JSCSEN 80(3)279–452(2015)

ISSN 1820-7421(Online)

Journal of the Serbian Chemical Society

Electronic version

Volume 80 : : 2015 : : 85 Years of the Journal

1930 Glasnik Hemijskog Društva Kraljevine Jugoslavije
Journal of the Chemical Society of the Kingdom of Yugoslavia

1947 Glasnik hemijskog društva Beograd
Journal of the Chemical Society of Belgrade

1985 Journal of the Serbian Chemical Society

VOLUME 80

No 3

BELGRADE 2015

Available on line at

www.shd.org.rs/JSCS/

The full search of JSCS is available through DOAJ DIRECTORY OF OPEN ACCESS JOURNALS www.doaj.org

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

J. Serb. Chem. Soc. Vol. 80, No. 3 (2015)

CONTENTS

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

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 279–304 (2015) UDC **Pseudomonas aeruginosa*:577.115+547.441:
JSCS–4716 661.185.000.57+66.0
Review

REVIEW

Rhamnolipid biosurfactant from *Pseudomonas aeruginosa* – from discovery to application in contemporary technology

MILENA G. RIKALOVIĆ[1], MIROSLAV M. VRVIĆ[1] and IVANKA M. KARADŽIĆ[2]*

[1]*Faculty of Chemistry, University of Belgrade, Studentski trg 12–16, Belgrade and* [2]*Department of Chemistry, Faculty of Medicine, University of Belgrade, Visegradska 26, Belgrade, Serbia*

(Received 27 June, revised 19 September, accepted 20 September 2014)

Abstract: The rhamnolipids will most likely be the next generation of biosurfactants to reach the market. They should follow closely after alkyl polyglycosides, already established in the biosurfactants market, and sophorolipids, which can be found in several cleaning agents. However, the greatest numbers of recent publications and patents among glycolipid biosurfactants have been dedicated to rhamnolipids. Produced mainly by *Pseudomonas aeruginosa*, rhamnolipids are mixtures of different rhamnolipid congeners, which show physico-chemical properties that differ from those of single congeners, with the most abundant structure in the mixture having the largest impact on the overall characteristics of the total mixture. Characteristics of biodegradability, low toxicity, production from renewable sources and antimicrobial (particularly antifungal) activity together make rhamnolipid biosurfactants particularly promising for broad commercial application. Although to date, bioremediation has been the major topic filed for patents utilizing rhamnolipids, the increasing number of patents for applications in cosmetics, agronomy and food industries, formulation of cleaners and nanotechnology indicates their future implementation in these fields.

Keywords: rhamnolipids; *Pseudomonas aeruginosa*; methods; production; application.

CONTENTS

* Corresponding author. E-mail: ikaradzic@med.bg.ac.rs
doi: 10.2298/JSC140627096R

1. INTRODUCTION

Research related to the characterization of microbial surfactants and their applications has been rapidly growing in recent years. The reasons for the increased attention in these naturally derived compounds are: diversity, environmentally friendly nature, biodegradability, low toxicity, excellent foaming properties, high selectivity and activity at extremes of temperature, pH and salinity, plus the fact that they are able to be produced from industrial wastes and by-products.[1] Biosurfactants have shown great potential for applications in environmental protection, crude oil recovery, health care and food-processing industries.[1–6]

Microbial surface active compounds are important for the physiology of the cells that produce them, as they are involved in cell motility (gliding and swarming motility as well as de-adhesion from surfaces), cell–cell interactions (biofilm formation, maintenance and maturation, quorum sensing, amensalism and pathogenicity), cellular differentiation, substrate accession (*via* direct interfacial contact and pseudosolubilization of substrates), as well as avoidance of toxic elements and compounds.[6]

Biosurfactants are divided into low-molecular-mass molecules, which efficiently lower surface and interfacial tension, and high-molecular-mass polymers, which are more effective as emulsion stabilizing agents.[6] The major classes of low-mass surfactants include glycolipids, lipopeptides and phospholipids, whereas high-mass surfactants include polymeric and particulate (vesicles and whole cells) surfactants.[6] *Pseudomonas* spp. are considered to be the most common producers of surface active compounds, synthesizing both classes of biosurfactants,[6,7] with *P. aeruginosa* being the preeminent rhamnolipid (RL) biosurfactant-producing bacteria species.[6,8]

After alkyl polyglycosides, bio-based surfactants that are already accepted on the market, and sophorolipids, which are in several cleaning agents formulations, RLs will probably be the next generation of biosurfactants to reach the market.[9] The properties of RLs, including biodegradability, low toxicity, the possibility for production from renewable sources, together with their antimicrobial and antifungal activity, make them particularly interesting for commercial application.[9]

Aside from the environmental use of RLs,[4,10,11] an increasing number of patents[9] indicate the potential successful application of rhamnolipids in cosmetics, agronomy, the food industry, cleaner formulations and nanotechnology. In this review, special attention is paid to RLs produced by *P. aeruginosa*, their properties and the potential for their use in different technological fields.

2. GLYCOLIPID BIOSURFACTANTS

2.1. The most researched glycolipid biosurfactants

The most well known biosurfactants are glycolipids, which are composed of carbohydrates linked by means of ether or ester bonds with either long-chain aliphatic acids or hydroxyaliphatic acids. Sophorolipids, trehalose lipids, mannosylerythritol lipids (MELs) and RLs are the most widely researched glycolipids.[2,12]

Sophorolipids are extracellularly produced by several yeast species, such as *Candida bombicola*, *C. apicola*, *Rhodotorula bogoriensis*, *Wickerhaminella domercqiae* and *C. batistae*.[13] These compounds contain two β-1,2 linked glucose units, with acetylated 6- and 6′-hydroxyl groups, connected to the lipid part by a glycosidic linkage.[14] Generally, sophorolipids occur as a mixture of macrolact-

ones and the free acid form of at least six to nine different sophorolipid congeners.[12] Applications of sophorolipids were reported in the biomedical field, in the synthesis of metal-bound nanoparticles in cosmetics and in the formulation of pharmaco-dermatological products.[13,15] Furthermore, these biosurfactants show potential for some recovery processes.[1]

Trehalose lipids are glycolipids containing trehalose as the sugar moiety, which is the basic component of the cell wall glycolipids in *Mycobacteria* and *Corynebacteria*.[16] Depending on the producing organism, trehalose lipids show differences in the size and structure of mycolic acid, the number of carbon atoms and the degree of unsaturation. The most reported trehalose lipid is trehalose 6,6′--dimycolate,[17] but different trehalose-containing glycolipids are known to be produced by several other microorganisms, such as *Arthrobacter*, *Nocardia*, *Rhodococcus* and *Gordonia*. Trehalose lipids have been applied in bioremediation to enhance the bioavailability of hydrocarbons. Recently, the increasing interest in trehalose lipids is related to their functions in cell membrane interaction and their potential as antitumor therapeutic agents.[17]

The mannosylerythritol glycolipids are reported as metabolites of yeasts strains, which belong to the genera *Pseudozyma* sp. and *Ustilago* sp., produced on soybean oil or *n*-alkane.[17] MELs are a mixture of partially acylated derivatives of 4-*O*-*β*-D-mannopyranosyl-D-erythritol, containing C_2, C_{12}, C_{14}, $C_{14:1}$, C_{16}, $C_{16:1}$, C_{18} and $C_{18:1}$ fatty acids as the hydrophobic groups.[18] Interesting applications of MELs have been reported in the biomedical field as antimicrobial, antitumor and immunomodulating molecules, in the biotechnological field for gene and drug delivery, and in cosmetic applications as skin moisturizers.[15,17,18]

RLs are glycosides, produced mainly by *P. aeruginosa* and by the genus *Burkholderia*, composed of one (for mono-rhamnolipids) or two (for di-rhamnolipids) rhamnose sugar moieties linked to one or two *β*-hydroxy fatty acid chains. RLs in which one or two molecules of rhamnose are linked to one or two molecules of *β*-hydroxydecanoic acid (Fig. 1) are the best-studied biosurfactant compounds. While the OH group of one of the acids is involved in a glycosidic linkage with the reducing end of the rhamnose disaccharide, the OH group of the second acid is involved in ester formation. RLs are the principal glycolipid produced by *P. aeruginosa*.[12] The major field of application of RLs is in bioremediation processes.[1,10,19] Additionally, these molecules, due to their physicochemical and biological properties, have great potential for applications in the biomedicine, pharmacology, cosmetics, food and agricultural industries.[12,15,17,19] RLs have also been used in the preparation of nanoparticles and microemulsions.[17]

2.2. Rhamnolipids

2.2.1. Discovery

The discovery of RLs dates back to 1946, when Bergström *et al.*[20,21] reported an oily glycolipid, which was named piolipic acid, composed of L-rhamnose and *β*-hydroxydecanoic acid,[22,23] produced by *P. pyocyaneus* (today *P. aeruginosa*), when cultivated on glucose. Further characterization of the structure by Jarvis and Johnson[23] showed that the isolated compound was composed of two *β*-hydroxydecanoic acids (connected *via* an ester bond) linked through a glycosidic bond to two rhamnose moieties. Additionally, Edwards and Hayashi,[24] using periodate oxidation and methylation, reported that the linkage between the two rhamnose moieties is an *α*-1,2-glycosidic linkage. Based on this, they chemically described the RL as 2-*O*-*α*-L-rhamnopyranosyl-*α*-L-rhamnopyranosyl-*β*-hydroxydecanoyl-*β*-hydroxydecanoate (di-RL structure). This was the first glycolipid discovered containing a link between a sugar and a hydroxylated fatty acid residue.[25] From 1972 to 2000, various RL structures produced by *P. aeruginosa* strains growing on different carbon sources (*n*-paraffins, glycerol, *n*-alkanes, glucose, *etc*.), were isolated and reported. The isolated RLs included all types of RL homologues (RL1, RL2, RL3 and RL4), and their number increased with the progress of analytical methods.[8]

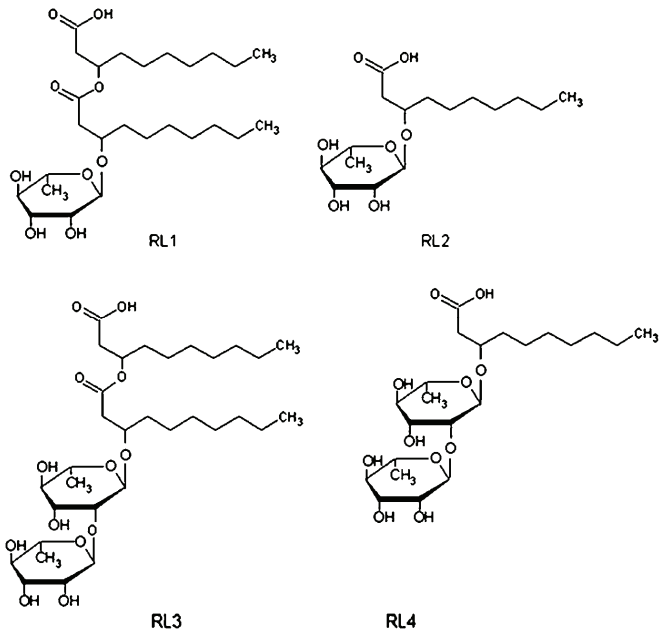


Fig. 1. Structure of rhamnolipid biosurfactants: mono-rhamno-di-lipidic structure (RL1), mono-rhamno-mono-lipidic structure (RL2), di-rhamno-di-lipidic structure (RL3) and di-rhamno-mono-lipidic structure (RL4).

Furthermore, some studies reported the complex composition of bacterially produced RL mixtures, including, for example, an RL mixture composed of 67 % di-rhamno-di-lipid, 22 % mono-rhamno-di-lipid, 9 % di-rhamno-mono-lipid, and less than 3 % of mono-rhamno-mono-lipid.[26] The development of sensitive, high throughput analytical techniques has led to the further discovery of a wide diversity of RL congeners and homologues (about 60) produced in different concentrations by various *Pseudomonas* species and other bacteria.[8]

2.2.2. *Pseudomonas* spp. as producers of rhamnolipids

For more than 60 years, much work has been undertaken to stimulate the development and establishment of RL production by *P. aeruginosa*.[5] When comparing the production of RLs by different *Pseudomonas* species, Onbasli and Aslim[27] found that *P. luteola* and *P. putida* gave higher RL yields than *P. aeruginosa*, *P. fluorescens* and *P. stutzerri*, but the composition and distribution of the RL homologues were not described for these strains, while the production of RLs by *P. chlororaphis* had lower yield than *P. aeruginosa*.[28] Furthermore, reports about RLs obtained from *Pseudomonas* species other than *aeruginosa* lack information about their surface properties, which is the most important indicator of surfactant quality.

3. PROPERTIES OF RHAMNOLIPID MIXTURES FROM *P. aeruginosa*

3.1. Diversity

Interest in RLs arises from two contrasting facts. On the one hand, RLs show relatively high surface activities, and are produced in relatively high yields after fermentation processes, which are relatively well understood. On the other hand, they are considered as virulence factors involved in the processes of *P. aeruginosa* pathogenesis, resulting in an investigation of RL biosynthesis in order to control their production and effects.[29]

P. aeruginosa produces various extracellular mono- and di-rhamnolipids. Two types of RLs (Fig. 1), termed RL1 and RL3, consisting of one or two L-rhamnose units and two units of ß-hydroxydecanoic acid, are the principal RLs. There are also RL2 and RL4 RLs, which consist of one or two L-rhamnose units and one unit of ß-hydroxydecanoic acid.[30]

3.2. Composition

Naturally produced RLs always appear as mixtures of different RL congeners, as was observed when produced by various strains of *P. aeruginosa*.[31–38] The complexity of the composition of RL mixtures was found to depend on various factors such as: origin of the bacterial strain, type of carbon substrate,[39–41] culture conditions,[42] age of the culture,[34] the *P. aeruginosa* strain itself,[34,38] as well as the method of RL isolation and purification.[43,44]

Some RL congeners are predominant in all RL-producing *P. aeruginosa* strains and are classified as the major RL structures, while others, produced only sometimes or with low abundance, are the minor RL structures. Both the major and the minor RL congeners contribute to the complete profile of RLs. In all studies of RL mixtures produced by *P. aeruginosa* strains, it was shown that, aside from their varying composition, these mixtures contain mono-RL and di--RL structures. In the lipid part, chain lengths from C_8 to C_{12} and Rha–C_{10}–C_{10} and Rha–Rha–C_{10}–C_{10} were the predominant congeners.[34,38,45,46] The ratio of total di- to mono-RL fractions for RL produced by *P. aeruginosa* is reported to range from 1 to 5.[10,38,47]

Despite the diversity of the reported RL structures, there are relatively few studies which both quantitatively and qualitatively analyzed RLs.[7,38,43,47–49] The composition of RL mixtures produced is important, since this defines the physicochemical properties of the products, which further impacts on their potential application.

3.3. Physicochemical properties

RL mixtures show physicochemical properties that are different from single RL congeners. Hence, the most abundant structure will have the biggest impact on the characteristics and dictate the behavior of the total RL mixture.

RLs can reduce the surface tension of water from 72 to 30 mN m^{-1} and the interfacial tension of water/oil to values < 1 mN m^{-1}.[33] RLs are weak acids due to the carboxylic acid moieties and are known to undergo aggregation in solution.[33] At concentrations above the critical micelle concentration (CMC), RLs form micelles, vesicles or lamella, depending on the solution properties.[50] The CMC values for RLs depend on the chemical composition of the congeners present in each RL mixture (the ratio and composition of the homologues, the presence of unsaturated bonds, the branching and length of the alkyl chain, or the size of the hydrophilic head group), and on their chemical environment, and was reported to range from 5 to 200 mg L^{-1}.[38,47,48,51] Furthermore, in a comparative study, a statistically significant increase in the CMC values was observed on lowering the ratio of total mono-/di-rhamnolipid and ratio of Rha–C_{10}–C_{10}/Rha--Rha–C_{10}–C_{10} ($P \leq 0.05$), indicating that mono-rhamnolipids start to form micelles at lower concentrations than di-rhamnolipids.[38]

RLs were shown to have higher solubilization capacity, which is expressed as the molar solubility ratio (*MSR*) for non-aqueous phase liquids (NAPL), as well as for solid-phase organics, such as polycyclic aromatic hydrocarbons (PAH) and hydrocarbon mixture than some of the commonly used synthetic surfactants (Tween, Triton and alkylbenzyl sulfonate).[3]

RLs are also an example of a readily degraded agent, as determined by the OECD 301 Ready Biodegradability Test.[52] Furthermore, invertebrate toxicity

tests performed in accordance with OECD 202 showed RL to have very low toxicity.[3] RLs also have high affinity for a variety of metals, including cadmium, copper, lanthanum, lead and zinc.[3]

4. METHODS FOR THE CHARACTERIZATION OF RHAMNOLIPID MIXTURES

4.1. Identification and quantification

The quest for a cost-efficient production of RL biosurfactant and alternatives, plus more efficient, RL-producing strains of bacteria is associated not only with economic aspects of their production, but also with new demands in the analysis of RL mixtures.[53] A large variety of analytical methods have been employed to identify and quantify the different RL species and RL production. These methods range from indirect analysis based on the physical properties of RLs (determination of surface tension and hemolytic test),[29,54] colorimetric measurement (cetyltrimethylammonium bromide agar test, anthrone method and orcinol test)[54–56] to spectroscopic analysis of sample structure by infrared and nuclear magnetic resonance spectroscopy and a sophisticated analysis of composition by mass spectrometry (MS).[34,43,46,57,58]

One current method for RL identification and quantification is high performance liquid chromatography (HPLC) analysis coupled with electrospray ionization mass spectrometry (ESI-MS).[33,34,43,53,57] Furthermore, fragmentation of the pseudomolecular ion using tandem MS analysis of the parent ion is often performed in order to provide additional structural information, such as discrimination of isomeric pairs with subtle structural variations.[43] Matrix assisted laser desorption ionization-time of flight MS (MALDI-TOF MS) approaches were recently developed for high-throughput screening of naturally occurring mixtures of RLs from *Pseudomonas* spp.[58]

4.2. Tensioactive properties

Methods based on the tensioactive properties of RL biosurfactant are typical for all surface-active compounds and involve the calculation of the value of the hydrophilic–lipophilic balance (*HLB*), the drop collapsing test, the oil spreading test, determination of the surface activity and the *CMC* values,[9,29] as well as the determination of the interfacial tension and the wetting properties.[59]

Determination of the *HLB* value, which varies between 0 and 20, enables the prediction of the behavior of surface-active compounds: 0–3 antifoaming agents, 4–6 W/O (water/oil) emulsifiers, 7–9 wetting agents, 8–18 O/W (oil/water) emulsifiers, 13–15 typical detergents and 10–18 solubilizers/hydrotropes. The *HLB* value can only be used as a preliminary guide, and further analysis by more exact and precise analytical techniques is required. According to Griffin,[60] the *HLB* can be calculated as $HLB = 20\times(MW_{HP}/MW_{SA})$, where MW_{HP} is the molecular

weight of the hydrophilic part and MW_{SA} is the molecular weight of the whole surface-active agent.[9]

The drop-collapsing test and oil-spreading test are methods for rapid screening of RL-producing bacterial strains.[29] Both tests are based on a similar approach, where the responses depend on the presence of RLs in culture supernatant. In the case of the drop-collapsing test, a sample of supernatant is applied to a polystyrene plate containing shallow wells covered with oil,[61] and spreads over the oil only if the sample of culture supernatant contains RLs. In the oil--spreading test, a drop of bacterial supernatant is added on top of an oil/water interface,[62] where the presence of RLs causes the oil to be repelled, resulting in the formation of a clearing zone, the diameter of which can be correlated with the activity of the tensioactive compounds in the supernatant.[29]

A more precise method for determining the surfactant properties of RLs in bacterial culture supernatants or of isolated RL biosurfactants is the direct measurement of surface activity and the determination of the *CMC* (specific to each surfactant), which is performed by measurement of the surface tension after sequential dilution of the solution.[29] However, the determination of the *CMC* suffers the drawbacks that it is time-consuming and not applicable for high--throughput screening.[29] Additionally, the results of these methods, as well as of all the previous indirect tests based on surface tension, could be affected by the potential presence of other tensioactive compounds.

One method to quantify solid–liquid adsorption is to measure the interfacial tension, which can be performed by tensiometry or by contact angle goniometry.[59,63] Surface tension and contact angle measurements as a function of surfactant concentration are directly related to the difference in the adsorption of surfactants on solid–vapor (S–V) and solid–liquid (S–L) interfaces. Contact angle measurements also aid the elucidation of the nature of interactions between a surfactant molecule and a solid surface.[64] Özdemir and Malayoglu[64] investigated the wetting behavior of a mixture of mono- and di-rhamnolipid and SDS (sodium dodecyl sulfate) molecules on glass, PET (polyethylene terephthalate) and gold surfaces by measuring the advancing contact angle, and elucidated the preferences of surfactant molecules adsorbed onto SL (solid–liquid)–SV (solid–vapor) and L–V (liquid–vapor) interfaces.

5. PRODUCTION OF RHAMNOLIPIDS FROM *P. aeruginosa*

5.1. Bottlenecks for production

The high cost of production, isolation and purification of RLs, and the low yield has determined that, despite 60 years of research in the area of RL production, the economic feasibility of these glycolipids is still pending.[65] Often, the amount and type of a raw material can contribute considerably to the production cost, as raw materials account for 10–30 % of the total production cost in most

biotechnological processes.[12] Solutions resulting in more economic production include the use of cheap and renewable resources, optimized and efficient bioprocesses, over-production by mutant or recombinant strains,[12,66] as well as the application of response surface methodology (RSM) as a modern concept for optimization of biotechnological processes.[67]

The major problems for large-scale applications of RLs are safety and yield issues related to their main producers - strains of *P. aeruginosa*. Despite the fact that in the UK, *P. aeruginosa* spp. are classified as type II opportunistic pathogens that are not highly infective, large-scale fermentation production would require compliance to special procedures by employees involved in the production.[68] However, commercial-scale production is already being undertaken in the USA, with no reported problems.[68]

Problems related to the pathogenic status of *P. aeruginosa* could be solved by two strategies: heterologous production of RLs in non-pathogenic bacteria and identification of potential new non-pathogenic RL producers, for which the RL yields are acceptable.[29,68] The heterologous production of RLs would bring about two major advantages as compared to their production by *P. aeruginosa*. The first is the increased safety when handling large amounts of culture broths, while the second is the possibility of constitutive RL production, in contrast to the very tightly regulated production in *P. aeruginosa*.[29] Although several attempts to obtain heterologous production of *Pseudomonas* RLs have been reported, none of them produced RLs at levels comparable with those of the best *P. aeruginosa* strains.[29] In terms of the commercial production of RLs, there is still vast potential for genetic optimization.[29] Other species and genera of bacteria were also found to be RL producers.[8] Most RL-producing species belong to the closely related genera *Pseudomonas* and *Burkholderia*.[47] The most prominent nonpathogenic RL producers from the genus *Pseudomonas* are *P. chlororaphis*, *P. alcaligenes* and *P. putida*, and from the genus *Burkholderia*, they are *B. glumae*, *B. plantarii*, *B. pseudomallei* and *B. thailandensis*. These species from the genus *Burkholderia* primarily produce Rha–Rha–C_{14}–C_{14}, but a number of other congeners, including mono-RLs, mostly Rha–C_{14}–C_{14}, were recently detected in cultures of *B. pseudomallei* and *B. thailandensis*.[47] Additionally, in the mentioned study, the RL mixture was further characterized and the results, such as ratio of di-RL to mono-RL, the major RL congeners and genetic background of RL biosynthesis, were compared to RL produced by *P. aeruginosa*. On the other hand, one non-pathogenic RL producer, *P. chlororaphis*, synthesizes exclusively mono-RL congeners, primarily Rha–C_{10}–C_{12} and Rha–C_{10}–$C_{12:1}$, under static conditions of fermentation, with a yield of 1 g L^{-1}.[28] RLs were also detected in cultures of many other genera and species of widely different taxonomical origins.[29] However, for most of these strains, structural determination of the putative RLs was not accomplished and, sometimes, the actual identification of the

producing strain was not firmly confirmed.[29] Due to this, their biotechnological potential is currently unknown.

Other major obstacles in the production of RL are related to the yields, the substrates needed to produce them, and the downstream processing required. At present, strains of *P. aeruginosa* produce 10–20 g L^{-1} of RL in the laboratory (although there have been some reports of hyperproducer strains), which is low in comparison to the yields obtained for other glycolipid biosurfactants, MEL and sophorolipids (100 g L^{-1}).[68] One advantage of RLs, and glycolipid biosurfactants generally, is the possibility for production from a range of renewable substrates, such as industrial waste. Additionally, complex regulation of RL biosynthesis in *P. aeruginosa* (environment conditions, nutrition factors and quorum sensing mechanism), which is not yet fully understood, prevents the development of hyperproducing strains, either by mutagenesis and selection or by genetic manipulation. Failure to achieve high yields may eventually preclude rhamnolipids from use in many possible applications.[68]

5.2. Commercial production

Rhamnolipids from *P. aeruginosa* are officially produced by only a few companies in the USA.[68] Small quantities of RL from the strain *P. aeruginosa* NY3 are produced by AGAE Technologies (www.agaetech.com). This company produces highly purified rhamnolipid as a mixture, which can be a regular or liquid product, or separate mono- and di-RL fractions (90 or 95 % purity). The Jeneil Biotech Company, which is generally a food additive producer, and Rhamnolipid Companies, Inc. (www.rhamnolipid.com), a specialist producer of these biosurfactants, both manufacture RL on a larger scale. Jeneil Biotech RL products range from the crudest preparation comprising fermentation broth with approximately (2 % RLs) to partially purified products (99 % RLs).[68] Rhamnolipid Companies, Inc. offers raw 99 % rhamnolipid in 1–25 % aqueous solutions and 1–25 % solutions with stabilizers, refined RL in 0.1 to 7 % in silicone, mineral oil, alcohol or other carriers, and refined powder mono- and di-RL.

5.3. Substrates for improved production

RLs are considered to be secondary metabolites and as such, their production coincides with the onset of the bacterial stationary phase.[3] Production of RLs on the cellular level, as well as under conditions for laboratory or industrial production, are not fully mastered, even though the biosynthesis of RLs and their regulation are partially explained.[10,69,70] Traditional engineering by random and targeted genetic alteration, process design, and recombinant strategies have not yet proven successful.[65] For enhanced process development, there is an urgent need for in-depth information concerning the regulation of RL production during

bioreactor cultivation of RL-producing bacteria, in order to design knowledge-based genetic and process engineering strategies.[65]

Reported data indicate that the production of RLs is possible from simple carbon sources, such as glycerol and glucose, or complex carbon sources (olive oil, sunflower oil) or wastes (crude whey, distillery waste, molasses, corn steep liquor, sunflower oil mill effluent, olive oil mill effluent (OOME), frying oil, soap stock) and minerals ($NaNO_3$, NH_4Cl) or combined nitrogen source ($NaNO_3$ and yeast extract or NH_4Cl and peptone).[10,29] This indicates that economically viable levels of biosurfactant production could be achieved using renewable resources for the carbon source, with special emphasis on the importance of the utilization of industrial by-products and agricultural wastes as cost-effective alternative substrates for microbial growth and biosurfactant production.[38,41,66,68–75]

The study of Dubey and Juwarkar[71] is an example of RL production from renewable sources, in which *P. aeruginosa* BS2 was cultivated on whey for 48 h, with an RL yield of 0.92 g L^{-1}. In other studies, Mercade *et al.*[74] cultivated *Pseudomonas* sp. JAMM in a medium with OOME (100 g L^{-1}) and $NaNO_3$ (2.5 g L^{-1}), which resulted in an RL yield of 14 g kg^{-1} OOME after 150 h of cultivation. Haba *et al.*,[75] using sunflower and olive frying oil as carbon sources and $NaNO_3$ as the nitrogen source, studied the production of biosurfactant by several *P. aeruginosa* strains. Most of the examined *P. aeruginosa* strains showed good growth on both sources and one of them produced a yield of 2.7 g L^{-1} of RL. In another study, Haba *et al.*[34] found *P. aeruginosa* AT10 produced an RL yield of 16.5 g L^{-1} on soybean residual fatty acids. Benincasa *et al.*[39] tested *P. aeruginosa* LBI for RL production by batch fermentation on soap stock as a carbon source; the maximal production of biosurfactant was 15.9 g L^{-1}. Rikalovic *et al.*[38] tested the effect of different carbon sources on RL production by several *P. aeruginosa* isolates. The best RL yield of 1.3 g L^{-1} was obtained using sunflower frying oil as a carbon source and peptone and NH_4Cl as nitrogen sources. In the study of Aparna *et al.*,[45] *Pseudomonas* sp. 2B produced 4.97 g L^{-1} of RL when cultivated on molasses, peptone and NH_4Cl.

6. APPLICATIONS OF RHAMNOLIPIDS

The main properties of biosurfactants, *i.e.*, detergency, foaming, emulsifying, demulsifying, solubilizing, wetting, thickening, metal sequestering, vesicle forming and phase dispersion properties, among others, can be exhibited by RLs.[5,76] All of these properties are associated with the amphiphilic character of the RL molecules, and confer upon them the ability to accumulate between fluid phases, thus reducing surface and interfacial tensions.[5] Environmental uses of RLs are currently considered as the major field for potential application of RLs,[3,46,77–80] but an increasing number of patents indicates the successful appli-

cation of RLs in the food industry, agriculture, the cosmetics industry, pharmacology, and in nanotechnology.[9] The applications of rhamnolipids, which will be further discussed, are summarized in Table I.

TABLE I. Fields for rhamnolipid applications

Field of application	Specific processes / purposes
Bioremediation	Organic flushing agents (PAH, NAPL, hydrocarbon mixtures Metal flushing agents (heavy metals Cu, Zn, Pb, Ni, Cr) Biodegradation of organics (petroleum and petroleum derivatives) Biodegradation of organics in metal–organic co-contaminated systems
Food industry	Multipurpose ingredients (decrease surface and interfacial tension, formation and stabilization of emulsions, improvement of stability, texture, volume and conservation of products) Source of L-rhamnose Surface conditioning
Cosmetic and pharmacy industries	Health care formulations Drug delivery systems Skin care products
Biomedicine	Antimicrobial and antiviral activity Cellular immunosuppression Wound healing, treatment prevention of gum disease and periodontal regeneration of ulcer Inhibition of growth of human breast cancer cell lines
Agronomy	Control of zoosporic plant pathogens by affecting motility, causes lysis, inhibition of spore germination and mycelium growth
Formulation of cleaners and wetting agents	Replacements for the synthetic compounds in liquids and powders cleaning formulations Superior wetting abilities compared to synthetic surfactants on different types of surfaces
Bio- and nanotechnologies	Synthesis of nanoparticles (metal nanoparticles) Drug delivery systems Formulation of microemulsions

6.1. Bioremediation

RLs have been studied and shown to have potential in bioremediation of organics, as organic flushing agents, as metal flushing agents, and in the biodegradation of organics in metal–organic co-contaminated systems.[3,10]

6.2. Biodegradation and uptake of hydrocarbons

Numerous studies confirmed that biosurfactants, especially RLs, could affect the biodegradation of hydrocarbons, both aliphatic and aromatic.[81] Furthermore, it was shown that the addition of RLs to pure cultures of *P. aeruginosa* enhanced the biodegradation of hexadecane, octadecane, *n*-paraffins and phenanthrene,[82–84]

as well as degradation in soil systems in the presence of hexadecane, tetradecane, pristane, creosote or hydrocarbon mixtures.[3] Additionally, besides the role of RLs in the biodegradation of hydrocarbons, reports showed that RLs facilitate the uptake of hydrocarbons by *P. aeruginosa*.[46,77,85,86]

Although some studies reported positive effects on the biodegradation of petroleum hydrocarbons in presence of RL biosurfactant, a lack of influence or even a negative effect of biosurfactant supplementation was observed just as frequently.[87] Some reports indicated that the potential reason for inhibition of degradation is that RLs are favored as the carbon source for bacterial metabolism.[52] Recently, it was observed that the presence of RLs, or other surfactant compounds, may induce changes in the microbial community, which in turn corresponded to differences in the degradation patterns.[87] Some earlier reports suggested mechanisms of hydrocarbon biodegradation facilitated by RL, and assumed that RLs, due to their physicochemical properties, increased the hydrocarbon solubility and bioavailability or that RLs interacted with the bacterial cell, making the cell surface more hydrophobic and easily accessible to hydrophobic substrates.[52] On the other hand, some recently published studies proposed three mechanisms of interaction between microorganisms and hydrocarbons: access to water-solubilized hydrocarbons, direct contact of cells with large oil drops and contact with pseudo-solubilized or emulsified oil,[88] as well as combinations of these interactions.[87] However, regardless of whether the biodegradation process is enhanced or inhibited, the effects are bacterial strain-specific in the sense of strain characteristics and response to environmental conditions.[83] Although much work was realized by many groups to explain the role of RLs, and biosurfactants generally, in the degradation of water-immiscible substrate, their significance and exact purpose in this process still remain unclear.

6.3. Flushing agents for organic pollution

Biodegradation of NAPL and soil-phase organics, such as PAH, is often a slow and non-feasible process.[3] The addition of surfactants to a flushing solution could enhance the flushing efficiency, either by mobilization or by an increase in the solubilization of these compounds.[3] Thus, to be effective, a surfactant must have good solubilization capacity and/or be able to reduce interfacial tension. RLs were shown to have an *MSR* for the model NAPL, hexadecane that was 20 times greater than the *MSR* for hexadecane alkyl benzyl sulfonate.[3] In studies examining the use of RLs for the removal of residual hexadecane from soil, it was shown that RL (20 % removal) was more effective than either SDS (negligible removal) or Tween 80 (6 % removal). Additionally, it was shown the optimal removal of NAPL compounds (60 %) could be achieved by altering the pH and ionic strength, thereby maximizing the reduction of the surface tension.[3] Similar results were obtained for RL solubilization of solid phase materials. For

example, the *MSR* of rhamnolipid–octadecane was ten and five times higher than the *MSR*s for Triton-X-114–octadecane and for Corexit 0600–octadecane, respectively.[3] Moreover, the *MSR* of rhamnolipid-phenanthrene was 1.7 to 2.8 times higher than for 13 different synthetic surfactants that were tested.[3] Furthermore, in a comparison of the removal of a hydrocarbon mixture (undecane, pentadecane, hexadecane, octadecane, pristane, naphthalene, phenanthrene and pyrene) from soil, RLs were more effective than Triton X-100 or Tween 60 for all hydrocarbon components.[3] Finally, RL-enhanced removal of phenanthrene, pyrene and polychlorinated biphenyls and a variety of PAH from soil were reported.[3]

6.4. Bioremediation of heavy metals

Juwarkar *et al.*[79] conducted column experiments to evaluate the potential of environmentally compatible RL biosurfactants produced by *P. aeruginosa* BS2, to remove Cd and Pb from artificially contaminated soil. Results showed that dirhamnolipid removed not only the leachable or available fraction of Cd and Pb, but also the bound metals, whereas tap water removed only the mobile fraction.[79] Washing contaminated soil with tap water revealed that only ≈ 2.7 % of Cd and 9.8 % of Pb in contaminated soil were in freely available or weakly bound forms and were able to be removed, whereas washing with RL had removed 92 % of Cd and 88 % of Pb after 36 h of leaching.[79] Wang and Mulligan[80] evaluated the feasibility of using RL foam to remove Cd and Ni from a sandy soil. Application of RL foam increased the efficiency and enabled the removal of 73.2 % and 68.1 % of Cd and Ni, respectively, whereas the RL solution alone flushed 61.7 % and 51 % of Cd and Ni, respectively.[80] Mulligan *et al.*[89] designed batch washing experiments to evaluate the feasibility of using biosurfactants to remove heavy metals from sediments. Thus, surfactant from *Bacillus subtilis*, RLs from *P. aeruginosa*, and sophorolipid from *Torulopsis bombicola* were evaluated on a sediment containing 110 mg kg^{-1} of Cu and 3300 mg kg^{-1} of Zn. A single washing with 0.5 % RL removed 65 % of the copper and 18 % of the zinc, whereas 4 % sophorolipid removed 25 % of Cu and 60 % of Zn.[89] Avramovic *et al.*[90] studied the chromium(VI) tolerance of *P. aeruginosa* NCAIM (P) B001380 and showed that the strain was chromium tolerant and had potential for application in heavy metal bioremediation.

6.5. Bioremediation of co-contaminated sites

Sandrin *et al.*[91] studied the effectiveness of RL biosurfactants in the remediation of a Cd and naphthalene co-contaminated site. They observed that reduced cadmium toxicity as a result of the addition of *P. aeruginosa* RL led to enhanced naphthalene biodegradation by *Burkholderia* sp. NCBI U37342.[91] These authors suggested that reduction of metal toxicity by RL might involve a

combination of RL complexation with cadmium and RL interaction with the cell surface to alter Cd uptake, resulting in enhanced rates of bioremediation. In another co-contaminant study, it was observed that the inhibition of phenanthrene mineralization in the presence of Cd was reduced by the pulsed addition of RL.[52] Dahrazma and Mulligan[78] reported higher rates of Cu and Ni removal from sediments by adding 1 % NaOH to a solution of RL. Efficient removal of Zn and Cu from co-contaminated soil with a 12.6 % oil content using RLs was also demonstrated.[92]

6.6. Food industry

Some properties of RLs, such as emulsion formulation and stabilization, as well as anti-adhesive and antimicrobial activity, make them interesting for the food industry as multipurpose ingredients.[4] Apart from their role as surface active agents, there are reports that RLs could have several other functions in food.[12,93] Some examples are an improvement of dough stability, texture, volume and conservation of bakery products, obtained by the addition of RL surfactants,[94] while some other authors suggested the use of RLs for improvement of properties of butter cream, croissants and frozen confectionery products.[93]

Finally, RLs could serve a source of L-rhamnose, a compound used commercially in the production of high quality flavor compounds. L-Rhamnose is a methyl pentose natural sugar, classified as one of the rarer sugars, and is found in several animal, plant and bacterial polysaccharides, as well in RLs. This compound was already successfully obtained by hydrolyzing RL surfactants produced by *P. aeruginosa*.[95] L-Rhamnose is a sugar that the Food and Drug Administration (FDA) has approved as a food additive and hence, it has found use in the flavor industry as a precursor for the production of 2,5-dimethyl-4-hydroxy--3(2*H*)-furanone, the high-quality flavor aroma furaneol (trademark of Firmenich SA, Geneva),[12] which resembles strawberry and raspberry. It is also the starting raw material in the reaction flavors developed during the preparation of various foods, such as bread, grilled meats, *etc*. Thus, there is a great deal of interest in obtaining commercial quantities of rhamnolipids to provide a source of L-rhamnose, which already has the above mentioned applications in the food industry.

6.7. Surface conditioning

Bacterial biofilms present on surfaces in the food industry are potential sources of contamination, which may lead to food spoilage and disease transmission, and thus, controlling the adherence of microorganisms to food contact surfaces is an essential step in providing safe and quality products to consumers.[96] The promising results from studies of the disruption of *Bordetella bronchiseptica* biofilm by RL[97] and reduction of adhesion of *Streptococcus salivarius* and *C. tropicalis* by RL[98] suggested a potential application of RLs for

surface conditioning in the food industry. Moreover, studies by Meylheuc *et al.*[99,100] showed inhibition of the adhesion of the pathogen *Listeria monocytogenes* to two types of surfaces typically used in the food industry using biosurfactant obtained from *P. fluorescens*, while Dagbert *et al.*[101] showed that the surfactant produced by *P. fluorescens* also has good potential as a corrosion inhibitor.

It is important to note that RLs, as products obtained from *P. aeruginosa*, which are considered to be opportunistic pathogens, still face some difficulties (particularly the long process required by regulations for the approval required by governmental agencies) related to application in food industry as food ingredients or integration of these biosurfactants in industrial processes on any large scale level.[4] This obstacle could, in the future, be prevented by using RL produced by, as already mentioned, nonpathogenic bacterial species, such as *P. chlororophasis*.

6.8. Cosmetic and pharmacy industries

Cosmetic surfactants perform detergency, wetting, emulsifying, solubilizing, dispersing and foaming effects.[15] Biodegradability, low toxicity and ecological acceptability, which, at the same time, are the benefits of a naturally derived surfactant that promises cosmetic safety are, therefore, in high demand. In particular, application of RLs in the field of cosmetics and pharmaceuticals as emulsifiers, penetrating agents and drug delivery systems is an emerging area of research.[102,103]

RLs are used in health care products in several different formulations,[15] for example, in insect repellents, antacids, acne pads, anti-dandruff products, contact lens solutions, deodorants, nail care products and toothpastes.[3,104] These formulations require surfactants with high surface activity and, in particular, emulsifying activities,[105] which is the essence of the texture consistency of these products.[48] Furthermore, requirements for the biological activities for cosmetics should expand the application of RLs, and a delivery system has been achieved, not only for emulsions but also for liposomes.[106] Patents for cosmetics containing RLs have been granted for anti-wrinkle and anti-aging products,[107] which were launched in several dosage forms as commercial skin care cosmetics[108] because of their skin compatibility and extremely low skin irritation.[48]

6.9. Biomedicine

Early on, the wide-ranging antimicrobial properties of RLs were noted. Interestingly, they were shown to be active against a large variety of bacteria, including both Gram-negative and Gram-positive species.[8] In several studies, the antimicrobial properties of mixtures of RL congeners produced by three different strains of *P. aeruginosa* were investigated.[31,33,34] The various RL combinations displayed antimicrobial activity against nearly all the tested Gram-positive spe-

cies, including *Staphylococcus*, *Mycobacterium*, and *Bacillus*, and significant activity against a number of Gram-negative species, with *Serratia marcescens*, *Enterobacter aerogenes*, and *Klebsiella pneumoniae* being especially sensitive.[29]

Rhamnolipids were also shown to affect cellular immunosuppression[109] and wound healing, treatment and prevention of gum disease and periodontal regeneration,[110] and to display differential effects on human keratinocytes and fibroblast cultures.[111] Moreover, Piljac *et al.*[112] reported the successful treatment of a decubitus ulcer with an ointment containing 0.1 % of a di-rhamnolipid.

Thanomsub *et al.*[113] tested the cytotoxic activity of a crude RL extract, Rha––Rha–C_{10}–C_{10} and Rha–Rha–C_{10}–C_{12}, produced by *P. aeruginosa* B189 isolated from a milk factory, against herpes simplex virus, insect and cancer cell lines. Rha–Rha–C_{10}–C_{10} exhibited significant inhibition of growth of human breast cancer cell lines (MCF-7), with minimum inhibitory concentration (*MIC*) of 6.25 μg mL^{-1}, Rha–Rha–C_{10}–C_{12} had an *MIC* of 50 μg mL^{-1} against insect cell line C6/36, while the crude RL extract showed no cytotoxic activity.[113] The potential mechanism of activity, regarding the structure of the biosurfactant, is a toxicological effect on the cell membrane permeability. Furthermore, Rha–Rha––C_{10}–C_{10} and Rha–Rha–C_{10}–C_{12} had no inhibition effect on the normal cell line (Vero cell) at concentrations up to 50 μg mL^{-1}. This confirmed the specific toxicity of these compounds to the cell lines used. However, the inhibitory mechanisms against these cell lines are as yet unknown and are under investigation.[113]

6.10. Agronomy

RLs also showed the ability to control certain zoosporic plant pathogens, including *Phytophthora cryptogea* and *Pythium* spp.[114–116] Purified mono- and di-RL, in concentrations ranging from 5 to 30 mg L^{-1}, caused cessation of motility and lysis of the entire zoosporic population in less than 1 min.[3] This observation led to the development of a RL-containing biofungicide formulation, used to prevent crop contamination by pathogenic fungi.[5] This product is considered to be non-mutagenic and of low acute toxicity to mammals. It was approved by the FDA for direct use on vegetables, legumes and fruit crops.[4] Dorey *et al.*[117] reported the role of RLs in triggering defense responses and protection against the fungus *Botrytis cinerea* in grapevines.[117] The authors showed that RLs inhibited spore germination and mycelium growth, thus efficiently protecting grapevines against the fungus by inducing the plant defense system. A product based on an aqueous RL solution (0.01 %) was claimed to act as a novel agent for stimulating the natural defense reactions of plants against pathogenic fungi.[5,117]

6.11. Formulation of cleaners and wetting agents

One of the major commercial domestic applications of biosurfactants is in the field of cleaning and laundry products. The interfacial chemistry created by

adsorbed surface-active molecules, of either biological or chemical origin, dominates the end-use properties of materials in many different applications.[59] The properties and efficacy of detergent formulations are critically dependent on the interfacial activity of various surfactants, which are present in their composition.[64] At present, the liquids and powders generally contain alkyl sulfonates, such as linear alkylbenzene sulfonates, but the glycolipid biosurfactants and among them rhamnolipids, produced by *P. aeruginosa*, are possible candidates to be used for the, at least, partial replacements of these synthetic compounds.[68] One of the major challenges in the use of RLs is the fact that they are produced as a mixture of different congeners, which affects physicochemical properties and behavior, as mentioned previously.

Bafghi and Fazaelipoor[118] investigated RLs in the formulation of a washing powder. The results showed that the biosurfactant was effective in removal of oil from the samples. The formulation presented in this study was also compared with some commercial powders for the removal of edible oil, chocolate and albumen stains. The results showed that the RL-inclusive formulation was comparable to the commercial powders in terms of stain removal. Biodegradability tests performed on pure RL and the RL-inclusive formulation confirmed the good biodegradability of this biosurfactant.

Özdemir and Malayoglu[64] investigated the wetting behavior of a mixture of mono- and di-rhamnolipid (in 1:1 ratio of mono:di-rhamnolipid) on glass, PET and gold surfaces by measuring the advancing contact angle, and elucidated the preferences of the surfactant molecules adsorbed onto SL–SV and L–V interfaces, with SDS as the reference surfactant. The study showed that at low concentrations of RL and reference surfactant, the contact angle varied in a certain range depending on the character of the surfactant interactions with the surface.[64] Moreover, on hydrophobic surfaces, the adhesion tension had a specific dynamic, depending on surfactant concentration, while on hydrophilic surfaces a steady decrease in adhesion tension was observed with both surfactant solutions.[64]

Costa *et al.*[59] also studied the wetting behavior of RLs produced by *P. aeruginosa* LBI grown on a waste oil substrate, and the chemical surfactant SDS, on glass, PET, poly(vinyl chloride) (PVC), poly(ε-caprolactone) (PCL) and a polymer blend (PVC–PCL) by measuring the contact angle of sessile drops. The comparison of the wetting profiles showed dynamic changes in the contact angle at low SDS and RL concentrations – the contact angle increased and when the concentration of the surfactant increased further, the contact angle decreased.[59] The results showed that RLs produced by *P. aeruginosa* LBI exhibited superior wetting abilities compared to SDS. This is the first work that evaluated the wetting properties of RLs on polymer blends.

6.12. Bio- and nano-technology

Biosurfactants have been increasingly attracting attention in the field of nanotechnology as a "green" alternative for high-performance nanomaterials.[17] During the last decade, unique properties of biosurfactants, including versatile self-assembling and biochemical properties, which do not usually occur among chemically derived surfactants, were studied and analyzed.[119–121] RLs, alone or in combination with other glycolipid biosurfactants, have potential roles as systems for drug delivery, synthesis of nanoparticles and formulation of microemulsions.

6.12.1. Drug delivery systems

In 1988, RL liposomes were patented as drug delivery systems, useful as microcapsules for drugs, proteins, nucleic acids, dyes and other compounds, as biomimetic models for biological membranes and as sensors for detecting pH variations.[17] These novel liposomes were described as safe and biologically decomposable, with suitable affinity for biological organisms, stable and with long service and shelf life.[17] Recently, in a study of Sharma *et al.*,[122] RLs and sophorolipids were mixed with lecithins to prepare biocompatible microemulsions in which the phase behavior was unaffected by changes in temperature and electrolyte concentration, making them desirable for cosmetic and drug delivery applications.[17]

6.12.2. Synthesis of nanoparticles

Another interesting aspect of the applications of RL is the synthesis of metal nanoparticles as an alternative (a more ecological technology) to traditional methods of production.[121] There are several reports with RL applications in this field. Kumar *et al.*[123] synthesized silver nanoparticles using purified RLs from *P. aeruginosa* BS-161R, which showed a broad spectrum of antimicrobial activity. Xie *et al.*[124] successfully synthesized silver nanoparticles in RL reverse micelles. Palanisamy and Raichur[125] demonstrated a simple and eco-friendly method for synthesis of spherical nickel oxide nanoparticles by a microemulsion technique using RLs as an alternative surfactant.

In two recent studies, RL biosurfactants were used as capping agents for the synthesis of ZnS nanoparticles.[126,127] Narayanan *et al.*[126] demonstrated a novel method for the synthesis of ZnS nanoparticles in aqueous medium and showed that RL biosurfactant has potential as an effective capping agent for the synthesis of uniform nanoparticles. Hazra *et al.*[127] reported a facile eco-friendly procedure for biosynthesis of RL capped ZnS nanoparticles, their structural characterization, biocompatibility, cytotoxicity assessment and their applicability as a nanophotocatalyst for the degradation of a textile azo dye. The results obtained explained the importance of environmentally friendly RLs as an effective and

inexpensive capping and stabilizing agent for the development of stable and biocompatible ZnS nanoparticles as nanophotocatalysts in the textile industry and for wastewater and effluent treatment.[127]

6.12.3. Microemulsions

Xie *et al.*[128] showed that RLs have potential for application in the formulation of microemulsions with medium-chain alcohols as co-surfactants. Furthermore, the same authors observed that the phase behavior and microstructure of these microemulsions were related to the conformational changes of the RL molecules at the oil/water interface.[128,129] In another study, RLs were successfully used as the surfactant to synthesize spherical nickel oxide nanoparticles by a microemulsion technique.[125]

Nguyen and Sabatini[130] focused their research on developing alcohol-free biosurfactant-based microemulsions. RL-based mixtures were found to have doubled the solubilization parameter as compared to sodium bis(2-ethylhexyl) sulfosuccinate/sodium dihexyl sulfosuccinate/sodium mono- and di-methylnaphthalene sulfonate at the same total molar concentration.[130] Additionally, these authors developed a phase diagram for surfactant mixtures containing methyl ester ethoxylate, RL and oleyl alcohol with limonene oil, which could be used as a guideline for selecting a surfactant system and surfactant ratio to formulate microemulsions with a given oil. The RL biosurfactant used by Nguyen and Sabatini[130] was the least hydrophobic type (fatty acid tails of C_8 chain length), since its tail length was the shortest within the typical range (the tail length usually varies from C_8 to C_{14}). Potential application of longer tail RL would make the surfactant system more hydrophobic and, as a result, either the optimum formulation at lower salinity for the studied oils would be achieved, or it would be easier to formulate microemulsions with oils that are more hydrophobic than the studied oils.[130] In a recent report, Nguyen *et al.*[121] formulated and evaluated microemulsions of lecithin/RL/sophorolipid biosurfactants with a range of oils.

7. CONCLUDING REMARKS AND FUTURE PERSPECTIVE

Next to the already established biosurfactants alkyl polyglycosides and sophorolipids, RLs shows the highest economic potential among all the currently investigated microbial biosurfactants.[9] The reasons for this are their environmental friendly properties (biodegradability, low aquatic toxicity, production from renewable resources or industrial wastes), as well as additional benefits of their physicochemical characteristics and biological origin. In spite of the several drawbacks discussed above, strains of *P. aeruginosa* are the most promising candidates for RL production, because they can metabolize a variety of carbon sources, including renewable sources, with good yields of RL surfactants and their RL mixtures contain all types of RL structures. In comparison to other gly-

colipid biosurfactants, RLs have a broader range of applications, from environmental to industrial spectra. The increasing commercial importance of RL is suggested by the number of companies that are official manufacturers of RL biosurfactant produced by *P. aeruginosa*, and the great number of related patents. However, in the near future, it seems that the most likely progress in the application of RLs will be in the field of bioremediation, biodegradation of hydrocarbons and removal of heavy metals. In addition to the research focused on commercial application of RLs in bioremediation, further investigation and understanding of the mechanisms of action of these compounds in the environment appears to be of importance.

Acknowledgment. This work was supported by Project III 43004, funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

ИЗВОД

РАМНОЛИПИДНИ БИОСУРФАКТАНТ ИЗ *Pseudomonas aeruginosa* – ОД ОТКРИЋА ДО ПРИМЕНЕ У САВРЕМЕНОЈ ТЕХНОЛОГИЈИ

МИЛЕНА Г. РИКАЛОВИЋ[1], МИРОСЛАВ М. ВРВИЋ[1] и ИВАНКА М. КАРАЦИЋ[1]

[1]*Хемијски факултет, Универзитет у Београду, Студентски трг 12–16, Београд и* [2]*Медицински факултет, Институт за хемију у медицини, Универзитет у Београду, Вишеградска 26, Београд*

Рамнолипиди су, највероватније, следећа генерација биосурфактаната која ће доминирати на тржишту сурфактаната. Налазе се одмах иза алкил-полигликозида, који су већ прихваћени на тржишту, и софоролипида, који се за сада налазе у неколико формулација за чишћење. Ипак, највећи број нових публикација и патената везаних за гликолипидне биосурфактанте се односи на рамнолипиде. Главни продуценти рамнолипида су *Pseudomonas aeruginosa.* Ови биосурфактанти су смеше различитих рамнолипидних структура, које показују физичко-хемијске особине различите од појединачних структура, при чему најзаступљеније рамнолипидно једињење има највећи утицај на укупне карактеристике смеше. Особине рамнолипида, као што су биоразградивост, ниска токсичност, продукција из обновљивих извора, антимикробна (посебно антифунгална) активност, заједно их чине потенцијално погодним за широку комерцијалну примену. До сада је главна примена рамнолипида била у области биоремедијације, а растући број патената везаних за примену рамнолипида у козметичкој, агро- и прехрамбеној индустрији, формулацијама за чишћење и нано-технологији указују да је њихова будућност имплементација у овим областима.

(Примљено 27. јуна, ревидирано 19. септембра, прихваћено 20. септембра 2014)

REFERENCES

1. M. Pacwa-Płociniczak, M. G. A. Płaza, Z. Piotrowska-Seget, S. S. Cameotra, *Int. J. Mol. Sci.* **12** (2011) 633
2. J. D. Desai, I. M. Bant, *Microbiol. Mol. Biol. Rev.* **61** (1997) 47
3. R. M. Maier, G. Soberon-Chávez, *Appl. Microbiol. Biotechnol.* **54** (2000) 625
4. M. Nitschke, S. G. V. A. O. Costa, *Trends Food Sci. Tech.* **18** (2007) 252
5. M. Nitschke, S. G. V. A. O. Costa, J. Contiero, *Process Biochem.* **46** (2011) 621
6. J. D. van Hamme, A. Singh, O. P. Ward, *Biotechnol. Adv*. **24** (2006) 604

7. H. Górna, Ł. Lawniczak, A. Zgoła-Grześkowiak, E. Kaczorek, *Bioresour. Technol.* **10** (2011) 3028
8. A. M. Abdel-Mawgoud, F. Lépine, E. Déziel, *Appl. Microbiol. Biotechnol.* **86** (2010) 1323
9. M. M. Müller, J. H. Kügler, M. Henkel, M. Gerlitzki, B. Hörmanna, M. Pöhnlein, C. Syldatk, R. Hausmann, *J. Biotechnol*. **162** (2012) 366
10. M. G. Rikalovic, M. M. Vrvic, I. M. Karadzic, in *Bioremediation: Challenges, Future and Prospectives*, J. B. Velázquez-Fernández, S. Muñiz-Hernández, Eds. Nova Science Publishers, New York, 2014, p. 299
11. P. Singh, S. S. Cameotra, *Trends Biotechnol*. **22** (2007) 142
12. K. Muthusamy, S. Gopalakrishnan, T. K. Ravi, P. Sivachidambaram, *Curr. Sci*. **94** (2008) 736
13. I. N. A. van Bogaert, W. Soetaert, in *Microbiology Monographs 20 – Biosurfactants: From Genes to Applications*, G. Soberón-Chávez, A. Steinbüchel, Eds., Springer, Berlin, 2011, p. 179
14. E. Rosenberg, E. Z. Ron, *Appl. Microbiol. Biotechnol*. **52** (1999) 154
15. N. Lourith, M. Kanlayavattanakul, *Int. J. Cosmetic Sci.* **31** (2009) 255
16. A. Franzetti, I. Gandolfi, G. Bestetti, T. J. P. Smyth, I. M. Banat, *Eur. J. Lipid Sci. Technol*. **112** (2010) 617
17. L. Fracchia, M. Cavallo, M. G. Martinotti, I. M. Banat, in *Biomedical Science, Engineering and Technology*, D. N. Ghista, Ed., InTech, Rijeka, 2012, p. 326
18. J. Arutchelvi, M. Doble, in *Microbiology Monographs 20 – Biosurfactants: From Genes to Applications*, G. Soberón-Chávez, A. Steinbüchel, Eds., Springer, Berlin, Germany, 2011, p. 145
19. I. M. Banat, A. Franzetti, I. Gandolfi, G. Bestetti, M. G. Martinotti, L. Fracchia, T. J. Smyth, R. Marchant, *Appl. Microbiol. Biotechnol.* **187** (2010) 427
20. S. Bergström, H. Theorell, H. Davide, *Arkiv Chem. Mineral Geol. 23A* **13** (1946) 1
21. S. Bergström, H. Theorell, H. Davide, *Arch. Biochem. Biophys*. **10** (1946) 165
22. G. Hauser, M. L. Karnovsky, *J. Bacteriol*. **68** (1954) 645
23. F. G. Jarvis, M. J. Johnson, *J. Am. Chem. Soc*. **71** (1949) 4124
24. J. R. Edwards, J. A. Hayashi, *Arch. Biochem. Biophys*. **111** (1965) 415
25. N. Shaw, *Microbiol. Mol. Biol. Rev*. **34** (1970) 365
26. S. Arino, R. Marchal, J. P. Vandecasteele, *Appl. Microbiol. Biotechnol.* **45** (1996) 162
27. D. Onbasli, B. Aslim, *J. Environ. Biol*. **30** (2009) 161
28. N. W. Gunther, A. Nunez, W. Fett, D. K. Y. Solaiman, *Appl. Environ. Microbiol*. **71** (2005) 2288
29. A. M. Abdel-Mawgoud, R. Hausmann, F. Lépine, M. M. Műller, E. Déziel, in *Microbiology Monographs 20 – Biosurfactants: From Genes to Applications*, G. Soberón-Chávez, A. Steinbüchel, Eds., Springer, Berlin, 2011, p. 13
30. A. Tahzibi, F. Kamal, M. M. Assadi, *Iran. Biomed. J.* **8** (2004) 25
31. A. Abalos, A. Pinazo, M. R. Infante, M. Casals, F. Garcia, A. Manresa, *Langmuir* **17** (2001) 1367
32. A. M. Abdel-Mawgoud, M. M. Aboulwafa, N. A. H. Hassouna, *Appl. Biochem. Biotechnol*. **157** (2009) 329
33. M. Benincasa, A. Abalos, I. Oliveira, A. Manresa, *Antonie Van Leeuwenhoek* **85** (2004) 1
34. E. Haba, A. Abalos, O. Jauregui, M. J. Espuny, A. Manresa, *J. Surfactants Deterg*. **6** (2003) 155
35. J. C. Mata-Sandoval, J. Karns, A. Torrents, *J. Chromatogr*. **864** (1999) 211

36. O. Pornsunthorntawee, P. Wongpanit, S. Chavadej, M. Abe, R. Rujiravanit, *Bioresour. Technol.* **99** (2008) 1589
37. M. Rikalovic, G. Gojgic-Cvijovic, M. M. Vrvic, I. Karadzic, *J. Serb. Chem. Soc.* **77** (2012) 27
38. M. G. Rikalovic, A. M. Abdel-Mawgoud, E. Déziel, G. Dj. Gojgic-Cvijovic, Z. Nestorovic, M. M. Vrvic, I. M. Karadzic, *J. Surfactants Deterg.* **16** (2013) 673
39. M. Benincasa, J. Contiero, M. A. Manresa, I. O. Moraes, *J. Food Eng.* **54** (2002) 283
40. M. Benincasa, F. R. Accorsini, *Bioresour. Technol*. **99** (2008) 3843
41. K. S. M. Rahman, T. J. Rahman, S. McClean, R. Marchant, I. M. Banat, *Biotechol. Prog.* **18** (2002) 1277
42. S. G. V. A. O. Costa, M. Nitschke, R. Haddad, M. N. Eberlin, J. Contiero, *Process Biochem*. **41** (2006) 483
43. E. Déziel, F. Lépine, D. Dennie, D. Boismenu, O. A. Mamer, R. Villemur, *Biochim. Biophys. Acta* **1440** (1999) 244
44. M. Nitschke, S. G. Costa, J. Contiero, *Appl. Biochem. Biotechnol.* **160** (2010) 2066
45. A. Aparna, G. Srinkethan, H. Smitha, *Colloids Surfaces, B* **95** (2012) 23
46. M. Nie, X. Yin, C. Ren, Y. Wang, F. Xu, Q. She, *Biotechnol. Adv*. **28** (2010) 635
47. D. Dubeau, E. Déziel, D. Woods, F. Lépine, *BMC Microbiol.* **9** (2009) 263
48. E. Haba, A. Pinazo, O. Jauregui, M. J. Espuny, M. R. Infante, A. Manersa, *J. Surfact. Deterg.* **6** (2003) 155
49. M. M. Müller, B. Hörmann, M. Kugel, C. Syldatk, R. Hausmann, *Appl. Microbiol. Biotechnol*. **89** (2011) 585
50. A. Lebron-Paler, J. E. Pemberton, B. Becker, W. H. Otto, C. K. Larive, R. M. Maier, *Anal. Chem*. **78** (2006) 7649
51. M. Nitschke, S. G. V. A. O. Costa, J. Contiero, *Biotechnol. Prog*. **21** (2005) 1593
52. P. Maslin, R. M. Maier, *Biorem. J.* **4** (2000) 295
53. M. Heyd, A. Kohnert, T.-H. Tan, M. Nusser, F. Kirschhöfer, G. Brenner-Weiss, M. Franzreb, S. Berensmeier, *Anal. Bioanal. Chem.* **391** (2008) 1579
54. I. Siegmund, F. Wagner, *Biotechnol. Tech*. **5** (1991) 265
55. E. V. Chandrasekaran, J. N. BeMiller, in *Methods in Carbohydrate Chemistry*, R. L. Whistler, Ed., Academic Press, New York, 1980, p. 89
56. J. E. Hodge, B. T. Hofreiter, in *Methods in Carbohydrate Chemistry*, R. L. Whistler, M. L. Wolfrom, Eds., Academic Press, New York, 1962, p. 380
57. E. Déziel, F. Lépine, S. Milot, R. Villemur, *Biochim. Biophys. Acta* **1485** (2000) 145
58. N. P. J. Price, K. J. Ray, K. Vermillion, T. M. Kuo, *Carbohydr. Res.* **344** (2009) 204
59. S. G. V. A. O. Costa, S. R. de Souza, M. Nitschke, S. M. M. Franchetti, M. Jafelicci., R. B. Lovaglio, J. Contiero, *J. Surfactants Deterg*. **12** (2009) 125
60. W. Griffin, *J. Soc. Cosmet. Chem*. **5** (1954) 235
61. D. K. Jain, D. L. Collins-Thompson, H. Lee, J. T. Trevors, *J. Microbiol. Methods* **13** (1991) 271
62. M. Morikawa, Y. Hirata, T. Imanaka, *Biochim. Biophys. Acta* **1488** (2000) 211
63. E. A. Vogler, *Langmuir* **8** (1992) 2005
64. G. Özdemir, U. Malayoglu, *Colloids Surfaces, B* **39** (2004) 1
65. M. M. Müller, R. Hausmann, *Appl. Microbiol. Biotechnol*. **91** (2011) 251
66. S. Maneerat, *J. Sci. Technol.* **27** (2005) 675
67. S. Mukherjee, P. Das, R. Sen, *Trends Biotechnol.* **24** (2006) 509
68. R. Marchant, I. M. Banat, *Trends Biotechnol*. **30** (2012) 558

69. D. H. Dusane, S. S. Zinjarde, S. Smita, V. P. Venugopalan, R. J. C. McLean, M. M. Weber, P. K. S. M. Rahman, *Biotechnol. Genet. Eng. Rev.* **27** (2010) 159
70. G. Soberon-Chavez, F. Lepine, E. Déziel, *Appl. Microbiol. Biotechnol*. **68** (2005) 718
71. K. Dubey K, A. Juwarkar, *World J. Microbiol. Biotechnol*. **17** (2001) 61
72. A. B. Moldes, A. M. Torrado, M. T. Barral, J. M. Domianguez, *J. Agric. Food Chem.* **55** (2007) 4481
73. Z. A. Raza, A. Rehman, M. S. Khan, Z. M. Khalid, *Biodegradation* **18** (2007) 115
74. M. E. Mercade, M. A. Manresa, M. Robert, M. J. Espuny, C. De Andres, J. Guinea, *Bioresour. Technol.* **43** (1993) 1
75. E. Haba, M. J. Espuny, M. Busquets, A. Manresa, *J. Appl. Microbiol*. **88** (2000) 379
76. P. D. Mukherjee, R. Sen, *Trends Biotechnol*. **24** (2006) 509
77. S. S. Cameotra, P. Singh, *Int. Biodeter. Biodegr*. **62** (2008) 274
78. B. Dahrazma, C. N. Mulligan, *Chemosphere* **69** (2007) 705
79. A. A. Juwarkar, K. V. Dubey, A. Nair, S. K. Singh, *Indian J. Microbiol*. **48** (2008) 142
80. S. Wang, C. N. Mulligan, *Water Air Soil Pollut*. **157** (2004) 315
81. R. M. Miller, in *Bioremediation: science and applications*, H. D. Skipper, R. F. Turco, Eds., Soil Science Society of America, Madison, WI, 1995, p. 322
82. G. S. Shreve, S. Inguva, S. Gunnam, *Mol. Mar. Biol. Biotechnol*. **4** (1995) 331
83. Y. M. Zhang, R. M. Miller, *Appl. Environ. Microbiol.* **60** (1994) 2101
84. Y. M. Zhang, R. M. Miller, *Appl. Environ. Microbiol.* **61** (1995) 2247
85. R. Beal, W. B. Betts, *J. Appl. Microbiol*. **89** (2000) 158
86. W. H. Noordman, D. B. Janssen, *Appl. Environ. Microbiol*. **68** (2002) 4502
87. Ł. Chrzanowski, Ł. Ławniczak, K. Czaczyk, *World J. Microbiol. Biotechnol*. **28** (2012) 401
88. A. Franzetti, I. Gandolfi, G. Bestetti, T. J. Smyth, I. M. Banat, *Eur. J. Lipid Sci. Tech.* **112** (2010) 617
89. C. N. Mulligan, R. N. Yong, B. F. Gibbs, *J. Hazard. Mat*. **85** (2001) 111
90. N. S. Avramovic, S. D. Nikolic-Mandic, I. M. Karadzic, *J. Serb. Chem. Soc.* **78** (2013) 639
91. R. T. Sandrin, M. C. Andrea, R. M. Maier, *Appl. Environ. Microbiol.* **66** (2000) 4585
92. C. N. Mulligan, C. N. Yong, B. F. Gibbs, *Environ. Prog.* **18** (1999) 31
93. N. Kosaric, *Food Technol. Biotechnol.* **39** (2001) 295
94. I. P. H. Van Haesendonck, E. C. A. Vanzeveren, (International application patent PCT), W.O. 2004/040984 (2004)
95. R. J. Linhardt, R. Bakhit, L. Daniels, F. Mayerl, W. Pickenhagen, *Biotechnol. Bioeng*. **33** (1989) 365
96. S. K. Hood, E. A. Zottola, *Food Control* **6** (1995) 9
97. Y. Irie, G. A. O'Toole, M. H. Yuk, *FEMS Microbiol. Lett.* **250** (2005) 237
98. L. R. Rodrigues, I. M. Banat, H. C. Mei, J. A. Teixeira, R. Oliveira, *J. Appl. Microbiol.* **100** (2006) 470
99. T. Meylheuc, C. J. van Oss, M. N. Bellon-Fontaine, *J. Appl. Microbiol*. **91** (2001) 822
100. T. Meylheuc, C. J. van Oss, M.N. Bellon-Fontaine, *Int. J. Food Microbiol*. **109** (2006) 71
101. C. Dagbert, T. Meylheuc, M. N. Bellon-Fontaine, *Electrochim. Acta* **51** (2006) 5221
102. A. Ortiz, J. A. Teruel, M. J. Espuny, A. Marqués, A. Manresa, F. J. Aranda, *Int. J. Pharm*. **325** (2006) 99
103. G. Piljac, V. Piljac, (U.S. Patent), 5,455,232 (1995)
104. V. Kleckner, N. Kosaric, in *Biosurfactants: Production, Properties, Applications*, N. Kosaric, Ed., Marcel Dekker, New York, 1993, p. 373

105. E. Vasileva-Tonkova, D. Galabova, E. Karpenko, A. Shulga, *Lett. Appl. Microbiol.* **33** (2001) 280
106. Y. Ishigami, Y. Gama, H. Nagahora, T. Hongu, M. Yamaguchi, (U.S. Patent), 4,902,512 (1988)
107. T. Piljac, G. Piljac, (Europan Patent Office), E.P. 1, 889,623 A2 (1999)
108. K. Desanto, (Aurora Advanced Beauty Labs St. Petersburg, FL), Word Patent 2008/013899, (2008)
109. G. Piljac, V. Piljac, (U.S. Patent) 5,466,675 (1995b)
110. T. Stipcevic, T. Piljac, J. Piljac, T. Dujmic, D. Piljac, (PCT Patent Application) WO 01/10447 A1 (2001)
111. T. Stipcevic, T. Piljac, R. R. Isseroff, *J. Dermatol. Sci.* **40** (2005) 141
112. A. Piljac, T. Stipcevic, J. Piljac-Zegarac, G. Piljac, *J. Cutan. Med. Surg.* **12** (2008) 142
113. B. Thanomsub, W. Pumeechockchai, A. Limtrakul, P. Arunrattiyakorn, W. Petchleelaha, T. Nitoda, H. Kanzaki, *Bioresour. Technol.* **98** (2007) 1149
114. K. De Jonghe, I. De Dobbelaere, R. Sarrazyn, M. Höfte, *Plant Pathol.* **54** (2005) 219
115. M. Perneel, L. D'hondt, K. De Maeyer, A. Adiobo, K. Rabaey, M. Höfte, *Environ. Microbiol.* **10** (2008) 778
116. A. L. Varnier, L. Sanchez, P. Vatsa, L. Boudesocque, A. Garcia-Brugger, F. Rabenoelina, A. Sorokin, J. H. Renault, S. Kauffmann, A. Plugin, C. Clément, F. Baillieul, S. Dorey, *Plant Cell Environ.* **32** (2009) 178
117. S. Dorey, F. Baillieul, C. Clement, A. Varnier, M. Terganier, (International Patent), WO 2007/045767 (2007)
118. M. K. Bafghi, M. H. Fazaelipoor, M. Hassan, *J. Surfactants Deterg.* **15** (2012) 679
119. D. Kitamoto, T. Morita, T. Fukuoka, M. Konishi, T. Imura, *Curr. Opin. Colloid Interface Sci.* **14** (2009) 315
120. D. Kitamoto, K. Toma, M. Hato, in *Handbook of nanostructured biomaterials and their applications in nanobiotechnology*, Vol. 1, H. Singh Nalwa, Ed., American Science Publishers, Valencia , CA, USA, 2005, p. 239
121. T. T. L. Nguyen, A. Edelen, B. Neighbors, D. A. Sabatini, *J. Colloid Interface Sci.* **348** (2010) 498
122. V. K. Sharma, R. A. Yngard, Y. Lin, *Adv. Colloid Interface Sci.* **145** (2009) 83
123. C. G. Kumar, S. K. Mamidyala, B. Das, B. Sridhar, G. S. Devi, M. S. Karuna, *J. Microbiol. Biotechnol.* **20** (2010) 1061
124. Y. Xie, R. Ye, H. Liu, *Colloids Surfaces, A* **279** (2006) 175
125. P. Palanisamy, A. M. Raichur, *Mater. Sci. Eng. C* **29** (2009) 199
126. J. Narayanan, R. Ramji, H. Sahu, P. Gautam, *IET Nanobiotechnol.* **4** (2010) 29
127. C. Hazra, D. Kundu, A. Chaudhari, T. Jana, *J. Chem. Technol. Biotechnol.* **88** (2013) 1039
128. Y. Xie, Y. Li, R. Ye, *J. Dispers. Sci. Technol.* **26** (2005) 455
129. Y. W. Xie, R. Q. Ye, H. L. Liu, *Colloids Surfaces, A* **292** (2007). 189
130. T. T. Nguyen, D. A. Sabatini, *J. Surfactants Deterg.* **12** (2009) 109.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 305–313 (2015)
JSCS–4717
UDC 547.587.51+547.77+542.913:537.5–962:615.28–188
Original scientific paper

Microwave-assisted synthesis of some new coumarin–pyrazoline hybrids and their antimicrobial activity

DONGAMANTI ASHOK*, BOMMIDI VIJAYA LAKSHMI, SIDDA RAVI and ARRAM GANESH

Department of Chemistry, Osmania University, Hyderabad, India

(Received 21 June, revised 15 October, accepted 20 October 2014)

Abstract: A series of pyrazolines were synthesized by Michael addition of chalcones with hydrazine hydrate in the presence of sodium acetate under conventional heating and microwave irradiation. The structures of the newly synthesized chalcones and pyrazolines were established based on IR, ^{1}H- and ^{13}C--NMR and mass spectral data. All the synthesized compounds were screened for their antimicrobial activity. Some of the compounds showed very good activity compared to standard drugs against all pathogenic bacteria and fungi.

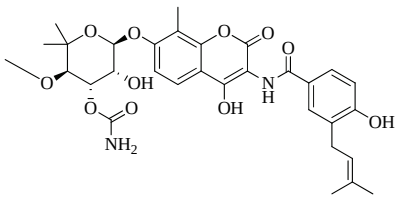
Keywords: coumarin; chalcones; Michael addition; microwave irradiation; pyrazolines.

INTRODUCTION

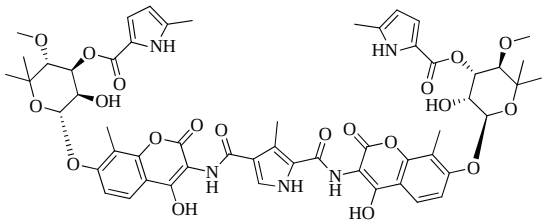
Coumarins are plant flavonoids widely distributed in nature. Coumarins are important oxygen-containing fused heterocycles used in drugs and dyes.[1] Natural coumarins are known to have antidiabetic activity,[2] anabolic, antioxidant and hepatoprotective activities.[3] Substituted coumarin derivatives were reported to have a variety of biological activities, such as anti-inflammatory,[4] antimicrobial,[5] antioxidant,[6] anticancer[7] and antiviral[8] activities. The potent antibiotics novobiocin and coumaromycin (Fig. 1) are coumarin derivatives. Compounds with a backbone of chalcones were reported to possess various biological activities, such as antimicrobial, anti-inflammatory, analgesic, antiplatelet, anti-ulcerative, antimalarial, anticancer,[9] antiviral, antileishmanial, antioxidant,[10] antitubercular,[11] antihyperglycemic, immunomodulatory, inhibition of the release of chemical mediators,[12] inhibition of leukotriene B4,[13] inhibition of tyrosinase[14] and inhibition of aldose reductase,[15] and estrogenic activities.[16] Licochalcone (Fig. 1) was found to exhibit antimalarial activity.[17] Chalcones are used as starting materials for the synthesis of various chemicals, including plastics, resins, pesti-

* Corresponding author. E-mail: ashokdou@gmail.com
doi: 10.2298/JSC140021101A

cides, dyes and pharmaceuticals.[18] Hence, the synthesis of chalcones has generated vast interest among organic as well as medicinal chemists.

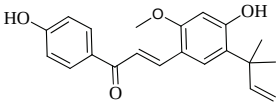


Novobiocin



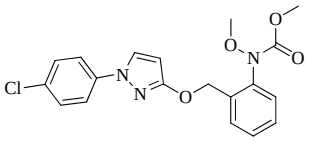
Coumermycin A_1

Coumarin core embedded drugs.



Licochalcone A

A drug based on chalcone motif



Pyraclostrobin

pyrazoline based antifungal agrochemical

Fig. 1. Biologically active compounds embedded with coumarin, chalcone and pyrazoline motifs.

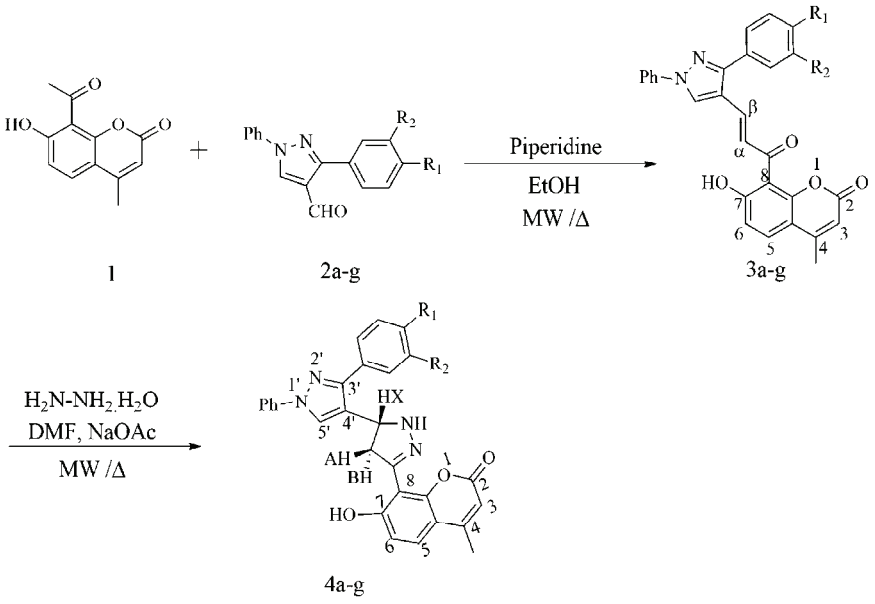
The existing literature is rich with progressive findings concerning the synthesis and pharmacological properties of pyrazolines. Pyrazolines were found to possess antimicrobial,[19] antibacterial,[20] anti-amoebic,[21,22] antidepressant,[23] anticonvulsant,[24] anti-inflammatory,[25,26] and antitumor activities. Pyraelestrobin (Fig. 1) is one of the most important fungicides on the active agrochemicals market.

Microwave irradiation has gained popularity in the past decade as a powerful tool for the rapid and efficient synthesis of a variety of compounds, resulting from the selective absorption of microwave energy by polar molecules.[27] The application of microwave irradiation provides enhanced reaction rates and improved product yields in organic synthesis and it is proving quite successful in the formation of a variety of carbon–heteroatom bonds. Recently, considerable efforts have been made in the design and realization of innovative synthetic protocols in organic synthesis, whereby a more eco-sustainable approach was adopted.[28–30] As part of an ongoing research program, the syntheses of new biologically active pyrazoline derivatives, substituted (*E*)-8-[3-(1,3-diphenyl-1*H*--pyrazol-4-yl)-1-oxo-2-propen-1-yl]-7-hydroxy-4-methyl-2*H*-1-benzopyran-2-one and substituted 8-(4′,5′-dihydro-1,3-diphenyl[4,5′-bi-1*H*-pyrazol]-3′-yl)-7-hyd-

roxy-4-methyl-2*H*-1-benzopyran-2-one, under conventional and non-conventional (microwave irradiation) methods, are reported herein.

RESULTS AND DISCUSSION

In this article, the syntheses are reported of new biologically active pyrazoline derivatives, substituted (*E*)-8-[3-(1,3-diphenyl-1*H*-pyrazol-4-yl)-1-oxo-2--propen-1-yl]-7-hydroxy-4-methyl-2*H*-1-benzopyran-2-one (**3a**–**g**) and substituted 8-(4′,5′-dihydro-1,3-diphenyl[4,5′-bi-1*H*-pyrazol]-3′-yl)-7-hydroxy-4-methyl-2*H*--1-benzopyran-2-one (**4a**–**g**), under conventional and non-conventional (microwave irradiation) methods. The synthesis of new derivatives of pyrazoline (**4a**–**g**) was realized as outlined in Scheme 1. The chalcones were prepared by reacting 8-acetyl-7-hydroxy-4-methylcoumarin (**1**) with substituted 1,3-diphenyl-1*H*-pyrazole-4-carboxaldehydes **2a**–**g** in presence of piperidine by conventional method as well as microwave irradiation using the Claisen–Schmidt condensation. The Michael addition of **3a**–**g** with hydrazine hydrate in DMF was realized by heating conventionally and irradiating with microwaves in the presence of sodium acetate affording the new pyrazoline derivatives **4a**–**g**.



Scheme 1. Synthetic route for the preparation of the coumarin–pyrazoline hybrids **4a**–**g**; **a**, R_1 = H, R_2 = H; **b**, R_1 = OCH_3, R_2 = H; **c**, R_1 = OCH_3, R_2 = OCH_3; **d**, R_1 = CH_3, R_2 = H; **e**, R_1 = F, R_2 = H; **f**, R_1 = Cl, R_2 = H; **g**, R_1 = Br, R_2 = H.

It was found that the synthesis of chalcones **3a**–**g** and pyrazolines **4a**–**g** by conventional method took a longer time and gave lower yields (Table I), when compared to the microwave irradiation technique in which the reaction proceeded

smoothly with excellent yields, within a few minutes for the pyrazolines **4a**–**g** and in 10–15 min for the chalcones **3a**–**g** (Table I).

TABLE I. Comparisons of the yields of the synthesized compounds **3a**–**g** and **4a**–**g**

Compd.	M.p., °C	Conventional method		MWI	
		t / h	Yield, %	*t* /min	Yield, %
3a	202	24	50	10	80
3b	242	20	51	15	80
3c	228	22	52	12	82
3d	206	24	56	14	86
3e	210	24	58	13	87
3f	215	20	50	14	86
3g	207	24	52	15	80
4a	233	2	52	1	80
4b	259	2.5	59	1.5	89
4c	256	3	52	2	80
4d	251	2	56	2	88
4e	253	2.5	56	1	88
4f	232	2	52	2	80
4g	226	3	54	2	86

Analytical and spectral data for compounds **3a**–**g** and **4a**–**g** are given in Supplementary material to this paper. The ^{1}H-NMR spectrum of chalcone **3a** showed characteristic signals at δ 8.04, 8.18 and 8.62 ppm corresponding to H_{α}, H_{β} and the pyrazole H, respectively. In the ^{13}C-NMR spectrum of chalcone **3a**, the carbonyl carbon appeared at δ 193.0 ppm. The mass spectra of **3a** showed the molecular ion peak at *m*/*z* 449 $[M+H]^+$. The ^{1}H-NMR spectrum of pyrazoline derivative **4a** displayed three characteristic signals due to the diastereotopic protons (H_A, H_B and H_X). The H_A proton, which is *cis* to H_X, resonated upfield at δ 3.78 ppm as a doublet of a doublet (*dd*), while the H_B proton, which is *trans* to H_X, resonated downfield at δ 4.09 ppm (*dd*). The H_X proton, which is vicinal to two methylene protons (H_A and H_B), was also observed as a doublet of a doublet at δ 5.14. The cyclization of chalcones into pyrazolines was further supported by the ^{13}C-NMR spectrum of **4a**, in which the C–H_A and C–H_X carbons of the pyrazoline ring resonated at δ 44.2 and 54.3 ppm, respectively. These values are in close agreement with the previously reported values for the pyrazoline carbons C–H_A and C–H_X.[31,32]. The combination of ^{1}H-NMR and ^{13}C-NMR provides strong evidence in support of structures assigned to pyrazoline derivatives. The mass spectrum of **4a** showed the molecular ion peak at *m*/*z* 463 $[M+H]^+$. These data were considered satisfactory for the structures assigned to compounds **3a**–**g** and **4a**–**g.**

Antibacterial activity

All the compounds were screened for their antibacterial activity against *Staphylococcus aureus*, *Bacillus subtilis*, *Pseudomonas aeruginosa* and *Escherichia coli* using ampicillin as the standard drug. The activity was determined using cup plate agar diffusion method by measuring the zone of inhibition in mm. The compounds were screened at the concentration of 50 μg mL^{-1} in DMSO. From the screening studies (Table II), it is evident that the synthesized compounds **3b**, **3c**, **4a**, **4b**, **4c** and **4g** showed good antibacterial activity against all the tested organisms. It was further observed that the electron rich chalcone **3c**, with both aryl rings possessing a methoxy substituent, showed the best activity, and closely followed by **3b**, which has only one methoxy substituent. This leads to the conclusion that electron rich chalcones showed higher activity. Furthermore, changing the halogens from F to Cl or Br, did not provide any significant change in the levels of activity against bacteria. However, in the case of the pyrazoles derived from the chalcones, the activity does not depend much on the electronic nature of the compounds. This is evident from the fact that **4a**, **4b**, **4c** and **4g**, all have approximately the same level of activity.

