mp = 143-145 ºC. IR (ATR) cm^{-1}: 1743, 1617, 1511, 1344, 1056, 741, 680; ^{1}H NMR (400 MHz, $CDCl_3$): δ 9.09 (s, 1H), 8.54 (d, 1H, *J* = 8.8 Hz), 7.54 (d, 1H, *J* = 8.8 Hz), 6.83 (s, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) *δ* 158.9, 147.2, 141.9, 136.3, 129.7, 126.3, 126.2, 119.9, 108.7; MS (EI): *m/z* 268.9 $[M]^+$, 224.9, 189.9, 174.0, 134.0.

*3-Bromo-1*H*-isochromene-1-thione (**2c**)*

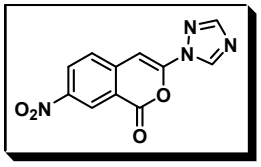


Compound as synthesised from **2a** following the general procedure for synthesis of thioisocoumarins. Flash chromatography (SiO_2, 9:0.5:0.5 v/v petroleum ether-diethyl ether-dichloromethane, Rf = 0.35) afforded the product **2c** (7 mg, 27 %) as a yellow amorphous solid,

* Corresponding author. E-mail: milena@pharmacy.bg.ac.rs

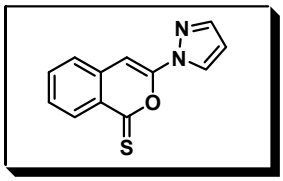
mp = 111-113 ºC; IR (ATR) cm^{-1}: 1613, 1546, 1468, 1327, 1298, 835, 879, 766; ^1H NMR (400 MHz, $CDCl_3$) δ 8.61 (d, 1H, J = 8.4 Hz), 7.71 (t, 1H, J = 7.6 Hz), 7.50 (t, 1H, J = 7.6 Hz), 7.32 (d, 1H, J = 7.6 Hz), 6.85 (s, 1H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 199.5, 135.4, 133.9, 133.1, 132.2, 129.7, 129.6, 125.2, 112.2; MS (EI): *m/z* 239.8 $[M]^+$, 225.9, 213.9, 160.9, 133.0, 105.9, 89.0; HRMS (ESI/Q-TOF) *m/z*calcd for $[C_9H_5BrOS + H^+]$: 240.9323, found 240.9322.

*7-Nitro-3-(1*H*-1,2,4-triazol-1-yl)-1*H*-isochromen-1-one (**3c**)*



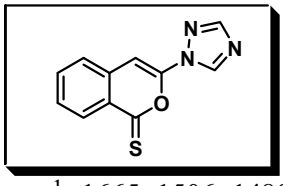
Compound was synthesised from 3-bromo-7-nitro-isocoumarine (**2b**) and 1, 2, 4-triazole following the general procedure for Buchwald-Hartwig reaction.[3] Flash chromatography (SiO_2, 9:1 v/v diethyl ether-petroleum ether, Rf = 0.29) afforded the product **3c** (10 mg, 48 %) as a pale yellow amorphous solid, mp = 208-211 ºC. IR (ATR) cm^{-1}: 1755, 1652, 1611, 1341, 1218, 985, 869, 665; ^1H NMR (400 MHz, $CDCl_3$): δ 9.16 (s, 1H), 8.88 (s, 1H), 8.58 (d, 1H, J = 8.8 Hz), 8.18 (s, 1H), 7.73 (d, 1H, J = 8.8 Hz), 7.15 (s, 1H); ^{13}C (100 MHz, $CDCl_3$) δ 157.3, 154.0, 147.0, 146.6, 142.3, 141.9, 129.9, 127.9, 126.3, 119.8, 90.9; MS (EI): *m/z* 258.0 $[M]^+$,230.0, 188.9, 149.0, 134.0, 88.0; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_{11}H_6N_4O_4 + H^+]$: 259.1467, found 259.0468.

*3-(1*H*-pyrazol-1-yl)-1*H*-isochromene-1-thione (**4a**)*



Compound was synthesised from **3a** following the general procedure for synthesis of thioisocoumarins. Flash chromatography (SiO_2, 8:1:1 v/v petroleum ether-diethyl ether-dichloromethane, Rf = 0.28) afforded the product **4a** (18 mg, 67 %) as a yellow needles, mp = 107-109 ºC. IR (ATR) cm^{-1}: 1657, 1551, 1480, 1457, 1288, 1160, 823, 760; ^1H NMR (400 MHz, $CDCl_3$) δ 8.64 (d, 1H, J = 8.0 Hz), 8.30 (s, 1H), 7.78 (s, 1H), 7.71 (t, 1H, J = 7.6 Hz), 7.48-7.42 (m, 2H), 7.15 (s, 1H), 6.52 (d, 1H, J = 1.2 Hz); ^{13}C NMR (101 MHz, $CDCl_3$) δ 197.2, 149.3, 143.4, 135.6, 133.2, 133.1, 129.3, 128.5, 127.8, 126.5, 108.5, 92.9; MS (EI): *m/z* 228.0 $[M]^+$, 199.0, 168.0, 146.0, 133.0, 89.0; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_{12}H_8N_2OS + H^+]$: 229.0436, found 229.0435.

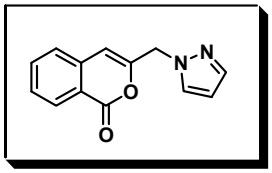
*3-(1*H*-1,2,4-triazol-1-yl)-1*H*-isochromene-1-thione (**4b**)*



Compound was synthesised from **3b** following the general procedure for synthesis of thioisocoumarins. Flash chromatography (SiO_2, 7:3 v/v petroleum ether-ethyl acetate, Rf = 0.35) afforded the product **4b** (50 %) as a yellow amorphous solid, mp = 60-63 ºC.IR (ATR) cm^{-1}: 1665, 1506, 1480, 1286, 1163, 1100, 1055, 996, 742, 685; ^1H NMR (400

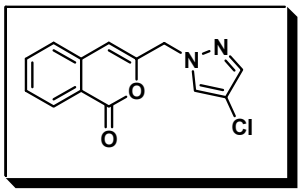
MHz, $CDCl_3$): δ 8.99 (*s*, 1H), 8.65 (*d*, 1H, J = 8.4 Hz), 8.19 (s, 1H), 7.77 (t, 1H, J = 7.6 Hz), 7.54-7.51 (m, 2H) 7.16 (s, 1H) ppm; ^{13}C (100 MHz, $CDCl_3$) δ 196.3, 153.1, 146.4, 141.7, 135.8, 133.3, 131.8, 129.9, 129.6, 126.9, 95.1 ppm;MS (EI): *m/z* 229.0 $[M]^+$, 200.9, 185.0, 146.9, 132.9, 120.0, 89.0; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_{11}H_7N_3OS + H^+]$: 230.0388, found 230.0388.

*3-Pyrazol-1-ylmethyl-isochromen-1-one (**6a**)*



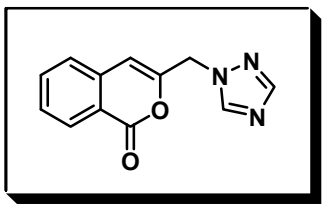
Compound **6a** was synthesised from 3-brom-methylisocoumarin and pyrazole following the general procedure. Flash chromatography (SiO_2, 7:3 v/v diethyl ether-petroleum ether, Rf = 0.29) afforded the product **6a** (45 %) as a colourless solid, mp = 110-111 °C. IR(ATR) cm^{-1}: 1721, 1283, 1051, 1021, 755, 687; ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.23 (d, 1H, J = 7.9 Hz), 7.70 – 7.64 (m, 1H), 7.59 (dd, 2H, J = 10.7, 1.7 Hz), 7.50 - 7.46 (m, 1H), 7.34 (d, 1H, J = 7.8 Hz), 6.33 (t, 1H, J = 2.0 Hz), 6.27 (s, 1H), 5.14 (s, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 161.8, 151.6, 140.4, 136.4, 134.9, 130.4, 129.6, 128.7, 125.9, 120.6, 106.5, 104.8, 52.9; MS (EI): *m/z* 226.0 $[M]^+$, 198.0, 184.0, 169.0, 143.0, 130.0, 117.0, 89.0; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_{13}H_{10}N_2O_2 + H^+]$: 227.0821, found 227.0821.

*3-(4-Chloro-pyrazol-1-ylmethyl)-isochromen-1-one (**6b**)*



Compound **6b** was synthesised from 3-brom-methylisocoumarin and 4-chloropyrazole following the general procedure. Flash chromatography (SiO_2, 6:4 v/v diethyl ether-petroleum ether, Rf = 0.28) afforded the product **6b** (81 %) as a colourless solid, mp = 160-163 °C. IR (ATR) cm^{-1}: 1722, 1400, 1378, 938, 833, 757, 688; ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.26 (d, 1H, J = 4.0 Hz), 7.72 (t, 1H, J = 8.0 Hz), 7.61-7.50 (m, 3H), 7.40, d, 1H, J = 8.0 Hz), 6.37 (s, 1H), 5.08 (s, 2H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 161.7, 150.5, 138.9, 136.1, 135.1, 129.8, 129.0, 128.2, 126.0, 120.7, 111.0, 105.4, 53.7; MS (EI): *m/z* 260.0 $[M]^+$, 232.0, 218.0, 169.0, 89.0; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_{13}H_9ClN_2O_2 + H^+]$: 261.0431, found 261.0432.

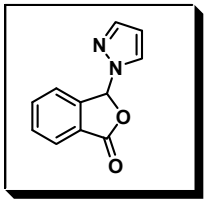
*3-[1,2,4]Triazol-1-ylmethyl-isochromen-1-one (**6c**)*



Compound **6c** was synthesised from 3-brom-methylisocoumarin and 1,2,4-triazole following the general procedure. Flash chromatography (SiO_2, 1:1 v/v ethyl acetate/diethyl ether, Rf = 0.21) afforded the product **6c** (61 %) as a colourless solid, mp = 153-157°C.IR (ATR) cm^{-1}: 1720, 1504, 1267, 1058, 1015, 768, 677; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.31 (s, 1H), 8.27 (d, 1H, J

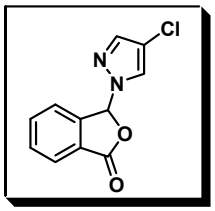
= 8.0 Hz), 8.01 (s, 1H), 7.73 (d, 1H, *J* = 8.0 Hz), 7.55 (t, 1H, *J* = 8.0Hz), 7.42 (d, 1H, *J* = 8.0 Hz), 6.46 (s, 1H), 5.20 (s, 2H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 161.4, 152.6, 149.4, 144.0, 135.9, 135.2, 129.9, 129.2, 126.0, 120.8, 106.0, 50.6; MS (EI): *m/z* 227.0 $[M]^+$, 198.9, 185.0, 169.9, 144.9, 117.0, 89.0; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_{12}H_9N_3O_2 + H^+]$: 228.0773, found 228.0773.

*3-Pyrazol-1-yl-3*H-*isobenzofuran-1-one* (***8a***)



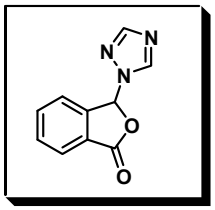
Compound **8a** was synthesised from 3-bromophthalide and pyrazole following the general procedure. Flash chromatography (SiO_2, 6:4 v/v diethyl ether-petroleum ether, Rf = 0.28) afforded the product **8a** (78 %) as a colourless solid, mp = 58-59 ºC. IR (ATR) cm^{-1}: 1768, 1437, 1283, 1061, 947, 732; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8,02 (d, 1H, *J* = 7.6 Hz), 7.85 (dt, 1H, *J* = 7.6 and 1.2 Hz), 7.71 (t, 1H, J = 7.6 Hz), 7.66 (d, 1H, 1.6 Hz), 7.54 (s, 1H), 7.51 (d, 1H, *J* = 0.4 Hz), 7.36 (d, 1H, *J* = 2.4 Hz), 7.35 (s, 1H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.8, 143.7, 141.9, 134.9, 131.3, 128.9, 126.9, 126.0, 123.7, 108.0, 87.5; MS (EI): *m/z* 200.0 $[M]^+$, 171.0, 166.0, 133.0, 105.0, 77.0; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_{11}H_8N_2O_2 + H^+]$: 201.0659, found 201.0659.

*3-(4-Chloro-pyrazol-1-yl)-3*H-*isobenzofuran-1-one* (***8b***)



Compound **8b** was synthesised from 3-bromophthalide and 4-chloropyrazole following the general procedure. Flash chromatography (SiO_2, 6:4 v/v petroleum ether-diethyl ether, Rf = 0.29) afforded the product **8b** (61 %) as a colourless solid, mp = 104-106 ºC. IR (ATR) cm^{-1}: 1773, 1430, 1284, 1060, 950, 732; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.02 (d, 1H, *J* = 7.2 Hz), 7.80 (t, 1H, *J* = 7.2 Hz), 7.73 (t, 1H, *J* = 7.2 Hz), 7.58 (s, 1H), 7.54 (d, 1H, *J* = 7.6 Hz), 7.31 (s, 1H), 7.25 (d, 1H, *J* = 6.4 Hz); ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.5, 142.9, 140.4, 135.2, 131.6, 126.8, 126.6, 126.1, 123.8, 112.8, 87.7; MS (EI): *m/z* 234.0 $[M]^+$, 190.0, 133.0, 105.0, 77.0; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_{11}H_7ClN_2O_2 + H^+]$: 235.0274, found 235.0271.

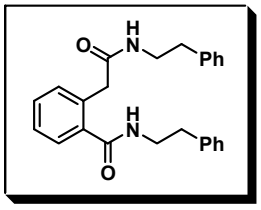
*3-[1,2,4]Triazol-1-yl-3*H-*isobenzofuran-1-one* (***8c***)



Compound **8c** was synthesised from3-bromophthalide and 1,2,4-triazole following the general procedure. Flash chromatography (SiO_2, 1:1 v/v diethyl ether-ethylacetate, Rf = 0.42) afforded the product **8c** (50 %) as a colourless solid, mp = 133-136 ºC. IR (ATR) cm^{-1}: 1769, 1432, 1301, 1139, 965, 739; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.15 (s, 1H), 7.98 (d, 1H, *J* = 7.6 Hz), 7.96 (s, 1H), 7.75 (dt, 1H, *J* = 7.6 and 1.2 Hz), 7.68 (t, 1H, *J* = 7.6 Hz), 7.48 (dd, 1H, *J* = 7.6 and 0.8 Hz), 7.30 (s, 1H); ^{13}C NMR (101 MHz,

$CDCl_3$) δ 167.1, 153.1, 143.4, 142.6, 135.3, 131.8, 126.4, 126.3, 123.6, 84.7; MS (EI): *m/z* 201.0 $[M]^+$, 172.0, 133.0, 105.0, 77.0; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_{10}H_7\ N_3O_2 + H^+]$: 202.0608, found 202.0608.

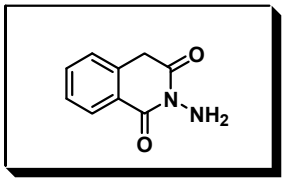
N-*Phenethyl-2-(phenethylcarbamoyl-methyl)-benzamide (**9**)*



To a solution of 3-bromoisocoumarin (23 mg, 0.10 mmol) in dichloromethane (2 mL) 2-phenylethylamine (44 mg, 0.36 mmol) was added at r.t. The mixture was stirred at r.t. overnight. Solvent was removed under reduced pressure and crude mixture was purified by flash chro00matography (SiO_2, diethyl ether, Rf = 0.35) to afford the product **9** (50 %) as a colourless solid, mp = 197-198 ºC. IR (ATR) cm^{-1}: 1637, 1535, 1319, 744, 695; ^1H NMR (400 MHz, $CDCl_3$) δ 7.41-7.16 (m, 15 H), 7.04 (d, 2H, *J* = 7.2 Hz), 6.59 (s, 1H), 3.70 (q, 2H, *J* = 6.8 Hz), 3.47 (s, 2H), 3.42 (q, 2H, *J* = 6.8 Hz), 2.94 (t, 2H, *J* = 6.8 Hz), 2.75 (t, 2H, *J* = 6.8 Hz); ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.1, 169.9, 139.1, 138.6, 135.9, 134.5, 130.8, 130.7, 128.8, 128.7 (2C), 128.4, 127.4, 127.1, 126.7, 126.2, 41.6, 41.1, 40.9, 35.6, 35.5; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_9H_{26}N_2O_2 + H^+]$: 387.2072, found 387.2072.

Reaction of compound **3b** with phenylethylamine also gave **9**, confirmed only by ^1H NMR spectrum of crude product.

*2-Amino-4*H*-isoquinoline-1,3-dione (**10**)*



To a solution of 3-bromoisocoumarin (26 mg, 0.12 mmol) in dichloromethane (2 mL) hydrazine hydrate (6.5 mg, 0.13 mmol) and K_2CO_3 (16 mg, 0.12 mmol) were added at room temperature. The mixture was stirred at r.t. for 1h. After completion of the reaction indicated by TLC, the mixture was diluted with DCM, washed with water and dried over $MgSO_4$. Solvent was removed under reduced pressure. The crude mixture was purified by flash chromatography (SiO_2, diethyl ether, Rf = 0.33) to afford the product **10** (56 %) as a colourless solid, mp = 48-150 ºC. IR (ATR) cm^{-1}: 1719, 1637, 1548, 1387, 1194, 919, 746; ^1H NMR (400 MHz, $CDCl_3$) δ 8.22 (d, 1H, *J* = 8.0 Hz), 7.61 (t, 1H, *J* = 7.6 Hz), 7.47 (t, 1H, *J* = 8.0 Hz), 7.30 (d, 1H, *J* = 7.6 Hz), 5.29 (s, 2H), 4.13 (s, 2H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.2, 161.8, 133.7, 133.3, 128.9, 127.9, 127.3, 124.6, 36.0; MS (EI): *m/z* 176.0 $[M]^+$, 161.0, 145.0, 132.0, 118.0, 104.0, 90.0, 77.0, 63.0; HRMS (ESI/Q-TOF) *m/z* calcd for $[C_9H_8\ N_2O_2 + H^+]$: 177.0664, found 177.0659.

REFERENCES

1. T. X. Nguyen, M. Abdelmalak, C. Marchand, K. Agama, Y. Pommier, M. Cushman, *J. Med. Chem*. **58** (2015) 3188 (https://doi.org/10.1021/acs.jmedchem.5b00136)

2. H. Duddeck, M. Kaiser, *Spectrochim. Acta, A* (1985) 913 (https://doi.org/10.1016/0584-8539(85)80224-5)
3. M. Simic, N. Paunovic, I. Boric, J. Randjelovic, S. Vojnovic, J. Nikodinovic-Runic, M. Pekmezovic, V. Savic, *Bioorg. Med. Chem. Lett*. **26** (2016) 235 (https://doi.org/10.1016/j.bmcl.2015.08.086)
4. F. J-J. Bihel, M. Hellal, J-J. Bourguignon, *Synthesis* **24** (2007) 3791 (https://doi.org/10.1055/s-2007-990893).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. *86* (7–8) 651–662 (2021)
JSCS–5451
Original scientific paper

Statistical optimization of bioethanol production from waste bread hydrolysate

KATARINA R. MIHAJLOVSKI[1]*, MARIJA MILIĆ[1], DANIJELA PECARSKI[2] and SUZANA I. DIMITRIJEVIĆ-BRANKOVIĆ[1]

[1]*University of Belgrade, Faculty of Technology and Metallurgy, Department for Biochemical Engineering and Biotechnology, Karnegijeva 4, Belgrade, Serbia and* [2]*Belgrade Academy of Professional Studies, Department Medical College of Professional Health, Cara Dušana 254, Belgrade, Serbia*

(Received 8 March, revised 20 April, accepted 22 April 2021)

Abstract: A recent trend in sustainable bioethanol production is the use of agricultural waste or food waste as an inexpensive and the most available feedstock. Bread waste is the major food waste that could be successfully used for the production of bioethanol. The aim of this study was to optimize ethanol production by the response surface methodology (RSM) using waste bread hydrolysate. Waste bread hydrolysate was obtained using crude hydrolytic enzymes that produce bacterial isolate Hymenobacter sp. CKS3. The influence of time of fermentation (24–72 h) and waste brewer's yeast inoculum (1–4 %) on ethanol production was studied. The optimal conditions, obtained by central composite design (CCD), were 48.6 h of fermentation and 2.85 % of inoculum. Under these conditions, a maximum of 2.06 % of ethanol concentration was reached. The obtained ethanol concentration was in good correlation, coefficient of 0.858, with yeast cell yield. The results obtained in this study imply that waste bread hydrolysate could be used as a biomass source for biofuel production with multiple benefits relating to environmental protection, reduction of production costs, and saving fossil fuels.

Keywords: waste bread; bioethanol; waste brewer's yeast; optimization; response surface methodology.

INTRODUCTION

During the last few decades, biofuels have gained more attention as an alternative to petroleum-based fuels.[1,2] Globalization and industrialization worldwide caused the depletion of fossil fuels that are major environmental pollutants. To reduce the negative environmental impact on climate changes caused by environ-

* Corresponding author. E-mail: kmihajlovski@tmf.bg.ac.rs
https://doi.org/10.2298/JSC210308032M

mental pollutions, researchers over the world are searching for alternative and renewable energy resources that could replace the use of fossil fuels.[3]

Bioethanol is one of the dominant biofuel worldwide and especially in the transportation sector.[4] Mixing bioethanol and gasoline leads to reduction of greenhouse gas and carbon dioxide emission.[5] It was estimated that by 2030, fuel ethanol will displace 10–20 % of the gasoline demand.[6]

Bioethanol is produced by fermentation of simple sugars present in biomass or sugars obtained by prior chemical or enzymatic treatment of the biomass. As the biomass or feedstock for bioethanol production, different agricultural raw materials that contain appreciable amounts of sugar like molasses or starch (corn, wheat, rice, potatoes, *etc*.) and lignocellulosic raw materials (straw, grasses, *etc*.) could be converted into sugar and then fermented into ethanol.[7,8] In Europe, starch-based materials are currently most utilized for bioethanol production[9] but starchy raw materials are very costly.[10] To overcome this problem, utilization of food waste, rather than waste bread, for bioethanol production is very attractive. Bread is the most used food in many countries worldwide but, unfortunately, it becomes a major part of food waste. Apart from its good nutritional value, bread is one of the most wasted foods. Although a good part of the waste bread is used as animal fodder, a high amount is still wasted annually.[11] Therefore, the utilization of this organic waste for the production of fermentable sugars could be an option to reduce the costs of bioethanol production. In addition, by generating a value-added product from waste biomass (waste bread), the principles of a circular economy are respected while solid waste disposal is reduced.[12]

Bread contains a significant amount of starch that is hydrolyzed to simple sugars using enzymes amylases.[13] Additionally, proteins in bread after enzymatic hydrolysis to peptides and amino acids could be used for yeast growth and fermentation to obtain bioethanol.[13,14]

To develop an economically feasible process of enzymatic hydrolysis and thus bioethanol production, it is necessary to replace the commercial enzymes with in-house produced enzymes by adequate microorganisms. Today, there is a large number of microorganisms, natural isolates, that produce industrially important enzymes.[15] These microorganisms can use different agricultural by-products for their growth and to produce desired enzymes which are then used in the process of hydrolysis. From a large number of unexplored genera, special attention is drawn to the genus *Hymenobacter*, the enzymatic potential of which has not been revealed. It has previously been shown that the CKS3 strain, previously isolated from soil, possesses the ability to produce crude enzyme amylase as well as other industrially important enzymes.[16] In line with this, this study presents a continuation of previous research. Thus, enzymatically treated waste bread, by crude CKS3 amylase, was used in the optimization of bioethanol production using waste brewer's yeast. For this purpose, a statistical tool, response surface

methodology (RSM), was used to describe the optimization of fermentation ethanol production. The parameters used in this optimization were the time of fermentation and yeast inoculum. Waste brewer's yeast, previously isolated from yeast biomass from the brewing industry, served as the inoculum for fermentation.

EXPERIMENTAL

Materials

Waste bread was kindly donated by the bakery "Skroz dobra pekara", Belgrade, Serbia. This waste bread (after shelf-life and without mold) was cut into cubes (2 cm×2 cm (±0.5 cm)).

Hydrolysis of waste bread

Enzymatic hydrolysis of waste bread was previously performed by crude enzyme amylase produced by *Hymenobacter* sp. CKS3, described in a previous work.[16] Hydrolysis was performed in an Erlenmeyer flask (300 mL) that contained 30 g of waste bread in 45 mL of 0.1 M acetate buffer pH 4.80. After sterilization (15 min, 120 °C), 75 mL of crude amylase was added at 50 °C. After 100 h of hydrolysis on an orbital shaker (200 rpm), the sample was centrifuged and the reducing sugars were quantified according to the method by Miller.[17]

Inoculum for ethanol fermentation

Waste brewer's yeast *Saccharomyces cerevisiae* was used for ethanol fermentation. The inoculum was prepared as described in a previous work by growing the yeast in malt extract broth (20 g L^{-1} malt extract) at 30 °C for 24 h.[18] A different amount of inoculum was used for ethanol fermentation according to the central composite design.

Optimization of ethanol production – central composite design

For ethanol production, waste bread hydrolysate was used. CCD was employed to evaluate the influence of two independent parameters: time of fermentation (A / h) and yeast inoculum concentration (B / %) on two dependent responses: ethanol production concentration ($Y1$ / %) and yeast cell yield ($Y2$, expressed as log (CFU / mL^{-1}). Each variable was studied at 3 different levels (Table I) and a set of 13 experiments were run (Table II).

TABLE I. Process parameters (ranges and levels) used in the CCD

Variable	–1	0	+1
A / h	24	48	72
B / %	1	2.5	4

The results showed that experimental data could be fitted with a quadratic polynomial model.

Analytical methods

The ethanol concentration was determined based on the density of alcohol distillate at 20 °C and expressed in wt. %.[19]

The number of viable yeast cells was determined by the indirect counting method – pour plate technique on malt extract agar plates. Serial dilutions of the samples were prepared, and after incubation at 30 °C for 48 h, the colonies were counted on malt extract agar plates. The changes in the viable cell number were calculated as:

$$\log (CFU / mL^{-1}) = \log ((CFU / mL^{-1})_1 - (CFU / mL^{-1})_0) \quad (1)$$

where $(CFU / mL^{-1})_1$ is the number viable cell per milliliter after fermentation and $(CFU / mL^{-1})_0$ is the number viable cell per milliliter before fermentation.

TABLE II. CCD of RSM for ethanol production with actual (experimental) values

Run	Variable		Experimental values	
	A / h	*B* / %	*Y1* / %	*Y2*, as log (*CFU* / mL^{-1})
1	24	1	0.24	8.87
2	48	2.5	2.04	9.69
3	72	4	1.11	9.36
4	48	2.5	1.96	9.72
5	24	2.5	1.65	9.49
6	48	2.5	2.12	9.58
7	48	4	1.63	9.48
8	72	2.5	1.74	9.66
9	48	2.5	1.92	9.75
10	24	4	1.34	9.53
11	48	1	0.82	9.23
12	48	2.5	2.03	9.66
13	72	1	0.75	9.47

RESULTS AND DISCUSSION

Optimization of ethanol production – central composite design

Yeast fermentation using waste bread hydrolysate is a process in which bioethanol, as a value-added product, is obtained. In order to obtain maximum ethanol concentration, it is necessary to optimize the production process using simultaneously different parameters with a small number of experiments. In line with this, RSM was used as a very powerful tool for understanding the interactions between the independent parameters on one or several responses.

Ethanol production was optimized using a waste bread hydrolysate with a reducing sugar yield of 19.89 g L^{-1}, obtained in a previous study.[16] A set of 13 experiments were performed according to a designed matrix for CCD (Table II). The equations that relate ethanol concentration (*Y1* / %) and yeast cell yield (*Y2* expressed as log (*CFU* / mL^{-1})) as dependent variables to the independent variables time of fermentation (*A* / h) and inoculum concentration (*B* / %) could be expressed as follows:

$$Y1 = 2.02 + 0.062A + 0.38B - 0.18AB - 0.34A^2 - 0.81B^2 \quad (2)$$

$$Y2 = 9.67 + 0.10A + 0.13B - 0.19AB - 0.077A^2 - 0.30B^2 \quad (3)$$

An analysis of variance ANOVA was performed to determine the significance of the regression model for the two responses (Table III). The robustness of the two models was determined by calculating the determination coefficient R^2 and for the first response – ethanol concentration (*Y1*), it was 0.9933 and for the second response – yeast cell yield (*Y2*), R^2 was 0.9710. The high values of both R^2 (values close to 1) suggested that the models were reliable and could explain

more than 99.33 and 97.10 %, respectively, of all variations. An adequate precision value higher than 4 (41.114 for *Y1* and 22.574 for *Y2*) indicate that the models could be used to predict the values for the responses.[20]

TABLE III. ANOVA table of the experimental results of the CCD (ethanol production); significant coefficient: $P < 0.05$

Source	*Y1*		*Y2*	
	F value	*P* value – *Prob* > *F*	*F* value	*P* value – *Prob* > *F*
Model	207.54	0.0001	46.95	0.0001
A / h	5.63	0.0493	20.83	0.0026
B / %	212.08	0.0001	37.03	0.0005
AB / h %	33.81	0.0007	51.46	0.0002
A^2 / h^2	79.41	0.0001	5.72	0.0480
B^2 / $\%^2$	448.83	0.0001	84.71	0.0001
Lack of fit	0.25	0.8602	0.25	0.8593
R^2	0.9933		0.9710	
Adjusted R^2	0.9885		0.9504	
Predicted R^2	0.9823		0.9291	
Coefficient of variation, %	4.28		0.56	
Adequate precision	41.114		22.574	

The relatively high adjusted determination coefficient adjusted R^2 (0.9885 for *Y1* and 0.9504 for *Y2*) refers to the significance of the tested models. Both response surface models for predicting ethanol production and yeast cell yield may be considered satisfactory.

The normal probability plot of the residuals *versus* externally studentized residuals for both models showed normally distributed points around a straight line, indicating that the models were suitable for use (Fig. 1a and b). The externally studentized residuals is a form of a Student's *t*-statistic, with the estimate of error varying between points, resulting from the division of a residual and its standard deviation.[20]

The regression analysis of the data (Table III) showed that ethanol production (*Y1*) was significantly affected by both tested parameters. The factors *A* and *B* were statistically significant ($P < 0.05$) as well as its quadratic coefficient and their interaction *AB* (time of fermentation-inoculum concentration). The influence of factors *A* and *B* on the response *Y1* (ethanol production) is presented on the perturbation plot (Fig. 2).

The quadratic coefficients A^2 and B^2 presented the maximum of the quadratic function indicating that ethanol production was highly influenced by these tested parameters. According to Eq. (2), parameter *B* (inoculum concentration) had the main influence on ethanol production, which describes a parabola with a pronounced maximum. With the increasing inoculum concentration up to 2.50 %, the ethanol production also increased. Higher inoculum concentration (2.5-2.85

%) also showed increasing ethanol production but after reaching a maximum of 2.85 %, the ethanol concentration decreased.

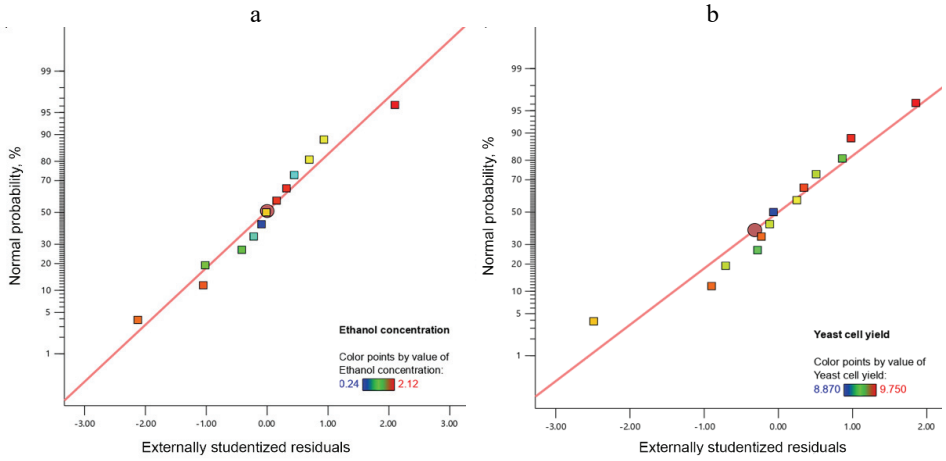


Fig.1. Normal probability plot of the residuals for: a) ethanol production (*Y1*) and b) yeast cell yield (*Y2*).

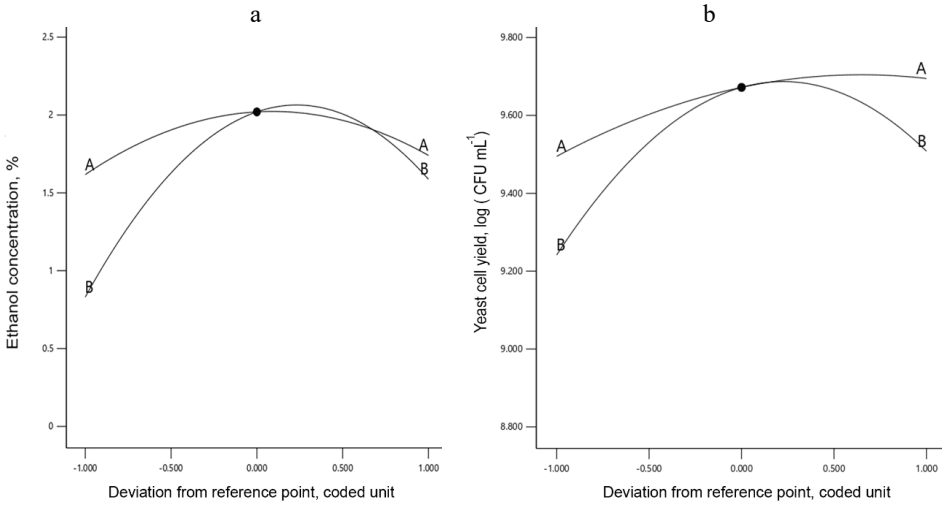


Fig. 2. Plot of the perturbation of time of fermentation (*A*) and inoculum concentration (*B*) on: a) ethanol production (*Y1*) and b) yeast cell yield (*Y2*).

The amount of inoculum for yeast fermentation is one of the crucial factors that influence ethanol production.[21] The increase in inoculum concentration leads to better utilization of reducing sugars by yeast, an increase in viable yeast biomass, and therefore the production of bioethanol.[21] However, further increase in

inoculum concentration above the optimum results in reducing the number of viable yeast cells that are directly related to ethanol production.[22] Similar to this result, Izmirlioglu and Demirci[23] reported that 3 % of yeast inoculum was optimal for obtaining maximum ethanol production from waste potato mash. Furthermore, for maximum ethanol production from carob extract, a 3 % inoculum of *S. cerevisiae* was used. However, a higher inoculum concentration (up to 10 %) was also reported in the literature.[24,25]

The time of fermentation (*A*) also had a statistically positive influence on ethanol production (Table III). Compared with parameter *B*, the time of fermentation, *A* had a slightly curved line but a positive influence on ethanol production (Fig. 2a). Increasing the time of fermentation (*A*), the ethanol concentration increased, reaching a maximum at 48 h after which the ethanol concentration decreased. Observing the interaction *AB* (time of fermentation and inoculum fermentation, Fig. 3), increasing the time of fermentation (*A*) and inoculum concentration (*B*), ethanol concentration also increased, reaching a maximum and after further increasing of both variables, the ethanol concentration was decreased. The close relationship of these factors on the ethanol production could be confirmed. Namely, the time of fermentation affected the growth of the yeast cell. If the fermentation time is short and inoculum is low, there are not enough cells to multiply and fermentation is incomplete, while longer fermentation time and higher inoculum than optimal values, gave too many cells. Furthermore, the obtained ethanol under these fermentation conditions had a toxic effect on yeast cells and caused cell death.[24]

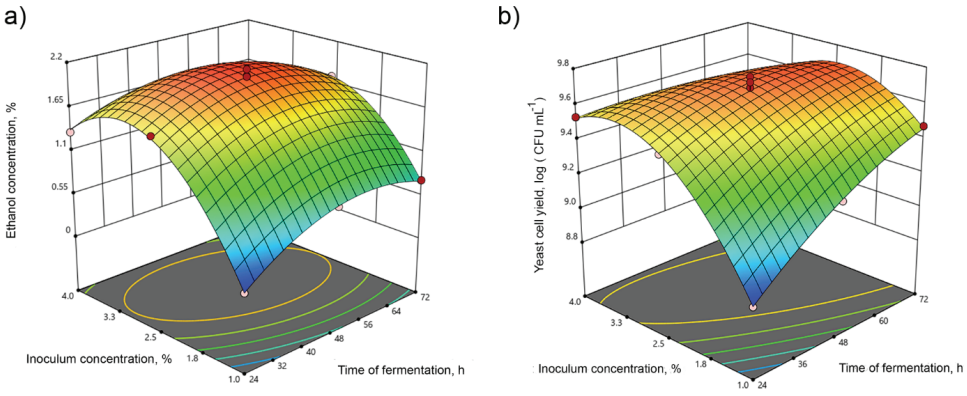


Fig. 3. The surface plot of time of fermentation and inoculum concentration on: a) ethanol production (*Y1*) and b) yeast cell yield (*Y2*).

During ethanol fermentation, the production of by-products (organic acids and higher alcohols) could decrease ethanol production by affecting some glycolytic intermediates in the corresponding metabolic pathways.[25] In addition, the

effect of "environmental stresses" on yeast cells could cause yeast cell inhibition and further ethanol production.[25]

On the industrial scale, a prolonged fermentation time could increase ethanol production costs due to the higher energy consumption.[21] A fermentation time longer than 48 h obtained in this study was reported in the work of Awolu and coworkers.[26] They obtained 1.8 % of ethanol after 7 days of fermentation of waste bread using baker's yeast *S. caerevisiae.*[26] A shorter fermentation time with optimal ethanol yield is preferable from an economic point of view.

Datta and coworkers[27] obtained a higher concentration of ethanol 5.4 % than 2.12 % obtained in this study. They used an enzyme mixture that contained glucoamylase and protease, produced by the fungus *Aspergillus niger*, for enzymatic hydrolysis of waste bread and obtained 145 g L^{-1} glucose. This concentration of glucose was higher than 19.89 g L^{-1} of reducing sugars obtained in waste bread hydrolysate which was used for yeast fermentation and thus the ethanol concentration was lower. In most the published studies, enzymatic hydrolysis of waste bread was performed using commercial enzymes. In this study, waste bread hydrolysate was obtained by crude hydrolytic enzymes, mainly amylase, that produced a bacterial isolate *Hymenobacter* sp. CKS3. Our previous study showed that waste bread hydrolysis to fermentable sugars by CKS3 crude amylase was not complete and the end products of starch (waste bread) hydrolysis were mainly dextrins with the addition of maltotriose, maltose, and glucose.[16] Therefore, other published studies reported higher values of obtained ethanol concentration. In line with this, 10 % of the ethanol from bread residues was reported in the work of Ebrahimi and co-workers[28] while 8.31 % of ethanol was obtained using wheat-rye bread hydrolysate.[29] In both published studies,[28,29] commercial amylases were used for waste bread hydrolyze that released a higher amount of reducing sugar that is necessary for yeast fermentation and obtained higher ethanol concentration than in the present study (2.12 %). However, the use of raw enzymes in hydrolysis is a cheaper and cleaner process. Moreover, it is worth mentioning that this is one of the first studies that deals with the statistical optimization of bioethanol production using waste bread hydrolysate obtained by the hydrolytic enzymes produced by *Hymenobacter* sp. CKS3.

At lower ethanol concentrations, higher growth of yeast cells was detected, while with the achievement of maximum concentrations of ethanol of 2.12 % a decreased number of yeast cells was noted. Even earlier, Lind and coworkers[30] reported that ethanol altered the polarity of the cell membrane thus inhibiting the viability of yeast cells.

According to the regression analysis, the yeast cell yield was affected by the time of fermentation and inoculum concentration. Both parameters, their interaction, and their quadratic coefficients were statistically significant. The perturbation plot (Fig. 2b) shows the influence of factors *A* and *B* on yeast cell yield.

The yeast cell yield gradually increased with increasing inoculum concentration but after reaching a maximum of 2.80 %, the yeast cell yield decreased (Fig. 2b). The time of fermentation had a positive influence on yeast cell yield (Fig. 2b). As could be noted from Table II (run 9), the maximum yeast cell yield (9.75log (*CFU* / mL^{-1})) was reached after 48 h of incubation and after that time, the yeast cell yield decreased. During fermentation, the yeast cells grow and multiply that lead to higher biomass (yeast cell yield).[31] The primary aim of a yeast cell is to reproduce and to produce cell biomass rather than produce ethanol.[32] Without significant yeast cell growth, ethanol cannot be produced.[32] On the other hand, yeast cells are sensitive to higher ethanol concentrations. As could be seen from Table II (run 6), after reaching a maximum concentration of produced ethanol (2.12 %), the yeast cell yield decreased. As mentioned before, a prolonged time of fermentation also had a negative effect on yeast cells, due to the accumulation of some toxic products.[24]

The high correlation coefficient of 0.890 between ethanol production (*Y1*) and yeast cell yield (*Y2*) showed that similar conditions affected both responses (Fig. 4). According to Bai[25] and Walker,[32] ethanol production is tightly linked with yeast cell growth.

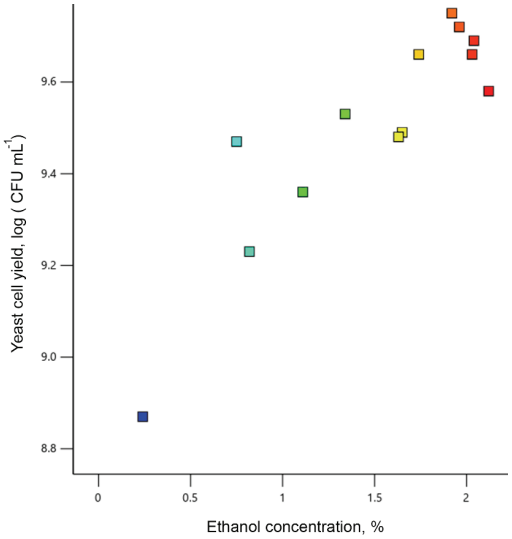


Fig. 4. Correlation between responses – ethanol production (*Y1*) and yeast cell yield (*Y2*).

Under the optimized conditions obtained from numerical optimization, the maximum ethanol production of 2.06 % was obtained after 48.68 h of fermentation with 2.85 % of inoculum. For this optimization, using software Design Expert 12, the criteria were to obtain maximum ethanol production with a selection of yeast cell yield "in range".

Validation of the models

As mentioned before, the main goal of this study was to optimize the conditions that favor maximum ethanol production. Validation of the experiments was performed under the optimized conditions obtained from the desirability function: 48.86 h of fermentation and 2.85 % of inoculum. For these parameters, additional experiments were performed. The obtained ethanol concentration 2.06±0.10 % and yeast cell yield, log (*CFU* / mL^{-1}) = 9.72±0.1, fitted within the 95 % prediction intervals (PI) for ethanol production (1.90–2.23) and yeast cell yield (9.55–9.83). These results demonstrated that both models were reliable and acceptable for use.

CONCLUSIONS

An increasing interest worldwide in an alternative source of energy consequently leads to the use of various waste substrates in bioethanol production. Valorization of waste bread for bioethanol production is very promising due to its large availability, non-cost and protection from environmental pollution. For industrial demand, it is important to use low-cost waste substrates and "in house" produced enzymes that are economically very accepted. In this study, a statistical design – CCD was used to optimize the culture conditions for yeast fermentation, using for the first time waste bread hydrolysate previously obtained using hydrolytic enzymes produced by *Hymenobacter* sp. CKS3. Under optimal conditions, the maximum of produced ethanol was 2.06 %. This is the first study that deals with the optimization of ethanol production on waste bread hydrolysate, previously obtained using hydrolytic enzymes produced by *Hymenobacter* sp. CKS3, in the biotechnological process of obtaining bioethanol with the aim to reduce the number of wastes. In addition, the enzymatic potential of a novel strain of *Hymenobacter* may contribute to the development of the production of biofuels.

Acknowledgements. This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Contract No. 451-03-9/2021-14/ /200135). We are grateful to "Skroz dobra pekara", Belgrade, Serbia, for providing the material (waste bread) for the experiments.

И З В О Д

ОПТИМИЗАЦИЈА ПРОИЗВОДЊЕ БИОЕТАНОЛА ПОМОЋУ ХЛЕБНОГ ХИДРОЛИЗАТА

КАТАРИНА Р. МИХАЈЛОВСКИ[1], МАРИЈА МИЛИЋ[1], ДАНИЈЕЛА ПЕЦАРСКИ[2] и СУЗАНА И. ДИМИТРИЈЕВИЋ-БРАНКОВИЋ[1]

[1]Универзитет у Београду, Технолошко–металуршки факултет, Катедра за биохемијско инжењерство и биотехнологију, Карнегијева 4, Београд и [2]Београдска академија професионалних студија, Висока здравствена школа струковних студија у Београду, Цара Душана 254, Београд

Недавни тренд у одрживој производњи биоетанола је коришћење пољопривредног отпада или отпада од хране као јефтине и лако доступне сировине. Отпадни хлеб, као главни отпад од хране, се може веома успешно користити за производњу биоетанола. Циљ овог рада био је оптимизација услова производње биоетанола етанолном фермен-

тацијом, односно ферментацијом квасца, на отпадном хлебном хидролизату, применом статистичке методе одзивне површине (RSM). Отпадни хидролизат хлеба је претходно добијен хидролизом отпадног хлеба помоћу сирових хидролитичких ензима, које производи бактеријски изолат *Hymenobacter* sp. CKS3. У овом раду, испитан је утицај времена ферментације отпадног пивског квасца (24-72 h) и концентрације инокулума отпадног пивског квасца (1–4 %) на производњу етанола користећи отпадни хлебни хидролизат. Оптимални услови, који су добијени применом централног композитног дизајна (CCD) у оквиру статистичке методе одзивне површине, били су 48,6 h ферментације и 2,85 % инокулума квасца. Под овим условима добијена је максимална концентрација етанола која износи 2,06 %. Ова концентрација етанола била је у добром коефицијенту корелације 0,858 са приносом ћелија квасца. Добијени резултати, у овој студији, указују на то да се хидролизат отпадног хлеба може веома успешно користити, као извор отпадне биомасе, за производњу биогорива. Оваква производња биогорива је еколошки и економски оправдана и утиче на смањење потрошње фосилних горива.

(Примљено 8. марта, ревидирано 20. априла, прихваћено 22. априла 2021)

REFERENCES

1. S. Tiwari, S. Jadhav, K. Tiwari, *Int. J. Environ. Sci. Technol.* **12** (2015) 3819 (https://doi.org/10.1007/s13762-014-0746-1)
2. S. Rezania, B. Oryani, J. Cho, A. Talaiekhozani, F. Sabbagh, B. Hashemi, P. F. Rupani, A. A. Mohammadi, *Energy* **199** (2020) 117457 (https://doi.org/10.1016/j.energy.2020.117457)
3. E. Demiray, S. E. Karatay, G. Dönmez, *Environ. Sci. Pollut. Res.* **26** (2019) 29366. (https://doi.org/10.1007/s11356-019-06020-1)
4. S. Nikolić, J. Pejin, L. Mojović, *Chem. Ind. Chem. Eng. Q.* **22** (2016) 1 (https://doi.org/10.2298/CICEQ151030001N)
5. H. S. Hafid, U. K. M. Shah, A. S. Baharuddin, A. B. Ariff, *Renew. Sustain. Energy Rev.* **74** (2017) 671 (https://doi.org/10.1016/j.rser.2017.02.071)
6. A. Walter, F. Rosillo-Calle, P. Dolzan, E. Piacente, K. B. da Cunha, *Biomass Bioenergy* **32** (2008) 730 (https://doi.org/10.1016/j.biombioe.2008.01.026)
7. P. Dinesha, S. Kumar, M. A. Rosen, *Environ. Sci. Pollut. Res.* **26** (2019) 8069 (https://doi.org/10.1007/s11356-019-04270-7)
8. J. Dodić, J. Grahovac, Z. Rončević, R. Pajović-Šćepanović, S. Dodić, B. Bajić, D. Vučurović, *J. Process. Energy Agric.* **22** (2018) 34 (https://doi.org/10.5937/JPEA1801034D)
9. C. A. Cardona, Ó. J. Sánchez, *Biores. Technol.* **98** (2007) 2415 (https://doi.opr/10.1016/j.biortech.2007.01.002)
10. J. R. Kwiatkowski, A. J. McAloon, F. Taylor, D. B. Johnston, *Ind. Crop Prod.* **23** (2006) 288 (https://doi.org/10.1016/j.indcrop.2005.08.004)
11. A. S. Demirci, I. Palabıyık, T. Gümüs, Ş. Özalp, *Waste Biomass Valor.* **8** (2017) 775 (https://doi.org/10.1016/j.indcrop.2005.08.004)
12. A. L. Young, *Environ. Sci. Pollut. Res*. **16** (2009) 117 (https://doi.org/10.1007/s11356-009-0106-8)
13. W. Pietrzak, J. Kawa-Rygielska, *Fuel* **134 (**2014) 250 (https://doi.org/10.1016/j.fuel.2014.05.081)
14. K. C. Thomas, W. Ingledew, *Appl. Environ. Microbiol.* **56** (1990) 2046 (https://doi.org/10.1007/BF02916454)

15. R. Ravindran, S. Hassan, G. Williams, A. Jaiswal, *Bioeng.* **5** (2018) 93 (https://doi.org/10.3390/bioengineering5040093)
16. K. Mihajlovski, M. Rajilić-Stojanović, S. Dimitrijević-Branković, *Renew. Energy* **152** (2020) 627 (https://doi.org/10.1016/j.renene.2020.01.101)
17. G. L. Miller, *Anal. Chem.* **31** (1959) 426 (https://doi.org/10.1021/ac60147a030)
18. K. Mihajlovski, Ž. Radovanović, M. Carević, S. Dimitrijević-Branković, *Fuel* **224** (2018) 591 (https://doi:10.1016/j.fuel.2018.03.135)
19. W. Horwitz, P. Chichilo, H. Reynolds, *Official methods of analysis of the Association of Official Analytical Chemists*, 1970 (https://www.cabdirect.org/cabdirect/abstract/19720492404)
20. D. C. Montgomery, *Design and analysis of experiments*, John Wiley & Sons, New York, 2008 (ISBN: 9781119113478)
21. M. A. Saeed, H. Ma, S. Yue, Q. Wang, M. Tu, *Environ. Sci. Pollut. Res.* **25** (2018) 28851 (https://doi.org/10.1007/s11356-018-2972-4)
22. T. W. Nagodawithana, K. H. Steinkraus, *Appl. Environ. Microbiol.* **31** (1976) 158 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC169741/pdf/aem00002-0022.pdf)
23. G. Izmirlioglu, A. Demirci, *Appl. Sci.* **2** (2012) 738 (https://doi.org/10.3390/app2040738)
24. S. H. Mohd Azhar, R. Abdulla, S. A. Jambo, H. Marbawi, J. A. Gansau, A. A. Mohd Faik, K. F. Rodrigues, *B. B. Reports* **10 (**2017) 52 (https://doi.org/10.1016/j.bbrep.2017.03.003)
25. F. Bai, W. Anderson, M. Moo-Young, *Biotechnol. Adv.* **26** (2008) 89 (https://doi.org/10.1016/j.biotechadv.2007.09.002)
26. O. O. Awolu, I. O. Ibileke, *Afr. J. Food Sci.* **5** (2011)148 (https://academicjournals.org/journal/AJFS/article-full-text-pdf/CBEEF3D3037)
27. P. Datta, S. Tiwari, L. Pandey, in *Utilization and Management of Bioresources*, Springer, Singapore, 2018 (https://doi.org/10.1007/978-981-10-5349-8)
28. F. Ebrahimi, M. Khanahmadi, S. Roodpeyma, M. J. Taherzadeh, *Biomass Bioenergy* **32** (2008) 333 (https://doi.org/10.1016/j.biombioe.2007.10.007)
29. J. Kawa-Rygielska, W. Pietrzak, A. Czubaszek, *Biomass Bioenergy* **44** (2012) 17 (https://doi.org/10.1016/j.biombioe.2012.04.016)
30. L. R. Lynd, H.-J. Ahn, G. Anderson, P. Hill, D. S. Kersey, T. Klapatch, *Appl.Biochem. Biotechnol.* **28** (1991) 549 (https://doi.org/10.1007/BF02922633)
31. M. E. Ojewumi, O. E. Kolawole, D. Oyekunle, O. S. Taiwo, A. Adeyemi, *J. Ecol. Eng.* **20** (2019) 35 (https://doi.org/10.12911/22998993/102614)
32. G. M. Walker, G. G. Stewart, *Beverages* **2** (2016) 30 (https://doi.org/10.3390/beverages2040030).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. *86* (7–8) 663–672 (2021)
JSCS–5452
Original scientific paper

Crystal structure of $K_3EuSi_2O_7$

SABINA Z. KOVAČ*, PREDRAG Z. DABIĆ and ALEKSANDAR S. KREMENOVIĆ

University of Belgrade, Faculty of Mining and Geology, Department of Mineralogy, Crystallography, Petrology and Geochemistry, Laboratory of Crystallography, Đušina 7, 11000Belgrade, Serbia

(Received 18 February, revised 15 March, accepted 31 March 2021)

Abstract: As part of research on the flux technique for growing alkali rare-earth elements (REE) containing silicates, tripotassium europium disilicate, $K_3EuSi_2O_7$, was synthesized and characterized by single-crystal X-ray diffraction. It crystallizes in the space group $P6_3/mcm$. In the crystal structure of the title compound, one part of the Eu cations are in a slightly distorted octahedral coordination and the other part are in an ideal trigonal prismatic coordination environment. The disilicate Si_2O_7 groups connect four EuO_6 octahedra and one EuO_6 trigonal prism. Three differently coordinated potassium cations are located between them. Silicates containing the larger rare earth elements usually crystallize in a structure that contains the rare-earth cation in both a slightly distorted octahedral and an ideal trigonal prismatic coordination environment.

Keywords: alkali rare-earth silicates; crystal structure; flux synthesis; single-crystal X-ray diffraction.

INTRODUCTION

Silicates containing rare-earth elements (REE silicates) have been the subject of intense research in recent decades. Many studies were performed on REE silicates because of their wide range of optical and magnetic properties combined with high thermal stability,[1–9] which make them very interesting and promising materials for different applications. REE silicates also exhibit broad structural diversity.[10] A highly adaptive framework can incorporate almost all REEs and this opens the possibility of synthesizing numerous structures and structural series of REE-containing silicates. Small but systematic variations in the cation radii among the lanthanides allow extensive cation substitutions.

Several ternary potassium–REE silicates with the stoichiometric composition $K_3REESi_2O_7$ that crystallize in the space groups $P6_3/mcm$ and $P6_3/mmc$ have hitherto been described: $K_3EuSi_2O_7$,[11] $K_3NdSi_2O_7$,[12] $K_3REESi_2O_7$ (REE = Gd–

* Corresponding author. E-mail: sabina.kovac@rgf.bg.ac.rs
https://doi.org/10.2298/JSC210218026K

–Lu),[13] $K_3ScSi_2O_7$,[14] $K_3YSi_2O_7$ (2 polymorphs),[6] $K_3SmSi_2O_7$[7] and $K_3ErSi_2O_7$ (2 polymorphs).[15]

During our investigations focused on the synthesis of new potassium-REE silicates, single crystals of $K_3EuSi_2O_7$ were obtained. $K_3EuSi_2O_7$ crystallizes in the hexagonal space group $P6_3/mcm$ (No. 193) and is isostructural with $K_3NdSi_2O_7$,[12] $K_3REESi_2O_7$ (REE = Gd–Yb),[13] $K_3YSi_2O_7$,[6] $K_3SmSi_2O_7$[7] and $K_3ErSi_2O_7$.[15] The reported compound was already known in the literature. Bondar and co-authors[11] stated the existence of this phase, however, they reported only the unit cell parameters (a = 9.98(1), c = 14.44(2) Å), space group ($P6_3/mcm$), chemical composition and described the morphology of crystals. Myers[16] outlined the successful synthesis of $K_3REESi_2O_7$ (REE = Eu, Y, Sc), but without any peer-reviewed data for $K_3EuSi_2O_7$. Thus, the crystal structure of this compound, $K_3EuSi_2O_7$, has not hitherto been described in the literature.

EXPERIMENTAL

Materials and measurements

All the materials were of analytical reagent grade and used as received without further purification. The crystal structure of $K_3EuSi_2O_7$ was analyzed by the single-crystal X-ray diffraction method. Diffraction experiments were conducted on a Gemini S (Oxford Diffraction) four-circle diffractometer equipped with a Mo-anode sealed tube and a Sapphire3 CCD detector. Diffraction data were processed with CrysAlis^{Pro}.[17] The synthesized products were tested using X-ray powder diffraction (XRPD). A small amount of material was finely ground and tested on a Philips PW1710 X-ray powder diffractometer at room temperature using Bragg–Brentano geometry and CuKα radiation. The diffractometer was operated at 40 kV and 30 mA, while the 2θ scan range was from 3 to 60°, with a step size of 0.02° and time per step of 1.25 s. PDXL 2 software[18] was used for analysis of the powder diffraction data. Scanning electron microscopy (SEM) was used for the morphological characterization. The analyses were performed on a single crystal using a JEOL JSM-6610LV instrument equipped with an Oxford INCA Energy 350 EDS detector.

Synthesis and crystallization

The following chemicals were used in the high-temperature flux synthesis experiment: Eu_2O_3 (Sigma–Aldrich, 99.99 %), SiO_2 (Merck, p.a.) and KF (Centrohem, 99 %). A total of 0.055 g of a mixture of Eu_2O_3 and SiO_2 in a molar ratio of 1:4 was added to 1 g of KF. The resulting mixture was homogenized in an agate mortar, transferred to a platinum crucible covered with a platinum lid and placed in a resistance-heated furnace. The mixture was heated to 1173 K at a rate of 200 K h^{-1} and kept for 10 h at that temperature. The temperature was lowered to 973 K at a rate of 2 K h^{-1} and then the furnace was switched off. After cooling to room temperature, the resulting material was manually removed from the crucible, washed with distilled water and dried in air. Bow-tie aggregate colorless crystals up to 300 μm in size embedded in a polycrystalline matrix were observed.

X-Ray structure determination

A selected elongated prismatic colorless single crystal was fixed on glass fiber using nail polish as glue and X-ray diffraction data were collected at room-temperature. Relevant information on the crystal data, data collection and structure refinement are compiled in Table I.

