Tetrahedral Homonuclear Tetrameric Species: Occurrence, Forms, Structures, Properties, and Perspectives

Kiralj, Rudolf

Source / Izvornik: **Kemija u industriji : Časopis kemičara i kemijskih inženjera Hrvatske, 2024, 73, 27 - 44**

Journal article, Published version
Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

https://doi.org/10.15255/KUI.2023.024

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:144:174538

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2025-02-22**

Repository / Repozitorij:

Digital Repository of Bjelovar University of Applied Sciences

https://doi.org/10.15255/KUI.2023.024
KUI-04/2024
Original scientific paper
Received August 1, 2023
Accepted October 7, 2023

Tetrahedral Homonuclear Tetrameric Species: Occurrence, Forms, Structures, Properties, and Perspectives

R. Kiralj *

Bjelovar University of Applied Sciences, Trg Eugena Kvaternika 4, 43 000 Bjelovar, Croatia

This work is licensed under a Creative Commons Attribution 4.0 International License

Abstract

Tetrahedral homonuclear tetrameric species are neutral or ionic tetrahedra of chemical elements, true molecular tetrahedra in strict geometrical sense. The aim of this work was to find all true elemental tetrahedra experimentally determined as stable and capable of being transferred from one medium/condition to another, to rationalize their structures and properties, and seek their technological significance. A database of elemental tetrahedra was built based on reliable literature, and a dataset of molecular descriptors was generated. Correlation analysis, hierarchical cluster analysis, principal component analysis, and some regression analysis were carried out. Twelve chemical elements from Groups 13 – 15 have tetrahedral species: anions E_4^{8-} (Al, Ga, In, Tl), anions E_4^{4-} (Si, Ge, Sn, Pb), neutral molecules E_4 (P, As, Sb), and cations E_4^{+} (P, As, Sb, Bi). These species appear in 39 forms, such as gas phase, solid Zintl phases, thin films, Zintl phase ammoniates, and cage compounds, among others. Zintl phases and forms of P_4 and As_4 are important for synthesis of new materials, chemical reactions, and other applications. Data compression has shown that the true tetrahedral species are a three-dimensional phenomenon, and that true tetrahedra can be distinguished from fictive tetrahedra. Grouping of the tetrahedra according to their periodic groups, charges, frequency, and diversity class values, tetrahedral size, and metallic character, were observed. Empirical rules for species size and principal components were established for true tetrahedra. Most descriptors are parabolic functions of the class variable. True tetrahedral species partly remain a puzzle and should be further investigated.

Keywords

Homonuclear tetrahedral molecules, structural databases, Zintl phases, cage compounds, statistical analysis, dimensionality reduction, empirical rules

All nature is at the disposal of humankind.
We are to work with it. For without, we cannot survive.
(St. Hildegard of Bingen).

1 Introduction

In the strict geometrical sense, a tetrahedron is a three-dimensional polyhedron consisting of four triangles (triangular faces) joined at their edges (six edges), and the edges are joined into vertices (four vertices).[1] Therefore, geometry, as a mathematical discipline, is concerned with the shape of a tetrahedron, not with its internal structure. Geometry supposes that the entire tetrahedron is made of homogeneous matter. In the world of small molecules, the closest to a geometrical tetrahedron are tetrahedral homonuclear tetrameric species (neutral molecules or ions), whose four atoms are situated at the vertices, and covalent chemical bonds are the edges of the triangular faces. Within such a tetrahedron, there are no atoms, and these molecules can be called true tetrahedral molecules. They are not so much known among chemists. Chemistry textbooks at the university level usually illustrate a tetrahedral molecule of white phosphorus P_4,[2–4] while molecules As_4 and Sb_4 are just mentioned as existing in the vapour phase,[2,3] and As_4 as molecules of yellow arsenic.[3] What in chemistry is usually called a tetrahedron (tetrahedral structure[5,6]), is a molecule having four atoms at the four vertices and the fifth atom at the centre of a fictive geometrical tetrahedron, while the chemical bonds exist only inside the tetrahedron, between the central atom and the vertex atoms. In such a case, edges of the tetrahedron do not exist, which is why the tetrahedron is fictive; in fact, the edges should be imagined as directional non-bonding interactions between two atoms bound to the central atom. Such tetrahedral molecules, starting with the tetrahedral carbon atom, were gradually accepted in chemistry during the 19th and 20th centuries,[7–9] especially in organic chemistry and inorganic stereochemistry. Molecules CH_4, CCl_4, and CMe_4 are didactical examples of such tetrahedra.[6,10,11]

This work deals with true tetrahedra of chemical elements: explores their existence and occurrence in various forms, rationalizes their structures and properties, and seeks whether materials containing such tetrahedra have some technological importance. In the methodological sense, the research follows the sequence: database – dataset – analysis – interpretation.[12,13] Firstly, the database of true tetrahedra was formed based on reliable literature sources. An appropriate dataset of molecular descriptors was then generated. Finally, the dataset was analysed by means of various statistical and chemometric methods, and interpreted in terms of both statistics and chemistry. An important issue resulting from the analyses is the set of conditions the chemical elements must satisfy to form stable tetrahedral species, capable of surviving for a significant period as independent species, and being able to transfer from one medium to another.

* Rudolf Kiralj, PhD
Email: rkiralj@vub.hr

2 Methods

2.1 Database and dataset formation for tetrahedral homonuclear tetramers

A database for tetrahedral homonuclear tetramers was created through an extensive literature search for known tetrahedral homonuclear tetrameric species (both neutral molecules and ions) that are stable and may exist independently from other species. The existence of such tetrameric species (regular or irregular tetrahedra) has been determined experimentally, and in special cases, theoretical calculations were supported by strong experimental evidence that indicated the existence of the species in question. This way, the set of periodic groups of chemical elements was defined as the block of elements whose members may exist as tetrahedral species.

Once the database had been formed, a corresponding dataset was generated from steric and electronic descriptors for chemical elements of the aforementioned groups and two neighbouring groups. The descriptors were from relatively novel and relevant literature. The dataset was created with the purpose of rationalising the existence of stable tetrahedral homonuclear tetramers throughout the periodic table, by means of statistical and chemometric analysis of relevant descriptors.

2.2 Statistical and chemometric analyses

The descriptors dataset was analysed for mutual correlations and their statistical significance using programming, numeric computing and graphics software *Scilab*[14], and online statistical software *Value from Pearson (R) Calculator.*[15] The dataset, in matrix form, was normalised using programmed autoscaling in *Scilab*,[14] where the values for each descriptor were centred by subtracting its mean value and then divided by its standard deviation. Additionally, the values were divided by the square root of $n-1$ (n is the number of data, *i.e.* the number of tetrahedral species). When such an autoscaled matrix is multiplied by its transpose, the correlation matrix is obtained,[16] and this fact was used in further programming in *Scilab*.[14] Each correlation coefficient, r, from the correlation matrix was then tested for its statistical significance using the online two-tailed t-test,[15] with degrees of freedom $df = n-2$. The t-parameter was defined by Eq. (1).

$$t = r\sqrt{\frac{n-2}{1-r^2}} \tag{1}$$

The online software for the t-test[15] automatically calculates the t-parameter and then the probability for the null-hypothesis, the hypothesis stating that the correlation coefficient is not statistically significantly different from zero.

A sub-dataset in matrix form was made from moderately to highly correlated descriptors: the matrix **X** had n rows (samples, tetrahedral species) and m columns (variables, selected descriptors). It was then processed using data compression methods, hierarchical cluster analysis (HCA) and principal component analysis (PCA),[17,18] which were carried out using chemometrics package *Pirouette*,[19] a multivariate data analysis software. The software performs automatic calculations once methods and conditions are selected, and the results are presented in the form of tables and graphics. For the data preprocessing method, autoscaling was selected, both for HCA and PCA, meaning that each column of the data matrix was mean-centred and divided by its respective standard deviation. HCA for samples (tetrahedral species) was conducted for all available connectivity schemes to determine which one would yield the best chemical clustering. Autoscaled data matrix is an m-dimensional space in which the objects (tetrahedral species) are points, among which Euclidean distances are calculated using *Pirouette*.[19] The samples are then linked into clusters according to a selected connectivity scheme (linking criterion), starting from two-membered clusters and finishing with all the samples as the final cluster. The graphics output, the dendrogram of samples, is based on the calculation of the similarity index, S_{AB},[17,18] among any pair of clusters A and B, as defined by Eq. (2):

$$S_{AB} = 1 - \frac{d_{AB}}{d_{max}} \tag{2}$$

where d_{AB} is the distance between the two clusters and d_{max} is the maximum distance among all clusters.

PCA is a data compression method in which m original variables (columns of the data matrix **X**) are transformed into m principal components (PCs) through a series of matrix equations. The new variables are linear combinations of the old variables, and are arranged in descending order of importance, indicating their contribution to the total variance. PCA in *Pirouette*[19] is based on SVD (singular value decomposition) and NIPALS (nonlinear iterative partial least squares) algorithms.[17] The final results of these procedures are variances and percent variances for all PCs, and coordinates of the old variables (loadings) and samples (scores) in the space of the PCs, in tabular and graphic form. However, *Pirouette*[19] also generates matrix $\hat{\mathbf{X}}$ as the reconstruction of the original matrix **X**, using a definite number k of PCs, so that the residual matrix $\hat{\mathbf{E}}$ is defined by Eq. (3).

$$\hat{\mathbf{E}} = \mathbf{X} - \hat{\mathbf{X}} \tag{3}$$

Furthermore, the residuals for the i-th sample form a row vector according to Eq. (4).

$$\hat{\boldsymbol{e}}_i = \boldsymbol{x}_i - \hat{\boldsymbol{x}}_i \tag{4}$$

The final measure of residuals for all n samples and k PCs is the predicted residual error sum of squares (PRESS), calculated as the sum of products of these vectors with their respective transpose vectors, as defined by Eq. (5).

$$\text{PRESS} = \sum_i^n \frac{\hat{\mathbf{e}}_i \hat{\mathbf{e}}_i^{\mathrm{T}}}{m-k} \tag{5}$$

PRESS and percent variance can be used to determine the optimal number of PCs, but they are not as sensitive as VPRESS (validated predicted residual error sum of squares), an analogue of PRESS. For this purpose, leave-one-out cross-validation was selected in *Pirouette*[19] and VPRESS

values were obtained for all numbers of PCs. Leave-one-out cross-validation is a procedure in which *n* reconstructions of the matrix **X** are obtained. In this process, the *i*-th sample is deleted from **X**, PCA is carried out, and the reconstructed $\hat{\mathbf{X}}_{cv}$ is stored. At the end of leave-one-out cross-validation, the average $\hat{\mathbf{X}}_{cv}$ is calculated, and the residuals for the *i*-th sample form a row vector obtained using Eq. (6).

$$\hat{\boldsymbol{e}}_{cvi} = \boldsymbol{x}_i - \hat{\boldsymbol{x}}_{cvi} \tag{6}$$

Finally, VPRESS is calculated using Eq. (7).

$$\text{VPRESS} = \sum_i^n \frac{\hat{\mathbf{e}}_{cvi}\hat{\mathbf{e}}_{cvi}^{\mathrm{T}}}{m-k} \tag{7}$$

HCA and PCA, also known as exploratory data analysis,[17,18] are very useful for chemical interpretation of chemometric and statistical results. Some examples of the application of HCA and PCA include quantitative relationships between molecular structure and measured biological activity[20,21] or physicochemical properties,[22] structural features of molecules,[12,13,23] molecular modelling problems,[24] and analysis of statistical results,[25] among others.

Other analyses for nonlinear relationships between some variables included parabolic (second-order polynomial) regressions, tested graphically by online software *Polynomial Regression Data Fit*,[26] which uses the system of normal equations in matrix form, and the equations are solved *via* Gauss-Jordan elimination. There are no estimated standard deviations calculated for regression coefficients. That was the reason why the regression analyses with matrix inversion were carefully carried out by programming in *Scilab*.[14] The vector $\boldsymbol{y}$ for a dependent variable and the matrix **X** for an independent variable (first column – ones, second column – values of the independent variable, third column – squares of these values) were defined for each regression. The matrix equation $\boldsymbol{y} = \mathbf{X}\boldsymbol{c}$, where $\boldsymbol{c}$ is the vector of regression coefficients, was solved with the set of matrix operations, and predicted $\hat{\boldsymbol{y}}$ was obtained, using Eqs. (8)–(12).

$$\mathbf{A} = \mathbf{X}^{\mathrm{T}}\boldsymbol{y} \tag{8}$$

$$\mathbf{B} = \mathbf{X}^{\mathrm{T}}\mathbf{X} \tag{9}$$

$$\mathbf{D} = \mathbf{B}^{-1} \tag{10}$$

$$\boldsymbol{c} = \mathbf{D}\mathbf{A} \tag{11}$$

$$\hat{\boldsymbol{y}} = \mathbf{X}\boldsymbol{c} \tag{12}$$

Estimated standard deviations for regression coefficients $\sigma(c_i)$ were calculated as diagonal elements of the variance – covariance matrix. This matrix is the product of matrix **D** with the estimated variance of residuals σ^2 when the residuals follow a normal distribution.[27] Starting with the column vector of residuals $\hat{\boldsymbol{\varepsilon}}$ and with $k = 2$ (number of variables in parabolic regression), the following expressions were programmed:

$$\hat{\boldsymbol{\varepsilon}} = \boldsymbol{y} - \hat{\boldsymbol{y}} \tag{13}$$

$$\sigma^2 = \frac{\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}\hat{\boldsymbol{\varepsilon}}}{n-k-1} \tag{14}$$

$$\sigma(c_i) = \sigma\sqrt{d_{ii}} \tag{15}$$

The coefficient of determination r^2 and the *F*-ratio were determined for *F*-test for a regression equation, and *t*-parameter was calculated for each regression coefficient for *t*-test, using expressions (Eqs. (16)–(19)):

$$\Delta = \boldsymbol{y} - \overline{\boldsymbol{y}} \tag{16}$$

$$r^2 = \frac{\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}\hat{\boldsymbol{\varepsilon}}}{\Delta^{\mathrm{T}}\Delta} \tag{17}$$

$$F = \frac{n-k-1}{k}\frac{r^2}{1-r^2} \tag{18}$$

$$t_i = \frac{c_i}{\sigma(c_i)} \tag{19}$$

where Δ is the deviation vector for $\boldsymbol{y}$, and $\overline{\boldsymbol{y}}$ is the vector containing *n* times the mean value $\overline{y}$ of $\boldsymbol{y}$. The statistical significance of the obtained regression equations was checked *via F*-test using online software *P-Value from F-Ratio Calculator (ANOVA)*,[28] with degrees of freedom equal to *k* for the numerator, and equal to *n*–*k*–1 for the denominator of the *F*-ratio. Two-tailed *t*-tests with *n*–*k*–1 degrees of freedom for all regression coefficients were carried out in online package *P Value Calculator (GraphPad)*.[29] Statistical significance level for all statistical tests in this work was $\alpha = 0.05$. In this manner, the null hypothesis was tested for each regression coefficient, asserting that a regression coefficient was not statistically significantly different from zero. This approach is more rigorous than the *F*-test, which tests the entire equation under the null hypothesis that the *F*-ratio is not statistically significantly different from zero. Confirmation of any null hypothesis means that the regression equation is not reliable.

