

Table of Contents

Dynamical Temperature from the Phase Space Trajectory ANDRÁS BARANYAI	1-3
Separation of Gases by Membranes: The Effects of Pollutants on the Stability of Membranes NÁNDOR NEMESTÓTHY	5-8
Selective Removal of Hydrogen Sulphide from Industrial Gas Mixtures Using Zeolite NaA TAMÁS KRISTÓF	9-15
Dynamic Simulator-Based Apc Design For A Naphtha Redistillation Column LÁSZLÓ SZABÓ, KLÁRA KUBOVICSNÉ STOCZ, LAURA SZABÓ, SÁNDOR NÉMETH, AND FERENC SZEIFERT	17-22
Replacement of Biased Estimators with Unbiased Ones in the Case of Student's t-Distribution and Geary's Kurtosis GERGELY TÓTH AND PÁL SZEPESVÁRY	23-27
Formation, Photophysics, Photochemistry and Quantum Chemistry of the Out-Of-Plane Metalloporphyrins ZSOLT VALICSEK, MELITTA P. KISS, MELINDA A. FODOR, MUHAMMAD IMRAN, AND OTTÓ HORVÁTH	29-36
Improvement Possibilities of Heterogeneous Photocatalysis with the Aim of In-Field Use ORSOLYA FÓNAGY, PÉTER HEGEDŰS, ERZSÉBET SZABÓ-BÁRDOS, ANNAMÁRIA DOBRÁDI, AND OTTÓ HORVÁTH	37-48
Adapting the SDEWES Index to Two Hungarian Cities VIKTOR SEBESTYÉN, VIOLA SOMOGYI, SZANDRA SZÓKE, AND ANETT UTASI	49-59
Surface Energy Heterogeneity Profiles of Carbon Nanotubes with a Copolymer-Modified Surface Using Surface Energy Mapping by Inverse Gas Chromatography FRUZSINA GERENCSÉR, NORBERT RIEDER, CSILLA VARGA, JENŐ HANCSÓK, AND ANDRÁS DALLOS	61-66
Initial Electrical Parameter Validation in Lead-Acid Battery Model Used for State Estimation BENCE CSOMÓS, DÉNES FODOR, AND GÁBOR KOHLRUSZ	67-71
Simulating Ion Transport with the NP+LEMC Method. Applications to Ion Channels and Nanopores. DÁVID FERTIG, ESZTER MÁDAI, MÓNIKA VALISKÓ, AND DEZSŐ BODA	73-84

IN MEMORIAM JÁNOS LISZI

This issue is dedicated to Professor János Liszi who was the Rector of the University of Veszprém in 1989-1995. The University acquired this name during his term in an era of severe political and economic challenges (today: University of Pannonia). Under his leadership the University took on these challenges and has been modernized to become an institution with a wider scope. He was the head of the Department of Physical Chemistry for many years. The Department was also reshaped during this period.

His early death was unexpected and mourned.

He influenced the life of many and left his mark on his Alma Mater in Veszprém, but he also had an impact on other institutions as the commemoration of László Rác below indicates. László Rác was the Head of the Chemistry Department in Eger when Professor Liszi joined to teach Physical Chemistry there.

It was an honor to organize and edit this issue to salute my respected supervisor.

Veszprém, 10 October 2017.

Dezső Boda
Issue Editor



THE EGER-CONNECTION OF PROFESSOR LISZI

He was a tough and righteous man towards himself just as towards his colleagues in Veszprém. I can say that the same was true regarding his relationship with the colleagues and students of the Chemistry Department at the Eszterházy Károly College in Eger. That is why we got to like him so much.

When a position for teaching Physical Chemistry became available in our Department, it was not easy to find the right person to fill the boots left by an excellent colleague who had held this position for 30 years. According to the urgent need to yield the proper knowledge to our students, our director indicated to Professor Liszi that there was a problem to be solved in Eger. He did not hesitate and drove 230 km to Eger right away. He assisted with our teaching work for many years.

A fruitful professional relationship had been established between Veszprém and Eger. We gladly accepted the support offered by Professor Liszi. Beyond his Physical Chemistry courses in the curriculum, he regularly gave seminars about various topics not only to our students but also to colleagues and students from other departments. An example that I witnessed characterizes his professional work. After a class, a student from the Physics programme approached him and said, "I will enter the Chemistry programme because this was so beautiful and logical". And he did. He became a Physics-Chemistry teacher and now works in the Physics Department.

Since 2015, our College has functioned as a University. I am sure that the support of Professor Liszi has contributed to the development of our institution considerably. He has also contributed to my personal development because he encouraged me to defend my habilitation thesis. He also contributed to the education of our team beyond the field of Physical Chemistry. During excursions to Tihany and at The Benedictine Order he acted as a guide and demonstrated his vast knowledge of History and classical culture.

The expertise of our Department, making wine, also belongs to classical culture. It was a pleasure to host him and his wife for wine tastings on several occasions. At such times, we discussed the science of wine-making. Time flew and we got onto the subject of the thermodynamics of multicomponent homogeneous systems after a few glasses of wine. Sadly the Blessing of St. John had to come early because he had an early start the following morning. The minimum 1000 m session in the swimming pool could not be left unperformed.

Together with him we profess that technical details can only come after the firm scientific basics have been laid down. He achieved a lot for higher education and research in Eger. We are grateful for his efforts

We respected and loved him not only for his knowledge, but also for his humanity.

Eger, 12 October 2017

László Rác

DYNAMICAL TEMPERATURE FROM THE PHASE SPACE TRAJECTORY

ANDRÁS BARANYAI*

Department of Theoretical Chemistry, Eötvös Loránd University, P. O. Box 32, Budapest, 1518, HUNGARY

In classical statistical mechanics the trajectory in phase space represents the propagation of a classical Hamiltonian system. While trajectories play a key role in chaotic system theory, exploitation of a single trajectory has yet to be considered. This work shows that for ergodic dynamical systems the dynamical temperature can be derived using phase space trajectories.

Keywords: temperature, statistical mechanics

A Hamiltonian system of N mass points is related to the microcanonical ensemble of statistical mechanics when the volume, V , and the energy, E , of the system are fixed. A microscopic state of the system is a point in phase space represented by a $6N$ -dimensional vector, $\Gamma \equiv (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$. The time evolution of this point is described as a trajectory in phase space. In microcanonical ensembles, since the energy is fixed, the evolution of the system is restricted to a $6N-1$ dimensional hypersurface. The logarithm of the area of this hypersurface yields the entropy, S , of the system. Since the entropy of a stable thermodynamic system is a monotonically increasing function of the internal energy, the accessible phase space increases with energy. This provides the opportunity to determine the thermodynamic derivative, $(\partial S / \partial E)_{V,N} = 1/T$, where T is the absolute temperature.

The trajectories within the limit of infinite time perfectly cover the $6N-1$ dimensional hypersurface according to the ergodic hypothesis. This means that the length of the trajectory for a period of time, τ , is proportional to the microcanonical entropy:

$$S \propto k \ln \left(\lim_{\tau \rightarrow \infty} \int_0^\tau dt \sqrt{\dot{\Gamma}^2} \right). \quad (1)$$

Eq.(1) is a line integral measuring the length of the path of the moving system in phase space. This equation can be rewritten as:

$$S \propto k \ln \left(\lim_{\tau \rightarrow \infty} \int_0^\tau dt \sqrt{\left(\frac{\partial H}{\partial \Gamma} \right)^2} \right) = k \ln \left(\lim_{\tau \rightarrow \infty} \int_0^\tau dt \left| \frac{\partial H}{\partial \Gamma} \right| \right). \quad (2)$$

Let us consider two Hamiltonian systems, H and H' , with only a minor difference between their energy parameters, E and E' . The distance covered in phase space by these systems is different because the two energies belong to different areas of the hypersurface to be covered by the trajectories. If the difference in entropies between the two systems is calculated, the quotient in the argument of the logarithm makes it possible to use equality instead of a mere proportionality.

$$\begin{aligned} S(E') - S(E) &= k \ln \left(\frac{\left(\lim_{\tau \rightarrow \infty} \int_0^\tau dt \left| \frac{\partial H'}{\partial \Gamma} \right| \right)}{\left(\lim_{\tau \rightarrow \infty} \int_0^\tau dt \left| \frac{\partial H}{\partial \Gamma} \right| \right)} \right) = \\ &= k \ln \frac{\left\langle \frac{\partial H'}{\partial \Gamma} \right\rangle_\tau}{\left\langle \frac{\partial H}{\partial \Gamma} \right\rangle_\tau} = k \ln \frac{\left\langle \frac{\partial H'}{\partial \Gamma} \right\rangle}{\left\langle \frac{\partial H}{\partial \Gamma} \right\rangle}. \end{aligned} \quad (3)$$

In Eq.(3) the time integral was replaced with the product of time and the time-average of the integrands. The Hamiltonian, H' , can be approximated as

$$H' \cong H + \frac{\partial H}{\partial \Gamma} d\Gamma. \quad (4)$$

Then

$$\frac{\partial H'}{\partial \Gamma} \cong \frac{\partial H}{\partial \Gamma} + \frac{\partial^2 H}{\partial \Gamma^2} d\Gamma. \quad (5)$$

*Correspondence: baranyai@chem.elte.hu

Using Eq.(5), Eq.(3) can be rewritten as

$$S(E') - S(E) \cong k \ln \left\langle \frac{\left| \frac{\partial H}{\partial \Gamma} + \frac{\partial^2 H}{\partial \Gamma^2} d\Gamma \right|}{\left| \frac{\partial H}{\partial \Gamma} \right|} \right\rangle \cong k \left\langle \frac{\frac{\partial^2 H}{\partial \Gamma^2} d\Gamma}{\left| \frac{\partial H}{\partial \Gamma} \right|} \right\rangle. \quad (6)$$

In the second equality of Eq.(6), the $\ln(1+x) \cong x$ approximation was exploited.

To obtain the temperature, division with the energy difference should be included:

$$\frac{S(E') - S(E)}{E' - E} \cong k \frac{\left\langle \frac{\frac{\partial^2 H}{\partial \Gamma^2} d\Gamma}{\left| \frac{\partial H}{\partial \Gamma} \right|} \right\rangle}{\left\langle \frac{\partial H}{\partial \Gamma} \right\rangle} = k \left\langle \frac{\frac{\partial^2 H}{\partial \Gamma^2}}{\left| \frac{\partial H}{\partial \Gamma} \right|^2} \right\rangle = \frac{1}{T}. \quad (7)$$

The dynamical temperature was derived first by Rugh as a time average of $\nabla \cdot (\nabla H / |\nabla H|^2)$ on the energy surface [1].

$$\frac{1}{kT} = \left\langle \frac{\nabla \cdot \nabla H}{|\nabla H|^2} \right\rangle. \quad (8)$$

Later, Butler *et al.* applied a similar approach and derived Eq.(8) only for the configuration part of the phase space and also checked the performance of the method numerically [2]. The essence of both derivations was to transform the phase-space vector, Γ , from system E into Γ' of system E' by a vector containing the Hamiltonian gradient. In this way a relationship was established between the phase-space points of the two systems.

Our result uses the average speed of evolution of trajectories in phase space, therefore, a Hamiltonian gradient is not required to connect the two sets of phase-space points. Expansion of the energy in Eq.(4) is sufficient to relate the two trajectories. Certainly, the results are identical which proves the validity of this alternative approach while, at the same time, it is a nice example of ergodicity: the method which uses the difference in area of hypersurfaces in an ensemble-average fashion leading to results that are identical to

the method that uses relative trajectory propagation in time.

The explicit form of Eq.(7) can be written as

$$\frac{1}{kT} = \left\langle \frac{\sum_{i=1}^N 3/m + \nabla_q^2 \Phi}{\sum_{i=1}^N \mathbf{p}_i^2 / m + |\nabla_q \Phi|^2} \right\rangle, \quad (9)$$

where Φ is the position-dependent potential in the Hamiltonian and m is the mass of a particle. Eq.(9) contains two terms that feature both in the numerator and denominator. The quotient of the first terms is the well-known kinetic temperature and the quotient of the second terms is the configurational temperature. In actual calculations these two temperatures can be calculated contrary to the full form of Eq.(9) which contains dimensional discrepancies in both the numerator and denominator. The configurational temperature calculation was suggested as an algorithmic check for Monte Carlo computer simulations where the Boltzmann temperature is an input parameter [2].

For completeness, the third, very simple method of deriving the configurational temperature is mentioned [3]. The previous derivations did not state whether the two temperatures, the kinetic and configurational, are equal. Using a trivial derivation it is shown that the two quotients are equal [3]. A further advantage of this method is that there is no need for the condition of $N \rightarrow \infty$ to accept the validity of a microcanonical result in terms of the canonical ensemble. The temperature from the momenta of the particles is

$$T = \frac{1}{3Nkm} \sum_{i=1}^N \mathbf{p}_i^2. \quad (10)$$

If this temperature remains constant over time:

$$\dot{T} = C \sum_{i=1}^N \mathbf{p}_i \dot{\mathbf{p}}_i = C \sum_{i=1}^N \mathbf{p}_i \mathbf{F}_i = 0, \quad (11)$$

where, for the sake of simplicity, the constant factor is denoted in front of the sum as C and $\mathbf{F}_i \equiv -(\partial \Phi / \partial \mathbf{q}_i)$ is the Newtonian force, it is also possible to write that

$$\begin{aligned} \ddot{T} &= C \sum_{i=1}^N (\dot{\mathbf{p}}_i \mathbf{F}_i + \mathbf{p}_i \dot{\mathbf{F}}_i) = \\ &= C \sum_{i=1}^N \left(\sum_{\alpha=x,y,z} F_{\alpha i}^2 + \frac{\partial \mathbf{F}_i}{\partial \alpha} \frac{p_{\alpha i}^2}{m} \right) = 0. \end{aligned} \quad (12)$$

At equilibrium in isotropic systems the different Cartesian directions are equivalent and there is no correlation between velocities and position-dependent quantities. Thus,

$$\frac{3m}{\langle p^2 \rangle} \equiv kT = \frac{\langle \nabla_{\mathbf{q}}^2 \Phi \rangle}{\langle |\nabla_{\mathbf{q}} \Phi|^2 \rangle}. \quad (13)$$

Eq.(13) can be written in this form when $\langle T \rangle$ is manipulated. The results would be the same. Therefore, if the expectation value of the temperature is constant, the kinetic and configurational temperatures are equivalent.

As for the temperature derivation from the phase-space trajectory it is important to note that the $\tau \rightarrow \infty$ condition is essential. Finite segments of trajectories do not cover the hypersurface completely and only contain minimal information about the temperature and force distributions of the system.

REFERENCES

- [1] Rugh, H.H.: Dynamical approach to temperature, *Phys. Rev. Lett.*, 1997 **78**(5), 772-774 DOI:10.1103/PhysRevLett.78.772
- [2] Butler, B.D.;Ayton, G.A.;Jepps, O.G.;Evans, D.J.: Configurational temperature: Verification of Monte Carlo simulations, *J. Chem. Phys.*, 1998 **109**, 6519-6522 DOI:10.1063/1.477301
- [3] Baranyai, A.: On the configurational temperature of simple fluids, *J. Chem. Phys.*, 2000 **112**, 3964-3966 DOI:10.1063/1.480995

SEPARATION OF GASES BY MEMBRANES: THE EFFECTS OF POLLUTANTS ON THE STABILITY OF MEMBRANES

NÁNDOR NEMESTÓTHY*

Research Institute of Bioengineering, Membrane Technology and Energetics, University of Pannonia, Egyetem u. 10, Veszprém, H-8200, HUNGARY

The long-term stability of membranes is determined mainly by their sensitivity to pollutants. Their stability was tested using a novel, multichannel measuring system, which is based on pressure differences. This measuring system is suitable to determine the changes in permeability of polymer membranes. The damaging effects of H₂S, BTX and n-dodecane were investigated in terms of polyimide gas separation membranes using nitrogen gas.

Keywords: multichannel test equipment, pressure difference, hydrogen sulphide, BTX

1. Introduction

Previously, it was thought that the stability of membranes is determined by the mechanical stress (shear) and the natural aging of polymers. Recently, however, it has been confirmed that their stability is limited mainly by the sensitivity of membranes to certain pollutants. These aggressive compounds, pollutants, e.g. chlorine, hydrogen sulphide, hydrocarbons, etc., may damage the structure of the polymer, thus its physical and chemical properties change, and consequently the permselectivity of the membrane changes, as well.

It is known that polymeric reverse osmosis membranes are sensitive to strong oxidising agents, especially chlorine compounds [1], therefore, intensive research has been conducted to avoid or at least reduce any damage [2-3]. Similar levels of membrane degradation are observed in proton-exchange membrane (PEM) fuel cells and batteries containing membranes, where oxidising compounds are in contact with the membranes, as well [4-5].

The stability of polymeric gas separation membranes has hardly been investigated. The long-term effects of H₂S on inorganic membranes has been studied by Australian researchers at low concentrations (50 ppm) [6]. However, H₂S may not only cause long-term, but immediate damage, mainly in the form of swelling, which strongly influences the gas transport properties of membranes. Koros and co-workers presented the effects of extremely high H₂S concentrations on a polymer membrane (50,000 - 100,000 ppm) [7].

In the field of membrane technology, sensitivity can be measured by a sort of effectiveness unit. The

product of the concentration and time period yields a value where the effectiveness of the membrane decreases from 95 to 90 and then to 70% of its original value (e.g. 1000 ppm*hour means that the membrane was exposed to 1000 ppm of pollutant for 1 hour, or 0.1 ppm of pollutant for 10,000 hours during the tests).

According to the literature these types of measurements have yet to be published for gas separation membranes, thus the aim of this work was to design, construct and operate a piece of test equipment that conducts reliable laboratory tests.

For the determination of stability, direct and indirect methods can be used to measure the gas volumes passing through the membranes. Direct methods are usually preferred, and – if the composition of the gas is known – are more exact than indirect ones. However, when the gas composition varies and small amounts of gases need to be measured, indirect methods are often more suitable. In this work an indirect method based on a pressure differential technique was chosen, where the pressure of a closed vessel is measured and the varying pressure yields information about the volume of the gas passing through the membrane.

During the investigation the effects of pollutants on the permeability of nitrogen was to be studied. The following pollutants were used:

- compounds containing sulphur at associated gas
- BTX mixture (benzene, toluene, xylene)
- heavy hydrocarbons

In this research the aim was to determine quantitatively the effects of pollutants on the membranes to define the tolerance range of particular membranes.

*Correspondence: nemesn@almos.uni-pannon.hu



Figure 1. The small modules constructed for the tests



Figure 3. The test system

2. Experimental

For this series of measurements, polyimide gas separation membranes (synthesised by UBE) were used. They were taken from a hollow fibre module and can accurately model the properties of industrial gas separation membranes. From the hollow fibres small modules containing 6 capillaries were constructed (Fig.1) and their ends were closed, thus their tests were carried out in a “sack” configuration.

In the design of the test system it was important that several parallel measurements should be conducted and the measuring channels combined with each other. The scheme of a measuring channel can be seen in Fig.2

The gas was introduced into the measuring system through valve V1 (which can be adjusted by valve V2 if necessary). Before measurements were taken the pressure of the vessel was checked by pressure transducer PT1. To start the test the pressure was adjusted by regulating valve PV1, which was checked by pressure transducer PT2. Then the membrane was installed into the thermostatic system in a way that ensured its mobility was not restricted, thus the permeation of gas could not influence the flux.

A photograph of the measuring system is shown in Fig.3. For the permeability measurements nitrogen gas from a cylinder was used (99.5%; Messer Hungarogáz Kft., Hungary).

The permeability of the membrane was determined from the pressure of the vessel and the transmembrane pressure measured on-line during the experiments. In the experiments, H₂S, methyl mercaptan and ammonia (compounds containing S or N), oleic acid, ethyl

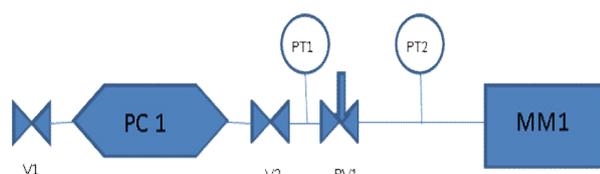


Figure 2. The scheme of the test system

alcohol, moreover, a benzene-toluene-xylene mixture (BTX) and n-dodecane (as a heavier hydrocarbon) were used as pollutants.

For the stability experiments the small membrane modules were put in a closed vessel (Fig.4) where the headspace was saturated with the given pollutant. The vessels were placed in a thermostatic incubator at 27 °C usually for between 1 and 7 days.

Certain materials (e.g. BTX) damaged the epoxy resin glue used to adhere the fibres of the membranes, therefore, these experiments were repeated using poly-ether-sulfone glue instead.

3. Results

The nitrogen permeability of the membranes was determined before and after the incubations.

In the preliminary experiments, the ammonia solution and methyl mercaptan severely damaged the surface of the membranes, thus no flux could be measured. Oleic acid and ethyl alcohol hardly influenced the flux, while BTX, H₂S and n-dodecane changed the permeability of nitrogen considerably.

For further investigation of the pollutants an experimental design was constructed, using appropriate



Figure 4: The membranes in the closed vessel

Table 1. The parameters of the experimental design

pollutant	C _{min} ppm	C _{av} ppm	C _{max} ppm	t _{min} d	t _{av} d	t _{max} d
H ₂ S	100.000	300.000	500.000	1	3.5	7
BTX mixt.	1.000	750	500	1	3.5	7
dodecane	1.000	5.500	10.000	1	3.5	7

statistical methods. The parameters selected were the concentrations of the pollutants (minimum, maximum and average) and the incubation time (minimum, maximum and average). The Statistica 8 computer program was applied to the design presented in Table 1.

Firstly the effect of H₂S on the nitrogen permeability of the membranes was measured. The experimental results are presented in Fig.5.

It can be seen that the permeability of nitrogen increased even when the concentrations of pollutants were low and rose by using higher concentrations and longer periods of exposure.

From the H₂S concentration and the incubation time it was possible to calculate a special parameter of exposure with the unit of ppm*h. Permeability was presented as a function of this parameter (Fig.6), where an almost linear relationship was observed.

The results suggest that the process can be described as a first-order reaction, which means that no safe limit can be determined where H₂S is regarded as harmless, on the contrary, it should be considered at all times.

The effect of the BTX mixture was studied using a similar methodology. The results are summarized in Fig.7. This figure shows that the permeability of nitrogen increased at low concentrations of the BTX mixture. At higher concentrations and over longer periods of time, no further significant changes were observed.

The exposure parameter was also calculated in the unit of ppm*h and permeability was presented as its function (Fig.8). The diagram can be described as a saturation-type curve.

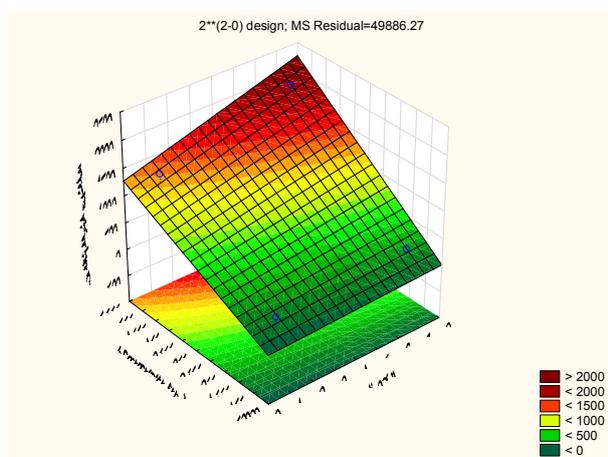


Figure 5. The effect of H₂S on the nitrogen permeability

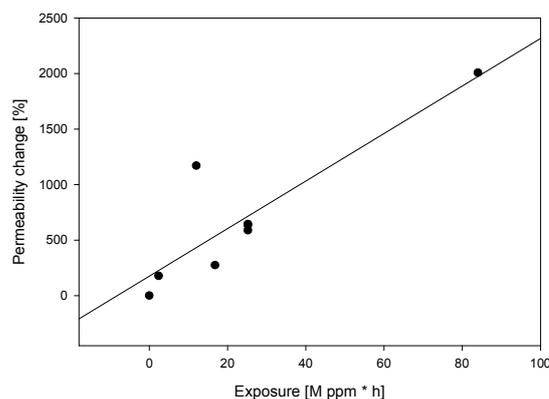


Figure 6. Permeability changes against exposure to H₂S

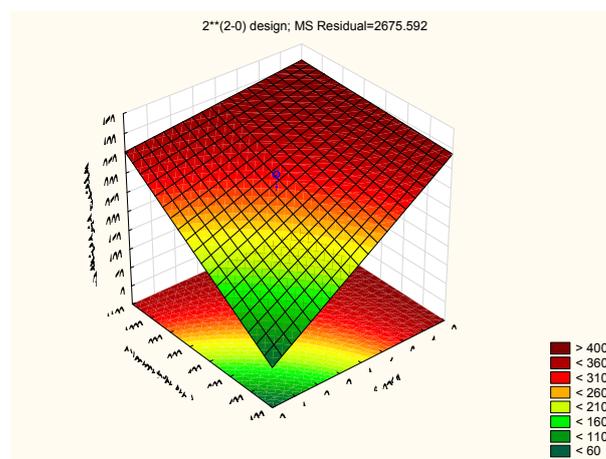


Figure 7: The effect of the BTX mixture

In the last series of experiments, the effect of exposure to n-dodecane was investigated experimentally. The results are presented in Fig.9.

n-dodecane caused – unlike H₂S and BTX – a reduction in the permeability of nitrogen even at low concentrations. The flux fell to zero at higher concentrations and over longer incubation periods.

Permeability was investigated as a function of exposure (Fig.10). The process can be described as a first-order reaction, thus the effect of heavier hydrocarbons, e.g. n-dodecane, should always be considered.

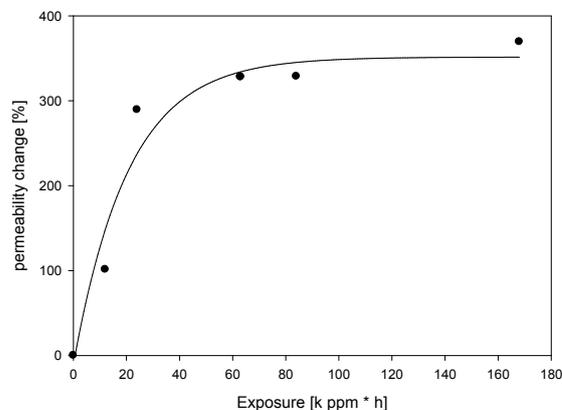


Figure 8: Permeability changes against BTX exposure

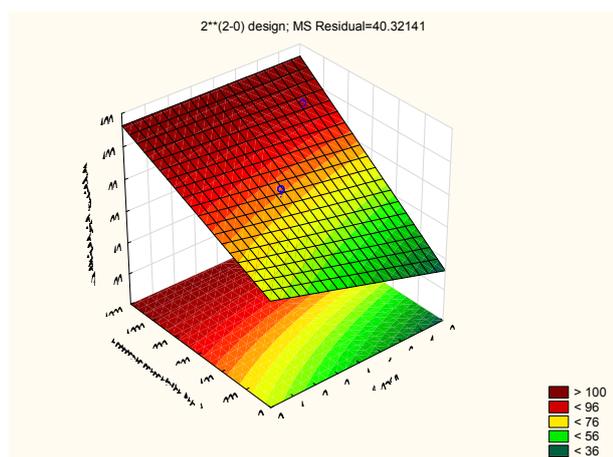


Figure 9: The effect of n-dodecane

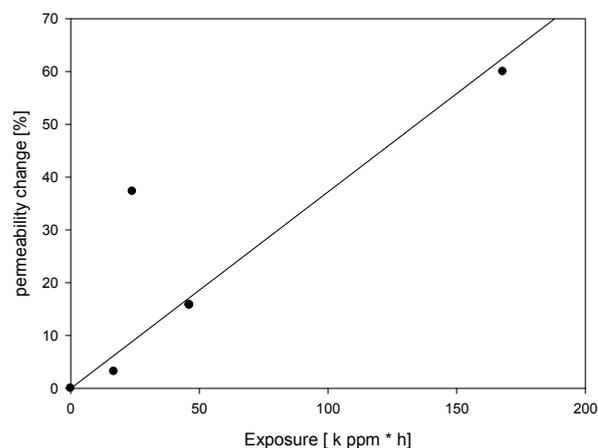


Figure 10. Change in permeability against exposure to n-dodecane

4. Conclusion

The long-term stability of polyimide gas separation membranes was tested against various pollutants: H₂S, a BTX mixture and n-dodecane. These compounds significantly affected the nitrogen permeability of the membranes which were described by using a special parameter of exposure. It was found that H₂S and the BTX mixture increased the permeability, while n-dodecane reduced the permeability of the membranes. Further investigations are planned to investigate the effect of other pollutants, moreover, to determine the permeability of additional gases, e.g. carbon dioxide, methane, etc.

Acknowledgement

We acknowledge the financial support of Széchenyi 2020 under the EFOP-3.6.1-16-2016-00015 project. This research was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

REFERENCES

- [1] Glater, J.; Hong, S.K.; Elimelech, M.: The search for a chlorine-resistant reverse osmosis membrane, *Desalination*, 1994 **95**(3), 325-345 DOI: 10.1016/0011-9164(94)00068-9
- [2] Zhang, Y.; Zhao, C.; Yan, H.; Pan, G.; Guo, M.; Na, H.; Liu, Y.: Highly chlorine-resistant multilayer reverse osmosis membranes based on sulfonated poly(arylene ether sulfone) and poly(vinyl alcohol), *Desalination*, 2014 **336**, 58-63 DOI: 10.1016/j.desal.2013.12.034
- [3] Gohil, J.M.; Suresh, A.K.: Chlorine attack on reverse osmosis membranes: Mechanisms and mitigation strategies, *J. Membr. Sci.*, 2017 **541**, 108-126 DOI: 10.1016/j.memsci.2017.06.092
- [4] Huang, X.; Pu, Y.; Zhou, Y.; Zhang, Y.; Zhang, H.: In-situ and ex-situ degradation of sulfonated polyimide membrane for vanadium redox flow battery application, *J. Membr. Sci.*, 2017 **526**, 281-292 DOI: 10.1016/j.memsci.2016.09.053
- [5] Lopicque, F.; Belhadj, M.; Bonnet, C.; Pauchet, J.; Thomas, Y.: A critical review on gas diffusion micro and macroporous layers degradations for improved membrane fuel cell durability, *J. Power Sources*, 2016 **336**, 40-53 DOI: 10.1016/j.jpowsour.2016.10.037
- [6] Uhlmann, D.; Smart, S.; Diniz da Costa, J.C.: H₂S stability and separation performance of cobalt oxide silica membranes, *J. Membr. Sci.*, 2011 **380**(1), 48-54 DOI: 1016/j.memsci.2011.06.025
- [7] Kraftschik, B.; Koros, W.J.; Johnson, J.R.; Karvan, O.: Dense film polyimide membranes for aggressive sour gas feed separations, *J. Membr. Sci.*, 2013 **428**, 608-619 DOI: 10.1016/j.memsci.2012.10.025

SELECTIVE REMOVAL OF HYDROGEN SULPHIDE FROM INDUSTRIAL GAS MIXTURES USING ZEOLITE NaA

TAMÁS KRISTÓF*

Department of Physical Chemistry, Institute of Chemistry, University of Pannonia, Egyetem u. 10., Veszprém, H-8200, HUNGARY

Hydrogen sulphide removal from simple gas mixtures using a highly polar zeolite was studied by molecular simulation. The equilibrium adsorption properties of hydrogen sulphide, hydrogen, methane and their mixtures on dehydrated zeolite NaA were computed by Grand Canonical Monte Carlo simulations. Existing all-atom intermolecular potential models were optimized to reproduce the adsorption isotherms of the pure substances. The adsorption results of the mixture, also confirmed by IAST calculations, showed very high selectivities of hydrogen sulphide to the investigated non-polar gases, predicting an outstanding performance of zeolite NaA in technological applications that target hydrogen sulphide capture.

Keywords: Hydrogen sulphide, Zeolite, Selectivity, Gas mixture, Molecular simulation

1. Introduction

Hydrogen sulphide is a highly toxic, acidic and corrosive substance. It is present naturally in landfills, natural and biogases, as well as in several synthesis gases. One of its main anthropogenic sources is the processing of crude oil in industrial refineries, where hydrodesulfurization (HDS) of a variety of streams (e.g. engine fuels) produces hydrogen sulphide-containing gas mixtures, which need to be purified. The economic removal of hydrogen sulphide is a long-standing task of the oil and gas industry. Adsorptive separation involves the use of solid substrates with a specific affinity with particular compounds of the mixture. Zeolites have been applied as catalysts in the petrochemical industry for a relatively long time and these materials can also be used for purification/separation purposes. Zeolites are crystalline aluminosilicates consisting of a three-dimensional framework of SiO_4 and AlO_4 tetrahedra of a highly regular porous structure [1]. The typical size of zeolitic micropores is similar to that of many small molecules. In contrast to various other adsorbents, zeolites generally endure high temperatures and pressures well, and can tolerate harsh chemical environments. The Si/Al ratio is a key factor in the application of zeolites. Zeolites with lower Si/Al ratios are more hydrophilic, whereas high-silica zeolites often possess fewer structural defects. These latter adsorbents are preferred in the separation of non-polar gases.

Zeolite NaA is a synthetic microporous zeolite which accommodates extraframework Na^+ ions. It exhibits an especially high affinity with small polar

molecules such as water. The adsorption and separation properties of zeolite NaA have already been examined in several experimental [2-6] and theoretical/simulation [7-16] works. In our laboratories, the adsorption characteristics of zeolite NaA and its performance as a drying agent by classical atomistic simulations [10, 12-14] were studied, and new intermolecular potential models for this zeolite [12-14] proposed. Our models were optimized for the study of the selective adsorption of water from its mixtures with less polar or non-polar molecules like simple alcohols, carbon monoxide, hydrogen and methane.

In this paper, the selective removal of hydrogen sulphide by zeolite NaA is investigated. Molecular simulation predictions for mixture adsorption from two- and three-component gas mixtures containing hydrogen sulphide (H_2S), hydrogen (H_2) and methane (CH_4) are presented.

2. Models and simulation details

Zeolite NaA [17-18] is of LTA framework type, the structure of which belongs to the Fm-3c space group with a lattice parameter of 2.4555 nm. The three-dimensional cubic arrangement of its framework atoms is comprised of three kinds of rings with four (4R), six (6R) or eight (8R) oxygen atoms. The interconnection of 4R and 6R rings forms nearly spherical cages (sodalite cages) and these cages are linked by oxygen bridges, shaping straight channels of supercages with a maximum diameter of about 1.2 nm. The standard type of zeolite NaA has a Si/Al ratio of 1.

In the present study, the unit-cell composition of the standard type of zeolite NaA was chosen: it consists

*Correspondence: kristoft@almos.uni-pannon.hu

Table 1. Lennard-Jones energy (ϵ), size parameters (σ) and partial charges (q) for the models used in this work (d is the bond length, k is the Boltzmann constant).

interaction site	σ / nm	$(\epsilon/k) /$ K	$q /$ electron charge	position in the structure/molecule
Na ⁺ (NaA)	0.250	100	0.60	random positions in supercages
Si (NaA)	0.230	22.0	2.40	experimental atomic positions [17]
Al (NaA)	0.240	16.5	1.80	experimental atomic positions [17]
O (NaA)	0.330	190	-1.20	experimental atomic positions [17]
S (H ₂ S)	0.360	122	-	$d_{S-H} = 0.134$ nm
H (H ₂ S)	0.250	50.0	0.21	H–S–H angle: 92°
X _S (H ₂ S)*	-	-	-0.42	$d_{S-X} = 0.03$ nm
C (CH ₄)	0.350	33.21	-0.24	$d_{C-H} = 0.109$ nm
H (CH ₄)	0.250	15.1	0.06	H–C–H angle: 109.47°
H (H ₂)	0.250	15.1	0.4829	$d_{H-H} = 0.0741$ nm
centre of mass (H ₂)	-	-	-0.9658	geometric centre of the linear H ₂ molecule

*off-atom site on the H–S–H angle bisector towards the hydrogen atoms

of 576 framework atoms, namely 96 silicon, 96 aluminium and 384 oxygen atoms. The framework atoms were fixed at the atomic positions measured in X-ray diffraction experiments [17] and the 96 non-framework Na⁺ ions were allowed to move. According to the Löwenstein rule that prohibits Al–O–Al linkages, each AlO₄ tetrahedron of this framework is connected to a SiO₄ tetrahedron.

Realistic and rigid all-atom intermolecular potential models, consisting of Lennard-Jones and Coulombic interaction sites, were used in the simulations. In these models, the interaction sites were fixed at their experimental atomic positions and assigned their Lennard-Jones energy (ϵ) and size (σ) parameters, as well as point charges (q). For zeolite NaA, the model that was developed earlier was modified slightly [14] by adding weak Lennard-Jones interaction sites with realistic size parameters [16, 19] to the (originally) pure Coulombic silicon and aluminium atoms, thus preserving the dominant role of oxygen atoms in the dispersion interactions of the framework. For hydrogen sulphide, a rigid four-site model proposed recently by Shah et al. [20] was adopted, in which the location of the charge parameter of the sulphur atom is offset on the H–S–H angle bisector towards the hydrogen atoms. An OPLS-AA model [21] was used for the methane molecules, and its Lennard-Jones H parameters were also applied to the H₂ molecules, with partial charges on the atomic sites and on the molecular centre of mass [22]. Table 1 lists the potential parameters of the above models. Instead of the generally accepted Lorentz-Berthelot combining rule, the unlike Lennard-Jones interactions were computed by the combining rule proposed by Waldman and Hagler [23]. This combining rule links the behaviour of the unlike

energy parameter ϵ_{ij} to the relative sizes of atoms i and j and yields somewhat smaller values for the parameters ϵ_{ij} and σ_{ij} when $\sigma_{ii} \neq \sigma_{jj}$. Song et al. [24] found that the experimental thermodynamic properties of pure methane can be reproduced more accurately using the Waldman-Hagler combining rule.

Gas adsorption simulations were carried out by the standard grand canonical Monte Carlo methodology [25]. Total pressure p and partial pressures in the gas phase were specified by the chemical potentials of the components; in a diluted gas, these can be calculated from the ideal gas law [12]. The long-range Coulomb interactions were treated with the Wolf method [26-27] using a convergence parameter of $\alpha = 2/r_c$ and cutoff radius of $r_c = L/2$ (L is the length of the simulation box). The simulations involved an equilibration period of 5×10^7 Monte Carlo moves and an averaging period of 2×10^8 moves, consisting of 70% molecular insertion/deletion and 30% molecular translation steps. Since the random insertion of molecules is unable to take into account the inaccessibility of the sodalite cages by multiatomic molecules (the physical diffusion pathways to them), creation of H₂S and CH₄ molecules inside these cages was blocked artificially by placing repulsive dummy atoms at the centres of the cages. As H₂ molecules are sufficiently small to pass through the windows of the sodalite cages, the insertion of H₂ molecules into these cages was permitted. In either case, the transition of molecules into sodalite cages via translational trial moves was not artificially prevented.

In addition to the adsorption loading, the isosteric heat of adsorption was calculated using the equation:

$$q = \left(\frac{\partial H^b}{\partial n^b} \right)_{p,T} - \left(\frac{\partial U^a}{\partial n^a} \right)_{V,T}, \quad (1)$$

where H^b and U^a stand for the residual enthalpy and residual internal energy, respectively, n is the number of moles of the substance in the adsorbed (a) or bulk (b) phases. In the grand canonical ensemble, the second part of the equation can be determined from the particle number fluctuations of the simulation and the cross-correlation of potential energy and particle number fluctuations [28-29]. Assuming the ideal gas adsorbates, the first part of the equation is equal to RT , where R is the gas constant and T is the temperature.

Predictions for mixture adsorption were also made using the ideal adsorbed solution theory (IAST) [30-31], which is an analogue of the ideal Raoult's Law. Using IAST, mixture adsorption loadings at a given T can be obtained from single-component adsorption loadings by determining the bulk pressure of each component p_i^0 at the same spreading pressure π of the adsorbed phase:

$$y_i^a = p y_i^b / p_i^0(\pi). \quad (2)$$

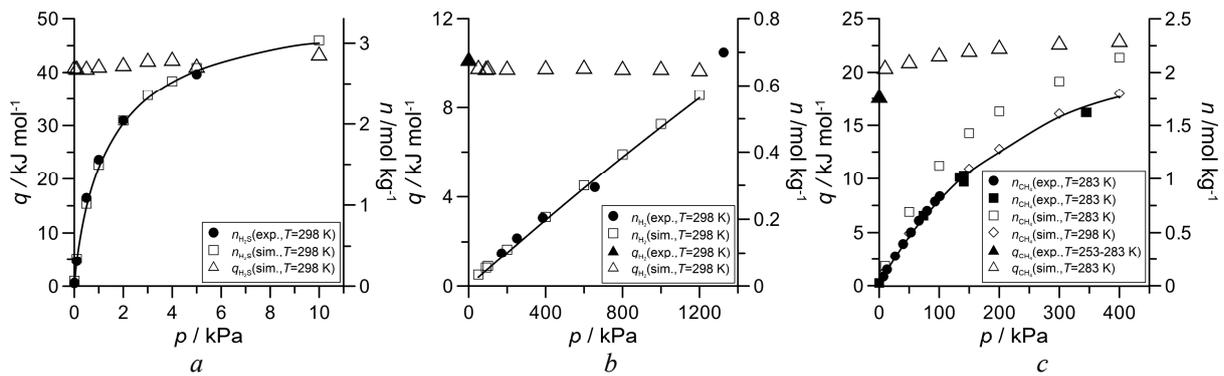


Figure 1. Equilibrium adsorption loading (n) and isosteric heat of adsorption (q) as a function of the bulk-gas pressure for pure hydrogen sulphide (a), hydrogen (b) and methane (c) on zeolite NaA at the temperatures indicated. The statistical uncertainties of the simulations results do not exceed the size of the symbols. The lines connecting simulation data at 298 K are only drawn to guide the eyes. sim.: simulation data; exp.: experimental data.

Here, y_i is the mole fraction of component i , and $p_i^0(\pi)$ is given implicitly by:

$$\pi(p_i^0) = \frac{RT}{A} \int_0^{p_i^0} n_i^a(p) d \ln p, \quad (3)$$

where A is the surface area of the adsorbent.

3. Results and discussion

The intermolecular potential models were tested by determining equilibrium adsorption isotherms for pure H_2S , H_2 and CH_4 on zeolite NaA. Experimental data at 298 K are available for H_2S [32] and H_2 [8] and nearly room-temperature ($T = 283$ K) data for CH_4 were taken from [33]. Fig. 1 shows that the models are largely able to reproduce the experimental adsorption data for these substances. The reproduction of the experimental isotherm is quite good for H_2S and H_2 . In the case of CH_4 , the extent of overestimation of the experimental results at 283 K is considered acceptable, given that the availability of transferable zeolite models that are appropriate as far as adsorption predictions are concerned for both polar and non-polar compounds is rather limited [16]. Furthermore, it is expected that the observed discrepancies between simulation and experimental results for this non-polar component are unable to cause significant errors in terms of mixture adsorption data, where the CH_4 content of the gas phase is low and the adsorption of H_2S is dominant.

The calculated isosteric heat of adsorption data together with available experimental results for H_2 and CH_4 [5] are also plotted in Fig. 1. The pressure-dependence of these data is weak. The general order of $q_{\text{H}_2\text{S}} > q_{\text{CH}_4} > q_{\text{H}_2}$ is in line with expectations, bearing in mind that the isosteric heat of adsorption at low loadings proves the strength of interaction between the zeolite framework and the adsorbate molecules. q values

are considerably higher for polar H_2S than for non-polar substances and the order of magnitude of the former indicates the significance of electrostatic interactions. The relation of $q_{\text{CH}_4} > q_{\text{H}_2}$ can be attributed to the greater polarizability of CH_4 molecules (this is implicitly included in the attracting Lennard-Jones terms of the potential model), and to that the explicitly modelled real quadrupole moment of H_2 molecules is very weak. Considering the greater uncertainties of these simulation results and that the experimental data for H_2 and CH_4 were obtained for zero coverage within a given temperature range, these simulation results also confirm the suitability of the models used in this study.

Equilibrium adsorption selectivities were predicted for typical hydrodesulfurization stream outlets of petroleum refinery units separated by zeolite NaA at near-atmospheric pressures. The studied gas streams were comprised of between 1 and 2% H_2S and ~95% H_2 ; the remaining hydrocarbon content (low alkanes) was represented by the presence of CH_4 in the model mixtures. For comparison, other compositions including very low and reasonably high H_2S contents, as well as low pressure ranges were also investigated. The raw simulation results for the H_2S - H_2 mixtures in comparison with IAST predictions shown in Fig. 2 illustrate well the dissimilar levels of adsorption of the two substances, with the exception of the nearly zero H_2S contents of the bulk mixture. The IAST calculations underestimate the simulation results for H_2 at higher pressures and on the whole overestimate the simulation results for H_2S at lower pressures (for visual reasons, data obtained within the very low pressure range are not presented in this figure). The most accurate estimations were achieved at 10 kPa, which is an impractical parameter for the present applications. Strictly speaking, the hypothesis of IAST which states that the different adsorbate molecules have access to the same adsorbent surface cannot be applied to microporous adsorbents such as zeolite NaA, where the accessible surface area depends on the size of the adsorbate. In light of this, the IAST predictions can be considered to be remarkably accurate.

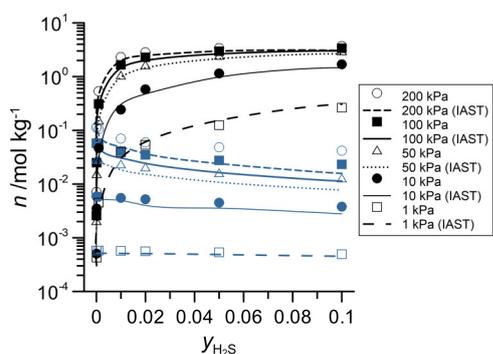


Figure 2. Comparison of the IAST predictions with simulation data for hydrogen sulphide (black) and hydrogen (blue). Equilibrium adsorption loading (n) as a function of the mole fraction of hydrogen sulphide ($y_{\text{H}_2\text{S}}$) in the binary gas phase mixture on zeolite NaA at 298 K and at the pressures indicated. The statistical uncertainties of the simulation results do not exceed the size of the symbols. (For interpretation of the references to colour in this figure, the reader is advised to refer to the online version of this article.)