An analytical absorption correction based on indexed crystal faces using the procedure of Clark & Reid[19] in combination with an empirical absorption correction using spherical harmonics (implemented in SCALE3 ABSPACK scaling algorithm) was applied during the data reduction. The $K_3EuSi_2O_7$ crystallizes in the space group $P6_3/mcm$. Since it is isostructural with $K_3ErSi_2O_7$,[15] atomic coordinates of this structural analogue were used as a starting model for structure refinement. When Er^{3+} was replaced with Eu^{3+}, the model converged rapidly, and the crystal structure was refined to $R = 3.5$ %. The structure was refined on F^2 by full-matrix least-squares techniques using SHELXL programs.[20] Selected bond lengths and angles are presented in Table II. All figures of the crystal structures were prepared using the Vesta program.[21]

TABLE I. Experimental details

Crystal data		Data collection	
Chemical formula	$K_3EuSi_2O_7$	Diffractometer	Gemini S (Oxford Diffraction)
M_r	437.44	Absorption correction	Analytical + empirical
Crystal system	Hexagonal	Measured reflections	10007
Space group	$P6_3/mcm$	Independent reflections	528
Temperature, K	298	Observed ($I > 2\sigma(I)$) reflections	392
a / Å	9.9512 (3)	R_{int}	0.077
c / Å	14.4480 (4)	$(\sin\theta/\lambda)_{max}$ / Å^{-1}	0.641
α / °	90	Refinement	
γ / °	120	$R[F^2 > 2\sigma(F^2)]$	0.035
V / Å^3	1239.05 (8)	$wR(F^2)$	0.069
Z	6	S	1.13
Radiation type	MoKα	No. of reflections	528
μ / mm^{-1}	9.39	No. of parameters	40
Crystal size, mm	0.27×0.04×0.04	$\Delta\rho_{max}$, $\Delta\rho_{min}$ / e Å^{-3}	1.01, –0.69

TABLE II. Selected geometric parameters for $K_3EuSi_2O_7$. Symmetry codes: $^{i}x-y, -y, z$; $^{ii}-y+1, x-y, z$; $^{iii}-x+1, -x+y, z$; $^{iv}y, -x+y, -z$; $^{v}x, x-y, -z$; $^{v}-x+1, -x+y, -z+1/2$; $^{vii}x-y, -y, -z+1/2$; $^{viii}y, x-1, z$; $^{ix}y, x-1, -z+1/2$; $^{x}-x+y, y, -z$; $^{xi}y, x, z$; $^{xii}-y, -x, -z$; $^{xiii}-x, -y, -z$; $^{xiv}x-y, x-1, -z$; $^{xv}-x+1, -y, -z$; $^{xvi}y, x, -z+1/2$; $^{xvii}x, y, -z+1/2$; $^{(xviii}y+1, x, z$

Bond	Bond length, Å	Bond	Bond length, Å
K1–O1	2.770(5)	K3–O2^{xii}	2.754(7)
K1–O1^{i}	2.770(5)	K3–O2^{xiii}	2.754(7)
K1–O2^{i}	2.962(3)	K3–O2	2.754(7)
K1–O2	2.962(3)	Eu1–O1^{iv}	2.282(4)
K1–O1^{ii}	3.006(6)	Eu1–O1^{i}	2.282(4)
K1–O1^{iii}	3.006(6)	Eu1–O1^{xiv}	2.282(4)
K1–O1^{iv}	3.049(5)	Eu1–O1^{iii}	2.282(4)
K1–O1^{v}	3.049(5)	Eu1–O1^{viii}	2.282(4)
K2–O1^{i}	2.923(4)	Eu1–O1^{xv}	2.282(4)
K2–O1^{vi}	2.923(4)	Eu2–O2^{vii}	2.334(7)
K2–O1^{vii}	2.923(4)	Eu2–O2	2.334(7)
K2–O1^{iii}	2.923(4)	Eu2–O2^{xi}	2.334(7)
K2–O1^{viii}	2.923(4)	Eu2–O2^{i}	2.334(7)

TABLE II. Continued

Bond	Bond length, Å	Bond	Bond length, Å
K2–O1^{ix}	2.923(4)	Eu2–O2^{xvi}	2.334(7)
K2–O3^{viii}	3.029(3)	Eu2–O2^{xvii}	2.334(7)
K2–O3	3.029(3)	Si1–O1^{iii}	1.616(5)
K2–O3^{iii}	3.029(3)	Si1–O1^{ii}	1.616(5)
K3–O2^{i}	2.754(7)	Si1–O2^{xviii}	1.622(7)
K3–O2^{x}	2.754(7)	Si1–O3	1.661(4)
K3–O2^{xi}	2.754(7)		
Bond	Bond angle, °	Bond	Bond angle, °
O1^{iii}–Si1–O1^{ii}	111.0(4)	O1^{ii}–Si1–O3	106.7(2)
O1^{iii}–Si1–O2^{xviii}	110.9(2)	O2^{xviii}–Si1–O3	110.5(4)
O1^{ii}–Si1–O2^{xviii}	110.9(2)	Si1–O3–Si1^{xvii}	136.7(6)
O1^{iii}–Si1–O3	106.7(2)		

RESULTS AND DISCUSSION

Description of the crystal structure

As a result of an extensive study of the flux technique for growing new alkali rare-earth elements (REE) containing silicates, $K_3EuSi_2O_7$ was synthesized and characterized by single-crystal X-ray diffraction. The compound crystallizes in the space group $P6_3/mcm$ and is isostructural with disilicates $K_3ErSi_2O_7$[15] and $K_3YbSi_2O_7$.[22] The crystal structure is comparable to the sorosilicate structure type reported earlier by Vidican and co-authors.[13]

The crystal structure of the title compound consists of Si_2O_7 groups and EuO_6 polyhedra that form a 3D framework with potassium ions in the holes (Fig. 1).

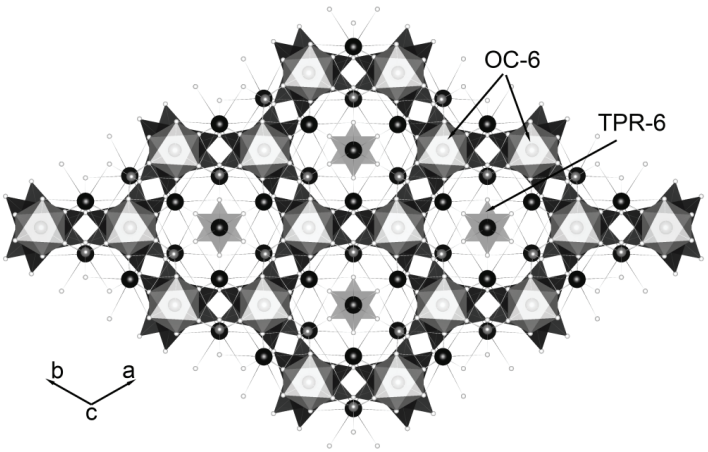


Fig. 1. The framework of the crystal structure of $K_3EuSi_2O_7$ along [001]. Color code: K atoms and polyhedra are black, EuO_6 polyhedra are grey (octahedra: OC-6; trigonal prisms: TPR-6), Si atoms are dark grey and O atoms are white.

The disilicate groups Si_2O_7 connect four EuO_6 octahedra (OC-6), as well as one EuO_6 trigonal prism (TPR-6, Fig. 2). Three differently coordinated potassium cations are located between them. The structure can also be described as regular alternation of two types of layers, which are parallel to the (001) plane: (I) octahedral layers and (II) sorosilicate layers, formed by a mixture of Si_2O_7 groups and K2 O_{6+3} polyhedra (Fig. 3).

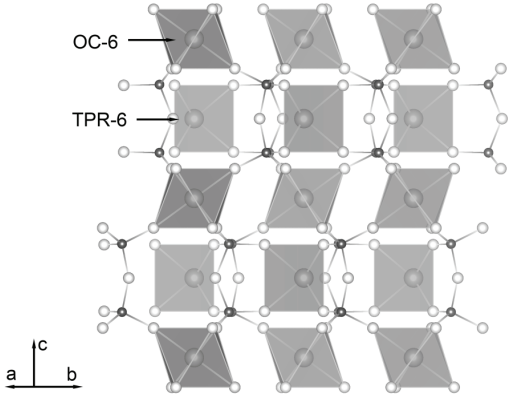


Fig. 2. The Si_2O_7 groups and EuO_6 polyhedra (octahedra: OC-6; trigonal prisms: TPR-6) in the crystal structure of $K_3EuSi_2O_7$ oriented approximately along [120] (*c*-axis is vertical) showing their connections. Potassium cations have been omitted for clarity. Color code: Eu atoms and polyhedra are grey, Si atoms are dark grey and O atoms are white.

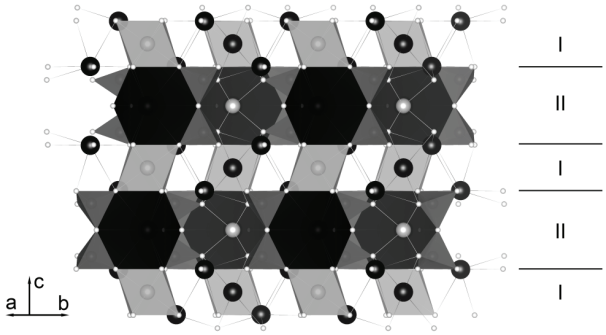


Fig. 3. Projection of structure along [120] showing two types of regularly alternating layers, which are parallel to the (001) plane: (I) octahedral layers and (II) sorosilicate layers. K atoms and polyhedra are black, Eu atoms and polyhedra are grey, SiO_4 polyhedra are dark grey and O atoms are white.

The structure is characterized by two crystallographically distinct six-coordinated sites occupied by Eu1 and Eu2: $Eu1O_6$ is an octahedron (OC-6) and $Eu2O_6$ is a trigonal prism (TPR-6). Two types of polyhedra around europium cations are distinguished by a small difference in the Eu–O distances: all Eu1–O distances are 2.282(4) Å, while all Eu2–O bond lengths are 2.334(7) Å. Each $Eu1O_6$ octahedron (OC-6) shares its six corners with six different Si_2O_7 groups, while each Er_2O_6 trigonal prism (TPR-6) also shares six corners, but only with three different Si_2O_7 groups.

Each Si_2O_7 group is connected to five different cations, *i.e.*, four Eu1 and one Eu2. The disilicate group consists of two SiO_4 tetrahedra sharing an O atom, denoted O3, with a bridging bond angle Si–O–Si of 136.7(6)°. The observed individual Si–O distances are within the acceptable range for silicate structures. The bridging Si–O bond (1.661(4) Å) is expectedly longer than the nonbridging Si–O bonds (1.616(5) and 1.622(7) Å). The O–Si–O bond angles range from 106.7(2) to 111.0(4)°.

The three potassium cations show three different coordination environments. The coordination sphere of K1 includes eight oxygen ions and forms a $K1O_{2+4+2}$ coordination polyhedron that could be described as a distorted hexagonal pyramid with a split apex (distances range from 2.770(5) to 3.049(5) Å). The K2 atom is coordinated by nine oxygen atoms with distances from 2.923(4) to 3.029(3) Å providing a regular tricapped trigonal prism ($K2O_{6+3}$). The coordination of the K3 atom (CN=6) could be described as an undistorted trigonal antiprism (K–O length = 2.754 (7) Å).

Bond-valence calculations showed that the bond valences are well balanced and the K–O, Eu–O and Si–O bond lengths are consistent with the presence of K^+, Eu^{3+}, Si^{4+} and O^{2-} in the structure. Only the bond-valence sum for Eu1 is oversaturated (3.42 v.u.), which could be attributed to the environment of the $Eu1O_6$ octahedron. All oxygen atoms from the $Eu1O_6$ octahedron are common to all adjacent polyhedra around the K1 and K2 atoms for which BVS shows slightly undersaturated values. This means that slightly deviating bonds are formed in order to satisfy the local valence disagreement due to structural disorder. These values are in accordance with the literature data for $K_3REESi_2O_7$[13] that contains the rare-earth cation in both slightly distorted octahedral and an ideal trigonal prismatic coordination environment. The results of bond-valence-sum calculations (VaList software;[23] bond valence parameters: Brown & Altermatt[24] and Brese & O'Keeffe[25]), are presented in Table III.

TABLE III. Bond valence sums for the cations and anions in $K_3EuSi_2O_7$

Atom	Bond valence, v. u.						Σv_{ij} / v. u.
	K1[a]	K2[a]	K3[a]	Eu1[a]	Eu2[a]	Si[b]	
O1	0.178×2	0.118×6		0.57×6		1.022×2	2.066
	0.094×2						
	0.084×2						
O2	0.106×2		0.186×6		0.495×6	1.005	1.898
O3		0.089×3				0.905	1.988
Σv_{ij} / v. u.	0.924	0.975	1.116	3.42	2.97	3.954	

[a]Brown & Altermatt[24]; [b]Brese & O'Keeffe[25]

X-Ray powder diffraction (XRPD)

Collected XRPD data are shown in Fig. 4. A diffraction pattern calculated using the CIF file obtained from the single-crystal data of $K_3EuSi_2O_7$ was overlaid. The experimental XRPD patterns are in good agreement with those calculated from the CIF.

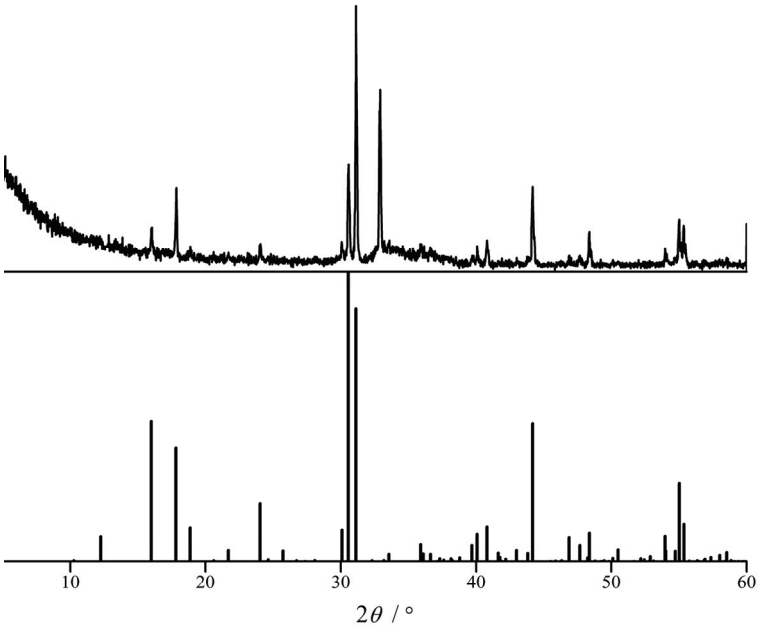


Fig. 4. X-ray powder diffraction pattern of $K_3EuSi_2O_7$. Upper pattern represents experimental data while lower pattern represents overlaid positions of Bragg's peaks calculated from $K_3EuSi_2O_7$ CIF.

Scanning electron microscopy (SEM)

The crystal morphology of a single $K_3EuSi_2O_7$ crystal obtained by SEM is shown in Fig. 5. An elongated prismatic crystal is clearly visible. The crystal

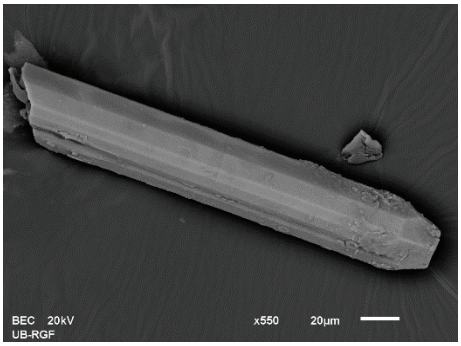


Fig. 5. A scanning electron microscope (SEM) picture of a $K_3EuSi_2O_7$ crystal.

appears to be homogeneous. Similar crystal morphology was also reported by Bondar and co-authors,[11] although their crystals were much thicker and less elongated.

CONCLUSIONS

Single crystals of $K_3EuSi_2O_7$ were prepared by high-temperature flux crystal growth and characterized by single-crystal X-ray diffraction, powder X-ray diffraction and SEM analysis. The compound crystallizes like disilicates containing larger REEs, *i.e.*, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb and Y, with the exception of Er and Y, which can crystallize as polymorphs in space groups $P6_3/mcm$ and $P6_3/mmc$. Their effective ionic radii ranged from 0.983 Å for $^{VI}Nd^{3+}$ to 0.868 Å for $^{VI}Yb^{3+}$ ($^{VI}Eu^{3+}$ = 0.947 Å).[26] The structure of these compounds contain the rare-earth cation in both a slightly distorted octahedral and an ideal trigonal prismatic coordination environment.

SUPPLEMENTARY MATERIAL

The crystallographic data for this paper have been deposited at the Cambridge Crystallographic Data Centre under deposition number CCDC 2062974.

Acknowledgements. The authors gratefully acknowledge financial support from the Ministry of Education, Science and Technological Development of the Republic of Serbia (Grant No. III45007 and III45015); the Ministry of Education, Science and Technological Development of the Republic of Serbia & the Austria Federal Ministry of Education, Science and Research (Project No. 451-03-02141/2017-09/14; WTZ project No. SRB 14/2018). The authors also thank the Laboratory for Scanning Electron Microscopy at the Department of Mineralogy, Crystallography, Petrology and Geochemistry, University of Belgrade, Faculty of Mining and Geology and the Laboratory of X-ray diffraction at the Department of Physics, University of Novi Sad, Faculty of Sciences.

ИЗВОД

КРИСТАЛНА СТРУКТУРА $K_3EuSi_2O_7$

САБИНА З. КОВАЧ, ПРЕДРАГ З. ДАБИЋ и АЛЕКСАНДАР С. КРЕМЕНОВИЋ

Универзитет у Београду, Рударско–геолошки факултет, Департман за минералогију, кристалографију, петрологију и геохемију, Лабораторија за кристалографију, Ђушина 7, 11000 Београд

У оквиру испитивања алкалних силиката елемената ретких земаља добијених методом флукса из растопа, синтетисан је трикалијум еуропијум дисиликат, $K_3EuSi_2O_7$. Кристална структура одређена је методом рендгенске дифракције на монокристалу. Кристалише у просторној групи $P6_3/mcm$. У кристалној структури испитиваног једињења, део Eu катјона је у благо дефорисаној октаедарској координацији, а други део је у идеалном тригонално-призматичном координационом окружењу. Дисиликатне групе Si_2O_7 повезују четири EuO_6 октаедра и једну EuO_6 тригоналну призму. Између њих се налазе три различито координисана катјона калијума. Силикати који садрже веће јоне ретких земаља имају структуру која садржи катјон ретке земље у два различита координациона окружења: у благо искривљеном октаедру и у идеалном тригонално-призматичном окружењу.

(Примљено 18. фебруара, ревидирано 15. марта, прихваћено 31. марта 2021)

REFERENCES

1. J. Felsche, *In The Crystal Chemistry of the Rare-Earth Silicates. Rare Earths. Structure and Bonding*, Vol. 13, Springer, Heidelberg, 1973 (https://doi.org/10.1007/3-540-06125-8_3)
2. A. Kitai, *In Luminescent Materials and Applications*, John Wiley & Sons Ltd., Chichester, 2008 (Online ISBN:9780470985687)
3. A. M. Latshaw, W. M. Chance, N. Trenor, G. Morrison, M. D. Smith, J. Yeon, D. E. Williams, H.-C. zur Loye, *CrystEngComm* **17** (2015) 4691 (https://doi.org/10.1039/C5CE00630A)
4. A. M. Latshaw, K. D. Hughey, M. D. Smith, J. Yeon, H.-C. zur Loye, *Inorg. Chem.* **54** (2015) 876 (https://doi.org/10.1021/ic502185b)
5. A. M. Latshaw, B. O. Wilkins, K. D. Hughey, J. Yeon, D. E. Williams, T. T. Tran, P. S. Halasyamani, H.-C. zur Loye, *CrystEngComm* **17** (2015) 4654 (https://doi.org/10.1039/C5CE00671F)
6. A. M. Latshaw, G. Morrison, K. D. zur Loye, A. R. Myers, M. D. Smith, H.-C. zur Loye, *CrystEngComm* **18** (2016) 2294 (https://doi.org/10.1039/C6CE00177G)
7. A. M. Latshaw, J. Yeon, M. D. Smith, H.-C. zur Loye, *J. Solid State Chem.* **235** (2016) 100 (https://doi.org/10.1016/j.jssc.2015.12.013)
8. G. Morrison, A. M. Latshaw, N. R. Spagnuolo, H.-C. Zur Loye, *J. Am. Chem. Soc.* **139** (2017) 14743 (https://doi.org/10.1021/jacs.7b08559)
9. B. R. Figueiredo, A. A. Valente, Z. Lin, C. M. Silva, *Micropor. Mesopor. Mat.* **234** (2016) 73 (https://doi.org/10.1016/j.micromeso.2016.07.004)
10. F. Liebau, *Structural chemistry of silicates: structure, bonding and classification*, Springer, Heidelberg , 1985, p. 347 (https://doi.org/10.1007/978-3-642-50076-3)
11. I. A. Bondar, T. F. Tenisheva, Y. F. Shepelev, N. A. Toropov, *Dokl. Akad. Nauk SSSR* **160** (1965) 1069 (http://www.mathnet.ru/links/df924f2db1305a2430f200e3f58341c7/dan30741.pdf)
12. M. S. Hwang, H. Y.-P. Hong, M. C. Cheng, Y. Wang, *Acta Cryst., C* **43** (1987) 1241 (https://doi.org/10.1107/S0108270187092308)
13. I. Vidican, M. Smith, M., H.-C. zur Loye, *J. Solid State Chem.* **170** (2003) 203 (https://doi.org/10.1016/S0022-4596(02)00029-4)
14. J. D. Napper, R. C. Layland, M. D. Smith, H. Loye, *J. Chem. Crystallogr.* **34** (2004) 347 (https://doi.org/10.1023/B:JOCC.0000028666.53348.fc)
15. P. Dabić, M. G. Nikolić, S. Kovač, A. Kremenović, *Acta Cryst., C* **75** (2019) 1417 (https://doi.org/10.1107/S2053229619011926)
16. A. Myers, *J. South Carolina Acad. Sci.* **12** (2014) 200 (https://scholarcommons.sc.edu/jscas/vol12/iss1/1)
17. *Rigaku Oxford Diffraction, CrysAlis^{Pro} Software system*, Rigaku Corporation, Oxford, 2018
18. *Rigaku PDXL 2: Integrated powder X-ray diffraction software. Version 2.8.3.0*, Rigaku Corporation, Tokyo, 2007 https://www.rigaku.com/en/service/software/pdxl
19. R. C. Clark, J. S. Reid, *Acta Cryst., A* **51** (1995) 887 (https://doi.org/10.1107/S0108767395007367)
20. G. M. Sheldrick, *Acta Cryst., C* **71** (2015) 3 (https://doi.org/10.1107/S2053229614024218)
21. M. Momma, F. Izumi, *J. Appl. Cryst.* **44** (2011) 1272 (https://doi.org/10.1107/S0021889811038970)

22. P. Dabić, V. Kahlenberg, B. Krueger, M. Rodić, S. Kovač, J. Blanuša, Z. Jagličić, Lj. Karanović, V. Petríček, A. Kremenović, *Acta Cryst., B*, under review
23. A. S. Wills, *VaList*, 2010, Program available from www.CCP14.ac.uk
24. I. D. Brown, D. Altermatt, *Acta Cryst., B* **41** (1985) 244 (https://doi.org/10.1107/S0108768185002063)
25. N. E. Brese, M. O'Keeffe, *Acta Cryst., B* **47** (1991) 192 (https://doi.org/10.1107/S0108768190011041)
26. R. D. Shannon, *Acta Cryst., A* **32** (1976) 751 (https://doi.org/10.1107/S0567739476001551).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. *86* (7–8) 673–684 (2021)
JSCS–5453
Original scientific paper

Use of GA-ANN and GA-SVM for a QSPR study on the aqueous solubility of pesticides

AMEL BOUAKKADIA[1,2]*, NOUREDDINE KERTIOU[1,2], RANA AMIRI[1], YOUSSOUF DRIOUCHE[1] and DJELLOUL MESSADI[1]

[1]*Environmental and Food Safety Laboratory, Department of Chemistry, Badji Mokhtar University – Annaba, BP. 12, 23000 Annaba, Algeria and* [2]*Abbes Laghrour University, Faculty of Sciences and Technology – Khenchela, BP 1252 Route de Batna, 40004 Khenchela, Algeria*

(Received 18 June, revised 19 August, accepted 6 October 2020)

Abstract: The partitioning tendency of pesticides, in this study herbicides in particular, into different environmental compartments depends mainly of the physicochemical properties of the pesticides itself. Aqueous solubility (S) indicates the tendency of a pesticide to be removed from soil by runoff or irrigation and to reach surface water. The experimental procedure for determining the aqueous solubility of pesticides is very expensive and difficult. QSPR methods are often used to estimate the aqueous solubility of herbicides. The artificial neural network (ANN) and support vector machine (SVM) methods, always associated with selection of a genetic algorithm (GA) of the most important variable, were used to develop QSPR models to predict the aqueous solubility of a series of 80 herbicides. The values of log S of the studied compounds were well correlated with the descriptors. Considering the pertinent descriptors, a Pearson correlation squared coefficient (R^2) of 0.8 was obtained for the ANN model with a structure of 5-3-1 and 0.8 was obtained for the SVM model using the RBF function for the optimal parameters values: C = 11.12; σ = 0.1111 and ε = 0.222.

Keywords: genetic algorithm; agrochemicals; descriptors; statistical methods.

INTRODUCTION

Pesticide distribution and fate in various environmental media and compartments is strongly influenced by the inherent properties of the compounds themselves, particularly by the basic physicochemical properties, such as solubility in water, vapor pressure (p_v) and partitioning coefficients between the organic matter (in soil or sediment) and water.[1]

* Corresponding author. E-mail: amelbouakkadia@yahoo.fr
https://doi.org/10.2298/JSC200618066B

The water solubility reflects the maximum amount of a chemical that will dissolve in pure water at a given temperature. The water solubility is one of the most important physicochemical properties in ecological hazard and exposure assessments, including environmental fate. The spatial and temporal movement (mobility) of a substance within and between the environmental compartments of soil, water and air depends largely on its solubility in water.

The knowledge of the solubility in water is essential when estimating biological degradation, bioaccumulation, hydrolysis, adsorption and the octanol/ /water partition coefficient. Highly water-soluble chemicals are potentially easier distributed by the hydrologic cycle, as they tend to have relatively low adsorption coefficients (*i.e.*, low adsorption to soil and sediment).[2] However, due to the complexity of analytical methods and the high cost of experiments, the application of theoretical predictive methods, which are fast, convenient and cost-effective, for preliminary assessment and estimation of aqueous solubility of pesticides have gained much attention.

Quantitative structure–property relationships (QSPR) models are found on the basis of the correlation between the experimental values of the physicochemical properties and descriptors reflecting the molecular structure of the compounds. In QSPR studies, a regression model of the form ($y = Xb + e$) may be used to describe a set of predictor variables (X) with a predicted variable (y) by means of a regression vector (b).[3]

The objective of the present study was to develop valid QSPR models for the aqueous solubility, S, of pesticides. Artificial neural network (ANN) and support vector machine (SVM) were applied for modeling the quantitative relationship between the aqueous solubility and the structural descriptors of pesticides. Previous to the generation of the ANN and SVM models, a genetic algorithm (GA) was used for descriptor selection. The strength and the predictive performance of the proposed models were verified using both internal (cross-validation and Y-scrambling) and external statistical validations.

EXPERIMENTAL

Data set

The data set in this study comprises diverse chemical classes of 80 pesticides. The data were collected from the literature,[2] and were compiled in the units of mg L^{-1}, and presented as the logarithm of S, the values of which ranged from -1.04576 to 5.90091.

The data set was randomly divided into a training set of 58 compounds and a test set of 22 compounds. The names of the compound and their aqueous solubility are given in Table S-I of the Supplementary material to this paper.

Software

Geometry optimization was performed using HyperChem 6.03.[4] Dragon software was utilized to calculate the molecular descriptors. ANN and SVM regression were performed in the Molegro software.[5]

Molecular optimization and descriptor generation

The molecular modeling software HyperChem 6.03[4] was used to represent the molecules and then, the semi-empirical AM1 method was used to obtain the final geometry. All calculations were performed under RHF formalism[6] without configuration interaction. The molecular structures were optimized using the Polak–Ribiere algorithm for criterion with a root mean square gradient of 0.001 kcal* mol^{-1}. The optimized geometries were then transferred to the Dragon computer software Version 5.3[7] to calculate 1201 descriptors belonging to different classes. Using the corresponding DRAGON software options, constant values of the descriptors (standard deviations less than 0.0001), which provide no information, were eliminated first and then those that were highly correlated ($R \geq 0.95$) as they convey redundant information. For each pair of correlated descriptors, the one with the highest cross-correlation with other descriptors was automatically eliminated. In this work, the genetic algorithm (GA) variable subset selection method in the MOBYDIGS release of Todischini[8] was used for the selection of the most relevant descriptors from the pool of remaining descriptors. These descriptors were used as inputs of ANN and SVM for the construction of QSPR models, as in similar work.[9]

Genetic algorithm

Modeling of the genetic process initiated the development of genetic algorithms, which could be exploited in a variety of optimization problems.[10,11] In this case, a potential solution is considered as an individual in a population. The value of the cost function associated with a measurement solution "adaptation" of the individual related to its environment. A genetic algorithm simulates the evolution, over several generations, of an initial population whose individuals are poorly adapted using genetic operators of reproductions and mutations. After a number of generations, the population consists of well adapted individuals, *i.e.*, the supposed "good" solutions to the optimization problem. From a statistical view point, the ratio of the number of samples (n) to the number of descriptors (m) should not be too low. Usually, it is recommended that $n/m \geq 5$.[12] The GA was stopped when increasing the model size did not significantly increase the Q^2 value.

Artificial neural networks

An artificial neural network is modeled on the biological neural networks found in the central nervous system of animals.[13] An artificial neural network is a mathematical and numerical method based on biological neural networks. An ANN consists of some connected neurons and process information. A network is made of one input layer, one output layer and may also consist of some hidden layers. Each layer is made of some neurons connected to other neurons in previous and subsequent layers. A neuron has an input, an output and a transfer function. The sigmoidal transfer function is one of the performed functions, expressed by the following equation:

$$a_j = \frac{1}{1+e^{-S_j}} \tag{1}$$

where a_j is the output of the j^{th} neuron and S_j is the input of the j^{th} neuron, produced by outputs of the previous layer. S_j is given as:

$$S_j = \sum_{i=1}^{n}\left(w_{ij}a_i\right) + b_j \tag{2}$$

* 1 kcal = 4184 J

where a_i are the outputs of the i^{th} neuron from the previous layer, w_{ij} presents the weights applied to the connection of the i^{th} and j^{th} neuron, and b_j is a bias number.[14]

An ANN is an adaptive network that changes its structure based on external or internal information that flows through the network during the learning (training) phase. Estimation of optimum weights and biases of network needs an algorithm, called the propagation method. Several kinds of propagation methods are available and back propagation (BP) is the easiest and simple one with enough reliability. BP and other usual propagation methods are explained completely in mathematical literatures.[15,16]

Artificial neural networks are capable of recognizing highly non-linear relationships contrary to MLR. When compared with the traditional statistical methods, the flexibility of an ANN enables more complex relationships in the experimental data to be discoverd.[17]

Support vector machine

SVM is a new and very promising classification and regression method developed by Vapnik.[18] SVMs were originally developed for classification problems; they can also be extended to solve nonlinear regression problems by the introduction of an ε-insensitive loss function. In the support vector regression, input x is first mapped into a higher dimensional feature space by using a kernel function, and then a linear model is constructed in this feature space. The kernel functions often used in SVM include linear, polynomial, radial basis function, and sigmoid function. The general form of the SVR-based regression function can be written as:[19,20]

$$f(x,w) = f\left(x,\alpha_i,\alpha_i^*\right) = \sum_{i=1}^{N}\left(\alpha_i - \alpha_i^*\right)K\left(x_i,\alpha_j\right) + b \tag{3}$$

where both α_i and α_i^* are Lagrange multipliers. According to the Karush–Kuhn–Tucker conditions, only a minority sample coefficients are non-zero values, the data points corresponding to them are called support vectors. These support vectors are the samples that can determine the hyper plane.[13,21] $K(x,x_i)$ is the Kernel function.[22] Any function satisfying the Mercer condition can be used as the Kernel function.[23] In this work, the Gaussian radial basis function (RBF) Kernel was used in the SVM as below:

$$K(x,x_i) = e^{-\|x-x_i\|^2/\sigma^2} \tag{4}$$

where σ^2 is the width of the Gaussian function, so C and σ that are the relative weights of the regression error and the kernel parameter of the RBF kernel, should be optimized by the user, to obtain the support vector. The parameters of SVMR were optimized by systemically changing their values in the training step and calculating the *RMSE* and accuracy of the model using 5-fold cross-validation. The optimized values of C, σ and ε were 11.12, 0.1111 and 0.222, respectively, obtained based on the minimum *RMSE* and maximum accuracy of the model.

A detailed description of the theory of SVM can be found in several excellent books and tutorials.[24,25]

An additional external validation according to Golbraikh and Tropsha[26] was applied solely to the test set. According to the recommended criteria of Tropsha *et al*., a predictive QSPR model must meet the following conditions:[27]

$$Q^2_{EXT} > 0.5 \tag{5}$$

$$R^2 > 0.6 \tag{6}$$

$$(R^2 - R^2_0)/R^2 < 0.6 \text{ and } 0.85 < k < 1.15 \tag{7}$$

or

$$(R^2–R'^2_0)/R^2 < 0.6 \text{ and } 0.85 < k' < 1.15 \quad (8)$$

where

$$R = \frac{\sum(y_i - \bar{y})(\tilde{y}_i - \bar{\tilde{y}})}{\sqrt{\sum(y_i - \bar{y})^2 \sum(\tilde{y}_i - \bar{\tilde{y}})^2}}, R_0^2 = 1 - \frac{\sum(y_i - y_i^{r0})^2}{\sum(y_i - \bar{y})^2}, R_0'^2 = 1 - \frac{\sum(\tilde{y}_i - \tilde{y}_i^{r0})^2}{\sum(\tilde{y}_i - \bar{\tilde{y}})^2},$$

$$k = \frac{\sum(y_i \tilde{y}_i)}{\sum(\tilde{y}_i)^2}, k' = \frac{\sum(y_i \tilde{y}_i)}{\sum(y_i)^2}$$

where R is the correlation coefficient between the calculated and experimental values in the test set; R_0^2 (calculated *vs*. observed values) and $R_0'^2$ (observed *vs*. calculated values) are the coefficients of determination; k and k' are slopes of regression lines through the origin of calculated *vs*. observed and observed *vs.* calculated, respectively; y_i^{r0} and $\tilde{y}_i^{r0}$ are defined as:

$$y_i^{r0} = k\tilde{y} \quad (5)$$

and

$$\tilde{y}_i^{r0} = k'y \quad (6)$$

respectively; and the summations are over all samples in the test set.

The reason to use R_0^2 and require k values that are close to 1 is that when actual *vs.* predicted properties are compared, an exact fit is required, not just a correlation.

RESULTS AND DISCUSSION

The GA parameters that were used in this study are: population size 100 and maximum generations 100. Five descriptors were selected by GA (Table S-II of the Supplementary material), which are: MATS8m, RNCG, AlogP2, MAXDN and Mor26u. Table I gives e short descriptions of these descriptors.

TABLE I. Description of the selected descriptors by GA

No	Symbol	Class	Meaning
1	MATS8m	2D autocorrelation indices	Moran autocorrelation – lag 8 / weighted by atomic masses
2	RNCG	Charge descriptors	Relative negative charge
3	AlogP2	Molecular properties	Squared Ghose–Crippen octanol–water partition coeff. log P^2
4	MAXDN	Topological descriptors	Maximal electrotopological negative variation
5	Mor26u	3D-MoRSE descriptors	3D-MoRSE – signal 26 / unweighted

Artificial neural network

The inputs of the ANN were a subset of the descriptors selection by genetic algorithm from a large set of descriptors. The input layer comprised five descriptors. Usually one hidden layer is enough. After several trials, a hidden layer with three neurons was selected; Fig. 1 explains this choice. The calculated sol-

ubility values (log *S*) constitute the output layer (*e.g.*, the NN architecture is 5-3-1). A sigmoid transfer function $1/(1+e^x)$ was used. After 212 epochs, a correlation coefficient of 0.86 (n = 58) between calculated and observed log (S / mg L^{-1}) was obtained.

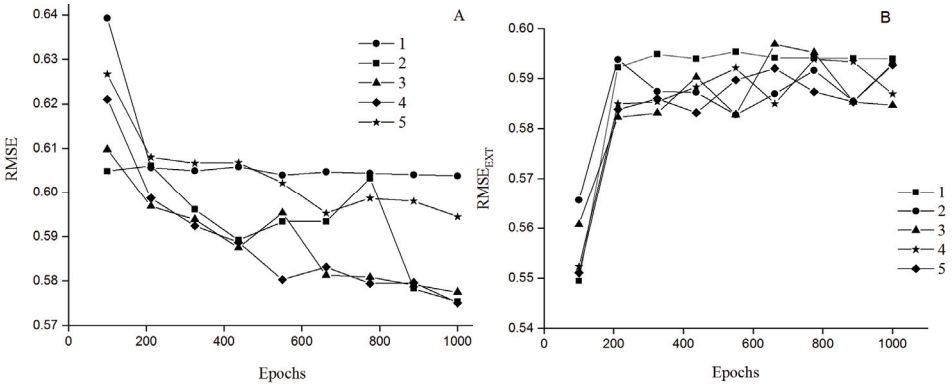


Fig. 1. A – RMSE and B – $RMSE_{EXT}$ values *vs.* number of neurons of the hidden layer and the number of epochs.

The general statistical parameters selected to evaluate the prediction ability of the constructed model are: root mean squared error for the training set (*RMSE*), root mean squared error for the external data set ($RMSE_{EXT}$), cross validated squared CC (Q^2) and the Pearson correlation squared (R^2).

Thus, a 5-3-1 architecture was obtained, the statistic parameters of the final optimal model are given in Table II.

TABLE II. Results and statistical parameters of GA-ANN and GA-SVM

Parameter	Set	GA-ANN	GA-SVM
Number of neurons	–	3	–
Epochs	–	212	–
C	–	–	11.12
E	–	–	0.222
Σ	–	–	0.1111
Q^2_{LOO}	–	0.7748	0.7125
R^2	Training set	0.8097	0.8403
	Validation set	0.7412	0.7068
RMSE	Training set	0.5968	0.5933
	Validation set	0.5823	0.5782

The plot shown in Fig. 2 indicates that there was a significative correlation between calculated and observed log *S* values (Table S-II of the Supplementary material), which shows a weak dispersion of the points around the first bisectrix, which confirm the acceptable performance especially for rough estimations.

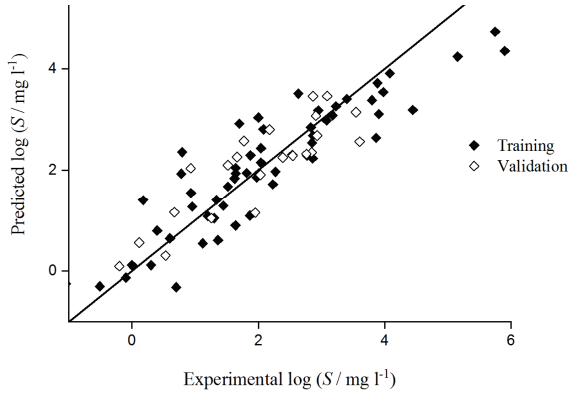


Fig. 2. Predicted values *vs.* experimental values for the training and validation sets.

Support vector machine

The quality of SVM for the regression depends on several parameters, namely, kernel type k, which determines the sample distribution in the mapping space, and its corresponding parameter σ, capacity parameter C, and ε-insensitive loss function. Optimization of the SVM parameters was performed by systematically changing their values in the training set and calculating the *RMSE* of the model using 5-fold cross validation.

First, with the value of C fixed and the epsilon value and σ value varied, a minimal *RMSE* corresponds to optimal values of ε and σ. Once ε and σ are optimized, the regularization parameter C that controls the trade-off between maximizing the margin and minimizing the training error is optimized. To find an optimal value of C, the *RMSE* of SVM models with different C values was calculated. The variation of *RMSE vs*. C values was plotted in Fig. 3. As shown in this figure, the optimal value of C was 11.12. The optimal values of the three parameters and the final optimal model were determined and are given in Table II.

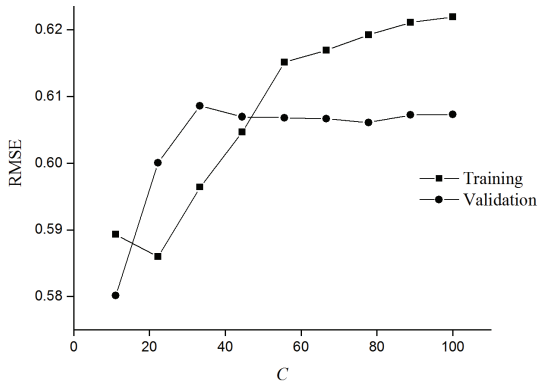


Fig. 3. Variation of *RMSE vs. C* values.

The predicted log *S* values are plotted against the experimental values (Table S-II of the Supplementary materials) in Fig. 4. The predicted values are, in general, in good agreement with the corresponding experimental values.

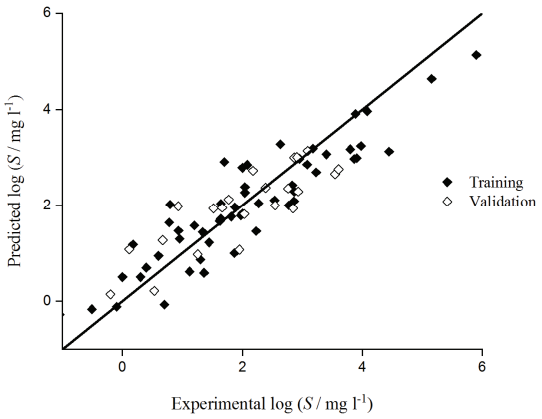


Fig. 4. Predicted values *vs*. experimental values for the training, validation sets.

The statistical parameters show that the models established a strong correlation between the five selected variables and the studied property, characterized by excellent parameters $Q^2_{LOO} > 0.5$.[28] in addition to a good standard error ($RMSE_{ANN}$ = 0.5968, $RMSE_{SVM}$ = 0.5933). All statistical parameters of the model are satisfactory and prove that the models are stable, robust and predictive. Then, the built models were used to predict the test set data.

Analysis of descriptors contribution in the ANN model and interpretation

To evaluate the influence of each descriptor on the calculated solubility, the relevance score was used. The relevance score is calculated by following all paths from the input neuron to the output neuron (including hidden layers). For each path, the product of all the connection weights (in absolute values) were added to the score. Afterwards, all relevance scores are normalized to be in the range between 0 and 100.[5]

The relative value of contributions of the five descriptors of the model was determined (Table III). These values of contributions allow the following classification: RNCG > AlogP2 > MAXDN > MATS8m > Mor26u. It should be noted that the difference in the descriptor contribution between any two descriptors used in the model is not significant, indicating that all of the descriptors are indispensable in generating the predictive model.

These values confirm the great effect of the RNCG and AlogP2 on the solubility. Another advantage of this method is the determination of the impact of each descriptor on the aqueous solubility.

The relative negative charge (*RNCG*), the first important descriptor, is the charge of the most negative atom divided by the total negative charge (Q_{neg}). The

charge density of the ions plays an important role in the interactions of these ions and the water molecule, which reflects the influence of the negative charge on the aqueous solubility. The second important descriptor, the squared Ghose–Crippen–Viswanadhan octanol–water partition coefficient (AlogP2) is calculated from a regression equation based on the hydrophobic character of the molecule. It reflects both the interactions of the solute with the bulk of the surrounding solvent (macroscopic or non-specific solvent effects) and the specific bonding between the solute and individual solvent molecules (microscopic or specific solvent effects).[29]

TABLE III. Relevance score

Index	Name	Relevance score
0	RNCG	100
1	ALOGP2	76
2	MAXDN	39
3	MATS8m	32
4	Mor26u	28

MAXDN, a topological descriptor, is the maximal electrotopological negative variation (MAXDN) is calculated as the maximum negative value of DI_i (topologic distance) in the molecule.

Mats8m is a 2D autocorrelation indices that is calculated by applying the Moran coefficient to the molecular graph.[7]

Mor26u is a descriptor of the 3D-MoRSE descriptor class, the 3D-MoRSE descriptors are 3D molecular representations of structure based on electron diffraction descriptor,[30,31] that are calculated by summing atomic weights viewed by a different angular scattering function. The values of these descriptor functions are calculated at 32 evenly distributed values of scattering angle (s) in the range of 0–31 Å from the three-dimensional atomic coordinates of a molecule. The 3D-MoRSE descriptor is calculated using following expression:

$$\text{Morsw} = \sum_{i=1}^{n\text{AT}-1} \sum_{j=i+1}^{n\text{AT}} w_i w_j (\sin(sr_{ij}) / sr_{ij}) \quad (16)$$

where s is the scattering angle, nAT is the number of atoms, r_{ij} is the interatomic distance between the i^{th} and the j^{th} atoms, w is an atomic property, including atomic number, masses, van der Waals volumes, Sanderson electronegativities and polarizabilities.

The statistical parameters obtained for the test set,[32] demonstrate the power of the predictivity of the models.

Comparison of the results with other modeling methods

The present results were compared with those obtained in previous publications using other modeling methods. The comparisons are summarized in Table

IV, which shows that the presented SVM model gives better predictions than most of the other methods, because fewer descriptors were used than in the the other models, and an almost similar result was obtained. In addition, the herein presented models were evaluated using different statistical parameters compared with the other model in the literature.[33]

TABLE IV. Comparison of the presented results with those obtained using other modeling methods

Reference	Method	Test set	Training set	Number of descriptors	R2tr	R2 test	*RMSE*
Our results	ANN	22	58	5	80.97	74.12	0.5968
	SVM	22	58	5	84.03	70.68	0.5933
Deeb and Goodarzi[3]	PLS	–	219	22	79.98	79.44	–
	PC-ANN	–	219	22	84.35	81.93	–
Bouakkadia *et al*.[29]	MLR	19	58	6	88.95	85.11	0.5200

The difference between this work and the previously published work on this data set is that the number of compounds in the validation set was not the same and also the two training and validation sets did not contain the same compounds, because the method of separation was not the same. Additionally, the descriptors selected by the genetic algorithm are different except for AlogP2.

CONCLUSIONS

A quantitative structure–property relationship analysis was performed on the logarithm of the solubility in water for 80 pesticide compounds using ANN and SVM. The built models clearly demonstrate good correlations between the structure and aqueous solubility of the studied compounds. Five descriptors were selected with genetic algorithm. The selected descriptors, *i.e*., MATS8m, RNCG, AlogP2, MAXDN and Mor26u, were found to be important factors controlling the aqueous solubility. Comparison between the ANN and SVM methods demonstrates that the performance of SVM model is better than that of ANN, but the ANN model is more general than SVM because of the great value of R^2_{test}.The proposed models will help identifying new pesticides and provide insight to guide their development and may be useful for predicting their solubility.

SUPPLEMENTARY MATERIAL

Additional data are available electronically at the pages of journal website: https://www.shd-pub.org.rs/index.php/JSCS/index, or from the corresponding author on request.

ИЗВОД

КОРИШЋЕЊЕ GA-ANN И GA-SVM ЗА QSPR СТУДИЈУ РАСТВОРЉИВОСТИ ПЕСТИЦИДА У ВОДИ

AMEL BOUAKKADIA[1,2], NOUREDDINE KERTIOU[1,2], RANA AMIRI[1], YOUSSOUF DRIOUCHE[1] и DJELLOUL MESSADI[1]

[1]*Environmental and Food Safety Laboratory, Department of Chemistry, Badji Mokhtar University – Annaba, BP. 12, 23000 Annaba, Algeria и* [2]*Abbes Laghrour University, Faculty of Sciences and Technology – Khenchela, BP 1252 Route de Batna, 40004 Khenchela, Algeria*

Тежња пестицида, у овој студији нарочито хербицида, за расподелу по различитим одељцима животне средине, зависи углавном од физичкохемијских особина самих пестицида. Растворљивост у води (S) указује на тенденцију пестицида да се уклоне испирањем или иригацијом да би завршили у површинским водама. Експериментални поступак за одређивање растворљивости пестицида у води је веома скуп и тежак. QSPR методе се често користе за процену растворљивости хербицида у води. Методе вештачке неуронске мреже (ANN) и векторске машине за подршку (SVM), сваки пут повезане са селекцијом помоћу генетичког алгоритма (GA) за избор најзначајније варијабле, биле су коришћене за развој QSPR модела за предвиђање растворљивости у води серије од 80 хербицида. Вредности $\log S$ проучаваних једињења добро су корелисане са дескрипторима. Разматрајући погодне дескрипторе, квадратни Пирсонов коефицијент (R^2) 0,8 добијен је за ANN модел за структуру 5-3-1, а 0,8 је добијен за SVM модел користећи RBF функцију за оптималне вредности параметара: $C = 11{,}12$; $\sigma = 0{,}1111$ и $\varepsilon = 0{,}222$.

(Примљено 18. јуна, ревидирано 19. августа, прихваћено 6. октобра 2020)

REFERENCES

1. P. Gramatica, A. Di Guardo, *Chemosphere* **47** (2002) 947 (https://doi.org/10.1016/S0045-6535(02)00007-3)
2. O. C. Hansen, *Quantitative Structure–Activity Relationships (QSAR) and Pesticides*, Pesticides Research No. 94, Teknologisk Institute, Taastrup, 2004 (https://www2.mst.dk/udgiv/publications/2004/87-7614-434-8/html/helepubl_eng.htm)
3. M. Zine, A. Bouakkadia, L. Lourici, D. Messadi, *J. Serb. Chem. Soc*. **84** (2019) 1405 (https://doi.org/10.2298/JSC190306059Z)
4. *Hyperchem*TM, Release 6.03 for Windows, Molecular Modeling system, 2000 (http://www.hyper.com)
5. *Molegro Data Modeller (*MDM), v.2.0. Copyright Molegro (2009) (https://www.scientific-computing.com/press-releases/molegro-data-modeller-v20)
6. I. N. Levine, *Quantum Chemistry*. 5thed., Prentice Hall, Hoboken, NJ, 2000, pp. 626–739 (ISBN-10:0136855121)
7. R. Todeschini, V. Consonni, M. Pavan, Dragon, *Software for the Calculation of Molecular Descriptors,* Release 5.3 for windows, Milano, 2006 (http://www.talete.mi.it)
8. R. Todeschini, D. Ballabio, V. Consonni, A. Mauri, M. Pavan, *MOBYDIGS*, *Software for Multilinear Regression Analysis and Variable Subset Selection by Genetic Algorithm*, Release 1.1 for Windows, Milano, 2009 (http://www.talete.mi.it/)
9. R. Amiri, D. Messadi, A. Bouakkadia, *J. Serb. Chem. Soc*. **85** (2020) 467 (https://doi.org/10.2298/JSC190610090A)
10. D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, London, 1989, pp. 89–94 (https://dl.acm.org/citation.cfm?id=534133)
11. A. Beheshti, E. Pourbasheer, M. Nekoei, S. Vahdani, *J. Saudi Chem. Soc.* **28** (2016) 282 (http://dx.doi.org/10.1016/j.jscs.2012.07.019)

12. J. Xu, H. Zhang, L. Wang, G. Liang, L. Wang, X. Shen, W. Xu, *Spectrochim. Acta, A* **76** (2010) 239 (https://doi.org/10.1016/j.saa.2010.03.027)
13. S. Jothilakshmi, V. N. Gudivada, *Handbook of Statistics,* Vol. 35, Ch. 10, Elsevier, Amsterdam, 2016, pp. 301–340 (http://dx.doi.org/10.1016/bs.host.2016.07.005)
14. M. Safamirzaei, H. Modarress, M. Mohsen-Nia, *Fluid Phase Equilib.* **266** (2008) 187 (https://doi.org/10.1016/j.fluid.2008.01.022)
15. S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice-Hall, New Delhi, 2006 (ISBN-13: 978-0023527616)
16. R. Rojas, *Neural Network*, Springer-Verlag, Berlin, 1996, pp. 229–261 (http://page.mi.fu-berlin.de/rojas/neural/)
17. O. Deeb, M. Drabh, *Chem. Biol. Drug. Des*. **76** (2010) 255 (https://dx.doi.org/10.4236/aces.2012.21010)
18. V. N. Vapnik, *Statistical Learning Theory*, John Wiley & Sons, New York,1998, pp. 375–473 (ISBN: 978-0-471-03003-4)
19. V. N. Vapnik, *The Nature of Statistical Learning Theory.* Springer, New York, 2000, pp. 267–287 (ISBN: 978-1-4757-3264-1)
20. C. J. Lu, T. S. Lee, C. C. Chiu, *Decision Support. Syst*. **47** (2009) 115 (https://doi.org/10.1016/j.dss.2009.02.001)
21. N. Cristianini, J. Shawe-Taylor, *An introduction to support vector machines and other kernel- based learning methods*, Publishing House of Electronics Industry, Beijing, 2005, pp. 93–122 (ISBN: 0521 780195).
22. R. Amiri**,** D. Messadi, A. Bouakkadia**,** L. Lourici, *Egypt. J. Chem.* **62** (2019) 1563 (https://doi.org/ 10.21608/ejchem.2019.4976.1446)
23. V. Vapnik, *The Nature of Statistical Learning Theory*, Springer*,* New York, 1995, pp. 70–71 (ISBN: 0-387-94559-8).
24. N. Cristianini, J. Shawe-Taylor, *An Introduction to Support Vector Machines*, Cambridge University Press, Cambridge, 2000, pp. 32–42 (ISBN: 0-521-78019-5).
25. B. Scholkopf, A. Smola, *Learning with Kernels*, MIT Press, Cambridge, MA, 2002, pp. 13–17 (ISBN: 0262194759)
26. A. Golbraikh, A. Tropsha, *J. Mol. Graph. Model*. **20** (2002) 269 (https://doi.org/10.1016/S1093-3263(01)00123-1)
27. N. Kertiou, A. Bouakkadia**,** D. Messadi, *Res. J. Pharm. Biol. Chem. Sci.* **8** (2017) 251 (https://www.rjpbcs.com/pdf/2017_8(6)/[29].pdf)
28. A. Bouakkadia**,**Y. Driouche, N. Kertiou, D. Messadi, *Int. J. Saf. Secur. Eng.* **10** (2020) 389 (https://doi.org/10.18280/IJSSE.100311)
29. A. Bouakkadia, H. Haddag, N. Bouarra, D. Messadi, *Synthèse: Revue Sci. Technol.* **32** (2016) 12 (https://www.ajol.info/index.php/srst/issue/view/13942)
30. J. Gasteiger, J. Sadowski, J. Schuur, P. Selzer, L. Steinhauer, V. Steinhauer, *J. Chem. Inf. Comput. Sci*. **36** (1996) 1030 (https://doi.org/10.1021/CI960343+)
31. J. Schuur, P. Selzer, J. Gasteiger, *J. Chem. Inf. Comput. Sci.* **36** (1996) 334 (https://doi.org/10.1021/CI950164C)
32. W. J. Wang, Z. B. Xu, W. Z. Lu, X. Y. Zhang, *Neurocomputing* **55** (2003) 643 (https://doi.org/10.1016/S0925-2312(02)00632-X)
33. O. Deeb, M. Goodarzi, *Mol. Phys.* **108** (2010) 181 (https://doi.org/10.1080/00268971003604575).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 86 (7–8) S131–S135 (2021) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Use of GA-ANN and GA-SVM for a QSPR study on the aqueous solubility of pesticides

AMEL BOUAKKADIA[1,2]*, NOUREDDINE KERTIOU[1,2], RANA AMIRI[1], YOUSSOUF DRIOUCHE[1] and DJELLOUL MESSADI[1]

[1]*Environmental and Food Safety Laboratory, Department of Chemistry, Badji Mokhtar University – Annaba, BP. 12, 23000 Annaba, Algeria and* [2]*Abbes Laghrour University, Faculty of Sciences and Technology – Khenchela, BP 1252 Route de Batna, 40004 Khenchela, Algeria*

J. Serb. Chem. Soc. 86 (7–8) (2021) 673–684

TABLE S-I. Pesticides compounds used in this study and their solubility

No	Name	log (S / mg L^{-1})
1	2-ethylamino-4-(isopropylamino)-6-(methylthio)-s triazine	2.27
2	methyl α-[(4,6-dimethoxypyrimidin-2-ylcarbamoyl) sulfamoyl]-*o*-toluate	2.08
3	S-Ethyldiisobutylcarbamothioate	1.64
4	ethyl 2-(4-chloro-6-methoxypyrimidin-2-ylcarbamoylsulfamoyl) benzoate	3.08
5	3-[4-(4-chlorophenoxy)phenyl]-1,1-dimethyl-urea	0.4
6	1-(2-chlorophenylsulfonyl)-3-(4-methoxy-6-methyl-1,3,5-triazin-2-yl)urea	4.45
7	3,6-dichloropyridine-2-carboxylic acid	5.16
8	(2-hydroxyethyl)ammonium 3,6-dichloropyridine-2-carboxylate	5.75
9	2-{[4-chloro-6-(ethylamino)-1,3,5-triazin-2-yl]amino}-2-methylpropanenitrile	2.23
10	*S*-ethyl *N*-cyclohexyl-*N*-ethylcarbamothioate	1.98
11	2,4-Dichlorophenoxyacetic acid	2.95
12	diethylammonium (2,4-dichlorophenoxy) acetate	5.9
13	methyl 2-(2,4-dichlorophenoxy)acetate	2
14	2,3,6-trichlorobenzoic acid	3.89
15	ethyl 3-phenylcarbamoyloxycarbanilate	0.95
16	(2R)-2-(2,4-dichlorophenoxy)propionic acid	2.77
17	3-[4-(4-methoxyphenoxy)phenyl]-1,1-dimethylurea	1.3
18	6-ethylthio-N^2,N^4-diisopropyl-1,3,5-triazine-2,4-diamine	1.2
19	3-(3,4-dichlorophenyl)-1,1-dimethylurea	1.62
20	*S*-ethyl dipropyl(thiocarbamate)	2.54
21	methyl 2-[(4-ethoxy-6-methylamino-1,3,5-triazin-2-yl)carbamoylsulfa-moyl]benzoate	1.7
22	ethyl (RS)-2-[4-(6-chloro-1,3-benzoxazol-2-yloxy)phenoxy] propionate	-0.1

* Corresponding author. E-mail: amelbouakkadia@yahoo.fr

TABLE S-I. Continued

23	butyl (RS)-2-{4-[5-(trifluoromethyl)-2-pyridyloxy]phenoxy}propionate	0
24	butyl (R)-2-{4-[5-(trifluoromethyl)-2-pyridyloxy]phenoxy}propionate	0.3
25	(*RS*)-1-methylheptyl 4-amino-3,5-dichloro-6-fluoro-2-pyridyloxyacetate	-1.05
26	*N*-(phosphonomethyl) glycine	4.08
27	(RS)-2-{4-[3-chloro-5-(trifluoromethyl)-2-pyridyloxy]phenoxy}propionic acid	1.64
28	3-(4-isopropylphenyl)-1,1-dimethylurea	1.81
29	3-cyclohexyl-1,5,6,7-tetrahydrocyclopentapyrimidine-2,4(3*H*)-dione	0.78
30	3-(3,4-dichlorophenyl)-1-methoxy-1-methylurea	1.88
31	4-(4-chloro-*o*-tolyloxy)butyric acid	1.64
32	(RS)-2-(4-chloro-*o*-tolyloxy)propionic acid	2.87
33	4-amino-4,5-dihydro-3-methyl-6-phenyl-1,2,4-triazin-5-on	3.23
34	2-(3,4-dichlorophenyl)-4-methyl-1,2,4-oxadiazolidine-3,5-dion	0.18
35	3-(3-chloro-4-methoxyphenyl)-1,1-dimethylurea	2.83
36	methyl 2-(4-methoxy-6-methyl-1,3,5-triazin-2-ylcarbamoylsulfamoyl)benzoate	3.98
37	(*RS*)-*N,N*-diethyl-2-(1-naphthyloxy)propionamide	1.87
38	S-propyl butyl(ethyl) thiocarbamate	2
39	4-amino-3,5,6-trichloropyridine-2-carboxylic acid	2.63
40	N^2,N^4-diisopropyl-6-methoxy-1,3,5-triazine-2,4-diamine	2.86
41	N^2,N^4-diisopropyl-6-methylthio-1,3,5-triazine-2,4-diamine	1.52
42	6-chloro-N^2,N^4-diisopropyl-1,3,5-triazine-2,4-diamine	0.93
43	*S*-benzyl dipropyl (thiocarbamate)	1.12
44	ethyl (2RS)-2-[4-(6-chloroquinoxalin-2-yloxy)phenoxy] propionate	-0.51
45	1-(4,6-dimethoxypyrimidin-2-yl)-3-(3-ethylsulfonyl-2-pyridylsulfonyl)urea	3.86
46	6-chloro-N^2,N^4-diethyl-1,3,5-triazine-2,4-diamine	0.79
47	1-(5-*tert*-butyl-1,3,4-thiadiazol-2-yl)-1,3-dimethylurea	3.4
48	3-*tert*-butyl-5-chloro-6-methyluracil	2.85
49	N^2-*tert*-butyl-N^4-ethyl-6-methylthio-1,3,5-triazine-2,4-diamine	1.34
50	methyl 3-(4-methoxy-6-methyl-1,3,5-triazin-2-ylcarbamoylsulfamoyl)thiophene-2-carboxylate	3.8
51	*S*-4-chlorobenzyl diethyl(thiocarbamate)	1.45
52	*S*-2,3,3-trichloroallyl diisopropyl(thiocarbamate)	0.6
53	methyl 2-[4-methoxy-6-methyl-1,3,5-triazin-2-yl(methyl)carbamoylsulfamoyl] benzoate	3.18
54	3,5,6-trichloro-2-pyridyloxyacetic acid	3.91
55	2-butoxyethyl 3,5,6-trichloro-2-pyridyloxyacetate	1.36
56	methyl 2-[4-dimethylamino-6-(2,2,2-trifluoroethoxy)-1,3,5-triazin-2-ylcarbamoylsulfamoyl] totoutoluate	2.04
57	1,1-dimethyl-3-(α,α,α-trifluoro-m-tolyl)urea	2.04
58	*iso*-octyl 4-chloro-*o*-tolyloxyacetate	0.7
59*	methyl 2-[4,6-bis(difluoromethoxy)pyrimidin-2-ylcarbamoylsulfamoyl] benzoate	2.39
60*	1-(2-methylcyclohexyl)-3-phenylurea	1.26
61*	6-chloro-N-ethyl-*N*'-isopropyl-1,3,5-triazin-2,4-diamine	1.52