3 Results and discussion

3.1 Database of stable tetrahedral homonuclear tetramers

Table 1 presents a list of tetrahedral molecules and ions as observed for Groups 13, 14, and 15 of chemicals elements, with relevant references.[30–79] No elements in other groups were found to form stable tetrahedra that could exist independently and be transferred to another medium. Transition metals, for example Mo,[80] Ru,[81] Pt,[82] and Au,[83] form tetrahedral units (clusters), but these are parts of more complicated species and cannot be isolated. Europium deposits on graphene exist in the form of separated tetrahedral clusters,[84] but these tetrahedra cannot be transferred to other media. Among the elements of Groups 13–15, the smallest atoms (B, C, and N) do not form tetrahedra (Table 1). Therefore, only 12 elements exist in tetrahedral form: in Group 13 – Al, Ga, In, and Tl as anions E_4^{8-}; in Group 14 – Si, Ge, Sn, and Pb as anions E_4^{4-}; and in Group 15 – P, As, Sb, and Bi as neutral molecules E_4 or cations E_4^{+}.

Table 1 – Experimentally detected tetrahedral homonuclear tetrameric species (neutral molecules and ions) in Groups 13–15
Tablica 1 – Eksperimentalno opažene tetraedrijske homonuklearne tetramerne vrste (neutralne molekule i ioni) u skupinama 13–15

Element	Species	Form	d_{XX}/Å	Ref.
B	none	---	---	---
Al	Al_4^{8-}	Zintl phase anion, crystalline state	2.72	30
Ga	Ga_4^{8-}	Zintl phase anion, crystalline state	2.67	31
In	In_4^{8-}	Zintl phase anion, crystalline state	3.11	32, 33
Tl	Tl_4^{8-}	Zintl phase anion, crystalline state	3.23	34, 35
C	none	---	---	---
Si	Si_4^{4-}	Zintl phase anion, crystalline state solvated, liquid ammonia solution interstitial hydride, crystalline state Zintl phase anion, melt Zintl phase anion, ammoniate, crystalline state	2.41 --- 2.43 --- ---	36, 37, 38, 39, 40 41 42 43 44
Ge	Ge_4^{4-}	Zintl phase anion, crystalline state Zintl phase anion, ammoniate, crystalline state	2.54 2.56	36, 39, 40, 45, 46 44
Sn	Sn_4^{4-}	Zintl phase anion, crystalline state solvated, liquid ammonia solution clathrate-Zintle phase intergrowth crystalline state Zintl phase anion, ammoniate, crystalline state	2.93 2.94 2.89 2.94	36, 47, 48 41, 49, 50, 51 52 44, 48
Pb	Pb_4^{4-}	Zintl phase anion, crystalline state Zintl phase anion, nanocrystals solvated, liquid ammonia solution Zintl phase anion, melt Zintl phase anion, ammoniate, crystalline	3.10 --- --- --- 3.09	36, 48, 53, 54, 55 56 49 57 44, 48
N	none	---	---	---
P	P_4 P_4^+	white phosphorus, crystalline state, unstable air-stable in organic cage, water solution gas phase, vapour of white/red phosphorus in organic supramolecular cage, crystalline state in organic polymeric cage, crystalline state in hydrophobic cage (carbon nanotube) in superfluid helium, quenched in intercalated system (fullerene), crystalline state gas phase	2.21 --- 2.21 2.16 2.17 --- --- --- ---	58, 59, 60 61 59, 62, 63 63, 64, 65 66 67 68 69 59, 70
As	As_4 As_4^+	gas phase in organic supramolecular cage, crystalline state in hydrophobic cage (carbon nanotube) yellow arsenic, solid state, unstable thin film, stable at very low temperatures gas phase	2.43 2.39 --- --- 2.42 ---	59, 60, 71, 72 63, 64 73 74, 75 75 59
Sb	Sb_4 Sb_4^+	thin film gas phase gas phase	2.68 --- ---	59, 76, 77 77, 78 59
Bi	Bi_4^+	gas phase	3.06	59, 79

Tetrahedral species of Group 13 exist as Zintl phase anions in crystalline state. There is more diversity among tetrahedral anions of Group 14: they exist in crystalline Zintl phases, the anions are found in ammonia solution (ammoniates), in molten Zintl phases, and in some other forms (Table 1). Fig. 1 shows two examples of Zintl phase anions, based on crystal structure data retrieved from the *Access Structures service of the Cambridge Structural Database* (CSD) and *Inorganic Crystal Structure Database* (ICSD),[85] and presented by means of a molecular graphics software *ViewerLite*.[86] Tl_4^{4-} in a pure Zintl phase[34] and Sn_4^{4-} anion in a Zintl phase ammoniate,[48] are both shown with their coordination spheres within van der Waals contact distances.

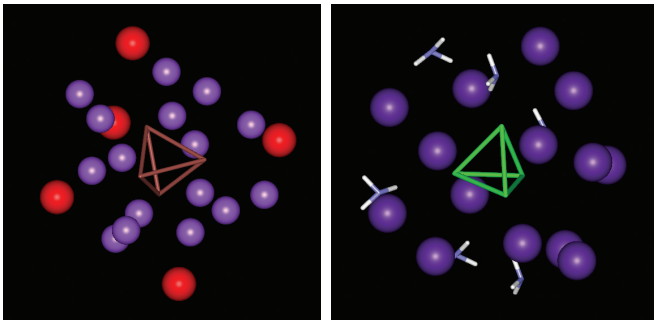

Fig. 1 – Two examples of tetrahedral Zintl phase anions within coordination spheres. Left: Tl_4^{4-} anion (brown) surrounded by Na^+ (purple) and Rb^+ (red) cations in the crystal structure of Na_7RbTl_4 (CSD: 2208674).[34] Right: Sn_4^{4-} anion (green) surrounded by Rb^+ cations (purple) and NH_3 molecules (N – blue, H – white) in the crystal structure of $Rb_4Sn_4{\cdot}2NH_3$ (ICSD: 416976).[48]

Slika 1 – Dva primjera tetraedrijskih aniona Zintlovih faza unutar koordinacijskih sfera. Lijevo: anion Tl_4^{4-} (smeđe) okružen kationima Na^+ (purpurno) i Rb^+ (crveno) u kristalnoj strukturi Na_7RbTl_4 (CSD: 2208674).[34] Desno: anion Sn_4^{4-} (zeleno) okružen kationima Rb^+ (purpurno) i molekulama NH_3 (N – plavo, H – bijelo) u kristalnoj strukturi $Rb_4Sn_4{\cdot}2NH_3$ (ICSD: 416976).[48]

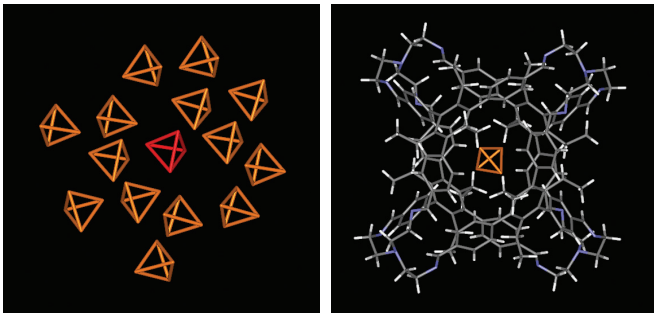

Fig. 2 – Two examples of tetrahedral P_4 molecules within coordination spheres. Left: central P_4 molecule (red) surrounded by other P_4 molecules (orange) in the crystal structure of white phosphorus (ICSD: 68326).[58] Right: P_4 molecule (orange) surrounded by a polyphenyl-polyamine-based supramolecule (C – grey, N – blue, H – white) in the crystalline state (CSD: PELSOR).[65]

Slika 2 – Dva primjera tetraedrijskih molekula P_4 unutar koordinacijskih sfera. Lijevo: središnja molekula P_4 (crveno) okružena drugim molekulama P_4 (narančasto) u kristalnoj strukturi bijelog fosfora (ICSD: 68326).[58] Desno: molekula P_4 (narančasto) okružena polifenolno-poliaminskom supramolekulom (C – sivo, N – plavo, H – bijelo) u kristalnom stanju (CSD: PELSOR).[65]

Among the Zintl phases cited in Table 1, one compound exhibits semi-metallic properties,[34] a few Zintl phases are useful in synthesis of diverse materials,[36-38] and some Zintl phases may serve as precursors for nanocrystalline elements, such as Si, Ge, and Pb.[37,56] Zintl phases have generally gained a great deal of attention due to their versatile properties and reactivity, leading to diverse applications:[87–90] synthesis of new materials (compounds, crystal structures, allotropes; and nanostructured materials); obtaining new materials through solution chemistry; catalysis; use as suitable thermoelectric, magnetic, insulator and semi-metal materials for various purposes; use in electrochemistry and photovoltaic technology, in energy storage and energy conversion technologies; surface modification of solids, and more.

Tetrahedral species of Group 15 include tetrahedra of Sb and Bi mainly in gas phase, and diverse tetrahedra of P and As (Table 1). Molecules P_4 and As_4 exist in several forms besides the gas phase form: in crystalline phase, in thin films, in cage or intercalated compounds, and some other forms. Capturing P_4 molecules by organic container molecules (supramolecular and polymeric)[61,63,65,66] transforms white phosphorus into non-toxic forms that are air-stable and water-soluble. In other words, this is an efficient way to store white phosphorus, especially for various chemical syntheses. Encapsulation of white phosphorus molecules P_4 by nanostructures, such as single-wall carbon nanotubes[67] and C_{60} fullerene molecules[69], is another way to obtain a stable form of phosphorus. Figure 2 shows two tetrahedral forms of phosphorus, in crystal structures obtained from the *Access Structures for the CSD and ICSD databases*,[85] and presented using the software *ViewerLite*:[86] P_4 in white phosphorus[58] and P_4 captured by an organic tetrahedral cage,[65] are both shown with their coordination spheres within van der Waals contacts.

Capturing As_4 molecules of yellow arsenic by organic supramolecular containers[63] resolves the problems of arsenic use in synthesis: toxicity, instability, and rapid polymerisation to grey arsenic which is unsuitable for chemical reactions. Single-wall carbon nanotubes have also shown to be effective in encapsulating As_4 molecules of yellow arsenic.[63] It is interesting to note that the carbon nanotubes act as catalysts in obtaining other, even new allotropes of arsenic or phosphorus.[67,73]

Tetrahedral forms of the twelve chemical elements (Table 1) are of technological importance. Studying them is not just a matter of theoretical considerations; it implies gaining a better understanding of the structures, properties, and reactivity of materials containing tetrahedral homonuclear species, which can have potential applications in chemical technology.

Table 1 contains the average tetrahedron d_{XX} bond lengths for a particular tetrahedral form, based on the references used. The d_{XX} values may aid in understanding how size of tetrahedra correlates with other properties of the elements across the periodic table. Notably, these values increase along the groups, and decrease along the periods, as expected. The trends are not always regular (compare Al and Ga). Small variations of the d_{XX} values occur among different forms of the same element.

3.2 Dataset of selected molecular descriptors

According to Table 1, it is clear that the first three groups of the *p*-block are characterised by tetrahedral species: all post-transitional metals (Al, Ga, In, Tl, Sn, Pb, Bi), some metalloids (Si, Ge, As, Sb), and a non-metal (P). What do all these elements have in common? A simple octet rule

may explain the need for three chemical bonds per atom in a tetrahedron species: 5 valence electrons in Group 15 are enough to form neutral tetrahedra, but only 4 valence electrons in Group 14 are not sufficient – 4 more electrons are necessary so that anions with charge 4− are formed. Furthermore, Group 13 has 3 valence electrons, meaning that 8 more electrons must be gained to form a tetrahedral anion with charge 8−. What about the neighbouring groups? The right neighbour, Group 16, contains more electronegative elements that should lose electrons so that a tetrahedral cation with charge 4+ may be formed, which is very unlikely. *Vernon*[91] has shown various relationships in the periodic table. For example, he pointed out the Be-Al diagonal relationship and the placement of Mg directly in front of Al. He also suggested the possibility of treating the main group elements as one block. As a result, for the left neighbour, comparable with Group 13, Group 2 of alkaline earth metals, an electropositive group, was considered. This group is not capable of capturing enough electrons to form a highly negative anion of charge 12−.

While the octet rule may appear quite reasonable, it is not adequate to rationalise the existence of stable elemental tetrahedra across the periodic table, including Groups 13–15. This is why descriptors of chemical elements for the following five groups were collected, and a dataset was created (Table 2): Q – charge of tetrahedral species (in electron units), χ – Allen electronegativity,[92] R_{at} – atomic radius[93] (at density cutoff 0.001 e Bohr^{-3}), χ_V – valence-state electronegativity,[94] D_e – experimental bond dissociation energy of diatomics,[94] R_0 – ionisation radius,[94] R_{vdW} – van der Waals radius,[95] R_{cov} – covalent radius,[96] χ_P – Pauling electronegativity,[97] and Class – the frequency and diversity class of tetrahedral species (carefully derived from the data in Table 1).

