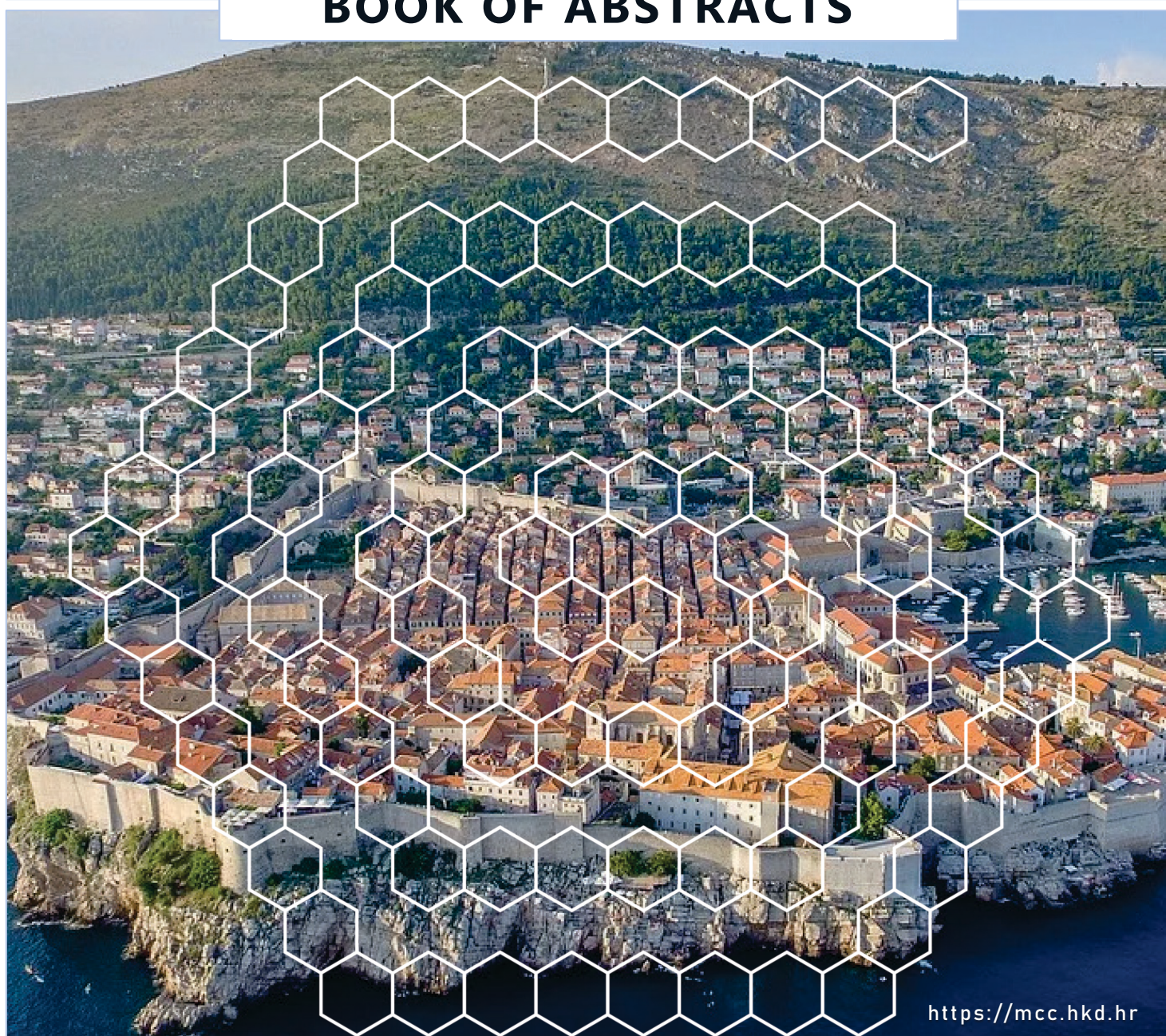


Math/Chem/Comp 2023 – 34th MC² Conference

Inter University Centre Dubrovnik, 5 – 9 June 2023

BOOK OF ABSTRACTS

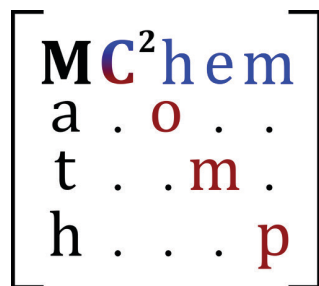


<https://mcc.hkd.hr>

The Math / Chem / Comp Conference is organised by
the Inter-University Centre, Dubrovnik, Croatian Chemical Society and the
Department of Chemistry, Faculty of Science, University of Zagreb, Croatia

IUC





Math/Chem/Comp 2023 – 34th MC² Conference
Inter University Centre Dubrovnik, 5 – 9 June 2023

BOOK OF ABSTRACTS

IMPRESSUM

ORGANIZER

Inter-University Centre, Dubrovnik
Croatian Chemical Society
Department of Chemistry, Faculty of Science, University of Zagreb, Croatia

PUBLISHED BY

Croatian Chemical Society

EDITORS

Hrvoj Vančik, Jerzy Cioslowski,
Danijel Namjesnik

DESIGN AND LAYOUT

Danijel Namjesnik

CONFERENCE VENUE

Inter-University Centre Dubrovnik
Don Frana Bulića 4
HR-20000 Dubrovnik, Croatia
<https://iuc.hr/>
Tel: +385 20 413 626, 413 627

<https://mcc.hkd.hr/>

ISBN: 978-953-8334-07-8

Zagreb, 2023

**The 34th International Course and Conference on the Interfaces among
Mathematics, Chemistry and Computer Sciences:
Mathematics, Chemistry, Computing (Math/Chem/Comp, MC²-34)**

Since 1986, the Math/Chem/Comp meetings have been bringing together researchers working on diverse subjects spanning mathematics, computer science, and chemistry. In continuation of this tradition, presentations concerning all aspects of current research on computer modeling, combinatorics, graph theory, and topology applied to any area of chemistry, physics, material and life sciences will be welcomed. The conference program will include invited plenary and special lectures workshops and posters. In particular, the MC²-34 meeting will revolve around the following topics:

1. Mathematical methods and computational algorithms in modeling of atoms, molecules, and chemical processes
2. Mathematical chemistry and chemical graph theory
3. Molecular modeling in practice (including molecular dynamics and aggregation)
4. Chemical experiments, industrial processes, and mathematical interpretation

The conference program will include invited lectures, contributed oral presentations, and posters. Moreover, we will offer **TWO Workshops** which will be held by Dr Nađa Došlić (Ruđer Bošković Institute, Zagreb, Croatia): *Photoinduced processes in molecules and how to follow them*, and Prof. Tomislav Došlić (Faculty of Civil Engineering, University of Zagreb, Croatia): *Chemical graph theory revisited*.