TABLE S-I. Continued

62*	5-bromo-3-(butan-2-yl)-6-methylpyrimidine-2,4(1*H*,3*H*)-dione	2.85
63*	(R)-1-(Ethylcarbamoyl)ethylcarbanilate	3.54
64*	isopropyl (3-chlorophenyl)carbamate	1.95
65*	4-(2,4-dichlorophenoxy)butanoicacid	1.66
66*	(2,4,5-trichlorophenoxy)acetic acid	2.18
67*	N^2-isopropyl-N^4-methyl-6-methylthio-1,3,5-triazine-2,4-diamine	2.76
68*	(2RS)-2-(2,4-dichlorophenoxy)propionic acid	2.54
69*	Haloxyfopethoxyethyl	0.11
70*	4-chloro-*o*-tolyloxyacetic acid	2.87
71*	(R)-2-(4-chloro-o-tolyloxy)propionic acid	2.93
72*	1-(1,3-benzothiazol-2-yl)-1,3-dimethylurea	1.77
73*	4-amino-6-*tert*-butyl-4,5-dihydro-3-methylthio-1,2,4-triazin-5-one	3.09
74*	methyl 3-(3-methylcarbaniloyloxy)carbanilate	0.67
75*	1-(4-methoxy-6-methyl-1,3,5-triazin-2-yl)-3-[2-(3,3,3-trifluoropropyl)phenylsulfonyl] urea	3.6
76*	N^2-*tert*-butyl-6-chloro-N^4-ethyl-1,3,5-triazine-2,4-diamine	0.93
77*	1-[2-(2-chloroethoxy)phenylsulfonyl]-3-(4-methoxy-6-methyl-1,3,5-triazin-2-yl) urea	2.91
78*	S-propyl dipropyl(thiocarbamate)	2.03
79*	(RS)-3-(3,5-dichlorophenyl)-5-methyl-5-vinyloxazolidine-2,4-dione	0.53
80*	2-isopropylideneaminooxyethyl (R)-2-[4-(6-chloroquinoxalin-2-yloxy)phenoxy] propionate	-0.2

*Validation set

TABLE S-II. Experimental, predicted values of log *S*, and descriptors values for the training and validation sets

N°	log (S_{exp} / mg l^{-1})	log (S_{pred} / mg l^{-1})		MATS8m	RNCG	ALOGP2	MAXDN	Mor26u
		ANN	SVM					
1	2.27	1.97007	2.04153	0.166	0.141	7.524	1.042	0.495
2	2.07918	2.80591	2.84972	–0.1	0.136	3.054	5.054	–0.081
3	1.64345	1.93482	2.03343	0	0.161	12.345	1.435	0.635
4	3.07918	2.98359	2.8582	–0.004	0.143	5.956	5.299	0.103
5	0.39794	0.800011	0.700217	–0.277	0.153	11.403	1.837	–0.136
6	4.4456	3.18637	3.12269	–0.042	0.162	3.456	5.049	–0.053
7	5.15534	4.23823	4.63397	0	0.333	5.064	2.85	–0.088
8	5.74819	4.72705	6.22588	0	0.422	1.409	1.403	0.05
9	2.23	1.72042	1.45922	0.123	0.134	5.456	2.006	–0.003
10	1.97772	1.85824	1.78546	0.13	0.167	11.136	1.387	0.349
11	2.94939	3.18143	2.97726	-0.15	0.262	7.907	2.717	-0.17
12	5.90091	4.34909	5.12943	0	0.349	0.232	0.406	0.139
13	2	3.03824	2.78796	0.049	0.251	9.381	2.129	–0.011
14	3.88649	3.71569	3.90686	0	0.33	11.748	2.854	–0.213
15	0.95424	1.27221	1.30116	0.685	0.115	10.775	2.278	–0.057
16	2.77085	2.2878	1.99897	–0.064	0.208	10.172	2.712	–0.126
17	1.30103	1.0429	0.869131	–0.452	0.14	7.268	1.834	–0.041

TABLE S-II. Continued

18	1.2	1.08567	1.57748	0.144	0.109	12.036	1.033	0.615
19	1.62325	1.83547	1.67145	–1.18	0.222	6.155	1.872	–0.088
20	2.53656	2.3036	2.1091	0	0.187	8.265	1.429	0.263
21	1.69897	2.91823	2.91128	–0.157	0.133	3.143	5.299	0.02
22	–0.09691	–0.127375	–0.117151	–0.107	0.132	22.584	2.076	–0.05
23	0	0.116869	0.503699	–0.135	0.099	28.423	5.696	0.176
24	0.30103	0.116869	0.503699	–0.135	0.099	28.423	5.696	0.176
25	–1.04576	–0.252303	–0.273216	–0.017	0.114	28.798	2.679	0.285
26	4.07918	3.90895	3.96613	0	0.242	4.273	4.902	–0.112
27	1.63749	0.900297	1.70448	–0.002	0.118	19.501	5.797	–0.174
28	1.81291	1.94637	1.76373	–0.595	0.168	5.506	1.771	0.243
29	0.77815	1.93203	1.64045	0.269	0.133	4.686	1.859	0.05
30	1.87506	2.29495	1.96685	–0.564	0.2	5.414	2.059	0.029
31	1.64345	2.04425	1.72174	–0.246	0.184	8.938	2.466	0.036
32	2.8657	2.68557	2.29279	0.842	0.183	9.067	2.653	–0.131
33	3.23	3.26032	2.69202	0.429	0.188	0.284	2.006	–0.032
34	0.17609	1.40374	1.18311	–0.665	0.188	8.466	2.406	–0.249
35	2.83123	2.84747	2.42529	–0.079	0.198	3.241	1.865	0.067
36	3.97772	3.53736	3.24333	–0.069	0.142	1.103	5.273	0.269
37	1.86923	1.0971	1.00116	0.023	0.137	10.733	1.805	0.059
38	2	1.93799	1.82598	0.141	0.173	11.096	1.419	0.319
39	2.63347	3.51021	3.27951	0	0.256	4.701	2.977	–0.203
40	2.86	2.23532	2.0885	0.516	0.13	6.565	1.36	0.449
41	1.52	1.67701	1.93075	0.154	0.131	9.738	1.044	0.593
42	0.934	1.54919	1.47091	0.151	0.134	8.487	1.476	0.25
43	1.12057	0.544881	0.618408	0.104	0.152	16.889	1.461	-0.001
44	-0.50864	-0.298238	-0.16316	0.023	0.12	21.98	2.07	-0.285
45	3.86332	2.64188	2.9698	-0.125	0.103	1.32	5.522	-0.083
46	0.792	2.3567	2.01921	0.178	0.164	4.658	1.47	0.153
47	3.39794	3.40434	3.06829	0.187	0.159	2.054	1.855	0.748
48	2.85126	2.54508	2.07957	0	0.151	1.016	2.111	0.074
49	1.34	1.40891	1.43875	0.154	0.126	8.691	1.305	0.301
50	3.79727	3.38015	3.17407	-0.06	0.136	1.007	5.219	0.183
51	1.44716	1.28912	1.2225	0.13	0.182	13.888	1.528	-0.038
52	0.60206	0.645302	0.943595	-1.126	0.186	15.682	1.683	0.242
53	3.17609	3.08267	3.19066	-0.075	0.143	1.578	5.285	-0.122
54	3.90849	3.10932	2.99134	-0.415	0.274	9.449	2.797	-0.114
55	1.36173	0.609083	0.590257	-0.068	0.159	20.456	2.21	0.174
56	2.04139	2.43526	2.38896	-1.27	0.167	4.387	5.654	-0.093
57	0.69897	-0.319298	-0.0707111	0.337	0.129	34.033	1.992	0.418
58	2.04139	2.14982	2.2675	-0.34	0.122	10.701	5.922	0.123
59*	1.52	2.10053	1.9355	0.163	0.143	6.43	1.473	0.362
60*	2.8451	2.35458	1.93893	0	0.146	1.99	2.009	0.096
61*	3.54407	3.14343	2.65957	0.368	0.135	0.899	2.281	0.475
62*	1.94939	1.15193	1.07316	-0.042	0.148	10.255	2.59	-0.233

TABLE S-II. Continued

63*	1.66276	2.2624	1.95081	-0.122	0.205	10.036	2.499	-0.034
64*	2.17609	2.79982	2.72653	-0.383	0.275	12.085	2.757	-0.172
65*	2.76	2.32061	2.35449	0	0.152	5.732	1.076	0.557
66*	2.54407	2.2878	1.99897	-0.064	0.208	10.172	2.712	-0.126
67*	0.11394	0.558332	1.07797	0.028	0.101	23.862	5.797	0.082
68*	2.8657	3.45844	2.9976	0.681	0.228	6.936	2.658	-0.058
69*	2.92942	2.68557	2.29279	0.842	0.183	9.067	2.653	-0.131
70*	1.77085	2.57556	2.12201	0.163	0.171	3.16	1.825	0.029
71*	3.09	3.46152	3.13894	0	0.177	2.15	1.947	0.688
72*	0.6721	1.15908	1.27212	0.767	0.116	11.696	2.277	-0.122
73*	3.60206	2.56352	2.75728	-0.086	0.137	8.116	5.697	-0.039
74*	0.929	2.03968	1.98117	0.151	0.136	7.511	1.479	0.494
75*	2.91116	3.07567	3.00836	-0.063	0.148	3.395	5.116	-0.003
76*	2.03342	1.90878	1.81662	0.141	0.173	11.55	1.412	0.338
77*	0.53148	0.304822	0.214497	-0.789	0.14	13.412	2.605	-0.164
78*	-0.20066	0.0957958	0.142756	-0.006	0.102	17.699	2.16	0.077
79*	2.38561	2.2525	2.3682	0.047	0.129	11.456	5.561	0.053
80*	1.25527	1.04243	0.975408	0.186	0.141	9.759	1.755	-0.189

*Validation set

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. *86* (7–8) 685–698 (2021)
JSCS–5454
Original scientific paper

New structure-based models for the prediction of normal boiling point temperature of ternary azeotropes

ZOHREH FARAMARZI[1], FATEMEH ABBASITABAR[2]*, HOSSEIN JALALI JAHROMI[1] and MAZIAR NOEI[1]

[1]*Department of Chemistry, Mahshahr Branch, Islamic Azad University, Mahshahr, Iran and* [2]*Department of Chemistry, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran*

(Received 18 February, revised 6 April, accepted 13 April 2021)

Abstract: Recently, development of the QSPR models for mixtures has received much attention. The QSPR modelling of mixtures requires the use of the appropriate mixture descriptors. In this study, 12 mathematical equations were considered to compute mixture descriptors from the individual components for the prediction of normal boiling points of 78 ternary azeotropic mixtures. Multiple linear regression (MLR) was employed to build all QSPR models. Memorized_ACO algorithm was employed for subset variable selection. An ensemble model was also constructed using averaging strategy to improve the predictability of the final QSAR model. The models have been validated by a test set comprised of 24 ternary azeotropes and by different statistical tests. The resulted ensemble QSPR model had $R^2_{training}$, R^2_{test} and q^2 of 0.97, 0.95, and 0.96, respectively. The mean absolute error (*MAE*), as a good indicator of model performance, were found to be 3.06 and 3.52 for training and testing sets, respectively.

Keywords: QSPR modelling; azeotropic mixture; memorized ant colony optimization; mixture descriptors; ensemble multiple linear regression analysis.

INTRODUCTION

Azeotropes are among mixtures which exibit nonideal solution behaviour. They both present challenges and opportunities related to separation processes. Azeotropes cannot be separated easily, however, some certain separations would be carried out by means of the formation of an azeotrope. The ordinary distillation cannot be used directly to separate azeotropes due to the absence of an interphase compositional differential. Several special distillation methods including azeotropic distillation,[1–3] extractive distillation,[4,5] and pressure swing distillation[6] have been used for the azeotrope separation. The understanding of the azeotropic properties including the boiling point is important for the preliminary

* Corresponding author. E-mail: fabbasitabar@gmail.com
https://doi.org/10.2298/JSC210218035F

evaluations aimed at designing of the distillation process. With this background, the use of computer programs for accurate estimation of boiling points of azeotropes becomes highly effective.

There are a variety of tools available for the prediction of mixture behaviour, including group contribution methods (*e.g.*, ASOG[7], UNIFAC and diverse modified UNIFAC versions[8,9]), methods derived from the regular solution theory,[10] linear solvation relationships (LSER),[11] statistical thermodynamics[12] and solvation thermodynamics.[13] These methods are time–consuming and expensive especially when dealing about ternary mixtures. The intermolecular interactions in these mixtures are very complex and cannot be precisely represented by quantum chemistry or molecular dynamics because the demand for CPU is still significatly high (taking many days or even weeks of the computation). Therefore, a better alternative seems to be the use of quantitative structure–property relationships (QSPR).

QSPR models using computational molecular descriptors have been developed to predict various molecular properties. Few QSPR studies have been carried out to estimate properties of mixtures.[14,15] Computing of mixture descriptors is the most important problem facing in the QSPR study of mixtures. One solution to this problem is the combination of the molecular descriptors calculated for the individual constitutes via a specific mathematical formula,[16-18] Previously, we developed QSPR models of the prediction of normal boiling points of binary azeotropes.[19] 22 different mixing rules were applied to compute mixture descriptors. It was shown that twelve mixing rules resulted in satisfied QSPR models. The aim of this study is to develop QSPR models for the prediction of the boiling points of ternary azeotropes. In order to compute the mixture descriptors, 12 different mathematical formulas are employed. To the best of our knowledge, this is the first report of the QSPR studied for ternary azeotropes in which mixture descriptors are calculated from those of individual constitutes.

MATERIALS AND METHODS

Normal boiling point data used in this study were taken from literature.[20] Experimental boiling points along with the molecular compositions of all azeotrope mixtures are given in Tables S-I and S-II of the Supplementary material to this paper. Azeotropic compositions of this data have been characterized by molar fractions. Molecular structures of compounds constituting each mixtures were drawn in 2D ChemDraw.[21] After creating the corresponding 3D structures in CS Chem3D Ultra, the geometry optimization was performed in the molecular operation software (MOE) using the AM1 procedure.[22] For each compound, to characterize the molecular structure, a total of 3475 descriptors including 251 MOE descriptors and 3224 Dragon descriptors were calculated[23] and then classified into 22 different descriptor blocks. These descriptors were firstly investigated to find and remove those having missing values or containing small variation. Next, the correlations between descriptors were calculated and from a set of highly correlated descriptors the one with the best correlation with the boiling

points of ternary azeotropes was maintained and the others were eliminated. After these pretreatments, 328 molecular descriptors were obtained for each individual constituent.

Mixture descriptors

Commercial softwares are able to calculate the molecular descriptors for individual compounds. To perform QSPR studies for the property of mixtures, the mixture descriptors can be computed from the molecular descriptors of individuals using different approaches.[18,24] In this study, twenty–two different types of mixture descriptors obtained from the molecular descriptors of individuals were used.[19] This 22 formulas contained both non-linear and linear mathematical formulas (Table I). From these 22 mathematical formulas, only twelve ones could be employed for ternary mixtures and those that involved minus in their mathematical formulas were ignored. These twelve mixture descriptors were classified into three distinct blocks: 1) those that not require any experimental data; 2) those in which mole fraction data is needed; 3) those which require information about potential energy and/or mole fraction. Information about the potential energies were extracted from MOE software based on the following illustration: consider a ternary azeotrope contains three individuals, namely, A, B and C. The aim is to compare potential energy of a system contains ternary azeotrope, *e.g.*, ABC to three systems each of which contains only one of individuals. In order to make better comparison, the potential energy of ABC system is separately compared to AAA, BBB and CCC. If the potential energy of ABC is close, for example, to that of AAA, it is then assumed that A has the highest contribution in the stability of the ternary mixture and therefore A should be weighted

TABLE I. Types of mixture descriptors, description and their formulas; the parameters c, x and D refer to potential coefficient, mole fraction and molecular descriptor, respectively

Descriptor symbol	Formula	Description
Centroid	$(D_A+D_B+D_C)/3$	Mean of descriptors of pure compounds
fmol–sum	$x_AD_A+x_BD_B+x_CD_C$	Sum of weighted descriptors by mol fractions
sqr–fmol	$x_A^2D_A+x_B^2D_B+x_C^2D_C$	Sum of weighted descriptors by square of mol fractions
root–fmol	$\sqrt{x_A}D_A+\sqrt{x_B}D_B+\sqrt{x_C}D_C$	Sum of weighted descriptors by root of mol fractions
sqr–fmol–sum	$(x_AD_A+x_BD_B+x_CD_C)^2$	Square sum of weighted descriptors by mol fractions
norm–cont	$\sqrt{(x_AD_A)^2+(x_BD_B)^2+(x_CD_C)^2}$	Norm of mole fraction contribution descriptors
poten–sum	$c_AD_A+c_BD_B+c_CD_C$	Sum of weighted descriptors by potential coefficients
sqr–poten	$c_A^2D_A+c_B^2D_B+c_C^2D_C$	Sum of weighted descriptors by square of potential coefficients
root–poten	$\sqrt{c_A}D_A+\sqrt{c_B}D_B+\sqrt{c_C}D_C$	Sum of weighted descriptors by root of potential coefficients
sqr–poten–sum	$(c_AD_A+c_BD_B+c_CD_C)^2$	Square sum of weighted descriptors by potential coefficients
poten–norm–cont	$\sqrt{(c_AD_A)^2+(c_BD_B)^2+(c_CD_C)^2}$	Norm of potential coefficients contribution descriptors
fmol–poten–sum	$x_Ac_AD_A+x_Bc_BD_B+x_Cc_CD_C$	Sum of weighted descriptors by both mol fractions and potential coefficients

more than the other two individuals in the mixture descriptor calculation process. The potential energy coefficients, which show the contribution of each individual component in the stability of a ternary mixture, were obtained from Eqs. (1)–(3):

$$c_A = \frac{(E_{ABC} - E_{AAA})^2}{(E_{ABC} - E_{AAA})^2 + (E_{ABC} - E_{BBB})^2 + (E_{ABC} - E_{CCC})^2} \quad (1)$$

$$c_B = \frac{(E_{ABC} - E_{BBB})^2}{(E_{ABC} - E_{AAA})^2 + (E_{ABC} - E_{BBB})^2 + (E_{ABC} - E_{CCC})^2} \quad (2)$$

$$c_C = \frac{(E_{ABC} - E_{CCC})^2}{(E_{ABC} - E_{AAA})^2 + (E_{ABC} - E_{BBB})^2 + (E_{ABC} - E_{CCC})^2} \quad (3)$$

where E_{ABC}, E_{AAA}, E_{BBB}, and E_{CCC} are the total potential energies, kcal* mol^{-1}, for the systems including ABC, AAA, BBB and CCC, respectively. MNDO Hamiltonian gives rough estimates of these potential energies. A brief description along with the mathematical formulas for all types of mixture descriptors is shown in Table I.

Model construction and descriptor selection method

In this study, QSPR models for the prediction of the boiling points of ternary azeotropes were established by the multiple linear regression (MLR). MLR has several advantages including its simple form, fast construction, and easily interpretation. However, the need for an effective tool to select an appropriate subset of descriptors is the main disadvantage of MLR. In this study, the memorized ant colony optimization algorithm as a powerful descriptor selection tool (memorized_ACO) was used.[25] To read more about this algorithm, readers are referred to our previous papers.[19,25-27] The reliability of the final models, constructed based on the descriptors selected by memorized_ACO, algorithm would be investigated by using internal and external validations. Also, the statistical parameters of $R^2_{training}$, R^2_{test}, and q^2 are used to compare the models. $R^2_{training}$ and R^2_{test} indicate the squared correlation coefficients of training and test sets, respectively. q^2 is the squared correlation coefficient obtained *via* leave-one-out cross-validation (LOO-CV) procedure. Moreover, 5-fold cross-validation (5-CV),[27] Monte Carlo cross-validation (MCCV)[28] and double-cross validation (D-CV)[29] are considered for further validation of the resulted QSPR models.

RESULTS AND DISCUSSION

In the present work, twelve types of mixture descriptors were used to characterize the ternary azeotropes. These mixture descriptors are different in their mathematical formulas and all are calculated from the individuals' descriptors. The applicability of the proposed mixture descriptors in developing QSPR models for the prediction of boiling points of ternary azeotropes were investigated. Table I represents the required information about these mixture descriptors.

The comparison between different types of mixture descriptors would be easily performed, provided that a linear regression is applied. Consequently, MLR combined to the memorized_ACO algorithm was used to establish the

* 1 kcal = 4184 J

QSPR models. Fig. 1 shows the calculated q^2 values for MLR models developed for all types of mixture descriptors.

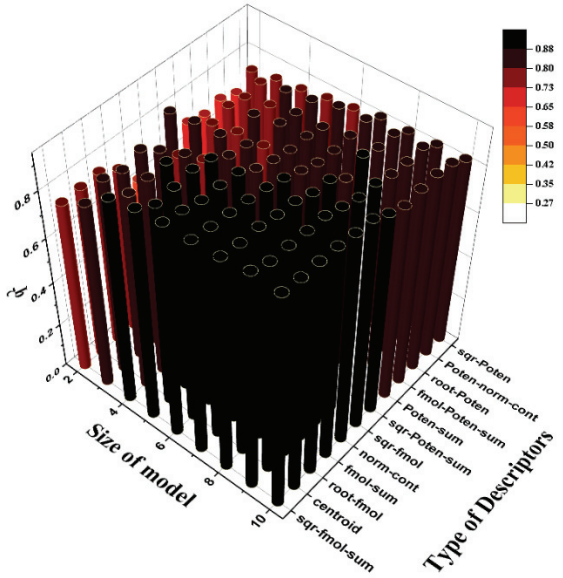


Fig. 1. q^2 values for QSPR models with different sizes constructed on the basis of different types of mixture descriptors.

It is worthy to note that MLR models with different sizes are generated for each type of mixture descriptor. That is, the momorized_ACO algorithm is run several times for each descriptor type while each time a specific number of descriptors is selected. MLR models of sizes 2–10 are investigated. Graphs related to $R^2_{training}$ and R^2_{test} are given in Figs. S-1 and S-2 of the Supplementary material. The statistical parameters for all types of QSPR models having different sizes are presented in Table S-III of the Supplementary material. More statistics for the each type of QSPR model can be found in Tables S-IV–S-XIII of the Supplementary material.

Analyzing Fig. 1 showed that all types of mixture descriptors proposed in this study are able in accessing predictive QSPR models. The values of $R^2_{training}$ and R^2_{test} of the best MLR models for each type of mixture descriptor were between 0.87 to 0.96 and 0.84 to 0.94, respectively. The value of q^2 was found to be in the range of 0.82–0.94. The training and the test sets had the values of MAE of 4.59 and 5.04 K, respectively, indicating the good prediction power of the best model. Based on the statistical parameters of the models, as it can be seen in Fig. 1, the goodness of the proposed mixture descriptors was as follows: sqr–fmol–sum > centroid > root–fmol > fmol–sum > norm–cont > sqr–fmol > sqr–poten–sum > poten–sum > fmol–poten–sum > root–poten > poten–norm–

–cont > sqr–poten. The best mixture descriptor in characterizing and developing predictive QSPR model for ternary azeotropic mixture was found to be sqr––fmol–sum. It should be emphasized again that all of the proposed mixture descriptors showed more or less similar results regarding the statistical parameters. For deep inspection, statistical parameters for QSPR models constructed on the basis of sqr–fmol–sum and sqr–poten are given in Tables II and III, respectively. The best QSPR model was constructed with six and seven descriptors for sqr––fmol–sum and sqr–poten descriptor types, respectively. Statistics of the resulted models for both types were good and found to be 0.96 and 0.87 for R^2_{training}, 0.94 and 0.84 for R^2_{test}, respectively. Statistical parameters for all other developed QSPR models are given in Tables S-IV–S-XIII.

TABLE II. Statistical parameters of the QSPR models based on sqr–fmol–sum descriptors; *F* stands for Fischer test

No. of descriptors	R^2_{training}	*RMSE*-Training	q^2	RMSE-CV	5-CV		MCCV		R^2_{test}	*RMSE*-Test	*F*
					q^2	*RMSE*	R^2_{training}	R^2_{test}			
2	0.79	9.37	0.76	9.98	0.75	10.15	0.79	0.73	0.74	11.29	94.42
3	0.85	7.83	0.82	8.63	0.82	8.70	0.85	0.82	0.85	6.70	95.46
4	0.92	5.72	0.91	6.24	0.90	6.38	0.92	0.86	0.93	4.16	142.36
5	0.95	4.70	0.93	5.34	0.93	5.49	0.95	0.92	0.92	4.51	169.46
6	**0.96**	**4.27**	**0.94**	**4.90**	**0.94**	**5.01**	**0.96**	**0.92**	**0.94**	**4.90**	**169.62**
7	0.96	3.93	0.94	4.87	0.94	5.07	0.96	0.93	0.94	4.53	169.12
8	0.96	3.84	0.95	4.61	0.94	4.83	0.97	0.93	0.95	3.83	152.15
9	0.98	3.07	0.96	3.81	0.96	4.04	0.98	0.96	0.97	3.64	209.62
10	0.97	3.39	0.96	4.26	0.95	4.48	0.97	0.93	0.97	2.96	150.12

TABLE III. Statistical parameters of the QSPR models based on sqr–poten descriptors

No. of descriptors	R^2_{training}	*RMSE*-Training	q^2	*RMSE*-CV	5-CV		MCCV		R^2_{test}	*RMSE*-Test	*F*
					q^2	*RMSE*	R^2_{training}	R^2_{test}			
2	0.77	9.80	0.74	10.40	0.73	10.65	0.77	0.74	0.75	10.81	84.16
3	0.79	9.31	0.76	10.05	0.75	10.26	0.79	0.73	0.78	10.19	62.67
4	0.82	8.72	0.77	9.82	0.75	10.08	0.81	0.77	0.81	9.69	54.16
5	0.83	8.42	0.78	9.42	0.78	9.65	0.83	0.77	0.81	10.09	46.31
6	0.84	8.09	0.80	9.11	0.79	9.33	0.85	0.79	0.80	10.28	41.58
7	**0.87**	**7.44**	**0.82**	**8.63**	**0.81**	**8.91**	**0.87**	**0.82**	**0.84**	**11.31**	**42.46**
8	0.87	7.31	0.82	8.65	0.81	9.02	0.88	0.80	0.87	9.16	44.19
9	0.88	7.14	0.82	8.51	0.81	8.79	0.88	0.82	0.84	9.08	39.86
10	0.88	7.03	0.83	8.48	0.80	9.36	0.89	0.80	0.87	12.31	31.58

It is worth to further analyze the best QSPR model that has been constructed on the basis of sqr–fmol–sum. For this model that contained six descriptors, MAEs of training and test sets were 3.39 and 3.69 K, respectively. More statistical parameters of this QSPR model are given in Table IV. The model is statistically sound since all *t* values are significant, *t* stands for *t*-student value at

95 % confidence interval. Variance inflation factors (VIFs) calculated for considered descriptors in the model are all less than 10, indicating the lack of collinearities among them.[30] The external validation of the model was further verified using Q^2_{F1}, Q^2_{F2}, Q^2_{F3}, r^2_m and CCC (concordance correlation coefficient).[31,32] For an acceptable model, the Q^2_{F1}, Q^2_{F2}, and Q^2_{F3} values should be greater than 0.6, r^2_m greater than 0.5, and CCC greater than 0.85. The values of Q^2_{F1}, Q^2_{F2}, Q^2_{F3}, $r^2_{m,}$ and CCC for the sqr–fmol–sum based QSPR model were found to be 0.92, 0.90, 0.94, 0.91, and 0.97, respectively. The corresponding values for other types of QSPR models are tabulated in Table S-XIV. To test the risk of chance correlation in this QSPR model, *y*-randomization test was employed. In this test, the vector containing information about the boiling points of ternary azeotropes is scrambled and then a MLR model with the same descriptors used is developed.

TABLE IV. Statistics for the best QSPR model based on sqr–fmol–sum descriptors

Descriptor	Definition	beta	*t*	*p*-value	*VIF*
Constant	–	349.75	561.81	$< 10^{-32}$	–
ZM2V	Second Zagreb index by valence vertex degrees	8.59	8.50	5×10^{-11}	2.58
X0A	Average connectivity index chi–0	–7.22	–5.00	8×10^{-6}	5.29
E2e	2nd component accessibility directional WHIM index / weighted by Sanderson electronegativity	5.25	4.24	1×10^{-4}	3.88
HATS3u	Leverage-weighted autocorrelation of lag 3 / unweighted	3.49	4.72	2×10^{-5}	1.38
R1u+	R maximal autocorrelation of lag 1 / unweighted	15.69	9.54	1×10^{-12}	6.85
vsurf_W1	Hydrophilic volume at −0.2 kcal mol^{-1}	23.16	13.04	3×10^{-17}	7.99

This iterative algorithm usually repeats 100 times. Next, the squared correla¬ti¬on coefficients for the models produced in this test are collected in a vector and compared to that for the best MLR model. The results are shown as bar plot in Fig. 2.

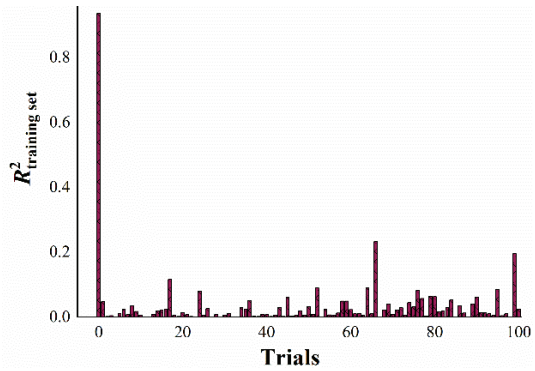


Fig. 2. *Y*-randomization test applied to the sqr–fmol–sum based QSAR model; the first bar shows the $R^2_{training}$ value for the original model.

Note that the first column represents R^2_{training} for the best QSPR model. It is clear that all of the models with scrambled property have poor quality. The same plots were provided for other types of mixture descriptors, given in Fig. S-3. These figures obviously showed that all the mixture descriptors proposed in this study had good ability to characterize mixture structures.

The presence of any outlier was checked through defining applicability domains (AD) of QSPR models. AD is a theoretical working space within which the property of a new compound can be accurately predicted. Several QSPR models with different type of descriptors that are capable to characterized ternary azeotropic mixtures were developed. Each QSPR model has been constructed on the basis of its own descriptor type, they would have different ADs. Different strategies have been introduced to define AD. The standardization approach reported by Roy et al. is the simplest one that was used in this study.[28,33] In this approach, the descriptor matrix of training set is autoscaled. The maximum and the minimum values of the autoscaled descriptors for each mixture are detected. If the minimum value exceeds 3, that compound should be treated as outlier. If the maximum value exceeds 3 and the minimum value is lower than 3, then a new measure (l) is to be computed for that compound by the following equation:

$$l = \bar{d} + 1.28 s_{\text{d}} \tag{4}$$

where $\bar{d}$ and s_{d} are the average and standard deviation of descriptor values for a compound, respectively. If the computed l value is more than 3, then that compound is an outlier.

No outliers within training set were detected for all types of QSPR models except for that constructed on the basis of sqr–poten–sum. Two mixtures (mixture numbers 6 and 39) were highlighted as outliers for the sqr–poten–sum model. Mixtures 6 and 39 contain water/carbon disulfide/ethanol and water/ethanol/hexane, respectively. The QSPR models had more complex behaviours regarding to the detected outs of domain. All samples in the test set were found to be within applicability domain for the root–sum based QSPR model whilst for other cases, one or two samples were detected out of domain. Mixture 7 was found to be out of applicability domain for seven QSPR models. This mixture contains water, carbon disulfide, and acetone. It is interesting that there are two mixtures in the data set in which carbon disulfide participated. They are mixtures 6 and 7. The former belongs to the training set and the later to the test set. Both of them were detected as outliers.

Plots of the actual versus predicted boiling points by all types of QSPR models for training and test sets are depicted in Fig. 3. According to this figure, a good correlation can be found between the experimental boiling points of ternary azeotropes and the predicted ones for all types of QSPR models developed in this study.

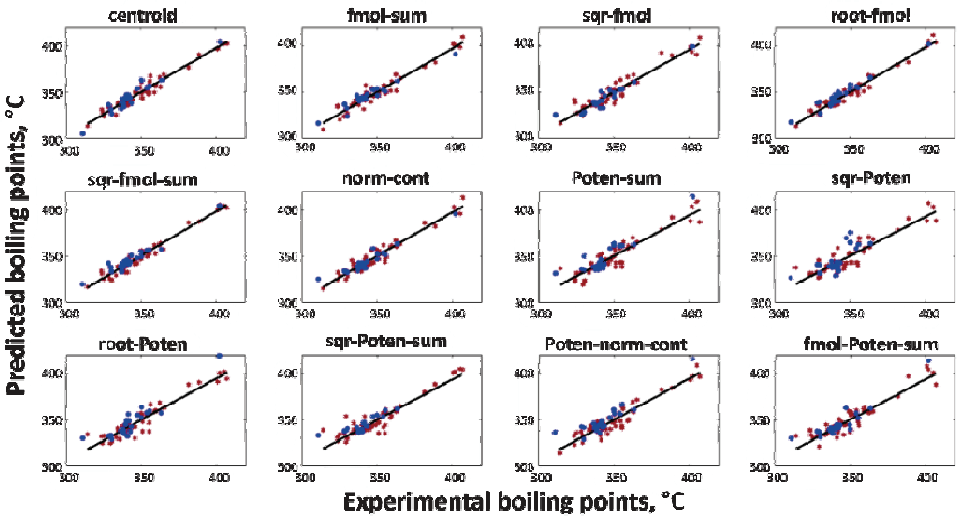


Fig. 3. Plot of observed *vs.* predicted boiling points by different QSPR models.

It was shown previously that prediction by an ensemble model would be more accurate than a single model.[19,34,35] Through a simple averaging strategy, an ensemble model was constructed using all of the QSPR models. R^2_{training}, R^2_{test} and q^2 of the ensemble model were found to be 0.96, 0.94, and 0.95, respectively. MAEs of training and test sets for the ensemble model were found to be 3.06 and 3.52, respectively. In comparison to the sqr–fmol–sum based QSPR model, that has been found to be the best QSPR model, the prediction by the ensemble model was improved.

A fit plot of experimental *versus* predicted boiling points of ternary azeotropes by the ensemble model for the training and test sets is shown in Fig. 4. The two outliers can be identified by visual inspection; one belongs to the training set

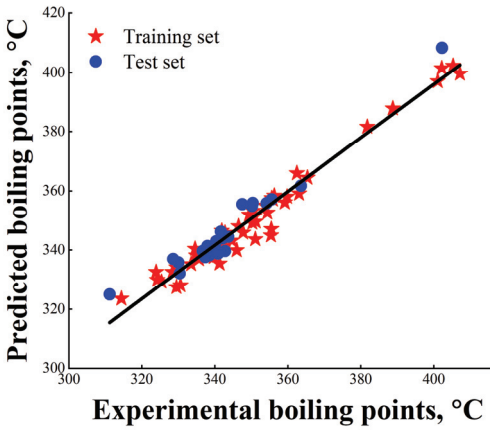


Fig. 4. Plot of observed *vs.* predicted boiling points by the ensemble QSPR model.

and the other to the test set. The constituent molecules of these outliers are water/ /carbon disulfide/acetone and water/carbon disulfide/ethanol. Both of them contain carbon disulfide. Since only these outliers contain carbon disulfide, maybe carbon disulfide can cause these mixtures to be outliers. Carbon disulfide has a linear structure and is a nonpolar compound with low basicity and zero acidity properties. Other statistical parameters of the ensemble model can be found in Table S-XIV.

To understand the predictive power of the resulted QSPR models better, the obtained models were compared with those published previously. Up to now, to the best of our knowledge, there is only one reported QSPR study for the prediction of boiling points of ternary azeotropes.[14] In this paper, the universal solvation equation and artificial neural network to predict boiling points of ternary azeotropes was used. MLR was also used for QSPR model development based on the universal solvation equation. Unfortunately, the MLR model over the whole data set of 78 azeotropes and no test set was considered successful. So, for better comparison the MLR models on the same training set were developed. The resulted statistical parameters along with those reported previously are given in Table V. The $R^2_{training}$ value for training set were reported 0.873 which was increased to 0.909 after the removal of two outliers. Since we used multiple linear regression with few independent variables, this comparison shows the preference of our models to the previously reported, in which non-linear fitting with large number of independent variables have been used.[14,36,37]

TABLE V. Comparison of statistics for the QSPR models with the previously reported

Ref.	Model	No. of descriptors	$R^2_{training}$	q^2	F	*RMSE*--Training	*RMSE*--CV
This study	fmol_sum	5	0.928	0.913	185.76	5.17	5.67
	sqr_fmol	6	0.910	0.890	119.54	5.79	6.39
	root_fmol	8	0.954	0.939	177.18	4.15	4.75
	sqr_fmol_sum	6	0.949	0.939	220.46	4.35	4.77
	norm_cont	6	0.932	0.912	162.37	5.02	5.72
	poten_sum	6	0.854	0.825	69.24	7.36	8.07
	sqr_poten	6	0.810	0.758	50.37	8.41	9.49
	root_poten	7	0.875	0.844	69.78	6.82	7.62
	sqr_poten_sum	5	0.856	0.836	85.85	7.31	7.81
	poten_norm_cont	6	0.865	0.832	75.94	7.08	7.90
	fmol_poten_sum	5	0.868	0.837	94.85	7.00	7.77
Oliferenk *et al.*[14]	Universal solvation equation	10	0.873	–	–	–	–

Details of the best model

The mathematical equation of the best QSPR model developed for each type of mixture descriptor is given in Table S-XIV. The best single QSPR model for

the prediction of boiling points of ternary azeotropes was found to be the one constructed on the basis of the sqr–fmol–sum mixture descriptor. This model consists of six descriptors including ZM2V, X0A, E2e, HATS3u, R1u+ and vsurf_W1>

$$T = 349.75 + 8.59(\pm 1.01)\text{ZMV2} - 7.22(\pm 1.44)\text{XOA} + 5.25(\pm 1.24)\text{E2e} + \\ + 3.49(\pm 0.74)\text{HATS3u} + 15.69(\pm 1.64)\text{R1u} + 23.16(\pm 1.78)\text{vsurf\_W1} \quad (5)$$

The vsurf_W1 has positive influence on the boiling points of ternary azeotropes since it pertinent to polarity, which means that the individual constitutes with more polarity result in a mixture with higher boiling point.[38] X0A which is corresponded to the molecular branching properties is a topological descriptor.[39] It has negative sign in Eq. (5), indicating that the more branching the individuals, the lower the boiling point of the resulted azeotrope.

ZM2V is the second Zagreb index by valence degrees. ZM2V is related to molecular branching. However, the sign of this descriptor in the QSPR model is positive because not only the molecular branching but also the valence degrees affect the amount of this descriptor. For example, ZM2V for 1-heptene and *n*-heptane were calculated as 26 and 20, respectively. The positive sign of ZM2V indicated that individual constitutes with unsaturated bonds cause to increase the boiling point of the resulted azeotrope.

E2e corresponds to 2nd component accessibility directional WHIM index/ /weighted by atomic Sanderson electronegativities. WHIM descriptors are built in such a way as to capture relevant 3D structural knowledge about molecular size, symmetry, shape, and distribution of atoms in relation to reference frames. Since E2e has positive sign in Eq. (5), it is concluded that by increasing the size and symmetry of the molecule, along with the presence of electronegative atoms, the boiling point increases.

HATS3u and R1u+ correspond to leverage–weighted autocorrelation of lag 3/unweighted and to R maximal autocorrelation of lag 1/unweighted, respectively. Both belong to the GETAWAY category. GETAWAYs are geometrical descriptors encode information on structural fragments and, to some extent, account for the information on molecular shape and size and for the specific atomic properties, as well. It was shown that they also encode information about the effective position of fragments and/or substituents in the molecular space.[28,40] Both of them have positive effects on the boiling points of ternary azeotropes.

CONCLUSIONS

Between mixtures, azeotropes exhibit non-ideal behaviour. Azeotropes cannot be separated using conventional distillation due to the lack of a combined phase differential. The estimation of boiling points of azeotropes for initial evaluation is very effective in the design of the distillation process.

In the present study, new and reliable boiling point prediction models for the ternary azeotropes were developed, based on different types of mixture descriptors. The mixture descriptors were calculated according to the molecular descriptors of the individual pure components involved in the mixture and their molar fraction using 12 different mixing rules. The subset of mixture descriptors was selected using memorized_ACO algorithm. Moreover, the multiple linear regression as a simple technique with easy interpretation was also employed to develop models based on selected mixture descriptors. The best QSPR model was found to be the one constructed on the basis of sqr–fmol–sum mixture descriptor. Due to the good predictive power of all models, all of them were considered in the construction of an ensemble model through a simple averaging strategy. The ensemble model prediction improved when compared to the best QSPR model. The validity of the proposed models was examined by the applying of internal and external validations together with the different statistical analysis. The results showed that the proposed models based on the mixed rules have a good ability to predict the boiling points of 78 ternary azeotropes studied. The proposed models for the prediction of the boiling points of ternary azeotropes were obtained using only the information about the molar fraction and the molecular structure of pure components and without any other information, including the boiling points of pure components of azeotropic mixtures.

SUPPLEMENTARY MATERIAL

Additional data are available electronically at the pages of journal website: https://www.shd-pub.org.rs/index.php/JSCS/index, or from the corresponding author on request.

Acknowledgments. This manuscript was extracted from the PhD Thesis of Zohreh Faramarzi. The financial support of this work, by Mahshahr branch–Islamic Azad University, is greatly appreciated.

ИЗВОД

НОВИ МОДЕЛИ ЗА ПРЕДВИЂАЊЕ ТЕМПЕРАТУРЕ НОРМАЛНИХ ТАЧКИ КЉУЧАЊА ТЕРНЕРНИХ АЗЕОТРОПА ЗАСНОВАНИ НА СТРУКТУРИ

ZOHREH FARAMARZI[1], FATEMEH ABBASITABAR[2], HOSSEIN JALALI JAHROMI[1] и MAZIAR NOEI[1]

[1]*Department of Chemistry, Mahshahr Branch, Islamic Azad University, Mahshahr, Iran и* [2]*Department of Chemistry, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran*

У новије време је много пажње привукао развој QSPR модела за смеше. QSPR моделовање смеша захтева коришћење одговарајућих дескриптора смеша. У овој студији је разматрано 12 математичких једначина за израчунавање дескриптора смеша из појединачних компоненти у циљу педвиђања нормалних тачки кључања за 78 тернерних азеотропских смеша. Вишеструка линеарна регресија (MLR) коришћена је за прављење свих QSPR модела. „Memorized_ACO“ алгоритам примењен је за одабир подскупа варијабли. Такође је конструисан ансамбл модел коришћењем стратегије упросечивања да би се побољшала моћ претсказивања коначног QSAR модела. Модели су валидирани тест скупом који садржи 24 тернерна азеотропа и путем различних статистичких тестова. Добијени QSPR ансамбл модел је имао R^2_{training}, R^2_{test}, и q^2 од 0,97, 0,95, односно 0,96.

Средња апсолутна грешка (*MAE*) је добар индикатор ваљаности модела и нађено је да је 3,06 и 3,52 за пробне тестове.

(Примљено 18. марта, ревидирано 6. априла, прихваћено 13. априла 2021)

REFERENCES

1. M. B. Franke, *Comput. Chem. Eng.* **89** (2016) 204 (https://doi.org/10.1016/j.compchemeng.2016.03.027)
2. Q.-K. Le, I. J. Halvorsen, O. Pajalic, S. Skogestad, *Chem. Eng. Res. Des.* **99** (2015) 111 (https://doi.org/10.1016/j.cherd.2015.03.022)
3. W. Li, L. Zhong, Y. He, J. Meng, F. Yao, Y. Guo, et al., *Ind. Eng. Chem. Res.* **54** (2015) 7668 (https://doi.org/10.1021/acs.iecr.5b00572)
4. Y. Wang, S. Liang, G. Bu, W. Liu, Z. Zhang, Z. Zhu, *Ind. Eng. Chem. Res.* **54** (2015) 12908 (https://doi.org/10.1021/acs.iecr.5b03666)
5. B. ZareNezhad, N. Hosseinpour, *Energy Convers. Manage.* **50** (2009) 1491 (https://doi.org/10.1016/j.enconman.2009.02.016)
6. Y. Tavan, S. Shahhosseini, *Energy Technol.* **4** (2016) 424 (https://doi.org/10.1002/ente.201500287)
7. K. Tochigi, D. Tiegs, J. Gmehling, K. Kojima, *J. Chem. Eng. Jpn.* **23** (1990) 453 (https://doi.org/10.1252/jcej.23.453)
8. S. M. Hosseini, M. M. Alavianmehr, D. Mohammad-Aghaie, F. Fadaei-Nobandegani, J. Moghadasi, *J. Ind. Eng. Chem.* **19** (2013) 769 (https://doi.org/10.1016/j.jiec.2012.10.013)
9. J. Gmehling, J. Li, M. Schiller, *Ind. Eng. Chem. Res.* **32** (1993) 178 (https://doi.org/10.1021/ie00013a024)
10. M. J. Hait, C. L. Liotta, C. A. Eckert, D. L. Bergmann, A. M. Karachewski, A. J. Dallas, et al., *Ind. Eng. Chem. Res.* **32** (1993) 2905 (https://doi.org/10.1021/ie00023a064)
11. S. Yousefinejad, F. Honarasa, H. Montaseri, *RSC Adv.* **5** (2015) 42266 (https://doi.org/10.1039/C5RA05930E)
12. A. Klamt, F. Eckert, *Fluid Phase Equilib.* **172** (2000) 43 (https://doi.org/10.1016/s0378-3812(00)00357-5)
13. D. E. Nanu, T. W. De Loos, *Mol. Phys.* **102** (2004) 235 (https://doi.org/10.1080/00268970410001655871)
14. A. A. Oliferenko, P. V. Oliferenko, J. S. Torrecilla, A. R. Katritzky, *Ind. Eng. Chem. Res.* **51** (2012) 9123 (https://doi.org/10.1021/ie202550v)
15. A. R. Katritzky, I. B. Stoyanova-Slavova, K. Tämm, T. Tamm, M. Karelson, *J. Phys. Chem., A* **115** (2011) 3475 (https://doi.org/10.1021/jp104287p)
16. I. Oprisiu, S. Novotarskyi, I. V. Tetko, *J. Cheminform.* **5** (2013) 4 (https://doi.org/10.1186/1758-2946-5-4)
17. V. Zare-Shahabadi, M. Lotfizadeh, A. R. A. Gandomani, M. M. Papari, *J. Mol. Liq.* **188** (2013) 222 (https://doi.org/10.1016/j.molliq.2013.09.037)
18. T. Gaudin, P. Rotureau, G. Fayet, *Ind. Eng. Chem. Res.* **54** (2015) 6596 (https://doi.org/10.1021/acs.iecr.5b01457)
19. Z. Faramarzi, F. Abbasitabar, V. Zare-Shahabadi, H. J. Jahromi, *J. Mol. Liq.* **296** (2019) 111854 (https://doi.org/10.1016/j.molliq.2019.111854)
20. Y. Demirel, *Thermochim. Acta* **339** (1999) 79 (https://doi.org/10.1016/s0040-6031(99)00211-7)
21. *ChemDraw Ultra 6.0 and Chem3D Ultra*, Cambridge Soft Corporation, Cambridge, MA
22. *MOE*, Chemical Computing Group Inc., Montreal (http://www.chemcomp.com)

23. R. Todeschini, V. Consonni, M. Pavan, *Dragon Software Version 2.1*, Chemometrics and QSAR Research Group, Milano, 2002
24. E. N. Muratov, E. V. Varlamova, A. G. Artemenko, P. G. Polishchuk, V. E. Kuz'min, *Mol. Inform.* **31** (2012) 202 (https://doi.org/10.1002/minf.201100129)
25. F. Abbasitabar, V. Zare-Shahabadi, *SAR QSAR Environ. Res.* **23** (2011) 1 (https://doi.org/10.1080/1062936x.2011.623316)
26. B. Hemmateenejad, M. Shamsipur, V. Zare-Shahabadi, M. Akhond, *Anal. Chim. Acta* **704** (2011) 57 (https://doi.org/10.1016/j.aca.2011.08.010)
27. V. Zare-Shahabadi, *Med. Chem. Res.* **25** (2016) 2787 (https://doi.org/10.1007/s00044-016-1666-z)
28. F. Abbasitabar, V. Zare-Shahabadi, *Chemosphere* **172** (2017) 249 (https://doi.org/10.1016/j.chemosphere.2016.12.095)
29. D. Baumann, K. Baumann, *J. Cheminform.* **6** (2014) 47 (https://doi.org/10.1186/s13321-014-0047-1)
30. D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. De Jong, P. J. Lewi, J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics Part A*, Elsevier, Amsterdam, 1997, pp. 286–288
31. S. Saaidpour, *Phys. Chem. Res.* **4** (2016) 61 (https://doi.org/10.22036/pcr.2016.11759)
32. F. Abbasitabar, V. Zare-Shahabadi, *Drug Res (Stuttgart)* **67** (2017) 476 (https://doi.org/10.1055/s-0043-108553)
33. K. Roy, S. Kar, P. Ambure, *Chemom. Intell. Lab. Syst.* **145** (2015) 22 (http://dx.doi.org/10.1016/j.chemolab.2015.04.013)
34. X. Bian, P. Diwu, Y. Liu, P. Liu, Q. Li, X. Tan, *J. Chemom.* **32** (2018) e2940 (https://doi.org/10.1002/cem.2940)
35. V. Zare-Shahabadi, F. Abbasitabar, M. Akhond, M. Shamsipur, *J. Braz. Chem. Soc.* **24** (2013) 1561 (http://dx.doi.org/10.5935/0103-5053.20130197)
36. S. Ma, S. Li, *Ind. Eng. Chem. Res.* **52** (2013) 543 (https://doi.org/10.1021/ie302909b)
37. A. A. Oliferenko, P. V. Oliferenko, J. S. Torrecilla, A. R. Katritzky, *Ind. Eng. Chem. Res.* **52** (2013) 545 (https://doi.org/10.1021/ie3033125)
38. S. Guariento, M. Tonelli, S. Espinoza, A. S. Gerasimov, R. R. Gainetdinov, E. Cichero, *Eur. J. Med. Chem.* **146** (2018) 171 (https://doi.org/10.1016/j.ejmech.2018.01.059)
39. O. Deeb, B. Hemmateenejad, *Chem. Biol. Drug Des.* **70** (2007) 19 (https://doi.org/10.1111/j.1747-0285.2007.00528.x)
40. R. Todeschini, V. Consonni, R. Mannhold, H. Kubinyi, G. Folkers. *Molecular Descriptors for Chemoinformatics: Volume I: Alphabetical Listing / Volume II: Appendices, References*, Wiley, New York, 2009, pp. 17–20.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 86 (7–8) S136–S144 (2021) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

New structure-based models for the prediction of normal boiling point temperature of ternary azeotropes

ZOHREH FARAMARZI[1], FATEMEH ABBASITABAR[2]*, HOSSEIN JALALI JAHROMI[1] and MAZIAR NOEI[1]

[1]*Department of Chemistry, Mahshahr Branch, Islamic Azad University, Mahshahr, Iran and* [2]*Department of Chemistry, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran*

J. Serb. Chem. Soc. 86 (7–8) (2021) 685–698

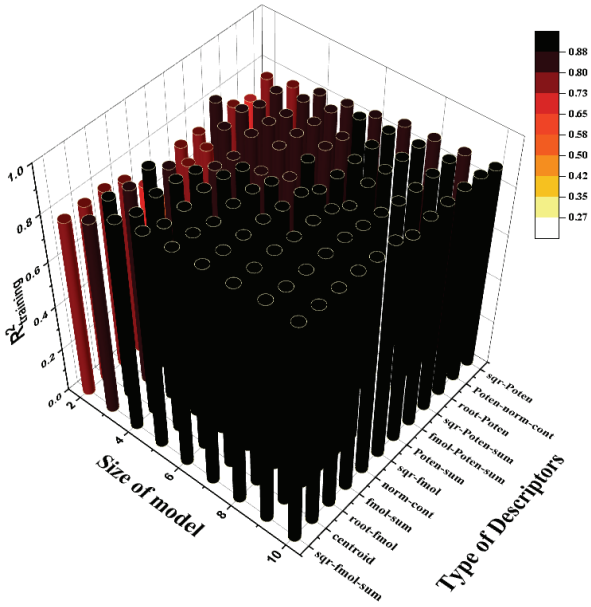


Fig. S-1. $R^2_{training}$ values for QSPR models with different sizes constructed on the basis of different types of mixture descriptors.

* Corresponding author. E-mail: fabbasitabar@gmail.com

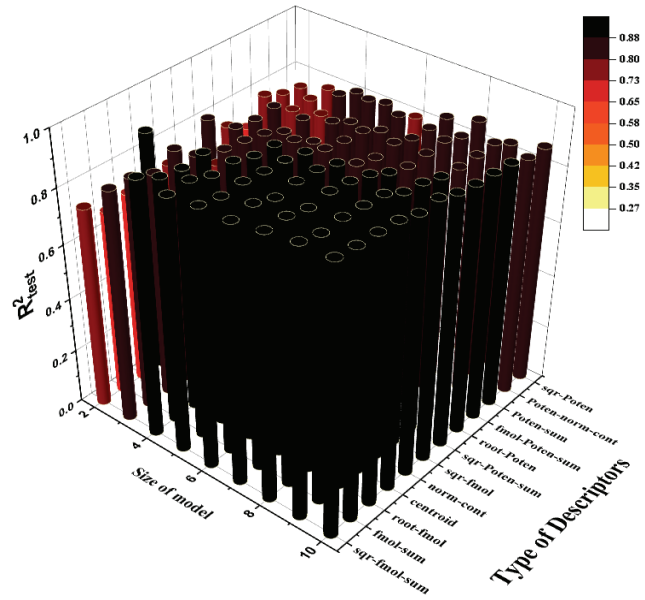


Fig. S-2. R^2_{test} values for QSPR models with different sizes constructed on the basis of different types of mixture descriptors.

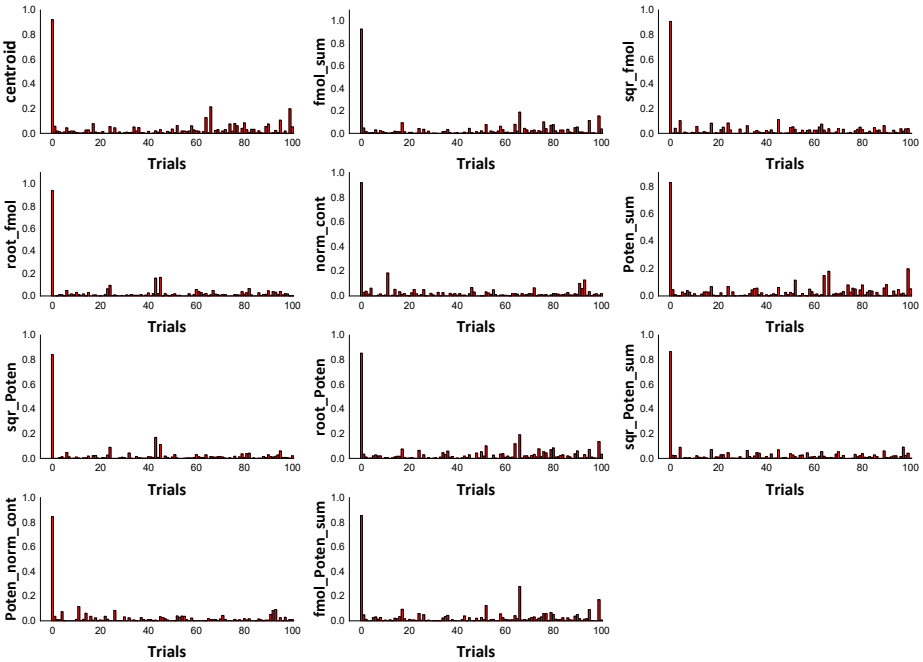


Fig. S-3. Y–randomization test applied to all QSAR model types; the first bar shows the $R^2_{training}$ value for corresponding original model.