Table 2 – Descriptors of tetrahedral species: charges, elemental properties, and frequency and diversity class
Tablica 2 – Deskriptori tetraedrijskih vrsta: naboji, elementarna svojstva te razred učestalosti i raznolikosti

Species	Q	χ	R_{at}/Å	χ_V	D_e/kJ mol^{-1}	R_0/Å	R_{vdW}/Å	R_{cov}/Å	χ_P	Class
Be_4^{12-}	−12	1.576	2.19	5.63	59	1.09	1.53	0.99	1.57	1
Mg_4^{12-}	−12	1.293	2.40	2.60	9	2.35	1.73	1.40	1.31	1
Ca_4^{12-}	−12	1.034	2.70	1.88	15	3.26	2.31	1.74	1.00	1
Sr_4^{12-}	−12	0.963	2.79	1.60	16	3.83	2.49	1.90	0.95	1
Ba_4^{12-}	−12	0.881	2.93	1.37	10	4.48	2.68	2.06	0.89	1
B_4^{8-}	−8	2.051	2.05	3.79	290	1.62	1.92	0.84	2.04	1
Al_4^{8-}	−8	1.613	2.39	2.35	133	2.61	1.84	1.24	1.61	2
Ga_4^{8-}	−8	1.756	2.33	1.86	138	3.29	1.87	1.23	1.81	2
In_4^{8-}	−8	1.656	2.46	1.73	100	3.55	1.93	1.42	1.78	4
Tl_4^{8-}	−8	1.789	2.42	1.61	63	3.82	1.96	1.44	1.80	4
C_4^{4-}	−4	2.544	1.90	3.83	596	1.60	1.70	0.75	2.55	1
Si_4^{4-}	−4	1.916	2.32	2.56	310	2.40	2.10	1.14	1.90	5
Ge_4^{4-}	−4	1.994	2.34	2.09	274	2.94	2.11	1.20	2.01	4
Sn_4^{4-}	−4	1.824	2.48	1.88	195	3.26	2.17	1.40	1.96	5
Pb_4^{4-}	−4	1.854	2.49	1.77	87	3.47	2.02	1.45	1.80	5
N_4	0	3.066	1.79	3.93	857	1.56	1.55	0.71	3.04	1
P_4	0	2.253	2.23	2.79	485	2.20	1.80	1.09	2.19	5
P_4^+	1	2.253	2.23	2.79	485	2.20	1.80	1.09	2.19	2
As_4	0	2.211	2.31	2.26	382	2.62	1.85	1.20	2.01	5
As_4^+	1	2.211	2.31	2.26	382	2.62	1.85	1.20	2.01	2
Sb_4	0	1.984	2.46	2.04	299	3.01	2.06	1.40	2.05	3
Sb_4^+	1	1.984	2.46	2.04	299	3.01	2.06	1.40	2.05	2
Bi_4^+	1	2.01	2.50	1.92	197	3.19	2.07	1.50	1.90	2
O_4^{4+}	4	3.610	1.71	4.23	498	1.45	1.52	0.64	3.44	1
S_4^{4+}	4	2.589	2.14	2.99	425	2.05	1.80	1.04	2.58	1
Se_4^{4+}	4	2.424	2.24	2.56	331	2.40	1.90	1.18	2.55	1
Te_4^{4+}	4	2.158	2.42	2.18	258	2.81	2.06	1.37	2.10	1
Po_4^{4+}	4	2.19	2.50	2.06	186	3.14	1.97	1.42	2.00	1

Table 3 – Correlation matrix (top and right), and probability matrix (bottom and left) for the dataset from Table 2
Tablica 3 – Korelacijska matrica (gore i desno) i vjerojatnosna matrica (dolje i lijevo) za skup podataka iz tablice 2

	Q	χ	R_{at}	χ_V	D_e	R_0	R_{vdW}	R_{cov}	χ_P	Class
Q		**0.7898**	**−0.4246**	0.0737	**0.6551**	−0.2849	−0.3304	**−0.4162**	**0.7643**	0.0281
χ	**<0.0001**		**−0.8622**	**0.5001**	**0.8478**	**−0.6398**	**−0.6913**	**−0.8229**	**0.9881**	−0.0496
R_{at}	**0.0243**	**<0.0001**		**−0.7788**	**−0.7865**	**0.8719**	**0.8651**	**0.9815**	**−0.8690**	0.1384
χ_V	0.7094	**0.0067**	**<0.0001**		**0.4702**	**−0.9227**	**−0.7409**	**−0.7877**	**0.4972**	−0.3538
D_e	**0.0002**	**<0.0001**	**<0.0001**	**0.0116**		**−0.6364**	**−0.5571**	**−0.7615**	**0.8443**	−0.0426
R_0	0.1417	**0.0002**	**<0.0001**	**<0.0001**	**0.0003**		**0.8235**	**0.8971**	**−0.6342**	0.2722
R_{vdW}	0.0859	**<0.0001**	**<0.0001**	**<0.0001**	**0.0021**	**<0.0001**		**0.8690**	**−0.6857**	0.1028
R_{cov}	**0.0276**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**		**−0.8296**	0.0663
χ_P	**<0.0001**	**<0.0001**	**<0.0001**	**0.0071**	**<0.0001**	**0.0003**	**0.0001**	**<0.0001**		−0.0382
Class	0.8871	0.8021	0.4825	0.0647	0.8296	0.1611	0.6027	0.7375	0.8470	

By inspecting correlations among these 10 descriptors in terms of the Pearson correlation coefficients and corresponding probabilities (Table 3), it becomes clear that the dataset is neither homogeneous (all high positive or negative correlations) nor heterogeneous (all low correlations). Variable Class has only low and statistically insignificant correlations, while all correlations among the other nine variables are moderate to high and statistically significant. Therefore, these variables should undergo procedures of data compression *via* dimensionality reduction.

3.3 Hierarchical Cluster Analysis

Selecting the best linkage in Hierarchical Cluster Analysis (HCA) is usually a trial-and-error method.[98] Among various methods (single, incremental, complete, centroid, median, group average, and flexible) available in *Pirouette*[19], the best clustering pattern of samples (*i.e.*, of the tetrahedra from Table 2) was obtained using the complete linkage clustering, known also as the farthest neighbour method (Fig. 3). Distances between cluster pairs in the complete linkage are calculated as the distances between the two farthest samples in the two clusters. The complete linkage has two peculiar characteristics:[17] generation of more compact clusters, and the maximum sensitivity to outliers.

Chemical interpretation of a HCA plot (dendrogram) for samples is possible when clusters and sub-clusters are inspected together with their similarity indices, and chemical characteristics (structural features, substance properties, reactivity, *etc.*) are assigned to samples in the dendrogram using suitable marks, colours, and symbols.[13,23–25] In Fig. 3, P*, As*, Sb*, and Bi* denote cations P_4^+, As_4^+, Sb_4^+ and Bi_4^+, respectively. Other species are E_4^{8-}, E_4^{4-}, and E_4, same as in Table 1. The d_{XX} values from Table 1 were used with less precision in Fig. 3, and for other elements, known bond lengths[99] were added with the same precision.

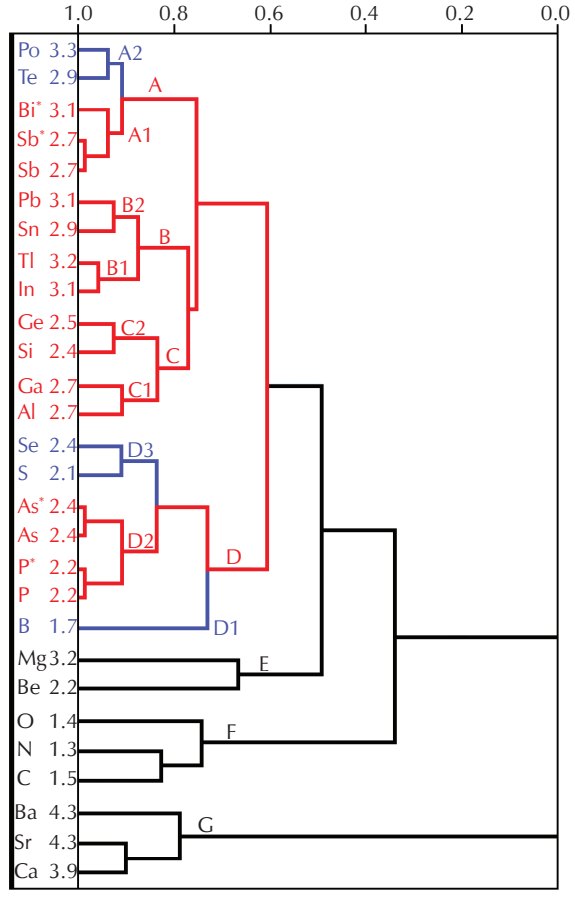

Fig. 3 – Dendrogram of samples (tetrahedral species) with complete linkage: red – real tetrahedra, blue – fictive tetrahedra not well-separated from the real tetrahedra, black – fictive tetrahedra well-separated from the real tetrahedra
Slika 3 – Dendrogram uzoraka (tetraedrijskih vrsta) s potpunom povezanošću: crveno – pravi tetraedri, plavo – fiktivni tetraedri koji nisu dobro razdvojeni od pravih tetraedara, crno – fiktivni tetraedri koji su dobro razdvojeni od pravih tetraedara

There are seven main clusters in the HCA dendrogram (clusters A–F, Fig. 3). Red clusters and sub-clusters are for elements having true tetrahedral species (Table 1), blue sub-clusters are for elements that are not well separable from the red sub-clusters (B, S, Se, Te, Po), and the black clusters are for other elements (C, N, O and Group 2). The blue sub-clusters and black clusters are fictive tetrahedra. The more similar the tetrahedra are in terms of their d_{xx} values, the more defined sub-clusters they tend to form indicated by higher similarity indices. This is why several two-membered sub-clusters appear at similarity indices above 0.9. Tetrahedra of Group 13 and Group 14 elements are mixed, so that heavier elements appear in cluster B as sub-clusters B1 (Group 13) and B2 (Group 14), and lighter elements are in cluster C as sub-clusters C1 (Group 13) and C2 (Group 14). This fact aligns with Table 1, where two groups are known for their Zintl phase anions. However, Group 15 elements are divided into two parts within clusters A and D: sub-cluster A1 contains heavier elements, while lighter elements are found in sub-cluster D2.

Why are blue sub-clusters not well-separated from the red ones, contrary to expectations (Fig. 3)? Firstly, HCA dendrogram is only a one-dimensional presentation of similarities and differences among samples, meaning that, in some cases, distinction among samples cannot be visualised. Secondly, there are certain structural similarities between the red and blue sub-clusters within the same cluster. Group 16 contains stable homonuclear tetrameric species: $S_4^{2+}$and Se_4^{2+} (square planar),[100,101] S_4 (trigonal),[102] Te_4^{2+} (square planar),[101,103] Se_4^{2-} and Se_4^{2-} (bent)[101] species (Po probably forms such species). This fact explains why Group 16 is divided into heavy elements (sub-cluster A2) and light elements (sub-cluster D3), and is mixed with Group 15. Furthermore, boron forms B_4 tetrahedra within simple molecules, such as B_4Cl_4 and $B_4{}^tBu_4$.[104–106] There are also free rhombus-shaped and square planar homonuclear species of B (B_4, B_4^+, B_4^-, B_4^{2-}).[107] This situation with boron explains why sub-cluster D1 is not well distinguished from sub-cluster D2.

Clusters A, B, C, and D make a macro-cluster at similarity index 0.61 (Fig. 3). The d_{XX} values for existing tetrahedral species (the red sub-clusters) range from 2.2 to 3.2 Å, which can be considered as a condition for existence of stable true tetrahedral species. The second empirical role is that small two-membered red sub-clusters have differences in d_{XX} values ranging from 0 to 0.2 Å, while blue sub-clusters have somewhat greater differences (0.3 and 0.4)

The black clusters are well separated from the macro-cluster and from each other (Fig. 3). The smallest d_{XX} values belong to cluster F, to the smallest atoms of C, N, and O. Tetrameric homonuclear species are known for these elements: tetracarbon C_4 (square planar and linear),[108,109] C_4^- (square planar),[109] tetranitrogen N_4 and N_4^+ (bent structures),[110] metastable tetraoxygen O_4 and O_4^+ (square planar structures),[111] and octaoxygen O_8 which appears to consist of two parallel square planar structures.[112] Carbon tetrahedron exists within a molecule of $C_4{}^tBu_4$.[113] Group 2 is divided into cluster E (light elements, hexagonal crystal structures[109]), and cluster G (heavy elements, face-centred and body-centred cubic structures[109]). The only "tetrahedra" to be discussed in such cases, are numerous imaginary species formed from four atoms belonging to two or three different layers of the crystal structure of a pure metallic element.

Group 17 was not included in this study. Two heavier elements, Br and I, are known for their tetrameric species: Br for Br_4^{2-} (linear),[114] and I for I_4^{2+} (square planar)[115,116] and I_4^{2-} (trigonal).[117] If included in the dataset, fictive tetrahedra of Br and I would probably mix with the red sub-clusters from Fig. 3.

3.4 Principal component analysis

Table 4 contains the basic statistics for Principal Component Analysis (PCA). For each number of principal components (PC) used, five quantities are shown: variance, percent variance, cumulative percent variance, VPRESS for validation, and PRESS for data reconstruction, using the selected number of PCs. The first three principal components, PC1, PC2, and PC3, which account for over 94 % of the total variance, were selected as sufficient for a model of reasonable quality and simplicity in terms of chemical in-

Table 4 – Basic PCA statistics for selection of essential principal components
Tablica 4 – Osnovna statistika PCA za odabir esencijalnih glavnih komponenata

PC	Variance	Percent variance	Cumulative percent variance	VPRESS	PRESS
PC1	181.082169	74.519409	74.519409	78.720467	61.917835
PC2	40.131596	16.515060	91.034470	36.402351	21.786240
PC3	**8.373278**	**3.445793**	**94.480263**	**28.569954**	**13.412962**
PC4	6.811586	2.803122	97.283386	16.357193	6.601376
PC5	3.928029	1.616473	98.899857	9.576053	2.673347
PC6	1.838977	0.756781	99.656639	3.463926	0.834369
PC7	0.480591	0.197774	99.854416	2.772244	0.353778
PC8	0.286725	0.117994	99.972412	0.304943	0.067052
PC9	0.067074	0.027602	100.000015	0.000000	0.000000

terpretation of loadings and scores. Further PCs make only a small contribution to the variance (below 3 %) and lead to model overfitting both in reconstruction and cross-validation, evident in reduced VPRESS and PRESS values.