Proceedings

The papers presented at the MC²-34 meeting can be submitted to [Croatica Chemica Acta](#) to be published as proceedings.

All contributions will be refereed. The participants are encouraged to submit their manuscripts before the meeting or at the latest upon arrival in Dubrovnik. Those unable to attend are invited to send their manuscripts to the director of the conference.

Hrvoj Vančik and Jerzy Cioslowski

CONFERENCE BOARD

Director: Hrvoj Vančik, Professor of Chemistry, University of Zagreb, Croatia

Co-director: Jerzy Cioslowski, Professor of Physics, University of Szczecin, Poland

SCIENTIFIC COMMITTEE

Ana Amić (Josip Juraj Strossmayer University of Osijek, Croatia), Vincenzo Aquilanti (University of Perugia, Italy), Attila G. Császár (Eötvös University, Budapest, Hungary), Tomica Hrenar (University of Zagreb, Croatia), Bono Lučić (Ruđer Bošković Institute, Zagreb, Croatia), Roland Mitrić (University of Würzburg, Germany), Andreas Savin (UPMC Sorbone University, Paris, France), Miquel Sola (University of Girona, Spain), Peter Stadler (University of Leipzig, Germany), Péter Surján (Eötvös Loránd University, Budapest, Hungary), Jesus Ugalde (Euskal Herriko University, Euskadi, Spain), Urban Bren (University of Maribor, Slovenia), Filip Prątnicki (University of Szczecin, Poland), Danijel Namjesnik (University of Zagreb, Croatia)

HONORARY COMMITTEE

Aleksandru Balaban, Ivan Gutman, Milan Randić, Tomislav Živković

CONTENTS

LECTURES

Miquel Solà

3D-aromaticity. Definition and Application to (car)boranes..... 12

Bartosz Trzaskowski

*Computational Methods to Study Electronic and Dynamic Properties of Molecular
Cages Encapsulating Silver Clusters 13*

Filip Prątnicki

Properties of the Natural Orbitals in Ground States of Two-electron Systems..... 14

Hrvoj Vančik

From Chemical Concepts to Iso-science 15

Jerzy Cioslowski

*One-electron Reduced Density Matrix Description of Long-range Interaction of
Two Hydrogen Atoms 16*

Gianluca Serra

A Graph-theoretical Analysis of the Effect of Boron-nitrogen Doping in Graphene Molecules.. 17

Henryk A. Witek

Toward Quantitative Raman Spectroscopy..... 18

Guillermo Restrepo

Chemical Space: What We Know and What We do not About it and its Evolution 19

Bono Lučić

*Improvement of Fingerprint-based Similarity Measures - in General, and Specifically in
Applications in Molecular Chemo/bio Sciences..... 20*

Urban Bren

*Conserved Water Networks Identification in Proteins Using Positional and
Orientational Data 21*

WORKSHOPS

Tomislav Došlić: *Chemical Graph Theory Revisited 24*

Nađa Došlić: *Photo-induced Processes in Molecules and How to Follow Them..... 25*

POSTERS

Ottorino Ori

Walking on Polyhexes: Size-depending Symmetry Between Tori & Klein Bottles 28

Marina Poljak, Jakov Borovec, Ines Primožič, and Tomica Hrenar

Optimizing Deep Neural Network Architecture for Fragrant Compounds Classification 29

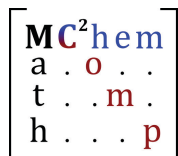
**Bono Lučić, Viktor Bojović, Nenad Židak, Davor Oršolić, Tomislav Šmuc, and
Christoph Steinbeck**

*Statistical and Chemical Relevance of Molecular Descriptors in
Structure-activity/Property Models* 30

AUTHORS AND PARTICIPANTS

AUTHORS 32

PARTICIPANTS 33



Math/Chem/Comp 2023 – 34th MC² Conference

Inter University Centre Dubrovnik, 5 – 9 June 2023

PROGRAMME

MONDAY, JUNE 5

9:30 – 10:00	REGISTRATION
10:00 – 10:30	CONFERENCE OPENING (Hrvoj Vančik / Jerzy Ciośłowski)
CHAIR: Jerzy Ciośłowski	
10:30 – 11:30	Miquel Solà: <i>3D-aromaticity. Definition and Application to (car)boranes</i>
11:30 – 12:30	Bartosz Trzaskowski: <i>Computational Methods to Study Electronic and Dynamic Properties of Molecular Cages Encapsulating Silver Clusters</i>
12:30 – 13:30	Filip Prątnicki: <i>Properties of the Natural Orbitals in Ground States of Two-electron Systems</i>
13:30 – 16:00	LUNCH BREAK
16:00 – 18:00	SCIENTIFIC DISCUSSIONS

TUESDAY, JUNE 6

CHAIR: Miquel Solà	
10:00 – 11:00	Hrvoj Vančik: <i>From Chemical Concepts to Iso-science</i>
11:00 – 12:00	Jerzy Ciośłowski: <i>One-electron Reduced Density Matrix Description of Long-range Interaction of Two Hydrogen Atoms</i>
12:00 – 13:00	Gianluca Serra: <i>A Graph-theoretical Analysis of the Effect of Boron-nitrogen Doping in Graphene Molecules</i>
13:00 – 15:00	LUNCH BREAK
15:00 – 18:00	SCIENTIFIC DISCUSSIONS