Table S-I. Composition of mixtures appeared in the training set together with the observed and calculated boiling points by different QSPR models

Mixture composition			Observed	Calculated													
A	B	C		centroid	fmol-sum	sqr-fmol	root-fmol	sqr-fmol-sum	norm-cont	poten-sum	sqr-poten	root-poten	sqr-poten-sum	poten-norm-cont	fmol-poten-sum	Ensemble
Water	Carbon disulfide	Ethanol	314.4	313.7	309.8	316.9	313.6	316.8	315.1	336.8	337.6	334.7	339.3	314.3	332.9	323.5
Acetic acid	Pyridine	O-xylene	405.3	401.9	398.6	390.2	411.3	403.4	395.4	408.5	402.6	401.7	405.0	408.7	400.5	402.3
Water	Nitromethane	Dodecane	356.3	367.6	349.4	353.9	358.6	356.7	348.8	365.1	369.6	366.7	358.0	361.8	344.0	358.4
Methanol	Methyl acetate	Cyclohexane	323.9	331.4	334.3	341.8	330.4	334.0	324.6	326.2	331.3	332.5	331.9	333.0	339.1	332.6
Acetic acid	Pyridine	Octane	388.8	390.2	381.9	389.3	382.1	386.9	381.9	391.4	390.3	390.5	388.1	384.4	396.9	387.8
Water	Acetonitrile	Ethanol	346.0	350.0	339.5	343.0	340.0	341.2	342.0	345.8	335.9	330.3	342.1	334.9	334.1	339.9
Water	C-tetrachloride	Tert-butanol	337.8	342.3	337.5	335.5	340.8	341.6	337.3	337.9	337.8	336.1	336.9	340.9	336.7	338.4
Acetic acid	Benzene	Cyclohexane	350.3	360.5	347.1	337.6	353.1	346.9	347.0	349.8	352.6	353.7	349.2	355.8	337.2	349.2
Water	Butanol	Nonane	363.1	355.8	351.8	368.5	355.3	363.9	360.3	359.9	359.7	358.2	359.2	358.6	358.1	359.1
Acetic acid	Pyridine	Decane	407.2	402.3	409.4	408.8	403.6	402.3	412.6	386.3	387.9	394.2	404.2	396.4	387.8	399.7
Water	Chloroform	Methanol	325.4	324.2	320.5	327.6	326.4	325.2	328.9	331.5	330.6	333.1	334.8	336.3	331.2	329.2
Butanol	Pyridine	Toluene	381.8	376.4	377.5	392.3	379.1	385.9	385.5	378.7	378.1	387.9	379.8	382.7	375.2	381.6
Water	Butanol	Hexane	334.6	336.5	337.5	339.6	339.8	338.7	341.5	335.9	339.2	341.1	346.4	343.6	345.0	340.4
Water	Nitromethane	Octane	350.8	344.3	351.4	351.9	348.0	354.2	351.6	355.6	351.9	355.7	349.6	348.7	355.1	351.5
Water	2-butanone	Benzene	342.0	339.3	346.2	340.2	345.8	342.8	343.4	341.1	339.7	342.8	344.0	348.8	341.3	343.0
Water	Pyridine	Octane	359.8	356.8	361.7	357.7	358.4	352.8	361.2	361.1	358.7	360.3	354.4	350.2	363.4	358.0
Water	C-tetrachloride	Ethanol	334.9	338.4	337.7	336.0	337.6	339.0	337.6	342.0	337.1	338.9	339.9	338.2	335.4	338.1
Water	Pyridine	Decane	365.4	368.9	364.7	360.4	367.2	357.1	363.4	364.8	368.3	362.9	365.0	371.6	359.3	364.5
Water	Nitromethane	Undecane	355.9	362.1	349.9	353.7	357.2	356.9	349.4	360.3	364.3	359.8	361.4	364.4	346.3	357.1
Water	Isopropyl alcohol	Toluene	349.4	355.2	353.4	352.1	349.7	347.0	357.3	349.2	352.6	356.6	344.6	347.3	359.2	352.0
Chloroform	Methanol	Methyl acetate	329.5	329.4	325.0	333.9	326.3	325.4	324.8	319.4	326.5	331.6	326.0	325.6	334.3	327.4
Water	Trichloro ethylene	Acetonitrile	340.1	340.4	342.9	338.7	337.2	339.8	341.4	350.1	344.6	343.9	344.9	336.9	343.1	342.0
Chloroform	Ethanol	Acetone	341.3	336.3	331.7	336.0	337.7	332.4	339.7	343.5	329.9	332.5	340.7	333.7	330.8	335.4
Water	Acetonitrile	Benzene	339.1	341.2	341.5	338.8	336.6	337.7	341.8	340.7	350.9	340.6	339.2	348.2	340.2	341.5
Water	Nitromethane	Propyl alcohol	355.4	348.4	349.1	347.8	350.8	352.7	343.9	338.3	342.7	337.5	335.7	343.6	348.7	344.9
Water	Butanol	Octane	359.2	349.6	351.7	366.5	351.0	361.5	360.0	357.9	354.4	359.0	352.5	349.2	360.3	356.1
Water	Isobutanol	Toluene	354.4	355.8	357.0	363.3	351.8	349.4	361.6	343.1	342.9	355.7	342.1	349.0	360.3	352.7
Isopropyl alcohol	2-butanone	Cyclohexane	342.0	343.9	343.2	336.1	337.5	343.1	337.0	350.0	333.5	346.6	344.4	328.5	336.6	340.0
Water	Trichloro ethylene	Propyl alcohol	344.7	341.8	347.9	339.2	345.6	344.6	340.7	346.2	342.6	342.1	339.8	340.1	347.7	343.2
Water	Butanol	Heptane	351.2	343.4	349.9	360.2	349.1	354.6	355.7	342.7	343.7	348.3	347.5	343.3	356.7	349.6
Water	C-tetrachloride	Sec-butanol	338.1	342.0	336.6	335.6	338.3	340.0	335.7	337.9	337.8	336.2	336.8	340.9	338.4	338.0
Acetic acid	Pyridine	Ethylbenzene	402.2	404.5	400.1	393.4	404.5	405.0	397.3	402.9	406.7	400.4	401.5	397.4	404.3	401.5
Water	Nitromethane	1-decene	355.6	360.6	351.2	354.6	357.2	356.1	352.9	361.5	362.0	357.0	357.8	365.7	354.1	357.6
Water	Tert-butanol	Cyclohexane	338.1	339.8	342.9	338.1	337.6	333.4	335.9	340.1	338.3	344.1	337.5	344.4	342.0	339.5
Water	Ethanol	Cyclohexane	335.7	335.9	341.6	336.8	335.1	336.3	334.3	337.3	340.0	329.4	336.1	340.5	339.5	336.9
Water	Nitromethane	Nonane	353.8	350.4	352.1	352.7	354.5	356.1	352.5	358.1	357.2	356.8	356.9	357.2	355.7	355.0
Water	Nitromethane	Isopropyl alcohol	351.1	350.3	346.0	342.1	351.2	350.5	346.4	338.2	342.7	337.7	335.6	343.6	339.9	343.7
Water	Chloroform	Acetone	333.5	325.5	335.5	340.6	339.0	329.4	345.4	332.2	334.2	331.2	335.7	339.6	334.1	335.2
Water	Ethanol	Hexane	329.5	334.3	330.7	325.1	335.0	337.7	334.1	333.5	345.5	325.0	354.0	333.3	321.4	334.1
Water	Isopropyl alcohol	2-butanone	346.5	349.4	354.4	354.5	352.5	346.2	356.3	341.8	339.6	342.6	343.0	352.0	344.5	348.1
Water	Trichloro ethylene	Isopropyl alcohol	342.5	343.7	346.6	339.1	344.9	343.7	341.4	346.2	342.6	342.1	339.8	340.1	345.7	343.0
Water	Nitromethane	3-pentanone	355.5	349.8	356.6	346.3	347.3	353.0	346.4	342.9	340.2	347.1	344.5	341.4	350.3	347.1
Chloroform	Methanol	Acetone	330.6	326.6	326.6	334.3	326.1	324.2	328.0	323.4	327.9	326.7	330.2	330.1	332.0	328.0
Methanol	Acetone	Cyclohexane	324.2	328.6	335.0	338.8	329.2	331.1	332.4	323.2	330.4	324.9	327.3	323.1	333.8	329.8
Water	Ethanol	Heptane	341.9	341.2	345.6	343.5	341.8	346.8	346.1	351.9	358.7	342.8	353.8	346.9	339.3	346.5
Acetic	Pyridine	Nonane	401.1	396.4	401.9	398.6	399.2	399.7	402.5	387.3	387.6	391.0	400.5	392.9	407.4	397.1

acid																
Water	Butanol	Butylacetate	362.5	365.8	368.9	361.5	368.2	362.0	366.9	369.8	370.3	368.4	362.0	369.3	357.9	365.9
Water	Ethanol	Acrylonitrile	342.6	345.2	357.2	343.6	347.7	349.0	349.5	337.0	338.1	340.5	339.4	346.5	340.5	344.5
Water	Nitromethane	1-heptene	342.8	341.9	349.7	348.9	343.1	350.2	348.0	343.2	343.7	344.3	341.0	345.4	350.7	345.8
Water	Propyl alcohol	Cyclohexane	339.7	335.2	343.0	337.6	335.7	332.5	333.7	335.4	339.2	336.7	338.3	341.6	342.9	337.6
Isopropyl alcohol	Benzene	Cyclohexane	342.2	339.1	341.0	335.9	337.0	341.4	339.0	345.9	339.6	340.1	340.5	348.5	338.8	340.6
Water	Nitromethane	1-octene	350.5	348.9	353.7	352.7	354.1	356.1	352.6	358.0	352.1	356.9	347.1	349.7	358.0	353.3
Water	Chloroform	Ethanol	328.4	333.9	324.1	326.9	337.0	329.2	334.3	333.2	332.4	333.1	335.7	337.4	332.7	332.5
Water	Ethanol	Triethylamine	347.8	349.2	345.7	347.7	348.5	342.1	343.3	345.7	342.5	355.7	344.9	340.3	345.1	345.9

Table S-II. Composition of mixtures appeared in the test set together with the observed and calculated boiling points by different QSPR models

Mixture composition			Observed	Calculated												
A	B	C		centroid	fmol-sum	sqr-fmol	root-fmol	sqr-fmol- sum	norm-cont	poten-sum	sqr-poten	root-poten	sqr-poten-sum	poten-norm-cont	fmol-poten-sum	Ensemble
Water	Carbon tetrachloride	Propyl alcohol	338.5	337.8	340.0	336.8	337.3	338.5	337.1	339.6	337.4	338.5	337.9	339.1	337.7	338.1
Water	Carbon tetrachloride	2-butanone	338.8	334.4	342.7	338.5	339.5	344.6	341.2	338.2	337.7	335.2	337.0	340.5	335.7	338.8
Water	Carbon disulfide	Acetone	311.2	305.3	316.6	325.6	317.8	319.8	325.2	331.3	326.4	330.5	333.6	336.7	330.2	324.9
Water	Nitromethane	1-hexene	328.6	336.6	336.9	325.9	338.6	342.4	333.8	336.2	340.8	333.1	338.2	342.8	337.8	336.9
Water	Nitromethane	Hexane	330.0	331.2	332.8	327.5	339.4	338.3	332.9	335.2	340.6	333.6	339.5	342.1	336.7	335.8
Water	Nitromethane	Heptane	343.7	338.1	347.8	346.0	348.5	350.6	344.4	341.3	342.9	343.0	343.2	343.6	344.6	344.5
Water	Nitromethane	1-nonene	354.2	354.8	352.7	353.1	355.8	356.2	353.2	361.0	358.1	357.8	353.8	356.7	357.1	355.9
Water	Nitromethane	Decane	355.6	356.4	350.9	353.4	354.9	356.3	351.5	361.1	362.5	359.5	360.5	366.7	351.9	357.1
Water	Acetonitrile	Triethylamine	341.7	347.1	343.1	351.0	346.6	344.1	336.7	346.7	343.2	356.3	347.4	347.5	345.0	346.2
Water	1,2 dichloroethane	Ethanol	340.9	342.6	342.0	332.8	339.8	346.2	337.8	339.7	328.8	332.6	346.5	336.2	340.5	338.8
Water	1,2 dichloroethane	Isopropyl alcohol	342.8	343.9	342.4	333.0	340.1	343.7	336.9	336.9	334.0	337.5	343.7	342.1	341.7	339.7
Water	Ethanol	Ethyl acrylate	350.2	363.3	354.0	344.9	353.7	354.4	346.6	363.7	364.6	352.4	355.6	361.0	344.9	354.9
Water	Ethanol	Chlorobenzene	350.4	362.4	346.4	349.4	358.0	349.6	352.5	361.6	375.3	363.7	355.1	357.0	340.0	355.9
Water	Ethanol	Benzene	338.0	343.3	343.9	339.2	343.9	341.6	342.6	337.6	344.3	340.1	338.5	339.6	341.3	341.3
Water	Ethanol	Toluene	347.5	353.9	352.6	352.3	349.2	348.2	358.3	358.2	369.4	355.3	353.8	352.5	362.9	355.6
Water	Isopropyl alcohol	Benzene	339.6	344.6	344.1	338.8	344.6	341.6	341.5	335.2	339.8	345.4	338.0	343.2	341.1	341.5
Water	Isopropyl alcohol	Cyclohexane	337.4	337.1	342.4	337.2	335.1	333.4	334.2	335.2	339.2	337.0	337.2	341.2	341.6	337.6
Water	Propyl alcohol	Benzene	340.1	342.7	344.6	339.2	345.9	342.3	341.6	335.7	339.8	345.1	339.9	344.5	341.4	341.9
Water	2-butanone	Cyclohexane	336.7	331.9	346.5	339.4	336.9	338.8	339.0	338.7	339.2	338.0	342.0	346.3	338.8	339.6
Water	Butanol	Benzene	342.1	345.6	344.6	340.8	350.5	341.4	342.2	339.3	338.4	349.2	344.8	349.0	341.4	343.9
Water	Tert-butanol	Benzene	340.4	347.2	344.6	339.7	346.9	342.3	342.7	342.9	337.8	349.3	337.5	343.8	341.3	343.0
Water	Pyridine	Nonane	363.6	362.9	361.9	360.9	365.5	357.0	363.2	362.5	363.5	357.4	361.9	361.0	362.5	361.7
Chloroform	Ethanol	Hexane	330.4	326.8	324.2	328.5	327.1	333.5	332.3	349.7	331.5	329.4	340.8	329.2	331.4	332.0
Acetic acid	Pyridine	P-xylene	402.3	404.2	390.7	398.8	401.6	404.0	395.1	413.7	421.6	418.9	422.6	415.6	413.5	408.4

Table S-III. $R^2_{training}$, R^2_{test}, and q^2 of MLR models with different sizes developed based on different mixture descriptors

Model size	centroid	fmol–sum	sqr–fmol	root–fmol	sqr–fmol–sum	norm–cont	poten–sum	sqr–poten	root–poten	sqr–poten–sum	poten–norm–cont	fmol–poten–sum
$R^2_{training}$												
2	0.79	0.77	0.57	0.79	0.79	0.69	0.75	0.77	0.75	0.82	0.73	0.75
3	0.88	0.89	0.79	0.86	0.85	0.83	0.80	0.79	0.81	0.85	0.82	0.83
4	0.89	0.90	0.84	0.9	0.92	0.88	0.82	0.82	0.83	0.85	0.84	0.86
5	0.91	0.93	0.9	0.92	0.95	0.92	0.83	0.83	0.86	0.87	0.86	0.87
6	0.95	0.94	0.91	0.93	0.96	0.94	0.87	0.84	0.86	0.88	0.89	0.88
7	0.95	0.95	0.93	0.95	0.96	0.94	0.88	0.87	0.89	0.89	0.89	0.88
8	0.96	0.95	0.93	0.95	0.96	0.95	0.89	0.87	0.9	0.91	0.9	0.92
9	0.96	0.96	0.94	0.96	0.98	0.95	0.93	0.88	0.9	0.91	0.9	0.91
10	0.97	0.96	0.95	0.96	0.97	0.95	0.91	0.88	0.92	0.92	0.91	0.92
R^2_{test}												

Model size	centroid	fmol–sum	sqr–fmol	root–fmol	sqr–fmol–sum	norm–cont	poten–sum	sqr–poten	root–poten	sqr–poten–sum	poten–norm–cont	fmol–poten–sum
2	0.88	0.68	0.55	0.71	0.74	0.73	0.79	0.75	0.75	0.81	0.76	0.71
3	0.86	0.85	0.74	0.82	0.85	0.81	0.78	0.78	0.82	0.83	0.78	0.82
4	0.9	0.89	0.86	0.88	0.93	0.87	0.80	0.81	0.81	0.84	0.84	0.82
5	0.91	0.93	0.89	0.9	0.92	0.89	0.81	0.81	0.82	0.87	0.86	0.85
6	0.92	0.92	0.91	0.93	0.94	0.92	0.83	0.80	0.87	0.88	0.84	0.86
7	0.94	0.95	0.91	0.92	0.94	0.93	0.84	0.84	0.85	0.87	0.84	0.87
8	0.95	0.96	0.92	0.94	0.95	0.94	0.87	0.87	0.87	0.89	0.86	0.87
9	0.95	0.95	0.91	0.96	0.97	0.95	0.88	0.84	0.88	0.9	0.88	0.87
10	0.95	0.96	0.94	0.96	0.97	0.93	0.89	0.87	0.9	0.92	0.87	0.9
q^2												
2	0.77	0.73	0.46	0.77	0.76	0.63	0.71	0.74	0.71	0.81	0.69	0.7
3	0.86	0.87	0.74	0.84	0.82	0.81	0.77	0.76	0.77	0.83	0.79	0.79
4	0.87	0.88	0.78	0.87	0.91	0.86	0.79	0.77	0.8	0.83	0.81	0.81
5	0.89	0.91	0.86	0.9	0.93	0.9	0.79	0.78	0.82	0.85	0.83	0.83
6	0.93	0.91	0.87	0.91	0.94	0.91	0.83	0.8	0.82	0.85	0.86	0.84
7	0.93	0.92	0.89	0.93	0.94	0.91	0.84	0.82	0.85	0.87	0.86	0.84
8	0.94	0.93	0.9	0.94	0.95	0.92	0.84	0.82	0.86	0.89	0.86	0.88
9	0.94	0.93	0.91	0.94	0.96	0.93	0.89	0.82	0.86	0.89	0.87	0.87
10	0.95	0.94	0.93	0.95	0.96	0.93	0.86	0.83	0.88	0.89	0.87	0.87

Table S-IV. Statistical parameters of the QSPR models based on centroid descriptors

No. of descriptors	$R^2_{training}$	RMSE-Training	q^2	RMSE-CV	5-CV[a]		MCCV[b]		R^2_{test}	RMSE-Test	F
					q^2	RMSE	$R^2_{training}$	R^2_{test}			
2	0.79	9.26	0.77	9.83	0.76	9.94	0.79	0.75	0.88	6.44	97.09
3	0.88	6.91	0.86	7.54	0.86	7.65	0.88	0.84	0.86	7.29	127.2
4	0.89	6.75	0.87	7.40	0.86	7.53	0.89	0.84	0.90	6.53	98.74
5	0.91	6.00	0.89	6.77	0.88	6.90	0.91	0.87	0.91	5.13	100.4
6	0.95	4.68	0.93	5.38	0.93	5.49	0.95	0.91	0.92	5.40	139.7
7	0.95	4.47	0.93	5.28	0.93	5.42	0.95	0.91	0.94	5.00	128.9
8	0.96	4.20	0.94	5.06	0.93	5.20	0.96	0.92	0.95	3.91	125.6
9	0.96	4.07	0.94	5.06	0.93	5.20	0.96	0.92	0.95	3.57	116.7
10	0.97	3.71	0.95	4.64	0.94	4.89	0.97	0.94	0.95	3.44	124.4

[a]The reported values are average of 100 repetitions of 5-fold CV
[b]The reported values are average of 100 repetitions of MCCV

Table S-V. Statistical parameters of the QSPR models based on fmol-sum descriptors

No. of descriptors	$R^2_{training}$	RMSE-Training	q^2	RMSE-CV	5-CV[a]		MCCV[b]		R^2_{test}	RMSE-Test	F
					q^2	RMSE	$R^2_{training}$	R^2_{test}			
2	0.77	9.74	0.73	10.56	0.73	10.59	0.77	0.73	0.68	9.18	85.46
3	0.89	6.75	0.87	7.31	0.87	7.46	0.89	0.83	0.85	6.27	134.32
4	0.90	6.38	0.88	7.02	0.88	7.20	0.90	0.85	0.89	5.21	112.04
5	0.93	5.31	0.91	5.99	0.91	6.10	0.93	0.89	0.93	5.09	130.96
6	0.94	5.06	0.91	6.04	0.91	6.22	0.94	0.89	0.92	4.66	118.40
7	0.95	4.68	0.92	5.72	0.91	6.12	0.95	0.91	0.95	4.04	117.12
8	0.95	4.57	0.93	5.47	0.92	5.62	0.95	0.90	0.96	3.57	123.10
9	0.96	4.29	0.93	5.27	0.93	5.46	0.96	0.91	0.95	4.45	120.31

10	0.96	3.93	0.94	4.90	0.94	5.11	0.96	0.93	0.96	3.96	110.68

[a]The reported values are average of 100 repetitions of 5-fold CV
[b]The reported values are average of 100 repetitions of MCCV

Table S-VI. Statistical parameters of the QSPR models based on sqr-fmol descriptors

No. of descriptors	$R^2_{training}$	RMSE-Training	q^2	RMSE-CV	5-CV[a]		MCCV[b]		R^2_{test}	RMSE-Test	F
					q^2	RMSE	$R^2_{training}$	R^2_{test}			
2	0.57	13.28	0.46	14.91	0.45	15.29	0.57	0.45	0.55	11.64	34.12
3	0.79	9.33	0.74	10.33	0.74	10.47	0.79	0.70	0.74	9.53	62.39
4	0.84	8.18	0.78	9.49	0.75	10.22	0.84	0.73	0.86	6.00	63.26
5	0.90	6.46	0.86	7.62	0.84	8.02	0.90	0.83	0.89	5.35	85.21
6	0.91	6.19	0.87	7.37	0.85	7.85	0.91	0.83	0.91	4.88	76.62
7	0.93	5.37	0.89	6.66	0.87	7.29	0.93	0.86	0.91	4.86	87.43
8	0.93	5.24	0.90	6.48	0.89	6.80	0.94	0.86	0.92	4.44	78.99
9	0.94	5.08	0.91	6.15	0.89	6.59	0.94	0.86	0.91	5.08	73.17
10	0.95	4.52	0.93	5.50	0.90	6.35	0.95	0.88	0.94	5.17	82.42

[a]The reported values are average of 100 repetitions of 5-fold CV
[b]The reported values are average of 100 repetitions of MCCV

Table S-VII. Statistical parameters of the QSPR models based on root-fmol descriptors

No. of descriptors	$R^2_{training}$	RMSE-Training	q^2	RMSE-CV	5-CV[a]		MCCV[b]		R^2_{test}	RMSE-Test	F
					q^2	RMSE	$R^2_{training}$	R^2_{test}			
2	0.79	9.24	0.77	9.69	0.77	9.75	0.79	0.74	0.71	8.85	97.83
3	0.86	7.47	0.84	8.08	0.84	8.17	0.86	0.81	0.82	7.32	106.71
4	0.90	6.27	0.87	7.23	0.87	7.35	0.90	0.84	0.88	5.71	116.50
5	0.92	5.77	0.90	6.47	0.89	6.62	0.92	0.86	0.90	5.13	109.24
6	0.93	5.42	0.91	6.07	0.91	6.25	0.93	0.88	0.93	4.45	102.21
7	0.95	4.67	0.93	5.48	0.92	5.77	0.95	0.90	0.92	4.70	117.74
8	0.96	4.09	0.94	4.94	0.94	5.01	0.96	0.93	0.94	4.96	133.34
9	0.96	4.16	0.94	5.13	0.93	5.32	0.96	0.90	0.96	3.69	111.86
10	0.96	3.84	0.95	4.74	0.94	4.88	0.97	0.93	0.96	3.92	116.17

[a]The reported values are average of 100 repetitions of 5-fold CV.
[b]The reported values are average of 100 repetitions of MCCV.

Table S-VIII. Statistical parameters of the QSPR models based on norm-count descriptors

No. of descriptors	$R^2_{training}$	RMSE-Training	q^2	RMSE-CV	5-CV[a]		MCCV[b]		R^2_{test}	RMSE-Test	F
					q^2	RMSE	$R^2_{training}$	R^2_{test}			
2	0.69	11.24	0.63	12.36	0.62	12.61	0.69	0.60	0.73	8.60	57.77
3	0.83	8.29	0.81	8.76	0.81	8.85	0.83	0.79	0.81	7.26	83.46
4	0.88	7.03	0.86	7.62	0.85	7.80	0.88	0.82	0.87	5.82	89.89
5	0.92	5.59	0.90	6.36	0.90	6.61	0.92	0.88	0.89	5.50	117.02
6	0.94	5.17	0.91	6.12	0.90	6.47	0.94	0.89	0.92	4.84	113.09
7	0.94	5.16	0.91	6.05	0.90	6.39	0.94	0.90	0.93	4.18	95.11
8	0.95	4.60	0.92	5.66	0.92	5.82	0.95	0.91	0.94	4.83	104.15

9	0.95	4.39	0.93	5.33	0.92	5.64	0.95	0.91	0.95	4.28	100.01
10	0.95	4.51	0.93	5.30	0.92	5.67	0.95	0.89	0.93	4.20	82.74

[a]The reported values are average of 100 repetitions of 5-fold CV
[b]The reported values are average of 100 repetitions of MCCV

Table S-IX. Sttistical parameters of the QSPR models based on Poten-sum descriptors

No. of descriptors	$R^2_{training}$	RMSE-Training	q^2	RMSE-CV	5-CV[a]		MCCV[b]		R^2_{test}	RMSE-Test	F
					q^2	RMSE	$R^2_{training}$	R^2_{test}			
2	0.75	10.23	0.71	10.85	0.71	10.95	0.74	0.69	0.79	9.39	75.11
3	0.80	9.17	0.77	9.79	0.76	9.92	0.80	0.75	0.78	10.12	65.10
4	0.82	8.54	0.79	9.32	0.78	9.47	0.83	0.77	0.80	10.62	56.99
5	0.83	8.45	0.79	9.27	0.79	9.41	0.83	0.78	0.81	8.73	45.85
6	0.87	7.38	0.83	8.29	0.83	8.50	0.87	0.82	0.83	8.22	51.52
7	0.88	6.94	0.84	8.05	0.84	8.24	0.89	0.83	0.84	7.91	49.77
8	0.89	6.87	0.84	8.07	0.83	8.34	0.89	0.84	0.88	7.61	43.60
9	0.93	5.48	0.89	6.82	0.86	7.87	0.93	0.87	0.87	7.37	62.19
10	0.91	6.14	0.86	7.56	0.85	7.89	0.91	0.84	0.89	6.82	48.63

[a]The reported values are average of 100 repetitions of 5-fold CV
[b]The reported values are average of 100 repetitions of MCCV

Table S-X. Statistical parameters of the QSPR models based on root-Poten descriptors.

No. of descriptors	$R^2_{training}$	RMSE-Training	q^2	RMSE-CV	5-CV[a]		MCCV[b]		R^2_{test}	RMSE-Test	F
					q^2	RMSE	$R^2_{training}$	R^2_{test}			
2	0.75	10.08	0.71	10.87	0.71	11.00	0.76	0.72	0.75	8.64	78.10
3	0.81	8.84	0.77	9.65	0.77	9.78	0.81	0.77	0.82	7.45	71.42
4	0.83	8.32	0.80	9.14	0.79	9.30	0.83	0.78	0.81	8.53	60.82
5	0.86	7.63	0.82	8.56	0.82	8.71	0.86	0.79	0.82	7.85	58.48
6	0.86	7.57	0.82	8.55	0.82	8.74	0.87	0.78	0.87	7.47	48.51
7	0.89	6.75	0.85	7.95	0.84	8.09	0.89	0.82	0.85	7.93	52.91
8	0.90	6.32	0.86	7.66	0.85	7.85	0.91	0.83	0.87	6.35	52.56
9	0.90	6.31	0.86	7.58	0.85	7.86	0.90	0.84	0.88	7.67	52.59
10	0.92	5.91	0.88	7.18	0.87	7.39	0.92	0.86	0.90	7.50	52.84

[a] The reported values are average of 100 repetitions of 5-fold CV
[b] The reported values are average of 100 repetitions of MCCV

Table S-XI. Statistical parameters of the QSPR models based on sqr-Poten-sum descriptors

No. of descriptors	$R^2_{training}$	RMSE-Training	q^2	RMSE-CV	5-CV[a]		MCCV[b]		R^2_{test}	RMSE-Test	F
					q^2	RMSE	$R^2_{training}$	R^2_{test}			
2	0.82	8.54	0.81	8.86	0.81	8.94	0.82	0.75	0.81	8.98	118.91
3	0.85	8.00	0.83	8.45	0.82	8.55	0.85	0.77	0.83	9.22	90.92
4	0.85	7.80	0.83	8.31	0.83	8.40	0.85	0.76	0.84	9.32	70.92
5	0.87	7.43	0.85	7.88	0.85	7.96	0.86	0.78	0.87	7.73	62.09
6	0.88	7.15	0.85	7.92	0.84	8.07	0.88	0.79	0.88	6.95	55.35
7	0.89	6.82	0.87	7.43	0.86	7.70	0.89	0.82	0.87	8.47	51.70
8	0.91	6.03	0.89	6.62	0.89	6.89	0.91	0.85	0.89	7.23	58.16

9	0.91	6.24	0.89	6.86	0.88	7.19	0.91	0.84	0.90	7.25	53.91
10	0.92	5.86	0.89	6.75	0.89	6.90	0.92	0.85	0.92	6.76	61.88

[a]The reported values are average of 100 repetitions of 5-fold CV
[b]The reported values are average of 100 repetitions of MCCV

Table S-XII. Statistical parameters of the QSPR models based on Poten-norm-cont descriptors

No. of descriptors	$R^2_{training}$	RMSE-Training	q^2	RMSE-CV	5-CV[a]		MCCV[b]		R^2_{test}	RMSE-Test	F
					q^2	RMSE	$R^2_{training}$	R^2_{test}			
2	0.73	10.64	0.69	11.32	0.68	11.46	0.72	0.66	0.76	11.32	67.40
3	0.82	8.70	0.79	9.35	0.79	9.42	0.82	0.78	0.78	8.63	74.18
4	0.84	8.11	0.81	8.84	0.80	8.98	0.84	0.77	0.84	8.20	64.58
5	0.86	7.70	0.83	8.47	0.82	8.63	0.86	0.78	0.86	7.68	57.24
6	0.89	6.81	0.86	7.69	0.85	8.01	0.89	0.82	0.84	8.56	61.78
7	0.89	6.65	0.86	7.58	0.85	7.95	0.89	0.82	0.84	8.49	54.77
8	0.90	6.44	0.86	7.57	0.85	8.06	0.90	0.81	0.86	8.39	50.26
9	0.90	6.34	0.87	7.38	0.83	8.53	0.90	0.80	0.88	7.64	45.30
10	0.91	6.22	0.87	7.46	0.78	10.10	0.91	0.80	0.87	8.01	41.52

[a]The reported values are average of 100 repetitions of 5-fold CV
[b]The reported values are average of 100 repetitions of MCCV

Table S-XIII. Statistical parameters of the QSPR models based on fmol-Poten-sum descriptors

No. of descriptors	$R^2_{training}$	RMSE-Training	q^2	RMSE-CV	5-CV[a]		MCCV[b]		R^2_{test}	RMSE-Test	F
					q^2	RMSE	$R^2_{training}$	R^2_{test}			
2	0.75	10.24	0.70	11.18	0.69	11.49	0.75	0.66	0.71	9.51	74.84
3	0.83	8.45	0.79	9.33	0.78	9.76	0.83	0.79	0.82	7.44	79.60
4	0.86	7.64	0.81	8.82	0.81	9.07	0.86	0.83	0.82	7.14	74.36
5	0.87	7.29	0.83	8.34	0.81	8.91	0.87	0.80	0.85	6.70	65.01
6	0.88	6.98	0.84	8.21	0.82	8.75	0.88	0.79	0.86	6.78	58.60
7	0.88	6.92	0.84	8.17	0.83	8.47	0.89	0.84	0.87	6.98	50.08
8	0.92	5.91	0.88	6.92	0.88	7.16	0.92	0.88	0.87	7.10	84.72
9	0.91	6.03	0.87	7.43	0.85	7.87	0.91	0.82	0.87	7.28	58.11
10	0.92	5.74	0.87	7.30	0.86	7.66	0.92	0.82	0.90	6.54	56.24

[a]The reported values are average of 100 repetitions of 5-fold CV
[b]The reported values are average of 100 repetitions of MCCV

Table S-XIV. QSPR equation formulas established by MLR analysis for different types of mixture descriptors along with their statistical parameters

mixture descriptor	QSPR equation	$R^2_{training}$	q^2	R^2_{test}	Q^2_{F1}	Q^2_{F2}	Q^2_{F3}	CCC	r^2_m
centroid	y = 349.75 + 7.35 (±1.094) Ss + 11.04 (±1.060) piPC05 - - 4.70(±0.986) MATS2e + 8.75 (±0.869) a_donacc - - 8.65 (±0.978) GCUT_SMR_0 + 6.65 (±1.302) vsurf_D7	0.91	0.89	0.91	0.90	0.88	0.93	0.96	0.90
fmol–sum	y = 349.75 -6.33 (±0.966) BAC + 4.47 (±1.485) E2e + + 31.26 (±2.404) HATS1u + 5.44 (±1.262) vsurf_HL2 + + 36.20 (±2.637) vsurf_W1	0.9	0.88	0.89	0.91	0.90	0.94	0.95	0.87

sqr–fmol	*y* = 349.75 -9.11 (±0.976) MSD + 9.92 (±1.107) piPC06 + + 5.79 (±1.016) DP17 -5.80 (±0.984) GCUT_PEOE_1 - - 2.49 (±1.202) PEOE_VSA+1 + 10.89 (±1.059) vsurf_HL2	0.9	0.86	0.89	0.92	0.91	0.94	0.95	0.82
root–fmol	*y* = 349.75 + 9.60 (±1.499) ZM2V -5.86 (±1.143) MSD + + 1.38 (±1.025) piPC06 -5.36 (±1.110) E1e -3.60 (±0.854) G2s + + 4.49 (±0.973) R1m + 5.61 (±0.815) vsurf_DD23 + + 10.29 (±1.079) vsurf_HL2	0.95	0.93	0.92	0.92	0.90	0.94	0.98	0.92
sqr–fmol–sum	*y* = 349.75 + 8.59 (±1.010) ZM2V -7.22 (±1.445) X0A + + 5.25 (±1.238) E2e + 3.49 (±0.739) HATS3u + + 15.69 (±1.645) R1u+ + 23.16 (±1.776) vsurf_W1	0.95	0.93	0.92	0.92	0.90	0.94	0.97	0.91
norm–cont	*y* = 349.75 + 7.06 (±0.918) piPC06 + 3.87 (±0.975) DP17 + 23.40 (±2.465) R1u+ + 3.37 (±0.875) FCASA - - 17.87 (±2.529) Q_VSA_FPPOS + 11.55 (±1.109) vsurf_HL2	0.92	0.9	0.89	0.92	0.91	0.94	0.95	0.87
poten–sum	*y* = 349.75 + 18.52 (±2.646) SRW10 -6.76 (±2.317) X5A + + 8.35 (±1.411) BIC2 -2.87 (±1.859) RDF045v + +7.47 (±1.658) E2e + 10.65 (±1.790) R7v+	0.83	0.79	0.81	0.77	0.73	0.84	0.92	0.75
sqr–poten	*y* = 349.75 + 8.44 (±1.964) piPC06 + 5.88 (±2.093) Mor11m - 4.92 (±1.589) E1s + 7.22 (±1.599) R7v+ + + 4.07 (±1.858) DCASA - 6.06 (±1.994) vsurf_EDmin3	0.83	0.79	0.81	0.64	0.58	0.74	0.88	0.70
root–poten	*y* = 349.75 + 3.10 (±2.067) SPI + 19.34 (±2.881) SRW01 - - 8.49 (±2.903) X5A -6.48 (±1.663) X2Av - -8.87 (±1.534) Mor32v + 7.70 (±2.123) R7v+ + + 3.07 (±1.447) Q_PC+	0.86	0.82	0.87	0.78	0.75	0.85	0.93	0.79
sqr–poten–sum	*y* = 349.75 + 10.13 (±3.632) EEig01x + 15.22 (±2.486) EEig02r - - 6.02 (±1.365) EEig06r + 1.77 (±1.165) JGI2 + 8.83 (±2.494) E2e	0.85	0.83	0.84	0.79	0.76	0.86	0.94	0.74
poten–norm–cont	*y* = 349.75 -19.70 (±2.358) SIC5 -9.62 (±1.468) EEig05d + + 4.80 (±1.560) GGI4 -4.14 (±1.046) L/Bw - 7.65 (±1.424) VAdjEq - 3.13 (±1.706) vsurf_IW1	0.86	0.83	0.86	0.75	0.71	0.82	0.93	0.78
fmol–poten–sum	*y* = 349.75 + 12.48 (±1.634) piPC06 + 3.52 (±1.611) GGI2 + + 14.84 (±1.446) DP17 -7.37 (±1.939) Mor26e + + 4.67 (±1.272) Q_PC+	0.86	0.81	0.82	0.85	0.82	0.89	0.93	0.75
Ensemble model	---	0.97	0.96	0.95	0.92	0.91	0.95	0.98	0.87

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. *86* (7–8) 699–710 (2021)
JSCS–5455
Original scientific paper

Organophosphorous pesticide removal from water by graphene-based materials – Only adsorption or something else as well?

VLADAN ANIĆIJEVIĆ[1], MARKO JELIĆ[2], ALEKSANDAR Z. JOVANOVIĆ[3], NEBOJŠA POTKONJAK[2], IGOR A. PAŠTI[3]*# and TAMARA D. LAZAREVIĆ PAŠTI[2]

[1]Military Technical Institute (VTI), Ratka Resanovića 1, 11132 Belgrade, Serbia, [2]Vinča Institute of Nuclear Sciences – National Institute of the Republic of Serbia, University of Belgrade, Mihajla Petrovića Alasa 12–14, Belgrade, Serbia and [3]University of Belgrade – Faculty of Physical Chemistry, Studentski trg 12–16, 11158 Belgrade, Serbia

(Received 8 January, revised 11 February, accepted 12 February 2021)

Abstract: The extensive use of pesticides requires innovative approaches to remove these compounds from the environment. Carbon materials are traditionally used as adsorbents for removing pesticides, and the development of the new sorts of carbon materials allows more advanced approaches in environmental applications. Using density functional calculations, we have predicted chemical reaction between the S(O)=P moieties of the organophosphates with point defects in graphene – single vacancies, Stone–Wales defects and epoxy--groups. The reaction was confirmed using ultra high performance chromatography for two graphene oxide samples and dimethoate as a representative of organophosphates. The exact reaction mechanism is still elusive, but it is unambiguously confirmed that no selective oxidation of dimethoate to more toxic oxo-analog occurs. The presented results can help to develop novel systems for the irreversible conversion of organophosphates to non-toxic compounds, without using aggressive chemical agents or external physical factors like UV radiation.

Keywords: graphene; pesticide; removal; adsorption; chemical reaction.

INTRODUCTION

The growing world population caused enlarged food production, which, on the other hand, requires different strategies to boost the yield. One of the strategies is to use different pesticides to protect crops and increase the amounts of obtained plant products. Today, different classes of pesticides exist, differing by their action and level of sophistication. However, less developed countries still use some older families of pesticides, which can have a significant impact, not

* Corresponding author. E-mail: igor@ffh.bg.ac.rs
# Serbian Chemical Society member.
https://doi.org/10.2298/JSC210108012A

only on their target species, but also on humans and the living world in general.[1] The mentioned fact particularly relates to the use of organophosphorus pesticides (OPs), which rely their action on the inhibition of acetylcholine-esterase (AChE).[2] This enzyme is present in mammals, and the extensive exposure to OPs can cause severe neurotoxic effects.[3]

To reduce the health and the environmental impact of OPs, there are different remediation strategies. These include either removing OPs *via* suitable substrates, or their degradation using chemical, physical or biological agents,[4–8] some of them being potentially rather aggressive. The adsorption of OPs on different substrates requires high surface are adsorbents, which have to be environmentally friendly and affordable for large scale use. Among many possible candidates, the carbon materials seem to be perfect due to the (generally) large surface area, the ability to tune the surface chemistry and low costs.[9–11] Typically, the activated carbons are used, but a fundamental understanding of OPs adsorption on carbon materials requires well-defined adsorbates, allowing one to resolve crucial factors in OPs adsorption on carbon materials.

Considering the facts stated above, graphene and graphene-based materials are excellent candidates, as model adsorbents for OPs removal. We have previously shown that different graphene-based materials display different adsorption properties towards the removing of OPs *via* adsorption from aqueous solutions.[4] Depending on the type of OP (aromatic or aliphatic), the removal of OPs will be governed by the concentration of defects and surface functionalities in graphene-based materials and the existence of extended sp^2-hybridized domains in the graphene basal plane.

The majority of our knowledge about the adsorption of OPs on carbon materials comes from experimental studies. However, theoretical analysis can provide us with valuable pieces of information as the model systems can be controlled at a high level. However, theoretical studies are rather scarce.[12] We believe that this is caused by the complexity of the systems and large sizes of OP molecules, making the direct electronic structure calculations heavy and time-consuming.

Nevertheless, we emphasize that the theoretical modelling can help us understand general trends, and certain approximations introduced in the model could allow us to perform the analysis, which could be very difficult, if not impossible, in the laboratory. In this contribution, we studied OPs adsorption on four different graphene model surfaces. As we have previously found that the S=P and O=P moieties of OPs dominate the adsorption on carbon materials, we simplify the models of OPs as $S{=}PH_3$ and $O{=}PH_3$. To control the electron density at S=P and O=P moieties, we replaced hydrogen atoms with highly electronegative fluorine atoms and consider $S{=}PF_3$ and $O{=}PF_3$ model OPs as well. In most cases, we found that the OPs adsorption is physical, but certain model molecules could also get chemically decomposed on the defective graphene surfaces. Chemical trans-

formations were also observed experimentally using UPLC analysis, while the preferential oxidation of thiophosphates into their highly toxic oxo-analogs was excluded by the direct evaluation of neurotoxicity of OP solution after the treatment with graphene oxide samples.

EXPERIMENTAL

Theoretical calculations

DFT calculations were done using QuantumESPRESSO code[13] at the GGA level.[14] The ultrasoft pseudopotentials were used, and the plane-wave basis set was expanded to 600 eV. The dispersion interactions were accounted for using the DFT-D3 approach of Grimme.[15] The graphene surfaces were modelled in a 4×4 cell, originally containing 32 carbon atoms. The vacuum size was set to 20 Å to prevent the coupling of periodic images along the *z*-axis of the cell. First Brillouin zone was sampled using the 10×10×1 Monkhorst–Pack *k*-point set.[16] The relaxation of all the atoms in the cell was allowed until the forces acting on atoms were below 0.01 eV Å^{-1}. Adsorption energies (E_{ads}) were calculated as:

$$E_{ads} = E_0(\text{OP+graphene}) - E_0(\text{graphene}) - E_0(\text{OP}) \quad (1)$$

where E_0(OP+graphene), E_0(graphene), and E_0(OP) stand for the total energies of: *i*) the graphene surface with adsorbed OP, *ii*) the clean graphene surface and *iii*) the isolate OP model molecule, respectively.

Materials

For the present study, we chose two commercial graphene oxide (GO) materials and dimethoate (DMT) as a representative of OPs. This combination of OP and adsorbents was justified because GO is not a very good adsorbent for removing DMT but possesses many different surface functional groups and defects. Hence, the possible chemical conversion of DMT *via* reactions with the surface functional groups is easier to observe, when compared to the situation when extremely effective adsorbents are used. Namely, in that scenario, all DMT would be adsorbed, and it would not be possible to see the reaction products. We specifically used GRAPHENEA GO,[17] denoted hereafter as GNA GO and Graphene Supermarket GO[18] denoted further as SM GO. GNA GO has 49–56 % of carbon and 41–50 % oxygen, with small amounts of hydrogen, nitrogen and sulphur. It also has a high concentration of epoxy groups, as confirmed by the manufacturer using XPS. SM GO has a higher carbon-to-oxygen ratio (79 to 20 %).

Adsorption measurements

The investigated adsorbents were dispersed in double-distilled water, and the desired amount of DMT stock solution (Pestanal, Sigma Aldrich, Denmark) was added to provide the targeted concentration of the adsorbent and DMT. The vessel containing the adsorbent+DMT mixture was then placed on a laboratory shaker (Orbital Shaker-Incubator ES-20, Grant-Bio) and left for 21 h at 25 °C to ensure reaching the equilibrium. After the equilibrium was reached, the mixture was centrifuged for 10 min at 14500 rpm, and the supernatant was filtered through the nylon filter membrane. The concentration of DMT after adsorption (C_{eq}) was determined using UPLC analysis. The control experiments were performed identically, but without the carbon adsorbents, and confirmed that there is no DMT degradation within the timeframes of the described experiments. From the described batch adsorption measurements, we calculated uptake for the investigated DMT on all the studied adsorbents, varying the

adsorbent dose and for fixed DMT concentration (C_0 = 5×10^{-4} M). The removal was calculated as:

$$\text{Removal} = 100(C_0 - C_{eq})/C_0 = 100C_{ads}/C_0 \quad (2)$$

UPLC analysis

The Waters ACQUITY ultra performance liquid chromatography (UPLC) system coupled with a tunable UV detector controlled by the Empower software was used. The chromatographic separations were run on an ACQUITY UPLC™ BEH C_{18} column, with the dimensions 1.7 μm, 100 mm× 2.1 mm (Waters). The DMT analysis was done under isocratic conditions, with the mobile phase consisting of acetonitrile (phase A, 20 vol.%) and 20 vol.% acetonitrile in water (phase B, 80 vol.%). The eluent flow rate was 0.25 mL min^{-1}, and the injection volume was 10 μL. The optical detection was done using a photodiode array detector. Under the described conditions, the retention time of DMT was 3.02±0.05 min.

Neurotoxicity measurements

Regarding their (neuro)toxicity, the physiological effects of treated DMT solutions were tested using AChE inhibition measurements. AChE activity was assayed according to Ellman's procedure.[19] The *in vitro* experiments were performed by the exposure of 2.5 IU commercially purified AChE from electric eel to the OP solutions obtained in the adsorption experiments (batch adsorption or filtration) at 37 °C in 50 mM PB pH 8.0 (final volume 0.650 mL). The enzymatic reaction was started by the addition of acetylcholine-iodide (AChI) in combination with DTNB as a chromogenic reagent and allowed to proceed for 8 min until stopped by 10 % sodium dodecyl sulfate (SDS). The enzymatic reaction product, thiocholine, reacts with DTNB and forms 5-thio-2-nitrobenzoate, whose optical adsorption was measured at 412 nm. It should be noted that in these measurements, the enzyme concentration was constant and set to give an optimal spectrophotometric signal. Physiological effects were quantified as AChE inhibition given as:

$$\text{AChE inhibition} = 100\frac{A_0 - A}{A_0} \quad (3)$$

where A_0 and A stand for the AChE activity in the absence of OP and the one measured after the exposure to a given OP.

RESULTS AND DISCUSSION

Model OPs

All the studied (O)SPH$_3$(F$_3$) have C_{3v} symmetry. When hydrogen atoms are replaced with the electron-withdrawing fluorine atoms, the O=P and S=P bonds were contracted. O=P bond, the length is reduced from 1.49 to 1.46 Å, while for the S=P bond, the length reduces from 1.94 to 1.88 Å. These results demonstrate that the idea to control the electron density at the O=P and S=P moieties by changing hydrogen with fluorine was correct.

Model graphene-based surfaces

To investigate the effects of different graphene-based surfaces on the adsorption processes of model OPs, we considered four different surfaces. These are pristine graphene surface, graphene surface with a single vacancy, graphene surface with Stone–Wales (SW) defect and graphene surface with one epoxy

group.[20] We do not describe these surfaces in detail, but here are provided their densities of states (DOS) and their optimized structures (Fig. 1).

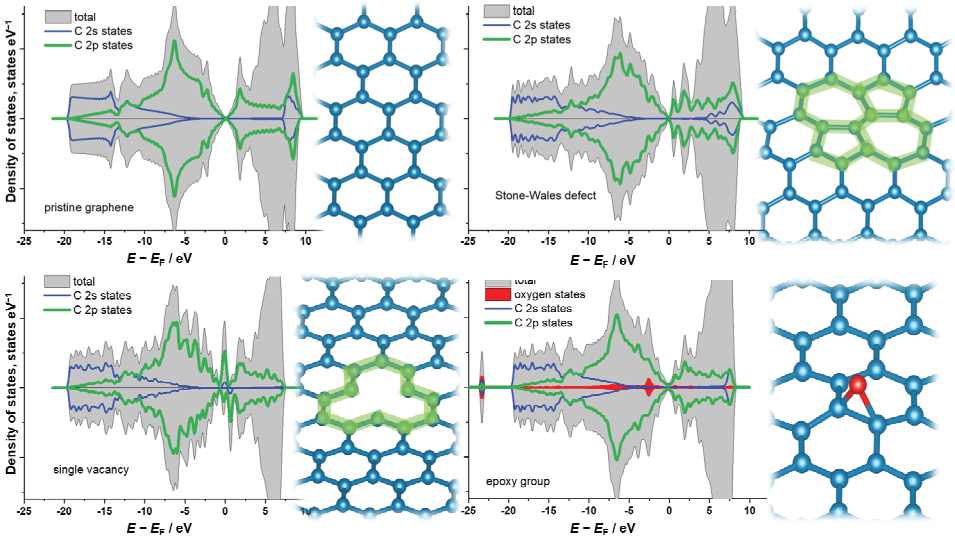


Fig. 1. Densities of states (DOS) and the optimized structures of considered model graphene surfaces. For graphene with the single vacancy and Stone–Wales (SW) defect, the defects are indicated in the corresponding structures.

The introduction of considered defects causes the large electronic structure disruptions, when compared to the electronic structure of pristine graphene. Graphene with a single vacancy remains flat, like pristine graphene. However, the dangling bonds are formed upon removing one C atom, and the corresponding states are located around the Fermi level.

There are pronounced DOS changes around the Fermi level for the SW defects, as the formation of the SW defect is preceded by the rehybridization of carbon atoms and the local bending. This local bending expands further to the entire basal plane for the studied concentration of SW defects. The addition of the epoxy group on the graphene basal plane causes sp^2 to sp^3 rehybridization of the carbon atoms bonded to oxygen. Hence, in this case, too, there is a local deformation of the basal plane. While one cannot expect that pristine graphene should show significant chemical reactivity, the introduction of defects can change the situation dramatically.

Interaction of model OPs with model graphene-based surfaces

The next step performed was the analysis of model OPs interactions with the described graphene-based surfaces. Molecules were soft-landed on the surfaces, and the systems were completely relaxed without any restrictions during the structural optimization. In line with our previous experimental findings,[9] we con-

cluded that the interaction is *via* S=P or O=P moiety. Moreover, the calculated adsorption energies indicated physisorption, when no chemical reaction between model OPs and the studied adsorbates occurs. The results are summarized in Table I.

TABLE I. Calculated adsorption energies (eV) for model OPs on four considered graphene-based surfaces. When "reaction" is stated, no structures with intact model OPs were observed upon the relaxation

Substrate	Surface			
	$S{=}PH_3$	$S{=}PF_3$	$O{=}PH_3$	$O{=}PF_3$
Pristine graphene	−0.18	−0.14	−0.16	−0.15
Graphene with Stone–Wales defect	−0.26	−0.19	−0.28	−0.19
Graphene with a single vacancy	Reaction	Reaction	−0.27	−0.22
Epoxy graphene	Reaction	Reaction	Reaction	Reaction

These few numbers presented in Table I tell a story of importance. As expected, the pristine graphene interacts weakly with model OPs. However, when the SW defect has been introduced, the interactions, which are still weak, become stronger. As for pristine graphene, the F-substituted molecules interact weaker with the SW graphene surface.

When the model OPs interacted with the graphene with the single vacancy, we observed that tio-forms always reacted with the graphene surface. However, for oxo-forms, we identified the cases where model OPs were intact after the adsorption. The adsorption energies calculated for these cases were similar to that in the case of SW graphene. This result showed that the oxo-forms were more stable than tio-forms, which is in accordance with our previous results about DMT oxidation and the product of the oxidation – omethoate, is more stable and more persistent to chemical modifications than its parental tio-form DMT. Finally, all the considered model OPs undergo chemical transformations when interacted with epoxy-group on graphene.

Physisorption of model OPs is clearly seen in DOS plots of studied systems where no model OPs react. Such a situation is shown for OPH_3 and OPF_3 on the SW graphene surface in Fig. 2.

The electronic states of OPH_3 and OPF_3 simply add up to the electronic states of the surface, which is almost unchanged when compared to the DOS of the clean surface (Fig. 1). Also, by comparing the DOSes presented in Fig. 2, one can see that the states of carbon atoms are almost identical for OPH_3 and OPF_3 adsorbates.

Except for the pristine graphene surface, some chemical reactions between the model OPs and the defects in the surfaces were observed in all the other cases, in addition to the physisorption events reported in Table I. This will not be described in many details, as it is not certain that all these reactions might not occur with a real OP molecule due to steric hindrances. However, the general

conclusions follow. For the epoxy graphene surfaces, the O atom from the epoxy function always reacts with model OPs. As a rule, this process is followed by incorporating O, S or P molecules into the graphene lattice, with the significant reordering of the hexagonal carbon network. For the SW graphene and the graphene with a single vacancy, these topological defects are reactive sites. The O, S and P atoms of the model OPs incorporate easily into the graphene lattice, and the result is the graphene doped with one of these heteroatoms or (S,P)-co-doped graphene. Such a scenario is presented in Fig. 3, where, upon the landing of SPF_3 on the SW graphene surface, both S and P are incorporated into the SW lattice site. The states of S and P get fully hybridized with the states of carbon atoms. As a side note, this seems like a rather effective way to dope the graphene lattice with S and P, which opens many possibilities for using such doped graphene.

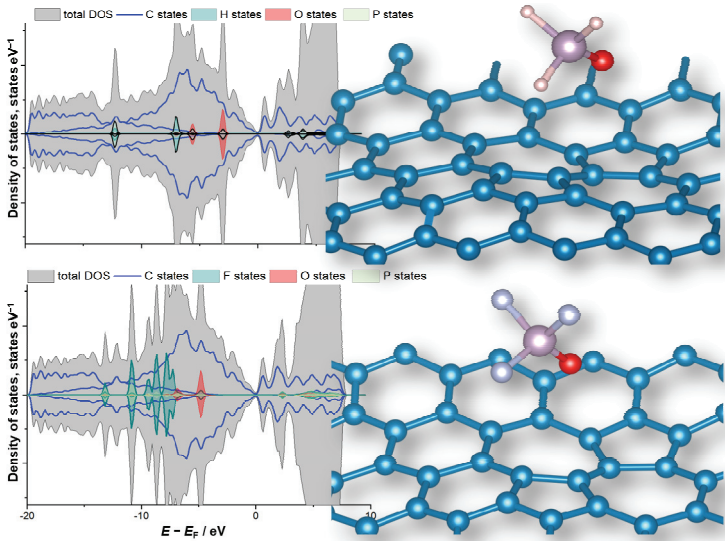


Fig. 2. Densities of states (DOS) and the optimized structures of model OPs interacting with the SW graphene surfaces.

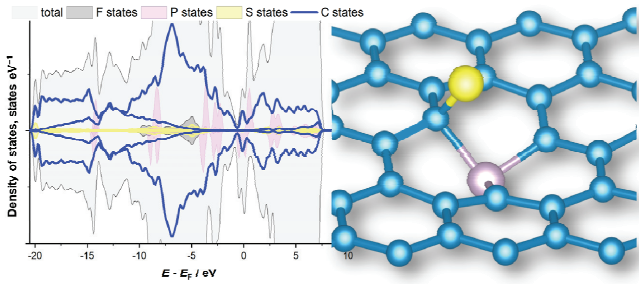


Fig. 3. Densities of states (DOS) and the optimized structure of SPF_3 upon the reaction with the SW graphene surface.

It is emphasized that the observed chemical reactions are spontaneous and, as a rule, disrupt the O=P or S=P moiety, which is directly responsible for the neurotoxicity of OPs.

Experimental results – UPLC analysis and neurotoxicity

To investigate the possible chemical reactions between OPs and graphene--based materials, we have performed a series of batch adsorption experiments using DMT and two commercial GO materials. The choice of GO as an adsorbent was explained previously. Upon the detailed analysis of UPLC chromatograms, we have identified one prominent signal that depends on the adsorbent mass (Fig. 4). This signal is located at a low retention time (RT = 0.877 min) and is well separated from the main DMT peak that appears at RT = 3.02 min. The signal at RT = 0.877 min is not present for the pure DMT solution and is not seen for the filtrate of pure GO adsorbents, which we have used as a background. The time evolution of this signal during the extended exposure to adsorbents is not very obvious, but it is visible. Moreover, there are slight variations in the retention time of the two chromatographic peaks. We suspect that the observed time variation can be understood as a balance between the chemical reaction giving the product with this specific signal in UPLC and its adsorption on GO, as well as the formation of more complex adducts.

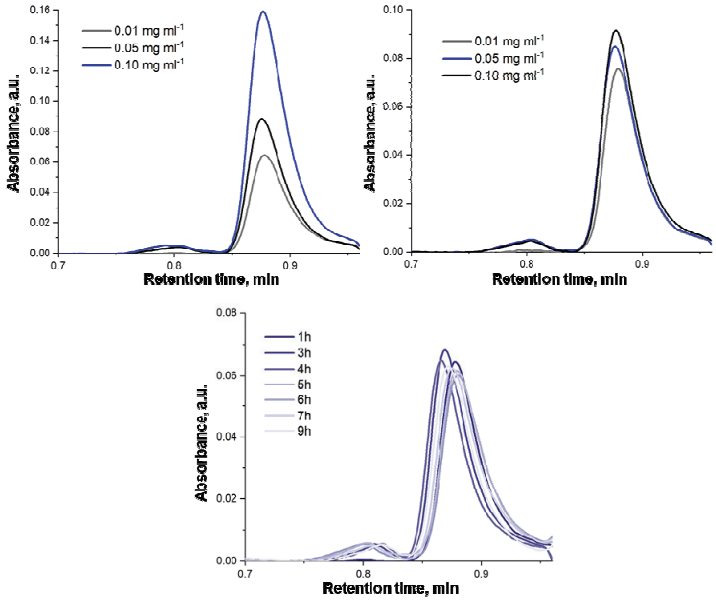


Fig. 4. UPLC chromatograms of DMT solution upon the batch adsorption experiment using SM GO (top left) and GNA GO (top right) for different masses of adsorbents. In the bottom row, the time evolution of the signal is presented for extended batch adsorption experiments using SM GO and DMT. The concentration of DMT was 5×10^{-4} mol dm^{-3}.

The observed signal is further analyzed, and the UV–Vis spectra were extracted (Fig. 5). This new signal shows broad adsorption, with the adsorption maximum at 199.2 nm. Contrary to that, a rather sharp absorption of DMT maximum is seen at 193.8 nm.

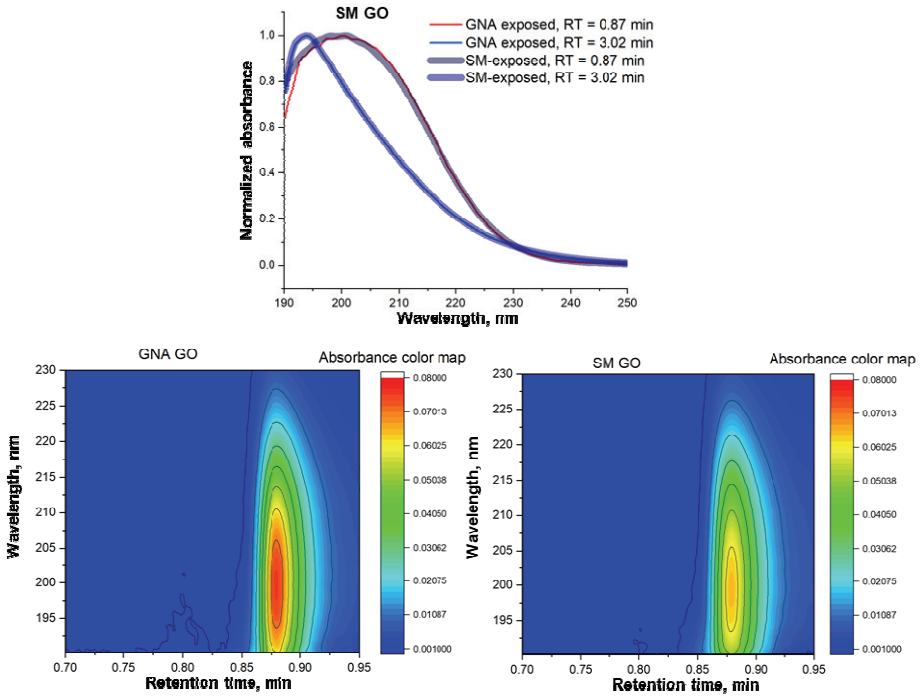


Fig. 5. Normalized UV spectra of the fraction corresponding to the retention time of 0.87 min and the UV spectra of DMT (*RT* = 3.02 min) upon the batch adsorption experiments using DMT ($C = 5\times10^{-4}$ mol dm^{-3}) and the concentration GO materials amounting to 0.01 mg ml^{-1} (upper row). In the bottom row, photodiode array signals (PDA) *versus* the retention time are provided for both GO used.

The only possible way to determine the exact mechanism of the DMT decomposition *via* GO substrates is the use of mass spectrometry coupled with UPLC. However, for the time being, this demanding task is left aside. Namely, we suspect that for other OPs one could expect different reaction mechanisms depending on the complexity of a given molecule. DMT is a rather simple aliphatic molecule, so deriving specific conclusions built up based on its case is risky. However, our main goal here was to communicate the chemical conversion of DMT with graphene oxide. Moreover, based on DFT calculation, we expected that the S=P moiety of OP molecules take part in chemical conversion and that this part of the molecule gets degraded. Hence, this reaction should reduce the

toxicity of the DMT solution after the treatment with a GO sample, being in favour for the remediation of OPs.

