Prediction of biological activity of compounds containing a 1,3,5-triazinyl sulfonamide scaffold by artificial neural networks using simple molecular descriptors

Eva Havránková [a, *], E.M. Peña-Méndez [b], Jozef Csöllei [a], Josef Havel [c, d]

[a] *Masaryk University, Faculty of Pharmacy, Department of Chemical Drugs, Palackého 1-3, CZ-612 42 Brno, Czech Republic*
[b] *Universidad de La Laguna (ULL), Facultad de Ciencias, Departamento de Química, Unidad Departamental de Química Analítica, 38201 La Laguna, Spain*
[c] *Masaryk University, Faculty of Science, Department of Chemistry, University Campus, Kamenice 753/5, CZ-625 00 Brno, Czech Republic*
[d] *International Clinical Research Center, St. Anne's University Hospital, Pekařská 53, 656 91 Brno, Czech Republic*

ARTICLE INFO

ABSTRACT

Simple molecular descriptors of extensive series of 1,3,5-triazinyl sulfonamide derivatives, based on the structure of sulfonamides and their physicochemical properties, were designed and calculated. These descriptors were successfully applied as inputs for artificial neural network (ANN) modelling of the relationship between the structure and biological activity. The optimized ANN architecture was applied to the prediction of the inhibition activity of 1,3,5-triazinyl sulfonamides against human carbonic anhydrase (hCA) II, tumour-associated hCA IX, and their selectivity (hCA II/hCA IX).

1. Introduction

Cancer is the second leading cause of death globally [1–3]. In 2016, cancer was responsible for over 9.6 million deaths, and the number of new cases is expected to rise by approximately 70% during the next 20 years.

Many tumour types, especially those with a low response to classical therapy, invasive growth, and metastatic spread, are closely associated with hypoxia (low concentration of oxygen in the microenvironment of the tumour, caused by changes in the metabolism of a tumour and rapid proliferation) [4–6]. In response to hypoxic conditions, hypoxia-inducible factor 1 (HIF-1) protein activates hypoxia-inducible genes, which are responsible for neoangiogenesis, improved glycolysis and energy production, upregulation of molecules connected with cell survival, and apoptosis and other mechanisms that help the tumour to survive [7–11]. One of the essential hypoxia-induced genes encodes human carbonic anhydrase (hCA) IX [7,8,12]. Transmembrane tumour-associated hCA IX is a zinc-based enzyme (containing a Zn^{2+} ion in the active site) that plays a crucial role in the pH regulation of the tumour microenvironment [13]. The high speed of proliferation of tumour cells leads to the production of increased amounts of metabolic products such as lactate, hydrogen protons, and bicarbonate ions [13,14]. The high concentration of these metabolites causes the more acidic extracellular environment of the tumour (pH 6.7–7.0) in comparison with healthy

tissue (pH 7.4) [7,8,14]. The enzyme hCA IX has the ability to catalyse the hydration of CO_2 to form bicarbonate anions and protons [7–9]. Bicarbonate anions are transported by anion transporters into the tumour cells (buffer intracellular microenvironment), and hydrogen protons contribute to the creation of the acidic extracellular environment. It has already been proved that the acidic tumour microenvironment enhances the growth and metabolic changes of tumour cells mentioned above but also that it has a crucial role in the spread of metastasis and resistance to classical chemotherapy or radiotherapy [7–9,14].

In general, hCA IX (and its inhibition) is an important target for research into new cancer chemotherapeutics. The employment of hCA IX inhibitors can enhance the efficiency of classical chemotherapeutics (i.e. coordination compounds of platinum) even against nonresponsive types of tumour (renal carcinoma) [13]. The inhibitory activity of many compounds, namely coumarins, carbamates, sulfamates or sulfonamides, against hCA IX has already been proved [15–18]. The sulfonamides containing the 1,3,5-triazine structural moiety are one of the most promising potential inhibitors of hCA IX (see Fig. 1). The research group of Prof. Supuran reported a series of sulfonamides incorporating 1,3,5-triazine moieties (**a**, **b**) in combination with alkylamino, alkoxy, and phenoxy structural moieties or hydroxy groups with an inhibition constant K_I in the range of 0.12–640 nM [19,20]. In subsequent work, Supuran et al. prepared and tested 1,3,5-triazinyl benzenesulfonamides
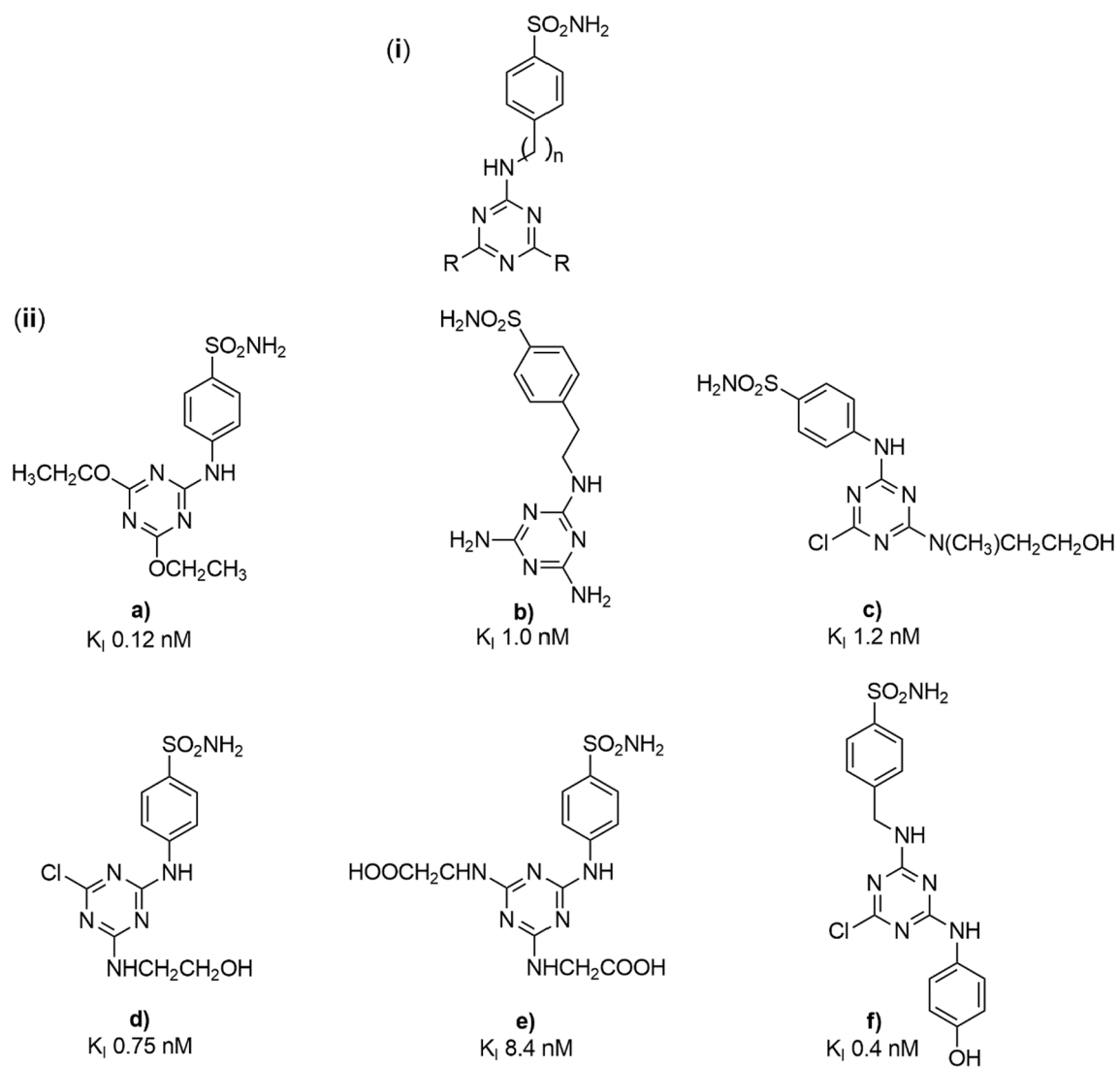
Fig. 1. (i) The general structure of hCA IX inhibitors; (ii) Examples of most active 1,3,5-triazinyl benzenesulfonamide derivatives studied in this paper [19–24].
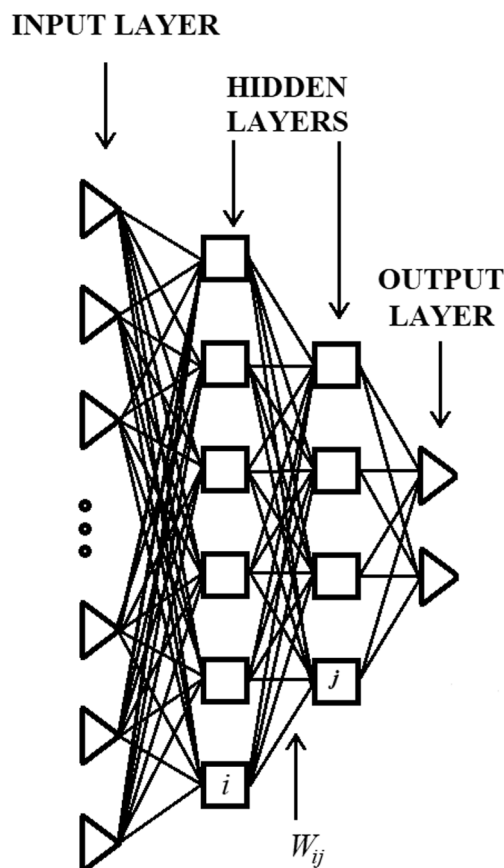
Fig. 2. The general architecture of feed-forward neural networks of multilayer perceptions. The w_{ij} are the weights belonging to the connection between the *i*-th and the *j*-th node. .
Adapted from [47]