WEDNESDAY, JUNE 7

CHAIR: Hrvoj Vančik	
10:00 – 11:00	Henryk A. Witek: <i>Toward Quantitative Raman Spectroscopy</i>
11:00 – 13:00	WORKSHOP 1 Tomislav Došlić: <i>Chemical Graph Theory Revisited</i>
13:00 – 15:00	LUNCH BREAK
15:00 – 17:00	WORKSHOP 2 Nađa Došlić: <i>Photo-induced Processes in Molecules and How to Follow Them</i>
19:00	CONFERENCE DINNER

THURSDAY, JUNE 8

CHAIR: Henryk A. Witek

10:00 – 11:00	Guillermo Restrepo: <i>Chemical Space: What We Know and What We do not About it and its Evolution</i>
11:00 – 12:00	Bono Lučić: <i>Improvement of Fingerprint-based Similarity Measures - in General, and Specifically in Applications in Molecular Chemo/bio Sciences</i>
12:00 – 13:00	Urban Bren: <i>Conserved Water Networks Identification in Proteins Using Positional and Orientational Data</i>
13:00 – 15:00	LUNCH BREAK
15:00 – 18:00	SCIENTIFIC DISCUSSIONS

FRIDAY, JUNE 9

10:00 – 12:00	POSTER SESSION
12:00 – 12:30	CONFERENCE CLOSING (Hrvoj Vančik / Jerzy Ciosłowski)

POSTER SESSION

P01	Ottorino Ori: <i>Walking on Polyhexes: Size-depending Symmetry Between Tori & Klein Bottles</i>
P02	Marina Poljak, Jakov Borovec, Ines Primožič, and Tomica Hrenar: <i>Optimizing Deep Neural Network Architecture for Fragrant Compounds Classification</i>
P03	Bono Lučić, Viktor Bojović, Nenad Židak, Davor Oršolić, Tomislav Šmuc, and Christoph Steinbeck: <i>Statistical and Chemical Relevance of Molecular Descriptors in Structure-activity/Property Models</i>

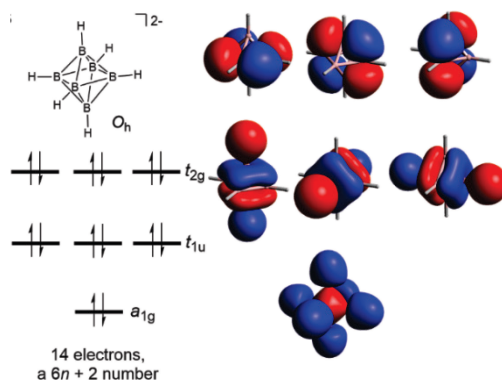
LECTURES

3D-AROMATICITY. DEFINITION AND APPLICATION TO (CAR)BORANES

Miquel Solà

Institut de Química Computacional and Departament de Química, Universitat de Girona,
C/ Maria Aurèlia Capmany 69, 17003 Girona, Catalonia, Spain
miquel.sola@udg.edu

Recently, it was discovered that a number of π -conjugated macrocycles with puckered or cage-like structures exhibited aromatic character according to both experiments and calculations. We look at their electronic structures and compare them to 2D-aromatic polycyclic aromatic hydrocarbons and three-dimensional (3D) aromatic compounds (such as *closo* boranes). We discover that the macrocycles investigated up to this point should be classified as 2D-aromatic with three-dimensional molecular structure (2D-aromatic-in-3D) and not as really 3D-aromatic using qualitative theory paired with quantum chemical calculations [1]. We discuss the requirements for a molecule to be consider 3D-aromatic. Among them, we conclude that they should have highly symmetric structures (or nearly so), leading to (at least) triply degenerate molecular orbitals, and for tetrahedral or octahedral molecules an aromatic closed-shell electronic structure with $6n + 2$ electrons. Next, we will analyze the aromaticity of *closo* and *nido* boranes and carboranes. Quantum chemical calculations reveal that the *nido* cage retains part of the aromatic character that characterizes the *closo* borane cages [1]. Then, we will elaborate on the aromaticity in benzenoids (2D aromatic) linked to carboranes (3D aromatic) [2]. After that we will apply the Clar's π -sextet model to explain the extrusion of benzenoids linked to cobaltabis(dicarbollide), $[3,3'\text{-Co}(1,2\text{-C}_2\text{B}_9\text{H}_{11})_2]^-$ [3]. Finally, we will show an example of double 3D-aromaticity.



REFERENCES

- [1] O. El Bakouri, D. W. Szczepanik, K. Jorner, R. Ayub, P. Bultinck, M. Solà, H. Ottosson, *J. Am. Chem. Soc.* **144** (2022) 8560-8575.
- [2] J. Poater, C. Viñas, M. Solà, F. Teixidor, *F. Nat. Commun.* **13** (2022) 3844.
- [3] J. Poater, C. Viñas, D. Olid, M. Solà, F. Teixidor, *Angew. Chem. Int. Ed.* **61** (2022) e202200672.

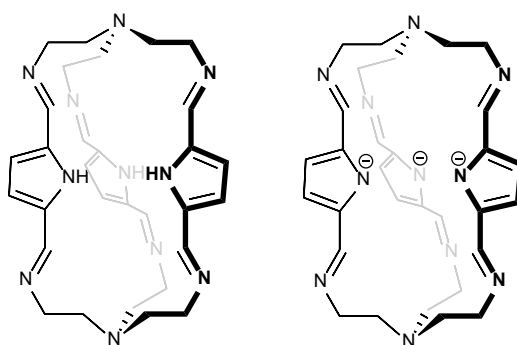
COMPUTATIONAL METHODS TO STUDY ELECTRONIC AND DYNAMIC PROPERTIES OF MOLECULAR CAGES ENCAPSULATING SILVER CLUSTERS

Bartosz Trzaskowski,^a Bartosz Szyszko^b

^a Centre of New Technologies, University of Warsaw, Banacha 2C St., 02-097 Warszawa, Poland
b.trzaskowski@cent.uw.edu.pl

^b Department of Chemistry, University of Wrocław, 14 F. Joliot-Curie St., 50-383 Wrocław, Poland
bartosz.szyszko@uwr.edu.pl

Tris(2-aminoethyl)amine (tren)-based iminocryptands are one of the archetypical cage systems used for selective ion recognition. Recently, a new molecular cage able to encapsulate various ions has been synthesized via the self-assembly reactions of tren and 2,5-diformylpyrrole in the presence of silver(I) [1]. The self-assembly of tren, 2,5-diformylpyrrole, and silver(I) resulted in two principal product types, which depend on the anion's size, coordination abilities, and basicity. Small, spherical halides favored cascade complexes incorporating the anion(s) entrapped between two metal centers. The triflate that was poorly coordinating and geometrically incompatible with the cavity dimensions enforced the formation of plenates - cages encapsulating silver(I) clusters.