It is noted that the amount of chemically converted DMT is rather small, so the main removal mechanism could still be adsorption. The relatively low efficiency of GO materials as adsorbents for DMT is discernible from the data presented in Table II. The removal is characterized by equilibrium, with up to 10 % of DMT being adsorbed/reacted on GO materials. While the performed UPLC analysis showed no DMT conversion to its oxo-analog, we have performed an additional check using the AChE inhibition measurements. The reduction of AChE inhibition is a joint effect of adsorption (which is highly dominant) and chemical degradation of DMT caused by GO samples as adsorbents. In all the cases, the reduced AChE inhibition after the batch adsorption measurements was confirmed, so DMT-oxon (omethoate) formation can be securely excluded. Namely, if only a small fraction of DMT is converted to oxon, the AChE inhibition would increase as the oxon is approximately 1000 times more toxic than DMT.[9]

TABLE II. The calculated equilibrium concentration (C_{eq}) and DMT uptakes for three different DMT concentrations in the batch adsorption experiments using SM GO and GNA GO as adsorbent. AChE inhibition is evaluated for the filtrates after the experiments. The uncertainty of reported results is within 5 %

DMT concentration, mol dm^{-3}	Adsorbent	C_{eq} / 10^{-4} mol dm^{-3}	Removal, %	AChE inhibition, %
10^{-3}	None	10.0	–	46
	SM GO	9.52	4.81	28
	GNA GO	9.59	4.1	29
10^{-4}	None	1.00	/	12
	SM GO	0.920	8.0	7
	GNA GO	0.942	5.8	8
10^{-5}	None	0.100	–	6
	SM GO	0.0908	9.2	0
	GNA GO	0.0889	11.1	0

CONCLUSION

Based on the DFT calculations of the interactions of different graphene-based surfaces and model organophosphorus molecules, a chemical reaction between point defects in graphene basal plane (single vacancy, Stone–Wales defect and epoxy group) and the S(O)=P moieties of organophosphorus compounds is expected. Using UPLC analysis, this reaction is confirmed for two commercial graphene oxide samples and for DMT as a representative of organophosphorous pesticides. While DMT adsorption is still a dominant mechanism of the pesticide removal from aqueous solutions on studied GO materials, although very small, this reaction also contributes to the reduction of neurotoxicity of the graphene oxide-treated samples. The presented results could lead to the develop-

ment of novel systems for the irreversible conversion of organophosphates to non-toxic compounds.

Acknowledgement. Authors acknowledge the financial support provided by the Serbian Ministry of Education, Science and Technological Development through the institutional funding of "VINČA" Institute of Nuclear Sciences – National Institute of the Republic of Serbia and the University of Belgrade – Faculty of Physical Chemistry.

ИЗВОД

УКЛАЊАЊЕ ОРГАНОФОСФОРНИХ ПЕСТИЦИДА ИЗ ВОДЕ КОРИШЋЕЊЕМ МАТЕРИЈАЛА НА БАЗИ ГРАФЕНА – САМО АДСОРЦИЈА ИЛИ ЈОШ НЕШТО?

ВЛАДАН АНИЋИЈЕВИЋ[1], МАРКО ЈЕЛИЋ[2], АЛЕКСАНДАР З. ЈОВАНОВИЋ[3], НЕБОЈША ПОТКОЊАК[2], ИГОР А. ПАШТИ[3] и ТАМАРА Д. ЛАЗАРЕВИЋ-ПАШТИ[2]

[1]Војно техничкиинститут (ВТИ), Ратка Ресановића 1, 11132 Београд, [2]Институт за нуклеарне науке Винча – Национални институт од значаја за Републику Србију, Универзитет у Београду, Михајла Петровића Аласа 12–14, Београд и [3]Универзитет у Београду – Факултет за физичку хемију, Студентски трг 12–16, 11158 Београд

Широка употреба пестицида захтева иновативне приступе уклањању ових једињења из животне средине. Угљенични материјали се традиционално користе као адсорбенти за уклањање пестицида, а развој нових врста угљеничних материјала омогућава напредније приступе у њиховој примени у животној средини. Користећи прорачуне базиране на теорији функционала густине, предвиђена је хемијска реакција између S(O)=P група органосфосфата са тачкастим дефектима у графену – појединачним ваканцијама, Стоне–Вејлс (Stone–Wales) дефеката и епоксидних група. Реакција је потврђена хроматографијом ултра високих перформанси за два узорка графен-оксида и диметоата као представника органофосфорних пестицида. Тачан реакциони механизам је још увек неразјашњен, али је недвосмислено потврђено да се не јавља селективна оксидација диметоата у токсичнији оксо-аналог. Презентовани резултати могу помоћи у развоју нових система за иреверзибилну конверзију органофосфорних пестицида у нетоксична једињења без употребе агресивних хемијских агенаса, или спољних физичких фактора попут UV зрачења.

(Примљено 8. јануара, ревидирано 11. фебруара, прихваћено 12. фебруара 2021)

REFERENCES

1. J. E. Casida, G. B. Quistad, *Chem. Res. Toxicol.* **17** (2004) 983 (https://doi.org/10.1021/tx0499259)
2. M. B. Colovic, D. Z. Krstic, T. D. Lazarevic-Pasti, A. M. Bondzic, V. M. Vasic, *Curr. Neuropharmacol.* **11** (2013) 315 (https://doi.org/10.2174/1570159X11311030006)
3. T. Lazarevic-Pašti, A. Leskovac, T. Momić, S. Petrovic, V. Vasic, *Curr. Med. Chem.* **24** (2017) 3283 (https://doi.org/10.2174/0929867324666170705123509)
4. T. Lazarević-Pašti, V. Anićijević, M. Baljozović, D. Vasić Anićijević, S. Gutić, V. Vasić, N. V. Skorodumova, I. A. Pašti, *Environ. Sci.: Nano* **5** (2018) 1482 (https://doi.org/10.1039/C8EN00171E)
5. T. Mitrović, S. Lazović, B. Nastasijević, I. A.Pašti, V. Vasić, T. Lazarević-Pašti, *J. Environ. Manage.* **246** (2019) 63 (https://doi.org/10.1016/j.jenvman.2019.05.143)
6. M. I. Badawy, M. Y. Ghaly, T. A. Gad-Allah, *Desalination* **194** (2006) 166 (https://doi.org/10.1016/j.desal.2005.09.027)

7. T. Matsushita, A. Morimoto, T. Kuriyama, E. Matsumoto, Y. Matsui, N. Shirasaki, T. Kondo, H. Takanashi, T. Kameya, *Water Res*. **138** (2018) 67 (https://doi.org/10.1016/j.watres.2018.01.028)
8. Y. Samet, L. Agengui, R. Abdelhedi, *J. Electroanal. Chem.* **650** (2010) 152 (https://doi.org/10.1016/j.jelechem.2010.08.008)
9. T. D. Lazarević-Pašti, I. A. Pašti, B. Jokić, B. M. Babić, V. M. Vasić, *RSC Adv.* **6** (2016) 62128 (https://doi.org/10.1039/C6RA06736K)
10. M. Valickova, J. Derco, K. Simovicova, *Acta Chim. Slovaca* **6** (2013) 25 (https://doi.org/10.2478/acs-2013-0005)
11. M. Vukčević, A. Kalijadis, B. Babić, Z. Laušević, M. Laušević, *J. Serb. Chem. Soc.* **78** (2013) 1617 (https://doi.org/10.2298/JSC131227006V)
12. S. M. Maliyekkal, T. S. Sreeprasad, D. Krishnan, S. Kouser, A. Kumar Mishra, U. V. Waghmare, T. Pradeep, *Small* **9** (2013) 273 (https://doi.org/10.1002/smll.201201125)
13. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin--Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, R. M. Wentzcovitch, *J. Phys. Condens. Matter* **21** (2009) 395502 (https://doi.org/10.1088/0953-8984/21/39/395502)
14. J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **77** (1996) 3865 (https://doi.org/10.1103/physrevlett.77.3865)
15. S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **132** (2010) 154104 (https://doi.org/10.1063/1.3382344)
16. H. J. Monkhorst, J. D. Pack, *Phys. Rev., B* **13** (1976) 5188 (https://doi.org/10.1103/PhysRevB.13.5188)
17. https://cdn.shopify.com/s/files/1/0191/2296/files/Graphenea_GO_4mgmL_Datasheet_201905.pdf?22 (accessed on January 3, 2021)
18. http://graphene-supermarket.com/images/P/Graphene%20Oxide-Data-Sheet-Graphene-Supermarket.pdf (accessed on January 3, 2021)
19. G. L. Ellman, K. D. Courtney, V. Andres jr., R. M. Featherstone, *Biochem. Pharmacol.* **7** (1961) 88 (https://doi.org/10.1016/0006-2952(61)90145-9)
20. G. Yang, L. Li, W. Bun Lee, M. Cheung Ng, *Engin. Struct. Mater.* **19** (2018) 613 (https://doi.org/10.1080/14686996.2018.1494493).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. *86* (7–8) 711–724 (2021)
JSCS–5456
Original scientific paper

Voltammetric quantification of the anesthetic drug propofol (2,6-diisopropylphenol) in pharmaceutical formulations on a boron-doped diamond electrode

ERTUĞRUL KESKİN[1]*, SHABNAM ALLAHVERDIYEVA[2], HANDE İZEM ÖZOK[3], ORUÇ YUNUSOĞLU[4] and YAVUZ YARDIM[3]**

[1]*Adıyaman University, Faculty of Pharmacy, Department of Analytical Chemistry, 02040 Adıyaman, Turkey,* [2]*Van Yüzüncü Yıl University, Faculty of Science, Department of Biochemistry, 65080 Van, Turkey,* [3]*Van Yüzüncü Yıl University, Faculty of Pharmacy, Department of Analytical Chemistry, 65080 Van, Turkey and* [4]*Van Yüzüncü Yıl University, Faculty of Medicine, Department of Pharmacology, 65080 Van, Turkey*

(Received 19 October 2020, revised 7 February, accepted 4 March 2021)

Abstract: In this paper, the detailed electrochemistry of propofol (PRO), which is one of the intravenous agents commonly used for sedative-hypnotic purposes, was examined. In cyclic voltammetry, the agent showed one irreversible and diffusion-controlled oxidation peak, resulting in the formation of a couple with a reduction and re-oxidation wave at less positive potentials. The effect of electrode pretreatment procedures on the electrochemical response of PRO was investigated using square wave voltammetry (SWV) and the optimum procedure was used to improve the signal response in subsequent studies. Quantification of PRO was realized based on the first oxidation peak using SWV. After optimization of all variables, the linear working range of PRO was found to be between 2.5 μg mL^{-1} (1.4×10^{-5} mol L^{-1}) and 160.0 μg mL^{-1} (1.1×10^{-3} mol L^{-1}, n = 15) with a detection limit 0.71 μg mL^{-1} (3.9×10^{-6} mol L^{-1}). No noteworthy interference effect was detected. Furthermore, the developed method was used for quantification of PRO in pharmaceutical samples.

Keywords: propofol; boron-doped diamond electrode; voltammetry; pharmaceutical formulation.

INTRODUCTION

The use of anesthetics agents for creating reversible unconsciousness was a huge breakthrough in medical procedures. Among these agents, intravenous ones are often used for sedative-hypnotic purposes in surgical operations due to their safety and rapid effectiveness. The propofol (PRO, 2,6-diisopropylphenol) agent

*,** Corresponding authors. E-mail: (*)keskinertugrul@gmail.com, (**)yavuzyardim2002@yahoo.com
https://doi.org/10.2298/JSC201019017K

in this group is one of the most preferred anesthetic agents in the western world due to advantages such as low drug interactions and relative benign side effects[1] (Fig. 1). Because of its high lipophilic character, it penetrates and distributes on the central nervous system, quickly. This allows more effective control over the onset of sedation and the thereafter recovery processes.

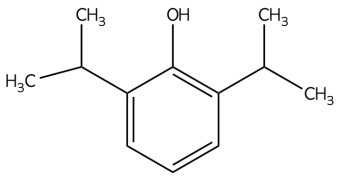


Fig. 1. Chemical structure of propofol.

The sedative and hypnotic activity of PRO is related to its concentration in plasma. So, the therapeutic concentration in plasma should be in the range of 0.5-1.5 μg mL^{-1} for sedation and 2-5 μg mL^{-1} for anesthesia.[2] On the other hand, it is a well-known fact that PRO is used for drug abuse as well as anesthetic purposes.[3] As a result of its misuse, death cases have been reported in different parts of the world, especially among healthcare professionals, such as doctors and nurses.[4] Therefore, a simple, fast, and sensitive analytical approach is required for the determination of PRO in pharmaceutical formulations. To date, many studies have been conducted for the determination of PRO in biological fluids, such as blood and urine. These studies included the use of various analytical methodologies, such as spectrophotometric,[5] high-performance liquid chromatography (HPLC),[6–9] gas chromatography–mass spectrometry (GC–MS),[10,11] liquid chromatography–mass spectrometry (LC–MS)[12–19] and capillary electrophoresis.[20] However, there are a limited number of studies in the scientific literature on the electrochemical analysis of PRO. In one of the non-quantitative studies, the researchers investigated the performance of differently produced pencil graphite electrodes (PGE) against electrode fouling in PRO monitoring.[21] In another study of the same research group, electrode cleaning procedures were developed to eliminate the fouling effect caused by propofol oxidation on rotating disk boron-doped diamond (BDD)/silica substrate material and PGE. Due to the low-adsorption capability of the BDD electrode, they used for quantitative analysis of PRO only at the PGE electrode following these cleaning procedures. After optimizing the experimental parameters, they reached a detection limit of 0.82 μM on PGE.[22] In an other electrochemical study, the performance of a bare GC electrode as substrate material was evaluated by using different electrochemical techniques. In this study, the obtained sensitivity levels based on both the reduction and oxidation peak of PRO were found as 5.5 and 3.2 μM, respectively.[23] In addition, Thiagarajan *et al.* reached a sensitivity of 0.08 μM thanks to anodized screen-printed carbon electrode (SPCE) formed as a result of pretreat-

ing the SPCE in the H_2SO_4 solution.[24] In addition, it the study in the literature reporting the electrochemical behavior of PRO in organic film-coated glassy carbon (GC) electrode should be highlighted.[25] However, this procedure has the difficulty of forming a uniform organic film on the electrode.

Electrochemical methods differ from other analytical methods in their speed, ease of application, reliability and simplicity. In addition, they offer considerable sensitivity and selectivity depending on the type of working electrode and the technique used.[26] Undoubtedly, another outstanding feature is that they clarify the redox behavior of the related species.[27] It is a known fact that the type of electrode used in the analysis performed by these methods, the pretreatment procedures, and the modification steps applied to the electrode play great roles in the electrochemical behavior of the analyte.[28–30]

A boron-doped diamond (BDD) electrode is a carbonaceous electrode material that offers distinctive features, such as wide working window (both in the cathodic and anodic direction), low ground current, low signal noise ratio, operation in corrosive environments and giving reproducible responses to electrochemical analysis.[31–34] Moreover, electroactive species that cannot be examined with conventional metal and other carbon electrodes can be analyzed with the wide working window of this electrode material.[35] However, a BDD electrode allows limited modification with modifying agents compared to electrode materials such as glassy carbon. The main reason for this is the low adsorption of both the analyte and the modifying agent to the surface of the BDD electrode material.[36]

As for the electrochemical pretreatment step, it is one of the most widely used techniques to increase the sensitivity of the BDD electrode. This procedure applies a certain potential to the BDD electrode in the anodic or cathodic direction for a certain period of time. These pretreatment steps applied to the electrode could be performed anodically or cathodically, anodically followed by cathodically or *vice versa*. Many electrochemical approaches based on BDD electrodes have increased the electrode sensitivity using such steps.[37–39]

In a recent paper, the performance of a BDD electrode together with PGE has been investigated in the effectiveness of electrode cleaning strategies in long-term PRO monitoring.[22] However, the quantitative analytical aspect of this study is based on PGE. As far as is known, no approach has been found in the literature for quantitative purposes in PRO analysis using the BDD electrode. In this article, a simple approach to the subject was implemented for quantitative analysis of PRO with just a simple pretreatment step on an unmodified BDD electrode. Its practical applicability was applied to commercial pharmaceutical dosages.

EXPERIMENTAL

Chemicals

PRO standard (Reagent Plus®, ≥99.82 %, liquid) was purchased from ChemScene LLC (USA) and used as received. A stock standard solution of 1 mg mL^{-1} of PRO was prepared by dissolving in methanol and storing in a volumetric flask in a refrigerator at 4–6 °C in order to avoid degradation when not in use. Analytical-grade reagents and purified water from a Millipore Milli-Q system (Millipore, resistivity ≥18.2 MΩ cm) were used for the preparation of Britton–Robinson (BR, 0.1 mol L^{-1}, pH 2–9), phosphate buffer (0.1 mol L^{-1}, pH 2.5 and 7.4), acetate buffer (0.1 mol L^{-1}, pH 4.7), $HClO_4$ (0.1 mol L^{-1}) and H_2SO_4 (0.1 mol L^{-1}) solutions. The working and calibration solutions of PRO were prepared from the stock solution before use with appropriate diluting with the selected supporting electrolyte. All analysis and voltammetric measurements were realized under laboratory conditions.

Apparatus and measurements

Electrochemical measurements were performed with an μAutolab Type III (Metrohm Autolab B.V., Utrecht, The Netherlands) controlled with GPES software (version 4.9). All square wave voltammograms were smoothed using a Savicky and Golay algorithm and baseline-corrected by the moving average algorithm filtering technique (peak width of 0.01 V). A conventional three-electrode setup in the electrochemical cell was used, consisting of a BDD working electrode (diameter of 3 mm, boron doping level of 1000 ppm, Windsor Scientific Ltd., Slough, UK), an Ag/AgCl/3 mol L^{-1} NaCl reference electrode (BAS, Model RE-1, USA) and a Pt counter electrode (BAS, MW-4130, USA). All potentials mentioned in the paper are refered to Ag/AgCl/3 mol L^{-1} NaCl reference electrode unless othervise stated. The pH was measured at 25 °C using a model WTW inoLab720 pH meter equipped with a combined glass electrode (Xylem, New York, USA).

At the beginning of each experiment day, an anodic pretreatment procedure was performed by applying a potential of 1.8 V to the BDD electrode in 0.5 mol L^{-1} H_2SO_4 solution for 180 s. After this process, a potential of –1.8 V was applied to the BDD electrode for the same period and in the same solution. As a result of these steps, the BDD electrode surface acquired a oxygen and hydrogen-terminated property.

First, the cyclic voltammetry (CV) technique was used to elucidate the electrochemical behavior of PRO and determine the reaction kinetics on the BDD electrode in the selected supporting electrolyte. Then, optimizations of the experimental parameters, such as supporting electrolyte and square wave voltammetric parameters, were conducted in order to improve the selectivity and sensitivity of the PRO analysis.

The quantitative analysis of PRO with the square wave voltammetry (SWV) technique was first started by immersing three electrodes into voltammetric cells containing PRO in the $HClO_4$ solution. After this step, anodic scanning was performed from 0 to 1.6 V. The quantitative analysis of PRO was carried out using an optimized 50 Hz frequency, 40 mV pulse amplitude, and 12 mV step potential values of the SWV technique.

Preparation of samples

The pharmaceutical formulations (Propofol-®Lipuro, labeled 10 mg mL^{-1} per ampoule) were obtained from a local hospital. A defined volume (10 μL) of this ampoule solution was transferred to 10 mL electrochemical cells containing 0.1 mol of L^{-1} $HClO_4$ solution. The PRO content in pharmaceutical formulation was determined by adding standard PRO concentrations to the drug sample.

RESULTS AND DISCUSSION

Electrochemical behavior of PRO on BDD electrode

The electrochemical behavior of PRO on the BDD electrode surface was examined by CV method. Three consecutive CVs for 100 μg mL^{-1} PRO between –0.7 and 1.7 V potentials were recorded in 0.1 mol L^{-1} $HClO_4$ solution at a potential scanning rate of 100 mV s^{-1}. As can be seen in Fig. 2A, upon first scanning in the oxidative direction, PRO exhibited one well-defined anodic peak, P_A, with a maximum at about 1.28 V. In the reverse scan, a reduction peak, P'_C, was obtained at about –0.27 V, which was accompanied by the appearance of a small additional anodic peak, P'_A, at a less positive potential than the main oxidation peak, P_A, during the second and subsequent scans. It should be noted that the peaks P'_C/P'_A did not appear in the CV curves if the potential was started at 0.0 V and ended at 1.7 V. As the number of scans increased, the peak height of P_A decreased, while the intensity of P'_C/P'_A increased very slightly. From these observations, it could be infered that P_A was an irreversible and P'_C/P'_A were a pair of redox peaks, which could be attributed to the products formed in the main electrooxidation step.

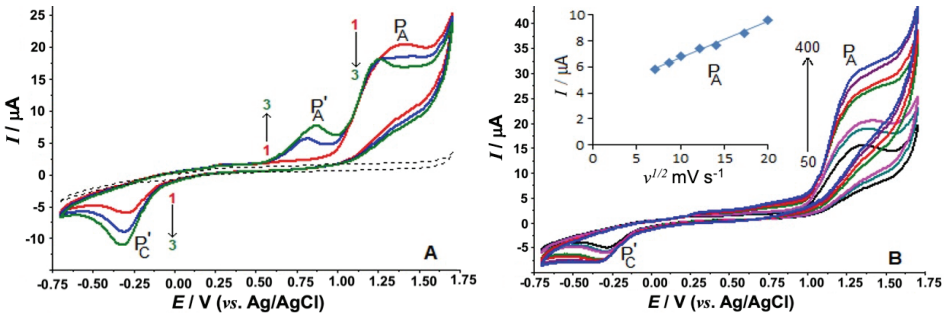


Fig. 2. The repetitive cyclic voltammograms at scan rate of 100 mV s^{-1} (A), and the cyclic voltammograms at different scan rates, 50, 75, 100, 150, 200, 300 and 400 mV s^{-1} (B), of 100 μg mL^{-1} propofol in 0.1 mol L^{-1} $HClO_4$ solution at the BDD; A: dashed lines represent background current; B: inset depicts the plot of peak current *vs.* square root of scan rate.

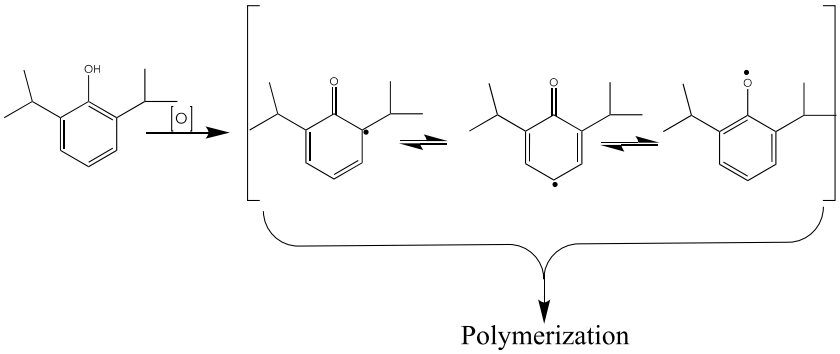
To determine the kinetics of the BDD electrode, the effect of the potential scan rates (v) on the anodic peak current response for 100 μg mL^{-1} PRO was examined by the CV technique 0.1 mol L^{-1} $HClO_4$ solution at 50–400 mV s^{-1} scan rates (Fig. 2B). The relationship between the square root scan rate ($v^{1/2}$) and anodic peak current (i_{pA}) of PRO was found to be linear according to the equation, $i_{pA} = (0.279\pm0.011)v^{1/2} + (3.92\pm0.188)$, $r = 0.994$. A similar linear correlation was found between log i_{pA} and log v, as shown in Eq. (1):

$$\log i_{pA} = (0.230\pm0.09)\log v + (0.370\pm0.019),\ r = 0.992 \quad (1)$$

These results could be evaluated as clear evidence that the kinetics of a BDD electrode are mainly diffusion controlled in the oxidation of PRO.

In order to ascertain the electron number (n) involved in the PRO oxidation process at a BDD electrode, the n value was determined from the CV voltammograms using the equation $\alpha n = 47.7/(E\text{p}–E\text{p}/2)$. In this study, the value of $E_\text{p}–E_\text{p}/2$ was 151 mV, thus the value of αn was calculated as 0.32. Generally, α (charge transfer coefficient) is assumed to be 0.5 in a totally irreversible electrode process. Hence, the n value was found to be 0.64 (≈1).

Although it is difficult to predict the general oxidation mechanism of PRO occurring on a BDD electrode surface, it is known that the mechanism that causes fouling on the electrode surface occurs through radical intermediate species polymerized similarly to the electrochemical oxidation of other phenolic compounds[21] and remain adsorbed at the electrode surface. Scheme 1 shows that electrochemical oxidation of PRO on the BDD electrode occurs first by the formation of phenoxy radicals and then by polymer formation.



Scheme 1. Possible redox mechanism of PRO at the BDD electrode in 0.1 mol L^{-1} $HClO_4$ solution.

Influence of the electrode pretreatment procedure

Before the voltammetric analysis of PRO, preliminary studies showed that the unpretreated BDD electrode could not be effective against passivation problems, especially at high PRO concentrations. This undesired situation caused the unpretreated BDD electrode to produce unsatisfactory results in terms of sensitivity and reproducibility. To overcome this and optimize the pretreatment procedures on the BDD electrode, the effectiveness of three different pretreatment procedures were investigated for 30 μg mL^{-1} PRO in $HClO_4$ solution by SWV. First, the performance of an anodic pretreated BDD electrode in voltammetric studies was examined. Secondly, the performance of a cathodic pretreated BDD electrode was investigated. Finally, anodic and cathodic pretreatment steps were applied sequentially to the BDD electrode and their performances in voltam-

metric analysis were recorded (Fig. 3). More sensitive and reproducible results were obtained in the analysis of PRO when the final pretreatment procedure was applied to the BDD electrode. For this reason, this procedure was chosen as the optimum pretreatment procedure and applied to the electrode at the beginning of each experiment day.

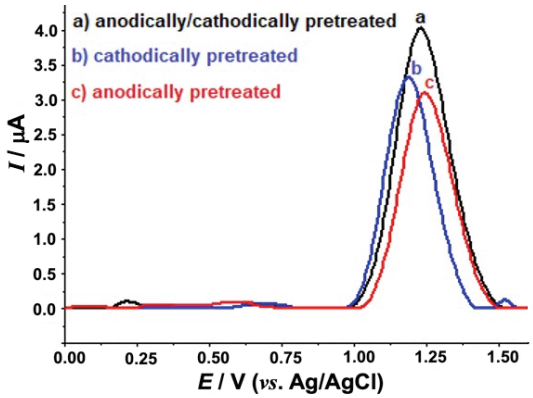


Fig. 3. SW voltammograms of 30 μg mL^{-1} propofol in 0.1 mol L^{-1} $HClO_4$ solution on BDD electrode after different electrochemical pretreatments. SWV parameters: frequency, 50 Hz; step potential, 8 mV; pulse amplitude, 30 mV.

Effect of pH

The effect of pH on oxidation peak currents and potentials of PRO was investigated by SWV on the BDD electrode using different supporting electrolytes at various pH values in order to obtain the best voltammetric response for analytical purposes., The baseline corrected SW voltammograms are depicted in Fig. 4A within the pH range 2.0–9.0 in BR buffer (0.1 mol L^{-1}) on 30 μg mL^{-1} PRO solution, within the potential range from 0 to 1.6 V.

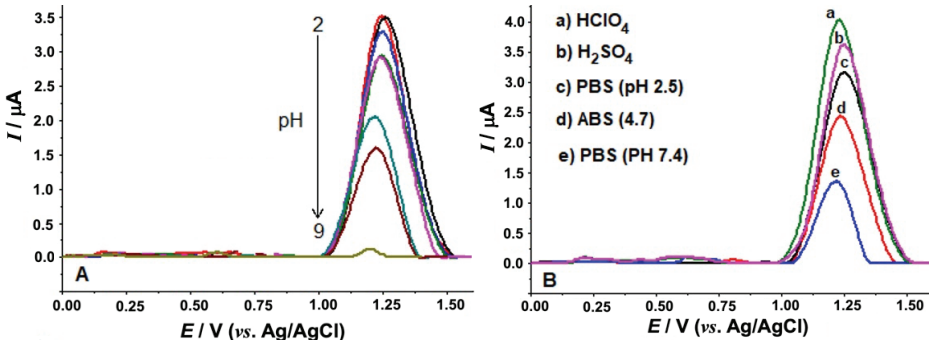


Fig. 4. SW voltammograms of 30 μg mL^{-1} propofol in BR buffer pH 2.0–9.0 (A), and in various supporting electrolytes at different pH values (B) on a BDD electrode. Other operating conditions were as indicated in Fig. 3.

Furthermore, it could be seen in Fig. 4A that PRO has one oxidation peak within the pH range 2.0–9.0 studied in the working potential range. In addition, it is clearly seen that with increasing solution pH, the potential of the oxidation peak becomes nearly pH-independent (from 1.25 to 1.22 V), indicating that no proton transfer steps occur before the rate-determining electron transfer step at these pH values. The SW voltammograms in various supporting electrolytes are depicted in Fig. 4B. Using 0.1 mol L^{-1} $HClO_4$, H_2SO_4, phosphate buffer (pH 2.5 and 7.4) and acetate buffer (pH 4.7), the oxidation peak potentials of PRO became nearly pH independent.

The evolution of peak currents with pH shows that this parameter reached the highest values in 0.1 mol L^{-1} $HClO_4$. Thus, ongoing studies will be performed in this solution.

Effect of SWV parameters

Another important factor affecting the sensitivity of PRO is optimization of the pulse parameters, such as frequency (f), pulse amplitude (ΔE_{sw}), and step potential (ΔE_s). This optimization step was realized by changing one of the pulse parameters while keeping the other two parameters constant and by recording the obtained signal. First, the f variable was examined in the range of 25–125 Hz, while ΔE_{sw} and ΔE_{ss} parameters were kept constant at 30 and 8 mV, respectively. So, the best sensitivity and peak shape for this pulse variable was recorded at 50 Hz. Next, the ΔE_{sw} value was changed between 30 and 60 mV, while the ΔE_{ss} and f values were kept constant at 8 mV and 50 Hz, respectively. The same optimization process was carried out by keeping f and ΔE_{sw} constant and examining the values of ΔE_{ss} between 6 and 12 mV. As a result, the best SWV instrumental parameters on the BDD electrode for 7.5 μg mL^{-1} PRO in 0.1 mol L^{-1} $HClO_4$ solution (not shown) were obtained at values f, 50 Hz; ΔE_s, 10 mV; ΔE_{sw}, 40 mV.

Analytical performance evaluation

The applicability of the developed SWV method based of the BDD electrode was tested under the optimized experimental and instrumental conditions. The increase in the anodic currents obtained as a function of the added PRO standards in the range from 2.5 (1.4×10^{-5}) to 160.0 μg mL^{-1} (1.1×10^{-3} mol L^{-1}) in 0.1 mol L^{-1} $HClO_4$ solution is shown in Fig. 5. The highly linear relationship between the SW voltammograms procured in response to these successively added PRO concentrations is given in Eq. (2):

$$i_p = (0.173\pm0.07\ C + (0.045\pm0.002)\ (r = 0.999, n = 15) \quad (2)$$

where i_p is the peak current, C the concentration, r the correlation coefficient, and n the number of experiments.

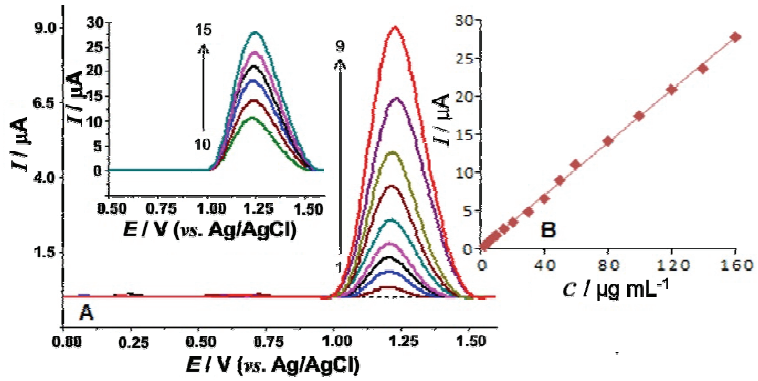


Fig. 5. SW voltammograms for propofol levels of (1–15) 2.5, 5.0, 7.5, 10, 15, 20, 30, 40, 50, 60, 80, 100, 120, 140 and 160 μg mL^{-1} in 0.1 mol L^{-1} $HClO_4$ solution (A); B: the corresponding calibration plot for the quantitation of propofol. SWV parameters: frequency, 50 Hz; step potential, 10 mV; pulse amplitude, 40 mV.

The developed methodology reached 0.71 μg mL^{-1} (3.98×10^{-6} mol L^{-1}) *LOD*, and 2.36 μg mL^{-1} (1.32×10^{-5} mol L^{-1}) *LOQ* levels with a simple pretreatment step applied to the electrode.

The *LOD* value was calculated according to the 3*s*/*m* formula where *s* and *m* are the standard deviation of ten consecutive measurements of the lowest concentration in the calibration range and the slope of the corresponding calibration curve, respectively. This *LOD* value proves that the sensitivity of the method is good enough and could be applied to real samples.

As for the comparison of the method with other techniques, Table I gives the performance of the BDD electrode with a limited number of other carbon-based electrochemical studies in the literature. The presented methods in Table I reached similar sensitivity levels except for the SPCE-based electrode.

TABLE I. Comparison of published electroanalytical methods for PRO detection; working electrodes: BDD, boron-doped diamond electrode; PGE, pencil graphite electrode; GCE, glassy carbon electrode; SPCE, screen printed carbon electrode; PVC/GCE, PVC membrane coated glassy carbon electrode; techniques**:** CV, cyclic voltammetry; DPV, differential pulse voltammetry; CSV, cathodic stripping voltammetry; CA, chronoamperometry; SWV, square-wave voltammetry

WE	Technique	E_p / V	Linear range, μM	*LOD* / μM	Analyzed sample	Reference
BDD, PGE	CV, DPV	0.4, 0.7	-	2.38-8.1	Serum	[22]
GCE	CV, CSV	0.70	0-30	5.5	–	[23]
SPCE	CV	0.56	0.09-0.9	0.08	Lab. samples	[24]
PVC/GCE	CA	-	0-56.6	0.03	–	[25]
BDD	SWV	1.25	14-1100	3.9	Pharm. formulation	This work

Furthermore, the proposed methodology has the level of sensitivity that could sense the level of PRO in plasma (between 1–60 μM) and the speed

required to analyze PRO in forensic cases. It is also very advantageous in terms of the stability of the electrode response in comparison with other electrodes. Moreover, another aspect that distinguishes the proposed method from the few electrochemical methods performed is that the PRO determination is realized at high positive potentials.

To test the precision the method, at a concentration of 2.5 μg mL^{-1} PRO, the intra-day (ten replicate) and inter-days (three days) reproducibility was examined under the same conditions and the relative standard deviation (*RSD*) values were calculated as 8.03 and 9.13 %, respectively. These values indicated that the BDD electrode produces sufficiently reproducible results in PRO measurements.

Effect of interfering compounds

Prior to the analyses of real samples, the selectivity of the proposed SWV protocol was also investigated in the presence of some species, such as lactose, sucrose, fructose, glucose, ascorbic acid, dopamine, uric acid and ions such as Ti^{4+}, Fe^{3+}, Zn^{2+}, Mg^{2+}, Ca^{2+}, K^{+}, Na^{+}, NO_3^{-}, Cl^{-}, SO_4^{2-} and some agents present in pharmaceutical formulations, such as starch, talc, cornstarch and magnesium stearate. The interfering effect on the PRO signal was examined at 1:1, 1:10 and 1:100 (PRO: interference species) molar ratios in supporting electrolyte containing 2.5 μg mL^{-1} PRO. The signal obtained from the mixture of PRO and the interfering species was compared with the signals obtained from the solution containing only PRO. The maximum concentration of a foreign substance that affected the PRO signal less than 7 % was described as the tolerance limit. Metal ions had no significant effect on the quantitative determination of PRO. This behavior could be explained by the emergence of metal ion signals in negative potentials but PRO signals in positive potential. The effects of interfering species such as starch, talc, cornstarch and magnesium stearate on the oxidation signal of PRO showed a negligible effect, even at over 100 times concentration. Species such as lactose, sucrose, fructose and glucose did not show a significant interfering effect for the quantitative determination of PRO even at 100 times the PRO concentration. However, it was not possible to analyze PRO together with dopamine and uric acid when their concentration was 10 times the concentration of PRO or ascorbic acid at concentrations 20 times higher than that of PRO. In order to analyze these compounds together, either a pre-separation method, such as chromatography or a chemometric method should be used. To conclude, the proposed voltammetric protocol is sufficiently selective and can perform PRO analysis in complex matrices.

Analytical application

The accuracy and practical usability of the proposed electroanalytical methodology was tested in commercially pharmaceutical formulations by using the multiple standard addition method. The sample preparation procedures are

described in the relevant section in detail. At this stage, it should be stressed that the sample was used directly without any dilution, pre-separation, filtering, or evaporation steps. Different volumes of standard PRO stock solution were transferred to the electrochemical cell containing the drug sample. As a result of these successive transfers, the final concentrations of PRO in the cell were adjusted to be 5, 7.5, 10, 15, 20, 30, 40 and 50 μg mL^{-1}. The overlapping voltammograms obtained from the drug sample after each added concentration are shown in Fig. 6. The appearing peak at approximately 1.23 V belonging to PRO oxidation increased linearly with increasing concentration of the added standards. Each Propofol-®Lipuro, ampoule was calculated to contain an average of 9.52 mg mL^{-1} PRO (*RSD* of 5.05 %) when evaluated based on these successive additions. This result also complies with the manufacturer's declaration claiming to contain 10 mg mL^{-1} per ampoule. As for testing the validity of the proposed method, this stage was assessed by the calibration curve and recovery method. The recovery values obtained by the SWV procedure are presented in Table II. These results proved that PRO analyzes can be performed safely without significant interference from commercial drug samples.

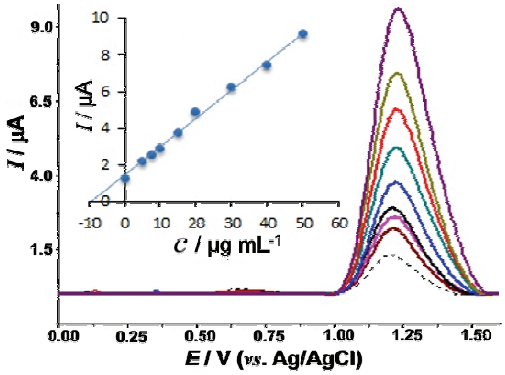


Fig. 6. SW voltammograms of the diluted drug sample (dashed line) and after standard additions of 5.0, 7.5, 10, 15, 20, 30, 40 and 50 μg mL^{-1} propofol (1–8) in 0.1 mol L^{-1} $HClO_4$ solution. Inset depicts the result of analysis by the standard addition method for the oxidation peak. Other operating conditions as indicated in Fig. 5.

TABLE II. Recovery values of pharmaceutical formulations samples spiked with standard solutions of propofol using the developed voltammetric method

c / μg mL^{-1}			Recovery ± *RSD*, %
Added[a]	Expected[a]	Found[a,b]	
0	–	9.52	–
5.0	14.52	15.20	104.7±6.07
7.5	17.02	16.37	96.2±5.43
10.0	19.52	18.64	95.5±5.13

[a]Concentration in the measured solution; [b]average of three replicate measurements

CONCLUSIONS

Within the scope of this article, the useability of a bare BDD electrode for the quantitative analysis of PRO was demonstrated. The electrochemical

behavior and electrode kinetics of PRO were revealed based on this electrode. As a result of the pretreatment procedures on the electrode, the change of PRO signals was examined and the optimum procedure was determined. To obtain the best voltammetric response, chemical variables, such as supporting electrolyte and pH, and puls variables, such as frequency, pulse amplitude, and step potential, were also optimized. Linearity was achieved between 2.5 and 160.0 μg mL^{-1} (1.4×10^{-5}–1.1×10^{-3} mol L^{-1}) under these optimized conditions while the *LOD* was calculated as 0.71 μg mL^{-1} (3.9×10^{-6} mol L^{-1}). The application of the method has been successfully demonstrated in the commercial drug formulation. The methodology proposed here is a good alternative to existing methods with its practical use, speed, and sensitivity.

И З В О Д

ВОЛТАММЕТРИЈСКА КВАНТИФИКАЦИЈА АНЕСТЕТИКА ПРОПОФОЛА (2,6-ДИИЗОПРОПИЛФЕНОЛ) У ФАРМАЦЕУТСКИМ ОБЛИЦИМА ПРИМЕНОМ БОРОМ ДОПИРАНЕ ДИЈАМАНТСКЕ ЕЛЕКТРОДЕ

ERTUĞRUL KESKİN[1], SHABNAM ALLAHVERDIYEVA[1], HANDE İZEM ÖZOK[3], ORUÇ YUNUSOĞLU[4] и YAVUZ YARDIM[3]

[1]*Adıyaman University, Faculty of Pharmacy, Department of Analytical Chemistry, 02040 Adıyaman, Turkey,* [2]*Van Yüzüncü Yıl University, Faculty of Science, Department of Biochemistry, 65080 Van, Turkey,* [3]*Van Yüzüncü Yıl University, Faculty of Pharmacy, Department of Analytical Chemistry, 65080 Van, Turkey и* [4]*Van Yüzüncü Yıl University, Faculty of Medicine, Department of Pharmacology, 65080 Van, Turkey*

У овом раду је испитано електрохемијско понашање пропофола (PRO), једног од често коришћених интравенозних агенаса који делује седативно-хипнотички. У цикличној волтаметрији, агенс показује један иреверзибилан и дифузионо контролисан оксидациони пик у директном, као и један редукциони пик у повратном скену. Даљом циклизацијом, до мање позитивног потенцијала, јавља се нови пик као последица ре-оксидације редуковане форме лека. Испитиван је утицај процедуре припреме електроде на одговор PRO добијен применом волтаметрије правоугаоних таласа (SWV), са циљем постизања оптималне процедуре ради побољшања сигнала одговора у наредним испитивањима. Квантификација PRO је изведена на основу првог оксидационог пика применом SWV. После оптимизације одређених параметара, добијен је линеаран радни опсег за PRO Између 2,5 μg mL^{-1} ($1,4\times10^{-5}$ mol L^{-1}) и 160,0 μg mL^{-1} ($1,1\times10^{-3}$ mol L^{-1}, $n = 15$), са границом детекције од 0,71 μg mL^{-1} ($3,9\times10^{-6}$ mol L^{-1}). Нису детектоване значајније интерференције. Осим тога, развијена метода је примењена за квантификацију PRO у фармацеутским узорцима.

(Примљено 19. октобра 2020, ревидирано 7 фебруара, прихваћено 4. марта 2021)

REFERENCES

1. K. Ode, *Anaesth. Intensive Care Med.* **20** (2019) 118 (http://dx.doi.org/10.1016/j.mpaic.2018.12.008)
2. X. S. Fan, F. X. Di, X. M. Feng, C. C. Li, C. Y. Bi, J. Li, J. Y. Yin, Y. C. Han, *Chinese J. Anal. Chem.* **48** (2020) e20056 (http://dx.doi.org/10.1016/S1872-2040(20)60011-1)
3. A. Maas, C. Maier, S. Iwersen-Bergmann, B. Madea, C. Hess, *J. Pharm. Biomed. Anal.* **146** (2017) 236 (http://dx.doi.org/10.1016/j.jpba.2017.08.035)

4. N. Ji Kwon, H. J. Kim, S. Cho, M. A. Lee, E. Han, *Forensic Sci. Int.* **306** (2020) 110070 (http://dx.doi.org/10.1016/j.forsciint.2019.110070)
5. I. Šrámková, C. G. Amorim, H. Sklenářová, M. C. B. M. Montenegro, B. Horstkotte, A. N. Araújo, P. Solich, *Talanta* **118** (2014) 104 (http://dx.doi.org/10.1016/j.talanta.2013.09.059)
6. M. H. Yeganeh, I. Ramzan, *J. Chromatogr., B* **691** (1997) 478 (http://dx.doi.org/10.1016/S0378-4347(96)00469-0)
7. H. Zhang, P. Wang, M. G. Bartlett, J. T. Stewart, *J. Pharm. Biomed. Anal.* **16** (1998) 1241 (http://dx.doi.org/10.1016/S0731-7085(97)00262-8)
8. C. A. J. Knibbe, V. S. Koster, V. H. M. Deneer, R. M. Stuurman, P. F. M. Kuks, R. Lange, *J. Chromatogr., B* **706** (1998) 305 (http://dx.doi.org/10.1016/S0378-4347(97)00571-9)
9. X. Cussonneau, E. De Smet, K. Lantsoght, J. P. Salvi, M. Bolon-Larger, R. Boulieu, *J. Pharm. Biomed. Anal.* **44** (2007) 680 (http://dx.doi.org/10.1016/j.jpba.2006.10.020)
10. F. Maurer, M. Geiger, T. Volk, D. I. Sessler, S. Kreuer, *J. Pharm. Biomed. Anal.* **143** (2017) 116 (http://dx.doi.org/10.1016/j.jpba.2017.05.042)
11. M. Y. M. Peeters, H. Kuiper, B. Greijdanus, J. van der Naalt, C. A. J. Knibbe, D. R. A. Uges, *J. Chromatogr., B* **852** (2007) 635 (http://dx.doi.org/10.1016/j.jchromb.2007.01.001)
12. L. Bajpai, M. Varshney, C. N. Seubert, D. M. Dennis, *J. Chromatogr., B* **810** (2004) 291 (http://dx.doi.org/10.1016/j.jchromb.2004.08.023)
13. F. Beaudry, S. A. Guénette, A. Winterborn, J. F. Marier, P. Vachon, *J. Pharm. Biomed. Anal.* **39** (2005) 411 (http://dx.doi.org/10.1016/j.jpba.2005.04.041)
14. S. Cohen, F. Lhuillier, Y. Mouloua, B. Vignal, P. Favetta, J. Guitton, *J. Chromatogr., B* **854** (2007) 165 (http://dx.doi.org/10.1016/j.jchromb.2007.04.021)
15. H. S. Kim, J. C. Cheong, J. Il Lee, M. K. In, *J. Pharm. Biomed. Anal.* **85** (2013) 33 (http://dx.doi.org/10.1016/j.jpba.2013.06.027)
16. L. K. Sørensen, J. B. Hasselstrøm, *J. Pharm. Biomed. Anal.* **109** (2015) 158 (http://dx.doi.org/10.1016/j.jpba.2015.02.035)
17. J. H. Kwak, H. K. Kim, S. Choe, S. In, J. S. Pyo, *J. Chromatogr., B* **1015**–**1016** (2016) 209 (http://dx.doi.org/10.1016/j.jchromb.2016.01.061)
18. A. Khedr, S. S. A. El-Hay, A. K. Kammoun, *J. Pharm. Biomed. Anal.* **134** (2017) 195 (http://dx.doi.org/10.1016/j.jpba.2016.11.051)
19. F. Maurer, T. Shopova, B. Wolf, D. Kiefer, T. Hüppe, T. Volk, D. I. Sessler, S. Kreuer, *J. Pharm. Biomed. Anal.* **150** (2018) 341 (http://dx.doi.org/10.1016/j.jpba.2017.12.043)
20. Y. Hui, K. Raedschelders, H. Zhang, D. M. Ansley, D. D. Y. Chen, *J. Chromatogr., B* **877** (2009) 703 (http://dx.doi.org/10.1016/j.jchromb.2009.01.030)
21. F. Stradolini, T. Kilic, A. Di Consiglio, M. Ozsoz, G. De Micheli, S. Carrara, *Electroanalysis* **30** (2018) 1363 (http://dx.doi.org/10.1002/elan.201700834)
22. F. Stradolini, T. Kilic, I. Taurino, G. De Micheli, S. Carrara, *Sensors Actuators, B* **269** (2018) 304 (http://dx.doi.org/10.1016/j.snb.2018.04.082)
23. J. Langmaier, F. Garay, F. Kivlehan, E. Chaum, E. Lindner, *Anal. Chim. Acta* **704** (2011) 63 (http://dx.doi.org/10.1016/j.aca.2011.08.003)
24. S. Thiagarajan, C. Y. Cheng, S. M. Chen, T. H. Tsai, *J. Solid State Electrochem.* **15** (2011) 781 (http://dx.doi.org/10.1007/s10008-010-1160-3)
25. F. Kivlehan, F. Garay, J. Guo, E. Chaum, E. Lindner, *Anal. Chem.* **84** (2012) 7670 (http://dx.doi.org/10.1021/ac3006878)
26. O. I. Lipskikh, E. I. Korotkova, Y. P. Khristunova, J. Barek, B. Kratochvil, *Electrochim. Acta* **260** (2018) 974 (http://dx.doi.org/10.1016/j.electacta.2017.12.027)

27. J. Xu, Y. Wang, S. Hu, *Microchim. Acta* (2016) (http://dx.doi.org/10.1007/s00604-016-2007-0)
28. M. Hanko, Ľ. Švorc, A. Planková, P. Mikuš, *J. Electroanal. Chem.* **840** (2019) 295 (http://dx.doi.org/10.1016/j.jelechem.2019.03.067)
29. Ľ. Švorc, K. Kalcher, *Sensors Actuators, B* **194** (2014) 332 (http://dx.doi.org/10.1016/j.snb.2013.12.104)
30. S. Allahverdiyeva, P. Talay Pınar, E. Keskin, O. Yunusoğlu, Y. Yardım, Z. Şentürk, *Sensors Actuators, B* **303** (2020) 127174 (http://dx.doi.org/10.1016/j.snb.2019.127174)
31. Ľ. Švorc, K. Borovská, K. Cinková, D. M. Stanković, A. Planková, *Electrochim. Acta* **251** (2017) 621 (http://dx.doi.org/10.1016/j.electacta.2017.08.077)
32. O. Sarakhman, L. Dubenska, Ľ. Švorc, *J. Electroanal. Chem.* **858** (2020) (http://dx.doi.org/10.1016/j.jelechem.2019.113759)
33. E. Keskin, S. Allahverdiyeva, E. Şeyho, Y. Yardim, *J. Serbian Chem. Soc.* **85** (2020) 923 (http://dx.doi.org/10.2298/JSC190906138K)
34. E. Keskin, Y. Yardim, A. Levent, Z. Şentürk, *Rev. Roum. Chim.* **64** (2019) 1063 (http://dx.doi.org/10.33224/rrch.2019.64.12.06)
35. D. F. Pereira, E. R. Santana, J. V. Piovesan, A. Spinelli, *Diam. Relat. Mater.* **105** (2020) 107793 (http://dx.doi.org/10.1016/j.diamond.2020.107793)
36. P. Samiec, Ľ. Švorc, D. M. Stanković, M. Vojs, M. Marton, Z. Navrátilová, *Sensors Actuators, B* **245** (2017) 963 (http://dx.doi.org/10.1016/j.snb.2017.02.023)
37. Ľ. Švorc, J. Sochr, M. Rievaj, P. Tomčík, D. Bustin, *Bioelectrochemistry* **88** (2012) 36 (http://dx.doi.org/10.1016/j.bioelechem.2012.04.004)
38. R. Trouillon, Y. Einaga, M. A. M. Gijs, *Electrochem. Commun.* **47** (2014) 92 (http://dx.doi.org/10.1016/j.elecom.2014.07.028)
39. P. T. Pinar, H. S. Ali, A. A. Abdullah, Y. Yardim, Z. Şentürk, *Marmara Pharm. J.* **22** (2018) 460.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. *86* (7–8) 725–738 (2021)
JSCS–5457
Original scientific paper

Response surface methodology for the study of interactions between components in a micellar system formulation

SOUHILA OMARI[1], MOHAMED NEDJHIOUI[1], NADJIA HAMIDI[1], OTHMANE BENKORTBI[2]* and NABIL BOUARRA[3]

[1]*Materials and Environment Laboratory (LME), Department of Chemical Engineering and Environment,Faculty of Technology, Yahia Feres University, 26000 Medea, Algeria,* [2]*Biomaterials and Transport Phenomena Laboratory (LBMPT), Department of Chemical Engineering and Environment, Faculty of Technology, Yahia Feres University, 26000 Medea, Algeria and* [3]*Centre de Recherche Scientifique et Technique en Analyses Physico–Chimiques, BP 384, Zone Industrielle Bou-Ismail, 42004 Tipaza, Algeria*

(Received 16 February, revised 28 April, accepted 29 April 2021)

Abstract: This work was aimed at the examination of the interaction of certain physicochemical properties on micellar systems constituting of a polymer (sodium alginate), two surfactants (CTAB and tween 80), and Algerian olive oil. Response surface modelling (RSM) was applied to study the combined effects of systems containing each type of surfactant. The monitoring of four independent parameters, *i.e.*, the interfacial tension (Y_1), the conductivity (Y_2), the viscosity (Y_3) and the turbidity (Y_4) as responses for the experimental design model, allowed the determination of the performance of the established models. Based on statistical analyzes, the coefficients R^2 and Q^2 for the interfacial tension, the conductivity, the viscosity and the turbidity are: 0.998 and 0.805; 0.982 and 0.742; 0.976 and 0.734, and 0.985 and 0.723, respectively. The obtained results indicate that these models showed a good predictive power for an optimal system composed of CTAB, Tween 80, AlgNa, and olive oil. For the CTAB/AlgNa and CTAB/olive oil systems, interfacial tension values of 33.85 and 34.39 mN m^{-1}, respectively, and maximum conductivity values of 4.126 and 4.064 mScm^{-1}, respectively, were obtained. For viscous compounds consisting of AlgNa/Olive Oil and AlgNa/Tween 80, maximum viscosity values of 202.5 and 196.6 mPa s, respectively, were obtained. For the same systems as those for viscosity, turbidity values of 300 and 304 NTU, respectively, were obtained.

Keywords: polymer; modelling; interfacial tension; conductivity; viscosity; turbidity.

* Corresponding author. E-mail: benkortbi_oth@yahoo.fr
https://doi.org/10.2298/JSC210216033O

INTRODUCTION

In recent years, the combined effects between a polymer and a surfactant have aroused great interest due to their multiple applications. In the petroleum industry, and more particularly, in improving the recovery of petroleum, the droplets of oils trapped in porous media could be displaced under the combined effect of viscous and interfacial forces. It is therefore important to have a system allowing very low interfacial tensions to be reached and to maintain high viscosities to be able to transport the oil present in the pores.[1,2] Solutions containing polymers and surfactants can give rise to molecular interactions liable to affect their rheological and physicochemical properties.[3] In addition, these interactions have characteristics that depend on the electrical charges of the polymer, the surfactant, the hydrophobicity, the conformation, the flexibility of the polymer, and the presence of additives.[4,5] Most studies in this area have focused on complexes of anionic surfactants with polymers.[6–8]

The interactions between the polymer and the surfactant could be investigated in two ways: first, the polymer is considered to be the substance influenced by the surfactant adsorbed on the sites of the polymer that disrupt the formation of the micelles. Second, the surfactant is considered to be the substance influenced by the polymer. In this case, the association of surfactant molecules with macromolecules facilitates the phenomenon of micellization.[9,10] The evolution of the physicochemical and rheological properties as a function of the chemical nature and the concentrations of components allows the establishment of relations between these factors and the responses of the system, such as conductivity, turbidity, and critical micellar concentration.

The use of the experiment plan method allows predictive models of the studied responses to be obtained as well as the optimal conditions with the minimum number of tests and a maximum of credibility. Given the assigned objective (determination of the effects of three constituents), the most adequate experimental planning strategy is based on a response surface modelling (RSM) by a two-order model taking into account all the interactions between the factors. In the industrial field, the use of the experimental design is constantly evolving and can be used to support the optimization of manufacturing and control processes, as well as the formulation of products. The method consists in gathering a set of statistical techniques intended to analyze the behaviour of an experimental system to understand and improve its functioning.[11]

The experimental plans allow the organization of the tests to be optimized, which makes it possible to obtain both a maximum of information with the minimum of number of experiments and with the best possible precision on the modelling of the results. This method is based on strict mathematical rules and requires a rigorous approach on the part of the experimenter.[12]

Several authors have used one factor at a time, which not only consumes time and increases costs, but also neglects the effect of the interaction between the factors. Although the traditional orthogonal method is able to consider several factors at once, it cannot give a function expression between the factors and the response values.

Response surface modeling (RSM) is a statistical method that uses quantitative data from appropriate experiments to determine several regression equations between factors and experimental results. The main advantage of this method compared to other methods of experimental statistical design is the reduced number of experimental tests necessary to evaluate several parameters and their interactions.[13]

This study aims to assess the effects of four factors, *i.e.*, a cationic surfactant, CTAB, non-ionic surfactant, Tween 80, Algerian olive oil and sodium alginate as a polymer on the physicochemical properties. Properties of aqueous solutions, such as turbidity, conductivity, viscosity, and turbidity, were used as responses for the experimental design model through a response surface methodology, namely D-optimal design.

EXPERIMENTAL

Experimental design

A D-optimal design was used to study the main effects of four independent factors such as the concentration of CTAB (X_1), the concentration of Tween 80 (X_2), olive oil (X_3) and the concentration of alginate sodium (X_4). The D-optimal criterion was developed to select the design points to minimize the variance associated with the estimates of the specified model coefficients.[14] This method is useful more than the central composite design method, which requires fewer experiments to be carried out, and it can also address the categorical factors included in the design of the experiments.[15,16] The high D value plans were constructed from the data by a computer algorithm. The variables were subsequently coded according to Eq. (1):

$$X_i = \frac{U_i - U_i^0}{\Delta U_i} \quad (1)$$

where X_i is the independent variable coded value; U_i independent variable: real value; U_i^0 independent variable: real value on the centre point; and ΔU_i is the step change value (Table S-I of the Supplementary material to this paper).

Materials and methods

The cationic surfactant (CTAB) was purchased from Fluka (Switzerland), the non-anionic surfactant (Tween 80), and the sodium alginate polymer (AlgNa) were supplied from Panreac Chimica (Spain). The sodium alginate used in this study was in the form of a white to creamy white powder; it is odorless, tasteless, and very soluble in water. The density of Algerian olive oil was 0.92 g cm^{-3}, it was mainly composed of 99 % fat divided into a fat-soluble fraction (triglycerides and 1 % fatty acids) and non-fat-soluble (secondary components).[17-19] The interfacial tensions and the critical micellar concentrations of the mixtures were measured at a room temperature of 22 °C with a Du Noüy tensiometer (model 70545). The turbidimetry was measured using a turbidimeter (turbo 550 IR, reference 600110) to determine the degree

of turbidity of a suspension in solution. The conductivity of the solutions was determined using a Hanna EC 214 type conductivity meter with a cell constant of 0.475 cm^{-1}. The Haak RVT5 type viscometer with a rotating mobile was used to determine the viscosity of Newtonian liquids. This device is equipped with six mobiles of different shapes and geometries, where each mobile is used within a well-defined range of viscosity and sheer speed. Polymer dispersions were prepared by dissolving the polymer in water with moderate stirring at room temperature. After 24 h, different amounts of surfactant and oil were added to the polymer solutions. The surfactant was dissolved under slow mixing of 30 rpm in a propeller mixer (Heidolph RZR 2020, Germany). Depending on the case, the surfactant concentrations were chosen to be equal to, greater than or less than the critical micellar concentration (*CMC*) of the surfactant. However, the polymer concentrations were chosen to give variations in the viscosimetric and turbidimetric properties of the solution.[20-22]

Establishment of the experimental matrix

The D-optimal matrix type experience was used. It responds to the error minimization strategy in the estimation of the coefficients and the overall error. The matrix contains 16 tests in different zones, to minimize the error and to estimate the standard deviation of the natural values. Table I shows the matrix of experiments based on this strategy, in which the factors (X_i) and the responses (Y_i) are defined as follows. X_1: mass concentration of CTAB, which varied between 0.01and 0.2 %; X_2: mass concentration of Tween 80, which varied between 0.01 and 0.04 %; X_3: mass concentration of sodium alginate, which varied between 0.3 and 0.8 %; X_4: mass concentration of the olive oil range of 0.1 and 0.3 %; Y_1: interfacial tension, mN m^{-1}; Y_2: conductivity, mS cm^{-1}; Y_3: viscosity, mPa s; Y_4: turbidity, NTU.