(**c**) substituted with ammonia, morpholine, anilines, *tert*-butylamine, 1,3-diaminopropane, and *N*-methyl-2-aminoethanol as nucleophiles [21]. The K_Is of the tested compounds were in the range of 2.4–34.1 nM. The series of 1,3,5-triazinyl benzenesulfonamides (**d**) studied by Supuran et al. was further extended by the incorporation of different amino acids, amino acid esters, aminoalcohols, and the mono-*tert*-butyl-dimethylsilyl derivative of ethylene glycol into the 1,3,5-triazinyl benzenesulfonamide precursor [22]. The K_Is of the studied compounds were in the range of 0.75–95 nM. Mikuš et al. prepared a systematic study of conjugates of 1,3,5-triazinyl benzenesulfonamides (**e**) with various proteinogenic and non-proteinogenic amino acids, with resulting K_Is in the range of 8.4–2592 nM [23]. Havránková et al. focused on 1,3,5-triazinyl benzenesulfonamide derivatives (**f**) containing piperazine and/or aminoalcohol structural moieties, with the K_Is of the compounds ranging from 0.4 to 307.7 nM [24]. Some of the tested inhibitors can act even at picomolar levels, such as 1,3,5-triazinyl benzenesulfonamide disubstituted with ethoxy moieties (K_I 120 pM) [19], 1,3,5-triazinylaminomethylbenzenesulfonamide monosubstituted with 4-aminophenol (K_I 400 pM) [24], etc. However, most of the tested 1,3,5-triazinyl sulfonamide derivatives demonstrate only moderate or low selectivity for hCA IX over physiologically relevant isoenzymes of hCA.

Low selectivity could be a significant complication in the use of inhibitors of hCA IX in clinical treatment through the risk of unwanted adverse effects and compensation mechanisms evolving in cancer cells as a result of their phenotypic plasticity [17]. Therefore, the invention of new, extremely selective inhibitors of hCA IX with very high activity is necessary.

The interaction of inhibitors with the enzyme is complex and is significantly different for each structure [25,26]. Therefore, the design

of new drugs is challenging and focused on innovative methods. Because of the complexity explained above, the design of new drugs preceding their synthesis is an important step in modern drug development, and the theoretical prediction of the biological activity of a new drug spares much time, funds, and effort. Quantitative structure–activity relationship (QSAR) quantum mechanical (QM) methods are often used, although their application in computational drug design presents some limitations [27–34]. Quantum descriptors (which are difficult to understand for most chemists and are calculated using special programs) must be used for the computation. Only a few examples of constitutional descriptors, which usually represent the number of atoms, rings, or double bonds in a molecule, have been used in QM methods, and always in combination with other types of descriptors (e.g. quantum chemical descriptors) [35–38].

An alternative approach is based on the application of methods based on artificial intelligence, namely artificial neural networks (ANNs), which belong to a group of techniques using artificial intelligence inspired by the neurobiology and architecture of the human brain [39,40]. The ANNs present many advantages over the QSAR QM methodology. An example of the advances of artificial-intelligence approaches (e.g. stepwise multiple linear regressions, genetic algorithms, simulated annealing techniques, and non-linear ANNs) is their successful application in the prediction of biological activities and other properties of sulfonamides and substituted sulfonamides [41–43]. The comparison of QSAR models based on combinations of genetic algorithms, stepwise multiple linear regressions, and ANN methods has been studied thoroughly, demonstrating that models based on ANNs have the best predictive ability [43].

Recently, during the finishing of this paper, a detailed computational inhibition activity study for triazinyl sulfonamide conjugates with amino acids has been published in which molecular descriptors were calculated using extensive quantum chemical calculations [44].

The aim of the present work is (i) to develop and examine a new series of simple molecular descriptors for sulfonamide derivatives (incorporating 1,3,5-triazine moieties); (ii) to verify their application potential for ANN modelling of the relationship *descriptors = f(biological activity);* and (iii) to evaluate the prediction ability of such descriptors when using ANNs. An advantage of this approach is that the descriptors are simple and do not require sophisticated computation and evaluation of descriptors by statistical programs, which significantly saves time.

2. Results and discussion

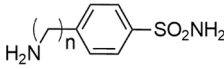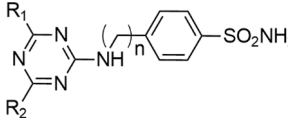We present here only a brief background of ANNs.

One of the most common types of ANN are feed-forward neural networks of multilayer perceptions (FFNN MLPs), which is often used for the prediction of biological activity [42,43,45–47]. The general expression of the typical architecture of FFNN MLPs is shown in Fig. 2.

The development of the ANN optimized model has three basic steps: training, verification, and prediction [42,43,45,47]. The determination of the appropriate number of nodes in each layer and the number of hidden layers is the most critical task in designing and optimizing FFNN architecture [42,43,45–47]. A useful criterion indicating whether a network structure is operating correctly during the training process is to minimize the value of the root mean square (RMS), calculated according to the following equation:

$$RMS = \sqrt{\frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(y_{ij} - out_{ij} \right)^2}{M - N}}$$
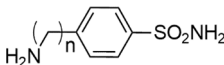
During the training phase, the optimal ANN structure and weighting (w_{ij}) coefficients are determined. The training phase is considered to be complete when the neural network model achieves the required statistical accuracy for the prediction of outputs and for maintaining the required RMS for a given sequence of inputs. The next step is verification

Table 1

Chemical structures of the 1,3,5-triazinyl sulfonamide derivatives studied in this paper.

n =	R$_1$ =	R$_2$ =	Reference
0, 1, 2	–	–	[19]

n =	R$_1$ =	R$_2$ =	Reference
0, 1, 2	—Cl	—Cl	[19]
0, 1, 2	—OH	—OH	[19]
0, 1, 2	—NH—CH$_3$	—NH—CH$_3$	[19]
0, 1, 2	—O—CH$_3$	—O—CH$_3$	[19]
0, 1, 2	—O—CH$_2$CH$_3$	—O—CH$_2$CH$_3$	[19]
1, 2	—NH$_2$	—NH$_2$	[20]
0, 1, 2	—NH—NH$_2$	—NH—NH$_2$	[20]
1, 2	—NH—CH$_2$CH$_3$	—NH—CH$_2$CH$_3$	[20]
0, 1, 2	—NH—CH(CH$_3$)$_2$	—NH—CH(CH$_3$)$_2$	[20]
0, 1, 2	—NH—CH$_2$CH$_2$CH$_3$	—NH—CH$_2$CH$_2$CH$_3$	[20]
0, 1, 2	—NH—CH$_2$CH$_2$CH$_2$CH$_3$	—NH—CH$_2$CH$_2$CH$_2$CH$_3$	[20]
0	—NH—CH$_2$CH$_2$N(CH$_2$CH$_3$)$_2$	—NH—CH$_2$CH$_2$N(CH$_2$CH$_3$)$_2$	[20]
0, 1	—NH—CH$_2$CH$_2$[N(CH$_2$CH$_2$)$_2$NH]	—NH—CH$_2$CH$_2$[N(CH$_2$CH$_2$)$_2$NH]	[20]
0, 1, 2	—N(CH$_3$)$_2$	—N(CH$_3$)$_2$	[20]
0, 1, 2	—N(CH$_2$CH$_3$)$_2$	—N(CH$_2$CH$_3$)$_2$	[20]
0, 1, 2	—N(CH$_3$)CH$_2$CH$_2$CH$_3$	—N(CH$_3$)CH$_2$CH$_2$CH$_3$	[20]
1	—Cl	—NH—CH$_2$COOH	[20]
2	—Cl	—NH—CH$_2$COOH	[20]
2	—Cl	—NH—CH$_2$COOCH$_3$	[20]
2	—Cl	—NH—CH$_2$CH$_2$COOH	[20]
0	—NH—CH$_2$COOH	—NH—CH$_2$COOH	[22]
0	—Cl	—NH—CH$_2$COOH	[22]
0, 1	—NH—CH$_2$COOCH$_3$	—NH—CH$_2$COOCH$_3$	[22]
0	—Cl	—NH—CH$_2$COOCH$_3$	[22]
0	—NH—CH$_2$CH$_2$COOH	—NH—CH$_2$CH$_2$COOH	[22]
0	—Cl	—NH—CH(CH$_3$)COOH	[22]
0	—Cl	—NH—CH(CH$_2$OH)COOH	[22]
0	—Cl	—NH—CH$_2$CH$_2$OH	[22]
0	—Cl	—NH—CH$_2$CH$_2$CH$_2$CH$_2$OH	[22]
1, 2	—NH—CH(CH$_3$)COOH	—NH—CH(CH$_3$)COOH	[22]
1, 2	—Cl	—NH—CH$_2$CH$_2$COOH	[22]
2	—NH—CH(CH$_2$OH)COOH	—NH—CH(CH$_2$OH)COOH	[22]
0	—Cl	—NH—C$_6$H$_5$	[21]
0	—Cl	—NH—C$_6$H$_4$—3-CH$_3$	[21]
0	—Cl	—N(CH$_3$)CH$_2$CH$_2$OH	[21]
0	—Cl	—NH—CH$_2$CH$_2$CH$_2$NH$_2$	[21]
0	—Cl	—NH—C$_6$H$_3$—3-Cl—4-F	[21]
0	—N(CH$_2$CH$_2$)$_2$O	—NH—C$_6$H$_5$	[21]
0	—N(CH$_2$CH$_2$)$_2$O	—NH—C$_6$H$_4$—3-CH$_3$	[21]
0	—NH$_2$	—NH$_2$	[21]
0	—N(CH$_2$CH$_2$)$_2$O	—N(CH$_2$CH$_2$)$_2$O	[21]
0	—NH—CH$_2$CH(CH$_3$)$_2$	—NH—CH$_2$CH(CH$_3$)$_2$	[21]
0	—N(CH$_3$)CH$_2$CH$_2$OH	—N(CH$_3$)CH$_2$CH$_2$OH	[21]
0	—NH—CH$_2$CH$_2$CH$_2$NH$_2$	—NH—CH$_2$CH$_2$CH$_2$NH$_2$	[21]
0, 1, 2	—NH—CH$_2$COOH	—NH—CH$_2$COOH	[23]
0, 1, 2	—NH—CH$_2$CH$_2$COOH	—NH—CH$_2$CH$_2$COOH	[23]
0, 1, 2	—NH—CH(COOH)CH(CH$_3$)$_2$	—NH—CH(COOH)CH(CH$_3$)$_2$	[23]
0, 1, 2	—NH—CH(COOH)CH$_2$CH(CH$_3$)$_2$	—NH—CH(COOH)CH$_2$CH(CH$_3$)$_2$	[23]
0, 1, 2	—NH—CH(COOH)CH(CH$_3$)CH$_2$CH$_3$	—NH—CH(COOH)CH(CH$_3$)CH$_2$CH$_3$	[23]
0, 1, 2	—NH—CH(COOH)CH$_2$CH$_2$—SCH$_3$	—NH—CH(COOH)CH$_2$CH$_2$—SCH$_3$	[23]
0, 1, 2	–1-proline	–1-proline	[23]
0, 1, 2	—NHCH(COOH)CH$_2$C$_6$H$_5$	—NHCH(COOH)CH$_2$C$_6$H$_5$	[23]
0, 1, 2	—NH—CH(COOH)CH$_2$COOH	—NH—CH(COOH)CH$_2$COOH	[23]
0, 1, 2	—NH—CH(COOH)CH$_2$CH$_2$COOH	—NH—CH(COOH)CH$_2$CH$_2$COOH	[23]
1, 2	—NH—CH$_2$—C$_6$H$_4$—4-SO$_2$NH$_2$	—NH—CH$_2$—C$_6$H$_4$—4-SO$_2$NH$_2$	[24]
1	—NH—CH$_2$—C$_6$H$_4$—4-SO$_2$NH$_2$	—NH—CH$_2$CH$_2$OH	[24]
1	—NH—CH$_2$—C$_6$H$_4$—4-SO$_2$NH$_2$	—NH—CH$_2$CH(OH)CH$_2$OH	[24]
1	—NH—CH$_2$CH$_2$—C$_6$H$_4$—4-SO$_2$NH$_2$	—NH—CH$_2$CH$_2$OH	[24]
1	—[N(CH$_2$CH$_2$)$_2$N]CH$_2$COOCH$_3$	—NH—CH$_2$CH$_2$—C$_6$H$_4$—4-SO$_2$NH$_2$	[24]
1	—[N(CH$_2$CH$_2$)$_2$N]COOCH$_3$	—NH—CH$_2$CH(OH)CH$_2$OH	[24]
1	—[N(CH$_2$CH$_2$)$_2$N]CH$_2$COOCH$_3$	—NH—CH$_2$CH(OH)CH$_2$OH	[24]
2	—NH—C$_6$H$_4$—4-SO$_2$NH$_2$	—NH—CH$_2$CH$_2$—C$_6$H$_4$—4-SO$_2$NH$_2$	[24]