TABLE II. Antibacterial activity (zone of inhibition, mm) of compounds **3a**–**g** and **4a**–**g**

Compd.	Gram-positive bacteria		Gram-negative bacteria	
	S. aureus	*B. subtilis*	*P. aeruginosa*	*E. coli*
3a	16	4	5	12
3b	28	13	13	35
3c	30	12	9	32
3d	23	5	4	10
3e	22	7	8	25
3f	10	8	2	23
3g	22	10	7	22
4a	29	11	10	27
4b	27	10	8	28
4c	28	11	8	26
4d	20	7	5	22
4e	18	10	6	18
4f	15	9	8	21
4g	28	11	9	28
Ampicillin	30	12	10	30

Antifungal activity

All the compounds were screened for their antifungal activity against *Aspergillus niger*, *Penicillium italicum* and *Fusarium oxysporum* using griseofulvin as the standard drug. The activity was determined using the cup plate agar diffusion method by measuring the zone of inhibition in mm. The compounds were screened at a concentration of 50 μg mL^{-1} in DMSO. From the screening studies

(Table III), it was evident that the synthesized compounds **3b**, **3c**, **3e, 4a, 4c** and **4g** showed good antifungal activity against all the tested organisms. The fluorinated chalcone **3e** showed the highest activity, followed by the electron rich substrates **3b** and **3c**. When the fluorine atom was replaced by bromo or chloro substituents, the activity against fungi decreased. Among the pyrazoles, the unsubstituted compound **4a** showed the highest activity against the fungi. The highly electron rich pyrazole **4c** showed activity which was comparable to that of **4a,** but on the lower side, indicating that substituents on the pyrazole are detrimental to the observed activity. Among the halogen derivates, the bromo substituted compound **4g** showed significantly higher activity compared to those of the fluoro- and chloro-substituted compounds, **4e** and **4f**, respectively.

TABLE III. Antifungal activity (zone of inhibition, mm) of compounds **3a–g** and **4a–g**

Compd.	Fungus		
	A. niger	*P. italicum*	*F. oxysporum*
3a	8	16	22
3b	12	21	25
3c	11	20	23
3d	9	20	26
3e	12	24	28
3f	9	21	23
3g	6	18	23
4a	15	25	28
4b	8	14	19
4c	14	21	24
4d	7	15	18
4e	6	12	14
4f	8	10	10
4g	14	21	23
Griseofulvin	12	20	25

EXPERIMENTAL

Materials

All used materials were obtained commercially, mostly from Sigma–Aldrich, and were used without further purification.

Equipment

The melting points were determined in open capillaries and are uncorrected. The purity of the compounds was checked by TLC on silica gel 60 F_{254} (Merck). The ^{1}H-NMR and ^{13}C--NMR spectra were recorded on a Bruker Avance II 400 spectrometer using TMS as an internal standard. The IR spectra were recorded in KBr on a Shimadzu FTIR 8400S spectrophotometer. The mass spectra were obtained on a Shimadzu GCMS-QP 1000 mass spectrometer. The microwave reactions were performed in a Milestone multiSYNTH microwave system.

*General procedure for the synthesis of (*E*)-8-[3-(1,3-diphenyl-1*H*-pyrazol-4-yl)-1-oxo-2--propen-1-yl]-7-hydroxy-4-methyl-2*H*-1-benzopyran-2-one analogues **3a**–**g***

Conventional method. A mixture of 8-acetyl-7-hydroxy-4-methylcoumarin **1** (1 mmol), substituted 1,3-diphenyl-1*H*-pyrazole-4-carboxaldehydes **2a**–**g** (1 mmol) in ethanol (20 mL) and a few drops of piperidine was taken into a round bottomed flask and stirred at room temperature for 20–24 h. Progress of the reaction was monitored by TLC. After completion of the reaction, the reaction mixture was diluted with cold water and acidified with dil. HCl. The precipitate that formed was filtered, dried and recrystallized from ethanol to afford pure chalcones **3a**–**g**.

Microwave irradiation. A mixture of 8-acetyl-7-hydroxy-4-methylcoumarin **1** (1 mmol), substituted 1-(1,3-diphenyl-1*H*-pyrazole-4-carboxaldehydes **2a**–**g** (1 mmol) in ethanol (2 ml) and few drops of piperidine was taken in a glass vial equipped with a cap and then subjected to microwave irradiation at 100 W for 10 to 15 min. Progress of the reaction was monitored by TLC. After completion of the reaction, the reaction mixture was diluted with cold water and acidified with dil. HCl. The precipitate that formed was filtered, dried and recrystallized from ethanol to afford the pure chalcone.

*General procedure for the synthesis of substituted 8-(4′,5′-dihydro-1,3-diphenyl[4,5′-bi-1*H*--pyrazol]-3′-yl)-7-hydroxy-4-methyl-2*H*-1-benzopyran-2-one **4a**–**g***

Conventional method. To a solution of chalcone **3a**–**g** (1 mmol) in DMF (5 mL) containing sodium acetate (1 mmol), hydrazine hydrate (1 mmol) was added and the reaction mixture was heated at 80–90 °C for 2 to 3 h. The progress of the reaction was monitored by TLC. After the completion of the reaction, ice water was added. The solid product that separated was filtered, washed with water, dried and recrystallized from MeOH:$CHCl_3$ (1:1).

Microwave method. To a solution of chalcones **3a**–**g** (1 mmol) in DMF (1 mL) containing sodium acetate (1 mmol), in a 10 mL glass vial equipped with a cap, hydrazine hydrate (1 mmol) was added and the mixture was then irradiated for 1–2 min at an irradiation power of 180 W. The progress of the reaction was monitored by TLC. After completion of the reaction, ice water was added. The solid product that separated was filtered, washed with water, dried and recrystallized from MeOH:$CHCl_3$ (1:1).

Biological assays

Antibacterial activity, The novel synthesized compounds **3a**–**g** and **4a**–**g** were screened for their antibacterial activity against different types of bacterial strains, *i.e*., Gram-negative bacterial strains of *Pseudomonas aeruginosa* (9027) and *Escherichia coli* (ATCC-8739), Gram-positive bacterial strains of *Bacillus subtilis* (ATCC-11778) and *Staphylococcus aureus* (ATCC-9144) at a concentration of 50 μg mL^{-1}.

The cultures were diluted with 5 % autoclaved saline and the final volume was adjusted to a concentration of approximately 10^5–10^6 CFU mL^{-1}. The synthesized compounds were diluted with acetone for the antibacterial biological assays. For agar disc diffusion method,[33] the liquid form of the test compound was soaked on to a disc (5 mm) and then allowed to air dry, such that the disc became completely saturated with the test compound. The saturated chemical discs were introduced onto the upper layer of the medium evenly loaded with the bacteria and incubated at 37 °C for 24 to 48 h for better inhibition of the bacteria. The zones of inhibition were measured after 24 to 48 h. All the experiments were performed in triplicate and the results are expressed as zone of inhibition in mm. The zones of inhibition of synthesized compounds **3a**–**g** and **4a**–**g** were compared with the zone of inhibition of the standard antibiotic ampicillin (50 μg mL^{-1}).

Antifungal activity. The antifungal activity of synthesized compounds **3a**–**g** and **4a**–**g** was tested against three pathogenic fungi, namely *Fusarium oxysporum* (ATCC-7601), *Aspergillus niger* (ATCC-9029) and *Penicillium italicum* (ATCC-10454), by the poison plate technique at a concentration of 50 μg mL^{-1}. Three kinds of fungi were incubated in potato dextrose agar medium (PDA) at 25±1 °C for 5 days to obtain new mycelium for the antifungal assay and then mycelia as disks of approximately 0.45 cm diameter cut from the culture medium were picked up with a sterilized inoculation needle and inoculated into the centre of a PDA plate. The test compounds were dissolved in acetone (10 mL) then added to the PDA medium (90 mL). The final concentration of compounds in the medium was adjusted to 50 μg mL^{-1}. The inoculated plates were incubated at 25±1 °C for 5 days. Acetone was diluted with sterilized distilled water and used as the control, while griseofulvin (50 μg mL^{-1}) was used as the standard control for each treatment. Three replicates of the experiments were performed. The radial growth of the fungal colonies was measured on the sixth day.

CONCLUSIONS

Two new series of compounds **3a**–**g** and **4a**–**g** were synthesized under conventional and microwave irradiation conditions. In microwave irradiation method, the reactions were completed in shorter times with better yields compared to the conventional method. All the new compounds were screened for their antimicrobial activities. It was observed that compounds **3b**, **3c**, **4a**, **4b**, **4c** and **4g** exhibited broad spectrum of antibacterial activity, and compounds **3b**, **3c**, **3e, 4a, 4c** and **4g** showed a broad spectrum of antifungal activity against all the tested strains compared to the standard drugs at their respective concentrations.

SUPPLEMENTARY MATERIAL

Analytical and spectral data for compounds **3a**–**g** and **4a**–**g** are available electronically from http://www.shd.org.rs/JSCS/, or from the corresponding author on request.

Acknowledgments. We are thankful to The Head of Department of Chemistry, Osmania University, Hyderabad, India, for providing the laboratory facilities and The Director of CFRD, Osmania University, for providing the spectral analysis facilities. One of the authors, B. Vijaya Lakshmi, is thankful to CSIR, New Delhi, India, for financial support in the form of CSIR-SRF.

ИЗВОД

СИНТЕЗА НОВИХ КУМАРИН–ПИРАЗОЛИН ХИБРИДА ОЗРАЧИВАЊЕМ МИКРОТАЛАСИМА И ИСПИТИВАЊЕ ЊИХОВЕ АНТИМИКРОБНЕ АКТИВНОСТИ

DONGAMANTI ASHOK, BOMMIDI VIJAYA LAKSHMI, SIDDA RAVI и ARRAM GANESH

Department of Chemistry, Osmania University, Hyderabad, India

Синтетисана је серија деривата пиразолина **4a**–**g** Мајкловом реакцијом халкона **3a**–**g** са хидразин-хидратом у присуству натријум-ацетата загревањем класичним поступком или микроталасима. Структуре нових халкона **3a**–**g** и пиразолина **4a**–**g** утврђене су на основу IC, ^{1}H- и ^{13}C-NMR спектроскопије и масене спектрометрије. Свим синтетисаним једињењима испитана је антимикробна активност. Неки од деривата показују добре активности у поређењу са стандардним лековима према свим сојевима патогених бактерија и гљива.

(Примљено 21. јуна, ревидирано 15. октобра, прихваћено 20. октобра 2014)

REFERENCES

1. S. Rajasekaran, G. K. Rao, S. P. N. Pai, A. Ranjan, *Int. J. ChemTech Res.* **3** (2011) 555
2. R. Sharma, V. Arya, *J. Chem. Pharm. Res.* **3** (2011) 204
3. R. D. H. Murrey, D. Medez, S. A. Brown, *The natural coumarins: occurrence, chemistry and biochemistry*, Wiley Interscience, New York, 1982, p. 1
4. A. I. Vogel, B. S. Furniss, A. J. Hannaford P. W. G. Smith, A. R. Tatchell, *Vogel's Textbook of Practical Organic Chemistry*, **5**th ed., Longman, Harlow, 1989, p. 1191
5. V. V. Mulwad, J. M. Shirodkhar, *Indian J. Heterocycl. Chem.* **11** (2002) 192
6. K. Manohar, G. Manjunath, K. Raviraj, *Indian J. Heterocycl. Chem.* **13** (2004) 201
7. V. Rajeswar Rao, K. Srimanth, P. Vijaya Kumar, *Indian J. Heterocycl. Chem.* **14** (2004) 141
8. Z. M. Nofal, M. I. El-Zahar, S. S Abd El-Karim, *J. Antimicrob. Agents Chemother.* **5** (2005) 483
9. H. N. Pati, H. L. Holt, R. L. Blanc Jr., J. Dickson, M. Stewart, T. Brown, *Med. Chem. Res.* **14** (2005) 19
10. C. L. Miranda, J. F. Stevens, V. Ivanov, M. McCall, B. Frei, M. L. Deinzer, *J. Agric. Food Chem.* **48** (2000) 3876
11. M. A. Ali, M. Shaharyar, A. A. Siddiqui, *Eur. J. Med. Chem.* **42** (2007) 268
12. H. H. Ko, L. T. Tsao, K. L. Yu, C. T. Liu, J. P. Wang, C. N. Lin, *Bioorg. Med. Chem.* **11** (2003) 105
13. A. M. Deshpande, N. P. Argade, A. A. Natu, J. Eckman, *Bioorg. Med. Chem.* **7** (1999) 1237
14. S. Khatib, O. Nerya, R. Musa, M. Shmnel, S. Tamir, J. Vaya, *Bioorg. Med. Chem.* **13** (2005) 433
15. F. Severi, S. Benvenuti, L. Costantino, G. Vampa, M. Melegari, L. Antolini, *Eur. J. Med. Chem.* **33** (1998) 859
16. Y. Kohno, S. Kitamura, S. Sanoh, K. Sugihara, N. Fujimoto, S. Ohta, *J. Pharmacol. Exp. Ther.* **33** (2005) 1115
17. B. Srinivasan, T. E. Johnson, R. Lad, C. Xing, *J. Med. Chem.* **52** (2009) 7228
18. D. L. Opdyke, *Food Cosmet. Toxicol.* **11** (1973) 1011
19. B. F. Abdel-Wahab, H. A. Abdel-Aziz, E. M. Ahmed, *Eur. J. Med. Chem.* **44** (2009) 2632
20. M. Abid, A. Azam, *Bioorg. Med. Chem. Lett.* **16** (2006) 2812
21. M. Abid, A. R. Bhat, F. Athar, A. Azam, *Eur. J. Med. Chem.* **44** (2009) 417
22. Y. R. Prasad, A. L. Rao, L. Prasoona, K. Murali, P. R. Kumar, *Bioorg. Med. Chem. Lett.* **15** (2005) 5030
23. Z. Ozdemir, H. B. Kandilci, B. Gumusel, U. Calıs, A. A. Bilgin, *Eur. J. Med. Chem.* **42** (2007) 373
24. I. G. Rathish, K. Javed, S. Ahmad, S. Bano, M. S. Alam, K. K. Pillai, S. Singh, V. Bagchi, *Bioorg. Med. Chem. Lett.* **19** (2009) 255
25. N. Gokhan-Kelekçi, S. Yabanoglu, E. Kupeli, U. Salgın, O. Ozgen, G. Ucar, E. Yeşilada, E. Kendi, A. Yeşilada, A. A. Bilgin, *Bioorg. Med. Chem.* **15** (2007) 5775
26. E. C. Taylor, H. H. Patel, *Tetrahedron* **48** (1992) 8089
27. C. O. Kappe, *Angew. Chem. Int. Ed. Engl.* **43** (2004) 6256
28. D. Ashok, A. Ganesh, *Indian J. Heterocycl. Chem.* **23** (2013) 7
29. D. Ashok, K. Sudershan, M. Khalilullah, *Green Chem. Lett. Rev.* **5** (2012) 121
30. D. Ashok, D. Shravani, *Tetrahedron Lett.* **49** (2008) 7227
31. R. Aggarwal, V. Kumar, S. P. Singh, *Indian J. Chem., B* **46** (2007) 1332
32. A. Ozdemir, G. Turan-Zitouni, Z. A. Kaplancikh, G. Revial, K. Guven, *Eur. J. Med. Chem.* **42** (2008) 403
33. H. J. Benson, *Microbiological Applications*, W. C. Brown Publications, Boston, MA, 1990.

Journal of the Serbian Chemical Society
JSCS@tmf.bg.ac.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) S58–S62 (2015) *Supplementary material*

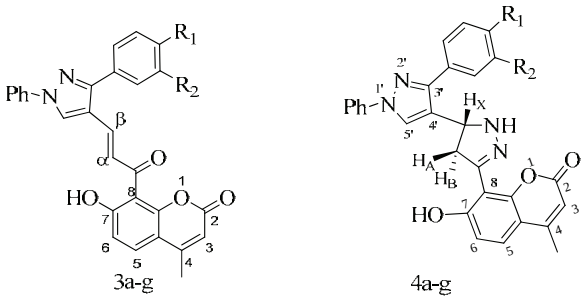
SUPPLEMENTARY MATERIAL TO

Microwave-assisted synthesis of some new coumarin–pyrazoline hybrids and their antimicrobial activity

DONGAMANTI ASHOK*, BOMMIDI VIJAYA LAKSHMI, SIDDA RAVI and ARRAM GANESH

Department of Chemistry, Osmania University, Hyderabad, India

J. Serb. Chem. Soc. 80 (3) (2015) 305–313



General structures of the synthesized compounds with atomic numbering.

ANALYTICAL AND SPECTRAL DATA FOR COMPOUNDS **3a**–**g** AND **4a**–**g**

*(E)-8-[3-(1,3-Diphenyl-1*H*-pyrazol-4-yl)-1-oxo-2-propen-1-yl]-7-hydroxy-4--methyl-2*H*-1-benzopyran-2-one (**3a**)*. Anal. Calcd. for $C_{28}H_{20}N_2O_4$: C, 74.94; H, 4.46; N, 6.21 %. Found: C, 74.99; H, 4.50; N, 6.25 %; IR (KBr, cm^{-1}): 3440 (OH), 1636 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm): 2.45 (3H, *d*, *J* = 1.004 Hz, CH_3), 6.20 (1H, *d*, *J* = 1.0 Hz, H_3), 6.96 (1H, *d*, *J* = 9.0 Hz, H_6), 7.33–7.37 (1H, *m*, Ar-H), 7.46–7.54 (5H, *m*, Ar-H), 7.68 (1H, *d*, *J* = 9.0 Hz, H_5), 7.73–7.75 (2H, *m*, Ar-H), 7.85–7.87 (2H, *m*, Ar-H), 8.04 (1H, *d*, *J* = 15.5 Hz, H_α), 8.18 (1H, *d*, *J* = 15.5 Hz, H_β), 8.62 (1H, *s*, pyrazole H), 13.94 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.3, 109.7, 110.9, 115.3, 118.6, 119.4, 122.8, 125.6, 126.3, 127.3, 128.7, 128.8, 129.5, 130.8, 134.2, 137.2, 137.5, 153.4, 154.8, 156.3, 153.5, 159.6, 163.6, 193.0; MS (*m*/*z*): 449 ($[M+H]^+$, 100 %).

*(E)-7-Hydroxy-8-{3-[3-(4-methoxyphenyl)-1-phenyl-1H-pyrazol-4-yl]-1-oxo--2-propen-1-yl}-4-methyl-2*H*-1-benzopyran-2-one (**3b**)*. Anal. Calcd. for $C_{29}H_{22}N_2O_5$: C, 72.79; H, 4.63; N, 5.85 %. Found: C, 72.81; H, 4.65; N, 5.87 %;

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

IR (KBr, cm^{-1}): 3444 (OH), 1635 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm): 2.45 (3H, *d*, CH_3, *J* = 1.0 Hz), 3.89 (3H, *s*, OCH_3), 6.20 (1H, *d*, *J* = 1.0 Hz, H_3), 6.96 (1H, *d*, *J* = 8.7 Hz, H_6), 7.05 (2H, *d*, *J* = 9.0 Hz, Ar-H), 7.33–7.37 (1H, *m*, Ar-H), 7.48–7.52 (2H, *m*, Ar-H), 7.67 (1H, *d*, *J* = 8.7 Hz, H_5), 7.69 (2H, *d*, *J* = = 9.0 Hz, Ar-H), 7.84–7.86 (2H, *m*, Ar-H), 8.03 (1H, *d*, *J* = 15.3 Hz, H_α), 8.16 (1H, *d*, *J* = 15.3 Hz, H_β), 8.60 (1H, *s*, pyrazole H), 13.97 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.3, 55.4, 110.9, 112.0, 114.3, 115.3, 118.4, 119.4, 120.3, 123.8, 124.1, 124.6, 125.3, 127.2, 129.5, 130.0, 130.8, 136.9, 139.4, 156.3, 156.6, 159.7, 160.1, 167.4, 196.1; MS (*m*/*z*): 479 ($[M+H]^+$, 100 %).

*(*E*)-8-{3-[3-(3,4-Dimethoxyphenyl)-1-phenyl-1H-pyrazol-4-yl]-1-oxo-2-propen-1-yl}-7-hydroxy-4-methyl-2*H*-1-benzopyran-2-one (**3c**)*. Anal. Calcd. for $C_{30}H_{24}N_2O_6$: C, 70.86; H, 4.76; N, 5.51 %. Found: C, 70.90; H, 4.78; N, 5.49 %; IR (KBr, cm^{-1}): 3447 (OH), 1633 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm): 2.45 (3H, *d*, *J* = 1.0 Hz, CH_3), 3.87 (3H, *s*, OCH_3), 3.89 (3H, *s*, OCH_3), 6.20 (1H, *d*, *J* = 1.0 Hz, H_3), 6.97 (1H, *d*, *J* = 9.0 Hz, H_6), 7.33–7.37 (1H, *m*, Ar-H), 7.46–7.54 (3H, *m*, Ar-H), 7.68 (1H, *d*, *J* = 9.0 Hz, H_5), 7.76–7.78 (2H, *m*, Ar-H), 7.86–7.88 (2H, *m*, Ar-H), 8.05 (1H, *d*, *J* = 15.5 Hz, H_α), 8.18 (1H, *d*, *J* = 15.5 Hz, H_β), 8.64 (1H, *s*, pyrazole H), 13.97 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.3, 55.2, 55.4, 110.8, 114.8, 118.6, 119.6, 121.8, 124.3, 125.4, 126.8, 127.1, 127.7, 128.1, 128.5, 129.5, 131.8, 132.5, 134.8, 137.2, 137.5, 153.4, 154.8, 156.3, 153.5, 159.6, 163.7, 193.0; MS (*m*/*z*): 509 ($[M+H]^+$, 100 %).

*(*E*)-7-Hydroxy-4-methyl-8-[3-(1-phenyl-3-p-tolyl-1H-pyrazol-4-yl)-1-oxo--2-propen-1-yl]-2*H*-1-benzopyran-2-one (**3d**)*. Anal. Calcd. for $C_{29}H_{22}N_2O_4$: C, 75.31; H, 4.79; N, 6.06 %. Found: C, 75.34; H, 4.82; N, 6.04 %; IR (KBr, cm^{-1}): 3436 (OH), 1637 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm): 2.43 (3H, *s*, CH_3), 2.45 (3H, *d*, *J* = 1.0 Hz, CH_3), 6.20 (1H, *d*, *J* = 1.0 Hz, H_3), 6.95 (1H, *d*, *J* = 9.0 Hz, H_6), 7.31–7.36 (3H, *m*, Ar-H), 7.48–7.52 (2H, *m*, Ar-H), 7.63 (2H, *d*, *J* = 8.0 Hz, Ar-H), 7.68 (1H, *d*, *J* = 9.0 Hz, H_5), 7.85 (2H, *d*, *J* = 8.0 Hz, Ar-H), 8.04 (1H, *d*, *J* = 15.5 Hz, H_α), 8.16 (1H, *d*, *J* = 15.5 Hz, H_β), 8.60 (1H, *s*, pyrazole H), 13.95 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.3, 21.37,109.6, 110.9, 112.0, 115.3, 118.5, 119.4, 125.4, 126.6, 127.2, 129.2, 129.5, 129.5, 129.6, 130.8, 138.6, 139.4, 153.4, 154.3, 154.8, 159.6, 167.4, 192.9; MS (*m*/*z*): 463 ($[M+H]^+$, 100 %).

*(*E*)-8-{3-[3-(4-Fluorophenyl)-1-phenyl-1H-pyrazol-4-yl]-1-oxo-2-propen-1--yl}-7-hydroxy-4-methyl-2*H*-1-benzopyran-2-one (**3e**)*. Anal. Calcd. for $C_{28}H_{19}FN_2O_4$: C, 72.10; H, 4.11; N, 6.01 %. Found: C, 72.12; H, 4.14; N, 6.06 %; IR (KBr, cm^{-1}): 3448 (OH), 1636 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / / ppm): 2.45 (3H, *d*, *J* = 1.0 Hz, CH_3), 6.20 (1H, *d*, *J* = 1.0 Hz, H_3), 6.95 (1H, *d*, *J* = 9.0 Hz, H_6), 7.05 (2H, *d*, *J* = 8.5 Hz, Ar-H), 7.35–7.36 (1H, *m*, Ar-H), 7.40––7.53 (2H, *m*, Ar-H), 7.69 (2H, *d*, *J* = 8.5 Hz, Ar-H), 7.71 (1H, *d*, *J* = 9.0 Hz, H_5), 7.84–7.86 (2H, *m*, Ar-H), 8.04 (1H, *d*, *J* = 15.3 Hz, H_α), 8.15 (1H, *d*, *J* =

= 15.3 Hz, H_β), 8.60 (1H, *s*, pyrazole H), 13.95 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.3, 110.7, 111.5, 114.4, 115.3, 118.4, 119.4, 120.3, 123.8, 124.1, 124.6, 125.2, 127.2, 129.5, 130.0, 130.8, 131.2, 135.6, 136.9, 139.4, 158.3, 159.6, 160.7, 164.1, 167.4, 196.1; MS (*m*/*z*): 467 ($[M+H]^+$, 100%).

*(*E*)-8-{3-[3-(4-Chlorophenyl)-1-phenyl-1H-pyrazol-4-yl]-1-oxo-2-propen-1--yl}-7-hydroxy-4-methyl-2*H*-1-benzopyran-2-one (**3f**)*. Anal. Calcd. for $C_{28}H_{19}ClN_2O_4$: C, 69.64; H, 3.97; N, 5.80 %. Found: C, 69.68; H, 4.01; N, 5.84 %; IR (KBr, cm^{-1}): 3445 (OH), 1636 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / / ppm): 2.46 (3H, *d*, *J* = 1.0 Hz, CH_3), 6.20 (1H, *d*, *J* = 1.0 Hz, H_3), 6.97 (1H, *d*, *J* = 8.7 Hz, H_6), 7.34–7.37 (3H, *m*, Ar-H), 7.49–7.51 (2H, *m*, Ar-H), 7.63 (2H, *d*, *J* = 8.5 Hz, Ar-H), 7.68 (1H, *d*, *J* = 8.7 Hz, H_5), 7.85 (2H, *d*, *J* = 8.0 Hz, Ar-H), 8.04 (1H, *d*, *J* = 15.5 Hz, H_α), 8.16 (1H, *d*, *J* = 15.5 Hz, H_β), 8.60 (1H, *s*, pyrazole H), 13.95 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.3, 110.9, 112.1, 115.4, 118.7, 119.2, 125.1, 125.7, 126.4, 127.4, 129.2, 129.5, 129.5, 129.6, 130.8, 138.6, 139.4, 142.8, 153.4, 154.3, 154.8, 159.6, 167.4, 192.9; MS (*m*/*z*): 483 ($[M+H]^+$, 100 %).

*(*E*)-8-{3-[3-(4-Bromophenyl)-1-phenyl-1H-pyrazol-4-yl]-1-oxo-2-propen-1--yl}-7-hydroxy-4-methyl-2*H*-1-benzopyran-2-one (**3g**)*. Anal. Calcd. for $C_{28}H_{19}BrN_2O_4$: C, 63.77; H, 3.63; N, 5.31 %. Found: C, 63.80; H, 3.67; N, 5.35 %; IR (KBr, cm^{-1}): 3441 (OH), 1654 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / / ppm): 2.45 (3H, *d*, *J* = 1.0 Hz, CH_3), 6.20 (1H, *d*, *J* = 1.0 Hz, H_3), 6.96 (1H, *d*, *J* = 9.0 Hz, H_6), 7.36–7.44 (5H, *m*, Ar-H), 7.68 (1H, *d*, *J* = 9.0 Hz, H_5), 7.73––7.75 (2H, *m*, Ar-H), 7.85–7.87 (2H, *m*, Ar-H), 8.04 (1H, *d*, *J* = 15.5 Hz, H_α), 8.18 (1H, *d*, *J* = 15.5 Hz, H_β), 8.62 (1H, *s*, pyrazole H), 13.94 (1H, *s*, OH); ^{13}C--NMR (100 MHz, $CDCl_3$, δ / ppm): 19.2, 109.62, 114.3, 115.2, 120.8, 123.2, 124.6, 126.8, 127.3, 127.7, 128.9, 129.5, 130.4, 134.3, 137.4, 137.3, 153.2, 154.6, 156.1, 153.5, 157.6, 162.8, 192.4; MS (*m*/*z*): 527 ($[M+H]^+$, 100 %).

*8-(4′,5′-Dihydro-1,3-diphenyl[4,5′-bi-1H-pyrazol]-3′-yl)-7-hydroxy-4-methyl--2*H*-1-benzopyran-2-one (**4a**)*. Anal. Calcd. for $C_{28}H_{22}N_4O_4$: C, 72.71; H, 4.79; N, 12.11 %. Found: C, 72.75; H, 4.82; N, 12.14 %; IR (KBr, cm^{-1}): 3333 (OH), 1597 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm): 2.41 (3H, *d*, *J* = 1.0 Hz, CH_3), 3.78 (1H, *dd*, *J* = 17.8 Hz, *J* = 9.0 Hz, H_A), 4.09 (1H, *dd*, *J* = 17.8 Hz, *J* = = 10.5 Hz, H_B), 5.14 (1H, *dd*, *J* = 9.0 Hz, *J* = 10.5 Hz, H_X), 5.97 (1H, *brs*, N–H, D_2O exchangeable), 6.12 (1H, *d*, *J* = 1.0 Hz, H_3), 6.97 (1H, *d*, *J* = 9.0 Hz, H_6), 7.29–7.31 (1H, *m*, Ar-H), 7.39–7.50 (6H, *m*, Ar-H), 7.69–7.74 (4H, *m*, Ar-H, H_5), 8.05 (1H, *s*, pyrazole H), 12.54 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.1, 44.2, 54.3, 105.4, 110.9, 112.3, 113.9, 119.2, 122.0, 125.9, 126.6, 128.2, 128.3, 128.8, 129.4, 132.8, 139.8, 151.5, 153.2, 155.5, 160.3, 161.8; MS (*m*/*z*): 463 ($[M+H]^+$, 100 %).

*8-[4′,5′-Dihydro-3-(4-methoxyphenyl)-1-phenyl[4,5′-bi-1H-pyrazol]-3′-yl]-7--hydroxy-4-methyl-2*H*-1-benzopyran-2-one (**4b**)*. Anal. Calcd. for $C_{29}H_{24}N_4O_4$:

C, 70.72; H, 4.91; N, 11.38 %. Found: C, 70.75; H, 4.94; N, 11.42 %; IR (KBr, cm^{-1}): 3335 (OH), 1599 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm): 2.42 (3H, *d*, *J* = 1.0 Hz, CH_3), 3.78 (1H, *dd*, *J* = 17.8 Hz, *J* = 9.0 Hz, H_A), 3.84 (3H, *s*, OCH_3), 4.10 (1H, *dd*, *J* = 17.8 Hz, *J* = 10.5 Hz, H_B), 5.13 (1H, *dd*, *J* = 9.0 Hz, *J* = 10.5 Hz, H_X), 5.83 (1H, *brs*, N–H, D_2O exchangeable), 6.13 (1H, *d*, *J* = 1.0 Hz, H_3), 6.97 (1H, *d*, *J* = 9.0 Hz, H_6), 6.99 (2H, *d*, *J* = 8.7 Hz, Ar-H), 7.42–7.54 (4H, *m*, Ar-H, H_5), 7.62 (2H, *d*, *J* = 8.7 Hz, Ar-H), 7.69–7.73 (2H, *m*, Ar-H), 8.04 (1H, *s*, pyrazole H), 13.92 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.3, 43.2, 53.7, 55.6, 109.7, 110.9, 115.3, 118.6, 119.4, 122.8, 125.6, 126.3, 127.3, 128.7, 128.8, 129.5, 130.8, 134.2, 137.2, 137.5, 153.4, 154.8, 156.3, 153.5, 159.6, 162.8; MS (*m*/*z*): 493 ($[M+H]^+$, 100 %).

*8-[3-(3,4-Dimethoxyphenyl)-4′,5′-dihydro-1-phenyl[4,5′-bi-1*H*-pyrazol]-3′-yl]--7-hydroxy-4-methyl-2*H*-1-benzopyran-2-one (**4c**)*. Anal. Calcd. for $C_{30}H_{26}N_4O_5$: C, 68.95; H, 5.02; N, 10.72 %. Found: C, 70.01; H, 5.06; N, 10.76 %; IR (KBr, cm^{-1}): 3340 (OH), 1596 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm): 2.41 (3H, *d*, *J* = 1.0 Hz, CH_3), 3.77 (1H, *dd*, *J* = 17.8 Hz, *J* = 9.0 Hz, H_A), 3.84 (3H, *s*, OCH_3), 3.87 (3H, *s*, OCH_3), 4.08 (1H, *dd*, *J* = 17.8 Hz, *J* = 10.5 Hz, H_B), 5.14 (1H, *dd*, *J* = 9.0 Hz, *J* = 10.5 Hz, H_X), 5.95 (1H, *brs*, N–H, D_2O exchangeable), 6.12 (1H, *d*, *J* = 1.0 Hz, H_3), 6.96 (1H, *d*, *J* = 9.0 Hz, H_6), 7.27–7.30 (1H, *m*, Ar-H), 7.39–7.50 (4H, *m*, Ar-H), 7.69–7.74 (4H, *m*, Ar-H, H_5), 8.06 (1H, *s*, pyrazole H), 12.55 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.0, 44.1, 52.3, 55.1, 55.4, 109.7, 110.9, 114.2, 114.4, 115.2, 116.3, 124.3, 112.3, 127.8, 130.5, 130.6, 131.4, 131.8, 132.7, 153.3, 158.5, 159.3, 160.4. MS (*m*/*z*): 523 ($[M+H]^+$, 100 %).

*8-[4′,5′-Dihydro-1-phenyl-3-p-tolyl-[4,5′-bi-1H-pyrazol]-3′-yl]-7-hydroxy-4--methyl-2*H*-1-benzopyran-2-one (**4d**)*. Anal. Calcd. for $C_{29}H_{24}N_4O_4$: C, 73.09; H, 5.08; N, 11.76 %. Found: C, 73.11; H, 5.11; N, 11.79 %; IR (KBr, cm^{-1}): 3336 (OH), 1599 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm): 2.41 (3H, *d*, *J* = = 1.0 Hz, CH_3), 2.43 (3H, *s*, CH_3), 3.77 (1H, *dd*, *J* = 17.8 Hz, *J* = 9.0 Hz, H_A), 4.09 (1H, *dd*, *J* = 17.8 Hz, *J* = 10.5 Hz, H_B), 5.14 (1H, *dd*, *J* = 9.0 Hz, *J* = 10.5 Hz, H_X), 5.30 (1H, *brs*, N–H, D_2O exchangeable), 6.12 (1H, *d*, *J* = 1.0 Hz, H_3), 6.95 (2H, *d*, *J* = 8.5 Hz, Ar-H), 6.99 (1H, *d*, *J* = 8.7 Hz, H_6), 7.42–7.54 (4H, *m*, Ar-H, H_5), 7.62 (2H, *d*, *J* = 8.5 Hz, Ar-H), 7.69–7.73 (2H, *m*, Ar-H), 8.04 (1H, *s*, pyrazole H), 13.92 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.2, 21.3, 44.8, 54.6, 108.4, 112.4, 113.2, 115.5, 121.2, 122.4, 125.3, 125.8, 126.2, 127.3, 129.71, 130.4, 131.3, 133.2, 151.0, 153.2, 159, 162.5; MS (*m*/*z*): 477 ($[M+H]^+$, 100 %).

*8-[3-(4-Fluorophenyl)-4′,5′-dihydro-1-phenyl[4,5′-bi-1*H*-pyrazol]-3′-yl]-7--hydroxy-4-methyl-2*H*-1-benzopyran-2-one (**4e**)*. Anal. Calcd. for $C_{28}H_{21}FN_4O_3$ C, 69.99; H, 4.41; F, 3.95; N, 11.66. Found: C, 70.02, H, 4.45, N, 11.69; IR (KBr, cm^{-1}): 3334 (OH), 1597 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm):

2.41 (3H, *d*, *J* = 1.0 Hz, CH_3), 3.77 (1H, *dd*, *J* = 17.8 Hz, *J* = 9.0 Hz, H_A), 4.08 (1H, *dd*, *J* = 17.8 Hz, *J* = 10.5 Hz, H_B), 5.13 (1H, *dd*, *J* = 9.0 Hz, *J* = 10.5 Hz, H_X), 4.97 (1H, *brs*, N–H, D_2O exchangeable), 6.11 (1H, *d*, *J* = 1.0 Hz, H_3), 6.95 (1H, *d*, *J* = 9.0 Hz, H_6), 6.97 (2H, *d*, *J* = 8.3 Hz, Ar-H), 7.40–7.48 (4H, *m*, Ar-H, H_5), 7.64 (2H, *d*, *J*=8.3 Hz, Ar-H), 7.69–7.74 (2H, *m*, Ar-H), 8.03 (1H, *s*, pyrazole H), 12.53 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.4, 44.2, 54.8, 109.7, 111.9, 115.3, 118.6, 119.4, 122.8, 125.6, 126.3, 127.3, 128.7, 128.8, 129.5, 130.8, 134.2, 137.2, 137.5, 153.4, 154.8, 156.3, 153.5, 159.6, 163.6; MS (*m*/*z*): 481 ($[M+H]^+$, 100 %).

*8-[3-(4-Chlorophenyl)-4′,5′-dihydro-1-phenyl[4,5′-bi-1*H*-pyrazol]-3′-yl]-7--hydroxy-4-methyl-2*H*-1-benzopyran-2-one (**4f**)*. Anal. Calcd. for $C_{28}H_{21}ClN_4O_3$: C, 67.67; H, 4.26; N, 11.27 %. Found: C, 67.70; H, 4.30; N, 11.30 %; IR (KBr, cm^{-1}): 3332 (OH), 1595 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm): 2.42 (3H, *d*, *J* = 1.0 Hz, CH_3), 3.75 (1H, *dd*, *J* = 17.8 Hz, *J* = 9.0 Hz, H_A), 4.06 (1H, *dd*, *J* = = 17.8 Hz, *J* = 10.5 Hz, H_B), 5.14 (1H, *dd*, *J* = 9.0 Hz, *J* = 10.5 Hz, H_X), 5.63 (1H, *brs*, N–H, D_2O exchangeable), 6.12 (1H, *d*, *J* = 1.0 Hz, H_3), 6.98 (1H, *d*, *J* = = 9.0 Hz, H_6), 7.18 (2H, *d*, *J* = 8.4 Hz, Ar-H), 7.37–7.42 (4H, *m*, Ar-H, H_5), 7.48 (2H, *d*, *J* = 8.4 Hz, Ar-H), 7.71–7.74 (2H, *m*, Ar-H), 8.02 (1H, *s*, pyrazole H), 12.58 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.2, 43.2, 54.8, 111.6, 116.3, 118.1, 119.8, 121.8, 125.8, 126.3, 127.3, 128.7, 128.9, 129.5, 130.8, 134.2, 137.2, 137.5, 153.4, 154.7, 156.3, 153.5, 157.6, 162.6; MS (*m*/*z*): 497 ($[M+H]^+$, 100 %).

*8-[3-(4-Bromophenyl)-4′,5′-dihydro-1-phenyl[4,5′-bi-1*H*-pyrazol]-3′-yl]-7--hydroxy-4-methyl-2*H*-1-benzopyran-2-one (**4g**)*. Anal. Calcd. for $C_{28}H_{21}BrN_4O_3$: C, 62.12; H, 3.91; N, 10.35 %. Found: C, 62.15; H, 3.94; N, 10.38 %; IR (KBr, cm^{-1}): 3330 (OH), 1596 (C=N); ^{1}H-NMR (400 MHz, $CDCl_3$, δ / ppm): 2.43 (3H, *d*, *J* = 1.0 Hz, CH_3), 3.80 (1H, *dd*, *J* = 17.8 Hz, *J* = 9.0 Hz, H_A), 4.08 (1H, *dd*, *J* = = 17.8 Hz, *J* = 10.5 Hz, H_B), 5.16 (1H, *dd*, *J* = 9.0 Hz, *J* = 10.5 Hz, H_X), 5.42 (1H, *brs*, N–H, D_2O exchangeable), 6.14 (1H, *d*, *J* = 1.0 Hz, H_3), 6.92 (1H, *d*, *J* = = 9.0 Hz, H_6), 7.18 (2H, *d*, *J*=8.3 Hz, Ar-H), 7.37–7.42 (4H, *m*, Ar-H, H_5), 7.51 (2H, *d*, *J* = 8.3 Hz, Ar-H), 7.68–7.72 (2H, *m*, Ar-H), 8.02 (1H, *s*, pyrazole H), 12.58 (1H, *s*, OH); ^{13}C-NMR (100 MHz, $CDCl_3$, δ / ppm): 19.3, 43.5, 54.3, 11.7, 110.0, 114.3, 119.0, 119.4, 122.4, 123.6, 125.3, 127.1, 128.4, 128.8, 129.1, 130.4, 134.1, 137.2, 137.5, 153.4, 154.6, 154.9, 155.5, 157.6, 164.6; MS (*m*/*z*): 541 ($[M+H]^+$, 100 %).

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 315–327 (2015)
JSCS–4718
UDC **Salmonella enterica*:547.587.51+
547.21.024:579.86:616.34:615.282
Original scientific paper

A small library of 4-(alkylamino)-3-nitrocoumarin derivatives with potent antimicrobial activity against gastrointestinal pathogens

NIKO S. RADULOVIĆ[1*#], ZORICA STOJANOVIĆ-RADIĆ[2], PREDRAG STOJANOVIĆ[3,4], NIKOLA STOJANOVIĆ[3], VIDOSLAV DEKIĆ[5#] and BILJANA DEKIĆ[5#]

[1]Department of Chemistry, Faculty of Science and Mathematics, University of Niš, Višegradska 33, 18000 Niš, Serbia, [2]Department of Biology and Ecology, Faculty of Science and Mathematics, University of Niš, Višegradska 33, 18000 Niš, Serbia, [3]Faculty of Medicine, University of Niš, Zorana Đinđića 81, 18000 Niš, Serbia, [4]Institute For Public Health Niš, Center of Microbiology, 18000 Niš, Serbia and [5]Department of Chemistry, Faculty of Science and Mathematics, University of Priština, Lole Ribara 29, 38220 Kosovska Mitrovica, Serbia

(Received 19 June, revised 22 August, accepted 29 August 2014)

Abstract: Due to the confirmed antimicrobial activity of both natural and synthetic coumarin derivatives, the present study was envisaged to provide further insight into the antimicrobial potential of coumarins through the screening of a designed library of nine 4-(alkylamino)-3-nitrocoumarins against a panel of 24 laboratory strains and resistant (isolates) bacterial and fungal pathogens. All compounds showed some degree of strain-selective activity that in some cases was very pronounced, reaching the value of 0.04 nmol mL^{-1} (*i.e.*, 12 ng mL^{-1}) for the minimal inhibitory concentration against *Candida albicans*. The observed activities were higher against Gram-negative strains, among which the most susceptible strain, among both ATCC strains and clinical isolates, was *Salmonella enterica* subsp. *enterica* serovar *Enteritidis*. These results indicate to a high potential of these coumarins as antimicrobials for the treatment of gastrointestinal and other infections caused by highly resistant microbial strains. Finally, a multivariate statistical analysis of the herein obtained and previous results on the antimicrobial activity of related selected coumarins was performed to allow an easier structure–activity discussion.

Keywords: 4-(alkylamino)coumarins; antimicrobial activity; *Candida albicans*; multivariate statistical analysis; *Salmonella enterica*.

* Corresponding author. E-mail: nikoradulovic@yahoo.com
# Serbian Chemical Society member.
doi: 10.2298/JSC140619085R

INTRODUCTION

The gastrointestinal (GI) tract represents an ideal place for the development of various microorganisms that can affect the (human) organism in both beneficial and adverse aspects, such as protection against infection, physiological functions, activation of the immune system, carcinogenesis, aging and alteration of the effect a therapeutic drug should have.[1] GI bacterial infections, which can range from mild (with few symptoms) to severe life-threatening disorders (*e.g.*, hemorrhagic colitis, intestine perforation and kidney failure) are one of the most frequent readons for patients to turn to physicians (one third of the people living in industrialized countries per year).[2] GI infections are mostly food-borne and result from inadequate cooking and inadequate storage of food, while some commensal bacteria can cause infections in immunocompromised patients (AIDS, cancer, *etc*.).[3] Hence, the search for new antimicrobial agents, renewed by the rapid and alarming development of multiresistant bacterial and fungal pathogens, has always been one of the most significant tasks of many medicinal chemists.

Coumarins are a widespread group of natural compounds present in the seeds, roots and leaves of many plant species. Their function in the plant organism is far from clear, although some authors believe that they have growth regulatory functions, or that they act as fungistatic and bacteriostatic agents.[4] This is substantiated by the extensive use of coumarin compound-containing herbal remedies in the traditional medicine of many nations due to their broad spectrum of pharmacological activities, including antibacterial and antifungal effects.[5–7]

Significant antimicrobial activity of a series of 4-aryl-substituted 3-nitrocoumarins was recently demonstrated by Debeljak and co-workers.[8] In connection with this, several arylamino- and (heteroarylamino)-3-nitrocoumarin derivatives were synthesized and their antimicrobial activity determined.[9–14] A number of these compounds showed strong activity in reducing the microbial growth comparable or even better than the activity of standard antibiotics.[9,13,14] Among them 4-(1-naphthylamino)-3-nitrocoumarin possessed the strongest antimicrobial potential as demonstrated by its action against *Salmonella enterica* in a low dose of 9 μg per disc.[15]

Additional fine-tuning of the antimicrobial activity of these coumarin derivatives could be accomplished by the introduction of an alkyl chain instead of the aromatic substituent on the nitrogen atom in position 4. In this way, the nucleophilicity and basicity would be increased that could lead to a stronger binding with potential biological target molecules. The previous statements and the fact that hitherto only the disc diffusion method was used to evaluate the activity of the synthesized coumarin derivatives, it was decided to synthesize a small library of coumarin derivatives possessing a 4-(alkylamino)-substituent on the 3-nitrocoumarin parent structure and screen the compounds for their *in vitro* antimicrobial activity against twenty four bacterial and fungal strains (both ATCC and

clinical isolates, including three important GI pathogens *Yersinia enterocolitica*, *Salmonella enterica* subsp. *enterica* serovar *Enteritidis* and *Shigella flexneri*) using a microdilution method. Finally, another task of this work was to compare previously published results concerning the antimicrobial activity of related coumarin derivatives and the herein obtained results using multivariate statistical analysis with the aim of finding possible structure–activity relationships.

EXPERIMENTAL

General

The melting points were determined on a Kofler hot-plate apparatus and are uncorrected. The HRMS(EI) spectra were recorded on a JEOL Mstation JMS 700 instrument (JEOL, Germany). The IR measurements (ATR – attenuated total reflectance) were performed using a Thermo Nicolet model 6700 FTIR instrument. The NMR spectra were recorded on a Varian Gemini 200 spectrometer (^{1}H-NMR at 200 MHz and ^{13}C-NMR at 50 MHz), using DMSO-d_6 as the solvent. The chemical shifts are expressed in δ (ppm) using TMS (Me_4Si) as the internal standard. Microanalyses of carbon, hydrogen and nitrogen were performed with a Carlo Erba 1106 microanalyser; the results agreed favorably with the calculated values. For TLC, silica gel plates (Kieselgel 60 F_{254}, Merck, Germany) were used. Visualization was affected by spraying the plates with 1:1 (*V*/*V*) aqueous solution of sulfuric acid and then heating. All the reagents and solvents were obtained from commercial sources (Aldrich, USA; Merck, Germany; Fluka, Germany) and used as received, except for the solvents that were purified by distillation.

Mass spectrometry

EI-MS and HRMS analyses were performed using a JEOL MStation JMS-700 mass spectrometer with an ionization energy of 70 eV, an ionization trap current of 300 μA and a source temperature of 230 °C. Perfluorokerosene (Sigma–Aldrich, USA) was used as the internal mass reference in the HRMS studies. The conversion of the mass reference list to a calibration was performed by the data system of the mass spectrometer. The corresponding range of mass measurements was set to include two standard peaks that encompassed the sample peak of interest. The mass resolution and scan speed used were 30,000 (10 % valley) and 60 s decade^{-1}, respectively. The accurate mass was calculated as the average of the values measured in 5–10 scans, determined from the mass centroids of $M^{+\bullet}$ and the other peaks. The error for each elemental composition data is given in units of mmu as calculated using the program installed in the data system.

*Synthesis of 4-chloro-3-nitrocoumarin (**3**)*

4-Hydroxycoumarin (**1**) was nitrated using 72 % HNO_3 in glacial AcOH to afford 4-hydroxy-3-nitrocoumarin (**2**).[16] The starting compound 4-chloro-3-nitrocoumarin (**3**) was prepared from **2** following the method of Kaljaj *et al*.[17] The preparation was carried out in the following manner: *N*,*N*-dimethylformamide (DMF, 2 mL, 26 mmol) was cooled to 10 °C in an ice bath. Under stirring, $POCl_3$ (4.0 g, 26 mmol) was added dropwise, and the obtained mixture was stirred for an additional 15 min. Then, the ice bath was removed and the reaction was left to proceed at room temperature for a further 15 min. Finally, a solution of **2** (5.4 g; 26 mmol) in DMF (12.5 mL) was added dropwise. After 15 min of stirring, the reaction was stopped by the addition of cold water (15 mL). The precipitated solid was collected by filtration and washed with saturated sodium bicarbonate solution and water. Recrystallization

from a mixture of benzene–hexane (1:1,*V*/*V*) yielded yellow crystals of **3** (5.1 g; 22.6 mmol) in 87 % yield, m.p. 162–163 °C. The procedure was repeated twice.

*General procedure for the synthesis of 4-(alkylamino)-3-nitrocoumarins **5a–i***

The solution of 4-chloro-3-nitrocoumarin (**3**, 1.0 g, 4.4 mmol) and the appropriate amine nucleophile **4a**–**i** (4.4 mmol) in ethyl acetate (10 ml) was refluxed in the presence of triethylamine (1 ml, 7.2 mmol) for 1–3 h. After cooling, the precipitated solid was filtered off and washed with ethyl acetate and water. Purity of the synthesized compounds was checked by TLC.

Antimicrobial activity

The synthesized compounds were tested against a panel of microorganisms including three Gram-positive (*Bacillus cereus* (isolate from food), *Bacillus subtilis* ATCC 6633 and *Staphylococcus aureus* ATCC 6538), nineteen Gram-negative (*Escherichia coli* ATCC 8739, *E. coli* (isolate from food), *Klebsiella pneumoniae* ATCC 10031, *Salmonella enterica* subsp. *enterica* serovar *Enteritidis* ATCC 13076, *S. enterica* (5 clinical isolates), *Shigella flexneri* (5 clinical isolates) and *Yersinia enterocolitica* (5 clinical isolates), one yeast (*Candida albicans* ATCC 10231) and one mold (*Aspergillus brasiliensis* ATCC 16404) strains.

Clinical isolates of *Y. enterocolitica*, *S. enterica* subsp. *enterica* serovar *Enteritidis* and *S. flexneri* (5 isolates per strain), together with the food isolates (*B. cereus* and *E. coli*) were obtained from the Institute of Public Health, Niš, Serbia and are stored in a private microbiological collection. The isolation of the mentioned strains was performed from stool (fecal) samples of patients with a diarrheal syndrome.

Testing of antimicrobial activity

The antimicrobial activity was evaluated using the broth microdilution method.[18] Minimum inhibitory concentration (*MIC*) determinations were performed by a serial dilution method in 96-well microtitre plates. Bacterial species were cultured at 37 °C in Müller–Hinton agar (MHA) and Sabouraud dextrose agar (SDA) was the culture medium for the yeast (30 °C). After 18 h of cultivation, bacterial suspensions were made in Mueller Hinton broth and their turbidity was standardized to 0.5 McFarland scale using a densitometer (DEN-1 McFarland densitometer, Biosan). The final density of the bacterial and the yeast inoculums corresponded to 5×10^5 CFU (colony forming units). A suspension of the mould (*A. niger*) was made in Sabouraud dextrose broth (SDB) and its turbidity was confirmed by viable counting in a Thoma chamber with 1×10^4 CFU as the final size of the inoculum. The inoculums were added to all wells and the plates were cultivated at 37 °C during 24 h (bacteria) or at 30 °C for 48 h (fungi). Tetracycline, chloramphenicol and nystatin served as positive controls, while the solvent (ethanol) was used as a negative control. One inoculated well was included, to allow control of the adequacy of the broth for organism growth. One non-inoculated well, free of antimicrobial agent, was also included to ensure medium sterility.

The bacterial growth was determined by adding 20 μL of a 0.5 % aqueous triphenyltetrazolium chloride (TTC) solution.[19] The MIC was defined as the lowest concentration of the tested compound that inhibited visible growth (red colored pellet on the bottom of the wells after the addition of TTC). The experiments were performed in triplicate and mean values are presented.

Statistical analyses

Agglomerative hierarchical clustering (AHC) was performed using the Excel program plug-in XLSTAT version 2011.3.02.[20] The method was applied utilizing the *MIC* values (in

μmol mL^{-1}) as original variables without any recalculation. AHC was determined using Pearson dissimilarity, where the aggregation criterion were simple linkage, unweighted pair-group average and complete linkage, the Euclidean distance, where the aggregation criterion were weighted pair-group average, unweighted pair-group average, and the Ward method.

RESULTS AND DISCUSSION

Chemistry

In the first reaction step, 4-hydroxycoumarin (**1**) was nitrated in glacial AcOH with HNO_3 to afford 4-hydroxy-3-nitrocoumarin (**2**). The starting 4-chloro-3-nitrocoumarin (**3**) was prepared from **2** in a reaction with phosphorus oxychloride and *N*,*N*-dimethylformamide.

The target compounds **5a**–**i** (three new – **5e**–**g**, three previously known – **5a**, **5c** and **5i**, and three commercially available but with no mention what so ever in the literature – **5b**, **5d** and **5h**), were prepared in the reaction of **3** and an alkylamine **4a**–**i**, respectively (1:1 mole ratio of starting materials) in ethyl acetate in the presence of two equivalents of triethylamine. The compounds were obtained in good yields (72–92 %). The synthesis scheme is presented in Fig. 1.`

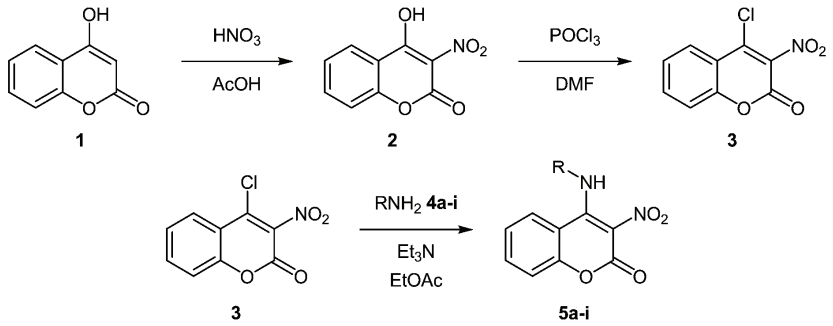


Fig. 1. Synthesis of 4-(alkylamino)-3-nitrocoumarin derivatives **5a**–**i**. (**4a**: butan-1-amine, **4b**: butan-2-amine, **4c**: 2-methylpropan-2-amine, **4d**: 3-aminopropan-1-ol, **4e**: hexan-1-amine, **4f**: octan-1-amine, **4g**: hexadecan-1-amine, **4h**: furan-2-ylmethanamine and **4i**: phenylmethanamine; R = **5a**: $CH_3(CH_2)_2CH_2$–; **5b**: $CH_3CH_2(CH_3)CH$–; **5c**: $(CH_3)_3C$–; **5d**: $HO(CH_2)_2CH_2$–; **5e**: $CH_3(CH_2)_4CH_2$–; **5f**: $CH_3(CH_2)_6CH_2$–; **5g**: $CH_3(CH_2)_{14}CH_2$–; **5h**: $C_4H_3OCH_2$– (furan-2-ylmethyl); **5i**: $C_6H_5CH_2$–.

The structures of the synthesized compounds were confirmed by analytical and spectral means (HRMS(EI), IR, ^{1}H- and ^{13}C-NMR). The physical, analytical and spectral data for compounds **5a**–**i** are given in the Supplementary material to this paper.

IR spectra of the synthesized compounds contained characteristic vibrations at 3030 to 3393 cm^{-1} corresponding to absorptions of the Ar–H and N–H bonds, and strong bands at 1677–1698 cm^{-1} corresponding to the C=O groups. The IR absorptions due to the presence of the 3-NO_2 group appeared at 1317–1364 and

1511–1556 cm^{-1}, and the C–H bonds showed characteristic vibrations at 2925––2960 cm^{-1}.

Protons from the coumarin moiety showed very similar chemical shifts in the ^{1}H-NMR spectra of compounds **5a**–**i**. Protons in the position 5 of the coumarin core were more downfield and showed higher chemical shifts (doublet of doublets at 8.35–8.52 ppm) than the other protons attached to that ring. Significant differences in the chemical shifts of the mentioned protons existed for compounds **5b**, **5c** and **5f** where the signals for H-5 appear at 7.89–8.00 ppm. All other protons on the coumarin moiety displayed much less variation in their chemical shifts. H-7 protons appeared as a doublet of triplets or only triplets at 7.69––7.75 ppm. H-6 and H-8, the signals of which overlapped at 7.28–7.50 ppm, were more shielded. The aromatic protons of the substituent groups in compounds **5h** and **5i** showed similar values of chemical shifts and appeared as a doublet of doublets at 7.64 ppm and a multiplet at 6.34–6.45 ppm, as well as a multiplet at 7.26–7.34 ppm, respectively. The remaining protons, from the alkyl groups bonded to the coumarin moiety, were as expected more upfield.

Antimicrobial activity

The created library was first evaluated for *in vitro* antimicrobial activity against a wide range of bacteria (four Gram-negative and three Gram-positive) as well as against a yeast and a mold species. The values of their minimal inhibitory concentrations (*MIC*, nmol mL^{-1}), determined in a microdilution assay, revealed the diverse effect of these compounds on the growth of microorganisms (all strains were susceptible to the presence of these compounds in their nutritive medium with an exception of *E. coli* (isolate) and *A. brasiliensis* (Table I)). The

TABLE I. Minimal inhibitory concentrations (*MIC* / nmol mL^{-1}) of the coumarins **5a**–**i**; NT – not tested, / – not sensitive in the range of the tested concentrations; AB – reference antibiotic

Bacterial strain		Compound										
		5a	**5b**	**5c**	**5d**	**5e**	**5f**	**5g**	**5h**	**5i**	AB[a]	AB[b]
Gram-negative[c]												
E. coli	ATCC 8739	3.0	12	–	–	11	2.45	1.8	–	5.3	3.5	NT
K. pneumoniae	ATCC 10031	6.0	1.5	6.0	–	–	–	0.91	–	5.3	0.88	NT
S. enterica	ATCC 13076	6.0	95	191	189	5.4	4.9	3.6	1.4	5.3	0.44	NT
Gram-positive												
B. cereus	Isolate from food	–	–	–	–	–	–	232	44	–	0.44	NT
B. subtilis	ATCC 6633	–	–	–	–	–	4.9	29	–	/	0.22	NT
S. aureus	ATCC 6538	6.0	6.0	6.0	189	1.3	1.2	3.6	5.4	21	0.22	NT
Fungal strain[d]												
C. albicans	ATCC 10231	0.34	0.05	3.0	11.8	0.04	0.08	0.44	1.4	0.64	NT	6.8

[a]Tetracycline; [b]nystatin; [c]the food isolate of *E. coli* was resistant to all compounds in the tested range of concentrations; [d]the mold *A. brasiliensis* (ATCC 16404) was resistant to all compounds in the tested range of concentrations

active inhibitory concentrations ranged from 0.04 to 232 nmol mL^{-1}, with **5g** being the most non-selective in its action. The best MIC value was observed for compound **5e** against *C. albicans*. The remaining compounds displayed antimicrobial activity with no special selectivity towards either Gram-positive or Gram-negative bacterial strains.

Among the Gram-positive strains, *B. cereus* was the most resistant strain, susceptible to only two out of the nine tested compounds, but at relatively high concentrations (**5h** and **5g**, Table I). *S. aureus* was the most sensitive bacterial strain inhibited by all tested compounds in a concentration range 1.2–189 nmol mL^{-1}. The activity of the compounds against Gram-negative bacteria was characterized by two extreme cases – *E. coli* (a food isolate which showed complete resistance to all the tested compounds, as well to the standard antibiotic) and *S. enterica* (an ATCC reference strain (13076) showing the highest susceptibility to the tested compounds with *MIC* values in the range (1.4 to 191 nmol mL^{-1}) similar to the ones for Gram-positive bacteria). This strong resistance of *E. coli* could be explained by its origin (food isolate, a probable cause of food-borne infections) and this was corroborated by the higher registered *MIC* values for an ATCC strain of *E. coli* (Table I). All compounds were also proven to possess good anticandidal activity (0.04–12 nmol mL^{-1}) but had no effect at the highest tested concentration on the growth of the other tested fungal organism (the mold *A. brasiliensis*) (see Table I).

Prompted by the results of this initial screening of antimicrobial activity, a selection of three compounds (based on their structural diversity, **5b** having a short *N*-alkyl substituent; **5g** having the longest *N*-alkyl substituent and among the most potent ones thus far and **5i** possessing a *N*-benzyl group) were further tested against fifteen bacterial strains representing gastrointestinal pathogenic bacteria isolated from human clinical samples (Table II). Table II also contains clinically relevant data on the patient diagnosed condition caused by the particular bacterial isolate.

Human sample isolates (five per strain) of *Y. enterocolitica*, *S. enterica* and *Sh. flexneri* showed themselves to be highly resistant to the influence of the tested coumarinic derivatives (most of which having *MIC* values higher than 1.9 nmol mL^{-1}). The differing effect on (clinical/food) isolates and ATCC strains (Tables I and II) was evident from the high resistance of the isolates to the control antibiotic used in this study as well. In the light of these facts, compound **5b** looks highly promising, since it was the only compound among the three tested, which showed inhibitory action against *S. enterica* with *MIC* values of 0.48 and 0.95 μmol mL^{-1}. Although this does not appear significant at first, this activity is more than relevant if one considers the fact that this strain was completely resistant to the highest tested concentration of the reference antibiotic (0.11 μmol mL^{-1}).

TABLE II. Effect of the selected coumarin derivatives (**5b**, **5g** and **5i**) on growth of three gastrointestinal pathogenic bacterial strains (isolates)

Bacteria	Isolate No.	Diagnosis	*MIC* / μmol mL^{-1}			
			5b	**5g**	**5i**	Chloram-phenicol
Yersinia enterocolitica	1	Acute enterocolitis	>1.9	>1.2	>1.7	0.11
	2		>1.9	>1.2	>1.7	0.11
	3		>1.9	0.58	0.85	0.027
	4		>1.9	0.29	>1.7	0.054
	5		>1.9	>1.2	>1.7	0.11
Salmonella enterica	1	Gastroenterocolitis	0.95	>1.2	>1.7	>0.11
	2		0.95	>1.2	>1.7	>0.11
	3		0.48	>1.2	>1.7	>0.11
	4		>1.9	>1.2	>1.7	>0.11
	5		0.95	>1.2	>1.7	>0.11
Shigella flexneri	1	Shigellosis	>1.9	0.87	>1.7	0.11
	2		0.95	0.29	1.3	0.11
	3		>1.9	0.87	>1.7	0.11
	4		>1.9	>1.2	>1.7	0.11
	5		0.95	0.87	>1.7	0.11

Hitherto, a substantial amount of data on the antimicrobial potential of a number of *N*-substituted 4-amino-3-nitrocoumarins has been accumulated.[13–15,21] Using multivariate statistical analyses (MVA), in an attempt to determine their similarities and selectivity against certain pathogenic strains, the relations between the *MIC* values of the herein tested coumarin derivatives and those previously published were analyzed.[13–15,21] The compounds taken into consideration (Fig. 2) for the statistical analysis were those that were previously tested

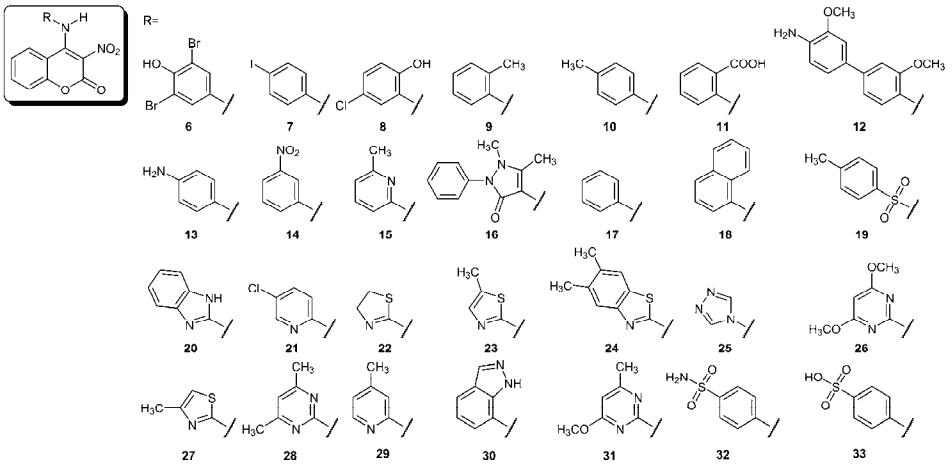


Fig. 2. Structures of coumarin derivatives (**6**–**33**) included in the multivariate statistical analyses.

against the same panel of microorganisms as the compounds tested in this work. Agglomerative hierarchical analysis using *MIC* values against seven bacterial and two fungal strains as variables resulted in the dendrogram shown in Fig. 3 with three clearly separated groups/classes of compounds (C1–C3). Group centroids (the average representative of a group) were also plotted against microorganism strains tested (Fig. 4) to give a better picture of the distinctions existing between the groups.

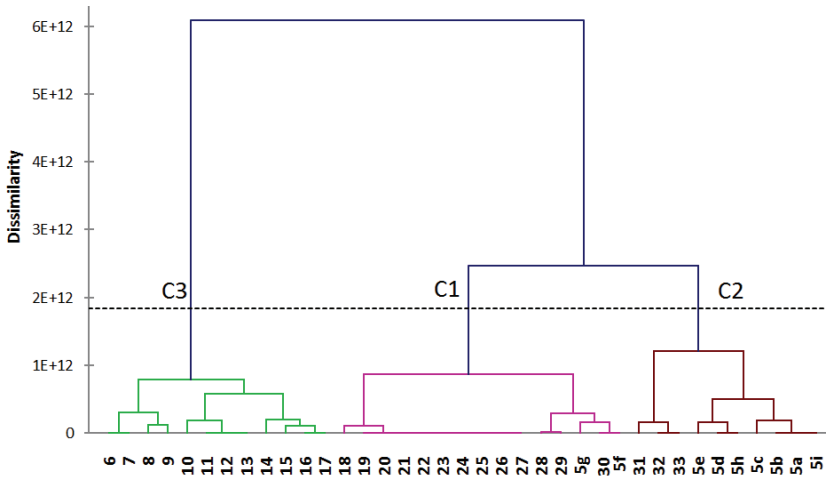


Fig. 3. Dendrogram (AHC analysis) representing the dissimilarity in the antimicrobial activity (*MIC* values of selected coumarin derivatives used as variables against five bacterial and two fungal strains) relationships of thirty-seven compounds obtained by Euclidean distance dissimilarity, using the aggregation criterion-the Ward method. Three groups of compounds were found: C1–C3.

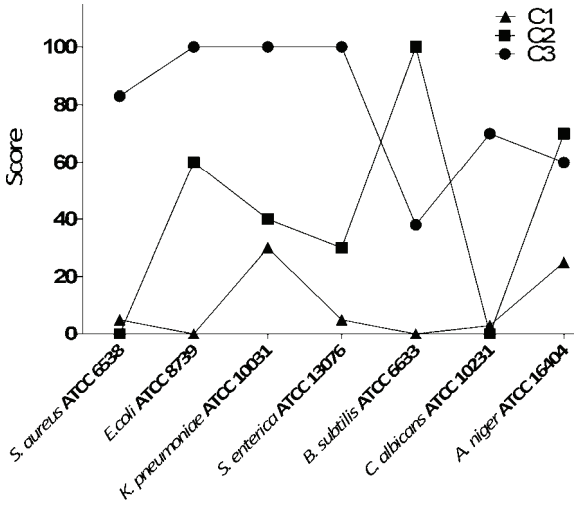


Fig. 4. Graphical representation of the scores (the lowest activity, *i.e.*, the highest value of *MIC*, was given the score 100 and the best activity value 0) of the centroids (classes C1–C3) from the performed AHC analysis (for details see Fig. 3).

Group C1 (Figs. 3 and 4) consisted of compounds that had rather non-selective activity with the lowest *MIC* values towards *E. coli, K. pneumoniae, S. enterica, B. subtilis, C. albicans* and *A. brasiliensis*. Among these microorganisms, *K. pneumoniae* and *A. brasiliensis* were the ones more resistant to the action of compounds from group C1. Compounds **5f** and **5g** from the present paper belong to this group. The second group, designated as C2, contained all other compounds prepared in this work (**5a**–**e**, **5h** and **5i**). Group C2 compounds demonstrated a greater degree of selectivity, highly affecting the growth of *S. aureus* and *C. albicans*, while remaining relatively active against the other strains in comparison to group C1. All the other statistically compared compounds were grouped into C3 according to their overall low impact on bacterial growth (*i.e.*, high *MIC* values) – the only exception being their action against *B. subtilis* (Fig. 4).

From all of the above mentioned, it could be conclude that group 1 (C1) substances were the most potent (in terms of both their *MIC* values and broad spectrum of bacteria that the compounds affect) and group 2 substances are the most selective ones (even though they have lower *MIC* values).

Chemically speaking, compounds **5f** and **5g** differed from the rest of the compounds prepared in this work by having the longest alkyl chains attached to the nitrogen in position 4. This structural feature, a long lipophilic moiety (octyl and hexadecyl groups in **5f** and **5g**, respectively), made them less selective in their antimicrobial action, suggesting that lipophilicity, *i.e.*, an interaction with the cell membrane, might play a crucial role in their activity. It would be natural then to expect, as observed, that such compounds have an effect on more different strains (be less selective) by acting on/disrupting their membranes. Along with these *N*-alkyl derivatives (**5f** and **5g**), other members of C1 were almost invariantly *N*-heteroaryl derivatives (**20**–**30**). These could be regarded as possible hybrids of the coumarin core with other pharmacophoric groups (*i.e.*, those heterocycles that are known to possess an antimicrobial action). This hybridization appears to have resulted in the onset of the strongest observed activity.

Group C2 compounds demonstrated that the presence of shorter alkyl substituents on N-4 lowers the activity but also that the nucleophilicity/basic nature of this nitrogen could be relevant for the overall noted activity. The weakly nucleophilic/basic anilino-type nitrogen, present in all compounds from group C3, might be the reason why these compounds showed much lower potency in inhibiting the growth of microorganisms when compared to the activity of the much more basic/nucleophilic (at N-4) compounds from C2. Not even the presence of a phenol group in compounds **6** and **8** (known to be the carrier of antimicrobial activity of many natural compounds, thymol, for example) had a significant impact on the activity of the *N*-aryl coumarin derivatives grouped in C3. Moreover, the size of the *N*-substituent seems not to be of importance in the appearance of antimicrobial activity, as exemplified by the low activity of compounds **12** and **18**.

Bearing all of this in mind, it could be concluded that the nature of the *N*-substituent at position 4 profoundly affects the antimicrobial potency of 3-nitrocoumarins. The respective activity decreases in the following order *N*-long alkyl chains ≈ *N*-heteroaryl > *N*-short alkyl chains >> *N*-aryl derivatives.

CONCLUSION

In the present paper, a wide range of activities, ranging from completely inactive to active against antibiotic-resistant strains, was registered for the studied 4-(alkylamino)-substituted 3-nitrocoumarins. The most important information, relevant for any possible future applications of these compounds, is the fact that the tested coumarins exhibited activity against strains isolated from fecal samples of the patients with acute diarrheal syndrome. Bacterial infections of the digestive tract are considered the major cause of morbidity and mortality, especially in children, the elderly and immunosuppressed patients, both in industrialized and developing countries.[22] These results demonstrate the substantial antimicrobial potential of these coumarins in the therapy of GI and other infections caused by highly resistant strains. Furthermore, the activity observed against *S. aureus* was quite significant, knowing that this bacterial strain can be multidrug resistant.[23] A multivariate treatment of the results of the antimicrobial assays obtained in this work and those previously attained enabled the tested compounds to be classified according to their overall activity, selectivity and structural requirements for the observed activities to be discussed. A general conclusion could be reached that both a greater lipophilicity of the *N*-substituent, and nucleophilicity and basicity of this nitrogen have beneficial effects on the antimicrobial activity of these compounds. Such results and conclusions allow and direct further studies focused on an even larger enhancement of the here obtained activity by additional chemical modification of these coumarin antimicrobials.

SUPPLEMENTARY MATERIAL

Physical, analytic and spectral data for the prepared derivatives are available electronically from http://www.shd.org.rs/JSCS/, or from the corresponding author on request.

Acknowledgements. This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Project Nos. 172061 and 45022).

ИЗВОД

МАЛА СИНТЕТСКА БИБЛИОТЕКА 4-(АЛКИЛАМИНО)-3-НИТРОКУМАРИНА СА ЈАКОМ АНТИМИКРОБНОМ АКТИВНОШЋУ ПРОТИВ ГАСТРОИНТЕСТИНАЛНИХ ПАТОГЕНА

НИКО С. РАДУЛОВИЋ[1], ЗОРИЦА СТОЈАНОВИЋ-РАДИЋ[2], ПРЕДРАГ СТОЈАНОВИЋ[3,4], НИКОЛА СТОЈАНОВИЋ[3], ВИДОСЛАВ ДЕКИЋ[5] и БИЉАНА ДЕКИЋ[5]

[1]Департман за хемију, Природно–математички факултет, Универзитет у Нишу, Вишеградска 33, 18000 Ниш, [2]Департман за биологију и екологију, Природно–математички факултет, Универзитет у Нишу, Вишеградска 33, 18000 Ниш, [3]Медицински факултет, Универзитет у Нишу, Зорана Ђинђића 81, 18000 Ниш, [4]Институт за јавно здравље, Центар за микробиологију, 18000 Ниш, и [5]Природно–математички факултет, Универзитет у Приштини, Лоле Рибара 29, 38220 Косовска Митровица

Инфекције изазване микроорганизмима са развијеном резистентношћу према антибиотицима тренутно представљају главни узрок озбиљних здравствених компликација код пацијената. Стога су истраживања која за циљ имају откриће нових и/или активнијих антимикробних агенаса веома важна. Имајући у виду потврђену антимикробну активност природних и синтетичких деривата кумарина, овај рад треба да пружи додатни увид у антимикробни потенцијал кумаринских једињења испитивањем активности библиотеке од девет нових 4-(алкиламино)-3-нитрокумарина против 24 соја микроорганизама, који представљају лабораторијске сојеве и резистентне (изолати) бактеријске и фунгалне патогене. Сва једињења су показала одређени степен селективне активности, која је у неким случајевима била нарочито изражена достижући минималну инхибициону концентрацију од 0,04 nmol mL^{-1} (12 ng mL^{-1}) против гљивице *Candida albicans*. Утврђена активност је већа према Грам-негативним сојевима, а најосетљивија, и међу ATCC сојевима и клиничким изолатима, је била *Salmonella enterica* subsp. *enterica* serovar *Enteritidis*. Ови резултати указују на потенцијал ових кумарина као антимикробних агенаса (већина је активна у концентрацијама реда nmol/ml) у третману гастроинтестиналних и других инфекција изазваних високо резистентним сојевима микроорганизама. Коначно, ради лакшег повезивања структуре и активности урађена је мултиваријантна статистичка анализа података добијених у овом раду и оних из претходних студија антимикробне активности структурно сличних кумарина.

(Примљено 19. јуна, ревидирано 22. августа, прихваћено 29. августа 2014)

REFERENCES

1. D. J. Hentges, in *Human Intestinal Microflora in Health and Disease*, D. J. Hentges, Ed., Academic Press, New York, 1983, p. 311
2. D. Gilliss, A. Cronquist, M. Carter, M, Tobin-D'Angelo, D. Blythe, K. Smith, S. Lathrop, G. Birkhead, P. Cieslak, J. Dunn, K. G Holt, J. J. Guzewich, O. L. Henao, P. Griffin, S. M. Crim, O. L. Henao, *Vital Signs: Incidence and Trends of Infection with Pathogens Transmitted Commonly Through Food – Foodborne Diseases Active Surveillance Network, 10 U.S. Sites, 1996–2010*, Morbidity and mortality weekly report 60, Centers for Disease Control and Prevention, Atlanta, GA, 2011, p. 749
3. H. Rotterdam, P. Tsang, *Hum. Pathol.* **25** (1994) 1123
4. R. D. H. Murray, J. Mendez, S. A. Brown, *The Natural Coumarins: Occurrence, Chemistry and Biochemistry*, Wiley, New York, 1982
5. R. O'Kennedy, R. D. Thornes, *Coumarins - Biology, Applications and Mode of Action*, Wiley, Chichester, 1997
6. I. Kostova, *Curr. Med. Chem.: Anti-Cancer Agents* **5** (2005) 29

7. F. Borges, F. Roleira, N. Milhazes, L. Santana, E. Uriarte, *Curr. Med. Chem.* **12** (2005) 887
8. Z. Debeljak, A. Škrbo, I. Jasprica, A. Mornar, V. Plečko, M. Banjanac, M. Medić-Šarić, *J. Chem. Inf. Model.* **47** (2007) 918
9. N. Radulović, G. Stojanović, R. Vukićević, V. Dekić, B. Dekić, R. Palić, *Monatsh Chem.* **137** (2006) 1477
10. S. V. Dekić, V. S. Dekić, B. R. Dekić, M. S. Dekić, *Chem. Pap.* **61** (2007) 233
11. S. V. Dekić, V. S. Dekić, B. Vučić, B. R. Dekić, M. S. Dekić, *Facta Univ., Ser.: Phys., Chem. Technol.* **5** (2007) 85
12. M. S. Dekić, B. R. Dekić, V. S. Dekić, S. V. Dekić, *J. Heterocycl. Chem.* **45** (2008) 295
13. B. Dekić, V. Dekić, N. Radulović, R. Vukićević, R. Palić, *Chem. Pap.* **64** (2010) 354
14. B. Dekić, N. Radulović, V. Dekić, R. Vukićević, R. Palić, *Molecules* **15** (2010) 2246
15. N. Radulović, G. Bogdanović, P. Blagojević, V. Dekić, R. Vukićević, *J. Chem. Crystallogr.* **41** (2011) 545
16. V. Savel'ev, O. Artamonova, V. Troitskaya, V. Vinokurov, V. Zagorevskii, *Khim. Geterotsikl. Soedin*. **7** (1973) 885
17. V. Kaljaj, M. Trkovnik, L. Stefanović-Kaljaj, *J. Serb. Chem. Soc.* **52** (1987) 183
18. *Clinical and Laboratory Standards Institute, Performance standards for antimicrobial susceptibility testing. Eleventh informational supplement*. Document M 100-S11, NCCLS, Wayne, PA, 2003.
19. N. Radulović, M. Dekić, Z. Stojanović-Radić, R. Palić, *Turk. J. Chem.* **35** (2011) 499
20. N. S. Radulović, M. S. Dekić, *Chem. Biodivers.* **10** (2013) 2042
21. V. Dekić, N. Radulović, R. Vukićević, B. Dekić, Z. Stojanović-Radić, R. Palić, *Afr. J. Pharm. Pharmacol.* **5** (2011) 371
22. B. Miljković-Selimović, T. Babić, P. Stojanović, *Sci. J. Faculty of Med. Niš* **27** (2010) 55
23. N. Radulović, P. Blagojević, Z. Stojanović-Radić, N. Stojanović, *Curr. Med. Chem.* **20** (2013) 932.

Journal of the Serbian Chemical Society
JSCS@tmf.bg.ac.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) S63–S66 (2015) *Supplementary material*

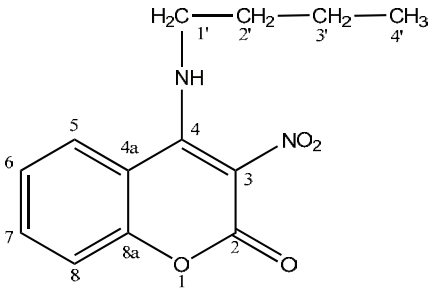
SUPPLEMENTARY MATERIAL TO

A small library of 4-(alkylamino)-3-nitrocoumarin derivatives with potent antimicrobial activity against gastrointestinal pathogens

NIKO S. RADULOVIĆ[1*#], ZORICA STOJANOVIĆ-RADIĆ[2], PREDRAG STOJANOVIĆ[3,4], NIKOLA STOJANOVIĆ[3], VIDOSLAV DEKIĆ[5#] and BILJANA DEKIĆ[5#]

[1]Department of Chemistry, Faculty of Science and Mathematics, University of Niš, Višegradska 33, 18000 Niš, Serbia, [2]Department of Biology and Ecology, Faculty of Science and Mathematics, University of Niš, Višegradska 33, 18000 Niš, Serbia, [3]Faculty of Medicine, University of Niš, Zorana Đinđića 81, 18000 Niš, Serbia, [4]Institute For Public Health Niš, Center of Microbiology, 18000 Niš, Serbia and [5]Department of Chemistry, Faculty of Science and Mathematics, University of Priština, Lole Ribara 29, 38220 Kosovska Mitrovica, Serbia

J. Serb. Chem. Soc. 80 (3) (2015) 315–327



Atomic numbering of the coumarin derivatives using **5a** as an example.