RESULTS AND DISCUSSION

Statistical analysis

First, a first-order experimental plan was created:

$$Y_i = b_0 + \sum_{i=1}^{3} b_i X_i \tag{2}$$

The results obtained were analyzed using first-order linear models. This design was rejected due to a mismatch between predicted and experimental values. The lack of fit is significant and the model error was significantly larger than the pure error (reproducibility). Using multiple regression analysis, the responses were correlated with the three design factors *via* a second-order polynomial:

$$Y_i = b_0 + \sum_{i=1}^{3} b_i X_i + \sum_{i=1}^{3} b_{ii} X_i^2 + \sum_{i=1}^{3} \sum_{j=2}^{3} b_{ij} X_i X_j \tag{3}$$

where b_0, b_i, b_{ii} and b_{ij} are constant regression coefficients of the model, while X_i, X_j are the independent variables. The statistical significance of the regression coefficients was determined by Fisher's analysis of variance, *F*-test and the proportion of variance explained by the model obtained was given by the multiple coefficients of determination, R^2.

Establishment of responses according to formulation factors

The values obtained from Q^2 and R^2 show that the established model could be predictive (Table S-II of the Supplementary material). The coefficients of the polynomial model were determined by the multilinear regression method (MLR).

Table S-II shows the values of R^2 and Q^2, the stress, the interfacial tension, and the conductivity. The quality of the results obtained after adjustments was determined by the coefficient of determination R^2 and the coefficient of prediction Q^2.

A value close to 1 means that the predicted values are in very good agreement with the observed values and a Q^2 value greater than 0.7 indicates that the model has good predictive power. The arrangements for D-optimal experiments are listed in Table S-III, which includes 16 sets of experiments.

The quadratic regression model for interfacial tension (Y_1), conductivity (Y_2), viscosity (Y_3) and turbidity (Y_4) in terms of coded factors are given as follows:

$$Y_1 = 35.56 + 0.57X_1 + 0.15X_2 - 1.125X_3 - 1.55X_4 - 0.16X_1X_2 + \\ + 0.18X_1X_3 - 0.012X_1X_4 + 0.88X_2X_3 + 0.71X_2X_4 - 0.78X_3X_4 \quad (4)$$

$$Y_2 = 386.87 + 13.12X_1 + 13.75X_2 + 18.75X_3 + 15.62X_4 - 74X_1X_2 - \\ - 6.24X_1X_3 + 3.12X_1X_4 + 6.87X_2X_3 - 1.41X_2X_4 - 2.50X_3X_4 \quad (5)$$

$$Y_3 = 162.43 + 9.68X_1 + 15.31X_2 + 24.93X_3 + 24.93X_4 - 4.4X_1X_2 + \\ + 0.93X_1X_3 + 1.56X_1X_4 + 2.81X_2X_3 - 0.06X_3X_4 \quad (6)$$

$$Y_4 = 271 + 15.25X_1 + 18.25X_2 + 26.87X_3 + 0.75X_4 - 3.12X_1X_2 - \\ - 0.62X_1X_3 + 2.3X_1X_4 - 2.87X_2X_3 + 14.75X_3X_4 \quad (7)$$

ANOVA analysis of the models used to estimate the interfacial tension, the conductivity, the viscosity, and the turbidity *versus* the concentrations of CTAB, Tween 80, olive oil and sodium alginate are represented in Table S-IV of the Supplementary material.

The statistical significance of the second-order model was assessed by the analysis of variance of the Fisher test (F-test), which revealed that this regression was statistically significant ($P < 0.001$) at the 95 % confidence level. For interfacial tension, the model presented the highest coefficient of determination ($R^2 = 0.998$) showing 99.8 % validity in the response.

For conductivity, the regression was less significant, it has a good coefficient of determination ($R^2 = 0.982$) showing 98.2 % of the validity of the response. The model used for the estimation of the viscosity showed the very strong significance of the model ($P = 0.001$) and presents a good coefficient of determination ($R^2 = 0.976$), indicating that only 2.4 % of the total variations were not explained by the model. For turbidity, the value of the adjusted coefficient of determination $R^2_{(adj)} = 0.976$ was also very high and indicates the great importance of the model. The values of Q^2 obtained were of the order of 0.723 for the turbidity and

0.805 for the interfacial tension, which makes it possible to conclude that the model has an acceptable predictive power.[11–16] The results obtained are given in Table S-IV.

Influence of factors on interfacial tension: IFT

The interaction of water-soluble polymers with anionic surfactants is generally controlled by surface tension and specific or equivalent conductivity measurements depending on the concentration of surfactant.[23] The IFT method is also used to explain the process of micellization of surfactant solutions, as well as the distribution of molecules in the presence of an additive, the surface activity, and the micelle formation of ionic surfactants in combination with a charged polymer and salt. The behavior in surface tension of a multi-component system can be obtained from classical thermodynamic relationships for interfacial properties. The formulation adopted is that due to Gibbs equation and is represented by:[24,25]

$$\mathrm{d}\gamma = -\sum \Gamma_i \mathrm{d}\mu_i \tag{8}$$

where γ is the surface tension or the interfacial tension; Γ_i, is the surface excess component, μ_i is the chemical potential of the component defined as follows:

$$\mu_i = \mu_i^{\ominus} + RT\ln a_i \tag{9}$$

where $\mu_i^{\ominus}$ is the standard chemical potential and a_i is the activity of the component i.

Eq. (10) was obtained assuming for dilute solution $a_i = C_i$ and the substitution of Eq. (9) in Eq. (8):

$$\mathrm{d}\gamma = -RT\sum \Gamma_i \,\mathrm{d}\ln C_i \tag{10}$$

In a mixed multicomponent system of constant composition, C can be written as:

$$C_1 = KC_2 = KC_3 = \ldots \tag{11}$$

Differentiation of the logarithm of C leads to:

$$\mathrm{dln}C_1 = \mathrm{dln}C_2 = \mathrm{dln}C_3 = \ldots \tag{12}$$

The Gibbs adsorption equation for a system containing three components becomes:

$$\mathrm{d}\gamma = -RT\,(\Gamma_{\mathrm{CTAB}} + \Gamma_{\mathrm{Tween80}} + \Gamma_{\mathrm{AlgNa}} + \Gamma_{\mathrm{olive\ oil}})\mathrm{dln}\ C_1 \tag{13}$$

For ionic compounds, complete dissociation of CTAB (CTAB = $CTA^+ + Br^-$, AlgNa = $Alg^- + Na^+$) was assumed, and the dissociation of Tween 80 and olive oil were negligited, hence:

$$\Gamma_{\mathrm{CTAB}} = \Gamma_{\mathrm{CTA}^+} + \Gamma_{\mathrm{Br}^-} \text{ and } \Gamma_{\mathrm{AlgNa}} = \Gamma_{\mathrm{AlgNa}} + \Gamma_{\mathrm{Alg}^-} \tag{14}$$

This hypothesis allows the adsorption process to be consider as being positive in nature. Thus, only the solute occupies the surface (the superficial excess of pure solvent (water)). The variation of Γ due to the variation in concentration of one of the components leads to the evaluation of the total excess:

$$\Gamma_{\text{tot}} = \Gamma_{\text{CTAB}} + \Gamma_{\text{Tween 80}} + \Gamma_{\text{AlgNa}} + \Gamma_{\text{olive oil}} \quad (15)$$

A sample of iso-response plots for IFT at different coded values of CTAB and AlgNa is shown in Fig. 1. Tween 80 and olive oil are kept at their zero levels: 0.25 % for the Tween 80 and 0.20 % for the olive oil. It could be observed that the IFT values decrease with increasing concentration of CTAB and increase with increasing concentration of sodium alginate.

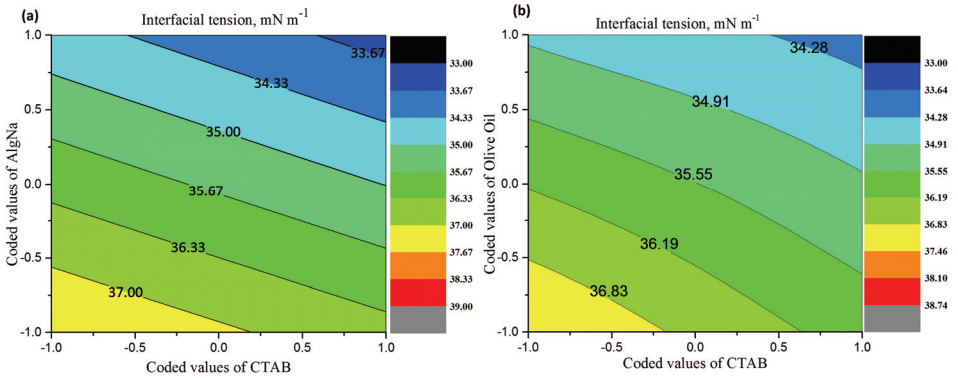


Fig. 1. Contour plots of independent variables: coded values of CTAB, AlgNa and olive oil on the interfacial tension.

Furthermore, the minimum of the IFT 33.85 mN m^{-1}, was obtained near the concentration of CTAB 0.8 g L^{-1} and for the concentration of AlgNa close to 0.125 in the coded values. The critical micellar concentration of CTAB decreased in the presence of AlgNa, and the number of aggregations of individual micelles increased with the polymer concentrations. Likewise, the addition of polymer to the CTAB solution allowed to decrease the concentration of critical aggregation (CAC) and to increase the number of sizes of the micellar aggregates which attach to the polymer coil. This indicates that the significant effect of the presence of a surfactant is to be expected in the solution with the polymer. On the other hand, an excess of sodium ions in solution should filter the electrostatic repulsions between the micellar aggregates attached to the polymer chain, thus reducing the degree of expansion of the latter.

Despite the electrolytic affinity of the dissolved AlgNa molecule, the presence of Na^+ does not affect an extension of the thickening behaviour of the solution.

Moreover, Fig. 1b shows the iso-responses of the surface tension at different coded values of CTAB and olive oil. The AlgNa and the Tween 80 were maintained at their zero levels: 0.55 and 0.025 %, respectively. The IFT values decreased with increasing concentrations of CTAB and olive oil (minimum IFT 34.39 mN m^{-1} was obtained for CTAB (0.5) and concentration of olive oil close to 0.75 in the coded values). The decrease of IFT could be explained in the same way as in Fig. 1a, at constant concentrations of AlgNa 0.55 % and Tween 0.025 %.

Influence of factors on conductivity

Conductivity measurements have been widely used to study the interactions between polymers and surfactants in an aqueous solution that are very important for the evaluation of electrostatic interactions in solutions when charged substances such as ionic surfactants, charged polymers, and electrolytes are used. This method was used by Goddard[26] to study the effect of salt on the interaction between the polymer (polyethylene oxide and SDS) and by Sovilj *et al.*[27] to study the influence of hydroxypropyl methylcellulose-SDS interactions.

The specific conductivity curves obtained from Eq. (5) are given in Fig. 2a and b. They represent the plots of isoresponses at different coded values of CTAB, Tween 80, AlgNa and olive oil. The obtained graphs show the effects of two factors while the other two factors are kept constant at their zero levels. Assuming that the total conductivity of the free ions is independent of any electrolyte present in the solution, the specific conductivity of each species at any calculated concentration is the sum of the conductivity of each ion present.Given this assumption, the specific conductivity of the solution containing total sodium σNa^+ is the sum of the conductivity of the charged polymer AlgNa, CTAB, Tween 80 and olive oil:

$$\sigma = \sigma_{CTA}{}^{+} + \sigma_{Br}{}^{-} + \sigma_{Tween\ 80} + \sigma_{Alg}{}^{-} + \sigma_{Na}{}^{+} + \sigma_{Oliveoil}{}^{+} \qquad (16)$$

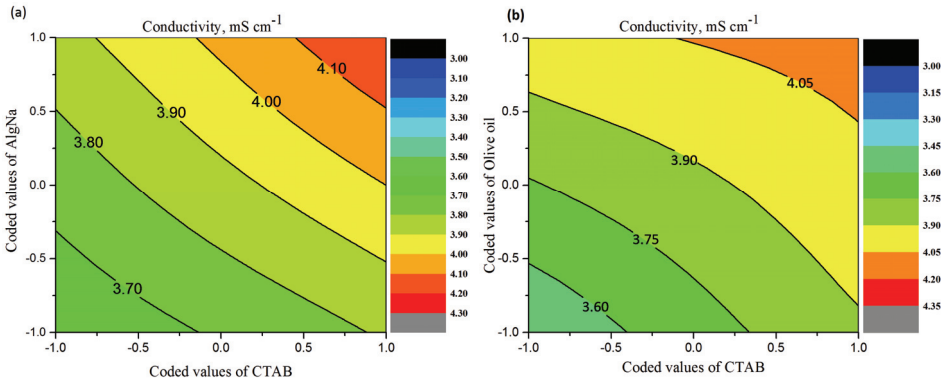


Fig. 2. Contour plots of independent variables: coded values of CTAB, AlgNa and olive oil on the conductivity.

In this context, only the total conductivity of the solution, σ, was obtained from the conductivity measurements. The effect of the factors on the conductivity is illustrated in Fig. 2. The curves obtained show the effects of the concentrations in coded values of CTAB and AlgNa on the conductivity when the concentrations of Tween 80 and olive oil are kept constant at their level zero. As expected, the presence of AlgNa (charged polymer) slightly increases the conductivity values with increasing CTAB at a constant concentration of Tween 80 (0.025 %) and olive oil (0.2 %).

By adding AlgNa to CTAB solutions at constant concentrations of Tween 80 and olive oil, the total conductivity of the solution strongly depends on the concentration of CTAB, as indicated by the model presented in Eq. (15). In this case, it could be supposed that complete dissociation of all the ionic species in solution occurred because Tween 80 and olive oil were at their constant concentrations.

The effects of the concentrations CTAB and olive oil, as coded values, on the conductivity when the concentrations of Tween 80 of (0.025 %) and AlgNa (0.55 %) were kept constant at their zero levels are shown in Fig. 2a and b. From these figures, it could be noted that the presence of olive oil with CTAB increased the conductivity in a less pronounced manner compared to the previous case. The conductivity varies from 3552 mS cm^{-1} for low concentrations to 4064 mS cm^{-1} for values close to the maximum concentrations of CTAB and olive oil. By adding olive oil to CTAB solutions at constant concentrations of Tween 80 and AlgNa, the total conductivity of the solution strongly depended on the charged surfactant (CTAB) and the charged polymer (AlgNa). As in the previous case, complete dissociation of all the ionic species in the solution was assumed because Tween 80 and the polymer were at their constant concentrations.

Effect of factors on viscosity

The behavior of the viscosity was mainly due to the presence of polymers and viscosimetric compounds, such as olive oil. Viscosity measurements are a practical way to study the hydrodynamic volume in the solution. The viscosity values of the solutions lower and higher than the concentrations of critical aggregations were determined and presented in the form of values of apparent viscosity. The surface response curves obtained from Eq. (6) for the viscosity at different coded values of CTAB, Tween 80, AlgNa, and olive oil are presented in Fig. 3a and b.

The viscosity values increased with increasing concentrations of AlgNa and Tween 80, as seen in Fig. 3a. The maximum viscosity value of 196.6 mPa s was obtained at concentrations of AlgNa and Tween 80 close to 0.8 and 0.4 in olive oil in incoded values corresponding to coded values of 0.2 % and CTAB of 0.105 %. Furthermore, this figure showes that the maximum value obtained of viscosity is less than the maximum value for the viscosity of 202.5 mPa s obtained in the

previous case. This could be explained by the high viscosity of olive oil compared to that of Tween 80.

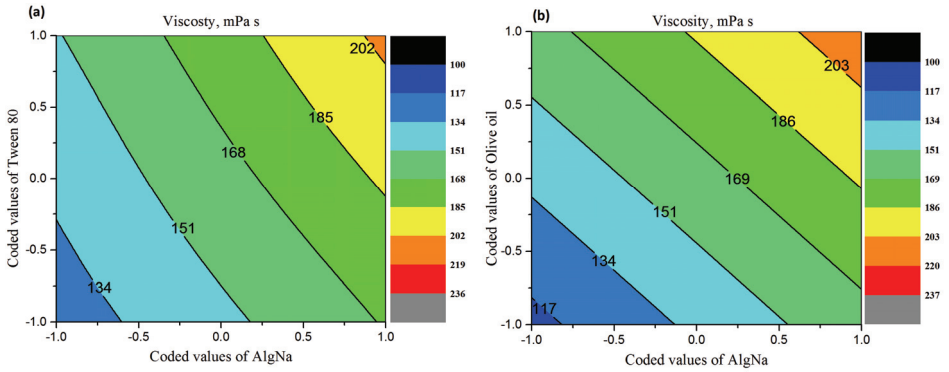


Fig. 3. Contour plots of independent variables: coded values of AlgNa, Tween 80 and olive oil, on the viscosity.

As shown in Fig. 3b, the apparent viscosity values increase with increasing concentrations of AlgNa and olive oil. The maximum viscosity value of 202.5 mPa s was obtained for olive oil concentration close to 0.75 in coded values at constant concentrations of CTAB of 0.105 % and of Tween 80 of 0.025 %.This behavior could be explained by the dispersion of the molecules of the charged polymer in the presence of the viscous oil, which shows significant changes in viscosity values with increasing concentrations of AlgNa and olive oil. In this region (maximum viscosity), the behavior could be attributed to the interactions of surfactant micelles with the polymer chains[28,29] and the associative interaction between the CTAB micelles and the polymer chains may be responsible for leading to the formation of another type of molecules called composite micelles.

Influence of factors on turbidity

Turbidity designates the content of the fluid in the materials that disturb it. In rivers, it is usually caused by suspended matter and colloidal particles that absorb, scatter, or reflect light. In eutrophic waters, it could also be bacteria and micro-algae. The transmittances of solutions containing various amounts of charged polymer, cationic, and nonionic surfactant in the presence of oil could be determined by the turbidimetric method.

The turbidity curves obtained from Eq. (7) are shown in Fig. 4a, which represents plots of iso-response at different coded values of AlgNa and olive oil, with the constants amounts of Tween 80 (0.025 %) and CTAB (0.105 %). The iso-response curves show the effects of these concentrations while the two others were kept constant. The curves show that the minimum turbidity of 200 NTU was obtained for the lowest concentrations of 0.3 % of AlgNa and of 0.1 of olive

oil. The turbidity increases with increase in the cloudiness of the solution, with the increase in the concentration indicated by the curves in Fig. 4a and the maximum turbidity of 300 NTU was obtained for the maximum concentrations of AlgNa and olive oil.

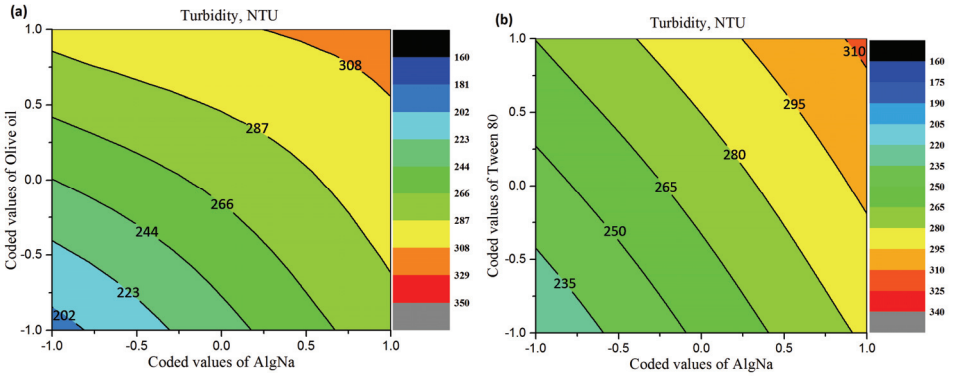


Fig. 4. Contour plots of independent variables: coded values of AlgNa, olive oil and Tween 80 on the turbidity.

The turbidity curves obtained from Eq (7) are illustrated in Fig. 4b, which represents the plots of isoresponse at different coded values of AlgNa and Tween 80 at constants olive oil of 0.2 % and CTAB of 0.105 %. The iso-response curves show the effect of these variable concentrations while the two others were kept constant. The curves show that the minimum turbidity of 200 NTU was obtained as expected at the lowest concentrations of 0.3 % of AlgNa and of 0.4 % of Tween 80. This is expected since the turbidity increases with increasing cloudiness of the solution. This is verified by the curves of Fig. 4b and by the obtained maximum value of the turbidity of 300 NTU.

CONCLUSIONS

The combined effects of the concentrations of CTAB, Tween 80, sodium alginate and olive oil on the physicochemical and viscosimetric properties (interfacial tension, conductivity, viscosity and turbidity) of aqueous solutions were studied using an RSM statistical experimental model to determine if interaction could occur. The study of polymer/surfactants/oil interactions allowed the following conclusions to be drawn:

The conductivity results showed that the tested factors strongly influence this property due to a loading of these factors whenever an increase in conductivity is observed. For the interfacial tension, the obtained values of R^2 and Q^2 were 0.998 and 0.805, respectively. However, the analysis of variance for the model used to estimate the conductivity gave an R^2 value of 0.982 (explaining 98.2 % of the validity of the response) and a Q^2 value of 0.742.

Analysis of the variance for the model obtained for viscosity presented a good R^2 value of 0.976, indicating that only 2.4 % of the total variations were not explained by the model and a Q^2 value of 0.734. For turbidity, the values of the adjusted coefficients of determination R^2 and Q^2 were 0.985 and 0.723, respectively. These values indicate that the model exhibits a good predictive power. For IFT, systems composed of CTAB/AlgNa and CTAB/olive oil, minimum values of 33.85 and 34.39 mN m^{-1} were obtained, respectively. For the same samples, the maximum conductivity values were of 4.126 and 4.064 mS cm^{-1}, respectively. In the case of systems composed of AlgNa/olive oil and AlgNa/Tween80, the maximum viscosity values were 202.5 and 196.6 mPa s, respectively. For the same systems, the maximum turbidity values were 300 and 304 NTU, respectively.

ИЗВОД

МЕТОДОЛОГИЈА ОДЗИВА ПОВРШИНЕ ЗА ПРОУЧАВАЊЕ ИНТЕРАКЦИЈА У ФОРМУЛАЦИЈИ МИЦЕЛАРНОГ СИСТЕМА

КАТАРИНА PSOUHILA OMARI[1], MOHAMED NEDJHIOUI[1], NADJIA HAMIDI[1], OTHMANE BENKORTBI[2] и NABIL BOUARRA[3]

[1]Materials and Environment Laboratory (LME), Department of Chemical Engineering and Environment, Faculty of Technology, Yahia Feres University, 26000 Medea, Algeria, [2]Biomaterials and Transport Phenomena Laboratory (LBMPT), Department of Chemical Engineering and Environment, Faculty of Technology, Yahia Feres University, 26000 Medea, Algeria и [3]Centre de Recherche Scientifique et Technique en Analyses Physico–Chimiques, BP 384, Zone Industrielle Bou-Ismail, 42004 Tipaza, Algeria

Циљ овог је у испитивање интеракције одређених физичко–хемијских својстава на мицеларним системима који се састоје од полимера (натријум-алгинат), два тензида (CTAB и Tween 80) и алжирског маслиновог уља. Моделирање одзива површине (RSM) примењено је за проучавање комбинованих ефеката система који садрже се врсту површински активних средстава. Праћење четири независна параметра као што су површински напон (Y_1), проводљивост (Y_2), вискозност (Y_3) и турбидитет (Y_4), и упоређивање са експерименталним моделом дизајна омогућили су да се утврде перформансе успостављених модела. На основу статистичких анализа, коефицијенти R^2 и K^2 за површински напон, проводљивост, вискозност и турбидитет су: 0,998 и 0,805; 0,982 и 0,742; 0,976 и 0,734 и 0,985 и 0,723, редом. Добијени резултати указују да су ови модели показали добро предвиђање за оптималан систем састављен од CTAB, Tween 80, AlgNa и маслиновог уља. За системе CTAB/AlgNa и CTAB/маслиново уље, вредности површинског напона од 33,85 и 34,39 mN m^{-1}. Такође су добијене максималне вредности проводљивости од 4,126 и 4,064 mS cm^{-1}. За вискозне системе који се састоје од AlgNa/маслиновог уља и AlgNa/Tween 80, добијене су максималне вредности вискозности од 202,5 и 196,6 mPa s. За исте системе, добијене су вредности турбидитета од 300 и 304 NTU.

(Примљено 16. марта, ревидирано 28. априла, прихваћено 29. априла 2021)

REFERENCES

1. A. Avarnas, I. Panagiotis, *J. Colloid Interface Sci.* **258** (2003) 102 (https://doi.org/10.1016/S0021-9797(02)00129-7)
2. N. Kamenka, I. Burgaud, C. Treiner, R. Zana, *Langmuir* **10**(1994) 3455 (https://doi.org/10.1021/la00022a016)

3. L. M. Smitter, J. Guédez, A. J. Müller, *J. Colloid Interface Sci.* **236** (2001) 343 (https://doi.org/10.1006/jcis.2001.7438)
4. T. Gilanyl, E. Wolfram, *Colloids Surfaces* **3** (1981) 181 (https://doi.org/10.1016/0166-6622(81)80077-7)
5. R. Barreiro-Iglesias, C. Alvarez-Lorenzo, A. Concheiro, *Int. J. Pharm*. **258** (2003) 165 (https://doi: 10.1016/s0378-5173(03)00182-0)
6. E. Minatti, D. Zanette, *Colloids Surfaces, A* **113** (1996) 237 (http://dx.doi.org/10.1016/0927-7757(96)03573-X)
7. T. Casgrove, S. J. Mears, T. Obey, L. Thompson, R. D. Welsey, *Colloids Surfaces, A* **149** (1999) 329 (https://doi.org/10.1016/S0927-7757(98)00301-X)
8. W. Guo, Y.W. Sun, G.S. Luo, Y.J. Wang, *Colloids Surfaces, A* **252** (2005) 71 (https://doi.org/10.1016/j.colsurfa.2004.10.013)
9. S. Puvvada, D. Blankschtein, *J. Chem. Phys*. **92** (1990) 371 (https://doi.org/10.1063/1.457829)
10. K. Holmberg, B. Jönsson , B. Kronberg , B. Lindman, *Surfactants and polymers in aqueous solution*, 2nd ed., John Wiley & Sons, LTD, New York, 2002 (ISBN: 0-471-49883-1)
11. A.I. Khuri, J.A. Cornell, *Response surfaces, design and analysis*, Marcel Dekker Inc., New York,1996, p. 536 (ISBN 978-0-367-40125-2)
12. D. C. Montgomery, *Design and analysis of experiments*, 3rd ed., John Willey & Sons Inc., New York, 1991, p. 688 (ISBN 978-1-119-49244-3)
13. K. G. S. Nair, R. Velmurugan, S. K. Sukumaran, *Bio. Nano Sci.* **10** (2020) 279 https://doi.org/10.1007/s12668-019-00713-0
14. J. Goupy, *Anal. Chim. Acta* **544** (2005) 184 (https://doi.org/10.1016/j.aca.2005.01.051)
15. X. Huang, P. Jiang, *Adv. Mater*. **27** (2005) 546 https://doi.org/10.1002/adma.201401310
16. W. G. Cochran, G. M. Cox, *Experimental designs*, 2nd ed., John Willey & Sons Inc., New York, 1990, p. 335 (ISBN978-0-471-54567-5)
17. M. Nedjhioui, J. P. Canselier, N. Moulai Mostefa, A. Bensmaili, A. Skender, *Desalination* **206** (2007) 589 (https://doi.org/10.1016/j.desal.2006.04.065)
18. M. Nedjhioui, N. Moulai Mostefa, A. Bensmaili, A. Morsli, *Desalination* **185** (2005) 543 (https://doi.org/10.1016/j.desal.2005.05.013)
19. M. Nedjhioui, J. P. Canselier, N. Moulai-Mostefa, A. Bensmaili, *J. Disper. Sci. Technol.* **30** (2009) 1331 (https://doi.org/10.1080/01932690902735538)
20. M. Nedjhioui, N. Moulai-Mostefa, A. Sellami, F.Toubal, *Desalin.Water. Treat.* **56** (2015) 2739 (https://doi.org/10.1080/19443994.2015.1012339)
21. M. Nedjhioui, N. Moulai-Mostefa, M. Tir, *Desalin. Water. Treat*. **55** (2015) 3704 (https://doi.org/10.1080/19443994.2014.940217)
22. N. Moulai-Mostefa, R. Khalladi, M. Nedjhioui, *Ann. Chim. Sci. Mat*. **32** (2007) 421 (https://doi: 10.3166/acsm.32.421-429)
23. A. Dal Bo, B. Schweitzer, A. C. Felippe, D. Zanette, B. Lindman, *Colloids Surfaces, A* **256** (2005) 171 (https://doi.org/10.1016/j.colsurfa.2005.01.017)
24. T. D. Blake, Y. D. Shikhmurzaev, *J. Colloid Interface Sci.* **253** (2002) 196 (https://doi.org/10.1006/jcis.2002.8513)
25. D. Langevin, *Adv. Colloid Interface Sci*. **89** (2001) 467 (https://doi.org/10.1016/S0001-8686(00)00068-3)
26. E. D. Goddard, *Colloids Surfaces* **19** (1986) 255 (https://doi.org/10.1016/0166-6622(86)80340-7)

27. V. J. Sovilj, L. B. Petrovic, *Carbohyd.Polym*. **64** (2006) 41 (https://doi.org/10.1016/j.carbpol.2005.10.030)
28. I. M. Harrison, F. Candau, R. Zana, *Colloid Polym. Sci.* **277** 1 (1999) 48 (https://doi.org/10.1007/s003960050366)
29. J. Merta, P. Stenius, E. Pirttinen, *J. Disper. Sci. Technol*. **20** (1999) 677 (https://doi.org/10.1080/01932699908943814).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 86 (7–8) S145–S147 (2021) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Response surface methodology for the study of interactions between components in a micellar system formulation

SOUHILA OMARI[1], MOHAMED NEDJHIOUI[1], NADJIA HAMIDI[1], OTHMANE BENKORTBI[2]* and NABIL BOUARRA[3]

[1]*Materials and Environment Laboratory (LME), Department of Chemical Engineering and Environment,Faculty of Technology, Yahia Feres University, 26000 Medea, Algeria,* [2]*Biomaterials and Transport Phenomena Laboratory (LBMPT), Department of Chemical Engineering and Environment, Faculty of Technology, Yahia Feres University, 26000 Medea, Algeria and* [3]*Centre de Recherche Scientifique et Technique en Analyses Physico–Chimiques, BP 384, Zone Industrielle Bou-Ismail, 42004 Tipaza, Algeria*

J. Serb. Chem. Soc. 86 (7–8) (2021) 725–738

TABLE S-I. Matrix of coded values X_i

Exp No.	Expname	Run order	Incl/Excl	Coded values X_i			
				CTAB	Tween 80	Olive oil	Alg Na
1	N1	12	Incl	-1	-1	-1	-1
2	N2	15	Incl	1	-1	-1	-1
3	N3	8	Incl	-1	1	-1	-1
4	N4	16	Incl	1	1	-1	-1
5	N5	4	Incl	-1	-1	1	-1
6	N6	5	Incl	1	-1	1	-1
7	N7	9	Incl	-1	1	1	-1
8	N8	7	Incl	1	1	1	-1
9	N9	11	Incl	-1	-1	-1	1
10	N10	2	Incl	1	-1	-1	1
11	N11	6	Incl	-1	1	-1	1
12	N12	1	Incl	1	1	-1	1
13	N13	14	Incl	-1	-1	1	1
14	N14	10	Incl	1	-1	1	1
15	N15	13	Incl	-1	1	1	1
16	N16	3	Incl	1	1	1	1

* Corresponding author. E-mail: benkortbi_oth@yahoo.fr

TABLE S-II. Coefficient of determination and of prediction obtained from the experimental results

	R^2	$R^2_{(adj)}$	Q^2	SD	RS	N
Interfacial tension	0.99	0.99	0.80	2.53	0.17	16
Conductivity	0.98	0.94	0.74	0.34	0.08	16
Viscosity	0.97	0.93	0.73	41.86	8.43	16
Turbidity	0.98	0.95	0.72	54.18	11.37	16

SD: standard deviation; RS: residual

TABLE S-III. Levels of independent variables in uncoded form and responses

Exp N°	Content, %				Interfacial tension, mN m^{-1}	Conductivity, mS cm^{-1}	Viscosity, mPa s	Turbidity, NTU
	CTAB	Tween 80	Olive oil	AlgNa				
1	0.01	0.01	0.10	0.30	37.90	3.00	100	150
2	0.20	0.01	0.10	0.30	37.00	3.55	110	184
3	0.01	0.04	0.10	0.30	38.40	3.65	115	195
4	0.20	0.04	0.10	0.30	36.50	3.80	125	251
5	0.01	0.01	0.30	0.30	35.40	3.85	125	268
6	0.20	0.01	0.30	0.30	34.70	3.90	165	290
7	0.01	0.04	0.30	0.30	39.00	3.95	175	300
8	0.20	0.04	0.30	0.30	38.00	4.00	185	315
9	0.01	0.01	0.10	0.80	38.00	3.45	120	250
10	0.20	0.01	0.10	0.80	36.80	3.90	165	280
11	0.01	0.04	0.10	0.80	35.50	3.85	180	278
12	0.20	0.04	0.10	0.80	33.40	4.25	185	305
13	0.01	0.01	0.30	0.80	32.00	3.95	187	285
14	0.20	0.01	0.30	0.80	31.50	4.25	205	315
15	0.01	0.04	0.30	0.80	32.90	4.20	220	320
16	0.20	0.04	0.30	0.80	32.00	4.35	237	350

TABLE S-IV. ANOVA analysis for the regression of the model representing interfacial tension, conductivity, viscosity and turbidity using coded values

	DF	SS	MS	F	p	SD
			Interfacial tension, mN m^{-1}			
Regression	10	95.97	9.59	336.75	0	3.10
Residual	5	0.14	0.030			0.17
Lack of fit	5					
Pure error	0					
Total	16	20.33	12.70			
Total corrected	15	96.11	6.41			2.53
R^2= 0.99						
$R^2_{(adj)}$ = 0.99						
Q^2 = 0.80						
			Conductivity, mS cm^{-1}			
Regression	10	1.72	0.17	26.94	0.001	0.41
Residual	5	0.032	0.006			0.08
Lack of fit	5					

TABLE S-IV. Continued

Pure error	0					
Total	16	241.22	15.10			
Total corrected	15	1.75	0.17			0.34
$R^2 = 0.98$						
$R^2_{(adj)} = 0.94$						
$Q^2 = 0.74$						
			Viscosity, mPa s			
Regression	10	1.72	0.17	26.94	0.001	0.41
Residual	5	0.032	0.006			0.08
Lack of fit	5					
Pure error	0					
Total	16	241.22	15.07			
Total corrected	15	1.75	0.12			41.86
$R^2 = 0.97$						
$R^2_{(adj)} = 0.93$						
$Q^2 = 0.73$						
			Turbidity, NTU			
Regression	10	43387.50	4338.75	33.55	0.001	65.87
Residual	5	646.50	129.30			11.37
Lack of fit	5					
Pure error	0					
Total	16	1. 21e006	76193.10			
Total corrected	15	44034	2935.60			54.18
$R^2 = 0,98$						
$R^2_{(adj)} = 0,95$						
$Q^2 = 0.72$						

DF: degree of freedom; SS: sum of squares; MS: mean square; *F*: Fisher test; *p*: probability; SD: standard deviation.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. *86* (7–8) 739–752 (2021)
JSCS–5458
Original scientific paper

Efficiency of different additives in the improvement of the oxidation stability of fatty acid methyl esters with different properties

MILICA RANKOV ŠICAR[1,2], RADOSLAV MIĆIĆ[3], MILAN TOMIĆ[4] and NATAŠA ĐURIŠIĆ-MLADENOVIĆ[1]*

[1]*University of Novi Sad, Faculty of Technology Novi Sad, Novi Sad, Serbia,* [2]*SP Laboratorija a.d., Bečej, Serbia,* [3]*University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia and* [4]*University of Novi Sad, Faculty of Agriculture, Novi Sad, Serbia*

(Received 9 February, pevised 8 April, accepted 12 April 2021)

Abstract: This study evaluates six formulations in improving the oxidation stability of different fatty acid methyl esters (MEs). Two MEs differed in the unsaturation levels as they were synthesized from different feedstock: a blend of soybean and sunflower oils, and waste cooking oil; they did not fulfill the requirements of the EN 14214 standard concerning their oxidation stability (≈0.6 h) and some impurities. The third MEs (SoSuME-EN) were fully compliant with the standard. Five formulations were phenolic-based, containing single or mixed antioxidant compounds of different molecular structures; one was amine-based. Different dosages of the formulations were added to the ME samples (corresponding to the addition range 50–48300 ppm). The MEs stability expressed as induction periods, *IP*, determined by the Rancimat method, were used for the calculation of stabilization factors, *SF*, indicating the efficiency of the applied formulation. The formulation containing 2-(1,1-dimethylethyl)-1,4-benzenediol was the most efficient concerning the lowest consumption rate and the highest *SF* achieved for the low quality ME. 2,2′-Methylene--bis-(4-methyl-6-*tert*-butylphenol) was linked with higher antioxidant potency than the amine-based formulation and the phenolic compounds with two bulky *tert*-butyl groups. Among the 4 selected phenolic additives, butylated hydroxytoluene and 2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol) had similar efficiency in SoSuME-EN (at ≈500 ppm they produced *SF* ≈ 2), while it took twice of this amount for mixed butylphenols to achieve the same effect.

Keywords: biodiesel; phenolic antioxidants; amine-based additive; oxidation stability; Rancimat induction period.

INTRODUCTION

Fatty acid methyl esters (MEs) are main constituents of biodiesel fuels, considered as one of the most promising petroleum diesel fuel substitutes. One of the

* Corresponding author. E-mail: natasadjm@tf.uns.ac.rs; natasa.mladenovic@uns.ac.rs
https://doi.org/10.2298/JSC210209029R

most important characteristics of biodiesel is oxidation stability, which describes its susceptibility to oxidation that results in degradation of the composition and consequently of the fuel properties.[1] In fact, poor oxidation stability is one of main shortages of biodiesel fuels compared to conventional diesel. This is why oxidation stability has been addressed in biodiesel quality standards, such as the European standard EN 14214 and the American standard ASTM D6751. According to the standard EN 14214 from 2012 that defines quality of MEs as automotive fuels, the minimum required value for the oxidation stability expressed as an induction period (or induction time), *IP*, determined by the Rancimat method (EN 14112) is 8 h, which is more rigorous than the previous requirement of 6 h (EN 14214 from 2005). The *IP* represents the time during which changes in the MEs composition induced under elevated temperature (110 °C) and intensive contact with air (accelerated oxidation) in the Rancimat apparatus are not significant for the fuel quality; after this period, serious changes in the composition greatly affect the applicability of the fuel. The longer the induction period is, the better is the oxidation stability. The ASTM D6751 standard specification for use of biodiesel as a blend component with middle distillate fuels is less demanding, setting 3 h as the minimum for oxidation stability.[1]

Degree of oxidative degradation is influenced by a number of factors: the most influential is the fatty acid composition of the MEs inherited from the oily feedstock, preferentially the relative contents of mono- and poly-unsaturated MEs. The trend of increasing oxidation rate among the most abundant unsaturated fatty acid with 18 C-atoms is as follows: oleic (having 1 unsaturated double bond, shortly marked C18:1) < linoleic (with 2 doubled bonds, marked C18:2) << linolenic (with 3 double bonds, C18:3);[2–4] The *IP* measured at 110 °C for the MEs of these fatty acids are: 2.5, 1 and 0.2 h, respectively.[5] A number of excellent reviews on different aspects of biodiesel oxidation have been published.[3,5–10] Apart of the fatty acid composition, the presence of minor components such as diacylglycerols, monoacylglycerols, FFAs, phospholipids, moisture and mineral impurities also affects the antioxidant behavior of biodiesel;[10–13] elevated temperature during storage as well as the exposed surface area between biodiesel and air might also increase the oxidative degradation.[1]

One of the most efficient methods for improving the stability and delaying the oxidative degradation of biodiesel is the addition of antioxidants. Antioxidants are chemicals that react with oxidative species more easily than the unsaturated MEs, leading to formation of a stable molecule, which is further resistant to the chain oxidation process.[3] Although effects of various antioxidants on the stability of MEs have often been investigated, a majority of the reported results refer to the addition of a single compound, whereas only few studies considered mixtures that resembled the commercially available formulations.[14–16] Namely, commercial formulations contain one or more active antioxidant compounds in

varying quantities; they have different physical properties (solid state or dissolved, density, viscosity, boiling point, *etc.*) that could affect the compatibility with MEs. Dosing and compatibility of available formulations with a particular biodiesel needs extensive laboratory trials, unless there are comparative information on efficiency of the formulations in improving the stability of fuels similar to the one of interest in a particular case, which can contribute to the more time- and cost-efficient trials. It has to be emphasized that there is no ideal antioxidant that prolongs storage stability of all kinds of biodiesel fuels. Additionally, combination of two or more antioxidants might induce a synergetic effect, enhancing *IP* of biodiesel above the sum of *IP* achievable with each antioxidant separately.[10] On the contrary, there are also examples when antioxidants added to biodiesel were associated with decrease in the oxidation stability due to some antagonistic effects.[16]

The main aim of this study was to test several antioxidant formulations in improving and protecting the stability of MEs of different origin and quality. By measuring *IP* before and after the addition of various formulation dosages (from rather low to massive ones) to MEs, their antioxidative efficiencies were compared, indicating the most efficient active compound(s) in term of the highest enhancement of the stability. Comparison of the results for MEs of varying qualities gave an insight into the antioxidant potencies over a wide range of concentrations in different substrates, also indicating indirectly the influence (or importance) of fuel impurities, the unsaturation level and/or starting *IP* on the protective power of the antioxidants. To the best of the authors' knowledge, this is the first study that compares mono- and multi-active compound-based formulations in different MEs. It contributes to the existing understanding of the protection power of antioxidants in different substrates and also provides valuable information for the selection and application of antioxidants in the fuel sector and other applications in which fatty compounds may be exposed to oxidative stress.

EXPERIMENTAL

Two different raw materials were used for the synthesis of the MEs: a blend (1:1 mass ratio) of domestically produced soybean and sunflower oils, and waste cooking oil collected from a restaurant at the Faculty of Agriculture, University of Novi Sad. The MEs produced from these two feedstock materials are marked as SoSuME and WCOME, respectively. The MEs were produced by a conventional alkali homogeneous transesterification in a laboratory glass reactor (methanol to oil mole ratio 8:1; 1 % KOH catalyst in oil) at a temperature ≈64 °C for 1 h. After the reaction, the product mixture was allowed to settle; the upper layer of MEs was collected, while the lower glycerol layer was discarded. The MEs were purified with distilled water and dried under vacuum. The schematic representation of the experimental setup is given in Fig. S-1a of the Supplementary material to this paper. The MEs were synthesized in successive batches and then homogenized by mixing all batches of a particular ME in one vessel in order to ensure that the total amount was enough for the planned number of experiments.

Before addition of the formulations, the MEs were characterized in accordance to the requirements of EN 14214 and the results are presented in Table S-I of the Supplementary material. The methods used for the characterization of the MEs are also listed in Table S-I. These analyses were performed in SP Laboratorija (Bečej, Serbia) accredited under the ISO/ /IEC 17025 standard. As could be seen, two quality levels of SoSuME were used: one in compliance with EN 14214 (marked as SoSuME-EN), having initial *IP* of 8.88 h, and the other (SoSuME) that failed the EN 14214 requirements for minimum oxidation stability (the initial *IP* was 0.60 h) and for the methyl ester content, as well as for the maximum content of alkali and earth-alkali metals, phosphorous and water. WCOME was also characterized with low starting oxidation stability (0.63 h) and elevated levels of metals but not so remarkably as in the case of SoSuME. In fact, SoSuME and WCOME were intentionally not purified satisfactory, and they were left for "aging" (under sunlight and in contact with air) for a few months before the addition of the antioxidants. In this way, low quality MEs were produced. Concerning the contents of mono- and poly-unsaturated C18 MEs determined by gas chromatography (GC), SoSuME and WCOME differed: WCOME had more than two times higher content of C18:2 esters (57 %) compared to SoSuME/SoSuME-EN (22 %), while SoSuME/ /SoSuME-EN (3.2 %) had 32 times higher content of C18:3 than WCOME (0.1 %); the content of C18:1 differed less between MEs (39 % in SoSuME/SoSuME-EN and 31 % in WCOME). In summary, the total C18 unsaturation level (calculated as the sum of the contents of the C18 unsaturated MEs determined by GC method listed in Table S-I) differed markedly between WCOME (88.1 %) and SoSuME (64.2 %). The significant difference was seen also in the content of C16:0 (30 % in SoSuME/SoSuME-EN and 7 % in WCOME). These differences indicated different susceptibilities of these MEs to oxidation. The oxidizability of MEs was calculated using the contents (%) of MEs of the C18:1, C18:2 and C18:3 fatty acids in the following manner:[5] oxidizability = (0.02 · C18:1 % + C18:2 % + 2 · C18:3 %)/100.

Six antioxidant formulations, marked from A to F, were available for testing. Five formulations were phenolic-based (A-E) and one was amine-based (F); five contained one active antioxidant compound (*i.e.*, mono-active compound-based formulations: A-D and F), while one represented a mixture of different phenolic compounds (*i.e.*, multi-active compound-based formulation: E). The main antioxidant compounds in the formulations are listed in Table I.

TABLE I. Composition of the commercial formulations A-F used in this study based on the Material Safety Data Sheets (MSDSs) provided by the producers

Formulation	Composition	Type
A	2,6-Di-*tert*-butyl-*p*-cresol (butylated hydroxytoluene – BHT >99.7 %)	Phenolic
B	2-*tert*-Butylhydroquinone (TBHQ, 10–20 %), butene, homopolymer (2.5–5 %), solvent naphtha, heavy aromatics (0.25–1 %)	Phenolic
C	2,2'-Methylene-bis-(4-methyl-6-*tert*-butylphenol) (20 %)	Phenolic
D	2,2'-Methylene-bis-(4-methyl-6-*tert*-butylphenol) (96–99 %)	
E	2,6-Di-*tert*-butylphenol (>75 %), 2,4,6-tri-*tert*-butylphenol (15 %), 2-*tert*-butylphenol (7 %), 2,4-di-*tert*-butylphenol (2 %), 2,5-di-*tert*--butylphenol (1.7 %), phenol 0.2 %, 4-*tert*-butylphenol (0.2 %)	Phenolic
F	Amine substituted resin (60-100 %), heavy aromatic naphtha (0–30 %), naphthalene (0–5 %), 1,2,4-trimethylbenzene (0–5 %)	Amine

The A additive was based on a compound with two bulky *tert*-butyl groups linked to an aromatic ring in the ortho positions (2,6-) to the OH-group. The main antioxidant compound

in B consisted of two OH groups linked to aromatic ring in the para position (1,4-) and one *tert*-butyl group in the ortho position (2-). The C and D formulations had the same main active compound (with two methyl-butylphenol rings bridged by a CH_2-group) but in different concentrations prepared by the producer. The E formulation had a unique composition compared to the other phenolic formulations as it represented a mixture of various phenols with one, two or three *tert*-butyl groups. The additive F contained mainly an amine-based resin. All of them were tested for improving the oxidation stability of SoSuME and WCOME, while four phenolic-based formulations, found to be less effective for these low quality MEs, were further assessed in maintaining the initially high stability of SoSuME-EN under the accelerated oxidation test. Some properties of the formulations are given in Table II. Different dosages of the formulations were added to 100 mL of MEs, as presented in Table III.

TABLE II. Properties of the commercial formulations A–F used in this study; properties except boiling points were taken from the MSDSs. Boiling points are not given in the MSDSs, but they were taken from Wikipedia pages, considering the main components of the tested formulations: BHT for A; TBHQ for B; and 2,6-di-*tert*-butylphenol for E; the boiling point for 2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol), the main compound in C and D, was taken from https://www.alfa.com/en/catalog/H36783/

Property	Formulation					
	A	B	C	D	E	F
Physical state	Solid	Liquid	Liquid	Solid	Liquid	Liquid
Density, kg dm^{-3}	1.03 at 20 °C	0.914 at 15 °C	0.89 at 20 °C	1.04–1.1 at 20 °C	0.94 at 20 °C	0.92 at 20 °C
Flash point, °C	127	62	180	185	114	64
Viscosity, kinematic, $mm^2\ s^{-1}$	–	605 at 20 °C 135 at 40 °C	–	–	–	157 at 100 °C
Solubility in water, mg dm^{-3}	0.76	0.03	Insoluble	0.007	0.4	–
Boiling point, °C	265	273	187	187	253	–

TABLE III. The mass concentrations (in ppm_m) of the main active compounds added to the methyl esters (MEs) in different amounts (in mL) of the formulations A–F; for liquid formulations, it is calculated by taking into account the volumes of the formulations added to 100 mL of MEs, the average value of the MEs densities given in Table S-I (886 kg m^{-3}), the densities of the formulations presented in Table II, and the percentage share of the main active compound. The measured mass of solid additives A and D were taken for the calculation of the relevant ppm_m concentrations

Formulation dosage in mL added to 100 mL of MEs	A	B	C	D	E	F
0.05	564	51–102	100	604	530	519
0.10	1129	102–203	201	1208	1061	1038
0.20	2257	203–406	402	2415	2122	2077
0.50	5643	508–1016	1005	6038	5305	5192
1.0	11287	1016–2032	2009	12077	10609	10384
3.0	33860	3047–6095	6027	36230	31828	31151
4.0	45147	4063–8126	8036	48307	42438	41535

Before the addition, the liquid formulations were heated up to 60 °C in order to reduce the viscosity and make easier pipetting.[17] In the case of the solid additives (A and D), the adequate quantity was first dissolved in toluene, added to 100 mL of MEs and mixed, followed by evaporation of the solvent. As ≈4 % is the limit for constituents other than MEs in accordance to the limit for the MEs content (96.5 %) requested by EN 14214, the higher dosages were not tested. Although massive dosages of the antioxidants (*i.e.*, those ≥1 %) could be considered neither economically nor technically feasible for the biodiesel industry,[18] they were applied in order to investigate their protective potency over a wide range of concentrations, particularly to comparatively investigate their ability to increase the low initial stability of poor quality substrates with different unsaturation levels.

The efficiency of additives was expressed as a stabilization factor, *SF*,[19] calculated by dividing *IP* obtained for the additized MEs (*i.e.*, MEs with added antioxidant formulation, IP_X) with the initial (starting) *IP* of the untreated (neat) MEs (*i.e.*, MEs without the additive, IP_0). The *IPs* were determined in accordance to the method EN 14112 using Metrohm 743 Rancimat equipment (Herisau, Switzerland). The principle of the method for the measurement of the *IPs* is illustrated in Fig. S-1b of the Supplementary material. Three grams of sample were placed in the heating block at 110 °C and intensively oxidized with an air flow (10 L h^{-1}), which resulted in formation of volatile products (mainly formic and acetic acids)[4] that were transferred to a flask containing 50 mL of distilled water. After some time, large quantities of volatiles dissolved in water induce a significant increase in the measured conductivity. The time corresponding to the inflection point in the oxidation curve (conductivity *vs.* time) is the *IP*. Each sample was analyzed in duplicate and the results on *IP* are presented as averages.

RESULTS AND DISCUSSION

Impacts of the antioxidant formulations added in different dosages on the *IPs* of studied MEs are presented in Fig. S-2 of the Supplementary material. The lowest dosages of the antioxidants that provided *IPs* above the limit of 8 h for the low quality MEs (SoSuME and WCOME) are listed in Table IV. The stabilization factors, *SF*, calculated for the tested formulations and the added amounts, are presented in Fig. S-3 of the Supplementary material. The highest *SF* and the corresponding *IP* achieved for each of the additized MEs are given in Table V.

TABLE IV. The lowest dosages of the formulations A–F (in mL) followed by the corresponding active antioxidant compound (in ppm_m) that provided oxidative stability of low quality methyl esters (SoSuME and WCOME) above the EN 14214 limit of 8 h; NA: the EN 14214 limit is not achieved (*IP* was < 8 h) at any of the tested dosages

Ester	A	B	C	D	E	F
SoSuME	1/11287	0.5/508-1016	1/2009	3/36230	1/10609	0.5/5192
WCOME	NA	1/1016-2032	NA	3/36230	NA	NA

Concerning MEs that were not in compliance with EN 14214, *i.e.*, SoSuME and WCOME, the lowest additions of 0.05, 0.10 and 0.20 mL, corresponding up to about 2500 ppm_m of the main antioxidant compounds (Table III), were insufficient to improve the very low starting *IP* to be in compliance with the standard thresholds (Fig. S-2), even though these dosages increased the initial *IP* by factor of 2–4 (Fig. S-3). It could be seen that the *IP* of these MEs increased above the

less stringent ASTM D6751 limit with the 0.50-mL dosages of the formulations (Fig. S-2), except for C and D. The dosage of 0.50 mL increased *IP* of SoSuME above the EN 14214 limit of 8 h only in the case of B and F (Fig. S-2). Taking into account the concentrations of the main active compounds in 0.50 mL of B and F (Table III), it could be concluded that the more potent antioxidant was the TBHQ present in B than the amine substituted resin in F. Looking further into the contents of the active compounds in other formulations needed to be added to SoSuME to overcome the 8-h limit (Table IV), it could be observed that the active compound of C (2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol)) also had marked antioxidant effect as its concentration was only 2 times that of TBHQ, while approximately 10–20 times more of the active compounds in A (2,6-di--*tert*-butyl-*p*-cresol) and E (2,6-di-*tert*-butylphenol and other phenols with one, two or three *tert*-butyls) were needed compared to TBHQ. Even though D contained the same active compound as C but in a higher amount, it showed very low antioxidant efficiency as a very large quantity was needed to be added to SoSuME to reach the 8-h limit (35–70 times more than in the case of TBHQ). As D is a solid substance that needed to be dissolved before addition to MEs (toluene was used), while C was a solution (20 %) of the same compound prepared by manufacturer (using biodiesel as the solvent), it could be only presumed that the method of the dissolution and/or mixing of the D solutions with MEs, particularly in the lower addition volumes, was the reason for its low efficiency. Furthermore, the lower boiling point of the active compound in D could be the reason for its higher losses during the solvent evaporation step after addition to MEs than for the active compound in A, which was also prepared by dissolving in toluene since it was provided in the solid state as D (Table II). Hence, the increasing quantities of the relevant active compounds necessary to enhance the *IP* of SoSuME above 8 h, ordered the formulations as B < C < F < E=A (D was excluded here due to the probable mixing problem that impaired its protective efficiency).

TABLE V. The highest stabilization factors (*SF*)/induction periods (*IP* in h) achieved for the antioxidants A–F added to the tested methyl esters SoSuME, WCOME and SoSuME-EN (numbers in parentheses are the antioxidant dosages in ppm_m that provided the specified values of *SF*/*IP*)

Formulation	SoSuME	WCOME	SoSuME-EN
A	74.2/44.5 (11287)	11.9/7.5 (11287)	11.1/98.3 (45147)
B	116.8/70.0 (1016-2032)	27.2/17.2 (1016-2032)	–
C	24.6/14.8 (2009)	7/4.4 (8036)	5.7/50.7 (8036)
D	43.2/25.9 (48307)	28.8/18.4 (48307)	18.0/159.5 (48307)
E	52.7/31.6 (10609)	6.4/4.0 (5305)	6.26/55.6 (42438)
F	22.1/13.3 (5192)	10.2/6.4 (5192)	–

The highest *IP* of SoSuME were not measured for the highest applied dosages of the formulations, except in the case of D (Fig. S-2, Table V). This *IP* was several dozen times higher than the starting *IP* of SoSuME and the formulations could be order according to the highest achieved *SF* as follows (Table V): B (116.8) > A (74.2) > E (52.7) > D (43.2) > C (24.6) > F (22.1). The main antioxidant concentrations that induced these highest *IP*/*SF* varied significantly (Table V): B and C reached their maximum antioxidative potency in SoSuME at similar concentration of up to ≈2000 ppm_m, whereas the maximum for A and E was at higher levels ≈11000 ppm_m. It could be stated that the most efficient antioxidants for SoSuME were TBHQ and 2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol), with the later applied as the original solution (prepared in biodiesel by the producer, C). Difference seen between C and D regardless of the same active compound they contained could be ascribed to different states in which they were provided by the producer that further implied different ways of the application.

In the case of WCOME, an *IP* above 8 h was achieved only with B and D (Table IV). In fact, TBHQ and 2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol) were the only compounds that enabled improvement of the oxidation stability of the highly unsaturated WCOME to be in compliance with EN14214, while BHT in A, mixed mono-, di- and tri-*tert*-butylphenols in E, and amine substituted resin in F failed. These two compounds were also successful in making SoSuME compliant with the EN standard at the lowest consumption rates, but WCOME required larger dosages. As the highest dosage of the C formulation (4 mL) that contained 8036 ppm_m of the same active compound as D did not increase the *IP* above 8 h (Fig. S-2), it could be assumed that the massive consumption rate of the D formulation was the consequence of both the lower ability of 2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol) to protect MEs with high unsaturation, and the possible problems of mixing with the MEs, previously also suspected in the case of SoSuME. Generally, it could be observed that all formulations were less efficient for WCOME compared to SoSuME, as the *IP* and *SF* of the additized WCOME were lower than in the case of SoSuME (Figs. S-2 and S-3). This could be ascribed to higher unsaturation level in WCOME and, as expected, its much higher oxidative instability than SoSuME: the oxidizability[5] for WCOME was double the one for SoSuME (0.59 and 0.29, respectively). This further implied that the unsaturation level has a more important effect on reducing stability than the level of impurities found in low quality MEs.

The highest efficiency of B could be explained by the unique dihydroxyl aromatic structure of TBHQ that more easily donates hydrogen atoms (radical) to free radicals of esters produced during the oxidation stress of MEs than monohydroxyl structures present in A and E, interrupting the propagation step of the free radicals chain more readily, and making it a more powerful and protective antioxidant.[5,20] Interestingly, 2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol)

in C and D also contains 2 OH-groups per molecule, but each of them is linked to one of two aromatic rings in the bisphenyl structure, which seems to render them less protective than 2 OH-groups linked to an aromatic ring in para-position, as in TBHQ. Besides the number of OH-groups in the molecule, presence of alkyl groups is also important for antioxidant behavior, as hydroxyl groups on benzene ring do not show antioxidant behavior unless its ortho or para positions were occupied by alkyl groups; in this way, the electron density on the OH-groups is increased, which makes the release of a hydrogen radical easy.[5] Both compounds, TBHQ and 2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol), contain one bulky butyl group in the ortho position to an OH-group. On the other hand, the main active compounds in A and E had two tertiary butyl groups in ortho position to OH-group that possibly induced more steric hindrance, suppressing the electron-release from the OH group and making A and E less efficient, especially if compared to TBHQ.[16] Apart of the favorable structure, the highest antioxidant efficiency of TBHQ could be additionally related to the products known to be formed during the antioxidative activity of TBHQ that may also have antioxidant properties, synergistically retarding the oxidation.[20] Besides the chemical structure, difference in the antioxidant efficiency of the phenolic compounds could also be attributed to physical properties of the formulations that might influence compatibility and mixing with the substrates as was suggested for the poor performances of D (originally provided as a solid substance).