(continued on next page)

Table 1 (*continued*)

$n =$	$R_1 =$	$R_2 =$	Reference
1	—Cl	—NH—CH$_2$—C$_6$H$_4$—4-SO$_2$NH$_2$	[24]
0, 1	—Cl	—NH—CH$_2$CH$_2$—C$_6$H$_4$—4-SO$_2$NH$_2$	[24]
0, 1	—Cl	— [N(CH$_2$CH$_2$)$_2$N]COOCH$_3$	[24]
0, 1	—Cl	— [N(CH$_2$CH$_2$)$_2$N]CH$_2$COOCH$_3$	[24]
0, 1	—Cl	— [N(CH$_2$CH$_2$)$_2$N]CH$_2$CH$_2$COOCH$_3$	[24]
1, 2	—Cl	—NH—C$_6$H$_4$—4-OH	[24]
1	— [N(CH$_2$CH$_2$)$_2$N]CH$_2$CH$_2$COOCH$_3$	—NH—C$_6$H$_4$—4-OH	[24]
1	—NH—C$_6$H$_4$—4-OH	—NH—CH$_2$CH(OH)CH$_2$OH	[24]
0	—Cl	—NH—CH$_2$CH$_2$CH$_2$OH	

[42,43,45–47]. During the verification process, the correctness and accuracy of the prediction are tested. Finally, the proposed optimized model of ANN can be used for the prediction of biological activity.

2.1. Dataset

Data published in the literature [19–24] were collected for a total of 129 1,3,5-triazinyl aminobenzene sulfonamide derivatives that had undergone biological testing. The research included determined the biological activity against the isozyme hCA II (one of the most physiologically relevant) and an isozyme of hCA IX (tumour associated). The selectivity ratio for the inhibition of isozyme IX over isozyme II was also provided in the studies. A data set containing information about the molecular structure of the compounds (Table 1), their inhibitory activity against hCA II and hCA IX, and their selectivity was built from these sources.

2.2. Descriptors

It has already been proved that descriptors based only on the structure are suitable for the prediction of different properties [32,39,41]. However, these descriptors need sophisticated calculation by specialized software [31,48]. Just a few examples of the use of only constitutional descriptors in prediction modelling have been published [49,50]. For instance, Keshavarz et al. recently published the use of constitutional descriptors besides with descriptors characterizing molecular fragments as inputs for a multiple linear regression method to predict the toxicity of some nitroaromatic compounds [50].

The constitutional descriptors are mostly used in combination with other types of descriptors (e.g., topological and/or quantum chemical descriptors) [35,36,38,51]. For example, Srivastava et al. developed a new classification model to predict the biofilm inhibitory potential of small molecules, using constitutional descriptors in combination with topological, connectivity, and MOE-type descriptors [52]. Moreover, the combination of constitutional, topological, geometrical, electrostatic, and quantum-chemical hypothetical descriptors was successfully used for prediction of the toxicity of binary mixtures of different chemicals [53]. The used computational methods were forward stepwise multiple linear regression and non-linear radial basis function neural networks [53].

In the present study, we examined descriptors, as selected in the literature cited in Table 1, coding the structure and also descriptors related to physicochemical properties.

In the first stage (**Phase 1**), the simplest way of coding was studied: a numerical code was given to each substituent (i.e. three substituents = three descriptors). However, this approach was found to be inappropriate, as a high RMS value equal to 358.6 was obtained.

In the next phase (**Phase 2**), four descriptors were assigned to each substituent. **Descriptor 1** describes the presence or absence of a heterocyclic skeleton, common to all compounds (*1*, triazine), **Descriptors 3** and **9** describe the atom through which the substituent is attached to the skeleton, **Descriptors 4** and **10** indicate if the main chain of the substituent is linear or cyclic, **Descriptors 6** and **12** determine the number of carbons in the main chain of the substituent, and **Descriptors 7** and **13** define substituents attached to the main chain. Despite the use of four descriptors per substituent, the RMS value was still too high (RMS = 254.3), indicating the low importance of or lack of information in the descriptors.

Therefore, in **Phase 3**, **Descriptor 2** was added, identifying the number of carbons in the linking chain between sulfonamide and triazine. The use of this descriptor reduced the resulting RMS to 91.6, indicating that the sulfonamide substituent is crucial for biological activity against hCA.

In **Phase 4**, in order to further reduce the RMS value, the structure was encoded using a higher number of descriptors. The result of **Phase 4** was the final set of descriptors coding the structure (**Descriptors 1–14**; see Table 2); the calculated RMS was 69.2.

An example of the coding for a specific structure is given in Fig. 3.

In **Phases 5** and **6**, the descriptors related to physicochemical properties were used in the computations. In **Phase 5**, polarizability α (**Descriptor 15**) and log *P* (*P* is the partition coefficient) (**Descriptor 16**) were added, leading to an RMS value of 47.3. In **Phase 6**, molecular weight MW (**Descriptor 17**), Parachor P (**Descriptor 18**), molar refractivity *A* (**Descriptor 19**), molar volume *V*m (**Descriptor 20**), and surface tension γ (**Descriptor 21**) were included. The RMS value obtained using the model with 21 descriptors was 9.1, showing a significant improvement; however, a completely satisfactory result was still not obtained.

In **Phase 7**, **Descriptor 22** (nPip, total number of piperazine rings in the structure), **Descriptor 23** (nBz, number of benzene rings), **Descriptor 24** (nHAcc, number of H acceptors), and **Descriptor 25** (nHDon, number of H donors) were therefore added. Although the application of descriptors in **Phase 7** led to a decrease in the RMS value to 2.1, the values of residuals were still high, and a large number of data defined as outliers were obtained (almost 35%).

This problem was solved by the addition of a new descriptor (**Phase 8**, **Descriptor 26**), which divides compounds in a dataset of six groups (Subsets). The introduction of this additional descriptor increased the number of inputs (and thus the information content of the data) and therefore enabled the better prediction of the biological activity of structurally similar compounds. The addition of **Descriptor 26** proved to be a significant step as it, resulted in lower values of residuals, and no outliers were observed.

In summary, the above-listed descriptors (i) code the structure, (ii) describe physicochemical properties, and/or (iii) have a different meaning (e.g. the number of H donors in the structure, division into the subgroups, etc.).

Table 2

Characterization of Descriptors used in this study. When the substituent was not present, the value was assigned to *0*.