PHYSICAL, ANALYTIC AND SPECTRAL DATA FOR THE PREPARED DERIVATIVES

*4-(Butylamino)-3-nitro-2*H*-chromen-2-one (**5a**)*. Yield: 86 %; color: yellow; m.p.: 165–167 °C (lit. 168–169 °C[1]); Anal. calcd. for $C_{13}H_{14}N_2O_4$: C, 59.54; H, 5.38; N, 10.68 %. Found C, 59.41; H, 5.28; N, 10.51 %; IR (neat, cm^{-1}): 3346––3180 (N–H and Ar–H), 2960 (C–H), 1683 (C=O), 1608 (C=C), 1520 and 1354 (NO_2), 1218, 1062, 744, 623; ^{1}H-NMR (200 MHz, DMSO-d_6, δ / ppm): 0.86 (3H, *t*, *J* = 7.0 Hz, H-4′), 1.22–1.38 (2H, *m*, H-3′), 1.56–1.70 (2H, *m*, H-2′), 3.10––3.16 (2H, *m*, H-1′), 7.38–7.49 (2H, *m*, H-6, H-8), 7.73 (1H, *dt*, *J* = 8.5, 1.5 Hz, H-7), 8.37 (1H, *dd*, *J* = 8.2, 1.5 Hz, H-5), 8.54 (1H, *brs*, N–H); ^{13}C-NMR (50

* Corresponding author. E-mail: nikoradulovic@yahoo.com

MHz, DMSO-d_6, δ / ppm): 13.6, 19.6, 30.6, 44.3, 114.3, 115.7, 117.5, 124.3, 124.7, 134.0, 147.0, 151.0, 155.5; HRMS(EI): M^+ ($C_{13}H_{14}N_2O_4$), 262.0972; requires 262.0954 (Δ = +1.8 mmu); EI-MS (*m/z*, (relative abundance, %)): 262 (M^+, 23), 216, 205, 159 (BP, 100), 131.

*4-(*sec-*Butylamino)-3-nitro-2*H*-chromen-2-one (**5b**)*. Yield: 72 %; color: orange; m.p.: 208–210 °C; Anal. Calcd. for $C_{13}H_{14}N_2O_4$: C, 59.54; H, 5.38; N, 10.68 %. Found C, 59.48; H, 5.32; N, 10.60 %; IR (neat, cm^{-1}): 3338–3062 (N–H and Ar–H), 2933 (C–H), 1677 (C=O), 1606 (C=C), 1556 and 1317 (NO_2), 1061, 916, 743; ^{1}H-NMR (200 MHz, DMSO-d_6, δ / ppm): 1.02 (3H, *t*, *J* = 6.4 Hz, H-3′), 1.52 (3H, *d*, *J* = 6.5 Hz, H-1″), 1.73–1.87 (2H, *m*, H-2′), 4.15–4.36 (1H, *m*, H-1′), 7.33–7.40 (2H, *m*, H-6, H-8), 7.69 (1H, *dt*, *J* = 7.5, 1.5 Hz, H-7), 7.89 (1H, *dd*, *J* = 8.5, 1.5 Hz, H-5), 9.99 (1H, *brs*, N–H); ^{13}C-NMR (50 MHz, DMSO-d_6, δ / ppm): 10.1, 21.3, 31.3, 55.3, 110.6, 113.1, 118.6, 124.2, 126.2, 134.9, 152.8, 153.2, 154.5; HRMS(EI): M^+ ($C_{13}H_{14}N_2O_4$), 262.0942; requires 262.0954 (Δ = –1.2 mmu); EI-MS (*m/z*, (relative abundance, %)): 262 (M^+, 16), 247, 216, 205, 159 (BP, 100).

*4-(*tert-*Butylamino)-3-nitro-2*H*-chromen-2-one (**5c**)*. Yield: 75 %; color: yellow; m.p.: 134–136 °C (lit. 136–138 °C[1]); Anal. Calcd. for $C_{13}H_{14}N_2O_4$: C, 59.54; H, 5.38; N, 10.68 %. Found C, 59.60; H, 5.37; N, 10.77 %; IR (neat, cm^{-1}): 3325–3036 (N–H and Ar–H), 2975 (C–H), 1698 (C=O), 1600 (C=C), 1543 and 1360 (NO_2), 1176, 1030, 762; ^{1}H-NMR (200 MHz, DMSO-d_6, δ / ppm): 1.59 (9H, *s*, CH_3), 7.28–7.40 (2H, *m*, H-6, H-8), 7.69 (1H, *dt*, *J* = 8.0, 1.0 Hz, H-7), 7.98 (1H, *dd*, *J* = 8.2, 1.0 Hz, H-5), 9.56 (1H, *brs*, N–H); ^{13}C-NMR (50 MHz, DMSO-d_6, δ / ppm): 31.7 (three C), 59.3, 114.4, 114.5, 117.9, 123.4, 128.8, 135.0, 152.6, 154.4, 155.2; HRMS(EI): M^+ ($C_{13}H_{14}N_2O_4$), 262.0947; requires 262.0954 (Δ = –0.7 mmu); EI-MS (*m/z*, (relative abundance %)): 262 (M^+, 10), 247, 216, 205, 159 (BP, 100).

*4-[(3-Hydroxypropyl)amino]-3-nitro-2*H*-chromen-2-one* (**5d**). Yield: 83 %; color: yellow; m.p.: 158–160 °C; Anal. Calcd. for $C_{12}H_{12}N_2O_5$: C, 54.55; H, 4.58; N, 10.60 %. Found C, 54.63; H, 4.60; N, 10.53 %; IR (neat, cm^{-1}): 3393––3216 (N–H and Ar–H), 2941 (C–H), 1695 (C=O), 1609 (C=C), 1551 and 1347 (NO_2), 1223, 1051, 879, 757, 637; ^{1}H-NMR (200 MHz, DMSO-d_6, δ / ppm): 1.76–1.89 (2H, *m*, H-2′), 3.25 (2H, *t*, *J* = 6.6 Hz, H-1′), 3.48 (2H, *t*, *J* = 5.8 Hz, H-3′), 4.77 (1H, brs, OH), 7.38–7.48 (2H, *m*, H-6, H-8), 7.73 (1H, *dt*, *J* = 8.4, 1.2 Hz, H-7), 8.31 (1H, *dd*, *J* = 8.2, 1.2 Hz, H-5), 8.57 (1H, *brs*, N–H); ^{13}C-NMR (50 MHz, DMSO-d_6, δ / ppm): 31.4, 42.7, 58.5, 114.3, 115.6, 117.6, 124.3, 124.7, 134.1, 147.4, 151.0, 155.5; HRMS(EI): M^+ ($C_{12}H_{12}N_2O_5$), 264.0757; requires 264.0746 (Δ = +1.1 mmu); EI-MS (*m/z*, (relative abundance %)): 264 (M^+, 7), 246, 200, 159 (BP, 100), 115.

*4-(Hexylamino)-3-nitro-2*H*-chromen-2-one* (**5e**). Yield: 92 %; color: yellow; m.p.: 116–118 °C; Anal. Calcd. for $C_{15}H_{18}N_2O_4$: C, 62.06; H, 6.25; N, 9.65 %.

Found C, 61.97; H, 6.15; N, 9.54 %; IR (neat, cm^{-1}): 3345–3216 (N–H and Ar–H), 2925 (C–H), 1683 (C=O), 1605 (C=C), 1542 and 1364 (NO_2), 1219, 1059, 753, 625; ^{1}H-NMR (200 MHz, DMSO-d_6, δ / ppm): 0.79–0.88 (3H, *m*, H-6′), 1.17–1.31 (6H, *m*, H-3′, H-4′, H-5′), 1.55–1.72 (2H, *m*, H-2′), 3.14 (2H, *t*, *J* = 7.6 Hz, H-1′), 7.39–7.49 (2H, *m*, H-6, H-8), 7.73 (1H, *dt*, *J* = 8.2, 1.0 Hz, H-7), 8.35 (1H, *dd*, *J* = 8.0, 1.0 Hz, H-5), 8.47 (1H, *brs*, N–H); ^{13}C-NMR (50 MHz, DMSO-d_6, δ / ppm): 13.9, 22.0, 26.0, 28.6, 30.8, 44.6, 114.3, 115.7, 117.6, 124.3, 124.7, 134.0, 147.0, 151.0, 155.4; HRMS(EI): M^+ ($C_{15}H_{18}N_2O_4$), 290.1276; requires 290.1267 (Δ = +0.9 mmu); EI-MS (*m/z*, (relative abundance %)): 290 (M^+, 10), 244, 219, 205, 159 (BP, 100).

*3-Nitro-4-(octylamino)-2*H*-chromen-2-one* (***5f***). Yield: 78 %; color: yellow; m.p.: 128–130 °C; Anal. Calcd. for $C_{17}H_{22}N_2O_4$: C, 64.13; H, 6.97; N, 8.80 %. Found C, 64.08; H, 7.04; N, 8.78 %; IR (neat, cm^{-1}): 3337–3048 (N–H and Ar–H), 2957 (C–H), 1685 (C=O), 1603 (C=C), 1539 and 1359 (NO_2), 1215, 1055, 916, 750; ^{1}H-NMR (200 MHz, DMSO-d_6, δ / ppm): 0.89 (3H, *t*, *J* = 6.6 Hz, H-8′), 1.29–1.52 (10H, *m*, H-3′, H-4′, H-5′, H-6′, H-7′), 1.79–1.93 (2H, *m*, H-2′), 3.82–3.91 (2H, *m*, H-1′), 7.32–7.40 (2H, *m*, H-6, H-8), 7.69 (1H, *dt*, *J* = = 8.2, 1.4 Hz, H-7), 8.00 (1H, *dd*, *J* =8.4, 1.4 Hz, H-5), 10.16 (1H, *brs*, N–H); ^{13}C-NMR (50 MHz, DMSO-d_6, δ / ppm): 13.9, 22.5, 26.5, 28.9 (two C), 30.3, 31.6, 48.5, 113.2, 118.5, 124.1, 126.6, 126.7, 135.0, 152.9, 154.5, 158.5; HRMS(EI): M^+ ($C_{17}H_{22}N_2O_4$), 318.1574; requires 318.1580 (Δ = –0.6 mmu); EI-MS (*m/z*, (relative abundance %)): 318 (M^+, 7), 272, 205, 159 (BP, 100).

*4-(Hexadecylamino)-3-nitro-*2H*-chromen-2-one* (***5g***). Yield: 81 %; color: yellow; m.p.: 120–122 °C; Anal. Calcd. for $C_{25}H_{38}N_2O_4$: C, 69.74; H, 8.90; N, 6.51 %. Found C, 69.82; H, 8.88; N, 6.49 %; IR (neat, cm^{-1}): 3336–3180 (N–H and Ar–H), 2955 (C–H), 1690 (C=O), 1609 (C=C), 1517 and 1327 (NO_2), 1066, 787, 747; ^{1}H-NMR (200 MHz, DMSO-d_6, δ / ppm): 0.85 (3H, *t*, H-16′), 1.22 (26H, *brs*, H-3′–15′), 1.51–1.64 (2H, *m*, H-2′), 7.39–7.48 (2H, *m*, H-6, H-8), 3.10 (2H, *t*, H-1′), 7.74 (1H, *dt*, *J* = 8.0, 1.0 Hz, H-7), 8.36 (1H, *dd*, *J* = 8.4, 1.0 Hz), H-5; ^{13}C-NMR (50 MHz, DMSO-d_6, δ / ppm): 14.1, 22.2, 26.2, 28.5, 28.8, 28.9, 29.0, 29.1 (five C), 31.4 (two C), 44.6, 96.8, 114.3, 117.6, 124.3, 124.7, 134.0, 147.0, 151.0, 155.5; HRMS(EI): M^+ ($C_{25}H_{38}N_2O_4$), 430.2823; requires 430.2832 (Δ = –0.9 mmu); EI-MS (*m/z*, (relative abundance %)): 430 (M^+, 4), 384, 205, 159 (BP, 100).

*4-[(Furan-2-ylmethyl)amino]-3-nitro-2*H*-chromen-2-one* (***5h***). Yield: 74 %; color: yellow; m.p.: 234–236 °C; Anal. Calcd. for $C_{14}H_{10}N_2O_5$: C, 58.74; H, 3.52; N, 9.79 %. Found C, 58.78; H, 3.61; N, 9.80 %; IR (neat, cm^{-1}): 3336––3130 (N–H and Ar–H), 1686 (C=O), 1605 (C=C), 1551 and 1326 (NO_2), 1015, 884, 750; ^{1}H-NMR (200 MHz, DMSO-d_6, δ / ppm): 3.42 (2H, *s*, CH_2), 6.34––6.45 (2H, *m*, H-4′, H-5′), 7.40–7.50 (2H, *m*, H-6, H-8), 7.64 (1H, *dd*, *J* = 2.6, 0.8 Hz, H-3′), 7.75 (1H, *dt*, *J* = 8.4, 1.4 Hz, H-7), 8.35 (1H, *dd*, *J* = 8.4, 1.4 Hz, H-5),

9.09 (1H, *brs*, N–H); ^{13}C-NMR (50 MHz, DMSO-d_6, δ / ppm): 40.0, 109.4, 111.0 (two C), 114.4, 117.7, 124.5, 124.9, 134.4, 143.5, 147.3, 149.0, 151.2, 155.3; HRMS(EI): M^+ ($C_{14}H_{10}N_2O_5$), 286.0577; requires 286.0590 (Δ = –1.3 mmu); EI-MS (*m/z*, (relative abundance %)): 286 (M^+, 37), 240, 205, 159 (BP, 100), 81.

*4-(Benzylamino)-3-nitro-2*H*-chromen-2-one* (***5i***). Yield: 86 %; color: yellow; m.p.: 182–184 °C (lit. 180–181 °C[1]); Anal. Calcd. for $C_{16}H_{12}N_2O_4$: C, 64.86; H, 4.08; N, 9.46 %. Found C, 64.79; H, 3.97; N, 9.36 %; IR (neat, cm^{-1}): 3336––3030 (N–H and Ar–H), 1682 (C=O), 1606 (C=C), 1511 and 1323 (NO_2), 1217, 1062, 738, 692; ^{1}H-NMR (200 MHz, DMSO-d_6, δ / ppm): 3.36 (2H, *s*, CH_2), 7.26–7.34 (5H, *m*, H-2′, H-3′, H-4′, H-5′, H-6′), 7.41–7.50 (2H, *m*, H-6, H-8), 7.75 (1H, *t*, *J* = 8.0 Hz, H-7), 8.42 (1H, *dd*, *J* = 8.2, 1.0 Hz, H-5), 9.21 (1H, *brs*, N–H); ^{13}C-NMR (50 MHz, DMSO-d_6, δ / ppm): 47.9, 114.4, 116.3, 117.6, 124.5, 124.8, 127.3 (two C), 127.8, 128.7 (two C), 134.2, 136.7, 147.3, 151.1, 155.3; HRMS(EI): M^+ ($C_{16}H_{12}N_2O_4$), 296.0813; requires 296.0797 (Δ = +1.6 mmu); EI-MS (*m/z*, (relative abundance %)): 296 (M^+, 44), 250, 205, 159 (BP, 100), 91.

REFERENCES

1. V. Savel’ev, O. Artamonova, V. Troitskaya, V. Vinokurov, V. Zagorevskii, *Khim. Geterotsikl. Soedin.* **7** (1973) 885.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 329–342 (2015)
JSCS–4719
UDC 547.233+546.763+546.121.13:534.1:531.121
Original scientific paper

Derivation of a new set of force field parameters for ammine complexes of chromium(III) containing halogenido ligands: modeling of the *trans*-influence of halogenido ligands

IVANA DJORDJEVIĆ[1], SONJA GRUBIŠIĆ[1]*, MILOŠ MILČIĆ[2] and SVETOZAR NIKETIĆ[1]

[1]*Center for Chemistry, ICTM, University of Belgrade, Njegoševa 12, 11001 Belgrade, Serbia and* [2]*Faculty of Chemistry, University of Belgrade, Studentski trg 16, 11001 Belgrade, Serbia*

(Received 3 September, revised 26 October, accepted 27 October 2014)

Abstract: An approach to model the *trans*-influence using partial atomic charges derived from the molecular electrostatic potential by means of the restrained electrostatic potential (RESP) fitting method is exemplified on a series of halogenido–ammine octahedral chromium(III) complexes. RESP charges incorporated in the present vibrationally optimized consistent force field account for second-order phenomena, improve the modeling and assignment of skeletal vibrations, and reproduce the trends in frequency shifts along the F, Cl, Br, I series. In addition, a supplementary statistical analysis is given for the Cr–halogen bonds in the crystal structures from the CSD.

Keywords: partial atomic charges; chromium(III) complexes; *trans*-influence; consistent force field; skeletal vibrations.

INTRODUCTION

Elucidating structures and properties of chemical species by computational methods has been a growing, fruitful and helpful research practice ever since the emergence of computational chemistry. Essentially two complementary paradigms were developed in previous decades: one stemming from the study of electronic structure based on the first principles of quantum mechanics (QM) and its methods, and the other based on modeling the molecular potential energy surface with empirically derived analytic functions together with a set of adjustable parameters – known collectively as a force field (FF).

An example of synergy between the two paradigms was actualized in the development of an improved and physically more realistic treatment of atomic charges in a force field based on the monopole approximation, by fitting the

* Corresponding author. E-mail: grubisic@chem.bg.ac.rs
doi: 10.2298/JSC030914105D

charges to the QM-derived molecular electrostatic potential. In this context, it was recently shown[1] that atomic charges derived using the restrained electrostatic potential (RESP) program[2–4] constitute a convenient choice for the optimization of a new generation force field for chromium(III) complexes, which attempts to overcome the limitations of the predefined (*i.e.*, atom-type pre-assigned) atom-centered charges by taking into account the molecular environment of the atoms. (In what follows, the label RESP denotes the computer program of the same name irrespective of whether or not the restraint function is used). For example, the RESP scheme allows each nitrogen ligator atom in $[Cr(NH_3)_5X]^{2+}$ to attain a charge that depends on the arrangement of all other ligator atoms, thus taking care of the mutual influences of the ligands (commonly referred to as *trans*- or *cis*-influences[5,6]) in the octahedral coordination sphere. The latter phenomena are seldom touched upon in present-day molecular modeling, with the notable exceptions of the comprehensive studies from the Deeth group on Werner-type complexes,[7] and from Meuwly *et al*. on organometallics.[8] The atomic point-charge assignment procedure presented in this work (or its potential extensions) could provide another possible way to address the *trans*-influence in a straightforward and simple manner.

Whereas previously basis sets of various complexity were investigated in QM computation of the molecular electrostatic potential (MEP) for $[Cr(NH_3)_{6-n}(Cl)_n]^{(3-n)+}$ complexes (n = 0,1,2,3),[1] as well as the effect of restraint weighted RESP charges,[4,9] in this work, QM derived atomic charges were applied in consistent force field (CFF) calculations on $[Cr(NH_3)_5X]^{2+}$ species (X = F, Cl, Br, I) in order to validate the applicability of the RESP approach for different halogenido ligands, to gauge the anticipated *trans*-influence of X along the halogen series, and to ascertain whether the RESP-derived charges for halogens can reproduce the experimental vibrational frequencies within the vibrationally optimized force field (VOFF) for coordination compounds.[10]

METHODOLOGY

All QM computations of MEP were realized using the GAMESS electronic structure program.[11] Following former[1] guidelines to avoid the influence of the basis set on the computed MEP values and the RESP-derived partial atomic charges, the restricted open Hartree–Fock (ROHF) method with three out of 12 previously explored basis sets, *i.e.*, the triple-ζ valence basis (TZV),[12] the effective core potential valence basis (SBKJC),[13,14] and the Dunning-type correlation consistent basis set with polarization function (cc-pVTZ),[15] was used for the metal ion. The three basis sets produced practically the same final results for the optimized geometries and vibrational frequencies. For all other atoms, the following were employed: the 6-31G(d) polarized basis set[16] for N, F, and Cl atoms, the 6-311G(d) basis[17,18] for Br and I atoms, and the unpolarized Pople 6-31G basis set[19] for the ammine H atoms. This choice of the computational level offered the best compromise between feasibility of the procedures, the consistency of the results, and computational economy.

The RESP program[4] to derive partial atomic charges from MEP was used as set out previously.[1] Atomic charges were fitted iteratively[9] and a satisfactory convergence was achieved in all cases (the residual sum of squares, χ^2, below 0.007 and the standard error of estimate, χ^2/n, below 0.003). Departure from the HF/6-31G*, traditionally employed in RESP procedures, was both possible and necessary in order to obtain physically realistic and consistent results, and was paralleled in other similar improvements of RESP strategies (*e.g.*, within AMBER[20,21] or CHARMM[22] schemes) in the last decade.

Force field calculations and the analysis of normal vibrations were performed with the consistent force field (CFF) program for transition metal complexes and VOFF.[10] Additional force field parameters related to halogeno atoms were introduced as required and the values are given in Table S-I of the Supplementary material to this paper. The parameters were systematically derived to archive the closest possible match of calculated and experimental stretching vibrations. The details and the components of the force field (potential functions for bond stretching, angle bending, torsional, non-bonded and electrostatic contributions) were particularized earlier.[10]

All geometry optimizations (energy minimizations) were performed using the Newton–Raphson analytical second derivatives method. The root mean squared deviation of the gradient at convergence was always $<10^{-7}$ kcal mol^{-1} Å^{-1}, and the experimental crystal structures were reproduced with an average deviation of 0.005 Å in the bond lengths and 0.5° in the bond angles.

Supporting density functional theory (DFT) computations (*cf.* Subsection: Modeling the *trans*-influence of X) were performed using Gaussian09,[23] with either the hybrid B3LYP functional,[24,25] the standard LANL2DZ-ECP basis[26] for Cr and I atoms, and 6-31G(d,p) basis[19] for H, N, C, F, Br and Cl atoms, or the S-VWN functional[27-29] with 6-31+G(d) basis[16] for all atoms. Geometries of $[Cr(NH_3)_5X]^{2+}$ were optimized without any symmetry restrictions in the gas phase and the minima were confirmed by frequency calculations.

RESULTS AND DISCUSSION

Crystal structures

A search for structural data on $[Cr(NH_3)_{6-n}(X)_n]^{(3-n)+}$ (X = F, Cl, Br, I) revealed only a modest number of published reports on their crystal structures, which prevented any value for the equilibrium metal–ligand (M–L) bond lengths to be inferred from, *e.g.*, a statistical distribution of the observed M–L distances. Angular deformations of octahedra were sufficiently small and/or induced externally, enabling the adoption of the values $\pi/2$ and π for the reference valence angle parameters in the force field for metal centers throughout this study.

Furthermore, since the only one complete series of experimentally characterized structures, namely the one with $n = 1$, is available, the present discussion will be restricted to the $[Cr(NH_3)_5X]^{2+}$ (X = F, Cl, Br, I) series. CFF modeling of the other series (with two or more coordinated halogenido ligands) did not reveal any new features, but corroborated all the findings presented herein.

Standard bond lengths were sought in the 1989 compilation[30] of the data from the Cambridge structural database (CSD) as apparently no updates were published in the meantime. The values relevant for this work (Table I; part A)

show that the number of Cr–X bonds characterized experimentally are mainly insufficient. However, taking into consideration all the crystal structures published after the issuing of the 1989 tables,[30] it was possible to generate more reliable statistic (see Table I; part B). For 124 Cr–F bonds, there were 12 outliers from two structures (occurring at > 4σ), and another two were eliminated on chemical criteria. For 120 Cr–Cl bonds, only five distances were eliminated on chemical criteria. In both cases, the remaining values represented acceptably unimodal distributions. For 30 Cr–Br bonds, the distribution of values appeared to be bimodal, but an elimination of 6 values from a total of two structures did decrease the standard deviation appreciably, so a further reduction was not considered necessary for the present purpose. The number of Cr–I bonds was still too low, but the values were grouped in two narrow intervals, one of which (with the longer bond lengths) was eliminated on chemical grounds. Distribution of all Cr–X bonds is presented as bar diagrams in Fig. S-1 of the Supplementary material.

TABLE I Metal–ligand bond lengths (Å) from X-ray diffraction data for octahedral chromium(III) complexes given as: unweighted mean (d), median (m), standard deviation (σ), lower quartile (q_L), upper quartile (q_U), and the number of observations (n)

Bond	d	m	σ	q_L	q_U	n	Excl.[a]
A. Data from CSD (1989)[b]							
Cr–F	1.870	1.872	0.027	1.847	1.892	5	–
Cr–Cl	2.335	2.318	0.055	2.303	2.370	9	–
	2.309	2.316	0.018	2.229	2.319	7	–
Cr–Br	2.577	–	–	–	–	1	–
Cr–I	2.669	–	–		–	1	–
	2.781	–	–	–	–	1	–
B. Data from CSD (2013)[c]							
Cr–F	1.881	1.882	0.024	1.812	1.960	112	12,(2)
	1.882	1.883	0.023	1.841	1.960	110	2, (2)
Cr–Cl	2.325	2.322	0.022	2.287	2.389	120	–
	2.323	2.322	0.018	2.287	2.372	115	5, (3)
Cr–Br	2.481	2.479	0.025	2.436	2.521	28	2, (1)
	2.471	2.470	0.018	2.436	2.502	24	4, (1)
Cr–I	2.687	2.662	0.044	2.648	2.750	7	–
	2.662	2.662	0.010	2.648	2.677	5	2, (1)

[a]Values shown: excluded Cr–X bonds, (excluded structures); [b]quoted from Orpen *et al.*;[30] [c]Cambridge structural database 5.35 (Nov. 2013. + updates)

In keeping with the common practice of CFF,[31,32] the reference bond-length parameters in the bond stretching potential functions were assigned values which reproduce, after geometry optimization, the mean values of experimental bond lengths – in this case from X-ray diffraction studies, corresponding to r_α values in the Kuchitsu notation.[33] Taking into consideration previous bond stretching

parameters from VOFF,[10] and the mean values of r_α as given in Table I (part B), the following values for reference bond-length parameters were assigned: Cr–F 1.883 Å, Cr–Cl 2.300 Å, Cr–Br 2.400 Å, Cr–I 2.620 Å, Cr–N 2.052 Å, and N–H 0.950 Å. In furtherance of the discussion on the *trans*-influence (see Subsection: Modeling the *trans*-influence of X), it should be emphasized that the reference bond-length parameters, as well as bond-stiffness constants, were identical for all Cr–N bonds.

The coordination polyhedra of all structures were assumed to possess the highest possible symmetry (O_h for CrN_6, and C_{4v} for CrN_5X), due to the absence of any experimental data that might indicate otherwise.

Atomic charges

The choice of the RESP program and details about its application in the derivation of partial atomic charges from the QM molecular electrostatic potential for octahedral complexes of chromium(III) was already justified.[1] In particular, the selection of the restraint weight,[9] needed for the treatment of atoms buried inside larger structures, was explored and it was found that the central metal atom does not quite behave as buried, enabling the use of the single-stage derivation mode (*viz*. the "ESP" option) of the RESP program.

The results for $[Cr(NH_3)_5X]^{2+}$ complexes (X = F, Cl, Br, I) are summarized in Table II. General trends reported previously[1] for $[Cr(NH_3)_{6-n}(Cl)_n]^{(3-n)+}$ (n = = 0,1,2,3) complexes are perceived here as expected: for example, a decrease of q(X) is followed by a small decrease of q(M) and a marked decrease of $q(N^{ax})$. Furthermore, a comparison of the partial atomic charges with other previously published data (Table S-II of the Supplementary material) shows that trends (*e.g.*, decrease of q(M) upon substitution by X, or decrease of q(X) in the halogen series from F to I) are in full agreement, notwithstanding the differences in the reported values, even in cases when they originate from the same laboratory (Table S-II of the Supplementary material).

TABLE II. RESP atomic charges (esu) for complexes $[Cr(NH_3)_5X]^{2+}$ (X = F, Cl, Br, I), calculated with three different basis sets for the metal atom; q = RESP atomic charges, eq = equatorial or *cis*(X,NH_3), ax = axial or *trans*(X,NH_3), TZV = triple-ζ valence basis, SBKJC = effective core potential valence basis, cc-pVTZ = Dunning-type correlation consistent basis set with polarization function

Basis for Cr(III)	q(Cr)	q(X)	q(N)	q(H)
		$[Cr(NH_3)_5(X)]^{2+}$ (X = F)		
TZV	1.6409	–0.5867	-1.0317^{eq}	0.4083^{eq}
			-1.0000^{ax}	0.3898^{ax}
SBKJC	1.6668	–0.5896	-1.0366^{eq}	0.4083^{eq}
			-1.0530^{ax}	0.4032^{ax}

TABLE II. Continued

Basis for Cr(III)	*q*(Cr)	*q*(X)	*q*(N)	*q*(H)
		$[Cr(NH_3)_5(X)]^{2+}$ (X = F)		
cc-pVTZ	1.6340	–0.5941	-1.0128^{eq}	0.4046^{eq}
			-1.0718^{ax}	0.4091^{ax}
		$[Cr(NH_3)_5(X)]^{2+}$ (X = Cl)		
TZV	1.1800	–0.4400	-0.9087^{eq}	0.3819^{eq}
			-0.5672^{ax}	0.2929^{ax}
SBKJC	1.1804	–0.4386	-0.9134^{eq}	0.3834^{eq}
			-0.5701^{ax}	0.2936^{ax}
cc-pVTZ	1.1333	–0.4364	-0.8826^{eq}	0.3770^{eq}
			-0.5899^{ax}	0.2999^{ax}
		$[Cr(NH_3)_5(X)]^{2+}$ (X = Br)		
TZV	0.8217	–0.3420	-0.7925^{eq}	0.3558^{eq}
			-0.1928^{ax}	0.2044^{ax}
SBKJC	0.8032	–0.3350	-0.7917^{eq}	0.3561^{eq}
			-0.1791^{ax}	0.2013^{ax}
cc-pVTZ	0.7739	–0.3394	-0.7682^{eq}	0.351^{eq}
			-0.2037^{ax}	0.2086^{ax}
		$[Cr(NH_3)_5(X)]^{2+}$ (X = I)		
TZV	0.5399	–0.2243	-0.6171^{eq}	0.3084^{eq}
			-0.1677^{ax}	0.2066^{ax}
SBKJC	0.5315	–0.2202	-0.6154^{eq}	0.3082^{eq}
			-0.1809^{ax}	0.2106^{ax}
cc-pVTZ	0.4843	–0.2193	-0.5904^{eq}	0.3034^{eq}
			-0.1827^{ax}	0.2127^{ax}

By default in the CFF program, the atom-centered effective point charges used to evaluate the Coulomb potential are chosen from the table of force field values based on the atom type. However, in this as well as in the previous study,[1] the RESP charges (as given in Table II) were directly assigned to individual atoms together with their atomic coordinates.

Vibrational spectra

For the 22-atom $[Cr(NH_3)_5(X)]^{2+}$ species there are 60 vibrational normal modes, which are classified into $12A_1 + 4A_2 + 7B_1 + 5B_2 + 16E$ in the point group C_{4v} defined by the non-hydrogen (skeletal) atoms of $[Cr(NH_3)_5X]^{2+}$. Symmetry coordinates can be described in terms of the following redundant set of 68 internal coordinates: ΔR_i (ν(Cr–L), i = 1,6; L = N,X), Δr^k_i (ν(N–H), i = 1,5; k = 1,3), $\Delta\alpha_{ij}$ (δ(L–Cr–L), i = 1,5; j = i+1,6; L = N,X), $\Delta\beta^k_i$ (δ(Cr–N–H), i = 1,5; k = 1,3), $\Delta\gamma^k_i$ (δ(H–N–H), i = 1,5; k = 1,3), and five torsional angles $\Delta\tau_i$ (i = 1,5) defining internal rotation of the ammine ligands about the Cr–N bonds. Eighteen internal coordinates (R,α) define 11 skeletal modes ($4A_1 + 2B_1 + B_2 + 4E$; total degeneracy = 15) of the CrN_5X moiety. The remaining 45 internal coordinates (r,

β, γ) describe the ammine ν(N–H), δ(H–N–H), and δ(Cr–N–H) vibrations. One of the possible ways to label the latter modes in the C_{4v} point group is based on two assumptions: the first, to impose the invariance of the hydrogens on the axial ammine ligand to all symmetry operations in C_{4v}, and the second, to apply the projection operator method onto the basis consisting of orthogonal linear combinations of internals derived from the treatment of ammonia molecule in the C_{3v} point group[34] instead of the internals (N–H, H–N–H, and Cr–N–H) themselves. In this way, it was possible to obtain the following five sets of symmetry coordinates for $[Cr(NH_3)_5X]^{2+}$ in C_{4v} (with the total degeneracies shown in parentheses): symmetrical N–H stretching and symmetrical H–N–H bending ("umbrella") modes (each spanning $2A_1 + B_1 + E$; (5)), antisymmetrical N–H stretching, antisymmetrical H–N–H bending, and NH_3 rocking (sometimes referred to as libration) modes (each spanning $A_1 + A_2 + B_1 + B_2 + 3E$; (10)) totaling to 40 optical modes. The inventory is completed with five torsional modes (transforming as $A_1 + A_2 + B_2 + E$; (5)) derived in the same way. The skeletal modes of CrN_5X and a possible distribution of all normal modes for $[Cr(NH_3)_5X]^{2+}$ in the point group C_{4v} are given in Figure S-2 and Table S-III of the Supplementary material, respectively.

Since the assignments over the entire spectral range are in general accord with the published analyses of vibrational spectra of $[Cr(NH_3)_5X]^{2+}$ species (as shown in Table III), the focus here will be on the shifts of low frequency intramolecular modes in the series F, Cl, Br, I, which was the key question addressed in this paper. The most prominent tendency was an expected decrease in the Cr–X stretching frequency with increasing mass of X (see Table IV), which is in full correspondence with the experimental data. Although a further refinement of the VOFF was possible, it was not considered opportune in this case for at least two reasons. First, in line with the consistent force field (CFF) paradigm,[31,32,37,38] the force field parameters are not equatable to the derived vibrational force constants; hence, their optimization is meaningful only with respect to a wider range of observables for a family of related chemical species. In addition, crystal lattice effects (hydrogen bonds, counter-ion interactions, *etc.*), which are observed in solid-state IR spectra, were not explicitly treated in the present modeling study. Therefore, they are included in the force field parameters in an indirect way, thus obviating the need for additional parameter optimization as long as the differences among experimental values (of vibrational frequencies) are of the same order of magnitude as the differences between experimental and calculated values. Supporting DFT calculations (Subsection: Modeling the *trans*--influence of X) yielded vibrational frequencies (Table S-IV of the Supplementary material) that are in general agreement with those shown in Tables III and IV, except for the $\nu_a(NH_3)$, which were placed slightly above the experimental values.

TABLE III. Comparison of experimental vibrational frequencies (cm^{-1}) for $[Cr(NH_3)_5X]^{2+}$ (X = F, Cl, Br, I) to the corresponding values calculated using RESP charges[1] and VOFF parameters[10]; key: ν = stretching, δ = bending (deformation), ρ = rocking, s = symmetric, a = antisymmetric

Principal assignment	Sym.	X = F		X = Cl		X = Br		X = I	
		Exp.[a]	Calc.	Exp.[a]	Calc.	Exp.[a]	Calc.	Exp.[a]	Calc.
$\nu_a(NH_3)$	B_1	–	3277	3277[b]	3275	–	3277	–	3277
		–	3270	–	3271	–	3270	–	3271
$\nu_s(NH_3)$	E	–	3162	–	3160	–	3161	–	3161
		–	3160	–		–		–	–
δ_a(H–N–H)	B_1	–	1627	–	1618	–	1613	–	1611
		–	1589	1600[b]	1583	–	1587	–	1586
δ_s(H–N–H)	E	–	1337	–	1303	–	1323	–	1318
		–	1323	1289[b]	1293	–	1314	–	1309
$\rho_a(NH_3)$	B_1	–	748	774[b]	758	–	765	–	759
		–	710	–	702	–	711	–	700
ν(Cr–X)	A_1	540[c]	534	302[d] 306[e]	305	204[d] 206[f]	187	184[d] 195[g]	177
ν(Cr–N)	E	470[c]	474	472[d,e]	469	474[d] 470[f]	470	463[d] 467[g]	469
	A_1	462[c]	439	460[d] 462[e]	448	463[d] 453[f]	449	457[d] 454[g]	450
	A_1	428[c]	387	433[d] 434[e]	407	438[f]	403	440[d] 442[g]	402
	B_1	399[c]	380	406	371	425[d] 427[f]	372	415[d] 417[g]	370
δ(N–Cr–N)	B_2	–	313	–	296	–	293	–	288
	E	275[c]	274	285[d]	281	287[d] 274[f]	271	260[d] 261[g]	269
	E	260[c]	257	255[d,e]	266	265[d] 259[f]	263	250[d] 248[g]	257
	E	238[c]	235	–	241	250[d] 248[f]	244	240[d] 238[g]	244
	B_1	–	190	200[d]	200	–	196	–	196
δ(N–Cr–X)	E	214[c]	217	–	–	–	–	–	–
	E	184[c]	181	170[d,e]	167	172[d] 171[f]	171	172[d] 168[g]	163
	E	–	–	124[d] 126[e]	127	122[d,f]	128	116[d] 112[g]	111
δ(N–Cr–N)	E	114[c]	120	114[d,e]	103	110 114[f]	117	–	–

[a]From Schmidtke and Rosner,[35] unless otherwise noted; [b]from:Tanaka *et al.*,[36] counter-ions: c = perchlorate, d = nitrate, e = chloride, f = bromide, g = iodide

TABLE IV. Principal assignment of the skeletal vibrations (cm^{-1}) for $[Cr(NH_3)_5X]^{2+}$ (X = F, Cl, Br, I)

No.	$[Cr(NH_3)_5(F)]^{2+}$			$[Cr(NH_3)_5(Cl)]^{2+}$			$[Cr(NH_3)_5(Br)]^{2+}$			$[Cr(NH_3)_5(I)]^{2+}$		
	Sym	PED	ν	Sym	PED	ν	Sym	PED	ν	Sym	PED	ν
41	A_1	S_4+S_5	534	A_1	S_4+S_5	305	A_1	S_4+S_5	187	A_1	S_4+S_5	177
42	E	$2S_3$	474	E	$2S_3$	469	E	$2S_3$	470	E	$2S_3$	471
43	E	$2S_3$	473	E	$2S_3$	469	E	$2S_3$	470	E	$2S_3$	470
44	A_1	$S_1+S_4+S_5$	439	A_1	$S_4–S_5–S_8$	448	A_1	$S_4–S_5–S_8$	449	A_1	$S_4–S_5–S_8$	450
45	A_1	$S_1–S_4–S_5$	387	A_1	$S_1–S_4+S_5$	407	A_1	$S_1–S_4+S_5$	403	A_1	$S_1–S_4+S_5$	402
46	B_1	S_2	380	B_1	S_2	371	B_1	S_2	372	B_1	S_2	370
47	B_2	S_6	313	B_2	S_6	296	B_2	S_6	293	A_1	S_8	298
48	E	S_7+S_{11}	289	E	S_7+S_{11}	281	A_1	$S_8+S_4+S_5$	293	B_2	S_6	288
49	E	S_7+S_{11}	274	E	S_7+S_{11}	266	E	S_7+S_{11}	271	E	S_7+S_{11}	269
50	E	S_7+S_{11}	257	E	S_7+S_{11}	241	E	S_7+S_{11}	263	E	S_7+S_{11}	257

TABLE IV. Continued

No.	$[Cr(NH_3)_5(F)]^{2+}$			$[Cr(NH_3)_5(Cl)]^{2+}$			$[Cr(NH_3)_5(Br)]^{2+}$			$[Cr(NH_3)_5(I)]^{2+}$		
	Sym	PED	ν	Sym	PED	ν	Sym	PED	ν	Sym	PED	ν
51	A_1	S_8	248	A_1	S_8	226	E	$S_{10}+S_{11}$	244	E	$S_{10}+S_{11}$	244
52	E	S_7+S_{11}	235	E	S_7+S_{11}	221	E	$S_{10}+S_{11}$	225	E	$S_{10}+S_{11}$	222
53	E	$S_{11}+S_7$	223	B_1	S_9	200	B_1	S_9	196	B_1	S_9	196
54	E	$S_{10}+S_{11}$	217	E	$S_{10}+S_{11}$	192	E	$S_{10}+S_{11}$	189	E	$S_{10}+S_{11}$	184
55	B_1	S_9	190	E	$S_{10}+S_{11}$	167	E	$S_{10}+S_{11}$	171	E	$S_{10}+S_{11}$	163
56	E	$S_{10}+S_{11}$	181	E	$S_{10}+S_{11}$	152	B_2	S_6	153	E	$S_{10}+S_{11}$	149
57	B_2	S_6	158	E	$S_{10}+S_{11}$	143	B_2	S_6	137	B_2	S_6	135
58	E	$S_{10}+S_{11}$	136	E	$S_{10}+S_{11}$	127	E	$S_{10}+S_{11}$	128	E	$S_{10}+S_{11}$	123
59	E	$S_{10}+S_{11}$	120	E	$S_{10}+S_{11}$	103	E	$S_{10}+S_{11}$	117	E	$S_{10}+S_{11}$	111
60	E	$S_{10}+S_{11}$	51	E	$S_{10}+S_{11}$	43	E	$S_{10}+S_{11}$	21	E	$S_{10}+S_{11}$	29

Variations in other skeletal modes (Cr–N stretching and angle bending on the central metal atom) were observed but the shifts were not distinctly regular along the series X = F, Cl, Br, and I (Table IV). A plausible explanation might lie in the fact that skeletal vibrations in $[Cr(NH_3)_5X]^{2+}$ are affected not only by the mass of the atom X, but also through at least two additional phenomena: a) by the capacity of X to produce a redistribution of electron density on ammine ligands through *d*-orbitals in a mechanism commonly referred to as *trans*-influence (Subsection: Modeling the *trans*-influence of X), and b) through intramolecular steric interactions that depend on the size of X.

Concerning the results listed in Table IV, it should be emphasized that C_{4v} symmetry labels, as well as the symmetry coordinates derived in the C_{4v} point group, were used to describe the low-frequency skeletal vibrations. As the present CFF calculations were performed on the complete 22-atom structure with full relaxation of all internal coordinates (*i.e.*, without any symmetry constraint), the present vibrational analysis pertains, in a strict sense, to $[Cr(NH_3)_5X]^{2+}$ structures of lower (or at most C_s) symmetry. Consequently, a significant coupling in potential energy distribution (PED) was to be expected, and for this reason, only a few principal PED[39] contributions were shown in Table IV for each calculated vibrational frequency. The most noteworthy is the strong coupling of ammine torsions with N–Cr–N bending modes of the CrN_5X skeleton (not shown in Table IV, but clearly evident in computer animations of molecular vibrations,[40,41] available from the authors upon request).

Modeling the trans-influence of X

It is noteworthy that the RESP program[4] used in this work offers a straightforward and very convenient way to control the averaging of atomic charges over stereochemically equivalent atoms. This feature helped in the demarcation of axial from equatorial ammine nitrogens, where the difference in q(N) resulted from the *trans*-influence of the X atom in $[Cr(NH_3)_5X]^{2+}$.

The *trans*-influence of X in $[Cr(NH_3)_5X]^{2+}$ is expected to manifest itself through elongation of the axial Cr–NH_3 bond, and concomitant lowering of its Cr–N stretching frequency. Both phenomena were noted in crystallographic and spectroscopic studies on octahedral complexes. The present CFF modeling of $[Cr(NH_3)_5X]^{2+}$ with VOFF established a lengthening of Cr–N bond *trans* to the halogenido ligand (*cf.* Fig. 1) for X = Cl, Br, and I on average by 0.011(5) Å in geometry optimization with RESP charges (obtained from ROHF computations using the cc-pVTZ basis set for Cr) while keeping all FF parameters and functions involving all N ligator the same. The *trans*-influence turned out to be unexpectedly weak in the case of $[Cr(NH_3)_5F]^{2+}$. These results are in qualitative agreement with the fact that the Cr–NH_3 bond *trans* to X is lengthened on average by 0.03 Å with respect to other Cr–NH_3 bonds in 57 crystal structures containing 161 Cr–NH_3 bonds retrieved from the most recent CSD,[42] as well as with the recent findings of Shibahara, Akashi *et al.* on pentaaminenitrosylchromium(III) complexes exhibiting a particularly strong *trans*-influence of the NO

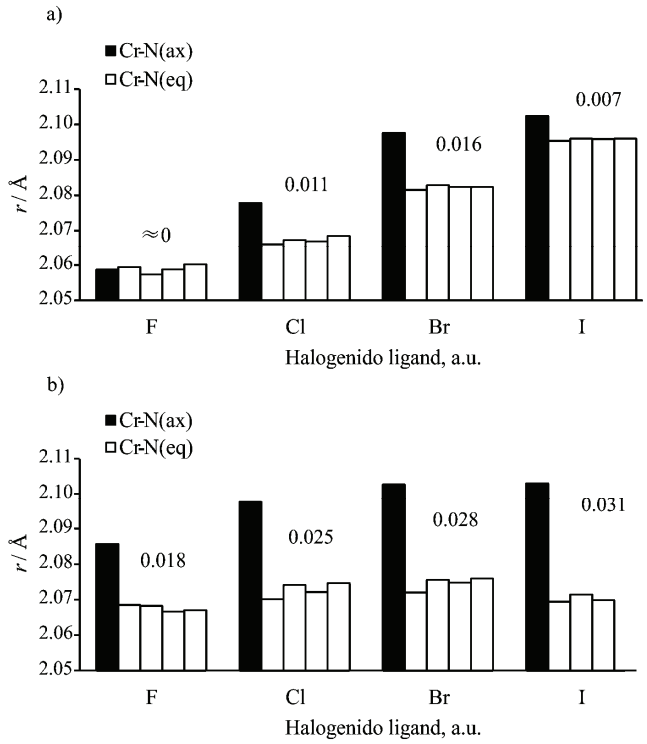


Fig. 1 Bond lengths (Å) for Cr–N^{ax} (*trans* to X) and Cr–N^{eq} in $[Cr(NH_3)_5X]^{2+}$ calculated with (a) CFF modeling and (b) DFT. DFT results were obtained with S-VWN/6-31+G(d). The values shown above the bars represent Δr_e, the difference between r_e^{trans} and the mean value of the four r_e^{cis} distances, which could be taken as a measure of the extent of *trans*-influence.

ligand.[43–46] Furthermore, in a broader sense, the present results are in certain conformity with the current views on ligand ordering based on their *trans*-influence in square-planar metal complexes.[47]

Differences in Cr–N bond lengths, as well as differences in their vibrations, are expected due to the presence of halogenido ligands and their *trans*-influence. Namely, as the axial Cr–N bonds in all molecules are shorter compared to equatorial ones, according to the general rule (shorter the bond - the larger the force constant – the greater the value of the wavenumber or vibration), the vibrations of Cr–N^{eq} bonds should have higher values compared to those of Cr–N^{ax} vibrations. If the individual contributions to the Cr–N vibrations are considered, the mentioned regularity is evident. The first calculated Cr–N vibrations (at the highest values of wavenumbers) originate exclusively from the equatorial Cr–N bonds, while the vibrations of the axial Cr–N bonds occur at lower values. For molecule $[Cr(NH_3)_5F]^{2+}$, the frequency at 474 cm^{-1} has a dominant PED contribution of Cr–N^{eq} bonds, while the frequency at 387 cm^{-1} has dominant PED contribution of the Cr–N^{ax} bond (Table III). The frequency at 469 cm^{-1} originates only from vibrations of the equatorial Cr–N bonds of $[Cr(NH_3)_5Cl]^{2+}$, but the vibrations of axial Cr–N occur at 448 cm^{-1}. The first calculated Cr–N vibration of the complex $[Cr(NH_3)_5Br]^{2+}$ at 470 cm^{-1} belongs to Cr–N^{eq} bonds, while Cr–N^{ax} bonds contribute to the second frequency at 449 cm^{-1}. Finally, for the molecule $[Cr(NH_3)_5I]^{2+}$, the calculated frequency at 469 cm^{-1} includes only vibrations of Cr–N^{eq} bonds, while frequency at 450 cm^{-1} has the highest contribution of vibrations of the Cr–N^{ax} bond.

Owing to the fact that structural data on $[Cr(NH_3)_5F]^{2+}$ are not available, and that most of the published crystal structures of $[Cr(NH_3)_5X]^{2+}$ with other halogenido ligands do not offer sufficiently conclusive evidence on the *trans*-influence of X, additional DFT calculations were performed for the whole series. First, it was observed that the recommended B3LYP functional[46] produces expected trends but yields excessively large metal–ligand distances. However, acceptable bond lengths were obtained with the S-VWN functional with basis sets as cited in the Section Methodology. The latter DFT results indeed revealed *trans*-influence for all four halogenido ligands, with reasonable Δr_e values (r_e^{trans} – $-r_e^{cis}$) for the Cr–N bonds, substantiating the trends resulting from the CFF modeling to the greatest extent possible in any comparison of QM and MM approaches.

The expected lowering of ν(Cr–N^{ax}) is more difficult to distinguish due to admixture of ν(Cr–X) and ν(Cr–N^{ax}) in both S_4 and S_5 symmetry coordinates (Table IV and Fig. S-2 of the Supplementary material) and additional coupling with ν_s(Cr–N^{eq}) (S_1), which was apparent in CFF modeling. It should be noted that this approach to model the *trans*-influence falls entirely within the force field (molecular mechanics) paradigm, *i.e.*, parametric representation of the potential

energy surface allowing for second-order stereochemical phenomena, and therefore does not propose an explanation for the origin of the *trans*-influence. As such, or with possible refinements or extensions (of which there are several), it may have some conceptual resemblance to the modeling of the *trans*-influence with VALBOND-TRANS force field.[8,48]

CONCLUSIONS

By extending the recently proposed scheme[1] for the derivation of partial atomic charges for chromium(III) complexes from QM computed MEP using the RESP procedure,[2] onto the mixed halogenido−ammine complexes with F, Cl, Br, and I ligands, and using such charges together with the vibrationally optimized force field (VOFF) for coordination compounds,[10] a reliable and physically realistic model was built, which was capable of reproducing experimental structures and vibrational frequencies in a consistent way throughout the series of $[Cr(NH_3)_{6-n}(X)_n]^{(3-n)+}$ species (n = 1; X = F, Cl, Br, I). The results for the other series (n > 1), not shown here due to the lack of experimental data, appear to be consistent with the ones presented.

In the course of this study it was necessary to define reference Cr–X bond distances, which were derived from a statistical analysis of the data from the CSD base that contained a 10-fold increase in measured Cr–X distances (a total of 254 relevant values) with respect to the 24 values used in the previous such compilation.[30]

The calculated vibrational frequencies with the present VOFF were in accordance with available experimental data over the whole frequency range (3500–100 cm^{-1}). In particular, the low-frequency skeletal vibrations were reassigned in the C_{4v} point group of the CrN_5X core, their shifts across the series (X = F, Cl, Br, I) were elucidated, and the presence of a strong coupling between some skeletal bending modes and torsional motion of ammine ligands was observed in the present zero torsional barrier approximation.

Finally, differences in the RESP derived partial atomic charges between the axial and equatorial ammine ligands in $[Cr(NH_3)_5X]^{2+}$, presumably due to second-order electronic effects, seem to offer a simple and straightforward way to incorporate *trans*-influence in force field modeling of coordination compounds, which manifested itself in the elongation of metal–ligand bonds in an expected way and to a plausible extent.

SUPPLEMENTARY MATERIAL

Additional VOFF parameters for $[Cr(NH_3)_5X]^{2+}$, the distribution of the Cr–X bond distances, a review of published atomic charges, the symmetry coordinates of $[Cr(NH_3)_5X]^{2+}$, the distribution of the symmetry vibrations and DFT calculated vibrational frequencies are available electronically from http://www.shd.org.rs/JSCS/, or from the corresponding author on request.

Acknowledgements. Financial support from the Serbian Ministry of Education, Science and Technological Development of the Republic of Serbia through Grant No. 172035 is gratefully acknowledged. We thank the Centre for Molecular and Biomolecular Informatics, Radboud University Nijmegen, for access to the Cambridge Crystal Structure Database (CSD).

ИЗВОД

РАЗВИЈАЊЕ НОВОГ СКУПА ПАРАМЕТАРА ПОЉА СИЛА ЗА ХАЛОГЕНИДО--АММИНСКЕ КОМПЛЕКСЕ ХРОМА(III): МОДЕЛИРАЊЕ *trans*-УТИЦАЈА ХАЛОГЕНИДО ЛИГАНАДА

ИВАНА ЂОРЂЕВИЋ[1], СОЊА ГРУБИШИЋ[1], МИЛОШ МИЛЧИЋ[2] и СВЕТОЗАР НИКЕТИЋ[1]

[1]Центар за хемију, ИХТМ, Универзитет у Београду, Његошева 12, 11001 Београд и [2]Хемијски факултет, Универзитет у Београду, Студентски трг 16, 11001 Београд

На серији октаедарских халогенидо-амминских комплекса хрома(III) приказан је нов приступ за моделирање *trans*-утицаја, користећи парцијална атомска наелектрисања изведене фитовањем из молекулског електростатичког потенцијала применом RESP методе. RESP наелектрисања заједно с приказаним вибрационо оптимизованим пољем сила објашњавају феномене другог реда, поспешују моделирање и асигнацију скелетних вибрација, и репродукују правилности у вибрационим помацима у низу халогенидо лиганада F, Cl, Br и I. Поред тога, дата је статистичка анализа за дужине Cr-халоген веза у кристалним структурама из Cambridge банке кристалографских података (CSD).

(Примљено 3. септембра, ревидирано 26. октобра, прихваћено 27. октобра 2014)

REFERENCES

1. I. S. Djordjević, S. R. Niketić, *Comput. Theor. Chem.* **1001** (2012) 20
2. F.-Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski, P. Cieplak, *Phys. Chem. Chem. Phys.* **12** (2010) 7821
3. F.-Y. Dupradeau, A. Pigache, T. Zaffran, P. Cieplak, *R.E.D. User's Manual and Tutorial*, version 2.0, http://q4md-forcefieldtools.org/RED/RED-II.pdf (accessed in 2005)
4. *RESP standalone*, version 2.3, http://q4md-forcefieldtools.org (accessed in 2011)
5. K. B. Yatsimirskii, *Pure Appl. Chem.* **49** (1977) 115
6. J. Gažo, R. Boča, E. Jóna, M. Kabešová, Ľ. Macášková, J. Šima, *Coord. Chem. Rev.* **43** (1982) 87
7. A. E. Anastasi, R. J. Deeth, *J. Chem. Theory Comput.* **5** (2009) 2339
8. I. Tubert-Brohman, M. Schmid, M. Meuwly, *J. Chem. Theory Comput.* **5** (2009) 530
9. C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **97** (1993) 10269
10. a) J.-H. Choi, S. R. Niketić, I. S. Djordjević, W. Clegg, R. W. Harrington, *J. Mol. Model.* **18** (2012) 2135; b) S. Grubisić, S. R. Niketić, D. D. Radanović, U. Rychlewska, B. Warzajtis, *Polyhedron* **24** (2005) 1701
11. M. Schmidt, K. K. Baldridge, J. A. Boatz, S. Elbert, M. Gordon, J. H. Jenson, S. Koeski, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem.* **14** (1993) 1347
12. A. K. Rappé, T. A. Smedley, W. A. Goddard III, *J. Phys. Chem.* **85** (1981) 2607
13. W. J. Stevens, M. Krauss, H. Basch, P. G. Jasien, *Can. J. Chem.* **70** (1992) 612
14. T. R. Cundari, W. J. Stevens, *J. Chem. Phys.* **98** (1993) 5555
15. N. B. Balabanov, K. A. Peterson, *J. Chem. Phys.* **123** (2005) 064107-1

16. M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, J. A. Pople, *J. Chem. Phys.* **77** (1982) 3654
17. L. A. Curtiss, M. P. McGrath, J. P. Blaudeau, N. E. Davis, R. C. Binning Jr., L. Radom, *J. Chem. Phys.* **103** (1995) 6104
18. M. N. Glukhovtsev, A. Pross, M. P. McGrath, L. Radom, *J. Chem. Phys.* **103** (1995) 1878
19. W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **56** (1972) 2257
20. Y. Duan, C. Wu, S. Chowdhur, M. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, *J. Comput. Chem.* **24** (2003) 1999
21. D. L. Mobley, É. Dumont, J. D. Chodera, K. A. Dill, *J. Phys. Chem B* **111** (2007) 2242
22. S. Patel, A. MacKerell Jr., *J. Comput. Chem.* **25** (2004) 1504
23. *Gaussian 09*, Revision D.01, Gaussian, Inc., Wallingford, CT, 2009
24. C. Lee, W. Yang, R. G. Parr, *Phys. Rev., B* **37** (1988) 785
25. A. D. Becke, *J. Chem. Phys.* **98** (1993) 1372
26. P. J. Hay, W. R. Wadt, *J. Chem. Phys.* **82** (1985) 270
27. S. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **58** (1980) 1200
28. J. C. Slater, *Phys. Rev.* **81** (1951) 358
29. P. A. M. Dirac, *Proc. R. Soc. London, Ser. A* **123** (1929) 714
30. A. G. Orpen, L. Brammer, F. H. Allen, O. Kennard, D. G. Watson, R. Taylor, *J. Chem. Soc., Dalton Trans.* **12** (1989) S1
31. S. Lifson, *in Dynamic Aspects of Conformational Changes in Biological Macromolecules*, C. Sadron, Ed., D. Reidel, Dordrecht, 1973, p. 421
32. S. Lifson, in *Supramolecular Structure and Function*, G. Pifat, J. H. Herak, Eds., Springer, New York, 1983, p. 1
33. K. Kuchitsu, in *Proceedings of Critical Evaluation of Chemical and Physical Structural Information*, Dartmouth College, Washington DC, 1974, p. 132
34. J. R. Ferraro, J. S. Ziomek, *Introductory Group Theory and Its Application to Molecular Structure*, Plenum Press, New York, 1975
35. H. H. Schmidtke, M. Rosner, *Inorg. Chem.* **28** (1989) 2510
36. N. Tanaka, M. Kamada, J. Fujita, E. Kyuno, *Bull. Chem. Soc. Jpn.* **37** (1964) 222
37. S. Lifson, A. Warshel, *J. Chem. Phys.* **49** (1968) 5116
38. A. Warshel, M. Levitt, S. Lifson, *J. Mol. Spectrosc.* **33** (1970) 84
39. S. Califano, *Vibrational States*, Wiley, London, 1976
40. *MOLEKEL 4.3*, Swiss Center for Scientific Computation, Manno, Switzerland, 2000
41. S. Portmann, H. P. Lüthi, *Chimia* **54** (2000) 766
42. S. R. Niketić, unpublished analysis of CSD data, 2013
43. H. Akashi, M. Mori, T. Shibahara, *Acta Crystallogr., E* **57** (2001) i75
44. H. Akashi, M. Mori, T. Shibahara, *Inorg. Chim. Acta* **331** (2002) 290
45. H. Akashi, T. Yamauchi, T. Shibahara, *Inorg. Chim. Acta* **357** (2004) 325
46. T. Shibahara, H. Akashi, M. Asano, K. Wakamatsu, K. Nishimoto, M. Mori, *Inorg. Chem. Commun.* **4** (2001) 413
47. F. A. Weinhold, C. R. Landis, *Valency and Bonding: A Natural Bond Orbital Donor–Acceptor Perspective*, Cambridge University Press, New York, 2005
48. J. Huang, M. Devereux, F. Hofmann, M. Meuwly, in *Computational Organometallic Chemistry*, O. Wiest, Y. Wu, Eds., Springer, Heidelberg, 2012, p. 19.

Journal of the Serbian Chemical Society
JSCS@tmf.bg.ac.rs • www.shd.org.rs/JSCS

J. Serb. Chem. Soc. 80 (3) S67–S73 (2015) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Derivation of a new set of force field parameters for ammine complexes of chromium(III) containing halogeno ligands: modeling of the *trans*-influence of halogenido ligands

IVANA DJORDJEVIĆ[1], SONJA GRUBIŠIĆ[1]*, MILOŠ MILČIĆ[2] and SVETOZAR NIKETIĆ[1]

[1]*Center for Chemistry, ICTM, University of Belgrade, Njegoševa 12, 11001 Belgrade, Serbia and* [2]*Faculty of Chemistry, University of Belgrade, Studentski trg 16, 11001 Belgrade, Serbia*

J. Serb. Chem. Soc. 80 (3) (2015) 329–342

TABLE S-I. Additional VOFF parameters for $[Cr(NH_3)_5X]^{2+}$ (X = F, Cl, Br, I) series; r_0 and θ_0 are the equilibrium bond length and bond angle, respectively; k_r and k_θ are the force constants for the bonds and bond angles, respectively (1 kcal = 4.184 J)

Bond stretching	k_r / kcal mol^{-1} Å^{-2}		r_0 /Å	
M–F	310		1.883	
M–Cl	160		2.300	
M–Br	170		2.400	
M–I	200		2.620	
Angle bending	k_θ / kcal mol^{-1} rad^{-2}		θ_0 / rad	
	cis	*trans*	*cis*	*trans*
N–M–N	30.00	30.00	1.570	3.142
N–M–F	15.00	15.00	1.570	3.142
N–M–Cl	20.00	20.00	1.570	3.142
N–M–Br	30.00	30.00	1.570	3.142
N–M–I	35.00	35.00	1.570	3.142
van der Waals	ε / kcal mol^{-1}		r^* / Å	
F···H	0.095		3.40	
F···N	0.100		3.70	
Cl···H	0.106		3.53	
Cl···N	0.115		3.85	
Br···H	0.122		3.68	
Br···N	0.133		4.00	
I···H	0.136		3.83	
I···N	0.150		4.10	

* Corresponding author. E-mail: grubisic@chem.bg.ac.rs

TABLE S-II. Comparison of the calculated partial atomic charges (esu); key: *q* = partial atomic charges, eq = equatorial or *cis*(X,NH_3), ax = axial or *trans*(X,NH_3), X=F, Cl, Br, I; Methods (as cited in the original publications): NPA = natural population analysis, MKS = Merz–Kollman–Singh scheme, NLDA = nonlocal density approximation, MEP = molecular electrostatic potential

Complex	*q*(M)	*q*(N)	*q*(H)	*q*(F)	*q*(Cl)	*q*(Br)	*q*(I)	Meth.	Ref.
Published data									
$[Co(NH_3)_6]^{3+}$	1.67	–1.19	0.47					NPA	1
$[Co(NH_3)_6]^{3+}$	1.55	–1.05	0.43					MKS	1
$[Co(NH_3)_6]^{3+}$	2.0								2
$[Co(NH_3)_6]^{3+}$	2.71								3
$[Co(NH_3)_5F]^{2+}$	1.6								2
$[Co(NH_3)_5F]^{2+}$	1.76								3
$[Cr(H_2O)_6]^{3+}$	1.403							NLDA	4
$[Cr(H_2O)_6]^{3+}$	0.722							MEP	4
$[Cr(H_2O)_6]^{3+}$	2.12								5
$[Cr(H_2O)_6]^{3+}$	2.2								2
$[CrF_6]^{3-}$	1.62							NPA	6
$[ZrF_6]^{2-}$	2.634			–0.77				MEP	7
$[ZrCl_6]^{2-}$	1.509				–0.58			MEP	7
$[ZrBr_6]^{2-}$	0.938					–0.49		MEP	7
$[ZrI_6]^{2-}$	0.440						–0.41	MEP	7
This work									
$[Cr(NH_3)_5F]^{2+}$	1.63	-1.01^{eq}	0.41^{eq}	–0.60					
$[Cr(NH_3)_5Cl]^{2+}$	1.13	-0.88^{eq}	0.38^{eq}		–0.44				
$[Cr(NH_3)_5Br]^{2+}$	0.77	-0.77^{eq}	0.35^{eq}			–0.34			
$[Cr(NH_3)_5I]^{2+}$	0.48	-0.59^{eq}	0.30^{eq}				–0.22		

CSD REFCODES

For structures with Cr–F bonds

ADOSOC, BAPPOX, BAVGOU, BAVGUA10, BIVVUX, BORWOU, BORWUA, BORWUA01, BUTDAV, DAFHOH, DEDCAQ, DIYFEW, DORZEP, DUCQIC, DULQAD, FEZXUD, FOJBAH, GAWBUC, IZOVEZ, KOMREJ, LEYDOJ, LEYDUP, LEYSEN, MPZFCR, OCOLIB, POFGEX, PUXPUT, QEXSES, SIXWOL, SUKHAH, TAZFCR, TETVUK, VAFDEN, VAFDIR, VIWMEU, VOLREU, YASHAB, YASHOP.

For structures with Cr–Cl bonds

ABILUT, AMZCCR, BIJZOK, BIWPIH, CEDCRC, COKJAN, CUKPON, CUSBUN, DOHBEH, EGULOI, EKOCEM, EMIWIG, EMUNIJ, FAMVUK10, FUHCIU, GADTOU, JAQPEW, JATTED, JATVAB, JOFYIM, JOFYOS, JOMNAA, JOMNAA01, JOMNEE, KUGMAA, KUGMEE, LIJRAX, MATHIZ, MATHOF, MATHUL, MOLPUZ, PADSIW, PAZDUQ, PUWFAO, PYMECR, QISXIA, QOWZOS, QOWZUY, QOXBAH, RAKGIT, RUNRIB, RUNRIB01,

TEBDIO, TEZSIA, TEZSOG, TEZSUM, TEZTAT, TEZTEX, ULAYUC, UWEKEN, WATREO, WOGRUG, YISLIV, YISLOB, ZALZER, ULAYUC01, WEBCAJ, ACUCAE.

For structures with Cr–Br bonds

BIWBEO, BOGDIL, CAGCIW, ESIMUO, IHEJOU, KUGMEE, MAMLER, PUDNUY, PUDPAG, PURCIO, UKEDEU, UKEDIY, VAGBIQ, VAQFAU, VAVNOV, VEFVEH, YIBWOV, TEDYIM.

For structures with Cr–I bonds

BOGDOR, CPMINC, DAGNUW, JEKKEP, NATZIS, NMCPNC, ODAXIA, OFAKAI, PPNCCO, RAJXUV, SESYAQ, UGUBAZ, UKEDOE, UKEDUK, UKEFAS, UKEFOG, UKEFUM, VAZWAU, YEKLUW.

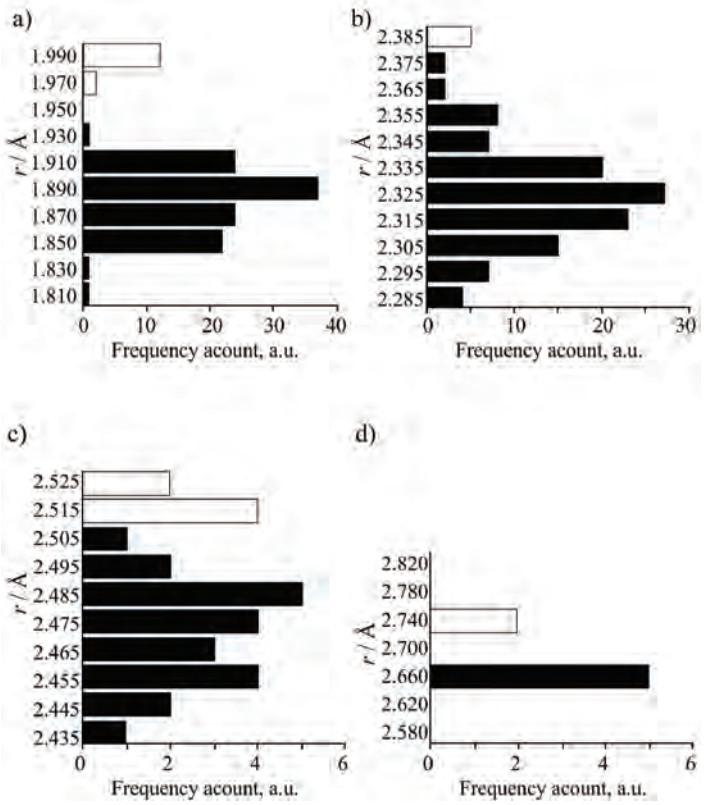


Fig. S-1. Distribution of chromium(III)–halogen distances in the crystal structures retrieved from CSD (2013.): a) for Cr–F bond lengths (Å); b) for Cr–Cl bond lengths (Å); c) for Cr–Br bond lengths (Å) and d) for Cr–I bond length (Å). The data used in the derivation of the statistics are shown in black; data eliminated for any reason (see text) are shown in white. Note the difference in the scale for the *x*-axis between the upper and lower pair of charts.

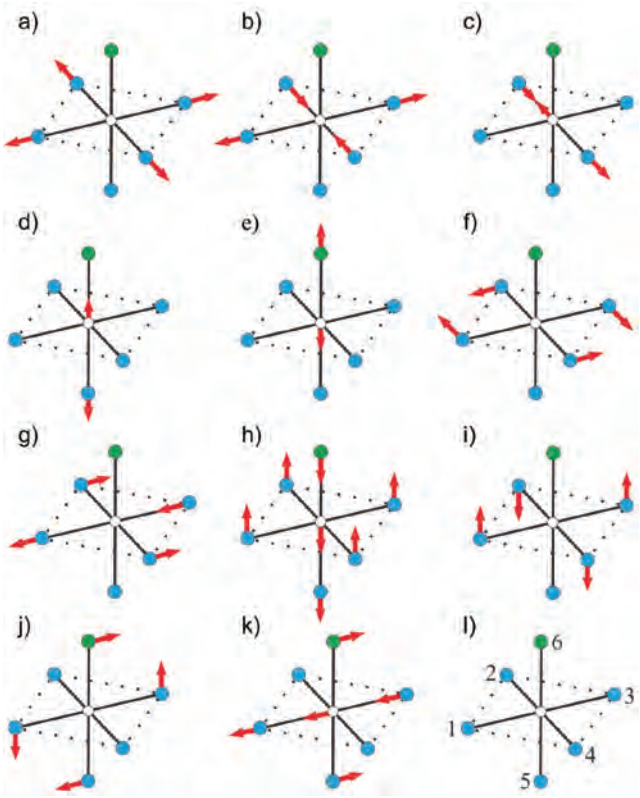


Fig. S-2. Symmetry coordinates for the normal vibrations of the C_{4v}-[CrN_5X] skeleton. Definition of internal coordinates for skeletal vibrations: Metal–ligand bonds: R_i (i = 1,6). Octahedral angles: α_{ij} (i, j = ligand atoms). a) $S_1(A_1)$ $\phi_1(A_1) = (1/2)\Delta(R_1 + R_2 + R_3 + R_4)$, b) $S_2(B_1)$ $\phi_2(B_1) = (1/2)\Delta(R_1 - R_2 + R_3 - R_4)$, c) $S_3(E)$ $\phi_3(E^a) = (1/2)^{1/2}\Delta(R_2 - R_4)$ $\phi_3(E^b) = (1/2)^{1/2}\Delta(R_1 - R_3)$, d) $S_4(A_1)$ $\phi_4(A_1) = \Delta R_5$, e) $S_5(A_1)$ $\phi_5(A_1) = \Delta R_6$, f) $S_6(B_2)$ $\phi_6(B_2) = (1/2)\Delta(\alpha_{12} - \alpha_{23} + \alpha_{34} - \alpha_{41})$, g) $S_7(E)$ $\phi_7(E^a) = (1/2)^{1/2}\Delta(\alpha_{12} - \alpha_{34})$ $\phi_7(E^b) = (1/2)^{1/2}\Delta(\alpha_{23} - \alpha_{41})$, h) $S_8(A_1)$ $\phi_8(A_1) = (1/8)^{1/2}\Delta(\alpha_{15} - \alpha_{16} + \alpha_{25} - \alpha_{26} + \alpha_{35} - \alpha_{36} + \alpha_{45} - \alpha_{46})$, i) $S_9(B_1)$ $\phi_9(B_1) = (1/8)^{1/2}\Delta(\alpha_{15} - \alpha_{16} - \alpha_{25} + \alpha_{26} + \alpha_{35} - \alpha_{36} + \alpha_{45} + \alpha_{46})$, j) $S_{10}(E)$ $\phi_{10}(E^a) = (1/2)\Delta(\alpha_{15} - \alpha_{16} - \alpha_{35} + \alpha_{36})$ $\phi_{10}(E^b) = (1/2)\Delta(\alpha_{25} - \alpha_{26} - \alpha_{45} + \alpha_{46})$, k) $S_{11}(E)$ $\phi_{11}(E^a) = (1/2)\Delta(\alpha_{15} + \alpha_{16} - \alpha_{35} + \alpha_{36})$ $\phi_{11}(E^b) = (1/2)\Delta(\alpha_{25} + \alpha_{26} - \alpha_{45} + \alpha_{46})$; l) atom numbering.

TABLE S-III. Distribution of symmetry vibrations of $[M(NH_3)_5X]^{2+}$ among the irreducible representations of the point group C_{4v} ($\Gamma^{VIB} = 12A_1 + 4A_2 + 7B_1 + 5B_2 + 16E$); ν = stretching, δ = bending (deformation), ρ = rocking, s = symmetric, a = antisymmetric

Motion	A_1	A_2	B_1	B_2	E
Skeletal[a] stretch	S_1: $\nu_s(MN)^{eq}$ S_4: $\nu(MN)^{ax}$ S_5: $\nu(MX)$		S_2: $\nu_{as}(MN)^{eq}$		S_3: $\nu_a(MN)^{eq}$
Skeletal[a] bend	S_8: $\delta(MN_5X)$		S_9: $\delta(NMN)^{eq}$	S_6: $\delta(NMN)^{eq}$	S_7: $\delta(NMN)^{eq}$ S_{10}: $\delta(MN_5X)$ S_{11}: $\delta(MN_5X)$
Sym. stretch	$\nu_s(NH)^{eq}$ $\nu_s(NH)^{ax}$		$\nu_s(NH)^{eq}$		$\nu_s(NH)^{eq}$
As. stretch	$\nu_{as}(NH)^{eq}$	$\nu_{as}(NH)^{eq}$	$\nu_{as}(NH)^{eq}$	$\nu_{as}(NH)^{eq}$	$\nu_{as}(NH)^{eq}$ $\nu_{as}(NH)^{eq}$ $\nu_{as}(NH)^{ax}$
Bend	$\delta_{as}(HNH)^{eq}$	$\delta_{as}(HNH)^{eq}$	$\delta_{as}(HNH)^{eq}$	$\delta_{as}(HNH)^{eq}$	$\delta_{as}(HNH)^{eq}$ $\delta_{as}(HNH)^{eq}$ $\delta_{as}(HNH)^{ax}$
Rock	$\rho_{as}(NH_3)^{eq}$	$\rho_{as}(NH_3)^{eq}$	$\rho_{as}(NH_3)^{eq}$	$\rho_{as}(NH_3)^{eq}$	$\rho_{as}(NH_3)^{eq}$ $\rho_{as}(NH_3)^{eq}$ $\rho_{as}(NH_3)^{ax}$
UMB[b]	$\delta_s(NH_3)^{eq}$ $\delta_s(NH_3)^{ax}$		$\delta_s(NH_3)^{eq}$		$\delta_s(NH_3)^{eq}$
Torsion	$\tau_{as}(NH_3)^{eq}$	$\tau_{as}(NH_3)^{ax}$		$\tau_{as}(NH_3)^{eq}$	$\tau_{as}(NH_3)^{eq}$

[a]S_i = Skeletal symmetry modes depicted in Fig. S-2; [b]UMB = Symmetrical NH_3 bending ≡ umbrella mode

TABLE S-IV. Calculated[a] vibrational frequencies (cm^{-1}) for $[Cr(NH_3)_5X]^{2+}$; ν = stretching, δ = bending (deformation), ρ = rocking, s = symmetric, a = antisymmetric

No.	Principal assignment	X = F			X = Cl			X = Br			X = I		
		Exp.	CFF	DFT	Exp.	CFF	DFT	Exp.	CFF	DFT	Exp.	CFF	DFT
1	$\nu_a(NH_3)$		3277	3388	3277	3275	3383		3277	3382		3277	3395
2	$\nu_a(NH_3)$		3273	3387		3272	3381		3272	3381		3272	3395
3	$\nu_a(NH_3)$		3271	3387		3272	3380		3271	3380		3271	3394
4	$\nu_a(NH_3)$		3271	3385		3271	3380		3271	3379		3271	3391
5	$\nu_a(NH_3)$		3270	3380		3270	3379		3270	3374		3270	3387
6	$\nu_a(NH_3)$		3270	3379		3270	3378		3270	3374		3270	3385
7	$\nu_a(NH_3)$		3269	3367		3270	3367		3270	3365		3270	3373
8	$\nu_a(NH_3)$		3269	3365		3270	3362		3270	3364		3270	3363
9	$\nu_a(NH_3)$		3268	3364		3268	3359		3269	3357		3269	3361
10	$\nu_a(NH_3)$		3267	3363		3268	3357		3268	3355		3269	3358
11	$\nu_s(NH_3)$		3166	3286		3164	3282		3165	3284		3165	3300
12	$\nu_s(NH_3)$		3162	3269		3161	3276		3162	3271		3162	3284
13	$\nu_s(NH_3)$		3160	3266		3160	3271		3161	3270		3161	3277
14	$\nu_s(NH_3)$		3160	3265		3160	3263		3161	3262		3161	3267
15	$\nu_s(NH_3)$		3160	3264		3160	3261		3160	3259		3160	3259
16	δ_a(H–N–H)		1634	1641		1631	1635		1624	1632		1620	1599
17	δ_a(H–N–H)		1627	1627		1618	1625		1613	1629		1611	1593

TABLE S-IV. Continued

No.	Principal assignment	X = F			X = Cl			X = Br			X = I		
		Exp.	CFF	DFT	Exp.	CFF	DFT	Exp.	CFF	DFT	Exp.	CFF	DFT
18	δ_a(H–N–H)		1616	1618		1612	1620		1607	1623		1604	1586
19	δ_a(H–N–H)		1616	1616		1612	1617		1606	1618		1604	1581
20	δ_a(H–N–H)		1590	1614		1589	1608		1587	1607		1586	1574
21	δ_a(H–N–H)		1589	1610		1588	1604		1586	1606		1585	1569
22	δ_a(H–N–H)		1586	1599	1600	1583	1598		1581	1600		1581	1561
23	δ_a(H–N–H)		1583	1599		1580	1589		1580	1585		1579	1555
24	δ_a(H–N–H)		1580	1596		1579	1579		1578	1584		1578	1545
25	δ_a(H–N–H)		1573	1576		1575	1573		1574	1580		1574	1538
26	δ_s(H–N–H)		1366	1354		1336	1356		1372	1356		1338	1341
27	δ_s(H–N–H)		1337	1327		1313	1332		1323	1335		1318	1314
28	δ_s(H–N–H)		1323	1317		1303	1323		1314	1321		1309	1298
29	δ_s(H–N–H)		1312	1315	1289	1293	1314		1300	1306		1297	1292
30	δ_s(H–N–H)		1305	1300		1287	1318		1290	1303		1289	1284
31	$\rho(NH_3)$		804	768	774	758	732		765	727		759	714
32	$\rho(NH_3)$		775	703		742	702		739	698		726	700
33	$\rho(NH_3)$		748	699		712	691		711	696		700	697
34	$\rho(NH_3)$		744	692		702	687		703	692		692	688
35	$\rho(NH_3)$		731	657		694	659		696	663		690	654
36	$\rho(NH_3)$		710	649		686	655		686	651		675	640
37	$\rho(NH_3)$		697	638		679	644		667	637		661	617
38	$\rho(NH_3)$		685	598		662	583		665	573		658	566
39	$\rho(NH_3)$		641	594		628	570		622	569		619	556
40	$\rho(NH_3)$		622	587		613	560		617	558		615	545
41	ν_a(Cr–X)	540	534	554	306	305	344	206	187	179	195	177	176
42	ν_a(Cr–N)	470	474	435		469	433		470	430		471	438
43	ν_a(Cr–N)	462	473	434	472	469	427	470	470	427	467	469	436
44	ν_s(Cr–N)	428	439	411	460	448	425	453	449	413	454	450	415
45	ν_s(Cr–N)	339	387	402	433	407	414	438	403	403	442	402	393
46	ν_a(Cr–N)		380	389	406	371	385	427	372	385	417	370	388
47	$\delta(CrN_5X)$		313	249		296	224		293	276		298	220
48	$\delta(CrN_5X)$		289	236	285	281	218		293	236		288	213
49	$\delta(CrN_5X)$	275	274	224	255	266	215	274	271	216	261	269	179
50	$\delta(CrN_5X)$	260	257	211		241	203	259	263	213	250	257	194
51	$\delta(CrN_5X)$		248	196		226	201	248	244	209	240	244	185
52	$\delta(CrN_5X)$	238	235	188		221	188		225	196		222	183
53	$\delta(CrN_5X)$		223	158	200	200	164		196	188		196	162
54	$\delta(CrN_5X)$	214	217	150		192	153		189	185		184	138
55	$\delta(CrN_5X)$		190	146	170	167	144	171	171	175	168	163	124
56	$\delta(CrN_5X)$	184	181	137		152	136		153	158		149	122
57	$\delta(CrN_5X)$		158	131		143	130		137	137		135	114
58	$\delta(CrN_5X)$		136	118	126	127	81	122	128	128		123	94
59	$\delta(CrN_5X)$	114	120	86	114	103	73	114	117	119	112	111	80
60	$\delta(CrN_5X)$		51	57		43	31		21	56		29	44

[a]Scaling factor for DFT (0.9833)[8]

REFERENCES

1. W. R. Fawcett, M. Hromadova, G. A. Tsirlina, R. R. Nazmutdinov, *J. Electroanal. Chem.* **498** (2001) 93
2. W. R. Fawcett, G. J. Chavis, M. Hromadová, *Electrochim. Acta* **53** (2008) 6787
3. R. R. Nazmutdinov, M.-Y. Rusanova, D. Van der Porten, G. A. Tsirlina, W. R. Fawcett, *J. Phys. Chem. C* **113** (2009) 2881
4. A. Bleuzen, F. Foglia, E. Furet, L. Helm, A. E. Merbach, J. Weber, *J. Am. Chem. Soc.* **118** (1996) 12777
5. M.-Y. Rusanova, G. A. Tsirlina, R. R. Nazmutdinov, W. R. Fawcett, *J. Phys. Chem., A* **109** (2005) 1348
6. C. Adamo, V. Barone, *Theor. Chem. Acc.* **105** (2000) 169
7. J. Rak, M. Gutowski, P. Dokurno, H. V. Thanh, J. Blazejowski, *J. Chem. Phys.* **100** (1994) 5810
8. M. W. Wong, *Chem. Phys. Lett.* **256** (1996) 391.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 343–353 (2015)
JSCS–4720
UDC 547.466.1+539.16:547.466+532.74
Original scientific paper

A quantitative structure–activity relationship (QSAR) study of peptide drugs based on a new descriptor of amino acids

JIAN-BO TONG*, JIA CHANG, SHU-LING LIU and MIN BAI

Shaanxi University of Science and Technology, Xi'an 710021, China

(Received 4 June, accepted 7 July 2014)

Abstract: The quantitative structure–activity relationship (QSAR) approach is used for finding the relationships between molecular structures and the activity of peptide drugs. In this work, the stepwise multiple regression method was employed to select the optimal subset of descriptors that have significant contributions to the drug activity of 21 oxytocin analogues, 48 bitter tasting threshold activities and 58 angiotensin-converting enzyme inhibitors. A new set of descriptor, SVWGM, was used for the prediction of the activity of the peptide drugs and then were used to build a model by the partial least squares method, The stability and generalization ability of the model was strictly analyzed by both internal and external validations, by cross-validation and external validation correlations.

Keywords: amino acids; peptide; quantitative structure–activity relationships; partial least square; SVWGM.

INTRODUCTION

As one of the hottest topic, peptide drugs, which perform important roles in all living systems, have been widely debated.[1,2] Most of experimental methods on peptide drugs were inefficient and expensive and hence, computational methods, such as quantitative structure–activity relationships (QSAR), which involve not only the key idea of pharmacology, but also the foundation of drug design, were brought into the spotlight. In a QSAR study of peptides, the major information on structure and function for peptides is contained in their amino acid sequence and hence it is crucial to characterize the sequence structure. The pioneering work on amino acid descriptors was undertaken by Sneath *et al.*[3] who used semi-quantitative physicochemical data to derive descriptors for the 20 naturally occurring amino acids. Kidera *et al.*[4–6] collected 188 properties of the 20 natural amino acids and employed factor analysis to obtain 10 orthogonal factors that were the most important for determining the three-dimensional structure

* Corresponding author. E-mail: jianbotong@aliyun.com
doi: 10.2298/JSC140604069T

of proteins. Soon afterwards, Hellberg *et al.*[5,7–12] extracted three principal components (PCs), or *z*-scales, from 29 physicochemical properties of coded amino acids by PC analysis (PCA). In recent years, a new series of amino acid indices based on quantum topological molecular similarity (QTMS) descriptors[12,13] were proposed for peptide QSAR studies and good results were obtained. The study of the structure-based approaches for peptides by Hou *et al.*[14–17] achieved satisfactory results. Since then, many quantitative descriptors for amino acid were proposed and successfully used in practice.[18,19] They were then applied to peptide QSAR analyses and showed good structural characterization ability.