Using a combination of density functional theory, tight binding approached and molecular dynamics we show that in the case of neutral pyrrole cages, Gibbs free energies of formation are a good measure of predicting the ratio of bound ions. For the charged pyrrole cages, we have identified relatively strong argentophilic interaction between Ag ions on the basis of the calculated bond indices and obtained molecular orbitals. Regarding the cage encapsulating Ag₄ cluster, the two minimum-geometry conformations have been estimated to differ only by 6.5 kcal/mol with an energy barrier of less than 1 kcal/mol, suggesting a very flexible structure, further supported by the results of the molecular dynamics. Finally, we have also estimated the energy barriers of formation of cages encapsulating Ag₁-Ag₅ species, showing relatively low energy barriers of formation of all of them, consistent with the experimental data.

REFERENCES

[1] A. Sarwa, A. Białońska, M. Garbicz, B. Szyszko, *Chem. – Eur. J.* **29** (2023) e202203850.

PROPERTIES OF THE NATURAL ORBITALS IN GROUND STATES OF TWO-ELECTRON SYSTEMS

Filip Prątnicki

Institute of Physics, University of Szczecin, Wielkopolska 15, 70-451 Szczecin, Poland

filip.pratnicki@usz.edu.pl

The natural spinorbitals (NOs) have been extensively studied over the last 50 years. These studies notwithstanding, many properties of NOs have not been sufficiently investigated. In particular, the fundamental question whether some of the NOs pertaining to Coulombic systems can be unoccupied has remained open. Addressing this gap in the present knowledge, the research reported in this speech focuses upon the analysis of the properties, trends and mathematical expressions concerning the NOs of the ground states of two-electron systems.

Limited reliability of the known numerical methods has been solved with the derivation of a new mathematical formalism based upon the regularized Krylov sequences, which has allowed the calculations of unprecedented accuracy on the electronic properties of the members of the helium isoelectronic series [1,2]. These calculations (carried out also for the H₂ molecule) uncovered the presence of the so called solitonic natural orbitals (SoNOs) characterized by approximate retention of their shapes upon variation of the control parameters in the Hamiltonian, substantial spatial localization, and asymptotic behaviour that differs from that of the regular NOs. The analysis of the evolution of these SoNOs with the nuclear charge Z in the case of the helium-like systems and the internuclear distance R in the case of the H₂ molecule has opened an avenue to the estimation of the nuclear charges at which particular SoNOs are supposed to turn into UNOs [3].

In addition to answering one of the longstanding fundamental questions concerning the existence of UNOs, the research reported in this speech has produced a plethora of highly useful numerical data [2-5].

REFERENCES

- [1] J. Cioslowski, F. Prątnicki, *J. Chem. Phys.* **149** (2018) 184107.
- [2] J. Cioslowski, F. Prątnicki, *J. Chem. Phys.* **153** (2020) 224106.
- [3] J. Cioslowski, F. Prątnicki, K. Strasburger, *J. Chem. Phys.* **156** (2022) 034108.
- [4] J. Cioslowski, F. Prątnicki, *J. Chem. Phys.* **150** (2019) 074111.
- [5] J. Cioslowski, F. Prątnicki, *J. Chem. Phys.* **151** (2019) 184107.

FROM CHEMICAL CONCEPTS TO *ISO*-SCIENCE

Hrvoj Vančik

Faculty of Science, Department of Chemistry, University of Zagreb, Horvatovac 102a, Zagreb, Croatia
vancik@chem.pmf.hr

Basic chemical theories, such as the theory of types that has later been transformed to structural theory, has emerged by searching for the *analogy* between chemical entities, i.e., between substances, or crystals, or even reactions. For naming the analogous chemical entities, the specific terminology with the prefix *iso-* was created: *isomers*, *isotopes*, *isomorphs*, *isostructural*, *isoelectronic*, etc. This prefix *iso* serves to collect a group of entities into specific category on the basis of particular common invariant property.

I argue, how this epistemological principle, which is one of the focal points in chemical conceptualism, could be generalized in science: *iso*-theories, *iso*-functionality, *iso*-models, *iso*-coding, etc. The universality of such an *iso*-approach is discussed within the frame of the complexity theory and the theory of systems.

REFERENCES

- [1] H. Vančik, "Philosophy of Chemistry" *Springer Nature*, **2021**, Dordrecht, New York, Berlin.
- [2] H. Vančik, "From Complexity to Systems" *Foundations of Chemistry* **2023**.
<https://doi.org/10.1007/s10698-022-09455-6>

ONE-ELECTRON REDUCED DENSITY MATRIX DESCRIPTION OF LONG-RANGE INTERACTION OF TWO HYDROGEN ATOMS

Jerzy Cioslowski

Institute of Physics, University of Szczecin, Wielkopolska 15, 70-451 Szczecin, Poland

jerzy.cioslowski@usz.edu.pl

The leading terms in the large- R asymptotics of the functional of the one-electron reduced density matrix for the ground-state energy of the H_2 molecule with the internuclear separation R are derived thanks to the solution of the phase dilemma at the $R \rightarrow \infty$ limit. At this limit, the respective natural orbitals (NOs) are given by symmetric and antisymmetric combinations of “half-space” orbitals with the corresponding natural amplitudes having the same absolute values but opposite signs. Minimization of the resulting explicit functional yields the large- R asymptotics for the occupation numbers of the weakly occupied NOs and the C_6 dispersion coefficient. The highly accurate approximates for the radial components of the p-type “half-space” orbitals and the corresponding occupation numbers (that decay like R^{-6}), which are available for the first time thanks to the development of the present formalism, have some unexpected properties.