Descriptor	Significance of descriptor	Symbol
Descriptor 1	presence or absence of a heterocyclic skeleton, common to all compounds	*1*, triazine
Descriptor 2	the type of sulfonamide group (distinguished by the number of the CH_2 groups) bonded to the triazine skeleton	*2*, in the absence of CH_2; *3*, one CH_2; *4*, two CH_2
Descriptor 3 and 9	through which atom the substituent is bonded to the skeleton	*5*, Cl, i.e. no substitution; *6*, O; *7*, N
Descriptor 4 and 10	indication if the main chain of the substituent is linear unbranched, linear branched, or cyclic	*8*, linear unbranched; *9*, linear branched; *10*, cyclic
Descriptor 5 and 11	the position of branching in the case of a linear branched chain or, in the case of a cyclic substituent, indicate the nature of the cycle	linear branched chain: *11*, C_1; *12*, N_3; *13*, N_0; *14*, C_2; *15*, C_1 and C_3; *16*, C_1 and C_2 cyclic substituent: *17*, piperazine substituted in position N4; *18*, morpholine; *19*, pyrrolidine
Descriptor 6 and 12	the number of carbons in a linear chain except for carbons of functional groups; for example, COOH, COOR, CH_3	*20*, $(C)_1$; *21*, $(C)_2$; *22*, $(C)_3$; *23*, $(C)_4$
Descriptor 7 and 13	substituent attached to the branching in positions indicated by **Descriptors 5** and **11**	*24*, CH_3; *25*, CH_2CH_3; *26*, $2\times$ (CH_3); *27*, OH; *28*, CH_2OH; *29*, Ph; *30*, $COOCH_3$; *31*, CH_2COOCH_3; *32*, $CH_2CH_2COOCH_3$; *33*, C_6H_5-3-Cl-4-F; *34*, C_6H_5-3-CH_3; *35*, C_6H_5-4-OH; *36*, $(C)_1$-COOH; *37*, $(C)_1$-COOH and $(C)_2$-CH_3; *38*, $(C)_1$-COOH and $(C)_3$-CH_3
Descriptor 8 and 14	the last structural moiety attached to the chain	*39*, CH_3; *40*, OH; *41*, COOH; *42*, $COOCH_3$; *43*, NH_2; *44*, SCH_3; *45*, C_6H_5; or *0* when no substituent was present
Descriptor 15	polarizability	α $[10^{-24}$ $cm^3]$
Descriptor 16	log of partition coefficient	log *P*
Descriptor 17	molecular weight	MW [g/mol]
Descriptor 18	value of Parachor	P $[cm^3]$
Descriptor 19	molar refractivity	A $[cm^3]$
Descriptor 20	molar volume	Vm $[cm^3/mol]$
Descriptor 21	surface tension	γ [dyn/cm]
Descriptor 22	total number of piperazine rings in a structure	nPip
Descriptor 23	total number of benzene rings	nBz
Descriptor 24	total number of H acceptors present in a structure	nHAcc
Descriptor 25	total number of H donors present in a structure	nHDon
Descriptor 26	divides the compounds into six groups	described with numbers 1 to 6

2.3. Search for optimized ANN architecture

In the first stage, we examined the use of the following ANN architecture: input, one hidden layer, one output. As an example, the construction of the 'optimized' network for **Subset 1** is described below. Fig. 4 presents the RMS as a function of nodes in the hidden layer and demonstrates that the optimal value of *n* (number of nodes in the hidden layer) is close to 4–5. The best correlation between predicted and experimental biological activity values in the training step was observed for five nodes in the hidden layer. Subsequently, the architecture of the network suitable for modelling the relationship between the structure of compounds and their biological activity was found to be (**26**, **5**, **1**).

2.4. Modelling and verification

From the plot of experimental vs. predicted inhibition activity, it can be clearly seen that the experimental and predicted inhibition activities fit well in both the training and verification phases (results for **Subsets 2** and **3** are shown in Figs. 5 and 6 as examples of modelling and verification, respectively).

2.5. Examination of the importance of individual descriptors

The importance of individual descriptors was subsequently examined by the leave-one-out method. One of individual descriptors (or series of descriptors) was taken out of calculation, and the value of the RMS was determined (Fig. 7). If the elimination of the descriptor was not reflected in a significant decrease in the RMS value, then its importance could be considered to be diminished, and it could be ignored.

When **Descriptors 1–14** were removed en masse from the modelling, the RMS value rose significantly (RMS = 1.88). On the contrary, when **Descriptors 15–25** were removed en masse from the modelling, the value of the RMS (RMS = 0.0039) was equal to that of a model containing these descriptors. However, for both cases, we could observe high values of residuals, implying that, in each group of descriptors (i.e. based only on structure and those based on physicochemical properties), there are some important descriptors, and both of the groups must be taken into consideration.

It was found that the elimination of **Descriptor 2** presented the most significant effect on the value of the RMS. This is to be expected, because **Descriptor 2** determines the type of sulfonamide group, which is the key substituent with respect to inhibition activity. Other important descriptors based on structure are **Descriptors 4** and **10,** which indicate if the main chain of the substituent is linear unbranched, linear branched, or cyclic. Similarly, elimination of **Descriptors 7**, **8**, **13**, and **14**, giving information about the functional groups of substituents, led to significantly higher RMS values. From the second group of descriptors (based on physicochemical properties and others), the most significant change in the RMS value was associated with **Descriptors 19** (molar refractivity), **21** (surface tension), and **24** (number of H acceptors). Molar refractivity correlates with molecular weight, density, and refractive index. Surface tension correlates with molecular weight, density, and molar volume. **Descriptor 24** gives the number of H acceptors present in a structure, which explains its importance because this number provides information about the tendency of the molecule to form hydrogen bonds in general, in particular with the hydroxy groups of CA.

After the evaluation of the importance of individual descriptors, a value of RMS = 0.2 was taken as the limit value. Thus, descriptors whose elimination from modelling led to RMS values lower than 0.2 were ignored (**Descriptors 1**, **3**, **6**, **9**, **12**, **13**, **15**, **17**, **18**, **22**, **23**, and **25**). The new model employing only **Descriptors 2**, **4**, **5**, **7**, **8**, **10**, **11**, **14**, **16**, **19**, **20**, **21**, and **24** was trained further. The RMS value of the model was then 0.1137, which is comparable to the RMS of the model using all of the descriptors. The graph of residuals (Fig. 8; example **Subset 1** of residuals for inhibition activity against hCA IX) and the correlation between experimental and predicted inhibition activity values (Fig. 9;
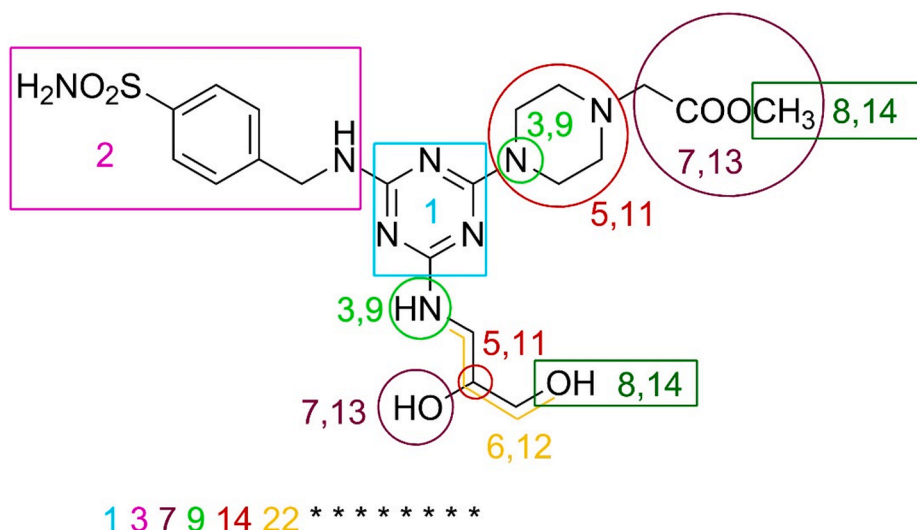
Fig. 3. Scheme of coding of the structure of drugs (code sequences are the numbers given below the illustrated structure).
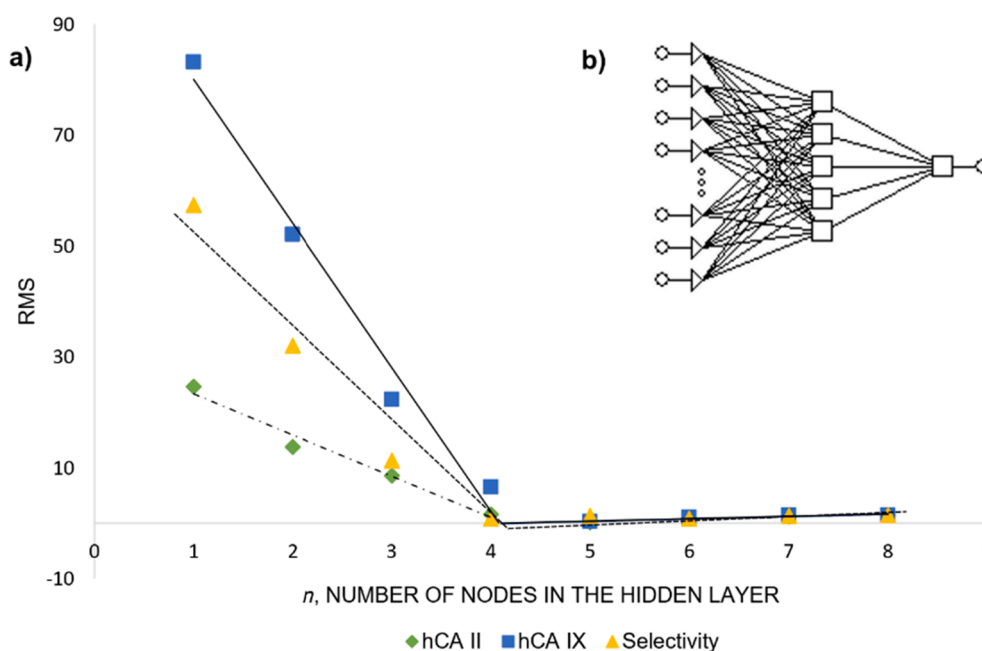

Fig. 4. a) Dependence of RMS as a function of the number of nodes (n) in the hidden layer. b) Design of the 'optimized' network.

example **Subset 1**) are in a good agreement with the results of the model using all of the descriptors.