PRINCIPLE AND METHODS

Molecular structural characterization

Molecular structural characterization (MSC) focuses on transforming molecular structural characteristics into a group of numerical codes, dedicated to minimizing information loss during this process. In this study, 99 weighted holistic invariant molecular (WHIM) descriptors (obtained as statistical indices of the atoms projected onto the 3 principal components obtained from weighted covariance matrices of the atomic coordinates, the meanings of which are listed in Table S-I of the Supplementary material to this paper), 74 geometrical descriptors (obtained from molecular geometry, the meanings are listed in Supplementary Table S-II of the Supplementary material) and 160 3D-molecule representations of structures based on electron diffraction (3D-MORSE) descriptors (calculated by summing atom weights viewed by a different angular scattering function, the meanings of which are listed in Supplementary Table S-III of the Supplementary material) for each single amino acid were generated *via* Dragon 5.4 software.

The original variable matrices $X_{20\times99}$ (Table S-IV of the Supplementary material), $X_{20\times74}$ (Table S-V of the Supplementary material) and $X_{20\times160}$ (Table S-VI of the Supplementary material were submitted to PCA, autoscaling was always applied. The PCA scores (Table S--VII of the Supplementary material) were extracted for all those components that were identified to be necessary to explain a minimum of 85 % of the total variance, also, the corresponding eigenvalue > 1. A total of 18 PCs (Table S-VIII of the Supplementary material) were selected from the PCA for each amino acid, which were treated as a new vector descriptor – SVWGM (principal component scores vector of WHIM descriptors, Geometrical descriptors and 3D-MORSE descriptors) to characterize a peptide. PCA was implemented by MATLAB R2010a software.

Variable selection and modeling

The selection of molecular descriptors most relevant for acute drug activity is a key problem involved in QSAR, as is the choice of a proper technique for the construction of models. At present, the stepwise multiple regression (SMR) method is an interesting and widely used variable selection method.[20] SMR was implemented by SPSS software for Windows 16.0.

The partial least squares (PLS) method was employed for developing a reliable and predictive QSAR model and for an estimation of its stability, which was strictly analyzed by the cross-validation correlation coefficient (Q^2_{LOO}) and correlation coefficient (R^2_{cum}). The PLS analyses were implemented using SIMCA-P 12.0 software.

RESULTS AND DISCUSSION

QSAR model for oxytocin

Oxytocin (OT), a neurohypophyseal hormone, stimulates contraction of the uterine myometrium at parturition and contracts myoepithelial cells during lactation. Twenty-one oxytocin analogues[21] varying at positions 2, 3 and 8 are often utilized to test the effectiveness of diverse kinds of amino acid descriptors for its oxytocin-activity denoted by *OA* (Table I). For each peptide, its structure was first quantified by 3×18 SVWGM scales. In order to prove the validity and stability of the model, the whole data set was systematically divided into two subsets. The dataset (n = 21) was divided into a training set of 18 samples and a test set (highlighted with "*" in Table I) of 3 samples. The training sets were then used for developing QSAR models employing the PLS technique, while the corresponding test set was utilized to evaluate the predictive ability of the obtained model.[12] In the PLS regression analysis, the Q^2_{LOO} values (Table S-IX of the Supplementary material) changed with the SMR-introduced variables, achieving the maximum (0.864) at the fifth step, meanwhile, the value of the correlation coefficient (R^2_{cum}) reached up to 0.944. The optimum PLS model with 5 variables (v_{24}, v_{17}, v_{36}, v_{51} and v_{23}) was then obtained, wherein only two PLS components were sufficient to account for 97.2 % of the variances of the Y variables. A plot of the observed values for OT *vs*. the predicted ones (listed in Table I) is presented in Fig. 1, from which it could be seen that the points were minutely scattered about the through the origin-passing diagonal and the model possessed a strong fitting capability. The results obtained from the PLS models using the SVWGM descriptors are given in Table II, which were superior to those from the literature (Table II). A scoring distribution scatter of the 21 samples in the first and second PLS principal component spaces is shown in Fig. 2, from which it could be seen that most samples fell within the Hotelling T^2 confidence ellipse[22,23] with a 95 % confidence interval, with the exception of $7^{\#}$.

The distance to the PLS model in the X space is described in Fig. 3. It can be seen that the normalized distance to X for none of sample was beyond the critical value of 2.148 (significance level = 5 %).

TABLE I. Amino acid sequences of 21 OT and their observed and predicted activities; Obs – observed values, Pred – predicted values, * – test set

No.	Peptide	Obs	Pred	No.	Peptide	Obs	Pred	No.	Peptide	Obs	Pred
1	YYL	2.00	2.54	8	SIK	1.00	1.03	15	YFL	4.30	4.22
2	FFL	3.52	3.60	9	YWK	1.00	0.82	16	YII	5.46	5.12
3	YWL	1.60	1.90	10	FIK	3.00	3.52	17	YFK	3.70	3.13
4	YLL	4.65	5.27	11	YIR	4.88	5.11	18	YYK*	1.00	1.45
5	YVL	4.77	4.95	12	YFH*	3.18	3.09	19	FFK	2.48	2.51
6	YIV*	5.30	4.26	13	YIL	5.65	5.24	20	YIK	4.89	4.15
7	FYK	1.00	0.83	14	FIL	4.50	4.62	21	YFR	4.30	4.10

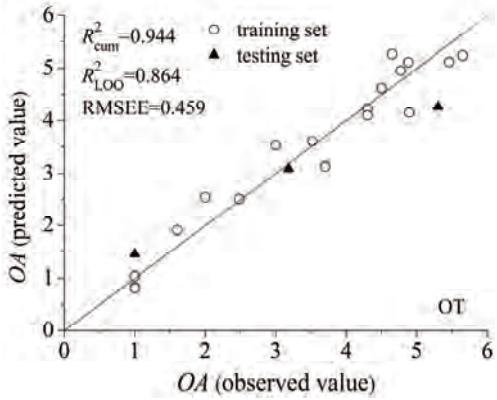


Fig. 1. Plot of predicted *vs.* observed values for 21 OT by the PLS model.

TABLE II. Comparison between the different QSAR models of OT; –: not determined

No.	Descriptors	Model	A[a]	R^2_{cum} [b]	Q^2_{LOO} [c]	*RMSEE*[d]
1	SVRDF	PLS	2	0.941	0.811	0.417
2	DISMAT 250	SMR	3	0.931	0.907	–
3	**SVWGM**	**PLS**	**2**	**0.944**	**0.864**	**0.459**
4	SVRG	PLS	2	0.955	0.908	0.368
5	SVRG	PLS	3	0.968	0.931	0.317

[a]principal components, [b]cumulative multiple correlation coefficient, [c]cumulative cross-validated, [d]root mean square error

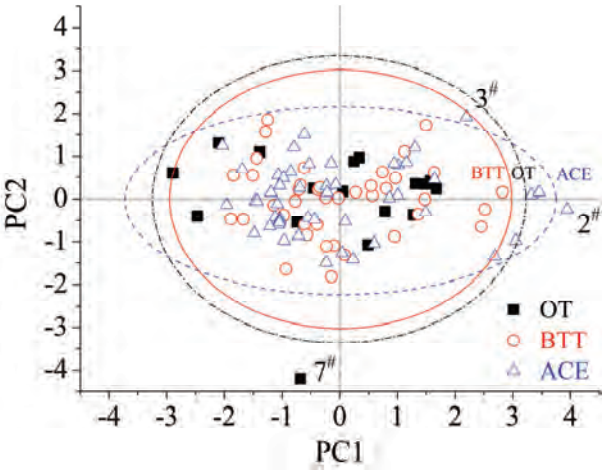


Fig. 2. Score scatter of the top two principle components for 21 OT, 48 BTT and 58 ACE.

A good statistical result often indicates a good robustness and high internal predictive power of a QSAR model. The cross-validation coefficient between the predicted and the observed values of the test set, Q^2_{ext}, was used to verify the external predictive ability of the PLS model. Recently, several novel methods for model validation were developed, such as the leave-several-out (LSO),[24]

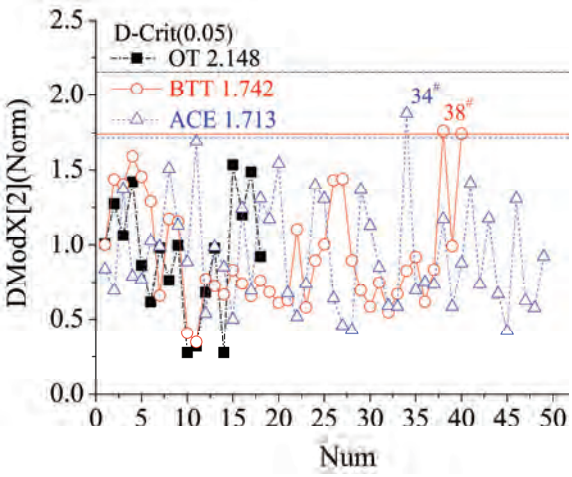


Fig. 3. Distance to the PLS model in *X* space of 21 OT, 48 BTT and 58 ACE.

Y-scrambling,[25] self-organizing mapping of molecular objects,[26] and external validation using division of a dataset into training and test sets.[25–27] In this work, the external predictive ability of the PLS model was calculated according to Eq. (1):[12,25]

$$Q^2_{\text{ext}} = 1 - \frac{\sum (Y_{\text{obs}} - Y_{\text{pred}})^2}{\sum (Y_{\text{obs}} - \overline{Y}_{\text{training}})^2} \quad (1)$$

where Y_{obs} and Y_{pred} indicate, respectively, the values of observed and predicted activities (the predicted values are listed in Table I), while $\overline{Y}_{\text{training}}$ indicates the mean activity value of the training set. The constructed model was then utilized to predict the test set, which gave Q^2_{ext} value of 0.863. All the results showed that the presented PLS model had an excellent prediction capability for drug design.

QSAR model for bitter tasting thresholds

As a classical sample set in QSAR studies, the 48 bitter tasting thresholds (BTT) reported by Collantes,[28] with their activities denoted by –log IC_{50} values (Table III), are often utilized to validate the efficiency of amino acid descriptors. Accordingly, each BTT could be characterized with 2×18 SVWGM scales, and using the same method, the 48 BTT were divided into 40 training samples and 8 test samples (highlighted with "*" in Table III). The optimum QSAR models of –log (IC_{50}) were created by the SMR–PLS method with 7 independent variables (v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6 and v_{12}). The results showed that the established model had a good predictive ability and strong robustness, with an R^2_{cum} value of 0.911, a Q^2_{LOO} of 0.818 (Table S-X of the Supplementary material), and Q^2_{ext} of 0.825. The plot of predicted value (predicted values are listed in Table III) against the observed data for BTT is presented in Fig. 4, from which it could be seen that

most samples were shown to be uniformly dispersed around the through the origin-passing diagonal. The results obtained from the PLS models using SVWGM descriptors are given in Table IV, showing that the herein obtained results were superior to those reported in the literature (Table IV). The scoring distribution scatter of the 48 samples in the spaces of the first and second PLS principal component is presented in Fig. 2. It could be seen that all samples fell within the Hotelling T^2 95 % confidence ellipse. The distance to the PLS model in the *X* space is shown in Fig. 3, from which it could be seen that the normalized distance to *X* for most sample was smaller than the critical value of 1.742 (significance level = 5 %), except for sample 38#.

TABLE III. Amino acid sequences of 48 BTT and their observed and predicted activities; Obs – observed values, Pred – predicted values, * – test set

No.	Peptide	Obs	Pred	No.	Peptide	Obs	Pred	No.	Peptide	Obs	Pred
1	GV	1.13	0.96	17	LL	2.35	2.55	33	IS	1.49	1.59
2	GL	1.68	1.54	18	LF*	2.75	2.88	34	IT	1.49	1.87
3	GI	1.70	1.60	19	LW	3.40	3.12	35	PA	1.32	1.39
4	GP	1.35	1.42	20	LY	2.46	2.67	36	PL*	2.22	2.09
5	GF	1.80	1.87	21	IG	1.68	1.40	37	PI	2.33	2.15
6	GW*	1.89	2.10	22	IA	1.68	1.73	38	PY	1.80	2.21
7	GY	1.77	1.66	23	IV	2.05	1.85	39	PF	2.80	2.42
8	AV	1.16	1.07	24	IL*	2.26	2.42	40	FG	1.77	1.75
9	AL	1.70	1.64	25	II	2.26	2.48	41	FL	2.87	2.78
10	AF	1.72	1.97	26	IP	2.40	2.31	42	FP*	2.70	2.67
11	VG	1.19	1.24	27	IW	3.05	2.99	43	FF	3.10	3.11
12	VA*	1.16	1.57	28	IN	1.49	1.69	44	FY	3.13	2.91
13	VV	1.71	1.69	29	ID	1.37	1.46	45	WE	1.56	1.50
14	VL	2.00	2.26	30	IQ*	1.49	1.17	46	WW	3.60	3.36
15	LG	1.72	1.53	31	IE	1.37	1.13	47	YL	2.40	2.54
16	LA	1.72	1.86	32	IK	1.65	1.84	48	SL*	1.49	1.68

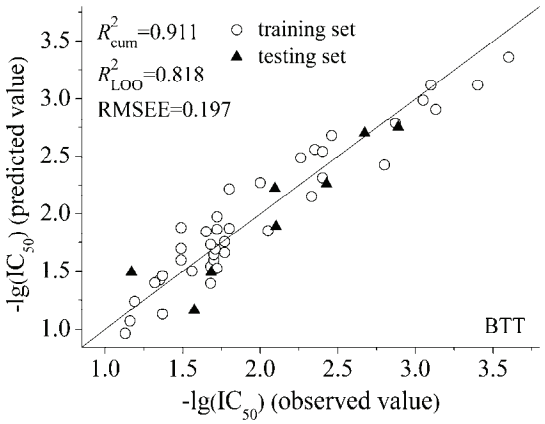


Fig. 4. Plot of predicted *vs*. observed values for 48 BTT by the PLS model.

TABLE IV. Comparison between QSAR models of BTT; –: not determined

No.	Descriptor	Model	A[a]	R^2_{cum} [b]	R^2_{LOO} [c]	*RMSEE*[d]
1	MEEV	MLR	3	0.649	0.57	0.37
2	MS–WHIM	PLS	3	0.704	0.633	–
3	MEEV	MLR	10	0.711	0.475	0.34
4	MEEV	MLR	3	0.735	0.677	0.32
5	MS–WHIM	PLS	3	0.754	0.71	0.32
6	SVRDF	PLS	1	0.766	0.86	0.209
7	MEEV	MLR	10	0.773	0.588	0.33
8	z scale	PLS	2	0.824	–	0.26
9	SSIA–HF	PLS	2	0.844	0.798	0.25
10	T-scales	PLS	2	0.845	0.786	0.39
11	c-scales	PLS	3	0.847	0.863	0.195
12	ISA–ECI	PLS	2	0.847	–	–
13	SSIA–AM1	PLS	1	0.85	0.837	0.25
14	SSIA–DFT	PLS	2	0.856	0.741	0.24
15	VSW	PLS	2	0.868	0.696	0.24
16	VSW	PLSd	2	0.873	0.751	0.23
17	VHSE	PLS	3	0.881	0.843	0.22
18	SVGETAWAY	PLS	2	0.887	0.753	0.216
19	SSIA–PM3	PLS	4	0.888	0.829	0.22
20	SZOTT	PLS	2	0.908	0.736	0.2
21	SVRG	PLS	3	0.91	0.863	0.195
22	**SVWGM**	**PLS**	**2**	**0.911**	**0.818**	**0.197**
23	SVHEHS	MLR	6	0.949	0.886	0.136
24	GRID	PLS	1	–	0.78	–

[a]principal components, [b]cumulative multiple correlation coefficient, [c]cumulative cross-validated, [d]root mean square error

QSAR model for angiotensin-converting enzyme inhibitors

The angiotensin-converting enzyme (ACE) plays a critical physiological role in raising blood pressure, the inhibition of ACE activity could lead to an overall antihypertensive effect.[29–31] Dipeptide sequences of 58 ACE inhibitors, a classical sample set for QSAR studies, with their activities expressed as log ($1/IC_{50}$) values,[10,32] was employed. The structure of each peptide was first quantified by 2×18 SVWMG scales. As shown in Table V, the dataset (n = 58) was divided into a training set of 49 samples and test set (highlighted with "*" in Table V) of 9 samples. The training set was then used for the development of QSAR models using the PLS technique, while the corresponding test set was utilized to validate the external prediction power of the developed model. In PLS regression analysis, Q^2_{LOO} (Table S-XI of the Supplementary material) changed with the SMR--introduced variables, achieving a maximum (0.799) at the seventh step, meanwhile, the correlative coefficient (Q^2_{cum}) was up to 0.865. The optimum PLS model with 7 variables (v_{23}, v_{34}, v_6, v_{28}, v_{30}, v_9 and v_{21}) was then obtained, wherein only two significant principal components were necessary to explain

TABLE V. Amino acid sequences of 58 ACE and their observed and predicted activities; Obs – observed values, Pred – predicted values, * – test set

No.	Peptide	Obs	Pred	No.	Peptide	Obs	Pred	No.	Peptide	Obs	Pred
1	VW	5.80	5.19	21	IG	2.92	2.93	41	GG	2.14	2.13
2	IW	5.70	5.58	22	GI	2.92	2.72	42	QG*	2.13	2.27
3	IY	5.43	4.48	23	GM	2.85	2.62	43	SG	2.07	2.30
4	AW	5.00	5.09	24	GA*	2.70	2.54	44	LG	2.06	2.57
5	RW	4.80	5.28	25	YG	2.70	2.39	45	GD	2.04	2.41
6	VY*	4.66	4.08	26	GL	2.60	3.10	46	TG	2.00	2.31
7	GW	4.52	4.79	27	AG	2.60	2.36	47	EG	2.00	2.00
8	VF	4.28	3.72	28	GH	2.51	2.40	48	DG*	1.85	1.99
9	AY	4.06	3.91	29	GR	2.49	2.41	49	PG	1.77	1.78
10	IP	3.89	4.02	30	KG*	2.49	2.27	50	LA	3.51	2.98
11	RP	3.74	3.71	31	FG	2.43	2.45	51	KA	3.42	2.68
12	AF*	3.72	3.55	32	GS	2.42	2.92	52	RA	3.34	3.03
13	GY	3.68	3.67	33	GV	2.34	2.46	53	YA	3.34	2.80
14	AP	3.64	3.46	34	MG	2.32	2.93	54	AA*	3.21	2.78
15	RF	3.64	3.81	35	GK	2.27	1.83	55	FR	3.04	2.74
16	VP	3.38	3.63	36	GE*	2.27	2.00	56	HL	2.49	3.13
17	GP	3.35	3.22	37	GT	2.24	2.39	57	DA	2.42	2.41
18	GF*	3.20	3.32	38	WG	2.23	2.14	58	EA	2.00	2.41
19	IF	3.03	4.12	39	HG	2.20	2.17				
20	VG	2.96	2.53	40	GQ	2.15	2.63				

93.0 % of the variance of *Y*. The model exhibited good prediction ability and stability, which was better than the models established by other methods (Table VI). The plot predicted by the optimum QSAR models and observed values are shown in Fig. 5 (the predicted values are listed in Table VI), where most samples were shown to be uniformly dispersed around the through the origin-passing diagonal and the model possessed a strong capability for fitting. The scoring distribution scatter of the 58 samples in the first and second PLS principal component spaces is presented in Fig. 2. It could be seen that all samples fell within the Hotelling T^2 95 % confidence level ellipse, with the exception of $2^{\#}$ and $3^{\#}$. The errors may be caused by the characterization method, experiment, or other reasons. No matter what kinds of reasons, there are only two samples with large errors, and the excellent results of this study should be recognized. The distance to the PLS model in the *X* space was described in Fig. 3. It could be seen that the normalized distance to *X* for most samples was smaller than the critical value of 1.713 (significance level = 5 %), except for sample $34^{\#}$.

In the optimum model, the external prediction power was evaluated using Eq. (1) and the Q^2_{ext} value reached 0.883. All the results showed that the presented PLS model could be effectively used to predict the acute activity of ACE.

TABLE VI. Comparison between QSAR models of ACE; –: not determined

No.	Descriptor	Model	A[a]	R^2_{cum} [b]	R^2_{LOO} [c]	*RMSEE*[d]
1	MEEV	MLR	10	0.711	0.475	0.34
2	MEEV	MLR	3	0.649	0.57	0.37
3	MS–WHIM	PLS	3	0.657	0.541	–
4	MS–WHIM	PLS	3	0.708	0.637	0.54
5	SSIA–DFT	PLS	2	0.734	0.678	0.52
6	MEEV	MLR	3	0.735	0.677	0.32
7	SSIA–AM1	PLS	1	0.769	0.699	0.49
8	z-scale	PLS	2	0.77	0.723	–
9	VHSE	PLS	1	0.77	0.745	0.48
10	MEEV	MLR	10	0.773	0.588	0.33
11	SVTV	PLS	1	0.789	0.767	0.45
12	VSTV	PLS	1	0.789	0.767	0.46
13	SSIA–PM3	PLS	4	0.789	0.773	0.47
14	SSIA–HF	PLS	2	0.797	0.76	0.46
15	T-scale	PLS	2	0.845	0.786	0.39
16	HSEHPCSV	PLS	1	0.846	0.835	0.396
17	V	PLS	2	0.849	0.783	0.4
18	VSW	PLSd	1	0.861	0.835	0.38
19	**SVWGM**	**PLS**	**2**	**0.865**	**0.799**	**0.386**
20	VSW	PLS	2	0.868	0.784	0.37
21	MHDV	PCR	19	0.878	0.753	0.35
22	SVWG	PLS	3	0.893	0.83	0.34
23	SZOTT	PLS	2	0.894	0.828	0.331
24	SVEEVA	PLS	2	0.894	0.839	0.332
25	SVHEHS	MLR	8	0.936	0.854	0.259

[a]principal components, [b]cumulative multiple correlation coefficient, [c]cumulative cross-validated, [d]root mean square error

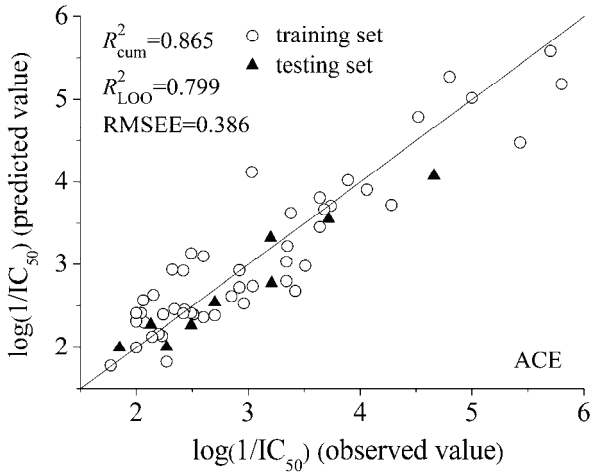


Fig. 5. Plot of predicted *vs.* observed values for 58 ACE by the PLS model.

CONCLUSIONS

In this study, a new set of descriptors-SVWGM was derived from principal component analyses of 99 WHIM descriptors, 74 geometrical descriptors and 160 3D-MORSE descriptors of 20 natural amino acids. On applying the SVWGM scales in the peptide QSAR studies for three kinds of classical peptide analogues, the results were similar to, or better than those found in the literature. The satisfactory results showed that the models had good prediction capability and favorable stability and the SVWGM descriptors could be well used to characterize molecular structure information and express the QSAR of peptide drugs.

SUPPLEMENTARY MATERIAL

The meanings and values of descriptors are available electronically from http:// //www.shd.org.rs/JSCS/, or from the corresponding author on request.

Acknowledgements. This work was supported by the National Natural Science Funds of China (21475081) and the Graduate Innovation Fund of Shaanxi University of Science and Technology.

И З В О Д

QSAR СТУДИЈА ПЕПТИДНИХ ЈЕДИЊЕЊА ЗАСНОВАНА НА НОВОМ ДЕСКРИПТОРУ АМИНОКИСЕЛИНА

JIAN-BO TONG, JIA CHANG, SHU-LING LIU и MIN BAI

Shaanxi University of Science & Technology, Xi'an, China

Метода квантитативних релација између структуре и активнисти (QSAR) примењена је за налажење везе између молекулске структуре и активности пептидних једињења. У овом раду коришћена је вишестепена вишеструка регресија да би се изабрао оптимални подскуп дескриптора, који имају значајан допринос активности за 21 окситоцински аналог, 48 једињења горког укуса и 58 инхибитора ангиотензин-конвертујућег ензима. Примењен је нови сет дескриптора SVWGM. Помоћу њих је добивен модел за предвиђање активности испитиваних супстанци. Модел је конструисан применом методе парцијалних најмањих квадрата, а затим је анализиран одговарајућим статистичким тестовима.

(Примљено 4. јуна, прихваћено 7. јула 2014)

REFERENCES

1. J. Jirácek, A. Yiotakis, B. Vincent, A. Lecoq, A. Nicolaou, F. Checler, *J. Biol. Chem.* **270** (1995) 21701
2. M. Marraud, A. Aubry, *J. Pept. Sci.* **40** (1996) 45
3. P. H. A. Sneath, *J. Theor. Biol.* **12** (1966) 157
4. A. Kidera, Y. Konishi, M. Oka, T. Ooi, H. A. Scheraga, *J. Protein Chem.* **4** (1985) 23
5. A. Kidera, Y. Konishi, T. Ooi, H. A. Scheraga, *J. Protein Chem.* **4** (1985) 265
6. K. Nakai, A. Kidera, M. Kanehisa, *Protein Eng.* **2** (1988) 93
7. S. Hellberg, L. Eriksson, J. Jonsson, F. Lindgren, M. Sjöström, B. Skagerberg, S. Wold, P. Andrews, *Int. J. Pept. Protein Res.* **37** (1991) 414
8. S. Hellberg, M. Sjöström, B. Skagerberg, S. Wold, *J. Med. Chem.* **30** (1987) 1126
9. S. Hellberg, M. Sjöström, S. Wold, *Acta Chem. Scand., B* **40** (1986) 135

10. S. Wold, L. Eriksson, S. Hellberg, J. Jonsson, M. Sjöström, B. Skagerberg, C. Wikström, *Can. J. Chem.* **65** (1987) 1814
11. J. Jonsson, L. Eriksson, S. Hellberg, M. Sjöström, S. Wold, *Quant. Struct.-Act. Relat.* **8** (1989) 204
12. L. Eriksson, J. Jonsson, S. Hellberg, F. Lindgren, B. Skagerberg, M. Sjostrom, S. Wold, *Acta Chem. Scand.* **44** (1990) 50
13. B. Hemmateenejad, S. Yousefinejad, A. R. Mehdipour, *Amino Acids* **40** (2011) 1169
14. T. J. Hou, Z. Xu, W. Zhang, W. A. McLaughlin, D. A. Case, Y. Xu, W. Wang, *Mol. Cell. Proteomics* **8** (2009) 639
15. T. J. Hou, Y. Y. Li, W. Wang, *Bioinformatics* **27** (2011) 1814
16. T. J. Hou, N. Li, Y. Y. Li, W. Wang, *J. Proteome Res.* **11** (2012) 2982
17. Z. Xu, T. J. Hou, N. Li, Y. Xu, W. Wang, *Mol. Cell. Proteomics* **11** (2012) O111
18. M. Sandberg, L. Eriksson, J. Jonsson, M. Sjöström, S. Wold, *J. Med. Chem.* **41** (1998) 2481
19. X. Y. Wang, J. Wang, Y. Lin, Y. Ding, Y. Q. Wang, X. M. Cheng, Z. H. Lin, *J. Mol. Model.* **17** (2011) 1599
20. A. Jain, D. Zongker, *IEEE Trans. Pattern Anal. Mach. Intell.* **19** (1997) 153
21. H. Mei, Y. Zhou, L. L. Sun, Z. L. Li, *Acta Phys.-Chim. Sin.* **20** (2004) 821
22. J. P. George, Z. Chen, P. Shaw, *World Acad. Sci. Eng. Technol.* **50** (2009) 970
23. T. Iwashita, *J. Statist. Plan. Inf.* **61** (1997) 85
24. J. Polanski, R. Gieleciak, A. Bak, *Comb. Chem. High. T. Scr.* **7** (2004) 793
25. A. Tropsha, P. Gramatica, V. K. Gombar, *QSAR Comb. Sci.* **22** (2003) 69
26. J. Polanski, A. Bak, R. Gieleciak, T. Magdziarz, *J. Chem. Inf. Model.* **46** (2006) 2310
27. A. Golbraikh, A. Tropsha, *J. Mol. Graphics Modell.* **20** (2002) 269
28. E. R. Collantes, W. J. Dunn, *J. Med. Chem.* **38** (1995) 2705
29. R. Natesh, S. L. Schwager, E. D. Sturrock, K. R. Acharya, *Nature* **421** (2003) 551
30. Z. P. Yu, B. Q. Liu, W. Z. Zhao, Y. G. Yin, J. B. Liu, F. Chen, *Food Chem.* **133** (2012) 315
31. M. A. Ondetti, D. W. Cushman, *Annu. Rev. Biochem.* **51** (1982) 283
32. F. F. Tian, P. Zhou, Z. L. Li, *J. Mol. Struct.* **830** (2007) 106.

Journal of the Serbian Chemical Society
JSCS@tmf.bg.ac.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) S74–S102 (2015) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

A quantitative structure–activity relationship (QSAR) study of peptide drugs based on a new descriptor of amino acids

JIAN-BO TONG*, JIA CHANG, SHU-LING LIU and MIN BAI

Shaanxi University of Science and Technology, Xi'an 710021, China

J. Serb. Chem. Soc. 80 (3) (2015) 343–353

TABLE S-I. The meanings of 99 WHIM descriptors

ID	Symbol	Meaning
1	L1u	1st component size directional WHIM index/unweighted
2	L2u	2nd component size directional WHIM index/unweighted
3	L3u	3rd component size directional WHIM index/unweighted
4	P1u	1st component shape directional WHIM index/unweighted
5	P2u	2nd component shape directional WHIM index/unweighted
6	G1u	1st component symmetry directional WHIM index/unweighted
7	G2u	2nd component symmetry directional WHIM index/unweighted
8	G3u	3rd component symmetry directional WHIM index/unweighted
9	E1u	1st component accessibility directional WHIM index/unweighted
10	E2u	2nd component accessibility directional WHIM index/unweighted
11	E3u	3rd component accessibility directional WHIM index/unweighted
12	L1m	1st component size directional WHIM index/weighted by atomic masses
13	L2m	2nd component size directional WHIM index/weighted by atomic masses
14	L3m	3rd component size directional WHIM index/weighted by atomic masses
15	P1m	1st component shape directional WHIM index/weighted by atomic masses
16	P2m	2nd component shape directional WHIM index/weighted by atomic masses
17	G1m	1st component symmetry directional WHIM index/weighted by atomic masses
18	G2m	2nd component symmetry directional WHIM index/weighted by atomic masses
19	G3m	3rd component symmetry directional WHIM index/weighted by atomic masses
20	E1m	1st component accessibility directional WHIM index/weighted by atomic masses
21	E2m	2nd component accessibility directional WHIM index/weighted by atomic masses
22	E3m	3rd component accessibility directional WHIM index/weighted by atomic masses

* Corresponding author. E-mail: jianbotong@aliyun.com

TABLE S-I. Continued

ID	Symbol	Meaning
23	L1v	1st component size directional WHIM index/weighted by atomic van der Waals volumes
24	L2v	2nd component size directional WHIM index/weighted by atomic van der Waals volumes
25	L3v	3rd component size directional WHIM index/weighted by atomic van der Waals volumes
26	P1v	1st component shape directional WHIM index/weighted by atomic van der Waals volumes
27	P2v	2nd component shape directional WHIM index/weighted by atomic van der Waals volumes
28	G1v	1st component symmetry directional WHIM index/weighted by atomic van der Waals volumes
29	G2v	2nd component symmetry directional WHIM index/weighted by atomic van der Waals volumes
30	G3v	3rd component symmetry directional WHIM index/weighted by atomic van der Waals volumes
31	E1v	1st component accessibility directional WHIM index/weighted by atomic van der Waals volumes
32	E2v	2nd component accessibility directional WHIM index/weighted by atomic van der Waals volumes
33	E3v	3rd component accessibility directional WHIM index/weighted by atomic van der Waals volumes
34	L1e	1st component size directional WHIM index/weighted by atomic Sanderson electronegativities
35	L2e	2nd component size directional WHIM index/weighted by atomic Sanderson electronegativities
36	L3e	3rd component size directional WHIM index/weighted by atomic Sanderson electronegativities
37	P1e	1st component shape directional WHIM index/weighted by atomic Sanderson electronegativities
38	P2e	2nd component shape directional WHIM index/weighted by atomic Sanderson electronegativities
39	G1e	1st component symmetry directional WHIM index/weighted by atomic Sanderson electronegativities
40	G2e	2nd component symmetry directional WHIM index/weighted by atomic Sanderson electronegativities
41	G3e	3rd component symmetry directional WHIM index/weighted by atomic Sanderson electronegativities
42	E1e	1st component accessibility directional WHIM index/weighted by atomic Sanderson electronegativities
43	E2e	2nd component accessibility directional WHIM index/weighted by atomic Sanderson electronegativities
44	E3e	3rd component accessibility directional WHIM index/weighted by atomic Sanderson electronegativities

TABLE S-I. Continued

ID	Symbol	Meaning
45	L1p	1st component size directional WHIM index/weighted by atomic polarizabilities
46	L2p	2nd component size directional WHIM index/weighted by atomic polarizabilities
47	L3p	3rd component size directional WHIM index/weighted by atomic polarizabilities
48	P1p	1st component shape directional WHIM index/weighted by atomic polarizabilities
49	P2p	2nd component shape directional WHIM index/weighted by atomic polarizabilities
50	G1p	1st component symmetry directional WHIM index/weighted by atomic polarizabilities
51	G2p	2nd component symmetry directional WHIM index/weighted by atomic polarizabilities
52	G3p	3rd component symmetry directional WHIM index/weighted by atomic polarizabilities
53	E1p	1st component accessibility directional WHIM index/weighted by atomic polarizabilities
54	E2p	2nd component accessibility directional WHIM index/weighted by atomic polarizabilities
55	E3p	3rd component accessibility directional WHIM index/weighted by atomic polarizabilities
56	L1s	1st component size directional WHIM index/weighted by atomic electrotopological states
57	L2s	2nd component size directional WHIM index/weighted by atomic electrotopological states
58	L3s	3rd component size directional WHIM index/weighted by atomic electrotopological states
59	P1s	1st component shape directional WHIM index/weighted by atomic electrotopological states
60	P2s	2nd component shape directional WHIM index/weighted by atomic electrotopological states
61	G1s	1st component symmetry directional WHIM index/weighted by atomic electrotopological states
62	G2s	2nd component symmetry directional WHIM index/weighted by atomic electrotopological states
63	G3s	3rd component symmetry directional WHIM index/weighted by atomic electrotopological states
64	E1s	1st component accessibility directional WHIM index/weighted by atomic electrotopological states
65	E2s	2nd component accessibility directional WHIM index/weighted by atomic electrotopological states
66	E3s	3rd component accessibility directional WHIM index/weighted by atomic electrotopological states
67	Tu	T total size index/unweighted

TABLE S-I. Continued

ID	Symbol	Meaning
68	Tm	T total size index/weighted by atomic masses
69	Tv	T total size index/weighted by atomic van der Waals volumes
70	Te	T total size index/weighted by atomic Sanderson electronegativities
71	Tp	T total size index/weighted by atomic polarizabilities
72	Ts	T total size index/weighted by atomic electrotopological states
73	Au	A total size index/unweighted
74	Am	A total size index/weighted by atomic masses
75	Av	A total size index/weighted by atomic van der Waals volumes
76	Ae	A total size index/weighted by atomic Sanderson electronegativities
77	Ap	A total size index/weighted by atomic polarizabilities
78	As	A total size index/weighted by atomic electrotopological states
79	Gu	G total symmetry index/unweighted
80	Gm	G total symmetry index/weighted by atomic masses
81	Gs	G total symmetry index/weighted by atomic Sanderson electronegativities
82	Ku	K global shape index/unweighted
83	Km	K global shape index/weighted by atomic masses
84	Kv	K global shape index/weighted by atomic van der Waals volumes
85	Ke	K global shape index/weighted by atomic Sanderson electronegativities
86	Kp	K global shape index/weighted by atomic polarizabilities
87	Ks	K global shape index/weighted by atomic electrotopological states
88	Du	D total accessibility index/unweighted
89	Dm	D total accessibility index/weighted by atomic masses
90	Dv	D total accessibility index/weighted by atomic van der Waals volumes
91	De	D total accessibility index/weighted by atomic Sanderson electronegativities
92	Dp	D total accessibility index/weighted by atomic polarizabilities
93	Ds	D total accessibility index/weighted by atomic electrotopological states
94	Vu	V total size index/unweighted
95	Vm	V total size index/weighted by atomic masses
96	Vv	V total size index/weighted by atomic van der Waals volumes
97	Ve	V total size index/weighted by atomic Sanderson electronegativities
98	Vp	V total size index/weighted by atomic polarizabilities
99	Vs	V total size index/weighted by atomic electrotopological states

TABLE S-II. The meanings of 74 geometrical descriptors

ID	Symbol	Meaning
1	W3D	3D-Wener index
2	J3D	3D-Balaban index
3	H3D	3D-Harary index
4	AGDD	Average geometric distance degree
5	DDI	D/D index
6	ADDD	Average distance/ distance degree
7	G1	Gravitational index G1
8	G2	Gravitational index G2(bond-restricted)
9	RGyr	Radius of gyration (mass weighted)
10	SPAN	Span R

TABLE S-II. Continued

ID	Symbol	Meaning
11	SPAM	Average span R
12	MEcc	Average eccentricity
13	SPH	Spherosity
14	ASP	Asphericity
15	FDI	Folding degree index
16	PJI3	3D petitjean shape index
17	L/Bw	Length-to-breadth ratio by WHIM
18	SEig	Absolute eigenvalue sum on geometry matrix
19	HOMA	Harmonic Oscillator Model of Aromaticity index
20	RCI	Jug RG index
21	AROM	Aromaticity index
22	HOMT	HOMA total
23	DISPm	d COMMA2 value /weighted by atomic masses
24	QXXm	Qxx COMMA 2 value /weighted by atomic masses
25	QYYm	Qyy COMMA 2 value /weighted by atomic masses
26	QZZm	Qzz COMMA 2 value /weighted by atomic masses
27	DISPv	d COMMA2 value /weighted by atomic van der waals volumes
28	QXXv	Qxx COMMA 2 value /weighted by atomic van der Waals volumes
29	QYYv	Qyy COMMA 2 value /weighted by atomic van der Waals volumes
30	QZZv	Qzz COMMA 2 value /weighted by atomic van der Waals volumes
31	DISPe	d COMMA2 value /weighted by atomic Sanderson electronegativities
32	QXXe	Qxx COMMA 2 value /weighted by atomic Sanderson electronegativities
33	QYYe	Qyy COMMA 2 value /weighted by atomic Sanderson electronegativities
34	QZZe	Qzz COMMA 2 value /weighted by atomic Sanderson electronegativities
35	DISPp	d COMMA2 value /weighted by atomic polarizabilities
36	QXXp	Qxx COMMA 2 value /weighted by atomic polarizabilities
37	QYYp	Qyy COMMA 2 value /weighted by atomic polarizabilities
38	QZZp	Qzz COMMA 2 value /weighted by atomic polarizabilities
39	G(N..N)	Sum of geometrical distance between N..N
40	G(N..O)	Sum of geometrical distance between N..O
41	G(N..S)	Sum of geometrical distance between N..S
42	G(N..P)	Sum of geometrical distance between N..P
43	G(N..F)	Sum of geometrical distance between N..F
44	G(N..Cl)	Sum of geometrical distance between N..Cl
45	G(N..Br)	Sum of geometrical distance between N..Br
46	G(N..I)	Sum of geometrical distance between N..I
47	G(O..O)	Sum of geometrical distance between O..O
48	G(O..S)	Sum of geometrical distance between O..S
49	G(O..P)	Sum of geometrical distance between O..P
50	G(O..F)	Sum of geometrical distance between O..F
51	G(O..Cl)	Sum of geometrical distance between O..Cl
52	G(O..Br)	Sum of geometrical distance between O..Br
53	G(O..I)	Sum of geometrical distance between O..I
54	G(S..S)	Sum of geometrical distance between S..S
55	G(S..P)	Sum of geometrical distance between S..P

TABLE S-II. Continued

ID	Symbol	Meaning
56	G(S..F)	Sum of geometrical distance between S..F
57	G(S..Cl)	Sum of geometrical distance between S..Cl
58	G(S..Br)	Sum of geometrical distance between S..Br
59	G(S..I)	Sum of geometrical distance between S..I
60	G(P..P)	Sum of geometrical distance between P..P
61	G(P..F)	Sum of geometrical distance between P..F
62	G(P..Cl)	Sum of geometrical distance between P..Cl
63	G(P..Br)	Sum of geometrical distance between P..Br
64	G(P..I)	Sum of geometrical distance between P..I
65	G(F..F)	Sum of geometrical distance between F..F
66	G(F..Cl)	Sum of geometrical distance between F..Cl
67	G(F..Br)	Sum of geometrical distance between F..Br
68	G(F..I)	Sum of geometrical distance between F..I
69	G(Cl..Cl)	Sum of geometrical distance between Cl..Cl
70	G(Cl..Br)	Sum of geometrical distance between Cl..Br
71	G(Cl..I)	Sum of geometrical distance between Cl..I
72	G(Br..Br)	Sum of geometrical distance between Br..Br
73	G(Br..I)	Sum of geometrical distance between Br..I
74	G(I..I)	Sum of geometrical distance between I..I

TABLE S-III. The meanings of 160 3D-MORSE descriptors

ID	Symbol	Meaning
1	Mor01u	3D-MoRSE -signal 01/unweighted
2	Mor02u	3D-MoRSE -signal 02/unweighted
3	Mor03u	3D-MoRSE -signal 03/unweighted
4	Mor04u	3D-MoRSE -signal 04/unweighted
5	Mor05u	3D-MoRSE -signal 05/unweighted
6	Mor06u	3D-MoRSE -signal 06/unweighted
7	Mor07u	3D-MoRSE -signal 07/unweighted
8	Mor08u	3D-MoRSE -signal 08/unweighted
9	Mor09u	3D-MoRSE -signal 09/unweighted
10	Mor10u	3D-MoRSE -signal 10/unweighted
11	Mor11u	3D-MoRSE -signal 11/unweighted
12	Mor12u	3D-MoRSE -signal 12/unweighted
13	Mor13u	3D-MoRSE -signal 13/unweighted
14	Mor14u	3D-MoRSE -signal 14/unweighted
15	Mor15u	3D-MoRSE -signal 15/unweighted
16	Mor16u	3D-MoRSE -signal 16/unweighted
17	Mor17u	3D-MoRSE -signal 17/unweighted
18	Mor18u	3D-MoRSE -signal 18/unweighted
19	Mor19u	3D-MoRSE -signal 19/unweighted
20	Mor20u	3D-MoRSE -signal 20/unweighted
21	Mor21u	3D-MoRSE -signal 21/unweighted
22	Mor22u	3D-MoRSE -signal 22/unweighted
23	Mor23u	3D-MoRSE -signal 23/unweighted

TABLE S-III. Continued

ID	Symbol	Meaning
24	Mor24u	3D-MoRSE -signal 24/unweighted
25	Mor25u	3D-MoRSE -signal 25/unweighted
26	Mor26u	3D-MoRSE -signal 26/unweighted
27	Mor27u	3D-MoRSE -signal 27/unweighted
28	Mor28u	3D-MoRSE -signal 28/unweighted
29	Mor29u	3D-MoRSE -signal 29/unweighted
30	Mor30u	3D-MoRSE -signal 30/unweighted
31	Mor31u	3D-MoRSE -signal 31/unweighted
32	Mor32u	3D-MoRSE -signal 32/unweighted
33	Mor01m	3D-MoRSE -signal 01/weighted by atomic masses
34	Mor02m	3D-MoRSE -signal 02/weighted by atomic masses
35	Mor03m	3D-MoRSE -signal 03/weighted by atomic masses
36	Mor04m	3D-MoRSE -signal 04/weighted by atomic masses
37	Mor05m	3D-MoRSE -signal 05/weighted by atomic masses
38	Mor06m	3D-MoRSE -signal 06/weighted by atomic masses
39	Mor07m	3D-MoRSE -signal 07/weighted by atomic masses
40	Mor08m	3D-MoRSE -signal 08/weighted by atomic masses
41	Mor09m	3D-MoRSE -signal 09/weighted by atomic masses
42	Mor10m	3D-MoRSE -signal 10/weighted by atomic masses
43	Mor11m	3D-MoRSE -signal 11/weighted by atomic masses
44	Mor12m	3D-MoRSE -signal 12/weighted by atomic masses
45	Mor13m	3D-MoRSE -signal 13/weighted by atomic masses
46	Mor14m	3D-MoRSE -signal 14/weighted by atomic masses
47	Mor15m	3D-MoRSE -signal 15/weighted by atomic masses
48	Mor16m	3D-MoRSE -signal 16/weighted by atomic masses
49	Mor17m	3D-MoRSE -signal 17/weighted by atomic masses
50	Mor18m	3D-MoRSE -signal 18/weighted by atomic masses
51	Mor19m	3D-MoRSE -signal 19/weighted by atomic masses
52	Mor20m	3D-MoRSE -signal 20/weighted by atomic masses
53	Mor21m	3D-MoRSE -signal 21/weighted by atomic masses
54	Mor22m	3D-MoRSE -signal 22/weighted by atomic masses
55	Mor23m	3D-MoRSE -signal 23/weighted by atomic masses
56	Mor24m	3D-MoRSE -signal 24/weighted by atomic masses
57	Mor25m	3D-MoRSE -signal 25/weighted by atomic masses
58	Mor26m	3D-MoRSE -signal 26/weighted by atomic masses
59	Mor27m	3D-MoRSE -signal 27/weighted by atomic masses
60	Mor28m	3D-MoRSE -signal 28/weighted by atomic masses
61	Mor29m	3D-MoRSE -signal 29/weighted by atomic masses
62	Mor30m	3D-MoRSE -signal 30/weighted by atomic masses
63	Mor31m	3D-MoRSE -signal 31/weighted by atomic masses
64	Mor32m	3D-MoRSE -signal 32/weighted by atomic masses
65	Mor01v	3D-MoRSE -signal 01/weighted by atomic van der Waals volumes
66	Mor02v	3D-MoRSE -signal 02/weighted by atomic van der Waals volumes
67	Mor03v	3D-MoRSE -signal 03/weighted by atomic van der Waals volumes
68	Mor04v	3D-MoRSE -signal 04/weighted by atomic van der Waals volumes

TABLE S-III. Continued

ID	Symbol	Meaning
69	Mor05v	3D-MoRSE -signal 05/weighted by atomic van der Waals volumes
70	Mor06v	3D-MoRSE -signal 06/weighted by atomic van der Waals volumes
71	Mor07v	3D-MoRSE -signal 07/weighted by atomic van der Waals volumes
72	Mor08v	3D-MoRSE -signal 08/weighted by atomic van der Waals volumes
73	Mor09v	3D-MoRSE -signal 09/weighted by atomic van der Waals volumes
74	Mor10v	3D-MoRSE -signal 10/weighted by atomic van der Waals volumes
75	Mor11v	3D-MoRSE -signal 11/weighted by atomic van der Waals volumes
76	Mor12v	3D-MoRSE -signal 12/weighted by atomic van der Waals volumes
77	Mor13v	3D-MoRSE -signal 13/weighted by atomic van der Waals volumes
78	Mor14v	3D-MoRSE -signal 14/weighted by atomic van der Waals volumes
79	Mor15v	3D-MoRSE -signal 15/weighted by atomic van der Waals volumes
80	Mor16v	3D-MoRSE -signal 16/weighted by atomic van der Waals volumes
81	Mor17v	3D-MoRSE -signal 17/weighted by atomic van der Waals volumes
82	Mor18v	3D-MoRSE -signal 18/weighted by atomic van der Waals volumes
83	Mor19v	3D-MoRSE -signal 19/weighted by atomic van der Waals volumes
84	Mor20v	3D-MoRSE -signal 20/weighted by atomic van der Waals volumes
85	Mor21v	3D-MoRSE -signal 21/weighted by atomic van der Waals volumes
86	Mor22v	3D-MoRSE -signal 22/weighted by atomic van der Waals volumes
87	Mor23v	3D-MoRSE -signal 23/weighted by atomic van der Waals volumes
88	Mor24v	3D-MoRSE -signal 24/weighted by atomic van der Waals volumes
89	Mor25v	3D-MoRSE -signal 25/weighted by atomic van der Waals volumes
90	Mor26v	3D-MoRSE -signal 26/weighted by atomic van der Waals volumes
91	Mor27v	3D-MoRSE -signal 27/weighted by atomic van der Waals volumes
92	Mor28v	3D-MoRSE -signal 28/weighted by atomic van der Waals volumes
93	Mor29v	3D-MoRSE -signal 29/weighted by atomic van der Waals volumes
94	Mor30v	3D-MoRSE -signal 30/weighted by atomic van der Waals volumes
95	Mor31v	3D-MoRSE -signal 31/weighted by atomic van der Waals volumes
96	Mor32v	3D-MoRSE -signal 32/weighted by atomic van der Waals volumes
97	Mor01e	3D-MoRSE -signal 01/weighted by atomic Sanderson electronegativities
98	Mor02e	3D-MoRSE -signal 02/weighted by atomic Sanderson electronegativities
99	Mor03e	3D-MoRSE -signal 03/weighted by atomic Sanderson electronegativities
100	Mor04e	3D-MoRSE -signal 04/weighted by atomic Sanderson electronegativities
101	Mor05e	3D-MoRSE -signal 05/weighted by atomic Sanderson electronegativities
102	Mor06e	3D-MoRSE -signal 06/weighted by atomic Sanderson electronegativities
103	Mor07e	3D-MoRSE -signal 07/weighted by atomic Sanderson electronegativities
104	Mor08e	3D-MoRSE -signal 08/weighted by atomic Sanderson electronegativities
105	Mor09e	3D-MoRSE -signal 09/weighted by atomic Sanderson electronegativities
106	Mor10e	3D-MoRSE -signal 10/weighted by atomic Sanderson electronegativities
107	Mor11e	3D-MoRSE -signal 11/weighted by atomic Sanderson electronegativities
108	Mor12e	3D-MoRSE -signal 12/weighted by atomic Sanderson electronegativities
109	Mor13e	3D-MoRSE -signal 13/weighted by atomic Sanderson electronegativities
110	Mor14e	3D-MoRSE -signal 14/weighted by atomic Sanderson electronegativities
111	Mor15e	3D-MoRSE -signal 15/weighted by atomic Sanderson electronegativities
112	Mor16e	3D-MoRSE -signal 16/weighted by atomic Sanderson electronegativities
113	Mor17e	3D-MoRSE -signal 17/weighted by atomic Sanderson electronegativities

TABLE S-III. Continued

ID	Symbol	Meaning
114	Mor18e	3D-MoRSE -signal 18/weighted by atomic Sanderson electronegativities
115	Mor19e	3D-MoRSE -signal 19/weighted by atomic Sanderson electronegativities
116	Mor20e	3D-MoRSE -signal 20/weighted by atomic Sanderson electronegativities
117	Mor21e	3D-MoRSE -signal 21/weighted by atomic Sanderson electronegativities
118	Mor22e	3D-MoRSE -signal 22/weighted by atomic Sanderson electronegativities
119	Mor23e	3D-MoRSE -signal 23/weighted by atomic Sanderson electronegativities
120	Mor24e	3D-MoRSE -signal 24/weighted by atomic Sanderson electronegativities
121	Mor25e	3D-MoRSE -signal 25/weighted by atomic Sanderson electronegativities
122	Mor26e	3D-MoRSE -signal 26/weighted by atomic Sanderson electronegativities
123	Mor27e	3D-MoRSE -signal 27/weighted by atomic Sanderson electronegativities
124	Mor28e	3D-MoRSE -signal 28/weighted by atomic Sanderson electronegativities
125	Mor29e	3D-MoRSE -signal 29/weighted by atomic Sanderson electronegativities
126	Mor30e	3D-MoRSE -signal 30/weighted by atomic Sanderson electronegativities
127	Mor31e	3D-MoRSE -signal 31/weighted by atomic Sanderson electronegativities
128	Mor32e	3D-MoRSE -signal 32/weighted by atomic Sanderson electronegativities
129	Mor01p	3D-MoRSE -signal 01/weighted by atomic polarizabilities
130	Mor02p	3D-MoRSE -signal 02/weighted by atomic polarizabilities
131	Mor03p	3D-MoRSE -signal 03/weighted by atomic polarizabilities
132	Mor04p	3D-MoRSE -signal 04/weighted by atomic polarizabilities
133	Mor05p	3D-MoRSE -signal 05/weighted by atomic polarizabilities
134	Mor06p	3D-MoRSE -signal 06/weighted by atomic polarizabilities
135	Mor07p	3D-MoRSE -signal 07/weighted by atomic polarizabilities
136	Mor08p	3D-MoRSE -signal 08/weighted by atomic polarizabilities
137	Mor09p	3D-MoRSE -signal 09/weighted by atomic polarizabilities
138	Mor10p	3D-MoRSE -signal 10/weighted by atomic polarizabilities
139	Mor11p	3D-MoRSE -signal 11/weighted by atomic polarizabilities
140	Mor12p	3D-MoRSE -signal 12/weighted by atomic polarizabilities
141	Mor13p	3D-MoRSE -signal 13/weighted by atomic polarizabilities
142	Mor14p	3D-MoRSE -signal 14/weighted by atomic polarizabilities
143	Mor15p	3D-MoRSE -signal 15/weighted by atomic polarizabilities
144	Mor16p	3D-MoRSE -signal 16/weighted by atomic polarizabilities
145	Mor17p	3D-MoRSE -signal 17/weighted by atomic polarizabilities
146	Mor18p	3D-MoRSE -signal 18/weighted by atomic polarizabilities
147	Mor19p	3D-MoRSE -signal 19/weighted by atomic polarizabilities
148	Mor20p	3D-MoRSE -signal 20/weighted by atomic polarizabilities
149	Mor21p	3D-MoRSE -signal 21/weighted by atomic polarizabilities
150	Mor22p	3D-MoRSE -signal 22/weighted by atomic polarizabilities
151	Mor23p	3D-MoRSE -signal 23/weighted by atomic polarizabilities
152	Mor24p	3D-MoRSE -signal 24/weighted by atomic polarizabilities
153	Mor25p	3D-MoRSE -signal 25/weighted by atomic polarizabilities
154	Mor26p	3D-MoRSE -signal 26/weighted by atomic polarizabilities
155	Mor27p	3D-MoRSE -signal 27/weighted by atomic polarizabilities
156	Mor28p	3D-MoRSE -signal 28/weighted by atomic polarizabilities
157	Mor29p	3D-MoRSE -signal 29/weighted by atomic polarizabilities
158	Mor30p	3D-MoRSE -signal 30/weighted by atomic polarizabilities
159	Mor31p	3D-MoRSE -signal 31/weighted by atomic polarizabilities
160	Mor32p	3D-MoRSE -signal 32/weighted by atomic polarizabilities

TABLE S-IV. The values of 99 WHIM descriptors of 20 amino acid

Amino acid	1	2	3	4	5	6	7	8	9	10
Ala A	2.076	1.214	0.482	0.55	0.322	0.213	0.213	0.213	0.427	0.51
Arg R	7.806	1.54	0.675	0.779	0.154	0.175	0.175	0.175	0.543	0.387
Asn N	2.115	1.541	0.908	0.463	0.338	0.197	0.236	0.197	0.579	0.503
Asp D	2.579	1.498	0.863	0.522	0.303	0.2	0.2	0.2	0.447	0.395
Cys C	2.627	1.089	0.638	0.603	0.25	0.208	0.208	0.208	0.43	0.424
Gln Q	4.74	1.209	0.66	0.717	0.183	0.188	0.219	0.188	0.509	0.448
Glu E	4.059	1.324	0.86	0.65	0.212	0.191	0.191	0.191	0.465	0.37
Gly G	2.36	0.727	0.168	0.725	0.223	0.231	0.231	0.231	0.482	0.446
His H	4.126	1.475	0.964	0.628	0.225	0.188	0.188	0.188	0.502	0.35
Ile I	3.163	1.692	1.022	0.538	0.288	0.183	0.183	0.183	0.49	0.461
Leu L	3.25	1.692	0.959	0.551	0.287	0.183	0.183	0.211	0.594	0.548
Lys K	6.389	1.219	0.862	0.754	0.144	0.179	0.179	0.179	0.55	0.35
Met M	5.175	1.355	0.866	0.7	0.183	0.188	0.188	0.188	0.491	0.367
Phe F	6.581	1.404	0.702	0.758	0.162	0.181	0.207	0.181	0.497	0.464
Pro P	2.17	1.347	0.775	0.506	0.314	0.197	0.197	0.197	0.497	0.553
Ser S	2.076	1.272	0.741	0.508	0.311	0.208	0.208	0.208	0.452	0.456
Thr T	2.884	1.397	0.624	0.588	0.285	0.197	0.197	0.197	0.471	0.36
Trp W	7.091	2.364	0.711	0.698	0.232	0.174	0.174	0.195	0.569	0.487
Tyr Y	6.392	1.859	0.716	0.713	0.207	0.179	0.179	0.179	0.523	0.448
Val V	3.033	1.424	0.784	0.579	0.272	0.191	0.191	0.191	0.433	0.527
	11	12	13	14	15	16	17	18	19	20
Ala A	0.306	1.835	0.891	0.292	0.608	0.295	0.213	0.213	0.213	0.352
Arg R	0.423	7.543	1.347	0.419	0.81	0.145	0.175	0.175	0.175	0.518
Asn N	0.455	2.38	1.404	0.428	0.565	0.333	0.197	0.197	0.197	0.791
Asp D	0.489	2.684	0.95	0.664	0.624	0.221	0.2	0.2	0.2	0.492
Cys C	0.428	3.449	0.708	0.323	0.77	0.158	0.208	0.208	0.208	0.767
Gln Q	0.408	4.594	1.12	0.304	0.777	0.176	0.188	0.188	0.188	0.637
Glu E	0.515	4.685	0.899	0.643	0.752	0.144	0.191	0.191	0.191	0.639
Gly G	0.202	1.814	0.661	0.045	0.72	0.262	0.231	0.231	0.231	0.29
His H	0.418	3.985	1.136	0.541	0.704	0.201	0.188	0.188	0.188	0.516
Ile I	0.549	3.146	1.111	0.712	0.633	0.224	0.183	0.211	0.183	0.502
Leu L	0.544	2.924	1.278	0.6	0.609	0.266	0.183	0.183	0.183	0.524
Lys K	0.508	6.63	0.917	0.556	0.818	0.113	0.179	0.179	0.179	0.598
Met M	0.428	4.902	0.849	0.646	0.766	0.133	0.188	0.188	0.188	0.489
Phe F	0.465	6.635	0.825	0.409	0.843	0.105	0.181	0.181	0.181	0.504
Pro P	0.451	2.747	0.862	0.325	0.698	0.219	0.197	0.197	0.197	0.774
Ser S	0.467	1.854	0.94	0.466	0.569	0.288	0.208	0.208	0.208	0.41
Thr T	0.428	2.703	1.067	0.418	0.646	0.255	0.197	0.197	0.197	0.421
Trp W	0.273	6.05	1.497	0.667	0.737	0.182	0.174	0.174	0.174	0.424
Tyr Y	0.349	6.215	1.161	0.577	0.782	0.146	0.179	0.179	0.204	0.53
Val V	0.526	2.932	0.983	0.402	0.679	0.228	0.191	0.191	0.191	0.405
	21	22	23	24	25	26	27	28	29	30
Ala A	0.272	0.102	1.672	0.982	0.347	0.557	0.327	0.213	0.213	0.213
Arg R	0.342	0.172	7.174	1.273	0.472	0.804	0.143	0.175	0.175	0.175
Asn N	0.357	0.08	1.985	1.204	0.68	0.513	0.311	0.197	0.197	0.217

TABLE S-IV. Continued

Amino acid	21	22	23	24	25	26	27	28	29	30
Asp D	0.166	0.296	2.274	1.094	0.656	0.565	0.272	0.2	0.2	0.2
Cys C	0.188	0.103	2.586	0.843	0.434	0.669	0.218	0.208	0.208	0.208
Gln Q	0.243	0.088	4.28	1.008	0.464	0.744	0.175	0.188	0.188	0.188
Glu E	0.224	0.27	3.762	0.967	0.639	0.701	0.18	0.191	0.191	0.224
Gly G	0.376	0.014	1.852	0.573	0.106	0.732	0.226	0.231	0.231	0.313
His H	0.221	0.135	3.912	1.084	0.691	0.688	0.191	0.188	0.188	0.188
Ile I	0.206	0.236	2.862	1.306	0.784	0.578	0.264	0.183	0.183	0.183
Leu L	0.284	0.211	2.86	1.362	0.703	0.581	0.277	0.183	0.183	0.183
Lys K	0.207	0.208	5.952	0.909	0.635	0.794	0.121	0.179	0.179	0.179
Met M	0.15	0.194	4.82	0.971	0.701	0.743	0.15	0.188	0.188	0.188
Phe F	0.178	0.143	5.965	1.035	0.486	0.797	0.138	0.181	0.181	0.181
Pro P	0.224	0.063	2.055	1.018	0.525	0.571	0.283	0.197	0.197	0.197
Ser S	0.266	0.181	1.726	1.014	0.539	0.526	0.309	0.208	0.208	0.208
Thr T	0.205	0.166	2.574	1.059	0.449	0.631	0.259	0.197	0.197	0.197
Trp W	0.198	0.262	6.203	1.762	0.562	0.727	0.207	0.174	0.174	0.174
Tyr Y	0.216	0.221	5.536	1.341	0.561	0.744	0.18	0.179	0.179	0.179
Val V	0.236	0.079	2.601	1.1	0.596	0.605	0.256	0.191	0.191	0.191
	31	32	33	34	35	36	37	38	39	40
Ala A	0.289	0.326	0.159	2.111	1.192	0.476	0.559	0.315	0.213	0.213
Arg R	0.459	0.267	0.213	7.869	1.562	0.669	0.779	0.155	0.175	0.175
Asn N	0.519	0.279	0.242	2.183	1.582	0.868	0.471	0.341	0.197	0.197
Asp D	0.348	0.208	0.286	2.673	1.445	0.861	0.537	0.29	0.2	0.2
Cys C	0.426	0.246	0.197	2.705	1.06	0.63	0.616	0.241	0.208	0.208
Gln Q	0.423	0.303	0.204	4.822	1.291	0.626	0.715	0.192	0.188	0.188
Glu E	0.402	0.209	0.294	4.264	1.297	0.867	0.663	0.202	0.191	0.191
Gly G	0.299	0.282	0.081	2.34	0.746	0.153	0.722	0.23	0.231	0.231
His H	0.467	0.193	0.223	4.12	1.501	0.932	0.629	0.229	0.188	0.188
Ile I	0.41	0.285	0.315	3.217	1.666	1.024	0.545	0.282	0.183	0.183
Leu L	0.465	0.367	0.291	3.255	1.701	0.949	0.551	0.288	0.183	0.183
Lys K	0.478	0.201	0.275	6.51	1.232	0.858	0.757	0.143	0.179	0.179
Met M	0.431	0.2	0.292	5.147	1.365	0.876	0.697	0.185	0.188	0.188
Phe F	0.408	0.25	0.221	6.751	1.36	0.695	0.767	0.154	0.181	0.181
Pro P	0.44	0.315	0.205	2.314	1.318	0.745	0.529	0.301	0.197	0.197
Ser S	0.319	0.291	0.249	2.089	1.251	0.73	0.513	0.307	0.208	0.208
Thr T	0.381	0.207	0.218	2.908	1.392	0.626	0.59	0.283	0.197	0.197
Trp W	0.436	0.27	0.172	7.065	2.311	0.749	0.698	0.228	0.174	0.174
Tyr Y	0.4	0.247	0.225	6.537	1.826	0.737	0.718	0.201	0.179	0.179
Val V	0.317	0.31	0.303	3.103	1.397	0.77	0.589	0.265	0.191	0.191
	41	42	43	44	45	46	47	48	49	50
Ala A	0.213	0.442	0.489	0.294	1.716	1.025	0.371	0.551	0.33	0.213
Arg R	0.175	0.554	0.411	0.419	7.174	1.305	0.507	0.798	0.145	0.175
Asn N	0.197	0.625	0.512	0.408	1.952	1.226	0.744	0.498	0.313	0.197
Asp D	0.2	0.482	0.371	0.488	2.265	1.164	0.686	0.551	0.283	0.2
Cys C	0.208	0.457	0.406	0.415	2.865	0.841	0.438	0.691	0.203	0.208
Gln Q	0.188	0.531	0.511	0.37	4.26	1.003	0.512	0.738	0.174	0.188

TABLE S-IV. Continued

Amino acid	41	42	43	44	45	46	47	48	49	50
Glu E	0.191	0.515	0.367	0.535	3.69	1.02	0.668	0.686	0.19	0.191
Gly G	0.231	0.474	0.469	0.167	1.929	0.587	0.123	0.731	0.222	0.231
His H	0.188	0.505	0.364	0.389	3.907	1.131	0.751	0.675	0.195	0.188
Ile I	0.183	0.505	0.447	0.555	2.885	1.376	0.818	0.568	0.271	0.183
Leu L	0.183	0.604	0.563	0.531	2.91	1.408	0.752	0.574	0.278	0.183
Lys K	0.179	0.571	0.364	0.505	5.917	0.948	0.676	0.785	0.126	0.179
Met M	0.188	0.487	0.384	0.445	4.922	0.967	0.74	0.742	0.146	0.188
Phe F	0.181	0.523	0.44	0.468	5.979	1.107	0.518	0.786	0.146	0.181
Pro P	0.197	0.557	0.528	0.397	2.005	1.081	0.577	0.547	0.295	0.197
Ser S	0.208	0.462	0.445	0.457	1.775	1.05	0.578	0.522	0.308	0.208
Thr T	0.197	0.478	0.353	0.431	2.614	1.1	0.477	0.624	0.263	0.197
Trp W	0.174	0.566	0.467	0.314	6.333	1.859	0.566	0.723	0.212	0.174
Tyr Y	0.179	0.551	0.446	0.385	5.555	1.433	0.573	0.735	0.189	0.179
Val V	0.191	0.45	0.499	0.49	2.638	1.167	0.63	0.595	0.263	0.191
	51	52	53	54	55	56	57	58	59	60
Ala A	0.213	0.213	0.302	0.358	0.182	2.032	0.955	0.268	0.624	0.293
Arg R	0.175	0.175	0.459	0.277	0.243	7.922	1.51	0.456	0.801	0.153
Asn N	0.197	0.197	0.496	0.296	0.3	2.73	1.628	0.336	0.582	0.347
Asp D	0.2	0.2	0.344	0.235	0.311	2.993	0.967	0.763	0.634	0.205
Cys C	0.208	0.208	0.531	0.241	0.201	3.181	0.818	0.361	0.73	0.188
Gln Q	0.188	0.188	0.415	0.309	0.247	5.576	1.22	0.243	0.792	0.173
Glu E	0.191	0.191	0.385	0.226	0.317	5.339	0.946	0.708	0.763	0.135
Gly G	0.231	0.231	0.324	0.296	0.11	1.891	0.779	0.027	0.701	0.289
His H	0.188	0.188	0.461	0.208	0.262	4.151	1.242	0.523	0.702	0.21
Ile I	0.183	0.183	0.415	0.313	0.339	3.509	1.111	0.688	0.661	0.209
Leu L	0.183	0.183	0.478	0.389	0.332	3.183	1.267	0.605	0.63	0.251
Lys K	0.179	0.179	0.472	0.216	0.31	7.233	0.975	0.565	0.824	0.111
Met M	0.188	0.188	0.451	0.198	0.325	4.939	0.995	0.68	0.747	0.15
Phe F	0.181	0.181	0.41	0.284	0.245	7.237	0.828	0.365	0.858	0.098
Pro P	0.197	0.197	0.423	0.356	0.254	3.391	0.846	0.222	0.761	0.19
Ser S	0.208	0.208	0.334	0.309	0.286	2.074	0.985	0.465	0.589	0.28
Thr T	0.197	0.197	0.392	0.225	0.25	2.992	1.177	0.438	0.649	0.255
Trp W	0.174	0.174	0.454	0.3	0.171	6.281	1.403	0.769	0.743	0.166
Tyr Y	0.179	0.179	0.4	0.277	0.229	6.809	1.199	0.593	0.792	0.139
Val V	0.191	0.191	0.327	0.353	0.34	3.29	1.063	0.286	0.709	0.229
	61	62	63	64	65	66	67	68	69	70
Ala A	0.279	0.279	0.279	0.445	0.323	0.081	3.773	3.017	3.001	3.778
Arg R	0.218	0.218	0.218	0.576	0.457	0.204	10.02	9.309	8.918	10.1
Asn N	0.24	0.24	0.24	1.063	0.483	0.046	4.564	4.212	3.869	4.633
Asp D	0.24	0.24	0.24	0.619	0.185	0.388	4.94	4.299	4.024	4.979
Cys C	0.263	0.263	0.263	0.639	0.299	0.122	4.354	4.48	3.863	4.395
Gln Q	0.231	0.231	0.231	0.872	0.23	0.054	6.61	6.378	5.751	6.739
Glu E	0.231	0.231	0.231	0.85	0.28	0.277	6.242	6.226	5.369	6.428
Gly G	0.301	0.301	0.218	0.6	0.196	0.082	8.686	7.868	7.487	8.806
His H	0.224	0.224	0.224	0.585	0.268	0.128	6.565	5.662	5.687	6.553

TABLE S-IV. Continued

Amino acid	61	62	63	64	65	66	67	68	69	70
Ile I	0.24	0.24	0.24	0.646	0.207	0.183	5.877	4.969	4.952	5.907
Leu L	0.24	0.24	0.24	0.638	0.239	0.215	5.901	4.802	4.925	5.904
Lys K	0.231	0.231	0.231	0.723	0.225	0.197	8.47	8.103	7.496	8.599
Met M	0.24	0.24	0.24	0.508	0.196	0.214	7.396	6.398	6.492	7.389
Phe F	0.218	0.251	0.218	0.6	0.196	0.082	8.686	7.868	7.487	8.806
Pro P	0.25	0.25	0.25	1.18	0.215	0.026	4.292	3.935	3.598	4.377
Ser S	0.263	0.263	0.263	0.559	0.298	0.173	4.088	3.261	3.279	4.07
Thr T	0.25	0.25	0.25	0.552	0.251	0.169	4.904	4.188	4.082	4.925
Trp W	0.204	0.204	0.204	0.469	0.177	0.334	10.17	8.214	8.527	10.13
Tyr Y	0.213	0.213	0.213	0.654	0.264	0.169	8.968	7.953	7.438	9.101
Val V	0.25	0.25	0.25	0.522	0.273	0.033	5.241	4.316	4.296	5.269
	71	72	73	74	75	76	77	78	79	80
Ala A	3.112	3.255	4.108	2.429	2.563	4.085	2.776	2.741	0.213	0.213
Arg R	8.986	9.888	18.34	13.89	13.12	18.61	13.66	16.27	0.175	0.175
Asn N	3.922	4.694	6.578	4.961	4.56	6.721	4.757	5.909	0.209	0.197
Asp D	4.115	4.723	7.381	4.966	4.697	7.409	4.988	5.917	0.2	0.2
Cys C	4.144	4.36	5.231	3.784	3.668	5.238	4.032	4.047	0.208	0.208
Gln Q	5.774	7.039	9.662	7.393	6.764	10.05	6.965	8.455	0.198	0.188
Glu E	5.378	6.993	10	7.798	6.664	10.35	6.912	9.504	0.191	0.191
Gly G	7.604	8.431	14.84	1.309	1.318	2.218	1.443	1.545	0.231	0.231
His H	5.788	5.917	11.48	7.296	7.693	11.42	8.199	7.979	0.188	0.188
Ile I	5.079	5.309	10.31	6.528	7.005	10.36	7.454	7.079	0.183	0.192
Leu L	5.069	5.054	10.24	6.258	6.864	10.24	7.342	6.723	0.192	0.183
Lys K	7.541	8.774	14.35	10.27	9.77	14.66	10.25	11.7	0.179	0.179
Met M	6.63	6.615	12.67	7.88	8.736	12.73	9.12	8.952	0.188	0.188
Phe F	7.604	8.431	14.84	8.522	9.581	14.82	10.29	8.94	0.189	0.181
Pro P	3.664	4.459	5.647	3.544	3.706	5.756	3.949	3.808	0.197	0.197
Ser S	3.403	3.524	5.12	3.046	3.227	5.052	3.497	3.465	0.208	0.208
Thr T	4.191	4.607	6.698	4.459	4.356	6.737	4.648	5.348	0.197	0.197
Trp W	8.758	8.453	23.48	14.09	15.41	23.36	16.41	14.72	0.181	0.174
Tyr Y	7.561	8.6	17.8	11.47	11.28	18.1	11.97	12.91	0.179	0.187
Val V	4.434	4.639	7.813	4.453	5.066	7.796	5.473	4.742	0.191	0.191
	81	82	83	84	85	86	87	88	89	90
Ala A	0.279	0.326	0.412	0.336	0.338	0.327	0.436	0.414	0.242	0.258
Arg R	0.218	0.668	0.715	0.707	0.669	0.698	0.702	0.451	0.344	0.313
Asn N	0.24	0.201	0.348	0.269	0.219	0.247	0.393	0.512	0.41	0.347
Asp D	0.24	0.283	0.437	0.348	0.305	0.326	0.45	0.444	0.318	0.28
Cys C	0.263	0.405	0.655	0.504	0.423	0.537	0.594	0.427	0.353	0.29
Gln Q	0.231	0.576	0.665	0.616	0.573	0.607	0.688	0.455	0.322	0.31
Glu E	0.231	0.475	0.629	0.551	0.495	0.529	0.645	0.45	0.378	0.301
Gly G	0.301	0.588	0.58	0.598	0.584	0.596	0.552	0.377	0.227	0.221
His H	0.224	0.443	0.556	0.532	0.443	0.512	0.552	0.423	0.291	0.294
Ile I	0.24	0.307	0.45	0.367	0.317	0.352	0.492	0.5	0.315	0.337
Leu L	0.24	0.326	0.413	0.371	0.327	0.361	0.445	0.562	0.34	0.375
Lys K	0.231	0.631	0.727	0.691	0.636	0.677	0.737	0.469	0.337	0.318

TABLE S-IV. Continued

Amino acid	81	82	83	84	85	86	87	88	89	90
Met M	0.24	0.55	0.649	0.614	0.545	0.614	0.62	0.429	0.277	0.308
Phe F	0.228	0.636	0.765	0.695	0.65	0.679	0.788	0.475	0.275	0.293
Pro P	0.25	0.259	0.547	0.357	0.293	0.321	0.641	0.5	0.354	0.32
Ser S	0.263	0.262	0.353	0.29	0.27	0.283	0.383	0.459	0.286	0.286
Thr T	0.25	0.382	0.468	0.446	0.386	0.435	0.474	0.42	0.264	0.269
Trp W	0.204	0.546	0.605	0.591	0.547	0.585	0.615	0.443	0.295	0.292
Tyr Y	0.213	0.569	0.672	0.616	0.577	0.602	0.688	0.44	0.322	0.291
Val V	0.25	0.368	0.519	0.408	0.383	0.392	0.564	0.495	0.24	0.31
	91	92	93	94	95	96	97	98	99	
Ala A	0.408	0.281	0.283	9.096	5.923	6.134	9.058	6.541	6.516	
Arg R	0.461	0.326	0.412	36.48	27.45	26.35	36.93	27.4	31.61	
Asn N	0.515	0.364	0.531	14.1	10.6	10.06	14.35	10.46	12.1	
Asp D	0.447	0.297	0.397	15.66	10.96	10.35	15.71	10.91	12.85	
Cys C	0.426	0.325	0.353	11.41	9.053	8.477	11.44	9.23	9.347	
Gln Q	0.47	0.323	0.385	20.06	15.46	14.51	20.69	14.93	17.15	
Glu E	0.472	0.309	0.469	20.86	16.73	14.36	21.57	14.81	20.08	
Gly G	0.37	0.243	0.284	5.777	3.882	3.961	5.725	4.222	4.281	
His H	0.419	0.31	0.327	23.92	15.41	16.31	23.74	17.3	16.59	
Ile I	0.503	0.356	0.345	21.66	13.99	14.89	21.76	15.78	15.07	
Leu L	0.566	0.4	0.364	21.41	13.3	14.53	21.39	15.49	14.22	
Lys K	0.48	0.333	0.382	29.54	21.75	20.7	30.14	21.59	24.46	
Met M	0.438	0.325	0.306	26.14	16.97	18.51	26.28	19.27	18.91	
Phe F	0.477	0.313	0.293	30.01	18.63	20.07	30.01	21.32	19.56	
Pro P	0.494	0.344	0.474	12.2	8.25	8.402	12.41	8.864	8.903	
Ser S	0.455	0.31	0.344	11.16	7.12	7.448	11.03	7.977	7.938	
Thr T	0.421	0.289	0.314	14.12	9.851	9.661	14.2	10.21	11.5	
Trp W	0.449	0.308	0.327	45.57	28.35	30.08	45.72	31.83	29.95	
Tyr Y	0.461	0.302	0.363	35.28	23.58	22.88	36.01	24.09	26.35	
Val V	0.48	0.34	0.276	16.44	9.926	11.07	16.4	11.85	10.38	

TABLE S-V. The values of 74 geometrical descriptors of 20 amino acid

Amino acid	1	2	3	4	5	6	7	8	9	10
Ala A	209.85	4.831	34.08	32.285	77.895	11.984	6.024	3.62	1.808	2.969
Arg R	1340.1	6.439	103.07	103.08	313.6	24.123	14.186	7.503	4.49	5.056
Asn N	400.83	5.79	54.002	47.156	129.08	15.186	10.862	5.677	2.442	2.965
Asp D	365.24	5.311	47.078	45.655	118.97	14.871	10.904	5.738	2.393	3.453
Cys C	261.67	4.892	37.263	38.381	93.191	13.313	8.069	4.449	2.285	3.051
Gln Q	651.02	5.934	68.797	65.102	185.19	18.519	11.626	6.239	3.259	4.285
Glu E	573	5.786	62.691	60.316	168.64	17.751	11.777	1.297	3.139	4.257
Gly G	112.57	3.913	21.436	22.514	45.863	9.173	4.721	3.077	1.448	2.741
His H	656.44	3.238	67.416	65.644	189.66	18.966	13.618	7.28	3.07	4.027
Ile I	760.03	6.831	83.806	69.094	222.15	20.196	10.101	5.299	3.018	3.586
Leu L	762.04	6.797	83.553	69.278	220.61	20.055	10.026	5.28	2.967	3.356
Lys K	1054.4	6.418	42.492	87.866	269.59	22.466	10.965	5.971	4.025	4.804
Met M	686.07	5.621	66.375	68.607	191.27	19.127	10.824	5.91	3.264	4.536

TABLE S-V. Continued

Amino acid	1	2	3	4	5	6	7	8	9	10
Phe F	981.29	3.314	83.241	85.3	256.27	22.284	13.931	7.395	3.883	4.898
Pro P	388.62	3.401	55.181	45.72	136.69	16.081	9.051	5.088	2.361	2.981
Ser S	254.55	5.031	38.337	36.357	90.225	12.889	7.715	4.33	1.949	3.066
Thr T	411.16	5.71	53.554	48.372	133.74	15.734	9.08	4.865	2.434	3.35
Trp W	1476.1	2.543	107.02	109.34	356.38	26.399	19.432	9.707	4.299	4.982
Tyr Y	1092.5	3.374	88.813	91.041	276.38	23.031	15.917	8.158	3.987	4.85
Val V	532.76	6.231	65.34	56.08	167.45	17.627	8.74	4.725	2.612	3.685
	11	12	13	14	15	16	17	18	19	20
Ala A	0.478	0.987	0.617	0.199	0.93	0.598	2.1	24.192	0	0
Arg R	0.441	0.998	0.798	0.519	0.937	0.967	5.6	50.44	0	0
Asn N	0.418	0.984	0.403	0.161	0.905	0.349	1.7	32.509	0	0
Asp D	0.465	0.969	0.476	0.194	0.938	0.499	2.8	30.58	0	0
Cys C	0.467	0.996	0.56	0.434	0.963	0.77	4.9	26.962	0	0
Gln Q	0.463	0.998	0.7	0.455	0.936	0.824	4.4	38.085	0	0
Glu E	0.473	0.991	0.587	0.396	0.943	0.818	5.2	36.537	0	0
Gly G	0.524	1	0.845	0.381	0.925	0.601	2.7	18.495	0	0
His H	0.449	0.991	0.56	0.317	0.956	0.73	3.5	39.695	0	0
Ile I	0.404	0.974	0.478	0.207	0.927	0.656	2.8	42.26	0	0
Leu L	0.391	0.979	0.512	0.186	0.921	0.488	2.3	42.613	0	0
Lys K	0.447	0.996	0.695	0.531	0.945	0.807	7.2	16.121	0	0
Met M	0.476	0.991	0.649	0.422	0.968	0.924	5.8	39.621	0	0
Phe F	0.461	0.998	0.758	0.587	0.977	0.766	8	46.298	0.969	1.4
Pro P	0.419	0.993	0.458	0.313	0.957	0.566	3.2	33.063	0	0
Ser S	0.463	0.968	0.456	0.14	0.929	0.591	2	26.024	0	0
Thr T	0.444	0.988	0.618	0.237	0.935	0.681	2.5	31.864	0	0
Trp W	0.43	0.994	0.79	0.373	0.984	0.689	4	55.378	0.908	1.42
Tyr Y	0.45	0.996	0.76	0.456	0.967	0.933	5.4	48.317	0.969	1.4
Val V	0.44	0.991	0.551	0.283	0.936	0.863	3	36.127	0	0
	21	22	23	24	25	26	27	28	29	30
Ala A	0	0	9.52	8.94	15.818	20.016	3.203	9.126	13.683	18.066
Arg R	0	0	5.07	25.892	115.41	128.82	0.889	24.416	106.95	118.11
Asn N	0	0	6.65	25.521	30.948	40.219	3.075	17.774	24.768	29.55
Asp D	0	0	5.398	17.96	37.146	40.194	1.729	15.481	25.878	29.735
Cys C	0	0	4.95	11.063	37.658	41.691	6.063	10.297	23.655	27.065
Gln Q	0	0	3.755	21.315	60.137	73.784	2.057	16.368	51.478	57.674
Glu E	0	0	3.762	19.35	65.428	67.79	0.806	16.788	45.949	49.278
Gly G	0	0	7.527	4.455	11.594	15.499	2.694	3.55	10.198	12.634
His H	0	0	5.286	22.59	57.811	65.904	6.272	21.096	54.274	58.816
Ile I	0	0	11.813	20.65	42.088	45.836	3.066	24.279	42.299	48.339
Leu L	0	0	8.735	21.601	37.431	45.877	1.356	23.985	41.33	48.979
Lys K	0	0	10.49	18.551	87.172	91.572	4.062	19.465	82.997	86.389
Met M	0	0	4.673	21.268	68.445	69.282	3.065	18.711	61.299	64.021
Phe F	0.992	5.813	5.748	17.032	97.253	102.19	5.382	21.317	90.361	98.039
Pro P	0	0	9.057	11.857	29.494	34.106	3.304	14.529	24.259	28.91
Ser S	0	0	6.939	12.891	20.124	24.059	4.034	11.395	16.587	20.034