A GRAPH-THEORETICAL ANALYSIS OF THE EFFECT OF BORON-NITROGEN DOPING IN GRAPHENE MOLECULES

Gianluca Serra,^a Giacomo Asnaghi,^b Giacomo Bettani,^c Edoardo Cardani,^d Gabriele Corno,^e Francesco Cutore,^f and Matteo Tommasini^g

^a Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; gianluca.serra@polimi.it

^b Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; giacomo.asnaghi@mail.polimi.it

^c Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; giacomo.bettani@mail.polimi.it

^d Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; edoardo.cardani@mail.polimi.it

^e Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; gabriele1.corno@mail.polimi.it

^f Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; francesco.cutore@mail.polimi.it

^g Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; matteo.tommasini@polimi.it

Graphene molecules, also known as polycyclic aromatic hydrocarbons (PAHs), are among the most studied materials for the realization of organic electronic devices. These molecules have topology-dependent optical and electronic properties, and can be synthesized through bottom-up approaches. One way of fine-tuning their properties is by formally substituting carbon–carbon (C–C) bonds with boron–nitrogen (B–N) atomic pairs: in such way, the electronic structure of the resulting BN-PAH is perturbed keeping the geometry of the molecule almost unchanged [1].

One particular electronic property that can be fine-tuned is the energy difference ΔE between the frontier molecular orbitals. As expected, the perturbation to ΔE in a molecule depends both on the number and on the spatial disposition of the B–N atomic pairs; however, the exact mechanism governing this effect is still not properly understood.

In this contribution, we investigate the effect of formal B–N doping by means of chemical graph theory, considering the colored hydrogen-depleted graphs of several BN-PAHs. In particular, we search for a correlation between the perturbation to ΔE induced by formal BN doping and the corresponding variation of path-based graph invariants.

REFERENCES

- [1] X. Chen, D. Tan, D.-T. Yang, *J. Mater. Chem. C* **10** (2022) 13499–13532.

TOWARD QUANTITATIVE RAMAN SPECTROSCOPY

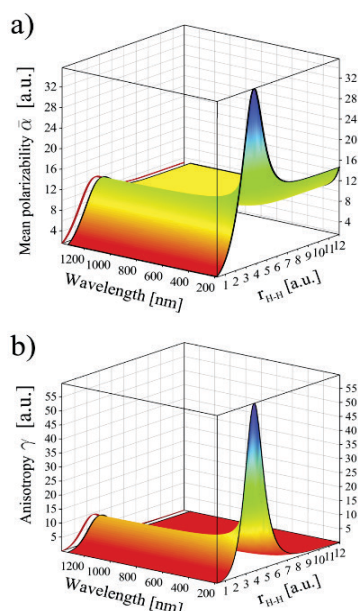
Henryk A. Witek, Ankit Raj, Yen-Bang Chao

Department of Applied Chemistry and Institute of Molecular Science, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

hwitek@nycu.edu.tw

Our goal is to standardize Raman spectroscopy by introducing simple molecular benchmark systems in order to establish absolute Raman intensity scale. This is achieved by computing accurate distance-dependent polarizability invariants of diatomic systems and averaging them over accurate rovibrational wave functions. The matrix elements determined in this fashion are used to calibrate the in-house built Raman spectrometer using the predicted intensity values of selected rovibrational Raman transitions. The calibrated spectrometer is subsequently used to determine the absolute Raman intensities of benchmark systems which are supposed to serve as a reference for calibration of the wide range of commercially available Raman spectrometers.

The current talk will mainly focus on theoretical aspects of the process described above. We will show that the calculations of polarizability invariants but for the simplest systems are exceedingly difficult and challenging due to the inherent shortcomings of the applicable wave functions. The computer programs for computing the rovibrational transition matrix elements over the polarizability tensor invariants for isotopes of molecular hydrogen are freely available at <https://github.com/ankit7540/H2-PolarizabilityMatrixElements>.



Dynamic (wavelength dependent) a) mean polarizability $\bar{\alpha}$ and b) polarizability anisotropy γ of H_2 , HD, and D_2 as a function of internuclear separation.

REFERENCES

- [1] A. Raj, H. Hamaguchi, and H. A. Witek, "Polarizability tensor invariants of H_2 , HD, and D_2 ", *J. Chem. Phys.* **148**, 104308 (11 pages) (2018).
- [2] A. Raj, H. A. Witek, and H. Hamaguchi, "Vibration-rotation interactions in H_2 , HD and D_2 : centrifugal distortion factors and the derivatives of polarizability invariants", *Mol. Phys.* **118**, e1632950 (13 pages) (2020).
- [3] A. Raj, C. Kato, H. A. Witek, and H. Hamaguchi, "Toward standardization of Raman spectroscopy: Wavenumber and intensity calibration using rotational Raman spectra of H_2 , HD, D_2 and vibration-rotation spectrum of O_2 as primary standards", *J. Raman Spectrosc.* **51**, 2066–2082 (2020)
- [4] A. Raj, H. A. Witek, and H. Hamaguchi, "Determination of accurate absolute Raman cross-sections of benzene and cyclohexane in the gas phase", *Asian J. Phys.* **30**, 321–335 (2021).
- [5] A. Raj, C. Kato, H. A. Witek, and H. Hamaguchi, "Accurate intensity calibration of multichannel detectors using Raman intensity ratios", *J. Raman Spectrosc.* **52**, 2038–2050 (2021).
- [6] A. Raj, Y.-B. Chao, and H. A. Witek, "Testing the limitations of harmonic approximation in the determination of Raman intensities", *Mol. Phys.* e2069613 (2022).

CHEMICAL SPACE: WHAT WE KNOW AND WHAT WE DO NOT ABOUT IT AND ITS EVOLUTION

Guillermo Restrepo

Max Planck Institute for Mathematics in the Sciences, Inselstr. 22, 04103, Leipzig, Germany
restrepo@mis.mpg.de

In this talk I will present some of the results on the exploration of the historical unfolding of the chemical space, understood as the collection of all substances and chemical reactions reported over the history of chemistry. Results include the finding of an exponential growth of the number of new substances, with doubling time of 16 years. The stability of this growth under social, technological and scientific changes. The detection of three statistical regimes of production of new chemicals that shed light on the inner workings of chemistry. I will present some upper bounds for the chemical space and the talk will be concluded with open chemo-mathematical questions that involve modelling the space, random hypergraphs and the geometry and curvature of the space.