However, the predictive ability of the ANN model using only descriptors **2, 4, 5, 7, 8, 10, 11, 14, 16, 19, 20, 21,** and **24** was rather weak (Table 3). This is to be expected, partly because each of the descriptors describes a different part of the molecule and partly because the biological activity is affected by the entire structure of the molecule.

As a result of the examination of the importance of individual descriptors, we conclude that to maintain the high accuracy of prediction, no descriptors can be eliminated. Even though the importance of some descriptors is weak, the use of them improve the accuracy of prediction.

2.6. Prediction of biological activity: proof of concept

The prediction of the biological activity of 129 compounds using ANN technology and newly developed molecular descriptors not requiring chemical quantum computation was successfully demonstrated in this work. In addition, the modelling of the biological activity of not-yet-synthesized compounds was enabled by the ANN database.

Structure-based molecular descriptors for Compound **1** (Fig. 10) were used for the ANN model, and the values of inhibition activity against hCA II and hCA IX and the selectivity (hCA II/hCA IX) were predicted. Compound **1** was then synthesized, and the inhibition activity was determined experimentally. The predicted and experimental values were in good agreement. The predicted value of inhibition activity (K_i) against hCA II was 42.82 nM, which was very close to the experimental value of 44.8 nM (RMS = −1.977). A slightly lower level of agreement between the predicted and experimental values was obtained for the prediction of inhibitory activity against hCA IX (predicted K_i = 239.8 nM, experimental K_i = 228.9 nM) but even this result is satisfying. The values for predicted selectivity (0.2281) and experimentally determined selectivity (0.19) were almost the same (RMS = 0.03807).
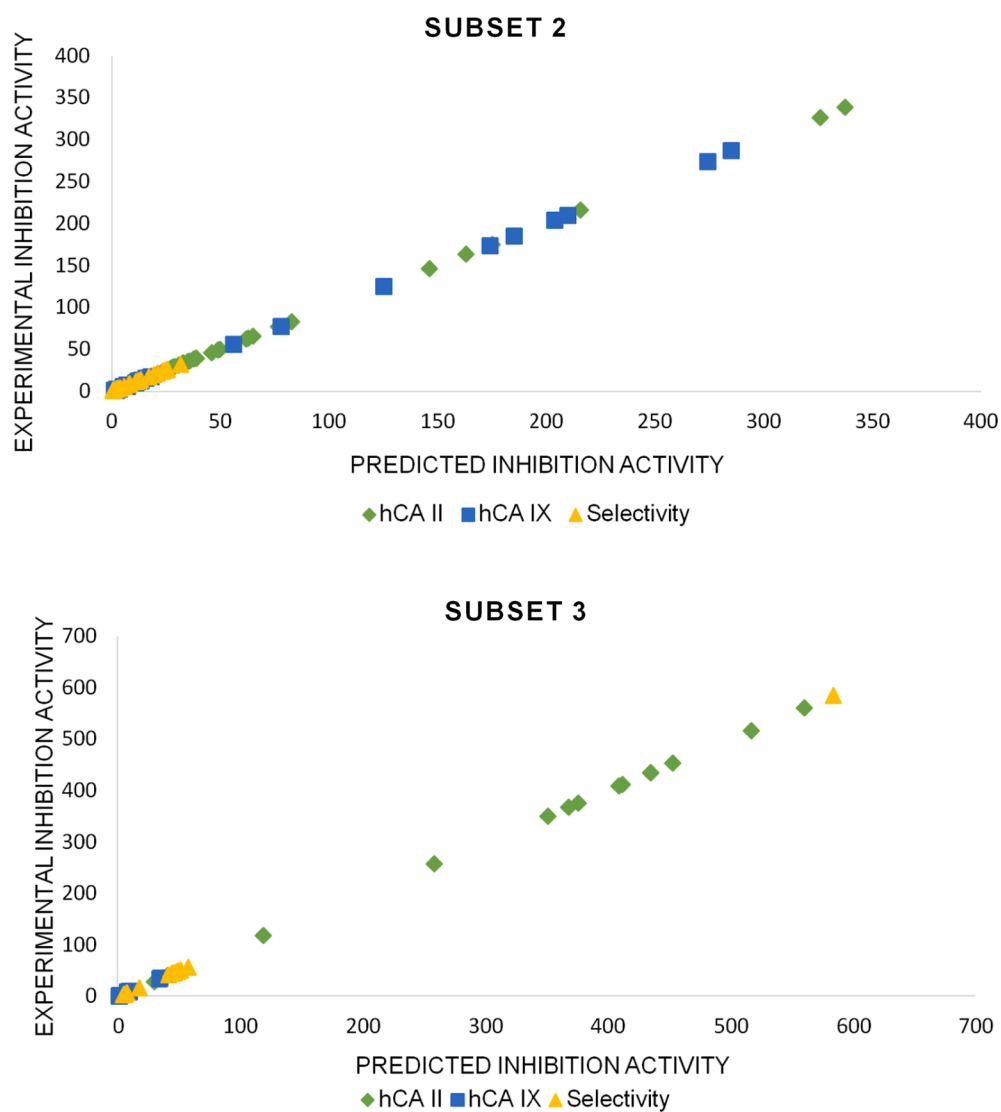
Fig. 5. Modelling: experimental vs. predicted inhibition activity for **Subset 2** and **Subset 3** as examples.
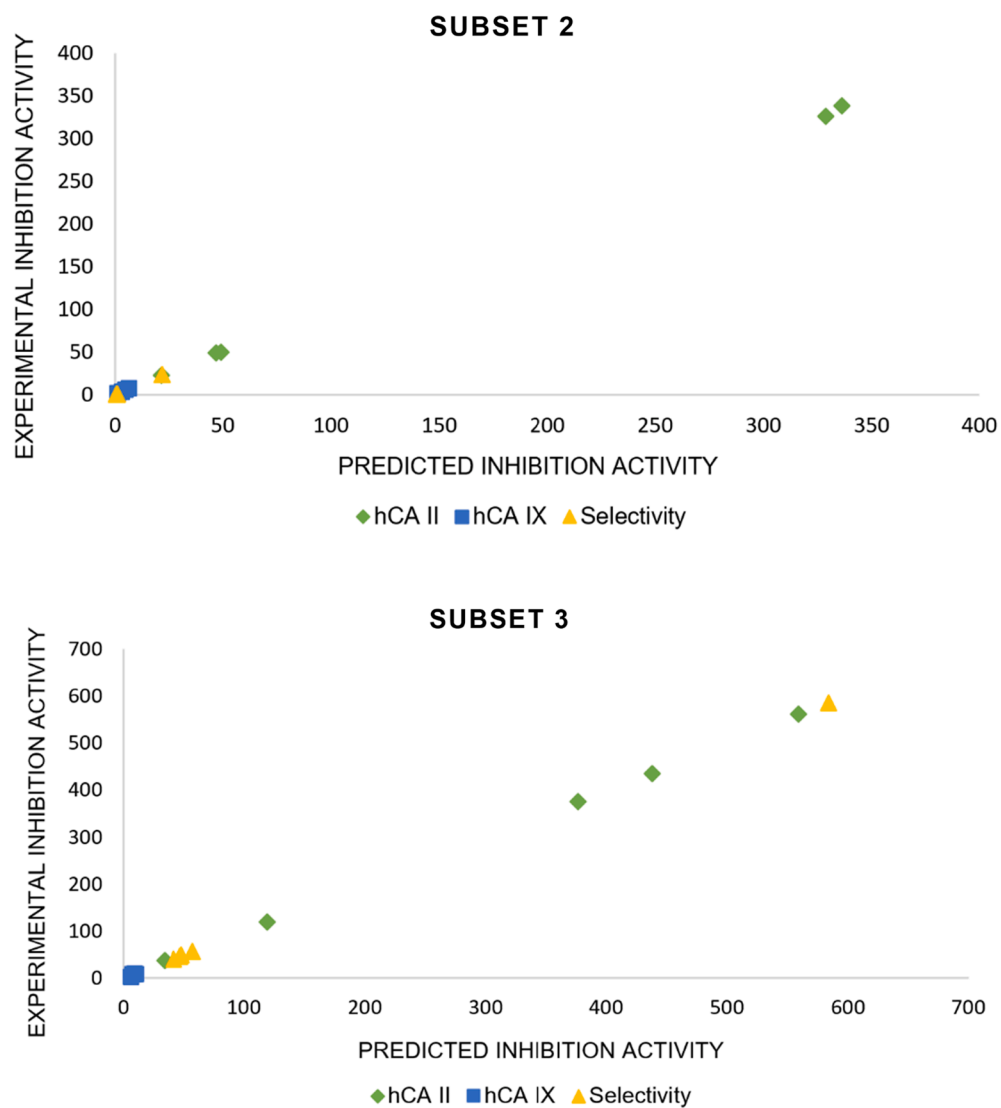
Fig. 6. Verification: experimental vs. predicted inhibition activity for **Subset 2** and **Subset 3** as examples.
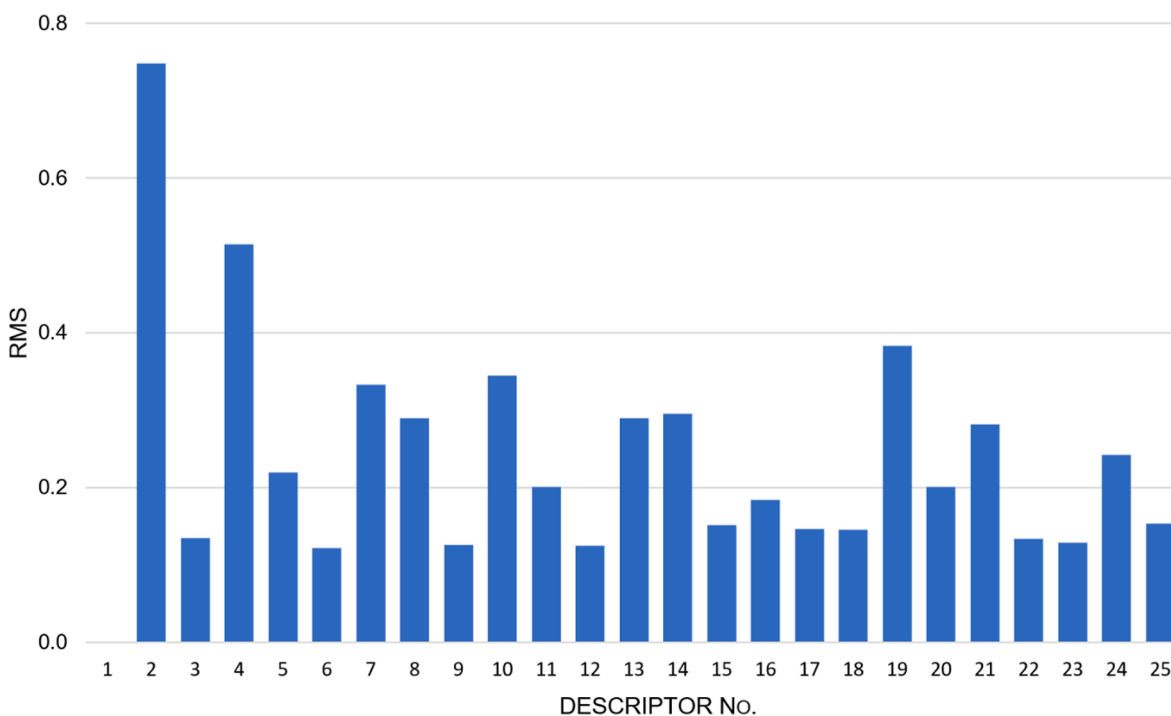
Fig. 7. Subset 1: Examination of the importance of individual descriptors.