Chemical interpretation of PCA results for selected PCs can be performed using two- or three-dimensional scores plots, loading plots or the loadings matrix, and marks and colours for various chemical characteristics of samples in the plots, as well as by labelling HCA clusters or some other groups in the PCA plots.[12,13,20–25]

The scores plot in three dimensions (Fig. 4) features tetrahedra with their Class value (Table 2): red (1 – for non-existing or fictive tetrahedra), green (2), blue (3), pink (4), and black (5). A higher Class value indicates greater frequency and diversity of materials in which the tetrahedra were identified. The scores plot reveals significant differences compared to the HCA dendrogram (Fig. 3). Firstly, all real tetrahedra (green, blue, pink, and black) are separated from all fictive tetrahedra (red), while in the HCA dendrogram, some mixing occurs. Secondly, the space of the real tetrahedra may be contoured by an ellipse-like shape in the plane of the projection; the real tetrahedra are close to Group 16 and boron. Thirdly, mixing within the ellipse-like shape is minimal – Groups 13 and 14 are not mixed as in HCA, except in one case – Pb is positioned among Group 13 elements.

The ellipse-like shape in Fig. 4 exhibits several distinctions among the tetrahedra, which are not apparent in HCA, and cannot be explained by the simple octet rule. There are three clusters in Fig. 4 separated by straight brown lines: Group 13 with Pb, Group 14 without Pb, and Group 15. Furthermore, Group 15 contains tetrahedra with no charge or small positive charge (1+), and the other two clusters comprise tetrahedra with high negative charges (4− and 8−). With respect to the Class values, there is some regularity within the ellipse-like shape: low Class values (green and blue) occupy mainly the central part of the shape, and high Class values (pink and black) are divided into left and right branches. There is also some intermixing of these areas around species of As and P. It is interesting to note that the metal – metalloid – non-metal character can also be observed within the ellipse-like shape and its surroundings. Species of metalloids As, Sb, Ge, and Si form a diagonal line in purple which continues to include red Te, Se, and B. To the right of the metalloid species, the non-metallic P species create a continuous space with other red non-metal species. To the left of the metalloid species are metal species of Sn, Pb, and metals of Group 13, and outside of the ellipse-like shape are species of other red metals of Group 2 and Po.

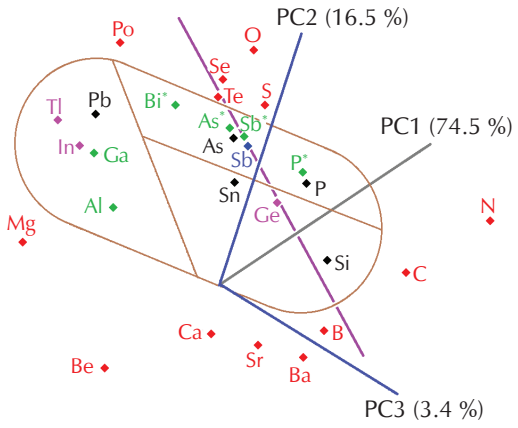

Fig. 4 – Tetrahedral species in the space defined by the first three principal components, coloured according to their Class values from 1 to 5: 1 – red, 2 – green, 3 – blue, 4 – pink, and 5 – black. Real tetrahedra are grouped in three parts of the ellipse-like shape. Tetrahedra of metalloids are positioned along the purple diagonal line.

Slika 4 – Tetraedrijske vrste u prostoru definiranom prvim trima glavnim komponentama, s bojanjem prema pripadnim Class vrijednostima od 1 do 5: 1 – crveno, 2 – zeleno, 3 – plavo, 4 – ružičasto i 5 – crno. Pravi tetraedri su grupirani u tri dijela elipsi sličnog lika. Tetraedri metaloida nalaze se uzduž purpurnog dijagonalnog pravca.

The scores plot (Fig. 4) and its chemical meaning confirm that indeed three principal components are required to describe the data from Table 2. How can we understand the PCA scores in terms of the original descriptors used? Table 4 contains the loadings matrix for the first three principal components. Loadings for each PC are divided into two groups: PC1 – electronic (italics), and steric (bold) descriptors; PC2 – valence state (bold-underlined), and other (plain) descriptors; PC3 – non-bonding (plain-underlined), and bonding interaction (italics-bold) descriptors. PC1 is a general PC, a stereoelectronic property, taking into account atomic size and electronic content. PC2 is a property which distinguishes valence state from ground state descriptors; in other words, PC2 accounts for electron acceptance/release or reduction/oxidation property. The meaning of PC3 according to the loadings (Table 5) is rather uncertain, due to its small contribution to the total variance (3.4 %). Two-dimensional scores plots may aid in understanding the nature of this PC.

Table 5 – Loadings for the selected principal components and descriptor types

Tablica 5 – Koeficijenti odabranih glavnih komponenata i vrste deskriptora

Descriptor	PC1 (74.5 %)	PC2 (16.5 %)	PC3 (3.4 %)
Q	*0.229113*	0.600628	***−0.217356***
χ	*0.355625*	0.288983	***−0.094717***
χ_V	*0.292143*	**−0.471973**	***0.190372***
χ_P	*0.354645*	0.279951	***−0.075934***
D_e	*0.327023*	0.245703	0.672473
R_{at}	**−0.375484**	0.110281	***−0.014299***
R_0	**−0.337997**	0.329011	***−0.142657***
R_{vdW}	**−0.331101**	0.231474	0.654964
R_{cov}	**−0.371639**	0.142279	***0.004964***

Two-dimensional PCA scores plots (Figs. 5–7, with the same colouring as in Fig. 4), illustrate the meanings of the PCs in terms of positions and grouping of samples (tetrahedral species). Green confidence ellipses represent the 95 % confidence level. Trends observed in Fig. 4 are not so well visible in Figs. 5–7 due to the loss of one PC.

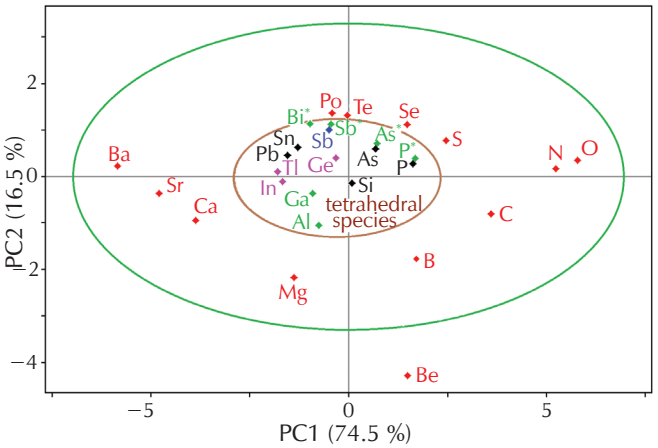

Fig. 5 – Tetrahedral species in the space defined by principal components PC1 and PC2, coloured according to their Class values from 1 to 5: 1 – red, 2 – green, 3 – blue, 4 – pink, and 5 – black

Slika 5 – Tetraedrijske vrste u prostoru definiranom glavnim komponentama PC1 i PC2, s bojanjem prema pripadnim Class vrijednostima od 1 do 5: 1 – crveno, 2 – zeleno, 3 – plavo, 4 – ružičasto i 5 - crno

In the PC1-PC2 space (Fig. 5), real tetrahedral species are well grouped within a central brown ellipse, clearly separated from all fictive tetrahedra. Upon a detailed analysis of Fig. 5, one can derive a condition for real tetrahedra, PC2 ≤ 1.25, and another condition derived from the inequality of the ellipse: PC1 + 4PC2 ≤ 6.34. Although the ellipse was drawn arbitrarily, it is useful for qualitative considerations. For a more rigorous and quantitative approach, many chemical elements should be included in PCA to construct an exact ellipse. When analysing the positions of samples with respect to the coordinate axes, it is evident that smaller atoms, which are more electronegative and non-metallic elements, are located at high positive values of PC1. The positive end of PC1 is also characterised by positive contributions of electronegativity descriptors, charge Q, and dimer dissociation energy D_e to PC1 (Table 5). In other words, high values of these descriptors describe the elements at the positive end of PC1. In contrast, at high negative values of PC1, there are larger atoms, more electropositive metallic elements, because this end of PC1 is determined by negative contributions of radii descriptors to PC1 (Table 5), meaning that atoms with small radii are placed at this end of PC1. Practically the same trend along PC1 exists within the central ellipse, clearly visible because PC1 contains the largest fraction of total variance (74.5 %). Tetrahedra of metalloids are positioned in the central and upper part of the small ellipse. Samples (tetrahedral species) at negative values of PC2 are usually of electron-deficient elements that either receive electrons in forming covalent bonds (B, C, Be) or tetrahedral anions (Al, Ga, In), or ionise (Mg, Ca, Sr, and Be as an anomalous sample). This trend is caused by the high negative contribution of the valence-state electronegativity χ_V to PC2 (Table 6); thus, the mentioned elements are well-described by greater values of χ_V. Tetrahedra at the positive end of PC2 are of elements which are mainly richer in valence electrons, so that they do not receive electrons to form tetrahedra (Sb, As, P) or other covalent structures (Group 16). The loadings matrix (Table 6) shows positive contribution of most descriptors, especially of charge Q to PC2. In other words, high values of these descriptors characterise the mentioned species at the positive end of PC2. PC2 contains a smaller fraction of total variance (16.5 %), which is why the samples' trends along it are somewhat less clear than in the case of PC1. Interesting trends can be noticed when inspecting whether the chemical elements along PC1 and along PC2 follow the same order as in their respective groups in the periodic table, *i.e.,* in the sense of increasing atomic numbers. Along PC1, from its positive end toward its negative end, all elements follow the same order as in the periodic table. The same can be said for PC2, when observing the elements from the negative end of PC2 to its positive end. Such arrangement of the elements means that Groups 2, 13–16 are positioned in diagonal directions, almost mutually parallel.

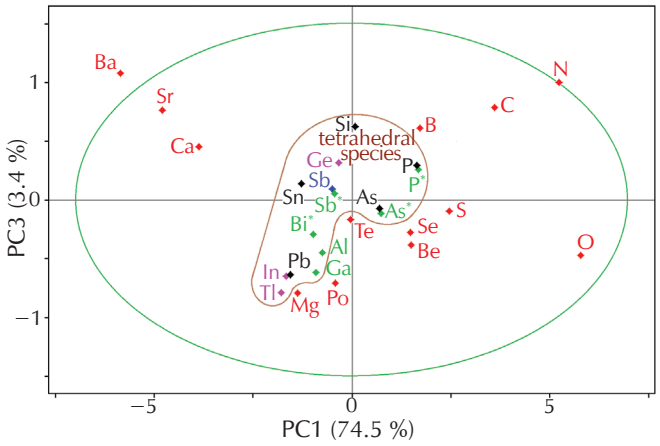

Fig. 6 – Tetrahedral species in the space defined by principal components PC1 and PC3, coloured according to their Class values from 1 to 5: 1 – red, 2 – green, 3 – blue, 4 – pink, and 5 – black

Slika 6 – Tetraedrijske vrste u prostoru definiranom glavnim komponentama PC1 i PC3, s bojanjem prema pripadnim Class vrijednostima od 1 do 5: 1 – crveno, 2 – zeleno, 3 – plavo, 4 – ružičasto i 5 - crno

Existing tetrahedral species in the PC1-PC3 space (Fig. 6) form a retort-shaped space, which is obviously less regular than the ellipse in Fig. 5. It is impossible to determine a condition for real tetrahedra from this figure. In Fig. 7, the samples are situated in a complicated kidney-shaped space, with mixing of real and red tetrahedra to its right. Trends along PC2, as mentioned previously for Fig. 5, are somewhat clearer in Fig. 7, because PC2 is now the horizontal axis defining the direction of the major axis of the

confidence ellipse. Notably, descriptors that account for non-bonding interactions (D_e and R_{vdW}) show high positive contributions to PC3, while other descriptors show small positive and negative contributions to this PC (Table 6). PC3 contains a small fraction of total variance (only 3.4 %), and trends along it are not clear in the scores plots. It may be said that PC3 discriminates elements with respect to the size of species that the elements form in compounds, in a very general sense. The size of species is always related to the intensity of non-bonding interactions. Samples at high positive values of PC3 represent two types of elements which form greater species (Figs. 6 and 7): very large cations (Ba, Ca, Sr), and complicated molecules with several homonuclear covalent bonds (B, C, N). At negative values of PC3 are elements that form smaller species, smaller cations (Mg, Tl, Po, In, Pb, Ga). Group 16 elements build less complicated molecules (only two covalent bonds per atom); thus, corresponding samples are placed at less negative values of PC3. Contrary to PC1 and PC2, PC3 introduces some distinction among the five groups of chemical elements. Moving from the positive to the negative end of PC3, Groups 13–15 maintain their order as in the periodic table, with the exception of Sb and As, which have swapped places. However, Group 2 and Group 16 are oriented contrary to their order in the periodic table, exhibiting some irregularities such as Mg and Be swapping places and O appearing between Se and Po. Therefore, PC3 can serve as a discriminating factor between Groups 13–15 (where real tetrahedra exist) and Groups 2 and 16 (where real tetrahedra do not exist).