REFERENCES

- [1] E. J. Llanos; W. Leal; D. H. Luu.; J. Jost; P. F. Stadler; G. Restrepo. *Proc. Natl. Acad. Sci.* **116** (2019) 12660–12665. J. Čížek, *J. Chem. Phys.* **45** (1966) 4256–4267.
- [2] G. Restrepo. *Digit. Discov.* **1** (2022) 568–585.
- [3] A. H. Lipkus. *J. Org. Chem.* **84** (2019) 13948–13956.
- [4] M. Fialkowski; K. J. M. Bishop; V. A. Chubukov; C. J. Campbell; B. A. Grzybowski. *Angew. Chem. Int. Ed.* **44** (2005) 7263–7269.
- [5] S. Szymkuć; T. Badowski; B. A. Grzybowski. *Angew. Chem. Int. Ed.* **60** (2021) 26226–26232.
- [6] W. Leal; G. Restrepo; P. F. Stadler; J. Jost. *Adv. Complex. Syst.* **24** (2019)
<https://doi.org/10.1142/S021952592150003X>

IMPROVEMENT OF FINGERPRINT-BASED SIMILARITY MEASURES - IN GENERAL, AND SPECIFICALLY IN APPLICATIONS IN MOLECULAR CHEMO/BIO SCIENCES

Bono Lučić,^a Viktor Bojović,^b Jadranko Batista,^c and Christoph Steinbeck^d

^a Ruđer Bošković Institute, NMR Centre, Zagreb, Croatia; lucic@irb.hr

^b Ruđer Bošković Institute, Centre for Informatics and Computing, Zagreb, Croatia

^c University of Mostar, Mostar, Faculty of Science and Education, Bosnia and Herzegovina

^d Friedrich-Schiller-University, Institute for Inorganic and Analytical Chemistry, Jena, Germany

In the era of digitisation of data, there is an increasing need for computational tools capable of comparing and analysing entities/objects based on their similarity, including in chemistry and life sciences [1]. Similarity analysis requires searching large databases of molecules. The problem of defining and calculating similarity is a challenge for which there is no unique (mathematical, algorithmic) solution. Therefore, various similarity measures (algorithms, tools, formulas) have been developed to calculate and analyse the similarity between a set of objects/entities/molecules stored in the database according to well-defined rules [2]. Based on the selected optimal values of the similarity measures, conclusions are drawn about the selection of a molecule for further analysis and research or its exclusion from further consideration. Among the similarity measures, the Tanimoto coefficient (Tc) is most commonly used in molecular similarity analysis applications. It is calculated on a set (among many existing) of binary fingerprint representations of a certain length (bits) according to the expression $Tc = TP / (TP + FN + FP)$ [3], where TP (true positive) is the number of TP matching structural properties between two molecules and, analogously, FN (false negative) is the number of matching non-existing structural properties in two strings of fingerprints as well as FN (false negatives) and FP (false positive). In recent comparative analyses, authors have shown good/best properties of Tc as a measure of similarity [2,4]. The view that molecules are considered similar if their $Tc > 0.85$ has gained acceptance and is widely used, although this is only defined (and partially valid) for a particular set (of Daylight) fingerprints. Furthermore, previous detailed combinatorial analyses have shown that the Tc measure is biased and favours the value $Tc = 0.34$ [3]. To address the question of whether it is justified to consider Tc as the best measure of molecular similarity and apply it to all sets of binary fingerprints of a given (or arbitrary) length, we performed analysis of random values of Tc, analogous to the approach introduced for calculating random model accuracy [5]. This analysis highlights the shortcomings of Tc as a measure of similarity when similarity is determined by exhaustive searches of large molecular databases. Ideas for necessary changes/improvements to a similarity measure such as Tc (and other similarity measures) in molecular modelling applications are analysed and presented, as well as the possibility of applying them to non-binary (i.e. multi-class) representations of objects in similarity searches and calculations.

Acknowledgements: This research is supported by the Croatian Ministry of Science and Education and the German Academic Exchange Service (DAAD) for the ChemInf4Biosciences project, and by the Croatian Government and the EU through the ERD Fund for the SCE BioProCro (KK.01.1.1.01).

REFERENCES

- [1] M. Duran-Frigola, *et al.*, *Nat. Biotechnol.* **38** (2020) 1087–1096.
- [2] H. Safizadeh, *et al.*, *J. Chem. Inf. Model.* **61** (2021) 4156–4172.
- [3] J.W. Godden, L. Xue, J. Bajorath, *J. Chem. Inf. Comput. Sci.* **40** (2000) 163–166.
- [4] D. Bajusz, A. Rácz, K. Héberger, *J. Cheminform.* **7** (2015) 20.
- [5] B. Lučić, *et al.*, *Croat. Chem. Acta* **92** (2019) 379–391.

CONSERVED WATER NETWORKS IDENTIFICATION IN PROTEINS USING POSITIONAL AND ORIENTATIONAL DATA

Urban Bren, Marko Jukič, Domagoj Fijan, and Jelena Tošović

Faculty of Chemistry and Chemical Technology, University of Maribor, Smetanova ulica 17,
SI-2000 Maribor, Slovenia
urban.bren@um.si

This work describes the development and testing of a method for the identification and classification of conserved water molecules and their networks from molecular dynamics (MD) simulations. The conserved waters in the active sites of proteins influence protein–ligand binding. Recently, several groups have argued that a water network formed from conserved waters can be used to interpret the thermodynamic signature of the binding site. We implemented a novel methodology to categorize water molecules extracted from the MD simulation trajectories using clustering approaches. The main advantage of our methodology as compared to the current state-of-the-art approaches is the inclusion of information on the orientation of hydrogen atoms to further inform the clustering algorithm and to classify the conserved waters into different subtypes depending on how strongly certain orientations are preferred. This information is vital for assessing the stability of water networks. The newly developed approach is described in detail as well as validated against known results from the scientific literature including comparisons with the experimental data on thermolysin, thrombin, and Haemophilus influenzae virulence protein SiaP as well as with the previous computational results on thermolysin. We observed excellent agreement with the literature and were also able to provide additional insights into the orientations of the conserved water molecules, highlighting the key interactions which stabilize them. The source code of our approach, as well as the utility tools used for visualization, are freely available on GitHub.