TABLE S-V. Continued

Amino acid	21	22	23	24	25	26	27	28	29	30
Thr T	0	0	10.611	15.099	31.106	36.883	4.392	13.513	26.964	32.336
Trp W	0.977	5.447	3.784	37.626	114.21	127.55	6.332	39.537	115	135.38
Tyr Y	0.992	5.813	4.54	27.607	102.19	110.19	3.83	27.763	88.396	99.797
Val V	0	0	8.755	13.959	33.239	37.012	3.116	16.986	31.98	37.008
	31	32	33	34	35	36	37	38	39	40
Ala A	0.37	22.353	34.687	44.253	0.069	10.115	14.97	19.721	0	6.33
Arg R	0.245	59.118	226.07	249.74	0.061	26.72	113.22	124.96	21.46	43.73
Asn N	0.322	43.754	54.465	67.051	0.192	19.139	25.986	30.633	3.03	19.05
Asp D	0.302	39.354	60.293	70.259	0.085	16.859	26.877	31.219	0	13.84
Cys C	0.35	24.471	48.318	54.515	0.957	11.792	28.741	32.782	0	6.22
Gln Q	0.148	39.923	112.7	126.65	0.131	17.421	54.397	60.074	5.64	22.29
Glu E	0.206	43.192	102.33	110.89	0.087	18.362	47.383	51.172	0	16.7
Gly G	0.336	9.463	26.238	32.496	0.087	3.87	11.154	13.68	0	6.37
His H	0.243	50.194	104.09	115.84	0.385	23.064	56.819	61.426	10.69	24.35
Ile I	0.414	59.397	93.552	107.65	0.034	27.41	46.235	53.148	0	6.31
Leu L	0.357	58.499	92.662	109.3	0.106	26.961	45.72	53.881	0	6.39
Lys K	0.363	50.531	177.96	187.03	0.162	21.915	88.958	92.572	6.86	20.57
Met M	0.405	45.462	121.93	131.86	0.634	21.372	70.101	72.539	0	6.33
Phe F	0.439	47.633	172.58	187.93	0.479	23.922	95.659	104.3	0	6.22
Pro P	0.385	35.72	52.9	62.738	0.144	16.524	25.718	30.738	0	6.22
Ser S	0.329	29.217	41.531	49.199	0.243	12.462	18.014	21.615	0	9.15
Thr T	0.411	35.555	62.266	75.733	0.152	14.887	29.099	34.93	0	9.91
Trp W	0.365	83.527	213.16	255.71	0.479	42.999	122.31	145.21	5.95	15.99
Tyr Y	0.27	62.87	178.21	204.88	0.344	30.514	92.946	106.03	0	14.09
Val V	0.311	41.562	74.292	86.206	0.149	19.263	35.049	40.779	0	6.24
	41	42	43	44	45	46	47	48	49	50
Ala A	0	0	0	0	0	0	0	0	0	0
Arg R	0	0	0	0	0	0	0	0	0	0
Asn N	0	0	0	0	0	0	7.32	0	0	0
Asp D	0	0	0	0	0	0	16.5	0	0	0
Cys C	3.33	0	0	0	0	0	0	9.71	0	0
Gln Q	0	0	0	0	0	0	11.79	0	0	0
Glu E	0	0	0	0	0	0	22.57	0	0	0
Gly G	0	0	0	0	0	0	0	0	0	0
His H	0	0	0	0	0	0	0	0	0	0
Ile I	0	0	0	0	0	0	0	0	0	0
Leu L	0	0	0	0	0	0	0	0	0	0
Lys K	0	0	0	0	0	0	0	0	0	0
Met M	4.8	0	0	0	0	0	0	10.42	0	0
Phe F	0	0	0	0	0	0	0	0	0	0
Pro P	0	0	0	0	0	0	0	0	0	0
Ser S	0	0	0	0	0	0	6.91	0	0	0
Thr T	0	0	0	0	0	0	6.66	0	0	0
Trp W	0	0	0	0	0	0	0	0	0	0
Tyr Y	0	0	0	0	0	0	0	0	0	0

TABLE S-V. Continued

Amino acid	41	42	43	44	45	46	47	48	49	50
Val V	0	0	0	0	0	0	0	0	0	0
	51	52	53	54	55	56	57	58	59	60
Ala A	0	0	0	0	0	0	0	0	0	0
Arg R	0	0	0	0	0	0	0	0	0	0
Asn N	0	0	0	0	0	0	0	0	0	0
Asp D	0	0	0	0	0	0	0	0	0	0
Cys C	0	0	0	0	0	0	0	0	0	0
Gln Q	0	0	0	0	0	0	0	0	0	0
Glu E	0	0	0	0	0	0	0	0	0	0
Gly G	0	0	0	0	0	0	0	0	0	0
His H	0	0	0	0	0	0	0	0	0	0
Ile I	0	0	0	0	0	0	0	0	0	0
Leu L	0	0	0	0	0	0	0	0	0	0
Lys K	0	0	0	0	0	0	0	0	0	0
Met M	0	0	0	0	0	0	0	0	0	0
Phe F	0	0	0	0	0	0	0	0	0	0
Pro P	0	0	0	0	0	0	0	0	0	0
Ser S	0	0	0	0	0	0	0	0	0	0
Thr T	0	0	0	0	0	0	0	0	0	0
Trp W	0	0	0	0	0	0	0	0	0	0
Tyr Y	0	0	0	0	0	0	0	0	0	0
Val V	0	0	0	0	0	0	0	0	0	0
	61	62	63	64	65	66	67	68	69	70
Ala A	0	0	0	0	0	0	0	0	0	0
Arg R	0	0	0	0	0	0	0	0	0	0
Asn N	0	0	0	0	0	0	0	0	0	0
Asp D	0	0	0	0	0	0	0	0	0	0
Cys C	0	0	0	0	0	0	0	0	0	0
Gln Q	0	0	0	0	0	0	0	0	0	0
Glu E	0	0	0	0	0	0	0	0	0	0
Gly G	0	0	0	0	0	0	0	0	0	0
His H	0	0	0	0	0	0	0	0	0	0
Ile I	0	0	0	0	0	0	0	0	0	0
Leu L	0	0	0	0	0	0	0	0	0	0
Lys K	0	0	0	0	0	0	0	0	0	0
Met M	0	0	0	0	0	0	0	0	0	0
Phe F	0	0	0	0	0	0	0	0	0	0
Pro P	0	0	0	0	0	0	0	0	0	0
Ser S	0	0	0	0	0	0	0	0	0	0
Thr T	0	0	0	0	0	0	0	0	0	0
Trp W	0	0	0	0	0	0	0	0	0	0
Tyr Y	0	0	0	0	0	0	0	0	0	0
Val V	0	0	0	0	0	0	0	0	0	0
	71	72	73	74						
Ala A	0	0	0	0						

TABLE S-V. Continued

Amino acid	61	62	63	64	65	66	67	68	69	70
Asn N	0	0	0	0						
Asp D	0	0	0	0						
Cys C	0	0	0	0						
Gln Q	0	0	0	0						
Glu E	0	0	0	0						
Gly G	0	0	0	0						
His H	0	0	0	0						
Ile I	0	0	0	0						
Leu L	0	0	0	0						
Lys K	0	0	0	0						
Met M	0	0	0	0						
Phe F	0	0	0	0						
Pro P	0	0	0	0						
Ser S	0	0	0	0						
Thr T	0	0	0	0						
Trp W	0	0	0	0						
Tyr Y	0	0	0	0						
Val V	0	0	0	0						

TABLE S-VI. The values of 160 3D-MORSE descriptors of 20 amino acid

Amino acid	1	2	3	4	5	6	7	8	9	10
Ala A	78	16.945	–0.943	–0.031	–1.478	–1.004	0.556	0.387	0.569	0.135
Arg R	325	27.716	–3.756	3.192	–4.551	–1.631	1.132	0.944	0.563	–0.323
Asn N	136	18.987	–0.236	0.051	–1.883	–1.672	1.042	0.623	0.238	0.419
Asp D	120	16.214	–0.831	0.233	–1.719	–1.436	1.27	0.206	0.373	0.133
Cys C	91	15.743	–1.442	0.12	–1.55	–1.034	0.743	0.473	0.406	0.064
Gln Q	190	21.665	–2.455	0.811	–3.009	–1.294	0.416	1.492	0.12	0.456
Glu E	171	19.606	–2.639	0.744	–2.419	–1.696	1.199	0.408	0.266	0.343
Gly G	45	13.001	–0.719	–0.047	–0.908	–1.159	0.752	0.228	0.255	0.179
His H	190	15.678	–0.496	0.655	–2.74	–1.134	1.222	1.82	–0.219	0.035
Ile I	231	20.727	–2.519	1.631	–3.19	–1.295	0.711	0.579	1.166	0.12
Leu L	231	20.07	–1.863	0.765	–2.925	–1.92	0.88	0.972	1.144	–0.069
Lys K	276	27.074	–5.673	2.853	–3.553	–1.577	1.072	0.899	0.675	0.229
Met M	190	19.917	–3.055	0.82	–2.259	–1.028	0.705	1.109	0.908	0.099
Phe F	253	20.049	–2.17	0.998	–3.235	–2.403	2.762	0.172	0.235	–0.258
Pro P	136	22.565	–2.379	0.157	–1.578	–1.901	1.621	0.287	0.318	0.311
Ser S	91	16.798	–0.475	–0.093	–1.333	–1.4	0.777	0.337	0.527	–0.029
Thr T	136	19.421	–1.768	0.69	–2.436	–1.11	0.444	0.762	0.742	–0.145
Trp W	351	20.721	–1.789	0.169	–2.74	–2.385	3.609	0.628	–0.417	0.027
Tyr Y	276	19.38	–1.418	1.528	–2.606	–2.614	2.676	0.186	0.532	–0.611
Val V	171	20.914	–1.088	0.131	–2.712	–0.97	0.228	0.852	0.964	0.136
	11	12	13	14	15	16	17	18	19	20
Ala A	–0.648	–0.414	0.161	0.308	0.316	–0.181	–0.036	–0.513	0.338	0.172
Arg R	–0.338	–0.857	–0.429	1.284	0.218	0.002	–0.243	–0.905	0.301	0.585
Asn N	–0.685	–0.413	–0.266	0.663	0.321	–0.112	–0.189	–0.246	–0.073	0.198

TABLE S-VI. Continued

Amino acid	11	12	13	14	15	16	17	18	19	20
Asp D	–0.281	–0.573	–0.163	0.666	–0.062	0.143	–0.081	–0.321	0.114	0.056
Cys C	–0.449	–0.228	–0.089	0.265	0.349	–0.136	–0.051	–0.345	0.042	0.33
Gln Q	–0.26	–1.281	0.312	0.931	0.097	–0.186	–0.165	–0.633	0.478	0.131
Glu E	–0.334	–0.934	0.147	0.661	–0.068	–0.022	–0.143	–0.529	0.4	0.189
Gly G	–0.353	–0.299	–0.044	0.255	0.207	–0.037	–0.008	–0.323	–0.035	0.187
His H	0.155	–1.109	0.309	0.076	0.535	0.279	–0.107	–0.922	0.345	0.433
Ile I	–0.986	–1.047	0.756	0.847	0.054	0.044	–0.391	–0.958	0.985	0.213
Leu L	–0.849	–1.531	0.89	1.038	–0.148	0.216	–0.842	–0.704	0.98	0.428
Lys K	–0.816	–1.444	0.291	1.297	–0.252	0.237	–0.395	–0.937	0.621	0.501
Met M	–1.089	–0.576	0.399	0.661	0.275	–0.153	–0.302	–0.616	0.649	0.571
Phe F	0.165	–1.165	0.249	0.817	0.097	0.327	–0.619	–0.826	0.325	1.094
Pro P	–0.685	–0.959	0.731	0.359	0.219	0.034	–0.756	–0.146	0.466	0.443
Ser S	–0.107	–0.572	–0.255	0.553	0.228	–0.008	–0.084	–0.348	–0.02	0.287
Thr T	–0.221	–0.849	0.11	0.541	0.253	0.079	–0.364	–0.391	0.295	0.227
Trp W	0.455	–1.198	–0.248	0.726	0.263	0.634	–0.698	–1.053	0.82	0.645
Tyr Y	0.275	–1.007	–0.058	0.564	0.375	0.093	–0.364	–0.77	0.221	0.617
Val V	–0.802	–0.934	0.47	0.768	0.108	–0.057	–0.47	–0.562	0.601	0.445
	21	22	23	24	25	26	27	28	29	30
Ala A	–0.106	0.158	–0.18	–0.056	–0.032	0.087	0.17	–0.14	–0.056	0.019
Arg R	–0.041	0.191	–0.3	–0.006	–0.018	0.106	0.441	–0.078	–0.061	–0.251
Asn N	0.167	0.122	–0.16	–0.126	–0.171	0.158	0.157	0.07	–0.14	–0.052
Asp D	0.169	0.051	0.037	–0.212	–0.024	0.079	0.009	0.089	–0.032	–0.015
Cys C	–0.105	0.208	–0.199	–0.128	0	0.104	0.045	–0.028	–0.026	0.055
Gln Q	–0.079	0.153	–0.053	–0.28	–0.091	0.215	0.137	0.012	0.133	–0.006
Glu E	–0.278	0.309	–0.242	–0.152	–0.04	0.06	0.097	0.017	–0.143	–0.069
Gly G	0.01	0.009	0.052	–0.177	0.01	0.017	0.058	0.019	0.063	–0.131
His H	–0.101	0.24	–0.316	–0.171	0.202	0.229	–0.086	–0.003	–0.11	–0.053
Ile I	–0.065	–0.177	–0.132	–0.312	0.281	0.191	0.193	–0.225	–0.357	0.188
Leu L	–0.29	–0.072	–0.041	–0.204	0.231	0.245	–0.002	–0.118	–0.465	0.313
Lys K	–0.476	0.488	–0.665	0.136	–0.034	0.41	0.127	–0.031	–0.326	–0.205
Met M	–0.282	0.198	–0.225	–0.215	0.21	0.065	0.186	–0.281	–0.165	0.098
Phe F	–0.803	0.468	–0.769	–0.053	0.631	–0.108	–0.057	–0.214	0.052	–0.062
Pro P	–0.579	0.403	–0.356	–0.14	0.193	0.138	0.152	–0.155	–0.381	0.185
Ser S	–0.036	0.019	0.151	–0.32	0.058	0.012	0.107	–0.071	0.093	–0.042
Thr T	–0.159	0.085	–0.151	0.008	–0.082	–0.203	–0.005	0.008	–0.101	–0.033
Trp W	0.069	–0.088	–0.639	–0.148	0.437	–0.003	–0.029	–0.097	–0.004	–0.151
Tyr Y	–0.25	0.369	–0.529	–0.211	0.462	0.109	–0.302	–0.086	0.136	–0.131
Val V	–0.374	0.098	–0.191	–0.066	–0.131	0.379	0.067	–0.153	–0.277	0.107
	31	32	33	34	35	36	37	38	39	40
Ala A	0.017	0.042	23.545	7.14	–0.808	–0.249	–0.826	–0.015	0.883	–0.365
Arg R	0.378	–0.057	97.692	11.139	–1.326	–0.784	–1.636	0.389	1.813	–0.635
Asn N	–0.015	–0.061	54.475	9.434	–0.585	–0.181	–1.776	0.293	1.137	–0.209
Asp D	–0.085	0.05	55.173	9.337	–0.665	–0.398	–1.395	–0.026	1.495	–0.465
Cys C	–0.115	0.003	43.358	6.797	–1.352	–0.042	–0.568	0.035	0.507	–0.348
Gln Q	–0.085	0.107	67.503	8.786	–1.697	–0.176	–1.57	0.264	1.427	–0.42

TABLE S-VI. Continued

Amino acid	31	32	33	34	35	36	37	38	39	40
Glu E	0.163	–0.035	68.296	8.636	–2.156	–0.486	–1.398	0.147	1.438	–0.559
Gly G	0.003	–0.013	16.067	5.861	0.466	–0.285	–0.684	–0.102	0.764	–0.14
His H	0.106	–0.075	76.626	9.474	–0.185	–1.376	–1.243	–0.281	1.841	–0.292
Ile I	0.284	–0.01	54.168	8.141	–0.794	–0.413	–1.316	0.473	0.979	0.708
Leu L	0.397	–0.193	54.168	7.653	–0.22	–0.778	–0.991	0.194	1.066	–0.736
Lys K	0.445	–0.107	67.928	9.252	–1.038	–0.535	–1.109	0.488	1.015	–0.681
Met M	0.244	–0.068	68.647	7.665	–2.776	–0.355	–0.545	0.599	0.292	–0.245
Phe F	0.267	–0.054	87.62	11.38	–1.573	–0.056	–2.002	0.54	1.834	0.974
Pro P	0.268	–0.099	40.976	9.56	–1.014	–0.807	–0.729	–0.057	1.377	–0.833
Ser S	–0.08	0.036	33.43	8.436	–0.884	–0.406	–0.965	0.056	1.061	–0.484
Thr T	0.158	–0.117	43.829	8.793	–1.035	–0.315	–1.103	0.12	1.157	–0.674
Trp W	–0.02	–0.015	135.935	11.872	–1.073	–1.115	–1.896	–0.197	2.571	–1.007
Tyr Y	0.159	–0.173	105.945	11.075	–1.371	–0.531	–2.208	0.448	2.385	–0.976
Val V	0.276	–0.051	42.596	8.325	–1.061	–0.131	–1.324	0.47	0.776	–0.548
	41	42	43	44	45	46	47	48	49	50
Ala A	–0.137	0.103	–0.051	0.077	–0.093	0.013	0.144	–0.151	0.028	–0.98
Arg R	–0.445	0.177	0.171	0.117	–0.517	0.261	0.377	–0.089	–0.097	–0.179
Asn N	–0.479	0.185	–0.14	0.323	–0.284	0.102	0.108	–0.198	0.037	0.135
Asp D	–0.533	0.318	–0.098	0.168	–0.104	–0.026	0.065	–0.068	0.026	0.029
Cys C	0.283	–0.145	–0.16	0.198	–0.014	–0.17	0.249	–0.124	–0.053	0.01
Gln Q	–0.231	0.132	–0.031	0.072	–0.202	0.283	0.057	–0.293	0.113	0.046
Glu E	–0.412	0.303	0.047	0.046	–0.055	0.056	0.043	–0.206	–0.025	0.117
Gly G	–0.28	0.115	–0.065	0.13	–0.047	–0.109	0.167	–0.099	0.027	–0.022
His H	–1.083	0.179	0.445	–0.01	–0.056	–0.984	0.247	0.296	–0.083	–0.254
Ile I	–1.083	0.179	0.445	–0.01	–0.056	–0.984	0.247	0.296	–0.083	–0.254
Leu L	0.063	0.134	0.07	–0.179	–0.218	0.434	–0.055	–0.083	–0.176	0.039
Lys K	–0.163	0.19	0.177	–0.365	–0.299	0.425	0.181	–0.056	–0.203	0.075
Met M	0.296	–0.141	–0.161	0.355	0.025	–0.366	0.188	0.017	–0.166	0.102
Phe F	–0.471	0.148	0.463	0.143	–0.505	–0.096	0.398	–0.021	–0.063	–0.284
Pro P	–0.18	0.389	–0.167	0.112	–0.31	0.234	0.232	–0.333	–0.035	0.108
Ser S	–0.22	0.164	–0.034	0.214	–0.356	0.155	0.071	–0.017	–0.048	–0.053
Thr T	–0.076	0.097	0.128	–0.089	–0.109	0.012	0.239	–0.171	–0.021	–0.018
Trp W	–1.183	0.272	0.479	0.438	–0.867	–0.327	0.577	0.315	–0.36	–0.337
Tyr Y	–0.587	0.157	0.348	0.347	–0.563	–0.316	0.636	–0.251	0.087	–0.378
Val V	–0.04	0.206	–0.046	–0.031	–0.236	0.34	0.005	–0.137	–0.076	0.015
	51	52	53	54	55	56	57	58	59	60
Ala A	0.082	–0.017	–0.111	0.097	–0.009	0.029	–0.059	–0.011	0.039	0.018
Arg R	0.172	–0.052	–0.059	0.033	0.083	0.063	–0.054	–0.1	0.162	–0.08
Asn N	–0.054	–0.014	–0.179	0.257	–0.062	0.022	–0.11	–0.008	0.104	0
Asp D	0.052	–0.098	–0.051	0.072	0.057	0.024	–0.124	0.024	–0.01	0.067
Cys C	0.125	–0.038	–0.152	0.151	–0.041	0.038	–0.064	0.027	–0.038	0.045
Gln Q	0.303	–0.28	–0.047	0.11	0.094	–0.064	–0.098	0.041	0.097	0.003
Glu E	0.11	–0.036	–0.172	0.117	–0.052	0.077	–0.167	0.038	0.014	0.088
Gly G	0.018	0.048	–0.098	0.04	0.016	0.03	–0.024	–0.059	0.029	0.038
His H	0.043	0.119	–0.021	0.05	–0.096	0.084	0.002	0.04	–0.053	0.023

TABLE S-VI. Continued

Amino acid	51	52	53	54	55	56	57	58	59	60
Ile I	0.043	0.119	–0.021	0.05	–0.096	0.084	0.002	0.04	–0.053	0.023
Leu L	0.324	–0.227	–0.092	0.092	0.093	–0.051	–0.099	0.032	0.119	–0.061
Lys K	0.193	–0.123	–0.222	0.206	0.016	0.085	–0.163	0.087	0.074	0.013
Met M	0.182	0.013	–0.214	0.314	–0.116	–0.048	–0.082	0.174	–0.057	–0.048
Phe F	0.152	0.4	–0.315	0.057	–0.142	0.141	0.139	–0.084	–0.099	0.009
Pro P	0.117	–0.02	–0.268	0.192	0.114	–0.15	–0.024	0.02	0.073	0.04
Ser S	0.103	–0.012	–0.045	–0.04	0.114	–0.06	0.06	–0.139	0.094	–0.014
Thr T	0.079	0.069	–0.2	0.108	–0.01	0.072	–0.098	0.012	–0.002	0.053
Trp W	0.19	0.248	0.032	–0.185	–0.286	0.254	0.048	–0.147	–0.157	–0.006
Tyr Y	0.003	0.325	–0.077	0.068	–0.207	0.069	0.161	–0.023	–0.169	0.047
Val V	0.222	–0.119	–0.151	0.144	0.009	0.045	–0.164	0.065	0.089	–0.047
	61	62	63	64	65	66	67	68	69	70
Ala A	–0.048	0.047	–0.021	–0.024	20.873	6.303	–0.446	–0.239	–0.614	0.113
Arg R	0.101	–0.041	0.057	0.045	92.957	10.834	–1.227	–0.069	–1.718	0.034
Asn N	–0.07	0.073	–0.048	–0.048	40.157	8.201	–0.456	–0.39	–0.877	–0.074
Asp D	–0.022	0.041	–0.049	–0.077	35.939	7.424	–0.517	–0.335	–0.855	0.009
Cys C	–0.02	0.127	–0.208	0.065	28.284	6.41	–0.77	–0.203	–0.615	0.028
Gln Q	–0.03	0.11	–0.09	–0.036	55.727	8.7	–1.05	–0.22	–1.179	0.042
Glu E	–0.082	0.113	–0.03	–0.071	50.739	8.289	–1.044	–0.217	–1.097	0.015
Gly G	–0.002	–0.007	–0.014	–0.043	11.858	4.46	–0.148	–0.289	–0.368	–0.183
His H	–0.057	0.151	–0.048	–0.11	65.203	8.596	–0.204	–0.779	–1.067	–0.237
Ile I	–0.107	0.06	0.097	–0.123	63.23	9.217	–0.779	–0.034	–1.449	0.154
Leu L	–0.109	0.091	0.061	–0.062	63.23	9.026	–0.547	–0.291	–1.189	–0.099
Lys K	–0.088	–0.004	0.114	–0.069	74.967	10.245	–1.451	–0.049	–1.314	0.196
Met M	0.049	0.075	–0.004	–0.09	57.446	8.017	–1.531	–0.099	–1.017	0.161
Phe F	0.058	0.103	–0.107	–0.137	92.58	11.443	–1.185	–0.197	–1.824	–0.087
Pro P	–0.202	0.13	0.042	–0.058	40.847	9.749	–1.02	–0.566	–0.554	–0.387
Ser S	0.015	0.037	–0.08	0.012	24.358	6.569	–0.448	–0.322	–0.604	–0.072
Thr T	–0.061	0.077	–0.052	–0.039	36.743	7.754	–0.713	0.176	–0.94	0.069
Trp W	0.134	0.065	–0.333	–0.075	137.702	12.289	–1.027	–1.072	–1.798	–0.389
Tyr Y	0.059	0.153	–0.116	0.21	99.748	10.75	–1.066	–0.356	–1.706	–0.061
Val V	0.073	0.088	–0.038	0.037	46.559	8.842	–0.718	–0.168	–1.212	0.141
	71	72	73	74	75	76	77	78	79	80
Ala A	0.432	–0.093	0.056	0.061	–0.063	–0.148	–0.043	0.176	0.119	0.119
Arg R	0.864	–0.079	–0.087	0.019	0.247	–0.327	–0.32	0.499	0.192	–0.106
Asn N	0.607	–0.091	–0.097	0.158	–0.07	–0.076	–0.228	0.315	0.073	–0.113
Asp D	0.61	–0.168	–0.072	0.112	0.007	–0.141	–0.146	0.27	0.025	–0.068
Cys C	0.316	–0.081	0.106	0.021	–0.079	–0.052	–0.087	–0.087	0.179	–0.129
Gln Q	0.562	0.009	–0.078	0.158	0.096	–0.382	–0.133	0.451	0.041	–0.209
Glu E	0.695	–0.203	–0.111	0.181	0.075	–0.328	–0.102	0.363	–0.001	–0.134
Gly G	0.369	–0.008	–0.048	0.044	–0.021	–0.054	–0.068	0.073	0.114	0.077
His H	1.045	0.019	–0.562	0.094	0.338	–0.305	–0.057	0.005	0.186	0.153
Ile I	0.546	–0.18	0.175	0.06	0.039	–0.522	0.014	0.574	–0.026	–0.138
Leu L	0.679	–0.239	0.283	0.027	0.035	–0.615	0.065	0.555	0.008	–0.147
Lys K	0.652	–0.108	–0.016	0.152	0.128	–0.688	–0.087	0.597	0.062	–0.079

TABLE S-VI. Continued

Amino acid	71	72	73	74	75	76	77	78	79	80
Met M	0.315	0.154	0.238	0.036	–0.293	–0.119	0.009	0.204	0.142	–0.126
Phe F	1.65	–0.562	–0.294	–0.073	0.527	–0.279	–0.322	0.158	0.296	0.094
Pro P	1.006	–0.456	0.022	0.309	–0.163	–0.306	0.003	0.298	0.188	–0.248
Ser S	0.447	–0.131	0.031	0.038	0.041	–0.128	–0.173	0.225	0.108	–0.099
Thr T	0.425	–0.107	0.063	0.048	0.069	–0.31	–0.058	0.263	0.124	–0.125
Trp W	2.291	–0.523	–0.856	0.148	0.572	–0.092	–0.563	–0.001	0.47	0.293
Tyr Y	1.737	–0.592	–0.298	–0.138	0.48	–0.123	–0.379	0.034	0.459	–0.042
Val V	0.403	–0.1	0.16	0.096	–0.042	–0.368	–0.072	0.501	–0.008	–0.127
	81	82	83	84	85	86	87	88	89	90
Ala A	–0.042	–0.078	0.141	0.02	–0.089	0.068	–0.027	–0.025	–0.007	0.022
Arg R	–0.177	–0.248	0.269	0.041	–0.085	0.109	–0.051	0.005	–0.05	0.04
Asn N	–0.074	0.018	0.018	0.038	–0.083	0.119	–0.011	–0.065	–0.065	0.062
Asp D	–0.065	–0.03	0.125	–0.05	–0.051	0.044	0.059	–0.057	–0.065	0.072
Cys C	–0.042	–0.06	0.103	0.037	–0.107	0.113	–0.044	–0.047	–0.01	0.051
Gln Q	–0.058	–0.084	0.279	–0.108	0.101	0.107	0.022	–0.109	0.066	0.122
Glu E	–0.1	–0.049	0.211	–0.081	–0.162	0.13	–0.029	–0.031	–0.059	0.058
Gly G	0.016	–0.065	0.047	0.065	–0.066	0.007	0.034	–0.032	0	–0.014
His H	–0.118	–0.303	0.177	0.134	–0.067	0.068	–0.124	0	0.051	0.085
Ile I	–0.208	–0.121	0.42	–0.133	–0.026	0.004	0.003	–0.132	0.045	0.085
Leu L	–0.273	–0.1	0.462	–0.102	–0.109	0.048	0.026	–0.129	0.035	0.12
Lys K	–0.236	–0.103	0.342	–0.019	–0.239	0.201	–0.093	0.007	–0.049	0.134
Met M	–0.206	–0.054	0.263	0.105	–0.148	0.149	–0.088	–0.09	0.028	0.104
Phe F	–0.228	–0.394	0.196	0.514	–0.351	0.107	–0.274	0.068	0.214	–0.011
Pro P	–0.178	0.017	0.231	0.039	–0.251	0.174	–0.002	–0.149	0.053	0.075
Ser S	–0.045	–0.068	0.087	0.045	–0.071	0.028	0.044	–0.069	0.016	–0.022
Thr T	–0.114	–0.065	0.186	0.011	–0.116	0.069	–0.024	–0.006	–0.045	0.071
Trp W	–0.413	–0.434	0.34	0.347	–0.002	–0.145	–0.297	0.102	0.174	–0.054
Tyr Y	–0.086	–0.434	0.152	0.356	–0.126	0.025	–0.203	–0.018	0.205	0.044
Val V	–0.188	–0.082	–0.313	–0.008	–0.168	0.125	–0.073	0.008	–0.101	0.154
	91	92	93	94	95	96	97	98	99	100
Ala A	0.055	–0.035	–0.056	0.027	0.031	0.005	82.837	17.763	–0.833	0.048
Arg R	0.162	–0.071	–0.036	–0.046	0.164	0.001	336.929	28.849	–3.378	3.075
Asn N	0.072	–0.003	–0.103	0.05	0.017	–0.029	149.531	20.016	–0.021	0.35
Asp D	–0.002	0.022	–0.048	0.002	0.043	–0.027	136.252	17.582	–0.698	0.462
Cys C	0.02	–0.022	–0.041	0.058	–0.039	0.027	97.269	16.618	–1.469	0.24
Gln Q	0.048	–0.036	–0.071	0.031	0.032	–0.006	203.734	22.159	–2.48	1.033
Glu E	0.033	0.013	–0.102	0.011	0.102	–0.031	188.222	20.43	–3.045	0.878
Gly G	0.023	0.019	–0.012	–0.023	0.008	0.006	49.721	14.26	–0.768	0.091
His H	–0.035	–0.012	–0.059	0.023	0.086	–0.073	201.631	16.419	–0.163	0.525
Ile I	–0.035	–0.012	–0.059	0.023	0.086	–0.073	201.631	16.419	–0.163	0.525
Leu L	0.054	–0.096	–0.181	0.1	0.192	–0.067	232.078	19.713	–1.482	0.796
Lys K	0.058	–0.029	–0.157	–0.019	0.223	–0.061	279.532	27.373	–5.107	2.638
Met M	0.022	–0.078	–0.066	0.05	0.106	–0.056	194.655	20.285	–2.995	0.836
Phe F	–0.117	–0.044	0.05	0.019	0.077	–0.155	256.714	20.944	–2.253	1.086
Pro P	0.052	–0.018	–0.222	0.118	0.127	–0.068	140.56	22.745	–2.214	0.1

TABLE S-VI. Continued

Amino acid	91	92	93	94	95	96	97	98	99	100
Ser S	0.075	–0.047	–0.003	0.005	0.004	0.012	100.633	18.38	–0.323	–0.044
Thr T	0.026	–0.024	–0.066	0.009	0.089	–0.04	145.892	20.31	–1.579	0.816
Trp W	–0.158	–0.022	0.077	0.028	–0.127	–0.088	358.067	21.217	–1.518	0.218
Tyr Y	–0.186	0.011	0.066	0.009	0.034	–0.174	287.473	20.412	–1.225	1.738
Val V	0.054	–0.095	–0.102	0.036	0.131	–0.002	174.016	20.95	–1.043	0.334
	101	102	103	104	105	106	107	108	109	110
Ala A	–1.678	–1.164	0.777	0.366	0.499	0.219	–0.694	–0.39	–0.147	0.238
Arg R	–4.895	–1.662	1.434	0.889	0.524	–0.24	–0.379	–0.823	–0.556	1.245
Asn N	–2.445	–1.747	1.241	0.696	0.161	0.472	–0.774	–0.331	–0.316	0.616
Asp D	–2.176	–1.617	1.61	0.164	0.272	0.25	–0.318	–0.584	–0.171	0.598
Cys C	–1.698	–1.21	0.932	0.444	0.413	0.089	–0.506	–0.26	–0.096	0.221
Gln Q	–3.454	–1.366	0.638	1.512	0.142	0.448	–0.299	–1.227	0.24	0.95
Glu E	–2.76	–1.855	1.496	1.376	0.178	0.487	–0.395	–0.888	0.159	0.526
Gly G	–1.143	–1.31	0.93	0.258	0.179	0.264	–0.408	–0.296	–0.027	0.18
His H	–2.887	–1.356	1.528	0.809	–0.377	0.118	0.174	–1.086	0.328	–0.086
Ile I	–3.356	–1.364	0.948	0.469	1.061	0.136	–0.918	–1.047	0.747	0.743
Leu L	–3.018	–2.089	1.126	0.908	1.093	–0.033	–0.79	–1.522	0.815	0.996
Lys K	–3.641	–1.706	1.28	0.888	0.607	0.33	–0.788	–1.476	0.26	1.194
Met M	–2.399	–1.17	0.944	1.046	0.825	0.178	–1.115	–0.547	0.363	0.541
Phe F	–3.407	–2.351	2.833	0.187	0.151	–0.157	0.144	–1.125	0.214	0.72
Pro P	–1.656	–1.952	1.755	0.274	0.229	0.338	–0.653	–0.915	0.659	0.352
Ser S	–1.62	–1.571	1.001	1.301	0.518	0.016	–0.09	–0.597	–0.335	0.575
Thr T	–2.693	–1.249	0.666	0.651	0.821	–0.177	–0.119	–0.965	0.12	0.449
Trp W	–2.897	–2.54	3.768	0.541	–0.505	0.036	0.459	–1.17	–0.333	0.605
Tyr Y	–3.045	–2.627	2.863	0.174	0.493	–0.529	0.283	–1.017	–0.14	0.471
Val V	–2.905	–1.019	0.354	0.805	0.886	0.212	–0.764	–0.937	0.424	0.717
	111	112	113	114	115	116	117	118	119	120
Ala A	0.397	–0.195	–0.011	–0.52	0.3	0.16	–0.086	0.179	–0.173	–0.058
Arg R	0.408	–0.018	–0.213	–0.94	0.235	0.546	0.043	0.173	–0.233	–0.035
Asn N	0.399	–0.122	–0.18	–0.237	–0.118	0.178	0.166	0.225	–0.22	–0.111
Asp D	–0.026	0.233	–0.108	–0.317	0.056	–0.116	0.24	0.081	0.036	–0.187
Cys C	0.389	–0.103	–0.055	–0.358	0.015	0.317	–0.089	0.235	–0.203	–0.113
Gln Q	0.165	–0.213	–0.123	–0.631	0.475	0.054	–0.01	0.178	–0.042	–0.297
Glu E	0.014	–0.028	–0.132	–0.514	0.334	0.187	–0.256	0.325	–0.237	–0.15
Gly G	0.279	–0.032	–0.004	–0.332	0.001	0.191	0.015	0.041	0.047	–0.173
His H	0.631	0.34	–0.082	–0.947	0.262	0.46	–0.067	0.261	–0.296	–0.188
Ile I	0.178	–0.01	–0.346	–0.921	0.909	0.178	–0.013	–0.155	–0.115	–0.328
Leu L	–0.085	0.211	–0.833	–0.665	0.925	0.37	–0.254	–0.032	–0.024	–0.208
Lys K	–0.103	0.237	–0.39	–0.898	0.541	0.465	–0.452	0.503	–0.611	0.111
Met M	0.338	–0.158	–0.273	–0.58	0.582	0.568	–0.249	0.242	–0.226	–0.216
Phe F	0.148	0.352	–0.605	–0.801	0.285	1.062	–0.763	0.453	–0.716	–0.05
Pro P	0.242	0.027	–0.73	–0.138	0.443	0.411	–0.575	0.439	–0.339	–0.16
Ser S	0.221	0.087	–0.095	–0.396	–0.036	0.264	0.013	0.005	0.197	–0.334
Thr T	0.356	0.096	–0.355	–0.388	0.223	0.246	–0.167	0.129	–0.142	0.021
Trp W	0.343	0.698	–0.717	–1.059	0.747	0.61	0.102	–0.095	–0.637	–0.11

TABLE S-VI. Continued

	111	112	113	114	115	116	117	118	119	120
Tyr Y	0.469	0.06	–0.341	–0.793	0.159	0.583	–0.191	0.412	–0.541	–0.198
Val V	0.168	–0.049	–0.0453	–0.556	0.561	0.388	–0.339	0.143	–0.185	–0.056
	121	122	123	124	125	126	127	128	129	130
Ala A	–0.059	0.073	0.168	–0.097	–0.055	0.02	–0.015	0.026	23.503	6.852
Arg R	–0.024	0.118	0.499	–0.064	0.007	–0.264	0.297	–0.034	103.614	11.716
Asn N	–0.197	0.111	0.21	0.089	–0.129	–0.058	–0.07	–0.06	43.369	8.672
Asp D	–0.098	0.071	0.005	0.138	–0.012	–0.008	–0.145	0.024	38.361	7.734
Cys C	–0.031	0.082	0.06	0.016	0.002	0.062	–0.182	0.135	35.365	7.289
Gln Q	–0.134	0.18	0.178	0.047	–0.098	–0.005	–0.166	0.113	61.289	9.482
Glu E	–0.103	0.042	0.112	0.06	–0.125	–0.043	0.098	–0.069	55.31	8.968
Gly G	–0.012	–0.009	0.085	0.024	0.085	–0.142	–0.014	–0.036	13.019	4.754
His H	0.182	0.197	–0.078	0.035	–0.111	–0.011	0.028	–0.091	70.101	8.839
Ile I	0.262	0.152	0.198	–0.188	–0.328	0.182	0.228	–0.032	73.57	10.225
Leu L	0.197	0.183	0.043	–0.106	–0.435	0.31	0.321	–0.167	73.57	10.049
Lys K	–0.069	0.376	0.165	0.006	–0.303	–0.221	0.392	–0.109	86.362	11.419
Met M	0.165	0.066	0.176	–0.226	–0.149	0.095	0.213	–0.078	71.445	9.327
Phe F	0.595	–0.132	–0.041	–0.189	0.062	–0.053	0.21	–0.072	102.665	12.055
Pro P	0.181	0.111	0.164	–0.124	–0.371	0.176	0.235	–0.101	46.05	10.611
Ser S	0.031	–0.026	0.123	–0.047	0.133	–0.052	–0.131	0.035	26.775	6.929
Thr T	–0.127	0.177	–0.013	0.065	–0.094	–0.007	0.08	0.124	41.162	8.368
Trp W	0.37	–0.028	–0.01	–0.068	0.021	–0.175	–0.099	–0.021	150.163	12.814
Tyr Y	0.429	0.082	–0.297	–0.054	0.157	–0.094	0.097	–0.213	109.356	11.268
Val V	–0.168	0.344	0.092	–0.105	–0.277	0.113	0.2	–0.044	53.778	9.748
	131	132	133	134	135	136	137	138	139	140
Ala A	–0.479	–0.227	–0.623	–0.184	0.395	–0.022	0.095	0.049	–0.091	–0.191
Arg R	–1.459	0.197	–1.882	–0.126	0.767	0.079	–0.031	–0.028	0.227	–0.421
Asn N	–0.516	–0.387	–0.795	–0.211	0.576	–0.022	–0.062	0.16	–0.084	–0.142
Asp D	–0.573	–0.323	–0.791	–0.092	0.567	–0.105	–0.021	0.089	0.004	–0.195
Cys C	–0.97	–0.198	–0.676	0.023	0.26	0.013	0.134	0.617	–0.146	–0.015
Gln Q	–1.15	–0.166	–1.187	–0.083	0.476	0.157	–0.075	0.175	0.086	–0.486
Glu E	–1.11	–0.16	–1.112	–0.116	0.657	–0.105	–0.079	0.17	0.058	–0.407
Gly G	–0.175	–0.287	–0.34	–0.246	0.347	0.026	–0.013	0.033	–0.03	–0.082
His H	–0.262	–0.635	–1.108	–0.29	0.955	0.109	–0.479	0.07	0.306	–0.386
Ile I	–0.901	0.103	–1.564	0.001	0.517	–0.033	0.226	0.072	–0.054	–0.589
Leu L	–0.687	–0.183	–1.3	–0.265	0.655	–0.081	0.333	0.011	–0.033	–0.718
Lys K	–1.791	0.177	–1.455	0.036	0.638	0.048	0.013	0.139	0.071	–0.756
Met M	–2.146	0.189	–1.096	0.08	0.205	0.289	0.377	–0.015	–0.426	–0.053
Phe F	–1.22	–0.157	–1.856	–0.332	1.713	–0.446	–0.252	–0.128	0.526	–0.399
Pro P	–1.108	–0.498	–0.585	–0.554	1.022	–0.353	0.054	0.298	–0.203	–0.397
Ser S	–0.451	–0.319	–0.57	–0.166	0.409	–0.062	0.073	0.013	0.038	–0.181
Thr T	–0.78	–0.133	–0.974	–0.024	0.367	0.015	0.095	0.026	0.045	–0.364
Trp W	–1.103	–1.028	–1.804	–0.554	2.353	–0.372	–0.805	0.115	0.574	–0.222
Tyr Y	–1.077	–0.276	–1.654	–0.292	1.739	–0.484	–0.237	–0.207	0.486	–0.23
Val V	–0.731	–0.154	–1.284	0.01	0.372	0.032	0.215	0.072	–0.086	–0.447

TABLE S-VI. Continued

Amino acid	141	142	143	144	145	146	147	148	149	150
Ala A	–0.012	0.203	0.118	–0.141	–0.055	–0.102	0.164	0.032	–0.09	0.067
Arg R	–0.265	0.565	0.157	–0.107	–0.195	–0.278	0.299	0.076	–0.101	0.122
Asn N	–0.208	0.358	0.072	–0.114	–0.086	–0.006	0.031	0.048	–0.062	0.097
Asp D	–0.133	0.315	0.013	–0.068	–0.073	–0.048	0.14	–0.045	–0.043	0.038
Cys C	–0.089	0.051	0.226	–0.138	–0.044	–0.087	0.116	0.059	–0.112	0.127
Gln Q	–0.072	0.494	0.029	–0.2	–0.087	–0.121	0.297	–0.073	–0.112	0.103
Glu E	–0.069	0.411	–0.021	–0.124	–0.117	–0.088	0.247	–0.004	–0.172	0.135
Gly G	–0.061	0.103	0.107	–0.078	–0.019	–0.079	0.055	0.071	–0.059	0.001
His H	–0.021	0.063	0.188	0.129	–0.123	–0.332	0.208	0.152	–0.081	0.076
Ile I	0.087	0.632	–0.043	–0.125	–0.244	–0.186	0.487	–0.104	–0.032	–0.016
Leu L	0.161	0.607	–0.007	–0.133	–0.329	–0.146	0.522	–0.059	–0.124	0.032
Lys K	–0.01	0.656	0.02	–0.072	–0.254	–0.165	0.39	0.024	–0.263	0.215
Met M	0.047	0.2	0.187	–0.14	–0.231	–0.121	0.325	0.145	–0.2	0.183
Phe F	–0.24	0.228	0.259	0.123	–0.283	–0.426	0.217	0.568	–0.387	0.132
Pro P	0.099	0.31	0.177	–0.223	–0.232	–0.005	0.267	0.065	–0.262	0.18
Ser S	–0.145	0.246	0.114	–0.109	–0.053	–0.073	0.082	0.066	–0.081	0.03
Thr T	–0.02	0.298	0.118	–0.121	–0.138	–0.087	0.214	0.021	–0.114	0.063
Trp W	–0.486	0.085	0.429	0.302	–0.445	–0.471	0.397	0.38	–0.01	–0.138
Tyr Y	–0.318	0.105	0.425	–0.012	–0.127	–0.452	0.181	0.382	–0.149	0.038
Val V	–0.001	0.537	–0.01	–0.128	–0.222	–0.118	0.352	0.032	–0.183	0.113
	151	152	153	154	155	156	157	158	159	160
Ala A	–0.043	–0.029	0.004	0.03	0.062	–0.054	–0.058	0.029	0.039	0.011
Arg R	–0.088	0.002	–0.036	0.064	0.167	–0.076	–0.062	–0.049	0.202	–0.01
Asn N	–0.013	–0.072	–0.064	0.079	0.066	–0.004	–0.11	0.046	0.029	–0.03
Asp D	0.054	–0.073	–0.049	0.079	–0.002	0.014	–0.052	0.002	0.049	–0.013
Cys C	–0.072	–0.065	0.004	0.081	0.002	–0.029	–0.045	0.074	–0.043	0.025
Gln Q	0.011	–0.121	–0.056	0.142	0.041	–0.04	–0.084	–0.049	0.046	0.003
Glu E	–0.046	–0.046	–0.035	0.062	0.036	–0.002	–0.108	0.002	0.125	–0.024
Gly G	0.033	–0.042	0.005	–0.004	0.021	0.011	–0.01	–0.026	0.011	0.011
His H	–0.139	–0.016	0.07	0.102	–0.038	–0.02	–0.066	0.006	0.109	–0.065
Ile I	–0.018	–0.15	0.088	0.154	0.025	–0.082	–0.2	0.082	0.21	–0.068
Leu L	0.002	–0.141	0.077	0.142	0.037	–0.107	–0.211	0.116	0.23	–0.076
Lys K	–0.149	0.012	–0.022	0.156	0.052	–0.043	–0.178	–0.024	0.261	–0.068
Met M	–0.104	–0.139	0.099	0.105	0.03	–0.125	–0.068	0.067	0.102	–0.057
Phe F	–0.325	0.051	0.251	–0.005	–0.119	–0.067	0.046	0.005	0.119	–0.151
Pro P	–0.044	–0.143	0.075	0.089	0.052	–0.039	–0.236	0.127	0.149	–0.072
Ser S	0.039	–0.083	0.025	–0.01	0.075	–0.059	–0.004	0.005	0.013	0.013
Thr T	–0.042	–0.009	–0.033	0.089	0.022	–0.036	–0.074	0.004	0.114	–0.044
Trp W	–0.322	0.068	0.216	–0.04	–0.154	–0.03	0.061	0.016	–0.086	–0.085
Tyr Y	–0.228	–0.037	0.236	0.059	–0.197	–0.007	0.065	–0.014	0.068	–0.167
Val V	–0.094	0	–0.083	0.18	0.045	–0.111	–0.119	0.04	0.165	–0.007

TABLE S-VII. The prominent principal component score and the corresponding eigenvalue of 99 WHIM, 74 geometrical and 160 3D-MORSE descriptors

No.	WHIM descriptors		Geometrical descriptors			3D-MORSE descriptors		
	Eigenvalue	%	No.	Eigenvalue	%	No.	Eigenvalue	%
1	50.53231	51.04274	1	24.0685	55.9734	1	51.4977	32.186
2	21.8993	22.1205	2	5.3808	12.5134	2	33.1044	20.6902
3	7.79161	7.87031	3	4.2917	9.9808	3	13.9686	8.7304
4	**6.45442**	**6.51961**	4	2.6995	6.2779	4	12.3945	7.7466
5	3.857	3.89596	5	**2.1097**	**4.9062**	5	9.3042	5.8151
6	2.10476	2.12602	6	1.4465	3.364	6	6.4255	4.0159
7	1.90928	1.92857	7	0.7745	1.8012	7	6.1609	3.8505
8	1.28317	1.29613	8	0.5868	1.3648	8	5.1864	3.2415
9	1.00451	1.01465	9	0.4141	0.9629	9	**4.2146**	**2.6341**
10	0.49921	0.50425	10	0.3348	0.7786	10	3.7662	2.3539
11	0.40984	0.41398	11	0.3145	0.7313	11	3.2758	2.0474
12	0.31501	0.31819	12	0.201	0.4675	12	2.7476	1.7173
13	0.29745	0.30046	13	0.144	0.3349	13	2.0126	1.2579
14	0.20545	0.20753	14	0.1035	0.2406	14	1.673	1.0456
15	0.18188	0.18372	15	0.0674	0.1567	15	1.2115	0.7572
16	0.12194	0.12317	16	0.0304	0.0706	16	1.0376	0.6485
17	0.08147	0.0823	17	0.015	0.0349	17	0.7839	0.4899
18	0.03159	0.03191	18	0.0101	0.0234	18	0.696	0.435
19	0.0198	0.02	19	0.0073	0.017	19	0.5391	0.337

TABLE S-VIII. SVWGM descriptors for amino acid

Amino acid	4 PCs of 99 WHIM descriptors				5 PCs of 74 GEOMETRICAL descriptors				
	1	2	3	4	5	6	7	8	9
Ala A	–11.6344	–1.89677	1.97792	–2.60584	–4.8673	4.8567	0.2935	–1.1984	–1.7882
Arg R	11.87107	–2.87032	2.74823	1.25733	2.8074	–5.7328	–10.0304	0.1795	–1.8715
Asn N	–5.34951	7.68317	4.11656	4.1737	–4.2352	4.4868	–1.581	–0.5684	2.9533
Asp D	–4.02676	2.99296	–3.35851	–3.77013	–3.675	5.6066	–1.3744	2.375	2.2126
Cys C	–5.65032	–2.87927	–2.99048	2.3435	–4.8303	6.3381	0.7404	–3.6652	–1.4237
Gln Q	2.17649	–2.40014	0.84542	3.57161	–4.7011	–0.8588	–3.5331	2.8551	0.7795
Glu E	2.36689	0.15236	–4.04847	0.80392	–2.5167	–0.6925	–1.4272	–0.5017	3.8605
Gly G	–11.7823	–13.6975	3.47023	0.20083	–4.6947	9.543	–0.1147	–0.1988	–0.4047
His H	2.33867	0.36099	–1.56522	–1.07602	4.3481	3.6123	1.8679	3.8021	0.8011
Ile I	0.41167	6.40422	–1.24429	–1.62207	–2.213	–5.2	3.6825	5.6186	–5.5525
Leu L	0.26884	8.11636	2.89694	0.98231	–3.1883	–9.1722	5.8668	4.8209	–0.3621
Lys K	9.00644	–2.09657	–3.35485	2.39215	–0.5855	–11.934	–4.8895	–1.5481	3.8189
Met M	4.36305	–1.66543	–3.97675	–1.02332	–3.9573	–4.0717	2.3255	–4.3088	–5.9397
Phe F	7.26373	–4.36604	–1.09058	1.62148	13.8796	–2.8069	2.0969	–6.5485	–0.7813
Pro P	–5.3069	3.18399	0.59462	4.27735	–2.7762	–3.1282	7.5571	–2.5701	7.0349
Ser S	–9.15545	2.32007	–0.49944	–2.26929	–3.3416	6.8606	–0.9756	0.8169	–1.6716
Thr T	–4.22009	–0.27234	–1.39095	–2.53752	–2.7966	0.8798	–1.07	–1.2773	–0.1541
Trp W	11.70248	0.16164	5.62047	–4.91875	19.2063	3.7382	–0.0853	6.6477	1.0446
Tyr Y	8.54042	–1.52583	1.74131	–1.28473	13.2974	2.2367	0.8869	–4.4618	–1.3706
Val V	–3.18401	2.29446	–0.49218	–0.51649	–5.1598	–4.5618	–0.2362	–0.2687	–1.1856

TABLE S-VIII. Continued

Amino acid	9 PCs of 160 3D–MORSE descriptors								
	10	11	12	13	14	15	16	17	18
Ala A	0.1084	–0.9456	2.594	–0.4204	–5.9082	0.7197	–0.4996	1.1106	1.2405
Arg R	–2.2739	1.9078	1.1935	–2.4422	7.4386	–3.1167	4.6443	–0.2724	0.8752
Asn N	–1.0546	0.4217	–0.9058	–0.1832	–3.2495	–2.6099	–1.4085	–0.0991	–1.3005
Asp D	1.3277	2.1233	–0.3749	2.3417	–3.4151	–1.5233	–0.4735	0.8486	–2.1593
Cys C	–1.5163	–2.8869	–1.281	–0.17	–3.4528	4.8887	0.7554	–2.7832	–1.1356
Gln Q	–1.1897	0.1375	–5.7141	1.1501	0.7412	–0.9023	3.0675	1.2869	–1.4086
Glu E	1.0923	–0.405	–1.4034	3.1204	–0.6074	–0.935	2.76	1.4287	–2.2504
Gly G	1.5034	0.8204	3.4931	0.5227	–6.8679	2.8835	1.4014	2.882	1.5407
His H	5.1879	–5.1697	–0.707	–3.4635	1.2963	–0.2267	0.2235	0.0099	–1.0526
Ile I	2.8169	1.2975	–0.3499	3.6691	–0.6107	–3.1206	–2.2368	–1.9152	1.3986
Leu L	–0.8706	3.2602	1.0827	–2.473	–0.7406	–3.4737	–1.9935	–1.5138	0.4502
Lys K	2.5185	–2.3203	1.6602	1.1263	3.019	–0.349	2.4237	–0.6669	3.2261
Met M	–3.4378	–3.8467	–2.0643	0.6833	1.3496	3.1447	1.4995	–4.1297	–0.9596
Phe F	0.7353	0.795	2.6529	2.9515	6.3729	3.7134	–1.7188	1.5139	1.1937
Pro P	–3.0942	0.809	–0.1377	–0.9751	–3.2208	0.2478	–1.5559	–0.0085	0.8157
Ser S	–2.4715	3.4412	1.2993	–0.512	–5.334	–0.3277	–1.1852	0.4623	–0.9085
Thr T	2.2859	0.1757	1.2979	–0.8649	–3.2382	–0.3181	–0.8146	–0.0062	1.1008
Trp W	–4.6764	–2.6452	0.9322	0.8924	11.0178	0.4087	–3.3601	0.4752	–0.9542
Tyr Y	3.1242	4.2243	–4.419	–2.4006	7.4664	1.5601	–1.4356	1.7758	–0.5768
Val V	–0.1157	–1.1943	1.1514	–2.5526	–2.0567	–0.6638	–0.0932	–0.399	0.8648

TABLE S-IX. Statistics of PLS on variables selected by SMR (OT analogues)

No.	Variable	l[a]	R^2_{cum}	Q^2_{LOO}
1	v_{24}	1	0.516	0.501
2	v_{24}, v_{17}	2	0.855	0.596
3	v_{24}, v_{17}, v_{36}	2	0.909	0.777
4	$v_{24}, v_{17}, v_{36}, v_{51}$	2	0.950	0.843
5	$\boldsymbol{v_{24}, v_{17}, v_{36}, v_{51}, v_{23}}$	**2**	**0.944**	**0.864**
6	$v_{24}, v_{17}, v_{36}, v_{51}, v_{23}, v_{14}$	3	0.955	0.647
7	$v_{24}, v_{17}, v_{36}, v_{51}, v_{23}, v_{14}, v_{46}$	3	0.948	0.657
8	$v_{24}, v_{17}, v_{36}, v_{51}, v_{23}, v_{14}, v_{46}, v_{40}$	2	0.954	0.668
9	$v_{24}, v_{17}, v_{36}, v_{51}, v_{23}, v_{14}, v_{46}, v_{40}, v_{27}$	2	0.953	0.601
10	$v_{24}, v_{17}, v_{36}, v_{51}, v_{23}, v_{14}, v_{46}, v_{40}, v_{27}, v_{19}$	2	0.952	0.747

[a]Principal components

TABLE S-X. Statistics of PLS on variables selected by SMR (BTT)

No.	Variable	l[a]	R^2_{cum}	Q^2_{LOO}
1	v_{32}	1	0.462	0.421
2	v_{32}, v_{1}	1	0.692	0.646
3	v_{32}, v_{1}, v_{34}	2	0.709	0.644
4	$v_{32}, v_{1}, v_{34}, v_{36}$	2	0.882	0.825
5	$v_{32}, v_{1}, v_{34}, v_{36}, v_{25}$	2	0.887	0.799
6	$v_{32}, v_{1}, v_{34}, v_{36}, v_{25}, v_{6}$	2	0.891	0.812

TABLE S-X. Continued

No.	Variable	l^a	R^2_{cum}	Q^2_{LOO}
7	**$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}$**	**2**	**0.911**	**0.818**
8	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}$	2	0.913	0.812
9	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}$	3	0.926	0.795
10	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}$	3	0.927	0.812
11	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}$	3	0.92	0.814
12	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}, v_7$	3	0.926	0.772
13	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}, v_7, v_{18}$	3	0.921	0.659
14	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}, v_7, v_{18}, v_{28}$	3	0.920	0.672
15	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}, v_7, v_{18}, v_{28}, v_{22}$	3	0.921	0.641
16	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}, v_7, v_{18}, v_{28}, v_{22}$ v_{21}	3	0.926	0.606
17	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}, v_7, v_{18}, v_{28}, v_{22}$, v_{21}, v_{24}	3	0.919	0.623
18	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}, v_7, v_{18}, v_{28}, v_{22}$, v_{21}, v_{24}, v_{35}	3	0.930	0.589
19	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}, v_7, v_{18}, v_{28}, v_{22}$, $v_{21}, v_{24}, v_{35}, v_{30}$	3	0.917	0.553
20	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}, v_7, v_{18}, v_{28}, v_{22}$, $v_{21}, v_{24}, v_{35}, v_{30}, v_5$	3	0.927	0.607
21	$v_{32}, v_1, v_{34}, v_{36}, v_{25}, v_6, v_{12}, v_{29}, v_{27}, v_{20}, v_{23}, v_7, v_{18}, v_{28}, v_{22}$, $v_{21}, v_{24}, v_{35}, v_{30}, v_5, v_9$	3	0.931	0.564

[a]Principal components

TABLE S-XI. Statistics of PLS on variables selected by SMR (ace)

No.	Variable	l^a	R^2_{cum}	Q^2_{LOO}
1	*v23*	1	0.637	0.633
2	*v23, v34*	1	0.703	0.691
3	*v23, v34, v6*	1	0.756	0.734
4	*v23, v34, v6, v30*	1	0.786	0.768
5	*v23, v34, v6, v30, v28*	2	0.835	0.795
6	*v23, v34, v6, v30, v28, v9*	2	0.852	0.786
7	***v23, v34, v6, v30, v28, v9, v21***	**2**	**0.865**	**0.799**
8	*v23, v34, v6, v30, v28, v9, v21, v29*	2	0.852	0.777
9	*v23, v34, v6, v30, v28, v9, v21, v29, v36*	3	0.862	0.748
10	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17*	3	0.858	0.757
11	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4*	3	0.874	0.708
12	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18*	3	0.873	0.694
13	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12*	3	0.882	0.719
14	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31*	3	0.883	0.710
15	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11*	3	0.884	0.711
16	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25*	3	0.883	0.709
17	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26*	3	0.883	0.702

TABLE S-XI. Continued

No.	Variable	l^a	R^2_{cum}	Q^2_{LOO}
18	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13*	3	0.876	0.686
19	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20,*	3	0.875	0.696
20	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19*	3	0.876	0.692
21	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2*	3	0.873	0.692
22	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3*	3	0.787	0.629
23	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3, v10*	3	0.87	0.685
24	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3, v10, v8*	3	0.873	0.664
25	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3, v10, v8, v5*	3	0.871	0.655
26	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3, v10, v8, v5, v7*	3	0.871	0.642
27	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3, v10, v8, v5, v7, v14*	3	0.878	0.625
28	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3, v10, v8, v5, v7, v14, v16*	3	0.879	0.62
29	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3, v10, v8, v5, v7, v14, v16, v15*	3	0.883	0.619
30	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3, v10, v8, v5, v7, v14, v16, v15, v33*	3	0.879	0.614
31	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3, v10, v8, v5, v7, v14, v16, v15, v33, v27*	3	0.88	0.613
32	*v23, v34, v6, v30, v28, v9, v21, v29, v36, v17, v4, v18, v12, v31, v11, v25, v26, v13, v20, v19, v2, v3, v10, v8, v5, v7, v14, v16, v15, v33, v27, v32*	3	0.88	0.629

[a]Principal components

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 355–366 (2015)
JSCS–4721
UDC 547.213+547.313.3–31:66.094.258:661.183.8+544.478
Original scientific paper

Higher propene yield by tailoring operating conditions of propane oxidative dehydrogenation over V_2O_5/γ-Al_2O_3

YOUSEF ZAYNALI and SEYED MEHDI ALAVI*

Chemical Engineering Department, Iran University of Science and Technology, P. O. Box 16765-163, Tehran, Iran

(Received 5 May, revised 25 June, accepted 27 June 2014)

Abstract: Supported vanadia catalysts were successfully synthesized using wet impregnation of γ-alumina to study propane oxidative dehydrogenation (POD). The prepared catalysts were characterized by X-ray diffraction analysis, specific surface area determination, and temperature programmed reduction tests. Over a broad temperature range (340 to 630 °C), the effects of vanadia loading (2.7, 5.4, and 9 wt. %) and propane to oxygen ratio (3:1 to 1:3) were thoroughly investigated on propane conversion as well as propene yield at atmospheric pressure. The results indicated that on increasing the vanadia content, the activity of catalyst increased while the selectivity to propene decreased monotonically. With increasing temperature from 340 to 630°C, the yield to propene showed an ascending behavior for all catalyst samples. The yield to propene exhibited a maximum on changing the propane to oxygen ratio from 3:1 to 1:3. The yield increased with increasing oxygen partial pressure in the feed until an equimolar ratio of propane and oxygen and then it declined with further increase in the oxygen partial pressure. A maximum propene yield of 17 % was experienced on the catalyst with 2.7 wt. % vanadia at a temperature of 550 °C.

Keyword: propane oxidative dehydrogenation; vanadia loading; propane to oxygen ratio; feed composition.

INTRODUCTION

Propene as a commodity product is an important feedstock in the chemical and petrochemical industries. Propene is used in the production of 2-propanol, acrylonitrile, propylene oxide and epichlorohydrin but its main use is in the production of polypropylene.[1] Conventional processes including thermal and catalytic cracking of naphtha and dehydrogenation of propane suffer from technical and economical drawbacks.[2–4] Furthermore, a consumption increase of about 5 % per year due to expansion of the polypropylene market[5,6] makes it crucial to

* Corresponding author. E-mail: alavi.m@iust.ac.ir
doi: 10.2298/JSC140505067Z

develop new routes and processes for propene production. Due to the inherent beneficial characteristics of non-equilibrium exothermic reactions, propane oxidative dehydrogenation (POD) seems to be a promising route to fill the gap between the increasing demand and the capacity of existing plants. However, despite these tempting benefits, the low yield of propene prevents the commercial development of POD.[7,8]

Vanadia was shown to be selective in the oxidative dehydrogenation of alkanes.[9–13] Thus, the impact of vanadia catalysts supported on metal oxides was studied for POD.[14–24] Focused on the dispersion and morphology of vanadia over different supports, Khodakov *et al.*[14] reported that among the different forms of vanadia, a polyvanadate monolayer is more selective in POD reactions. The diversity in the structure of vanadia depends on its dispersion, which mainly arises from loading as well as support characteristics, such as surface area and acid/base properties.[2,24,27–30] Martra *et al.*[2] studied the acid/base properties of vanadia impregnated on Al_2O_3, H–Na/Y zeolite, MgO, SiO_2, TiO_2 and ZrO_2. In a similar work but with more emphasis on the effect of the support, Dinse *et al.*[31] examined CeO_2, TiO_2, Al_2O_3, ZrO_2 and SiO_2 with low loadings of vanadia. Although the catalysts with acidic supports showed higher conversions in comparison with more basic ones, they had lower selectivity.[8] In attempts to enhance the reducibility of the active sites and increase the selectivity to propene, some researchers modified the acid/base properties of the surface by adding different alkali and alkaline earth metals (Li,[32] Na,[32] K[8,18,32,33] and Mg[34]).

Besides the aforementioned factors, the operating conditions and feed composition are decisive parameters in product distribution/catalytic performance of POD reactions.[34–36] POD is a set of consecutive reactions of oxydehydrogenation of propane and further oxidation of propene to CO and CO_2 according to the Mars van Kerevelen (MvK) mechanism.[37–40] Whether the reaction proceeds under propane rich (propane to oxygen ratio greater than stoichiometry) or propane deficient (propane to oxygen ratio less than stoichiometry) atmosphere, the controlling step changes and so does the product distribution. To the best of our knowledge, not so many studies thoroughly explore the effect of the propane to oxygen ratio (C_3/O_2) and reaction conditions on overall propene yield. Most of the studies were under stoichiometric or propane rich atmospheres[2,3,14,16,19,20,24,41–44] and in relatively fewer studies were the results obtained under oxygen-rich condition.[15,17,21,45,46] Given the very different conditions and even inconsistent results in various studies, it is not possible to reach a concrete understanding of the roles of the operating conditions and feed composition from the published works. According to Creaser *et al.*[41] at low propane concentrations or low propane conversions, a propane-rich atmosphere is beneficial for higher propene yield but no data at higher propane concentrations or conversions were

reported. On the contrary, the results of Jibril *et al.*[45] showed that higher yields of propene could be achieved at moderately propane deficient feeds.

The aim of present study was to examine the effect of the operating conditions and feed composition over V_2O_5/γ-alumina catalysts to achieve high propene yields. A comprehensive study was performed to explore the effects of vanadia loading over γ-alumina, temperature and feed composition on the overall propene yield, propane conversion and propene selectivity.

EXPERIMENTAL

Catalyst preparation

All catalysts were prepared by wet impregnation of γ-alumina (Merck) support (average particle diameter of 131 μm and density of 2250 kg m^{-3}) with aqueous solution of ammonium monovanadate (Merck). The impregnation was performed in a rotary evaporator (Biby Strilin Ltd., RE200) at reduced pressure and low temperature. The impregnated samples were dried overnight at 110 °C, further treated by heating up to 600 °C during 3 h, and then calcined at this temperature for 6 h. Three catalysts, VAl-01, VAl-02 and VAl-03, with different vanadia loadings of 2.7, 5.4 and 9 wt. %, respectively, approximately 8 to 10 g of each, were prepared following this procedure.

Characterization

The surface areas of the catalysts and support were determined by nitrogen adsorption––desorption at liquid nitrogen temperature (77 K), using a ChemBET-3000 (Quanta Chrome instrument). To ensure a clean and dry surface, the samples were first degassed for 1 h at 300 °C and 0.15 mbar. The surface areas were calculated using the Brunauer, Emmett and Teller (BET) method.

The X-ray diffraction (XRD) patterns were recorded on an X-Pert Philips diffractometer on powdered samples. The scans were collected within the 2θ range 10–110° in 0.01° steps with a 1 s step time using CuK_α radiation.

The reducibility of the catalysts was studied by temperature-programmed reduction (TPR) with a ChemiSorb 2750 (Micromeritics) apparatus. 50 mg of the catalysts were oxidized in an O_2–He mixture (20 vol. % O_2) with a flow rate of 30 cm^3 min^{-1} at 550 °C for 30 min and then cooled to 50 °C. Successively, the samples were degassed with nitrogen (20 cm^3 min^{-1}) for 2 h at 250 °C to remove the non-bound oxygen. After cooling to 25 °C, TPR tests were performed under a 20 mL min^{-1} flow of a 5 vol. % H_2/Ar mixture through a micro-reactor heating at a constant rate of 10 °C min^{-1} up to 950 °C.

Activity test

A one meter long quartz tube with an internal diameter of 6 mm, installed in a furnace, was used as a reactor. For every test, 100 mg of catalyst diluted with SiC was loaded in the reactor and feed mixture was introduced to catalytic bed at a total flow rate of 100 cm^3 min^{-1} at standard temperature and pressure (STP). The feed ratio was varied from N_2:O_2:C_3 = 6:3:1 (lean in propane) to 6:1:3 (rich in propane) to investigate the effect of feed composition. A thermometer was installed under the catalytic bed to monitor precisely the temperature of the reaction bed. After quenching and removal of water, the flow exiting the reactor was analyzed with an online Thermofinigan GC (model No. KAV00109) equipped with HayeSep Q, MolSieve 13X columns and in line thermal conductivity detector (TCD) for the detection of permanent gasses, *i.e.*, oxygen, nitrogen, and carbon oxides and Q-Plot column in line with a

flame ionization detector (FID) to detect the hydrocarbons, *i.e.*, propane and propene. The conversion, selectivity, and yield were calculated according to the following equations:

$$\text{Propane conversion: } X_{\text{propane}} = \frac{\left(F_{\text{propane,i}} - F_{\text{propane,e}}\right)}{F_{\text{propane,i}}} \times 100 \,, \quad (1)$$

$$\text{Propene selectivity: } S_{\text{propene}} = \frac{F_{\text{propene}}}{\left(F_{\text{propane,i}} - F_{\text{propane,e}}\right)} \times 100 \,, \quad (2)$$

$$\text{Propene yield: } Y_{\text{propene}} = \frac{\left(S_{\text{propene}} \times X_{\text{propane}}\right)}{100} \quad (3)$$

where $F_{\text{propane,i}}$, $F_{\text{propane,e}}$ and $F_{\text{propene,e}}$ stand for the molar flow rate of propane at the inlet, the molar flow rate of propane at the outlet and the molar flow rate of propene at the outlet, respectively.

RESULTS AND DISCUSSION

The surface areas of the catalysts and the calculated VO_x surface coverage, θ_s (fraction of a theoretical monolayer), are given in Table I as a function of the V_2O_5 content. The coverage was calculated using 4.98×10^{14} V_2O_5 molecules cm^{-2} required for full coverage of the surface and the initial BET area of the used γ-Al_2O_3.[32] Theoretically, a V_2O_5 loading of 18.1 wt. % is sufficient to form a monolayer of vanadia over the γ-Al_2O_3 support. As can be seen from Table I, in all cases, the vanadia loading was well below that theoretically required for monolayer coverage; thereby, no drastic changes were observed in the BET results. This could be attributed to the good dispersion of vanadia on the γ-alumina surface. A vanadia loading of 2.7 wt. % slightly increased the surface area but a further increase in the vanadia loading led to an approximately 10 % decrease in the surface area. It could be stated that at first, on introduction of vanadia species to the surface of the alumina support, new pores were formed that led to a slight increase in the surface area. With further increase in the surface density of vanadia, although new pores and surfaces were formed, due to blockage of pore opening at the surface and exclusion of the interior surface area, the total surface area decreased.

TABLE I. Vanadium content, surface area (S_{BET}), calculated VO_x surface density (θ_s), and T_{TPR} for the studied supported vanadium oxide catalysts

Sample	Loading, wt. %	S_{BET} / m^2 g^{-1}	θ_s / %	T_{TPR} / °C
γ-Alumina	–	120.15	–	–
VAl-01	2.7	122.24	15	537
VAl-02	5.4	104.39	30	528
VAl-03	9	104.20	50	522

According to the XRD patterns presented in Fig. 1, only peaks related to the support were detectable and no V_2O_5-related peaks could be observed. Small

V_2O_5 crystallites may be formed on the surface but they are not detectable by the X-ray diffraction method due to their poor crystallinity, small unit cell size and low concentration.[14,31] In accordance with the BET results, it was confirmed that a monotonous dispersion of V_2O_5 species was obtained.

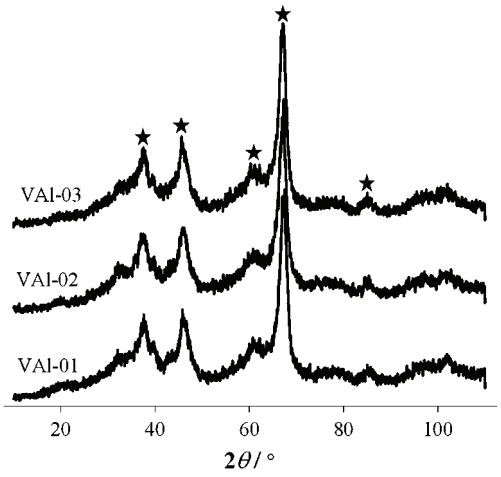


Fig. 1. XRD patterns of the catalyst samples (★: γ-alumina).

The TPR results are presented in Fig. 2. All catalysts start to be reduced at about 350 °C. On increasing the vanadia content, the maximum reduction temperature peak shifted slightly to the left, from 537 °C for VAl-01 to 528 °C and 522 °C for VAl-02, and VAl-03, respectively, indicating better reducibility and higher activity.

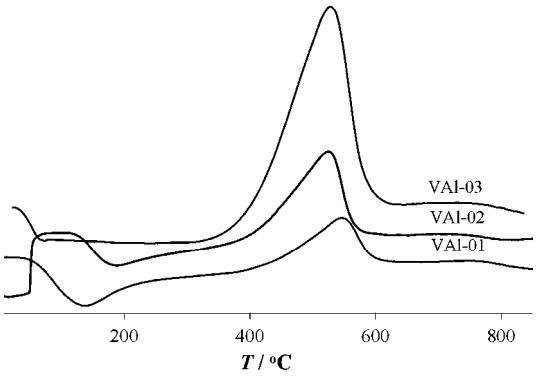


Fig. 2. TPR patterns of the catalyst samples.

The conversion profiles for the three different catalysts at the stoichiometric feed ratio (N_2:O_2:C_3 = 7:1:/2) are shown in Fig. 3a. As the vanadia content of the catalyst increased, the reaction temperature decreased. This could be interpreted as increasing activity with increasing vanadia loading. This result is consistent with the result of the TPR test. On the other hand, as the vanadia content inc-

reased, the selectivity to propene decreased (Fig. 3b). The large surface area of γ-alumina improves the dispersion of active sites over the surface with sufficient distance between them to depress the extent of deep oxidation of species to CO_x. As the loading of vanadia increased so did the probability of the formation of adjacent Brønsted-containing active sites.[2,19] This may be due to the formation of micro-crystallites of vanadia on the surface of γ-alumina on increasing the loading.[14,20,31] Consequently, deep oxidation to CO_x leads to a decreased selectivity towards propene.

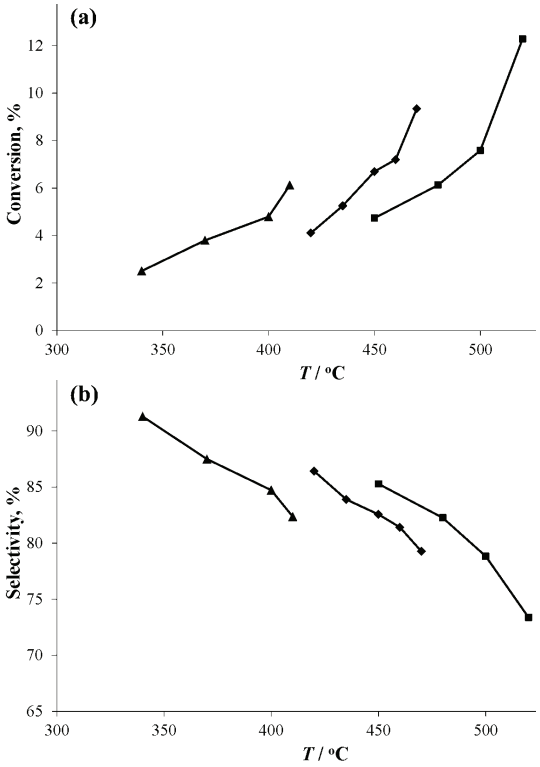


Fig. 3. a) Propane conversion and b) propene selectivity over the studied catalysts. Reaction conditions: gas hourly space velocity = 1.67×10^{-2} m^3 kg^{-1} s^{-1}, feed composition: N_2:O_2:C_3 = = 8:1:1. ■) VAl-01, ◆) VAl-02 and ▲) VAl-03.

A thorough investigation of the conversion profile *vs.* temperature showed that as the temperature increases a shift from low conversions to high conversions occurs. The same transient behavior was observed for selectivity to propene but with the transition from high selectivities to low selectivities (Fig. 4). All three catalysts showed same behavior at different feed compositions from propane rich, stoichiometric, to propane deficient conditions. The temperature at which the transition occurred, T_{tran}, for each catalyst was almost constant under different feed conditions but on increasing the vanadia loading, T_{tran} decreased. This could be related to the higher activity at higher catalyst loadings. Irrespective of whether the feed was reducing (propane rich, high C_3/O_2 ratio) or oxi-

dizing (propane deficient, low C_3:O_2 ratio), the conversion and selectivity showed different trends above T_{tran}. The performances of all three catalysts were

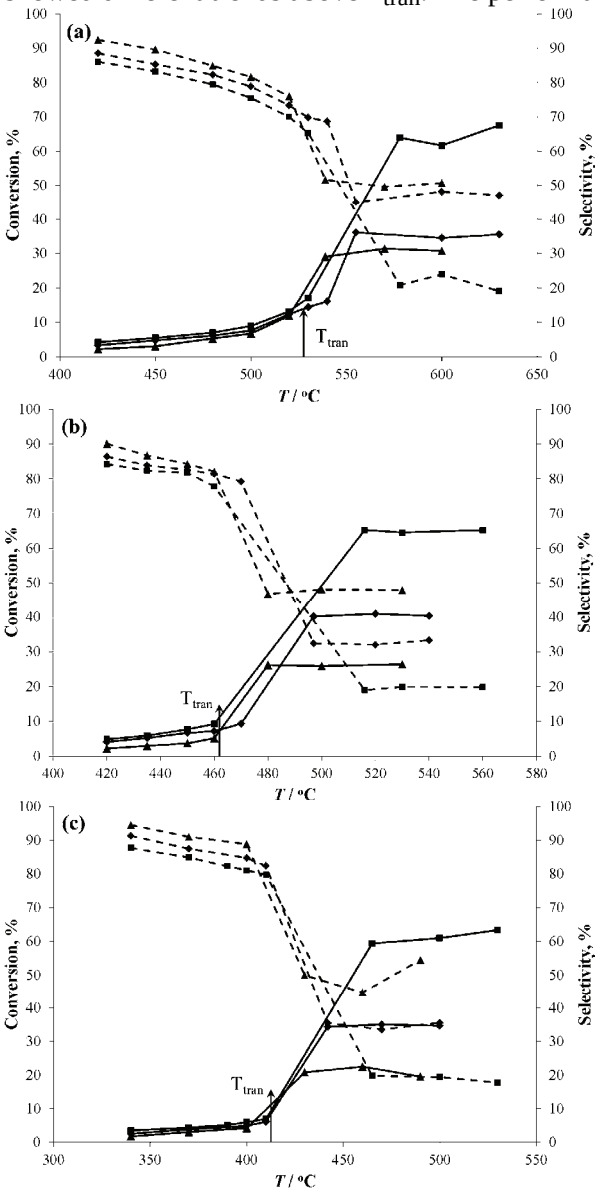


Fig. 4. The effect of the feed composition (N_2:O_2:C_3) on propane conversion (solid line) and propene selectivity (dashed line) over the catalyst samples: a) VAl-01, b) VAl-02 and c) VAl--03). Reaction condition: GHSV = 1.67×10^{-2} m^3 kg^{-1} s^{-1}; N_2:O:C_3 – ■) 7:2:1, ◆) 8:1:1 and ▲) 7:1:2.

similar at low C_3:O_2 ratios and they show same conversion and selectivity. Whilst with increasing C_3:O_2 ratio, the different catalysts show distinguishable performances. As the feed composition approached the high C_3:O_2 ratio, the step change caused by increasing temperature was hindered.

The yield profile *vs*. temperature is shown in Fig. 5. The yield increased with temperature, underwent a sudden increase at T_{tran} and then remained almost constant. Although VAl-01 exhibited the lowest activity amongst catalyst samples, it showed the highest yield as the temperature increased above T_{tran}. Investigation of Y_{tran} (yield above T_{tran}) at different feed compositions indicated the existence of a maximum value for yield on increasing C_3/O_2 (Fig. 6a). The reaction feed compositions varied from N_2:O_2:C_3 = 6:1:3 to 6:3:1, but except for VAl-02, all tests were performed in the range N_2:O_2:C_3 = 7:1:2 to 7:2:1. In comparison with the UEL (upper explosion limit) of propane (9.5 %), all propane mole fractions in reaction mixtures were above the UEL except for N_2:O_2:C_3 = = 6:3:1. The low yield of propene above T_{tran} at a C_3:O_2 ratio of 1:3 could be due to operating at temperatures above the AIT (auto ignition temperature) of propane (470 °C), which resulted in homogeneous combustion. As C_3:O_2 reaches 1:2, a sharp increase in yield was observed. The highest yield was attained at a C_3:O_2 of 1:1 and the yield decreased with further increasing of C_3:O_2. The highest yield occurred under propane deficient conditions in the range of 1:2 < < C_3:O_2 < 2:1 (Fig. 6b). VAl-01 showed the highest yield amongst the three catalysts so that a maximum yield of 17 % could be achieved at a C_3:O_2 ratio of 1:1 above 550 °C.

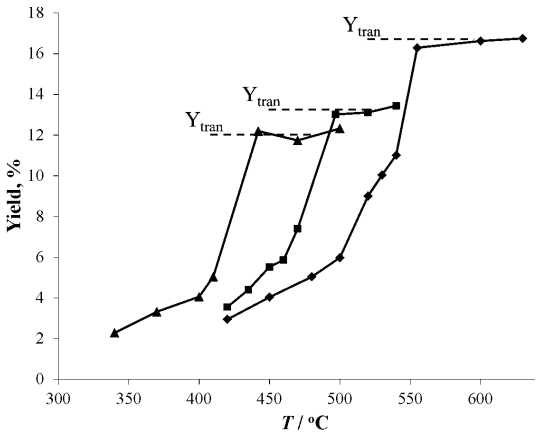


Fig. 5. Propene yield variation with temperature. Reaction conditions: GHSV = 1.67×10^{-2} m^3 kg^{-1} s^{-1} and feed composition: N_2:O_2:C_3 = 8:1:1. ◆) VAl-01, ■) VAl-02 and ▲) VAl-03).

The yield to propene is the result of propane conversion and propene selectivity. Thus, to investigate yield variations, the separate effects of temperature and partial pressure of species on the reaction rates of the POD network must be taken into consideration. As stated, it is believed that the POD reactions occur *via* the MvK mechanism in which abstraction of a hydrogen from propane is the rate determining step; in other words, the activation energy of the selective oxidation of propane to propene is greater than the activation energy of deep oxidation of propene.[15,47] Therefore, the selective oxidation of propane is more sensitive to

temperature so that higher temperatures elevate the selectivity to propene at constant conversion. On the other hand, a higher partial pressure of the reacting spe-

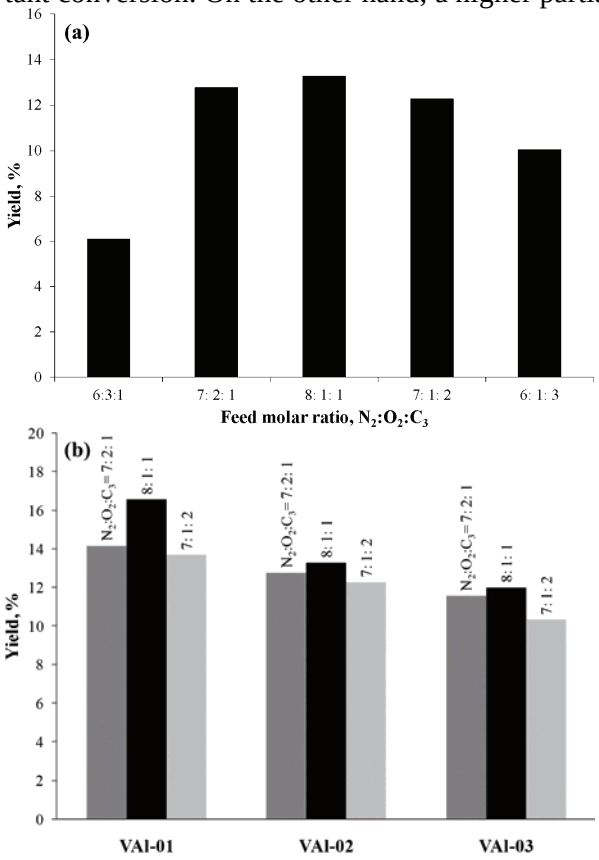


Fig. 6. a) Variation of the maximum propene yield with feed ratio in the case of the VAl-02 catalyst and b) comparison of the maximum propene yield amongst the catalyst samples at different feed ratios. Reaction condition: gas hourly space velocity = = 1.67×10^{-2} m^3 kg^{-1} s^{-1}.

cies accelerates the reaction rate and conversion. Since oxygen is consumed in both selective and deep oxidation, operation at higher oxygen partial pressure (C_3:$O_2 \leq$ 1:2) eases the re-oxidation of the reduced surface and consequently encourages the undesired deep oxidation of the produced propene, which means reduced selectivity. Although under stoichiometric/propane rich conditions (C_3:$O_2 \geq$ 2:1), selectivity to propene is relatively higher, the catalyst surface is in a reduced state which limits conversion. It seems that at C_3:O_2 = 1:1, surface reduction and re-oxidation are balanced in such a way that the propane oxydehydrogenation reaction, which determines conversion, and further deep oxidation of the produced propene, which controls selectivity, are leveled. Hence, the overall yield at C_3:O_2 = 1:1 was superior to those obtained under other conditions.