The calculated equilibrium selectivities are defined as:

$$S = \frac{n_{\text{H}_2\text{S}}^a / n_j^a}{n_{\text{H}_2\text{S}}^b / n_j^b}, \quad (4)$$

where n_j stands for the equilibrium number of moles of H_2 in the investigated two-component mixtures or the sum of the equilibrium numbers of moles of H_2 and CH_4 in the three-component mixtures, as plotted in Fig. 3. On the whole, this zeolite exhibits an exceptional level of selectivity of H_2S to the other substances; this is not surprising given the significant differences between the equilibrium adsorption loadings of the pure components (cf. Fig. 1). In the case of the two-component gas mixtures, the tendency of the data satisfies the criterion that at the low-pressure limit the selectivity as defined here should be independent of the composition of the bulk-gas mixture (it is the quotient of the ratio of the single-particle partition function of the two substances in the adsorbed phase and the ratio of their free-particle partition functions [34-35]). Because of technical reasons, at lower pressures the uncertainties of the selectivity data are relatively large as these data are calculated from simulation results at very low zeolite loadings. At higher pressures the separation efficiency of this zeolite is somewhat weaker. This and the fact that the change with pressure is less intense at lower H_2S contents suggest the existence of a ‘crowding’ effect, which inhibits more strongly the sorption of the larger molecule, H_2S . In the case of the investigated three-component mixtures, the overall picture is similar, but the selectivity values are smaller. This makes sense since the competitive effect of the additional component, CH_4 , for the adsorption sites is stronger. Yet, the values far in excess of 1000 obtained for the

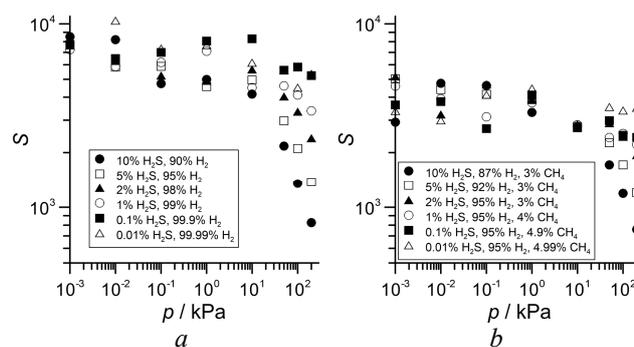


Figure 3. Equilibrium adsorption selectivity (S) on zeolite NaA at 298 K as a function of the bulk-gas pressure for binary (a) and ternary (b) gas mixtures with the compositions indicated.

typical hydrodesulfurization streams (1-2% H_2S and ~95% H_2) are compelling.

From the mixture adsorption data, once again it was verified that electrostatic effects control the adsorbent-adsorbate interactions with this zeolite, which implies that the amount of adsorption of pure H_2 and CH_4 should always be small. Adsorbed mole number data showed that the presence of the non-polar components does not affect the sorption of H_2S in the adsorbed phase. This conclusion is also supported by the heat of adsorption data (not presented) calculated by assuming one-component mixtures (i.e. using Eq.(1)). These data turned out to be simply the amount-weighted average of the q values of pure components and are very close to $q_{\text{H}_2\text{S}}$ at the given pressure. On the other hand, the degree of adsorption of the non-polar substances is reduced by the presence of H_2S . Selectivities under real conditions (at 50, 100 and 200 kPa and with realistic H_2S contents; 1, 2 and 5%) shown in Fig. 4 make this fact obvious. Here, S values obtained from mixture adsorption simulations significantly exceed their counterparts calculated for an ideal case of independent adsorption (i.e. by substituting into Eq.(4) the pure-component adsorption loadings determined at pressures that are equal to the partial pressures of the mixture components). Extensive non-ideality in the adsorbed phase can also be seen from the comparative failure of IAST (which utilizes the assumption that the adsorbed mixture is an ideal solution) to predict the simulation results accurately at near-atmospheric pressures.

It is remarkable that the selectivity of H_2S to the two non-polar substances decreases as the temperature and partial pressure of H_2S in the bulk gas increase. As the adsorption loading of the zeolite rises, steric hindrance plays an increasingly important role, and sorption of the larger H_2S molecules reduces to a greater extent. The impact of an increase in temperature is as expected, e.g. from the higher $q_{\text{H}_2\text{S}}$ values, but the magnitude of decrease in selectivity with temperature changes significantly as a function of the partial pressure of H_2S . Data lines at the two investigated temperatures seem to converge to similar values at higher partial pressures, because the drop in the sorption of H_2S as the temperature increases already becomes

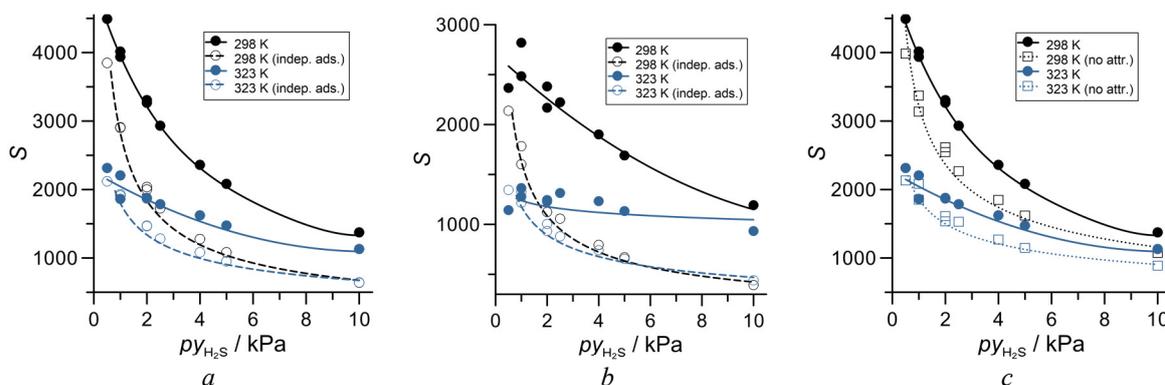


Figure 4. Equilibrium adsorption selectivity (S) as a function of the partial pressure of hydrogen sulphide in the bulk gas for selected binary (a) and ternary (b) gas mixtures on zeolite NaA at 298 K and 323 K, and at bulk gas pressures of 50 kPa, 100 kPa, and 200 kPa. Comparison of the mixture selectivity data with the selectivity data for independent adsorption (indep. ads.; panels a, b) and with mixture selectivity data calculated without adsorbate-adsorbate attractions (no attr., panel c). (For interpretation of the references to colour in this figure, the reader is referred to the web version of the article.)

less significant at higher loadings. The two panels of Fig.4 also illustrate the above-mentioned difference between the data of the binary (a) and ternary (b) mixtures, i.e. the numerical values are smaller for the ternary mixtures. Besides this, the scatter of the points is larger in panel b, because the ratio of n_{H_2} to n_{CH_4} in the bulk gas unavoidably changes as the partial pressure of H_2S increases (for compositions, see Fig.3). Finally, panel c in Fig.4 illustrates the influence of adsorbate-adsorbate attraction on the adsorption characteristics. It was simulated by eliminating the Coulomb potential and the attractive part of the Lennard-Jones potential, but retaining its soft-sphere repulsion potential component, when calculating the instantaneous adsorbate-adsorbate pair interactions in the adsorbed phase. The observed reduction in $n_{\text{H}_2\text{S}}$ and selectivity is sizable enough to establish that like-like attraction is an important factor in the adsorption of H_2S .

4. Conclusion

In this work, molecular simulation predictions for the adsorption of H_2S from simple non-polar gas mixtures of technological interest (hydrodesulfurization stream outlets in petroleum refinery units) on zeolite NaA were presented. The realistic all-atom intermolecular potential models adopted for the computations were validated by comparing the calculated isotherms of the pure substances with available experimental adsorption data. What is especially noticeable here is the matching of the experimental and simulated adsorption loadings for H_2S that was achieved. The investigated zeolite exhibited a remarkable ability to capture H_2S , from either binary or ternary mixtures with non-polar gases, namely H_2 and CH_4 . The interactions between the polar H_2S molecules and the hydrophilic zeolite framework were found to be particularly favourable, and the mixture-adsorption loadings for H_2S essentially agreed with the corresponding pure component loadings (with the exception of the very low H_2S contents of the bulk-

gas mixture). The reverse is true when considering the adsorption of the non-polar gaseous components under technological conditions (at near-atmospheric pressures and with a small proportion of H_2S in the bulk); their bindings to the inner surface sites of the zeolite were suppressed by H_2S . These results can be of practical importance in terms of selectivity. Selectivities of H_2S to the non-polar substances were generally higher at lower H_2S partial pressures in the bulk gas, and well over 1000 for the range of H_2S contents of the typical hydrodesulfurization streams. The obtained order of magnitude of the isosteric heat of adsorption data and the large decrease in selectivity with increasing temperature suggest that electrostatic interactions play a more pronounced role in the selective removal of H_2S by zeolite NaA and the effect of size has only a limited impact. In association with this, it was also revealed that H_2S - H_2S attraction contributes to the preferred adsorption of this substance.

Acknowledgement

Present article was published in the frame of the project GINOP-2.3.2-15-2016-00053 (“Development of engine fuels with high hydrogen content in their molecular structures (contribution to sustainable mobility)”). We gratefully acknowledge the financial support of the Hungarian National Research Fund (OTKA K124353). The author would like to thank Tamás Kovács and Zoltán Ható (Department of Physical Chemistry, Institute of Chemistry, University of Pannonia) for their assistance in terms of data analysis.

REFERENCES

- [1] Auerbach, S.M.; Carrado, K.A.; Dutta, P.K. (Eds.): Handbook of Zeolite Science and Technology (Marcel Dekker, New York) 2003 ISBN: 0-8247-4020-3

- [2] Xu, X.; Yang, W.; Liu, J.; Chen, X.; Lin, L.; Stroh, N.; Brunner, H.: Synthesis and gas permeation properties of an NaA zeolite membrane, *Chem. Commun.* 2000 **0**, 603-604 DOI: 10.1039/B000478M
- [3] Aoki, K.; Kusakabe, K.; Morooka, S.: Separation of gases with an A-type zeolite membrane, *Ind. Eng. Chem.* 2000 **39**, 2245-2251 DOI: 10.1021/ie990902c
- [4] Okamoto, K.; Kita, H.; Horii, K.; Tanaka, K.; Kondo, M.: Zeolite NaA membrane: preparation, single-gas permeation, and pervaporation and vapor permeation of water/organic liquid mixtures, *Ind. Eng. Chem.* 2001 **40**, 163-175 DOI: 10.1021/ie0006007
- [5] Zhu, W.; Gora, L.; van den Berg, A.W.C.; Kapteijn, F.; Jansen, J.C.; Moulijn, J.A.: Water vapour separation from permanent gases by a zeolite-4A membrane, *J. Membrane Sci.* 2005 **253**, 57-66 DOI: 10.1016/j.memsci.2004.12.039
- [6] Yamamoto, T.; Kimb, Y.H.; Kimb, B.C.; Endoa, A.; Thongprachana, N.; Ohmoria, T.: Adsorption characteristics of zeolites for dehydration of ethanol: Evaluation of diffusivity of water in porous structure, *Chem. Eng. J.* 2012 **181-2**, 443-448 DOI: 10.1016/j.cej.2011.11.110
- [7] Lee, S.H.; Moon, G.K.; Choi, S.G.; Kim, H.S.: Molecular dynamics simulation studies of zeolite-A. 3. Structure and dynamics of Na⁺ ions and water molecules in a rigid zeolite-A, *J. Phys. Chem.* 1994 **98**, 1561-1569 DOI: 10.1021/j100057a006
- [8] Akten, E.D.; Siriwardane, R.; Sholl, D.S.: Monte Carlo Simulation of Single- and Binary Component Adsorption of CO₂, N₂, and H₂ in Zeolite Na-4A, *Energy & Fuels* 2003 **17**, 977-983 DOI: 10.1021/ef0300038
- [9] Furukawa, S.; Goda, K.; Zhang, Y.; Nitta, T.: Molecular simulation study on adsorption and diffusion behavior of ethanol/water molecules in NaA zeolite crystal, *J. Chem. Eng. Japan* 2004 **37**, 67-74 DOI: 10.1252/jcej.37.67
- [10] Kristóf, T.; Csányi, É.; Rutkai, G.; Merényi, L.: Prediction of adsorption equilibria of water-methanol mixtures in zeolite NaA by molecular simulation, *Mol. Sim.* 2006 **32**, 869-875 DOI: 10.1080/08927020600934179
- [11] Cosoli, P.; Ferrone, M.; Pricl, S.; Fermeglia, M.: Hydrogen sulphide removal from biogas by zeolite adsorption, Part I-II, *Chem. Eng. J.* 2008 **145**, 86-92 and 93-99 DOI: 10.1016/j.cej.2008.07.034 & 10.1016/j.cej.2008.08.013
- [12] Rutkai, G.; Csányi, É.; Kristóf, T.: Prediction of adsorption and separation of water-alcohol mixtures with zeolite NaA, *Microporous and Mesoporous Mat.* 2008 **114**, 455-464 DOI: 10.1016/j.micromeso.2008.01.044
- [13] Csányi, É.; Kristóf, T.; Lendvay, Gy.: Potential model development using quantum chemical information for molecular simulation of adsorption equilibria of water-methanol (ethanol) mixtures in zeolite NaA-4, *J. Phys. Chem. C* 2009 **113**, 12225-12235 DOI: 10.1021/jp902520p
- [14] Csányi, É.; Ható, Z.; Kristóf, T.: Molecular simulation of water removal from simple gases with zeolite NaA, *J. Mol. Model.* 2009 **18**, 2349-2356 DOI: 10.1007/s00894-011-1253-7
- [15] Sun, Y.; Han, S.: Diffusion of N₂, O₂, H₂S and SO₂ in MFI and 4A zeolites by molecular dynamics simulations, *Mol. Sim.* 2015 **41**, 1095-1109 DOI: 10.1080/08927022.2014.945082
- [16] Vujic, B.; Lyubartsev, A.P.: Transferable force-field for modelling of CO₂, N₂, O₂ and Ar in all silica and Na⁺ exchanged zeolites, *Modelling Simul. Mater. Sci. Eng.* 2016 **24**, 045002 (1-26) DOI: 10.1088/0965-0393/24/4/045002
- [17] Pluth, J.J.; Smith, J.V.: Accurate redetermination of crystal structure of dehydrated zeolite A. Absence of near zero coordination of sodium. Refinement of Si, Al-ordered superstructure, *J. Am. Chem. Soc.* 1980 **102**, 4704-4708 DOI: 10.1021/ja00534a024
- [18] Mikula, A.; Król, M.; Kolezynski, A.: Periodic model of an LTA framework, *J. Mol. Model.* 2015 **21**, 275 (1-9) DOI: 10.1007/s00894-015-2820-0
- [19] Bai, P.; Tsapatsis, M.; Siepmann, J.I.: TraPPE-zeo: Transferable Potentials for Phase Equilibria Force Field for All-Silica Zeolites, *J. Phys. Chem. C* 2013 **117**, 24375-24387 DOI: 10.1021/jp4074224
- [20] Shah, M.S.; Tsapatsis, M.; Siepmann, J.I.: Development of the Transferable Potentials for Phase Equilibria Model for Hydrogen Sulfide, *J. Phys. Chem. B* 2015 **119**, 7041-7052 DOI: 10.1021/acs.jpcc.5b02536
- [21] Kaminski, G.; Duffy, E.; Matsui, T.; Jorgensen, W.: Free Energies of Hydration and Pure Liquid Properties of Hydrocarbons from the OPLS All-Atom Model, *J. Phys. Chem.* 1994 **98**, 13077-13081 DOI: 10.1021/j100100a043
- [22] Darkrim, F.; Levesque, D.: Monte Carlo simulations of hydrogen adsorption in single-walled carbon nanotubes, *J. Chem. Phys.* 1998 **109**, 4981-4984 DOI: 10.1063/1.477109
- [23] Waldman, M.; Hagler, A.: New combining rules for rare gas van der Waals parameters, *J. Comput. Chem.* 1993 **14**, 1077-1084 DOI: 10.1002/jcc.540140909
- [24] Song, W.; Rossky, P.J.; Maroncelli, M.: Modelling alkane+perfluoroalkane interactions using all-atom potentials: Failure of the usual combining rules, *J. Chem. Phys.* 2003 **119**, 9145-9162 DOI: 10.1063/1.1610435
- [25] Gubbins, K.E.; Quirke, N. (Eds.): *Molecular Simulation and Industrial Applications: Methods, Examples and Prospects* (Gordon & Breach, Amsterdam) 1997 ISBN: 9056990055, 9789056990053
- [26] Wolf, D.; Keblinski, P.; Phillpot, S.R.; Eggebrecht, J.: Exact method for the simulation of Coulombic systems by spherically truncated, pairwise r⁻¹ summation, *J. Chem. Phys.* 1999 **110**, 8254-8282 DOI: 10.1063/1.478738
- [27] Demontis, P.; Spanu, S.; Suffritti, G.B.: Application of the Wolf method for the evaluation of Coulombic interactions to complex condensed matter systems: Aluminosilicates and water, *J. Chem. Phys.* 2001 **114**, 7980-7988 DOI: 10.1063/1.1364638

- [28] Nicholson, D.; Parsonage, N.G.: Computer Simulation and the Statistical Mechanics of Adsorption (Academic Press, London) 1982 ISBN: 0125180608, 9780125180603
- [29] Karavias, F.; Myers, A.L.: Isothermic heats of multi-component adsorption: thermodynamics and computer simulations, *Langmuir* 1991 **7**, 3118-3126 DOI: 10.1021/la00060a035
- [30] Myers, A.; Prausnitz, J.M.: Thermodynamics of mixed-gas adsorption, *AIChE J.* 1965 **11**, 121-127 DOI: 10.1002/aic.690110125
- [31] Simon, C.M.; Smit, B.; Haranczyk, M.: pyIAST: Ideal adsorbed solution theory (IAST) Python package, *Comp. Phys. Commun.* 2016 **200**, 364-380 DOI: 10.1016/j.cpc.2015.11.016
- [32] Cruz, A.J.; Pires, J.; Carvalho, A.P.; Carvalho, M.B.: Physical Adsorption of H₂S Related to the Conservation of Works of Art: The Role of the Pore Structure at Low Relative Pressure, *Adsorption* 2005 **11**, 569-576 DOI: 10.1007/s10450-005-5614-3
- [33] Mohr, R.J.; Vorkapic, D.; Rao, M.B.; Sircar, S.: Pure and Binary Gas Adsorption Equilibria and Kinetics of Methane and Nitrogen on 4A Zeolite by Isotope Exchange Technique, *Adsorption* 1999 **5**, 145-158 DOI: 10.1023/A:1008917308002
- [34] Tant, Z.; Gubbins, K.E.: Selective Adsorption of Simple Mixtures in Slit Pores: A Model of Methane-Ethane Mixtures in Carbon, *J. Phys. Chem.* 1992 **96**, 845-854 DOI: 10.1021/j100181a059
- [35] Battisti, A.; Taioli, S.; Garberoglio, G.: Zeolitic imidazolate frameworks for separation of binary mixtures of CO₂, CH₄, N₂ and H₂: A computer simulation investigation, *Microporous and Mesoporous Mat.* 2011 **143**, 46-53 DOI: 10.1016/j.micromeso.2011.01.029

DYNAMIC SIMULATOR-BASED APC DESIGN FOR A NAPHTHA REDISTILLATION COLUMN

LÁSZLÓ SZABÓ,^{1*} KLÁRA KUBOVICSNÉ STOCZ,¹ LAURA SZABÓ,¹ SÁNDOR NÉMETH,²
AND FERENC SZEIFERT²

¹Department of Technology and Development, MOL Plc., Olajmunkás u 2., Százhalombatta, H-2443, HUNGARY

²Department of Process Engineering, University of Pannonia, Egyetem u 10., Veszprém, H-8200, HUNGARY

In this simulation study the operation of a naphtha redistillation column (a column with two feeds and three products) was analyzed with the application of Aspen HYSYS[®] software. The simulator, structure of local controllers and the product-quality estimators were based on the data of an industrial column in the Danube Refinery. The aim of the analysis was to identify the dynamic and steady-state effects of heating and cooling as well as the sidestream of product qualities. The relationship between the tray temperatures and the quality of the products was also identified and inferential calculations were made. Based on the identified relationships, a two-level hierarchical control structure was developed. On the lowest level of the hierarchy are the local controllers of flowrates, liquid levels, pressure and duty. The inferential calculations are important components of the controller which serve as the controlled variables at the coordination level. The inputs of the estimators are the process data of the column, e.g. temperature, pressure and flowrate. On the top level of the control hierarchical structure the quality of the products are controlled by manipulating the setpoint of the local controllers. Based on the analysis of the Controlled Variable – Manipulated Variable relationship, closed-loop quality control was achieved with PID controllers. The result of the analysis may form the basis of Advanced Process Controller implementation.

Keywords: distillation, multilevel control, MIMO, inferential, naphtha distillation, plant data

1. Introduction

Distillation is the most widely used separation technique in the chemical and petrochemical industries. Around 95 % of the separation tasks of components are performed using distillation in the chemical industry, and the distillation units use approximately 3% of the total energy consumption of the world [1]. Therefore, the improvement of distillation controllers and processes is important because of their widespread use and huge levels of energy consumption. Distillation is a complex task because all of the equipment are Multiple Input, Multiple Output (MIMO) objects from a control point of view with strong interactions [2]. The performance of distillation controllers directly affects the productivity rates, utility usage and product quality. During the operation, attaining a sufficient level of product quality is the final goal. Hence, the precise calculation of the qualities is very important from an economic point of view.

In this work, the multilevel control structure of the naphtha redistillation column has been analyzed (Fig. 1). Furthermore, a simulator with the same local control

structure as the existing distillation column in the Danube Refinery was created. The model was adapted according to the data of the sensors and laboratory measurements. The controlled quality values were provided by inferential calculations based on archived data of the unit. During the analyses a control structure was created which could be suitable for the quality control of the column, and also the resulting models can be used in the implementation of Advanced Process Controllers (APC).

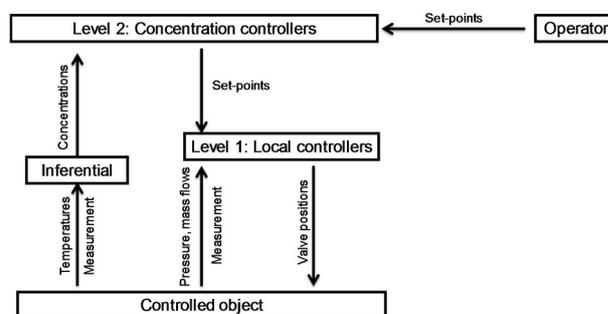


Figure 1. Structure of the control system

*Correspondence: lszabosz@mol.hu

Table 1. Disturbances

Value type	Noise	Delay	Lag
Mass flowrate	1.0 %	15 s	0 s
Pressure	0.2 %	0 s	0 s
Temperature	0.3 %	30 s	60 s
Level	1.0 %	60 s	0 s

2. Experimental

For the analyses of dynamic responses a piece of commercial simulation software was used. Data evaluation and parameter identification of the inferential calculations were performed with R programming language.

2.1. Separation task

The column has two feeds, the light naphtha feedstock possesses an initial boiling point of 12 °C and a final boiling point of 150 °C, while the boiling range of the heavy naphtha feedstock stretches from 57 to 197 °C. The task of the column is to separate the two feeds into three different naphtha products. The specification of the product streams are the following: distillate final boiling point should be between 58 and 65 °C, the final boiling point of the side product should be between 87 and 147 °C, and the initial boiling point of the bottom product should be between 99 and 103 °C. The simulator was implemented using Aspen HYSYS® software. The structure of the simulator is shown in Fig.2.

2.2. Simulator

The implementation of the column was achieved over two main steps. In the first step, a steady-state simulator of the column was developed. The structure of the simulator follows the process flow diagram in Fig.2. The column is divided into two parts, the light feedstock is loaded into the upper section (Column 107), and the heavy feedstock is fed into the lower section (Column 102). The towers, air coolers and heat exchangers were simulated with rigorous blocks and the model was adapted to data of the sensors and laboratory measurements. In the second step, the dynamic simulator was created based on the steady-state model. During the implementation of the dynamic simulator the calculation blocks were defined from the sizes of the actual equipment, e.g. the parameters of the trays in the tower. In some cases, pseudo pieces of equipment had to be defined to achieve a more realistic simulator, e.g. the bottom of the columns were modelled with a vertical drum to create a more detailed model.

2.3. Measurement Disturbances

Before determining the parameters of the controller, noise and time delays were applied to the calculated values. The disturbances were determined based on our

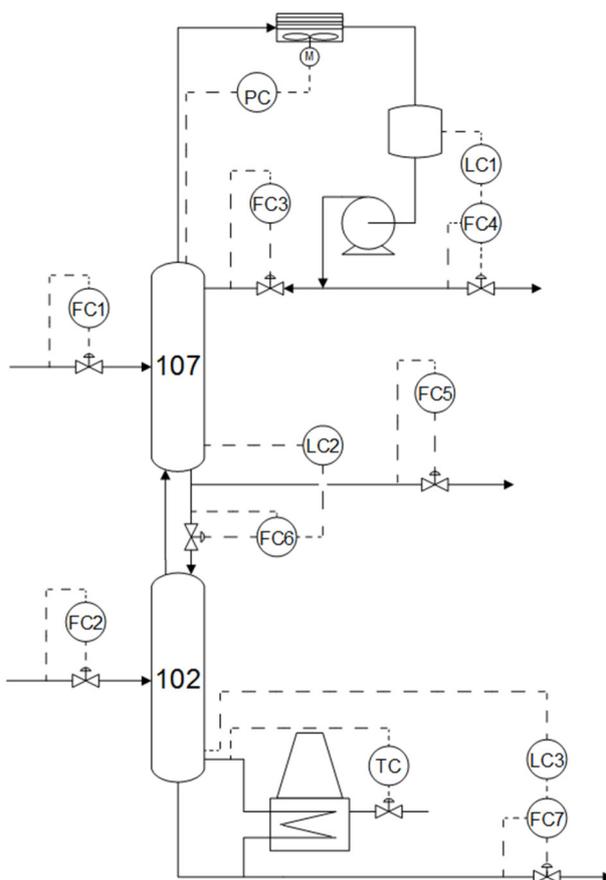


Figure 2. Structure of the simulator

previous experience. Transfer function blocks were used to simulate measuring instruments. Their parameters can be seen in Table 1. These signals were used later as controlled values.

2.4. Control Structure of Column

Fig.1 represents the overall structure of controllers. On the lower level of the hierarchy the local controllers can be found. On this level, the controllers maintain the operating conditions (pressure and liquid level controllers) of the column and eliminate the disturbances (mass flowrate controllers) of the environment. The controller output variables are the valve positions and the process variables are the mass flowrates, liquid levels and pressure.

The structure of the local controllers can be seen in Fig.2, which is the same as the structure of the controller in the real column. The FC1 and FC2 controllers eliminate the fluctuation of the feed mass flowrate. The PC and the TC controllers compensate for weather changes and the fluctuation of fuel gas. The liquid level controllers (LC1, LC2 and LC3) were necessary to ensure normal operation because the refluxes and the boil-up flowrate would be zero if there is no liquid phase in the reflux drum and at the bottom of the columns [3]. PI controllers were applied at the local level.

The degree of separation is based on the sustenance of the temperature difference between

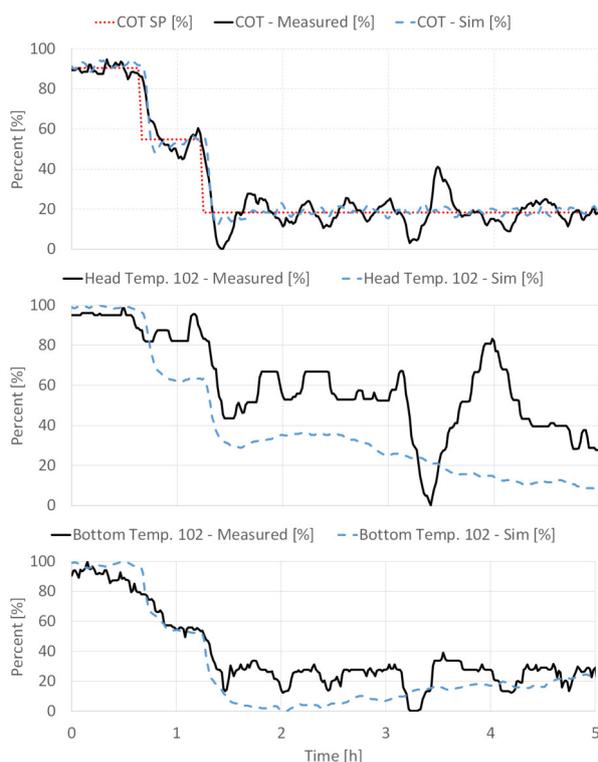


Figure 3. Validation of the Simulator

products. The main target of the operation is to achieve specified product qualities. The provision of online quality values for the control is an important task in the industry. One of the feasible solutions is to apply inferential calculations. Generally the inputs of these calculations are temperatures, because the quality of products is strongly correlated with the tray temperatures of the columns. Consequently, the temperature and quality controls serve the same purpose. Therefore, in this paper the quality of the products was controlled at the top level of the control hierarchy. The controller outputs on this level were the setpoints of the local controllers (FC3, FC5 and TC) and the controlled values were the outputs of the inferential calculations.

3. Results and Analysis

3.1. Validation of the Simulator

In the first analysis, the dynamic behavior of the simulator was validated. (The steady-state accuracy was validated during the construction of the model.) The model was compared to the historical data of the real column.

As Fig.3 shows, the coil outlet temperature (COT) of the furnace was changed over two steps. During this process the top temperature of Column 107 was controlled. The head and bottom temperatures of Column 102 were analyzed, because the effect of the COT was significant on these parameters. The operating point of the real column was different from the

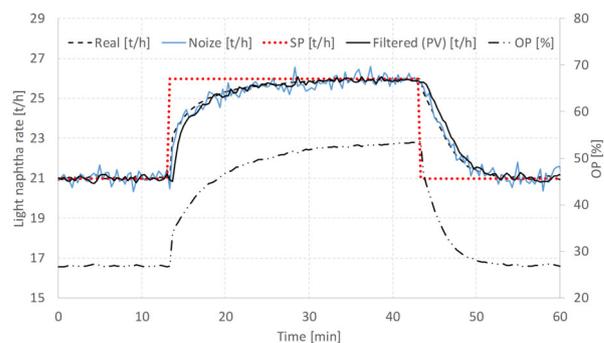


Figure 4. FC1 controller

simulated one, however, in this analysis only its dynamic behavior was evaluated. Therefore, in the simulator the setpoint of the COT was changed in the same direction and by the same amount as the measured data. The signals were normalized in order to compare the trends easily. It can be seen that the curves possess the same correlation in both the simulated and real columns. The dynamics of the responses were also identical in both cases, thus the simulator was found reliable for further analyses.

3.2. Analysis of Local Controllers

For the tuning of the local controllers previous experiences were drawn on [4]. It was assumed that the controllers exhibit no significant interactions so they can be tuned independently. Before determining the parameters of the controllers, filters were applied, hence, the dynamics of the loops were approximated to the real control loops.

The control loops were tuned in the following order: flowrate (FC1-7), the liquid level (LC1-3) and pressure (PC) controllers. In the last step the parameters of the coil outlet temperature controller (TC) were determined. For the loops, Honeywell equation B [5-6] was used, and the proportional and integrating parameters were determined using the built-in autotuner tool. After automatic tuning, the loops were tested and fine tuning was carried out to avoid fluctuation of the loops. The behavior of the FC1 is presented in Fig.4. As the figure shows, the controllers follow the setpoint without exceeding it, which ensures a robust local controller system. The raw calculated signal, the signal loaded with noise and the filtered (controlled) signals are also presented in Fig.4. As the results show, the control of the raw values is stable and is the first signal to reach a setpoint.

3.3. Inferential Calculations

The quality estimators were created for the final boiling points of the distillate and the side product and for the initial boiling point of the bottom product. For the inferential fitting the laboratory and process data was used; the data was collected over a period of one year. In the case of the process data, all relevant measurement points of the column were considered (flowmeters,

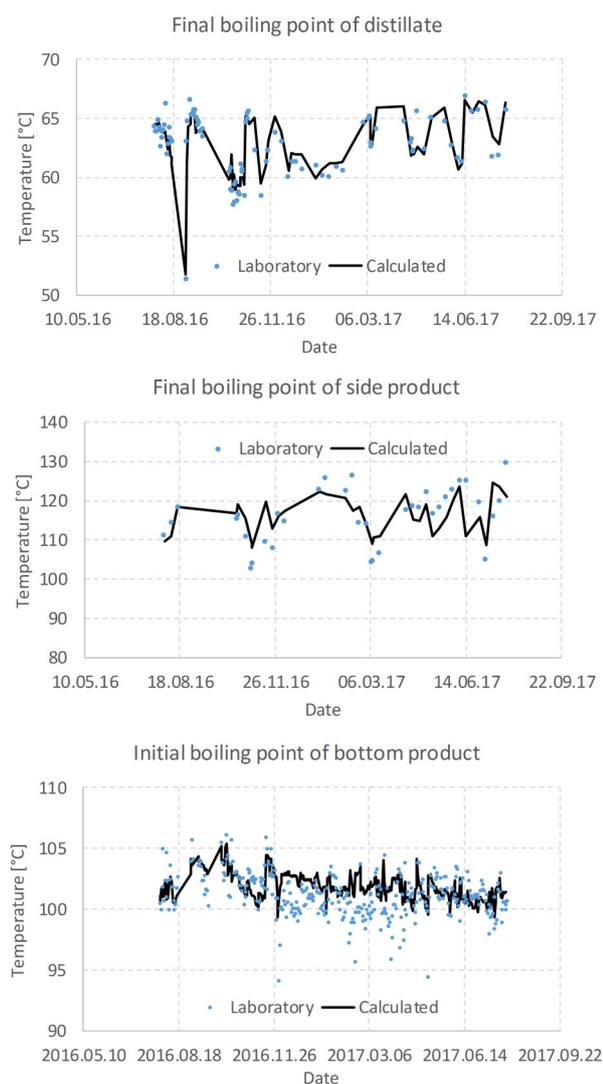


Figure 5. Performance of inferential calculations

temperature and pressure measurements). One hour-long periods were aggregated from the online data before each laboratory sample time. From these values pressure compensated temperatures (PCT) and ratios of the mass flowrates were also calculated.

The inputs of the equations (see Table 2) were determined using correlation analysis, thus the qualities were calculated from the most relevant values. The parameters of the equations were calculated using linear regression. The results of the fittings are presented in Fig. 5.

Table 2. Input parameters of inferential calculations

Inferential	Input 1	Input 2
Final boiling point of distillate	PCT of 107 head temperature	
Final boiling point of side product	PCT of 107 bottom temperature	<i>Light naphtha</i> <i>Light feed</i>
Initial boiling point of bottom product	PCT of 102 bottom temperature	

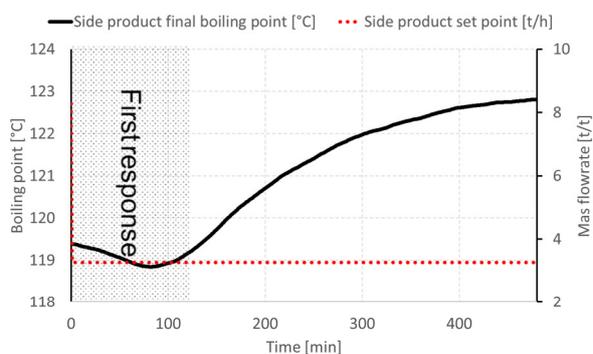


Figure 6. Open loop test CV2 – MV2

3.4. Structure of Quality Controllers

According to the separation task the controlled variables (CV) are: distillate final boiling point (CV1), side-product final boiling point (CV2) and bottom product initial boiling point (CV3). The manipulated variables are the following: setpoints of FC3 (MV1), FC5 (MV2) and TC (MV3). The structure of the quality controllers was determined with the help of the relative gain array (RGA) [7]. To create the array, a step test was performed for each manipulated variable (MV) in the simulator. The RGA is presented in Table 3. The elements of the matrix evaluate a possible loop with the given MV – CV pair. If the element of the array is negative, closing other loops will reverse pairing gain. This configuration will become unstable if the other loops operate. If the element of the RGA exceeds one, interaction with other loops has an inhibiting effect. Closing other loops decreases pairing gain. If the value is between 0 and 1 then pairing gain will increase by closing the other loops; the interaction is the greatest at 0.5. In the case of this pairing the loop should be tuned when the other loops are automated. If the RGA value is 1 then the other loops have no effect on the pairing gain. In the case of the CV1 the evident manipulated variable is MV1. The other two can be paired in two ways. According to the RGA, the obvious pairing should be CV2 – MV2 / CV3 – MV3. However, there is a primary effect between MV2 and CV2, and it possesses a negative gain. This behavior could make the control loop instable, or the settling time of the controlled object would be large. The dynamic responses are presented in Fig. 6. In this analysis the MVs were stepped during the first minute. Only the local controllers were in closed loops during the analyzed period.

Table 3. Relative Gain Array

	MV1	MV2	MV3
CV1	7.7	-1.2	-5.5
CV2	-6.9	1.9	6.0
CV3	0.2	0.3	0.5

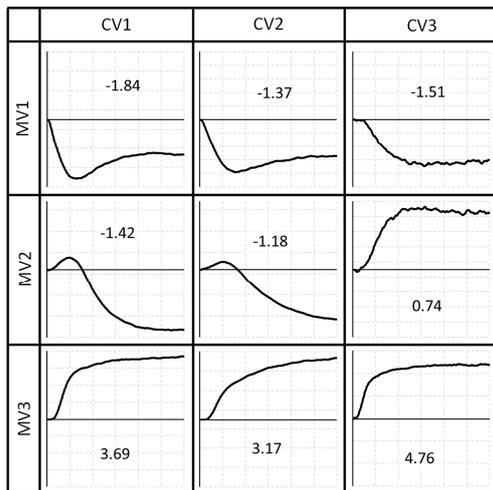


Figure 7. APC models

The following final control pairings were applied: CV1 – MV1, CV2 – MV3 and CV3 – MV2. According to this analysis quality control can be achieved using simple PID control loops. Therefore, the analysis of the closed-loop quality control in the following chapter was performed using only PID loops. This analysis was also the basis of an APC implementation. The initial APC models can be seen in Fig.7.

3.5. Operation of Quality Controllers

In this analysis the operation of the quality controllers was analyzed in the case of a disturbance. The mass flowrate of the light naphtha feed was changed at 5 o'clock. During the analyzed period all of the local and quality controllers were in closed loops. According to the results in Fig.8, the controller system can compensate for this change. In the figure the values of the MVs were normalized between their minimum and maximum limits.

In the case of CV1 this quality did not violate the specification limit, and the controlled variable was close to the setpoint after 2.5 hours. This parameter possessed the smallest peak of the three CVs at approximately 1.3 °C. CV2 exhibited a peak at approximately 2 °C, nevertheless, this was well within the specification limit so it did not violate it. The settling time was about 7.5 hours. CV3 exhibited a peak at approximately 3 °C which violated the specification limit. After 2.5 hours the CV once again exceeded the limit. The settling time was ~15 hours. As can be seen the outputs (OP) were moved after the settling of the CVs, but because of the long settling time the whole time period was not presented.

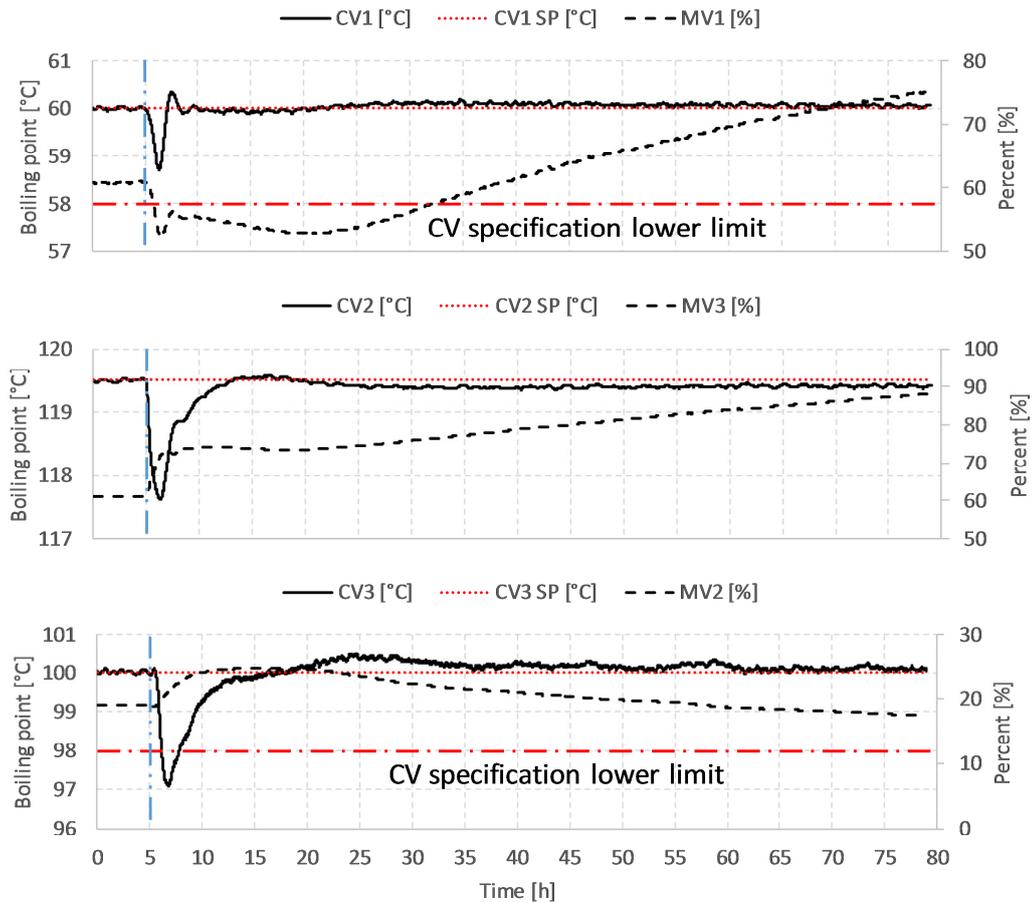


Figure 8. Operation of Quality Controllers

4. Conclusion

For the analysis of the naphtha redistillation column at MOL Plc. the dynamic simulator of the unit was adapted to the data of the real equipment. The steady-state accuracy and the dynamic behavior of the model were evaluated. The simulator is reliable for the analysis of the column.

For the operation of the distillation column a two-level control structure was implemented in the simulator. On the lower level of the control hierarchy controllers were created which facilitated normal operating conditions of the column (pressure and liquid-level controllers) and eliminated the disturbances of the environment (mass flowrate controllers). To ensure quality signals, inferential calculations were created based on the historical data of the real column. The inputs of the calculations were mainly the temperature measurements of the columns. Hence, the controlled temperatures were chosen during the correlation analysis of the inferential calculations; on the top level of the control hierarchy the controlled variables were the outputs of the quality estimators.

The manipulated variables were the setpoints of the local controllers. The pairing of the manipulated and controlled variables should be defined carefully. The first aspect of the pairing was the gain between the variables, however, the dynamic responses were also considered. To eliminate the interaction and to improve the performance of the quality controller, decoupling [8] or a model predictive controller [9] was recommended.

The simulator-based column analysis can be a good foundation for an Advanced Process Controller. The CV-MV pairing and resulting single input-single output models can be used as initial models in Advanced Process Controllers.

Furthermore, the simulator can be used for the analysis of column operation issues in the plant. The resulting inferential calculations have already been implemented and are available for the operators to facilitate quality control.

Acknowledgements

We acknowledge the financial support of Széchenyi 2020 under the project EFOP-3.6.1-16-2016-00015.

REFERENCES

- [1] Mansour, K.; Mansour, E.: Rigorous optimization of heat-integrated and Petlyuk column distillation configurations based on feed conditions, *Clean Techn. Environ. Policy*, 2009 **11**(1), 107–113 DOI: 10.1007/s10098-008-0171-6
- [2] Skogestad, S.; Morari, M.: The dominate time constant for distillation columns, *Computers and Chemical Engineering*, 1987 **11**(6), 607–617 DOI: 10.1016/0098-1354(87)87006-0
- [3] Buckley, P.S.: *Techniques of Process Control*, (Wiley, London, United Kingdom) 1964 ISBN-10:0882757776
- [4] Szabó, L.; Németh, S.; Szeifert, F.: Three-Level Control of a Distillation Column, *Engineering*, 2012 **4**(10), 675–681 DOI: 10.4236/eng.2012.410086
- [5] Honeywell: Application Module Algorithm Engineering Data, 1-800 343-0228
- [6] AspenTech: Aspen HYSYS Unit Operations Reference Guide, Version Number: V10
- [7] Skogestad, S.: Dynamics and Control of Distillation Columns: A Tutorial Introduction, *Chemical Engineering Research and Design*, 1997 **75**(6), 539–562 DOI: 10.1205/026387697524092
- [8] Benz, S.J.; Scennal, N.J.: An Extensive Analysis on the Start-up of a Simple Distillation Column with Multiple Steady States, *The Canadian Journal of Chemical Engineering*, 2002 **80**(5), 865–881 DOI: 10.1002/cjce.5450800510
- [9] Luyben, W.L.: *Process Modeling, Simulation, and Control for Chemical Engineers* (McGraw-Hill, New York, USA) 1990 ISBN-10:0070391599

REPLACEMENT OF BIASED ESTIMATORS WITH UNBIASED ONES IN THE CASE OF STUDENT'S t -DISTRIBUTION AND GEARY'S KURTOSIS

GERGELY TÓTH* AND PÁL SZEPESVÁRY

Institute of Chemistry, Faculty of Science, Eötvös Loránd University, 1/A Pázmány Péter sétány, Budapest, H-1117, HUNGARY

The use of biased estimators can be found in some historically and up to now important tools in statistical data analysis. In this paper their replacement with unbiased estimators at least in the case of the estimator of the population standard deviation for normal distributions is proposed. By removing the incoherence from the Student's t -distribution caused by the biased estimator, a corrected t -distribution may be defined. Although the quantitative results in most data analysis applications are identical for both the original and corrected t -distributions, the use of this last t -distribution is suggested because of its theoretical consistency. Moreover, the frequent qualitative discussion of the t -distribution has come under much criticism, because it concerns artefacts of the biased estimator. In the case of Geary's kurtosis the same correction results $(2/\pi)^{1/2}$ unbiased estimation of kurtosis for normally distributed data that is independent of the size of the sample. It is believed that by removing the sample-size-dependent biased feature, the applicability domain can be expanded to include small sample sizes for some normality tests.

Keywords: unbiased estimator, normal distribution, Anscombe-Glynn test, Jarque-Bera test, Bonett-Seier test

1. Introduction

The Student's t -distribution [1] is one of the most widely used statistical functions in experimental practice. It is well documented that experts active in analytical, physical and clinical chemistry, biology, agriculture, ecology, economy as well as forensic science or even legal representatives apply this tool to formulate solid statements, e.g. population means, to differentiate between two sample means, etc., which are in most cases necessarily based on a limited number of observations. Fortunately, scientists are in this respect soundly supported by a lot of textbooks, standards, software, software systems, and last but not least trained in elements of statistics. However, it should be a moral obligation to be aware of the evolution of this routinely used method, its principles and often tacitly supposed assumptions. Although not crucial in daily use, it is worthwhile to know that besides Student's t -distribution there are different functions which may be suitable for the determination of percentiles in the same way as the t -distribution. They may differ mainly in terms of alternate estimators for the population mean and population standard deviation. An attractive variant is presented in this work emphasizing its theoretically fully consistent feature on the contrary to the Student's t -distribution.

The Student's t -distribution corresponds to a ratio of normally distributed random variables to chi-distributed random variables (see the references in the historical review of Zabell [2]). Chi-distributed variables postulate normally distributed data as well. Gosset [1] used in his definition an estimate of the standard deviation which is biased relating the population standard deviation and even the variance in the case of normally distributed random variables as was shown earlier by Helmert [3-4]. Amazingly, when Fischer proposed a transformation of Gosset's original z variable to $t = z\sqrt{N-1}$ [5-6], he chose an estimate for the standard deviation which is also biased relating σ . A corrected t_c -distribution is proposed that fulfils all theoretical requirements and yields a more normal distribution-like shape. It is consequently based on normal sample data and uses an unbiased estimator.