REFERENCES

- [1] M. Jukič, J. Konc, D. Janežič, U. Bren, *ACS Med. Chem. Lett.* **11** (2020) 877–882.
- [2] J. Tošović, D. Fijan, M. Jukič, U. Bren, *J. Chem. Inf. Model.* **62** (2022) 6105–6117.

WORKSHOPS

CHEMICAL GRAPH THEORY REVISITED

Tomislav Došlić

University of Zagreb Faculty of Civil Engineering, Kačićeva 26, 10000 Zagreb, Croatia
tomislav.doslic@grad.unizg.hr

Chemical Graph Theory is the name given to an eclectic collections of concepts, methods and techniques aiming at studying chemical compounds by applying methods of discrete mathematics, in particular graph theory. It is based on the assumption that important information on physico-chemical properties of various classes of compounds can be acquired by studying the topological and metric properties of the corresponding graphs and extracting the information coded there. Several subdisciplines evolved over the course of the last couple of decades, some of them from scratch, some based on already existing areas of graph theory. I survey here, from a necessarily personal point of view, recent developments in three such subdisciplines: topological indices, spectral techniques, and large matchings, packings, and related structures.

PHOTO-INDUCED PROCESSES IN MOLECULES AND HOW TO FOLLOW THEM

Nada Došlić

Ruđer Bošković Institute, Department of Physical Chemistry, Bijenička 54, 10000 Zagreb, Croatia
nadja.doslic@irb.hr

Trajectory surface hopping (TSH) methods [1-3] are computational tools used to study photoinduced nonadiabatic processes both in the gas and condensed phases. In TSH, nuclear trajectories are propagated classically on a single Born-Oppenheimer surface, whereas hops between electronic states are dictated by a stochastic algorithm. As TSH methods are usually implemented on-the-fly, they bypass the main limitations of quantum nuclear dynamics - the requirement for precalculated global potential energy surfaces and the exponential scaling of computational costs with the number of nuclear degrees of freedom.[4] TSH simulations are also increasingly used in computational spectroscopy, where the development of trajectory-based protocols for the simulation of spectroscopy signals [5] has made a direct comparison with experiment possible.[6]

Most importantly, TSH methods have made simulations of photo-induced processes more accessible to non-experts. This, however, comes at a price. While it is relatively simple to run surface hopping calculations, the interpretation of the results is never straightforward.[7] In this course I will present different TSH methods, validate their accuracy with respect to quantum nuclear dynamics [3,4] and discuss their advantages and limitations. In particular, I focus on various simulation protocols and provide an (inherently incomplete) set of "best practices" that allows users to avoid common errors and obtain an accurate chemical insight into the process under study.

REFERENCES

- [1] J. C. Tully, *J. Chem. Phys.* **93**, (1990) 1061
- [2] E. Subotnik *et al.*, *Annu. Rev. Phys. Chem.* **67**, (2016) 387
- [3] W. Xie *et al.*, *J. Chem. Phys.* **150**, (2019) 154119
- [4] J. Coonjobeechary *et al.*, *Philos. Trans. Royal Soc. A* **380**, (2022) 20200386
- [5] M. F. Gelin *et al.*, *J. Chem. Theory Comput.* **17**, (2021) 2394–2408
- [6] T. Piteša *et al.*, *J. Chem. Theory Comput.* **17**, (2021) 5098
- [7] O. Travníkova *et al.*, *J. Am. Chem. Soc.* **144**, (2022) 21878

POSTERS

WALKING ON POLYHEXES: SIZE-DEPENDING SYMMETRY BETWEEN TORI & KLEIN BOTTLES

Ottorino Ori

Actinium Chemical Research srl, Rome, Italy

ottorino.ori@gmail.com

By switching the polyhexes topology from Tori to Klein bottles (KB), it has been observed [1] the general mechanism of *eccentricity shrinkage* consisting in the relationship $\underline{\varepsilon} < \bar{\varepsilon}$ between the eccentricities of Tori $\bar{\varepsilon}$ and Klein bottles $\underline{\varepsilon}$ with the same size. Surprisingly, for certain *critical* sizes, a newly topological symmetry has been discovered making both systems transitive and indistinct. Under this condition, the translation invariance typical of the graphenic Tori (periodic graphene) becomes a topological property of the Klein bottle polyhexes too, puzzling us on the possible real chemical synthesis of KB-shaped sp^2 chemical structures. This is an intriguing perspective, worthy of further theoretical and experimental (quantum) investigation.

REFERENCES

- [1] Putz, O. Ori, *Symmetry* **12(8)** (2020) 1233. <https://doi.org/10.3390/sym12081233>

OPTIMIZING DEEP NEURAL NETWORK ARCHITECTURE FOR FRAGRANT COMPOUNDS CLASSIFICATION

Marina Poljak,^a Jakov Borovec,^b Ines Primožič,^b and Tomica Hrenar^b

^a Saponia d.d., M. Gupca 2, HR-31000 Osijek, Croatia

Marina.Poljak@saponia.hr

^b Department of Chemistry, University of Zagreb Faculty of Science, Horvatovac 102a, 10000 Zagreb, Croatia

jakov.borovec@student.pmf.hr, ines.primozic@chem.pmf.hr, tomica.hrenar@chem.pmf.hr

A set of 82 fragrant compounds was investigated by various spectroscopy methods. This set includes 6 main types of perfumery odor notes [1] and an attempt will be made to develop a classification model. The data matrix consisting of spectral data was decomposed using the principal component analysis (PCA) and the quality of PCA models was evaluated by determining the optimal number of principal components for the representation in the reduced space. For building a classification model for fragrant compounds, a deep reinforcement learning protocol implemented in the program *moonee* [2] will be used. However, building such a model is far from trivial and depends on many control and extraneous variables, starting from a proper reduced space coverage, training dataset selection, correct setting in the learning method, optimization of deep neural network architecture, *etc.*

In order to optimize the architecture of the deep neural network that will be used to build the classification model, the implemented full combinatorial algorithm [2] will be used. The scanning of different deep neural network architectures will be performed in a progressive manner where each tested architecture will be built with a different number of hidden layers and a different number of nodes in these layers. An optimal architecture for building of classification model will be selected based on various criteria such as accuracy of the classification, root-mean-square error in the classification, numerical stability of the model and total time of training.