The CA inhibition assay is described in detail in the Supporting Information.

3. Experimental

3.1. Coding of the structure of drugs and determination of descriptors

The structures of the compounds analyzed were coded according to the key presented in Table 2. An example of code sequences is given in Fig. 3. **Descriptors 15–21**, which describe the physical properties of the molecules, were calculated from the structures using the ChemSketch program [54].

3.2. Instrumentation

All calculations were performed on an AMILO Pro computer (Fujitsu Siemens Computers; Intel® Core™2 CPU, T5500@1.66 GHz, 1.66 GHz, 3.00 GB RAM, Operating system: Microsoft Windows XP Professional,
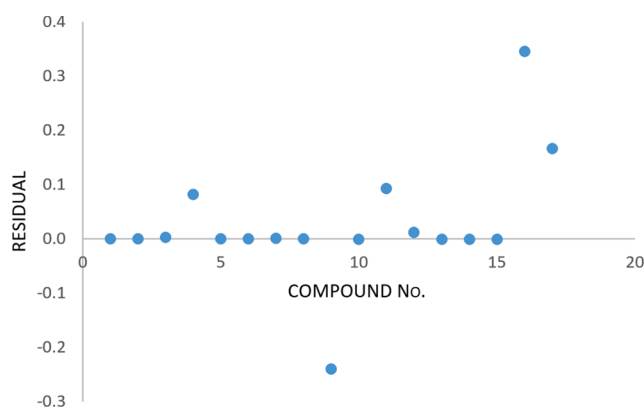

Fig. 8. Example (**Subset 1**) of residuals for the modelling of inhibition activity against hCA IX after the elimination of redundant descriptors (**Descriptors 1, 3, 6, 9, 12, 13, 15, 17, 18, 22, 23,** and **25**).
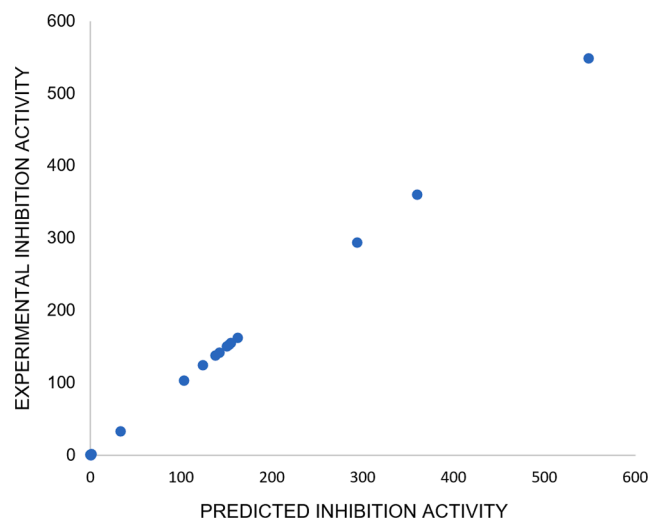

Fig. 9. Example (**Subset 1**) of the modelling of ANNs after the elimination of redundant descriptors: experimental vs. predicted inhibition activity against hCA IX.

Table 3

Example (**Subset 1**): Verification of inhibition activity against hCA IX after the elimination of redundant descriptors.

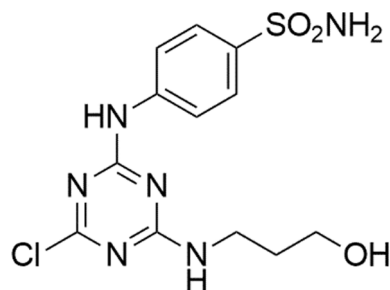| Compound	Predicted inhibition activity	Experimental inhibition activity	Residual
6	141.2	138	3.17
3	72.46	33	39.46
10	12.15	153	140.85

Fig. 10. Structure of compound **1.**

Table 4
Optimized network assessment: the function of the RMS as a function of nodes in the hidden layer.

Subset 1 – hCA II		Subset 1 – hCA IX		Subset 1 – Selectivity	
Hidden layers	RMS	Hidden layers	RMS	Hidden layers	RMS
1	24.6	1	83.2	1	57.5
2	13.8	2	52.1	2	32.0
3	8.6	3	22.3	3	11.3
4	1.6	4	6.5	4	0.8
5	0.1	5	0.2	5	1.3
6	0.9	6	1.0	6	0.8
7	1.2	7	1.4	7	1.3
8	1.5	8	1.5	8	1.5

Version 2002, and Service Pack 2). The Trajan 4.0 software package [55] was used for the calculations. For the calculation of physico-chemical properties (Descriptors 15–21), ChemSketch from ACD/Labs [54] was applied.

3.3. Search for optimized ANN architecture

As an example, the construction of the 'optimized' network for **Subset 1** is described below. The RMS values were determined as a function of nodes in the hidden layer (see Table 4). The optimal number of nodes in the hidden layer was determined from the graph presented in Fig. 4. Subsequently, the architecture of the network suitable for modelling the relationship between the structure of compounds and

their biological activity was found to be (**26, 5, 1**).

3.4. Modelling and verification: statistical data

Statistical data for the modelling of individual subsets in the training phase are given in Table 5.

3.5. Modelling and verification: calculation of residual values

For each of the compounds of the dataset, the residual value was calculated according to the following equation:

Residual = predicted value − experimental value

Residual values calculated for **Subsets 2** and **3** (as examples) in the training and verification processes, respectively, are recorded in Figs. 11 and 12.

3.6. Synthesis of 4-({4-Chloro-6-[(3-hydroxypropyl)amino]-1,3,5-triazin-2-yl}amino) benzenesulfonamide

3.6.1. General information

All reagents were purchased from commercial suppliers (Sigma-Aldrich, Darmstadt, Germany) and used as supplied, without further purification. The catalytic system of Cu(I)-supported on a weakly acidic resin was purchased from www.entwickchemicals.com.

The reaction was monitored by TLC performed on precoated Silica Gel 60 F254 plates (Merck, Darmstadt, Germany). Methanol was used as an eluent; UV light (254 and 356 nm) and ninhydrin reagent were used for the detection of spots at 180 °C. NMR spectra were recorded on a DRX 500 Avance (Bruker Biospin, Billerica, USA) spectrometer using TMS as an internal standard. The FTIR spectrum was obtained on an Alpha II FTIR Spectrometer (Bruker, Billerica, USA) equipped with the ATR module. The melting point (uncorrected) was recorded on a Kofler's block Boetius Rapido PHMK 79/2106 (Wagetechnik, Dresden, Germany) with a temperature gradient of 4 °C/min^{-1}.

3.6.2. Synthetic procedure

The starting compound 4-{[(4,6-dichloro-1,3,5-triazin-2-yl)amino] methyl}benzene-1-sulfonamide was prepared according to the literature [19].