The unique composition of the additive F (the only amine-based additive among the tested ones) influenced the oxidation stability of the treated low quality MEs in a rather different manner in comparison to the phenolic additives. It seemed to be moderately efficient in the stability enhancement of SoSuME compared to other phenolic-based formulations. At the level of 0.50 mL that contained 5192 ppm_m of the active compound (Tables IV and V, Fig. S-2), an *SF* of 22.1 was achieved (Fig. S-3), which was enough to surpass the threshold of 8 h required by EN14214. However, the same dosage failed to increase *IP* of WCOME above the lower limit. Further increase of the dosages actually fully destabilized SoSuME and WCOME as the *IP* was measured to be ≈0. This might be interpreted in the following way: interaction of large dosages of the formulations with the low quality MEs constituents induced an almost instantaneous increase of the conductivity in the measuring cell of the Rancimat apparatus. This observation will be further explained hereafter in comparison to the results obtained for high additions to SoSuME-EN. It is known that the antioxidant function of an aromatic amine is superior to that of phenols as many more radicals per molecule could be trapped than in the case of phenols and therefore, the usually the required quantity of amine-based antioxidants is lower than for the phenolic ones,[10] which was the case in this study if less efficient phenolic formulations A and E were taken into account. Additionally, it is worth noting that the addition

of amine-based antioxidants leads to the introduction of nitrogen into the biodiesel composition, diminishing one of its advantages over fossil fuels of being fuel without heteroatoms such as S and N. This was the reason why it was excluded from the further testing performed in this study.

It is known that antioxidants efficiently function in an optimal range of concentrations between the critical and the finite saturation levels beyond which no further improvement in the antioxidant activity could be achieved.[10] For the low quality MEs tested in this study, it was found that there was a dosage that induced maximum ("peak") of *IP* for all the tested phenolic formulations except for D (probably because of the mentioned problem of its mixing), Fig. S-2. This maximum *IP* might be regarded as the "saturation" level for SoSuME and WCOME. After the observed "peaks", the reduced *IP*, *i.e.*, antioxidant power was seen with further increase of dosages (Fig. S-2) that could be considered as counter-intuitive behavior. In some cases, the high dosages even resulted in full disturbance of the oxidation stability manifested as $IP \approx 0$ (Fig. S-2). The null value of *IP* was not seen in the additized good quality MEs (SoSuME-EN). It might be only suspected that all the antioxidants added at the higher dosages displayed some pro-oxidative behavior in the presence of impurities in SoSuME and WCOME, which was more pronounced for monoaromatic phenolic and amine structures in the MEs with higher unsaturation level (WCOME), but this observation requires further study.

As the B formulation was found to be the most efficient antioxidant for the low quality MEs (Tables IV and V), the formulation was presumed to be highly likely the most protective also in the case of SoSuME-EN. Thus it was excluded from the further analysis of the phenolic antioxidants, and only A, C, D and E that were found to be less effective for low quality MEs, were added to SoSuME-EN in order to examine their capability to retain oxidative degradation of biodiesel with an initial *IP* above the EN threshold. The following levels led to a doubling of the initial *IP* of SoSuME-EN: 500–1000 ppm_m of BHT in A and of mixed butylphenols in E, and ≈400 ppm_m (≈2400 ppm_m) of 2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol) in C (and D). Thus, the efficiency of BHT, 2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol), and mixed butylphenols added to high quality MEs was similar (500–1000 ppm_m); D might be excluded from this observation due to already mentioned possible problem of the mixing.

Generally, the *SF* achieved for SoSuME-EN were lower than for SoSuME and WCOME (Fig. S-3). The *SF* calculated for the low dosages were in the narrow range up to 5.3, whereas more pronounced differences among *SF* (2.7–116.8) were seen for the higher dosages (1, 3 and 4 mL), probably because differences in antioxidation mechanisms and specific interactions of the formulation with the substrate were more pronounced at higher concentrations. Additionally, the Rancimat method has its own limitations in measuring long *IP*, when the

results may not be accurately obtained because of water evaporation in the conductivity cell.[20] Different trends between *IPs* and the dosages could be observed among the formulations added to SoSuME-EN. Namely, almost the same linear trend of the *IP* increase with the dosage was seen for additions of A and E (Fig. S-2); for C and D, a slight increase of *IP* was observed up to the 0.20-mL dosage followed by the sharp trend of the *IP* increase with the higher dosages. The saturation levels as maximum *IP* ("peak") were not seen for any of the tested formulations in SoSuME-EN, contrary to those observed for low quality MEs (Fig. S-2), which may be the consequence of specific interactions of impurities in the latter with active compounds in higher levels during accelerated oxidation.

Literature data also evidenced high potency of TBHQ. The comprehensive study of Mittelbach and Schober on the efficiency of 10 synthetic antioxidants (besides 10 mixtures of natural antioxidants also included in the study) added at 1000 ppm to four different MEs (rapeseed oil-, waste frying oil-, sunflower oil-, and tallow-based) also revealed TBHQ to be among the most efficient synthetic antioxidants, producing the greatest enhancement of the *IP*; *i.e*., at the level of 1000 ppm of TBHQ, the *SF* obtained for undistilled waste frying oil-based biodiesel with an initial oxidation stability of 5.9 h was about 5.[19] The widely used BHT was also tested in that study, but it was not as efficient as TBHQ: at a concentration of 1000 ppm, it led to an *SF* of ≈2 for the same biodiesel,[19] coinciding with the finding for the same dosage of the A formulation added to SoSuME-EN in the present study. The relative inefficiency of BHT compared to TBHQ was suspected to be the consequence of partially evaporation during measurements at 110 °C, failing to protect the MEs from oxidative degradation.[19] Again, the higher efficiency of TBHQ than BHT was proven in the study of Ryu: TBHQ at 300 ppm and BHT at 1000 ppm provided sufficient resistance to oxidation of soybean MEs (initial *IP* of 1.36 h) taking into account the previous EN 14214 requirement of 6 h; TBHQ at a dosage of 2000 ppm enhanced the *IP* over 40 h, while BHT did not provide *IP* over 10 h at the same dosage.[21] Yang *et al.* investigated the efficiency of several antioxidants with dosages up to 8000 ppm to improve very low starting *IP* (0.7 h) of the soybean MEs.[22] They also reported TBHQ as more efficient than BHT: at 3000 ppm TBHQ improved initial *IP* to ≈6 h, while BHT did not increase *IP* to 6 h even at 8000 ppm. Domingo *et al*. studied the effect of BHT and TBHQ on soybean oil ethyl esters with very low initial *IP* of 0.16 h at dosages ranging from 200 to 8000 ppm: the highest efficiency was observed for TBHQ at 8000 ppm with *SF* of 52.52; at the same dosage, BHT provided a *SF* of 35.59.[23,24] Lapuerta *et al*. found that BHT amounts of ~4000 ppm and ≈1500 ppm in biodiesel from used cooking oil and soybean oil (initial *IP* of 0.7 and 3.61 h), respectively, were enough to achieve the limit of 8 h at 110 °C.[18]

CONCLUSIONS

The range of the results confirmed that MEs of different compositions required different dosages of the analyzed formulations, *i.e*., the potency of an antioxidant, including its consumption rate and the efficiency, differed among substrates. With respect to MEs with very low initial oxidation stability, the results showed the advantage of TBHQ over the other tested phenolic antioxidants, while the amine-based formulation showed moderate protection potency in the narrowest range of the applied concentrations. These findings implied that for low quality substrates, molecules with two OH groups per molecule could be linked with higher antioxidant potency than those with one OH-group and two bulky *tert*-butyl groups or amines. Higher unsaturation levels required higher dosages of the antioxidants to induce the 8-h threshold. Concerning the high quality MEs, BHT and 2,2′-methylene-bis-(4-methyl-6-*tert*-butylphenol) ensured $SF \approx 2$ with the lowest quantity among the four tested phenolic formulations. As the main differences between SoSuME and SoSuME-EN were the impurity levels (alkali and earth-alkali metals, phosphorous, water) and the initial *IP*, the different trends observed for them in the *IP* increase with the antioxidant levels ("peaks" in the curve of *IP vs*. dosage observed for SoSuME, but not for SoSuME-EN) suggested the possible role of impurities in pro-oxidative behavior of the higher antioxidant dosages under accelerated oxidation, but this observation requires further studies.

SUPPLEMENTARY MATERIAL

Additional data are available electronically at the pages of journal website: https://www.shd-pub.org.rs/index.php/JSCS/index, or from the corresponding author on request.

Acknowledgement. This paper was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (451-03-68/2020-14/200134; 451-03-9/2021-14/200134; 451-03-68/2020-14/200117)**.**

ИЗВОД

ЕФИКАСНОСТ АНТИОКСИДАНТНИХ ФОРМУЛАЦИЈА У МЕТИЛ ЕСТРИМА МАСНИХ КИСЕЛИНА РАЗЛИЧИТИХ КАРАКТЕРИСТИКА

МИЛИЦА РАНКОВ ШИЦАР[1,2], РАДОСЛАВ МИЋИЋ[3], МИЛАН ТОМИЋ[4] и НАТАША ЂУРИШИЋ-МЛАДЕНОВИЋ[1]

[1]Универзитет у Новом Саду, Технолошки факултет Нови Сад, Булевар цара Лазара 1, 21000 Нови Сад, [2]СП Лабораторија а.д., Индустријска 3, 21220 Бечеј, [3]Универзитет у Новом Саду, Технички факултет "Михајло Пупин", Ђуре Ђаковића бб, 23000 Зрењанин и [4]Универзитет у Новом Саду, Пољопривредни факултет, Трг Доситеја Обрадовића 8, 21000 Нови Сад

У раду је процењен утицај шест формулација на побољшање оксидационе стабилности различитих метил-естера масних киселина (ME). Два испитана ME се разликују у нивоима засићености, јер су синтетизовани из различитих сировина: мешавине соје и сунцокретовог уља (SoSuME) и од отпадног уља за кување (WCOME); они не испуњавају услове стандарда EN 14214 у погледу оксидационе стабилности (~0,6 h) и неких примеса. Трећи ME (SoSuME-EN) је у потпуности усаглашен са стандардом EN 14214. Пет антиоксидативних формулација је на бази фенола, са једном или више активних суп-

станци различите молекулске структуре, док је једна на бази амина. Формулације су додате у 7 нивоа (у опсегу садржаја активних супстанци 50-48300 ppm) Стабилност ME је изражена као период индукције, *IP*, одређен Ранцимат методом, и коришћен за израчунавање фактора стабилизације, *SF*, који указује на ефикасност примењене формулације. Формулација која садржи TBHQ била је најефикаснија у погледу најмањих количина и највећих *SF* постигнутих код ME лошијег квалитета. 2,2′-метилен-бис-(4-метил-6-*терц*-бутилфенол) је показао бољу антиоксидациону моћ од аминске формулације, као и фенолних формулација са две терц-бутил групе. Од 4 изабрана фенолна адитива, BHT и 2,2′-метилен-бис-(4-метил-6-терц-бутилфенол) су показали сличну ефикасност у SoSuME-EN (при ~500 ppm постигнут је $SF \approx 2$), док је била потребна двоструко већа количина формулација са мешаним бутилфенолима за постизање истог ефекта.

(Примљено 9. фебруара, ревидирано 8. априла, прихваћено 12. априла 2021)

REFERENCES

1. M. Tomić, N. Đurišić-Mladenović, R. Mićić, M. Simikić, L. Savin, *Fuel* **235** (2019) 269 (https://doi.org/10.1016/j.fuel.2018.07.123)
2. M. J. Ramos, C. M. Fernández, A. Casas, L. Rodríguez, A. Pérez, *Bioresour. Technol.* **100** (2009) 261 (http://doi.org/10.1016/j.biortech.2008.06.039)
3. S. Jain, M. P. Sharma, *Renew. Sustain. Energy Rev.* **14** (2010) 667 (https://doi.org/10.1016/j.rser.2009.10.011)
4. G. Karavalakis, S. Stournas, *Energy Fuels* **24** (2010) 3682 (https://doi.org/10.1021/ef1004623)
5. S. Agarwal, S. Singhal, M. Singh, S. Arora, M. Tanwer, *ACS Sustain. Chem. Eng.* **6** (2018) 11036 (https://doi.org/10.1021/acssuschemeng.8b02523)
6. J. Pullen, K. Saeed, *Renew. Sustain. Energy Rev.* **16** (2012) 5924 (https://doi.org/10.1016/j.rser.2012.06.024)
7. Z. Yaakob, B. N. Narayanan, S. Padikkaparambil, K. Surya Unni, M. P. Akbar, *Renew. Sustain. Energy Rev.* **35** (2014) 136 (https://doi.org/10.1016/j.rser.2014.03.055)
8. R. Kumar Saluja, V. Kumar, R. Sham, *Renew. Sustain. Energy Rev.* **62** (2016) 866 (https://doi.org/10.1016/j.rser.2016.05.001)
9. M. M. Rashed, M. A. Kalam, H. H. Masjuki, H. K. Rashedul, A. M. Ashraful, I. Shancita, A. M. Ruhul, *RSC Adv.* **5** (2015) 36240 (https://doi.org/10.1039/C4RA14977G)
10. K. Varatharajan, D. S. Pushparani, *Renew. Sustain. Energy Rev.* **82** (2018) 2017 (https://doi.org/10.1016/j.rser.2017.07.020)
11. B. Chen, D. McClements, E. Decker, *Crit. Rev. Food Sci. Nutr.* **51** (2011) 901 (https://doi.org/10.1080/10408398.2011.606379)
12. W. Liu, G. Lu, G. Yang, Y. Bi, *Fuel* **242** (2019) 133 (https://doi.org/10.1016/j.fuel.2018.12.132)
13. G. Knothe, K. R. Steidley, *Fuel Process. Technol.* **177** (2018) 75 (https://doi.org/10.1016/j.fuproc.2018.04.009)
14. I. van der Westhuizen, W. W. Focke, *Fuel* **219** (2018) 126 (https://doi.org/10.1016/j.fuel.2018.01.086)
15. M. Serrano, M. Martínez, J. Aracil, *Fuel Process. Technol.* **116** (2013) 135 (https://doi.org/10.1016/j.fuproc.2013.05.011)
16. W. W. Focke, I. Van Der Westhuizen, A. B. L. Grobler, K. T. Nshoane, J. K. Reddy, A. S. Luyt, *Fuel* **94** (2012) 227 (https://doi.org/10.1016/j.fuel.2011.11.061)
17. I. Tasić, M. D. Tomić, A. L. Aleksić, N. Đurišić-Mladenović, F. L. Martinović, R. D. Mićić, *Hem. Ind.* **73** (2019) 103 (https://doi.org/10.2298/HEMIND190117009T)

18. M. Lapuerta, J. Rodríguez-Fernández, A. Ramos, B. Álvarez, *Fuel* **93** (2012) 391 (https://doi.org/10.1016/j.fuel.2011.09.011)
19. M. Mittelbach, S. Schober, *J. Am. Oil Chem. Soc.* **80** (2003) 817 (https://doi.org/10.1007/s11746-003-0778-x).
20. J. Zhou, Y. Xiong, Y. Shi, *Energy Fuels* **30** (2016) 10534 (https://doi.org/10.1021/acs.energyfuels.6b02199)
21. K. Ryu, *Bioresour. Technol.* **101** (2010) S78 (https://doi.org/10.1016/j.biortech.2009.05.034)
22. Z. Yang, B. P. Hollebone, Z. Wang, C. Yang, M. Landriault, *Fuel Process. Technol.* **106** (2013) 366 (https://doi.org/10.1016/j.fuproc.2014.05.033)
23. I. M. Rizwanul Fattah, H. H. Masjuki, M. A. Kalam, M. A. Hazrat, B. M. Masum, S. Imtenan, A. M. Ashraful, *Renew. Sustain. Energy Rev.* **30** (2014) 356 (https://doi.org/10.1016/j.rser.2013.10.026)
24. A. Domingos, E. Saad, W. Vechiatto, H. Wilhelm, L. Ramos, *J. Braz. Chem. Soc.* **18** (2007) 416. (https://doi.org/10.1590/S0103-50532007000200026).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 86 (7–8) S148–S151 (2021) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Efficiency of different additives in the improvement of the oxidation stability of fatty acid methyl esters with different properties

MILICA RANKOV ŠICAR[1,2], RADOSLAV MIĆIĆ[3], MILAN TOMIĆ[4] and NATAŠA ĐURIŠIĆ-MLADENOVIĆ[1]*

[1]University of Novi Sad, Faculty of Technology Novi Sad, Novi Sad, Serbia, [2]SP Laboratorija a.d., Bečej, Serbia, [3]University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia and [4]University of Novi Sad, Faculty of Agriculture, Novi Sad, Serbia

J. Serb. Chem. Soc. 86 (7–8) (2021) 739–752

TABLE S-I. Properties of methyl esters obtained in this study from the soybean and sunflower oils blend (SoSuME and SoSuME-EN) and waste cooking oil (WCOME) in comparison with the properties required by the EN14214 standard

Property	EN 14214	SoSuME	WCOME	SoSuME-EN	Analytical method
Fatty acid methyl ester content, wt. %	min 96.5	86.75	99.75	98.02	GC/FID, EN14103
Linolenic acid methyl ester content, wt. %	max 12	3.28	0.08	3.24	
Polyunsaturated methyl esters (≥4 double bonds) content, wt. %	max 1	<0.02	<0.02	0.19	
Methanol content, wt. %	max 0.2	0.08	0.07	0.08	GC/FID/HSS, EN14110
Na+K, mg kg^{-1}	max 5	31.66	12.54	4.50	ICP/MS, VM/MET 883
Ca+Mg, mg kg^{-1}	max 5	57.47	7.49	2.91	
Water content, mg kg^{-1}	max 500	1029	764	449	Karl Fisher, ISO 12937
Oxidation stability at 110 °C, h	min 8	0.60	0.63	8.88	Rancimat, EN 14112
Cold filter plugging point, °C	Guidance	-13	-12	-12	ASTM D 6371
Acid number, mg KOH g^{-1}	max 0.5	0.39	0.80	0.23	Volumetric, EN 14104
Iodine number, g I_2 / 100 g	120	126	120	115	Volumetric, EN 14111
Phosphorous, mg kg^{-1}	max 4.00	23.03	1.52	2.13	Spectrophotometry, VM/ MET 659
Density at 15 °C, kg m^{-3}	860-900	885	887	885	Areometry, EN ISO 3675

* Corresponding author. E-mail: natasadjm@tf.uns.ac.rs; natasa.mladenovic@uns.ac.rs

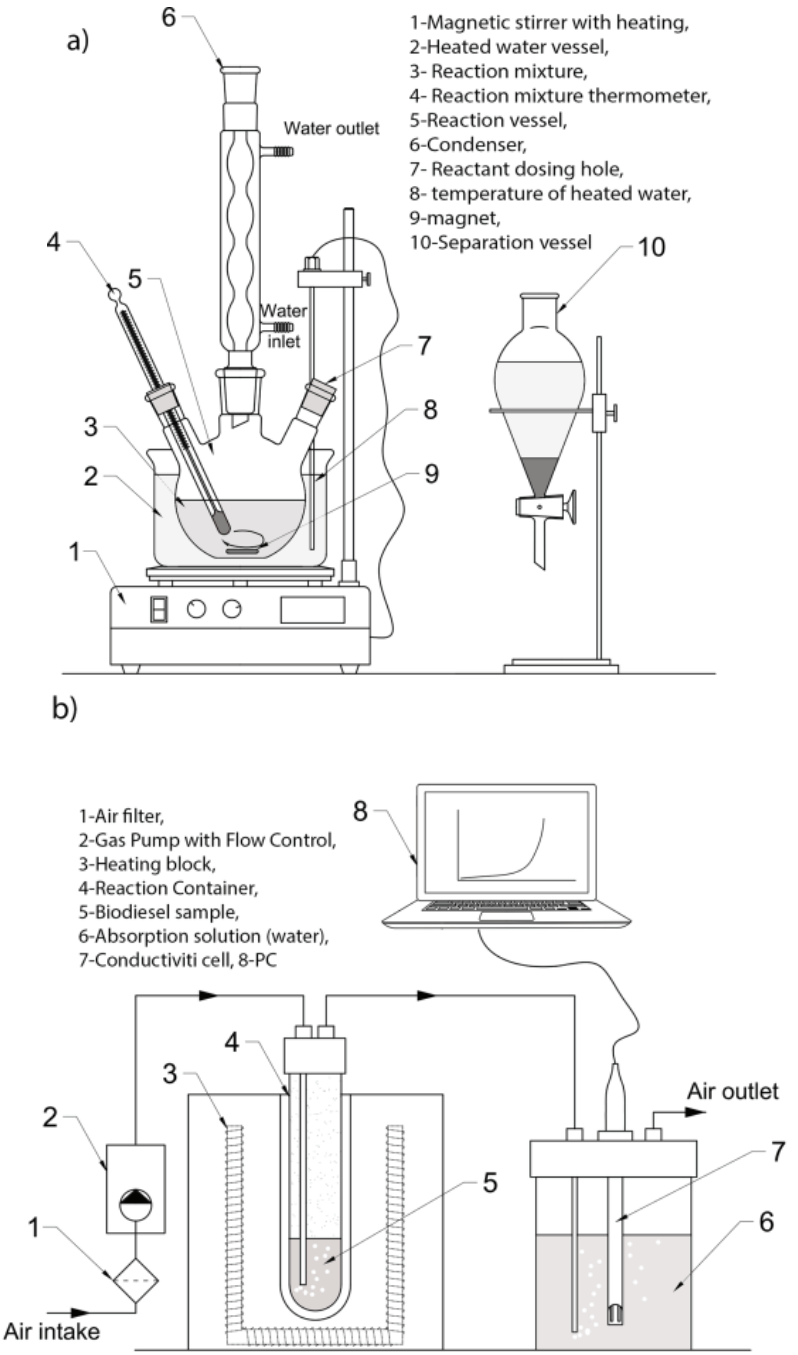


Fig. S-1. Schematic representation of the experimental setup for: (a) production of MEs by transesterification and separation from glycerol, and (b) induction period measurement.

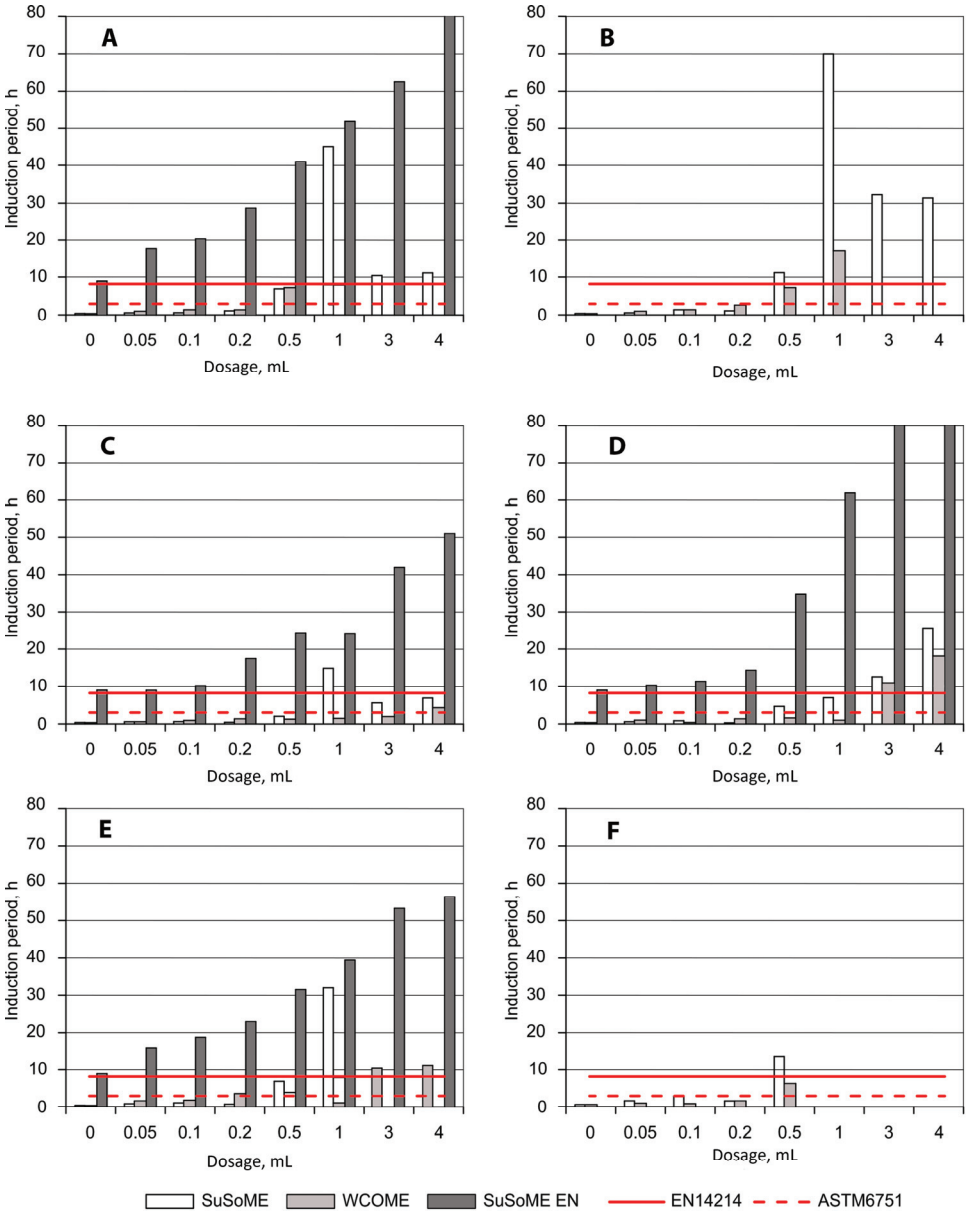


Fig. S-2. Effects of various dosages of antioxidant formulations (marked A-F, Tables III and IV) on the oxidation stability (expressed as induction period) of SoSuME, WCOME and SoSuME-EN.

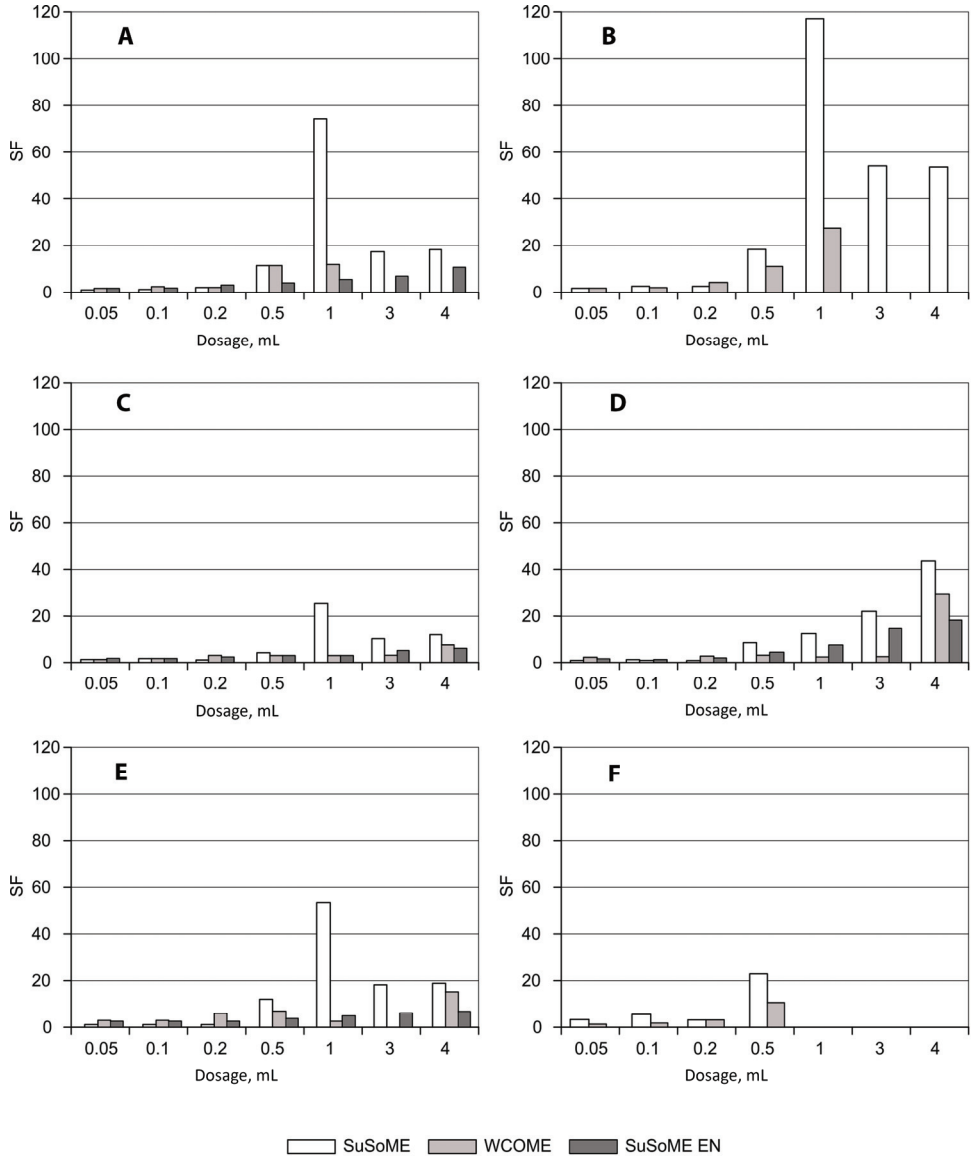


Fig. S-3. Stabilization factors, *SF*, calculated for each tested formulation (A-F, Table III) added in different dosages to SoSuME, WCOME and SoSuME-EN.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. *86* (7–8) 753–766 (2021)
JSCS–5459
Original scientific paper

Ambient air particles: The use of ion chromatography and multivariate techniques in the analysis of water-soluble substances

ŽAKLINA N. TODOROVIĆ[1], JELENA M. RADULOVIĆ[1], IVANA D. SREDOVIĆ IGNJATOVIĆ[2], LJUBIŠA M. IGNJATOVIĆ[3]*# and ANTONIJE E. ONJIA[4]

[1]*Anahem Laboratory, Mocartova 10, 11160 Belgrade, Serbia,* [2]*University of Belgrade, Faculty of Agriculture, Nemanjina 6, 11080 Belgrade, Serbia,* [3]*University of Belgrade, Faculty of Physical Chemistry, Studentski trg 12–16, 11000 Belgrade, Serbia and* [4]*University of Belgrade, Faculty of Technology and Metallurgy, Karnegieva 4, 11000 Belgrade, Serbia*

(Received 26 August, revised 22 November, accepted 30 November 2020)

Abstract: Seventeen water-soluble substances (of sodium, ammonium, potassium, magnesium, calcium, formate, methanesulfonate, glyoxylate, chloride, nitrite, nitrate, glutarate, succinate, malate, malonate, sulfate and oxalate) in 94 samples of particle matter in the ambient air, collected over ten months, in a suburb of Belgrade (Serbia), were determined by ion chromatography. To apportion the sources of the air pollution, the log-transformed data were processed by applying multivariate techniques. Principal component and factor analysis identified three main factors controlling the data variability: stationary combustion processes with the highest loadings of oxalate, malonate and malate; landfill emission and secondary inorganic aerosol characterized by high levels of ammonium, nitrate and sulfate; a contribution of mineral dust composed of magnesium, calcium and chloride. The hierarchical cluster analysis pointed out a differentiation of the samples into five groups belonging to different variables inputs. For the classification of ambient air samples using nine selected ions, the recognition ability of linear discriminant analysis, *k*-nearest neighbors, and soft independent modeling of class analogy were 87.0, 94.6, and 97.8 %, respectively. Time-series analysis showed that the traffic emission is more pronounced in winter in contrast to the mineral dust influence, while the effect of waste combustion exhibits no trend.

Keywords: organic acids; pollution sources; PCA; emission factors; time-series.

INTRODUCTION

Over the last few decades, the analysis of ambient air particles by means of different analytical techniques and different receptor modeling methods has received increasing attention.[1,2]

* Corresponding author. E-mail: ljignjatovic@ffh.bg.ac.rs
# Serbian Chemical Society member.
https://doi.org/10.2298/JSC200826077T

The complimentary use of inductively coupled plasma mass spectrometry (ICP-MS) for trace elements, and gas chromatography–mass spectrometry (GC––MS) for volatile and semi-volatile organics has been the traditional approach to the chemical analysis of ambient air samples.[3,4] Of many other analytical techniques that are being used, high-performance liquid chromatography coupled to mass spectrometry (HPLC–MS) due to its good sensitivity and resolution power is receiving more and more popularity in trace organics quantification.[5] On the other hand, ion chromatography (IC) gained an important role as a technique of choice for the analysis of water-soluble ions in atmospheric air particles.[1,2,6–9]

Since water-soluble substances make up a third or more of the particulate mass of ambient air particles, these ions have been studied intensively in recent years.[10–14] Those investigations provided essential information on the physicochemical properties of ambient air particles, the mechanism of their formation and source of pollution. The surface acidity and hygroscopicity properties of ambient air particles govern their chemical and optical behaviors. These particles in ambient air mainly consist of inorganic ions.

Besides the mentioned inorganic matter, an organic component, *i.e.*, a considerable amount of low-molecular weight (LMW) organic acids, are present in the water-soluble part of a typical ambient air particle.[15,16] Initially, this was proved by means of GC–MS, which is a suitable technique after a derivatization step that has to be performed in order to obtain these acids in their volatile form.[17] The derivatization process is often troublesome, and it was shown that the application of IC for LMW organic acids analysis is more advantageous than GC–MS.[18]

An optimized IC method can readily be applied for the analysis of major cations (sodium, ammonium, potassium, magnesium, calcium) and inorganic anions (chlorides, nitrates, sulfates, oxalates, phosphates, bromides), even if they are present at a very low level in aqueous solution.[19–21]

Here, an isocratic IC in the cation-exchange mode was applied for the analysis of five common cations, whereas twelve inorganic anions and LMW organic acids were simultaneously analyzed by anion-exchange gradient IC.

In order to obtain a better insight into ambient air particles at the studied site, the IC results were evaluated by different multivariate pattern recognition methods.

Multivariate analysis of water-soluble inorganic matter, measured together with PAHs, metals and reactive oxygen species, has been extensively studied.[2,10,22] As mentioned above, this approach requires the use of several analytical techniques. Although there are many studies in which inorganic ions together and LMW organic acids were analyzed, there is a lack of literature on pattern recognition studies on such datasets.[6,8–10] In addition, different conventional multivariate techniques produce different results, even when used together.[23] In this work, common non-supervised pattern recognition methods: principal

component analysis (PCA), factor analysis (FA), and hierarchical cluster analysis (HCA) were compared, while supervised pattern recognition methods, such as linear discriminant analysis (LDA), *k*-nearest neighbors (*k*-NN) and soft independent modeling of class analogy (SIMCA), were evaluated for their recognition ability. These methods assign the output values according to some specific machine learning algorithms. The method is supervised, provided that a training dataset is used. On the other hand, in an unsupervised method, it is possible to perform pattern recognition even if no dataset is present.[23]

EXPERIMENTAL

The sampling procedure is presented in the Supplementary material to this paper.

Standard solutions. Individual stock standard solutions (1000 mg dm^{-3}) were supplied by AccuStandard, Merck, or Fluka. For some ions, *i.e.*, most organic acid salts (glyoxylates, methanesulfonates, glutarates, succinates, malonates), there are no commercially available stock standard solution (1000 mg dm^{-3} or less). In this case, standards were made by dissolving an appropriate amount of an analyte in deionized water. The stock standards were stored at 4 °C, whereas composite working standards at lower analyte concentrations, in a select range similar to the expected analyte concentrations in the samples, were prepared before analysis. The concentrations were calculated from 5 points standard calibration curves using the external standard method. As a quality control measure, duplicates, blanks and spiked samples were also analyzed.

Anion IC system. A Dionex DX-500 ion chromatography system consisting of a gradient pump, ED40 electrochemical detector operating in the conductivity mode (detector current setting: 50 mA), injection valve fitted with 400 μL sample loop, ASRS-ultra self-regenerating suppressor (external water mode), an AS11 (250 mm×4 mm) analytical column, an AG11 (50 mm×4 mm) guard column, and a Peaknet 5.1 chromatography workstation. A mixture of HPLC grade methanol and hydroxide ions, at a flow rate of 2.0 ml min^{-1}, was used as the eluent. The following gradient program was applied: 1) time: 0.0 min, KOH: 0.175 mM, CH_3OH: 5.0 %, 2) time: 6.0 min, KOH: 0.575 mM, CH_3OH: 5.0 %, 3) time: 26.0 min, KOH: 8.575 mM, CH_3OH: 17.0 % and 4) time: 30.0 min, KOH: 10.175 mM, CH_3OH: 5.0 %.

Cation IC system. A Dionex DX-500 ion chromatograph with a conductometric detector, a CS12 (250 mm×4 mm) analytical column, a CG12 (50 mm×4 mm)×guard column, a CSRS--ultra self-regenerating suppressor operating in the recycling mode, and a Peaknet 5.1 chromatography workstation. The eluent was 20 mM methanesulfonic acid (isocratic). The injection sample loop was 100 μL. The system was operated at room temperature with a flow rate of 1.0 cm^3 min^{-1}.

Data analysis. The statistical data analyses in this work was performed by means of Minitab, SPSS and PLS Toolbox for Matlab software packages. Details of PCA, FA, HCA, *k*-NN and SIMCA procedure are given elsewhere.[23]

RESULTS AND DISCUSSION

Ion chromatography analysis

Typical IC chromatograms of ambient air particles extract are shown in Figs. 1 and 2. In the cation-exchange mode, five common cations (sodium, ammonium, potassium, magnesium and calcium) are easily separated in less than 10

min. In some samples, lithium-ion was detected at trace level, but it was not quantified in this study. Ordinarily, lithium is eluted as the first ion and fairly separated from sodium.

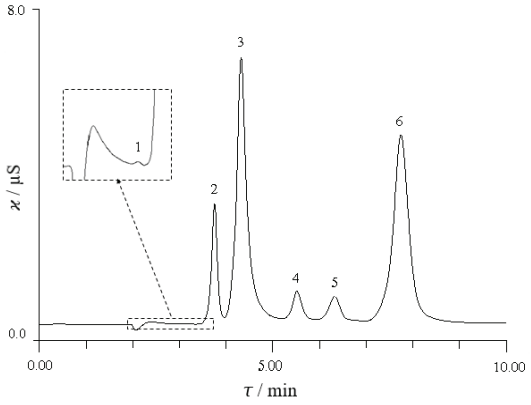


Fig. 1. Typical cation chromatogram of an ambient air particles sample extract. Column: Dionex CS12. Eluent: 20 mM methanesulfonic acid (isocratic). Peaks: 1) lithium, 2) sodium, 3) ammonium, 4) potassium, 5) magnesium, 6) calcium.

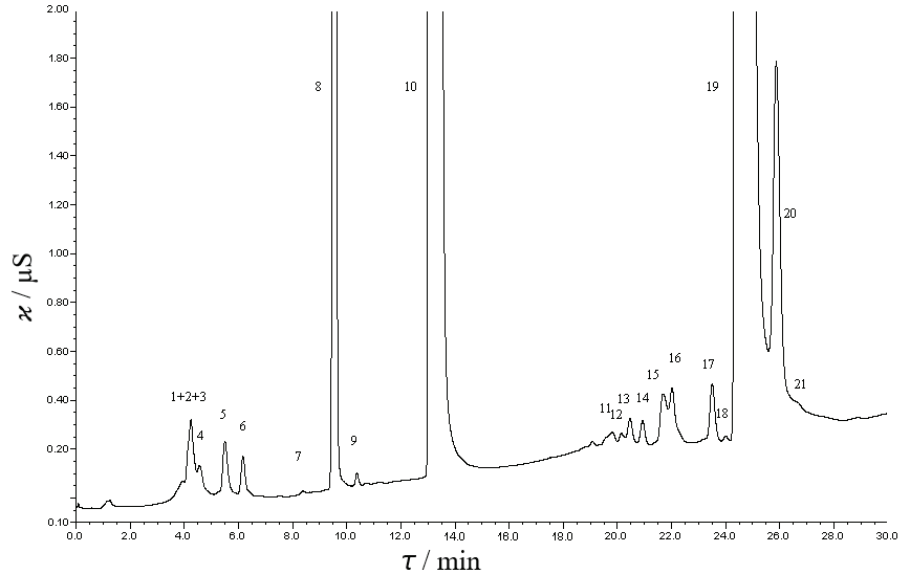


Fig. 2. Anion chromatogram of the same ambient air particles sample extract, as in Fig. 1. Column: Dionex AS11. Eluent: see the experimental part for details. Peaks: 1) fluoride, 2) lactate, 3) acetate, 4) glycolate, 5) formate, 6) methanesulfonate, 7) glyoxylate, 8) chloride, 9) nitrite, 10) nitrate, 11) glutarate, 12) pimelate, 13) succinate, 14) malate, 15) carbonate, 16) malonate, 17) sulfite, 18) fumarate, 19) sulfate, 20) oxalate, 21) azelate.

For any source apportioning method to be applied in an air pollution study, the number of measured chemical substances (variables) should be as high as possible. Preferably, there should be no missing values in the data set.

A few studies have made the apportionment of sources of air pollution even when using only major ions data.[13,14,24] Although, it is known that not enough tracer information is carried by major ions only. Furthermore, LMW organic acids are shown to be important tracers for different particle emission sources.[15]

In this work, for the separation of anions, a multi-step gradient provided simultaneous separation of both inorganic ions (chloride, nitrite, nitrate, sulfate) and ionic form of LMW organic acids (formate, glyoxylate, methanesulfonate, glutarate, succinate, malate, malonate, oxalate). In addition to these anions, several other anions (bromide, phosphate, sulfite, benzoate, caproate, pimelate, fumarate, phthalate, acetate, lactate, glycolate, pyruvate, tartrate, azelate, *etc*. that were detected in some samples, are not included in this statistical evaluation. The reason was incomplete measurement results for these rarely detected anions in most samples so that some of these anions were often either below the quantification limit of the applied method (pyruvate, tartarate, bromide, phtalate, caproate, fumarate) or there were co-elution problems (fluoride/lactate/acetate, acetate/ /glycolate) or problem with their stability (sulfite, azelate). Thus, 17 ions were quantified in all samples.

Descriptive statistics

After the completion of IC measurements, the final data matrix 94×17 (observations *x* variables) was formed. In the first step of data examination, the Ryan–Joiner Test was used for testing the hypothesis of normal distribution.[23] In all statistical tests, the significance level was 0.05. As shown for oxalate in Fig. S-2a (Supplementary material), the plotted points do not approximate a straight line. The departures from linearity indicate that the actual data distribution deviates from a normal distribution. In contrast, the log-transformed data (Fig. S-2b) appear to be well fitted on the normal probability plot with R and p values well above 0.99 and 0.1, respectively. All other variables follow a similar distribution; however, for the sake of brevity, their graphs are not displayed. For further statistical evaluation, the transformed values were used. The data matrix was tested for outliers by applying the Rosner Test.[23] Two samples with high sulfate and ammonia concentrations were discarded as outliers. Therefore, explorative multivariate analysis was realized on the final 92×17 data matrix.

The arithmetic means, minimum values, maximum values, and standard deviations for the concentrations of all substances in the studied ambient air particles were calculated (Table I). It can be seen that the variables could be considered as major ions (nitrate, sulfate, ammonium, sodium, chloride, calcium), trace ions (formate, glyoxylate, methanesulfonate, nitrite, glutarate, succinate, malate, malonate), and those ones at medium to a high level (potassium, magnesium, oxalate). Skewness is a measure of the lack of symmetry varied between 0.5 and 2.5, which means that this data distribution is more or less right-skewed.

Since uniform distribution for different analytes is very rare, kurtosis values are used to estimate the size of tailings. This dataset has both low and high kurtosis, *i.e.*, ranged from –0.5 to 7.6.

TABLE I. Descriptive statistics of the data for the studied water-soluble substances in ambient air particles in Mirijevo outskirt of Belgrade in the period 09.09.2013–25.06.2014

No.	Variable	c / ng m^{-3}						
		Mean	Median	Skewness	Kurtosis	St. deviation	Max.	Min.
1.	Sodium	617	581	1.1	1.1	254	1382	229
2.	Ammonium	2186	1707	1.0	0.4	1598	6700	304
3.	Potassium	167	154	1.1	1.3	67	408	69
4.	Magnesium	144	123	1.2	0.9	86	406	27
5.	Calcium	929	685	1.4	1.8	780	3781	94
6.	Formate	16.7	14.8	1.0	0.2	8.2	38.0	5.2
7.	Methanesulfonate	25.5	18.7	1.9	4.1	23.3	118	2.8
8.	Glyoxylate	9.5	7.9	1.0	0.2	5.7	25.8	1.2
9.	Chloride	1217	869	1.0	0.3	1030	4282	54.9
10.	Nitrite	20.9	18.3	0.5	–0.5	11.2	47.7	4.4
11.	Nitrate	3765	3405	0.8	0.2	2019	9837	634
12.	Glutarate	7.2	5.4	2.0	5.2	5.8	34.0	0.9
13.	Succinate	18.0	15.0	2.4	7.6	14.6	86.7	4.0
14.	Malate	15.8	8.9	2.5	6.4	19.0	94.2	1.6
15.	Malonate	19.5	9.2	1.8	2.5	25.7	108	0.1
16.	Sulfate	3230	2873	1.0	0.5	1736	8221	1105
17.	Oxalate	83.2	79.6	0.7	–0.2	42.7	193	17.1

Unsupervised pattern recognition

These data were subjected to several common multivariate pattern recognition methods, which are described in detail elsewhere.[23,25,26]

Principal component analysis (PCA) is a method of removing the highly inter-correlated nature of variations in atmospheric concentrations. Thurston and Spengler suggest that the dataset must have many more samples than analytes if stable PCA results are to be derived.[27] Their requirements: $n \geq m + 50$ (where n is the number of samples; m is the number of analytes) is fulfilled in this work.

When PCA was applied to the autoscaled data matrix, three principal components (PCs) appeared to account for 84 % of the variance in the data. The eigenvalue and cumulative variance as a function of the number of PCs are shown in Fig. S-3 (Supplementary material). In this case, the descriptor space dimensionality is reduced to three since only the first three PCs were retained. These PCs meet the Kaiser criterion to have the eigenvalues higher than one.[23] All other PCs show eigenvalues lower than one, with an exponential trend in the decrease.

In order to highlight the relations between variables, the data set of concentration measurements was subjected to factor analysis (FA). This method was

applied with PCA extraction of the correlation matrix and varimax orthogonal rotation. The FA method identified three significant factors (Table II): the first (24.6 % of variance) comprises organic ions oxalate, malonate, malate with high loadings and glutarate, succinate, methanesulfonate, formate, nitrite, and sulfate with relatively lower loadings.

TABLE II. Varimax rotated factor loadings (values > 0.5 are in bold)

No.	Variable	Factor 1	Factor 2	Factor 3
1.	Sodium	0.257	0.223	0.192
2.	Ammonium	0.225	**0.909**	–0.210
3.	Potassium	0.140	0.146	0.288
4.	Magnesium	–0.255	–0.217	**0.871**
5.	Calcium	–0.396	–0.183	**0.853**
6.	Formate	0.449	0.214	–0.234
7.	Methanesulfonate	**0.501**	0.381	–0.199
8.	Glyoxylate	0.092	0.431	0.138
9.	Chloride	–0.453	0.195	**0.687**
10.	Nitrite	0.493	0.323	–0.359
11.	Nitrate	0.068	**0.953**	–0.029
12.	Glutarate	**0.585**	0.491	–0.299
13.	Succinate	**0.596**	**0.650**	–0.243
14.	Malate	**0.837**	0.165	–0.396
15.	Malonate	**0.800**	0.241	–0.302
16.	Sulfate	0.396	**0.791**	–0.017
17.	Oxalate	**0.842**	0.295	–0.295

All organic ions, except succinate and glyoxylate, have higher loadings in this factor than in the second one. Oxalate and malonate are markers for combustion processes.[8,28] Hence, this factor identifies exposure to emissions from fuel combustion. Although, other biogenic sources could also contribute to the production of these two acids. Mainly composed of ammonium, nitrate, sulfate, and succinate, the second factor accounts for a further 22.7 % of the variance. The association between sulfate and nitrate suggests an anthropogenic origin. Waste landfill emissions that occur in the vicinity of the sampling site, as well as a secondary inorganic aerosol, might be responsible for this factor. The third factor with 16.8 % of the variance was composed of magnesium, sodium, chloride, and with a lower loading of potassium. This source profile strongly suggests that their concentrations are affected by mineral dust from both natural and anthropogenic sources. Except for potassium (loading = 0.536) and calcium (loading = 0.565), no significant loading value was obtained for any variable on Factors 4 and 5, which are responsible for 6.9 and 6.6 % of the total variance, respectively. It is relevant that some ions, such as methanesulfonate, nitrite, glutarate, sulfate, succinate, obtained rather high scores both on the first and second factor, indicating more than one possible significant emission source.

Hierarchical cluster analysis (HCA) was made on the variables that had previously been standardized.[29] An amalgamation rule employed the Ward method to link the objects, while for the measure the proximity between samples, the squared Euclidean distance was selected.[30] The dendrogram is shown in Fig. 3. As a result of applying HCA to the factor score matrix, the ambient air particles split into two main groups. Furthermore, each group could be subdivided into three or two sub-groups, respectively. There are, therefore, five clusters (C1, C2, C3, C4 and C5), the composition of which depends on the inputs of variables.

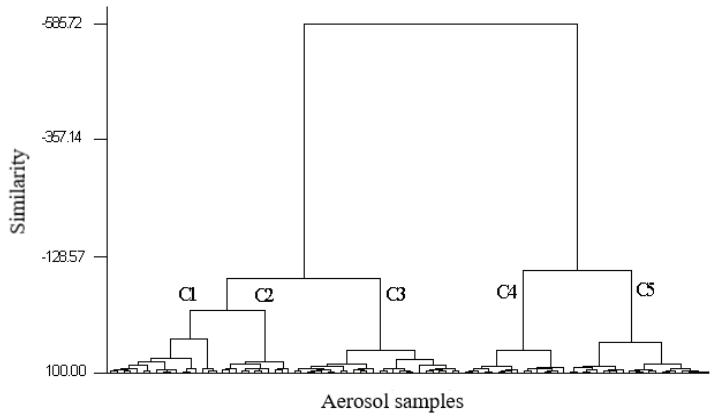


Fig. 3. Dendrogram of cluster analysis of ambient air particles.

Cluster C1 consists of the samples characterized by high score values for the combustion factor, whereas the second group of samples (C2) has high levels in the landfill emission and secondary inorganic aerosol factor. The samples with medium–low scores of both the combustion and waste emission/secondary inorganic aerosol factors (C1 and C2) are grouped in cluster C3. These three groups (C1, C2, and C3) form one of the two main groups that all samples can be split. The fourth cluster (C4) comprises samples with the highest levels of the variables (magnesium, calcium, chloride) for the mineral dust factor, while medium–high score values for the previous factor (C4) are characteristic for the samples in C5.

Supervised pattern recognition

In order to evaluate the possibility of classification of ambient air particles into these five groups, three common supervised pattern recognition techniques were employed, linear discriminant analysis (LDA), k-nearest neighbors (*k*-NN), and soft independent modeling of class analogy (SIMCA).[31,32]

First, the samples were classified using all variables. Then, nine variables with the first three highest loadings in each Factor in Table II (malate, malonate, oxalate, ammonium, nitrate, sulfate, magnesium, calcium, chloride) were selected

as key features for classification. In all cases, the leave-one-out method of cross--validation was used.

Linear discriminant analysis (LDA) is a very simple supervised classification technique that reduces the dimensionality of the features in a similar way as PCA. LDA makes projection axes for separating the classes by minimizing the within-class variance and maximizing the between-class variance. In this case, LDA produced a classification success rate of 83.7 % for all samples and all variables. The separation of ambient air particles by two discriminant functions is shown in Fig. 4. When performing LDA with the reduced number of variables, 87.0 % of grouped samples were correctly classified.

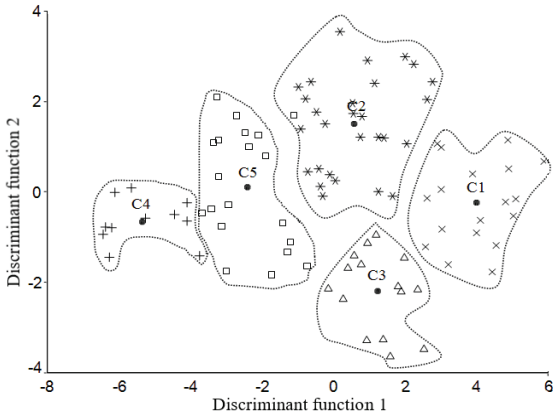


Fig. 4. Discriminant analysis of ambient air particles (solid symbols are the group centroids).

One of the simplest classifiers is the *k*-NN algorithm. It classifies an input feature vector by determining the k closest training vectors according to a distance metric. Then, this input vector is assigned to a class to which the majority of its k nearest neighbors belong. Here, the metric employed was the Euclidean distance.

As in the case of LDA, *k*-NN and SIMCA were performed using both all and selected variables. The obtained results are given in Table III. Using both methods, the recognition ability for the five classes was highly satisfactory; the percentages of classification obtained were 93.5–100 %.

TABLE III. Recognition and prediction ability (%) of LDA, *k*-NN and SIMCA to classify ambient air particles; (A) – all variables, (B) – nine selected variables

Ability for:		Method		
		LDA	*k*-NN	SIMCA
Recognition	(A)	83.7	93.5	100
	(B)	87.0	94.6	97.8
Prediction	(A)	52.8	94.6	78.3
	(B)	56.5	92.4	73.9

Another approach that can give a more realistic error rate is to split the data into two parts, one part to create the discriminant function or model, and the other part as a testing set. Randomly splitting an equal number of samples (46 samples) into these two sets, the *k*-NN method produces good prediction power (94.6 and 92.4 %), whereas LDA (52.8 and 56.5 %) and SIMCA (78.3 and 73.9 %) give a much lower percentage of the correctly classified samples than previously, for both all variables and nine variables, respectively.

Time-series analysis

To examine the trends of the concentration changes from each emission source, over a ten-month period (fall 2013 to summer 2014), the concentrations of indicator ions (with the highest loadings) for various emission sources were plotted by the moving averages for the samples in Fig. 5.

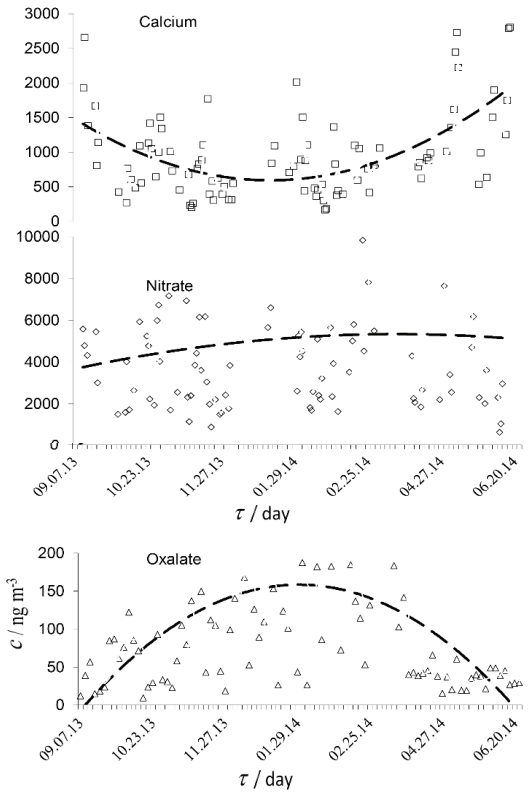


Fig. 5. Variation of ambient air particles concentrations of oxalate, nitrate, and calcium in the Mirijevo outskirt of Belgrade during the period 2013 to 2014. The dot lines indicate the quadratic fit.

The time-trends of these variables were also fitted by quadratic curves. Variables in Factor 1 show an affinity to change with time, increasing from low values in the fall to a maximum during the winter, followed by a slight decrease through the spring in 2014. This trend resembled that reported elsewhere.[21,33]

The variables in Factor 3 showed the opposite trends over the same time period, whereas those in Factor 2 have no trend. Similar seasonal behavior for the ions comprising the latter two factors was obtained by other authors.[22,24,33] A significant fluctuation is present with all analyzed variables. No monthly periodicity was revealed. The reason for seasonal variation could be found in the local climate and weather conditions. Hence, dust resuspension (Factor 1) is more intense in the summer, with many dry days with no snow coverage. On the other hand, the fuel combustion (Factor 3) is more intense in winter days. No seasonal changes in the municipally waste landfill activities was observed, causing no trend in Factor 2. Generally, these results could be explained by the occurrence of atmospheric transport on a local scale.

CONCLUSIONS

The combined use of ion chromatography and multivariate pattern recognition methods can provide valuable information for an ambient air particle research study. Some water-soluble substances, such as bromide and phosphate, which are traditionally measured in ambient air particle samples by NAA, XRF, PIXE, ICP-MS, and used as important source markers, can also be quantified by IC, but the applied IC separation in this work should to be additionally optimized. The obtained data chemical set showed the characteristic of log-normal distribution, while multivariate methods revealed a distinctive grouping of variables and observations. Of the three common supervised pattern recognition methods (LDA, *k*-NN and SIMCA), the classification prediction ability was only satisfactory when applying k-NN. Some variables do not contribute to the classification. The emissions from combustion processes, landfill emission together with secondary inorganic aerosol, and a contribution of mineral dust were identified as the three dominant factors that have an effect on pollutant levels at the studied site. The landfill emission/secondary inorganic aerosol undergoes no change in seasons, whereas the other two other factors show different behavior during winter and summer. In general, a considerable high level of sulfate, nitrate, ammonium, and LMW organic acids in the ambient air particles at the site pointed to a strong anthropogenic influence.

SUPPLEMENTARY MATERIAL

Additional data are available electronically at the pages of journal website: https://www.shd-pub.org.rs/index.php/JSCS/index, or from the corresponding author on request.

Acknowledgements. This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Contract No. 451-03-68/2020-14/200135).

ИЗВОД

ЧЕСТИЦЕ У АМБИЈЕНТАЛНОМ ВАЗДУХУ: ПРИМЕНА ЈОНСКЕ ХРОМАТОГРАФИЈЕ И МУЛТИВАРИЈАНТНИХ ТЕХНИКА АНАЛИЗЕ ВОДЕНО-РАСТВОРНИХ СУПСТАНЦИ

ЖАКЛИНА Н. ТОДОРОВИЋ[1], ЈЕЛЕНА М. РАДУЛОВИЋ[1], ИВАНА Д. СРЕДОВИЋ ИГЊАТОВИЋ[2], ЉУБИША М. ИГЊАТОВИЋ[3] и АНТОНИЈЕ Е. ОЊИА[4]

[1]Анахем Лабораторија, Моцартова 10, 11160 Београд, [2]Универзитет у Београду, Пољопривредни факултет, Немањина 6, 11080 Београд, [3]Универзитет у Београду, Факултет за физичку хемију Студентски трг 12–16, 11 000 Београд и [4]Универзитет у Београду, Технолошко–металуршки факултет, Карнегијева 4, 11000 Београд

Методом јонске хроматографије одређен је садржај 17 супстанци растворљивих у води (које садрже натријум, амонијум, калијум, магнезијум, калцијум, формате, метансулфонате, глиоксилате, хлориде, нитрите, нитрате, глутарате, сукцинате, малате, малонате, сулфате и оксалате) у укупно 94 узорка честица у амбијенталном ваздуху, који су узорковани током десет месеци у предграђу Београда. Да би се одредили извори загађења ваздуха, добијени резултати су најпре трансформисани у њихове логаритме, а потом анализирани применом мултиваријантних техника. Анализа главних компоненти и факторска анализа идентификовали су три главна фактора који контролишу варијабилност података: процес стационарног сагоревања са највећим оптерећењем оксалата, малоната и малата, емисија са депоније и секундарни неорганski аеросол које карактеришу високи нивои амонијум-јона, нитрата и сулфата и допринос минералне прашине која се састоји од материја магнезијума, калцијума и хлорида. Примена хијерархијске анализе кластера указала је на диференцијацију узорака у пет група које припадају различитим профилима аналита. За класификацију узорака амбијенталног ваздуха помоћу девет одабраних јона, способност препознавања методе линеарне дискриминантне анализе, методе *k*-најближих суседа и методе меког независног моделовања аналогне класе износили су 87,0, 94,6 и 97,8 %. Анализа временских серија показала је да је фактор емисије из саобраћаја у току зиме израженији, за разлику од фактора минералне прашине. Ефекат стационарног сагоревања не показује тренд.