REFERENCES

- [1] P. Kraft, J. A. Bajgrowicz, C. Denis, G. Fráter, *Angew. Chem. Int. Ed.* **39** (2000) 2980–3010.
- [2] T. Hrenar, *moonee*, Code for Manipulation and Analysis of Multi- and Univariate Big Data, rev. 0.68268, 2023.

STATISTICAL AND CHEMICAL RELEVANCE OF MOLECULAR DESCRIPTORS IN STRUCTURE-ACTIVITY/PROPERTY MODELS

Bono Lučić,^{a,*} Viktor Bojović,^b Nenad Židak,^{a,c,#} Davor Oršolić,^d
Tomislav Šmuc,^d and Christoph Steinbeck^e

^a Ruđer Bošković Institute, NMR Centre, Zagreb, Croatia,

^b Ruđer Bošković Institute, Centre for Informatics and Computing, Zagreb, Croatia

^c ten23 health, 3930 Visp, Switzerland

^d Ruđer Bošković Institute, Department of Electronics, Zagreb, Croatia

^e Institute for Inorganic and Analytical Chemistry, Friedrich-Schiller-University, Jena, Germany

* lucic@irb.hr

external collaborator - PhD student

To better understand the relationship between the structure of a molecule and its properties, molecular descriptors are defined and used. They are the result of the development of chemical science (theory), physical concepts, but also mathematics and computer science. Nowadays, there are many programmes to calculate thousands of molecular descriptors starting from 1D, 2D or 3D structures of chemical compounds [1]. The quality of the descriptor can be observed in the physicochemical sense, but also in the statistical sense, on a set of molecules with similar activity or property. The quality of the descriptor defined in this way largely determines the quality of the quantitative structure-activity relationship (QSAR) models. Their development was stimulated by OECD activities to develop regulatory QSAR models for human health and environmental applications [1,2]. The results of the assessment of the monotonicity/variability of different molecular descriptors calculated on the OCHEM platform [3] in modelling the inhibition of tyrosine kinases by 1038 chemical compounds [4] are presented. The first set of models was developed with 1D and 2D descriptors and the second set of models included 2D and 3D descriptors in addition to 1D, for which variability was investigated with respect to the conformational change of the input molecule representing a random event in the model. A comparison of the models of these two groups revealed that before including 3D descriptors in QSAR models, it is necessary to analyse the justification of their introduction in the model, taking into account the nature of the problem, the quality of the 3D descriptors themselves and their reliability.

Acknowledgements: This research is supported by the Croatian Ministry of Science and Education and the German Academic Exchange Service (DAAD) for the ChemInf4Biosciences project, and by the Croatian Government and the EU through the ERD Fund - the Competitiveness and Cohesion Operational Programme (KK.01.1.1.01) The Scientific Centre of Excellence for Marine Bioprospecting - BioProCro.

REFERENCES

- [1] OECD, 2007, https://read.oecd-ilibrary.org/environment/guidance-document-on-the-validation-of-quantitative-structure-activity-relationship-q-sar-models_9789264085442-en, May 20, 2023.
- [2] M. Zeeman, et al., *SAR QSAR Environ. Res.* 3 (1995) 179–201.
- [3] I. Sushko, et al. *Journal of Computer-Aided Molecular Design* **25** (2011) 25, 533–554.
- [4] R. Ancuceanu, B. Tamba, C. S. Stoicescu, M. Dinu, *Int. J. Mol. Sci.* 21 (2019) 19.

**AUTHORS
AND
PARTICIPANTS**

AUTHORS

Asnaghi, Giacomo	17	Oršolić, Davor.....	30
Batista, Jadranko.....	20	Poljak, Marina.....	29
Bettani, Giacomo	17	Prątnicki, Filip	14
Bojović, Viktor	20, 30	Primožič, Ines	29
Borovec, Jakov	29	Raj, Ankit.....	18
Bren, Urban	21	Restrepo, Guillermo.....	19
Cardani, Edoardo	17	Serra, Gianluca	17
Chao, Yen-Bang	18	Solà, Miquel.....	12
Cioslowski, Jerzy.....	16	Steinbeck, Christoph.....	20, 30
Corno, Gabriele	17	Szysko, Bartosz.....	13
Cutore, Francesco	17	Šmuc, Tomislav	30
Došlić, Nađa	25	Tommasini, Matteo.....	17
Došlić, Tomislav.....	24	Tošović, Jelena	21
Fijan, Domagoj	21	Trzaskowski, Bartosz.....	13
Hrenar, Tomica.....	29	Vančik, Hrvoj	15
Jukić, Marko	21	Witek, Henryk A.....	18
Lučić, Bono	20, 30	Židak, Nenad.....	30
Ori, Ottorino	28		

PARTICIPANTS

Bren, Urban

Cioslowski, Jerzy

Došlić, Nađa

Došlić, Tomislav

Lučić, Bono

Namjesnik, Danijel

Ori, Ottorino

Poljak, Marina

Prątnicki, Filip

Restrepo, Guillermo

Serra, Gianluca

Solà, Miquel

Trzaskowski, Bartosz

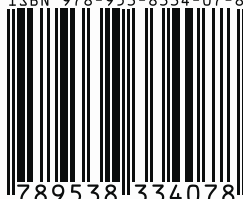
Vančik, Hrvoj

Witek, Henryk

Židak, Nenad

M**C**²**h****e****m**
a . **o** . .
t . . **m** .
h . . . **p**

ISBN 978-953-8334-07-8



9 789538 334078