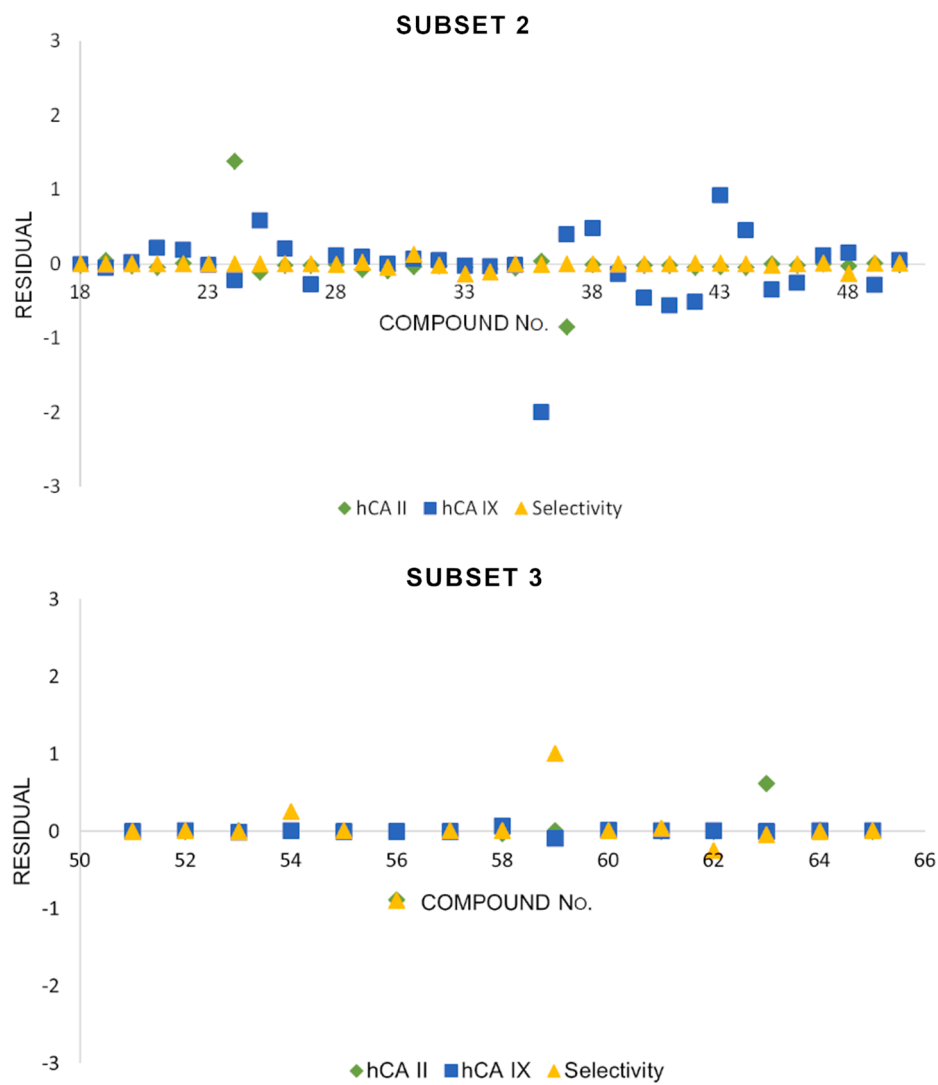Compound **4-({4-Chloro-6-[(3-hydroxypropyl)amino]-1,3,5-**

Table 5
Statistical data for the modelling of individual subsets.

Subset	Output	Mean[a]	S.D.[b]	Error Mean[c]	Error S.D.[d]	S.D. Ratio[e]
1	hCA II	94.94	90.18	0.0121	0.1278	0.001417
	hCA IX	139.20	148.00	0.2044	0.4043	0.002731
	Selectivity	68.71	178.30	0.02782	0.6789	0.003807
2	hCA II	72.94	82.70	−0.004847	0.2899	0.003505
	hCA IX	52.42	87.20	−0.03131	0.4714	0.005406
	Selectivity	10.09	9.87	−0.004079	0.04421	0.004477
3	hCA II	320.10	179.40	−0.01723	0.2873	0.001601
	hCA IX	10.61	10.16	0.001609	0.02436	0.002397
	Selectivity	73.09	142.60	0.008393	0.3718	0.002606
4	hCA II	47.22	34.63	0.002473	0.07370	0.002128
	hCA IX	13.76	10.08	0.001438	0.02943	0.002921
	Selectivity	6.76	8.903	−0.003487	0.03093	0.003474
5	hCA II	3.15	0.46	−0.03313	0.1824	0.3972
	hCA IX	0.86	0.73	0.02046	0.06651	0.09056
	Selectivity	0.86	0.73	−0.01441	0.2713	0.3694
6	hCA II	72.39	82.30	−0.1031	0.5444	0.006615
	hCA IX	90.29	105.80	−0.002762	0.2084	0.001969
	Selectivity	3.176	4.53	−0.005323	0.04026	0.008881

[a] Mean, the average of a data.
[b] S.D., standard deviation.
[c] Mean Error.
[d] Standard Error of S.D.
[e] S.D. Ratio, the error: data standard deviation ratio.

Fig. 11. Residuals for ANN modelling: results for **Subset 2** and **Subset 3** as examples.
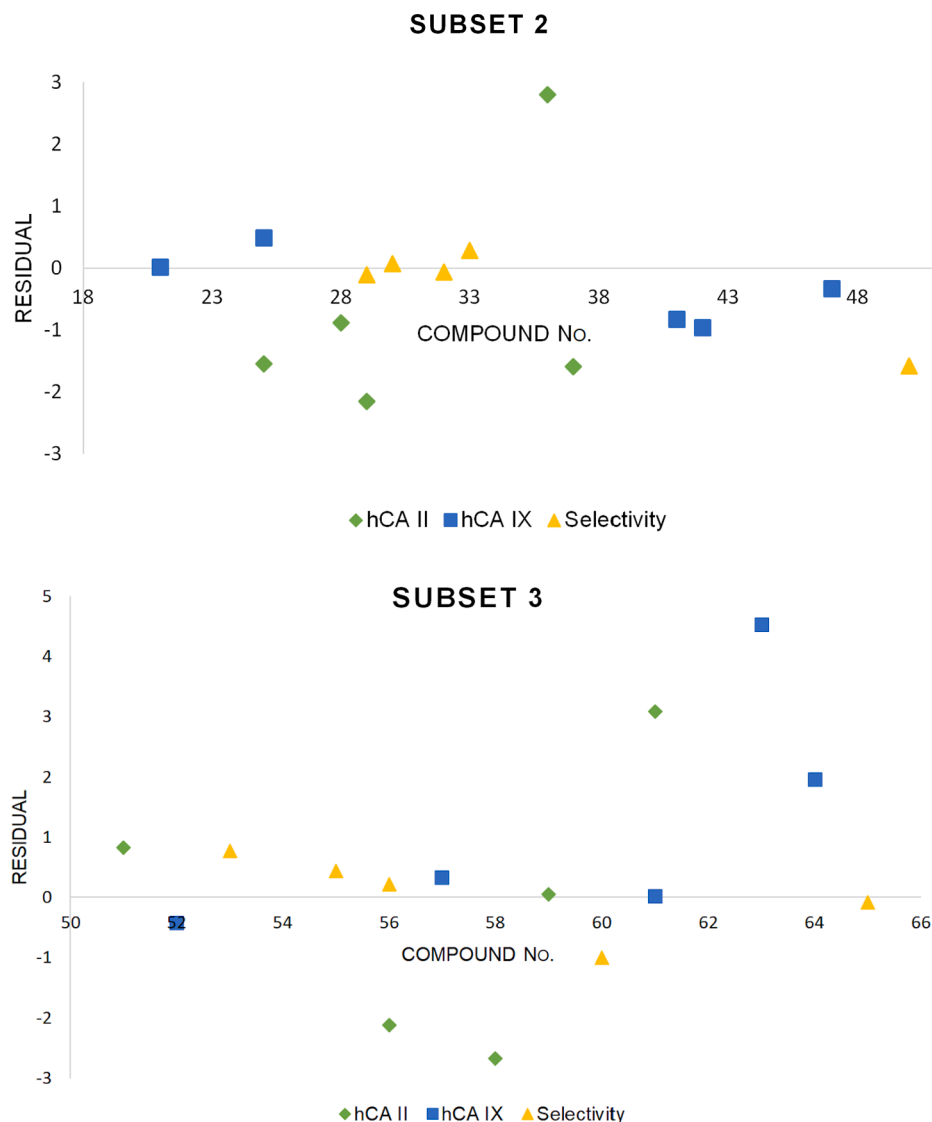
SUBSET 2

SUBSET 3

Fig. 12. Residuals for verification: results for **Subset 2** and **Subset 3** as examples.

triazin-2-yl}amino) benzenesulfonamide was prepared according to the methodology published in [24,56]. Starting dichlorotriazinyl benzenesulfonamide (1 mmol) was dissolved in 10 mL of DMF. One mmol of solid anhydrous potassium carbonate was added gradually, and the mixture was stirred for 10 min. Then, 1 mmol of 3-aminopropanole was added portion-wise. Finally, 2.5% mol of supported Cu(I) ions were added into the reaction mixture. The reaction was stirred at 35 °C until the maximum conversion of starting reactants was achieved (monitored by TLC). After the completion of a reaction, the catalyst and salt were filtered off. Crushed ice was then added into the solution, and the precipitate formed was collected by filtration. The crude product was dissolved in hot acetone and precipitated by the addition of isopropyl alcohol.

4-({4-Chloro-6-[(3-hydroxypropyl)amino]-1,3,5-triazin-2-yl}amino)benzenesulfonamide (1): $C_{12}H_{15}ClN_6O_3S$; 358.80 g/mol; white solid; yield 95.5%; mp 223–225 °C; ^1H NMR: δ_H *(ppm, 500 MHz, DMSO-d$_6$)* 7.76 (2H, d, $J = 8.1$ Hz, CH$_{Ar}$), 7.51 (2H, d, $J = 8.1$ Hz, CH$_{Ar}$), 3.44–3.41 (2H, m, C$\underline{H}_2$—OH), 3.32–3.29 (2H, m, NH—C$\underline{H}_{2alc}$),

1.58–1.56 (2H, m, CH$_{2alc}$); ^{13}C NMR: δ_C *(ppm, 125 MHz, DMSO-d$_6$)* 168.2 (C—Cl), 165.8 (C—N$_{alc}$), 162.7 (C—N$_{sulf}$), 144.7 (C), 143,3 (C—SO$_2$NH$_2$), 128.1 (CH$_{Ar}$), 126.8 (CH$_{Ar}$), 58.3 (CH$_2$—OH), 41.2 (NH—CH$_{2alc}$), 32.9 (CH$_{2alc}$); *IR: ν_{max} (cm^{-1})* 3353 (OH, NH, NH$_2$), 3258, 2952 (CH$_{Ar}$), 2926 (CH$_{aliph}$), 1604 (C=C$_{Ar}$), 1434, 1391, 1153 (SO$_2$NH$_2$)

4. Conclusions

A new series of simple molecular descriptors for 1,3,5-triazinyl sulfonamide derivatives was developed and evaluated. The descriptors enabled the successful employment of ANN technology in modelling the relationship between the structures of more than one hundred 1,3,5-triazinyl sulfonamide derivatives and their biological activity. This approach is simple and avoids extensive, sophisticated calculation by specialized software and the necessity to evaluate descriptors using statistical programs. The ANN architecture developed enables the prediction of the biological activity for unknown (not-yet-synthesized)

compounds, which brings substantial savings in time, effort, and funds.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bioorg.2020.104565.