A similar correction can be applied to Geary's kurtosis that is the ratio of the mean deviation to the standard deviation [7]. In this work the use of the correction is proposed in order to eliminate the sample-size dependency of the mean Geary's kurtosis on normally distributed data.

Finally, some remarks are made on statistical tests based on sample-size dependent values in order to extend their applicability to small sample sizes.

*Correspondence: toth@chem.elte.hu

2. Theoretical Background

The square root of the mean of the square of the centred observations of the random variable, Y ,

$$s_N = \sqrt{\frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N}} \quad (1)$$

may be an estimator of the population standard deviation, σ , of a sample. N is the sample size and \bar{y} denotes the sample mean. Bessel pointed long before to the bias of s_N and proposed:

$$s = \sqrt{\frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N - 1}} \quad (2)$$

Although s^2 is already an unbiased estimator of the variance, σ^2 , irrespective of the distribution of Y , the statistics s_N and s are both biased in terms of the standard deviation, σ . However, there is a correction [5-6, 8-9]:

$$c_4(N) = \sqrt{\frac{2}{N-1}} \frac{\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{N-1}{2}\right)}, \quad (3)$$

which applies to s by transforming it, in the case of normally distributed variables, into an unbiased statistic:

$$s_c = \frac{s}{c_4(N)}. \quad (4)$$

$c_4(N)$ follows from Helmert's papers or Cochran's theorem [10]. For normally distributed Y , $\sqrt{N-1}(s/\sigma)$ exhibits a chi distribution with $N-1$ degrees of freedom. $c_4(N)$ is the expected value of s/σ . There is a $c_2(N)$ value as well, that is equal to the expected value of s_N/σ . The correcting effect of $c_4(N)$ may be considerable for low values of N (Table 1).

Table 1. Selected $c_4(N)$ corrections

N	$c_4(N)$
2	0.7979
3	0.8862
4	0.9213
5	0.9400
10	0.9727
20	0.9869
30	0.9914

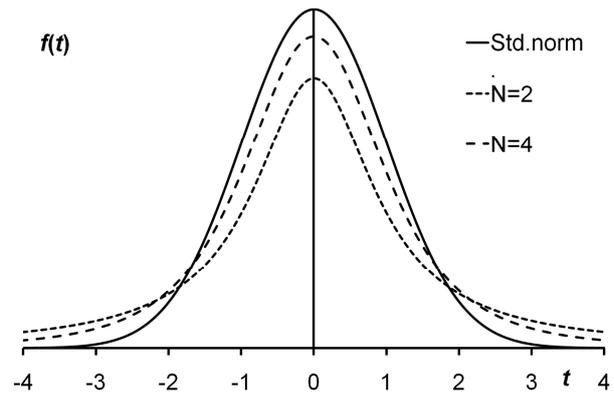


Figure 1. Student's t -distribution compared with the standard normal distribution

The correction term is frequently used in Statistical Process Control (SPC) to define $\pm 3\sigma$ intervals and process standard deviations that are determined for samples of different sizes.

3. Results and Discussion

3.1. Discussion of the density functions of the Student's t -distribution

Let Y be an independent $N(\mu, \sigma^2)$ variable. The

$$t = \frac{\bar{y} - \mu}{s/\sqrt{N}} \quad (5)$$

statistic exhibits the usual Student t -distribution with $N-1$ degrees of freedom. The density function (Eq.(6)) can be derived as the quotient of normal and chi-distributed random variables. The independency of the nominator and denominator can be shown and it is also proved by a series of theorems that t defined in Eq.(5) follows a Student's t -distribution with $N-1$ degrees of freedom (see the references in [2]):

$$f(t) = \frac{1}{\sqrt{\pi(N-1)}} \frac{\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{N-1}{2}\right)} \left(1 + \frac{t^2}{N-1}\right)^{-\frac{N}{2}} \quad (6)$$

The distribution sketched in Fig.1 is widely known and needs no comments except for its flaw: it is based on the biased statistic (Eq.(2)) for σ .

By replacing the standard deviation of the sample, s , in Eq.(5) with the unbiased equivalent one, s_c , corrected by $c_4(N)$, results in a new value:

$$t_c = \frac{\bar{y} - \mu}{s_c / \sqrt{N}}, \quad (7)$$

and in the density function:

Table 2. Sample-size dependence of Geary's and Pearson's kurtosis. The $c_4(N)$ -corrected Geary's and the size-corrected Pearson's kurtosis [11] provided the ∞ limit values for all sample sizes within the statistical uncertainty of the 10^5 random samples.

N	Geary's kurtosis	Pearson's kurtosis
3	0.9004	1.5000
4	0.8659	1.8005
5	0.8489	1.9996
6	0.8385	2.1437
7	0.8319	2.2501
8	0.8267	2.3340
9	0.8233	2.4006
10	0.8200	2.4557
20	0.8085	2.7141
50	0.8021	2.8820
100	0.7999	2.9408
∞	0.7979	3.0000

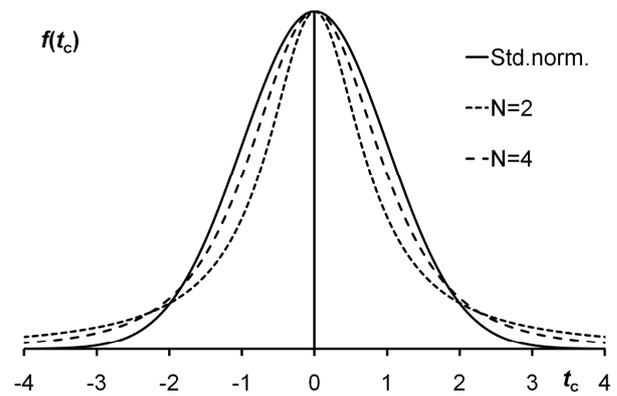


Figure 2. Corrected Student's *t*-distribution compared with the standard normal distribution

$$f(t_c) = \frac{1}{c_4(N)} \frac{\Gamma\left(\frac{N}{2}\right)}{\sqrt{(N-1)\pi} \Gamma\left(\frac{N-1}{2}\right)} \left(1 + \frac{t_c^2}{c_4^2(N-1)}\right)^{-\frac{N}{2}}$$

$$= \frac{1}{\sqrt{2\pi}} \left(1 + \frac{t_c^2}{c_4^2(N-1)}\right)^{-\frac{N}{2}} \quad (8)$$

which is shown in Fig.2. As expected, the corrected Student's *t*-distribution (Student's t_c -distribution) consists of more distinct peaks, it exhibits fatter tails than the standard normal distribution, but rather oddly at its maximum, i.e. when $t_c = 0$, the value $f(t_c)$ does not depend on N and is represented by the value $1/\sqrt{2\pi}$, the same as for those of normal distribution. The function $f(t_c)$ and the Student's *t*-distribution without a doubt differ in this respect.

The obvious differences between the Student and modified Student distributions do not complicate their daily usage. The confidence intervals calculated using s and t or s_c and t_c estimators and distributions, respectively,

$$\bar{y} \pm \frac{s_c t_c^{\alpha/2, (N-1)}}{\sqrt{N}} \quad \text{and} \quad \bar{y} \pm \frac{s t^{\alpha/2, (N-1)}}{\sqrt{N}} \quad (9)$$

do not differ, because

$$s_c t_c^{\text{crit}(\alpha/2, N-1)} = \frac{s}{c_4(N)} t^{\text{crit}(\alpha/2, N-1)} c_4(N)$$

$$= s t^{\text{crit}(\alpha/2, N-1)} \quad (10)$$

Evidently, corresponding estimators and functions should be used.

3.2. The effect of the correction to Geary's kurtosis

Geary's kurtosis [7], w_N , is the ratio of the mean absolute deviation (MAD) to the standard deviation (Eq.(11)). It is an alternative to Pearson's kurtosis based on the fourth moment. The expected value of Geary's kurtosis depends on the sample size even for normally distributed data [7]. The mean values of 10^5 random samples from standard normal distributions are shown in Table 2.

$$w_N = \frac{\frac{1}{N} \sum_{i=1}^N |y_i - \bar{y}|}{s_N} \quad (11)$$

If the $c_4(N)$ corrections (Table 1) are applied during the calculation of the nominator of the ratio as $w_{N, \text{corr}} = w_N c_4(N)$, the expected value of the kurtosis is $(2/\pi)^{1/2}$ for all sample sizes. This means that the platykurtic and the leptokurtic features of a sample can be found without searching for the size-dependent dividing value in tables.

3.3. Sample-size bias in statistical tests

Geary's kurtosis and its transformed values are used in normality tests due to their enhanced sensitivity to some leptokurtic deviations from normality [12]. Contrary to the case of the Student's *t*-distribution, where the correction has no effect on the *t*-test, here the effect of the correction is not cancelled. Generally, the size dependence decreases the performance of the tests for small sample sizes. This feature is interpreted by users as a recommendation that the tests are unsuitable for small sample sizes. In the same way, neglect of sample-size dependence is applicable in tests where Pearson's kurtosis is used. The calculated mean value of Pearson's kurtosis is shown in Table 2 but its convergence is rather weak to the theoretical value of 3. It should be noted here that the sample-size unbiased estimator of kurtosis can be easily calculated [11].

Table 3. Sample-size dependence of five normality tests based on unbiased or sample-size-dependent biased estimators. The numbers show the ratio of the rejected null hypotheses to all trials at a significance level of 0.05 from 10^5 random samples.

<i>N</i>	<i>Shapiro-Wilk</i>	<i>D'Agostino</i>	<i>Anscombe-Glynn</i>	<i>Bonett-Seier</i>	<i>Jarque-Bera</i>
4	0.0502				
5	0.0521			0.0373	
6	0.0477		0.0189	0.0087	
7	0.0492		0.0343	0.0282	
8	0.0505		0.0365	0.0348	
9	0.0507	0.0538	0.0369	0.0390	0.0024
10	0.0498	0.0528	0.0394	0.0417	0.0058
20	0.0497	0.0525	0.0406	0.0421	0.0089
50	0.0498	0.0500	0.0466	0.0470	0.0241
100	0.0499	0.0493	0.0533	0.0490	0.0368

In *Table 3* the type-I error of some normality tests calculated on 10^5 standard normal samples is shown. In the calculation the ‘moments’ package in R was used [13]. *Table 3* contains the ratio of the samples to all samples where the H_0 hypothesis of normality was rejected at the significance level of 0.05. The Shapiro-Wilk method [14] uses the ratio of two unbiased estimators of the variance, and is suitable for all data sizes. The skewness test of D’Agostino [15] slightly overestimates the number of rejected cases. It is based on the normalized third-order central moment definition of skewness, where the expected value is estimated without bias. The Anscombe-Glynn [16] test applies Pearson’s kurtosis without size correction using a biased estimation of normally distributed data. The Bonett-Seier [12] test shown here uses Geary’s kurtosis and the Jarque-Bera test [17] combines skewness and Pearson’s kurtosis. The performance of the last three tests was rather weak for small sample sizes, whereof one cause might be the lack of correction for small sample sizes even for normally distributed data. These tests are usually only recommended for medium and large sample sizes. The correction should extend the applicability domain to small sample sizes. Of course, the type-I error of normally distributed data is only one narrow aspect of a test, detailed analysis should be performed to investigate the effect of correction on many distributions, like, e.g. in ref. [12].

4. Conclusion

Nowadays, data are evaluated by computers and biased estimators can be replaced by unbiased ones, even if their calculation schemes are complex.

It has been shown that, in terms of Student’s t -distributions, to decide upon the confidences of statistics one has two functions which are completely equivalent as far as practical applicability is concerned. They can, however, be distinguished theoretically. The assertion that only the unbiased estimator should be recognized as the correct one implies the use of the corrected Student’s t -distribution, $f(t_c)$. In that case the known

shape of the Student’s t -distribution may be labelled as an artefact and the usual application of the Student’s t -distribution as a production of “correct numbers by an incoherent theory”.

In the case of Geary’s kurtosis, the correction removes the sample-size dependence from the expected value. This change of distinguishing platykurtic or leptokurtic features of the sample is simpler than using the original version of Geary’s kurtosis. Furthermore, subtracting $(2/\pi)^{1/2}$ results in a number to be interpreted in a similar way to the excess kurtosis obtained by subtracting 3 from the Pearson’s kurtosis.

As a further study, the use of unbiased/sample-size-dependent corrections to extend the applicability domain to small sample sizes in the case of normality tests is recommended. It is believed that the use of biased estimators was acceptable before the age of computers and a systematic change to unbiased ones might be necessary in terms of statistics and standards with regard to industrial processes.

Acknowledgement

The authors are sincerely grateful for the valuable comments of Prof. S. Kemény and other participants during the scientific discussion after the presentation concerning Student’s t -distribution at the SCAC 2010 conference in Budapest in September 2010.

REFERENCES

- [1] Student (Gosset, W.S.): The Probable Error of a Mean, *Biometrika* 1908 **6**(1), 1–25 DOI: 10.2307/2331554
- [2] Zabell, S.L.: On Student’s 1908 Article “The Probable Error of a Mean”, *J. Am. Stat. Assoc.* 2008 **103**(481), 1–7 DOI: 10.1198/016214508000000049
- [3] Helmert, F.R.: Über die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler und über einige damit in Zusammenhang stehende Fragen, *Z. Math. Phys.* 1876 **21**, 192–218
- [4] Helmert, F.R.: Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehles directer Beobachtungen gleicher Genauigkeit, *Astronomische Nachrichten* 1876 **88**(8-9), 113–131 DOI: 10.1002/asna.18760880802
- [5] Fischer, R.A.: Applications of “Student’s” distribution, *Metron* 1925 **5**, 90–104
- [6] Fischer, R.A.: Statistical methods for research workers, (Oliver & Boyd, Edinburgh and London) 1925.
- [7] Geary, R.C.: Moments of the ratio of the mean deviation to the standard deviation for normal samples, *Biometrika* 1936 **28**(3/4), 295–307 DOI: 10.2307/2333953
- [8] Grubbs, F.E.; Weaver, C.L.: The best unbiased estimate of population standard deviation based on group ranges, *J. Am. Stat. Assoc.* 1947 **42**(238), 224–241 DOI: 10.1080/01621459.1947.10501922

- [9] Vincze, I.: Matematikai statisztika ipari alkalmazásokkal, (Műszaki Kiadó, Budapest), 1975 pp. 58, 89, 165–168 ISBN 963-10-0472-4
- [10] Cochran, W.G.: The distribution of quadratic forms in a normal system, with applications to the analysis of covariance, *Math. Proc. Cambridge* 1934 **30**(2), 178–191 DOI: 10.1017/S0305004100016595
- [11] Joanes, D.N.; Gill, C.A.: Comparing measures of sample skewness and kurtosis, *J. Roy. Stat. Soc. D-Sta.* 1998 **47**(1), 183–189 DOI:10.1111/1467-9884.00122
- [12] Bonett, D.G.; Seier, E.: A test of normality with high uniform power, *Comput. Stat. Data An.* 2002 **40**(3), 435–445 DOI: 10.1016/S0167-9473(02)00074-9
- [13] Komsta, L.; Novomestky, F.: Moments R package, version 0.14, 2015 at <http://cran.r-project.org>, accessed in October 2017
- [14] Shapiro, S.S.; Wilk, M.B.: An Analysis of Variance Test for Normality (Complete Samples), *Biometrika* 1965 **52**(3/4), 591–611 DOI: 10.1093/biomet/52.3-4.591
- [15] D'Agostino, R.B.: Transformation to Normality of the Null Distribution of G_1 , *Biometrika* 1970 **57**(3), 679–681 DOI: 10.1093/biomet/57.3.679
- [16] Anscombe, F.J.; Glynn, W.J.: Distribution of the kurtosis statistic b_2 for normal samples, *Biometrika* 1983 **70**(1), 227–234 DOI: 10.1093/biomet/70.1.227
- [17] Jarque, C.M.; Bera, A.K.: Efficient tests for normality, homoscedasticity and serial independence of regression residuals, *Economic Letters* 1980 **6**(3), 255–259 DOI: 10.1016/0165-1765(80)90024-5

FORMATION, PHOTOPHYSICS, PHOTOCHEMISTRY AND QUANTUM CHEMISTRY OF THE OUT-OF-PLANE METALLOPORPHYRINS

ZSOLT VALICSEK,^{1*} MELITTA P. KISS,¹ MELINDA A. FODOR,¹ MUHAMMAD IMRAN,²
AND OTTÓ HORVÁTH¹

¹Department of General and Inorganic Chemistry, Institute of Chemistry, Faculty of Engineering, University of Pannonia, Egyetem u. 10., H-8200 Veszprém, HUNGARY

²Department of Chemistry, Baghdad-ul-Jadeed Campus, The Islamia University of Bahawalpur, 63100 Bahawalpur, PAKISTAN

Among the complexes of porphyrins, special attention has been paid to those possessing out-of-plane (OOP) structures, for the formation of which the size, as well as the coordinative character of the metal center are responsible. In these coordination compounds, the central atom cannot fit coplanarly into the cavity of the ligand, hence, it is located above the porphyrin plane, distorting it. Equilibria and kinetics of the complex formation, spectrophotometric, photophysical and primary photochemical properties of post-transition and lanthanide OOP metalloporphyrins were investigated, in addition electronic structural calculations were performed; hence, the general OOP characteristics were determined. Meanwhile, few doubtful questions have attempted to be answered concerning the categorization of metalloporphyrins, the borderline case complexes and hyperporphyrins.

Keywords: out-of-plane metalloporphyrins, formation kinetics, UV-Vis spectrophotometry, photochemistry, borderline case complexes

1. Introduction

Porphyrins and their derivatives play important roles in several biochemical systems. Four pyrroles are connected through methylened bridges, forming the porphyrin ring. Its planar structure with an extended conjugated π -electron system provides aromatic characteristics and a special coordination cavity for the binding of metal ions of suitable radius [1-2].

Metalloporphyrins are the central parts of naturally important compounds, e.g., magnesium(II) chlorins in bacteriochlorophylls and chlorophylls; iron(II) protoporphyrin in hemoglobin; and iron(III) protoporphyrins in myoglobin, cytochromes, oxidase, peroxidase, catalase, and oxoanion reductase enzymes. Ringed tetrapyrroles provide strong chelating effect which can promote the hyperaccumulation of rare metal ions in living cells, and also in abiotic environments, e.g. in kerogens.

In porphyrins the conjugation favors a planar structure. However, peripheral substituents or the metal center (originating from its size or axial ligand) can cause geometrical distortion. This certainly affects the functions of enzymes, as well as the biosynthesis of metalloporphyrins. In chemical research, due to distortion, redox potentials, basicity, reactivity, catalytic

activity and coordinative abilities of porphyrins can be modified. Also, due to the distortion, the degree of symmetry decreases, resulting in characteristic spectral changes in various ranges of the electromagnetic spectrum. The most frequent types of distortions are dome, saddle, ruffled and wave (chair-like) [3].

Overcrowded substitution on the periphery [3-4] or insufficiently short metal-nitrogen bonds due to the shrinkage of the coordination cavity can cause the ruffled or saddled deformation [2, 5-7]. If, however, the M-N bonds are significantly longer than half the length of the diagonal N-N distance in the free-base porphyrin, dome deformation can occur. This happens if the radius of the metal center exceeds the critical value of about 75-90 pm (depending on the type of porphyrin ligand) or square planar coordination is not preferred. Such metal ions are too big to fit into the ligand cavity. Hence they are located above the plane of the pyrrolic nitrogens; forming sitting-atop (SAT) or out-of-plane (OOP - see Fig.1) complexes, displaying thermodynamic instability, kinetic lability, typical photophysical features and photochemical reactivity [8-9].

In this work, we review our recent results regarding the formation, structure and photoinduced behavior of water-soluble OOP metalloporphyrins. These complexes with a diverse range of metal ions can be more simply produced in aqueous systems than in organic solvents. In this regard one of the most suitable free-base ligands is the anionic 5,10,15,20-tetrakis(4-sulfonatophenyl)porphyrin (H_2TSPP^{4-} - see Fig.1) due

*Correspondence: valicsek@almos.uni-pannon.hu

to its negative charge promoting the coordination of positively charged metal ions. Besides, this ligand is the most frequently used reagent among the free-base porphyrins [1].

2. Experimental

Analytical grade tetrasodium 5,10,15,20-tetrakis(4-sulfonatophenyl)porphyrin ($C_{44}H_{26}N_4O_{12}S_4Na_4 \cdot 12H_2O = Na_4H_2TSPP \cdot 12H_2O$) (Sigma–Aldrich) and simple metal salts such as nitrate, sulfate, chloride or perchlorate were used for the experiments. The solvent was double-distilled water purified with a Millipore Milli-Q system. The pH of the majority of the metalloporphyrin solutions was adjusted to 8 by application of a borate buffer, whilst maintaining the ionic strength at a constant value of 0.01 M. In a few cases, the pH was regulated to 6, and the ionic strength to 1 M, by an acetate buffer, to hinder hydrolysis.

The absorption spectra were recorded and the spectrophotometric titrations were monitored by using a Specord S-100 or a Specord S-600 diode array spectrophotometer. For the measurement of fluorescence spectra, a Perkin Elmer LS-50B or a Horiba Jobin Yvon FluoroMax-4 spectrofluorometer was applied. The latter piece of equipment supplemented with a time-correlated single photon counting (TCSPC) accessory was utilized to determine fluorescence lifetimes, too. UV-Vis spectrophotometric data (molar absorption, fluorescence quantum yields and lifetimes) of the free-base porphyrin were used as references for the determination of those of metalloporphyrin complexes [1].

For the determination of photochemical properties via continuous irradiations, a piece of AMKO LTI photolysis equipment (containing a 200W Xe–Hg lamp and a monochromator) was applied.

For the electronic structural calculations, the B3LYP Density Functional Theory (DFT) method and the LANL2DZ basis set were used. On the basis of our earlier quantum chemistry experiences, the sulfonatophenyl substituents exhibit negligible effects on the coordination of the metal center in the cavity; thus, the anionic porphyrin (H_2TSPP^{4-}) can be modeled on the unsubstituted porphyrin (H_2P) [4, 10].

3. Results and discussion

3.1. UV-Vis spectrophotometry

Porphyrins and their derivatives belong to the strongest light-absorbing materials (both natural and artificial), therefore, ultraviolet-visible spectrophotometry is one of the most fundamental, in addition, most informative spectroscopic techniques in porphyrin chemistry. They possess two $\pi\pi^*$ electronic transitions in the visible range of the electromagnetic spectrum: B- or Soret band at about 350–500 nm, usually with a molar absorbance of $10^5 \text{ M}^{-1}\text{cm}^{-1}$ (Fig.2), and Q bands at 500–750 nm

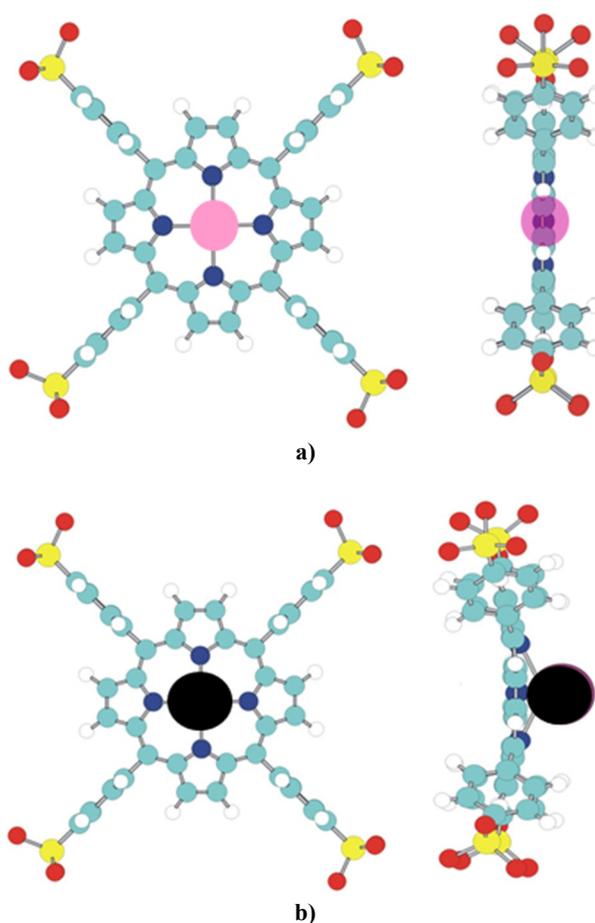
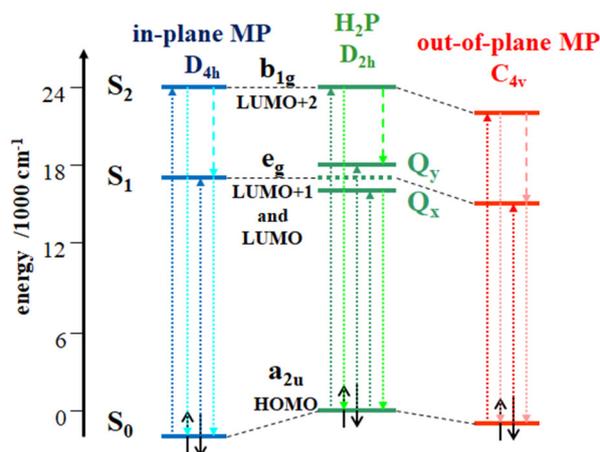


Figure 1. Structure of an in-plane metalloporphyrin {MTSPP=metallo-5,10,15,20-tetrakis(4-sulfonatophenyl)porphyrin} (a); and that of an out-of-plane complex (b) [9].

generally with intensities of one order of magnitude less. These latter bands in free-base ligands split due to the presence of protons on two diagonally situated pyrrolic nitrogens, to be more precise, as a result of the reduced symmetry (because of the disappearance of the four-fold rotation axis) compared to the metallated or deprotonated forms. This split is not detectable in the Soret range, hence, these two types of bands in the visible region are remarkably different [1].

In the Soret region, compared to the corresponding free-base ligands, the typical in-plane metalloporphyrins (e.g. Fe^{3+} , Au^{3+} , Cu^{2+} , Pd^{2+}) exhibit blueshifts because the atomic orbitals of their metal centers which are covalently bonded in the plane can overlap more strongly with the highest occupied molecular orbitals (HOMO) of the ligand, resulting in a stronger reduction in energy; whereas the lowest unoccupied molecular orbitals (LUMO) do not change. Thus, the energy gaps between the excited and ground states become greater. In the OOP complexes, the atomic orbitals of the more weakly bonded metal ions (e.g. Cd^{2+} , Hg^{2+} , Tl^{3+}) may slightly affect the unoccupied MOs and to a lesser extent the occupied ones, leading to a reduction of the energy gaps, i.e. an increase in the corresponding wavelengths (Scheme 1) [1, 9].



Scheme 1. Simplified energy level diagram for the change of the porphyrin's molecular orbitals in different types of complexes [9].

Beside electronic factors, due to the rigidity of the porphyrins' ringed structure, steric effects also influence the spectra: the redshift of absorption bands is one of the most common spectroscopic consequences of the non-planarity of porphyrin [3]. Octabrominated free-base porphyrin, $\text{H}_2\text{TSPBr}_8^{4-}$, was applied to investigate the spectrophotometric effects of the macrocycle's highly saddle-distorted structure (Fig.2). In porphyrins with aryl substituents, this distortion can lead to the extension of delocalization by the twisting of aryl substituents from a nearly perpendicular orientation closer to the porphyrin plane (Fig.1) [4].

The larger, post-transition metal ions, e.g. thallium(I), lead(II) and bismuth(III) ions, can cause a similarly large redshift of the porphyrins' absorption bands. Since their complexes possess the most highly dome-distorted structures, also a ruffled-like deformation of the periphery superposes on this high degree of doming. Considering the spectral effects (bathochromic or not quite exactly hyperchromic effects), the complexes possessing such highly redshifted absorption bands used to be referred to as hyperporphyrins; depending on the highest occupied electron subshell of the metal center, p- or d-type hyperporphyrins. Previously in terms of this categorization of metalloporphyrins, only the electronic effects of the metal ion (through its electron configuration) were taken into consideration and not steric (distorting) effects [1]. Nevertheless, in the typical d-type hyperporphyrins, e.g. the low-spin chromium(III), manganese(III), nickel(II) and cobalt(III) porphyrins, the radius of the metal center, and thus, the metal-nitrogen bonds are too short, resulting in the contraction of the coordination cavity, along with the ruffled deformation of the macrocycle [2, 5-7].

3.2. Equilibrium and kinetics of complex formation

Porphyrins are peculiar ligands in terms of complexation due to their planar, cyclic, rigid, aromatic,

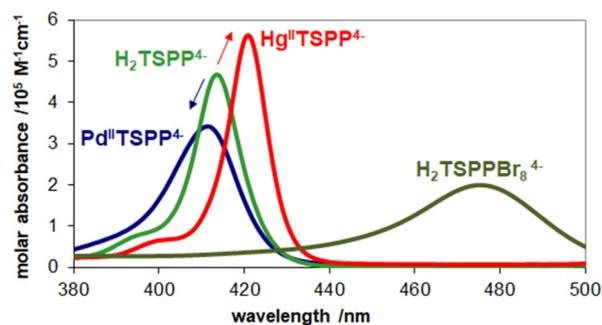


Figure 2. Absorption spectra of the free base ($\text{H}_2\text{TSP}^{4-}$); the highly distorted, octabrominated free base ($\text{H}_2\text{TSPBr}_8^{4-}$); a typical in-plane ($\text{Pd}^{\text{II}}\text{TSP}^{4-}$); and a typical out-of-plane metalloporphyrin ($\text{Hg}^{\text{II}}\text{TSP}^{4-}$) within the Soret range [1].

tetradentate, as well as protonated structure. The formation of an out-of-plane complex of a large metal ion is usually at least two orders of magnitude faster than that of an in-plane one since a smaller metal ion is not able to coordinate to all four pyrrolic nitrogens of the reaction intermediate, in the cavity of which the two protons also remain $\{\text{H}_2\text{-P-M}\}$. Therefore, dissociation of the metal ion is more favorable than that of the protons. Besides the insertion of a smaller metal ion, its dissociation from the in-plane complex of the end-product may be kinetically hindered due to the rigidity of the macrocycle [1].

Formation of the in-plane complexes used to be enhanced by the addition of a small concentration of a metal ion with larger ionic radius (e.g. Cd^{2+} , Hg^{2+} , Pb^{2+}) to the solution of the smaller one because the insertion of the larger metal ion into the ligand cavity is much faster. However, the OOP complex is considerably less stable. In its dome-distorted structure, two diagonal pyrrolic nitrogens are more accessible from the other side of the ligand, owing to the enhancement of their sp^3 hybridization, hence, the metal center can be easily exchanged for the smaller one [1-2, 9].

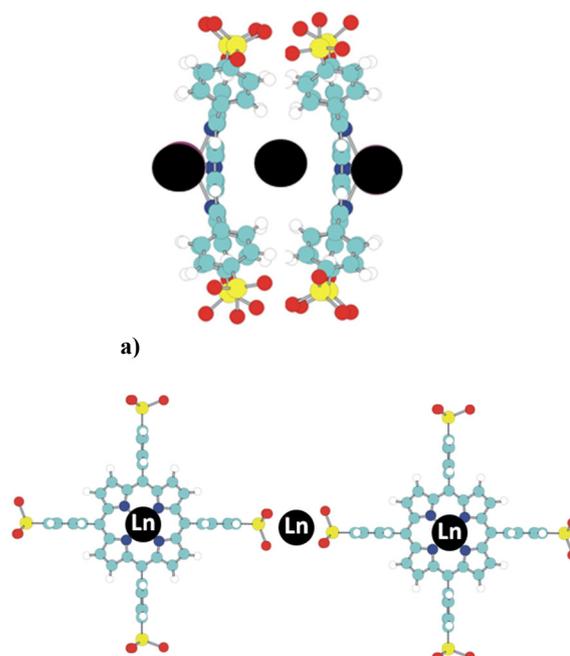
This accessibility makes the realization of dinuclear out-of-plane monoporphyrins (2:1 complexes) possible if the metal ion possesses a low (single) positive charge and is large, i.e. its charge density is small enough, e.g. mercury(I), silver(I) and thallium(I) ions [1, 8-9].

Moreover, the out-of-plane position of the metal center, together with the dome-distorted structure (owing to the twisting of aromatic substituents from a nearly perpendicular position closer to the porphyrin plane) may promote the formation of so-called sandwich complexes of various compositions, in which two metal ions can coordinate to one macrocycle, and, reversely, one metal ion can concomitantly coordinate to two ligands, (Fig.3) [1-2, 9]. Lanthanide(III) ions form typical examples of metallo-oligoporphyrins because they are inclined to form complexes of higher coordination number (8-12). However, they are hard Lewis acids, hence, their insertion into the coordination cavity of the softer N-donor porphyrin ligand is a slow and complicated process in aqueous solutions. This

phenomenon partly originates from the stability of their aqua complexes. Due to the consequence of their Pearson-type hard character, they coordinate rather to the peripheral substituents of porphyrin (instead of the pyrrolic nitrogens), i.e. to the ionic group ensuring water-solubility if they possess similarly hard O-donor atoms (e.g. carboxy or sulfonatophenyl groups). At lower temperatures, under kinetic control, the early lanthanide(III) ions are not able to coordinate into the cavity, rather to the periphery; resulting in the formation of the tail-to-tail dimer of free-base ligands (as the tail used to be referred to as the periphery). Higher temperatures and thermodynamic control are also necessary for the insertion of metal ions into the cavity produced by four pyrrolic nitrogens; resulting in the formation of typical metalloporphyrin complexes. After the discovery of the possible coordination bonds between lanthanide ions and sulfonato substituents, the formation of lanthanide bisporphyrins may be realized as a tail-to-tail dimerization of two metallo-monoporphyrin complexes through a metal bridge; deviating from the head-to-head connection as in the case of typical sandwich complexes (head refers to the cavity; see Fig.3). On the basis of our previous experiences, the coordination position of lanthanide ions was influenced by the change in temperature [1-2], [11-13].

During the investigation of the formation of “typical p-type hyperporphyrin” complexes (e.g. Tl^+ , Pb^{2+} , Bi^{3+}), the species possessing highly redshifted absorption bands are the end-products of metalation only in hydrophobic solvents, since they can appear in aqueous solutions as intermediates with shorter or longer lifetimes depending on the metal ion. The absorption spectra of the end-products of these transformation reactions are very similar (less redshifted) to those of the typical, common out-of-plane metallo-monoporphyrins (e.g. Hg^{II} -porphyrin in Fig.2). This phenomenon may be accounted for to the considerable coordination ability or the polarizing effect of water molecules, which can enable the complex to overcome the kinetic energy barrier towards the formation of the more stable structure, in which the metal center is located closer to the ligand plane, resulting in decreases in distortion, as well as redshift. Furthermore, “hyperporphyrins” can appear as intermediates in smaller amounts during the formation of typical, common out-of-plane metallo-monoporphyrins as well [1, 14].

In the case of “d-type hyperporphyrin” complexes (e.g. Mn^{3+} , Co^{3+} , Ni^{2+}), the low-spin and ruffled complex with a contracted cavity can exist in a spin-isomerization equilibrium with the high-spin and planar forms, which not only exhibits less redshift, but rather blueshifted absorption bands compared to those of the free-base ligand. This reaction can be influenced by the strength of the M-N bonds (owing to the electronic effects of peripheral or axial substituents), due to the size of the coordination cavity (owing to the substitution or saturation of methylenide bridges or pyrrolic carbons) [1-2, 5-7].



b) Figure 3. Potential structures of 3:2 bisporphyrin: (a) head-to-head or (b) tail-to-tail [2].

3.3. Photophysics

Porphyrins represent one of the most interesting groups of compounds in terms of photophysical properties and biological significance. Due to their rigid structure and aromatic electronic system, they display two types of fluorescence: beside their relatively strong singlet-1 fluorescence in the range of 550–800 nm, weak and rare singlet-2 luminescence is observable between 400 and 550 nm upon excitation of the Soret band (*Scheme 1*) [1].

The quantum yields of S_2 -fluorescence are about 3 orders of magnitude lower than those of S_1 -fluorescence in the case of free-base porphyrins, especially ~1200-fold for H_2TSP^{4+} (6.3×10^{-5} vs. 7.5%). However, this ratio decreases with metalation, mainly in the case of the formation of out-of-plane complexes. Since the structure of S_2 -excited porphyrins may be close to that of the dome-distorted OOP complexes that are already in the electronic ground state. Another consequence of this structural similarity (namely small Stokes shift) is that the directions of the shifts of S_2 -fluorescence bands invert (according to Soret absorption) between the in-plane (redshifted) and out-of-plane (blueshifted) complexes when compared to the free-base ligand [1].

Singlet-1 fluorescence bands exhibit blueshifts in both types (in-plane and out-of-plane) of metalloporphyrins, as a consequence of the aforementioned split in free-base ligands because of the presence of two protons, as well as their reduced symmetry (*Scheme 1*). Furthermore, almost all complexes exhibit similarly large Stokes shifts, as well as lifetimes and quantum yields. In the case of in-plane

metal centers, the spin-orbit coupling, as an electronic quenching effect, may be dominant. Whereas for typical out-of-plane metalloporphyrins, the distortion, as a steric effect, can enhance their non-radiative decay.

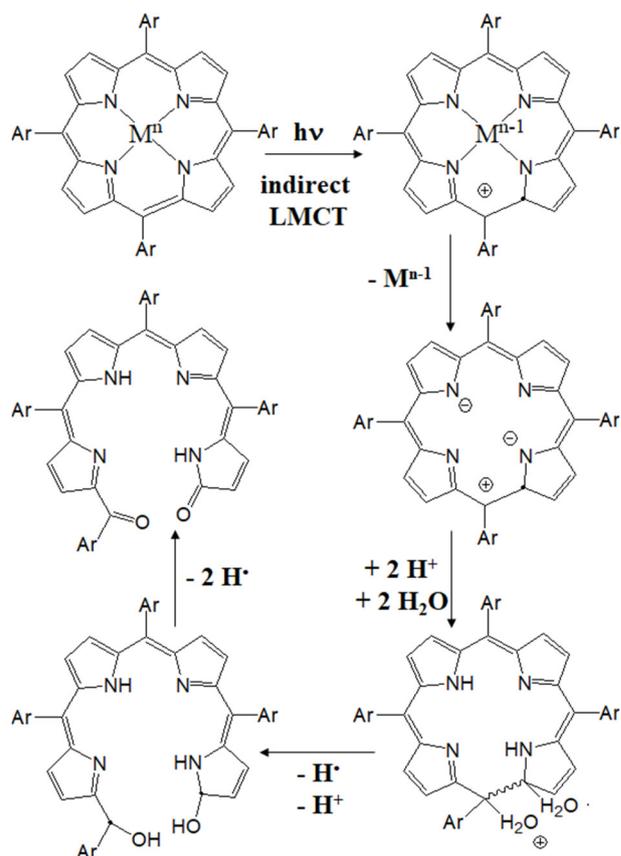
The highly distorted (d- and p-type) hyperporphyrins, the paramagnetic in-plane complexes (e.g. $\text{Fe}^{\text{III}}\text{TSP}^3$), as well as the head-to-head-type OOP bisporphyrins {e.g. $\text{Hg}^{\text{II}}_3(\text{TSP})_2^{6-}$ } do not exhibit significant levels of luminescence at room temperature. Conversely, the paramagnetic out-of-plane complexes (e.g. $\text{Ln}^{\text{III}}\text{TSP}^3$) possess similar fluorescence properties to that of the diamagnetic ones because a paramagnetic metal ion can cause the disappearance of fluorescence by spin-orbit coupling only if it is located in the plane. In the OOP position, it is not able to perturb as efficiently the molecule orbitals of the macrocycle that result in the common absorption and emission out-of-plane characteristics [1-2, 9].

Lanthanide(III) bisporphyrins $\{\text{Ln}^{\text{III}}_3(\text{TSP})_2^{3-}\}$ have many similarities in terms of absorption and emission properties to those of monoporphyrin complexes ($\text{Ln}^{\text{III}}\text{TSP}^3$). These may only originate from the very weak π - π interactions between the macrocycles in the tail-to-tail-type aggregations (Fig.3) [1-2, 11-13].

3.4. Photochemistry

Porphyrin derivatives are the main components of photosynthesis, synthetically as well. Since the overall quantum yield of fluorescence and intersystem crossing resulting in the formation of triplet states is in excess of 95%, merely a slight proportion of excitation energy is dissipated as heat from singlet states. This ratio is the major reason why porphyrins are efficient in terms of optical sensations and photosensitizations. Free-base and kinetically inert in-plane metalloporphyrins may be appropriate candidates to be applied in photocatalytic systems based on outer-sphere electron transfer. D-type hyperporphyrins can be particularly promising from this viewpoint owing to their distorted structure which may enhance the (photo)redox reactivity of these coordination compounds. In the presence of a suitable electron acceptor (methylviologen, MV^{2+}) and donor (e.g. triethanolamine, TEOA), these complexes proved to be efficacious photocatalysts that transfer electrons between the ground-state reactants through an outer-sphere mechanism, generating the $\text{MV}^{\cdot+}$ radical cation. This system can be applied for the production of hydrogen from water [2, 5-7].

Contrarily, the inner-sphere photoredox reactions are characteristic of the out-of-plane metalloporphyrins because of this special coordination (Scheme 2): an irreversible photoinduced charge-transfer from the ligand to the metal center (ligand-to-metal charge transfer, LMCT) improves the efficiency of charge separation, which allows their utilization as catalysts in cyclic processes for the synthesis of chemicals capable of conserving light energy, hopefully in terms of the photochemical cleavage of water. Due to photoinduced LMCT the charge of the metal center decreases and its size increases, overall its charge density diminishes,



Scheme 2. Simplified demonstration of the mechanism for the inner-sphere photoredox reaction of an out-of-plane metalloporphyrin [8].

hence, the coordinative bonds can easily split. The reduced metal ion can leave the cavity, primarily in polar solvents, and induce further redox reactions. The latter processes strongly depend on the stability of the reduced metalion in the actual medium. The oxidized and metal-free (cationic) radical of porphyrin is a very strong base: it is immediately protonated and forms the free-base radical, which is a long-lived and rather strong electron acceptor, especially in deaerated solutions. Since it would only oxidize water to oxygen at higher pHs, a slightly more efficient reducer (such as alcohols or aldehydes of low molecular weight) is needed, from which useful byproducts can be produced in terms of photocatalytic hydrogen generation. In the absence of any electron donor that promotes the regeneration of the porphyrin, it undergoes the primary photochemical processes; an overall four-electron oxidation involving a ring-cleavage, the end-product of which is a dioxo-tetrapyrrole derivative (bilindione). This ring-opening process can be followed by spectrophotometry owing to the disappearance of the Soret band, as well as the typical change in the region of Q bands [2, 4, 8-16].

Photochemical quantum yields of this ring-opening reaction (without regeneration) are about 2-3 orders of magnitude higher for the out-of-plane complexes ($10^{-4} - 10^{-2}$) than for the free-base and in-plane metalloporphyrins ($10^{-6} - 10^{-5}$). In addition, in the case of out-of-plane complexes, photoinduced dissociation in

the absence of a redox reaction can occur, originating from their lability, and structural transformations to another complex form or conformer were observed in some cases as a photoinduced change of the type or measure of distortion (e.g. d- and p-type hyperporphyrins) [2, 4, 8-10, 14-16].

Besides the typical post-transition metal ions, lanthanide(III) ions were also applied for out-of-plane coordination because their contraction makes the fine-tuning of the out-of-plane distance possible, and their high negative redox potentials promote the photoinduced cleavage of water. Photochemical activities of their complexes confirm that the redox potentials of the metal centers are not the main determining factor, rather their out-of-plane distances [2, 11-13].

Deviating from the OOP complexes of post-transition metal ions, another stable photoproduct was observable during the photolysis of lanthanide(III) porphyrins. It displays a typical absorption band in the Q range (at ~600 nm), which may be assigned as a charge transfer between the metal ion and open-chain, dioxo-tetrapyrrole derivative (bilindione, see *Scheme 2*). Its oxo-groups, as donor atoms, may coordinate with the lanthanide ions, as a consequence of their similar Pearson-type hard characteristics, contrary to the softer post-transition metal ions [2, 11].

During the photolysis experiments, only small differences appeared between lanthanide(III) mono- and bisporphyrin complexes, which might confirm a special type of aggregation through the peripheral sulfonato substituents with weak π - π interactions (tail-to-tail, see *Fig. 3*) [2, 11-13]. Deviating from these observations, the differences are much more significant in the case of the most typical, post-transition metallo-bisporphyrin compared to the monoporphyrim equivalent; namely between $\text{Hg}^{\text{II}}_3(\text{TSPP})_2^{6-}$ and $\text{Hg}^{\text{II}}\text{TSPP}^{4-}$. The overall quantum yield is ~2 orders of magnitude higher and the photoinduced dissociation of a metal ion became the dominant reaction in the head-to-head sandwich complex as a consequence of the strong π - π interactions [9-10].

3.5. Quantum chemical calculations

The main aims of our electronic structural calculations were to determine the primary consequences for the out-of-plane position of a metal center, and confirm the experimentally observed correlation between the UV-Vis spectral shifts and the coordination position of the metal center (in-plane or out-of-plane). In the light of these aspects, the unsubstituted porphin (H_2P , $\text{C}_{20}\text{H}_{14}\text{N}_4$) was used as a model for the calculations, instead of the tetrakis(sulfonatophenyl)porphyrin ($\text{H}_2\text{TSPP}^{4-}$, $\text{C}_{44}\text{H}_{26}\text{N}_4\text{O}_{12}\text{S}_4^{4-}$). On the basis of the few comparative calculations that were conducted, the phenyl-, as well as the sulfonatophenyl substituents have negligible effects on the coordination of the metal center in the cavity. However, they can significantly influence the formation of bisporphyrin complexes, even in the case of head-to-head structures [4, 10].

According to our quantum chemical experience, the value of the critical radius became ~100 pm instead of the experimentally suggested ~75-90 pm as a consequence of the significant expansion of the coordination cavity to coplanarly incorporate the metal ions. The proportion of borderline cases, i.e. complexes with questionable structures (somewhere between in-plane and out-of-plane), increased with further post-transition metal ions (e.g. Ag^{2+} [15], Cd^{2+} [4], Tl^{3+} [16]) that possess ionic radii of ~90-95 pm. Calculated bond lengths (M-N) and atomic distances (N-N) considerably deviate from the expected ones supposed on the basis of the values of the deprotonated porphyrins (P^{2-}). To describe this phenomenon, an axial ligand was applied to these metal centers to extract them out of the cavity. Consequently, expansion stopped, and the out-of-plane distance increased dramatically together with the degree of dome distortion and redshifts of absorption bands. From this point of view, two possible explanations can be supposed for the borderline-case complexes: the experimentally observed common OOP characteristics may originate from this expansion, tension; and small perturbations (e.g. the axial coordination in the calculation or photoexcitation in the experiments) may facilitate the metal center to adopt an out-of-plane position, too. Another possibility is that the method of calculation strongly prefers planar structures.