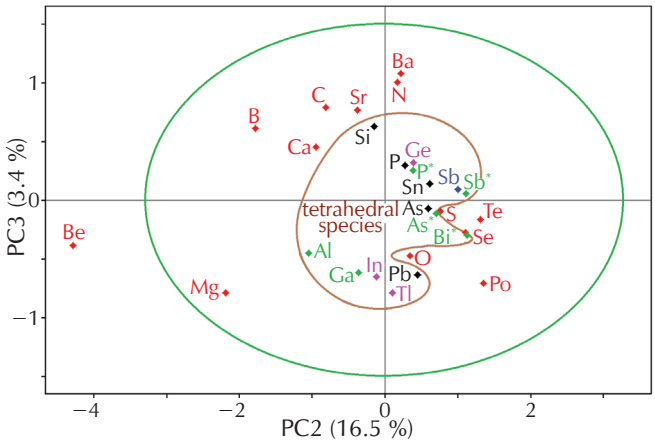

Fig. 7 – Tetrahedral species in the space defined by principal components PC2 and PC3, coloured according to their Class values from 1 to 5: 1 – red, 2 – green, 3 – blue, 4 – pink, and 5 – black

Slika 7 – Tetraedrijske vrste u prostoru definiranom glavnim komponentama PC2 i PC3, s bojanjem prema pripadnim Class vrijednostima od 1 do 5: 1 – crveno, 2 – zeleno, 3 – plavo, 4 – ružičasto i 5 – crno

3.5 Parabolic regression modelling

Table 3 shows that variable Class has statistically no significant correlations with the other nine variables. The relationships between Class and other descriptors were reinvestigated for eventual non-linearities, with special emphasis on Groups 13–15. The same was done with the relationship between Class and the selected PCs. The results are summarised in Table 6, and they are surprising and unexpected. Parabolic (second-order polynomial) regression models $y = a + bx + cx^2$ with estimated standard deviations σ_a, σ_b, and σ_c for regression coefficients, were constructed, and showed some statistical significance for relationships between Class and eight descriptors and PC1, only for

Table 6 – Parabolic regression modelling of PC1 and descriptors as functions of variable Class
Tablica 6 – Paraboličnо regresijsko modeliranje PC1 i deskriptora kao funkcija varijable Class

Descriptor	r	F	p	c, σ_c	b, σ_b	a, σ_a	t_c, p_c	t_b, p_b	t_a, p_a
χ	**0.6696**	**6.1**	**0.0115**	3.228 0.353	−0.840 0.271	0.120 0.043	**9.135** **<0.0001**	**−3.094** **0.0074**	**2.800** **0.0134**
χ_V	**0.8917**	**29.1**	**<0.0001**	5.735 0.461	−2.355 0.354	0.334 0.056	**12.446** **<0.0001**	**−6.656** **<0.0001**	**5.986** **<0.0001**
χ_P	**0.6699**	**6.1**	**0.0115**	3.126 0.330	−0.771 0.253	0.109 0.040	**9.474** **<0.0001**	**−3.042** **0.0082**	**2.733** **0.0154**
D_e	**0.6429**	**5.3**	**0.0183**	967.76 213.28	−469.70 163.81	66.92 25.83	**4.538** **0.0004**	**−2.867** **0.0117**	**2.590** **0.0205**
R_{at}	**0.8442**	**18.6**	**<0.0001**	1.445 0.150	0.606 0.115	−0.086 0.018	**9.636** **<0.0001**	**5.262** **<0.0001**	**−4.711** **0.0003**
R_0	**0.8191**	**15.3**	**0.0002**	−0.118 0.542	2.024 0.416	−0.288 0.066	-0.219 0.8299	**4.865** **0.0002**	**−4.391** **0.0005**
R_{vdW}	**0.6257**	**4.8**	**0.0241**	1.477 0.172	0.293 0.132	−0.039 0.021	**8.586** **<0.0001**	*2.221* *0.0421*	−1.852 0.0839
R_{cov}	**0.8317**	**16.8**	**0.0001**	0.193 0.183	0.724 0.141	−0.103 0.022	1.054 0.3087	**5.136** **0.0001**	**−4.647** **0.0003**
PC1	**0.7984**	**13.2**	**0.0005**	7.883 1.575	−5.368 1.210	0.758 0.191	**5.005** **0.0002**	**−4.438** **0.0005**	**3.976** **0.0012**

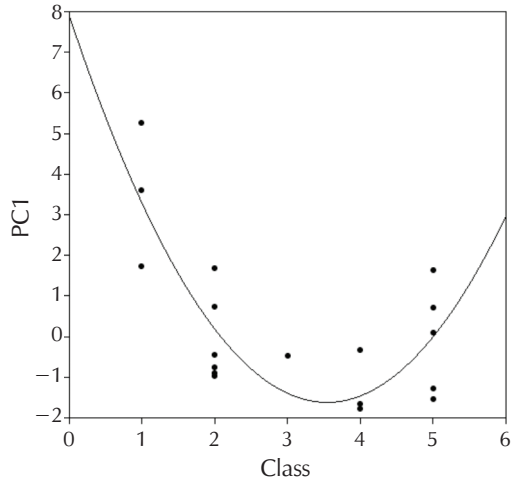

Fig. 8 – Parabolic regression model for PC1 as a function of variable Class

Slika 8 – Parabolični regresijski model za PC1 kao funkciju varijable Class

Groups 13–15. Using other PCs and Groups 2 and 16 did not yield sensible non-linear models. It was not expected that Class would be the independent variable, but on the contrary, the modelled variable. *F*-test parameters are: *r* – Pearson correlation coefficient, *F* – *F*-ratio, and *p* – corresponding probability. Student test parameters: t_a, t_b, and t_c are *t*-parameters for *a*, *b*, and *c*, respectively, and p_a, p_b, and p_c are the corresponding probabilities. Values of *t*-parameters and probabilities in Table 6 are typed bold in case of statistical significance, and in italics in case of borderline statistical significance.

All nine parabolic models are statistically very or extremely significant in the *F*-test (Table 6). However, six out of nine models are statistically very or extremely significant in the *t*-test for all regression coefficients: the models for three electronegativity variables and the models for PC1, atomic radius R_{at} and for bond dissociation energy D_e. The models for the other three radii descriptors fail in the *t*-test for only one regression coefficient. These facts indicate that non-linear relationships between Class and the nine descriptors are important and thus, cannot be ignored. All radii descriptors as functions of Class have maxima, while other descriptors and PC1 have minima (Fig. 8). *Scilab*[14] was used to plot the data points for Class, PC1, and for the squared function from regression analysis. Upon detailed analysis of Fig. 8, one obtains a condition for PC1 for real tetrahedra: PC1 ≤ 1.69, meaning that Class > 1.

Class was generated from data in Table 1, and was based on literature: Class is a measure of both frequency and diversity of tetrahedral forms reported in literature. However, its modelled relationships (Table 6) indicate that Class reflects some intrinsic, basic property for Groups 13–15, well-correlated to the general principal component PC1 and eight descriptors, and not correlated to PC2 and PC3. The Class – PC1 relationship includes in fact three quantities - Class, square of Class, and PC1, which again confirms that the real tetrahedra are well-described in a three-dimensional space.

4 Conclusion

This work deals with identification of chemical elements that occur as stable tetrahedral homonuclear tetramers (neutral molecules and ions) and their forms of appearance. Furthermore, rationalisation of tetrahedral structures and properties with certain limits to distinguish these elements from the other elements is given. A database for tetrahedral homonuclear species was formed with a dataset of elemental descriptors. The dataset was then analysed using correlation analysis, hierarchical cluster analysis, principal component analysis, and parabolic regression with appropriate statistical tests. The results can be summarised as follows.

1) In literature, tetrahedral species were experimentally identified for 12 chemical elements. These elements are from Groups 13–15 and from Periods 3-6: Al, Ga, In, Tl, Si, Ge, Sn, Pb, P, As, Sb, and Bi.
2) These elements compose 39 forms of tetrahedral occurrence (gas phase, solid Zintl phase, thin films, Zintl phase ammoniates, cage compounds, and other).
3) A simple octet rule for *p*-elements explains that only Groups 13–15 with 3–5 valence electrons may form such tetrahedra.
4) Hierarchical cluster analysis distinguishes well the 12 elements from Group 2, C, N, and O, but does not separate them from Group 16 and B.
5) The average tetrahedral bond length should be 2.2 Å < d_{XX} < 3.2 Å.
6) Principal component analysis with parabolic regression yield approximate conditions for principal components: PC1 ≤ 1.69, PC2 ≤ 1.25, PC1 + 4PC2 ≤ 6.34.
7) Principal component analysis with two and three principal components explains rather well various properties and groupings of the tetrahedra. PC3 discriminates Groups 13–15 from Groups 2 and 16 in terms of the elements' order with respect to their order in the periodic table.
8) The tetrahedra are shown to be a three-dimensional phenomenon, because three quantities are required to describe the set of tetrahedra.
9) The frequency and diversity Class is shown to reflect a basic property on which PC1 and eight descriptors depend in parabolic regression, which should be further investigated.
10) Class, its square, and PC1 reconfirm that real tetrahedral species are a three-dimensional phenomenon.

Many of the tetrahedral forms belong to materials of diverse technological importance. Therefore, gaining a deeper understanding of the structures and properties of these tetrahedra at a more advanced level is crucial, and more such studies should be conducted in the future.

List of abbreviations and symbols
Popis kratica i simbola

a – free regression coefficient in parabolic regression
– slobodan regresijski koeficijent u paraboličnoj regresiji

α – significance level
– razina značajnosti

$\mathbf{A}$ – matrix obtained as $\mathbf{X}^{\mathrm{T}}\boldsymbol{y}$ in regression analysis
– matrica dobivena kao $\mathbf{X}^{\mathrm{T}}\boldsymbol{y}$ u regresijskoj analizi

b – regression coefficient for the linear independent variable in parabolic regression
– regresijski koeficijent za linearnu nezavisnu varijablu u paraboličnoj regresiji

$\mathbf{B}$ – matrix obtained as $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ in regression analysis
– matrica dobivena kao $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ u regresijskoj analizi

c – regression coefficient for the squared independent variable in parabolic regression
– regresijski koeficijent za kvadrat nezavisne varijable u paraboličnoj regresiji

$\boldsymbol{c}$ – vector of regression coefficients in regression analysis, obtained as the product **DA**
– vektor regresijskih koeficijenata u regresijskoj analizi, dobiven kao produkt **DA**

c_i – i-th regression coefficient c_i in regression model
– i-ti regresijski koeficijent c_i u regresijskom modelu

Class – frequency and diversity class of tetrahedral species
– razred učestalosti i raznolikosti tetraedrijskih vrsta

CSD – Cambridge Structural Database, a crystallographic database for organic compounds
– kristalografska baza podataka za organske spojeve Cambridge Structural Database

$\mathbf{D}$ – matrix obtained as $\mathbf{B}^{-1}$ in regression analysis
– matrica dobivena kao $\mathbf{B}^{-1}$ u regresijskoj analizi

Δ – vector of deviations of observed values of the dependent variable $\boldsymbol{y}$ from its mean value
– vektor odstupanja opaženih vrijednosti zavisne varijable $\boldsymbol{y}$ od svoje srednje vrijednosti

d_{AB} – Euclidean distance between clusters A and B in HCA
– euklidska udaljenost između klastera A i B u HCA

D_{e} – experimental bond dissociation energy of diatomics
– eksperimentalna energija disocijacije veze diatomskih molekula

df – degrees of freedom
– broj stupnjeva slobode

d_{ii} – i-th element of the main diagonal of the matrix **D** in regression analysis
– i-ti element glavne dijagonale matrice **D** u regresijskoj analizi

d_{max} – maximum Euclidean distance among clusters in HCA
– najveća euklidska udaljenost među klasterima u HCA

Δ^{T} – transpose of the vector Δ
– transponirani vektor vektora Δ

d_{XX} – average bond length in a tetrahedral species
– prosječna duljina veze u tretraedrijskoj vrsti

$\hat{\mathbf{E}}$ – residual matrix in PCA
– rezidualna matrica u PCA

$\hat{\boldsymbol{\varepsilon}}$ – vector of residuals in regression analysis
– vektor reziduala u regresijskoj analizi

$\hat{\mathbf{e}}_{\mathrm{cv}i}$ – i-th row vector of the residuals from cross-validation in PCA, defined as $\boldsymbol{x}_i - \hat{\boldsymbol{x}}_{\mathrm{cv}i}$
– i-ti redčani vektor reziduala iz unakrsne validacije u PCA, definiran kao $\boldsymbol{x}_i - \hat{\boldsymbol{x}}_{\mathrm{cv}i}$

$\hat{\mathbf{e}}_{\mathrm{cv}i}^{\mathrm{T}}$ – transpose of the vector $\hat{\mathbf{e}}_{\mathrm{cv}i}$
– transponirani vektor vektora $\hat{\mathbf{e}}_{\mathrm{cv}i}$

$\hat{\mathbf{e}}_i$ – i-th row vector of the residuals matrix in PCA
– i-ti redčani vektor rezidualne matrice u PCA

$\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}$ – transpose of the vector of residuals in regression analysis
– transponirani vektor vektora reziduala u regresijskoj analizi

$\hat{\mathbf{e}}_i^{\mathrm{T}}$ – tranpose of the vector $\hat{\mathbf{e}}_i$
– transponirani vektor vektora $\hat{\mathbf{e}}_i$

E_4^{8-} – tetrahedral anions in Group 13
– tetraedrijski anioni u Grupi 13

E_4^{4-} – tetrahedral anions in Group 14
– tetraedrijski anioni u Grupi 14

E_4 – tetrahedral molecules in Group 15
– tetraedrijske molekule u Grupi 15

E_4^{+} – tetrahedral cations in Group 15
– tetraedrijski kationi u Grupi 15

F – F-ratio in F-test
– F-omjer u F-testu

F-test – Fisher's test
– Fisherov test

HCA – hierarchical cluster analysis
– hijerarhijska klasterska analiza

ICSD – Inorganic Crystal Structure Database, a crystallographic database for inorganic compounds
– kristalografska baza podataka Inorganic Crystal Structure Database za anorganske spojeve

k – number of principal components in data reconstruction in PCA; number of variables in regression analysis
– broj glavnih komponenata u rekonstrukciji podataka u PCA; broj varijabli u regresijskoj analizi

m – number of variables (descriptors) or the number of columns in a data matrix
– broj varijabli (deskriptora) ili broj stupaca u matrici podataka

n – number of samples (tetrahedral species) or the number of rows in a data matrix
– broj uzoraka (tetraedrijskih vrsta) ili broj redova u matrici podataka