(Примљено 26. августа, ревидирано 22. новембра, прихваћено 30. новембра 2020)

REFERENCES

1. A. Agarwal, A. Satsangi, A. Lakhani, K. M. Kumari, *Chemosphere* **242** (2020) 125132 (https://doi.org/10.1016/j.chemosphere.2019.125132)
2. M. Dos Santos, D. L. Dawidowski, E. Gautier, P. Smichowski, *Microchem. J.* 91 (2009) 133 (https://doi.org/10.1016/j.microc.2008.09.001)
3. K. Swami, C. D. Judd, J. Orsini, K. X. Yang, L. Husain, *Fresenius J. Anal. Chem.* **369** (2001) 63 (https://doi.org/10.1007/s002160000575)
4. N. N. Naing, K. B. Yeo, H. K. Lee, *J. Chromatogr., A* **1612** (2020) 460646 (https://doi.org/10.1016/j.chroma.2019.460646)
5. J. Xu, J. He, H. Xu, D. Ji, C. Snape, H. Yu, C. Jia, C. Wang, J. Gao, *RSC Adv.* **8** (2018) 34136 (https://doi.org/10.1039/C8RA04991B)
6. T. Fosco, M. Schmeling, *Environ. Monit. Assess.* **130** (2007) 187 (https://doi.org/10.1007/s10661-006-9388-1)
7. A. Arias, R. J. N. Bettencourt da Silva, M. F. G. F. C. Camoes, C. M. R. R. Oliveira, *Talanta* **104** (2013) 10 (https://doi.org/10.1016/j.talanta.2012.11.024)
8. S. Karthikeyan, S. W. See, R. Balasubramanian, *Anal. Lett.* **40** (2007) 793 (https://doi.org/10.1080/00032710601017920)

9. Y. I. Tsai, L. Y. Hsieh, T. H. Weng, Y. C. Ma, S. C. Kuo, *Anal. Chim. Acta* **626** (2008) 78 (https://doi.org/10.1016/j.aca.2008.07.041)
10. T. Zhang, J. J. Cao, X. X. Tie, Z. X. Shen, S. X. Liu, H. Ding, Y. M. Han, G. H. Wang, K. F. Ho, J. Qiang, W. T. Li, *Atmos. Res.* **102** (2011) 110 (https://doi.org/10.1016/j.atmosres.2011.06.014)
11. L. Mu, L. Zheng, M. Liang, M. Tian, X. Li, D. Jing, *Aerosol Air Qual. Res.* **19** (2019) 2396 (https://doi.org/10.4209/aaqr.2019.03.0109)
12. A. Mihajlidi-Zelić, D. Đorđević, D. Relić, I. Tošić, Lj. Ignjatović, M. Stortini, A. Gambaro, *Open Chem.* **13** (2015) 245 (https://doi.org/10.1515/chem-2015-0010)
13. D. Đorđević, T. Mihajlidi-Zelić, D. Relić, Lj. Ignjatović, J. Huremović, A.M. Stortini, A. Gambaro, *Atmos. Environ.* **46** (2012) 309 (https://doi.org/10.1016/j.atmosenv.2011.09.057)
14. H. Wang, J. An, M. Cheng, L. Shen, B. Zhu, Y. Li, Y. Wang, Q. Duan, A. Sullivan, L. Xia, *Chemosphere* **148** (2016) 526 (https://doi.org/10.1016/j.chemosphere.2016.01.066)
15. K. Kawamura, S. Bikkina, *Atmos. Res.* **170** (2016) 140 (https://doi.org/10.1016/j.atmosres.2015.11.018)
16. A. Röhrl, G. Lammel, *Chemosphere* **46** (2002) 1195 (https://doi.org/10.1016/S0045-6535(01)00243-0)
17. K. Kawamura, *Anal. Chem.* **65** (1993) 3505 (https://doi.org/10.1021/ac00071a030)
18. K. Kawamura, L. A. Barrie, D. Toom-Sauntry, *Atmos. Environ.* **44** (2010) 5316 (https://doi.org/10.1016/j.atmosenv.2010.08.051)
19. Ž. Todorović, Lj. Rajaković, A. Onjia, *Hem. Ind.* **71** (2017) 27 (https://doi.org/10.2298/HEMIND151107014T)
20. Ž. Todorović, Lj. Rajaković, A. Onjia, *J. Serb. Chem. Soc.* **81** (2016) 661 (https://doi.org/10.2298/JSC150927022T)
21. D. Čičkarić, I. Deršek-Timotić, A. Onjia, Lj. Rajaković, *J. Serb. Chem. Soc.* **70** (2005) 995 (https://doi.org/10.2298/JSC0507995C)
22. E. I. Tolis, D. E. Saraga, G. Z. Ammari, E. I. Gkanas, T. Gougoulas, C. C. Papaioannou, A. K. Sarioglou, E. Kougioumtzidis, A. Skemperi, J. G. Bartzis, *Cent. Eur. J. Chem.* **12** (2014) 643 (https://doi.org/10.2478/s11532-014-0531-5)
23. A. Onjia, *Chemometric approach to the experiment optimization and data evaluation in analytical chemistry*, University of Belgrade – Faculty of Technology and Metallurgy, Belgrade, 2016 (ISBN 978-86-7401-338-0)
24. H. Wang, D. Shooter, *Atmos. Environ.* **35** (2001) 6031 (https://doi.org/10.1016/S1352-2310(01)00437-X)
25. S. Dragović, A. Onjia, *Appl. Radiat. Isot.* **65** (2007) 218 (https://doi.org/10.1016/j.apradiso.2006.07.005)
26. S. Razić, A. Onjia, *Am. J. Enio. Viticult.* **61** (2010) 506 (https://doi.org/10.5344/ajev.2010.10002)
27. G. D. Thurston, J. D. Spengler, *Atmos. Environ.* **19** (1985) 9 (https://doi.org/10.1016/0004-6981(85)90132-5)
28. S. R. Souza, P. C. Vasconcellos, L. R. F. Carvalho, *Atmos. Environ.* **33** (1999) 2563 (https://doi.org/10.1016/S1352-2310(98)00383-5)
29. L. Slavković, B. Škrbić, N. Miljević, A. Onjia, *Environ. Chem. Lett.* **2** (2004) 105 (https://doi.org/10.1007/s10311-004-0073-8)
30. J. H. Ward, *J. Am. Stat. Assoc.* **58** (1963) 236 (https://doi.org/10.1080/01621459.1963.10500845)

31. B. Kowalski, C. Bender, *J. Am. Chem. Soc.* **94** (1972) 5632 (https://doi.org/10.1021/ja00771a016)
32. S. Wold, *Pattern Recog.* **8** (1976) 127 (https://doi.org/10.1016/0031-3203(76)90014-5)
33. M. Possanzini, V. Di Palo, A. Cecinato, C. Balducci, *Anal. Lett.* **34** (2001) 957. (https://doi.org/10.1081/AL-100103605).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 86 (7–8) S152–S155 (2021) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Ambient air particles: The use of ion chromatography and multivariate techniques in the analysis of water-soluble substances

ŽAKLINA N. TODOROVIĆ[1], JELENA M. RADULOVIĆ[1], IVANA D. SREDOVIĆ IGNJATOVIĆ[2], LJUBIŠA M. IGNJATOVIĆ[3]* and ANTONIJE E. ONJIA[4]

[1]Anahem Laboratory, Mocartova 10, 11160 Belgrade, Serbia, [2]University of Belgrade, Faculty of Agriculture, Nemanjina 6, 11080 Belgrade, Serbia, [3]University of Belgrade, Faculty of Physical Chemistry, Studentski trg 12–16, 11000 Belgrade, Serbia and [4]University of Belgrade, Faculty of Technology and Metallurgy, Karnegieva 4, 11000 Belgrade, Serbia

J. Serb. Chem. Soc. 86 (7–8) (2021) 753–766

The site. A total of 94 daily samples of ambient air particles were collected between 9th of September 2013 and 25th of June 2014 at Mirijevo, the outskirts of Belgrade, Serbia (44°47'36.7"N 20°32'09.3"E). Located to the south–east of the city, the sampling site is characterized by the following environment: (1) a suburban residential and commercial area (near the commercial area, retail stores, houses, and restaurants), (2) a heating plant located about 0.6 km from the sampling site; (3) the city landfill Vinča, located 4.7 km from the sampling site; (4) a petrochemical industry and oil refinery located in the industrial area of Pančevo at a distance of 11.5 km from the sampling site (Fig. S-1).

Sample handling. Ambient air particles were collected on quartz filters with a TCR Tecora (Echo Hi-Vol) sampler. All filters had a diameter of 102 mm, and the sampling was performed continuously for 24 h at a flow rate of 225 dm^3 min^{-1}. After the sampling, the filters were stored in a freezer until extraction. A portion of each loaded filter (1.0 cm^2 punch) was transferred to a clean 15 cm^3 polystyrene tube and extracted with 10 ml of deionized water (Millipore, 18.2 MΩ). All extracted filters were mechanically shaken on a Fisher Scientific model 361 orbital shaker for 20 min, and the resulting extracts were analyzed immediately by ion chromatography using the full-loop injection mode. A recovery experiment was performed by spiking a blank filter with an appropriate standard solution, while the extraction efficiency was checked by triple extraction of the same part of the filter. In both experiments, the yield was in the range 80–100 % and the *RSD* was less than 10 % for all ions.

* Corresponding author. E-mail: ljignjatovic@ffh.bg.ac.rs

Fig. S-1. Map of the sampling site and potential air pollution sources: (2) the heating plant in the vicinity; (3) the city landfill Vinča; (4) the petrochemical industry in Pančevo.

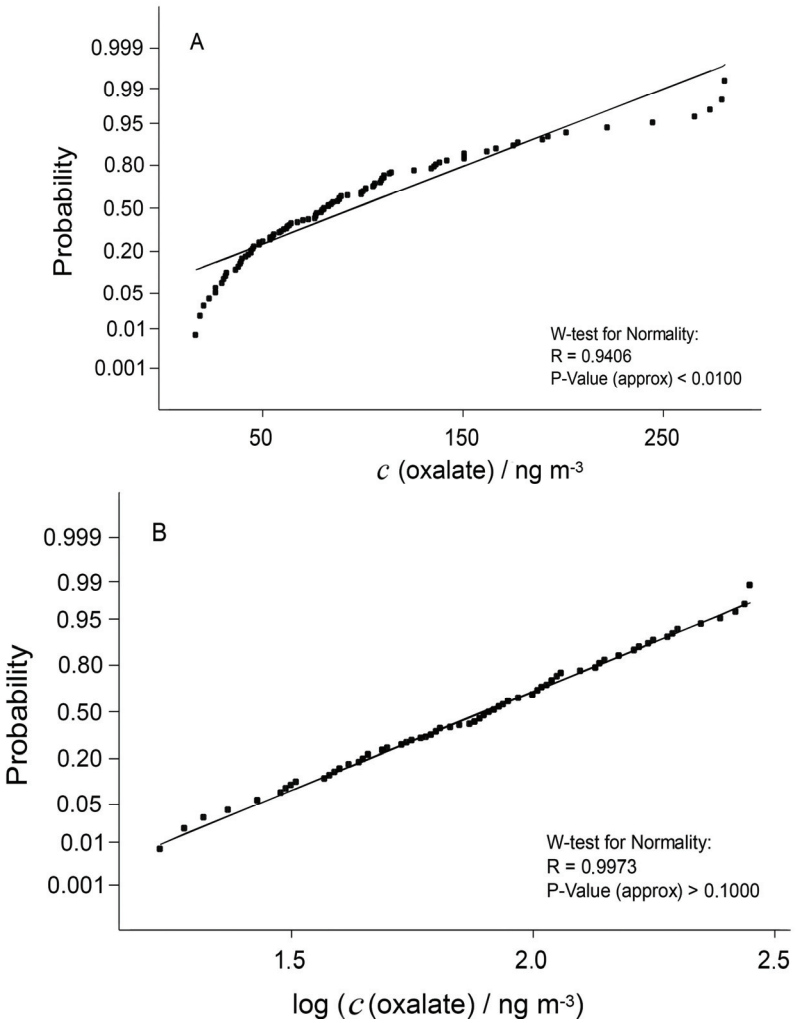


Fig. S-2. Normal (A) and lognormal (B) probability plots for the distribution of oxalate in atmospheric ambient air particles.

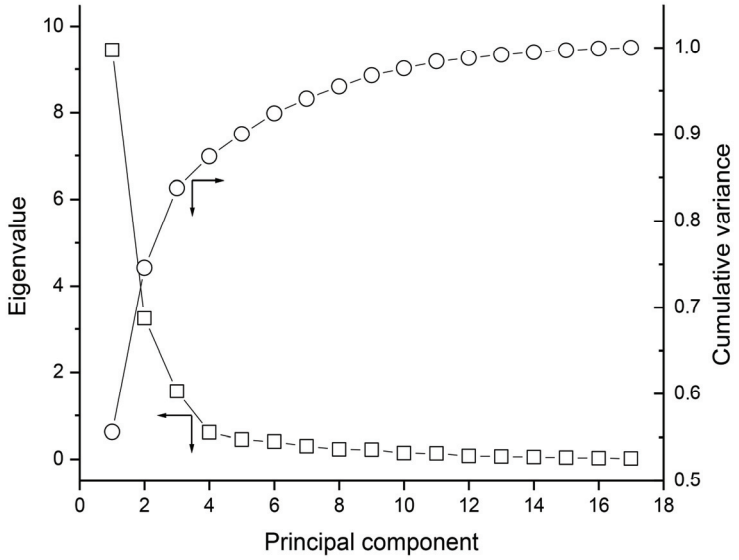


Fig. S-3. Eigenanalysis of the correlation matrix.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc.86 (7–8) 767–780 (2021)
JSCS–5460
Original scientific paper

A prudent approach for the removal of copper (II) and cadmium (II) ions from aqueous solutions using indigenous *Mactra aequisulcata* shells

SYED MUHAMMAD REHAN ULLAH[1,2]*, ERUM ZAHIR[2] and MUHAMMAD ASIF ASGHAR[3]

[1]*Department of Chemical Oceanography & Marine Environment, National Institute of Oceanography, St-47, Block-1 Clifton, Karachi−75600, Sindh−74200, Pakistan,* [2]*Department of Chemistry, University of Karachi, Karachi−75270, Sindh−74200, Pakistan and* [3]*Food and Feed Safety Laboratory, Food and Marine Resources Research Centre, PCSIR Laboratories Complex, Shahrah-e-Salimuzzaman Siddiqui, Off University Road, Karachi−75280, Sindh−74200, Pakistan*

(Received 28 December 2020, revised 22 March, accepted 4 April 2021)

Abstract: In this study, the raw seashells of *Mactra aequisulcata*, of the class Bivalvia and the phylum Mollusca were employed as an adsorbent to remove copper (II) and cadmium (II) metal ions from aqueous solution. The characterization of shells was performed using FTIR, EDX and SEM, BET isotherm and point of zero charge (pH_{pzc}). Batch experiments were performed to assess various factors on the biosorption efficiency. Maximum adsorption of both metals ions conveniently found at 0.4 g L^{-1} adsorbent dose, pH 5 and at 303 K. Maximum biosorption capacities for Cu^{2+} and Cd^{2+} were 59.57 and 38.12 mg g^{-1}, respectively. The Langmuir isotherm model was found to be the best fit for the acquired equilibrium data. Thermodynamic and kinetic parameters showed that the process was feasible, exothermic and followed pseudo-first order.

Keywords: adsorbent; heavy metals; pollutants; remediation; wastewater.

INTRODUCTION

The environmental condition for humans and other living beings worldwide has become more inhospitable than it has ever been before.[1] The industrial activities have posed series threats to the environment. Technological advancements and rapid consumption have led to various genres of pollution.[2] The discharge of industrial effluent bearing heavy metals causes great adversities due to their presence and ability to accumulate and toxic effects on living things.[3] As far as environmental threat is concerned, toxic heavy and radionuclides metals are of main concern and ought to be nullified. Toxic metal ions can cause minor

* Corresponding author. E-mail: rehan-syed@outlook.com
https://doi.org/10.2298/JSC201228028U

discomforts as well as life threatening illnesses by irreversible damage of vital body systems. Eco-toxicologically, most toxic metals are cadmium, chromium, lead and mercury. Heavy metals can bioaccumulate as well as biomagnify in the food chain. So, their toxicities are more noticeable in the animals that are in high trophic levels.[4] Currently, various techniques are used to remediate waste water from these toxic metals. Adsorption processes have been proven to be a better technique as compared with other techniques, namely electrochemical, oxidation, membrane filtration, ozonization and complexation. It has also greater importance due to operational easiness, relatively low cost, and simplicity of design in the decoloration process.[5] Mollusca shells are mostly composed of aragonite and calcite type calcium carbonate, that's around 95–99 % of entire shell quantity, and also contains a small amount of organic matter.[6] The seashells as an agent to remediate wastewater from heavy metals have also been employed, rather than geological calcium carbonate.[7] Biosorption augmentation through pre-treatment processes, such as with organic chemicals,[5] thermal processing[8] and acid treatment[9] have also been successfully opted. *Mactra aequisulcata*, species of mussels, found in marine habitat, has aragonite as its major constituent[10] and biogenic aragonite is considered to be a suitable biosorbent for heavy metal adsorption[9]. Moreover, the aforementioned species has not been reported for the remediation of heavy metals in the aqueous medium as well from industrial effluents, and not any mussels' species locally, for that matter. Considering the aforementioned approaches, the study was designed to assess the biosorption capability of pulverized *Mactra aequisulcata* shells to remove Cu^{2+} and Cd^{2+} from the aqueous phase. Optimal adsorption process was assessed by varying in adsorbent dose, interaction time, metals initial concentration, pH and temperature. The biosorption mechanism was conducted in terms of isotherm models, kinetics and thermodynamics studies.

EXPERIMENTAL

Materials

The chemicals employed in this research were of analytical grade. Millipore Q5 Direct water purification system was used to prepare deionized water with conductivity of 0.06 μS cm^{-1}. *Mactra aequisulcata* (Class Bivalvia, Phylum Mollusca, 100 units) were sampled from the Clifton Beach, Karachi city of Pakistan and stored in a pre-washed plastic ziplocks. The seashells were washed with tap water, then immersed into 0.5 % of HCl solution for 30 min, for the removal of impurities adhered on the surface of shells. The shells were washed with DI-H_2O repeatedly, drying in an oven at 80 °C and pulverized to powder form using 500 g electric grain grinder mill. The pulverized adsorbent was kept in polyethylene airtight vessels in ambient condition. The preparation of working stock solutions of Cu^{2+} and Cd^{2+} was done with $CuSO_4$ $5H_2O$ and $CdCl_2$ (>99 %, Merck, Germany), respectively. The solution pH were altered to the desired level with 0.1 M HCl and NaOH.

Biosorption of metal ions

For metals biosorption, a known quantity of pulverized seashell was used with 50 mL of Cu^{2+} and Cd^{2+} solutions in an Erlenmeyer flask, separately. The outcome of the adsorbent dose (0.2, 0.4, 0.6, 0.8 and 1.0 g L^{-1}), initial metals ions concentration (5, 10, 25, 50 and 100 mg L^{-1}), initial pH of metal solutions (2, 3, 4, 5 and 6), interaction time (10, 20, 30, 40 and 60 min) and temperature (303, 308, 313, 318 and 323 K) were studied. After each experiment, the solutions were filtered using Whatman filter paper no. 42. Controlled sample without the addition of biosorbent, was also quantified to assess the interference of filter papers, and no interference was observed. The filtrate was used to determine the heavy metal ions concentrations using atomic absorption spectrometer. The percentage of metal adsorption and maximum biosorption capacity, q in mg g^{-1}, were measured using the following equations:

$$ADS = 100\frac{C_0 - C_e}{C_0} \quad (1)$$

$$q = \frac{V(C_0 - C_e)}{m} \quad (2)$$

where, C_0 and C_e in mg L^{-1}, correspond to the initial and equilibrium concentrations of the metals ions, respectively, m is the mass of the biosorbent in mg and V is volume of the solution in L.

Instrumentation

The Cu^{2+} and Cd^{2+} in the samples were quantified with atomic absorption spectrometry (PinAACle 900T, Perkin Elmer). Eutech Cyber Scan PCD 650 was used for the pH measurements. An electric grain grinder mill was used to grind the seashells. The availability of functional groups in the ground seashell samples was determined with Fourier transform infrared spectrophotometry (FTIR, Perkin Elmer, Singapore). The surface area was evaluated through Autosorb-1 Quantachrome, Asiawin. The morphology of the surface and diameter of ground seashell were examined using SEM (JSM 6380A, JEOL Ltd, Japan). The deduction of elemental composition was carried with EDX spectroscopy (EX-54175 JMU, JEOL Ltd, Tokyo, Japan).

Determination of pH_{pzc}

The point of zero charge (pH_{pzc}) of the pulverized *M. aequisulcata* shells was determined by taking six separate Erlenmeyer flasks containing 50 ml 0.1 M NaCl having pH 2, 4, 6, 8, 10 and 12. The pH of the solutions were adjusted by adding 1 M HCl or 1 M NaOH. To each flask, 0.1 g pulverized shell powder was added and then the capped flasks were shaken at 150 rpm for 4 h. The suspensions were filtered with the Whatman filter paper no. 42. The difference between initial pH (pH_i) and the final pH (pH_f) was plotted with pH_i. The intersection of the plot and pH_i (x-axis) was deemed as the pH_{pzc}.

Biosorption isotherm models

For isotherm studies, an adsorbent dose of 0.4 g L^{-1} of pulverized seashell was used with 50 mL of Cu^{2+} and Cd^{2+} solutions in Erlenmeyer flasks, separately; with initial adsorbate concentrations of 5, 10, 25, 50 and 100 mg L^{-1}, initial pH of 5, at 303 K, and interactions time of 20 and 30 min for Cu^{2+} and Cd^{2+}, respectively. The isotherm explains how metal ions concentrations and sorbed metal amount interact on the solid surface at their equilibrium phases. In this work, the biosorption specifics were studied by Langmuir, Freundlich and Temkin isotherm models. Langmuir isotherm details about the monolayer adsorption and states the abs-

ence of lateral interaction between the adsorbate molecules. The nonlinear Langmuir isotherm model is expressed as follows:[11]

$$q_e = \frac{q_m b C_e}{1 + b C_e} \tag{3}$$

where, C_e in mg L^{-1} signifies the equilibrium concentration of the metal ions, q_e in mg g^{-1}, corresponds to the equilibrium adsorption capacity, q_m in mg g^{-1} and b in L mg^{-1}, are the Langmuir constants relating to adsorption capacity and energy, respectively.

The favourability of adsorption process can be predicted by the assessment of Langmuir essential characteristic denoted as the separation factor, R_L and its equation is illustrated as:[12]

$$R_L = \frac{1}{1 + b C_0} \tag{4}$$

The Freundlich model, nonlinear form is applicable to multilayer adsorption, relating to energetic surface heterogeneity and can be assessed from the following equation:[13]

$$q_e = K_f C_e^{1/n} \tag{5}$$

where K_f / mg g^{-1} (L $mg^{-1})^{1/n}$ and $1/n$ correspond to the Freundlich constants of adsorption capacity and heterogeneity factor, respectively.

The following equation illustrates the Temkin isotherm nonlinear model (Eq. 6):

$$q_e = B \ln (A_T C_e) \tag{6}$$

where, B in J mol^{-1}, is the constant of heat of sorption at equilibrium and A_T in L g^{-1}, being the binding constant. The Temkin isotherm illustrates that the heat of adsorption of every molecule in the layer decreases directly with increased surface coverage.[14] Dubinin–Radushkevich model considers the adsorption of subcritical vapours on microporous adsorbent followed by the mechanism of pore filling, and is applicable to detail the mechanism of adsorption on heterogeneous surfaces with the Gaussian energy distribution.

The nonlinear Dubinin–Radushkevich model is represented as[15]:

$$q_e = q_s e^{-K_{DR} \varepsilon^2} \tag{7}$$

$$\varepsilon = RT \ln \left(1 + \frac{1}{C_e} \right) \tag{8}$$

where, q_s in mg g^{-1} is the Dubinin–Radushkevich isotherm model constant relating to maximum adsorption capacity; K_{DR}, mol^2 kJ^{-2}, is Dubinin–Radushkevich isotherm constant; ε is the Polanyi potential, R / J mol^{-1} K^{-1}; and T / K are the gas constant and absolute temperature, respectively.

Biosorption kinetics

For kinetics studies, an adsorbent dose of 0.4 g L^{-1} of pulverized seashell was used with 50 mL of Cu^{2+} and Cd^{2+} solutions in Erlenmeyer flasks, separately; with initial metal ions concentration of 100 mg L^{-1}, initial pH of 5, at 303 K, and over the period of 10, 20, 30, 40 and 60 min. The two most important nonlinear kinetic models: *i*) pseudo-first order (Eq. (7)) and *ii*) pseudo-second order (Eq. (8)) are described by the equations as follows:

$$q_t = q_e (1 - e^{-k_1 t}) \tag{9}$$

$$q_t = \frac{q_e^2 k_2 t}{q_e k_2 t + 1} \quad (10)$$

k_1 / min^{-1} and k_2 / g mg^{-1} min^{-1} are the rate constants of pseudo-first order and pseudo-second order, respectively; and q_t /mg g^{-1} is the amount of metal ion adsorbed at any given time t.

Biosorption thermodynamics

The nature of biosorption system was studied with thermodynamics analysis through van't Hoff equation. The following equations were applied for the deduction in the change in Gibbs energy, ΔG / kJ mol^{-1}, entropy, ΔS / kJ mol^{-1} K^{-1} and enthalpy, ΔH / kJ mol^{-1}:

$$\ln K_c = \left(\frac{\Delta S}{R}\right) - \left(\frac{\Delta H}{R T}\right) \quad (11)$$

$$\Delta G = RT \ln K_c \quad (12)$$

where, K_c represents the constant of thermodynamic equilibrium and R corresponds to the Universal gas constant and T / K is the absolute temperature.

RESULTS AND DISCUSSION

Analysis and characterization of seashell

The Cu^{2+} and Cd^{2+} concentrations in the seashells were found to be well below the lower limit of quantitation (0.1 and 0.01 mg L^{-1} for Cu^{2+} and Cd^{2+}, respectively) of the utilized method. The results signified that the biosorbents would not interfere during the biosorption process as they do not have heavy metals in them. The functional groups in seashell powder sample were deduced by FTIR spectroscopy as illustrated in (Fig. 1). The various peaks appeared in the spectrum at 1789, 1455, 1076, 854, and 726 cm^{-1} and considered the specific features of CO_3^{2-} in $CaCO_3$. The aragonite content of $CaCO_3$ in the sample was confirmed with the observance of a signal at 1076 cm^{-1}.[16]

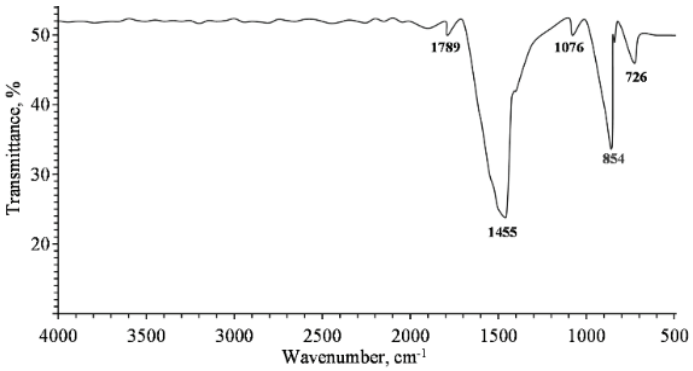


Fig. 1. FTIR spectrum of *M. aequisulcata* seashell powder.

The BET isotherm demonstrated that the biosorbent bears a small BET specific surface area (1.77 m^2 g^{-1}) and median pore width of 134.1 nm. Similarly, researchers[17,18] have also deducted that $CaCO_3$-rich biosorbents possess

low S_{BET} values. Hence, here, the surface area cannot be considered as a determining factor for the biosorption of cadmium and copper[9]. The morphology of the surfaces of pulverized, Cu^{2+} and Cd^{2+} laden seashell powders, were observed using SEM analysis and presented in Fig. 2a–c. It was evident that the surface of seashell powder (Fig. 2a) was heterogeneously distributed with slight grains and spaces. The copper ions SEM image (Fig. 2b) shows some sort of similarity with the adsorbent SEM image with tightly packed particles, and the presence of secondary solids on the adsorbent surface, signifying the occurrence of surface precipitation. The similarity with the SEM image of the adsorbent could possibly be due to precipitation of Cu^{2+} with bicarbonates in the aqueous medium since the $CaCO_3$ dominant adsorbent tend to leach its constituents ions in the solvent medium, of which carbonates transform into bicarbonates.[19] The SEM image of post-biosorption of Cd^{2+} (Fig. 2c) can be categorized as powder layered, showing precipitation on the surface. Cadmium adsorption on calcium carbonate compounds usually follows surface precipitation since ionic radii of the divalent calcium and cadmium are similar, as shown in the SEM image.[18]

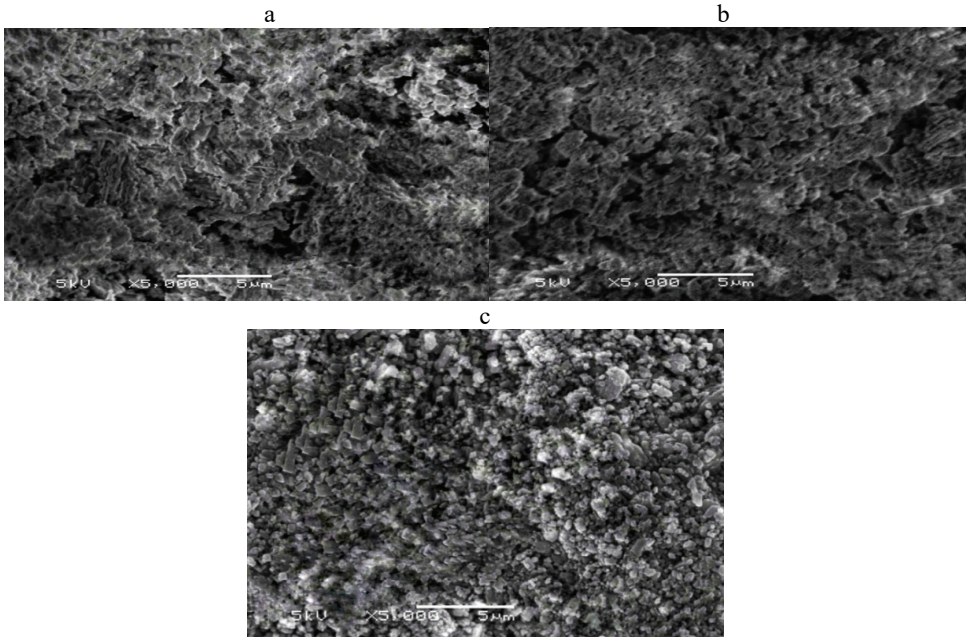


Fig. 2. SEM images of: a – seashell powder, b – Cu^{2+} and c – Cd^{2+} laden seashell powder.

The pre and post adsorption elemental profiles of seashell powder are shown in Fig. 3a–c. The seashell powder EDX spectra (Fig. 3a) showed carbon, oxygen and calcium were the major constituents. The calcium-loaded adsorbent spectra (Fig. 3b) showed the addition of copper in the elemental profile signifying the

adsorption on the adsorbent, and the slight decrease in weight and atomic percentage of calcium indicate the surface precipitation of Cu^{2+}. The Cd^{2+} laden adsorbent EDX spectra (Fig. 3c) illustrated a small change in the elemental profiles of calcium and other elements, with only the addition of cadmium signal, indicating the aggregate formation, *i.e.*, $(Cd,Ca)CO_3$.[9]

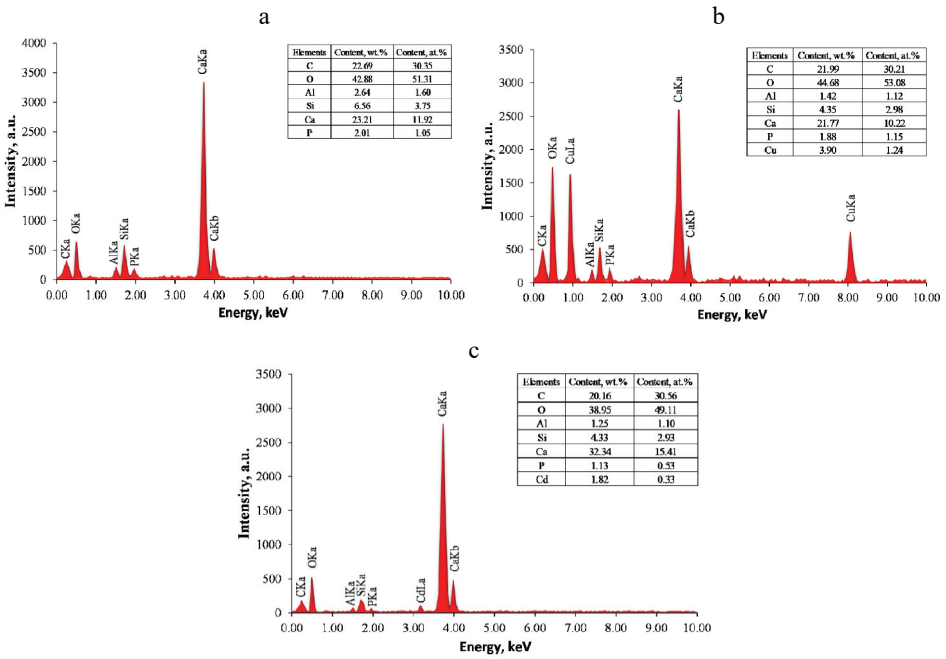


Fig. 3. EDX spectra of: a – seashell powder, b – Cu^{2+} and c – Cd^{2+} laden seashell powder.

The pH_{pzc} value of the adsorbent was found to be 9.25, signifying that the adsorbent surface was mostly positive. This result suggests that positively charged Cu^{2+} and Cd^{2+} couldn't easily adsorb onto the adsorbent through ion exchange and thus complexation or surface precipitation could be considered as the primary adsorption mechanism.[20]

Adsorption study of metal ions using seashell

The Cu^{2+} and Cd^{2+} adsorption was performed with *M. aequisulcata* seashell. The biosorption efficacy was augmented by altering the metal ions concentrations, temperature of the process, pH and contact time.

Interaction time. Fig. 4 illustrates the plot of contact time, *t* in min, and percentage adsorption of Cu^{2+} and Cd^{2+} with *M. aequisulcata* shells. The plot shows that the metal uptake was rapid and the increase in biosorption capacity in the initial 20 min for Cu^{2+} and 30 min for Cd^{2+}, afterward no considerable variation was observed until 60 min. The fast process was due to involvement of the

biosorbent external surface in the metal ions adsorption. The next phase was slower as several of the external places were engaged.[2]

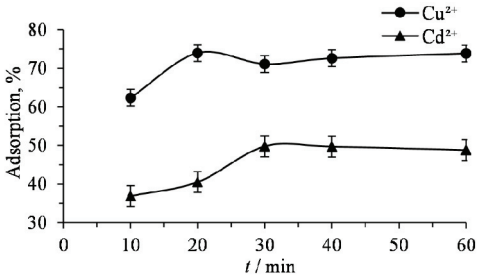


Fig. 4. Effect of contact time (t) on adsorption of Cu^{2+} and Cd^{2+}.

pH. A principal parameter for the assessment of the behavior of adsorbent––adsorbate is evaluation of the effect that changing pH confers onto adsorption. This was carried out from pH value 2 to 6. Fig. 5 is the plot between pH and percentage adsorption of Cu^{2+} and Cd^{2+} using pulverized seashells. The plots show that the metal ions uptake was very low at pH 2 but biosorption capacity increased till pH 5.0 and was constant after that, till pH 6.0 for both metals ions. The lower biosorption capacities were found in intense acidic pH ranges for both metal ions, reason being the protonation of the active places of the biosorbent in acidic media, therefore, decrease in metal ions biosorption was observed.[21]

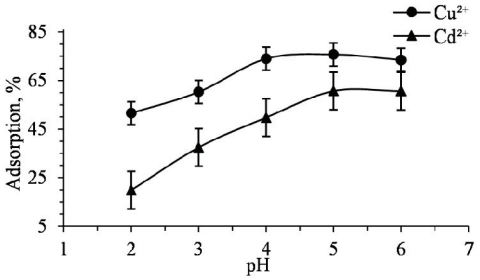


Fig. 5. Effect of initial pH on adsorption of Cu^{2+} and Cd^{2+}.

Adsorbent dose. Fig. 6 is the plot of adsorbent dose, g L^{-1}, and percentage adsorption of Cu^{2+} and Cd^{2+} using pulverized seashells. The plot shows that inc-

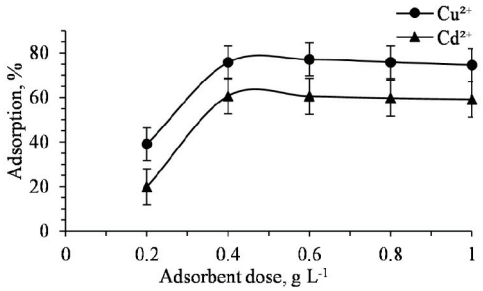


Fig. 6. Effect of adsorbent dose, on adsorption of Cu^{2+} and Cd^{2+}.

rease in adsorbent dose from 0.2–1.0 g L^{-1} led to increase in biosorption efficiency, however, increase in adsorbent dose after 0.4 g L^{-1} resulted in the decline in biosorption efficiency, this is due to the unsaturation of biosorption sites, particle interfaces, *i.e.*, accumulation and rarer energetic places at greater biosorbent doses.[2]

Biosorption isotherm models

The values obtained from calculating the equations (Eqs. (3–8)) are shown in Table I.

TABLE I. Langmuir, Freundlich, Temkin and Dubinin–Radushkevich isotherm models for the biosorption of Cu^{2+} and Cd^{2+} using *Mactra aequisulcata* seashells

Model	Parameter	Cu^{2+}	Cd^{2+}
	q_{exp} / mg g^{-1}	59.57	38.12
Langmuir isotherm	R^2	0.960	0.976
	q_{cal} / mg g^{-1}	59.98	40.14
	b_L / L mg^{-1}	0.227	0.37
Freundlich isotherm	R^2	0.864	0.801
	K_f / mg g^{-1} (L $mg^{-1})^{1/n}$	19.04	16.85
	n	3.69	4.86
Temkin isotherm	R^2	0.900	0.886
	A_T / L g^{-1}	2.25	7.76
	B / J mol^{-1}	231.22	286.56
Dubinin–Radushkevich isotherm	R^2	0.850	0.866
	q_{cal} / mg g^{-1}	46.00	36.10
	K_{DR} / mol^2 kJ^{-2}	1 x 10^{-6}	4 x 10^{-7}

Whereas, the biosorption isotherms are presented in Figs. 7 and 8 for Cu^{2+} and Cd^{2+}. The equilibrium statistics were more suited to the Langmuir isotherm when compared to Freundlich, Temkin and Dubinin–Radushkevich isotherm

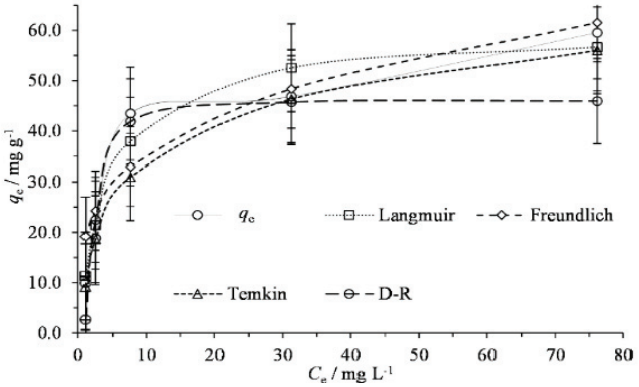


Fig. 7. Langmuir, Freundlich, Temkin and Dubinin–Radushkevich isotherm models for Cu^{2+} on seashell (*Mactra aequisulcata*) powder.

since the regression coefficient, R^2 values for Cu^{2+} and Cd^{2+} acquired by the Langmuir isotherm model were greater than other isotherm models, indicating monolayer adsorption and chemisorption.[2] The experimental q_{max} values were observed as 59.98 and 40.14 mg g^{-1} for Cu^{2+} and Cd^{2+}, respectively, according to the Langmuir isotherm model (Table I). These values were close with the biosorption capacities as calculated by Eq. (1) and found as 59.57 and 38.12 mg g^{-1} for Cu^{2+} and Cd^{2+}, respectively. The R_L results were found to be 0.042 for Cu^{2+} and 0.027 for Cd^{2+}. The R_L values of over 0 and less than 1, for both metals ions, signify the favourability of biosorption process and Langmuir isotherm model.[12]

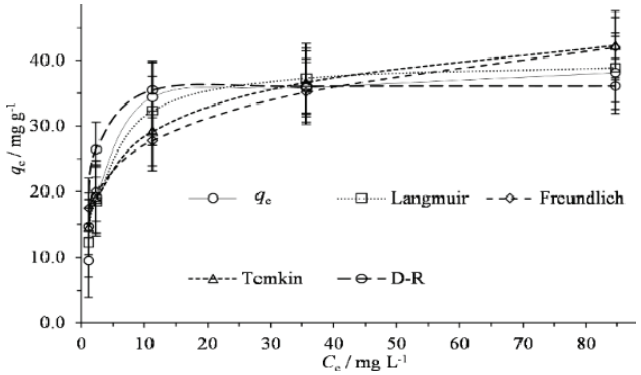


Fig. 8. Langmuir, Freundlich, Temkin and Dubinin–Radushkevich isotherm models for Cd^{2+} on seashell (*Mactra aequisulcata*) powder.

Biosorption kinetics

The biosorption rate is an essential factor to assess its reaction route. The adsorption data were assessed with the kinetic models using their respective equations (Eqs. (9) and (10)). The coefficients of correlation and calculated parameters are illustrated in Table II. The data (Table II) showed that pseudo-first order kinetic model described the adsorption processes for Cu^{2+} and Cd^{2+} better than the pseudo-second order model. The R^2 values of the pseudo-first order model were 0.961 and 0.970 and k_1 values were 0.160 and 0.08 min^{-1} corresponding to Cu^{2+} and Cd^{2+}, respectively. The small values of k_1 for both metals ions refer to readily occupied adsorbent sites. The kinetic pseudo-first order model proposes fast adsorption in the initial stage and limiting step afterwards, referring to chemisorption. The metal ions uptake was rapid in the first 20 min for Cu^{2+} and 30 min for Cd^{2+} due to their high concentrations, the process slowed down gradually, since the precipitation involves nucleation and the growth of crystals on the adsorbent sites that were less available than before; a limiting step, due to rearrangement of atoms and bonds.[21] The q_{max} values were also in accordance with pseudo-first order kinetic model deduced q_e values, *i.e.*,

59.86 and 39.54 mg g^{-1} for Cu^{2+} and Cd^{2+} and the R^2 values for the pseudo--order model being 0.820 and 0.911 for Cu^{2+} and Cd^{2+}, respectively.

TABLE II. Kinetic parameters of Cu^{2+} and Cd^{2+} using *Mactra aequisulcata* seashells

Kinetic model	Parameter	Cu^{2+}	Cd^{2+}
	q_{exp} / mg g^{-1}	59.57	38.12
Pseudo-first order	R^2	0.961	0.970
	k_1 / min^{-1}	0.160	0.08
	q_{cal} / mg g^{-1}	59.86	39.54
Pseudo-second order	R^2	0.820	0.911
	k_2 / g mg^{-1} min^{-1}	0.005	0.002
	q_{cal} / mg g^{-1}	64.46	47.74

Biosorption thermodynamics

The thermodynamic results of the biosorption are accessible in Table III. The negative values of ΔG values illustrating the spontaneity and feasibility of biosorption process.[22] The attraction of the adsorbate towards the biosorbent is evaluated by K_c values and the maximum values for each ion were found at 303 K. Hence, all biosorption process were accomplished at this condition. The ΔG statistics raised with enhanced temperatures, demonstrating the prompt and more spontaneous biosorption process at lower temperatures. The ΔH values were –9.77 and –5.18 kJ mol–^{1} for Cu^{2+} and Cd^{2+}, respectively. The negative results showing that the biosorption was exothermic, at 303–323 K. The negative values of entropy showing a reduction in the disorderliness in ions biosorbed on the biosorbent in the solution. Accordingly, there were no further accessible ions contributing to the entropy in the solution.[2]

TABLE III. Thermodynamic parameters at different temperatures for the biosorption of Cu^{2+} and Cd^{2+} using *Mactra aequisulcata* seashells

Ion	T / K					ΔH / kJ mol^{-1}	ΔS / kJ mol^{-1} K^{-1}
	303	308	313	318	323		
	ΔG / kJ mol^{-1}						
Cu^{2+}	–3.39	–3.27	–3.17	–3.06	–2.97	–9.77	–0.02
Cd^{2+}	–1.74	–1.70	–1.64	–1.60	–1.50	–5.18	–0.01

Table IV illustrates the biosorption capacities of similar genre of biosorbents. The table also shows that the used biosorbent has the potential to be used as an efficient biosorbent due to its relatively better biosorption capacities.

CONCLUSION

This study intends to gauge intrinsic efficacy of the seashells for its diverse application. The easily found indigenous seashell, *Mactra Aequisulcata* of class Bivalvia was selected and successfully employed as a biosorbent to remove Cu^{2+}

TABLE IV. An overview of biosorption capacities of different biosorbents of similar class.

Adsorbent	Material/polymorph	q_e / mg g^{-1}	Reference
	Cu^{2+}		
Anadara inaequivalvis shells	Aragonite ($CaCO_3$)	330.2	2
Mactra aequisulcata shells	Aragonite ($CaCO_3$)	59.60	This study
Chinonecetes opilio (crab shell)	Protein, ash, lipid and chitin	55.90	23
Bivalve *mollusc* shells	$CaCO_3$	38.93	24
Arca shell	Biomass	17.64	25
	Cd^{2+}		
Apple snail shell	Aragonite ($CaCO_3$)	81.3	18
Mactra aequisulcata shells	Aragonite ($CaCO_3$)	38.1	This study
Mussel shells	Aragonite ($CaCO_3$)	28.6	9
Corbicula fluminea shell	Aragonite ($CaCO_3$)	4.03	26
Crab shell	$CaCO_3$	3.43	27
Oyster shell	$CaCO_3$	3.42	28

and Cd^{2+} from aqueous phase. The maximum biosorption capacities were 59.6 and 38.1 mg g^{-1} at pH 5.0 for Cu^{2+} and Cd^{2+}, respectively. The acquired equilibrium data were best fitted with the Langmuir isotherm model. The biosorption showed spontaneity, followed pseudo-first order kinetics and was exothermic in nature. The kinetic study revealed quick uptake equilibrium at 20 and 30 min for Cu^{2+} and Cd^{2+}, respectively. Surface precipitation was deemed to be the primary adsorption mechanism. In conclusion, *M. aequisulcata* shells have the potential of a proficient biosorbent for the Cu^{2+} and Cd^{2+} removal from aqueous phase. Consequently, it can be used as economically and ecologically sutable biosorbent for heavy metals removal with reasonable biosorption capacity and abundance at industrial level.

SUPPLEMENTARY MATERIAL

Additional data are available electronically at the pages of journal website: https://www.shd-pub.org.rs/index.php/JSCS/index, or from the corresponding author on request.

ИЗВОД

ПОУЗДАН ПРИСТУП УКЛАЊАЊУ ЈОНА БАКРА (II) И КАДМИЈУМА (II) ИЗ ВОДЕНИХ РАСТВОРА КОРИШЋЕЊЕМ АУТОХТОНИХ ШКОЉКИ *Mactra aequisulcata*

SYED MUHAMMAD REHAN ULLAH[1,2], ERUM ZAHIR[2] и MUHAMMAD ASIF ASGHAR[3]

[1]Department of Chemical Oceanography & Marine Environment, National Institute of Oceanography, St-47, Block-1 Clifton, Karachi–75600, Sindh–74200, Pakistan, [2]Department of Chemistry, University of Karachi, Karachi–75270, Sindh–74200, Pakistan и [3]Food and Feed Safety Laboratory, Food and Marine Resources Research Centre, PCSIR Laboratories Complex, Shahrah-e-Salimuzzaman Siddiqui, Off University Road, Karachi–75280, Sindh–74200, Pakistan

Сирове шкољке *Mactra aequisulcata*, класе Bivalvia и врсте Mollusca коришћене су као адсорбенс за уклањање металних јона бакра (II) и кадмијума (II) из воденог раствора. Карактеризација љуштура извршена је помоћу FTIR, EDX и SEM методе, као и BET-изотерме и тачке нултог наелектрисања (pH_{PZC}). Изведени су шаржни експерименти за процену различитих фактора на ефикасност биосорпције. Максимална адсорпција

оба метална јона нађена је при дози адсорбенса од 0,4 g L^{-1}, при pH 5 и 303 K. Максимални капацитет биосорпције за Cu^{2+} и Cd^{2+} био је 59,57 и 38,12 mg g^{-1}, редом. Утврђено је да се модел Ленгмирове изотерме показао као најприкладнији за добијене равнотежне податке. Термодинамички и кинетички параметри показали су да је процес изводљив, егзотерман и да прати псеудо-први ред.

(Примљено 28. децембра 2020, ревидирано 22. марта, прихваћено 4. априла 2021)

REFERENCES

1. K. Nahar, M. A. K. Chowdhury, M. A. H. Chowdhury, A. Rahman, K. M. Mohiuddin, *Environ. Sci. Pollut.* Res. **25** (2018) 7954 (https://doi.org/10.1007/s11356-017-1166-9)
2. S. K. Bozbaş, Y. Boz, *Process Saf. Environ. Prot.* **103** (2016) 144 (https://doi.org/10.1016/j.psep.2016.07.007)
3. D. Gola, A. Malik, Z. A. Shaikh, T. R. Sreekrishnan, *Environ. Process.* **3** (2016) 1063 (https://doi.org/10.1007/s40710-016-0176-9)
4. O. Abdi, M. Kazemi, *J. Mater. Environ. Sci.* **6** (2015) 1386 (https://www.jmaterenvironsci.com/Document/vol6/vol6_N5/164-JMES-1454-2015-Abdi.pdf)
5. B. Bano, E. Zahir, *Water Sci. Technol.* **73** (2015) 1301 (https://doi.org/10.2166/wst.2015.590)
6. R. R. Crichton, *Biological Inorganic Chemistry,* Elsevier, Milan, 2007, p. 330 (https://doi.org/10.1016/B978-0-444-52740-0.X5024-8)
7. Y. Du, F. Lian, L. Zhu, *Environ. Pollut.* **159** (2011) 1763 (https://doi.org/10.1016/j.envpol.2011.04.017)
8. E. Jeon, S. Ryu, S. Park, L. Wang, D. C. W. Tsang, *J. Clean. Prod.* **176** (2017) 54 (https://doi.org/10.1016/j.jclepro.2017.12.153)
9. H. T. Van, L. H. Nguyen, V. D. Nguyen, X. H. Nguyen, T. H. Nguyen, T. V. Nguyen, S. Vigneswaran, J. Rinklebe, H. N. Tran, *J. Environ. Manage.* **241** (2019) 535 (http://hdl.handle.net/10453/128739)
10. *Mactra aequisulcata* Sowerby Iii 1894 - Encyclopedia of Life https://eol.org/pages/46471883, (28th January, 2021)
11. I. Langmuir, *J. Am. Chem. Soc.* **40** (1918) 136 (https://doi.org/10.1021/ja02242a004)
12. N. Ayawei, S. S. Angaye, D. Wankasi, E. D. Dikio, *Open J. Phys. Chem.* **5** (2015) 56
13. (https://doi.org/10.4236/ojpc.2015.53007)
14. H. Freundlich, *Z. Phys. Chem.* **57** (1907) 385 (https://doi.org/10.1515/zpch-1907-5723)
15. N. Ayawei, A. N. Ebelegi, D. Wankasi, *J. Chem.* **2017** (2017) (https://doi.org/10.1155/2017/3039817)
16. Z. Berizi, S. Y. Hashemi, M. Hadi, A. Azari, A. H. Mahvi1, *Water Sci. Technol.* **74** (2016) 1235 (https://doi.org/10.2166/wst.2016.320)
17. A. Shafiu Kamba, M. Ismail, T. A. Tengku Ibrahim, Z. A. B. Zakaria, *J. Nanomater.* **2013** (2013) (https://doi.org/10.1155/2013/398357)
18. D. Alidoust, M. Kawahigashi, S. Yoshizawa, H. Sumida, M. Watanabe, *J. Environ. Manage.* **150** (2015) 103 (https://doi.org/10.1016/j.jenvman.2014.10.032)
19. B. Zhao, J. E. Zhang, W. Yan, X. Kang, C. Cheng, Y. Ouyang, *Desalin. Water Treat.* **57** (2016) 23987 (https://doi.org/10.1080/19443994.2016.1140078)
20. A. Sdiri, T. Higashi, *Appl. Water Sci.* **3** (2013) 29 (https://doi.org/10.1007/s13201-012-0054-1)
21. H. N. Tran, S. You, H. Chao, *Waste. Manage. Res.* **34** (2016) 129 (https://doi.org/10.1177/0734242x15615698)

22. D. M. Veneu, C. L. Schneider, M. B. De Mello, O. Galvão, C. Cunha, L. Yokoyama, *Environ. Technol.* **39** (2018) 1670 (https://doi.org/10.1080/09593330.2017.1336574)
23. W. Wei, Q. Wang, A. Li, J. Yang, F. Ma, S. Pi, D. Wu, *Sci. Rep.* **6** (2016) 1 (https://doi.org/10.1038/srep31575)
24. A. Hossain, S. R. Bhattacharyya, G. Aditya, *Wat. Res.* **35** (2014) 3551 (https://doi.org/10.1016/S0043-1354(01)00099-9)
25. Y. Liu, C. Sun, J. Xu, Y. Li, *J. Hazard. Mater.* **168** (2009) 156 (https://doi.org/10.1016/j.jhazmat.2009.02.009)
26. S. Dahiya, R. M. Tripathi, A. G. Hegde, *Bioresour. Technol.* **99** (2008) 179 (https://doi.org/10.1016/j.biortech.2006.11.011)
27. F. A. Ismail, A. Z. Aris, *Environ. Sci. Pollut. Res.* **21** (2014) 344 (https://doi.org/10.1007/s11356-013-1906-4)
28. C. Zhou, X. Gong, W. Zhang, J. Han, R. Guo, *Water Environ. Res.* **89** (2017) 817 (https://doi.org/10.2175/106143017X14902968254854)
29. Y. Gao, *Asian J. Chem.* **25** (2013) 8537 (http://dx.doi.org/10.14233/ajchem.2013.14828A).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 86 (7–8) S156–S157 (2021) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

A prudent approach for the removal of copper (II) and cadmium (II) ions from aqueous solutions using indigenous *Mactra aequisulcata* shells

SYED MUHAMMAD REHAN ULLAH[1,2]*, ERUM ZAHIR[2] and MUHAMMAD ASIF ASGHAR[3]

[1]Department of Chemical Oceanography & Marine Environment, National Institute of Oceanography, St-47, Block-1 Clifton, Karachi−75600, Sindh−74200, Pakistan, [2]Department of Chemistry, University of Karachi, Karachi−75270, Sindh−74200, Pakistan and [3]Food and Feed Safety Laboratory, Food and Marine Resources Research Centre, PCSIR Laboratories Complex, Shahrah-e-Salimuzzaman Siddiqui, Off University Road, Karachi−75280, Sindh−74200, Pakistan

J. Serb. Chem. Soc. 86 (7–8) (2021) 767–780

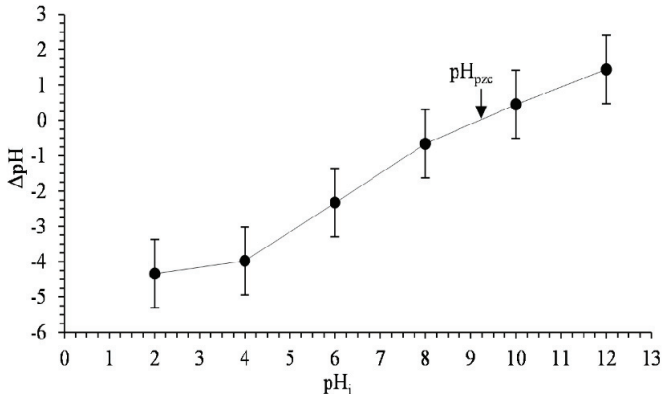


Fig. S-1. Point of zero charge of *Mactra aequisulcata* seashell powder.

* Corresponding author. E-mail: rehan-syed@outlook.com

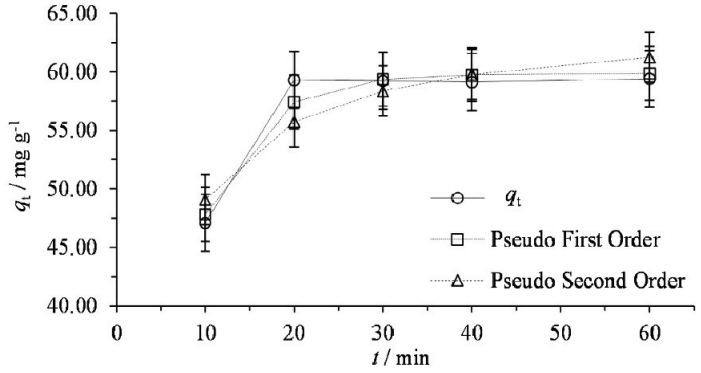


Fig. S-2. Pseudo-first order and Pseudo-second order plots of Cu^{2+} metal ions on seashell (*Mactra aequisulcata*) powder.

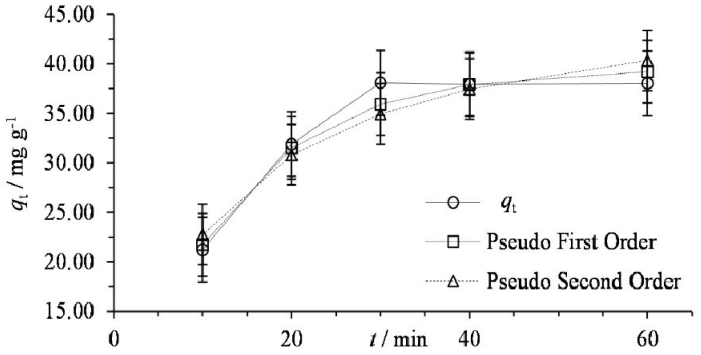


Fig. S-3. Pseudo-first order and Pseudo-second order plots of Cd^{2+} metal ions on seashell (*Mactra aequisulcata*) powder.

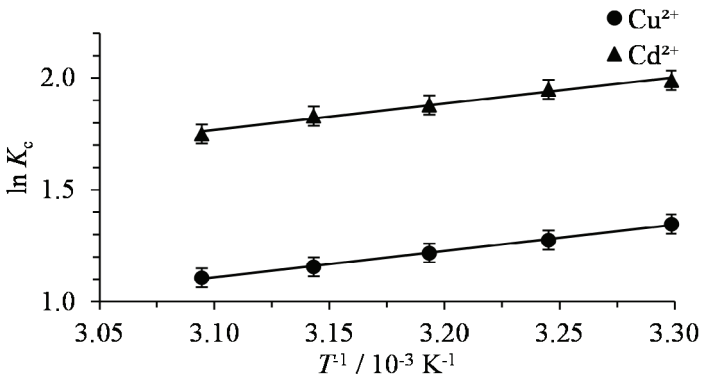


Fig. S-4. Thermodynamic equilibrium constant, K_c, and inverse temperature, T^{-1} / 10^{-3} K^{-1}, for biosorption of Cu^{2+} and Cd^{2+} ions.