In our time-dependent density functional theory (TD-DFT) calculations, the correlation found between the measured and calculated shifts associated with the position of the metal center was not totally linear, but nevertheless acceptable. The main exceptions were the borderline cases (high-spin Mn^{2+} , Fe^{2+} and Zn^{2+}) and the d-type hyperporphyrins (because their structures were determined to be totally planar), as well as the p-type hyperporphyrins (because a ruffled-like deformation did not superpose on their dome-like structure).

The regression of correlation was much worse within the Soret band than in the case of the Q bands. The Soret band was also split in the calculations, which cannot be detected experimentally. On the basis of further experimental observations and doubts in the literature, the validity of the theoretical model in use at present is questionable. Hence, the development of a more suitable one is in progress.

4. Conclusion

In conclusion, it can be declared that the categorization of metalloporphyrins was complemented by the role of their distortion, which is primarily responsible for their spectral features, whereas the electronic structure of their metal centers is a secondary factor, with a considerable level of emphasis on the in-plane complexes. The position of the metal center (in-plane or out-of-plane) in the monoporphyrim complexes, as well as the type (head-to-head or tail-to-tail) of the bisporphyrin complexes can be determined on the basis of their UV-Vis absorption and emission properties.

Hyperporphyrin spectra can appear, owing to the peripheral substitution (octabromination) of free-base ligands. Furthermore, the high degree of redshift may disappear during the spin isomerization of d-type metalloporphyrins or the transformation of p-type ones. Consequently, the real origin cannot be an electronic but rather a steric effect, namely the measure of distortion, which can confirm the absence of their fluorescence.

In terms of photochemical activity, several dissimilarities were found between the in-plane and out-of-plane metalloporphyrins; the most remarkable of them was the mechanism of their photoredox reactions: outer-sphere electron transfer is typical of the previous ones, while the inner-sphere equivalent is most prevalent for the latter ones. As a further consequence of the OOP position of the metal center, photoinduced dissociation and transformation reactions can occur within their complexes.

In our electronic structural calculations, the number of borderline-case complexes expanded, on the basis of which common OOP characteristics that can be experimentally observed may acquire a novel explanation.

Acknowledgement

This research was supported by the Széchenyi 2020 Fund under the GINOP-2.3.2-15-2016-00016 and EFOP-3.6.1-16-2016-00015 projects.

Assistance with quantum chemical calculations provided by Professor György Lendvay (Research Centre for Natural Sciences, Hungarian Academy of Sciences) is gratefully acknowledged.

Finally, this manuscript is dedicated to the memory of Professor János Liszi, who, as the head of the Doctoral School for Chemistry at the University of Veszprém, praised the corresponding author's PhD dissertation of a similar title in 2007.

REFERENCES

- [1] Valicsek, Z.; Horváth, O.: Application of the electronic spectra of porphyrins for analytical purposes: the effects of metal ions and structural distortions, *Microchem. J.*, 2013 **107**, 47–62 DOI: 10.1016/j.microc.2012.07.002
- [2] Horváth, O.; Valicsek, Z.; Fodor, M.A.; Major, M.M.; Imran, M.; Grampp, G.; Wankmüller, A.: Visible light-driven photophysics and photochemistry of water-soluble metalloporphyrins, *Coord. Chem. Rev.*, 2016 **325**, 59–66 DOI: 10.1016/j.ccr.2015.12.011
- [3] Shelnut, J.A.; Song, X.-Z.; Ma, J.-G.; Jia, S.-L.; Jentzen, W.; Medforth, C.J.: Nonplanar porphyrins and their significance in proteins, *Chem. Soc. Rev.*, 1998 **27**, 31–41 DOI: 10.1039/A827031Z
- [4] Valicsek, Z.; Horváth, O.; Lendvay, G.; Kikaš, I.; Škorić, I.: Formation, photophysics, and photochemistry of cadmium(II) complexes with 5,10,15,20-tetrakis(4-sulfonatophenyl)porphyrin and its octabromo derivative: the effects of bromination and the axial hydroxo ligand, *J. Photochem. Photobiol. A*, 2011 **218**, 143–155 DOI: 10.1016/j.jphotochem.2010.12.014
- [5] Fodor, M.A.; Horváth, O.; Fodor, L.; Grampp, G.; Wankmüller, A.: Photophysical and photocatalytic behavior of cobalt(III) 5,10,15,20-tetrakis(1-methylpyridinium-4-yl)porphyrin, *Inorg. Chem. Commun.*, 2014 **50**, 110–112 DOI: 10.1016/j.inoche.2014.10.029
- [6] Fodor, M.A.; Horváth, O.; Fodor, L.; Vazdar, K.; Grampp, G.; Wankmüller, A.: Photophysical and photochemical properties of manganese complexes with cationic porphyrin ligands: Effects of alkyl substituents and micellar environment, *J. Photochem. Photobiol. A*, 2016 **328**, 233–239 DOI: 10.1016/j.jphotochem.2016.06.011
- [7] Major, M.M.; Horváth, O.; Fodor, M.A.; Fodor, L.; Valicsek, Z.; Grampp, G.; Wankmüller, A.: Photophysical and photocatalytic behavior of nickel(II) 5,10,15,20-tetrakis(1-methylpyridinium-4-yl) porphyrin, *Inorg. Chem. Commun.*, 2016 **73**, 1–3 DOI: 10.1016/j.inoche.2016.09.001
- [8] Horváth, O.; Valicsek, Z.; Harrach, G.; Lendvay, G.; Fodor, M.A.: Spectroscopic and photochemical properties of water-soluble metalloporphyrins of distorted structure, *Coord. Chem. Rev.*, 2012 **256**, 1531–1545 DOI: 10.1016/j.ccr.2012.02.011
- [9] Horváth, O.; Huszánk, R.; Valicsek, Z.; Lendvay, G.: Photophysics and photochemistry of kinetically labile, water-soluble porphyrin complexes, *Coord. Chem. Rev.*, 2006 **250**, 1792–1803 DOI: 10.1016/j.ccr.2006.02.014
- [10] Valicsek, Z.; Lendvay, G.; Horváth, O.: Equilibrium, photophysical, photochemical and quantum chemical examination of anionic mercury(II) mono- and bisporphyrins, *J. Phys. Chem. B*, 2008 **112**(46), 14509–14524 DOI: 10.1021/jp804039s
- [11] Imran, M.; Szentgyörgyi, C.; Eller, G.; Valicsek, Z.; Horváth, O.: Peculiar photoinduced properties of water-soluble, early lanthanide(III) porphyrins, *Inorg. Chem. Commun.*, 2015 **52**, 60–63 DOI: 10.1016/j.inoche.2014.12.016
- [12] Kiss, M.P.; Imran, M.; Szentgyörgyi, C.; Valicsek, Z.; Horváth, O.: Peculiarities of the reactions between early lanthanide(III) ions and an anionic porphyrin, *Inorg. Chem. Commun.*, 2014 **48**, 22–25 DOI: 10.1016/j.inoche.2014.08.001
- [13] Valicsek, Z.; Eller, G.; Horváth, O.: Equilibrium, photophysical and photochemical examination of anionic lanthanum(III) mono- and bisporphyrins: the effects of the out-of-plane structure, *Dalton Trans.*, 2012 **41**, 13120–13131 DOI: 10.1039/C2DT31189E
- [14] Valicsek, Z.; Horváth, O.; Patonay, K.: Formation, photophysical and photochemical properties of water-soluble bismuth(III) porphyrins: the role of the charge and structure, *J. Photochem. Photobiol. A*, 2011 **226**, 23–35 DOI: 10.1016/j.jphotochem.2011.10.011

- [15] Harrach, G.; Valicsek, Z.; Horváth, O.: Water-soluble silver(II) and gold(III) porphyrins: the effect of structural distortion on the photophysical and photochemical behavior, *Inorg. Chem. Commun.* 2011 **14**, 1756–1761 DOI: 10.1016/j.inoche.2011.08.003
- [16] Valicsek, Z.; Horváth, O.: Formation, photophysics and photochemistry of thallium(III) 5,10,15,20-tetrakis(4-sulphonatophenyl)porphyrin: New supports of typical sitting-atop features, *J. Photochem. Photobiol. A*, 2007 **186**, 1–7 DOI: 10.1016/j.jphotochem.2006.07.003

IMPROVEMENT POSSIBILITIES OF HETEROGENEOUS PHOTOCATALYSIS WITH THE AIM OF IN-FIELD USE

ORSOLYA FÓNAGY,¹ PÉTER HEGEDŰS,¹ ERZSÉBET SZABÓ-BÁRDOS,¹ ANNAMÁRIA DOBRÁDI,²
AND OTTÓ HORVÁTH^{1*}

¹Department of General and Inorganic Chemistry, University of Pannonia, Egyetem u. 10., H-8200 Veszprém, HUNGARY

²Institute of Materials Engineering, University of Pannonia, Egyetem u. 10., H-8200 Veszprém, HUNGARY

Heterogeneous photocatalysis can be successfully applied for the degradation of organic pollutants, although the efficiency of this method is insufficient. It can be increased by modification of the catalyst with precious metals (e.g. silver) or heterogeneous catalysis can be combined with other oxidative procedures such as ozonation. Another obstacle to the in-field use of the method is that separating the catalyst from the liquid phase is difficult. This problem can be eliminated by the immobilization of the catalyst. The poly(vinyl alcohol)-TiO₂ immobilized catalyst has been prepared for this purpose. During TiO₂-based photocatalysis, active oxygen species such as the hydroxyl radical, superoxide radical and hydrogen peroxide are produced. The formation rate of the oxidative •OH radicals was also determined in the case of the previously mentioned techniques. As a scavenger of this radical, coumarin was added.

Keywords: TiO₂-based photocatalysis; Triton X-100; Benzenesulfonic acid; Silver-deposited TiO₂; Immobilization; Thermal decomposition of PVA

1. Introduction

Nowadays, a great proportion of anthropogenic pollutants cannot be mineralized by traditional biological and physicochemical wastewater treatment procedures. Therefore, important developments in chemical wastewater treatment technologies and in their utilizations are required. For economic, technological and healthcare reasons, in terms of both type and quantity, it is important to use as few chemical additives as possible. Their applicability against remarkably different pollutants using minimal levels of energy consumption is even more important. Advanced Oxidation Processes satisfy these challenges [1]. A common characteristic of these processes is the generation of oxidative radicals, predominantly hydroxyl radicals, using solar radiation or other kinds of energy. Hydroxyl radicals are able to oxidize a great variety of organic compounds, therefore, heterogeneous photocatalysis has become an intensively studied field of research.

In the pharmaceutical industry, benzenesulfonic acid is mostly used for producing other speciality chemicals. A variety of pharmaceutical drugs are prepared as salts of benzenesulfonic acid, known as besilates. Triton X-100

with an average $n \approx 9.5$ is one of the most widely applied man-made non-ionic surfactants. Non-ionic surfactants are more stable than ionic tensides and not sensitive to pH and electrolytes. Besides the hydrophilic polyethoxy chain, it also contains a hydrophobic octylphenyl group. Both compounds can hardly be degraded by biological treatments [2–4]. Hence, a more efficient degradation method like heterogeneous photocatalysis is indispensable for the total mineralization of these pollutants.

1.1. Photodegradation of model compounds

The titanium dioxide-mediated photocatalyzed degradation of benzenesulfonic acid (later abbreviated as BS) was investigated by monitoring the absorption and emission spectral changes, sulfate concentration, pH, as well as the total organic carbon (later abbreviated as TOC) content in both argon-saturated and aerated systems.

During the degradation of benzenesulfonic acid, a characteristic change in the $\pi \rightarrow \pi^*$ transition (typical of aromatic systems) in the absorption spectrum of BS ($\lambda_{\text{max}}=262 \text{ nm}$, $\epsilon=439 \text{ M}^{-1} \text{ cm}^{-1}$) can be observed. At first, there is a slight increase in the 260–270 nm range, while two new shoulders appear at around 280 and 300 nm and as the period of irradiation progresses the level of absorbance decreases. The reason for the increase in absorbance observed in the spectrum of this aromatic surfactant is that the absorbances of the intermediates

*Correspondence: horvath.otto@mk.uni-pannon.hu

formed in the initial period of the process are higher than that of the starting compound. On the basis of the characteristic molar absorbances of 4-hydroxybenzenesulfonic acid (4-HBS) ($\lambda_{\text{max}}=277$ nm, $\epsilon=868$ M⁻¹ cm⁻¹) and 2,5-dihydroxybenzenesulfonic acid ($\lambda_{\text{max}}=302$ nm, $\epsilon=3300$ M⁻¹ cm⁻¹) as commercially available standards, the appearance of a shoulder at 280–300 nm suggests the formation of hydroxylated intermediates. The results indicate that the initial step of the degradation is hydroxylation of the starting surfactant, resulting in the production of hydroxy- and dihydroxybenzenesulfonates [5–7]. Therefore, liquid chromatography-mass spectrometry analysis was utilized for the detection of intermediates.

The sulfonate group of the substrate to be degraded can easily be oxidized by oxygen-containing reactants generated in these systems. Sulfate ions were detected in the irradiated reaction mixtures, which means that the hydroxylation reactions were accompanied by desulfonation. The initial rate of this process was lower in each case than that for the transformation (i.e. hydroxylation) of the starting compound (BS). These results indicate that the first step of mineralization is hydroxylation; desulfonation takes place afterwards.

During the photocatalytic oxidation of BS, especially in the first 180 min, a continuous decrease of pH was observed in both the argon-saturated and the aerated systems, i.e. strong acidification occurred. Probably, there is a direct correlation between the formations of hydrogen and sulfate ions, due to the following reaction (Eq.(1)), which can take place between the sulfonate group and hydroxyl radical, the main oxidizing agent in these systems:



The hydroxy species did not decay during the irradiation in the absence of dissolved oxygen. In the aerated system desulfonation and hydroxylation was much more efficient, moreover, a significant decrease in TOC took place during the initial stage. Further hydroxylation resulted in the cleavage of the aromatic system, through the formation of polyhydroxy derivatives, followed by ring-opening, which led to the production of aldehydes and carboxylic acids. Total mineralization was realized by the end of the photocatalysis. It has been proved that in this photocatalytic procedure the presence of dissolved oxygen is indispensable for the cleavage of the aromatic ring because hydroxyl radicals photochemically formed in the deaerated system alone are not able to break the C-C bond [5].

The degradation mechanism of Triton X-100 was also investigated [8]. In the case of this surfactant, hydroxyl radicals can attack in three ways; on the ethoxy chain, on the aromatic ring or on the alkyl chain. According to the absorption spectrum (no shift in the longer wavelength (275 nm) band and no new shoulders), no hydroxylation of the aromatic ring took place during the photocatalytic degradation. After the

total disappearance of the starting surfactant (after approx. 60 min), an appreciable degree of absorbance arose at 275 nm, which indicates that aromatic intermediates were formed during the first hour. In liquid chromatographic experiments, Triton X-100 produced a multiple-peaked chromatogram. The retention time of these non-ionic surfactant components is in strong correlation with the length of the ethoxy chain. Intermediates formed from various lengths of ethoxy chains and others with aromatic rings were detected in the reaction mixture after irradiation started. The concentration of the shorter-chain molecules increased during the initial stage (in the first 10 min) of the irradiation, indicating that the fragmentation of the longer polyethoxy chains of the starting components initially increased the concentration of those with shorter ones, which was followed by a gradual decay. The component with a relative molecular weight of 206 Dalton could be clearly identified. Based on its structure, it was identified as the alkyl phenol part of the original Triton X-100. Its concentration peaked at around the 50th minute of irradiation. This is the result of the total cleavage of the polyethoxy chains without any oxidation of the rest of the original tenside molecules. Based on the results so far, it can be stated that hydroxyl radicals attack the ethoxy side-chain and subsequently the alkyl group, afterwards, the ring opens.

1.2. Challenges of heterogeneous photocatalysis

The obstacle to the practical application of heterogeneous photocatalysis is that its efficiency is relatively low. Fortunately there are several ways to improve its mineralization effect: by the combination of heterogeneous catalysis with other oxidative procedures such as ozonation [6, 9–10] or by the modification of the catalyst with precious metals. The latter solution helps promote charge separation by accumulating the electrons on the surface of the silver particles [11–12].

The other obstacle is that the catalyst is in a suspended form, therefore, its separation from the liquid phase is difficult and makes the technology more expensive. This problem can be eliminated if the catalyst is immobilized. There are several methods for immobilization, for example, the catalyst particles can be fixed directly into polymers. Between a polymer and the catalyst usually a physical contact is formed, but in some cases, such as with poly(vinyl alcohol), a chemical bond can be formed. Unfortunately, this method is not free of disadvantages. The fact that immobilization decreases the active surface area of the catalyst has to be taken into consideration, thus, it will become less efficient [13–16].

Irradiation in the near UV range generates electron-hole pairs in the TiO₂ nanoparticles. Under aerobic conditions the electrons in the conduction band are scavenged by adsorbed oxygen molecules with a relatively high quantum yield, producing O₂•⁻ ions, which are readily protonated in acidic media. Besides,

the holes, in the absence of reducing species, oxidize the surface-adsorbed water or hydroxide ions to hydroxyl radicals. The $\bullet\text{OH}$ radical can also be formed by thermal reactions following the primary electron transfer processes that occur at the surface of semiconductor particles or in the liquid phase [13,17].

The hydroxyl radical is often assumed to be the major reactant responsible for the photooxidation of various organic molecules. The first step of the mineralization of benzenesulfonic acid is hydroxylation of the aromatic ring, while the degradation of Triton X-100 is initialized by hydroxyl radicals, which attack the ethoxyl side chain (Fig.1).

Due to the high reactivity and short lifetime of $\bullet\text{OH}$, the direct detection of this species is difficult. Appropriate methods such as steady-state fluorescence were applied to measure the quantity of scavenged hydroxyl radicals. Recent studies have proved that several non- or weakly luminescent test molecules, such as coumarin, produce strongly luminescent compounds (7-hydroxycoumarin) with hydroxyl radicals. Thus, these molecules can be utilized for detecting and measuring $\bullet\text{OH}$ radicals produced by UV excitation of TiO_2 particles in heterogeneous systems [18-19].

The main goal of our work was to investigate the effect of silver doping and immobilization of the catalyst applied for photocatalytic degradations. In this work, coumarin was used as a scavenger to determine the formation rate of hydroxyl radicals under various circumstances.

2. Experimental section

2.1. Materials

For the experiments, the following materials were used: benzenesulfonic acid ($\text{C}_6\text{H}_5\text{O}_3\text{S}$ - Alfa Aesar), 4-hydroxybenzenesulfonic acid ($\text{C}_6\text{H}_4(\text{OH})\text{O}_3\text{S}$ - Alfa Aesar), silver nitrate (AgNO_3 - Aldrich), Degussa P25 TiO_2 : 25 \pm 5% rutile, 75 \pm 5% anatase with a specific surface area of 50 $\text{m}^2 \text{g}^{-1}$ (now called Evonik AEROXIDE® TiO_2 P25), coumarin ($\text{C}_9\text{H}_6\text{O}_2$ - VWR), Triton X-100 ($\text{C}_{14}\text{H}_{22}\text{O}(\text{C}_2\text{H}_4\text{O})_{9.5}$ - Alfa Aesar), and poly(vinyl alcohol) ((later abbreviated as PVA) $M=146000-186000 \text{ g mol}^{-1}$, hydrolysis: 99+% - Aldrich). The materials used for the HPLC and IC measurements were the following: methanol (CH_3OH - Sigma-Aldrich), tetra-n-butylammonium bromide ($(\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_2)_4\text{NBr}$ - Sigma-Aldrich), sodium sulfate (Na_2SO_4 - VWR), 37% cc. hydrochloric acid (HCl - VWR) and sodium hydroxide (NaOH - VWR). High purity water, used in this study as a solvent, was double distilled and then purified using a Milli-Q system.

2.2. Analytics

For the analysis, 4 cm^3 samples were taken, the solid phase was removed by filtration using Millipore Millex-LCR PTFE 0.45 μm filters. The absorption spectral

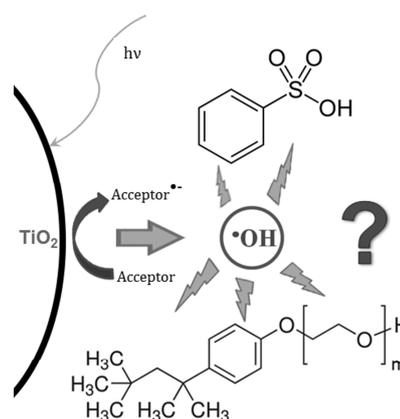


Figure 1. The role of $\bullet\text{OH}$ radicals in the degradation of the model compounds.

changes of the reaction mixtures irradiated were followed using a PerkinElmer Lambda 25 UV-Vis spectrophotometer, while the emission spectra were recorded using a PerkinElmer LS 50B spectrofluorometer. Both measurements were carried out in 1 cm quartz cuvettes. Mineralization was followed by measuring the total organic carbon (TOC) concentration, utilizing a Thermo Electron Corporation TOC TN 1200 apparatus. The pH of the aqueous phase of the reaction mixture was determined by a SP10T electrode connected to a Consort C561 instrument.

HPLC analyses of the benzenesulfonic acid samples were carried out using an Agilent 1100 instrument. A water:methanol (in a volume ratio of 95:5) solvent mixture containing 0.1 % (v/v) formic acid was the mobile phase with a flow rate of 1 $\text{cm}^3 \text{min}^{-1}$. During the experiments a 100 \times 2 mm Synergi Hydro-RP C18 (Phenomenex, Torrance, CA, USA) column packed with 2.5 μm particles was used. The column was thermostated at 30 $^\circ\text{C}$. The sample volume injected was 1 μl , photometric detection was applied at wavelengths of 262, 277 and 302 nm.

Degradation of Triton X-100 was followed by the 1290 Infinity UHPLC system (Agilent Technologies Inc., Santa Clara, CA, USA). The mobile phase was 65:35 methanol:water for five minutes of analysis, then it was gradually increased to 75:25 over the following five minutes. The signal was detected with a DAD detector. The column used during the experiments was a 100 \times 2 mm Synergi Hydro-RP C18 (Phenomenex, Torrance, CA, USA) column packed with 2.5 μm particles. The column was thermostated at 50 $^\circ\text{C}$. The flow rate of the eluent was 1 $\text{cm}^3 \text{min}^{-1}$.

During the study of the degradation mechanism an Agilent 6890N gas chromatograph with an Agilent DB-5ms Ultra Inert column (30 $\text{m}\times$ 0.25 $\text{mm}\times$ 250 μm) was used for the separation. As a detector, an Agilent 5973 N-type mass spectrometer was used.

The surface of immobilized catalysts was tested using a Philips/FEI XL30 scanning electron microscope at an accelerating voltage of 20 kV, using backscattered electron imaging. The composition was measured by an EDAX Genesis energy dispersive X-ray analyzer. Prior to testing, the samples were sputter coated by a 10 nm

thick layer of gold. Elements of the qualitative composition do not overlap with the gold, therefore, the gold coating did not disturb the analysis.

2.3. Experimental circumstances

The concentration of benzenesulfonic acid to be degraded was 10^{-3} M in each experiment, while that of the catalyst was 1 g dm^{-3} in all cases. For homogenization, the suspension was stirred for 20 minutes in the reactor before irradiation. The silver-deposited TiO_2 was prepared by photoreduction from a mixture initially containing 10^{-4} M AgNO_3 and 1 g dm^{-3} TiO_2 as a photocatalyst. After 1 hour of photoreduction the catalyst, brown in color, was separated by vacuum filtration then dried at room temperature. To re-suspend the catalyst powder before each experiment an ultrasonic bath was applied in order to achieve a dispersion of particles sufficiently small in size. During subsequent ultrasonic treatment, the substrate to be degraded was added to the system only after a period of time, so that the $\cdot\text{OH}$ radicals generated during the sonication could recombine and not tamper with the experimental results.

A detailed description of the production of immobilized catalyst can be found in our previous article [20]. The PVA- TiO_2 catalyst was prepared by using a PVA with a molecular weight of 146 000-186 000 and a 99+% degree of hydrolysis and Degussa P-25 TiO_2 . Immobilization was carried out using a solution-casting method [21]. The PVA- TiO_2 mixture was stirred at 90 then 60°C , subsequently a viscous solution was obtained, which was then poured into a petri dish and dried. The average weight of the immobilized catalyst was 0.47 g with a PVA content of 40%. Before its use, it was subjected to a 2 hour-long thermal treatment in an argon atmosphere to increase its stability. The Triton X-100 concentration was $2 \cdot 10^{-4}$ M with 1.5 g dm^{-3} TiO_2 . It is important to note that when comparing the efficiencies of the catalyst in various forms, the amount of TiO_2 immobilized in PVA (0.3 g) and used as a suspension in 200 cm^3 of reaction mixture were both 1.5 g dm^{-3} .

Photochemical experiments were carried out using two different reactors and light-source setups.

One was a laboratory-scale reactor with an effective volume of 2.5 dm^3 , equipped with a 40 W light tube located in the middle of the reactor. It emitted most of its energy at wavelengths exceeding 300 nm ($\lambda_{\text{max}}=350 \text{ nm}$, i.e. within the UVA range) [5]. The photon flux of the light source was $I_0 = 4.3 \cdot 10^{-6} \text{ mol photon dm}^{-3} \text{ s}^{-1}$, which was measured by the actinometry of trisoxalatoferate(III). The heterogeneous reaction mixture was circulated by using a peristaltic pump through the reactor and a buffer vessel. Compressed air was bubbled through the reaction mixtures from gas bottles, to facilitate stirring and (with its O_2 content) as an electron acceptor. O_3 was produced by a LAB2B ozone generator, and introduced into the same airstream. The ozone concentration was determined by iodometry, using sodium iodide as a

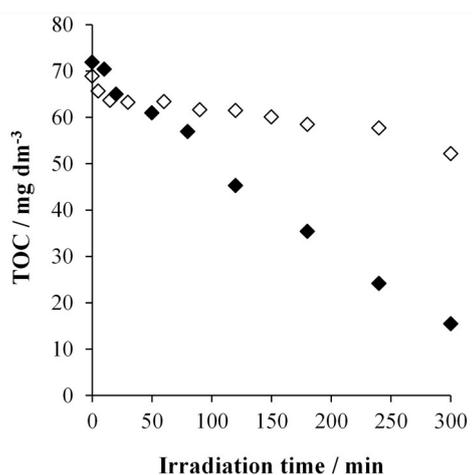


Figure 2. Change in the TOC concentration in the case of two different light sources used for irradiation.

$c(\text{BS})_0=10^{-3}$ M, air: $40 \text{ dm}^3 \text{ h}^{-1}$, $c(\text{TiO}_2)=1 \text{ g dm}^{-3}$

◇ – air+ TiO_2 +solar simulator ◆ – air+ TiO_2 +UV lamp

reagent and sodium thiosulfate for the titration of the iodine formed. The ozone dosage was estimated to be 0.35 mM min^{-1} .

The other setup consisted of one small reactor made of borosilicate glass (its volume was 250 cm^3) [20]. The catalyst film was placed on a perforated glass platform (foil holder) standing on a glass frit, through which compressed air was bubbled from a gas bottle at a flow rate of $10 \text{ dm}^3 \text{ h}^{-1}$, ensuring a constant oxygen concentration. The 200 cm^3 reaction mixture (Triton X-100 solution or distilled water in the case of pre-treatment) was circulated by a magnetic stirrer. It was irradiated from above using an Oriol LCS-100 solar simulator resulting in a light intensity of 72 mW cm^{-2} on the surface of the reaction mixture.

Fig. 2 shows the change in TOC during the photodegradation of benzenesulfonic acid under various experimental conditions. The band-gap energy of TiO_2 is 3.2 eV, indicating that the catalyst can be excited by light of 360 nm in wavelength. The difference between the two light sources used is clearly visible. The radiation intensity of the solar simulator in the UV range is very low, so the total organic carbon content of the initial solution decreased by only 10 mg over 360 minutes, while total mineralization was achieved using the UV lamp within this period.

3. Results

3.1. Increase in photocatalytic efficiency using various methods

As mentioned previously, the efficiency of heterogeneous photocatalysis can be increased in multiple ways: by combination with other oxidative procedures such as ozonation and by modification of the catalyst surface using noble metals.

Since the first reaction in the mechanism of photocatalytic oxidative degradation of benzenesulfonic

acid is hydroxylation of the aromatic ring, the determination of the rates of formation of hydroxyl radicals as the main oxidant in these systems was indispensable as far as the elucidation of the mechanism under various circumstances was concerned. Coumarin was used as an $\bullet\text{OH}$ scavenger.

An intensive band at 278 nm and a shoulder at around 300 nm is visible in the absorption spectrum of coumarin. The level of absorbance quickly decreased under irradiation (Fig.S1), while the concentration of the fluorescent 7-hydroxycoumarin rapidly increased during the initial period, and subsequently decreased gradually, i.e. degraded (Fig.S2). By depicting the intensity of the emission line at 453 nm as functions of the irradiation time, the initial rate of formation of 7-hydroxycoumarin, which is proportional to that of the hydroxyl radicals formed, can be determined from the initial gradients of the increasing sections, so it is easy to compare the ability of each process to generate hydroxyl radicals [18].

Before measuring the formation of 7-hydroxycoumarin in the photocatalytic oxidative systems that apply ozone, blind probes were used (Fig.3). The reaction between coumarin and ozone in the dark (at 10^{-4} M coumarin and 0.35 mM min^{-1} ozone concentrations) yielded a negligible spectral change after 2 hours. Although ozone itself is a strong oxidant, a very low rate of formation of 7-hydroxycoumarin ($0.32 \text{ INT min}^{-1}$) was observed under these experimental conditions. Upon UV irradiation of this probe, the rate of formation slightly increased (to $1.84 \text{ INT min}^{-1}$).

Ozone can be absorbed over a wide range, from infrared to vacuum UV, but is optimal at 254 nm. The light source applied in this work emits most of its radiation in the 300-380 nm range, where the molar absorbances of O_3 are rather low. Thus, it can poorly promote the dissociation of ozone via the homolytic cleavage of a bond.

Contrary to the blind probe with ozone, significant spectral changes were observed in irradiated reaction

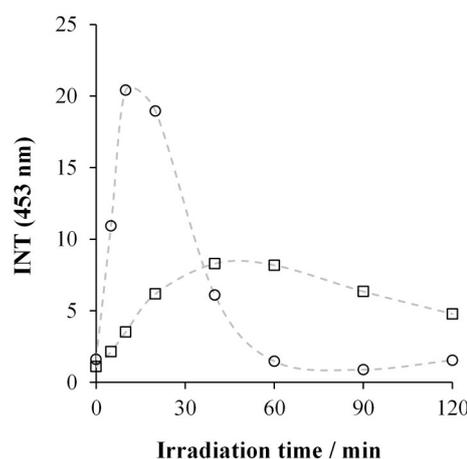


Figure 3. Change in the luminescence intensity of 7-hydroxycoumarin formed ($\lambda_{\text{exc}}=332 \text{ nm}$). $c(\text{coumarin})_0=10^{-4} \text{ mol dm}^{-3}$, air: $40 \text{ dm}^3 \text{ h}^{-1}$, $c(\text{O}_3)=0.35 \text{ mM}$, $\ell=1 \text{ cm}$
 ○ - O_3+UV
 □ - O_3

mixtures containing catalysts. The surface of the catalyst was also modified with silver and combined with ozonation, then the difference between the effects of the two methods of improvement was investigated by determining the rate of formation of oxidative $\bullet\text{OH}$ radicals (Fig.4).

The combination of heterogeneous photocatalysis with ozonation resulted in a 2.5 fold increase in the rate of formation (from $7.42 \text{ INT min}^{-1}$ to $19.4 \text{ INT min}^{-1}$). A further increase was achieved by the application of surface-modified (silver-deposited) titanium dioxide (to $21.9 \text{ INT min}^{-1}$). These results are in accordance with the fact that silver deposited on the surface of the catalyst diminishes the probability of the recombination of the photogenerated electron-hole pair, i.e. promotes charge separation [11].

3.1.1. Degradation of benzenesulfonic acid

Our research group has been studying the combination

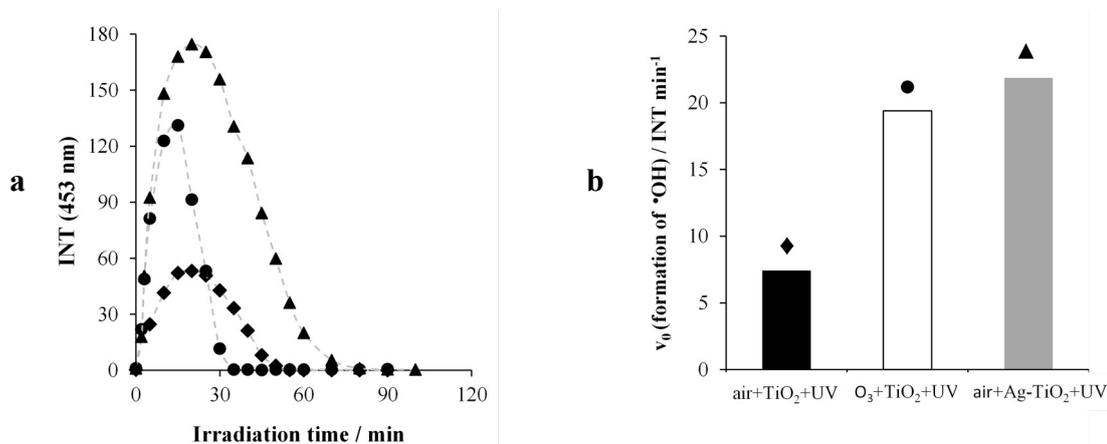


Figure 4. Change in the luminescence intensity of 7-hydroxycoumarin produced ($\lambda_{\text{exc}}=332 \text{ nm}$) (a) and rates of formation of $\bullet\text{OH}$ radicals under various circumstances (b). $c(\text{coumarin})_0=10^{-4} \text{ mol dm}^{-3}$, air: $40 \text{ dm}^3 \text{ h}^{-1}$, $c(\text{O}_3)=0.35 \text{ mM}$, $c(\text{catalyst})=1 \text{ g dm}^{-3}$, $\ell=1 \text{ cm}$
 ◆ - air+TiO₂+UV ● - O₃+TiO₂+UV ▲ - air+Ag-TiO₂+UV

Table 1. Comparison of the initial rates*.

	TOC	Benzenesulfonic acid
	[mg dm ⁻³ min ⁻¹]	[mM min ⁻¹]
air+ Ag-TiO ₂ +UV	1.58·10 ⁻¹	2.3·10 ⁻²
air+TiO ₂ +UV	2.24·10 ⁻¹	6.5·10 ⁻³
O ₃ +TiO ₂ +UV	5.92·10 ⁻¹	1.3·10 ⁻²

*The error of the rate determinations is ±10%.

of heterogeneous photocatalysis and ozonation in detail. It has been established that a joint application of these procedures results in synergic effects [6].

Another possibility of enhancing the efficiency of photocatalysis is the modification of the catalyst surface by the deposition of silver. Our results in this work are compared with previous observations [17].

The spectral changes during the irradiations of suspensions containing Ag-TiO₂ agree with earlier results. The degree of light absorption increases during the initial period (until the 50th minute), then it decreases (Fig.S3). The fine-band structure gradually disappears, while new bands arise at longer wavelengths (277 and 302 nm), suggesting the formation of isomers of hydroxy- and dihydroxybenzenesulfonic acid [6].

Characteristic changes can also be observed in the emission spectra of the samples in these experiments: an increase in the intensity of luminescence for 60 minutes, then a decrease accompanied by a red-shift of the 290 nm band (characteristic of the starting compound) and the appearance of a new emission band at about 330 nm (Fig.S4). The band-shift becomes considerable (46.5 nm) after the first hour of irradiation: from 286 nm to 333 nm. This finding also confirms the formation of hydroxy and dihydroxy intermediates with $\lambda_{em,max.}=305$ nm and 352 nm, respectively.

The changes in the concentrations of benzenesulfonic acid and the hydroxylated intermediates were followed by HPLC analyses (Fig.5). Of the three possible hydroxy isomers, the absolute concentration could only be determined for the commercially available 4-hydroxybenzenesulfonic acid (4-HBS, $r_t=0.86$ min).

Although the concentration of the substrate gradually decreases, the content of the benzenesulfonic acid in the reaction mixture is practically zero by the 180th minute of irradiation. On the basis of the concentration vs. time plot, the rate of decay of the model compound was determined by polynomial fit: $2.1 \cdot 10^{-2}$ mM min⁻¹. Regarding the heterogeneous catalytic experiments (untreated TiO₂ catalyst), $8 \cdot 10^{-3}$ mM min⁻¹ was the degradation rate of the starting surfactant, while $1.4 \cdot 10^{-2}$ mM min⁻¹ was measured in the case of combination with ozonation (Table 1) [6].

Mineralization of the model compound was followed by TOC (total organic carbon content) measurements. Taking the difference between the TOC value of the reaction mixture (TOC of the solution) and that of the benzenesulfonic acid (TOC of BS, calculated from its actual concentration), the TOC value regarding

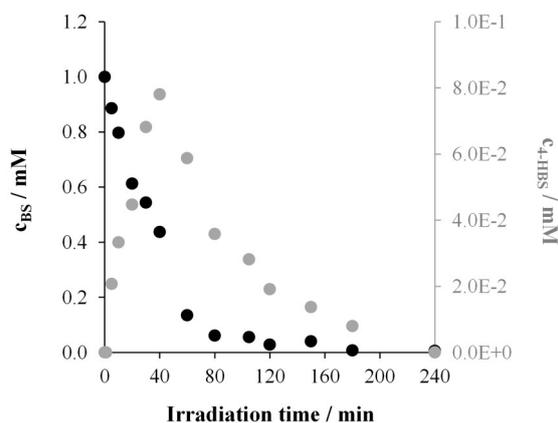


Figure 5. Transformation of the initial compound and change in concentration of the 4-HBS intermediate. $c(\text{BS})_0=10^{-3}$ M, air: $40 \text{ dm}^3 \text{ h}^{-1}$, $c(\text{Ag-TiO}_2)=1 \text{ g dm}^{-3}$

the intermediates in the solution can be obtained (Fig.6). This figure also displays the TOC values belonging to 4-HBS (calculated from its actual concentrations, see Fig.5)

The concentration of the starting material and its TOC value rapidly decrease in the heterogeneous photocatalytic experiments involving silver-modified TiO₂, while the amount of intermediates (predominantly hydroxy and dihydroxy derivatives) gradually increases. The initial rate of decrease in TOC for the reaction mixture was $1.58 \cdot 10^{-1} \text{ mg dm}^{-3} \text{ min}^{-1}$.

The first step of the degradation of the model compound is hydroxylation. The highest initial rate of decay of BS was observed in the case of the system containing the silver-modified catalyst (Fig.7, Table 1). These results are in good agreement with those obtained for the experiments using coumarin, an efficient scavenger of $\bullet\text{OH}$. The highest initial rate of formation of hydroxyl radicals was obtained on the Ag-TiO₂

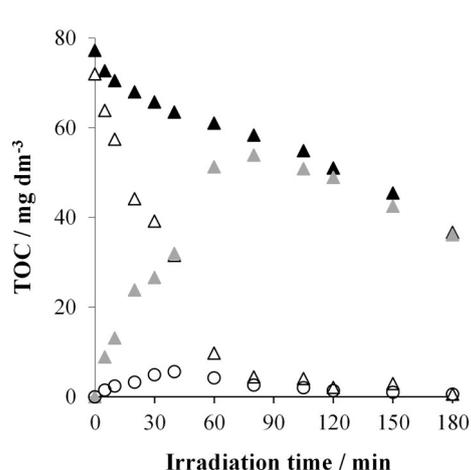


Figure 6. Change in TOC contents during the treatment.

$c(\text{BS})_0=10^{-3}$ M, air: $40 \text{ dm}^3 \text{ h}^{-1}$, $c(\text{Ag-TiO}_2)=1 \text{ g dm}^{-3}$
 ▲ – TOC of the solution (measured)
 △ – TOC of BS (calculated)
 ▲ – TOC of intermediates (calculated)
 ○ – TOC of 4-HBS (calculated)

catalyst.

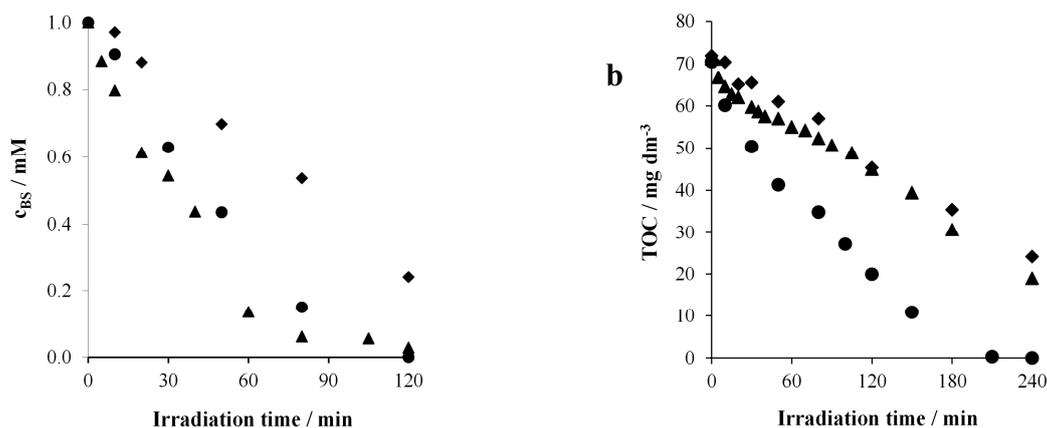


Figure 7. Change in concentrations of BS (a) and TOC (b) under various circumstances.

$c(\text{BS})_0=10^{-3}$ M, air: $40 \text{ dm}^3 \text{ h}^{-1}$, $c(\text{O}_3)=0.35$ mM, $c(\text{catalyst})=1 \text{ g dm}^{-3}$

◆ – air+TiO₂+UV • – O₃+TiO₂+UV ▲ – air+Ag-TiO₂+UV

The cleavage of the aromatic ring is indispensable in terms of the mineralization of the model compound, besides the fragmentation of the hydrocarbon chain. Szabó-Bárdos *et al.* rendered it probable that ring-opening takes place in the reaction of hydroxylated intermediates with reactive oxygen-containing species ($\text{O}_2^{\cdot-}/\text{HO}_2^{\cdot}$, $\text{O}_2(^1\Delta\text{g})$) [5]. Deviating from the decrease in the BS concentration, the lowest rate of the decrease in TOC, an indicator of mineralization, was observed in the case of the silver-modified catalyst, while the highest value (3.7 times higher than the previous one) was obtained for the combined procedure (Fig. 7, Table I).

Based on our results, the application of heterogeneous photocatalysis with silver-modified TiO₂ only accelerates the formation of hydroxyl radicals, while its combination with ozonation also increases the amount of other oxidative radicals ($\text{O}_2^{\cdot-}/\text{HO}_2^{\cdot}$, $\text{O}_2(^1\Delta\text{g})$).

3.2. Immobilization of the catalyst with poly(vinyl alcohol)

The thermal treatment of the immobilized catalysts, the preparation of which was thoroughly described in

Section 2.3, resulted in significant physical and chemical changes; on the one hand, their thickness varied (Fig. 8), on the other hand, their color turned from their original white to brown (Fig. 9). Song *et al.* described similar results [22].

The thickness of the foils prepared was tested by scanning electron microscopy prior to and after the thermal treatment. Following the treatment, the average thickness decreased from about 75.6 μm to about 69.6 μm (Fig. 8). The shrinkage of about 6 μm can be attributed to the evaporation of water content.

It has been established in our earlier studies that during the thermal treatment the polymer underwent dehydroxylation and cracking processes; compounds containing double and triple bonds as well as aromatic rings appeared on the surface of the foil, increasing its degree of light absorption [20]. Irradiation of the immobilized catalysts in the absence of the model compound led to the dissolution of the side-products of the cracking processes, and they also underwent heterogeneous photocatalytic degradation due to the presence of TiO₂. Accordingly, stearic acid and its derivatives containing double bonds such as linoleic acid (9,12-octadecadienoic acid (Z,Z)- and oleic acid (9-octadecanoic acid) were detected by GC-MS analysis of

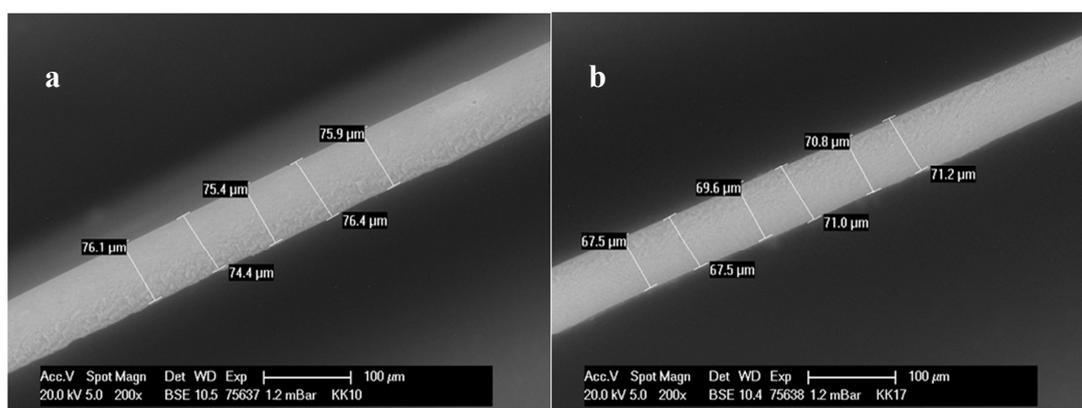


Figure 8. Cross sections of scanning electron micrographs of PVA-TiO₂ foils before (a) and after (b) thermal treatment.

the liquid phase. During the photocatalytic degradation, organic acids with shorter hydrocarbon chains were also produced, e.g. tetradecanoic acid (C14), caprylic acid (C8) and malic acid (C4) [20]. Although thermal treatment improves the stability of the foil, it restricts its efficiency because a considerable proportion of the oxidative radicals formed during photocatalysis are consumed in the reactions with the pollutants produced during the thermal treatment. Taking all these factors into consideration, how the temperature of the thermal treatment affects the stability of the immobilized catalyst was also studied.

The foils heated at different temperatures were submerged in 200 cm³ of distilled water, then irradiated for 8 hours. The change in TOC of the liquid phase is shown in *Fig.10/a*, while the reductions in the foil masses are displayed in *Fig.10/b*. For the interpretation of the results, the increase in the temperature of the solution due to irradiation also has to be taken into account, because it enhances the solubility of poly(vinyl alcohol), i.e. the amount of the dissolved PVA, which increases the actual TOC value. The organic materials (PVA and the products of the thermal treatment) dissolved off the foil which also underwent photocatalytic degradation in the presence of TiO₂. Both processes determine the change in TOC in the liquid phase, along with the mass reduction of the foil. The stability of the foils treated at lower temperatures (e.g. at 60°C) is rather modest, the PVA is dissolved in water, the TOC value of the liquid phase increased to 130 mg dm⁻³ by the end of the experiment, and the mass of the foil decreases considerably. The stability of the foils treated at higher temperatures (e.g. at 140°C or 200°C) significantly increased, while the TOC value of the liquid phase only grew slightly (to 34 mg dm⁻³ at 140°C, to 18 mg dm⁻³ at 200°C), in a similar fashion to the weight loss of the foils (140°C - Δm=93 mg, 200°C - Δm=24 mg).