CONCLUSION

The support, active metal, and promoter have determining roles in the performance of catalysts in terms of the achievement of higher conversions and selectivity towards the desired product. However, these are not all that can affect the performance of catalysts. Tuning the operating temperature and feed composition could also help to increase the overall yields.

Increasing the vanadia content leads to higher activity of catalyst as concluded from the higher conversion at lower temperatures. However, as the activity increases with vanadia content, the selectivity to propene declines. By adjusting the feed composition at propane deficient conditions and operation at higher temperatures, the overall yield increases. The highest propene yield of about 17 % was achieved using VAl-01 with 2.7 wt. % content of vanadia and operating at a C_3:O_2 ratio of 1:1 at temperatures above 550 °C.

И З В О Д

ДОБИЈАЊЕ ВЕЋЕГ ПРИНОСА ПРОПЕНА ПОДЕШАВАЊЕМ РАДНИХ УСЛОВА ОКСИДАТИВНЕ ДЕХИДРОГЕНИЗАЦИЈЕ ПРОПАНА НА V_2O_5/γ-Al_2O_3

YOUSEF ZAYNALI и SEYED MEHDI ALAVI

Chemical Engineering Department, Iran University of Science and Technology, P. O. Box 16765-163, Tehran, Iran

Чврсти катализатор ванадијум-пентоксид је успешно синтетисан импрегнацијом γ-глинице ради испитивања оксидативне дехидрогенизације пропана. Добијени катализатор је окарактерисан следећим методама: XRD, BET и TPR. У широком опсегу температура (340 до 630 °C), испитиван је утицај различитих количина ванадијум-пентоксида (2,7, 5,4, и 9 мас. %) и односа кисеоник:пропан (3:1 до 1:3) на конверзију пропана, као и на принос пропена на атмосферском притиску. Добијени резултати указују на то да са порастом садржаја ванадијум-пентоксида расте активност катализатора док селективност за пропен монотоно опада. Принос пропена расте са порастом температуре од 340 до 630 °C у случају свих испитиваних катализатора. Принос пропена достиже максимум при промени односа пропан:кисеоник од 3:1 до 1:3. Принос расте са порастом парцијалног притиска кисеоника до еквимоларног односа пропан:кисеоник, а затим опада са даљим порастом парцијалног притиска кисеоника. Максимални принос пропена од 17 % је добијен у случају катализатора који садржи 2,7 мас. % ванадијум-пентоксида на температури од 550 °C.

(Примљено 5. маја, ревидирано 25. јуна, прихваћено 27. јуна 2014)

REFERENCES

1. B. Susan, *The Merck Index*, 12th ed., Merck, Rahway, NJ, 1996, p. 1348
2. G. Martra, F. Arena, S. Coluccia, F. Frusteri, A. Parmaliana, *Catal. Today* **63** (2000) 197
3. H. Dai, A. T. Bell, E. Iglesia, *J. Catal.* **221** (2004) 491
4. S. N. Koc, G. Gurdag, S. Geissler, M. Guraya, M. Orbay, M. Muhler, *J. Mol. Catal., A* **225** (2005) 197
5. M. C. Abello, M. F. Gomez, O. Ferretti, *Appl. Catal., A* **207** (2001)421
6. B. Y. Jibril, S. M. AL-Zahrani, A. E. Abasaeed, R. Hughes, *Catal. Commun.* **4** (2003)579

7. B. Y. Jibril, M. C. Al-Kinany, S. H. Al-Khowaiter, S. A. Al-Drees, H. A. Al-Megren, M. A. Al-Dosari, R. H. Al-Rasheed, S. M. Al-Zahrani, A. E. Abasaeed, *Catal. Commun.* **7** (2006) 79
8. A. Al-Dosari, R. H. Al-Rasheed, S. M. Al-Zahrani, A. E. Abasaeed, *Catal. Commun.* **7** (2006) 79
9. E. Rombi, D. Gazzoli, M. G. Cutrufello, S. De Rossi, I. Ferino, *Appl. Surf. Sci.* **256** (2010) 5576
10. B. Grzybowska-Świerkosz, *Appl. Catal., A* **157** (1997) 409
11. M. A. Bañares, M. V. Martínez-Huerta, X. Gao, J. L. G. Fierro, I. E. Wachs, *Catal. Today* **61** (2000) 295
12. O. Rubio, J. Herguido, M. Menéndez, *Chem. Eng. Sci.* **58** (2003) 4619
13. Z. Zhao, Y. Yamada, A. Ueda, H. Sakurai, T. Kobayashi, *Catal. Today* **93**–**95** (2004) 163
14. D. Shee, G. Deo, *J. Mol. Catal., A* **308** (2009) 46
15. A. Khodakov, B. Olthof, A. T. Bell, E. Iglesia, *J. Catal.* **181** (1999) 205
16. P. Moggia, M. Devillers, P. Ruiz, G. Predieri, D. Cauzzi, S. Morselli, O. Ligabue, *Catal. Today* **81** (2003) 77
17. S. Sugiyama, T. Hashimoto, Y. Tanabe, N. Shigemoto, H. Hayashi, *J. Mol. Catal., A* **227** (2005) 255
18. M. Sarzi-Amade, S. Morselli, P. Moggi, A. Maione, P. Ruiz, M. Devillers, *Appl. Catal., A* **284** (2005) 11
19. A. Klisinska, K. Samson, I. Gressel, B. Grzybowska, *Appl. Catal., A* **309** (2006) 10
20. E. V. Kondratenko, M. Cherian, M. Baerns, *Catal. Today* **112** (2006) 60
21. O. Schwarz, D. Habel, O. Ovsitser, E. V. Kondratenko, C. Hessd, R. Schomäcker, H. Schubert, *J. Mol. Catal., A* **293** (2008) 45
22. R. Sasikala, V. Sudarsan, T. Sakuntala, Jagannath, C. Sudakar, R. Naik, S. R. Bharadwaj, *Appl. Catal., A* **350** (2008) 252
23. I. V. Mishakov, A. A. Vedyagin, A. F. Bedilo, V. I. Zaikovskii, K. J. Klabunde, *Catal. Today* **144** (2009) 278
24. S. Arias-Perez, R. Garcıa-Alamilla, M. G. Cardenas-Galindo, B. E. Handy, S. Robles-Andrade, G. Sandoval-Robles, *Ind. Eng. Chem. Res.* **48** (2009) 1215
25. P. Gruene, T. Wolfram, K. Pelzer, R. Schlögl, A. Trunschke, *Catal. Today* **157** (2010) 137
26. M. Høj, T. Kessler, P. Beato, A. D. Jensen, J. D. Grunwaldt, *Appl. Catal., A* **472** (2014) 29
27. H. Zhang, S. Cao, Y. Zou, Y. M. Wang, X. Zhou, Y. Shen, X. Zheng, *Catal. Commun.* **45** (2014) 158
28. S. Chen, F. Ma, A. Xu, L. Wang, F. Chen, W. Lu, *Appl. Surf. Sci.* **289** (2014) 316
29. M. Sun, J. Zhang, C. Cao, Q. Zhang, Y. Wang, H. Wan, *Appl. Catal., A* **349** (2008) 212
30. J. Zhang, M. Sun, C. Cao, Q. Zhang, Y. Wang, H. Wan, *Appl. Catal., A* **380** (2010) 87
31. A. Dinse, B. Frank, C. Hess, D. Habel, R. Schomäcker, *J. Mol. Catal., A* **289** (2008) 28
32. A. A. Lemonidou, L. Nalbandian, I. A. Vasalos, *Catal. Today* **61** (2000) 333
33. G. G. Cortez, J. L. S. Fierro, M. A. Bañares, *Catal. Today* **78** (2003) 219
34. M. Machli, E. Heracleous, A. A. Lemonidou, *Appl. Catal., A* **236** (2002) 23
35. J. Soler, M. L. Nietoj, J. Herguido, M. Menendez, J. Santamaria, *Catal. Lett.* **50** (1998) 25
36. M. Alfonso, M. Menendez, J. Santamaria, *Catal. Today* **56** (2000) 247
37. M. M. Barsan, F. C. Thyrion, *Catal. Today* **81** (2003) 159
38. K. Routray, K. R. S. K. Reddy, G. Deo, *Appl. Catal., A* **256** (2004) 103
39. T. V. M. Rao, G. Deo, *React. Kin. Catal.* **53** (2007) 1538
40. M. A. Vannice, *Catal. Today* **123** (2007) 18

41. D. Creaser, B. Andersson, R. R. Hudgins, P. L. Silveston, *Chem. Eng. Sci.* **54** (1999) 4563
42. S. Sugiyama, Y. Hirata, K. Nakagawa, K. I. Sotowa, K. Maeharad, Y. Himeno, W. Ninomiya, *J. Catal.* **260** (2008) 157
43. O. Ovsitser, R. Schomäcker, E. V. Kondratenko, T. Wolfram, A. Trunschke, *Catal. Today* **192** (2012) 16
44. L. Wang, W. Chu, C. Jiang, Y. Liu, J. Wen, Z. Xie, *J. Nat. Gas Chem.* **21** (2012) 43
45. B. Y. Jibril, A. Atta, S. A. Al-Dress, M. C. Al-Kinany, H. A. Al-Megren, *J. Eng. Res.* **9** (2012) 46
46. A. H. S. Kootenaei, J. Towfighi, A. Khodadadi, Y. Mortazavi, *Appl. Surf. Sci.* **298** (2014) 26
47. A. Dinse, S. Khennache, B. Frank, C. Hess, R. Herbert, S. Wrabetz, R. Schlögl, R. Schomäcker, *J. Mol. Catal.* **307** (2009) 43.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 367–376 (2015)
JSCS–4722
UDC 546.723–138.001:539.24+
537.621:535.4–34
Original scientific paper

Wet milling *versus* co-precipitation in magnetite ferrofluid preparation

LÁSZLÓ ALMÁSY[1], DORINA CREANGA[2]*, CLAUDIA NADEJDE[2], LÁSZLÓ ROSTA[1], EKATERINA POMJAKUSHINA[3] and MANUELA URSACHE-OPRISAN[2]

[1]*Wigner Research Centre for Physics, P. O. Box 49, Budapest – 1525, Hungary,* [2]*Faculty of Physics, "Al. I. Cuza" University, 11A Carol I Bd., RO-700506, Iasi, Romania and* [3]*Laboratory for Development and Methods, PSI, CH-5232 Villigen, Switzerland*

(Received 13 March 2014, accepted 21 May 2014)

Abstract: Various uses of ferrofluids for technical applications continuously raise interest in the improvement and optimization of the preparation methods. This paper deals with the preparation of finely granulated magnetite particles coated with oleic acid in hydrocarbon suspensions following either chemical co-precipitation from iron salt precursors or wet milling of micron size magnetite powder with the goal to compare the benefits and disadvantages of each method. Microstructural measurements showed that both methods gave similar-sized magnetite particles of 10–15 nm. The wet-milled magnetite suspension had a higher saturation magnetization than that obtained in the relatively rapid co-precipitation synthesis. The different efficacies of ferrophase incorporation into kerosene could be related to the different mechanisms of oleic acid bonding to the surface of the nanoparticles. Comparative data showed that wet milling represents a viable alternative to the traditional co-precipitation method since despite of longer processing time, the impact of chemicals on the environment and remnant water in the final product could be avoided.

Keywords: magnetite nanoparticles; kerosene suspensions; X-ray diffraction; TEM; magnetometry; IR spectra.

INTRODUCTION

Magnetic iron oxide nanoparticles and their fluid dispersions have been of scientific and technological interest since the 1960s. One of the most widely used and explored technique for the synthesis of submicron magnetic particles is the co-precipitation method.[1–8] Based on this simple and efficient procedure, condensing divalent and trivalent iron salts in reactions with hydroxide bases, magnetite crystalline nanoparticles form with a relatively broad size distribution. Due to their magnetic moments, these particles tend to aggregate. Agglomeration

* Corresponding author. E-mail: dorina.creanga@gmail.com
doi: 10.2298/JSC140313053A

could be prevented by using electrostatic or steric stabilization. Oleic acid was proven to be the best surfactant that forms a stabilizing shell around the magnetite particles due to its highest affinity to the surface of superfine magnetite compared to other surfactants.[5–8] The particle size and morphology can be changed by varying the parameters of synthesis, such as temperature, pH, concentration of components[2–3] or by using different surfactants for the stabilization in organic solvents.[4] A hydrophilic surface coating is used when aqueous dispersions are required for biomedical applications.[4,9,10]

Although small particles can be easily achieved by chemical precipitation, it has been shown that, in practice, the mechano–chemical method of wet milling is also a simple, efficient and environment-protective method for obtaining finely granulated magnetic ferrite particles.[11–14] In spite of the possible drawback of long grinding times, the mechano–chemical technique offers certain advantages which may place it as a useful method for obtaining nanometric particles in colloid suspensions, in particular, ferrofluids.

In this study, soft magnetic colloids were prepared by the two methods, and their microstructural and magnetic properties are comparatively discussed herein.

EXPERIMENTAL

Synthesis by co-precipitation

Ferrous sulphate heptahydrate (99 %), ferric chloride hexahydrate (99 %), 25 % ammonium hydroxide, oleic acid (90 %), hexane (95 %) and kerosene reagent grade were purchased from Sigma–Aldrich and used without further purification. Distilled water was deionised prior to use.

The first colloidal sample, S1, was obtained by chemical precipitation of magnetite in a multistep process (Fig. S-1 of the Supplementary material to this paper), starting from stoichiometric amounts of aqueous solutions of ferric and ferrous salts that were condensed by reaction with excess base (*i.e.*, ammonium hydroxide NH_4OH) to form magnetite nanoparticles.

Following the Massart procedure,[15] 300 mL of 1 mol L^{-1} $FeSO_4 \cdot 7H_2O$ solution and 300 mL of 2.8 mol L^{-1} $FeCl_3 \cdot 6H_2O$ solution were heated at a temperature of 80–85 °C (under atmospheric conditions) and then 200 mL of 25 % NH_4OH solution – as a precipitation agent – was added drop-wise under continuous stirring. A 20 % excess amount of NH_4OH was used to ensure the formation of magnetite nanoparticles and no other iron oxides. About 30 min after the co-precipitation reaction, phase separation was observed. Following magnetic decantation and repeated washing with deionised water, to remove salt residues, the ferrophase was further washed with ethanol to remove water traces, and subsequently filtered on a paper filter. The main chemical reactions could be given as:

$$2FeCl_3 + 6H_2O \rightarrow 6\ Cl^- + 6H^+ + 2Fe(OH)_3;$$

$$FeSO_4 + 2H_2O \rightarrow SO_4^{2-} + 2H^+ + Fe(OH)_2;$$

$$8NH_4^+ + 6Cl^- + SO_4^{2-} \rightarrow 6NH_4Cl + (NH_4)_2SO_4;$$

$$2Fe(OH)_3 + Fe(OH)_2 + 8OH^- + 6H^+ + 2H^+ \rightarrow Fe_3O_4 + 12H_2O$$

Thus, the mixed iron oxide Fe_3O_4 resulted from intermediate iron hydroxides in the aqueous reaction medium. Next, the magnetite particles were coated with oleic acid 1:6 (*V*/*V*) dispersed in hexane, the mixture being heated up to 80 °C under continuous stirring. After hexane evaporation, kerosene was added as the dispersion fluid, with continuous stirring for 60 min.

Wet milling preparation procedure

The second sample, S2, was obtained by mechano–chemical processing of micrometric magnetite particles (Fig. S-2 of the Supplementary material) in a vibrating ball mill.

The raw material was industrial magnetite powder (residual product from charcoal processing), with an average particle diameter of 1 μm, prepared by hydrothermal method, that was milled within a laboratory device with four sealed hardened steel cups, of about 700 mL each containing 4.5 and 13.5 mm diameter steel balls; mixing was realized by using a 380V/50Hz engine with a power supply of 1.1 kW at rotation speed about 900 rev min^{-1}. About 400 mL kerosene/oleic acid in the ratio of 6:1 (*V*/*V*) were put in every cup over an excessive amount of magnetite and wet milled for about three thousand hours (daily cycles of 12 h milling / 12 h resting). The oleate-coated ferrophase suspended in kerosene was collected from the top of the steel cups after 24 h decantation, while excess ferrophase remained as sludge at the bottom. Further centrifugation was realized at 3500 rev min^{-1} for 10 min.

Physical characterization

Microstructure and magnetic properties of the magnetic colloid samples were characterized by transmission electron microscopy (TEM), powder X-ray diffraction (XRD), magnetometry, and infrared spectrophotometry (FT-IR).

The TEM measurements were performed on a Philips CM100 Biotwin instrument, with 80 kV voltage and 180,000 magnification. The original colloid samples were diluted with toluene, a few droplets were deposited on a carbon-coated copper grid and left to dry over night protected from environmental particle contamination.

The X-Ray powder diffraction patterns were recorded on a Bruker D8 Advance instrument using CuK_α radiation in Bragg–Brentano geometry, with a scanning interval of 0.02°. The fluid samples were deposited on glass plates and dried. After drying, the remaining surfactant gave rise to strong background scattering, thus long measuring times of 14 h were required.

Magnetization investigation was carried out by a Gouy magnetometer[16] at room temperature using an air-tight nonmagnetic cylindrical sample holder of 3 mm in diameter and 25 mm in height, placed perpendicularly to the magnetic field.

The volume fraction of the ferrophase was calculated from the density values of the ferrofluid, kerosene and magnetite. Densities were measured at room temperature by pycnometric method.

The infrared absorption spectra were recorded in KBr discs using Bruker Vertex 70 FT-IR spectrometer.

RESULTS AND DISCUSSION

Electron microscopy

TEM images are shown in Fig. 1a and b. Mainly, quasi-spherical particles were seen in all recorded TEM images for both samples. The surfactant coating gave negligible contrast compared to iron oxide, hence the size of the observed

particles corresponded to the size of the magnetite cores. Better uniformity of the nanoparticles spreading on the support grid was obtained for the wet-milled sample, while groups of closely disposed particles were observed for the co-precipitated sample.

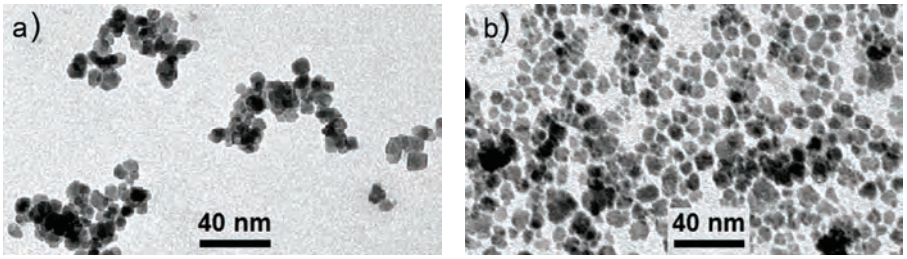


Fig. 1. TEM images of samples: a) S1 and b) S2.

Using several images for every sample, histograms of the particle size distribution were constructed by the manual counting of over a thousand particles (Fig. 2a and b). Both graphs have asymmetric shapes, the distribution tails extending up to 25 nm, while most of the particles were between 5 and 15 nm in diameter. The largest frequency corresponds to 10 nm diameter in sample S1 and to 7 nm in sample S2.

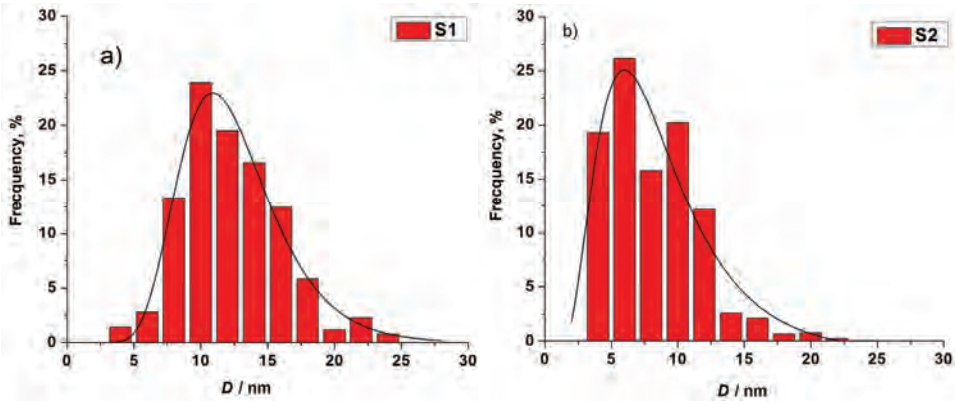


Fig. 2. Size distributions of magnetic nanoparticles obtained by a) the co-precipitation route S1 and b) the mechano-chemical route S2. The solid lines are the fits to the log–normal size distribution.

Log–normal size distribution curves were found to fit fairly well each data set. The mean values of the physical diameter D_{TEM} as well as the standard deviation σ_{TEM} of the fitted log–normal distributions are given in Table I.

TABLE I. Physical and magnetic parameters of the ferrophase in the two prepared ferrofluids

Sample	ϕ / %	D_{XRD} / nm	σ_{XRD} / nm	D_{TEM} / nm
S1	2.3	9.1	0.30	11.7
S2	5.6	8.7	0.30	7.6
	σ_{TEM} / nm	M_S / kA m^{-1}	D_M / nm	σ_M / nm
S1	0.28	3.6	7.7	0.39
S2	0.38	13.9	5.4	0.44

X-ray diffraction

The X-ray diffraction data (Fig. 3) for both samples revealed the typical spinel structure of magnetite with a lattice parameter 0.835 nm, as obtained by Rietveld refinement. No traces of impurities were registered. The line shape was modelled as originating from isotropic particle size broadening using the Fullprof program suite. The effective crystallite sizes of 9.1 and 8.7 nm were estimated for S1 and S2 and they were in agreement with the average size of the particle populations measured by TEM.

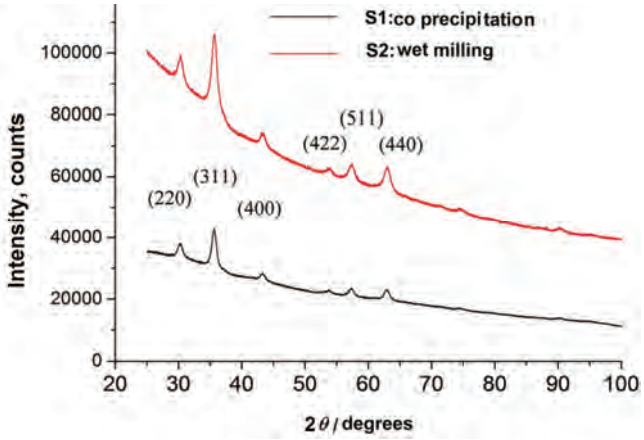


Fig. 3. X-ray diffraction patterns of the dried samples S1 and S2.

Pycnometry

The ferrophase volume fraction, ϕ, was calculated from the densities of the constituent materials according to formula:

$$\phi = \frac{(\rho_F - \rho_L)}{(\rho_S - \rho_L)}$$

where ρ_F, ρ_L and ρ_S are the densities of the ferrofluid, the carrier liquid (kerosene) and of the solid phase (magnetite), respectively. The results revealed that the wet milling procedure resulted in a higher efficiency of magnetite incorporation in kerosene; the volume percentage of magnetite in the wet milling ferro-

fluid S2 being higher than in the co-precipitation fluid S1, *i.e.*, 5.6 compared to 2.3 %.

Magnetometry

Superparamagnetic behaviour without hysteresis loop in the magnetization curves was observed by recording the sample magnetization, M, as a function of the applied magnetic field H up to 500 kA m^{-1}, as shown in Fig. 4. A higher saturation magnetization was obtained for the ferrofluid prepared by wet milling, in accordance with the higher ferrophase volume fraction determined by the pycnometric method. Analyzing the magnetization data, the magnetic moments of the individual particles could be obtained.

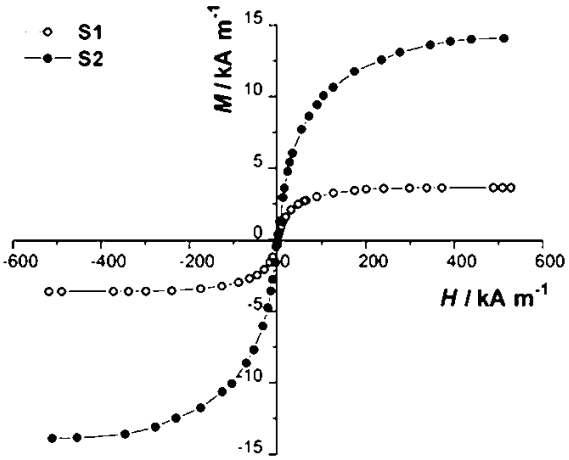


Fig. 4. Magnetization data for the fluid samples S1 and S2 at room temperature.

The magnetization data were modelled using the Langevin function:

$$M / M_S = L(x) = \coth x - \frac{1}{x}$$

where $x = \mu H/k_B T$, in which H is the applied field, k_B is the Boltzmann constant and $\boldsymbol{\mu}$ is the magnetic moment of an individual particle, given by $\boldsymbol{\mu} = m_S \pi D_M^3/6$, where D_M is the particle magnetic size and m_S is the saturation magnetization of bulk magnetite (0.48×10^6 A m^{-1}). The magnetic response of the population of particles obeying a log–normal size distribution was fitted to the measured data; integration was performed over the whole range of particle sizes.[3,4,11] This approach allowed the size distribution to be described by two parameters: the mean value of the particle magnetic size D_M and the width σ_M of the distribution curve. It can be seen (Table I) that the D_M values were smaller than the D_{TEM} ones. The non-magnetic surface layer of the nanoparticles usually results in a smaller magnetic size than the physical size by 1–2 nm.[4,6] The dimensional parameters of the nanoparticles, the saturation magnetizations of the fluids and

the ferrophase volume fractions are presented in Table I. The three characterization techniques show that the two samples contained magnetic nanoparticles with rather similar shape and size distributions, the width of the latter being about 0.4, which is typical for magnetite particles prepared by co-precipitation.[3,4]

Vibrational spectroscopy

The infrared spectra of S1 and S2 together with the spectra of magnetite powder and oleate are shown in Fig. 5a and b. The magnetite skeleton vibrations are seen at the lowest wave numbers, *i.e.*, toward 500 cm^{-1} but also at around 3500 cm^{-1} (Fig. 5a). In the spectrum of the S1 sample prepared by the co-precipitation method, the vibration band at about 3500 cm^{-1} could also arise from OH stretching mode of water molecules (Fig. 5b), this being a consequence of possible traces of water from the technological step of ferrophase synthesis in aqueous medium. The relative diminution of the magnetite band at 3500 cm^{-1} when passing from magnetite powder (Fig. 5a) to S1 and S2 magnetite suspensions (Fig. 5b) could be due to oleate ion bonding to the iron ions of the ferrophase. Moreover, if the intensity of the 3500 cm^{-1} vibrational mode remained relatively higher in the S1 co-precipitation sample compared to the S2 wet milling sample (Fig. 5b), this could be the consequence of the possible lower efficiency of oleate ion bonding in co-precipitation method as well as of the higher content of traces of water arising from the technological step of ferrophase preparation in aqueous medium (missing in the case of the mechano–chemical procedure).

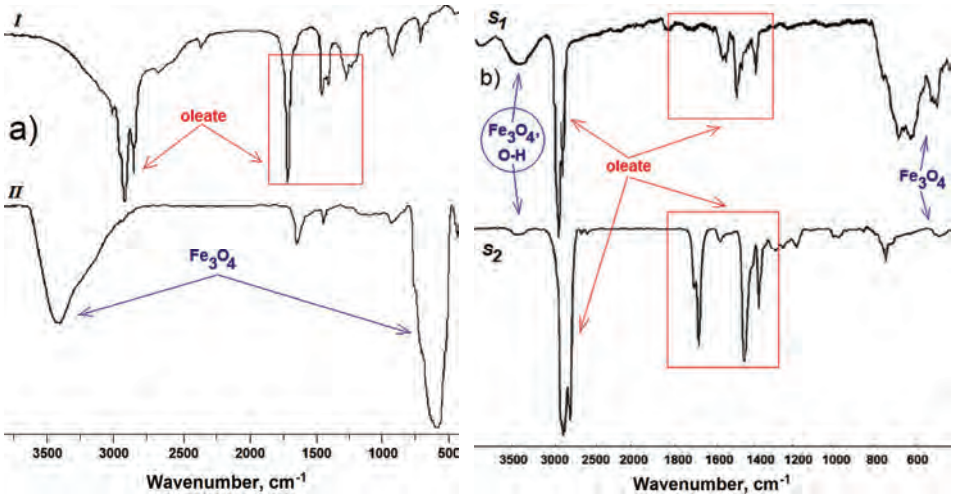


Fig. 5. Vibration spectra of a) the ferrofluid components: magnetite and oleate and b) the samples S1 and S2.

It could be assumed that the stretching vibrations of the C–H bonds of the oleate ion, both symmetric and asymmetric ones, at about 2950 cm^{-1} (Fig. 5a) as well as those from 1500–1700 cm^{-1} range (carboxyl stretching and bending vibrations) were slightly shifted toward smaller wavenumbers for the colloidal suspensions (Fig. 5b) as a consequence of the magnetite core – organic shell interactions. In addition, the relatively higher intensity of these vibration bands in S2 compared to S1 suggests the higher efficiency of ferrophase coating in the mechano–chemical procedure when water traces are theoretically missing.

The band with the highest wavenumber from this group could be assigned to the deformation vibration of hydroxyl from the carboxyl group, the corresponding interaction being formed during oleate attachment. The remarkably higher relative intensity of this band in the spectrum of the S2 sample compared to that in the spectrum of S1 suggests that possibly non-bound oleic acid remained in the final sample. The amount of the excess surfactant could be decreased by improving the preparation protocol of the wet milling procedure to include an additional dialysis step – which will be included in future investigations.

Mechanisms of co-precipitation and the mechano–chemical synthesis

In the case of conventional chemical route, the intimate contact of surfactant molecules with nanoparticles depends in principal on two factors: heat supply and stirring efficacy in the sample volume. However, neither of these is easy to control: during intensive heating, the hydrocarbon dispersion liquid evaporates and needs to be carefully added in new small aliquots, while the stirrer makes inhomogeneous contacts with the parts of the suspension volume during stirring. An inert gas atmosphere is required to avoid oxygen reaction with iron during the long stirring times.

Regarding the mechanochemical method, it should be mentioned first that the high kinetic energy of the milling balls favours the occurrence of reactions at temperatures lower than those required for conventional syntheses.[17,18] During the milling procedure, the innumerable ball–particle collisions favour chemical adsorption of surfactant each time at a different small site (close to 10^{-9} m^3)[19] of a given particle. Inelastic collisions between the particles and the balls fractures the particles leading to the formation of fresh interfaces between the reaction phases by means of dynamic deformation, fracturing, and cold welding, while energies of almost 1 J could develop during intervals of 10^{-4}–10^{-5} s.[18,19]

According to Guler and Evin,[20] the milling intensity is an important parameter in the occurring reactions. Maurice and Courtney[21] developed a computational modelling stating that the deformation size in a powder particle that is trapped between balls could be expressed as a function of the distance from the particle centre to the contact area, the relative impact speed, the ball density and the hardness of the particle. The frictional heat dissipated over collision areas is

transferred from balls to particles that empowered the exposed iron ions to react with local surfactant molecules. The ball surface could also be affected but the steel balls release only iron ions; in this respect, it could be hypothesized that such iron ions would also attach to oleate ions so that free oleate vibrational modes are overlapped with those of oleate–iron and oleate–magnetite particles, affecting the intensities in the infrared absorption spectrum.

CONCLUSIONS

Taking into account all above results, it may be concluded that the two preparation methods yield magnetic colloid particles having rather similar quasi-spherical shape and sizes of around 10 nm. The volume fraction of the ferrophase and the saturation magnetization of the colloidal fluids were different: the two times higher volume fraction and the slightly different size distribution of the particles resulted in more than three times higher saturation magnetization in the sample prepared by wet milling. These results show that the mechano–chemical method may be preferred in certain practical situations. With this method, no synthesis of magnetite is required, unlike in synthesis methods involving milling, in which mechanical milling is used for the formation of the ferrite phase from precursors. From this point of view, the process is simpler, as no chemical transformations occur except the bonding of the surfactant molecules to the magnetite surface. In certain cases, the long milling times could be an acceptable disadvantage, outbalanced by the higher efficiency of ferrophase incorporation by coating with oleate ions, and easy control of the particle size by varying the milling time.

SUPPLEMENTARY MATERIAL

Schematic presentations of the synthesis of magnetite are available electronically from http://www.shd.org.rs/JSCS/, or from the corresponding author on request.

Acknowledgements. This work was partially realized with the financial support received from the National University Research Council CNCSIS, Romania (grant PNII–IDEI-2021, no. 474/2009), and the National Office for Research and Technology, Hungary (grant NAP-VENEUS05). The financial support of the Adolphe Merkle Foundation is highly appreciated.

ИЗВОД

МОКРО МЛЕВЕЊЕ НАСПРАМ КОПРЕЦИПИТАЦИЈЕ – СИНТЕЗА МАГНЕТИТ-ФЕРОФЛУИДА

LÁSZLÓ ALMÁSY[1], DORINA CREANGA[2], CLAUDIA NADEJDE[2], LÁSZLÓ ROSTA[1], EKATERINA POMJAKUSHINA[3] и MANUELA URSACHE-OPRISAN[2]

[1]*Wigner Research Centre for Physics, POB 49, Budapest – 1525, Hungary,* [2]*Faculty of Physics, "Al. I. Cuza" University, 11A Carol I Bd., RO-700506, Iasi, Romania и* [3]*Laboratory for Development and Methods, PSI, CH-5232 Villigen, Switzerland*

Различите техничке примене ферофлуида доводе до пораста интересовања за унапређење и оптимизацију метода синтезе. У овом раду испитивана је припрема фино гранулисаних честица магнетита превучених олеинском киселином у суспензијама угљово-

дника након хемијске копреципитације из соли гвожђа као прекурсора и након мокрог млевења праха магнетита микронских димензија са циљем да се упореде предности и ограничења ове две методе. Микроструктурна мерења су показала да обе методе дају магентит сличне величине честица, 10–15 nm. Вишу вредност магнетизације при сатурацији показује суспензија магнетита добијена мокрим млевењем у поређењу са релативно брзом синтезом копреципитацијом. Различита ефикасност инкорпорације ферофазе у керозин може потицати од различитог механизма везивања олеинске киселине на површини наночестица. Упоредна анализа показује да мокро млевење представља практичну алтернативну традиционалној копреципитацији јер се, и поред дужег времена процесирања, може се избећи хемијски утицај на животну средину као и трагови воде у финалном производу.

(Примљено 13. марта, прихваћено 21. маја 2014)

REFERENCES

1. R. E. Rosensweig, *Ferrohydrodynamics*, Cambridge University Press, Dover Publications, Inc., Mineola, New York, 1985
2. I. Nyirő-Kosa, D. Csakberenyi Nagy, M. Posfai, *Eur. J. Min.* **21** (2009) 293
3. K. Parekh, *Indian J. Pure Appl. Phys.* **48** (2010) 581
4. M. V. Avdeev, D. Bica, L. Vekas L, V. L. Aksenov, A. V. Feoktystov, O. Marinica, L. Rosta, V. M. Haramus, R. Willumeit, *J. Colloid Interf. Sci.* **334** (2009) 37
5. L. Zhang, R. He, R., H. C. Gu, *Appl. Surf. Sci.* **253** (2006) 2611
6. V. B. Barbeta, R. F. Jardim, P. K. Kiyohara, F. B. Effenberger, L. M. Rossi, *J. Appl. Phys.* **107** (2010) 073913
7. P. Dutta, S. Pal, M. S. Seehra, N. Shah, G. P. Huffman, *J. Appl. Phys.* **105** (2009) 07B501
8. L. Vekas, D. Bica, M. V. Avdeev, *China Particuology* **5** (2007) 43
9. R. Omidirad, F. H. Rajabi, B. V. Farahani, *J. Serb. Chem. Soc*. **78** (2013) 1609
10. M. Racuciu, D. E. Creanga, A. Airinei, D. Chicea, V. Badescu, *Mater. Sci. – Poland* **3** (2010) 609
11. G. F. Goya, *Solid State Commun.* **130** (2004) 783
12. E. Petrovsky, M. D. Alcala, J. M. Criado, T. Grygar, A. Kapička, J. Šubrt, *J. Magn. Magn. Mater.* **210** (2000) 257
13. N. Ghazanfari, A. Kılıc, S. Ozcan, H. Sözeri, H. Özkan, A. Gencer, *AIP Conf. Proceedings* **929** (2007) 133
14. M. M. Can, S. Ozcan, A. Ceylan, T. Firat, *Mater. Sci. Eng., B* **172** (2010) 72
15. R. Massart, *IEEE Trans. Magn.* **17** (1981) 1247
16. C. D. H. Williams, S. R. Hoon, J. S. Thorp, *J. Mater. Sci. Lett.* **5** (1986) 832
17. F. Urakaev, V. V. Boldyrev, *Powder Technol.* **107** (2000) 197
18. T. H. Courtney, D. Maurice, *Scripta Mater.* **34** (1996) 5
19. P. Butyagin, *Colloids Surfaces, A* **160** (1999) 107
20. O. Guler, E. Evin, *Optoelectron. Adv. Mat.* **6** (2012) 183
21. D. Maurice, T. H. Courtney, *Metall. Mater. Trans., A* **26** (1995) 2437.

Journal of the Serbian Chemical Society
JSCS@tmf.bg.ac.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) S103–S104 (2015) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Wet milling *versus* co-precipitation in magnetite ferrofluid preparation

LÁSZLÓ ALMÁSY[1], DORINA CREANGA[2]*, CLAUDIA NADEJDE[2], LÁSZLÓ ROSTA[1], EKATERINA POMJAKUSHINA[3] and MANUELA URSACHE-OPRISAN[2]

[1]*Wigner Research Centre for Physics, P. O. Box 49, Budapest – 1525, Hungary,* [2]*Faculty of Physics, "Al. I. Cuza" University, 11A Carol I Bd., RO-700506, Iasi, Romania and* [3]*Laboratory for Development and Methods, PSI, CH-5232 Villigen, Switzerland*

J. Serb. Chem. Soc. 80 (3) (2015) 367–376

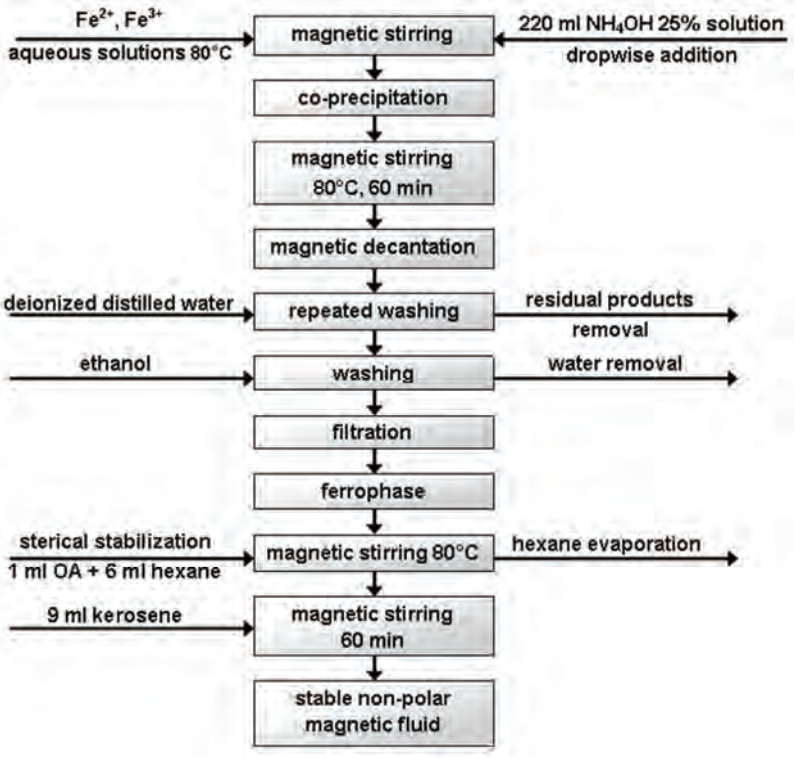


Fig. S-1. Co-precipitation synthesis of magnetite colloid by stabilization in kerosene with oleate surfactant.

* Corresponding author. E-mail: dorina.creanga@gmail.com

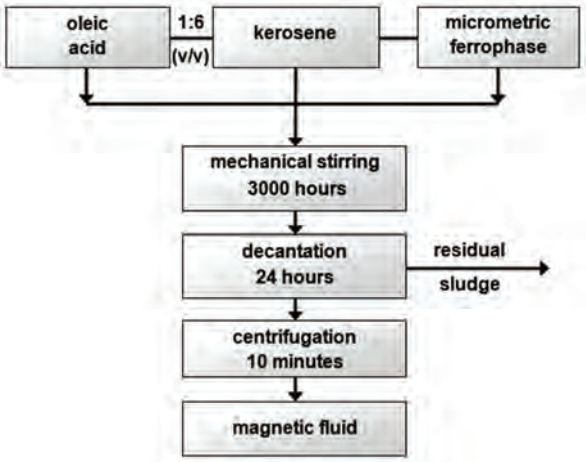


Fig. S-2. Simultaneous size reduction and coating in a one step mechano-chemical procedure.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 377–389 (2015)
JSCS–4723
UDC 662.75+662.61+66.023+544.4
Original scientific paper

Euler–Euler granular flow model of the combustion of liquid fuels in a fluidized reactor

STEVAN NEMODA[1]*, MILICA MLADENOVIĆ[1], MILIJANA PAPRIKA[1], DRAGOLJUB DAKIĆ[2], ALEKSANDAR ERIĆ[1] and MIRKO KOMATINA[3]

[1]*"Vinča" Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia,* [2]*Innovation Center, Faculty of Mechanical Engineering, University of Belgrade, Belgrade, Serbia and* [3]*Faculty of Mechanical Engineering, University of Belgrade, Belgrade, Serbia*

(Received 30 January, revised and accepted 27 March 2014)

Abstract: This paper deals with the numerical simulation of liquid fuel combustion in a fluidized reactor using two-fluid Eulerian–Eulerian fluidized bed modeling incorporating the kinetic theory of granular flow (KTGF) to gas and solid phase flow prediction. The comprehensive model of the complex processes in a fluidized combustion chamber incorporates, besides the prediction of gas and particular phase velocity fields, the energy equations for the gas and solid phase and the transport equations of conservation of chemical species with the source terms due to the conversion of chemical components. Numerical experiments showed that the coefficients in the model of inter-phase interaction drag force have a significant effect, and they have to be adjusted for each regime of fluidization. A series of numerical experiments was performed with combustion of liquid fuels in a fluidized bed (FB), with and without significant water content. The given estimations were related to the unsteady state, and the modeled period corresponds to the passing time of the flow through the reactor column. The numerical experiments were conducted to examine the impact of the water content in a liquid fuel on the global FB combustion kinetics.

Keywords: computational fluid dynamics model; combustion; fluidized bed; two-fluid model; non-conventional fuel.

INTRODUCTION

Bubbling fluidized bed reactors with gas–solid particles is usually employed in industrial operations, such as energy production and petrochemical processes. Lately these reactors were repeatedly used for the thermal disintegration (incineration) of industrial waste and by-products. There are many benefits of fluidized bed combustion of unconventional fuels (with high amounts of water and other

* Corresponding author. E-mail: snemoda@vinca.rs
doi: 10.2298/JSC140130029N

ballast matter). The high thermal capacity of the fluidized bed, the thermal conductivity, and the intensity of heat transfer between the inert bed material and the fuel facilitates a stable combustion process of a wide variety of unconventional fuels, accompanied by a low sensitivity to fuel quality. The zone of intensive combustion occupies a relatively small volume, because most of the fuel burns in the bed itself, and burning-off in the "splash" zone and above the bed. In addition, FB furnaces operate at lower temperatures (≈ 850 °C) that are optimal from the aspect of decreased NO_x emission in the flue gases. Moreover, these furnaces are favorable from the viewpoint of the efficiency of in-bed desulfurization,[1] when it is necessary. For these reasons, this technology is recommended by the EU for waste matter combustion.

Experimental methods and numerical simulations are equally employed in research and development in the fields of energy and process engineering. The numerical models provide great opportunities for saving resources and time in the development of facilities and technologies in these fields. However, it should be noted that the numerical tools for simulation of complex processes, such as fluidized bed (FB) combustion, are not completely developed due to difficulties in describing complex two-phase flow and the specificity of heat and mass transfer in bubbling fluidized bed (BFB). In addition, many of existing numerical methods are not compliant to engineering needs, because they are complex and require expensive computer equipment, and are therefore not suitable for development and engineering applications.

There are two main approaches for computational fluid dynamics (CFD) modeling of gas–solid hydrodynamics. The first one is the Lagrangian–Eulerian modeling approach, also called discrete particle modeling (DPM), which solves the equations of motion individually for each particle, whereby the continuous phase is modeled using an Eulerian framework and the trajectories of the particles are simulated within a Lagrangian framework.[1–3] In large systems of particles, the Lagrangian–Eulerian model requires powerful computational resources because of the number of the equations that are to be solved. The second approach for modeling gas–solid flows is Eulerian–Eulerian modeling,[4–7] also called granular flow modeling (GFM), which assumes that both phases can be considered as fluids and takes the interpenetrating effect of each phase into consideration by using drag models. Consequently, the application of a proper drag model in Eulerian–Eulerian modeling is very important. The Eulerian two-fluid approach is an extension of the fluid dynamics formulation of single phase to multiphase flow. Particles in gas–solid flow may be treated as magnified molecules, and the analogy of their behavior to gas molecules is the reason for the wide use of the kinetic theory of granular flow (KTGF) for modeling the motion of particles. The KTGF approach is based on the concept of granular temperature. The granular temperature measures these random oscillations of the

particles and it is defined as the average of the three variances of the velocities of a particle. A full mathematical description of the kinetic theory is provided elsewhere.[8] In spite of detailed mathematical modeling of the complex processes in FB, the drag laws used in two-fluid models are semi-empirical in nature. Therefore, it is essential to use a drag law that correctly predicts the minimum fluidization conditions where the particles are in a state of suspension because of the balance between interfacial drag and body forces. The inter-phase interaction drag force model by Syamlal–O'Brien[9] is often used; the coefficient between fluid and solid (granular) phase in that model depends only on the phase void fraction, but not on the fluidization conditions.

In the present study, the GFM approach was chosen for the simulation because it requires less computational time and capacities. Here, the presented numerical experiments are related to engineering issues, for which detailed computations are not necessary. The different coefficients of the Syamlal–O'Brien inter-phase interaction drag force model were applied for the numerical simulation of a fluidized reactor, depending on the conditions of the fluidization.[10] The fluidization conditions were varied by applying different bubble fluidization intensities, bed particle sizes, gas composition, and temperature conditions. The investigation also includes numerical experiments of different regimes of liquid fuels combustion in a FB furnace for combustion of non-conventional liquid fuels. In this part, special attention was paid to the influence of the water content in the fuel on the position of the intense combustion zone and the global reaction rate.

Euler–Euler granular model of the fluidized bed

The Euler–Euler fluidized bed modeling approach considers the gas and FB dense phase (gas–particle system under conditions of the minimum fluidization) as two fluids with different characteristics. In the transport equations for transfer of momentum of the effective fluid (the FB dense phase), fluid–particle interactions under conditions of the minimum fluidization velocity were modeled, as well as the interaction between the particles themselves. In the Eulerian–Eulerian approach, all phases have the same pressure and that is the pressure of the continuous-primary phase. This model solves the continuity and momentum equations for each phase, and tracks the volume fraction. Furthermore, an additional transport equation for the granular temperature (which represents the fluctuating energy of the solids) is solved, and the bulk and shear viscosity of the solids are determined using the kinetic theory of gases on granular flow.

Moreover, it is necessary to define the coefficients for calculating the inter-phase interaction term. For modeling the interactions between gas and particle phases, within the suggested Euler–Euler granular approach to fluidized bed modeling, the routines incorporated in the modules of the commercial CFD software package Fluent 6.3.26 were used. This code allows for the presence of

several phases within one control volume of the numerical grid, by introducing the volume fraction of each phase. The solid phase represents a granular layer made of spherical particles of uniform diameter. The mass and momentum conservation equations are solved for each phase separately.

The basic and constitutive equations of the two-fluid granular model of a fluidized bed are given in Supplementary material to this paper.

NUMERICAL TESTS OF THE EULER–EULER GRANULAR MODEL FOR DIFFERENT FLUIDIZATION CONDITIONS

Despite rigorous mathematical modeling of the associated physics, the drag laws used in the model continue to be semi-empirical in nature. The semi-empirical procedure is proposed primarily for the prediction of drag law coefficients that correspond to real minimum fluidization conditions. The constants 0.8 and 2.65 in the coefficient *B* of the Syamlal–O'Brien inter-phase interaction drag force model (Eqs. (12) and (13) of the Supplementary material) are not universal, particularly when it comes to the fluidization regimes with a multi-component fluid and under non-isothermal conditions. For this reason, the presented numerical experiments were performed with 2D simulations of the fluidization columns with different particle diameters and for various temperatures and gas composition for fluidization, in order to analyze the influence of the various fluidization conditions on the choice of most suitable values of the coefficient B constants.

For the basic multiphase calculations, a professional package (Fluent 6.3.26) was used, but the inter-phase interaction drag force model (Eqs. (12) and (13) of the Supplementary material) were included in the calculation with their own code using the UDF options of Fluent.

The numerical solving of the governing equations of the Fluent Euler–Euler granular model (Eqs. (1)–(6) of the Supplementary material) was performed by the control volume method, whereby the coupling and correction of the velocity and pressure for multiphase flows are realized with the Phase Coupled SIMPLE (PCSIMPLE) algorithm. The discretisation of the convective terms was performed with the second-order upwind scheme.

The calculations were non-stationary, with a time step of 1 ms, which allowed a relatively quick convergence with a maximum of 100 iterations per time step, whereby the convergence criterion between two iterations was set to 10^{-3}. The number of time steps, *i.e.*, the total simulation time, was determined by the time required for the fluid to pass through the entire reactor space. The computational domain consisted of the two zones: layer of particles in the fluidized bed and the free flow above the fluidized bed. The entire numerical grid consisted of more than 13000 nodes.

The inter-phase interaction drag force, Eqs. (12) and (13) of the Supplementary material, were included in the numerical simulation process by the specialized subroutines in the "C" programming language ("user defined functions"), with which the user is able to upgrade individual parts of the core Fluent code.

The proposed Euler–Euler granular model of the fluidized bed was applied on the three fluidization regimes, the basic characteristics of which are presented in Table I. In all cases, the theoretical value of the fluidization number was approximately three. The default constants of 0.8 and 2.65 in the coefficient *B* (Eq. (13) of the Supplementary material) had to be changed for all three regimes. Accordingly, the appropriate constants used for the simulation of regime 1 were 0.282 and 9.077, respectively. The same constants, used for simulation of

regimes 2 and 3 had the values of 3.2 and 0.6625, respectively. The results of numerical simulations of treated fluidization regimes are presented in Fig. 1, which presents the distribution of the solid volume fraction during the development period of the bubbled fluidization.

TABLE I. Numerical simulation cases

Regime	Reactor dimensions m^2	FB height m	Particle size mm	Particle material density, kg m^{-3}	Fluidization gas	Inlet fluidization gas velocity m s^{-1}	T K
1	1×0.15	0.15	0.3	2600	Air	0.25	300
2	0.23×0.4	0.3	0.8	2400	Air	1	300
3	0.23×0.4	0.3	0.8	2400	Air multi-component	1	1200

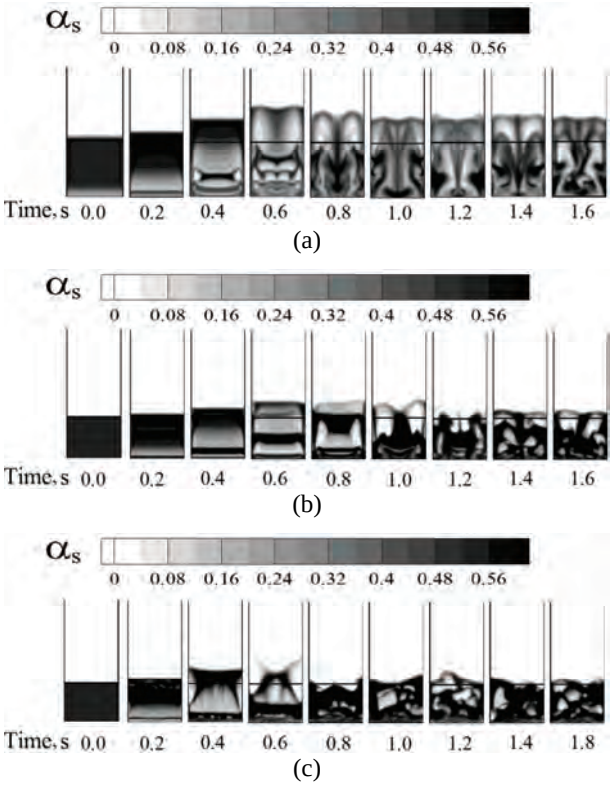


Fig. 1. Development of the solid volume fraction distribution for fluidization conditions in regimes: a) 1, b) 2 and c) 3.

2D simulation of combustion of liquid fuels in a fluidized combustion chamber

A comprehensive model of the complex processes in fluidized combustion chamber may be formed by upgrading of the proposed isothermal and single component Euler–Euler granular model of the fluidized bed, including the energy equation and the transport equations

of chemical species conservation with the source terms due to the conversion of chemical components. The additional equations for energy and conservation of chemical components are presented in Eqs. (1a), (1b) and (2):

Energy equation of the gas phase:

$$\frac{\partial}{\partial t}(\alpha_{\mathrm{g}}\rho_{\mathrm{g}}c_{p,\mathrm{g}}T_{\mathrm{g}})+\nabla(\alpha_{\mathrm{g}}\rho_{\mathrm{g}}\vec{u}_{\mathrm{g}}c_{p,\mathrm{g}}T_{\mathrm{g}})=\nabla\left(\frac{k_t}{c_{p,\mathrm{g}}}\nabla T_{\mathrm{g}}\right)+$$
$$+\nabla\left(\sum_i \alpha_{\mathrm{g}}\rho_{\mathrm{g}}D_{i,m}c_{p,i}T_{\mathrm{g}}\nabla Y_i\right)-h\left(T_{\mathrm{s}}-T_{\mathrm{g}}\right) \quad (1a)$$

Energy equation of the solid phase:

$$\frac{\partial}{\partial t}(\alpha_{\mathrm{s}}\rho_{\mathrm{s}}c_{p,\mathrm{s}}T_{\mathrm{s}})+\nabla(\alpha_{\mathrm{s}}\rho_{\mathrm{s}}\vec{u}_{\mathrm{s}}c_{p,\mathrm{s}}T_{\mathrm{s}})=\nabla\left(\frac{k_{\mathrm{s}}}{c_{p,\mathrm{s}}}\nabla T_{\mathrm{s}}\right)+h\left(T_{\mathrm{s}}-T_{\mathrm{g}}\right) \quad (1b)$$

Conservation equations for chemical components:

$$\frac{\partial}{\partial t}(\alpha_{\mathrm{g}}\rho_{\mathrm{g}}Y_i)+\nabla(\alpha_{\mathrm{g}}\rho_{\mathrm{g}}\vec{u}_{\mathrm{g}}Y_i)=\nabla\left(\alpha_{\mathrm{g}}\rho_{\mathrm{g}}D_{i,m}\nabla Y_i\right)+R_i \quad (2)$$

The energy balance equations for the two phases are connected through the interphase volumetric heat transfer coefficient (h), which was given by Gunn.[11] The granular conductivity coefficient has been determined using Syamlal and Gidaspow formulation.[12] Special attention was paid to the determination of the effective thermal conductivity of the gas and solid phase mixture.

The source term R_{i} in Eq. (2) corresponds to the chemical conversion rate of component i. The source term of the reaction rate for component i was defined using the Arrhenius expression:

$$R_i=k_{\mathrm{o},i}\exp\left(-\frac{E_{\mathrm{a},i}}{RT_{\mathrm{g}}}\right)\left[Y_i\right]^{a_1}\left[Y_j\right]^{a_2} \quad (3)$$

As the presented work is an integral part of extensive research of unconventional fuel combustion and industrial wastes incineration, diesel fuel ($C_{10}H_{22}$) with various water contents was used as the fuel in the processed numerical experiments. Accordingly, the list of chemical reactions that figure in treated test cases of the liquid fuels combustion in the FB, and the corresponding coefficients in the Arrhenius equation, Eq. (3), are given in Table II. The list includes two reactions that are not strictly chemical: the reactions of fuel devolatilization and the water evaporation. The basic coefficients for the kinetics reactions for the numerical experiments (Table II) were taken from the literature.[13,14]

TABLE II. Chemical reactions list

No.	Reaction	$-k_0$	E_{a} / J kmol^{-1}	a_1	a_2
1	$C_{10}H_{22(\mathrm{liquid})} \rightarrow C_{10}H_{22(\mathrm{vapor})}$	11.2	1.7E6	1	–
2	$C_{10}H_{22} + 10.5\ O_2 = 10CO + 11H_2O$	2.857E10	1.25E8	0.25	1.5
3	$CO + 0.5O_2 = CO_2$	1.0E12	1.0E8	1	1
4	$H_2O_{(\mathrm{liquid})} \rightarrow H_2O_{(\mathrm{vapor})}$	22.39	1.7E6	1	–

It should be noted that an actual quasi-three-phase system was treated here, that includes the liquid phase (water and fuel in liquid state at the inlet), the gas phase (air with flue gases: volatiles and steam) and the solid phase (inert bed particles). The numerical system was not really a three-phase one, since equations for only two phases were used: the fluid phase, consisting of the liquid fuel and water with gases (steam, volatiles, flue gasses and air) and the particle phase, which consisted of a granular layer (bed) of particles.

The presented numerical experiments were applied primarily for verification of the assumptions concerning the impact of the water content of a liquid fuel on the FB combustion intensity. In the experiments with a pilot furnace with liquid fuel feeding into the FB, withdrawal of the intense combustion zone was observed towards the areas below the bed surface during the combustion of liquid fuels with a significant water content, compared to those without water or with a low water content.[15,16] The emulsion of fuel (various oils) and water were prepared in special homogenizers (heated vessel with a stirrer and heated lines to the nozzle feeder). The detailed procedure is described elsewhere.[15,16] Based on these experimental results, it was concluded that the water in the fuel increased the global reaction rate of combustion in the fluidized bed furnace. This phenomenon is associated with a sudden transition of water in the fuel to steam while entering the heated FB (≈ 900 °C), which causes an expansion of the input jet and improved mixing of the fuel with oxidizer (air), as well as breakage of the liquid (still not vaporized) fuel into small droplets thereby increasing of the contact surface area of the fuel with the environment. The first attempts at numerical simulation of this phenomenon were reported earlier.[17]

The test case of numerical simulation of the processes in a fluidized combustion chamber was performed on the fluidization reactor of height 2.3 m and width of 0.4 m, as is shown in the schematic view of the reactor (Fig. 2). The modeled granular bed consisted of particles of diameter 0.8 mm and density of 2600 kg m^{-3}, where the height of the bed in the bulk condition was 0.3 m.

Fig. 2. Schematic view of the geometry of the numerically simulated fluidization reactor.

The fuel entered through a vertical nozzle placed axially on the bottom of the reactor. The nozzle for the introduction of the fuel was placed at a height of 0.05 m. Air for

fluidization was introduced annularly as is shown in Fig. 2. The inlet temperature of the air and fuel was ambient (300 K).

As was already mentioned, diesel ($C_{10}H_{22}$) with various contents of water was selected to be the test fuel. The presented numerical experiments were performed with 0.0, 0.1, 0.2, 0.3 and 0.4 mass fractions of water in the fuel. The inlet mass flow rate of air and the pure diesel fuel were 0.11627 and 0.002586 kg s^{-1}, respectively, in all calculations, which corresponded to an excess air ratio of λ = 3. In the cases when the fuel contained from 0.1 to 0.4 mass fraction of water, the inlet total fuel mass flows were 0.00287, 0.00323, 0.00370 and 0.00431 kg s s^{-1}.

THE RESULTS OF THE NUMERICAL SIMULATION OF LIQUID FUEL COMBUSTION IN THE FLUIDIZATION FURNACE

For the numerical experiments of the fluidized bed (FB) combustion, the calculation procedure and the numerical method were the same as for the cases treated in the numerical tests. The numerical grid shown in Fig. 2, consisting of 13130 nodes of which 3430 nodes were used for the granular bed zone, was employed.

The computational procedure consisted of two steps. In the first step of the calculation, the fluidization parameters were formed, according to the calculations presented in the numerical tests (Fig. 1c). The calculations were multicomponent and non-isothermal (the gas phase was heated up to 1200 K). The matrix values of the variables calculated in the first computing step were employed as the initial conditions for the second step of the calculation process. Moreover, in the second step, the boundary conditions were changed introducing the inlet fuel flow and the equations of chemical species with the source terms due to activation of the chemical reactions. Each of the steps comprised two simulation periods of approximately 2.5 s. After this period, it could be considered that the analyzed processes had entered a quasi-stationary state, because further changes were minor. Accordingly, in the second computing step, the processes were simulated from the beginning of the fuel introduction into the heated FB, under the developed fluidization conditions, up to the end of the period of 2.5–3 s.

The gas temperature distributions calculated by numerical simulation of diesel fuel combustion in fluidized bed under the specified conditions without moisture are shown in Fig. 3. for the period starting from 600 ms after fuel introduction into the heated fluidized bed. The temperature field in the zone of intense reaction in fluidized combustion chamber harmonically changed in time; hence, quasi-stationary processes could be assumed in this type of furnace.

The calculated temperature profiles along the fluidized reactor height for all five mass fractions of water from 0.0 to 0.4) are shown in Fig. 4. The temperature profiles were averaged over time for the period ranging from 2 to 3 s from the beginning of fuel feeding into the heated FB. The changes in the radial temperature profiles *versus* time were not symmetric, but stochastic; hence, the vertical

profiles in Fig. 4 represent the mass-weighted average temperature for the cross sections along the reactor height. The ordinate in Fig. 4 is the normalized temperature, defined as the ratio of the temperature and the maximal temperature (T_{max}) in corresponding conditions, while the abscissa represents the dimensionless height of the furnace, which is defined as ratio between the height of the reactor and the height of the fixed bed (H_{fb}). The temperature peaks and more rapid attaining of the maximum temperature in the given diagrams were shifted more towards the middle of the bed in the regimes with the fuel that contained water. It could be concluded that the water content in the fuel affects, to some degree, a withdrawal of the intense combustion zone in fluidized bed furnace towards the lower zones. This also pointed to the fact that with fuel that contained water, the location of combustion was lower, *i.e.*, that more efficient combustion was achieved. The registered effect of more a more rapid reaching of higher temperatures during combustion was more pronounced when the mass fractions of water in the fuel had values of 0.1 and 0.2, whereas for the higher moisture content in the fuel, this effect was less pronounced. This indicated the presence of a dual effect of the fuel moisture content, *i.e.*, on the one hand, the sudden expansion of water vapor improved the mixing of fuel and oxidizer, while the local heat balance was reduced on the other.

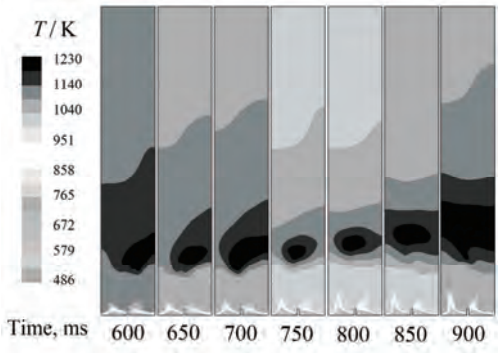


Fig. 3. The change in time of the temperature field in the fluidized combustion chamber.

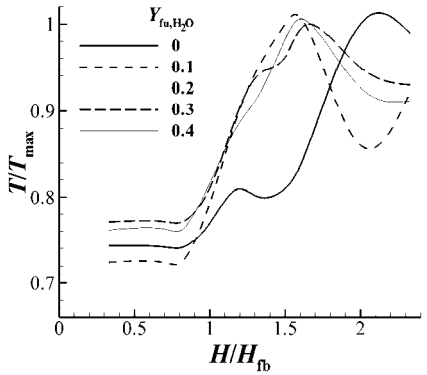


Fig. 4. Calculated temperature profiles along the fluidized reactor height for different mass fraction of water in the fuel.

The effect of the changing of combustion reaction rate due to the presence of water in the fuel could also be analyzed by using the diagram in Fig. 5, which shows the averaged values (in the combustion zone area) of the summary reaction kinetic rate depending on the water fraction in the fuel. These calculation results correspond to the conditions after a period of 2 s from the beginning of fuel feeding. Figure 5 shows that the mean reaction kinetic rate of diesel-fuel oxidation had higher values in the cases when the fuel contained water, whereby the dependency of the reaction rate on the moisture content in the fuel increased suddenly when the fraction of moisture was 0.1. The noted effect of the influence of fuel water content on the averaged fuel oxidation kinetic rate could be explained by the expansion of steam in the simulated FB, which in a number of nodes contributed to favorable (stoichiometric and over-stoichiometric) mixtures of fuel and oxidizer being obtained. On the other hand, higher values of the moisture content in the fuel led to a decrease in the local temperature of the FB, and then the reaction rates (according to the Arrhenius expression) were slightly lower.

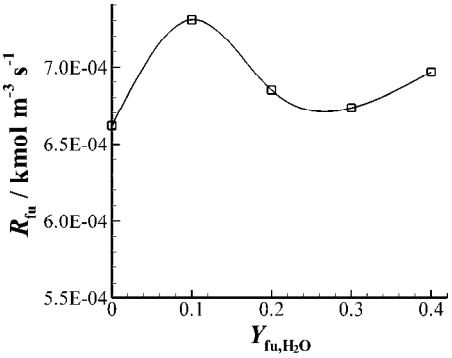


Fig. 5. Averaged kinetic rate of the summary reaction of diesel-fuel oxidation depending on the water content in the fuel.

CONCLUSIONS

A comprehensive 2D numerical model of the bubbled fluidized bed with combustion of liquid fuels has been proposed. The developed numerical model is based on Eulerian–Eulerian granular flow modeling with the kinetic theory of granular flow. It includes the following basic governing equations: the continuity equations of the solid and gas phase, the momentum conservation equations of the gas and solid phase, the energy equations of the gas and the solid phase and conservation equations for the chemical components.

For a more realistic simulation of events in the fluidized system of gas–solid particles that include combustion processes, it is necessary to adjust the constants of coefficient *B* (Eqs. (12) and (13) of the Supplementary material) depending on fluidization conditions and fluid and solid parameters. For this purpose, the proposed Euler–Euler granular model of the fluidized bed was applied on three different fluidization regimes.

The presented numerical experiments were primarily applied to an investigation of the impact of the water content in a liquid fuel on the intensity of FB combustion. The presented numerical experiments were performed with water mass fractions of 0.0, 0.1, 0.2, 0.3 and 0.4 in diesel fuel. Within the analysis of the results of the numerical simulation of FB combustion of the fuel with different water contents, the temperature profiles and the mean kinetic ratio of diesel fuel oxidation were considered.

The calculated vertical temperature profiles along the fluidized reactor height showed that the water content in the fuel affects, to some degree, the withdrawal of the intense combustion zone in the fluidized furnace towards the lower zones. Moreover, the averaged kinetic rate of diesel-fuel oxidation was higher when the fuel contained water. These two effects could be explained by an increase in the mixing intensity of the reactants and oxidizer due to steam expansion in the FB. Furthermore, the effect of the fuel moisture content had a dual nature: the sudden expansion of steam enhanced the mixing of fuel and oxidant, but also reduced the local thermal equilibrium; thus, the observed effect was less pronounced when the values of the water mass fraction in the fuel were higher than 0.1. Based on the analyzed results of the numerical experiments, it is could be concluded that liquid fuels burn efficiently in a fluidized bed, even if they contain a relatively large portion of water. When it comes to diesel fuel, with water mass fraction of 10 % had the most favorable kinetic characteristics, *i.e.*, fuel with 10 % of water by mass required less volume for the process of combustion in the FB.

SUPPLEMENTARY MATERIAL

The basic and constitutive equations of the two-fluid granular model of a fluidized bed are available electronically from http://www.shd.org.rs/JSCS/, or from the corresponding author on request.

NOMENCLATURE

a_1, a_2	rate exponent for the first and second reactant
C_D	drag coefficient
$c_{p,g}$, $c_{p,s}$	gas and solid specific heats
$D_{i,m}$	mass diffusion coefficient for species i
d_p, d_s	particle mean diameter
E_a	activation energy
e_s	restitution coefficient
g	gravity acceleration
g_{0s}	radial distribution function
h	heat transfer coefficient with specific surface
H	height
H_{fb}	height of the fixed bed
I	unity matrix
I_{2D}	second invariant of deviator of the strain rate tensor
K_{gs}	gas/solid momentum exchange

k_t thermal conductivity
k_o pre-exponential coefficient
$k_{\Theta s}$ diffusion coefficient for granular energy
p pressure
R universal gas constant
R_{fu} summary reaction rate
$\overline{\overline{S}}_k$ strain rate tensor
T absolute temperature
T_{max} maximal temperature in the regime
$\vec{u}$ instantaneous velocity vector
Y_i species mass fraction
Y_{fu,H_2O} water mass fraction in the fuel

Greek symbols

α phase void fraction
ρ density
λ bulk viscosity
$\overline{\overline{\tau}}$ phase stress–strain tensor
Θ_s granular temperature
$\mu_{s,kin}$ kinetic viscosity
$\mu_{s,coll}$ collisional viscosity
$\mu_{s,fr}$ frictional viscosity
ϕ angle of internal friction for the particle
ϕ_{gs} transfer rate of kinetic energy
$\gamma_{\Theta s}$ collisional dissipation energy

Indexes

b fluidized bed
g gas
s solid
i educts
j products

Acknowledgment. The authors wish to thank the Ministry of Education, Science and Technological Development of the Republic of Serbia for financing the project “Improvement of the industrial fluidized bed facility, within the scope of technology for energy efficient and environmentally feasible combustion of various waste materials in fluidized bed”, Project No. TR33042.

ИЗВОД

ДВОФЛУИДНИ ОЈЛЕР–ОЈЛЕРОВСКИ МОДЕЛ ГРАНУЛАРНОГ ТОКА САГОРЕВАЊА ТЕЧНОГ ГОРИВА У РЕАКТОРУ С ФЛУИДИЗОВАНИМ СЛОЈЕМ

СТЕВАН НЕМОДА[1], МИЛИЦА МЛАДЕНОВИЋ[1], МИЛИЈАНА ПАПРИКА[1], ДРАГОЉУБ ДАКИЋ[2], АЛЕКСАНДАР ЕРИЋ[1] и МИРКО КОМАТИНА[3]

[1]Институт за нуклеарне науке „Винча“, Универзитет у Београду, Београд, [2]Иновациони центар Машинског факултета, Универзитет у Београду, Београд и [3]Машински факултет, Универзитет у Београду, Београд

У раду је предложен модел нумеричке симулације сагоревања течних горива у флуидизованом слоју (FB), који се заснива на дво-флуидном Ојлер–Ојлеровском при-

ступу моделирања флуидизованог слоја уз одређивање поља брзина гаса и честица у двофазним грануларним токовима заснованог на аналогији са кинетичком теоријом гасова (RTGF). Свеобухватан модел комплексних процеса у флуидизационом ложишту подразумева, поред одређивања поља брзине гасне и честичне фазе, инкорпорирање енергетских једначина гасне и честичне фазе, као и транспортних једначина хемијских компоненти са изворним члановима који потичу од конверзије компонената. Нумерички експерименти показују да избор коефицијената у изразима за силе трења приликом интеракције фаза има изузетан значај и да се он мора спровести посебно за сваки значајно различит режим флуидизације. Урађене су серије нумеричких експеримената са симулацијом процеса сагоревања у FB, са и без значајног садржаја воде у гориву. Прорачуни су нестационарни, а моделирани временски период одговара времену за које гас прође целу висину реактора. Изложени нумерички експерименти су првенствено намењени за испитивање утицаја садржаја влаге у гориву на кинетику реакција у флуидизованом ложишту.

(Примљено 30. јануара, ревидирано и прихваћено 27. марта 2014)

REFERENCES

1. D. Gera, M. Gautam, Y. Tsuji, T. Kawaguchi, T. Tanaka, *Powder Technol.* **98** (1998) 3847
2. C. Ibsen, E. Helland, B. Hjertager, T. Solberg, L. Tadrist, R. Occelli, *Powder Technol.* **149** (2004) 29
3. G. A. Bokkers, M. van Sint Annaland, J. A. M. Kuipers, *Powder Technol.* **140** (2004) 176
4. A. Di Renzo, F. P. Di Maio, *Chem. Eng. Sci.* **62** (2007) 116
5. H. Enwald, E. Peiran, A. E. Almstedt, B. Leckner, *Chem. Eng. Sci*. **54** (1999) 311
6. L. Cammarata, P. Lettieri, G. D. M. Micale, D. Colman, *Int. J. Chem. Reactor Eng.* **1** (2003) 48
7. Y. Behjat, S. Shahhosseini, S. H. Hashemabadi, *Int. Commun. Heat Mass Transfer* **35** (2008) 357
8. D. Gidaspow, *Multiphase Flow and Fluidization: Continuum and Kinetic Theory Descriptions*, Academic Press, San Diego, CA, 1994, p. 30–320
9. M. Syamlal, T. J. O'Brien, *Int. J. Multiphase Flow* **14** (1988) 473
10. M. Syamlal, W. Rogers, T. J. O'Brien, *MFIX Documentation Theory Guide*, U.S. Department of Energy Office of Fossil Energy Morgantown Energy Technology Center, Morgantown, WV, 1993, p. 7
11. D. J. Gunn, *Int. J. Heat Mass Transfer* **21** (1978) 467
12. M. Syamlal, D. Gidaspow, *AIChE J.* **31** (1985) 127
13. O. G. Penyazkov, K. L. Sevrouk, V. Tangirala, N. Joshi, in *Proceedings of the Fourth European Combustion Meeting*, Vienna, Austria, 2009
14. J. F. Brennan, J. S. Shapiro, E. C. Watton, *J. Chem. Educ*. **51** (1974) 276
15. M. R. Mladenović, D. V. Dakić, S. Đ. Nemoda, R. V. Mladenović, A. M. Erić, M. J. Paprika, M. S. Komatina, in *Proceedings of 15th Symposium on Thermal Science and Engineering of Serbia*, Sokobanja, Serbia, 2011, CD-ROM, p. 490
16. M. R. Mladenović, M. J. Paprika, D. V. Dakić, S. Đ. Nemoda, A. M. Erić, M. S. Komatina, *Termotehnika* **38** (2012) 11 (in Serbian)
17. S. Nemoda, M. Mladenović, D. Dakić, M. Komatina, A. Erić, D. Đurović, In *Proceedings of International Conference Power Plants*, Zlatibor, Serbia, 2012, p. 1150.

Journal of the Serbian Chemical Society
JSCS@tmf.bg.ac.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) S105–S108 (2015) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Euler–Euler granular flow model of the combustion of liquid fuels in a fluidized reactor

STEVAN NEMODA[1]*, MILICA MLADENOVIĆ[1], MILIJANA PAPRIKA[1], DRAGOLJUB DAKIĆ[2], ALEKSANDAR ERIĆ[1] and MIRKO KOMATINA[3]

[1]*"Vinča" Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia,* [2]*Innovation Center, Faculty of Mechanical Engineering, University of Belgrade, Belgrade, Serbia and* [3]*Faculty of Mechanical Engineering, University of Belgrade, Belgrade, Serbia*

J. Serb. Chem. Soc. 80 (3) (2015) 377–389

TWO-FLUID GRANULAR MODEL OF A FLUIDIZED BED**

The basic and constitutive equations of the two-fluid granular model of a fluidized bed are given in Supplementary material to this paper. could be described by the following set of expressions:[1]

Continuity equation of the gas phase:

$$\frac{\partial}{\partial t}(\alpha_g \rho_g) + \nabla(\alpha_g \rho_g \vec{u}_g) = 0 \tag{1}$$

Continuity equation of the solid phase:

$$\frac{\partial}{\partial t}(\alpha_s \rho_s) + \nabla(\alpha_s \rho_s \vec{u}_s) = 0 \tag{2}$$

Momentum conservation equation of the gas phase:

$$\frac{\partial}{\partial t}(\alpha_g \rho_g \vec{u}_g) + \nabla(\alpha_g \rho_g \vec{u}_g \vec{u}_g) = -\alpha_g \nabla p + \nabla \cdot \bar{\bar{\tau}}_g + \alpha_g \rho_g \vec{g} + K_{gs}\left(\vec{u}_g - \vec{u}_s\right) \tag{3}$$

Momentum conservation equation of the solid phase:

$$\begin{aligned} &\frac{\partial}{\partial t}(\alpha_s \rho_s \vec{u}_s) + \nabla(\alpha_s \rho_s \vec{u}_s \vec{u}_s) = \\ &= -\alpha_s \nabla p + \nabla \bar{\bar{\tau}}_s + \alpha_s \rho_s \vec{g} + K_{gs}\left(\vec{u}_g - \vec{u}_s\right) \end{aligned} \tag{4}$$

* Corresponding author. E-mail: snemoda@vinca.rs
** Nomenclature is given in the basic paper.

where: $\vec{u}_g$ and $\vec{u}_s$ are the instantaneous velocity vectors for the gas and solid phase, respectively; α_g and α_s are the void fraction of the gas and solid phase, respectively.

The stress tensor of the gas and the granular phases are given by Eqs. (5) and (6), respectively:

$$\bar{\bar{\tau}}_g = 2\mu_g \bar{S}_g + (\lambda_g - \frac{2}{3}\mu_g)\nabla\vec{u}_g \bar{\bar{I}} \quad (5)$$

$$\bar{\bar{\tau}}_s = -p_s \bar{\bar{I}} + 2\alpha_s\mu_s \bar{S}_s + \alpha_s(\lambda_s - \frac{2}{3}\mu_s)\nabla\vec{u}_s \bar{\bar{I}} \quad (6)$$

The strain rate tensor is defined as:

$$\bar{S}_k = \frac{1}{2}\left(\nabla\vec{u}_k + (\nabla\vec{u}_k)^T\right), \text{ k = g,s} \quad (7)$$

The pressure of the granular phase is:[2]

$$p_s = 2\rho_s\Theta_s(1+e_s)\alpha_s^2 g_{0s} \quad (8)$$

The radial distribution function, g_{0s}, for the Syamlal model is equal to:

$$g_0(\alpha_s) = \frac{1}{1-\alpha_s} + \frac{3\alpha_s}{2(1-\alpha_s)^2} \quad (9)$$

where e_s is the restitution coefficient and Θ_s is the granular temperature.

The viscosity of the granular phase consists of the solids shear viscosity μ_s and the bulk viscosity λ_s. The solids bulk viscosity λ_s is a measure of resistance of solid particles to expansion/compression and, according to the Lun *et al.* model,[1] it is defined as:

$$\lambda_s = \frac{4}{3}\alpha_s\rho_s d_s g_{0s}(1+e_s)\left(\frac{\Theta_s}{\pi}\right)^{1/2} \quad (10)$$

The shear viscosity is the result of translator motion (kinetic viscosity, $\mu_{s,kin}$), mutual particle collisions (collision viscosity, $\mu_{s,coll}$) and frictional viscosity ($\mu_{s,fr}$): $\mu_{s,kin} = \mu_{s,kin} + \mu_{s,coll} + \mu_{s,fr}$.

According to the Syamlal model,[2] the kinetic viscosity is:

$$\mu_{s,kin} = \frac{\alpha_s d_s \rho_s (\Theta_s\pi)^{1/2}}{12(2-\eta)}\left[1 + \frac{8}{5}\eta(3\eta-2)\alpha_s g_{0s}\right] \quad (11)$$

$$\eta = \frac{(1+e_s)}{2}$$

and for the collision viscosity, the following expression is applied:

$$\mu_{\mathrm{s,coll}} = \frac{8}{5}\alpha_{\mathrm{s}}\rho_{\mathrm{s}}d_{\mathrm{s}}g_{0\mathrm{s}}\eta\left(\frac{\Theta_{\mathrm{s}}}{\pi}\right)^{1/2}$$

According to the Schaeffer model,[3] the frictional viscosity is defined as:

$$\mu_{\mathrm{s,fr}} = \frac{p_{\mathrm{s}}\sin\phi}{2\sqrt{I_{2\mathrm{D}}}}$$

where p_{s} is the granular phase (solids) pressure, ϕ is the angle of internal friction for the particle and $I_{2\mathrm{D}}$ is the second invariant of the deviator of the strain rate tensor.