NIPALS – nonlinear iterative partial least squares algorithm
– nelinearna iterativna metoda parcijalnih najmanjih kvadrata

Q – charge of a tetrahedral species (in electron units)
– naboj tetraedrijske vrste (u elektronskim jedinicama)

p – probability in F-test
– vjerojatnost u F-testu

p_{a} – probability in t-test for regression coefficient a
– vjerojatnost u t-testu za regresijski koeficijent a

p_{b} – probability in t-test for regression coefficient b
– vjerojatnost u t-testu za regresijski koeficijent b

p_{c} – probability in t-test for regression coefficient c
– vjerojatnost u t-testu za regresijski koeficijent c

PCA – principal component analysis
– analiza glavnih komponenata

PC – principal component
– glavna komponenta

PC1 – first principal component
– prva glavna komponenta

PC2 – second principal component
– druga glavna komponenta

PC3 – third principal component
– treća glavna komponenta

PDB – Protein Databank, a structural database for proteins and other macromolecules
– strukturna banka podataka za proteine i druge makromolekule Protein Databank

PRESS – predicted residual error sum of squares for reconstruction of the matrix **X** in PCA
– zbroj kvadrata prediktivnih rezidualnih pogrešaka u rekonstrukciji matrice **X** u PCA

r – Pearson correlation coefficient
– Pearsonov koeficijent korelacije

r^2 – coefficient of determination
– koeficijent determinacije

R_0 – ionisation radius

– ionizacijski radijus
R_{at} – atomic radius (at density cutoff 0.001 e Bohr^{-3})
– atomski radijus (kod granice gustoće od 0.001 e Bohr^{-3})
R_{cov} – covalent radius
– kovalentni radijus
R_{vdW} – van der Waals radius
– van der Waalsov radijus
S_{AB} – similarity index for a pair of clusters A and B in HCA
– indeks sličnosti za par klastera A i B u HCA
σ_a – estimated standard deviation for regression coefficient *a*
– procijenjeno standardno odstupanje za regresijski koeficijent *a*
σ_b – estimated standard deviation for regression coefficient *b*
– procijenjeno standardno odstupanje za regresijski koeficijent *b*
σ_c – estimated standard deviation for regression coefficient *c*
– procijenjeno standardno odstupanje za regresijski koeficijent *c*
σ^2 – estimated variance of residuals in regression analysis
– procijenjena varijancija reziduala u regresijskoj analizi
$\sigma(c_i)$ – estimated standard deviation for the *i*-th regression coefficients c_i
– procijenjeno standardno odstupanje za *i*-ti regresijski koeficijent c_i
SVD – singular value decomposition algorithm
– dekompozicija matrice na singularne vrijednosti
t-test – Student's test
– Studentov test
t – *t*-parameter for correlation coefficient *r*
– *t*-parametar za koeficijent korelacije *r*
t_a – *t*-parameter for regression coefficient *a* in parabolic regression
– *t*-parametar za regresijski koeficijent *a* u paraboličnoj regresiji
t_b – *t*-parameter for regression coefficient *b* in parabolic regression
– *t*-parametar za regresijski koeficijent *b* u paraboličnoj regresiji
t_c – *t*-parameter for regression coefficient *c* in parabolic regression
– *t*-parametar za regresijski koeficijent *c* u paraboličnoj regresiji
t_i – *t*-parameter for the *i*-th regression coefficient c_i in regression model
– *t*-parametar za *i*-ti regresijski koeficijent c_i u regresijskom modelu
VPRESS – predicted residual error sum of squares for validation in PCA
– zbroj kvadrata prediktivnih rezidualnih pogrešaka u validaciji u PCA
x – independent variable in a regression model
– nezavisna varijabla u regresijskom modelu
$\mathbf{X}$ – data matrix of independent variables in PCA or in regression analysis
– matrica podataka za nezavisne varijable u PCA ili u regresijskoj analizi
$\hat{\mathbf{X}}$ – reconstructed data matrix in PCA
– rekonstruirana matrica podataka u PCA
χ – Allen electronegativity
– Allenova elektronegativnost
$\hat{\mathbf{X}}_{cv}$ – average data matrix from crossvalidation in PCA
– prosječna matrica podataka iz unakrsne validacije u PCA
$\hat{\boldsymbol{x}}_{cvi}$ – *i*-th row vector of the matrix $\hat{\mathbf{X}}_{cv}$ from crossvalidation in PCA
– *i*-ti redčani vektor matrice $\hat{\mathbf{X}}_{cv}$ iz unakrsne validacije u PCA
$\boldsymbol{x}_i$ – *i*-th row vector of the data matrix $\mathbf{X}$ in PCA
– *i*-ti redčani vektor matrice podataka $\mathbf{X}$ u PCA
$\hat{\boldsymbol{x}}_i$ – *i*-th row vector of the reconstructed data matrix $\hat{\mathbf{X}}$ in PCA
– *i*-ti redčani vektor rekonstruirane matrice podataka $\hat{\mathbf{X}}$ u PCA
χ_P – Pauling electronegativity
– Paulingova elektronegativnost
$\mathbf{X}^T$ – transpose of the matrix $\mathbf{X}$
– transponirana matrica matrice $\mathbf{X}$
χ_V – valence-state electronegativity
– elektronegativnost valentnog stanja
y – dependent (modelled) variable in a regression model
– zavisna (modelirana) varijabla u regresijskom modelu
$\overline{y}$ – mean value of the dependent variable *y*
– srednja vrijednost zavisne varijable *y*
$\boldsymbol{y}$ – vector of observed values of the dependent variable in regression analysis
– vektor opaženih vrijednosti zavisne varijable u regresijskoj analizi
$\overline{\boldsymbol{y}}$ – vector containing *n* times the mean value of *y*
– vektor koji *n* puta sadrži srednju vrijednost od *y*
$\hat{\boldsymbol{y}}$ – vector of predicted values of the dependent (modelled) variable in regression analysis
– vektor proračunatih vrijednosti zavisne (modelirane) varijable u regresijskoj analizi

References
Literatura

1. URL: https://mathworld.wolfram.com/Tetrahedron.html (28. 7. 2023.).
2. *C. E. Housecroft, A. G. Sharpe*, Inorganic Chemistry, 5th Ed., Pearson, Harlow, 2018, pp. 510–513.
3. *P. Atkins, T. Overton, J. Rourke, M. Weller, F. Armstrong, M. Hagerman,* Shriver & Atkins' Inorganic Chemistry, 5th Ed., Freeman, New York, 2010, p. 376.
4. *G. Rayner-Canham, T. Overton,* Descriptive Inorganic Chemistry, 5th Ed., Freeman, New York, 2010, p. 390.
5. *M. North,* Principles and Applications of Stereochemistry, Routledge, London, 1998, p. 3.
6. *M. Nógrádi,* Stereochemistry: Basic concepts and applications, Pergamon Press, London, 1981, pp. 6–8.
7. *D. F. Larder,* Historical Aspects of the Tetrahedron in Chemistry, J. Chem. Educ. **44** (1967) 661–666, doi: https://doi.org/10.1021/ed044p661.
8. *P. Cintas,* On the Origin of Tetrahedral Carbon: A Case for Philosophy of Chemistry?, Found. Chem. **4** (2002) 149–161, doi: https://doi.org/10.1023/A:1016063004499.
9. *G. Natta,* M. Farina, Stereochemistry, Longman, London, 1972, pp. 18–21.
10. *D. G. Morris,* Stereochemistry, Tutorial Chemistry Texts, Vol. 1, Royal Society of Chemistry, London, 2002, pp. 1–2.
11. *T. L. Brown, H. E. LeMay Jr., B. E. Bursten, C. J. Murphy, P. M. Woodward, M. W. Stoltzfus, M. W. Lufaso,* Chemistry: The Central Science, 14th Ed., Pearson, London, 2017, pp. 339–341.
12. *R. Kiralj, M. M. C. Ferreira,* Predicting Bond Lengths in Planar Benzenoid Polycyclic Aromatic Hydrocarbons: A Chemometric Approach, J. Chem. Inf. Comput. Sci. **42** (2002) 508–523, doi: https://doi.org/10.1021/ci010063g.
13. *R. Kiralj, M. M. C. Ferreira,* On Heteroaromaticity of Nucleobases. Bond Lengths as Multidimensional Phenomena, J. Chem. Inf. Comput. Sci. **43** (2003) 787–809, doi: https://doi.org/10.1021/ci0200523.

14. *Scilab 5.5.2*, Scilab Enterprises S. A. S., Orsay, 2015.
15. URL: https://www.socscistatistics.com/pvalues/pearsondistribution.aspx (23. 7. 2023.).
16. URL: https://cran.r-project.org/web/packages/LearnPCA/vignettes/Vig_07_Functions_PCA.pdf (2. 10. 2023.).
17. *M. M. C. Ferreira*, Quimiometria: conceitos, métodos e aplicações, Editora da UNICAMP, Campinas, SP, 2015, pp. 107–250.
18. *M. M. C. Ferreira*, Multivariate QSAR, J. Braz. Chem. Soc. **13** (2002) 742–753, doi: https://doi.org/10.1590/S0103-50532002000600004.
19. *Pirouette 4.5 rev 1*, Infometrix, Inc., Bothell, WA, 2014.
20. *R. Kiralj, M. M. C. Ferreira*, A priori molecular descriptors in QSAR: a case of HIV-1 protease inhibitors. I. The chemometric approach, J. Mol. Graph. Model. **21** (2003) 435–448, doi: https://doi.org/10.1016/S1093-3263(02)00201-2.
21. *R. Kiralj, M. M. C. Ferreira*, Chemometric analysis of the multidrug resistance in strains of *Penicillium digitatum*, SAR QSAR Environ. Res. **19** (2008) 55–70, doi: https://doi.org/10.1080/10629360701844118.
22. *R. Kiralj, Y. Takahata,* Chemometric modeling of core-electron binding energies, Struct. Chem. **17** (2006) 525–538, doi: https://doi.org/10.1007/s11224-006-9104-3.
23. *M. M. C. Ferreira*, *R. Kiralj*, QSAR study of β-lactam antibiotic efflux by the antibacterial multidrug resistance pump AcrB, J. Chemom. **18** (2004) 242–252, doi: https://doi.org/10.1002/cem.867.
24. *R. Kiralj, M. M. C. Ferreira*, Combined Computational and Chemometric Study of 1*H*-Indole-3-Acetic Acid, Int. J. Quantum Chem. **95** (2003) 237–251, doi: https://doi.org/10.1002/qua.10693
25. *R. Kiralj, M. M. C. Ferreira*, Is your QSAR/QSPR descriptor real or trash?, J. Chemom. **24** (2010) 681–693, doi: https://doi.org/10.1002/cem.1331.
26. URL: https://arachnoid.com/polysolve/ (27. 7. 2023.).
27. URL: https://www.statlect.com/fundamentals-of-statistics/normal-linear-regression-model (2. 10. 2023.).
28. URL: https://www.socscistatistics.com/pvalues/fdistribution.aspx (27. 7. 2023.).
29. URL: https://www.graphpad.com/quickcalcs/pValue1/ (27. 7. 2023.).
30. *M. Wendorff, C. Röhr*, $Sr_{14}[Al_4]_2[Ge]_3$: Eine Zintl-Phase mit isolierten $[Ge]^{4-}$- und $[Al_4]^{8-}$-Anionen, Z. Naturforsch. B **62** (2007) 1227–1234, doi: https://doi.org/10.1515/znb-2007-1001.
31. *M. L. Fornasini, F. Merlo*, The crystal structure of $Ca_{11}Ga_7$, Z. Kristallogr. **187** (1989) 111–115, doi: https://doi.org/10.1524/zkri.1989.187.14.111.
32. *S. C. Sevov, J. D. Corbett*, Synthesis, Characterization, and Bonding of Indium Cluster Phases: $Na_{15}In_{27.4}$, a Network of In_{16} and In_{11} Clusters; Na_2In with Isolated Indium Tetrahedra, J. Solid State Chem. **103** (1993) 114–140, doi: https://doi.org/10.1006/jssc.1993.1084.
33. *M. Kirchner, W. Schnelle, F. R. Wagner, R. Kniep, R. Nie*, $(A_{19}N_7)[In_4]_2$ (A = Ca, Sr) and $(Ca_4N)[In_2]$: Synthesis, Crystal Structures, Physical Properties, and Chemical Bonding, Z. Anorg. Allg. Chem. **631** (2005) 1477–1486, doi: https://doi.org/10.1002/zaac.200500029.
34. *V. F. Schwinghammer, M. Janesch, N. Korber, S. Gärtner,* Na_7RbT_{l4} – A New Ternary Zintl Phase Containing $[Tl_4]^{8-}$ Tetrahedra, Z. Anorg. Allg. Chem. **648** (2022) e202200332, doi: https://doi.org/10.1002/zaac.202200332.
35. *Z.-C. Dong, J. D. Corbett*, $Na_{23}K_9Tl_{15.3}$: An Unusual Zintl Compound Containing Apparent Tl_5^{7-}, Tl_4^{8-}, Tl_3^{7-}, and Tl_5^- Anions, Inorg. Chem. **35** (1996) 3107–3112, doi: https://doi.org/10.1021/ic960014z.
36. *C. Liu, Z.-M. Sun*, Recent advances in structural chemistry of Group 14 Zintl ions, Coord. Chem. Rev. **382** (2019) 32–56, doi: https://doi.org/10.1016/j.ccr.2018.12.003.
37. *L. M. Scherf, O. Pecher, K. J. Griffith, F. Haarmann, C. P. Grey, T. F. Fässler,* Zintl Phases $K_{4-x}Na_xSi_4$ ($1 \leq x \leq 2.2$) and K_7NaSi_8: Synthesis, Crystal Structures, and Solid-State NMR Spectroscopic Investigations, Eur. J. Inorg. Chem. (2016) 4674–4682, doi: https://doi.org/10.1002/ejic.201600735.
38. *M. Beekman, S. M. Kauzlarich, L. Doherty, G. S. Nolas*, Zintl Phases as Reactive Precursors for Synthesis of Novel Silicon and Germanium-Based Materials, Materials **12** (2019) 1139, doi: https://doi.org/10.3390/ma12071139.
39. *J. Witte, H. G. Schnering, W. Klemm,* Die Kristallstruktur von NaSi und NaGe, Z. Anorg. Allg. Chem. **327** (1964) 260–273, doi: https://doi.org/10.1002/zaac.19643270319.
40. *H. Bürger, R. Eujen,* Infrarot-Spektren der Tetraeder-Ionen Si_4^{4-} und Ge_4^{4-}, Z. Anorg. Allgem. Chem. **894** (1972) 19–25, doi: https://doi.org/10.1002/zaac.19723940105.
41. *M. Neumeier, F. Fendt, S. Gärtner, C. Koch, T. Gärtner, N. Korber, R. M. Gschwind,* Detection of the Elusive Highly Charged Zintl Ions Si_4^{4-} and Sn_4^{4-} in Liquid Ammonia by NMR Spectroscopy, Angew. Chem. Int. Ed. **52** (2013) 4483–4486, doi: https://doi.org/10.1002/anie.201209578.
42. *V. F. Kranak, D. E. Benson, L. Wollmann, M. Mesgar, S. Shafeie, J. Grins*, Hydrogenous Zintl Phase $Ba_3Si_4H_x$ ($x = 1-2$): Transforming Si_4 "Butterfly" Anions into Tetrahedral Moieties, Inorg. Chem. **54** (2015) 756–764, doi: https://doi.org/10.1021/ic501421u.
43. *A. Courac, Y. Le Godec, C. Renero-Lecuna, H. Moutaabbid, R. Kumar, C. Coelho-Diogo, C. Gervais, D. Portehault*, High-Pressure Melting Curve of Zintl Sodium Silicide Na_4Si_4 by In Situ Electrical Measurements, Inorg. Chem. **58** (2019) 10822–10828, doi: https://doi.org/10.1021/acs.inorgchem.9b01108.
44. *C. Lorenz, S. Gärtner, N. Korber*, Ammoniates of Zintl Phases: Similarities and Differences of Binary Phases A_4E_4 and Their Corresponding Solvates, Crystals **8** (2018) 276, doi: https://doi.org/10.3390/cryst8070276.
45. *W. Carrillo-Cabrera, R. Cardoso Gil, M. Somer, Ö. Persil, H. G. von Schnering*, $Na_{12}Ge_{17}$: A Compound with the Zintl Anions $[Ge_4]^{4-}$ and $[Ge_9]^{4-}$ Synthesis, Crystal Structure, and Raman Spectrum, Z. Anorg. Allg. Chem. **629** (2003) 601–608, doi: https://doi.org/10.1002/chin.200328005.
46. *S. Stefanoski, G. J. Finkelstein, M. D. Ward, T. Zeng, K. Wei, E. S. Bullock, C. M. Beavers, H. Liu, G. S. Nolas, T. A. Strobel*, Zintl Ions within Framework Channels: The Complex Structure and Low-Temperature Transport Properties of Na_4Ge_{13}, Inorg. Chem. **57** (2018) 2002–2012, doi: https://doi.org/10.1021/acs.inorgchem.7b02914.
47. *G. Frisch, C. Hoch, C. Röhr, P. Zönnchen, K.-D. Becker, D. Niemeier,* Alkaline Metal Stannide-Stannates: 'Double Salts' with Zintl Sn_4^{4-} and Stannate SnO_3^{4-} Anions, Z. Anorg. Allg. Chem. **629** (2003) 1661–1672, doi: https://doi.org/10.1002/zaac.200300124.
48. K. *Wiesler, K. Brandl, A. Fleischmann, N. Korber,* Tetrahedral $[Tt_4]^{4-}$ Zintl Anions Through Solution Chemistry: Syntheses and Crystal Structures of the Ammoniates $Rb_4Sn_4{\cdot}2NH_3$, $Cs_4Sn_4{\cdot}2NH_3$, and $Rb_4Pb_4{\cdot}2NH_3$, Z. Anorg. Allg. Chem. **635** (2009) 508–512, doi: https://doi.org/10.1002/zaac.200801249.
49. *T. F. Fässler*, Zintl Ions: Principles and Recent Developments, Structure and Bonding 140, Springer, Heidelberg, 2011, doi: https://doi.org/10.1007/978-3-642-21181-2.