The results clearly indicate that by increasing the temperature of the thermal treatment, the stability of the foil is gradually improved, due to efficient crosslinking [21]. The thermally treated immobilized catalysts cannot be applied for photocatalytic purposes, because they undergo degradation themselves during the irradiation, increasing the TOC value of the solution

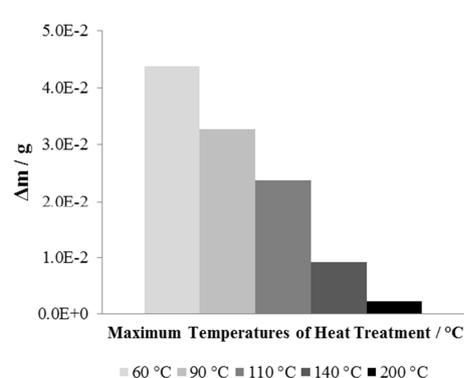
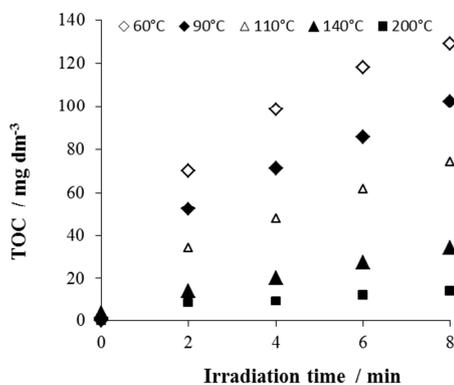
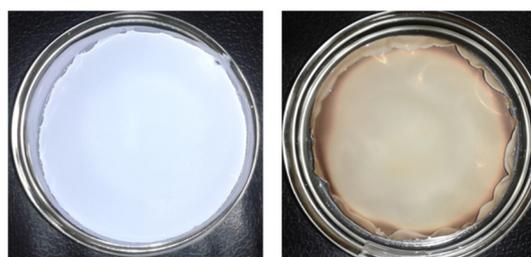


Figure 10. Change in the TOC content of the liquid phase after various heat treatments (a) and the weight loss of foils after 8 hours of irradiation (b) 200 cm³ distilled water, air: 10 dm³ h⁻¹, c(TiO₂)=1.5 g dm⁻³



Before heat treatment After heat treatment

Figure 9. Immobilized catalysts before and after heat treatment.

phase. Besides, the brownish organic compounds formed in the cracking processes diminish the catalytic efficiency because they occupy the active sites on the catalyst surface. Hence, a compromise has to be made between stability and efficiency. Efforts were made to mineralize the products of cracking during a pre-treatment. Before the photocatalytic experiments, the thermally treated foils were cleaned by three sequential 8-hour-long pre-treatments, to remove the pollutants formed during the cracking processes. Since the pollutants formed whilst heated at the higher temperature (200°C) were rather difficult to remove, 140°C was applied for the thermal treatment of the foils in further experiments of ours [20].

3.2.1. Efficiency of •OH production during pre-treatments of immobilized catalysts

To detect the hydroxyl radicals formed during the irradiation of TiO₂ particles embedded in the polymer, 10⁻⁴ M of coumarin as a scavenger was applied. Following the thermal treatment to enhance the stability of the foils, the undesirable products of this procedure have to be removed. In the first cycle of this process, when the foil is yet brown, the immobilized photocatalyst starts to work upon irradiation, but oxidative radicals formed in this process are mostly involved in the degradation of the products of cracking. The initial rate for the formation of 7-hydroxycoumarin just slightly exceeds the value observed in blind experiments (in the absence of a catalyst); 3·10⁻² INT min⁻¹ (blind), 4·10⁻² INT min⁻¹ (1st cycle) (*Fig.11*).

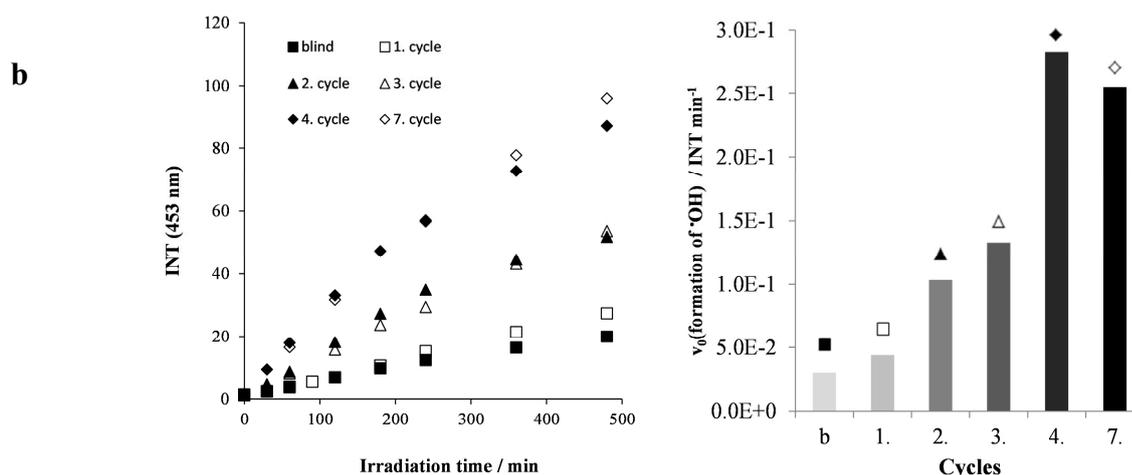


Figure 11. Change in the intensity of luminescence of 7-hydroxycoumarin formed.
 $c(\text{coumarin})_0 = 10^{-4} \text{ M}$, air: $10 \text{ dm}^3 \text{ h}^{-1}$, $\lambda_{\text{exc}} = 332 \text{ nm}$

The rate of formation of the hydroxyl radicals significantly increases during the 2nd cycle; the emission intensity at the end of this cycle ($v_0 = 10^{-1} \text{ INT min}^{-1}$) is more than twice as high as in the previous one. The results obtained during the 3rd cycle are similar to those during the 2nd one. Practically, during the three sequential 8-hour-long periods of irradiation, the polluting compounds that originate from the thermal treatment are totally degraded, hence, the active sites on the surface of the catalyst become more accessible, resulting in a much higher rate for the formation of oxidative radicals (in the 4th cycle $v_0 = 2.5\text{--}2.8 \cdot 10^{-1} \text{ INT min}^{-1}$, see Fig. 11). In other applications, the radical producing efficiency of the catalyst does not change significantly, i.e. the foil can be applied at constant efficacy. The lower value determined during the 7th cycle suggests a slow rate of degradation of the immobilized catalyst, which has to be taken into consideration during its application.

3.2.2. Degradation of the model compound

The photocatalytic degradation of the model compound was investigated following the cleaning procedure. The average degree of degradation of Triton X-100 in three sequential cycles was 62%; the deviations might originate from measurement errors (Fig. 12). This value is significantly lower than that observed in the case of the suspended catalyst (90%). Although the catalyst concentrations with regard to the volume of solution were equal (1.5 g dm^{-3}), immobilization considerably diminished the specific surface of the catalyst, leading to a substantial decrease in its efficiency. Mineralization of Triton X-100 takes place via cleavage-producing intermediates. The cleavage of the alkyl-phenyl part of the starting components of Triton X-100 took place from the beginning of the irradiation period. As complete intermediates originated from such a cleavage, ethylene glycol and dioxolane derivatives were detected by GC-MS. At the beginning of the irradiation the edge of the starting molecules or the initially generated intermediates mineralize to CO_2 . Therefore, the decrease in TOC is much less than

expected from the concentration change. The concentration changes of Triton X-100 components and the typical intermediates are in accordance with our previous results using a similar system that applied a suspended catalyst (see section 1.1). Remarkably, GC-MS analyses did not detect any intermediates which might derive from the PVA support. This suggested that three pre-treatment cycles of irradiation and rinsing were sufficient for a satisfactory level of removal of the derivatives formed during the thermal treatment.

3.2.3. Surface morphology of foils - SEM analysis

The behavior of the active surface of immobilized photocatalysts after repeated use raises an important issue. The surfaces of titanium dioxide containing PVA before and after thermal treatment are shown in Figs. 13/a and 13/b, respectively. Some TiO_2 -rich islands (Fig. 14) are not fully covered by the PVA film, these are visible on the surface. Otherwise, no significant changes are observed after the thermal treatment. However, simulated solar radiation (repeated usage) induces visible changes in the morphology of the

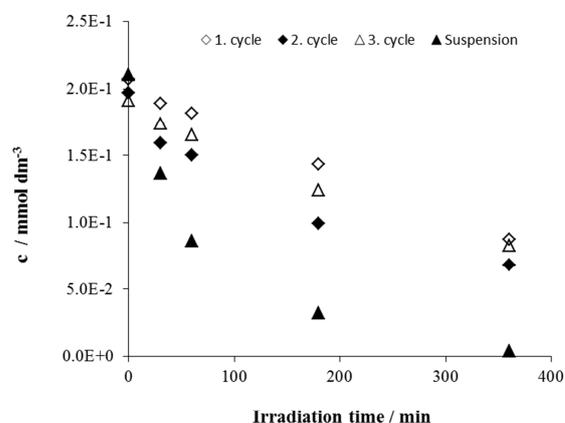


Figure 12. Change in concentration change of Triton X-100 during the irradiation.
 $c(\text{TX-100})_0 = 2 \cdot 10^{-4} \text{ M}$, air: $10 \text{ dm}^3 \text{ h}^{-1}$, $c(\text{TiO}_2) = 1.5 \text{ g dm}^{-3}$

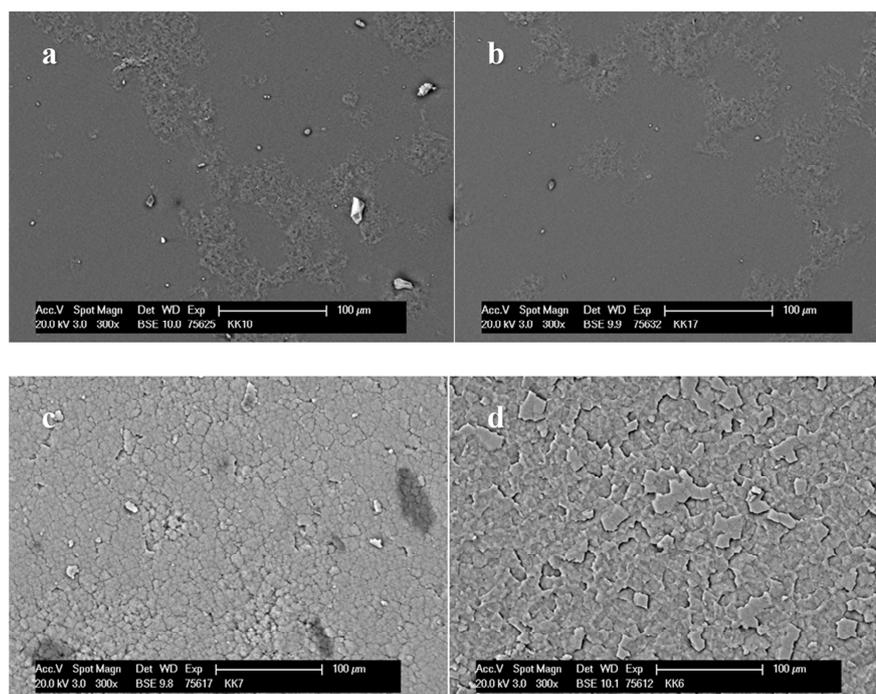


Figure 13. Scanning electron micrographs of foil surfaces.
a) before and b) after thermal treatment c) after 4 irradiation cycles d) after 7 cycles of irradiation

surface. Such a period of radiation partially or totally removes the PVA layer covering the TiO_2 agglomerates. This process is enhanced after prolonged use: after four cycles the homogeneously distributed particles are already visible (Fig. 13/c). Extended use - seven cycles - will enhance grain boundaries by removing more of the PVA matrix, indeed some of the particles become loosely bounded to the rest of the agglomerate (Fig. 13/d). The observed changes in thickness strongly correlate with the degradation of films (Fig. 15). After four cycles, a significant decrease in thickness can be observed: the foil thickness decreases to about $58.5 \mu\text{m}$ (from $69.6 \mu\text{m}$). This corresponds to a decrease of about $2.7 \mu\text{m} / \text{cycle}$. As for the thickness measured after seven cycles, the last three cycles only represent a decrease of about $0.9 \mu\text{m}/\text{cycle}$. These observations imply a more intense degree of degradation during the initial stage.

4. Conclusion

Heterogeneous photocatalytic procedures are widely applied for the degradation of organic pollutants. The degradation of benzenesulfonic acid and Triton X-100 by heterogeneous photocatalysis was investigated under various circumstances. The formation of hydroxyl radicals as one of the major reactants responsible for the photooxidation of these molecules was also measured in the case of both setups of the catalyst.

Silver deposition on the surface of the TiO_2 catalyst significantly increased the initial degradation (hydroxylation) rate of BS, due to the hindered recombination of the photogenerated electron-hole pairs, increasing the rate of $\bullet\text{OH}$ production. However,

mineralization of the intermediates was not accelerated by this method, deviating from the combination of the TiO_2 -based photocatalysis with ozonation. The degradation efficiency in the latter procedure indicated a synergic effect on the mineralization rate as a consequence of the enhanced production of oxidative agents other than $\bullet\text{OH}$ radicals [6]. These reactive oxidants play key roles in the ring-opening processes needed for the mineralization of aromatic intermediates. Quantitative determinations of such reactants (e.g. $\text{O}_2^{\bullet-}$ and $\text{O}_2(^1\Delta_g)$) to gain a deeper understanding of the mineralization mechanism are in progress. Our results indicate that the methods applied for the enhancement of the degradation efficiency in heterogeneous photocatalysis affect the generation of the oxidizing agents in different ways, which ought to be taken into consideration in further studies.

In general, suspensions of the photocatalyst are used in these procedures. Thus, at the end of these processes the semiconductor particles have to be separated from the cleaned solution. This operation can be taken out by immobilization of the catalyst, which was achieved by using poly(vinyl alcohol) in this work.

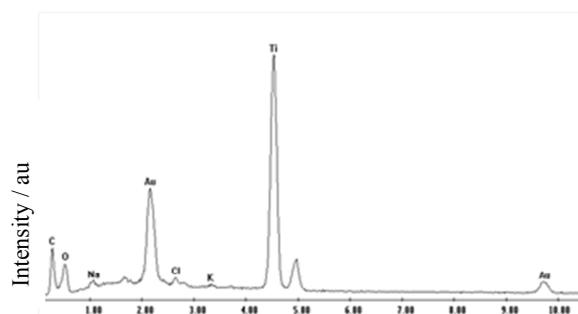


Figure 14. Energy dispersive X-Ray spectra of the TiO_2 island.

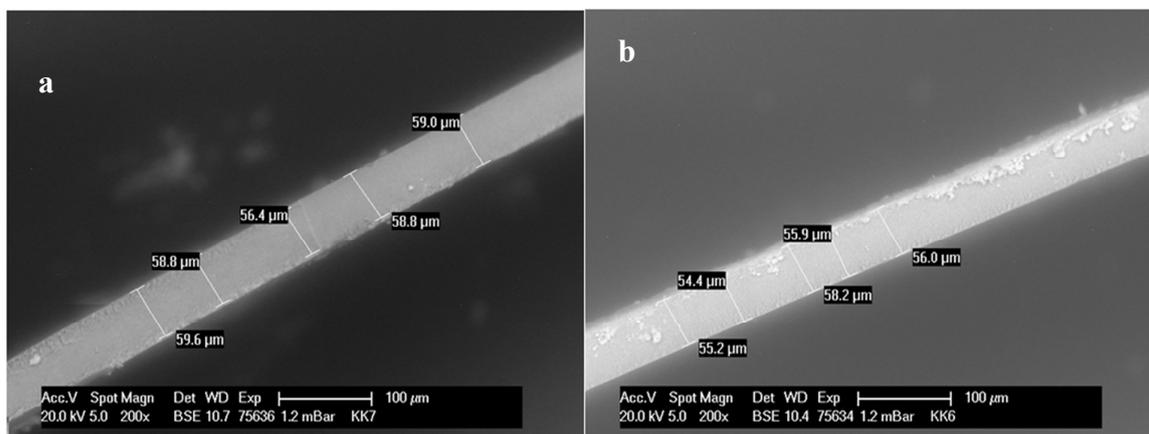


Figure 15. Scanning electron micrographs of cross sections of film. a) after 4 cycles of irradiation b) after 7 cycles of irradiation.

The stability of the PVA support proved to be too low for practical use. The stability of the PVA-based foils could be considerably improved by thermal treatment. However, before photocatalytic application the water-soluble products of the heating process had to be removed from the surface of the PVA-TiO₂ foil. According to the efficiency of •OH formation, three cycles of irradiation and rinsing proved to be sufficient for this purpose. The foil pre-treated by this method could be applied for the photocatalytic degradation of a frequently used non-ionic detergent, Triton X-100. When using the immobilized catalyst, it should be noted that the foil catalyst is degraded, its thickness decreases and the catalyst particles become more and more accessible on the surface.

Acknowledgements

This work was supported by Széchenyi 2020 under the projects GINOP-2.3.2-15-2016-00016 and EFOP-3.6.1-16-2016-00015.

REFERENCES

- [1] Dewil, R.; Mantzavinos, D.; Poullos, I.; Rodrigo, M. A.: New perspectives for Advanced Oxidation Processes, *J. Environ. Manage.*, 2017 **195** 93–99 DOI: 10.1016/j.jenvman.2017.04.010
- [2] Hack, N.; Reinwand, C.; Abbt-Braun, G.; Horn, H.; Frimmel, F. H.: Biodegradation of phenol, salicylic acid, benzenesulfonic acid, and iomeprol by *Pseudomonas fluorescens* in the capillary fringe, *J. Contam. Hydrol.*, 2015 **183** 40–54 DOI: 10.1016/j.jconhyd.2015.10.005
- [3] Mohan, P. K.; Nakhla, G.; Yanful, E. K.: Biokinetics of biodegradation of surfactants under aerobic, anoxic and anaerobic conditions, *Water Res.*, 2006 **40**(3) 533–540 DOI: 10.1016/j.watres.2005.11.030
- [4] Wyrwas, B.; Dymaczewski, Z.; Zgoła-Grześkowiak, A.; Szymański, A.; Frańska, M.; Kruszelnicka, I.; Ginter-Kramarczyk, D.; Cyplik, P.; Ławniczak, Ł.; Chrzanowski, Ł.: Biodegradation of Triton X-100 and its primary metabolites by a bacterial community isolated from activated sludge, *J. Environ. Manage.*, 2013 **128** 292–299 DOI: 10.1016/j.jenvman.2013.05.028
- [5] Szabó-Bárdos, E.; Markovics, O.; Horváth, O.; Törő, N.; Kiss, G.: Photocatalytic degradation of benzenesulfonate on colloidal titanium dioxide, *Water Res.*, 2011 **45**(4) 1617–1628 DOI: 10.1016/j.watres.2010.11.045
- [6] Zsilák, Z.; Szabó-Bárdos, E.; Fónagy, O.; Horváth, O.; Horváth, K.; Hajós, P.: Degradation of benzenesulfonate by heterogeneous photocatalysis combined with ozonation, *Catal. Today*, 2014 **230** 55–60 DOI: 10.1016/j.cattod.2013.10.039
- [7] Zsilák, Z.; Fónagy, O.; Szabó-Bárdos, E.; Horváth, O.; Horváth, K.; Hajós, P.: Degradation of industrial surfactants by photocatalysis combined with ozonation, *Environ. Sci. Pollut. Res.*, 2014 **21**(19) 11126–11134 DOI: 10.1007/s11356-014-2527-2
- [8] Hegedűs, P.; Szabó-Bárdos, E.; Horváth, O.; Horváth, K.; Hajós, P.: TiO₂-mediated photocatalytic mineralization of a non-ionic detergent: Comparison and combination with other advanced oxidation procedures, *Materials (Basel)*, 2015 **8**(1) 231–250 DOI: 10.3390/ma8010231
- [9] Rodríguez, E. M.; Márquez, G.; León, E. A.; Álvarez, P. M.; Amat, A. M.; Beltrán, F. J.: Mechanism considerations for photocatalytic oxidation, ozonation and photocatalytic ozonation of some pharmaceutical compounds in water, *J. Environ. Manage.*, 2013 **127** 114–124 DOI: 10.1016/j.jenvman.2013.04.024
- [10] Agustina, T. E.; Ang, H. M.; Vareek, V. K.: A review of synergistic effect of photocatalysis and ozonation on wastewater treatment, *J. Photochem. Photobiol. C Photochem. Rev.*, 2005 **6**(4) 264–273 DOI: 10.1016/j.jphotochemrev.2005.12.003

- [11] Szabó-Bárdos, E.; Czili, H.; Horváth, A.: Photocatalytic oxidation of oxalic acid enhanced by silver deposition on a TiO₂ surface, *J. Photochem. Photobiol. A.*, 2003 **154** 195–201 DOI: 10.1016/S1010-6030(02)00330-1
- [12] Ma, J.; Guo, X.; Zhang, Y.; Ge, H.: Catalytic performance of TiO₂@Ag composites prepared by modified photodeposition method, *Chem. Eng. J.*, 2014 **258** 247–253 DOI: 10.1016/j.cej.2014.06.120
- [13] Chong, M. N.; Jin, B.; Chow, C. W. K.; Saint, C.: Recent developments in photocatalytic water treatment technology: A review, *Water Res.*, 2010 **44**(10) 2997–3027 DOI: 10.1016/j.watres.2010.02.039
- [14] Gar Alalm, M.; Tawfik, A.; Ookawara, S.: Enhancement of photocatalytic activity of TiO₂ by immobilization on activated carbon for degradation of pharmaceuticals, *J. Environ. Chem. Eng.*, 2016 **4**(2) 1929–1937 DOI: 10.1016/j.jece.2016.03.023
- [15] Bouarioua, A.; Zerdaoui, M.: Photocatalytic activities of TiO₂ layers immobilized on glass substrates by dip-coating technique toward the decolorization of methyl orange as a model organic pollutant, *J. Environ. Chem. Eng.*, 2017 **5**(2) 1565–1574 DOI: 10.1016/j.jece.2017.02.025
- [16] Veréb, G.; Ambrus, Z.; Pap, Z.; Mogyorósi, K.; Dombi, A.; Hernádi, K.: Immobilization of crystallized photocatalysts on ceramic paper by titanium(IV) ethoxide and photocatalytic decomposition of phenol, *React. Kinet. Mech. Catal.*, 2014 **113**(1) 293–303 DOI: 10.1007/s11144-014-0734-y
- [17] Kondrakov, A. O.; Ignatev, A. N.; Lunin, V. V.; Frimmel, F. H.; Bräse, S.; Horn, H.: Roles of water and dissolved oxygen in photocatalytic generation of free OH radicals in aqueous TiO₂ suspensions: An isotope labeling study, *Appl. Catal. B Environ.*, 2016 **182** 424–430 DOI:10.1016/j.apcatb.2015.09.038
- [18] Czili, H.; Horváth, A.: Applicability of coumarin for detecting and measuring hydroxyl radicals generated by photoexcitation of TiO₂ nanoparticles, *Appl. Catal. B Environ.*, 2008 **81**(3–4) 295–302 DOI:10.1016/j.apcatb.2008.01.001
- [19] Ishibashi, K. I.; Fujishima, A.; Watanabe, T.; Hashimoto, K.: Detection of active oxidative species in TiO₂ photocatalysis using the fluorescence technique, *Electrochem. commun.*, 2000 **2**(3) 207–210 DOI:10.1016/S1388-2481(00)00006-0
- [20] Hegedűs, P.; Szabó-Bárdos, E.; Horváth, O.; Szabó, P.; Horváth, K.: Investigation of a TiO₂ photocatalyst immobilized with poly(vinyl alcohol), *Catal. Today.*, 2017 **284** 179–186 DOI:10.1016/j.cattod.2016.11.050
- [21] Lei, P.; Wang, F.; Gao, X.; Ding, Y.; Zhang, S.; Zhao, J.; Liu, S.; Yang, M.: Immobilization of TiO₂ nanoparticles in polymeric substrates by chemical bonding for multi-cycle photodegradation of organic pollutants, *J. Hazard. Mater.*, 2012 **227–228** 185–194 DOI:10.1016/j.jhazmat.2012.05.029
- [22] Song, Y.; Zhang, J.; Yang, H.; Xu, S.; Jiang, L.; Dan, Y.: Preparation and visible light-induced photo-catalytic activity of H-PVA/TiO₂ composite loaded on glass via sol-gel method, *Appl. Surf. Sci.*, 2014 **292** 978–985 DOI: 10.1016/j.apsusc.2013.12.090

ADAPTING THE SDEWES INDEX TO TWO HUNGARIAN CITIES

VIKTOR SEBESTYÉN, VIOLA SOMOGYI,* SZANDRA SZŐKE, AND ANETT UTASI

Institute of Environmental Engineering, University of Pannonia, 10 Egyetem str., Veszprém,
H-8200, HUNGARY

Numerous cities aim to mitigate their contribution to climate change and provide a liveable environment in the context of sustainable development. In order to measure these efforts, benchmarking performance would be a good solution. Methods for environmental analysis have their limitations when it comes to evaluating a city and other aggregated indicators focus on certain aspects of a sustainable or liveable settlement. The SDEWES Index was used for benchmarking several cities of different sizes in terms of metrics related to energy, water and environmental systems successfully thus it was chosen to compare the performance of Veszprém and Zalaegerszeg, two environmentally conscious Hungarian county seats of roughly the same size and population. The SDEWES Index consists of 7 dimensions, namely energy consumption, industrial profile with CO₂ emissions, CO₂-saving measures, R&D, renewable energy potential and utilization, water and environmental quality, and social environment and sustainability policy. Each dimension is composed of 5 indicators that provide information on sustainable development of energy, water and environmental systems in cities. Using the SDEWES Index the strengths and weaknesses of the two cities are highlighted, locating those key parameters where improvement can be achieved. Both for Veszprém and Zalaegerszeg progress could be realized concerning energy-saving measures and the proportion of green areas could be increased. To improve the method and facilitate a more comprehensive comparison of cities of differing sizes, data should be provided concerning the territory or population. Also, the definition and inclusion of a worst and best case scenario that takes into account the parameters would be advantageous in terms of a comparison. These were named 'horror' and SDEWES cities by the authors, respectively.

Keywords: SDEWES Index, Sustainability, City development, Sustainable Energy Action Plan, City sample

1. Introduction

Sustainability is a key issue when it comes to the development of cities. In 2014, 54% of the world's population lived in urban areas and according to the prognosis, the proportion will be as high as 66% by 2050 [1]. While the number and population of megacities is on the rise, 43% of urban dwellers lived in settlements consisting of less than 300,000 inhabitants in 2014 (in Europe the corresponding data was 58%) and only a modest decrease is estimated by 2030 [2].

Looking at these numbers, it is easy to notice that urbanised areas have a huge impact on achieving sustainable development. Various cities have started to address this issue and several indicators and comparisons were created to measure specific aspects of sustainable development. A few of these numerous examples are listed in this paper.

The City Development Index [3] studies the municipalities from social and governance aspects. The Global Power City Index (GPCI) has ranked 40 metropolises since 2008 [4] considering 70 individual indicators regarding the environment, liveability and

R&D among others. The Green City Index is focused on the environmental sustainability of large cities [5]. Carbon footprints of twelve metropolitan areas [6] and the San Francisco Bay Area [7] were determined. Several others are listed by López-Ruiz et al. [8] but usually smaller towns do not fall within the scope of these benchmarks.

Environmental analysis, on the other hand, facilitates the evaluation of the impacts of human activities (different actions, projects or investments) with regard to the local environment, economy and society. In this way it provides information on the status quo and helps the practical implementation of sustainable development by focusing attention on the points to be improved [9]. Several methods have been developed to carry out the procedure: checklists [10], the matrix technique [11], the network approach [12], GIS-based methods [13] and quantitative methods [14] may be used to evaluate environmental impacts. Aggregating methods such as the Global Pollution Index (I_{PG}) [15] may be suitable up to a point in providing a comprehensive sustainability analysis as they are only based on immission values. Several multiple-criteria decision-making (MCDM) techniques designed to assist with decision-making, e.g. the Analytic Hierarchy Process (AHP) [16] or the Technique for Order Preference by Similarity to Ideal Solution systems (TOPSIS) [17] combined with Simple

*Correspondence: somogyiv@uni-pannon.hu

Additive Weighting (SAW) [18], can be used as well [19]. These hierarchical methods (TOPSIS, SAW and AHP) rank the examined parameters which may be useful when deciding between options of individual investments but are problematic in terms of adapting them to the decision-making process with regard to development strategies of the settlements. The Sustainable Development of Energy, Water and Environment Systems (SDEWES) City Sustainability Index was developed to overcome the disadvantages and limitations of other measures with regard to benchmarking the performance of cities in terms of energy, water and environment systems. So far a list of 58 cities assessed by the SDEWES Index can be accessed online on the SDEWES Centre homepage [20]. Also, articles concerning the benchmarking of 12 South East European cities (such as Athens and Belgrade) [21], 22 Mediterranean port cities (e.g. Barcelona and Venice) [22], and a further 18 South East European cities (including Budapest and Pécs) [23] were published, and the inventory will no doubt be expanded upon in the near future.

In this paper two cities were evaluated by using the composite SDEWES Index and to test the method itself. Veszprém and Zalaegerszeg are two Hungarian county seats of roughly the same size and population, both are environmentally conscious, and are aiming to become environmentally friendly, liveable and sustainable cities.

Veszprém is near Lake Balaton with a population of around sixty-two thousand people and a surface area of 126.9 km². The city has won the Climate Star award[†] [24] and aims to become an eco-city. In its Energy Strategy [25] the following objectives were set by 2026:

- 20% of the energy demand should be satisfied by renewable energy resources while the energy renovation of public and residential buildings should be 70% complete resulting in a reduction in greenhouse gas (GHG) emissions of 25%;
- 35% of transportation has to be conducted by means of public transport with environmentally friendly vehicles that are less than 10 years old and 10% of the vehicle-kilometres should be undertaken by bicycles;
- increasing the proportion of green areas to 25 m²/capita and 60% of rainwater should be reused in some way.

Zalaegerszeg is situated in the west of the country and consists of sixty-two thousand inhabitants and a surface area of 100 km². Since the millennium its urban development strategies have focused on becoming a "Sustainable City". In the strategy formulated in 2014

[26] a major goal was to improve energy efficiency by 20% while producing more than 20% of its energy using local renewable resources by 2030. This would result in a reduction of 36% in terms of energy-related costs. Taking 2005 as a base the GHG emissions should be reduced by 20% by 2030 while the particle pollution PM₁₀ is planned to be mitigated by 10% by 2023. Besides achieving these indicators, the city council aims to create and strengthen its image of being an environmentally conscious, modern and sustainable green city.

2. Experimental

The SDEWES Index consists of 7 dimensions and 35 main indicators (*Table 1*). The indicators of each dimension are explained in detail in Ref. [21]; only those that need further clarification or some sort of adjustment due to problems concerning the access of data are highlighted in this paper.

The data for each indicator were normalised according to the Min-Max method [27]. Depending on whether the lower or higher values are more desirable, either Eq. (1) or (2) is used [21-22]. An example of the first case, i.e. when lower values are favourable, would be CO₂ emissions, while the normalised data for the number of local universities would be calculated by Eq. (2). Since the leader ($i_{x,y}(C_j) = \max(i_{x,y})$) equals 1 and the laggard ($i_{x,y}(C_j) = \min(i_{x,y})$) 0, if only two cities are compared and the values are identical, the denominator would become 0. To avoid this $I_{x,y}(C_j)$ should be set to 0 for such cases. Another solution would be to include further cities in the benchmark.

$$I_{x,y}(C_j) = \frac{(i_{x,y}(C_j) - \max(i_{x,y}))}{(\min(i_{x,y}) - \max(i_{x,y}))} \quad (1)$$

$$I_{x,y}(C_j) = \frac{(i_{x,y}(C_j) - \min(i_{x,y}))}{(\max(i_{x,y}) - \min(i_{x,y}))} \quad (2)$$

where:

- I – normalised value of the indicator,
- x – dimension number,
- y – indicator number within a dimension,
- C_j – j^{th} city,
- i – data input before normalization.

Value aggregation is done according to

$$\text{SDEWES}(C_j) = \sum_{x=1}^7 \sum_{y=1}^5 \alpha_x I_{x,y}(C_j), \quad (3)$$

where $\sum_{x=1}^7 \alpha_x = 1$ and α_x is the weight of the x^{th} dimension. The SDEWES Index of the j^{th} city is calculated by a double summation, where α_1 and α_5 are 0.22 since these dimensions include energy and CO₂ emissions data. Other dimensions are weighed less ($\alpha_x=0.11$) as they are not directly related to the sustainable energy action plan [21-22].

[†] The Climate Star award was founded by the Climate Alliance with the aim of demonstrating how climate protection initiatives can be implemented from the grass roots up [24]. Cities with initiatives in the fields of sustainable energy, mobility, consumption, urban and regional development, and citizen involvement may apply for the call in four categories.

Table 1. The Dimensions and Indicators [21-23].

		Dimensions						
		D1:	D2:	D3:	D4:	D5:	D6:	D7:
		Energy Consumption and Climate	Penetration of Energy and CO ₂ -Saving Measures	Renewable Energy Potential and Utilization	Water and Environmental Quality	CO ₂ Emissions and Industrial Profile	City Planning and Social Welfare	R&D, Innovation and Sustainability Policy
Indicators	Energy consumption of buildings [MWh]	Sustainable Energy Action Plan (SEAP)	Solar energy potential [Wh/m ² /day]	Domestic water consumption [m ³ /capita]	CO ₂ emission of buildings [t CO ₂]	Price of a public transport ticket [EUR]	R&D and innovation policy orientation	
	Energy consumption of transport [MWh]	Combined heat and power-based district (H/C)	Wind energy potential [m/s]	Water quality index [/100]	CO ₂ emissions of transport [t CO ₂]	Urban form and protected sites	National patents in clean technologies	
	Total energy consumption per capita [MWh/capita]	Energy savings in end-usage (buildings)	Geothermal energy potential [mW/m ²]	Annual mean PM ₁₀ concentration [μg/m ³]	Average CO ₂ intensity [t CO ₂ /MWh]	GDP per capita [PPP\$ national]	Local public/private universities	
	Heating Degree Days (HDD) [day °C]	Density of the public transport network	Renewable energy usage for electricity [%]	Ecological footprint [gha/capita]	Number of CO ₂ -intense industries	Inequality adjusted well-being (HPI)	National <i>h</i> -index of scientific publications	
	Cooling Degree Days (CDD) [day °C]	Efficient public-lighting armatures	Biofuel share in transport [%]	Biocapacity [gha/capita]	Airport Carbon Accreditation (levels)	Tertiary education rate (national)	Reduction Target for CO ₂ emission reduction (2020)	

2.1. Data of Veszprém and Zalaegerszeg

As of October 2017, neither of the cities are signatories of the Covenant of Mayors movement, therefore, alternative sources of data had to be found. Besides the sources suggested by the developer of the index, the energy and integrated city development strategies were used to retrieve data. Necessary changes and the simplification of the original method is explained in detail below.

2.1.1. Energy Consumption and Climate (D1):

The energy consumption of buildings (municipal, residential and commercial) and transportation (public, private and the municipal vehicle fleet) are indicators on their own (Table 2) but also included in terms of the total energy use per capita [21]:

$$\frac{E}{P(C_j)} = \frac{(\sum_{b=1}^3 E_b + \sum_{t=1}^3 E_t + E_g + E_d)}{P(C_j)} \quad (4)$$

where

- E – total energy consumption (MWh),

- $P(C_j)$ – population of the j^{th} city (capita),
- E_b – energy consumption of buildings (1: municipal, 2: residential and 3: commercial) (MWh),
- E_t – energy consumption of transport (1: public, 2: private and 3: municipal vehicle fleet) (MWh),
- E_g – energy consumption of public lighting (MWh),
- E_d – energy consumption of industry (MWh).

The energy consumption of transport was calculated based on the number of vehicles registered according to the energy strategies of the cities [25-26] by presuming an average mileage of 15,000 km/year and average consumption of 7.5 l/100 km [25]. The energy content of diesel and gasoline was assumed to be 10.83 kWh/l and 8.89 kWh/l, respectively. Data for commercial buildings are only included in the total energy consumption indicator as the consumption of the service sector and industries was not collected separately by the cities. As for the municipal vehicle fleet, due to a lack of data for Veszprém, this had to be neglected for both towns. It has to be noted that the data for Veszprém were from between 2007 and 2009 as stated in the strategy while for Zalaegerszeg information

Table 2. The data of Energy Consumption and Climate (D1)

Indicator	Veszprém	Zalaegerszeg
Energy consumption of buildings [MWh]	309,393 [25]	313,434 [26]
Energy consumption of transport (MWh)	225,479 [25]	217,228 [26]
Total energy consumption [MWh/capita] (in brackets: population)	19.50 (61,721) [25]	11.65 (59,499) [26]
Number of Heating Degree Days (HDD)	2,890 [28]	2,850 [28]
Number of Cooling Degree Days (CDD)	1,619 [28]	1,607 [28]

could only be obtained from between 2012 and 2013. In later documents, only improvements are mentioned, newly obtained data on overall consumption is not stated. Monitoring the energy consumption of public buildings and lighting is an issue for both cities that needs to be solved.

2.1.2. Penetration of Energy and CO₂-saving Measures (D2):

Neither of the cities have a Sustainable Energy Action Plan (SEAP) as of 2017 [31], therefore, both received zero for the first indicator (Table 3) though Veszprém is currently in the process of creating its Sustainable Energy and Climate Action Plan (SECAP). In the case of Veszprém, a cogeneration plant was recently installed [29] while there are only plans for such a system in Zalaegerszeg (though one district heating system operates using geothermal energy) [30].

Energy savings have been accomplished and are continuously implemented in both cities by renovating public buildings and installing photovoltaic systems, e.g. on the flat roofs of a grammar and primary school in Veszprém and on the Mayor's office in Zalaegerszeg (the performance of which can be accessed online from the webpage of the city). Nonetheless, there is no building with net zero CO₂ emissions, that is why both cities received 1 point for the 'energy savings in end-usage' indicator. The difference in size of the cities does not necessitate different types of public transport; both Zalaegerszeg and Veszprém have local bus routes that are operated by the same regional bus company. Further points could have been allocated for tram and subway lines (2 for existing, 1 for planned) and an extra point would have been given to the city with the longest tram/subway network [20]. LED technology is considered an efficient public lighting solution (1 point) and an

Table 3. The data of Penetration of Energy and CO₂-Saving Measures (D2)

Indicator	Veszprém	Zalaegerszeg
Sustainable Energy Action Plan (SEAP)	0 [31]	0 [31]
Combined heat and power-based district heating/cooling system	2 [25]	1 [26]
Energy savings in end-usage (buildings)	1 [25]	1 [26]
Density of public transport network	1 [29]	1 [30]
Efficient public lighting armatures	2	1

additional point can be gained if solar energy is used to power armatures. Recent investments were made in both cities to improve the energy efficiency of public lighting after the introduction of the cited strategies.

2.1.3. Renewable Energy Potential and Utilization (D3):

The renewable energy potential is highly dependent on the location, topology and geology of the area but the local government can have a strong influence on the utilisation of these resources. While regional data could be gathered for the potentials, national data [32] had to be used for the share of renewable sources in terms of electricity production and biofuel use in transport indicators because there was no reliable local information for Veszprém (Table 4). For Zalaegerszeg biogas from the regional municipal wastewater treatment plant is converted to provide the local buses with liquid fuel. Based on a presentation [33] the tanked volume is known for 2015, therefore, the biofuel utilization in terms of transportation was modified accordingly. To attain an accurate comparison the national value for 2015 was considered in the case of Veszprém.

Table 4. The data of the Renewable Energy Potential and Utilization (D3)

Indicator	Veszprém	Zalaegerszeg
Solar energy potential [Wh/m ² /day]	3,425 [25]	3,014 [26]
Wind energy potential [m/s]	4.921 [34]	3.505 [34]
Geothermal energy potential [mW/m ²]	60 [35]	90 [35]
Renewable energy usage for electricity [%]	8.76 [32]	8.76 [32]
Biofuel utilization in terms of transport [%]	4.15 [32]	5.46 [32-33]

Table 5. The data of Water and Environmental Quality (D4).

Indicator	Veszprém	Zalaegerszeg
Domestic water consumption [m ³ /capita]	7 [38]	7 [38]
Water quality index [/100]	92 [37]	92 [37]
Average air quality PM ₁₀ [µg/m ³]	23.59 [39]	29.60 [39]
Ecological footprint [gha]	2.9 [40]	2.9 [40]
Biocapacity [gha]	2 [40]	2 [40]

2.1.4. Water and Environmental Quality (D4):

There were no available local data for the domestic blue water footprint, ecological footprint and biocapacity, therefore, national values were applied in the calculation. Air quality is only described in terms of the PM₁₀ concentration [20]. The water quality index was ambiguous as the articles [21-22] referred to the indicator as drinking water quality but the Water Quality Index (WATQI) refers to natural water quality [36]. The index relies on the global database of the United Nations GEMS/Water Programme and includes five indicative parameters: dissolved oxygen, pH, conductivity, total nitrogen and total phosphorous. Unfortunately the WATQI of countries are only available for 2008 [37], from 2012 the water quality index was replaced with access to sanitation and drinking water in terms of the aggregated Environmental Performance. It has to be noted that in the case of Hungary the drinking water is supplied from underground reservoirs that are only linked indirectly to surface waterbodies while in other countries these serve as direct sources of drinking water. Thus using water quality indices for inland water bodies may be good indicators of safe access to water.

2.1.5. Emissions and Industrial Profile (D5):

As in the case of the first dimension the emission values of the commercial buildings and municipal vehicle fleet were unavailable, therefore, these could not be included in the calculation. Information on CO₂-intense industries was gathered by going through an online company database [41]. The indicator carbon accreditation of airports became zero for both cities for different reasons: Veszprém has no airport and the one near Zalaegerszeg has no accreditation. Since the first case means no emissions while in the second case the existing emissions are not measured, the purpose of the indicator is not fully achieved. The original aim was to include the emissions of the airports in some way in the SDEWES Index as the SEAPs do not take them into consideration [21].

Table 6. The data of CO₂ Emissions and Industrial Profile (D5)

Indicator	Veszprém	Zalaegerszeg
CO ₂ emissions of buildings [t CO ₂]	87,882 [25]	82,579 [26]
CO ₂ emissions of transport [t CO ₂]	71,711 [25]	70,282 [26]
Average CO ₂ emissions [t CO ₂ /MWh]	0.393	0.292
Number of CO ₂ intense industries	4 [41]	4 [41]
Carbon Accreditation of Airport [levels]	0	0 [42]

2.1.6. City Planning and Social Welfare (D6):

Two indicators need further explanation (*Table 7*). The prices of public transport were introduced in Ref. [21] instead of the share of public transport in terms of total passenger kilometres [22], the latter not being accessible in all cases. The more a single ticket costs, the less likely people will choose public transport. On the other hand, easy access to public transport should result in positive externalities such as cleaner air and less traffic jams, by and large a more liveable city. The urban form and protected sites indicator is an aggregation of several factors (*Table 8*): compact city form (whether it is mono- or polycentric), urban green areas and surrounding green corridors are evaluated.

To determine the compactness of the cities, the energy consumption of transport compared against population density was chosen, as a compact city can be described as of high population density [47] and because of the short distances cars are less likely to be used. Thus the smallest value received 3 points while the highest received 1. Urban green spaces were examined in

Table 7. The data of City Planning and Social Welfare (D6)

Indicator	Veszprém	Zalaegerszeg
Price of public transport ticket [EUR] (1 EUR = 310 HUF)	1.07 [43]	1.10 [43]
Urban form and protected sites	1	2
GDP per capita [PPP\$ national]	25,068.9 [44]	25,068.9 [44]
Inequality adjusted well-being (HPI)	4.3 [45]	4.3 [45]
Tertiary education rate (national) [%]	21 [46]	21 [46]

Table 8. Data for Grading Urban Form and Municipal Management

	Veszprém	Zalaegerszeg
Urban form and protected sites	1	2
Compact city form	1	3
monocentric	x	x
polycentric		
population density [capita/km ²]	486.38	580.99
Urban green spaces	1	2
urban park intensity [m ² /capita]	21 [29]	34.6 [30]
Green corridors	1	1
protected sites	x	x
national park/Ramsar	x	x

comparison with the Hungarian county seats [48]: 0-30 m²/capita: 1 point, 30-50 m²/capita: 2 points and over 50 m²/capita: 3 points. Green corridors were also assessed on a county basis [49] instead of using the suggested GIS-based method [20]. The categories were determined from 1-3 by only taking the green corridor areas of Hungarian counties into consideration.

2.1.7. R&D, Innovation and Sustainability Policy (D7):

Results for the seventh dimension are listed in Table 9. The number of public and private universities yielded an unexpected result for the two cities in question. Universities seated in the town and those where only a faculty is based there were equally counted. If only those universities that are seated in the said city were taken into consideration, Veszprém would have 2 versus 0 in Zalaegerszeg. Additional points were given if the university was listed in the Scimago Institutions Rankings [53]. The energy strategy of Veszprém envisions a 25% CO₂ emissions reduction by 2026, the basis being 2007 [25], while Zalaegerszeg aims to achieve a 36% reduction by 2030, compared to 2012 [26]. To facilitate a comparison, goals for 2020 were calculated by linear interpolation.

3. Results

After processing the necessary calculations, the SDEWES Indices of both Veszprém and Zalaegerszeg were 1.54. As is clear in Fig. 1, the values are integers and, except for one case (D5), Veszprém achieved better or equal results. It also has to be noted that on several occasions the difference between the data was very

Table 9. The data of the R&D, Innovation and Sustainability Policy (D7)

Indicator	Veszprém	Zalaegerszeg
R&D and innovation policy orientation	3 [50]	2 [50]
National patents in clean technologies	2.5 [51]	2.5 [51]
Number of public/private universities (city)	3 [52-53]	5 [52-53]
<i>h</i> -index of scientific publications	301 [54]	301 [54]
Reduction Target for CO ₂ Emissions (2020) [%]	18 [25]	16 [26]

small. Still, the better city was awarded with 1 and the worse value with zero in the normalisation process.

In order to eliminate this problem, a third city was included in the benchmark. Ohrid was chosen as the size of this historical Macedonian town is similar to the other two and all data were available from Kilkis [21]. Also, this was the only city in this comparison that had no SEAP. As an alternative solution the 'average South East European (SEE) city' from the same article was included to put the two Hungarian cities to the test to see where they would be in the ranking of the SEE cities of that sample.

The inclusion of these two examples changed the order of the cities (Table 10). While the results of both Veszprém and Zalaegerszeg improved, Zalaegerszeg gained more from the inclusion of another city from a different country in the benchmark. The reason for the improvement of the indices is that in several cases the national data had to be included in the calculation and hence, the indicator became 0.

While both Veszprém and Zalaegerszeg gained points from increasing the sample size, the accumulated increase was larger for Zalaegerszeg (10.03 compared to 8.12). The arrows show in which direction the indicators changed. In the case of Veszprém, data for D3 and D6 decreased but not significantly, for Zalaegerszeg there was no change in D3 and only a slight decline in D5. The SDEWES Index results are similar, the difference between the highest and lowest values is 0.37. Nonetheless, heterogeneity exists with regard to the individual indicators. Results for each dimension are visualised in Fig. 2.