The last term in Eqs. (3) and (4) is a consequence of the inter-phase interaction drag force, where the coefficient between the fluid and solid (granular) phase, according to the Syamlal–O'Brien model,[4] is:

$$K_{\mathrm{gs}} = \frac{3\alpha_{\mathrm{g}}\alpha_{\mathrm{s}}\rho_{\mathrm{g}}}{4u_{\mathrm{r,s}}^2 d_{\mathrm{s}}}C_{\mathrm{D}}\left|\vec{u}_{\mathrm{s}} - \vec{u}_{\mathrm{g}}\right|,\ C_{\mathrm{D}} = \left(0.63 + \frac{4.8}{\sqrt{\mathrm{Re}_{\mathrm{s}}/u_{\mathrm{r,s}}}}\right)^2$$

$$\mathrm{Re}_{\mathrm{s}} = \frac{\rho_{\mathrm{g}}d_{\mathrm{s}}\left|\vec{u}_{\mathrm{s}} - \vec{u}_{\mathrm{g}}\right|}{\mu_{\mathrm{g}}} \tag{12}$$

The terminal velocity coefficient for the solid phase, $u_{\mathrm{r,s}}$, was determined as:

$$u_{\mathrm{r,s}} = 0.5\left(A - 0.06\mathrm{Re}_{\mathrm{s}} + \sqrt{(0.06\mathrm{Re}_{\mathrm{s}})^2 + 0.12\mathrm{Re}_{\mathrm{s}}(2B - A) + A^2}\right)$$

$$A = \alpha_{\mathrm{g}}^{4.14},\ B = \begin{cases} = 0.8\alpha_{\mathrm{g}}^{1.28} & \alpha_{\mathrm{g}} \le 0.85 \\ = \alpha_{\mathrm{g}}^{2.65} & \alpha_{\mathrm{g}} > 0.85 \end{cases} \tag{13}$$

The granular temperature, starting from the equations of conservation of fluctuating granular energy, is:

$$\frac{3}{2}\left[\frac{\partial}{\partial t}(\rho_{\mathrm{s}}\alpha_{\mathrm{s}}\Theta_{\mathrm{s}}) + \nabla(\rho_{\mathrm{s}}\alpha_{\mathrm{s}}\vec{u}_{\mathrm{s}}\Theta_{\mathrm{s}})\right] = \left(-\rho_{\mathrm{s}}\bar{\bar{I}} + \bar{\bar{\tau}}\right)_{\mathrm{s}} : \nabla\vec{u}_{\mathrm{s}} + \\ + \nabla(k_{\Theta\mathrm{s}}\nabla\Theta_{\mathrm{s}}) - \gamma_{\Theta\mathrm{s}} + \phi_{\mathrm{gs}} \tag{14}$$

The diffusion coefficient or conductivity of the granular temperature, according to Syamlal,[2] is:

$$k_{\Theta_{\mathrm{s}}} = \frac{15\alpha_{\mathrm{s}}\rho_{\mathrm{s}}d_{\mathrm{s}}\sqrt{\Theta_{\mathrm{s}}\pi}}{4(41 - 33\eta)}\left[1 + \frac{12}{5}\alpha_{\mathrm{s}}g_{0\mathrm{s}}\eta^2(4\eta - 3)\right] \tag{15}$$

The granular energy dissipation due to the inelastic collisions was defined by Lun *et al.*,[1] as follows:

$$\gamma_{\Theta_s} = \frac{12(1-e_s)g_{0s}}{d_s\sqrt{\pi}}\rho_s\alpha_s\Theta_s^{3/2} \quad (16)$$

The exchange of kinetic energy between the phases was determined as:

$$\phi_{gs} = -3K_{gs}\Theta_s \quad (17)$$

REFERENCES

1. C. K. K. Lun, S. B. Savage, D. J. Jeffrey, N. Chepurniy, *J. Fluid Mech.* **140** (1984) 223
2. M. Syamlal, W. Rogers, T. J. O'Brien, *MFIX Documentation Theory Guide*, U.S. Department of Energy Office of Fossil Energy Morgantown Energy Technology Center, Morgantown, WV, 1993, p. 7
3. O. G. Penyazkov, K. L. Sevrouk, V. Tangirala, N. Joshi, in *Proceedings of the Fourth European Combustion Meeting*, Vienna, Austria, 2009
4. M. Syamlal, T. J. O'Brien, *Int. J. Multiphase Flow* **14** (1988) 473.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 391–405 (2015)
JSCS–4724
UDC 622–034.3+54.05+66.061:
621.928.17+622'17(497.11)
Original scientific paper

Bioleaching of copper from samples of old flotation tailings (Copper Mine Bor, Serbia)

SRĐAN STANKOVIĆ[1]*, IVANA MORIĆ[2], ALEKSANDAR PAVIĆ[2], SANDRA VOJNOVIĆ[2], BRANKA VASILJEVIĆ[2] and VLADICA CVETKOVIĆ[1]

[1]*Faculty of Mining and Geology, University of Belgrade, Đušina 7, Belgrade, Serbia and* [2]*Institute for Molecular Genetics and Genetic Engineering, University of Belgrade, Vojvode Stepe 444a, Belgrade, Serbia*

(Received 11 April, revised 4 July, accepted 23 September 2014)

Abstract: Bioleaching of samples taken from depths of 10, 15, and 20 m from old flotation tailings of the Copper Mine Bor was conducted in shaken flasks using the extremely acidic water of Lake Robule as lixiviant. The yields of copper after five weeks of the bioleaching experiments were 68.34±1.21 % for the 15-m sample, 72.57±0.57 % for the 20-m sample and 97.78±5.50 % for the 10-m sample. The obtained results were compared to the results of acid leaching of the same samples and it was concluded that bioleaching was generally more efficient for the treatment of samples taken from depths of 10 and 20 m. The content of pyrite in the 20-m sample, which contained the highest amount of this mineral, was reduced after bioleaching. The benefits of this approach are recovery of substantial amounts of copper, reducing the environmental impact of flotation tailings and the application of abundant and free water from the acidic Robule Lake as lixiviant. The obtained results showed that bioleaching could be more efficient than acid leaching for copper extraction from flotation tailings with higher sulfide contents.

Keywords: bioleaching; flotation tailings; Lake Robule; Bor.

INTRODUCTION

The Bor Mining and Smelting Company, one of the largest industrial facilities in the Republic of Serbia, is the only producer of copper, gold and silver in the country. The Copper Mine Bor has been operating since 1903 and since then, approximately 650 million tons of waste, including both overburden and flotation tailings, were generated.

In the early years of exploitation, the ore from the Copper Mine Bor was generally richer in copper.[1] The abundance of copper in combination with the

* Corresponding author. E-mail: srdjan.stankovic@gmail.com
doi: 10.2298/JSC140411097S

relatively inefficient technology used for flotation resulted in the production of flotation tailings containing significant amounts of copper. It is estimated that the old flotation tailings of the Copper Mine Bor, used from 1933 to 1987, contains about 27×10^6 tons of tailings with an average copper content in the range 0.2–0.3 wt. %, a gold content of 0.3–0.6 g t^{-1} and a silver content of 2.5–3 g t^{-1}.[2–6] Hence, the best estimate of the total copper content in the old flotation tailings of the Copper Mine Bor ranges between 54000 and 81000 t.

Previous detailed analyses of the mineral and chemical composition of samples of the old flotation tailings[5,6] showed that the copper content in the dump could reach 0.5 %. The analysis also revealed that the old flotation tailings consists mainly of quartz, silicates, carbonates (77 %), pyrite (21.57 %) and other metal sulfides (1.43 %). The most abundant copper sulfides are covellite (0.21 %), chalcopyrite (0.16 %), enargite (0.14 %) and chalcocite (0.04 %). The flotation tailings is finely ground, with 55.3 % of the particles being smaller than 0.074 mm.[6] The average liberation of pyrite is around 91 % that, in the presence of water and atmospheric oxygen, represents an important reactive potential for the mobilization of heavy metal ions and the generation of acid mine drainage for many years to come.[5]

Currently, the Bor Mining and Smelting Company is processing low-grade ores, since the average content of copper in the ores from the open pit mines is around 0.3 %.[7] Therefore, the old flotation tailings in Bor are, potentially, a valuable source of copper for the Company.

As copper sulfides are acid soluble, copper from flotation tailings can be leached by sulfuric acid solutions.[4–6] Another approach for the recovery of copper from tailings would include the use of acidophilic iron oxidizing microorganisms. This process is referred to as “bioleaching”.[8] It is generally known that many acidophilic microorganisms provide metabolic energy by oxidizing ferrous iron, generating ferric iron:

$$2FeSO_4(aq) + H_2SO_4(aq) + 0.5O_2(g) \rightarrow Fe_2(SO_4)_3(aq) + H_2O(l) \quad (1)$$

Ferric iron is a powerful oxidizing agent that is able to oxidize metal sulfides, such as pyrite and copper sulfides.[4]

$$2FeS_2(s) + 7Fe_2(SO_4)_3(aq) + 8H_2O(l) \rightarrow 15FeSO_4(aq) + 8H_2SO_4(aq) \quad (2)$$

Acidophilic microorganisms very efficiently oxidize the ferrous iron released from pyrite and regenerate ferric iron, accelerating the process of metal sulfide oxidation.[9] Oxidation of the most important copper sulfides (covellite, enargite, chalcocite and chalcopyrite, respectively) is presented in Eqs. (3)–(6):[4]

$$CuS(s) + Fe_2(SO_4)_3(aq) \rightarrow CuSO_4(aq) + S(s) + 2FeSO_4(aq) \quad (3)$$

$$2Cu_3AsS_4(s) + 11Fe_2(SO_4)_3(aq) + 8H_2O(l) \rightarrow$$
$$\rightarrow 6CuSO_4(aq) + 2H_2AsO_4(aq) + 5H_2SO_4(aq) + 8S(s) + 22FeSO_4(aq) \quad (4)$$

$$Cu_2S\ (s) + Fe_2(SO_4)_3\ (aq) \rightarrow 2CuSO_4(aq) + S\ (s) + 4FeSO_4(aq) \quad (5)$$

$$CuFeS_2(s) + Fe_2(SO_4)_3(aq) \rightarrow CuSO_4(aq) + 2S(s) + 5FeSO_4(aq) \quad (6)$$

Some species of acidophilic microorganisms are able to oxidize sulfur, producing sulfuric acid and lowering the pH of the environment. Oxidation of copper sulfides produces soluble copper that can be recovered from the solution by application of solvent extraction/electrowinning technology.[8]

Open pit acidic lakes are widespread in mining areas and they present the potential source of acidophilic microorganisms that could be used for bioleaching of copper. The Robule Lake is an extremely acidic water body that is located near the town of Bor, at the foot of the overburden of the open pit. The input of acid mine drainage from the overburden had caused the lake water to become highly acidic and deep red in color due to the high concentrations of ferric iron. The redox potential of the lake water measured in July 2012 was +850 mV.[11] Korać and Kamberović[9] reported that the concentration of suspended solids in a water sample collected in February 2005 was 199.6 mg L^{-1}. Previous studies confirmed the presence of acidophilic bacteria in the Lake water.[10,11] The latter study[11] showed that the most abundant bacteria identified in a sample of the Lake water were the heterotrophic bacteria *Acidiphilium cryptum* and the autotrophic iron-oxidizers *Leptospirillum ferrooxidans*, with a relatively small number of autotrophic iron- and sulfur-oxidizers *Acidithiobacillus ferrooxidans*.

Besides Lake Robule, there are also other sources of acidic wastewaters in the Bor Mine area, which could be used for leaching copper from waste material. For instance, the wastewater streams originating from the open pit Bor, copper refinery and smelting plants are also characterized by low pH values and high concentrations of iron.[9]

The aim of this work was to investigate the possible applications of the acidic water of Lake Robule in the process of bioleaching of copper from the flotation tailings.

EXPERIMENTAL

Samples

The flotation tailings samples were collected from 20 m-deep drill holes drilled in 2007. Chemical and mineral compositions of the selected samples were previously determined.[5,6] The samples used in this work were taken from depths of 10, 15 and 20 m, and contained 5100 mg kg^{-1}, 3300 mg kg^{-1} and 4300 mg kg^{-1} of copper, respectively.

The sample of the water from the Lake Robule was collected on June 12th, 2013. The water temperature and pH of the sample were measured on site using a Hanna Instruments HI 98312 device.

Copper analysis

The total copper concentration in the water sample was measured using a modified method described by Anwar *et al.*[12] Copper (II) was reduced to Cu(I) by the addition of 200 μL of 10 % hydroxylamine solution to 100 μL of water sample. The solution was mixed well

and incubated for 5 min at room temperature. Tartrate buffer (1 mL of 0.5 M HCl added to 100 mL of 0.5 M sodium tartrate and pH adjusted to 5.5) was added to the sample and the solution was mixed. Then, 500 μL of phosphate buffer (87.7 mL of 0.2 M $NaH_2PO_4 \cdot 2H_2O$ with 13.3 mL of 0.2 M $Na_2HPO_4 \cdot 12H_2O$) and 100 μL of 0.1 % bicinchoninic acid (Sigma Chemicals, USA) diluted in tartrate buffer were added and the sample well mixed. Finally, 0.8 mL of distilled water was added, mixed and after 10 min at room temperature, the absorbance at 562 nm was measured.

Iron analysis

Amounts of soluble ferrous and ferric iron were determined spectrophotometrically by the phenanthroline assay[13] and thiocyanate assay,[14] respectively. Briefly, the supernatants were collected after centrifuging 1 mL of bioleaching cultures for 10 min at 13000 rpm. For the determination of Fe(II) ions, the samples were prepared by mixing 50 μL of supernatant with 20 μL 1 M Na-acetate and 150 μL 5 mM 1,10-phenantroline-monohydrate and, finally, filled to 1 mL with deionized water. After 15 min incubation at room temperature, the absorbance was read at 510 nm. The Fe(II) concentration was determined based on the standard curve prepared using $FeSO_4 \cdot 7H_2O$. For the Fe(III) determination, 500 μL of appropriate dilution of supernatant was mixed with 500 μL of 1 M NaSCN and incubated for 15 min at room temperature. The concentration of ferric ions was determined by comparing the absorbance readings at 490 nm to standard curve prepared with $FeCl_3 \cdot 6H_2O$.

Bioleaching experiments

Water from the Lake Robule (100 mL was poured into each of three 250-mL conical flasks and 5 g of different samples of flotation tailings were added to each flask. The flasks were labeled S-10, S-15 and S-20, referring to the depth from which the samples were collected. Abiotic controls were prepared in a similar manner, except the samples of flotation tailings were incubated with 100 mL of sulfuric acid solution. Sulfuric acid was dissolved in distilled water, final pH of the solution was 2.5 in order to match the pH value of the Lake water. The control samples were labeled as CS-10, CS-15 and CS-20.

An additional control contained solely 100 mL of water from the Lake and was labeled LW. All samples and controls were incubated for five weeks at 30 °C under constant rotation of 150 rpm. One mL of sample from each flask was taken once a week for determination of the concentrations of copper, ferric and ferrous ions and pH, as described. Distilled water was added in order to compensate the evaporated water. All measurements were performed in triplicate.

X-Ray powder diffraction analysis

The samples were analyzed in the Laboratory for Crystallography at the Faculty of Mining and Geology, University of Belgrade, on a Philips PW 1710 X-ray powder diffractometer. Experimental conditions were as follows: U = 40 kV, I = 30 mA, CuK_α = 1.54178 Å, 2θ range of examination 4–90°, step 0.02° and time constant 1 s (per step). The source of X-rays was a copper anticathode with a graphite monochromator. The minerals were identified by comparing the obtained relative intensities I/I_{max} and d values with data from the JCPDS database. Semi-quantitative analysis of the diffraction data was preformed using computer software Powder Cell (Federal Institute for Materials Research and Testing, BAM, Berlin, Germany).

Statistical analysis

Values of copper concentrations released over time from the same flotation tailings sample by acid leaching and bioleaching were compared using the *t*-test. The effects of the solutions used for bioleaching on copper recovery were determined by one-factor ANOVA at a threshold level of $P = 0.05$ and by the Bonferonni test. All statistics analyses were performed using SPSS 20 (SPSS Inc., Chicago, IL) software.

RESULTS

Physical and chemical properties of the water from Lake Robule

The color of the lake water is deep red due to the presence of a high concentration of ferric iron in the water. The temperature and pH of the water measured on site were 21 °C and 2.53, respectively. The concentrations of copper and iron were determined in the laboratory the day after sampling. The concentration of iron in the water sample was 250±1.54 mg L^{-1} and that of copper was 47.24± ±0.610 mg L^{-1}.

Bioleaching experiment

Changes in the concentrations of copper, and ferrous and ferric irons during the bioleaching experiment are presented in Fig. 1. Since the lake water itself contained copper, the concentration of extracted copper was calculated by subtracting the copper concentration measured in flask with the lake water only (flask LW) from the total copper concentration measured in the leaching solution (flasks S-10, S-15 and S-20). The yields of copper after five weeks of bioleaching are presented in Table I.

TABLE I. Copper contents and the yield of copper after five weeks of bioleaching

Parameter	Sample			
	S-10	S-15	S-20	Mean
Cu content, mg kg^{-1}	5100	3300	4300	4233
Cu bioleached, mg L^{-1}	249±14	113±2.1	156±1.2	173±14
Yield of Cu, %	98±5.5	68±1.2	73±0.6	82±8.2

According to one-way ANOVA (leaching solution) and the Bonferonni test ($P < 0.001$), the extraction of copper from flotation tailing samples collected at depths of 10 and 20 m was significantly more efficient during bioleaching in comparison to acid leaching (Fig. 1). The differences in concentrations of copper recovered from the 15-m samples by bioleaching and acid leaching were not statistically significant ($P \geq 0.05$, Fig. 1). Notably, there was no significant increase of the released copper concentration during acid leaching over the period from the second to fifth week ($P \geq 0.05$, *t*-test). On the other hand, statistical analysis (Bonferroni test) showed that the concentrations of copper obtained after five weeks of bioleaching of the 10-m and 20-m samples were significantly higher than the concentrations measured in the previous four weeks.

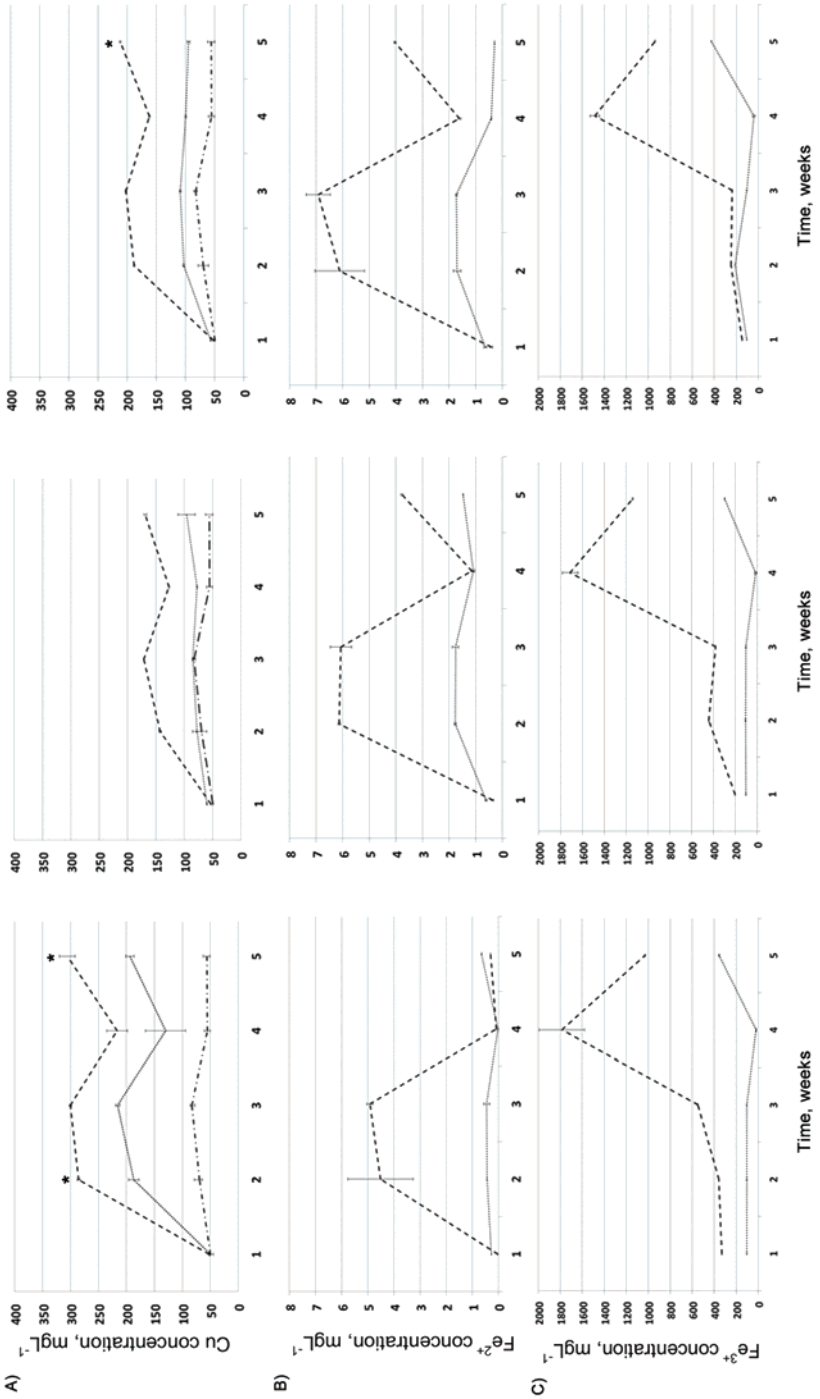


Fig. 1. Changes in concentrations of copper, and ferrous and ferric iron measured during bioleaching of the samples taken from depths of 10, 15 and 20 m. The values on the X-axis represent the duration of the treatment in weeks. The concentrations of copper that were significantly higher in the flasks with the lake water than in the flasks with sulfuric acid solutions of pH 2.5 are marked with an asterisk (*).

The changes in pH values during bioleaching with lake water, and leaching with sulfuric acid solution pH 2.5 are presented in Table II.

TABLE II. Changes in pH during the bioleaching experiment

Sample	Week				
	1	2	3	4	5
S-10	2.54	2.58	2.43	2.24	2.22
S-15	2.53	2.50	2.33	2.26	2.20
S-20	2.55	2.88	2.36	2.33	2.31
CS-10	2.50	3.44	3.27	3.02	3.34
CS-15	2.50	2.80	2.65	2.67	2.68
CS-20	2.50	2.88	2.63	2.70	2.77

X-Ray powder diffraction analysis

Results of the X-ray powder diffraction analysis of the flotation tailing samples before and after treatment are shown in Figs. 2 and 3, respectively. The results of the quantitative analysis of the diffraction data are presented in Table III. The most abundant mineral in all samples was quartz (SiO_2). Pyrite (FeS_2) and secondary copper-containing mineral langite ($Cu_4[(SO_4)(OH)_6]\cdot 2H_2O$) were also detected in all samples.

DISCUSSION

The aim of the study was to test the potential of acid water from Lake Robule for bioleaching of copper from the old flotation tailings of the Copper Mine Bor. In order to evaluate this, changes in concentrations of copper, and ferrous and ferric iron were monitored for five weeks in the bio- and acid- leaching experiments.

The obtained results showed that changes in the concentrations followed a similar pattern regardless of which sample of flotation tailings was used (Fig. 1). Within the first two weeks of bioleaching, the copper concentration in all three samples rose rapidly, and continued to rise during the third week but more slowly. Although decreases in the copper concentrations were observed during the fourth week, in the last week of the bioleaching experiment, the concentration of copper increased and reached the maximal level (Fig. 1A). The increase in concentrations of ferrous iron during bioleaching (Fig. 1B), which coincided with the initial rise in the copper concentrations, indicates that leaching of copper during the first three weeks was, at least partly, mediated by the presence of ferric iron that originated from the lake water. On the other hand, the increase in concentrations of ferric iron were slower until the fourth week, when a steep increase in its concentrations (Fig. 1C) and a decrease in ferrous iron concentrations (Fig. 1B) were detected. Although changes in the concentrations of copper, ferrous and ferric iron in the acid leaching experiments showed a certain similarity with those during bioleaching, differences were clearly visible in the fourth week, especially in the ferric iron concentrations (Fig. 1C). These results suggest

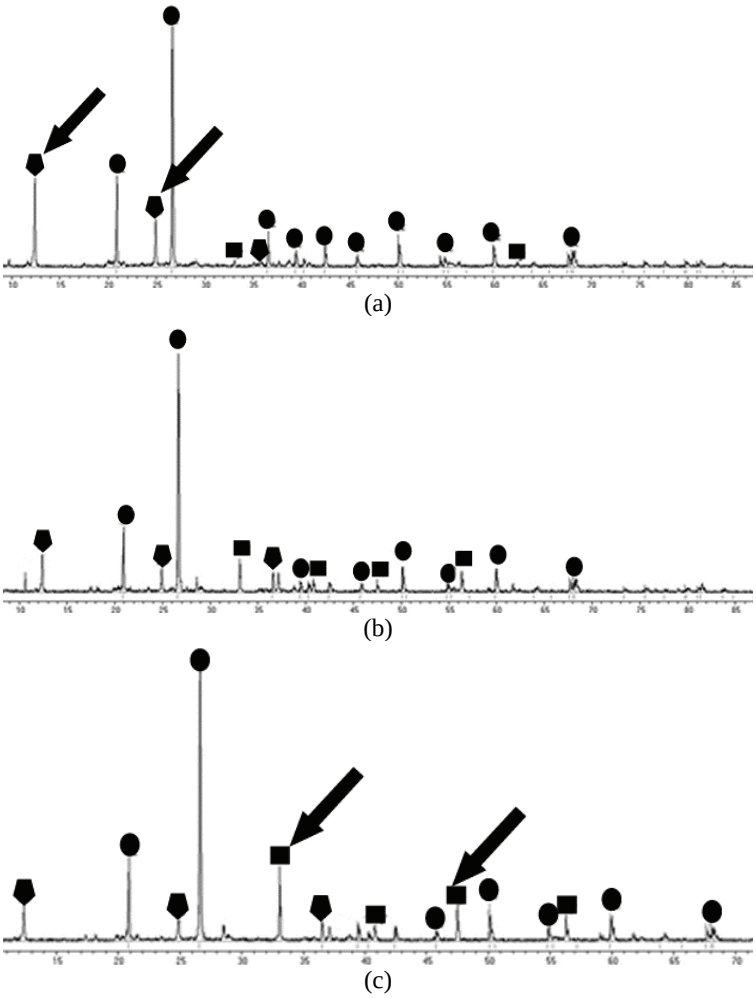
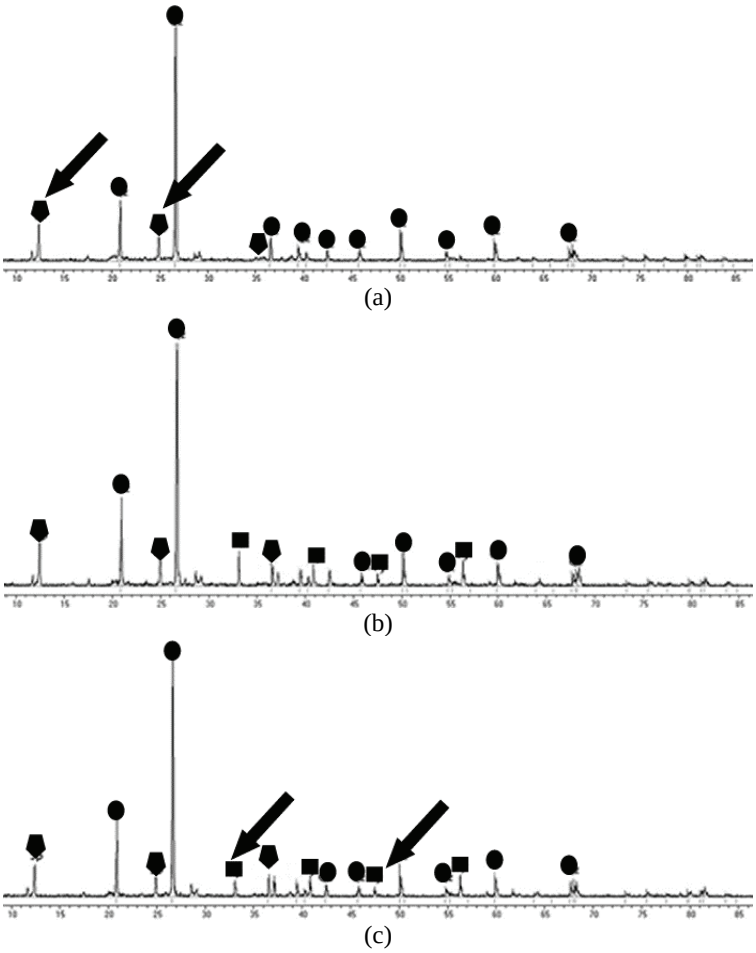


Fig. 2. X-Ray powder diffraction analysis of the tailings samples before treatment with lake water. Changes in peak heights that appear to be significant are marked with arrows. The samples of flotation tailings taken from depths of: a) 10, b) 15 and c) 20 m.

that bacterial influence on the oxidation of sulfide minerals in the first three weeks was lower than in the last two weeks, when the bacteria exhibited their maximal activity.

During the process of bioleaching, acidophilic iron oxidizers, such as *Acidithiobacillus ferrooxidans* or *Leptospirillum ferrooxidans*, attach to the mineral surface and produce a biofilm of exopolysaccharides, which facilitates the oxidation of pyrite and other sulfide minerals. After attachment, the bacteria in the biofilm multiply rapidly[15–18]. Zeng *et al*. tracked a number of attached and planktonic moderately thermophilic acidophiles during bioleaching of chalcopyrite.[18]



(a)

(b)

(c)

Fig. 3. X-Ray powder diffraction analysis of the tailings samples after treatment with lake water. Changes in peak heights that appear to be significant are marked with arrows. The samples of flotation tailings taken from depths of: a) 10, b) 15 and c) 20 m.

TABLE III. Results of the semi-quantitative analysis of the X-ray powder diffraction data of the samples of flotation tailings before and after bioleaching. Content of the mineral in the sample, %. Abbreviations: BT – before treatment, AT – after treatment

Sample	Component		
	Quartz	Pyrite	Langite
S-10 BT	76	3	22
S-10 AT	86	2	12
S-15 BT	74	10	16
S-15 AT	76	11	13
S-20 BT	72	20	8
S-20 AT	80	11	9

The authors reported that during bioleaching, the number of attached cells reached its maximum on the 20th day and remained stable until the 36th day, when it began to decline. On the other hand, the number of free cells was lower in comparison to attached cells during the first three weeks of the experiment. Subsequently, the number of free cells began to increase rapidly and significantly outnumbered the attached cells. The attached cells oxidized chalcopyrite causing the release of ferrous iron that was rapidly oxidized to ferric iron by the free bacterial cells, stimulating the growth of planktonic iron oxidizing bacteria and causing an increase in the concentration of ferric iron in the solution. After approximately 30 days of the experiment, the number of free cells started to decrease. Similarly, Zao *et al*. reported that the number of free cells during bioleaching of chalcopyrite with a pure culture of *Acidithiobacillus ferrooxidans* was the highest on the 50th day of experiment, as well as concentration of ferric iron and positive redox potential.[17] The concentration of ferrous iron increased until the 50th day, when a drop in its concentration was detected. After 50 days of bioleaching, the number of free cells in the solution began to decline, causing, on one hand, a decline in the concentrations of ferric iron and the redox potential, and, on the other hand, an increase in the concentrations of ferrous iron.[18]

These studies and the present results indicate that after three weeks of the bioleaching experiment presented in this paper, the number of attached cells probably reached its maximum, causing a rapid dissolution of pyrite and other sulfide minerals. Simultaneously, the ferrous iron that was released into solution was quickly oxidized to ferric iron by the fast growing free bacterial cells, stimulating further oxidation of pyrite and other metal sulfides. The number of free cells probably reached its maximum during the fourth week of the experiment, and after this period, their number started to decline.

The efficiency of bioleaching depends strongly upon the minerals that make up the treated material. Bioleaching of copper oxide minerals, such as tenorite (CuO), cuprite (Cu_2O), malachite ($Cu_2(CO_3)(OH)_2$) and chrysocolla ($CuSiO_3 \cdot 2H_2O$) are the most efficient.[8] Generally, bioleaching of copper sulfides is more difficult and requires more time than the bioleaching of copper oxides.[8] Copper sulfides, such as covellite (CuS), bornite (Cu_5FeS_4) and chalcocite (Cu_2S), can be successfully oxidized during bioleaching and the yield of copper is usually substantial.[20] On the other hand, the formation of a passivation layer of jarosites during the bioleaching of chalcopyrite($CuFeS_2$) prevents its oxidation in the early stage of the bioleaching process and hence, the yield of copper after bioleaching of chalcopyrite at ambient temperature is significantly lower in comparison to the bioleaching of covellite, chalcocite and bornite.[17,20] Dew *et al*. studied the bioleaching of various types of sulfide minerals and reported leaching rates of minerals from high to low: chalcocite, bornite, cubanite, covellite, enargite, carrolitte and chalcopyrite.[21] The most abundant copper sulfide mineral in the old flotation tail-

ings of the Copper Mine Bor is covellite, followed by chalcopyrite, enargite and chalcocite.[4–6]

The yield of copper after five weeks of bioleaching was 68.34±1.21 % for the S-15 sample, 73±0.6 % for the S-20 sample and 98±5.5 % for the S-10 sample (Table I). The unusually high content of copper in the S-10 sample (5100 mg kg^{-1}) is the consequence of precipitation of copper sulfate in the upper layers of the flotation tailings, which have been in contact with air and water for decades. Stevanović reported that the surface layers of the flotation tailings dump at depths of 0 and 1 m contain only 260 and 320 mg kg^{-1} of copper, respectively.[5] Content of copper in the dump is the highest at depths of 10 m and 20 m.[4–6] The exposure of the surface layers of the dump to air and water for years have initiated chemical and biological processes that led to almost complete oxidation of pyrite and other sulfide minerals (Table III, Fig. 2). X-Ray powder diffraction analysis revealed the presence of the secondary copper-containing mineral langite $Cu_4[(SO_4)(OH)_6]\cdot 2H_2O$ in all samples (Fig. 2). Langite usually occurs in mine dumps in which copper sulfate forms because of oxidation of the copper sulfides.[22] The content of this mineral was highest in the S-10 sample, and was significantly reduced after bioleaching (Table I, Fig. 1). Dissolution of the copper sulfate and langite in acidic solution was, presumably, the reason for a very high yield of copper, which was detected during both bioleaching and acid leaching of the S-10 sample (Fig. 1; Table I). Certain amounts of copper were also released from copper oxides that were formed during the oxidation of copper sulfides.[4–6] The content of copper in the 15-m sample was lower in comparison to the 20-m sample (3300 and 4100 mg kg^{-1}, respectively). This finding was expected assuming that the content of copper in the mined ore had been gradually decreasing over the years, and that the flotation technology improved and become more efficient with time. According to the data presented in this paper and data published by other authors,[4–6] the old flotation tailings can be divided into three zones: oxidized zone with the lowest content of copper (0–5 m) where sulfide minerals are almost completely oxidized, zone of precipitation (5–15 m) where precipitation of copper sulfate occurred, and unoxidized zone (15–30 m) where the oxidation rate of sulfide minerals was the lowest (Fig. 4).

The amount of pyrite in the 15-m sample was higher in comparison to that in the 10-m sample, and the content of langite was lower (Table III), which implies a lower oxidation rate of sulfide minerals in this layer. The content of pyrite and langite in the 15-m sample remained unchanged after bioleaching (Fig. 3, Table III). The highest amount of pyrite and lowest amount of langite were detected in the 20-m sample. Deeper sections of the old flotation tailings were in contact with air and water for a shorter period than the tailings in the upper sections of the dump. Due to a lower oxidation rate of sulfide minerals, the content of pyrite in the deeper sections is higher than its content in the upper sections of the

tailings and, consequently, the content of langite in the 20-m sample was much lower in comparison to that in the 10-m sample (Table III). After the bioleaching experiment, the content of pyrite in the 20-m sample was reduced (Fig. 3, Table III), as the process of bioleaching generates ferric iron that is able to oxidize pyrite, and thus reduce the amount of this mineral in the tailings.

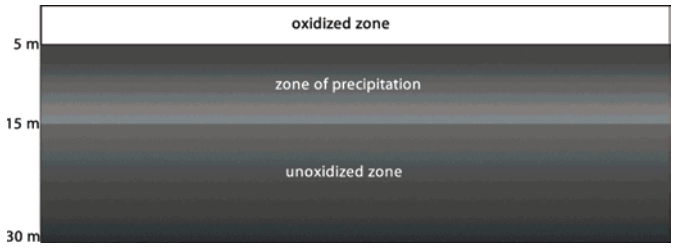


Fig. 4. Stratification of the old flotation tailings dump based on the content of copper. Shades of grey indicate copper concentrations in different zones.

Stevanović *et al*. performed an acid leaching experiment on the old flotation tailings samples in columns with or without the addition of oxidizing agents, with a 1:1 sample to acid solution ratio.[5] The highest yield of copper during this acid leaching experiment was demonstrated for the 10-m sample and it was 89.87 %. Maximal yields of copper detected after acid leaching with the addition of oxidizing agents were 73.97 and 48.12 % for the 15-m and the 20-m samples, respectively.[5] The degree of copper leaching from the 20-m sample was lower than those from the 10-m and 15-m samples,[5] which was not surprising, since the results of X-ray powder diffraction (Fig. 2) showed that the content of sulfide minerals was the highest in the 20-m sample.

The yield of copper detected after bioleaching of the 20-m sample reported in this paper was 73±0.6 % (Table I). By comparing the results obtained after acid leaching of the flotation tailings[4–6] and the results obtained by bioleaching of these samples presented in this paper, it could be concluded that bioleaching was more efficient than acid leaching for the treatment of the flotation tailings with higher sulfide mineral contents, which is located in the deep layers of the dump. In order to confirm further this conclusion, the next stage of the research will focus on the bioleaching of flotation tailings samples in columns. These column bioleaching tests will also provide an estimation of amounts of copper that could be recovered from the flotation tailings during large-scale heap bioleaching.

Lake Robule is marked as an environmental “hot spot” in the Bor mining area. In order to stop further pollution, treatment of the acidic water that constantly flows from the Lake is urgently required. The most efficient approach for such treatment of highly polluted acidic waters is the application of technology based on H_2S produced by sulfur-reducing bacteria in sulfidogenic bioreactors.

Hydrogen sulfide reacts with metal cations in wastewater thereby forming metal sulfides. Due to their low solubility, the metal sulfides precipitate and their selective precipitation is accomplished by controlling the pH of the solution.[23,25] The precipitated metal sulfides could be used for the production of pure metals, such as copper and zinc. This technology successfully removes As, Sb, Cu, Fe, Ni, Zn, Co, Cd, Mn and Pb from acidic waste waters.[23] After neutralization of the acidity, the treated water is pure enough to be released into the environment.[23] Heap bioleaching using acidic lake water as lixiviant, coupled with the above described technology, could turn the expensive process of wastewater purification into a profitable operation. The main disadvantage of this technology is the substantial initial capital investment.[23]

CONCLUSIONS

Basic bioleaching experiments in shaken flasks showed that significant amounts of copper could be recovered from samples taken from the old flotation tailings of the Copper Mine Bor (around 80 % on average), by application of water from the extremely acidic metal-rich water body of Lake Robule as lixiviant. It was also shown that after bioleaching of the 20-m sample, the content of pyrite in the sample was reduced. This last finding is especially important if one considers the fact that reducing the amount of pyrite could decrease the negative environmental impact of the tailings. It is an obvious advantage of bioleaching over acid leaching, since protons alone are not able to break the chemical bonds in the pyrite molecule; hence, this mineral is unaffected by acid leaching.

In general, there are at least three benefits of the presented approach: the recovery of substantial amounts of copper, a reduction of the environmental impact of the Lake of the tailings, and the use of abundant and free water from an extremely acidic lake.

Acknowledgments. The authors are grateful to Dr. Zoran Stevanović from the Institute of Mining and Metallurgy in Bor for providing samples of the old flotation tailings from the Copper Mine Bor. SS would like to thank Prof. Dr. Ljiljana Karanović from the University of Belgrade, Faculty of Mining and Geology, for conducting the X-ray powder diffraction analysis. This study was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, Projects Nos. 176016 and 173048.

ИЗВОД

БИОЛУЖЕЊЕ УЗОРАКА ФЛОТАЦИОНЕ ЈАЛОВИНЕ ИЗ СТАРОГ ФЛОТАЦИОНОГ ЈАЛОВИШТА РУДНИКА БАКРА БОР

СРЂАН СТАНКОВИЋ[1], ИВАНА МОРИЋ[2], АЛЕКСАНДАР ПАВИЋ[2], САНДРА ВОЈНОВИЋ[2], БРАНКА ВАСИЉЕВИЋ[2] и ВЛАДИЦА ЦВЕТКОВИЋ[1]

[1]Универзитет у Београду, Рударско–геолошки факултет, Ђушина 7, Београд и [2]Универзитет у Београду, Институт за молекуларну генетику и генетички инжењеринг, Војводе Степе 444а, Београд

У раду је описан експеримент биолужења узорака флотационе јаловине узетих са дубина од 10, 15 и 20 m из старог флотационог јаловишта Рудника бакра Бор. Као лужни

раствор је употребљена вода из екстремно киселог језера Робуле (Бор). Принос бакра након пет недеља лужења је био: 68±1,2 %, након третмана узорка са дубине од 15 m, 73±0,6 %, за узорак са дубине од 20 m и 98±5,5 %, за узорак са дубине од 10 m. Најмањи садржај сулфидних минерала је детектован у узорку са дубине од 10 m, а највећи у узорку узетом са дубине од 20 m. Резултати биолужења су упоређени са резултатима лужења истих узорака раствором сумпорне киселине. Применом методе биолужења добијен је значајно већи принос бакра из узорака са дубине од 10 и 20 m у односу на лужење истих узорака раствором сумпорне киселине. Такође, садржај пирита у узорку са дубине од 20 m је смањен након биолужења. Предности примене технологије биолужења за добијање бакра из флотационе јаловине су: екстракција значајне количине бакра из третираних узорака, смањење утицаја третиране јаловине на животну средину и употреба воде из екстремно киселог језера Робуле, које има у изобиљу. Резултати указују на то да би биолужење могло бити ефикасније у односу на киселинско лужење за екстракцију бакра из дубљих слојева старог флотационог јаловишта, које карактерише висок садржај сулфидних минерала.

(Примљено 11. априла, ревидирано 4. јула, прихваћено 23. септембра 2014)

REFERENCES

1. V. Simić, *Rudarski glasnik* **1** (1965) 27 (in Serbian)
2. M. Bugarin, Z. Stevanović, V. Gardić, V. Marinković, *Mining Eng.* **4** (2011) 11 (in Serbian)
3. G. Bogdanović, M. Trumić, M. Trumić, D. V. Antić, *Reciklaža i održivi razvoj* **4** (2011) 37 (in Serbian)
4. M. M. Antonijević, M. D. Dimitrijević, Z. O. Stevanović, S. M. Serbula, G. D. Bogdanović, *J. Hazard. Mater.* **158** (2008) 23
5. Z. Stevanović, M. Antonijević, G. Bogdanović, M. Bugarin, V. Trujić, R. Marković, D. Nedeljković, *Carpath. J. Earth Env.* **8** (2013) 29
6. Z. Stevanović, *PhD Thesis*, University of Belgrade, Technical Faculty in Bor, 2012 (in Serbian)
7. Bor Mining and Smelting Company official website, *Geology and Resources*, , http:// //rtb.rs/rtb-bor-doo/geologija-i-resursi/ (in Serbian, accessed Apr, 2014)
8. H. R. Watling, *Hydrometallurgy* **84** (2006) 81
9. M. Korać, Ž. Kamberović, *Metall. Mater. Eng.* **13** (2007) 41
10. V. P. Beškoski, P. Papić, V. Dragišić, V. Matić, M. M. Vrvić, *Adv. Mater. Res.* **71–73** (2009) 105
11. S. Stanković, I. Morić, A. Pavić, B. Vasiljević, D. B. Johnson, V. Cvetković, *J. Serb. Chem. Soc.* **79** (2014) 729
12. M. A. Anwar, M. Iqbal, M. A. Qamar, M. Rehman, A. M. Khalid, *World J. Microb. Biot.* **16** (2000) 135
13. P. Komadel, J. Stucki, *Clays Clay. Miner.* **36** (1998) 379
14. C. L. Luke, *Anal. Chim. Acta* **36** (1966) 122
15. T. Rohwerder, T. Gehrke, K. Kinzler, W. Sand, *Appl. Microbiol. Biotechnol.* **63** (2003) 239
16. H. Tribusch, J. Rojas-Chapana, in *Biomining*, D.E. Rawlings, D.B. Johnson, Eds., Springer-Verlag, Berlin, 2007, p. 263
17. Y. Dong, H. Lin, X. Xu, S. Zhou, *Hydrometallurgy* **140** (2013) 42
18. W. Zeng, S. Tan, M. Chen, G. Qiu, *Hydrometallurgy* **106** (2010) 46
19. X. Zao, R. Wang, X. Lu, J. Lu, C. Li, J. Li, *Miner. Eng.* **53** (2013) 184

20. M. Pradhan, K. C. Nathsarma, K. Srinavasa Rao, L. B. Sukla, B. K. Mishra, *Miner. Eng.* **21** (2008) 355
21. D. W. Dew, C. Van Burren, K. McEwan, C. Bowker, *J. South. Afr. Inst. Min. Metall.* **11** (2000) 409
22. A. Livingstone, H. G. Macpherson, B. Jackson, *Mineral. Mag.* **40** (1976) 893
23. Paques official website: *THIOTEC^{TM} metal recovery technology*, http://en.paques.nl/ /products/other/thioteqmetal (accessed Apr, 2014)
24. A. E. Lewis, *Hydrometallurgy* **104** (2010) 222
25. E. Jameson, O. F. Rowe, K. B. Hallberg, D. B. Johnson, *Hydrometallurgy* **104** (2010) 488.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 407–419 (2015)
JSCS–4725
UDC 582.26+66.081.000.57:669.24/.26+669.3/.5+661.183.2:544.4
Original scientific paper

Biosorption characteristics of *Spirulina* and *Chlorella* cells for the accumulation of heavy metals

ANIKÓ KŐNIG-PÉTER[1], FERENC KILÁR[1,2,3], ATTILA FELINGER[2,3] and TÍMEA PERNYESZI[2,3]*

[1]*University of Pécs, Faculty of Medicine, Institute of Bioanalysis, Szigeti út 12, H-7624 Pécs, Hungary,* [2]*University of Pécs, Faculty of Science, Department of Analytical and Environmental Chemistry Ifjúság útja 6, H-7624 Pécs, Hungary and* [3]*University of Pécs, Analytical Chemistry and Geoanalytical Research Group, Szentágothai Research Center, Ifjúság útja 20, H-7624 Pécs, Hungary*

(Received 21 March, revised 28 May 2014, accepted 2 June 2014)

Abstract: The biosorption of heavy metals by dried *Chlorella vulgaris* and *Spirulina platensis–Spirulina maxima* cells was studied under various experimental conditions. The effects of biosorbent dosage, pH, adsorption time, temperature, initial metal concentration on the biosorption were examined. A biosorption process can be divided into two parts: the first part follows zero-order and the second part pseudo second-order kinetics. Characterization of biosorption equilibrium was evaluated employing the Langmuir and Dubinin–Radushkevich models using non-linear regression. The optimum pH range was found to be 5.0−6.0 for Pb(II) and 4.0−6.0 for Cu(II) and Cd(II) adsorption. Based on the experimental data, the maximum adsorption capacities for Pb(II), Cd(II) and Cu(II) were 144, 161 and 138 mg g^{-1} by *Chlorella* cells and 370, 201 and 165 mg g^{-1} by *Spirulina* cells. The corresponding values for activated carbon were 86, 134 and 43 mg g^{-1}, respectively.

Keywords: algae; kinetics; isotherm; Langmuir; Dubinin–Radushkevich; activated carbon.

INTRODUCTION

Heavy metal pollution is one of the most important environmental problems today.[1] Cadmium, copper and lead are heavy metals that pose serious health hazards through entry into the food chain by anthropogenic pathways.[2] Most industrial technologies for their removal are not cost effective if the concentration of the heavy metal pollution is under 100 mg L^{-1}.[3] The disadvantages of the conventional methods are long operation time or the use of other chemicals, which could introduce toxic or carcinogenic intermediates into waterbodies.

* Corresponding author. E-mail: ptimea@gamma.ttk.pte.hu
doi: 10.2298/JSC140321060P

Certain natural materials of biological origin, such as bacteria, fungi, yeast and algae possess metal-sequestering properties and therefore could be used to rapidly decrease the concentration of heavy metal ions with high efficiency. These biosorbents are ideal candidates for the treatment of high-volume and low-concentration complex wastewaters contaminated with heavy metal ions.[4,5]

Living and dead biomass of algae cells can be used to decrease environmental heavy metal pollution. Several algae were tested for their ability to adsorb heavy metals.[6–11] The use of dead biomass is favourable because it tolerates high concentrations of toxic ions, a nutrient supply unnecessary and the culture conditions are not limiting.[6] Algae are widely used for several purposes, such as food for human consumption, a source of phycocolloids, β-carotane, atraxanthine and chlorophyll.[7] The most important criteria for developing algal biosorbents to decrease harmful metal contents in soil and water are the followings: ability for rapid adsorption of huge amounts of metal ions, availability in large quantities because of worldwide cultivation and relatively cheap production.[12,13]

Spirulina and *Chlorella* cells were used by many researchers to test their ability to remove heavy metal ions from polluted water.[7,12,14,15] The use of various conditions for the adsorption resulted in different adsorption properties. Mehta and Gaur compared the different adsorption features of living and dead biomass of *Chlorella vulgaris* cells.[6] Chojnacka and co-workers examined different morphological types of *Spirulina* species to compare the binding characteristics of these algae in case of chromium, cadmium and copper.[16] Fang *et al.* described the functional groups that could play a role in the surface adsorption of copper and cadmium by *Spirulina* cells with the help of FTIR and X-ray analyses. Al-Rub studied the effect of competitive cations on copper adsorption by *Chlorella* cells.[13] However, there is a lack of comparisons with often used adsorbents, such as activated carbon, in the literature.

The present study considers the adsorption properties of two different alga strains: green microalgae *C. vulgaris* and blue–green alga (cyanobacteria) *S. platensis–S. maxima* mixture. Their properties were compared with respect to the effect of pH, temperature and biosorbent dosage on Pb(II), Cd(II) and Cu(II) biosorption, with an evaluation of the kinetics and equilibrium capacities of the biomasses. The biosorption experiments were also performed using activated carbon. The obtained results showed that these biosorption systems using algal cells represent promising alternatives for the removal of heavy metal ions from polluted aqueous environments.

EXPERIMENTAL

Materials and adsorbent

C. vulgaris and *S. platensis–S. maxima* cells were purchased (Czech Academy of Sciences) in the dried form. Activated carbon was purchased from VWR Prolabo Ltd. (USA). The heavy-metal test solutions containing Pb(II), Cd(II) and Cu(II) ions were prepared from

reagent-grade metal of hydrated $Pb(NO_3)_2$, $Cd(NO_3)_2$ and $CuCl_2$ (Fluka, Germany) in the concentration range of 25–500 mg L^{-1}.

Determination of surface charge

Dried *C. vulgaris* or *S. platensis–S. maxima* biomass (1 g L^{-1}) was suspended in 25 mL distilled water in the pH range 2–11. The required pH was regulated with 0.1 M NaOH and 0.1 M HCl solutions. The specific surface charge was determined with a Zeta Sizer (Zetasizer Nano-Z, Malvern Instruments Ltd., Malvern, UK). The values were determined after the evaluation of 25 to 50 measurements.

Effect of pH on biosorption

The pH effect was tested using single-metal solutions containing 50 mg L^{-1} Pb(II), Cd(II) and Cu(II) ion. After 24 h of incubation, samples were taken from the suspensions, and the biomass was separated from the heavy metal solutions at 10 000 rpm for 5 min. The concentration of heavy-metals in supernatants was evaluated by atomic absorption spectrometry (AAS) (Perkin–Elmer 2380) at 217 nm for Pb(II), at 228 nm for Cd(II) and at 324.8 nm for Cu(II).

Effect of adsorbent dose on biosorption

Different dosages of algal biosorbents (0.25, 0.5, 1.0 and 2.0 g L^{-1}) were used at constant pH to find the optimal concentration. Solutions with 50 mg L^{-1} initial Pb(II) ion concentration were used under batch conditions. The heavy metal content in the supernatant was measured by AAS. The efficiency of the removal was calculated using the following equation:

$$\text{Removal efficiency} = \frac{100(c_0 - c_e)}{c_0} \quad (1)$$

c_0 and c_e are the metal concentrations initially and in the equilibrium (mg L^{-1}), respectively.

Adsorption dynamics

Batch sorption studies were investigated with 50 mg L^{-1} initial heavy metal concentration. Samples were taken at desired time periods and the residual Pb(II), Cd(II) or Cu(II) concentrations were determined.

Equilibrium evaluation based on biosorption isotherms

The biomass *C. vulgaris* and *S. platensis–S. maxima* cells (1 g L^{-1}) were suspended in heavy metal solutions, which were gently agitated at room temperature. For determination of adsorption isotherms, Pb(II), Cd(II) and Cu(II) solutions were used in the concentration range of 25–500 mg L^{-1} at a biosorbent dosage of 1.0 g L^{-1}. After incubation for 24 h, samples were taken from the suspensions and the heavy metal content in the supernatant was measured by AAS. The metal uptake per gram of adsorbent, q (mg g^{-1}), was calculated using the equation:

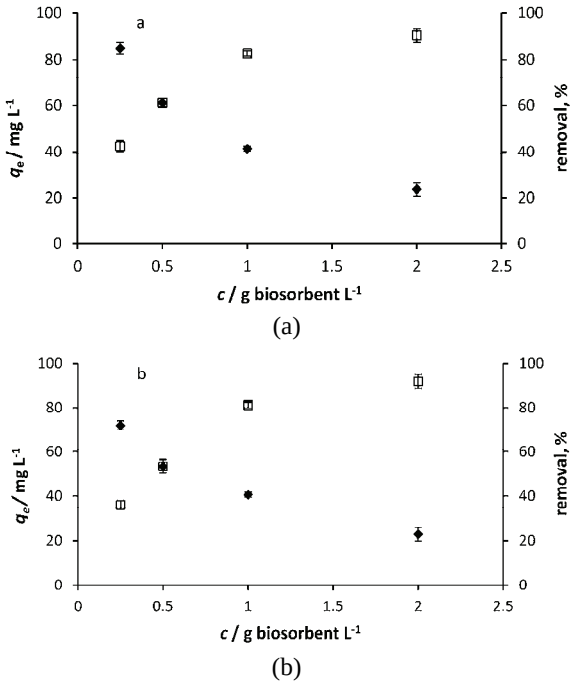
$$q = \frac{(c_0 - c_e)V}{m} \quad (2)$$

where V is the volume of the solution (L) and m is the mass of biosorbent (g). All experiments were performed in triplicate.

RESULTS AND DISCUSSION

Effect of biosorbent dose

Biosorbent concentration is one of the most important factors affecting biosorption. In this study, it was varied between 0.25 and 2.0 g L^{-1}. The effect of the biomass loading of the adsorption of Pb(II) by *Spirulina* and *Chlorella* cells is shown in Fig. 1, from which it could be seen that the value of biomass concentration influenced the uptake capacity. The uptake capacity increased[17] with decreasing dosage of the biosorbent, which may be due to interference between the binding sites on the biosorbent. A similar relationship was reported in the literature.[7,12,13] Such an adsorption behaviour could be described by the Kroeker relation. With increasing weight of biosorbent in the solution of defined volume and concentration, the specified amount adsorbed by the adsorbent decreased. By both organisms, a dosage of 1.0 g L^{-1} resulted in an efficiency of more than 80 %. On increasing the biosorbent concentration further (2.0 mg L^{-1}), the adsorbed amount increased by only 10 %. In the further experiments, an alga concentration of 1.0 g L^{-1} was used as the optimum for biosorption.



(a)

(b)

Fig. 1. Adsorption capacities of: a) *Spirulina* and b) *Chlorella* cells, in dependence on the biomass concentration at concentration of 50 mg lead ions. □ represents the removal in percent, and ◆ is the adsorption capacity.

Effect of pH on biosorption

The pH is an important factor in biosorption processes. Based on FTIR spectra analysis, the most dominant role in the binding play the carboxyl groups

from proteins and carboxylated polysaccharides.[18] These groups have the highest affinity for metal ions because they are deprotonated in the pH range of the highest adsorption. Other groups, such as phosphate, hydroxyl and amino groups, present in the surface area of alga cells can play only a minor role; their pK_a values allow deprotonation only at higher pH values.[16] The zeta potential determination of alga cell surfaces shows that they are negatively charged above pH 3 by both species (Fig. 2). Parallel to these findings, the adsorption profile by *Spirulina* and *Chlorella* cells are shown in Fig. 3 for different pH values. In the cases of Pb(II), Cd(II) and Cu(II) biosorption on *Chlorella* cells and for Pb(II) and Cu(II) adsorption on *Spirulina* cells, the uptake capacities were the lowest at pH 3 and remained nearby constant between pH 4−6. The adsorption of Cd(II) on *Spirulina* cells was not affected by the pH value. Above pH 6, Cu(II) begins to precipitate, which made the adsorption more difficult.

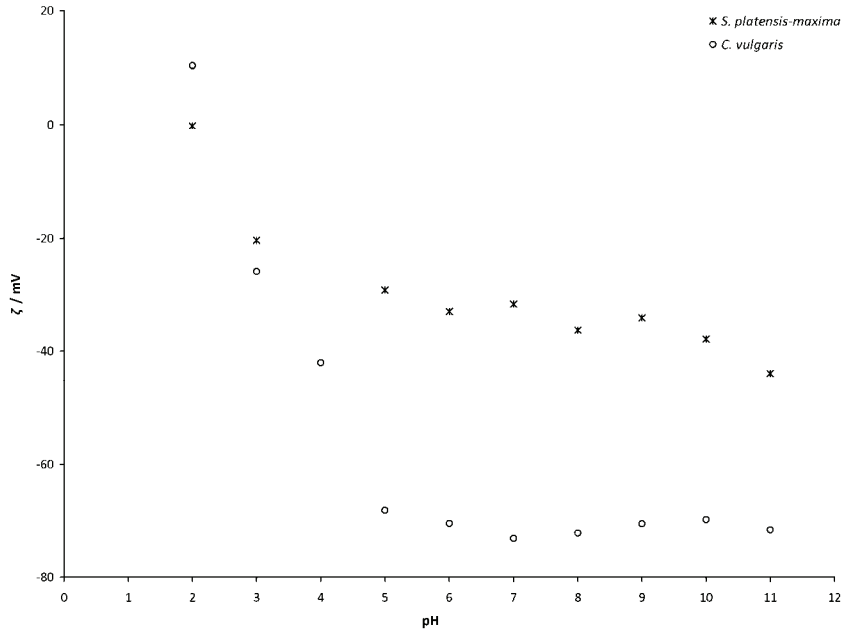


Fig. 2. Zeta potential values of *S. platensis*−*S. maxima* and *C. vulgaris* cells in the pH range of 2−11 for a biomass concentration of 1 g L^{-1} at 295 K.

The experimental results indicated that the biosorption process was effective in the pH range 5–6 for Pb(II) and pH 4−6 for Cd(II) and Cu(II) adsorption and thus no pH adjustment was necessary, because the initial pH values of all the metal solutions were within these ranges.

Effect of time on the biosorption

The adsorption of the heavy metal ions on the active sites of several types of biomass is generally considered very rapid.[2] This process may be independent of the initial metal concentration and the first period of the adsorption could be described with zero order adsorption kinetics. The practical aspect of the rapid kinetics is that a small reactor can ensure the cost-effective use.[14] Based on the morphology of the examined organisms, the heavy metal ions adsorb to the sheath

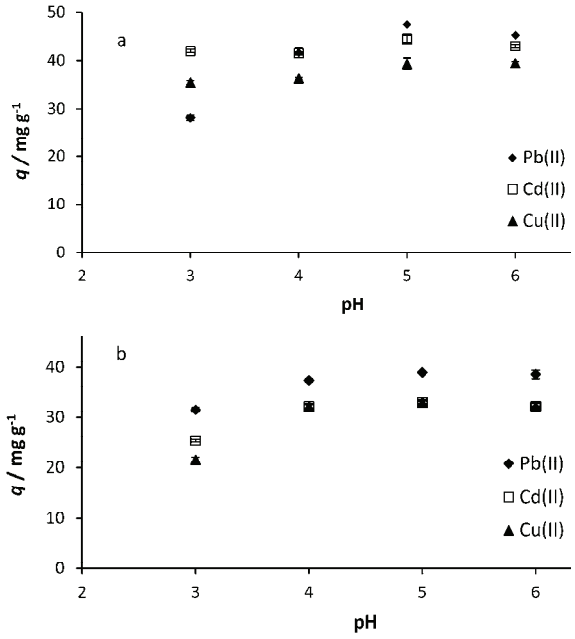


Fig. 3. Effect of pH on lead, cadmium and copper ions biosorption by: a) *Spirulina* and b) *Chlorella*, at 50 mg L^{-1} of heavy metal concentration and 1.0 g L^{-1} of algal cell concentration.

of the cells, very likely mainly by electrostatic attraction and complexation.[12] For the second phase of the adsorption, the chemisorption mechanism seems to be the most important force. The rate limiting step is the biosorption of positively charged heavy metal ions onto the cell surface.[19] Based on this finding, pseudo-second order kinetics were used to evaluate the kinetic parameters, because it takes into account the interaction between adsorbent and adsorbate.[2,6,15] The correlation coefficients for the second-order kinetic model were close to 1.0 for all cases. Other mechanisms examined in the kinetic evaluation were not so significant and hence the roles of physical adsorption and intraparticle diffusion were limited.[13,16] The zero order kinetic constants (k) and the pseudo-second order kinetic constants ($k_{2,ad}$), together with the calculated ($q_{eq,cal}$) and experimentally determined ($q_{eq,exp}$) adsorbed amounts of lead, cadmium and copper ions are presented in Table I.

After the first period of 8–10 min of adsorption (zero order kinetics), the equilibrium was established in 20–25 min (pseudo-second order kinetics) for each observed heavy metal ion, and no more heavy metal ions were adsorbed after this time. The data presented in Fig. 4 show that 70–90 % of the soluble ions were removed from the system. This adsorption profile may be characteristic for monolayer coverage. Similar findings were reported by Areco *et al*. for the adsorption of different heavy metals by *Avena fatua*.[20] Aksu reported the kinetic and thermodynamic parameters of Ni(II) adsorption by *C. vulgaris*.[14]

TABLE I. Constants for zero-order and pseudo second-order adsorption kinetic for the adsorption of lead, cadmium and copper ions by *Spirulina* and *Chlorella* cells

Biosorbent	Heavy metal	Zero-order kinetics	Pseudo-second order kinetics			
		k mg L^{-1} min^{-1}	$k_{2,ad}\times10^2$ g mg^{-1} min^{-1}	$q_{eq,cal}$ mg g^{-1}	R^2	$q_{eq,exp}$ mg g^{-1}
S. plat.– S. max.	Pb(II)	8.55	29	41.3	0.99	41.2
	Cd(II)	8.84	2.2	47.5	0.98	47.8
	Cu(II)	7.00	3.9	37.7	0.99	37.4
C. vulgaris	Pb(II)	7.57	5	40.3	0.99	40.6
	Cd(II)	6.48	52	32.7	0.99	32.4
	Cu(II)	5.9	8	31.3	0.99	31.2

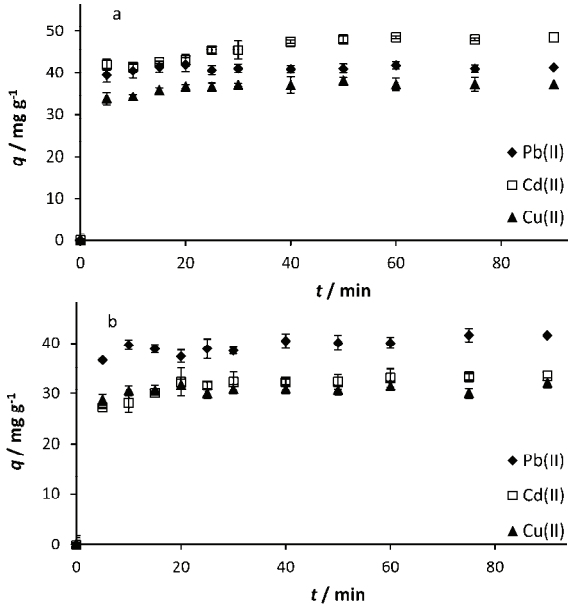


Fig. 4. A plot of the adsorbed amount *vs*. time for: a) *Spirulina* and b) *Chlorella* cells, at 50 mg L^{-1} initial lead, cadmium and copper ions concentrations.

Equilibrium studies of biosorption

The adsorption equilibrium relationship at a given temperature is referred to as an adsorption isotherm. Equations widely used to describe adsorption iso-

therms are the Langmuir, Freundlich, Dubinin–Radushkevich two-parameter models and Redlich–Peterson and Sips three-parameter models.[2,15,19,21,22] They describe the relationship between solute concentration remaining in the solution and the amount of solute adsorbed per unit mass of adsorbent. In this study, the Langmuir isotherm model was used because it is based on monolayer coverage over the cell wall, which could be inferred from the adsorption rate observations. It assumes identical surface sites that can accommodate one adsorbed particle. A number of studies demonstrated that despite these simplifications of the adsorption process, the Langmuir isotherm[2,15,16,20] could be used to interpret the biosorption data.[2] The model is given by Eq. (3):

$$q_e = \frac{q_{max} b c_e}{1 + b c_e} \quad (3)$$

where b is the adsorption equilibrium constant including the affinity of the binding sites (L mg^{-1}), c_e and q_e are unadsorbed metal ions in solution and adsorbed metal ions on the biosorbent at equilibrium, respectively, q_{max} is the maximum amount of metal ion per unit weight of adsorbent necessary to form a complex monolayer on the surface of the adsorbent (mg g^{-1}). The experimental values compared with the calculated data are presented in Table II. The highest b values (0.089 and 0.17) are obtained for Pb(II) adsorption for *Spirulina* and *Chlorella* cells, these values indicate steep initial slope and high affinity. Lower b values were calculated for Cd(II) and Cu(II) adsorption. According to this, the fitted curves have less steep initial slope.

TABLE II. The Langmuir and Dubinin–Radushkevich isotherm constants for lead, cadmium and copper ions adsorption by *Spirulina* and *Chlorella* cells at 295 K

Biosorbent	Heavy metal	Langmuir isotherm				Dubinin–Radushkevich isotherm		
		Q_{max} mg g^{-1}	B L mg^{-1}	R^2	$Q_D \times 10^3$ mol g^{-1}	E kJ mol^{-1}	R^2	Q_{exp} mg g^{-1} (mmol)
S. plat.– S. max.	Pb(II)	413	0.089	0.97	4.9	13.4	0.86	370 (1.8)
	Cd(II)	298	0.019	0.99	9.2	9.5	0.97	201 (1.8)
	Cu(II)	262	0.0039	0.99	9.6	8	0.99	165 (2.6)
C. vulgaris	Pb(II)	151	0.17	0.97	1.5	15.1	0.85	144 (0.7)
	Cd(II)	280	0.0085	0.99	8.7	8.5	0.95	161 (1.4)
	Cu(II)	233	0.0044	0.98	8.5	8	0.95	138 (2.2)

The parameters of this model were compared with the values calculated from the Dubinin–Radushkevich model,[21,22] which can describe the porous structure of a sorbent. It is more universal than the Langmuir isotherm model as it assumes a heterogeneous surface and not identical binding sites. The advantage of this model is that the mean energy of sorption can be determined and the magnitude of this value could be used to estimate the type of the sorption reaction. E values between 8.0−16 kJ mol^{-1} are in the range of an ion-exchange mechanism.[21,22]

This isotherm is quite often used to describe adsorption equilibrium.[21,22] The isotherm is given by Eqs. (4)–(6):

$$Q_0 = Q_D \exp\left(B_D \varepsilon_D^2\right) \tag{4}$$

$$\varepsilon_D = RT \ln\left(1 + \frac{1}{c_e}\right) \tag{5}$$

where Q_D characterizes the monolayer saturation capacity (mol g^{-1}), B_D is the Dubinin–Radushkevich model constant (mol^2 kJ^{-2}) and ε_D is the Polanyi potential. The mean adsorption energy can be calculated as follows:

$$E = \frac{1}{\sqrt{2B_D}} \tag{6}$$

The calculated E values are given in Table II, which shows that the values of E were 13.4 and 15.1 kJ mol^{-1} for Pb(II), 9.5 and 8.5 kJ mol^{-1} for Cd(II) and 8.0 kJ mol^{-1} for Cu(II) adsorption by *Spirulina* and *Chlorella* cells, respectively. The obtained E values were within the range of those denoting an ion-exchange mechanism was operative. The q_{max} values calculated from both models were much higher than the experimental values, which could indicate that the highest adsorbed amount was not reached. The fitted curves for the Langmuir and Dubinin–Radushkevich isotherms are shown in Figs. 5 and 6, respectively, for adsorption by *Spirulina* and *Chlorella* cells.

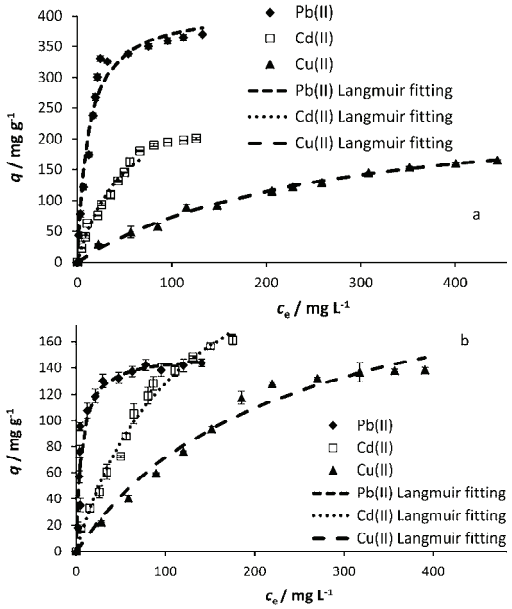


Fig. 5. Adsorption isotherms and Langmuir fitting of the adsorption of lead, cadmium and copper ions by: a) *Spirulina* and b) *Chlorella* cells, at 295 K.

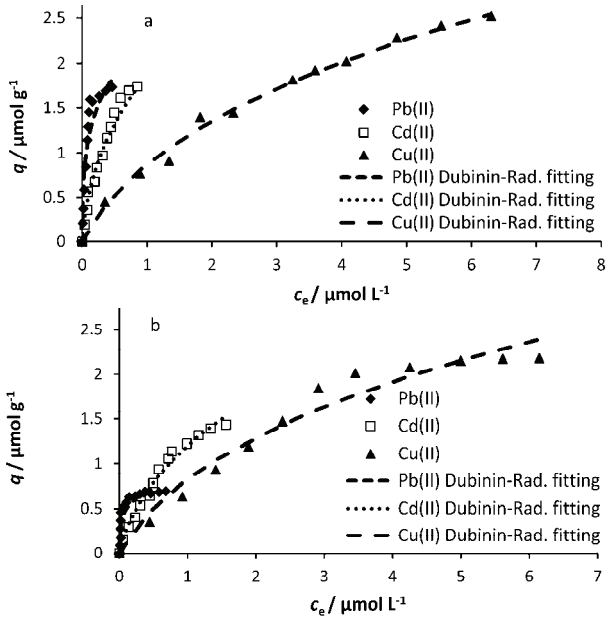


Fig. 6. Adsorption isotherms and Dubinin–Radushkevich fitting of the adsorption of lead, cadmium and copper ions by: a) *Spirulina* and b) *Chlorella* cells, at 295 K.

Effect of temperature

In most cases, the effect of temperature in the range of 295–313 K was negligible. The adsorptions did not change or showed only a slight increase with decreasing temperature. This finding was in good agreement with previously reported data.[23,24] A biosorption process may be slightly exothermic due to physical adsorption, however the contribution of physical adsorption in the overall process was maximum 4 %.[16] It could be assumed based on the obtained results that the heavy metal binding was mainly chemisorption based on ion-exchange. As an example, the isotherm for the adsorption of lead ions by *Chlorella* cells can be seen in Fig. 7.

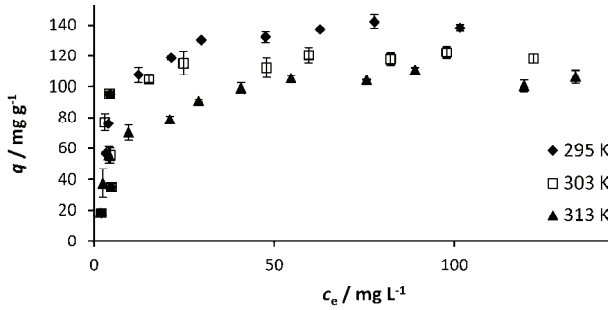


Fig. 7. Adsorption isotherms of lead ion adsorption by *Chlorella* cells at 295, 303 and 313 K.

Comparison with activated carbon

The experimental values for Pb(II), Cd(II) and Cu(II) adsorption on activated carbon were 86, 134 and 43 mg g^{-1}, respectively. The green alga *Chlorella* removed about 70 % more Pb(II) and 20 % more Cd(II) than activated carbon, and 4 times more Cu(II). The *Spirulina* strains were excellent biosorbents; they adsorb about 50 % more Cd(II) and 4 times more Pb(II) and Cu(II) than activated carbon. The adsorbed amounts of heavy metals by *Chlorella* cells, *Spirulina* cells and activated carbon are summarized in Table III. There is only little information about the comparison of biosorbents for heavy metal ions and activated carbon.[1,23]

TABLE III. Amounts of heavy metals adsorbed by *Chlorella*, *Spirulina* cells and activated carbon

Sorbent	Heavy metal							
	Pb(II)		Cd(II)		Cu(II)		Zn(II)	
	mg	mmol	mg	mmol	mg	mmol	mg	mmol
S. plat.–S. max.	370	1.8	201	1.8	165	2.6	128	2.0
C. vulgaris	144	0.7	161	1.4	138	2.2	127	1.9
Activated carbon	86	0.4	134	1.2	43	0.7	27	0.4

CONCLUSIONS

The biosorption of lead, cadmium and copper ions by dried *C. vulgaris* and *S. platensis–S. maxima* cells were studied using the batch technique. The adsorption was effective over a wide range of algal concentrations. The biosorption system containing 1.0 g L^{-1} adsorbent was optimal, and at this composition about 80 % of the heavy metal ions could be removed. An optimum pH range (pH 4–6) was determined for the biosorption process, which confirms the applicability of this system for wastewater treatment both at low and high heavy metal concentrations. In case of each metal by both organisms, the uptake capacity was the lowest at pH 3, and remained nearly constant between pH 4–6. Equilibrium was reached in 20–25 min for each heavy metal ion, and after this time, no more heavy metal ions were adsorbed. 70–90 % of the soluble ions was removed from the system. The biosorption process could be divided into two steps: the first part follows zero-order, the second part pseudo second-order kinetics. The equilibrium evaluation was described with adsorption models given by Langmuir and Dubinin–Radushkevich, and both models exhibited good fitting to the experimental values. The adsorption order was Pb < Cd < Cu for *Chlorella*, and Pb ≈ Cd < Cu for *Spirulina*. *Chlorella* and *Spirulina* algae biomass should be considered for use to decontaminate water from heavy metals.

Acknowledgement. T. P. and F. K. gratefully acknowledge the support for this research from TAMOP-4.2.2.A11/1/KONV-2012-0065 and OTKA 10067 (Hungarian research grants).

ИЗВОД

БИОСОРПЦИОНЕ КАРАКТЕРИСТИКЕ ЋЕЛИЈА *Spirulina* И *Chlorella* ПРИ АКУМУЛИРАЊУ ТЕШКИХ МЕТАЛА

ANIKÓ KŐNIG-PÉTER[1], FERENC KILÁR[1,2,3], ATTILA FELINGER[2,3] и TÍMEA PERNYESZI[2,3]

[1]*University of Pécs, Faculty of Medicine, Institute of Bioanalysis, Szigeti út 12, H-7624 Pécs, Hungary,* [2]*University of Pécs, Faculty of Science, Department of Analytical and Environmental Chemistry, Ifjúság útja 6, H-7624 Pécs, Hungary и* [3]*University of Pécs Analytical Chemistry and Geoanalytical Research Group, Szentágothai Research Center, Ifjúság útja 20, H-7624 Pécs, Hungary*

Биосорпција тешких метала осушеним ћелијама *Chlorella vulgaris* и *Spirulina platensis–Spirulina maxima* изучавана је у разним експерименталним условима. Испитани су утицаји дозе биосорбента, pH, времена адсорпције, температуре и почетне концентрације метала на биосорпцију. Процес биосорпције се може поделити на два дела: први део се одвија по кинетичком закону за реакцију нултог реда, а други део по кинетичком закону за реакцију псеудо-другог реда. Карактеристике биосорпционе равнотеже су одређене моделима Ленгмира и Дубињин–Радушкевича, коришћењем нелинеарне регресије. Нађено је да је оптимални опсег pH 5,0 до 6,0 за адсорпцију Pb(II) и 4,0 до 6,0 за Cu(II) и Cd(II). Максимални капацитети адсорпције за Pb(II), Cd(II) и Cu(II) су били 144, 161 и 138 mg g^{-1} на ћелијама *Chlorella*, а 370, 201 и 165 на ћелијама *Spirulina*, према експерименталним резултатима. Те вредности за активни угаљ су биле 86, 134 и 43 mg g^{-1}, редом.

(Примљено 21. марта, ревидирано 28. маја, прихваћено 2. јуна 2014)

REFERENCES

1. M. A. Hashim, S. Mukhopadhyay, J. N. Sahu, B. Sengupta, *J. Environ. Manage*. **92** (2011) 2355
2. J. Febrianto, A. N. Kosasih, J. Sunarso, Y. H. Ju, N. Indraswati, S. Ismadji, *J. Hazard. Mat*. **162** (2009) 616
3. S. S. Ahluwalia, D. Goyal, *Biores. Technol*. **98** (2007) 2243
4. C. Mack, B. Wilhelmi, J. R. Duncan, J. E. Burgess, *Biotechnol. Adv*. **25** (2007) 264
5. T. A. Kurniawan, G. Y. S. Chan, W.-H. Lo, S. Babel, *Sci. Total Environ.* **366** (2006) 409
6. S. K. Mehta, J. P. Gaur, *Ecol. Eng*. **18** (2001) 1
7. S. V. Gokhale, K. K. Jyoti, S. S. Lele, *Biores. Technol*. **99** (2008) 3600
8. K. Vijayaraghavan, D. Prabu, *J. Hazard. Mater*. **137** (2006) 558
9. M. T. Cabrita, J. Raimundo, P. Pereira, C. Vale, *Mar. Environ. Res*. **87**–**88** (2013) 37
10. J. L. Zhou, P. L. Huang, R. G. Lin, *Environ. Pollut*. **101** (1998) 67
11. I. Moreno-Garrido, O. Campana, L. M. Lubian, J. Blasco, *Mar. Pollut. Bull*. **51** (2005) 823
12. C. Solisio, A. Lodi, P. Torre, A. Converti, M. Del Borghi, *Biores. Technol*. **97** (2006) 1756
13. F. A. A. Al-Rub, M. H. El-Naas, I. Ashour, M. Al-Marzouqi, *Process Biochem.* **41** (2006) 457
14. Z. Aksu, *Process Biochem.* **38** (2002) 89
15. A. Celekli, M. Yavuzatmaca, H. Bozkurt, *J. Hazard. Mater*. **173** (2010) 123
16. K. Chojnacka, A. Chojnacki, H. Gorecka, *Chemosphere* **59** (2005) 75
17. F. Veglio, F. Beolchini, *Hydrometallurgy* **44** (1997) 301
18. L. Fang, C. Zhou, P. Cai, W. Chen, X. Rong, K. Dai, W. Liang, J.-D. Gu, Q. Huang, *J. Hazard. Mater*. **190** (2011) 810

19. K. Vijayaraghavan, Y. S. Yun, *Biotechnol. Adv*. **26** (2008) 266
20. M. M. Areco, L. Saleh-Medina, M. A. Trinelli, J. L. Marco-Brown, M. D. Afonso, *Colloids Surfaces, B* **110** (2013) 305
21. K. Vijayaraghavan, T. V. N. Padmesh, K. Palanivelu, M. Velan, *J. Hazard. Mater*. **133** (2006) 304
22. M. F. Ahmad, S. Haydar, T. A. Quraishi, *Int. Biodeter. Biodegr*. **83** (2013) 119
23. I. Han, M. A. Schlautman, B. Batchelor, *Water Environ. Res.* **72** (2000) 29.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 421–433 (2015)
JSCS–4726
UDC 547.211:66.011:510.644:519.233.5
Original scientific paper

Modeling of methane emissions using the artificial neural network approach

LIDIJA J. STAMENKOVIĆ[1#], DAVOR Z. ANTANASIJEVIĆ[2*#], MIRJANA Đ. RISTIĆ[1#], ALEKSANDRA A. PERIĆ-GRUJIĆ[1#] and VIKTOR V. POCAJT[1#]

[1]University of Belgrade, Faculty of Technology and Metallurgy, Karnegijeva 4, 11120 Belgrade, Serbia and [2]University of Belgrade, Innovation Center of the Faculty of Technology and Metallurgy, Karnegijeva 4, 11120 Belgrade, Serbia

(Received 2 April, revised 26 August, accepted 7 November 2014)

Abstract: The aim of this study was to develop a model for forecasting CH_4 emissions at the national level, using artificial neural networks (ANN) with broadly available sustainability, economical and industrial indicators as their inputs. ANN modeling was performed using two different types of architecture; a backpropagation neural network (BPNN) and a general regression neural network (GRNN). A conventional multiple linear regression (MLR) model was also developed in order to compare the model performance and assess which model provides the best results. ANN and MLR models were developed and tested using the same annual data for 20 European countries. The ANN model demonstrated very good performance, significantly better than the MLR model. It was shown that a forecast of CH_4 emissions at the national level using the ANN model could be made successfully and accurately for a future period of up to two years, thereby opening the possibility to apply such a modeling technique, which could be used to support the implementation of sustainable development strategies and environmental management policies.

Keywords: national emission; general regression neural network; backpropagation neural network; multiple linear regression.

INTRODUCTION

Due to higher economic and industrial development, air pollution has become a serious problem in many countries and regions in recent years; therefore, it has become essential to track the emission of pollutants that have a detrimental effect on the environment and human health. As a part of a global effort to reduce emission of air pollutants, the Convention on Long Range Transboundary Air Pollution (LRTAP),[1] the United Nation Framework Convention on

* Corresponding author. E-mail: dantanasijevic@tmf.bg.ac.rs
# Serbian Chemical Society member.
doi: 10.2298/JSC020414110S

Climate Change (UNFCCC)[2] and their associated protocols, imply an obligation for all countries to submit reports on estimated current and future emissions of air pollutants. To achieve this commitment, it is necessary to implement the application of suitable methods and models for estimating and simulating a range of emission scenarios.

Methane is one of greenhouse gases that contribute to climate changes and it is emitted into the atmosphere from various sources: transport and natural gas exploitation, landfills, coalmines, wastewater treatment, rice fields, the petrochemical industry, composting, metallurgical industry, farms, *etc*.[3]

The main sources of emission data for air pollutants, including CH_4, are emission inventories, which currently makes them a key foundation for air quality modeling and analysis. The values of emissions in emission inventories are a compilation of a large number of input parameters (more than 300 emissions sources), in which the method of calculating the total emissions is dependent on the employed emission model. According to the methodology proposed in the EMEP/CORINAIR Emission Inventory Guidebook provided by the European Environmental Agency (EEA),[4,5] emissions are calculated by multiplying the emission factors (*EF*) with the activity rate (*A*), a statistical parameter for the respective source. In an emission inventory, the values of the parameters are determined as best estimates and therefore the input data to the inventories could be classified as uncertain, which in turn, means that the reported data could also be assumed uncertain.[6,7] Numerous studies have discussed uncertainties in emission inventories.[6–11]

There are different approaches to modeling emission estimates and estimating CH_4 emissions has also been the subject of many studies in which different modeling approaches were used.[5,12–17] In order to predict emission at the national level, all the aforementioned approaches require the availability of a large amount of country-specific information for calculating emission factors and activity data. Thus, an important limitation for the applicability of existing models could be a lack of specific input information necessary for the calculation of emission factors for a specific country. This lack of input information is especially typical when considering developing countries, which often have the most severe problem with pollution and emission control, and it could therefore result in particularly high uncertainties in the estimated emissions.

Artificial neural networks (ANNs) are adaptable systems that can determine relationships between different sets of data.[18] ANNs can be described as multivariate non-linear statistical methods where selected input variables are presented to the network and output variables (one or more) are derived. ANNs learn from examples and this particular modeling approach has proved to be a useful tool for many problems related to air quality and forecasting of emissions.[19–24] In recent years, due to their flexibility, considerable progress has been made in the appli-

cation of ANNs for predicting CH_4 for the chosen area. Some examples of studies include modeling and prediction of ventilation methane emissions of U.S. longwall mines,[18] prediction of methane productions in anaerobic treatment of molasses wastewater,[25] and prediction of methane emissions from wetland ecosystems.[26]

This paper presents the development and evaluation of an ANN model for the prediction of CH_4 emissions at national levels, using sustainability, economic and industrial indicators as inputs. Two different ANN architectures, back-propagation neural network (BPNN) and general regression neural network (GRNN), were evaluated and their results compared with each other, as well as with the results of a multiple linear regression (MLR) model.