50. *F. Fendt, C. Koch, S. Gärtner, N. Korber,* Reaction of Sn_4^{4-} in liquid ammonia: the formation of $Rb_6[(\eta^2\text{-}Sn_4)Zn(\eta^3\text{-}Sn_4)]\cdot 5NH_3$, Dalton Trans. **42** (2013) 15548–15550, doi: https://doi.org/10.1039/c3dt51932e.
51. *W. Klein, C. B. Benda, T. Henneberger, B. J. L. Witzel, T. F. Fässler*, Investigations on the Solubility of Sn_4-Cluster Compounds in Liquid Ammonia, Z. Anorg. Allg. Chem. **647** (2021) 2047–2054, doi: https://doi.org/10.1002/zaac.202100239.
52. *S. Bobev, S. C. Sevov,* Synthesis and Characterization of $A_3Na_{10}Sn_{23}$ (A = Cs, Rb, K) with a New Clathrate-Like Structure and of the Chiral Clathrate $Rb_5Na_3Sn_{25}$, Inorg. Chem. **39** (2000) 5930–5937, doi: https://doi.org/10.1021/ic0007714.
53. *C. Röhr*, $K_{19}Pb_8O_4(OH)_3$ und KPb: Ein Oxid-Hydroxid und eine Zintl-Phase mit $[Pb_4]^{4-}$-Anionen, Z. Naturforsch. B **50** (1995) 802–808, doi: https://doi.org/10.1515/znb-1995-0519.
54. *M. Baitinger, K. Peters, M. Somer, W. Carrillo-Cabrera, Yu. Grin, R. Kniep, H. G. von Schnering*, Crystal structure of tetrarubidium *tetrahedro*-tetraplumbide, Rb_4Pb_4 and of tetracaesium *tetrahedro*-tetraplumbide, Cs_4Pb_4, Z. Kristallogr. NCS **214** (1999) 455–456, doi: https://doi.org/10.1515/ncrs-1999-0427.
55. *S. Bobev, S. C. Sevov,* Isolated deltahedral clusters of lead in the solid state: synthesis and characterization of Rb_4Pb_9 and $Cs_{10}K_6Pb_{36}$ with Pb_9^{4-}, and $A_3A'Pb_4$ (A=Cs, Rb, K; A'=Na, Li) with Pb_4^{4-}, Polyhedron **21** (2002) 641–649, doi: https://doi.org/10.1016/S0277-5387(01)01016-6.
56. *Z. Dang, J. Shamsi, Q. A. Akkerman, M. Imran, G. Bertoni, R. Brescia, L. Manna,* Low-Temperature Electron Beam-Induced Transformations of Cesium Lead Halide Perovskite Nanocrystals, ACS Omega **2** (2017) 5660−5665, doi: https://doi.org/10.1021/acsomega.7b01009.
57. *G. A. de Wijs, G. Pastore, A. Selloni, W. van der Lugt,* First-principles molecular-dynamics simulation of liquid CsPb, J. Chem. Phys. **103** (1995) 5031–5040, https://doi.org/10.1063/1.470590.
58. *A. Simon, H. Borrmann, H. Craubner,* Crystal Structure of Ordered White Phosphorus (β-P), Phosphorus Sulfur **30** (1987) 507–510, doi: https://doi.org/10.1080/03086648708080631.
59. *L.-S. Wang, B. Niu, Y. T. Lee, D. A. Shirley, E. Ghelichkhani, E. R. Grant*, Photoelectron spectroscopy and electronic structure of clusters of the group V elements. II. Tetramers: Strong Jahn–Teller coupling in the tetrahedral ^2E ground states of P_4^+, As_4^+, and Sb_4^+, J. Chem. Phys. **93** (1990) 6318–6326, doi: https://doi.org/10.1063/1.459698.
60. *S. Ebel, H. T. Dieck, H. Walther, J. Krizek*, Photoelectron Spectra of Group V Compounds, As_4 and As_2, Inorg. Chim. Acta **53** (1981) L101–L102, doi: https://doi.org/10.1016/S0020-1693(00)84756-8.
61. *P. Mal, B. Breiner, K. Rissanen, J. R. Nitschke*, White Phosphorus Is Air-Stable Within a Self-Assembled Tetrahedral Capsule, Science **324** (2009) 1697–1699, doi: https://doi.org/10.1126/science.1175313.
62. *J.-L. M. Abboud, M. Herreros, R. Notario, M. Esseffar, O. Mó, M. Yáñez*, A New Bond from an Old Molecule: Formation, Stability, and Structure of P_4H^+, J. Am. Chem. Soc. **118** (1996) 1126–1130, doi: https://doi.org/10.1021/ja9500668.
63. *C. Schwarzmaier, A. Schindler, C. Heindl, S. Scheuermayer, E. V. Peresypkina, A. V. Virovets, M. Neumeier, R. Gschwind, M. Scheer,* Stabilization of Tetrahedral P_4 and As_4 Molecules as Guests in Polymeric and Spherical Environments, Angew. Chem. Int. Ed. **52** (2013) 10896–10899, doi: https://doi.org/10.1002/anie.201306146.
64. *D. Yang, J. Zhao, L. Yu, X. Lin, W. Zhang, H. Ma, A. Gogoll, Z. Zhang, Y. Wang, X.-J. Yang, B. Wu,* Air- and Light-Stable P_4 and As_4 within an Anion-Coordination-Based Tetrahedral Cage, J. Am. Chem. Soc. **139** (2017) 5946−5951, doi: https://doi.org/10.1021/jacs.7b01890.
65. *T. Jiao, L. Chen, D. Yang, X. Li, G. Wu, P. Zeng, A. Zhou, Q. Yin, Y. Pan, B. Wu, X. Hong, X. Kong, V. M. Lynch, J. L. Sessler, H. Li*, Trapping White Phosphorus within a Purely Organic Molecular Container Produced by Imine Condensation, Angew. Chem. Int. Ed. **46** (2017) 14545−14550, doi: https://doi.org/10.1002/anie.201708246.
66. *W. Choi, H. Ohtsu, Y. Matsushitab, M. Kawano*, Safe P_4 reagent in a reusable porous coordination network, Dalton Trans. **45** (2016) 6357-6360, doi: https://doi.org/10.1039/c6dt00860g.
67. *M. Hart, E. R. White, J. Chen, C. M. McGilvery, C. J. Pickard, A. Michaelides, A. Sella, M. S. P. Shaffer, C. G. Salzmann*, Encapsulation and Polymerization of White Phosphorus Inside Single-Wall Carbon Nanotubes, Angew. Chem. Int. Ed. **28** (2017) 8144−8148, doi: https://doi.org/10.1002/anie.201703585.
68. *S. Albertini, F. Hechenberger, S. Kollotzek, L. Tiefenthaler, P. Martini, M. Kuhn, A. Menzel, E. J. Al Maalouf, H. Schobel, M. Mahmoodi-Darian, P. Scheier,* Phosphorus cluster cations formed in doped helium nanodroplets are different, Int. J. Mass Spectrom. **459** (2021) 116472, doi: https://doi.org/10.1016/j.ijms.2020.116472.
69. *I. W. Locke, A. D. Darwish, H. W. Kroto, K. Prassides, R. Taylor, D. R. M. Walton*, Phosphorus/buckminsterfullerene intercalation compound, $C_{60}(P_4)_2$, Chem. Phys. Lett. **225** (1994) 186−190, doi: https://doi.org/10.1016/0009-2614(94)00626-1.
70. *R.-B. Huang, Z.-Y. Liu, H.-F. Liu, L.-H. Chen, Q. Zhang, C.-R. Wang, L.-S. Zheng, F.-Y. Liu, S.-Q. Yu, X.-X. Ma*, Collision-induced dissociation of mass-selected phosphorus cluster cations, Int. J. Mass Spectrom. Ion Processes **151** (1995) 55−62, doi: https://doi.org/10.1016/0168-1176(95)04291-1.
71. *Y. Morino, T. Ukaji, T. Ito,* Molecular Structure Determination by Gas Electron Diffraction at High Temperatures. I. Arsenic, Bull. Chem. Soc. Jpn. **39** (1966) 64−71, doi: https://doi.org/10.1246/bcsj.39.64.
72. *J. Eiduss, R. Kalendarev, A. Rodionov, A. Sazonov, G. Chikvaidze,* Some Aspects of the Solid State Physics of Yellow Arsenic, Phys. Stat. Sol. B **193** (1996) 3–23, doi: https://doi.org/10.1002/pssb.2221930102.
73. *M. Hart, J. Chen, A. Michaelides, A. Sella, M. S. P. Shaffer, C. G. Salzmann*, One-dimensional Arsenic Allotropes: Polymerization of Yellow Arsenic Inside Single-wall Carbon Nanotubes, Angew. Chem. Int. Ed. **5**7 (2018) 11649–11653, doi: https://doi.org/10.1002/anie.201805856.
74. *C. Schwarzmaier, M. Sierka, M. Scheer,* Intact As_4 Tetrahedra Coordinated Side-On to Metal Cations, Angew. Chem. Int. Ed. **52** (2013) 858–861, doi: https://doi.org/10.1002/anie.201208226.
75. *M. F. Daniel, A. J. Leadbetter,* The structure of vapour-deposited films of molecular arsenic (As_4), Philos. Mag. B, **44** (1981) 509–529, doi: https://doi.org/10.1080/01418638108222585.
76. *T. M. Bernhardt, B. Stegemann, B. Kaiser, K. Rademann,* Crystalline Structures of Sb_4 Molecules in Antimony Thin Films, Angew. Chem. Int. Ed. **42** (2003) 199–202, doi: https://doi.org/10.1002/anie.200390077.
77. *B. Stegemann, B. Kaiser, K. Rademann,* Subsequent layer growth of supported nanoparticles by deposition of Sb_4 clusters onto $MoS_2(0001)$, New J. Phys. **4** (2002) 89, doi: https://doi.org/10.1088/1367-2630/4/1/389.
78. *M. Gu, C. Li, Y. Ding, K. Zhang, S. Xia, Y. Wang, M.-H. Lu,*

H. Lu, Y.-F. Chen, Direct Growth of Antimonene on C-Plane Sapphire by Molecular Beam Epitaxy, Appl. Sci. **10** (2020) 639, doi: https://doi.org/10.3390/app10020639.