Both Hungarian towns performed well concerning D1 (energy consumption and climate), D4 (water and environmental quality) and D7 (R&D, innovation and sustainability policy), while there is room for improvement in the fields of CO₂-saving measures and city planning. The average SEE city, on the other hand, possesses lower values regarding energy consumption and environmental and water quality.

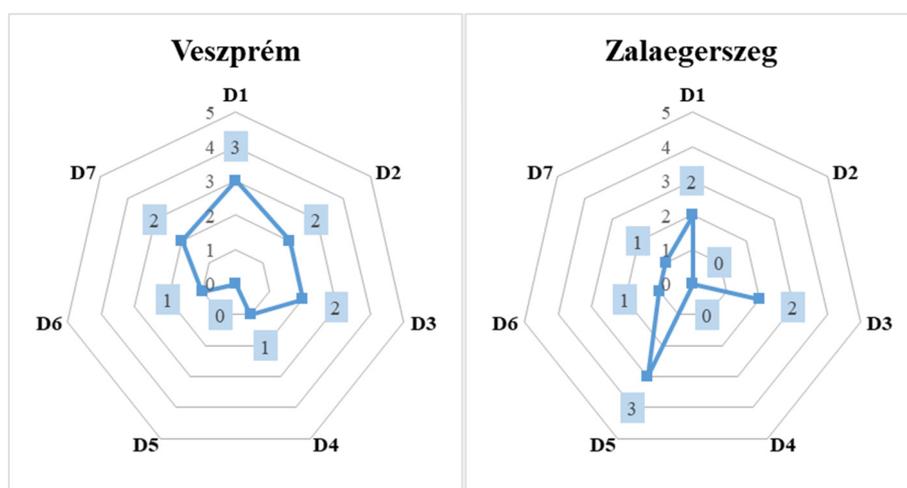


Figure 1. Results of the first comparison on radar charts.

4. Conclusion

The process of gathering data revealed that both Veszprém and Zalaegerszeg need to collect and measure data related to energy efficiency and other indicators of sustainable development more precisely. Creating a database of detailed information on energy use, CO₂ emissions and use of renewable sources which is regularly updated would help to achieve the ambitious goal of becoming a sustainable city within a relatively short timespan. Also, the development of a SEAP or SECAP and becoming a member of the Covenant of Mayors would be advantageous and for which Veszprém has started taking steps.

Based on the dimensions of the SDEWES Index, Veszprém needs to improve in terms of D3 and D6. The individual indicators highlight that the energy consumption of public transport could be reduced, based on the example of Zalaegerszeg, and also utilization of renewable energy should be improved. In

terms of city planning and social welfare the number and area of urban parks can be increased more easily than that of protected sites. Establishing green areas is included within the urban development strategy of the city [29], so improvements may be expected in terms of this indicator. Also, progress in developing a compact city form is anticipated based on the plans to reform the public transport system and relocate the central bus station to next to the railway station [29].

In the case of Zalaegerszeg, dimensions D2, D3 and D6 are lower. Energy-saving measures could be improved by constructing a cogeneration plant to improve the penetration of district heating, using solar panels in public lighting, and also using the wastewater heat to facilitate the full utilization of biogas as a liquid biofuel [55]. In the case of D3, the potential of solar and wind energy cannot be increased and no information concerning the local use of renewable energy resources in terms of electricity was found. As for D6 the same suggestions as in the case of Veszprém can be made to increase the values of the individual indicators. It has to

Table 10. Results of calculating the SDEWES Index

	Veszprém	Zalaegerszeg	Ohrid	Average SEE city
D1	3.53 ↑	3.93 ↑	3.78	1.43
D2	2.45 ↑	0.95 ↑	2.00	3.50
D3	1.99 ↓	2.00 -	1.60	3.12
D4	4.17 ↑	3.43 ↑	2.45	1.16
D5	2.28 ↑	2.79 ↓	3.00	2.45
D6	0.98 ↓	1.95 ↑	4.00	3.65
D7	3.72 ↑	3.97 ↑	1.08	1.75
SDEWES	2.74 ↑	2.83 ↑	2.71	2.47

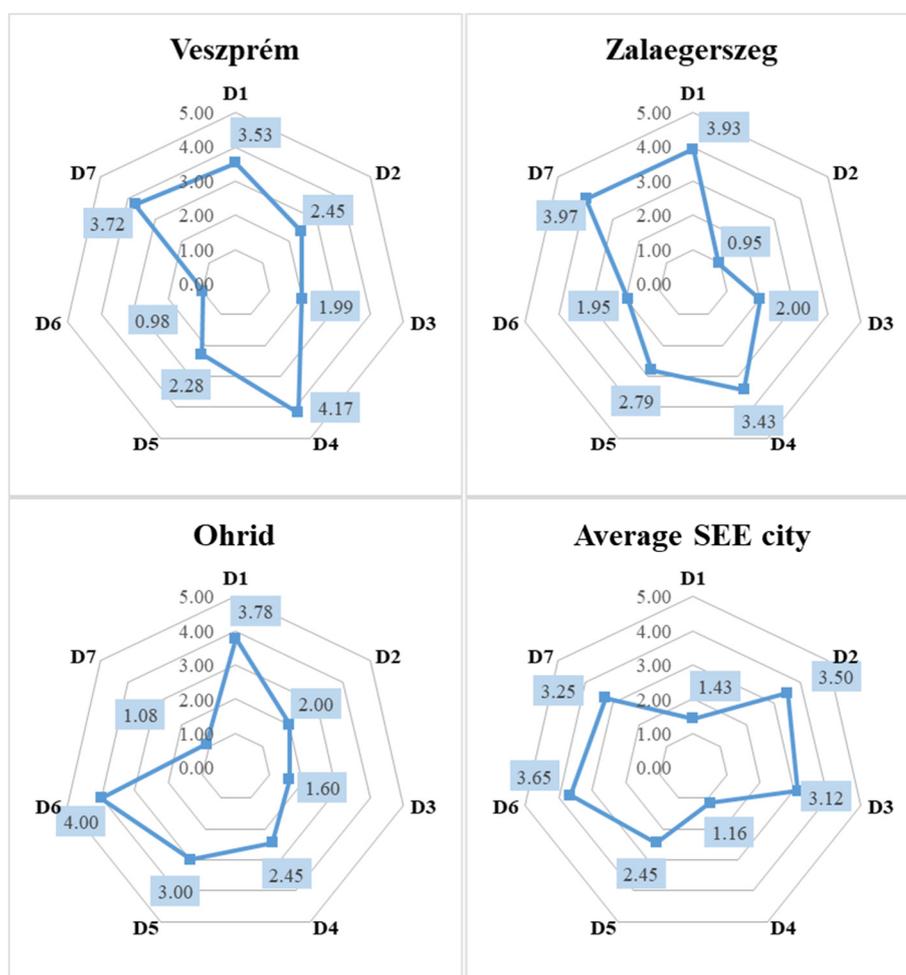


Figure 2. Results of the second comparison on radar charts.

be pointed out though that the use of the ticket price for public transport resulted in an unjust outcome: since in Ohrid there is no means of local public transport, this indicator became zero, which could also mean that transportation is free. Therefore, Ohrid will always receive the highest value in terms of the process of normalisation as long as there are no local buses in the city. In terms of the process of evaluating the two Hungarian county seats, the benchmarking method was assessed as well. Without a doubt, the SDEWES Index has its benefits. It uses environmental, economic and social indicators, gives credit to CO₂ reduction goals and also considers the possible use of renewable resources. Also, human resources are included

presuming that higher education and research and development seek to achieve sustainability.

On the other hand, the authors identified some drawbacks, too. The first and the fifth dimensions both focus on energy consumption and CO₂ emissions. Since these data strongly correlate with each other, the inclusion of both measures leads to redundancy. Also, these two dimensions are weighed more than the others, therefore, energy-related information outweighs other aspects of sustainability. Furthermore, some of the parameters favour smaller cities over larger ones and vice versa. For example if the absolute values of energy consumption of a small city and a capital are compared, the small city will undoubtedly achieve a better result.

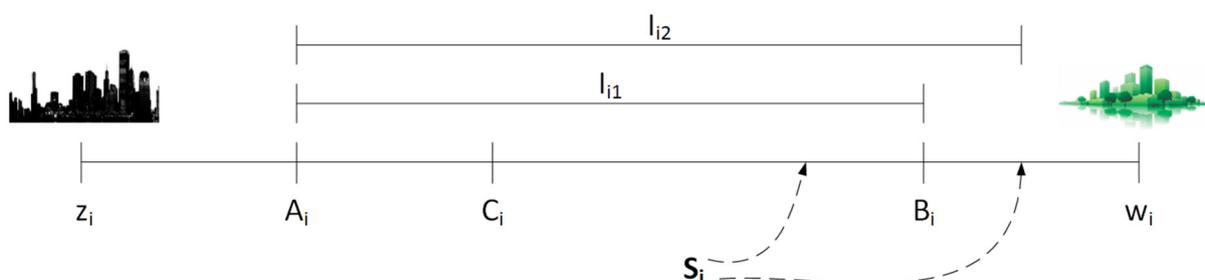


Figure 3. The evaluation intervals of SDEWES Index indicators

An example of the opposite would be a city with an accredited airport (ACA 3) as opposed to a town with no airport (small towns do not always have airports). Similarly in a capital, where subway and tram lines are at one's disposal, the density of the public transport network would be high while it would be uneconomical to have trams in a smaller town where the bus lines are sufficient. To overcome the problem of favouring results of cities of different sizes, using data which is proportional to area or population is suggested wherever possible.

Another problem was that the scoring of the qualitative indicators such as the urban form and protected sites was not always clear. If the SDEWES Index is to be used widely then these calculations have to be made transparently and be well documented.

Due to the nature of the Min-Max method, small differences may be magnified and large differences may diminish. Also, as the index requires certain data that can only be obtained on a national level, the comparison between cities in the same country is somewhat limited. A solution to this problem may be to include towns from different countries and to choose a range (by including more than two cities) in such a way that provides balanced scales in terms of the indicators.

While normalisation facilitates the inclusion of values on different scales, the Min-Max method makes it difficult to compare the results of two sets of cities. At present, the extremes are defined by the individual parameters of the cities chosen to be included in the benchmark (Fig. 3). In terms of another comparison with a different batch of cities (that can have a common set as with the previous version) the two evaluation intervals may not be equal ($I_{i1} \neq I_{i2}$ since $S_i \in I_{i2}$ and $S_i \notin I_{i1}$). As the two extremities are different in the two benchmarks, the results cannot be compared to each other. In the case of extremities a change in the order might appear as was the case in this paper.

Including the average of a different batch of cities (given that the average of any parameter is not equal to that of either city) may resolve the limitations of the Min-Max method but only momentarily. Since the ranking is dynamic and changes as the cities develop, the average SDEWES Index of a previous time period will not provide relevant information with regard to the current situation concerning the sample from which the average city was created and neither on the sample of two cities one wanted to expand.

Besides the obvious solution of having a sample size of at least three cities, the authors suggest the following: two artificial sets of parameters should be created to serve as absolute extremes of the SDEWES Index. The worst case scenario is referred to as the 'horror city' and the best case scenario is named 'SDEWES city' after the Index itself.

The legend for Fig. 3 is as follows:

- A_i, B_i, C_i – The measured/real indicator values of the analysed cities in the first calculation.
- S_i – The measured/real indicator value of the city to be included in the second calculation.

- I_{i1} – The evaluation interval, when the new city's value is between the other indicator values.
- I_{i2} – The evaluation interval, when the new city's value falls outside of the other indicator values.
- z_i – The theoretical minimum value of indicator i , the 'horror city'.
- w_i – The theoretical maximum value of indicator i , the 'SDEWES city'.

To resolve this problem with regard to the evaluation intervals changing from time to time, the minimum and maximum values of each indicator must be determined in a way that the examined cities could be included in the evaluation intervals:

$$I_{i1} = I_{i2} = I_{ij} \quad (5)$$

$$S_i \in I_{i2}, S_i \in I_{i1} \dots S_i \in I_{ij} \quad (6)$$

$$z_i \leq S_i \leq w_i \quad (7)$$

Defining these utopian and negative examples requires careful examination of the indicators. Some parameters are dependent on the geographical location while others need to follow a realistic optimal and unfavourable alternative, for example, the tertiary education rate may be zero in the worst case scenario but it is arguable whether 100% would be favourable from the viewpoint of urban management. Further studies are needed to define the 'horror' and 'SDEWES' cities of the SDEWES benchmarking method.

REFERENCES

- [1] United Nations Department of Economic and Social Affairs Population Division: *World Urbanization Prospects: The 2014 Revision, Highlights*. (United Nations, New York), 2014 ISBN 978-92-1-151517-6
- [2] United Nations Department of Economic and Social Affairs, Population Division: *World Urbanization Prospects: The 2014 Revision*. (United Nations, New York), 2015
- [3] Global Urban Observatory: *Global Urban Indicators Database*. United Nations Human Settlements Programme (UN - Habitat), 2002
- [4] Yamato, N.; Sasaki, K.; Hamada, Y.; Ito, K.; Wong, Y.Y.: *Global Power City Index 2015, Summary*. (Institute for Urban Strategies - The Mori Memorial Foundation, Tokyo, Japan), 2015
- [5] Economist Intelligence Unit: *European Green City Index*. (Siemens AG, Munich, Germany), 2009
- [6] Sovacool, B.K.; Brown, M.A.: Twelve metropolitan carbon footprints: A preliminary comparative global assessment, *Energy Policy*, 2010 **38**, 4856-4869 DOI: 10.1016/j.proeng.2017.07.146

- [7] Christopher, J.M.; Kammen, D.M.: A Consumption-Based Greenhouse Gas Inventory of San Francisco Bay Area Neighborhoods, Cities and Counties: Prioritizing Climate Action for Different Locations, *Bay Area Air Quality Management District*, UC Berkeley, 2015
escholarship.org/uc/item/2sn7m83z, Accessed: 15th August 2016.
- [8] López-Ruiz, V.R.; Alfaro-Navarro, J.L.; Nevado-Pena, D.: Knowledge-city index construction: An intellectual capital perspective, *Expert Syst. Appl.*, 2014 **41**, 5560-5572 DOI: 10.1016/j.eswa.2014.02.007
- [9] Toro, J.; Requena, I.; Duarte, O.; Zamorano, M.: A qualitative method proposal to improve environmental impact assessment, *Environ. Impact Assess. Rev.*, 2013 **43**, 9-20
DOI: 10.1016/j.eiar.2013.04.004
- [10] Eccleston, C.H.: *The EIS Book: Managing and Preparing Environmental Impact Statements*. (CRC Press), 2013 ISBN 9781466583641
- [11] Canter, L.W.: *Environmental impact assessment*, second edition. (McGraw-Hill, Inc., USA), 1996 ISBN 978-0071141031
- [12] Öberg, C.; Hüge-Brodin, M.; Björklund, M.: Applying a network level in environmental impact assessments, *J. Bus. Res.*, 2012 **65**, 247-255
DOI: 10.1016/j.jbusres.2011.05.026
- [13] Speizer, F.; Magyar, I.; Enisz, K.: Municipal environmental-monitoring system, *Hung. J. Ind. Chem.*, 2010 **38**(1), 63-66
- [14] Utasi, A.; Yuzhakova, T.; Sebestyén, V.; Németh, J.; Robu, B.; Rédey, Á.; Lakó, J.; Fráter, T.; Ráduly, I.; Ráduly, L.; Popita, G.: Advanced Quantitative Environmental Impact Assessment Method, *Environ. Eng. Manag. J.*, 2013 **12**, 305-310
- [15] Shaid, A.: Methodological Limitations of Determining Global Pollution Index as a Tool for Environmental Impact Assessment and a Proposed Extension, *Environment and Ecology Research*, 2014 **2**, 240-247 DOI: 10.13189/eer.2014.020603
- [16] Saaty, R.W.: The analytic hierarchy process - what it is and how it is used, *Math. Modelling*, 1987 **9**(3), 161-176 DOI: 10.1016/0270-0255(87)90473-8
- [17] Ching-Lai, H.; Kwangsun, Y.: *Multiple Attribute Decision Making - Methods and Applications: A State-of-the-Art Survey*. (Springer-Verlag, Berlin, Heidelberg), 1981 ISBN 978-3-642-48318-9
- [18] MacCrimon, K.R.: *Decision Making Among Multiple-Attribute Alternatives: a Survey and Consolidated Approach* RAND memorandum. (The Rand Corporation, Santa Monica, California), 1968
- [19] Herva, M.; Roca, E.: Review of combined approaches and multi-criteria analysis for corporate environmental evaluation, *J. Clean. Prod.*, 2013 **39**, 355-371 DOI: 10.1016/j.jclepro.2012.07.058
- [20] SDEWES Index [Online]
www.sdewes.org/sdewes_index.php, Accessed: 23rd October 2017.
- [21] Kılıkış, Ş.: Sustainable development of energy, water and environment systems index for Southeast European cities, *J. Clean. Prod.*, 2016 **130**, 222-234 DOI: 10.1016/j.jclepro.2015.07.121
- [22] Kılıkış, Ş.: Composite index for benchmarking local energy systems of Mediterranean port cities, *Energy*, 2015 **92**, 622-638
DOI: 10.1016/j.energy.2015.06.093
- [23] Kılıkış, Ş.: Benchmarking South East European Cities with the Sustainable Development of Energy, Water and Environment Systems Index, *J. Sustain. Dev. Energy Water Environ. Syst.*, In Press, 2017
DOI: 10.13044/j.sdewes.d5.0179
- [24] Holler, A.; Mader-Hirt, D.I.L. Eds.: *Climate Star 2012, The European Award for Local Climate Protection Initiatives*. (Land Niederösterreich, Gruppe Raumordnung, Umwelt und Verkehr, St. Pölten, Austria), 2012
old.klimabuendnis.org/fileadmin/inhalte/dokumente/2012/Readme_ClimateStar2012_en.pdf, Accessed: 14th February 2016.
- [25] *Energy Strategy of Veszprém for 2010-2025* (Veszprém Megyei Jogú Város energetikai stratégiája 2010-2025, in Hungarian), Veszprém, 2011
- [26] *Zalaegerszeg – Ecocity, Renewable Energy Strategy* (Zalaegerszeg – Ököváros, Megújuló Energia Stratégia, in Hungarian), Zalaegerszeg, 2014
- [27] OECD-JRC, *Handbook on Constructing Composite Indicators: Methodology and User Guide*. (OECD, Paris), 2008 ISBN 978-92-64-04345-9
- [28] Degree Days, [Online]. www.degreedays.net/, Accessed: 3rd April 2016.
- [29] *Integrated urban development strategy of Veszprém* (Veszprém Megyei Jogú Város integrált településfejlesztési stratégia, in Hungarian), Veszprém, 2014
- [30] *Integrated urban development strategy of Zalaegerszeg* (Zalaegerszeg Megyei Jogú Város integrált településfejlesztési stratégia 2014-2020, in Hungarian), Zalaegerszeg, 2014
- [31] Covenant of Mayors for Climate & Energy, [Online].
www.covenantofmayors.eu/actions/sustainable-energy-actionplans_en.html, Accessed: 20th November 2017
- [32] International Energy Agency, [Online].
www.iea.org/statistics/statisticsearch/report/?country=HUNGARY=&product=balances, Accessed: 9th November 2017
- [33] Böcskei, Zs.: Biomethane production in Zalaegerszeg (presentation), Workshop on procedures enabling better use of the energy content of sewage and sewage sludge, Budapest, 22nd June 2017
- [34] National Renewable Energy Laboratory, Solar and Wind Energy Resource Assessment, [Online].
maps.nrel.gov/swera/. Accessed: 26th March 2016
- [35] Horváth, F.; Bada, G.; Windhoffer, G.; Csontos, L.; Dombrádi, E.; Dövényi, P.; Fodor, L.; Grenerczy, Gy.; Sikhegyi, F.; Szafián, P.; Székely, B.; Timár, G.; Tóth, L.; Tóth, T.: Atlas of the present-day geodynamics of the Pannonian basin: Euroconform maps with explanatory text (A Pannonmedence jelenkori geodinamikájának atlasza: Euro-konform térképsorozat és magyarázó, in Hungarian), *Magyar Geofizika*, 2006 **47**(4), 133-137

- [36] Carr, G.M.; Rickwood, C.J.: Water quality: Development of an Index to assess country performance. *UNEP GEMS/Water Programme 1-24*, 2007
- [37] Esty, D.C.; Levy, M.A.; Kim, C.H.; de Sherbinin, A.; Srebotnjak, T.; Mara, V.: *2008 Environmental Performance Index*, New Haven: Yale Center for Environmental Law and Policy, 2008
- [38] Mekonnen, M.M.; Hoekstra, A.Y.: National water footprint accounts: The green, blue and grey water footprint of production and consumption, *Value of Water Research Report*, 2011 **50**, 1-50
- [39] Hungarian Air Quality Network (Országos Légszennyezettségi Mérőhálózat, in Hungarian), Ministry of Agriculture, [Online]. www.vevegominoseg.hu, Accessed: 3rd April 2016
- [40] Global Footprint Network, [Online]. www.footprintnetwork.org/en/index.php/GFN/page/trends/hungary/, Accessed: 28th April 2016
- [41] Company database (Cégbank, a cégadatbázis, in Hungarian), [Online]. www.cegbank.hu, Accessed: 11th April 2016
- [42] Airport carbon accreditation, [Online]. www.airportcarbonaccredited.org/airport/participants/europe.html, Accessed: 6th May 2016
- [43] North-west Hungary Transportation Center Ltd. (Északnyugat-magyarország Közlekedési Központ Zrt., in Hungarian), [Online]. www.enykk.hu, Accessed: 3rd April 2016
- [44] The World Bank, [Online]. data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD, Accessed: 4th April 2016
- [45] Abdallah, S.; Michaelson, J.; Shah, S.; Stoll, L.; Marks, N.: *The Happy Planet Index 2012 Report - A global index of sustainable well-being*, London, 2012 ISBN 978 1 908506 17 7
- [46] OECD, Country Note - Education at a Glance 2013: Hungary, 2013 DOI: 10.1787/eag-2017-50-en
- [47] Benton, E.; Jenks, M.; Williams, K.: *The Compact City: A Sustainable Urban Form?* (Routledge), 2003 ISBN: 978-0419213000
- [48] Hungarian Central Statistical Office, *Statistical Yearbook of Hungary 2013* (Központi Statisztikai Hivatal, Magyar Statisztikai Évkönyv 2013, in Hungarian), Budapest, 2014
- [49] Hungarian Central Statistical Office, [Online]. Available: www.ksh.hu/docs/hun/xstadat/xstadat_eves/i_ur005.html, Accessed: 28th April 2016
- [50] Hungarian Central Statistical Office, *Research and development 2014: Statistical Reflections* (Kutatás-fejlesztés 2014 Statisztikai tükör, in Hungarian), Budapest, 2015
- [51] European Patent Office, [Online]. www.epo.org/, Accessed: 3rd April 2016
- [52] Site of National Higher Education Information Centre, [Online]. www.felvi.hu, Accessed: 5th May 2016
- [53] Scimago Institutions Ranking, [Online]. <http://www.scimagoir.com/webvisibility.php?rankingttype=webvisibility&indicator=Website%20Size§or=&country=HUN&display=table&page=2&year=2008.>, Accessed: 5th May 2016
- [54] Scimago Journal & Country Rank, [Online]. <http://www.scimagojr.com/countryrank.php>, Accessed: 4th April 2016
- [55] Baranyák, Z.; Kontra, J.; Kovács, A.; Havas, M.; Jani, I.; Mayer, Z.; Pálfi, Sz.; Garbai, L.; Kovács, K.: *Zalaegerszeg Smart City 2050*, Goodwill Consulting, 2016

SURFACE ENERGY HETEROGENEITY PROFILES OF CARBON NANOTUBES WITH A COPOLYMER-MODIFIED SURFACE USING SURFACE ENERGY MAPPING BY INVERSE GAS CHROMATOGRAPHY

FRUZSINA GERENCSÉR,¹ NORBERT RIEDER,¹ CSILLA VARGA,² JENŐ HANCSÓK,²
AND ANDRÁS DALLOS^{1*}

¹ Department of Physical Chemistry, University of Pannonia, 10 Egyetem str., Veszprém, H-8200, HUNGARY

² MOL Department of Hydrocarbon and Coal Processing, University of Pannonia, 10 Egyetem str., Veszprém, H-8200, HUNGARY

The effectiveness and quantitative control of the surface transition of multi-walled carbon nanotubes (MWCNTs) was characterized by inverse gas chromatography (IGC). The surface energy profile of carbon nanotubes compatibilized with an olefin-maleic-anhydride-ester-amide (OMAEA)-type coupling agent was determined by a surface energy analyzer (SEA). The surface energetic heterogeneity with energy distributions of dispersive and specific (acid-base) components of the surface energy of the MWCNTs were determined at various surface coverages. The results of the surface energy mapping showed that surface treatment significantly reduced the dispersive surface energy of MWCNTs and increased the specific surface energy. Furthermore, the surface modification enhanced its Lewis basic character and simultaneously decreased the acidic character of MWCNTs. It has been demonstrated that the surface treatment modified the heterogeneity profiles of the energetic surface of the carbonaceous nanomaterials.

Keywords: carbon nanotubes, surface treatment, inverse gas chromatography, surface energy analysis

1. Introduction

Carbon nanotubes (CNTs) can serve as excellent candidate materials for uses in numerous industrial applications because of their considerable advantages. CNTs are one of the best reinforcing constituents for nanocomposites [1] and hopefully catalytic metal-support in heterogeneous catalysis [2]. CNTs could replace the common catalyst supports of Ni/Mo-catalysts used in the production of fuel components of engine fuels with high hydrogen contents in their molecular structures [3]. Carbon nanotube-supported Co/Mo-catalysts with different Co/Mo atomic ratios were successfully used in the hydrocracking reaction of the vacuum residue of crude oil from Gudao oil field [4]. In the Fischer-Tropsch process (FTP), CNTs that supported transition metal catalysts are used to increase catalytic activities. An excellent study of FTP on Co catalysts supported by CNTs was reported by Tavasoli et al. [5]. Chen et al. [6] demonstrated that Fe nanoparticles encapsulated in CNTs are promising catalyst in FTP to synthesize light olefins. The catalytic consequence of hydrothermal liquefaction of microalgae to produce bio-oil over CNT-

supported transition metal (Co, Ni, Pt) catalysts was reported by Chen et al. [7].

To change the wettability and chemical character of the CNTs or to avoid agglomeration in nanocomposites, the CNT surfaces are often exposed to surface functionalization [2] and modification processes using polyfunctional anchoring, capping, and coupling agents [8]. Research has shown that metal-support bindings can be strengthened by functional groups that are covalently bonded (grafted) to the support. Functionalized carbon nanotube-supported Pt nanoparticles were applied with favourable results in terms of selective olefin hydrogenation [9]. Because CNTs adsorb molecules well, functionalized CNTs are attractive chromatographic stationary phases for separation of normal and isoalkanes and aromatic compounds in the development of alternative fuels with high hydrogen/carbon ratios [10].

However, non-covalent functionalization using coupling agents or compatibilizers does not perturb the structure of the carbon nanotubes, establishes proper interactions between carbon nanotubes and the polymer matrix, and prevents the formation of nanotube agglomerates [11]. In terms of the properties of the reinforced composites of CNTs, the couplings between the nanotubes and the matrix are important beside the mechanical properties of the building parts [12]. These interactions depend on the surface properties and energies of the two materials. The surfaces of chemically derivatized CNTs were investigated by means of various ana-

*Correspondence: dallos@almos.vein.hu

lytical methods, e.g. thermal analysis [9-10,13-14], infrared spectroscopy (IR) [10,13], transmission electron microscopy (TEM) [13,16], Raman [16] and atomic force microscopy [10], and inverse gas chromatography (iGC).

iGC is a precise analytical method which is suitable for determining the surface energetic characteristics of the CNTs [13,15-17]. iGC was used for the characterization of the chemical character of the surface and was utilized to measure dispersive and specific surface energies, of numerous CNT substances [18]. The quantitative characterization of surface functionalization by surface energy mapping is of great importance. However, previous papers have presented surface energy values for functionalized CNTs over unclear surface coverages without energetic profiles and surface energy distribution functions, which, therefore, could not give correct information on the surface of the CNTs.

In this study, the dispersive, specific (acid-base) components of the surface energy with their heterogeneity charts and energy probability density functions of untreated and compatibilized MWCNTs are presented. A comparative quantitative characterization of the effectiveness and quantitative control of surface treatment is given. The exclusive energy scaling of the surfaces of the MWCNTs by energy heterogeneity charts with surface energy probability density functions over wide surface coverages is the new approach and main novelty of this paper.

2. Experimental and Methods

2.1. Samples and Measurements

Multi-walled carbon nanotubes (MWCNTs) were manufactured at 973 K by the chemical vapour deposition (CVD) process over a Fe/Co bimetallic catalyst at the Department of Chemical Engineering Science (University of Pannonia, Veszprém, Hungary) [19]. Their diameter was between 10 and 20 nm and their average length was above 30 μm .

An olefin-maleic-anhydride-ester-amide (Fig. 1) copolymer (OMAEA) was used as a compatibilizer. The coupling agent was synthesized at the Department of MOL Hydrocarbon and Coal Processing (University of Pannonia, Veszprém, Hungary). The surface of MWCNTs was covered by the compatibilizer from a hydrocarbon solution of the coupling agent while the mixture was stirred for 1 hour at 333 K. The solvent was subsequently evaporated and the treated MWCNTs were dried at 383 K for 2 hours in air [11].

The surface energies of as-received and compatibilized samples of MWCNTs were measured by a Surface Energy Analyzer (iGC-SEA, Surface Measurement Systems Ltd., Alpertown, UK) over a series of surface coverages from $(n/n_m) = 0.005$ to $(n/n_m) = 0.030$. iGC samples were produced by filling 20-25 mg of CNTs into silanized Pyrex glass tubes of I.D. = 3 mm under a vacuum and moderate vibration. The samples of MWCNTs were stabilized in the column with plugs of

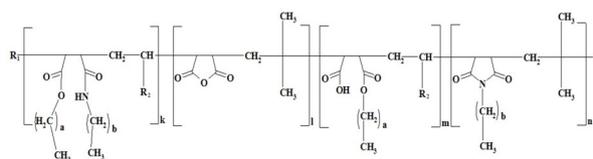


Figure 1. Structure of the olefin-maleic-anhydride-ester-amide copolymer (OMAEA) coupling agent, where R1: alkyl chain with length of the olefinic monomer; R2: alkyl chain with R1–2 carbon number; a, b: 2–21; k: 0.2–2; l: 1–7; m: 1–7 and n: 0.3–2 [11]

silanized glass wool. The samples were preconditioned in the column at the actually measured temperature for 60 minutes before each measurement. The iGC experiments were carried out at a column temperature of 353 K, with a Helium carrier gas flow of 10 cm^3/min . Methane gas was used as a dead-time marker using a flame ionization detector; and n-hexane, n-heptane, n-octane, n-nonane, chloroform and toluene as test compounds. The surface energy values were estimated using the specific retention volumes of the test compounds [20].

The specific retention volumes were obtained from the adjusted retention times:

$$V_w = t_R' \dot{V}_c / m_{sp} \quad (1)$$

The mean flow rate of the carrier gas in the column, \dot{V}_c , was evaluated as given in Ref. [20].

2.2. Theoretical Methodologies

The dispersive component of the surface energy (γ_s^d) and its heterogeneity profile of samples of MWCNTs were calculated using the Dorris-Gray method [21] over different surface coverages:

$$\gamma_s^d = \frac{1}{4\gamma_{\text{CH}_2}} \left(\frac{RT \cdot \ln(V_{w,nC+1}/V_{w,nC})}{N \cdot a_{\text{CH}_2}} \right)^2 \quad (2)$$

When plotting $RT \cdot \ln(V_{w,nC})$ against carbon number, nC , for the n -alkane probes, a straight line is generated from the gradient from which the dispersive free energy of the sample surfaces of the MWCNTs, γ_s^d , can be calculated.

The specific (Lewis acid-base) surface energy γ_s^{ab} of samples of MWCNTs was calculated from the basic component (γ_s^-) and the acidic component (γ_s^+) of the surface energy:

$$\gamma_s^{\text{ab}} = 2 \cdot \sqrt{\gamma_s^+ \gamma_s^-} \quad (3)$$

The basic and acidic components of the surface energy were obtained from the specific parts of free enthalpy changes of adsorption $\Delta G_{\text{ads},i}^{\text{ab}}$ of polar probes i :

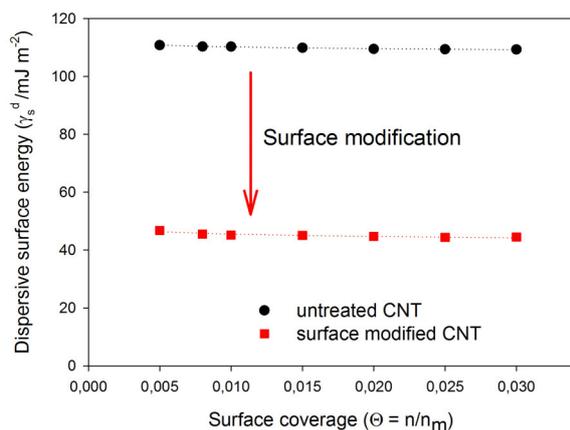


Figure 2. Dispersive surface energy profiles of untreated samples of MWCNTs and those treated with an olefin-maleic-anhydride-ester-amide copolymer (OMAEA) coupling agent compatibilized at 353 K and over various surface coverages (the dotted correlation lines are only to improve visualization).

$$\Delta G_{\text{ads},i}^{\text{ab}} = 2N \cdot a_i \cdot \left(\sqrt{\gamma_i^+ \gamma_s^-} + \sqrt{\gamma_i^- \gamma_s^+} \right) \quad (4)$$

applying the van Oss-Chaudhury-Good theory [22] with the Della Volpe scale [23]. The specific free energy changes of adsorption of the polar probes were obtained as suggested by Donnet et al. [24].

3. Results and Analysis

3.1. Experiments

The dispersive surface energy profiles of the untreated samples of MWCNTs and those treated with the olefin-maleic-anhydride-ester-amide copolymer (OMAEA) coupling agent compatibilized at 353 K and over low surface coverages (n/n_m) are presented in Fig. 2. The energy profiles show reasonable devaluation of dispersive surface energy for MWCNTs after surface modification detected by n -alkane molecular probes: the dispersive surface energy (γ_s^d) of the MWCNTs decreased to half of its initial value.

The untreated samples of MWCNTs exhibited dispersive surface energies of $\sim 110 \text{ mJ/m}^2$ at 353 K, which is comparable to the values reported by other researchers studying carbon nanotubes [13-15] and graphitic carbon materials [25]. The relatively high values of the dispersive surface energy of untreated MWCNTs can be attributed to a strong nonpolar interaction potential to build physical long-range Keesom, Debye, and London attractions, which explains their high tendency to agglomerate [14]. However, the anchoring of olefin-maleic-anhydride-ester-amide (OMAEA) up on the MWCNTs surface caused a marked decrement in dispersive part of surface energy from $\sim 110 \text{ mJ/m}^2$ to $\sim 48 \text{ mJ/m}^2$ at 353 K. The large drop in the value of γ_s^d of

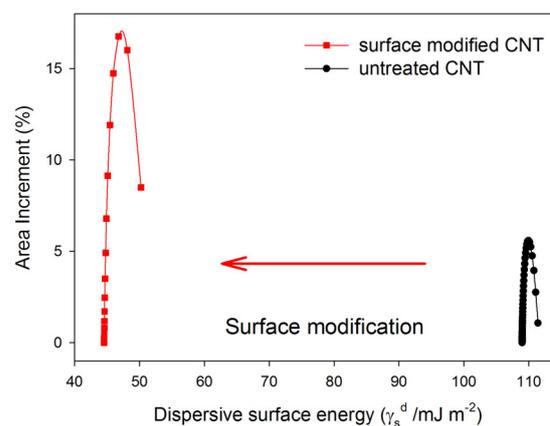


Figure 3. Dispersive surface energy probability functions of untreated samples of MWCNTs and those treated with an olefin-maleic-anhydride-ester-amide copolymer (OMAEA) coupling agent compatibilized at 353 K (the solid correlation lines are only to improve visualization).

surface-treated MWCNTs shows that the dispersive surface energy of MWCNTs has been obviously altered by the coupling agent.

The surface treatment also affected the dispersive surface energy heterogeneity profile of the MWCNTs. The surface energy mapping of the samples of MWCNTs indicated that the dispersive components of surface energies of untreated samples of MWCNTs are almost constant within the region of low surface coverage. Consequently, the surface of the untreated MWCNTs can be considered quasi-homogeneous. However, the dispersive surface energy heterogeneity profiles of the treated MWCNTs prove that the copolymer-modified MWCNT surface is energetically slightly heterogeneous, because the dependence of γ_s^d on surface coverage is relatively strong within the region of low surface coverage.

In addition, the distributions of the dispersive surface energies (Fig. 3) obtained by point-by-point integration of dispersive surface energy profiles over the investigated range of the surface coverage support in a more illustrative manner also results in an increase in the dispersive surface energy heterogeneity. The dispersive surface energy probability function of the MWCNTs became more spread out after modification of the surface indicated a greater degree of energetic surface inhomogeneity.

The specific surface energy (γ_s^{ab}) profiles of the untreated samples of MWCNTs and those treated with OMAEA compatibilized at 353 K and over various surface coverages (n/n_m) are presented in Fig. 4. The untreated samples of MWCNTs possess specific surface energy of $\sim 10 \text{ mJ/m}^2$ at 353 K, which value is near to that given by Lou et al. (8.84 mJ/m^2) for pristine carbon nanotubes at 373 K and over undefined degrees of surface coverage [14]. The quantitative surface energy analysis obtained by iGC-SEA methodology demonstrated that surface treatment of MWCNTs resulted in

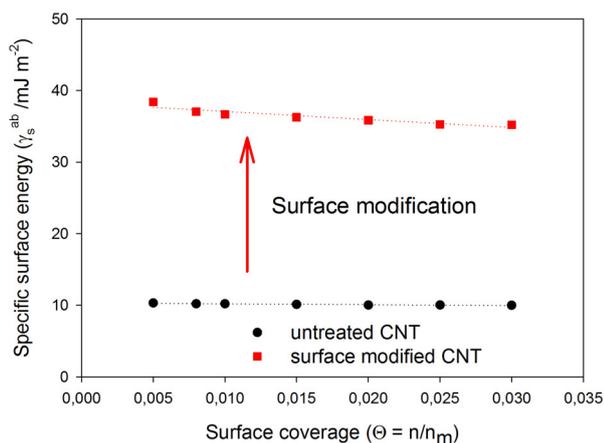


Figure 4. Specific surface energy profiles of untreated samples of MWCNTs and those treated with an olefin-maleic-anhydride-ester-amide copolymer (OMAEA) coupling agent compatibilized at 353 K and over various surface coverages (the dotted correlation lines only improve visualization).

significant changes in surface energies: the specific surface energy of CNT surfaces increased more than four-fold, from $\sim 10 \text{ mJ/m}^2$ to $\sim 41 \text{ mJ/m}^2$.

Furthermore, the dependence of γ_s^{ab} on surface coverage is pronounced for the compatibilized samples of MWCNTs which indicates that energetic heterogeneity attributed to chemical heterogeneity and the existence of electron donor-acceptor atomic groups on the surface. However, the quasi-constant specific component of surface energy for untreated MWCNTs suggests an energetically homogeneous surface and the absence of high specific energy surface sites.

The specific surface energy values of compatibilized MWCNTs are much higher than those of the untreated surfaces, declared the enhanced connection between compatibilized MWCNTs and polar analytes. The observed diversity in specific surface energies is accomplished from the adsorbed polar atomic clusters: namely the moderately electron-withdrawing maleic anhydride groups; and the electron-donating ester and amide groups with different nucleophilic or electrophilic characteristics.

The specific surface energy distributions in Fig. 5 represent the heterogeneity of the samples of MWCNTs and reveal that the untreated MWCNTs exhibited γ_s^{ab} values which varied from 9.9 to 10.3 mJ/m^2 . This small variation in the specific surface energy demonstrated a fairly energetically homogeneous surface for untreated MWCNTs. However, as surface treatment increased the concentrations of polar clusters on the surface of MWCNTs, the modified surface exhibited great variations in γ_s^{ab} (from 37.3 to 42.9 mJ/m^2), implying that the compatibilized MWCNTs are surface energetic heterogeneous.

The surface treatment also modified the chemical characteristics of the MWCNTs. The acid-base surface

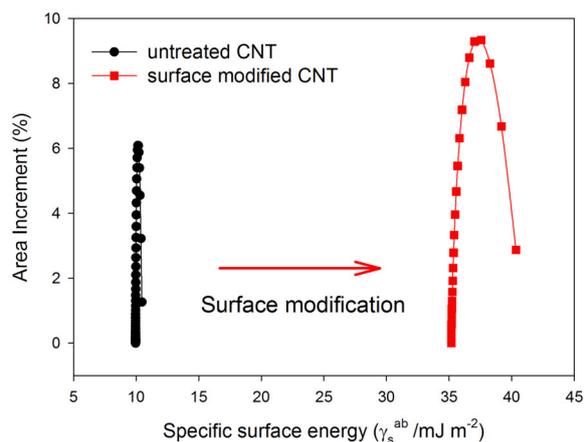


Figure 5. Specific surface energy probability functions of untreated samples of MWCNTs and those treated with an olefin-maleic-anhydride-ester-amide copolymer (OMAEA) coupling agent compatibilized at 353 K (the solid correlation lines only improve visualization).

energy mapping of the samples of MWCNTs (Figs. 6 and 7) indicate that both the electron-accepting and donating abilities of the MWCNTs were raised appreciably after compatibilization of untreated MWCNTs consistent with the adsorption of electron-withdrawing and electron-donating atomic groups on the surface. The untreated samples of MWCNTs possess base surface energy of $\sim 16 \text{ mJ/m}^2$ at 353 K, which value is near to that given by Lou et al. (12.97 mJ/m^2) for pristine carbon nanotubes at 373 K and over undefined surface coverage [14].

The surface energy mapping using the iGC-SEA methodology confirms that the surface treatment of MWCNTs raised the basic component (γ_s^-) of surface energy of MWCNTs: it resulted in a sevenfold increase from $\sim 16 \text{ mJ/m}^2$ to $\sim 112 \text{ mJ/m}^2$. The acidic component of surface energy γ_s^+ of untreated samples of MWCNTs was measured as $\sim 1.7 \text{ mJ/m}^2$ at 353 K, which is in good agreement with that reported by Lou et al. (1.51 mJ/m^2) for pristine carbon nanotubes at 373 K and over undefined surface coverages [14]. The surface modification of MWCNTs resulted in a more than twofold increase in the value of the acidic component (γ_s^+) of the surface energy of MWCNTs from $\sim 1.7 \text{ mJ/m}^2$ to $\sim 3.7 \text{ mJ/m}^2$.

However, the profiles of acid-base surface energy heterogeneity of the treated MWCNTs also prove that the copolymer-modified MWCNT surface became energetically more heterogeneous, because the dependencies of γ_s^- and γ_s^+ on surface coverage are relatively stronger than those of the untreated MWCNTs. The presence of non equi-energetic active surface centers exposes that the surfaces of the compatibilized MWCNTs are not energetically uniform for specific acid-base interactions and the surface treatment considerably altered the ability of MWCNTs to connect with molecular species by specific interactions.

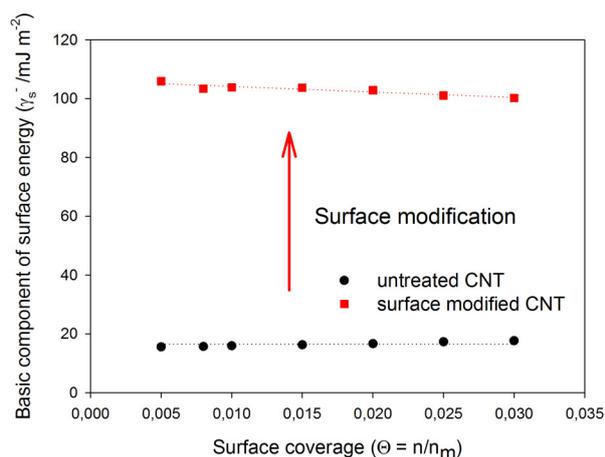


Figure 6. Basic component profiles of surface energy of untreated samples of MWCNTs and those treated with an olefin-maleic-anhydride-ester-amide copolymer (OMAEA) coupling agent compatibilized at 353 K and over various surface coverages (the dotted correlation lines only improve visualization).

The larger change in the base component of the surface energy (γ_s^-) after compatibilization with a polyalkenyl-poly-maleic-anhydride-ester-amide additive indicates a higher concentration of electron-donating ester and amide groups on the surface. The large value of γ_s^- (relative to γ_s^+) implies a more basic characteristic and donor properties of the surface of the treated MWCNTs.

4. Conclusion

The experimental data demonstrated that iGC is a useful methodology of characterizing the variation in the surface characteristics of MWCNTs after non-covalent functionalization. The exclusive energy scaling of the SEA methodology using energy heterogeneity charts with surface energy probability density functions over wide surface coverages presents profitable additional information on the differences in terms of the nature, homogeneity and heterogeneity of surface energies resulting from surface transformations. The multilateral surface energy analysis of SEA presents a quantitative control of the effectiveness of surface treatment and demonstrates the importance of the dependence of surface energy analysis on coverage.

Acknowledgement

The present work was published within the framework of the project GINOP-2.3.2-15-2016-00053.

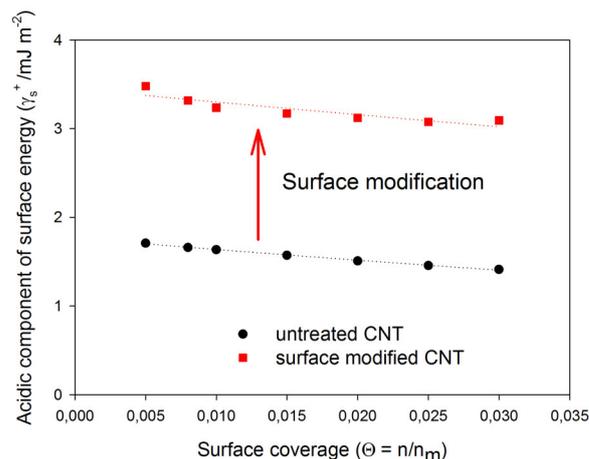


Figure 7. Acidic component profiles of surface energy of untreated samples of MWCNTs and those treated with an olefin-maleic-anhydride-ester-amide copolymer (OMAEA) coupling agent compatibilized at 353 K and over various surface coverages (the dotted correlation lines only improve visualization).