The main difference between conventional inventory based models and the proposed ANN model is that the ANN model requires a substantially smaller number of input parameters, which have significantly broader availability. Due to this factor, the ANN model could be applied to the forecasting of CH_4 emissions when counties and regions do not have a complete set of input parameters available, which is usually essential for other models based on activity levels and emission factors.

MODEL DESCRIPTION AND TESTING

Input and output data

An important step in developing ANN models is the selection of the input variables that have the most significant impact on the output.[27] Economic activities largely contribute to the total emission of methane, most contributing sectors being industrial processes, agriculture, and production and energy consumption.[3]

Several studies have already shown that ANN models, based on economic and sustainability indicators, can achieve good accuracy in forecasting pollutant emissions at the national level, *e.g.*, for greenhouse gases and PM10 emissions.[28–30] The Gross Domestic Product (GDP) is a broadly available country-specific parameter that was chosen for an input parameter as a general indicator of the size of the economy. The GDP has already demonstrated very good applicability as an input parameter for ANN models.[29,30]

Municipal waste generation and its disposal onto or into land were also chosen as a model input, since methane is produced from landfills by the anaerobic biodegradation of wet waste.[31] Since rice paddies are characterized by a high moisture content and relatively high organic carbon levels, as well as prolonged anaerobic conditions during rice growth, they are one of the major anthropogenic sources of methane;[17,32] therefore, the land used for rice cultivation was also selected as an input.

Livestock is another important source of methane emissions, which originate mainly from enteric fermentation and manure management of ruminant production systems.[33] The contribution of livestock to methane emissions was factored by selecting the number of cattle as the input parameter. In addition, since methane is a component of natural gas, primary natural gas production is directly connected with the emissions of methane and was consequently included as an input parameter. Data for all of these input variables at the national level, for every country included in this study, were published on Eurostat.[34]

Annual methane emissions at the national level were taken as the output variable, and the data for every country included in this study was published on EDGAR (the Emissions Database for Global Atmospheric Research).[35] EDGAR is a joint project of the European Commission JRC (Joint Research Centre) and the Netherlands Environmental Assessment Agency (PBL) and it provides global, past and present day anthropogenic emissions of greenhouse gases and air pollutants by country and on a spatial grid.

The variables chosen for inputs and the related output of the models, along with their units after normalization and data sources, are presented in Table I. The selected input variables were normalized per capita and/or per average GDP value of the EU27, in order to enable comparison of countries of different sizes.

TABLE I. List of input and output variables

Input variable	Unit[a]	Source of data
Gross domestic product (*GDP*)	–[b]	Eurostat[34]
Waste deposit onto or into land (*WDL*)	kg pc[c]	
Municipal waste generation (*MWG*)	kg pc	
Land use – rice (*LR*)	ha pc	
Number of cattle (*NC*)	cattle pc	
Primary production of gas (*PPG*)	toe pc[d]	
Output variable		
CH_4 (emission)	kg pc	Edgar[35]

[a]After normalization; [b]unit less because of normalization per GDP value of EU27; [c]per capita; [d]tons of oil equivalents

The ANN CH_4 emission models were trained, validated and tested with available data for 20 European countries for the period ranging from the years 2000 to 2008, since data after 2008 was not available. Three types of datasets were used:

A training set. The group of data with which the ANN was trained. Data from the years 2000 to 2006 was used for training; an example of the training data set for the year 2006 is presented in Table S-I of the Supplementary material to this paper.

A validation dataset. The group of data provided to the ANN in the learning phase for the evaluation of modeling error, for effectively updating the best thresholds and weights,[36] and for preventing overtraining of the ANN. Approximately 15 % of data randomly selected from the training data set was used as the validation dataset.

A test dataset. A set of data newly presented to the ANN, thus used to evaluate the ANN's generalization capability. In this case, the data from years 2007 and 2008 was used as the test dataset.

ANN architecture

Two different ANN architectures were compared: the backpropagation neural network (BPNN), as the most frequently and widely applied ANN architecture, and the general regression neural network (GRNN), since it has already demonstrated good results in the modeling of emissions.[29,30]

Details of the two ANN architectures are described in Supplementary material to this paper.

Performance indicators

In order to compare the performance of the different models, in this study the following statistical parameters were used:

The root mean squared error (*RMSE*):

$$RMSE = \left[\overline{\left(C_P - C_O\right)^2}\right]^{1/2} \quad (1)$$

The mean absolute error (*MAE*):

$$MAE = \frac{1}{n}\sum\left|C_P - C_O\right| \quad (2)$$

The correlation coefficient (*R*):

$$R = \frac{\overline{(C_O - \overline{C_O})(C_P - \overline{C_P})}}{\sigma_{C_O}\sigma_{C_P}} \quad (3)$$

The index of agreement (*IA*):

$$IA = 1 - \frac{\overline{\left(C_P - C_O\right)^2}}{\left[\left|C_P - \overline{C_O}\right| + \left|C_O - \overline{C_O}\right|\right]^2} \quad (4)$$

The percentage of predictions within the factor of 1.25 from the observed values ($FA_{1.25}$):

$$0.8 < \frac{C_P}{C_O} < 1.25 \quad (5)$$

where C_P and C_O are the predicted and observed values, respectively; σ_O and σ_P are the standard deviation of observations and predictions, respectively; the over bars refer to the average of all values.

These statistical parameters were previously used for the evaluation of emission models and other modeling techniques.[18,29] *MAE* and *RMSE* measure the residual errors, thus providing a general understanding about the difference between the observed and modeled values. *R* provides a variability measure of the data reproduced in the model.[37] *IA* provides a correlation between the modeled and measured values and incorporates the error between those values. $FA_{1.25}$ represents the proportion of data for which the model results are in "proximity" with the measured values. The significance of these statistical metrics is not equal and in this study, *IA* and $FA_{1.25}$ were used as the key statistical parameters for evaluating the models.

RESULTS

In this study, two types of ANN architectures, BPNN and GRNN, were used to create models for the prediction of CH_4 emissions at the national level. Along with the development of the ANN models, for comparison purposes, a multiple linear regression (MLR) model was developed employing the same data as was used in the ANN models. The MLR model was created using SPSS 19 software[38] and the obtained MLR coefficients are presented in the following equation:

$$\begin{aligned} CH_4 \text{ emission} &= 0.0480 + 0.0001WDL - 0.0001MWG + \\ &+0.0922NC - 6465.1928LR + 0.0008PPG + 0.0131GDP \end{aligned} \quad (6)$$

where *WDL* is waste deposit onto or into land, *MWG* is municipal waste generation, *NC* is the number of cattle, *LR* is land use – rice, *PPG* is primary production of gas and *GDP* is the gross domestic product.

The values of the performance indicators of the developed models for the training and testing phase are given in Table II. A comparison of the actual and predicted values of CH_4 emission for the training data (years 2000–2006) is shown in Fig. 1.

TABLE II. Performance indicators values for the GRNN, BPNN and MLR models; pc – per capita

Model	Performance indicators for test set				
	IA	$FA_{1.25}$ / %	*MAE* / kg pc	*RMSE* / kg pc	*R*
Training data					
BPNN	1.00	94	3.4	5.0	0.97
GRNN	0.97	92	3.6	7.0	0.94
MLR	0.83	64	11.3	14.0	0.75
Test data					
BPNN	1.00	90	5.6	7.7	0.95
GRNN	0.98	93	4.8	7.3	0.96
MLR	0.90	58	11.9	252.5	0.83

The dataset for the years 2007–2008 is the test dataset, which is a set of "new" data, previously not presented to the models within the training phase. The test dataset, which is presented in Table S-II of the Supplementary material, was used to evaluate the capability of the models to make good predictions. The results of the BPNN, GRNN and MLR models applied on the test dataset are presented in Fig. 2 and the values of statistical performance indicators are presented in Table II. The actual CH_4 emissions, the GRNN predictions and the relative errors for the years 2007 and 2008 are presented in Table III.

DISCUSSION

Based on the results given in Table II and shown in Fig. 1 that were generated by the models in the training phase, the BPNN and GRNN led to much better results than the MLR. As can be seen, the results given by the ANN models could be regarded as very good.

The testing results showed that the use of the GRNN architecture led to slightly better results than the BPNN and considerably better results than the MLR (Fig. 2). The most important statistical performance indicators $FA_{1.25}$ and *IA* showed much better values for the test dataset in the case of the ANN models, *i.e.*, for the GRNN model, $FA_{1.25}$ was 93 % and for the BPNN, it was 90 %, while for the MLR, it was only 58 %. Moreover, for the GRNN model, *IA* was 0.98, for the BPNN, it was 1.00, while for the MLR, *IA* was 0.90. It could be concluded that the GRNN model was optimally trained since it demonstrated very good results

with similar values of $FA_{1.25}$ and the error performance indicators derived for the training and test dataset. In contrast, the BPNN model had similar but slightly worse results for the test dataset compared to the training data (Table II), which means that the BPNN model may benefit from an adjustment in the design in order to optimize the prediction results.

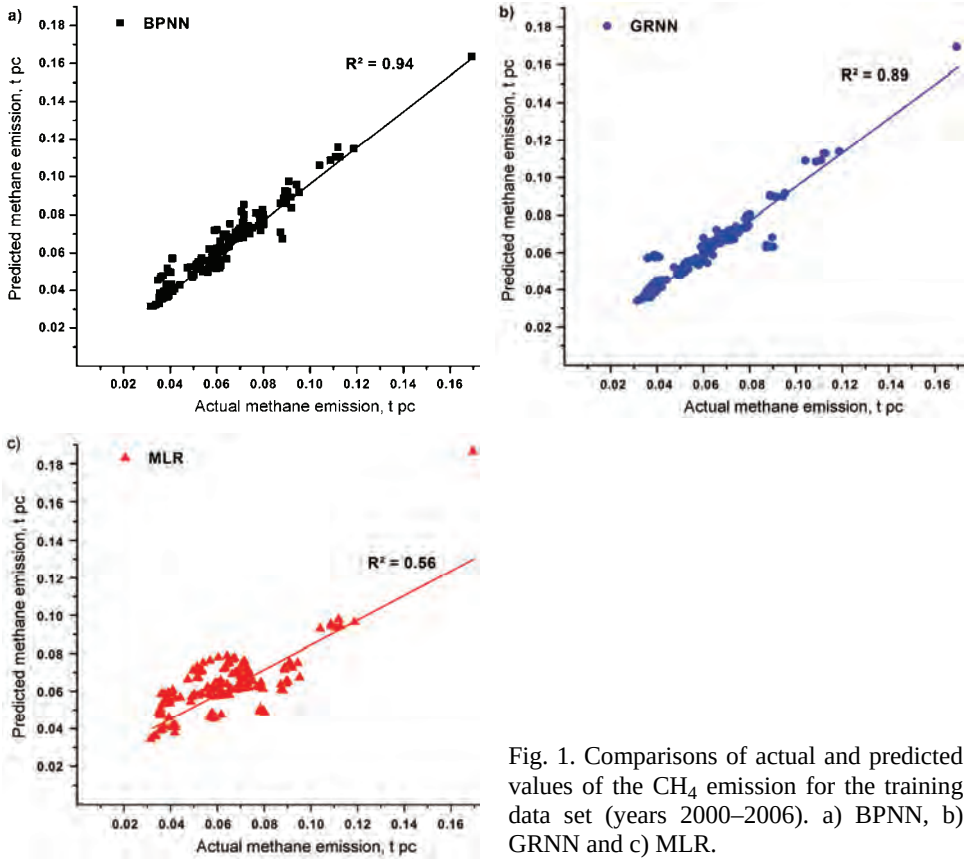


Fig. 1. Comparisons of actual and predicted values of the CH_4 emission for the training data set (years 2000–2006). a) BPNN, b) GRNN and c) MLR.

The GRNN model made 75 % of its predictions successfully, with the relative error of less than 10 %, another 10 % of its predictions had a relative error within the range of 10–20 %, while only 15 % of its predictions had a relative error higher than 20 %. In the case of Hungary, Bulgaria and The Netherlands, the predictions of the GRNN model had a relative error within the range of 10–20 %, while in the case of Slovakia, Poland and Greece, the relative error was higher than 20 % for both test years (Table III).

Uncertainties related to the input parameters used in the emission inventory approach (the activity statistics and emission factors) are considerable, ranging from 20–30 to up to 50 % in the case of Austria, which leads to the uncertainties

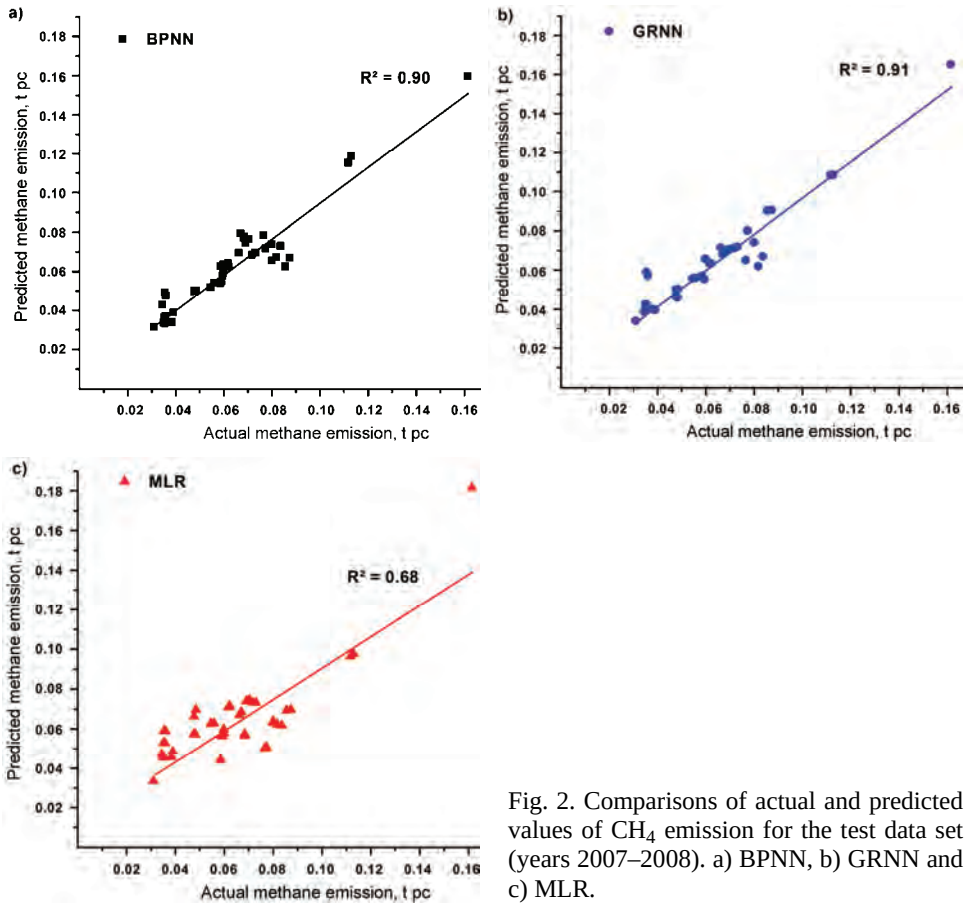


Fig. 2. Comparisons of actual and predicted values of CH_4 emission for the test data set (years 2007–2008). a) BPNN, b) GRNN and c) MLR.

TABLE III. Actual and CH_4 emission predicted by the GRNN model with relative errors (RE) for the years 2007 and 2008; pc – per capita

Country	Year 2007			Year 2008		
	Actual emission kg pc	*GRNN* kg pc	*RE* %	Actual emission kg pc	*GRNN* kg pc	*RE* %
Bulgaria	77.28	80.17	3.74	76.54	65.12	–14.92
Czech Republic	55.80	55.94	0.28	54.45	55.73	2.37
Denmark	68.33	69.30	1.42	68.39	69.32	1.36
Estonia	79.95	74.12	–7.29	79.97	74.22	–7.19
Greece	35.44	42.54	20.02	34.87	42.49	21.85
Spain	38.87	39.53	1.71	38.37	39.96	4.15
France	62.16	63.34	1.90	61.67	63.32	2.67
Latvia	66.19	71.49	8.01	66.93	68.16	1.84
Lithuania	71.69	71.22	–0.65	72.98	71.92	1.46
Luxembourg	111.64	108.51	–2.81	112.85	108.51	–3.84

TABLE III. Continued

Country	Year 2007			Year 2008		
	Actual emission kg pc	*GRNN* kg pc	*RE* %	Actual emission kg pc	*GRNN* kg pc	*RE* %
Hungary	35.73	39.35	10.13	35.12	40.18	14.39
Netherlands	59.80	65.75	9.95	59.77	65.87	10.21
Austria	48.20	45.97	–4.64	47.54	46.72	–1.72
Poland	83.57	67.05	–19.77	81.71	62.03	–24.09
Portugal	58.45	56.48	–3.39	58.44	57.12	–2.25
Romania	58.91	55.93	–5.06	59.35	55.34	–6.75
Slovenia	70.33	70.88	0.78	69.00	70.43	2.07
Slovakia	35.86	57.23	59.60	35.40	59.08	66.90
Finland	87.25	90.57	3.80	85.49	90.24	5.55
UK	48.39	50.26	3.86	47.61	50.12	5.27

in the estimated methane emissions of about 20–40 %.[9] Since the GRNN model uses significantly fewer input variables and provides predictions with the relative error lower than 20 % in case of most countries, the ANN approach could be considered as a viable alternative for the prediction of annual methane emissions on the national level, especially in cases when the emission inventory approach cannot be applied because of lack of data.

The obtained deviation between the actual data and predicted values for all countries for which the obtained relative error was higher than 10 % could be related to the quality of the input and output data used for the creation and evaluation of the model.

In the case of Slovakia, the data for waste deposit onto or into land (WDL) for the year 2001 was estimated. In addition, the data for municipal waste generation (MWG) was collected from the National Waste Catalogue until 2001 but from 2002 onwards from The European List of Waste.[34] The usage of estimated and non-standard training data could result in the large relative error in the year 2007 and 2008 (Table III). The relative error for The Netherlands for the year 2008 was 10.21 %, and this model error could be related to the ban of the direct disposal of mixed municipal waste in 2003, which led to a decrease in *WDL* from 50 kg pc in 2002 to only 11 kg pc in 2004.[39]

In the case of Greece, the training data of the input parameter *GDP* for the year 2005 was uncertain because of the inaccurate reporting of the country's economic performance by the government. The influence of this parameter was probably the reason for higher attained values of the relative error for both test years.

The *MWG* data for Poland and Hungary were estimated, for the year 2003 in the case of Hungary and from 2005 to 2008 in the case of Poland, which could explain the higher relative error values obtained (14.39 % for Hungary and –24.09 % for Poland).

In the case of Bulgaria, the GRNN model showed a relative error of –14.92 % for 2008. In fact, the landfill site in Sofia was closed between October 2005 and December 2007, which led to a temporary storage of 2.5 % of the collected municipal waste in 2005 and 10 % in 2006 and 2007. These amounts were not included in the figures for municipal waste generated and also caused a drop in the municipal waste landfilled in 2006 and 2007.[34] Considering all these factors, changes in MWG for the training years 2005 and 2006 could be a reasonable explanation for the higher relative error for the year 2008.

CONCLUSIONS

This study shows the development of a model for predicting emissions of CH_4 at the national level, using artificial neural networks (ANNs) and widely available sustainability, economic and industrial indicators as their inputs. Sustainability and the economic parameters considered to contribute the most to CH_4 emissions were used as input parameters for the models: gross domestic product (GDP) as a measure of industrial activity, waste generation and disposal, surface of land under fields of rice, number of cattle and the primary production of gas. Two ANN architectures were used, a back propagation neural network (BPNN) and a general regression neural network (GRNN), and both models were developed and tested using annual data for European countries for the period from the year 2000 to 2008, and subsequently compared with a corresponding multiple linear regression (MLR) model.

The performance evaluation of the models was realized using multiple statistical performance indicators. Based on the results of the evaluation, it could be concluded that the ANN models provide good and reliable predictions of CH_4 emissions, which were significantly better than the conventional MLR model. It was shown that the GRNN model demonstrates similar, but somewhat better predictions in comparison to the BPNN model.

The key advantage of the ANN approach in comparison with conventional emission inventory-based models is that the ANN models use a drastically smaller number of input parameters. In addition, emission inventory-based models use strictly defined input parameters the determination of which requires substantial field studies, while the proposed ANN approach uses widely available sustainability, economic and industrial parameters as its inputs. Furthermore, the ANN models allow simulation of different CH_4 emission scenarios by changing the values of input variables, for example to gain foresight because of proposed regulatory emission reduction actions. A developed ANN model could be applied for the prediction of emissions of CH_4, which could be very helpful in the implementation of sustainable development strategies and environmental management policies.

Further research is planned to expand the model to include other environmental quality indicators, such as emission of ammonia, ozone precursors, and in the application of new techniques for the optimization of the inputs, such as principal component analysis, correlation analysis and genetic algorithms.

SUPPLEMENTARY MATERIAL

Details of the two ANN architectures, s well as training and test datasets, Tables S-I and S-II, are available electronically from http://www.shd.org.rs/JSCS/, or from the corresponding author on request.

Acknowledgement. The authors are grateful to the Ministry of Education, Science and Technological Development of the Republic of Serbia, Project No. 172007, for financial support.

И З В О Д

МОДЕЛОВАЊЕ ЕМИСИЈЕ МЕТАНА ПРИМЕНОМ ВЕШТАЧКИХ НЕУРОНСКИХ МРЕЖА

LIDIJA J. STAMENKOVIĆ[1], DAVOR Z. ANTANASIJEVIĆ[2], MIRJANA Đ. RISTIĆ[1], ALEKSANDRA A. PERIĆ-GRUJIĆ[1] и VIKTOR V. POCAJT[1]

[1]*Универзитет у Београду, Технолошко–металуршки факултет, Карнегијева 4, 11120 Београд и* [2]*Универзитет у Београду, Иновациони центар Технолошко–металуршког факултета, Карнегијева 4, 11120 Београд*

У овом раду приказан је развој модела, заснованог на вештачким неуронским мрежама (*Artificial Neural Network* – ANN), за предвиђање вредности емисије метана на националном нивоу, при чему су као улазне променљиве коришћени економски и индустријски индикатори и индикатори одрживог развоја. Тестиране су две различите ANN архитектуре: неуронска мрежа са пропагацијом грешке уназад (*Backpropagation Neural Network* – BPNN) и неуронска мрежа са општом регресијом (*General Regression Neural Network* – GRNN). Развијен је и линеарни модел заснован на вишеструкој линеарној регресији (*Multiple Linear Regression* – MLR) са којим су упоређене перформансе наведених ANN модела. ANN и MLR модели су развијени и тестирани коришћењем података за 20 земаља Европске уније за период од 2000. до 2008. године. Поређење предвиђања ANN модела са актуелним вредностима CH_4 емисије је показало да се овом методологијом емисије метана на националном нивоу могу успешно и прецизно одредити до две године унапред, при чему је ANN модел знатно прецизнији у поређењу са MLR моделом. Наведени ANN модел се може користити као подршка при доношењу националних стратегија одрживог развоја.

(Примљено 2. априла, ревидирано 26. августа, прихваћено 7. новембра 2014)

REFERENCES

1. United Nations Economic Commission for Europe (UNECE), *Protocol to Abate Acidification, Eutrophication and Ground-level Ozone*, UNECE, Gothenburg, 1999
2. United Nations, *United Nations Framework Convention on Climate Change*, in *Proceedings of the Convention on Climate Change on the Work of the Second Part of its Fifth Session*, UN, New York, 1992
3. U.S. Environmental protection agency Website. http://www.epa.gov/methane/ /sources.html. (Accessed 15 April 2013)

4. European Environment Agency (EEA), *Joint EMEP/CORINAIR Atmospheric Emission Inventory Guidebook*, 3rd ed., European Environment Agency, Copenhagen, 2007
5. J. Barrancos, S. Briz, D. Nolasco, G. Melián, G. Padilla, E. Padrón, I. Fernández, N. Pérez, P. A. Hernández, *Atmos. Environ*. **74** (2013) 10
6. W. Winiwarter, K. Rypdal, *Atmos. Environ*. **35** (2001) 5425
7. K. Rypdal, *Environ. Sci. Policy* **5** (2002) 233
8. S. J. Lindley, D. E. Conlan, D. W. Raper, A. F. R. Watson, *Atmos. Environ*. **34** (2000) 375
9. K. Rypdal, W. Winiwarter, *Environ. Sci. Pol.* **4** (2001) 107
10. W. Schöpp, Z. Klimont, R. Suutari, J. Cofala, *Environ. Sci. Policy* **8** (2005) 601
11. M. El-Fadel, M. Zeinati, N. Ghaddar, T. Mezher, *Energy Policy* **29** (2001) 1031
12. S. Kumar, S. A. Gaikwad, A. V. Shekdar, P. S. Kshirsagar, R. N. Singh, *Atmos. Environ*. **38** (2004) 3481
13. T. Li, Y. Huang, W. Zhang, C. Song, *Ecol. Modell.* **221** (2010) 666
14. M. Amann, I. Bertok, J. Borken-Kleefeld, J. Cofala, C. Heyes, L. Höglund-Isaksson, Z. Klimont, B. Nguyen, M. Posch, P. Rafaj, R. Sandler, W. Schöpp, F. Wagner, W. Winiwarter, *Environ. Modell. Software* **26** (2011) 1489
15. L. Höglund-Isaksson, *Atmos. Chem. Phys.* **12** (2012) 9079
16. B. Khoshnevisan, S. Rafiee, M. Omid, M. Yousefi, M. Movahedi, *Energy* **52** (2013) 333
17. K. Chenga, S. M. Ogleb, W. J. Partonb, G. Pana, *Ecol. Modell.* **261–262** (2013) 19
18. C. Ö. Karacan, *Int. J. Coal. Geol.* **73** (2008) 371
19. S. A. Abdul-Wahab, S. M. Al-Alawi, *Environ. Modell. Software* **17** (2002) 219
20. G. Corani, *Ecol. Modell.* **185** (2005) 513
21. U. Brunelli, V. Piazza, L. Pignato, F. Sorbello, S. Vitabile, *Atmos. Environ*. **41** (2007) 2967
22. U. Brunelli, V. Piazza, L. Pignato, F. Sorbello, S. Vitabile, *Build. Environ.* **43** (2008) 304
23. K. P. Singh, S. Gupta, A. Kumar, S. P. Shukla, *Sci. Total Environ.* **426** (2012) 244
24. A. Russo, F. Raischel, P. G. Lind, *Atmos. Environ*. **79** (2013) 822
25. K. Yetilmezsoy, F. I. Turkdogan, I. Temizel, A. Gunay, *Int. J. Green Energy* **10** (2013) 885
26. X. Zhu, Z. Q. Zhuang, M. Glagolev, L. Song, *Global. Biogeochem. Cycl.* **27** (2013) 592
27. M. Arhami, N. Kamali, M. M. Rajabi, *Environ. Sci. Pollut. Res.* **20** (2013) 4777
28. D. M. Radojević, V. V. Pocajt, I. G. Popović, A. A. Perić-Grujić, M. Dj. Ristić, *Energ. Source, A* **35** (2013) 733
29. D. Z. Antanasijević, V. V. Pocajt, D. S. Povrenović, M. Dj. Ristić, A. A. Perić-Grujić, *Sci. Total. Environ.* **443** (2013) 511
30. D. Antanasijević, M. Ristić, A. Perić-Grujić, V. Pocajt, *Int. J. Greenh. Gas. Con.* **20** (2014) 244
31. R. Couth, C. Trois, S. Vaughan-Jones, *Int. J. Greenh. Gas. Con.* **5** (2011) 1443
32. J. Schimel, *Nature* **403** (2000) 375
33. A. A. Aljaloud, T. Yan, A. M. Abdukader, *Anim. Feed Sci. Technol.* **166**–**167** (2011) 619
34. Eurostat Website. http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/ /search_database (Accessed 15 April 2013)
35. Edgar Website, http://edgar.jrc.ec.europa.eu/overview.php?v=42. (Accessed 15 April 2013)
36. P. Viotti, G. Liuti, P. Di Genova, *Ecol. Modell.* **148** (2002) 27
37. S. I. V. Sousa, F. G. Martins, M. C. M. Alvim-Ferraz, M. C. Pereira, *Environ. Modell. Software* **22** (2007) 97

38. *IBM SPSS Statistics for Windows*, Version 19, IBM Corp., Armonk, NY, 2010
39. *Municipal waste statistics*, Eurostat Website, http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Municipal_waste_statistics (Accessed 20 January 2014).

Journal of the Serbian Chemical Society
JSCS@tmf.bg.ac.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) S109–S112 (2015) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Modeling of methane emissions using the artificial neural network approach

LIDIJA J. STAMENKOVIĆ[1#], DAVOR Z. ANTANASIJEVIĆ[2*#], MIRJANA Đ. RISTIĆ[1#], ALEKSANDRA A. PERIĆ-GRUJIĆ[1#] and VIKTOR V. POCAJT[1#]

[1]University of Belgrade, Faculty of Technology and Metallurgy, Karnegijeva 4, 11120 Belgrade, Serbia and [2]University of Belgrade, Innovation Center of the Faculty of Technology and Metallurgy, Karnegijeva 4, 11120 Belgrade, Serbia

J. Serb. Chem. Soc. 80 (3) (2015) 421–433

DETAILS OF ANN ARCHITECTURES

BPNN is a three-layer feed-forward network trained from input data using an error backpropagation algorithm. This is a supervised network, *i.e.*, trained with both inputs and outputs. A three-layer BPNN network with standard connections is suitable for almost all problems of this nature.[1] The learning process of a BP network consists of two iterative steps: forward computing of the data stream and backward propagation of error signals. During the forward computing, the original data are transmitted from the input layer to the output layer through the hidden processing layer and the neurons of each layer can only affect the neurons of the next layer. In this process, the synaptic weights are all adjusted in accordance with the error signals. With these two steps being performed iteratively, the error, *i.e.*, the difference between the network output and desired output, can be minimized using the delta rule.[2]

The number of neurons in the input and output layer is equal to the number of input and output parameters that are used in the BPNN model, while the number of neurons in the hidden layer is often computed as the sum of half of the inputs plus the output neurons and the square root of the number of training patterns. The number of neurons in the hidden layer in the employed BPNN model was 15, the logistic activation function was used, and the learning rate and momentum parameters were set to a value of 0.1. The training time was 30 min during which 73952 learning epochs were performed. The schematic representation of the BPNN model is presented in Fig. 1a.

A GRNN is also a feed forward neural network with supervised training based on the non-linear regression theory. It consists of an input layer, a hidden layer that can be divided into a pattern layer and a summation layer and an output layer. GRNNs work by measuring how far a given sample pattern is from the patterns in the training set in *N*-dimensional space, where *N* is the number of inputs in the problem. The training of the GRNN is quite different from the training used in BP neural networks, being completed after presentation of each input–output vector pair from the training data set to the GRNN input layer only once. The resulting parameter of the GRNN training is a smoothing factor, which determines how tightly the

* Corresponding author. E-mail: dantanasijevic@tmf.bg.ac.rs

network matches its predictions to the data in the training patterns.[1] The number of neurons in the input and output layer is equal to the number of input and output parameters that were used in the ANN model. The number of neurons in the pattern sub-layer is equal to the number of training cases used for model training, *i.e.*, 160 neurons. The input data was scaled using a linear function in range of 0 to 1, and the smoothing factor was determined using a genetic algorithm.[3] The schematic representation of the GRNN is shown in Fig. 1b.

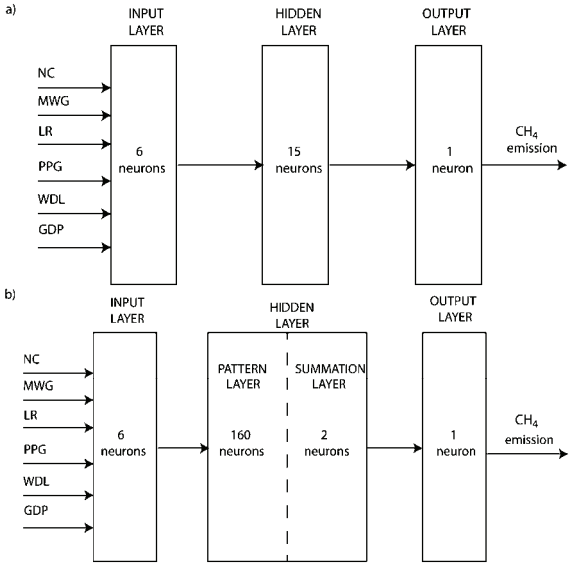


Fig. S-1. Schemas of the used ANN architectures: a) BPNN and b) GRNN.

For neural network design and training, the software tool NeuroShell 2 was used.[4]

TRAINING AND TEST DATA SETS

TABLE S-I. Training dataset for the year 2006

Country	*GDP*	*WDL* kg pc	*MWG* kg pc	*NC* cattle pc	*LR*×10^{-7} ha pc	*PPG* toe pc	CH_4 emission kg pc
Bulgaria	0.14	357	461	0.082	5.83	0.048	78.05
Czech Republic	0.49	199	296	0.136	0	0.014	55.77
Denmark	1.70	37	740	0.291	0	1.718	68.83
Estonia	0.42	278	399	0.182	0	0.000	78.68
Greece	0.79	385	442	0.061	20.8	0.002	35.41
Spain	0.95	355	594	0.141	24.3	0.001	39.25
France	1.20	194	536	0.299	27.4	0.017	62.68
Latvia	0.30	293	412	0.164	0	0.000	65.68
Lithuania	0.30	357	391	0.246	0	0.000	71.30
Luxembourg	3.03	130	683	0.397	0	0.000	108.65
Hungary	0.38	376	468	0.070	2.38	0.236	36.81
Netherlands	1.40	15	622	0.225	0	3.391	60.85

TABLE S-I. Continued

Country	*GDP*	*WDL* kg pc	*MWG* kg pc	*NC* cattle pc	*LR*×10^{-7} ha pc	*PPG* toe pc	CH_4 emission kg pc
Austria	1.32	25	653	0.243	0	0.189	48.64
Poland	0.30	236	321	0.138	0	0.102	87.27
Portugal	0.64	297	463	0.133	24.0	0.000	58.38
Romania	0.19	292	389	0.136	2.59	0.442	58.41
Slovenia	0.65	361	431	0.227	0	0.001	70.65
Slovakia	0.35	234	301	0.094	0	0.033	36.12
Finland	1.33	286	494	0.177	0	0.000	88.82
UK	1.36	352	586	0.171	0	1.192	49.67

TABLE S-II. Test data set

Country	*GDP*	*WDL* kg pc	*MWG* kg pc	*NC* cattle pc	*LR*×10^{-7} ha pc	*PPG* toe pc
Bulgaria	0.16	389	433	0.080	8.59	0.031
Czech Republic	0.51	205	293	0.133	0	0.016
Denmark	1.67	37	790	0.284	0	1.518
Estonia	0.48	291	449	0.179	0	0.000
Greece	0.80	357	447	0.061	23	0.002
Spain	0.94	347	583	0.148	21.6	0.000
France	1.18	194	543	0.300	2.59	0.014
Latvia	0.37	323	378	0.175	0	0.000
Lithuania	0.34	369	401	0.233	0	0.000
Luxembourg	3.12	126	695	0.406	0	0.000
Hungary	0.40	341	457	0.070	2.58	0.199
Netherlands	1.40	13	629	0.234	0	3.328
Austria	1.32	20	596	0.241	0	0.192
Poland	0.33	239	322	0.142	0	0.102
Portugal	0.64	299	468	0.136	25.4	0.000
Romania	0.23	284	379	0.131	3.9	0.428
Slovenia	0.68	341	439	0.239	0	0.001
Slovakia	0.41	240	309	0.093	0	0.020
Finland	1.36	267	506	0.171	0	0.000
UK	1.35	323	570	0.166	0	1.068
Bulgaria	0.18	441	474	0.075	10.2	0.020
Czech Rep.	0.59	197	305	0.131	0	0.016
Denmark	1.71	32	830	0.292	0	1.646
Estonia	0.49	248	391	0.177	0	0.000
Greece	0.83	372	452	0.061	24.1	0.001
Spain	0.96	287	556	0.133	21.1	0.000
France	1.20	171	541	0.303	2.56	0.013
Latvia	0.40	311	332	0.167	0	0.000
Lithuania	0.39	368	408	0.229	0	0.000
Luxembourg	3.23	123	697	0.406	0	0.000
Hungary	0.42	333	454	0.070	2.49	0.200
Netherlands	1.45	8	624	0.244	0	3.651

TABLE S-II. Continued

Country	*GDP*	*WDL* kg pc	*MWG* kg pc	*NC* cattle pc	*LR*×10^{-7} ha pc	*PPG* toe pc
Austria	1.36	19	599	0.240	0	0.158
Poland	0.38	228	320	0.146	0	0.097
Portugal	0.65	332	515	0.136	24.8	0.000
Romania	0.26	301	392	0.125	4.6	0.418
Slovenia	0.74	339	457	0.234	0	0.001
Slovakia	0.48	250	328	0.090	0	0.016
Finland	1.40	265	521	0.171	0	0.000
UK	1.17	287	544	0.162	0	1.025

REFERENCES

1. S. A. Kalogirou, *Prog. Energ. Combust.* **29** (2003) 515
2. M. Cai, Y. Yin, M. Xie, *Transport Res. D-TrE* **14** (2009) 32.
3. *Neuroshell 2 Help*, *GRNN Training Criteria*, Ward Systems Group Inc., http://www.wardsystems.com/manuals/neuroshell2/index.html?idxhowuse.htm (Accessed 20 August 2014)
4. *Neuroshell 2*, Ward Systems Group Inc., Frederick, MD, 1993.

Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS
J. Serb. Chem. Soc. 80 (3) 435–452 (2015)
JSCS–4727
UDC 371.124:54+011.892+373.5
Original scientific paper

Chemistry curricular knowledge of secondary school teachers

BILJANA TOMAŠEVIĆ*# and DRAGICA TRIVIĆ#

Faculty of Chemistry, University of Belgrade, P. O. Box 158, 11001 Belgrade, Serbia

(Received 2 October, revised 7 December 2014, accepted 8 December 2014)

Abstract: In the course of this research, the extent of chemistry teachers' professional knowledge related to the structure, contents and application of chemistry curricula and their components was investigated. The research comprised 119 teachers from 69 secondary schools (25 grammar schools and 44 vocational secondary schools). The questions in the questionnaire referred to general curriculum knowledge, knowledge of chemistry curriculum and the views/assessments of the teachers concerning the necessary changes in the curricula currently in effect. The teachers' answers showed that the most important components of the curriculum for their work are the goals and operative tasks/outcomes. The results indicated that information in the curriculum components exits that remains unused although it is relevant for a certain level of planning. Among the teachers in the sample, higher percentages of those with an appropriate teachers' training programme applied information from the curriculum within the teaching process through demonstration methods and problem solving. The research that was conducted provides a basis for defining the indicators for monitoring the level of teachers' capability to apply curricular knowledge in their practice. Such indicators are important for creating teaching situations and teachers' activities within the framework of initial education and continuing professional development of teachers.

Keywords: curricular knowledge; curricular components; annual work plan; monthly work plan; lesson plan; chemistry teaching.

INTRODUCTION

The needs of contemporary society require a constant redefining of the education of the young. In keeping with this, what also undergoes changes are requirements concerning teachers' competences (knowledge and understanding, skills, dispositions), which would enable an efficient realisation of the expected outcomes of the education of the young.[1] The said competences should be formed through initial teacher education (ITE) and further developed through

* Corresponding author. E-mail: bsteljic@chem.bg.ac.rs
# Serbian Chemical Society member.
doi: 10.2298/JSC1401002121T

continuing professional development (CPD). They represent a complex combination of knowledge, skills and value judgements.[2] Competences comprise teachers' knowledge about learning, the goals of education, outcomes, resources and the broader social context of teaching and education.[1,3]

Pedagogical knowledge of chemistry is part of chemistry teachers' competences that could be defined as a combination of pedagogy and the content of chemistry.[4–6] Today, the professional development of teachers also requires a technological content to pedagogical knowledge, which is necessary for the application of contemporary technologies in the teaching process.[7,8] Knowledge of science curricula, together with knowledge of students' understanding of science, knowledge of assessment of scientific literacy, knowledge of instructional strategies, shapes a teacher's orientation to science teaching.[9] This encompasses knowledge of two categories: the mandated objectives and tasks, and specific curriculum-related programmes and materials. The first category encompasses knowledge of the objectives and tasks that pupils are to accomplish within the framework of a subject, as well as knowledge of the manner of accomplishing them within the framework of the given topics during the course of the school year. This knowledge also includes a teacher's knowledge about vertical curricula (the curricula of the previous and the following school year). Awareness of horizontal connections among different subjects is also required, as is knowledge of the curricula and attendant materials of specific additional programmes in special areas and scientific topics relevant to teaching. This expresses the need for teachers to continually become informed about development projects, the results of which would improve the teaching practice, and about reforms that are periodically enacted. The extent to which teachers will successfully and appropriately introduce the most important results into their teaching practice depends on this.

Knowledge of the curricula and being trained to interpret them are considered to be important components of a teacher's knowledge. There are different models of the structure of a teacher's knowledge and they all emphasise the necessity of educating teachers in the sphere of the curriculum.[10–13]

In their model, Barnett and Hodson[14] included the following components of the knowledge of science teachers: academic and research knowledge, pedagogical content knowledge, professional knowledge and classroom knowledge. Knowledge of curricular documents and curriculum planning were viewed as components of professional knowledge. Professional knowledge comprises knowledge of curriculum documents, the duties of teachers, union matters, information about school administration and procedures for communicating with parents.

The curricular knowledge refers to the obligation of a teacher to be acquainted with the existence and functions of this document. Curricular knowledge serves to organise, present and adjust the contents of the curriculum, teaching topics, problems and issues related to the varying interests and abilities of the

pupils. Curricula are continually being changed, and a teacher cannot get all the necessary instructions for his/her future work through education. Teachers' education does not represent the acquisition of a certain number of ready-made solutions, but constitutes the development of a complex body of knowledge that is to be applied in resolving specific practical problems in accordance with the curricula currently in effect.[15] Researchers agree that subject matter plays the key role in a meaningful linking of the components of a teacher's knowledge.[16,17]

The results of research of the formation of curricular knowledge in pre-service and in-service teachers do not shed light on the ability of teachers to translate specific curricular material into practice. Research for the most part encompasses the most frequent teaching topics in the curricula and the necessary corpus of knowledge, explaining the most frequently encountered teachers' and students' fallacies (misconceptions) that make certain contents difficult to understand and clarify. Still, each teacher who possesses a more or less formed knowledge of the curriculum teaches according to the curriculum which he/she should adjust, transform and realise in the classroom in accordance with a great number of contextual factors. The forming of curricular knowledge may depend on a great number of factors, the initial education of teachers, the type of school where they work and the length of their professional experience.[18]

A survey that should show how teachers use the chemistry curricula for the planning and realisation of their teaching practice was conducted. Components of the Carlsen model of teacher knowledge, general curriculum knowledge and knowledge of chemistry curriculum (specific knowledge of science curriculum) were monitored.[13] General curriculum knowledge presupposed knowledge of the role of certain components and the curriculum as a whole. Knowledge of the chemistry curriculum presupposed monitoring how teachers translate the chemistry curriculum into classroom activities.

General curricular knowledge[4] and the knowledge of chemistry curricula[9] presuppose a teacher's knowledge of the objectives and tasks that students are supposed to accomplish within the framework of a subject. Using the curriculum, a teacher should know and interpret this document, both from the viewpoint of general pedagogical knowledge and from the viewpoint of knowledge of chemistry contents.

The purpose of the conducted survey was to investigate the curricular knowledge of chemistry teachers in grammar schools and secondary vocational schools. Monitored were indicators referring to:

1. general curriculum knowledge (knowledge of the purpose of the curriculum),

2. knowledge of chemistry curriculum (knowledge and application of information from the chemistry curriculum) and

3. views about the necessary changes in the current curricula.

The analysis of the data obtained from the questionnaire was supposed to establish:

– how teachers asses the usefulness of the curriculum and some of its components for various phases of planning (Q6, Q7, Q8, Q9),

–- how they assess the usefulness of the curriculum for the realisation of the most important and most frequent teaching situations, and which components of the curriculum they single out as the most important ones (Q10, Q11, Q12) and

– which curriculum components, as they see it, require changes (Q13).

Moreover, whether their replies were influenced by their initial education of teachers, the type of school where they work and the length of their professional experience influence were investigated.

DATA COLLECTION

The research sample

The questionnaire constructed for the purpose of this survey was completed by 119 teachers working in grammar schools and secondary vocational schools. The survey was conducted in 2013. Of the participating teachers, 41 work in grammar schools, while 78 work in secondary vocational schools. These schools were located in 41 cities and towns in Serbia**.** Since no reliable data concerning the total number of chemistry teachers in Serbia could be obtained, the response rate was calculated based on the number of schools in the sample.[19] According to official data,[20] there are 494 secondary schools in Serbia. The teachers in the sample came from 25 grammar schools and 44 secondary vocational schools (14.0 % of the total number of secondary schools in Serbia).

The instrument and the procedures

The questionnaire, given in the Supplementary material to this paper, completed by the teachers consisted of 13 questions (quoted along with the survey results). The first five questions referred to personal information about the respondents.

The questionnaire comprised closed-type questions and one open-type question. The closed-type questions were of the two-option response variety, multiple-choice questions and those with a Likert scale. Three closed-type questions were to be answered by ticking a box in the table.

The questions were formulated based on an analysis conducted beforehand concerning the structural components contained in the current curricula. The questionnaire required the teachers' assessment concerning the applicability of the current curricula, the importance and usefulness of the information mediated by the curricula in their entirety, and also through their individual components. Researched were the way teachers translate information from the curriculum into corresponding classroom activities and how they select the appropriate contents and methods.

The first version of the questionnaire was given to four expert chemistry teachers. Based on their comments and suggestions, a new version was prepared. This questionnaire was presented in a seminar in which twenty chemistry teachers participated. Based on their responses, the clarity of the formulations of the items was improved, as well as the order of the items in the questionnaire. Subsequently, the final version of the present questionnaire was constructed.

Data analysis

The response frequency, arithmetical mean and standard deviation were determined. The statistical significance of the differences between the answers given by the teachers from different schools, with different initial education and working experience, were determined by means of the χ^2-test of independence as a measure. When dealing with two variable categories (*e.g.*, grammar schools and vocational schools), the value of the Yates Correction for Continuity was monitored. In the cases where more than 20 % of the frequencies were lower than 5, the values were determined through the Exact Statistics and Monte Carlo methods. For the interpretation of the χ^2-test results, the 95 % confidence level was selected. We used IBM SPSS Statistics (Version 20.0) was used for the for the χ^2-tests. When no statistically significant differences between the answers given by teachers from various categories were established, the results in this study were presented for the entire sample.

RESULTS AND DISCUSSION

Respondents

The structure of the sample of the teachers surveyed is presented in Table I. Of the overall number of the teachers surveyed, 41 (34.5 %) work in grammar schools (GS), while 78 (65.5 %) work in secondary vocational schools (SVS). The majority of the teachers (42.9 %) had between 10 and 20 years of working experience. Only 9.2 % of the teachers were prepared for working in a school through initial education, whereas the others completed one of the non-teaching courses of studies at the Faculty of Chemistry or some related faculty, *e.g.*, the Faculty of Technology and Metallurgy. Nine of the teachers (7.6 %) had obtained some form of post-graduate education (specialisation, Master's Degree). The sample comprised 84.0 % of women. A little less than half of the teachers included in the sample belonged to the age group between 40 and 50 (44.5 %). The second largest group in the sample were teachers between 50 and 60 years of age (27.7 %).

TABLE I. General characteristics of the respondents (N = 119)

Type of school	N (%)	Years of work as teachers	N (%)	Initial education	N (%)
Grammar School (GS)	41 (34.5)	Less than 5	11 (9.3)	Faculty of Chemistry, Teacher-training programme (FC-TTP)	11 (9.2)
Secondary Vocational School (SVS)	78 (65.5)	5–10	13 (10.9)	Faculty of Chemistry, Non-teacher--training programme (FC-NTTP)	84 (70.6)
		10–20	51 (42.9)	Other non-teacher-training faculties (ONTTP)	24 (20.2)
		20–30	28 (23.5)		
		More than 30	16 (13.4)		

General curriculum knowledge and knowledge of the chemistry curriculum

Curricular knowledge presupposes knowledge of the purpose and the manner of implementing this document as a basis for the overall planning and realisation of teaching.[9] When asked for what level of planning they use the curricula, the majority of the teachers replied that they use them for annual planning (65.5 %), and for monthly planning (43.5 %). Viewing the sample as a whole, only a few teachers use the curricula when planning individual lessons. However, a statistically significant difference was established when the answers of teachers with different initial education were compared ($\chi^2(2.119) = 15.84$, $p = 0.00$). Only 18.2 % of the teachers who were prepared for teaching through their initial education (FC-TTP) actually use the curriculum for preparing a lesson plan. A greater percentage (34.2 %) of the teachers who completed non-teaching courses of studies at the Faculty of Chemistry (FC-NTTP) use the curriculum for preparing individual lessons. Of the teachers who completed non-teaching courses of studies at other faculties (ONTTP), 77.3 % use the curricula for preparing individual lessons.

The teachers who attended the teaching and non-teaching programmes of studies at the Faculty of Chemistry mostly work in grammar school (87.2 %). There were no statistically significant differences concerning their answers to questions about their use of the curricula for various levels of planning. As regards the answers of the teachers working in SVS, depending on their initial education (teaching and non-teaching programmes), there were statistically significant differences in their use of the curricula for preparing annual plans ($\chi^2(2.73) = 6.64$, $p = 0.04$) and planning individual lessons ($\chi^2(2.73) = 0.45$, $p = 0.00$). When making annual plans, the curricula are mostly used by those teachers who were trained for teaching (100 %), followed by those who attended a non-teaching course of studies at the Faculty of Chemistry (68.9 %) and the other teachers (50 %). As regards the planning of individual lessons, the curriculum is used the least by those teachers who attended a teacher training programme at the Faculty of Chemistry (12.5 %), followed by those who completed a non-teaching course of study at the same faculty (35.6 %), and 80.0 % of the teachers who graduated from other faculties.

The results obtained indicate that the curriculum, as a document prescribing the obligatory contents (teaching topics) and the number of lessons for their realisation, is used the most for macro planning.

When asked which components, *i.e.*, which data are useful to teachers in their work, more than half singled out the operative tasks/outcomes (Fig. 1). This was followed, in terms of the frequency of the answers, by the goals and tasks of chemistry (46.2 %). One-third of the teachers encompassed by the sample found the contents of topics a useful component of the curriculum. The curriculum component designated as instructions for the realisation of a topic was singled

out as informative by a small number of the teachers, who found the component the manner of realising the curriculum more useful. The low frequency of selecting these two components was probably connected with the fact that, within the current curricula, neither of them provides specific but generalised instructions. The fewest teachers (6.7 %) singled out the curriculum component pertaining to additional work with students as useful. In other words, apart from the obligatory segments of teaching, teachers find very little information in the curricula of use for additional work with those students who, in keeping with their interests and knowledge, should be offered additional contents.

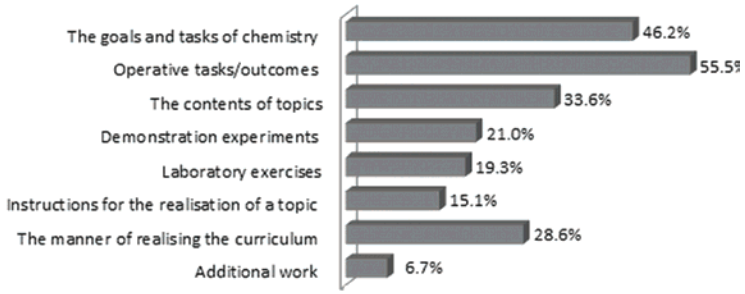


Fig. 1. The importance of information mediated through curriculum components (N = 119).

Most of the teachers surveyed (58.0 %) have used the curriculum in an equal measure during the course of their teaching career even though it remains unchanged. Compared to the initial years of their work, 11.8 % of the teachers feel a greater need to use the curriculum, whereas 26.1 % of the teachers tend to use it increasingly less over time. This also refers to the curricula that have remained unchanged for years.

The ninth question in the curriculum linked the levels of planning (annual, monthly, individual lessons) and curriculum components. Among the curriculum components used by the teachers for particular levels of planning, most of them (64.7 %) singled out the goals and tasks of chemistry for annual planning, whereas 58.8 % of them singled out the component designated as operative tasks/ /outcomes for the monthly planning of teaching and for planning individual lessons (Fig. 2). The component instructions for the realisation of a topic is used by the least for annual planning, and is useful to a greater number of the teachers (46.2 %) for planning a lesson. Similarly, the component list of demonstration experiments is of greater importance to a larger number of the teachers for planning a lesson than for monthly and annual planning. In addition, the component the manner of realising the curriculum, according to the answers supplied by the teachers, is the more useful for planning a lesson (38.7 %) and for monthly planning (35.3 %) than for yearly planning (25.2 %). The reason why these components are not used by the majority of the teachers has more to do with the fact

that the currently available information is not relevant for the monthly level of planning, and less with the functional applicability of the teachers' general curriculum knowledge.

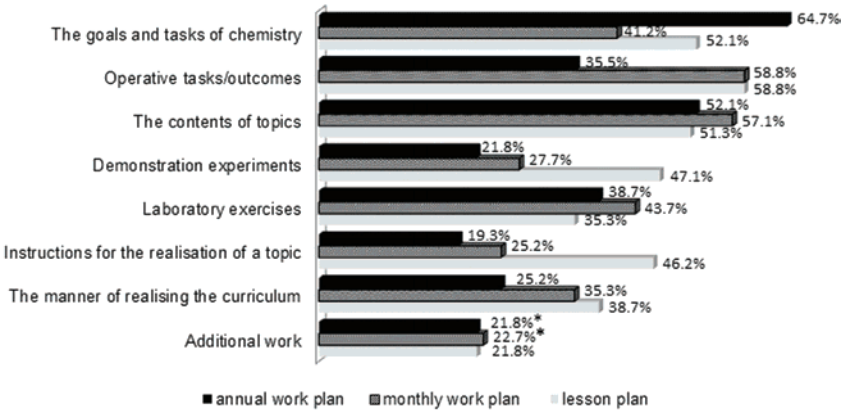


Fig. 2. The applicability of curriculum components for planning teaching.

The teachers' answers indicate that their curriculum knowledge does not enable them to recognise which components from the curriculum may be of importance for annual planning, especially when it comes to planning the overall funds required for the realisation of chemistry teaching in the course of a school year. Thus, for example, more than 60 % of the teachers do not use the information about demonstration experiments and laboratory exercises from the curriculum in their annual planning in order to assess the overall needs for laboratory equipment and substances required for the realisation of the experimental part of chemistry teaching during the course of the school year. In this respect, there was no statistically significant difference between the answers of those teachers who were initially educated for teaching and those who were not.

The current curricula neither develop some teaching units in detail, nor do they do so when it comes to the manner of work to be applied in some lessons. Lesson planning encompasses devising students' activities and planning the means that enable the outcomes envisaged by the curriculum for each segment of a lesson to be achieved. No statistically significant differences in connection with the use of the curriculum for the purpose of planning a lesson were found among the teachers with different initial education working in grammar schools. However, there is a statistically significant difference between the answers of the teachers with different initial education working in secondary vocational schools concerning the use of the curriculum for annual planning and lesson planning. Actually, all the teachers who were initially educated for working in a school use the curriculum for annual planning, and in this respect they differ from the other

teachers in a statistically significant manner. As opposed to this, the least number of them use the curriculum for lesson planning. The greater degree of confidence of these teachers and their greater autonomy when it comes to deciding what a lesson should be like are probably partly a result of their initial education and the previously reviewed approaches to the realisation of curricular contents. On the other hand, the teachers who did not have that form of initial education, when faced with new curricula (and the greatest changes over the past 15 years have occurred in the domain of vocational school curricula) and new requirements, feel a greater need for guidance provided by the curriculum. Apart from the curriculum, the teachers who were not prepared for teaching through initial education rely on the textbook when preparing a lesson and the textbook is often considered to constitute additional material.[21,22]

As already indicated, in their annual planning, the teachers do not use the components demonstration experiments and laboratory exercises for planning the overall equipment required for the realisation of chemistry teaching during a school year, whereas a number of them use the component laboratory exercises in their monthly planning, when they establish how much time is required for realising the laboratory exercises. The component demonstration experiments become important for a large number of teachers at the level of lesson planning.

The curriculum component that pertains to additional work with students is used in different ways by the teachers working in GS and SVS at the level of preparing the annual work plan ($\chi^2(1.119) = 9.33$, $p = 0.00$) and the monthly work plan ($\chi^2(1,119) = 8.15$, $p = 0.00$). In grammar schools, this component is used by 39.0 % of the teachers for preparing their plans at both levels of planning. In secondary vocational schools, this component is of importance for preparing the annual work plan for 12.8 % of the teachers and 14.1 % find it important for preparing the monthly work plan. There is a significant difference between the number of the teachers working in grammar schools and secondary vocational schools who use this component in their annual and monthly planning. The instructions for additional work are more important to the teachers working in grammar schools, which is explained by the fact that there are a greater number of students there who are high achievers, interested in further education in the domain of natural sciences and in various kinds of additional activities (competitions, projects, *etc.*). On the other hand, the programme does not provide any recommendations for additional support for such students who are faced with a lack of success in their learning.

Curricular knowledge, generally speaking, and knowledge of the chemistry curriculum are contained within the competence of chemistry teachers required to transform the curricula specific to chemistry into real teaching situations in the classroom. The aim of Q10 was to investigate which information from the curriculum guides this process. When answering this question, the teachers assessed

the usefulness of information from the curriculum in terms of selecting the teaching/learning method, devising activities and teaching situations. The teachers assessed the degree of usefulness on a scale of 1 to 5 (1 – not at all, 2 – negligibly small, 3 – small, 4 – mostly and 5 – completely). The calculated mean values and standard deviations are given in Table II. The teaching situations in Table II are classified based on the decreasing mean values pertaining to Q10. The calculated mean values ranged from 4.08 to 2.88. The results show that the teachers are best guided by the curriculum in the case of activities aimed at explaining and defining new concepts, whereas it provides the least support in organising the preparation of projects.

Answering Q11, the teachers specified for every activity which components from the curriculum they use for planning and realising the given activity. The percentages relating to the answers to this question are presented in Table II.

The answers given by teachers with different initial education differed statistically to a significant degree with regards to the translation of certain information from the curriculum into some of the classroom activities under consideration (the values for the Cramer V indicated that the initial teacher education had some influence, Table III). The information offered in the curriculum pertaining to these activities is mostly used by the teachers who were educated at the Faculty of Chemistry to be chemistry teachers. It can be seen from the results that the teachers who attended a teacher training programme translate information from the curriculum into teaching situations aimed at systematising the curriculum contents, acquisition of curriculum contents through problem solving, demonstration of experiments and other teaching aids to a greater degree than those who did not.

When it comes to demonstrating experiments and organising laboratory work, most of the teachers use the curriculum components that explicitly refer to these activities. Among the teachers in the sample, higher percentages of those who had partaken in an appropriate teachers training programme applied information from the curriculum within the teaching process through the demonstration method and problem solving. This is indicated by the answers of the teachers with different initial education concerning translation of curriculum information into these classroom activities (Table III). A necessary segment of chemistry teachers' pedagogical chemical knowledge is the knowhow and skills required for laboratory work.[23] Preparing and designing an experiment, a hands-on practical or experiment plan, implementing and evaluating a systematic and effective experiment are singled out as the most important competences of chemistry teachers.[24] For this reason, it is necessary to assess the extent to which teachers are capable of realising such curriculum requirements in practice.

The teachers consider traditional activities such as presenting curriculum contents and monitoring and checking students' achievements in accordance with

TABLE II. Teachers' answers to questions about the applicability of curriculum components for the realisation of teaching situations – questions Q10 and Q11 (N = 119)

Teaching situations	X	SD	Components of the curriculum							
			Goals and tasks of chemistry	Operative tasks/ out-comes	Con-tents	Demonstration experiments	Laboratory exercises	Instructions for realisation of a topic	Manner of realising curriculum	Additional work
2) Explaining and defining new concepts	4.08	0.91	26.1	41.2	39.5	23.5	19.3	19.3	12.6	8.4
6) Verifying the degree to which the subject matter taught in class has been learnt	3.96	0.88	26.9	41.2	21.8	10.9	13.4	5.9	17.6	5.9
7)Systematisation of the course contents	3.95	0.76	31.9	42.0	21.8	10.9	9.2	8.4	18.5[a]	3.4
1) Introducing a topic to students	3.93	0.84	49.6	26.9	37.8	21.8	14.3	19.3	14.3	4.2
13) Monitoring and assessing students' work	3.91	0.94	18.5	37.0	16.8	7.6	10.9	13.4	15.1	8.4
4) Organising laboratory exercises	3.77	1.15	16.8	17.6	10.1	19.3	46.2	15.1	15.1	5.9
3) Demonstration of experiments	3.75	1.03	11.8[a]	17.6	16.8	52.9	19.3	10.9	7.6	5.0
15) Using additional sources of knowledge (literature, the Internet...)	3.70	1.00	16.0	25.2	21.0	9.2	8.4	11.8	14.3	36.1
14) Assigning homework	3.70	0.92	16.0	31.1	24.4	4.2[a]	7.6	10.1	10.9	8.4
5) Demonstrating teaching aids	3.67	0.90	10.9	12.6[a]	18.5	19.3[a]	12.6	23.5	13.4	3.4
9) Organising individual work	3.61	1.01	10.1	16.8[a]	10.9	13.4	26.1	20.2	18.5	10.1
10) Acquisition of curriculum contents through problem solving	3.57	0.98	18.5	21.0	15.1	14.3	20.2	17.6	25.2[a]	20.2
8) Organising group work	3.56	0.94	12.6	17.6	14.3	14.3	26.9	22.7	26.1	10.1
12) Adjusting work to students with special needs	2.99	1.29	18.5	16.0	18.5	5.9	4.2	13.4	18.5	16.8
11) Project work	2.88	1.12	15.1	12.6	13.4	6.7	5.0	11.8	15.1	31.1

[a]Statistically significant differences between answers of teachers with different initial education

TABLE III. Answers of teachers with different initial education concerning translation of curriculum information into classroom activities

Activity/component	FC-TTP %	FC-NTTP %	ONTTP %	χ^2	p	Cramer V
Systematisation of curriculum contents /the manner of realisation of the curriculum	54.5	16.7	8.3	13.32	0.00	0.30
Acquisition of curriculum through problem solving/the manner of realising the curriculum	54.7	26.2	8.3	8.69	0.01	0.27
Demonstration of experiments/goals and tasks of chemistry	27.3	7.1	20.8	6.18	0.05	0.23
Demonstration of teaching aids/operative tasks - outcomes	36.4	13.1	0.0	9.12	0.01	0.28
Demonstration of teaching aids/demonstration of experiments	36.4	27.4	4.2	6.70	0.04	0.04

the established goals and outcomes of learning to be more applicable in practice than these activities that serve to fulfil contemporary requirements in teaching. This pertains to the application of contemporary work methods, the inclusion and presentation of new scientific achievements, adjustment to the context that refers to local and global levels, and also to the recognition of students' individual abilities. In such situations, more experienced teachers adjust their teaching more successfully, taking into consideration their students' abilities. However, experience is not a decisive factor when it comes to how much the competences of teachers are developed.[25] Thus younger teachers find more support in the curriculum for presenting the role of chemistry in various professions, which may be connected with the fact that they are better informed about the contemporary trends in science.

Apart from this, it was investigated how the teachers transform information from the curriculum to other activities that characterise good teaching practices, which are important both for chemistry and for all the other subjects being taught. The results of the teachers' answers are given in Table IV. The activities and results are classified based on the decreasing mean value for answers to question Q12.

The teachers who work in different kinds of schools and have different initial education did not manifest significant statistical differences in their assessments of the influence of the curriculum on the characteristics of their teaching practice. However, there are statistically significant differences between the assessments of the teachers with different working experience concerning the *recognition of individual abilities, predilections and the needs of students* ($\chi^2(1.104) = 20.02$, $p = 0.49$) and seeing the role of chemistry in certain professions ($\chi^2(1.107) = 29.15$, $p = 0.02$). According to the values for the Cramer V

TABLE IV. Teachers' answers to questions about the applicability of the curriculum for realisation of good teaching practices (N=119)

Chemistry curricula enable	1 Not at all	2 Negligibly small	3 Small	4 Mostly	5 Completely	X	SD
9) Students to continue education and for higher levels of education	0.0	3.7	5.6	50.0	40.7	4.28	0.73
1) Planning activities according to clearly defined teaching goals	0.0	3.8	7.7	62.5	26.0	4.11	0.70
18) Checking students' knowledge according to clearly defined learning outcomes	1.9	10.6	11.5	42.3	33.7	3.95	1.03
6) Linking curriculum contents with other natural sciences subjects	0.0	12.3	6.6	55.7	25.5	3.94	0.90
10) Students' continual self-education	0.0	8.7	14.4	53.8	23.1	3.91	0.85
16) Monitoring and evaluation of students' achievements	0.0	10.3	11.2	57.0	21.5	3.90	0.86
14) Seeing the role of chemistry in certain professions	0.9	10.3	14.0	51.4	23.4	3.86[a]	0.93
4) Acquisition of knowledge in keeping with students' age and previous knowledge	1.0	9.5	14.3	56.2	19.0	3.83	0.88
17) Application of clearly defined evaluation criteria	1.9	19.2	8.7	36.5	33.7	3.81	1.16
3) Teaching process featuring students in an active role	1.8	12.8	12.8	48.6	23.9	3.80	1.01
5) Acquisition of all necessary competences, knowledge, skills, views, values	1.0	9.5	16.2	56.2	17.1	3.79	0.87
7) Recognising the individual abilities, predilections and needs of students	1.0	13.5	19.2	39.4	26.9	3.78[a]	1.02
12) Inclusion and presentation of new knowledge and achievements	0.0	14.3	14.3	53.3	18.1	3.75	0.92
13) Adjusting the contents to subjects and grades	3.8	11.3	13.2	51.9	19.8	3.73	1.03
2) Application of contemporary work methods	0.9	14.2	15.1	54.7	15.1	3.69	0.93
15) Adjusting the contents to the technical, cultural and general level of social development	2.8	12.1	20.6	54.2	10.3	3.57	0.93
11) Adjusting the curriculum contents to the needs of the local environment	3.8	21.7	25.5	34.9	14.2	3.34	1.09
8) Inclusion of children with special needs	12.5	17.3	31.7	18.3	20.2	3.16	1.29

[a]Statistically significant differences between answers of teachers with different working experience

(0.50 and 0.52), this is considered to constitute a great influence.[26] The teachers with 20 years of working experience and more are of the opinion that the curricula mostly or completely provides a framework suitable for recognising the individual abilities of students through the teaching process. All the teachers with 5 to 10 years of working experience are of the opinion that the curricula mostly or completely enable the role of chemistry through certain professions to be seen.

Teachers' views on the necessary changes in the curricula

The teachers' views on the changes necessary in the curricula were also investigated. The teachers' competences, which include curriculum knowledge, should enable a critical view of the key documents for the realisation of teaching. For this reason, teachers are engaged during the course of reforms in order to give proposals for new curricula.[27] Within the framework of this survey, a detailed analysis of the curriculum was not expected, but an assessment of whether the information provided in the curriculum components is sufficient and specific enough for the planning and realisation of classroom work (Table V).

TABLE V. Teachers' answers (in %) to questions concerning changes in the contents of curriculum components (N = 119)

Curriculum components	Concretisation/reformulation/ developing the existing contents	Adding new elements	Excluding the existing contents
Goals and tasks of chemistry	71.4a	10.9a	3.4
Operative tasks/outcomes	58.0	26.1	3.4
Contents	37.8	27.7	17.6
Demonstration of experiments	41.2	38.7	2.5
Laboratory exercises	42.0	37.8	1.7
Instructions for the realisation of a topic	48.7	29.4	2.5
The manner of realising the curriculum	52.9	26.9	1.7
Additional work	37.8	39.5	5.0

[a]Statistically significant differences between answers of teachers with different working experience

There was a statistically significant difference between the views of the teachers with different working experience to the effect that, in the current curricula, it is necessary to concretise/reformulate/develop the existing goals and tasks of chemistry ($\chi^2(4.119) = 10.78$, $p = 0.03$) and to add new goals and tasks ($\chi^2(4.119) = 10.56$ $p = 0.03$). The teachers with less than 5 years working experience were the most in favour of the concretisation of the goals and tasks of chemistry (90.9 %). The percentage of such answers decreased the longer was the teachers' working experience, reaching up to 50.0 % among the oldest teachers. The teachers with more than 10 years of working experience differed from their younger colleagues in their request for adding new goals and tasks of chemistry.

The percentage of such answers increased with increasing working experience of the teachers (from 7.8 % to 31.2 %). These results indicate that the curricular knowledge depends on the length of the teachers' professional experience.

Curricular knowledge and an active attitude towards the curriculum on the part of teachers encompass their critical view of the structure and contents of the curriculum. Critical and analytical reviewing of the curriculum contributes to a greater ability of teachers to make the necessary decisions in their teaching on the basis of information from the curricular.[28] Moreover, in this way one obtains important feedback from teachers with practical experience about the quality of the curriculum, which is of importance for their future advancement. Teachers should understand the curriculum material as their professional means of work, and should learn, through education and professional development, about curricula and from them.[29]

CONCLUSIONS

The competences of chemistry teachers, a set of knowledge and skills needed to perform an activity, comprise, among other things, curricular knowledge. Translation of information from the curriculum into corresponding teaching situations, tasks for students, *i.e.*, activities in the classroom, requires knowledge of the nature of the contents being taught to students and the nature of the problem of forming certain concepts, then knowledge of the characteristics of certain methods of teaching and learning, the characteristics of the age group of the students that a teacher is working with, their interests and possible adjustments. The translation is also connected with planning the resources required for the realisation of teaching according to the curriculum (the materials for experiments, special teaching aids, printed and electronic materials, *etc*.). The entire teaching process depends on how teachers translate curriculum requirements into a form that students can understand and acquire.

The results of this survey show that the teachers are not aware of all the roles and the importance of certain curriculum components. Furthermore, it was perceived that information from certain components remains unused even though it is relevant to a certain level of planning (for example, about demonstration experiments and laboratory exercises, for the purpose of the annual planning of the overall funds required for teaching during one school year).

Teachers' professionalism encompasses autonomy and responsibility. The important question here is what autonomy means in relation to the curriculum, how a teacher uses information from the curriculum as a framework for various adjustments that teaching requires in a particular context (according to previous knowledge, interests and other needs of students, the available means at school, the requirements of the profession that students are preparing for, the needs of the local environment and society in general). Autonomy comprises deciding which

methods to use in order to realise teaching in keeping with the envisioned contents, the goals and outcomes in the curriculum. It also comprises deciding on how to set up a demonstration experiment, how to organise laboratory practice (individual or group work), and the like. A teacher's responsibility is reflected in his/her acknowledgement of information from the curriculum, professional assessment of the relevance of the information from some components for the planning and realisation of teaching.[30] Within the sample of teachers encompassed by this survey, it turned out that the goals and operative tasks/outcomes are the most important curriculum component for the teachers' work.

A well prepared chemistry teacher should apply and evaluate the curriculum. He/she attains the ability to do so by acquiring general curriculum knowledge, knowledge of the chemistry curriculum, and by forming a professional attitude towards the curriculum. The acquisition of such knowledge and the development of an active and professional attitude towards the curriculum should be included in teachers' initial education and further developed through programmes of continual professional development. The indicators of teachers' curricular knowledge are their understanding of the purpose and role of this document in the planning of teaching, the manner of applying the information contained in the curriculum and the evaluation of the quality of the curriculum. An important issue for the interpretation of the curriculum are a teacher's beliefs: beliefs about the goals or purposes of science teaching, beliefs about the nature of science and beliefs about science teaching and learning.[31]

Research has shown that there is not enough literature that could help teachers to understand the documents, instructions and materials according to which they should work,[32] and that there are few papers aimed at the manner of realising a great number of the prescribed standards.[33–35]

The survey realised provides a basis for defining indicators for monitoring the ability of teachers to apply curricular knowledge in their practice. Such indicators are important for creating tasks in the initial education of teachers, through which curricular knowledge is developed, tasks for teachers' professional development and for monitoring teachers' progress and planning activities which could improve curriculum knowledge.

A limitation of this survey is that monitoring was based on a teacher's personal assessment that could lead to attempts to give an answer that is considered desirable. Future surveys will include research methods of direct assessment of the degree and application of a chemistry teacher's knowledge of curricula.

SUPPLEMENTARY MATERIAL

The Questionnaire is available electronically from http://www.shd.org.rs/JSCS/, or from the corresponding author on request.

Acknowledgements. This paper represents a result of work within the project “Scientific Theory and Practice in Society: Multidisciplinary, Educational and Inter-generational Perspectives”, Project No. 179048, the realisation of which was financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

И З В О Д

ЗНАЊЕ СРЕДЊОШКОЛСКИХ НАСТАВНИКА О НАСТАВНИМ ПРОГРАМИМА ХЕМИЈЕ

БИЉАНА ТОМАШЕВИЋ и ДРАГИЦА ТРИВИЋ

Хемијски факултет Универзитета у Београду, п.пр. 158, 11001 Београд

У овом истраживању испитивано је професионално знање наставника хемије о структури, садржају и примени наставних програма хемије и њихових компоненти. Истраживањем је обухваћено 119 наставника из 69 средњих школа (25 гимназија и 44 средње стручне школе). Питања упитника односила су се на опште знање о наставним програмима, знање о наставним програмима хемије и на ставове о неопходним изменама у актуелним програмима. Одговори наставника показују да су за њихов рад најзначајније компоненте програма циљеви и оперативни задаци/исходи. Уочено је да информације из одређених компоненти остају неискоришћене иако су релевантне за одређени ниво планирања. Међу наставницима у узорку, они који имају иницијално образовање за наставничку професију у већем проценту су примењивали информације из наставног програма у реализацији наставе демонстрационом методом и методом учења путем решавања проблема. Изведено истраживање пружа основ за дефинисање индикатора за праћење оспособљености наставника да примењују знање о наставним програмима у својој пракси. Такви индикатори су значајни у креирању иницијалног образовања наставника и програма за њихов професионални развој.

(Примљено 2. октобра, ревидирано 7. децембра, прихваћено 8. децембра 2014)

REFERENCES

1. European Commission, *Education and Training, Supporting teacher competence development for better learning outcomes*, http://ec.europa.eu/education/policy/school/ teacher--training_en.htm (accessed 5/8/2014)
2. R. Deakin Crick, in *Getting involved: Global citizenship development and sources of moral values*, F. Oser, W. Veugelers, Eds., Sense Publishers, Rotterdam, 2008, p. 31
3. F. Caena, U. Margiotta, *Eur. Educ. Res. J.* **9** (2010) 317
4. L. S. Shulman, *Educ. Researcher* **15** (1986) 4
5. O. De Jong, *Eur. J. Teach. Educ.* **23** (2000) 127
6. A. N. Geddis, *Int. J. Sci. Educ.* **15** (1993) 673
7. Y. Shwartz, D. Katchevitch, *Chem. Educ. Res. Pract.* **14** (2013) 312
8. R. Blonder, M. Jonatan, Z. Bar-Dov, N. Benny, S. Rap, S. Sakhnini, *Chem. Educ. Res. Pract.* **14** (2013) 269
9. S. Magnusson, J. Krajcik, H. Borko, in *Examining Pedagogical Content Knowledge: The Construct and its Implications for Science Education*, J. Gess-Newsome, N. Lederman, Eds., Kluwer Academic, Dordrecht, 1999, p. 95
10. O. N. Kaya, *Int. J. Sci. Educ.* **31** (2009) 961
11. L. S. Shulman, *Harvard Educ. Rev.* **57** (1987) 1
12. M. Z. Hashweh, *Teachers Teaching: Theory Practice* **11** (2005) 273

13. W. S. Carlsen, in *Examining Pedagogical Content Knowledge: The Construct and its Implications for Science Education*, J. Gess-Newsome, N. Lederman Eds., Kluwer Academic, Dordrecht, 1999, p. 133
14. J. Barnett, D. Hodson, *Sci. Educ.* **85** (2001) 426
15. S. K. Abell, *Int. J. Sci. Educ.* **30** (2008) 1405
16. H. Borko, *Educ. Researcher* **33** (2004) 3
17. J. H. Van Driel, O. De Jong, N. Verloop, *Sci. Teac. Educ.* **86** (2002) 572
18. J. H. Van Driel, A. M. W. Bulte, N. Verloop, *Learn. Instr.* **17** (2007) 156
19. M. Drechsler, J. Van Driel, *Chem. Educ. Res. Pract.* **10** (2009) 86
20. *Statistical Office of the Republic of Serbia*, http://webrzs.stat.gov.rs (accessed 1/12/2013)
21. R. Bucat, *Chem. Educ. Res. Pract.* **5** (2004) 215
22. K. Padilla J. Van Driel, *Chem. Educ. Res. Pract.* **12** (2011) 367
23. J. Bond-Robinson, *Chem. Educ. Res. Pract.* **6** (2005) 83
24. J. Copriady, *Mediterr. J. Social Sci.* **5** (2014) 312
25. P. J. Friedrichsen, S. K. Abell, E. M. Pareja, P. L. Brown, D. M. Lankford, M. J. Volkmann, *J. Res. Sci. Teach.* **46** (2009) 357
26. F. J. Gravetter, L. B. Wallnau, *Statistics for the Behavioral Sciences*, Wadsworth, Belmont, CA, 2004
27. S. A. Al-Amoush, S. Markic, I. Eilks, *Chem. Educ. Res. Pract.* **13** (2012) 314
28. C. J. Beyer, E. A. Davis, *Sci. Educ.* **96** (2012) 130
29. P. Grossman, C. Thompson, *Teach. Teach. Educ.* **24** (2008) 2014
30. S. Park, J. S. Oliver, *Res. Sci. Educ.* **38** (2008) 261
31. P. Friedrichsen, J. H. Van Driel, S. K. Abell, *Sci. Educ.* **95** (2011) 358
32. E. A. Davis, D. Petish, J. Smithey, *Rev. Educ. Res.* **76** (2006) 607
33. P. Adams, G. H. Krockover, *Sci. Educ.* **81** (1997) 29
34. S. Lynch, *J. Res. Sci. Teach.* **34** (1997) 3
35. S. A. Southerland, J. Gess-Newsome, *Sci. Educ.* **83** (1999) 131.

J. Serb. Chem. Soc. 80 (3) S113–S118 (2015) *Supplementary material*

SUPPLEMENTARY MATERIAL TO

Chemistry curricular knowledge of secondary school teachers

BILJANA TOMAŠEVIĆ*# and DRAGICA TRIVIĆ#

Faculty of Chemistry, University of Belgrade, P. O. Box 158, 11001 Belgrade, Serbia

J. Serb. Chem. Soc. 80 (3) (2015) 435–452

THE QUESTIONNAIRE

1

In which type of school do you teach?

a) Grammar School
b) Secondary Vocational School

2

How many years have you taught?

a) Less than 5
b) 5-10
c) 10-20
d) 20-30
e) More than 30

3

Sex:

a) Male
b) Female

4

How old are you?

a) 25-30
b) 31-40
c) 41-50
d) 51-60
e) Above 60

5

What are your academic qualifications?

_____
_____
_____

* Corresponding author. E-mail: bsteljic@chem.bg.ac.rs

6

In which levels of planning do you use the chemistry curriculum?

a) When preparing the annual work plan
b) When preparing a monthly work plan
c) When preparing a class scenario in writing

7

What kind of information contained in the curriculum is the most important to you for the realisation of your teaching plan?

a) The goals and tasks of chemistry
b) Operative tasks/outcomes
c) The contents of topics
d) Demonstration experiments
e) Laboratory exercises
f) Instructions for the realisation of a topic
g) The manner of realising the curriculum
h) Additional work

8

Has your need to use the curriculum (during periods when there were no changes to it) changed as your working experience increased?

a) I have used the curriculum to an equal degree all the time
b) I use the curriculum less now than when I started working
c) I use the curriculum more now than when I started working

9

Mark the curriculum components which you use for the different levels of planning.

Levels of planning	Components of curriculum							
	The goals and tasks of chemistry	Operative tasks/ outcomes	The contents of topics	Demonstration experiments	Laboratory exercises	Instructions for the realisation of a topic	The manner of realising the curriculum	Additional work
The annual work plan								
The monthly work plan								
The lesson plan								

10

To what extent are you guided by information contained in the curriculum in realising the following teaching situations?

1 – Not at all 2 – Negligibly small 3 – Small 4 – Mostly 5 – Completely

1) Introducing a topic to student

2) Explaining and defining new concepts

3) Demonstration of experiments

4) Organising laboratory exercises

5) Demonstrating teaching aids

6) Verifying the degree to which the subject matter taught in class has been learned

7)Systematisation of the course contents

8) Organising group work

9) Organising individual work

10) Acquisition of curriculum contents through problem solving

11) Project work

12) Adjusting work to students with special needs

13) Monitoring and assessing students' work

14) Assigning homework

15) Using additional sources of knowledge (literature, the Internet...)

11

Mark the curriculum components through which you obtain the most information in realising the following teaching situations?

Teaching situations	**Components of curriculum**							
	The goals and tasks of chemis-try	Opera-tive tasks/ outcomes	The con-tents of topics	Demon-stration experi-ments	Labora-tory exercises	Instruct-ions for the realisation of a topic	The manner of realising the curri-culum	Addi-tional work
1) Introducing a topic to students								
2) Explaining and defining new concepts								
3) Demonstration of experiments								

4) Organising laboratory exercises								
5) Demonstrating teaching aids								
6) Verifying the degree to which the subject matter taught in class has been learned								
7) Systematis-ation of the course contents								
8) Organising group work								
9) Organising individual work								
10) Acquisition of curriculum contents through problem solving								
11) Project work								
12) Adjusting work to students with special needs								
13) Monitoring and assessing students' work								
14) Assigning homework								
15) Using additional								

sources of knowledge (literature, the Internet...)								

12

To what extent does the curriculum offer you possibilities to realise the following.

1 – Not at all 2 – Negligibly small 3 – Small 4 – Mostly 5 – Completely

1) Planning activities according to clearly defined teaching goals

2) Application of contemporary work methods

3) Teaching process featuring students in an active role

4) Acquisition of knowledge in keeping with students' age and previous knowledge

5) Acquisition of all necessary competences, knowledge, skills, views, values

6) Linking curriculum contents with other natural sciences subjects

7) Recognising the individual abilities, predilections and needs of students

8) Inclusion of children with special needs

9) Students to continue education and for higher levels of education

10) Students' continual self-education

11) Adjusting the curriculum contents to the needs of the local environment

12) Inclusion and presentation of new knowledge and achievements

13) Adjusting the contents to subjects and grades

14) Seeing the role of chemistry in certain professions

15) Adjusting the contents to the technical, cultural and general level of social development

16) Monitoring and evaluation of students' achievements

17) Application of clearly defined evaluation criteria

18) Checking students' knowledge according to clearly defined learning outcomes

13

What changes in the curriculum of chemistry would you be most important for teaching chemistry?

Curriculum components	Concretisation/ reformulation/ developing the existing contents	Adding new elements	Excluding the existing contents
Goals and tasks of chemistry			
Operative tasks / outcomes			
Contents			
Demonstration of experiments			
Laboratory exercises			
Instructions for the realisation of a topic			
The manner of realising the curriculum			
Additional work			