79. *R. Kelting, A. Baldes, U. Schwarz, T. Rapps, D. Schooss, P. Weis, C. Neiss, F. Weigend, M. M. Kappes,* Structures of small bismuth cluster cations, J. Chem. Phys. **136** (2012), 154309, doi: https://doi.org/10.1063/1.3703014.
80. *M. A. Shestopalov, A. Y. Ledneva, S. Cordier, O. Hernandez, M. Potel, T. Roisnel, N. G. Naumov, C. Perrin*, Tetrahedral Mo_4 Clusters as Building Blocks for the Design of Clathrate-Related Giant Frameworks, Angew. Chem. **123** (2011) 7438–7441, doi: https://doi.org/10.1002/ange.201101986.
81. *R. Gautier, F. Chérioux, G. Süss-Fink, J.-Y. Saillard,* Electron Deficiency in Tetrahedral Transition-Metal Clusters: Electronic Structure and Magnetic Properties of $[Ru_4(\eta^6\text{-}C_6H_6)_4(\mu_3\text{-}H)_4]^{2+}$, Inorg. Chem. **42** (2003) 8278–8282, doi: https://doi.org/10.1021/ic030053s.
82. *P. Hu, Q. Luo, Q.-S. Li, Y. Xie, R. B King, H. F. Schaefer,* Construction of the Tetrahedral Trifluorophosphine Platinum Cluster $Pt_4(PF_3)_8$ from Smaller Building Blocks, Inorg. Chem. **53** (2014) 5300−5310, doi: https://doi.org/10.1021/ic5005243.
83. *E. Zeller, H. Bernda, H. Schmidbaur,* Tetrahedral Gold Cluster $[Au]^{2+}$: Crystal Structure of $([(^tBu)_3PAu]_4)^{2+}(BF^{4-})_2{\cdot}CHCl_3$, Inorg. Chem. **32** (1993) 3203−3204, doi: https://doi.org/10.1021/ic00067a002.
84. *D. F. Förster, T. O. Wehling, S. Schumacher, A. Rosch, T. Michely,* Phase coexistence of clusters and islands: europium on graphene, New J. Phys. **14** (2012) 023022, doi: https://doi.org/10.1088/1367-2630/14/2/023022.
85. URL: https://www.ccdc.cam.ac.uk/structures/ (23. 7. 2023.).
86. *ViewerLite 4.2*, Accelrys, Inc., San Diego, CA, 2001.
87. *S. M. Kauzlarich*, Special Issue: Advances in Zintl Phases (Editorial), Materials **12** (2019) 2554, doi: https://doi.org/10.3390/ma12162554.
88. *Y. Wang, C. Zhang, X. Wang, J. Guo, Z.-M. Sun, H. Zhang*, Site Selective CO_2 Reduction over Highly Dispersed Ru-SnO_x Sites Derived from $[Ru@Sn_9]^{6-}$ Zintl Cluster, ACS Catal. **10** (2020) 7808–7819, doi: https://doi.org/10.1021/acscatal.0c01253.
89. *S. Geier, R. Jung, K. Peters, H. A. Gasteiger, D. Fattakhova-Rohlfing, T. F. Fässler,* A wet-chemical route for macroporous inverse opal Ge anodes for lithium ion batteries with high capacity retention, Sustain. Energy Fuels **2** (2018) 85–90, doi: https://doi.org/10.1039/c7se00422b.
90. *S. Scharfe, F. Kraus, S. Stegmaier, A. Schier, T. F. Fässler*, Zintl Ions, Cage Compounds, and Intermetalloid Clusters of Group 14 and Group 15 Elements, Angew. Chem. Int. Ed. **50** (2011) 3630–3670, doi: https://doi.org/10.1002/anie.201001630.
91. *R. E. Vernon,* The location and composition of Group 3 of the periodic table, Found. Chem. **23** (2021) 155–197, doi: https://doi.org/10.1007/s10698-020-09384-2.
92. *P. Karen, P. McArdle, J. Takats,* Comprehensive definition of oxidation state (IUPAC Recommendations 2016), Pure Appl. Chem. **88** (2016) 831–839, doi: https://doi.org/10.1515/pac-2015-1204.
93. *M. Rahm, R. Hoffmann, N. W. Ashcroft,* Atomic and Ionic Radii of Elements 1–96, Chem. Eur. J. **22** (2016) 14625–14632, doi: https://doi.org/10.1002/chem.201602949.
94. *J. C. A. Boeyens,* The Periodic Electronegativity Table, Z. Naturforsch. B **63** (2008) 199–209, doi: https://doi.org/10.1515/znb-2008-0214.
95. *M. Mantina, A. C. Chamberlin, R. Valero, C. J. Cramer, D. G. Truhlar,* Consistent van der Waals Radii for the Whole Main Group, J. Phys. Chem. A **113** (2009) 5806–5812, https://doi.org/10.1021/jp8111556.
96. URL: https://www.rsc.org/periodic-table (17. 7. 2023.)
97. *L. R. Murphy, T. L. Meek, A. L. Allred, L. C. Allen,* Evaluation and Test of Pauling's Electronegativity Scale, J. Phys. Chem. A **104** (2000) 5867–5871, doi: https://doi.org/10.1021/jp000288e.
98. URL: https://stats.stackexchange.com/questions/195446/choosing-the-right-linkage-method-for-hierarchical-clustering (23. 7. 2023.).
99. URL: https://webelements.com/ (23. 7. 2023.).
100. *B. Meyer,* Elemental Sulfur, Chem. Rev. **76** (1976) 367–388, doi: https://doi.org/10.1021/cr60301a003.
101. *N. N. Greenwood, A. Earnshaw*, Chemistry of the Elements, 2nd Ed., Butterworth Heinemann, Oxford, 1998, pp. 761–764, doi: https://doi.org/10.1016/C2009-0-30414-6.
102. *M. W. Wong, R. Steudel,* Structure and spectra of tetrasulfur S_4 – an ab initio MO study, Chem. Phys. Lett. **379** (2003) 162–169, doi: https://doi.org/10.1016/j.cplett.2003.08.026.
103. *J. Beck, M. Kasper, A. Stankowski,* Tellurium Polycations Stabilized by Halogenobismutates - Syntheses and Crystal Structures of $Te_4[Bi_6C1_{20}]$ and $Te_4[Bi_2Br_8]$, Chem. Ber. **130** (1997) 1189–1192, doi: https://doi.org/10.1002/cber.19971300904.
104. *M. Atoji, W. N. Lipscomb,* The Crystal and Molecular Structure of B_4Cl_4, Acta Cryst. **6** (1953) 547–550, doi: https://doi.org/10.1107/S0365110X53001472.
105. *P. T. Brain, A. J. Downs, M. Fanfarillo, M. J. Goode, A. G. Massey, D. W. H. Rankin, H. E. Robertson,* The Molecular Structure of Tetraboron Tetrachloride in the Gas Phase, Determined by Electron Diffraction, J. Mol. Struct. **19** (1989) 163–170, doi: https://doi.org/10.1016/0022-2860(89)87015-2.
106. *D. Hnyk, T* symmetrical gaseous tetra-tert-butyltetraboratetrahedrane: an electron diffraction study, Polyhedron **16** (1997) 603–606, doi: https://doi.org/10.1016/0277-5387(96)00320-8.
107. *A. N. Alexandrova, A. I. Boldyrev, H.-J. Zhai, L.-S. Wang,* All-boron aromatic clusters as potential new inorganic ligands and building blocks in chemistry, Coord. Chem. Rev. **250** (2006) 2811–2866, doi: https://doi.org/10.1016/j.ccr.2006.03.032.
108. URL: https://webbook.nist.gov/cgi/inchi?ID=C12184804&Mask=20 (24. 7. 2023.).
109. *D. W. Arnold, S. E. Bradforth, T. N. Kitsopoulos, D. M. Neumark,* Vibrationally resolved spectra of C_2–C_{11} by anion photoelectron spectroscopy, J. Chem. Phys. **95** (1991) 8753–8764, doi: https://doi.org/10.1063/1.461211.
110. *F. Cacace, G. de Petris, A. Troiani,* Experimental Detection of Tetranitrogen, Science **295** (2002) 480–481, doi: https://doi.org/10.1126/science.1067681.
111. *F. Cacace, G. de Petris, A. Troiani,* Experimental Detection of Tetraoxygen, Angew. Chem. **113** (2001) 4186–4189, doi: https://doi.org/10.1002/1521-3773(20011105)40:21<4062::AID-ANIE4062>3.0.CO;2-X.
112. *L. F. Lundegaard, G. Weck, M. I. McMahon, S. Desgreniers, P. Loubeyre,* Observation of an O_8 molecular lattice in the ε phase of solid oxygen, Nature **443** (2006) 201–204, doi: https://doi.org/10.1038/nature05174.
113. *H. Irngartinge, A. Goldmann, R. Jahn, M. Nixdorf, H. Rodewald, G. Maier, K.-D. Malsch, R. Emrich,* Tetra-*tert*-butyltetrahedrane – Crystal and Molecular Structure, Angew. Chem. Int. Ed. **23** (1984) 993–994, doi: https://doi.org/10.1002/anie.198409931.
114. *M. C. Aragoni, M. Arca, F. A. Devillanova, M. B. Hurst-*

house, S. L. Huth, F. Isaia, V. Lippolis, A. Mancini, H. Ogilvie, Self-assembly of supramolecular architectures based on polybromide anions: crystal structure of $[(H_4tppz^{4+})(Br^-)_2(Br_4^{2-})]$[tppz = tetra(2-pyridyl)pyrazine], Inorg. Chem. Commun. **8** (2005) 79–82, doi: https://doi.org/10.1016/j.inoche.2004.11.014.

115. *R. Faggiani, R. J. Gillespie, R. Kapoor, C. J. L. Lock, J. E. Vekris,* Preparation and Solid-state and Solution Studies of Three Compounds of the Tetraiodine Dication I_4^{2+}: $I_4^{2+}(AsF_6^-)_2$, $I_4^{2+}(SbF_6^-)_2$, and $I_4^{2+}(Sb_3F_{14}^-)(SbF_6^-)$, Inorg. Chem. **2**7 (1988) 4350–4355, doi: https://doi.org/10.1021/ic00297a00.

116. *R. J. Gillespie, R. Kapoor, R. Faggiani, C. J. L. Lock, M. Murchie, J. Passmore,* The I_4^{2+} cation. *X*-Ray Crystal Structures of $(I_4^{2+})(AsF_6^-)_2$ and $(I_4^{2+})(Sb_3F_{14}^-)(SbF_6^-)$, J. Chem. Soc., Chem. Commun. (1983) 8–9, https://doi.org/10.1039/C39830000008.

117. *V. K. Sharma, M. Stark, N. Fridman, Y. G. Assaraf, Z. Gross,* Doubly Stimulated Corrole for Organelle-Selective Anti-Tumor Cytotoxicity, J. Med. Chem. **65** (2022) 6100–6115, https://doi.org/10.1021/acs.jmedchem.1c02085.

SAŽETAK

Tetraedrijske homonuklearne tetramerne vrste: pojavnost, oblici, strukture, svojstva i perspektive

Rudolf Kiralj

Tetraedrijske homonuklearne tetramerne vrste su neutralni ili ionski tetraedri kemijskih elemenata, pravi molekulski tetraedri u striktno geometrijskom smislu. Cilj ovog rada je naći sve prave tetraedre elemenata, eksperimentalno određene kao stabilne i sposobne da se premjeste u drugi medij/uvjete, zatim razjasniti njihovu strukturu i svojstva te istražiti njihov tehnološki značaj. Izrađena je baza podataka za tetraedre elemenata na osnovi pouzdane literature te je sačinjen skup podataka za molekulske deskriptore. Provedena je korelacijska analiza, hijerahijska klasterska analiza, analiza glavnih komponenata i neke regresijske analize. Dvanaest kemijskih elemenata iz grupa 13 – 15 imaju tetraedrijske vrste: anione E_4^{8-} (Al, Ga, In, Tl), anione E_4^{4-} (Si, Ge, Sn, Pb), neutralne molekule E_4 (P, As, Sb) i katione E_4^+ (P, As, Sb, Bi). Ove vrste pojavljuju se u 39 oblika, kao što su plinska faza, čvrste Zintlove faze, tanki filmovi, amonijati Zintlovih faza i kavezni spojevi uz ostale oblike. Zintlove faze kao i oblici P_4 i As_4 važni su za sintezu novih materijala, kemijske reakcije i druge primjene. Sažimanje podataka pokazalo je da su prave tetraedrijske vrste trodimenzionalni fenomen i da se pravi tetraedri mogu razlikovati od fiktivnih tetraedara. Uočeno je grupiranje tetraedara s obzirom na periodne grupe, naboje, razrede učestalosti i raznolikosti, veličinu tetraedara i metalni karakter. Empirijska pravila za veličinu vrsta i za glavne komponente ustanovljena su za prave tetraedre. Većina deskriptora je parabolična funkcija varijable razred. Prave tetraedrijske vrste ostaju djelomično još zagonetka te ih treba više proučavati.

Ključne riječi
Homonuklearne tetraedrijske molekule, strukturne baze podataka, Zintlove faze, kavezni spojevi, statistička analiza, smanjenje dimenzionalnosti, empirijska pravila

Veleučilište u Bjelovaru,
Trg Eugena Kvaternika 4, 43 000 Bjelovar

Izvorni znanstveni rad
Prispjelo 1. kolovoza 2023.
Prihvaćeno 7. listopada 2023.