SYMBOLS

a_{CH_2}, a_i	cross sectional area of an adsorbed methylene group and of probe i
$\Delta G_{\text{ads},i}^{\text{ab}}$	specific free enthalpy of adsorption
m_{sp}	mass of the adsorbent in the column
N	Avogadro's number
R	gas constant
T	temperature
t_R'	adjusted retention time
\dot{V}_c	mean flow rate of the carrier gas
V_w	specific retention volume
$\gamma_s^{\text{ab}}, \gamma_s^{\text{d}}$	specific and dispersive parts of the surface energy of solid sample material
γ_{CH_2}	surface energy of a methylene group
γ_s^+, γ_s^-	acid-base components of the surface energy of solid sample material
γ_i^+, γ_i^-	acid-base components of the surface tension of polar liquid probe i
$\Theta = n/n_m$	surface coverage

REFERENCES

- [1] Dresselhaus, M.S.; Dresselhaus, G.; Avouris, P.: Carbon nanotubes: synthesis, structure, properties, and applications (Springer-Verlag GmbH, Berlin, Germany) 2001 DOI: 10.1007/3-540-39947-X

- [2] Yan, Y.; Miao, J.; Yang, Z.; Xiao, F.X.; Yang, H.B.; Liu, B.; Yang, Y.: Carbon nanotube catalysts: recent advances in synthesis, characterization and applications, *Chem. Soc. Rev.*, 2015 **44**, 3295–3346 DOI: 10.1039/C4CS00492B
- [3] Sági, D.; Holló, A.; Varga, G.; Hancsók, J.: Co-hydrogenation of fatty acid by-products and different gas oil fractions, *Journal of Cleaner Production*, 2017 **161**, 1352–1359 DOI: 10.1016/j.jclepro.2017.05.081
- [4] Li, C.; Shi, B.; Cui, M.; Shang, H.J.; Que, G.H.: Application of Co-Mo/CNT catalyst in hydrocracking of Gudao vacuum residue, *J. Fuel. Chem. Technol.*, 2007 **35**(4), 407–411 DOI: 10.1016/S1872-5813(07)60026-7
- [5] Tavasoli, A.; Sadagiani, K.; Khorashe, F.; Seifkordi, A.A.; Rohani, A.A.; Nakhaeipour, A.: Cobalt supported on carbon nanotubes — A promising novel Fischer–Tropsch synthesis catalyst, *Fuel Process. Technol.*, 2008 **89**(5), 491–498 DOI: 10.1016/j.fuproc.2007.09.008
- [6] Chen, X.; Deng, D.; Pan, X.; Bao, X.: Iron catalyst encapsulated in carbon nanotubes for CO hydrogenation to light olefins, *Chinese J. Catal.*, 2015 **36**(9), 1631–1637 DOI: 10.1016/S1872-2067(15)60882-8
- [7] Chen, Y.; Mu, R.; Yang, M.; Fang, L.; Wu, Y.; Wu, K.; Liu, Y.; Gong, J.: Catalytic hydrothermal liquefaction for bio-oil production over CNTs supported metal catalysts, *Chem. Eng. Sci.*, 2017 **161**, 299–307 DOI: 10.1016/j.ces.2016.12.010
- [8] Varga, Cs.; Bartha, L.: A novel route for injection moulding of long carbon fibre-reinforced LLDPE, *J. Reinf. Plast. Comp.*, 2014 **33**(20), 1902–1910 DOI: 10.1177/0731684414549443
- [9] Chen, P.; Chew, L.M.; Xia, W.: The influence of the residual growth catalyst in functionalized carbon nanotubes on supported Pt nanoparticles applied in selective olefin hydrogenation, *J. Catal.*, 2013 **307**, 84–93 DOI: 10.1016/j.jcat.2013.06.030
- [10] Speltini, A.; Merli, D.; Quartarone, E.; Profumo, A.: Separation of alkanes and aromatic compounds by packed column gas chromatography using functionalized multi-walled carbon nanotubes as stationary phases, *J. Chromatogr. A*, 2010 **1217**(17), 2918–2924 DOI: 10.1016/j.chroma.2010.02.052
- [11] Szentes, A.; Varga, Cs.; Horváth, G.; Bartha, L.; Kónya, Z.; Haspel, H.; Szél, J.; Kukovecz, Á.: Electrical resistivity and thermal properties of compatibilized multi-walled carbon nanotube/polypropylene composites, *Express Polym. Lett.*, 2012 **6**(6), 494–502 DOI: 10.3144/expresspolymlett.2012.52
- [12] Varga, Cs.; Miskolczi, N.; Bartha, L.; Lipóczi, G.; Falussy, L.: Improving the compatibility of man-made fibre reinforced composites, *Hun. J. Ind. Chem.*, 2008 **36**(1–2), 137–142
- [13] Zhang, X.; Yang, D.; Xu, P.; Wang, C.; Du, Q.: Characterizing the surface properties of carbon nanotubes by inverse gas chromatography, *J. Mater. Sci.*, 2007 **42**(17), 7069–7075 DOI: 10.1007/s10853-007-1536-7
- [14] Luo, Y.; Zhao, Y.; Cai, J.; Duan, Y.; Du, S.: Effect of amino-functionalization on the interfacial adhesion of multi-walled carbon nanotubes/epoxy nanocomposites, *Materials and Design*, 2012 **33**, 405–412 DOI: 10.1016/j.matdes.2011.04.033
- [15] Menzel, R.; Tran, M.Q.; Menner, A.; Kay, C.W.M.; Bismarck, A.; Shaffer, M.S.P.: A versatile, solvent-free methodology for the functionalisation of carbon nanotubes, *Chem. Sci.*, 2010 **1**(5), 603–608 DOI: 10.1039/C0SC00287A
- [16] Menzel, R.; Lee, A.; Bismarck, A.; Shaffer, M.S.P.: Inverse gas chromatography of as-received and modified carbon nanotubes, *Langmuir*, 2009 **25**(14), 8340–8348 DOI: 10.1021/la9000607s
- [17] Diaz, E.; Ordóñez, S.; Vega, A.: Characterization of nanocarbons (nanotubes and nanofibers) by inverse gas chromatography, *J. Phys. Conf. Ser.*, 2017 **61**(1), 904–908 DOI: 10.1088/1742-6596/61/1/180
- [18] Menzel, R.; Lee, A.; Bismarck, A.; Shaffer, M.S.P.: Deconvolution of the structural and chemical surface properties of carbon nanotubes by inverse gas chromatography, *Carbon*, 2012 **50**(10), 3416–3421 DOI: 10.1016/j.carbon.2012.02.094
- [19] Szentes, A.; Horváth, G.: Role of catalyst support in the growth of multi-walled carbon nanotubes, *Hun. J. Ind. Chem.*, 2008 **36**(1–2), 113–117
- [20] Kondor, A.; Dallos, A.: Adsorption isotherms of some alkyl aromatic hydrocarbons and surface energies on partially dealuminated Y faujasite zeolite by inverse gas chromatography, *J. Chromatogr. A*, 2014 **1362**, 250–261 DOI: 10.1016/j.chroma.2014.08.047
- [21] Dorris, G.M.; Gray, D.G.: Adsorption of n-alkanes at zero surface coverage on cellulose paper and wood fibers, *J. Colloid Interf. Sci.*, 1980 **77**(2), 353–362 DOI: 10.1016/0021-9797(80)90304-5
- [22] van Oss, C.J.; Good, R.; Chaudhury, M.K.: Additive and nonadditive surface tension components and the interpretation of contact angles, *Langmuir*, 1988 **4**(4), 884–891 DOI: 10.1021/la00082a018
- [23] Della Volpe, C.; Sibioni, S.: Some reflection on acid-base solid surface free energy theories, *J. Colloid Interf. Sci.*, 1997 **195**(1), 121–136 DOI: 10.1006/jcis.1997.5124
- [24] Donnet, J.B.; Park, S.J.; Balard, H.: Evaluation of specific interactions of solid surfaces by inverse gas chromatography. A new approach based on polarizability of the probes, *Chromatographia*, 1991 **31**, 435–440 DOI: 10.1007/BF02262385
- [25] Donnet, J.B.; Park, S.J.: Surface characteristics of pitch-based carbon fibers by inverse gas chromatography method, *Carbon*, 1991 **29**(7), 955–961 DOI: 10.1016/0008-6223(91)90174-H

INITIAL ELECTRICAL PARAMETER VALIDATION IN LEAD-ACID BATTERY MODEL USED FOR STATE ESTIMATION

BENCE CSOMÓS, DÉNES FODOR,* AND GÁBOR KOHLRUSZ

Department of Automotive Mechatronics, Institute of Mechanical Engineering, University of Pannonia, Egyetem u. 10., Veszprém, H-8200, HUNGARY

The paper presents a current impulse-based excitation method for lead-acid batteries in order to define the initial electrical parameters for model-based online estimators. The presented technique has the capability to track the SoC (State of Charge) of a battery, however, it is not intended to be used for online SoC estimations. The method is based on the battery's electrical equivalent Randles' model [1]. Load current impulse excitation was applied to the battery clamps during discharge while the voltage and current was logged. Based on the Randles' model, a model function and a fit function were implemented and used by exponential regression based on the measured data. The diffusion-related non-linear characteristic of the battery was approximated by a capacitor-like linear voltage function for speed and simplicity. The initial capacitance of this bulk capacitor was estimated by linear regression on measurements recorded in the laboratory. Then, the RC parameters of the equivalent battery model were derived from exponential regression on transients during each current impulse cycle. The battery model with initial RC parameters is suitable for model-based online observers.

Keywords: Battery, SoC, Exponential regression, Randles' model, Load current impulse

1. Introduction

In our daily lives, the number of mobile devices and utilities that can operate without grid connections is increasing. Even though lithium batteries possess better performance properties and energy indicators, lead-acid batteries are still cheaper, significantly present in commercial applications and almost fully recyclable. Therefore, any developments in lead-acid battery systems are still of interest.

According to Ref. [1], several methods exist to estimate a battery's State of Charge (SoC) and State of Health (SoH) but model-based prediction is the most widespread because of its reliability and robustness. Model-based methods, as the name suggests, need a valid, properly detailed electric battery model. The Randles' model as a standard battery model is very popular in the contexts of lead-acid and lithium-ion batteries because of its cost-effectiveness and the similarities of both types. By similarity it is meant that the same model can be reasonably used for the parameter estimation of both battery types [2-3].

Some additions to the standard Randles' model can be made if more details in electrochemistry are required such as diffusion in the bulk and porosity amongst others.

The model requires values of initial resistance (R) and capacitance (C). The more accurate the initial

parameters of the model, the faster and more reliable the convergence of a model-based predictor to the actual state, that is, the actual SoC.

The scope of the present work is to identify the initial values of RC components (parameters) by evaluating the voltage impulse responses excited by load currents in the time domain.

2. Battery model for impulse excitation

In this paper, a standard Randles' model [7-12] was analyzed that consists of charge-transfer resistance, R_{ct} , battery serial resistance, R_s , double-layer capacitance, C_{dl} , and bulk capacitance, C_b (Fig.1). The voltage references of the capacitors and currents in Fig.1 were set for discharge.

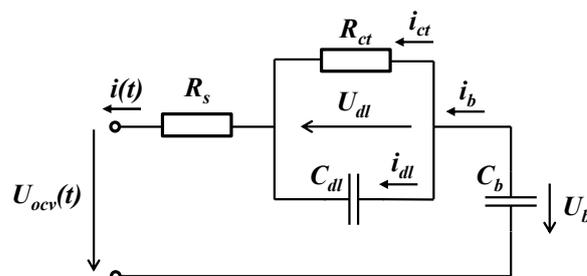


Figure 1. Randles' battery model applied to a discharging battery pack

*Correspondence: fodor@almos.vein.hu

2.1. State-space model and model function

By neglecting the intermediate mathematical steps and rearranging the state-space model, the system can be written in the following form:

$$\frac{d}{dt} \begin{bmatrix} U_{dl} \\ U_b \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_{dl}R_{ct}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_{dl} \\ U_b \end{bmatrix} + \begin{bmatrix} \frac{1}{C_{dl}} \\ -\frac{1}{C_b} \end{bmatrix} i \quad (1)$$

$$U_{ocv} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} U_{dl} \\ U_b \end{bmatrix} - R_s i. \quad (2)$$

By solving the output in Eq.(2) in terms of the time domain using a current impulse as the system input, the output function in Eq.(2) can be expressed by Eq.(3) that can be considered as the model function of the system. This form of the output equation serves as a basis for creating a fit function on measured voltage data and then, derive R and C values from the fit parameters.

For clarity, each of the terms of the output equation are grouped by alphabetical letters:

$$U_{ocv}(t) = Ae^{-\frac{t}{\tau}} + Bt + D \quad (3)$$

where tags A , B and C can be written according to

$$A = -U_{dl0} + R_{ct}i(t) \quad (4)$$

$$B = \frac{i(t)}{C_b} + \frac{U_{b0}}{t} \quad (5)$$

$$D = -(R_s + R_{ct})i(t) \quad (6)$$

where U_{ocv} is the battery's open-circuit voltage, τ is the system's time constant, t is the measurement time, and $i(t)$ is the impulse load current. Since the proposed method is based on load and relaxation cycles that follow each other during the analysis, U_{b0} represents the initial voltage of the bulk capacitor while U_{dl0} is the initial voltage of the double-layer capacitor at the beginning of each impulse cycle.

It should be noted that changes in current during each impulse cycle can be neglected as a result of working in the short time-constant region of the discharge curve, thus the current can be considered as a constant.

The battery model shown in Fig.1 is prepared for short-time transient analysis. Even though the model can be used and remains valid for modelling discharge processes that last for several hours, that is, for long-time transients, the accuracy of the model becomes poor under such circumstances. The reason for this is that the battery model presented excludes the diffusion effect that can even be observed by the initial valley-like voltage drop and later as a circle-like voltage response on the long-time discharge voltage curve (Fig.2). Such an exclusion was made because the scope of the current

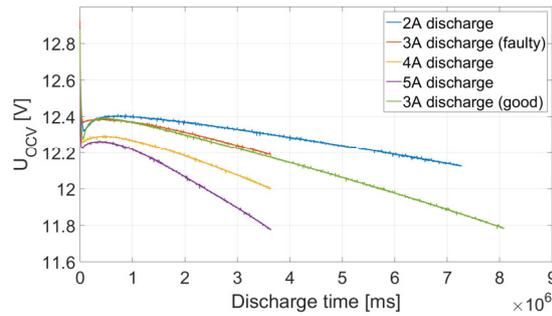


Figure 2. Discharge characteristics of a 15Ah AGM battery excited by different load currents

work is to focus on short-time voltage responses to avoid excessive measurement time intervals and computational resources. Consequently, the battery model was optimized for fast SoC detection by short-time battery checks.

2.2. Determining the initial capacity of C_b

In Eq.(1), the U_b voltage represents the equilibrium voltage of the battery, therefore, it is related to the battery's main charge and as a result its SoC. If the battery is excited by small $C/10$ - $C/30$ currents, U_b and U_{ocv} can be considered equal. Instead of U_b , U_{ocv} can be measured at the battery terminals. Therefore, the relationship between U_{ocv} and SoC is important and can be determined from laboratory OCV measurements.

A small discharge current was applied to the battery terminals under controlled conditions until the battery's OCV reached the factory's minimum voltage threshold from a fully charged state. The voltage, current and temperature were logged while the SoC could be calculated by the simple Coulomb-counting method during the process and saved in a lookup table. Then, the lookup table could be used to determine a discrete relationship between U_{ocv} and SoC in itself. The U_{ocv} - SoC characterisation method can be extended by the regression on lookup data in order to create a continuous U_{ocv} - SoC relationship. A linear function, such as a capacitor-like regression of U_{ocv} - SoC characteristics can be legitimate if the battery's excitation current is small, i.e. between $C/10$ and $C/15$ and is not discharged under 20-25% of SoC. This could be the case when low-power devices are considered as loads.

In the case of plain discharge, the SoC changes can be basically tracked by the basis of the B term like in Eq.(5) as conducted in the Coulomb-counting method [1]. In the presented battery model, the B term provides information on the long-term state of the battery and requires initial parameters such as C_b and U_{b0} . The former represents the battery's initial capacity, the latter is related to the battery's initial voltage at the beginning of each impulse cycle. Right before the first current impulse, U_{b0} is equal to the battery's equilibrium voltage since a relaxation time of between 30 minutes and one hour is sufficient for chemical processes to decay.

The initial value of C_b is crucial because it has a strong influence on both the initial voltage drop and the gradient of the long-time discharge voltage curve. In the proposed model, a simplified, that is, capacitor-like formula was used for initial battery capacity estimation so it approximates the battery non-linear discharge characteristic linearly. The introduction of a regression error of a few percent, however, can lead to an easier and faster determination of the initial capacity C_b .

The calculation of the battery's initial capacity was realized according to [4]. The fully charged battery at room temperature needed to be slowly discharged by a C/15 current until its OCV voltage reached the factory recommended minimum voltage threshold. Once the discharge had finished, 2 hours of relaxation had to be observed in order to return the battery back to its almost equilibrium state. Then, a C/15 slow charge had to proceed until the battery's OCV voltage reached the factory recommended maximum voltage threshold. In both cases, the battery's OCV voltage, current and temperature were recorded (Fig.3/a).

After averaging the discharge and charge-voltage measurements, linear regression was conducted on it according to

$$U_{ocv|C/15}(t) = \frac{1}{C_b} Q_b(t) + U_{b0} \quad (7)$$

where $Q_b(t)$ can be estimated by $Q_b(t) = i_0 t$.

Eq.(7) can be identified by the standard form of the linear curve $y=mx+b$. In Eq.(7), C_b is the battery's initial capacity, $Q_b(t)$ is the actual charge of the battery and U_{b0} is the actual voltage of C_b at the beginning of the impulses. By rearranging Eq.(7), C_b can be determined. Even though charge/discharge currents and SoC levels are constrained, the ambient temperature should be controlled as well to provide a constant temperature during the test periods. The value of C_b was calculated as 37766 F at 22°C using C/15 load currents.

3. Exponential regression to derive RC model parameters

The evaluation of measurement data and the comparison of measurements and simulations were realized using Matlab. The RC parameters in the model shown in Fig.1 and later in an OrCAD circuit were derived by an exponential regression using an appropriate fit function on the measured voltage data. The regression error between the measured and modelled characteristics can be minimized if the fit function follows the form of the model function, that is, both of them implement similar dynamics and the physical background of the inspected system. Therefore, the fit function can be written as

$$\hat{U}_{ocv}(t) = \hat{A}e^{-\frac{t}{\hat{\tau}}} + \hat{B} \quad (8)$$

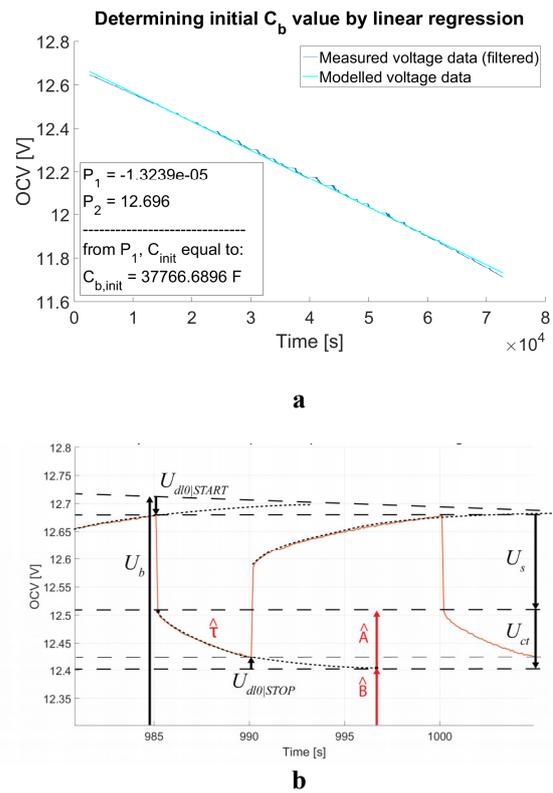


Figure 3. (a) Linear regression on the average of C/15 discharge and charge voltage data to derive initial value of C_b . Initial cut-off has been filtered. (b) References of model function and fit function

in a form that is similar to Eq.(3). The hat sign means that the form of Eq.(8) is similar to Eq.(3), but uses a different reference system.

When using Eq.(8) attention must be paid to the determination of RC parameters. Eq.(8) gives voltage references with respect to the ground, that is the x-axis, while Eq.(3) yields the references U_{dl} and U_s which correspond to U_b . References used by Eqs.(3) and (8) must be matched to derive correct RC parameters (Fig.3/b).

It can be seen that Eq.(8) neglects the linear term from Eq.(3). Since the effect of continuous slow discharge of the battery, which is linked to the C_b bulk capacitance in the battery model, possesses a time-constant several orders of magnitude higher than the $C_{dl}-R_{ct}$ subsystem, it cannot be observed during the short impulse cycles. However, it should be considered during the whole discharge process so Eqs.(3) and (8) need to be used together to estimate the SoC during long-term discharge processes.

According to Refs. [5-6], it is practical to evaluate only the discharge component of the voltage responses during every load cycle. This method is also supported by the fact that the dynamic behaviour of the battery for both load and relaxation states can be described by the same battery model since both responses originate from the same battery structure. Because load curves are only needed, the measured voltage should be separated from global data and then, concatenated right after each other. Since voltage data has been prepared in this way,

Table 1. Derived RC values of the battery model from regression

Derived parameter name	Average value
R_{ct}	0.032 Ohm
C_{dl}	92 F
R_s	0.056 Ohm
SoC after discharge	87.9 %

regression can be made during every discharge impulse cycle simultaneously.

The relaxation time during each impulse was set to provide enough time for the battery's double-layer capacitance to almost fully discharge. It is practical because it simplifies Eq. (8) by reducing the effect of the U_{dl0} term in Eq. (4).

The load time was set taking into consideration the time constant of the $C_{dl} - R_{ct}$ subsystem. According to experiments, it is within the range of 1 - 4s. The calculation method of the term \hat{b} in (8) is based on the calculation of the limit of the exponential term in Eq. (8). Practical experiments have proven that maintaining a load cycle of 4τ in length can speed up and correct the regression.

The RC values change during discharge. Due to offline voltage response evaluation, it is not possible to optimise these parameters according to load conditions. Therefore, the average RC values can be used for the whole period of discharge though this introduces a slight misalignment between the measured and modelled voltage characteristics.

By applying Eq. (8) on the prepared measurement data, the average RC parameters shown in Table 1 were derived.

4. Model implementation and validation of the impulse excitation method in OrCAD

The aim of the OrCAD implementation was to validate the proposed battery model and the applicability of the impulse excitation method for RC parameter determination. The equivalent circuit model shown in Fig. 1 was transformed into an electric circuit. The duration of the simulation was set to 1 hour and the time step was equal to the sample time of the real measurement, namely 100 ms.

4.1. Initialization of the OrCAD simulator and RC elements

The initial voltages of the capacitors and the proper directions of OrCAD elements should be set carefully. OrCAD uses a reference system of its own that influences the current references of each component. This should be considered while placing an element in the circuit editor and when comparing the electric circuit references to the Randles' model.

The initial voltages of the capacitors, U_{dl} and U_b , were set to describe the battery's discharge process and

Table 2. Initial values of simulation parameters that need to be set before a run

Simulation parameter	Initial value
Simulation step time	100 ms
Simulation duration	1 hour
U_{dl0}	0 V
U_{b0}	12.7 V
Relaxation time / cycles	10 s
Load time / cycles	5 s
Load current	3 A (with a 4 Ohm load)

thus also follow the voltage references of the Randles' model, shown in Fig. 1. The values of initial capacitor voltages were derived from the chemical background and assumptions. C_{dl} can be considered fully discharged through R_{ct} after the battery had relaxed (no load) for 2 hours so its initial voltage, U_{dl} , was set to 0. The bulk capacitance of the battery, C_b , reflects the lengthy time constant as well as diffusion-related processes and is related to the main charge storage capability of the battery. If the relaxation time is sufficiently long, the battery reaches an equilibrium state and its OCV becomes equal to U_b . An optimal relaxation time that is sufficiently long was derived from preceding measurements of discharge/charge cycles by applying different load currents and relaxation times. Based on experiments, a relaxation time of 2 hours was applied and as a result the OCV could be considered equal to U_b . Using these assumptions, the initial voltage, U_b , was set to 12.7 V and U_{dl} to 0 V.

The introduction of the load current as an impulse excitation can be achieved through a switched power resistor element which is setup in a similar way to the real arrangement (Fig. 3/b). The switching routine of the OrCAD element was tuned in accordance with real load-relaxation cycles that were applied to the real testbench. The resistor sets the load current that discharges the battery.

During this work, a load current of 3A was used with a load of 5s and relaxation cycles that lasted 10s. All of the initial values are summarized in Table 2.

4.2. Comparing the simulation with the battery model

In Fig. 4/a, the results of measured and simulated voltage responses are shown. The validity of the model was analyzed by the comparison of measured and simulated voltages. According to the setup, the comparison is performed within a time frame of 5,600 s. The blue curve that represents the simulated data has a longer tail than the red one because filtering needed to be performed on the measured data to cut the initialization process at the beginning of the testbench.

The zoomed-in segment shows the fitness of the simulated voltage response. The difference between the curves is relatively small, around 0.05 V to be precise. This error occurred because average RC values were used in the model instead of online tuned ones.

5. Conclusions

This work shows a current impulse-based excitation method that can be used to either track SoC changes during moderate discharge or find proper RC model values for model-based algorithms. The method is founded on an equivalent model-based approach that can implement the dynamic behaviour of a battery without using excessive chemical equations.

The technique uses offline analysis of a battery, therefore, it is not able to reasonably track SoC changes in real-time. This technique is not intended to estimate the SoC of batteries by itself, indeed, it estimates good set points for online estimators such as Kalman filters or other model-based observers.

The advantages of the presented approach are its rapid nature and simplicity while minimising the error of SoC estimation. The disadvantages of this method are its offline nature and related consequences.

This technique could potentially be applied to embedded systems and commercially.

6. Acknowledgement

This research was supported by the EFOP-3.6.1-16-2016-00015 project. The project was supported by the Hungarian Government and co-financed by the European Social Fund. The project was supported by the European Union and co-financed by the European Social Fund (EFOP- 3.6.2-16-2017-00002).

REFERENCES

- [1] Piller, S.; Perrin, M.; Jossen, A.: Methods for state-of-charge determination and their application, *Journal of Power Sources*, 2001 **96**(1), 113–120 DOI: 10.1016/S0378-7753(01)00560-2
- [2] Spagnol, P.; Rossi, S.; Savaresi, S.M.: Kalman filter SoC estimation for Li-ion batteries, *IEEE International Conference on Control Applications*, 28–30 Sept. 2011, Denver, USA, ISSN 1085-1992
- [3] Devarakonda, L.; Wang, H.; Hu, T.: Parameter identification of circuit models for lead-acid batteries under non-zero initial conditions, *American Control Conference*, 4–6 June 2014, Portland, USA
- [4] Platt, G.L.: *Battery modeling*, Vol. I., 2015 ISBN-13: 978-1-63081-023-8
- [5] Banaei A.; Fahimi B.: Real time condition monitoring in Li-ion batteries via battery impulse response, *IEEE Vehicle Power and Propulsion Conference (VPPC)*, 1–3 Sept. 2010, Lille, France ISSN 1938-8756
- [6] Banaei A.; Khoobroo A.; Fahimi B.: Online detection of terminal voltage in Li-ion batteries via impulse response, *IEEE Vehicle Power and Propulsion Conference, 2009. VPPC '09*, 7–10 Sept. 2009, Dearborn, USA ISSN 1938-8756
- [7] Kularatna, N.: Dynamics and modeling of rechargeable batteries, *IEEE Power Electronics Magazine*, 2014 **1**(4), 23–33 DOI: 10.1109/MPEL.2014.2361264
- [8] Lee, Y.-D.; Park, S.-Y.; Han, S.-B.: Online embedded impedance measurement using high-power battery charger, *IEEE Transactions on Industry Applications*, 2015 **51**(1), 498–508 DOI: 10.1109/TIA.2014.2336979
- [9] Kiani, M.: High resolution state of charge estimation in electrochemical batteries, *Applied Power Electronics Conference and Exposition (APEC), Twenty-Eighth Annual IEEE*, 17–21 March 2013, Long Beach, USA
- [10] Kujundzic, G.; Ileš, Š.; Matuško, J.; Vašak, M.: Optimal charging of valve-regulated lead-acid batteries based on model predictive control, *Applied Energy*, 2017 **187** 189–202 DOI: 10.1016/j.apenergy.2016.10.139
- [11] Bhangu B.S.; Bentley, P.; Stone, D.A.; Bingham C.M.: Observer techniques for estimating the state-of-charge and state-of-health of VRLABs for hybrid electric vehicles, *Vehicle Power and Propulsion 2005 IEEE Conference*, 2005 DOI: 10.1109/VPPC.2005.1554646
- [12] Wang, Y.; Fang, H.; Zhou, L.; Wada, T.: Revisiting the state-of-charge estimation for lithium-ion batteries, *IEEE Control Systems Magazine*, 2017 **37**(4), 73–96 DOI: 10.1109/MCS.2017.2696761

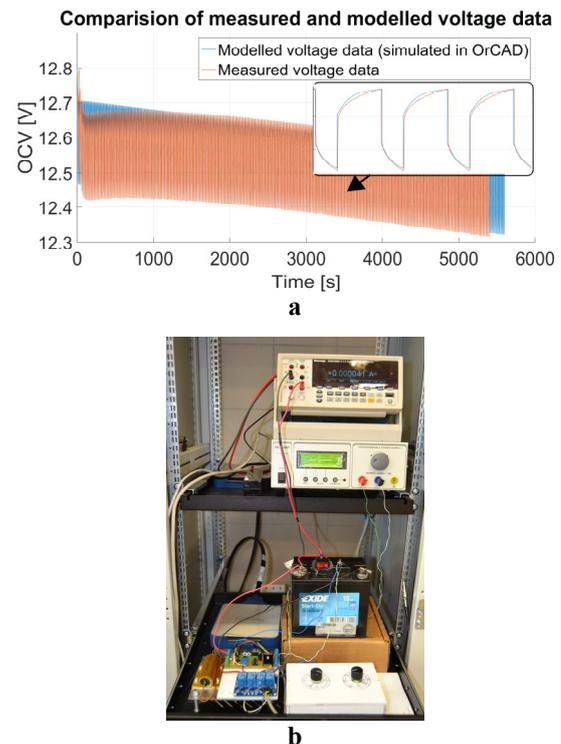


Figure 4. (a) Validation of modelled (OrCAD) and measured voltage data in Matlab environment (b) Battery testbench used during the analysis

SIMULATING ION TRANSPORT WITH THE NP+LEMC METHOD. APPLICATIONS TO ION CHANNELS AND NANOPORES.

DÁVID FERTIG^{*1}, ESZTER MÁDAI¹, MÓNICA VALISKÓ¹, AND DEZSŐ BODA^{1,2}

¹Department of Physical Chemistry, University of Pannonia, Egyetem u. 10., Veszprém, H-8200, HUNGARY

²Institute of Advanced Studies Kőszeg (iASK), Chernel u. 14., Kőszeg, H-9730, HUNGARY

We describe a hybrid simulation technique that uses the Nernst-Planck (NP) transport equation to compute steady-state ionic flux in a non-equilibrium system and uses the Local Equilibrium Monte Carlo (LEMC) simulation technique to establish the statistical mechanical relation between the two crucial functions present in the NP equation: the concentration and the electrochemical potential profiles (Boda, D., Gillespie, D., *J. Chem. Theor. Comput.*, 2012 **8**(3), 824–829). The LEMC method is an adaptation of the Grand Canonical Monte Carlo method to a non-equilibrium situation. We apply the resulting NP+LEMC method to ionic systems, where two reservoirs of electrolytes are separated by a membrane that allows the diffusion of ions through a nanopore. The nanopore can be natural (as the calcium selective Ryanodine Receptor ion channel) or synthetic (as a rectifying bipolar nanopore). We show results for these two systems and demonstrate the effectiveness of the NP+LEMC technique.

Keywords: ion transport, Nernst-Planck, Monte Carlo simulation, nanopore, ion channel

1. Introduction

We dedicate this paper to honoring Professor János Liszi, who was the supervisor of one of the authors (DB) and respected senior to another (MV), supporting to the careers of both. We (DB and VM) publish this paper together with our students (DF and EM) to demonstrate that the lesson of Professor Liszi — one of the most important goals of senior scientists is to pave the way to the junior ones — has not been left unconsidered.

The goal of this work is to present a computer simulation technique for computing steady-state ion transport that was developed and applied in our research group with the essential help of Dirk Gillespie [1]. The method is based on the Nernst-Planck (NP) transport equation coupled to the Local Equilibrium Monte Carlo (LEMC) simulation technique and is described in *Sec.(2)* in detail.

The method, called NP+LEMC, was applied for various problems in the last couple of years. These problems include particle transport through model membranes [1, 2], ion channels [3–5], and nanopores [6–8]. The basic goal is to have a computationally efficient simulation technique using reduced models of steady-state systems, where particles (ions being the focus) diffuse as a result of a maintained driving force that is a concentration gradient and/or an electrical potential gradient (that is, an electrochemical potential gradient).

We tend to regard these kind of systems as nanode-

vices that yield a macroscopic output signal (electrical current, for example) as a response to an input signal (voltage, for example). Reduced models have the advantage of including those degrees of freedom of the many-particle system that are essential in reproducing device behavior, namely, the relationship of the output and input signals [6].

We present two case studies here to demonstrate both the effectiveness (and also possible source of errors) of the method and the power of reduced models. In the Ryanodine Receptor (RyR) calcium release ion channel, we needed to model only those amino acids of the channel protein that are inside or near the selectivity filter and hang into the ionic pathway. In the bipolar nanopore, we needed to reproduce the variation of the electric field along the pore axis properly. The most important reduction in the degrees of freedom, however, is that we model water as a dielectric continuum. The effect of using implicit solvent instead of an explicit solvent model was discussed in our previous work [6] through comparisons to molecular dynamics (MD) simulations.

Our computational method is not the only one that is able to determine ionic current through biological and synthetic nanopores using reduced models. While the Brownian Dynamics (BD) simulation method [9–14] is the obvious candidate to simulate this problem, it has disadvantages from the point of view of sampling the flux of ions. The Poisson-Nernst-Planck (PNP) theory [11, 15–18] also uses the NP equation to compute flux, but uses a mean-field approach, the Poisson-Boltzmann

*Correspondence: fertig.david92@gmail.com

(PB) theory, to describe the statistical mechanical relationship between the concentration and electrochemical potential profiles.

To use something more state-of-the-art to provide this relationship beyond the PB level is the essence of our approach. We suggest the LEMC technique, which is a particle simulation method based on a three-dimensional model producing all the ionic correlations missing from PB. Gillespie et al. proposed a different technique, in which a density functional theory (DFT) was coupled to the NP equation [19, 20] and used to method (NP+DFT) to describe the behavior of the same ion channel that we study here, the RyR ion channel [21–23]. DFT is a continuum theory, but a very sophisticated one, where ionic correlations are included through a series expansion of the free energy functional and the finite size of ions is taken into account through the fundamental measure theory [24]. Its accuracy was demonstrated by computing electrical double layers in extreme conditions and comparing to Monte Carlo (MC) simulations [25, 26]. The disadvantage of this method is that it can be used efficiently only in one dimension. This reduction in dimensionality, however, is often feasible provided that we reduce the model in an intelligent way.

2. Modeling and methodology

A good model should be able to explain complex phenomena in a simplified way while it stays in agreement with experimental data. The nanopore systems (both natural and synthetic) studied in this work have some common features. The primitive model of electrolytes, that represents ions as charged hard spheres, is used. The interaction potential between two ions is defined by Coulomb's law in a dielectric material:

$$u_{ij}(r) = \begin{cases} \infty & \text{if } r < R_i + R_j \\ \frac{1}{4\pi\epsilon_0\epsilon} \frac{q_i q_j}{r} & \text{if } r \geq R_i + R_j \end{cases} \quad (1)$$

where R_i and R_j are the radii, q_i and q_j are the charges of the different ionic species, ϵ_0 is the permittivity of vacuum, $\epsilon = 78.5$ is the relative permittivity of the solvent, and r is the distance between two ions. Solvent (water, in this work) is not modeled explicitly; it is rather represented by two response functions. Energetically, solvent acts as a dielectric background that screens the electric field of ions by just dividing by ϵ in the Coulomb-potential (Eq.(1)). Dynamically, solvent molecules collide with ions and impede their diffusion; we take that effect into account by a diffusion coefficient function, $D_i(\mathbf{r})$.

The ions of the electrolyte diffuse through a pore between two bulk containers (see Fig.1). The pore penetrates a membrane that separates the two containers. The pore and the membrane are also modeled in a reduced way; they are confined by hard walls with which the hard sphere ion cannot overlap. Other details of the

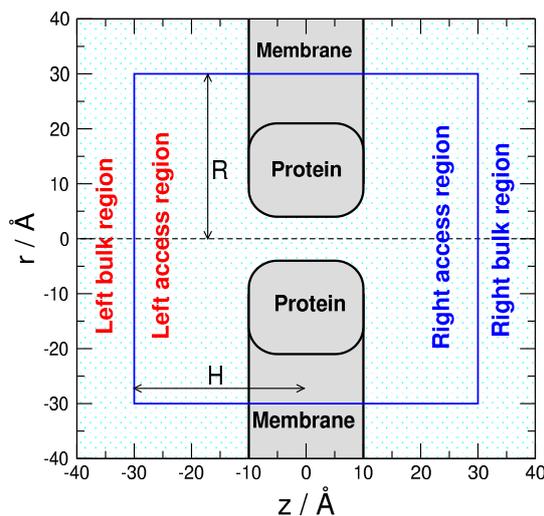


Figure 1: Simulation cell used in the application of the NP+LEMC method. The model is rotationally symmetric: the three-dimensional model is obtained by rotating the figure about the z -axis. The domain confined by the blue line is the non-equilibrium transport region (the solution domain of the NP+LEMC system) that includes the pore and the two access regions. The electrochemical potential in this region is not constant, but changes from one constant value to another (the values in the two bulk regions) in a monotonic manner to provide the driving force of ionic transport. Parameters R and H characterize system size.

pores (amino acid side chains and surface charges) will be presented in Sec.(3) for the respective ion channel and nanopore systems. Ion transport is steady-state; concentrations and electrical potentials, therefore, are kept fixed at the boundaries of the two baths on the two sides of the membrane (blue line in Fig.1).

We assume that ion transport is described by drift-diffusion. Our procedure is a hybrid method that separates the configuration degrees of freedom of ions (ion positions) from the kinetic degrees of freedom (ion velocities). In BD simulations they are treated together and ionic current is measured by counting ions that pass through the pore. The BD method has the disadvantage of weak sampling of flux especially at low concentrations (as in our ion channel example) and in cases, when current is limited by ionic depletion zones (as in our bipolar nanopore example).

Separation of the two parts of the phase space is a usual practice in equilibrium statistical mechanics, where the kinetic part is described by the ideal gas equations, while the excess quantities that produce all the peculiarities beyond ideal-gas behavior are from the potential energy, that, in turn, depends only on the configuration coordinates [27–30]. Out of equilibrium, however, this separation is not so obvious. Classical mechanics is still valid, so particles' motion can be described by Newton's equations of motion (MD simulations can be used). In the con-

tinuum solvent framework, Langevin's equation is used in BD simulations.

The meaning of various thermodynamic quantities, however, loses the solid ground that it has in equilibrium. There is no well-established non-equilibrium statistical mechanics finding its way into textbooks despite the fact that many authors made considerable advances in this field [31–36]. The most problematic quantities are those containing entropic effects, such as the chemical potential, μ_i . In the case of charged particles, we use the term electrochemical potential, but we talk about the same thing. This is a crucial quantity, because its homogeneity defines thermodynamic equilibrium (let us assume that the other two intensive parameters, temperature and pressure, are constant). If μ_i is not constant, species i will diffuse until it becomes constant according to the second law of thermodynamics. The gradient of $\mu_i(\mathbf{r})$, therefore, is the driving force of particle diffusion.

The empirical transport equation that is most widely used to describe transport of charged particles (electrodifusion) is the NP equation,

$$\mathbf{j}_i(\mathbf{r}) = -\frac{1}{kT} D_i(\mathbf{r}) c_i(\mathbf{r}) \nabla \mu_i(\mathbf{r}), \quad (2)$$

where $D_i(\mathbf{r})$ is the diffusion coefficient profile of ionic species i , $c_i(\mathbf{r})$ is the concentration profile, $\mu_i(\mathbf{r})$ is the electrochemical potential profile, $\mathbf{j}_i(\mathbf{r})$ is the particle flux density, k is Boltzmann's constant and $T = 298.15K$ is the temperature. The resulting $\mathbf{j}_i(\mathbf{r})$ must satisfy the continuity equation:

$$\nabla \cdot \mathbf{j}_i(\mathbf{r}) = 0. \quad (3)$$

The NP equation separates the configuration and kinetic parts of the phase space in a way that it provides flux as a function of three profiles that, in turn, depend only on configuration coordinates. Thus, we reduced our statistical mechanical task to averaging over states in the configuration space just as we did in the case of equilibrium statistical mechanics. If we treat $D_i(\mathbf{r})$ as a parameter, the task is reduced to finding the proper relation between $c_i(\mathbf{r})$ and $\mu_i(\mathbf{r})$. Defining local concentration in a non-equilibrium situation is the same as in equilibrium: we compute the average number of particles in a small volume element and divide it with the volume of that element (though we will also compute concentration in a different way on the basis of the Potential Distribution Theorem [37], see later).

Electrochemical potential, on the other hand, is more problematic. The trick in this case is that we assume local equilibrium (LE). If we divide the simulation cell into subvolumes B^α , we can characterize this subvolume with the electrochemical potential, μ_i^α , and assume that this value is constant in B^α . We designate the μ_i^α and c_i^α values to the mass centers of the B^α volume elements. The assumption of LE is also present in other methods (though not necessarily stated explicitly), where the $c_i(\mathbf{r})$ vs. $\mu_i(\mathbf{r})$ relationship is considered, such as in

the PNP theory or in the NP+DFT method of Gillespie et al. [19, 20].

The main proposal of our technique was to use an MC method for this purpose. We assumed that the subvolumes are open systems and they are in LE with a constant volume (V^α), temperature (T), and electrochemical potential (μ_i^α). Thus, we suggested to use Grand Canonical Monte Carlo (GCMC) simulations, where particle insertion/deletions are attempted in the simulation cell. The acceptance probability depends on which subvolume we try to insert to (or where is the particle that we try to delete): $\min(1; p_{i,\chi}^\alpha(\mathbf{r}))$, where

$$p_{i,\chi}^\alpha(\mathbf{r}) = \frac{N_i^\alpha!(V^\alpha)^\chi}{(N_i^\alpha + \chi)!} \exp\left(-\frac{\Delta U(\mathbf{r}) - \chi \mu_i^\alpha}{kT}\right). \quad (4)$$

Here, N_i^α is the number of ions of type i in subvolume B^α before insertion/deletion, $\Delta U(\mathbf{r})$ is the change of the system's potential energy during particle insertion to position \mathbf{r} (or deletion from there), $\chi = 1$ for insertion, and $\chi = -1$ for deletion. The difference between this method and equilibrium GCMC is that μ_i is space-dependent and the acceptance criterion is referred to a given subvolume instead of the whole simulation cell. The effect of the surrounding of subvolume B^α , however, is taken into account in the simulation through the energy change that includes all the interactions from other subvolumes, not only interactions between ions in B^α .

Although it is tempting to view the subvolume B^α as a distinct thermodynamic system with its own ensemble of states and to include the effect of other subvolumes as an external constraint, this is not the case. The ensemble of states belongs to the whole system, because ion configurations in subvolume B^α should be collected for every possible ion configurations of all the other subvolumes. Therefore, the independent variables of this ensemble are T and $\{V^\alpha, \mu_i^\alpha\}$, where α and i run over the volume elements and particle species, respectively. For comparison, the variables in global equilibrium are T , V , and μ_i , where V is the total volume and μ_i does not depend on space.

We solve the NP+LEMC system in an iterative way. The electrochemical potential is adjusted until conservation of mass ($\nabla \cdot \mathbf{j}_i(\mathbf{r}) = 0$) is satisfied. The procedure can be summarized as

$$\mu_i^\alpha[n] \xrightarrow{\text{LEMC}} c_i^\alpha[n] \xrightarrow{\text{NP}} \mathbf{j}_i^\alpha[n] \xrightarrow{\nabla \cdot \mathbf{j}=0} \mu_i^\alpha[n+1]. \quad (5)$$

The electrochemical potentials for the next iteration, $\mu_i^\alpha[n+1]$, are computed from the results of the previous iteration, $c_i^\alpha[n]$, on the basis of the divergence-theorem (also known as Gauss-Ostrogradsky's theorem). The continuity equation is converted to a surface integral:

$$0 = \int_{B^\alpha} \nabla \cdot \mathbf{j}_i(\mathbf{r}) dV = \oint_{S^\alpha} \mathbf{j}_i(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) da, \quad (6)$$

where volume B^α is bounded by surface S^α and $\mathbf{n}(\mathbf{r})$ denotes the normal vector pointing outward at position

r of the surface. Every \mathcal{S}^α surface is divided into $\mathcal{S}^{\alpha\beta}$ elements. Along these elements, \mathcal{B}^α and \mathcal{B}^β are adjacent cells. It is assumed that the concentration, the gradient of the electrochemical potential, the flux density and the diffusion coefficient are constant on a surface $\mathcal{S}^{\alpha\beta}$. They are denoted by hat: $\hat{c}_i^{\alpha\beta}$, $\nabla\hat{\mu}_i^{\alpha\beta}$, $\hat{\mathbf{j}}_i^{\alpha\beta}$, and $\hat{D}_i^{\alpha\beta}$. The $\hat{c}_i^{\alpha\beta}$ values are obtained from the values c_i^α and c_i^β via linear interpolation. The $\nabla\hat{\mu}_i^{\alpha\beta}$ values are also obtained from μ_i^α and μ_i^β assuming linearity.

Thus the integral in Eq.(6) for a given surface \mathcal{S}^α is replaced by a sum over the surface elements that constitute \mathcal{S}^α :

$$0 = \sum_{\beta, \mathcal{S}^{\alpha\beta} \in \mathcal{S}^\alpha} \hat{\mathbf{j}}_i^{\alpha\beta} \cdot \mathbf{n}^{\alpha\beta} a^{\alpha\beta}, \quad (7)$$

where $a^{\alpha\beta}$ is the area of surface element $\mathcal{S}^{\alpha\beta}$ and $\mathbf{n}^{\alpha\beta}$ is the outward normal vector in the center of $\mathcal{S}^{\alpha\beta}$. The iteration procedure is described by the following steps:

1. An appropriately chosen initial set of electrochemical potentials is chosen ($\mu_i^\alpha[1]$; in general: $\mu_i^\alpha[n]$, where $[n]$ denotes the n th iteration).
2. Using these $\mu_i^\alpha[n]$ parameters as inputs, LEMC simulations are performed. The resulting concentrations are denoted by $c_i^\alpha[n]$.
3. The flux computed from the $\{\mu_i^\alpha[n], c_i^\alpha[n]\}$ pair usually does not satisfy Eq.(7). The next set of electrochemical potential is calculated by assuming that the $\{\mu_i^\alpha[n+1], c_i^\alpha[n]\}$ pair *does* satisfy Eq.(7). If we write the value of $\hat{\mathbf{j}}_i^{\alpha\beta}$ as given by the NP equation into Eq.(7), we obtain

$$0 = \sum_{\beta} \hat{D}_i^{\alpha\beta} \hat{c}_i^{\alpha\beta}[n] \nabla \hat{\mu}_i^{\alpha\beta}[n+1] \cdot \mathbf{n}^{\alpha\beta} a^{\alpha\beta}, \quad (8)$$

where β runs over all the surface elements $\mathcal{S}^{\alpha\beta}$ that constitute \mathcal{S}^α . Eq.(8) is a system of linear equations; the unknown variables are denoted by $\mu_i^{\alpha, \text{CAL}}[n+1]$, where CAL refers to the fact that these values come from calculations by solving Eq.(8).

4. To achieve faster and more robust convergence in the case of large driving forces, the electrochemical potential used in the $(n+1)$ th iteration is mixed from the values calculated in the $(n+1)$ th iteration from Eq.(8) and the values mixed in the n th iteration:

$$\begin{aligned} \mu_i^{\alpha, \text{MIX}}[n+1] &= b_i \mu_i^{\alpha, \text{CAL}}[n+1] \\ &+ (1 - b_i) \mu_i^{\alpha, \text{MIX}}[n], \end{aligned} \quad (9)$$

where b_i is a mixing parameter that determines the ratio of mixing. If the parameter is close to 1, faster iteration can be achieved, however, it may result in the system fluctuating between local minima. Smaller b_i values prevent this fluctuation at the price of making the convergence slower.

5. The input of the $(n+1)$ th LEMC simulation is $\mu_i^{\alpha, \text{MIX}}[n+1]$.

The system of linear equations contains the boundary conditions for the electrochemical potential and concentration via the boundary elements that are at the system's boundaries (blue line in Fig.1). If the $\mathcal{S}^{\alpha\beta}$ face is on the system's boundary, the values of $\hat{c}_i^{\alpha\beta}$ and $\hat{\mu}_i^{\alpha\beta}$ are those prescribed on that face. The electrochemical potentials on the left (L) and right (R) boundaries are computed from

$$\mu_i^{\text{L}} = kT \ln c_i^{\text{L}} + \mu_i^{\text{EX, L}} + q_i \Phi^{\text{L}} \quad (10)$$

and

$$\mu_i^{\text{R}} = kT \ln c_i^{\text{R}} + \mu_i^{\text{EX, R}} + q_i \Phi^{\text{R}}, \quad (11)$$

where $\mu_i^{\text{EX, L}}$ and $\mu_i^{\text{EX, R}}$ are the excess chemical potentials in the absence of an external field (determined by the Adaptive GCMC method of Malasics et al. [38]), while $q_i \Phi^{\text{L}}$ and $q_i \Phi^{\text{R}}$ are the interactions with the applied electrical potentials. Prescribing Φ^{L} and Φ^{R} on the system's boundary means that we use an electrostatic Dirichlet boundary condition. Voltage is defined as $U = \Phi^{\text{L}} - \Phi^{\text{R}}$. Ultimately, we have boundary conditions for ion concentrations on the two sides of the membrane and for the voltage (ground is on the right in this study).

The energy change ΔU contains not only the interactions between particles (and interactions of particles with the pore), but also the interaction with an external electrical potential, $\Phi^{\text{appl}}(\mathbf{r})$. This applied potential is calculated by solving Laplace's equation

$$\nabla^2 \Phi^{\text{appl}}(\mathbf{r}) = 0 \quad (12)$$

for the system (inside the blue line) with the prescribed Dirichlet boundary condition on the system's boundaries (Φ^{L} and Φ^{R} on the left and right blue line, respectively). In the portion of the blue line inside the membrane, we interpolate between Φ^{L} and Φ^{R} . We solve this equation with the Induced Charge Computation method [39, 40].

In the present geometry, the elementary cells are $\Delta z \times \Delta r$ rectangles in the (z, r) plane. In three-dimensional space, these correspond to concentric rings with the z -axis in their centers.

The concentration in an elementary cell \mathcal{B}^α can be computed from dividing the average number of ions in the cell with the volume of the cell. This route is disadvantageous if ion concentrations are small. An alternative method is based on the Potential Distribution Theorem [37] that practically corresponds to the Widom particle insertion method [41, 42]. The total excess chemical potential (containing also the interaction with the applied field) is $\mu_i^\alpha - kT \ln c_i^\alpha$ and can be computed as

$$\exp[-(\mu_i^\alpha/kT - \ln c_i^\alpha)] = \langle \exp[-\Delta U_i^\alpha/kT] \rangle, \quad (13)$$

where ΔU_i^α is the energy change associated with the insertion of a test particle of species i into a randomly chosen position in the elementary cell α . This is the same energy computed in the particle insertion step of the LEMC technique, therefore, it does not require additional computational cost. The concentration can be calculated from

Eq.(13) because μ_i^α is known (that is the input of the simulation) and the ensemble average on the right hand side can be obtained from the LEMC insertions. This route provides a more accurate value for c_i^α , because this sampling is not discrete as just counting particles, but rather continuous, because we always gain information from the simulation at every ion insertion via the energy ΔU_i^α . The route of counting particles, however, might be better, when concentrations are very large in the volume element. In the case of long enough simulations, the two methods give identical results (within a statistical error).

Because of this statistical error, the iteration does not converge to an exact value of the ionic flux density, but it fluctuates around a limiting value. The final solution is obtained from a running average over iterations. Longer LEMC simulations and more iterations result in a more reliable outcome. The net flux of the diffusing ions through the cross section of the pore is calculated via averaging:

$$J_i(z) = 2\pi \int_0^{R(z)} r \mathbf{j}_i(z, r) \cdot \mathbf{n}_z dr \quad (14)$$

where $R(z)$ denotes the radius of the pore at coordinate z ($|z| < H_{\text{memb}}/2$, where H_{memb} is the thickness of the membrane) and \mathbf{n}_z is the unit vector along the z -axis. The electrical current of an ion is then

$$I_i = q_i J_i, \quad (15)$$

and the total current is

$$I = \sum_i q_i J_i. \quad (16)$$

3. Results and discussion

3.1 Calcium release channel

There are two important classes of calcium channels that have been our focus. The L-type calcium channel is found in excitable cell membranes [43] such as those of nerve cells and muscle cells (both cardiac and skeletal). These are strongly Ca^{2+} selective channels, meaning that the presence of only a μM Ca^{2+} in the bulk decreases the Na^+ current to half of its value compared to its value in the absence of Ca^{2+} . This is called micromolar Ca^{2+} affinity (or Ca^{2+} block) because a very small amount of Ca^{2+} is enough to block the channel [44].

The price of high Ca^{2+} selectivity is low Ca^{2+} current through this channel. The Ca^{2+} ions flow into the muscle cell when the L-type calcium channels open in response to an electrical signal (action potential) in a small quantity that is not sufficient to initiate muscle contraction. The solution of Nature for this problem is an amplification mechanism.

A large amount of Ca^{2+} ions are stored in organelles, the sarcoplasmic reticulum (SR), that are situated inside the muscle cells. There are calcium release channels (also

known as RyR calcium channels) embedded in the membrane of the SR that are opened by the Ca^{2+} ions provided by the L-type channel (in cardiac muscle) or via a direct link between the two types of calcium channels (in skeletal muscle). The RyR channels provide the large Ca^{2+} flux that is necessary for muscle contraction by binding to tropomyosins on the actin filaments and allowing myosin to climb up the filament.

These channels are wider (also, less Ca^{2+} selective) than the L-type channels. A large amount of experimental data is available for this channel [45–47]. On the basis of this database, Gillespie et al. developed a model for the RyR channel [21,22]. Although the model was a one-dimensional reduction, studied with a one-dimensional NP+DFT, they were able to reproduce hundreds of experimental current-voltage curves.

Later, when the NP+LEMC method became available, we developed [4] a three-dimensional version of the model of Gillespie et al. The model is similar to that used by Boda et al. [3, 4, 40, 48–54] for the L-type calcium channel inspired by the “space-charge competition mechanism” proposed by Nonner et al. [55]. From the point of view of ion selectivity and permeation, the important part of the channel is the selectivity filter, the bottleneck of the pore. This region is lined by four P-loops that are parts of the four membrane-spanning subunits. These loops contribute four glutamic acids (in the L-type calcium channel) to the filter-lining region (the EEEE locus). The COO^- groups of the carboxyl side chains are thought to reach into the pore lumen and interact with the passing ions [44]. In the case of the RyR channel, they are four aspartates (D4899, see Fig.2). Gillespie et al. [21–23, 46, 47] identified additional E and D amino acids in the two vestibules on the cytosolic and luminal sides (Fig.2).

It has been shown [3, 4, 40, 48–55] that the strong Ca^{2+} selectivity can be reproduced by a reduced model where eight half-charged oxygen atoms ($\text{O}^{1/2-}$) of the four COO^- groups force a strict competition between Ca^{2+} and Na^+ ions for space in the crowded filter. The exact positions of the $\text{O}^{1/2-}$ ions are not known and even irrelevant from the point of view of the selectivity of the model. What matters is that they attract the cations into the filter with their charge, and, at the same time, they exert a hard sphere exclusion and make it difficult for the cations to find space in the filter. Therefore, it proved to be sufficient to model the structural groups of the filter as mobile $\text{O}^{1/2-}$ ions that are confined to the filter region. The profiles shown in Fig.2 are results of the simulations and show how the eight $\text{O}^{1/2-}$ ions distribute in a given segment of the channel. There are 32 such $\text{O}^{1/2-}$ ions, plus point charges placed on a ring on the luminal side. The geometrical features (filter radius is 5.5 Å, filter length is 15 Å, and radii of left and right vestibules are 22 and 9 Å, respectively) are also designed on the basis of Gillespie’s model that, in turn, are based on fitting to experimental data.

This model was designed and used unchanged in ev-

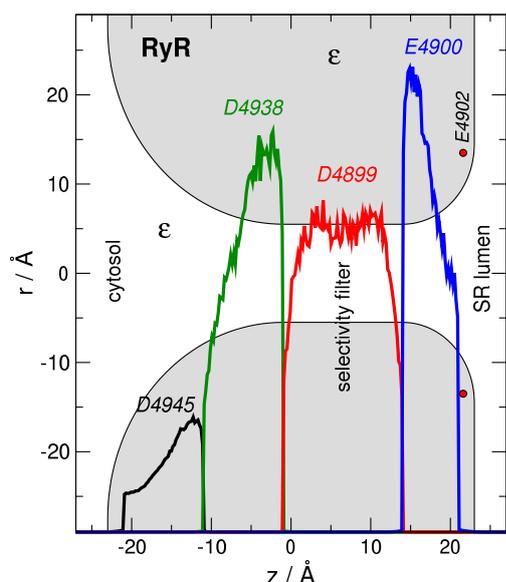


Figure 2: The three-dimensional model [4] of the RyR calcium channel based on the model of Gillespie et al. [21,22]. Each curve shows the distribution of eight $O^{1/2-}$ ions that are confined to the given region but otherwise free to move there. The charges of the E4902 amino acids are represented as point charges ($-1/2e$) at fixed positions on a ring. The thickness of the membrane is $H_{\text{memb}} = 46 \text{ \AA}$. The dielectric constant is $\epsilon = 78.5$.

ery calculation. What remains to be specified are the diffusion coefficient profiles of the various ionic species. These profiles depend only on z , $D_i(z)$; we assume no variation in the radial dimension. The values outside the channel and in the selectivity filter, D_i^B and D_i^F , respectively, are constant. In the vestibules at the two entrances of the channel, the diffusion constant profiles are interpolated between the D_i^B and D_i^F values in a way that their value changes in proportion with the cross section of the channel.

The bulk values, D_i^B , are taken from experiments; they are $1.334 \cdot 10^{-9}$, $7.92 \cdot 10^{-10}$, and $2.032 \cdot 10^{-9} \text{ m}^2\text{s}^{-1}$ for Na^+ , Ca^{2+} , and Cl^- , respectively. The values inside the cylindrical selectivity filter are fitted to two experimental data points: 250 mM symmetric NaCl at 100 mV for Na^+ , while added 10 mM luminal CaCl_2 at -100 mV for Ca^{2+} . The resulting values ($1.27 \cdot 10^{-10}$ and $1.27 \cdot 10^{-11}$ for Na^+ and Ca^{2+} , respectively) are fixed and never changed. The Cl^- value is irrelevant, because Cl^- ions do not carry significant current. These values have been obtained for a moderate system size ($H = 54 \text{ \AA}$, $R = 48 \text{ \AA}$). Currents, and, therefore, diffusion coefficients fitted to currents can depend on system size (see later).

Here, we show results for a NaCl- CaCl_2 mixture, where there is 250 mM Na^+ on both sides of the membrane, while there is 4 μM Ca^{2+} on the cytosolic (left) and 50 mM Ca^{2+} on the luminal (right) side. The voltage is changed from -150 mV to 150 mV with the ground on the right. The current vs. voltage curves are shown

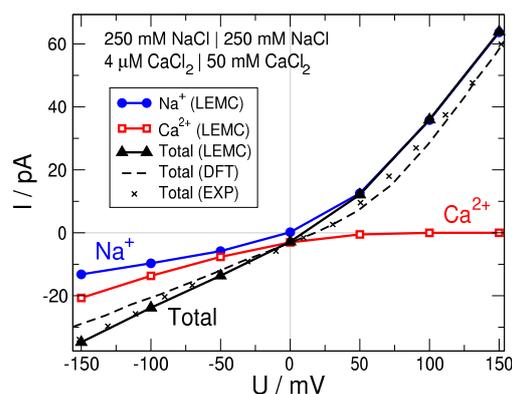


Figure 3: Currents vs. voltage curves as obtained from experiment [21], the model of Gillespie et al. [21, 22] obtained from DFT coupled to the NP equation, and from the NP+LEMC method. In the case of the NP+LEMC method, the currents carried by Na^+ and Ca^{2+} are also shown.

in Fig.3. Total currents are shown in black as obtained from experiments (\times symbols), Gillespie's NP+DFT calculations (dashed line), and our NP+LEMC calculations (solid line with filled triangles).

Agreement with experiment is very good using both models. The slope of the $I - U$ curve is larger for positive voltages than for negative voltages. More details and understanding can be gained from the current curves for the separate ions, Na^+ and Ca^{2+} (blue and red curves, respectively). At positive voltages, the Ca^{2+} current is practically zero, so the total current is carried by Na^+ ions. At negative voltages, on the other hand, both Na^+ and Ca^{2+} ions contribute to the current.

The explanation of this behavior can be drawn from concentration and electrochemical potential profiles shown in Fig.4. Let us discuss the Na^+ -profiles first (blue curves). The driving force for the Na^+ ions is the voltage because Na^+ concentration is the same on the two sides. The difference between Na^+ currents at -100 and 100 mV voltages, therefore, is the result of the different competition between Na^+ and Ca^{2+} ions at the two voltages. To understand the difference in this competition, we need to understand the behavior of Ca^{2+} ions.

At 100 mV, the driving force for Ca^{2+} ions is small because the concentration difference balances the electrical potential difference (see the Ca^{2+} electrochemical potential profile in the right panel of Fig.4B). This results in a small Ca^{2+} current. Ca^{2+} concentration is small in the left bulk, therefore, a Ca^{2+} depletion zone is formed in the left vestibule and in the selectivity filter of the channel (beware the logarithmic scale of the concentration axis). That depletion zone results in a decreased concentration of Ca^{2+} ions compared to Na^+ ions inside the channel.

In the case of -100 mV, on the other hand, the Ca^{2+} depletion zone in the left vestibule is absent because Ca^{2+} ions arrive from the right and "fill up" the left

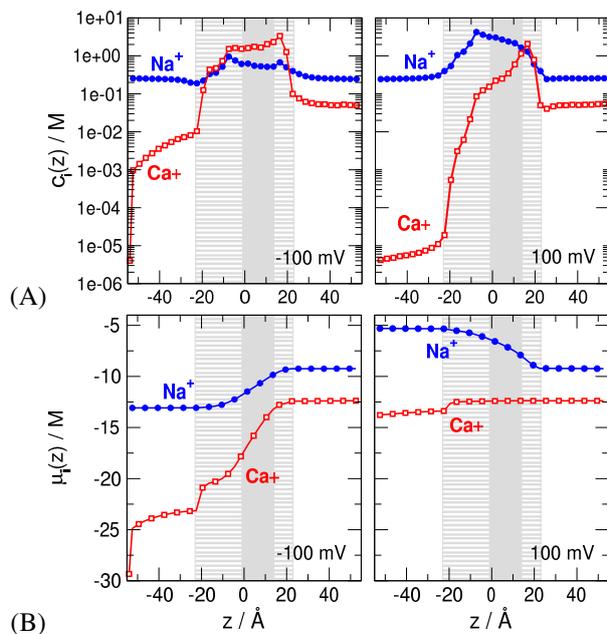


Figure 4: (A) Concentration and (B) electrochemical potential profiles of Na^+ and Ca^{2+} for two opposing voltages: -100 mV (left panels) and 100 mV (right panels). The striped parts indicate the cytosolic and luminal vestibules, while the gray area is the selectivity filter.

vestibule. There is a more balanced competition between Ca^{2+} and Na^+ ions in the channel and Ca^{2+} ions can use the free-energetic advantage that they have over Na^+ [22, 53]. This means that Ca^+ ions are favored by the crowded selectivity filter, because they provide twice the charge (compared to Na^+ ions) to balance the charge of $\text{O}^{1/2-}$ ions while occupying about the same space (their diameters are similar: 1.98 and 1.9 \AA for Ca^{2+} and Na^+ , respectively). The finite size of the ions plays a crucial role in the selectivity mechanism. This kind of selectivity could not be produced with the PNP theory.

Because Ca^{2+} concentration is large in the left vestibule of the channel, it drops quickly to the $4 \mu\text{M}$ value at the left boundary of the solution domain. This introduces a severe system-size dependence into the calculations in this case that is analyzed in Fig.5. Small Ca^{2+} concentration on the left hand side corresponds to a large Debye-length in the double layer at left hand side of the membrane (in the left access region). That double layer should fit into the simulation cell (as it does in the case of a larger cell, see black curves in Fig.5).

If the cell is too small (see red curves in Fig.5), there is not enough space for the Ca^{2+} concentration to reach the $4 \mu\text{M}$ limiting value at the left boundary of the cell. The electrochemical potential cannot reach its limiting value either (see the bottom panel). The NP+LEMC calculation, however, provides a solution in this case too, because the layer near the left outer boundary of the cell takes care of the missing access region in an averaged manner. The concentration has a small value in the layer,

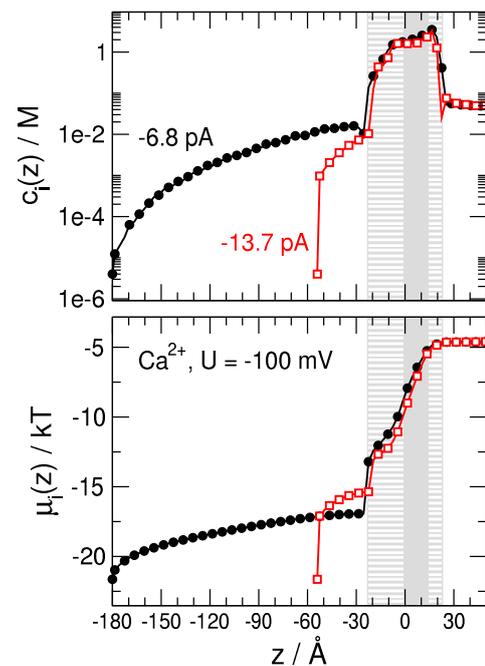


Figure 5: Concentration (top panel) and electrochemical potential (bottom panel) profiles of Ca^{2+} for -100 mV. The results of simulations for two system sizes are shown: $H = 54 \text{ \AA}$ (red) and $H = 180 \text{ \AA}$ (black).

while the electrochemical potential profile drops abruptly (large driving force). An appropriate value of the $c_i \nabla \mu_i$ product, therefore, is provided by the self-consistent solution so that the continuity equation is satisfied.

This solution, however, is approximate. A large portion of the access region with considerable resistance is taken into account in an averaged way by the layer near the system edge. This introduces an error into the value of the Ca^{2+} current that is indicated in Fig.5 for both cases. The good behavior of the current-voltage curves compared to experiments and DFT calculations is due to the fit of the Ca^{2+} diffusion coefficient, $D_{\text{Ca}^{2+}}^F$, inside the filter. That value was fitted for negative voltage at the given system size balancing the system-size error.

This result points out the importance of system size in the case of small ionic concentrations, but it also shows the role of the diffusion coefficient profile in NP+LEMC calculations. In confined geometries, the $D_i(\mathbf{r})$ profile, although it has a strong relation to the mobility of ions, is primarily an adjustable parameter that is fitted to experiments (as in the present case) or to MD results [6]. The value of D_i^F takes into account interactions that are absent in the reduced model or accounts for resistances of regions that are absent in the model.

In the case of the RyR channel, for example, the diffusion coefficient profile includes effects of the parts of the ion channel not included in the model: the real RyR channel is much larger than the 46 \AA portion modeled here. That region also tunes the total current, but it is not selective. The selectivity filter and its close neigh-

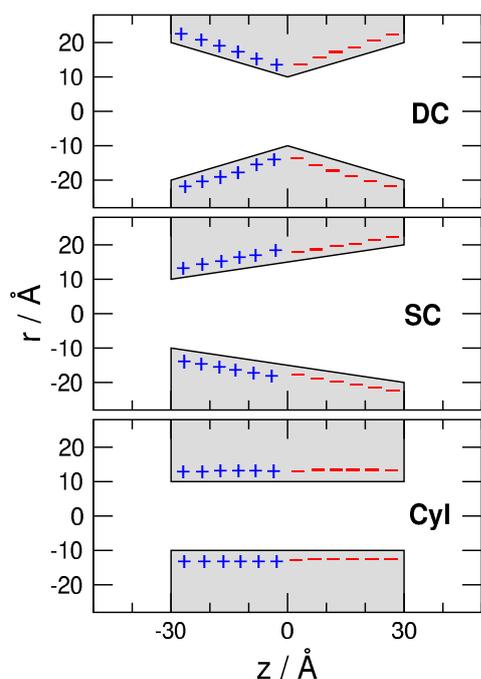


Figure 6: Bipolar nanopore geometries: Cyl: cylindrical; SC: single conical; DC: double conical. The pore is positively charged on the left hand side ($-30 < z < 0$ Å), while negatively charged on the right hand side ($0 < z < 30$ Å) keeping the surface charge fixed ($\sigma = \pm 0.1$ C/m²). Minimal and maximal pore radii are 10 and 20 Å, respectively.

borhood modeled here determines ion selectivity and is able to reproduce complex behavior such as anomalous mole fraction experiments discussed previously by Gillespie [21–23] and Boda [4].

3.2 Rectifying bipolar nanopores

Ion channels are natural nanopores with stable well-defined structures that are very narrow at their selectivity filters (often below 1 nm in radius) to make them suitably selective. The disadvantages, however, are considerable. The structures are often unknown. They are difficult to handle experimentally. Their manipulation is cumbersome with point mutations.

Synthetic nanopores, therefore, quickly gained attention due to the fact that they have special properties compared to those of micropores. These special properties arise because the screening length of the electrolyte is comparable to the radius of the nanopore. This fact gives nanopores properties that resemble those of ion channels.

One advantage of synthetic nanopores is that are relatively easy to fabricate [56–63]. They are either solid state nanopores using ion-beam or electron-beam sculpting in, for example, silicon compound membranes, or they are track etched into polymer membranes. Two basic properties of such nanopores are their geometry (shape) and the pattern of surface charge on the pore wall.

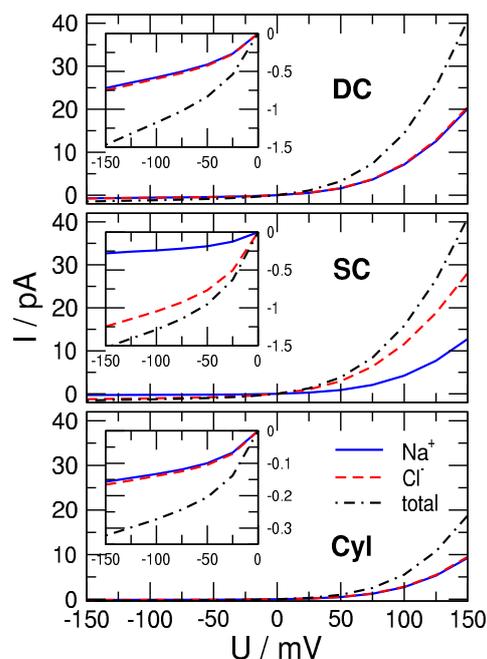


Figure 7: Current-voltage relations for the three nanopore geometries (Cyl, SC, and DC from bottom to top). Currents carried by Na⁺ (solid blue), Cl⁻ (dashed red), and their sum (dot-dashed black) are shown as a function of voltage. The insets magnify the results for negative voltages.

In our previous works, we studied the bipolar nanopore, where the surface charge is positive on the left hand side of the pore, while it is negative on the right hand side [6, 7]. These nanopores rectify ionic current, meaning that they let a much larger amount of ions through at a given sign of voltage than at the opposite sign. In those papers, we used a cylindrical geometry for the nanopore and focused on the effect of the charge pattern, pore radius, and concentration.

Here, we discuss the effect of pore geometry on the rectification properties of a bipolar nanopore. We performed NP+LEMC calculations for three different geometries, the cylindrical (Cyl), single conical (SC), and double conical (DC) shown in Fig.6. Simulations for different voltages have been performed from -150 mV to 150 mV. The concentration of the electrolyte was 0.1 M on both sides. The electrolyte is a 1:1 system (let us call it NaCl), but the diameters of the cations and anions are the same in order to avoid effects from ion size asymmetry. For the same reason, the same diffusion coefficients were used for the two ions.

This makes it possible to focus on the balance of charge and geometrical asymmetries. If the nanopore's shape is symmetric, for example, the cations and anions carry the same amount of current (Cyl and DC).

Figure 7 shows the current-voltage relations. Rectification is observed in all three geometries: current is much larger at positive than at negative voltages. The rectifi-

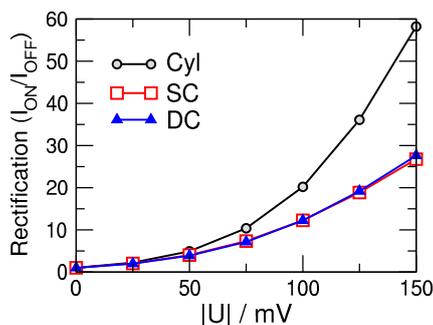


Figure 8: Rectification (defined as the absolute value of the current ratio in the ON and OFF states) as a function of the absolute value of the voltage for the three nanopore geometries.

cation is defined as the ratio of currents at positive (ON state) and negative voltages (OFF state). The results are shown in Fig.8.

The basic explanation of rectification can be depicted from the concentration profiles (Fig.9). There are depletion zones for cations in the positively charged half region (left), while there are depletion zones for anions in the negatively charged half region (right). In the OFF state (negative voltage) these depletion zones are deeper than in the ON state (positive voltage). More detailed explanation have been given in our previous works [6,7]. Here we focus our discussion to the effect of pore shape.

Total currents are larger in the SC and DC geometries due to their wider entrances. Interestingly, total currents are the same in the SC and DC geometries despite the quite different pore shapes. Whether this is a coincidence or it has a deeper explanation requires further investigation.

What is different in the SC and DC geometries is the partitioning of the total current between Na^+ and Cl^- . While Na^+ and Cl^- currents are the same in the DC geometry due to the symmetric shape, they are different in the SC geometry. The current carried by Na^+ ions is smaller than the current carried by Cl^- ions (middle panel of Fig.7). This is reflected in concentration profiles (see middle panel of Fig.9). The explanation is that the tip of the SC nanopore (left entrance) is positively charged so the depletion zones of Na^+ ions there is deeper than the depletion zone of Cl^- ions on the other side where the pore is wider. Deeper depletion zones result in smaller currents. Therefore, Na^+ current is smaller in the whole voltage range, but more so in the OFF state at negative voltages.

Rectification is largest in the Cyl geometry because the depletion zones are the deepest in that geometry for both ions. The deepest points of the depletion zones are formed around $|z| = 10 \text{ \AA}$. There are different effects that form the concentration profiles inside the pore. In the left region, for example, (1) the positive surface charge on the pore wall repels Na^+ ions, (2) the negative surface charge

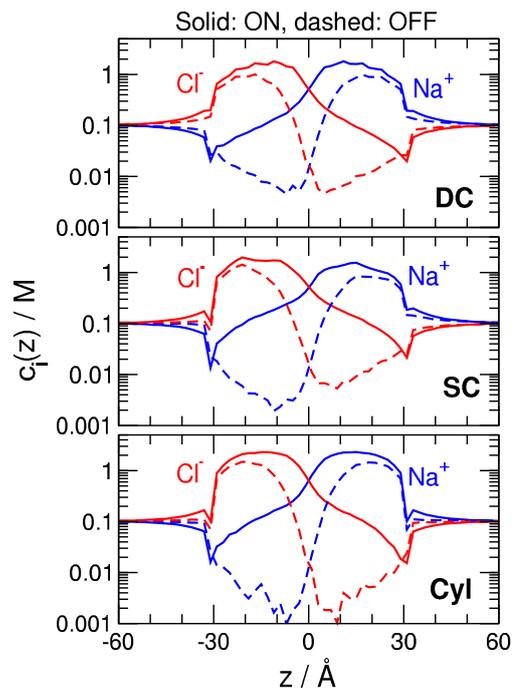


Figure 9: Concentration profiles of Na^+ (blue) and Cl^- (red) in the ON (solid lines) and OFF (dashed lines) states for the three nanopore geometries (Cyl, SC, and DC from bottom to top).

on the right hand side attracts them, (3) the 0.1 M bulk region acts as a source for the ions, (4) applied potential in the OFF state drives Na^+ ion to the right, where they have a peak, and (5) Cl^- ions (that have a peak on the left) attract them. The balance of all these effects forms the deep depletion zone of Na^+ at $z \approx -10 \text{ \AA}$ in the OFF state. Because the Cyl geometry has the smallest radius at the $|z| \approx 10 \text{ \AA}$ positions, this geometry provides the deepest depletion zone as a result of the dominant effect from the above list, the effect of repelling surface charge.

4. Summary

We presented the NP+LEMC method that is a hybrid technique, harvesting the advantageous properties of both the NP transport equation (fast calculation of flux) and LEMC particle simulation method (correct calculation of ionic correlations). We applied the method to compute ionic currents through reduced models of an ionic channel and a bipolar nanopore. Reduced models have the advantage of fast calculation (LEMC would not be feasible for an explicit-water model) and intellectual focus. With these models, we can concentrate on those properties of the system that are important to reproduce its behavior as a device.

For the RyR calcium channel, for example, the reduced model is able to reproduce complex selectivity behavior in agreement with experiments by modeling only the “important” amino acids in a reduced way

[21, 22]. For the bipolar nanopore, the reduced model using implicit water is able to reproduce MD results for an explicit-water model [6]. With further methodological and model development, we intend to simulate nanodevices as close to their real size as possible.

Acknowledgement

The financial support of the National Research, Development and Innovation Office (NKFIH K124353) is gratefully acknowledged. The present article was published in the frame of the Project No. GINOP-2.3.2-15-2016-00053 and supported by the UNKP-17-4 (to M.V.) and UNKP-17-1 (to E.M.), the New National Excellence Program of the Ministry of Human Capacities. We thank Dirk Gillespie for helpful discussions.

REFERENCES

- [1] Boda, D., Gillespie, D.: Steady state electrodiffusion from the Nernst-Planck equation coupled to Local Equilibrium Monte Carlo simulations, *J. Chem. Theor. Comput.*, 2012 **8**(3), 824–829, DOI: 10.1021/ct2007988
- [2] Ható, Z., Boda, D., Kristóf, T.: Simulation of steady-state diffusion: Driving force ensured by Dual Control Volumes or Local Equilibrium Monte Carlo, *J. Chem. Phys.*, 2012 **137**(5), 054109, DOI: 10.1063/1.4739255
- [3] Boda, D., Kovács, R., Gillespie, D., Kristóf, T.: Selective transport through a model calcium channel studied by Local Equilibrium Monte Carlo simulations coupled to the Nernst-Planck equation, *J. Mol. Liq.*, 2014 **189**, 100, DOI: 10.1016/j.molliq.2013.03.015
- [4] Boda, D.: Monte Carlo Simulation of Electrolyte Solutions in Biology: In and Out of Equilibrium, *Annual Reports in Computational Chemistry*, 2014, vol. 10, chap. 5 (Elsevier), 127–163, DOI: 10.1016/b978-0-444-63378-1.00005-7
- [5] Ható, Z., Boda, D., Gillespie, D., Vrabec, J., Rutkai, G., Kristóf, T.: Simulation study of a rectifying bipolar ion channel: detailed model versus reduced model, *Cond. Matt. Phys.*, 2016 **19**(1), 13802, DOI: 10.5488/cmp.19.13802
- [6] Ható, Z., Valiskó, M., Kristóf, T., Gillespie, D., Boda, D.: Multiscale modeling of a rectifying bipolar nanopore: explicit-water versus implicit-water simulations, *Phys. Chem. Chem. Phys.*, 2017 **17**(27), 17816–17826, DOI: 10.1039/C7CP01819C
- [7] Matejczyk, B., Valiskó, M., Wolfram, M.T., Pietschmann, J.F., Boda, D.: Multiscale modeling of a rectifying bipolar nanopore: Comparing Poisson-Nernst-Planck to Monte Carlo, *J. Chem. Phys.*, 2017 **146**(12), 124125, DOI: 10.1063/1.4978942
- [8] Má dai, E., Valiskó, M., Dallos, A., Boda, D.: Simulation of a model nanopore sensor: Ion competition underlies device behavior, *J. Chem. Phys.*, 2017 **147**(24), 244702, DOI: 10.1063/1.5007654
- [9] van Gunsteren, W., Berendsen, H.: Algorithms for Brownian dynamics, *Mol. Phys.*, 1982 **45**(3), 637–647, DOI: 10.1080/00268978200100491
- [10] Turq, P., Lantelme, F., Friedman, H.L.: Brownian dynamics: Its application to ionic solutions, *J. Chem. Phys.*, 1977 **66**(7), 3039–3044, DOI: 10.1063/1.434317
- [11] Corry, B., Kuyucak, S., Chung, S.H.: Tests of continuum theories as models of ion channels. II. Poisson-Nernst-Planck theory versus Brownian dynamics, *Biophys. J.*, 2000 **78**(5), 2364–2381, DOI: 10.1016/s0006-3495(00)76781-6
- [12] Chung, S.H., Allen, T.W., Hoyles, M., Kuyucak, S.: Permeation Of Ions Across The Potassium Channel: Brownian Dynamics Studies, *Biophys. J.*, 1999 **77**(5), 2517–2533, DOI: 10.1016/s0006-3495/99/77087-6
- [13] Chung, S.H., Hoyles, M., Allen, T., Kuyucak, S.: Study Of Ionic Currents Across A Model Membrane Channel Using Brownian Dynamics, *Biophys. J.*, 1998 **75**(2), 793–809, DOI: 10.1016/s0006-3495(98)77569-1
- [14] Berti, C., Furini, S., Gillespie, D., Boda, D., Eisenberg, R.S., Sangiorgi, E., Fiegna, C.: A 3-D Brownian Dynamics simulator for the study of ion permeation through membrane pores, *J. Chem. Theor. Comput.*, 2014 **10**(8), 2911–2926, DOI: 10.1021/ct4011008
- [15] Zheng, Q., Chen, D., Wei, G.W.: Second-order Poisson-Nernst-Planck solver for ion transport, *J. Comp. Phys.*, 2011 **230**(13), 5239–5262, DOI: 10.1016/j.jcp.2011.03.020
- [16] Schuss, Z., Nadler, B., Eisenberg, R.S.: Derivation of Poisson and Nernst-Planck equations in a bath and channel from a molecular model, *Phys. Rev. E*, 2001 **64**(3), 036116, DOI: 10.1103/physreve.64.036116
- [17] Nonner, W., Eisenberg, B.: Ion permeation and glutamate residues linked by Poisson-Nernst-Planck theory in L-type calcium channels, *Biophys. J.*, 1998 **75**(3), 1287–1305, DOI: 10.1016/s0006-3495(98)74048-2
- [18] Pietschmann, J.F., Wolfram, M.T., Burger, M., Trautmann, C., Nguyen, G., Pevarnik, M., Bayer, V., Siwy, Z.: Rectification properties of conically shaped nanopores: consequences of miniaturization, *Phys. Chem. Chem. Phys.*, 2013 **15**(3), 16917–16926, DOI: 10.1039/C3CP53105H
- [19] Gillespie, D., Nonner, W., Eisenberg, R.S.: Coupling Poisson-Nernst-Planck and density functional theory to calculate ion flux, *J. Phys.: Cond. Matt.*, 2002 **14**(46), 12129–12145, DOI: 10.1088/0953-8984/14/46/317
- [20] Gillespie, D., Nonner, W., Eisenberg, R.S.: Density functional theory of charged, hard-sphere fluids, *Phys. Rev. E*, 2003 **68**(3), 031503, DOI: 10.1103/physreve.68.031503
- [21] Gillespie, D., Xu, L., Wang, Y., Meissner, G.: (De)constructing the ryanodine receptor: Modeling

- ion permeation and selectivity of the calcium release channel, *J. Phys. Chem. B*, 2005 **109**(32), 15598–15610, DOI: 10.1021/jp052471j
- [22] Gillespie, D.: Energetics of Divalent Selectivity in a Calcium Channel: The Ryanodine Receptor Case Study, *Biophys. J.*, 2008 **94**(4), 1169–1184, DOI: 10.1529/biophysj.107.116798
- [23] Gillespie, D., Giri, J., Fill, M.: Reinterpreting the anomalous mole fraction effect: The Ryanodine receptor case study, *Biophys. J.*, 2009 **97**(8), 2212–2221, DOI: 10.1016/j.bpj.2009.08.009
- [24] Rosenfeld, Y.: Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing, *Phys. Rev. Lett.*, 1989 **63**(9), 980–983, DOI: 10.1103/physrevlett.63.980
- [25] Gillespie, D., Valiskó, M., Boda, D.: Density functional theory of the electrical double layer: the RFD functional, *J. Phys.-Cond. Matt.*, 2005 **17**(42), 6609–6626, DOI: 10.1088/0953-8984/17/42/002
- [26] Valiskó, M., Gillespie, D., Boda, D.: Selective adsorption of ions with different diameter and valence at highly-charged interfaces, *J. Phys. Chem. C*, 2007 **111**(43), 15575–15585, DOI: 10.1021/jp073703c
- [27] McQuarrie, D.A.: *Statistical Mechanics* (University Science Books, Sausalito), 2000, ISBN: 9781891389153
- [28] Allen, M.P., Tildesley, D.J.: *Computer Simulation of Liquids*, Oxford Science Publ (Clarendon Press, New York), 1989, ISBN: 9780198556459
- [29] Frenkel, D., Smit, B.: *Understanding molecular simulations* (Academic Press, San Diego), 1996, ISBN: 9780122673702
- [30] Hansen, J.P., McDonald, I.R.: *Theory of Simple Liquids* (Elsevier, Academic Press), 2006 ISBN: 9780080455075
- [31] Prigogine, I.: *Introduction to Thermodynamics of Irreversible Processes* (John Wiley and Sons, New York), 1968, ISBN: 9780470699287
- [32] Kreuzer, H.J.: *Nonequilibrium Thermodynamics and its Statistical Foundations*, Monographs on the physics and chemistry of materials (Clarendon Press, Oxford), 1983, ISBN: 9780198513759
- [33] de Groot, S.R., Mazur, P.: *Non-Equilibrium Thermodynamics*, Dover Books on Physics (Dover, New York), 2013, ISBN: 9780486153506
- [34] Tschoegl, N.W.: *Fundamentals of Equilibrium and Steady-State Thermodynamics* (Elsevier, Amsterdam), 2000, ISBN: 9780080532110
- [35] Attard, P.: *Thermodynamics and Statistical Mechanics, Equilibrium by Entropy Maximisation* (Academic Press, London), 2002, ISBN: 9780120663217
- [36] Evans, D.J., Morriss, G.: *Statistical Mechanics of Nonequilibrium Liquids* (Cambridge University Press, New York), 2008, ISBN: 9781139471930
- [37] Beck, T.L., Paulaitis, M.E., Pratt, L.R.: *The Potential Distribution Theorem and Models of Molecular Solutions* (Cambridge University Press, Cambridge), 2006, ISBN: 9781139457637
- [38] Malasics, A., Boda, D.: An efficient iterative grand canonical Monte Carlo algorithm to determine individual ionic chemical potentials in electrolytes, *J. Chem. Phys.*, 2010 **132**(24), 244103, DOI: 10.1063/1.3443558
- [39] Boda, D., Gillespie, D., Nonner, W., Henderson, D., Eisenberg, B.: Computing induced charges in inhomogeneous dielectric media: Application in a Monte Carlo simulation of complex ionic systems, *Phys. Rev. E*, 2004 **69**(4), 046702, DOI: 10.1103/physreve.69.046702
- [40] Boda, D., Valiskó, M., Eisenberg, B., Nonner, W., Henderson, D., Gillespie, D.: The effect of protein dielectric coefficient on the ionic selectivity of a calcium channel, *J. Chem. Phys.*, 2006 **125**(3), 034901, DOI: 10.1063/1.2212423
- [41] Widom, B.: Some Topics in the Theory of Fluids, *J. Chem. Phys.*, 1963 **39**(11), 2808–2812, DOI: 10.1063/1.1734110
- [42] Widom, B.: Structure of interfaces from uniformity of the chemical potential, *J. Stat. Phys.*, 1978 **19**(6), 563–574, DOI: 10.1007/bf01011768
- [43] Hille, B.: *Ion channels of excitable membranes* (Sinauer Associates, Sunderland), 2001, ISBN: 9780878933211
- [44] Sather, W.A., McCleskey, E.W.: Permeation and selectivity in calcium channels, *Ann. Rev. Physiology*, 2003 **65**(1), 133–159, DOI: 10.1146/annurev.physiol.65.092101.142345
- [45] Gao, L., Balshaw, D., Xu, L., Tripathy, A., Xin, C., Meissner, G.: Evidence for a Role of the Luminal M3-M4 Loop in Skeletal Muscle Ca²⁺ Release Channel (Ryanodine Receptor) Activity and Conductance, 2000 **79**(2), 828–840, DOI: 10.1016/s0006-3495(00)76339-9
- [46] Wang, Y., Xu, L., Pasek, D.A., Gillespie, D., Meissner, G.: Probing the Role of Negatively Charged Amino Acid Residues in Ion Permeation of Skeletal Muscle Ryanodine Receptor, 2005 **89**(1), 256–265, DOI: 10.1529/biophysj.104.056002
- [47] Xu, L., Wang, Y., Gillespie, D., Meissner, G.: Two Rings of Negative Charges in the Cytosolic Vestibule of Type-1 Ryanodine Receptor Modulate Ion Fluxes, 2006 **90**(2), 443–453, DOI: 10.1529/biophysj.105.072538
- [48] Boda, D., Valiskó, M., Eisenberg, B., Nonner, W., Henderson, D., Gillespie, D.: Combined effect of pore radius and protein dielectric coefficient on the selectivity of a calcium channel, *Phys. Rev. Lett.*, 2007 **98**(16), 168102, DOI: 10.1103/physrevlett.98.168102
- [49] Gillespie, D., Boda, D.: The anomalous mole fraction effect in calcium channels: A measure of preferential selectivity, *Biophys. J.*, 2008 **95**(6), 2658–2672, DOI: 10.1529/biophysj.107.127977

- [50] Boda, D., Valiskó, M., Henderson, D., Eisenberg, B., Gillespie, D., Nonner, W.: Ion selectivity in L-type calcium channels by electrostatics and hard-core repulsion, *J. Gen. Physiol.*, 2009 **133**(5), 497–509, DOI: 10.1085/jgp.200910211
- [51] Malasics, M., Boda, D., Valiskó, M., Henderson, D., Gillespie, D.: Simulations of calcium channel block by trivalent ions: Gd^{3+} competes with permeant ions for the selectivity filter, *Biochim. et Biophys. Acta - Biomembranes*, 2010 **1798**(11), 2013–2021, DOI: 10.1016/j.bbamem.2010.08.001
- [52] Rutkai, G., Boda, D., Kristóf, T.: Relating binding affinity to dynamical selectivity from dynamic Monte Carlo simulations of a model calcium channel, *J. Phys. Chem. Lett.*, 2010 **1**(14), 2179–2184, DOI: 10.1021/jz100718n
- [53] Boda, D., Giri, J., Henderson, D., Eisenberg, B., Gillespie, D.: Analyzing the components of the free energy landscape in a calcium selective ion channel by Widom's particle insertion method, *J. Chem. Phys.*, 2011 **134**(5), 055102, DOI: 10.1063/1.3532937
- [54] Boda, D., Henderson, D., Gillespie, D.: The role of solvation in the binding selectivity of the L-type calcium channel, *J. Chem. Phys.*, 2013 **139**(5), 055103, DOI: 10.1063/1.4817205
- [55] Nonner, W., Catacuzzeno, L., Eisenberg, B.: Binding and selectivity in L-type calcium channels: A mean spherical approximation, *Biophys. J.*, 2000 **79**(4), 1976–1992, DOI: 10.1016/s0006-3495(00)76446-0
- [56] Siwy, Z., Fulinski, A.: Fabrication of a synthetic nanopore ion pump, *Phys. Rev. Lett.*, 2002 **89**(19), 198103, DOI: 10.1103/physrevlett.89.198103
- [57] Siwy, Z., Apel, P., Baur, D., Dobrev, D.D., Korchev, Y.E., Neumann, R., Spohr, R., Trautmann, C., Voss, K.O.: Preparation of synthetic nanopores with transport properties analogous to biological channels, *Surf. Sci.*, 2003 **532**, 1061–1066, DOI: 10.1016/s0039-6028(03)00448-5
- [58] Siwy, Z., Apel, P., Dobrev, D., Neumann, R., Spohr, R., Trautmann, C., Voss, K.: Ion transport through asymmetric nanopores prepared by ion track etching, *Nuclear Instruments & Methods In Phys. Research Section B-beam Interactions With Materials Atoms*, 2003 **208**, 143–148, DOI: 10.1016/s0168-583x(03)00884-x
- [59] Howorka, S., Siwy, Z.: Nanopores: Generation, Engineering, and Singly-Molecule Applications, in Handbook of Single-Molecule Biophysics (P. Hinterdorfer, A. van Oijen, eds.), Advances in Chemical Physics, chap. Chapter 11 (Springer), 2009 293–339, DOI: 10.1007/978-0-387-76497-9_11
- [60] Guo, W., Tian, Y., Jiang, L.: Asymmetric Ion Transport through Ion-Channel-Mimetic Solid-State Nanopores, *Acc. Chem. Res.*, 2013 **46**(12), 2834–2846, DOI: 10.1021/ar400024p
- [61] Gibb, T., Ayub, M.: Solid-State Nanopore Fabrication, in Engineered Nanopores for Bioanalytical Applications (Elsevier BV), 2013 121–140, DOI: 10.1016/b978-1-4377-3473-7.00005-4
- [62] Guan, W., Li, S.X., Reed, M.A.: Voltage gated ion and molecule transport in engineered nanochannels: theory, fabrication and applications, *Nanotechnology*, 2014 **25**(12), 122001, DOI: 10.1088/0957-4484/25/12/122001
- [63] Zhang, H., Tian, Y., Jiang, L.: Fundamental studies and practical applications of bio-inspired smart solid-state nanopores and nanochannels, *Nano Today*, 2016 **8**(3), 1470–1478, DOI: 10.1016/j.nantod.2015.11.001