References

[1] Cancer: Fact sheet, 2019. http://www.who.int/news-room/fact-sheets/detail/cancer (accessed May 15, 2019).

[2] F. Trudu, F. Amato, P. Vaňhara, T. Pivetta, E.M. Peña-Méndez, J. Havel, Coordination compounds in cancer: Past, present and perspectives, J. Appl. Biomed. 13 (2015) 79–103, https://doi.org/10.1016/j.jab.2015.03.003.

[3] G.M. Cooper, The Cell: A Molecular Approach, second ed., ASM Press, Washington, 1996.

[4] G. Höpfl, O. Ogunshola, M. Gassmann, HIFs and tumors—causes and consequences, Am. J. Physiol. Regul. Integr. Comp. Physiol. 286 (2004) R608–R623, https://doi.org/10.1152/ajpregu.00538.2003.

[5] L.M. Sherwood, E.E. Parris, J. Folkman, Tumor angiogenesis: therapeutic implications, N. Engl. J. Med. 285 (1971) 1182–1186, https://doi.org/10.1056/NEJM197111182852108.

[6] J.T. Erler, C.J. Cawthorne, K.J. Williams, M. Koritzinsky, B.G. Wouters, C. Wilson, et al., Hypoxia-mediated down-regulation of bid and bax in tumors occurs via hypoxia-inducible factor 1-dependent and -independent mechanisms and contributes to drug resistance, Mol. Cell. Biol. 24 (2004) 2875–2889, https://doi.org/10.1128/MCB.24.7.2875-2889.2004.

[7] S. Pastorekova, J. Pastorek, Cancer-related carbonic anhydrase isozymes, in: C. T. Supuran, A. Scozzafava, J. Conway (Eds.), Carbonic Anhydrase-Its Inhibitors and Activators, CRC, Boca Raton, FL, USA, 2004, p. 253.

[8] S. Pastorekova, S. Parkkila, J. Pastorek, C.T. Supuran, Carbonic anhydrases: current state of the art, therapeutic applications and future prospects, J. Enzyme Inhib. Med. Chem. 19 (2008) 199–229, https://doi.org/10.1080/14756360410001689540.

[9] E. Švastová, A. Hulíková, M. Rafajová, M. Zat'ovičová, A. Gibadulinová, A. Casini, et al., Hypoxia activates the capacity of tumor-associated carbonic anhydrase IX to acidify extracellular pH, FEBS Lett. 577 (2004) 439–445, https://doi.org/10.1016/j.febslet.2004.10.043.

[10] C. Wykoff, N.J. Beasley, P.H. Watson, K.J. Turner, J. Pastorek, et al., Hypoxia-inducible expression of tumor-associated carbonic anhydrases, Cancer Res. 60 (2000) 7075–7083.

[11] D. Luo, Z. Wang, J. Wu, C. Jiang, J. Wu, The role of hypoxia inducible factor-1 in hepatocellular carcinoma, Bio. Med Res. Int. 2014 (2014) e409272.

[12] N. Robertson, C. Potter, A.L. Harris, Role of carbonic anhydrase IX in human tumor cell growth, survival, and invasion, Cancer Res. 64 (2004) 6160–6165, https://doi.org/10.1158/0008-5472.CAN-03-2224.

[13] B. Mahon, M. Pinard, R. McKenna, Targeting carbonic anhydrase IX activity and expression, Molecules 20 (2015) 2323–2348, https://doi.org/10.3390/molecules20022323.

[14] M. Benej, S. Pastorekova, J. Pastorek, J. Carbonic Anhydrase IX: Regulation and Role in Cancer, in: S.C. Frost, R. McKenna, (Eds.) Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications, Subcellular Biochemistry: Springer, Netherlands, 2014, pp. 199–219.

[15] C.T. Supuran, Carbonic anhydrase inhibition and the management of hypoxic tumors, Metabolites 7 (2017) 48–60, https://doi.org/10.3390/metabo7030048.

[16] C.T. Supuran, Carbonic anhydrase inhibitors as emerging agents for the treatment and imaging of hypoxic tumors, Expert Opin. Invest. Drugs 27 (2018) 963–970, https://doi.org/10.1080/13543784.2018.1548608.

[17] S. Pastorekova, R.J. Gillies, The role of carbonic anhydrase IX in cancer development: links to hypoxia, acidosis, and beyond, Cancer Metast. Rev. 38 (2019) 65–77, https://doi.org/10.1007/s10555-019-09799-0.

[18] J. Pastorek, S. Pastorekova, Hypoxia-induced carbonic anhydrase IX as a target for cancer therapy: from biology to clinical use, Semin. Cancer Biol. 31 (2015) 52–64, https://doi.org/10.1016/j.semcancer.2014.08.002.

[19] V. Garaj, L. Puccetti, G. Fasolis, J.-Y. Winum, J.-L. Montero, A. Scozzafava, et al., Carbonic anhydrase inhibitors: synthesis and inhibition of cytosolic/tumor-associated carbonic anhydrase isozymes I, II, and IX with sulfonamides incorporating 1,2,4-triazine moieties, Bioorg. Med. Chem. Lett. 14 (2004) 5427–5433, https://doi.org/10.1016/j.bmcl.2004.07.087.

[20] V. Garaj, L. Puccetti, G. Fasolis, J.-Y. Winum, J.-L. Montero, A. Scozzafava, et al., Carbonic anhydrase inhibitors: novel sulfonamides incorporating 1,3,5-triazine moieties as inhibitors of the cytosolic and tumour-associated carbonic anhydrase isozymes I, II and IX, Bioorg. Med. Chem. Lett. 15 (2005) 3102–3108, https://doi.org/10.1016/j.bmcl.2005.04.056.

[21] A.K. Saluja, M. Tiwari, D. Vullo, C.T. Supuran, Substituted benzene sulfonamides incorporating 1,3,5-triazinyl moieties potently inhibit human carbonic anhydrases II, IX and XII, Bioorg. Med. Chem. Lett. 24 (2014) 1310–1314, https://doi.org/10.1016/j.bmcl.2014.01.048.

[22] F. Carta, V. Garaj, A. Maresca, J. Wagner, B.S. Avvaru, A.H. Robbins, et al., Sulfonamides incorporating 1,3,5-triazine moieties selectively and potently inhibit carbonic anhydrase transmembrane isoforms IX, XII and XIV over cytosolic isoforms I and II: solution and X-ray crystallographic studies, Bioorg. Med. Chem. 19 (2011) 3105–3119, https://doi.org/10.1016/j.bmc.2011.04.005.

[23] P. Mikuš, D. Krajčiová, M. Mikulová, B. Horváth, D. Pecher, V. Garaj, et al., Novel sulfonamides incorporating 1,3,5-triazine and amino acid structural motifs as inhibitors of the physiological carbonic anhydrase isozymes I, II and IV and tumor-associated isozyme IX, Bioorg. Chem. 81 (2018) 241–252, https://doi.org/10.1016/j.bioorg.2018.08.005.

[24] E. Havránková, J. Csöllei, D. Vullo, V. Garaj, P. Pazdera, C.T. Supuran, Novel sulfonamide incorporating piperazine, aminoalcohol and 1,3,5-triazine structural motifs with carbonic anhydrase I, II and IX inhibitory action, Bioorg. Chem. 77 (2018) 25–37, https://doi.org/10.1016/j.bioorg.2017.12.034.

[25] C.T. Supuran, Carbonic anhydrases: novel therapeutic applications for inhibitors and activators, Nat. Rev. Drug Discov. 7 (2008) 168–181, https://doi.org/10.1038/nrd2467.

[26] C.T. Supuran, J.-Y. Winum, Carbonic anhydrase IX inhibitors in cancer therapy: an update, Future Med. Chem. 7 (2015) 1407–1414, https://doi.org/10.4155/fmc.15.71.

[27] N. Yorulmaz, O. Oltulu, E. Eroğlu, Development of selective QSAR models and molecular docking study for inhibitory activity of sulfonamide derivatives against carbonic anhydrase isoforms II and IX, J. Mol. Struct. 1163 (2018) 270–279, https://doi.org/10.1016/j.molstruc.2018.02.107.

[28] M. Jaiswal, P.V. Khadikar, A. Scozzafava, C.T. Supuran, Carbonic anhydrase inhibitors: the first QSAR study on inhibition of tumor-associated isoenzyme IX with aromatic and heterocyclic sulfonamides, Bioorg. Med. Chem. Lett. 14 (2004) 3283–3290, https://doi.org/10.1016/j.bmcl.2004.03.099.

[29] E. Eroglu, H. Türkmen, A DFT-based quantum theoretic QSAR study of aromatic and heterocyclic sulfonamides as carbonic anhydrase inhibitors against isozyme, CA-II, J. Mol. Grap. Model. 26 (2007) 701–708, https://doi.org/10.1016/j.jmgm.2007.03.015.

[30] M.N. Deodhar, P.L. Khopade, M.G. Varat, Sulfonamide based β -carbonic anhydrase Inhibitors: 2d QSAR study, ISRN Med. Chem. 2013 (2013) 1–8, https://doi.org/10.1155/2013/107840.

[31] G. Melagraki, A. Afantitis, H. Sarimveis, O. Igglessi-Markopoulou, C.T. Supuran, QSAR study on para-substituted aromatic sulfonamides as carbonic anhydrase II inhibitors using topological information indices, Bioorg. Med. Chem. 14 (2006) 1108–1114, https://doi.org/10.1016/j.bmc.2005.09.038.

[32] E. Eroğlu, H. Türkmen, S. Güler, S. Palaz, O. Oltulu, A DFT-based QSARs study of acetazolamide/sulfanilamide derivatives with carbonic anhydrase (CA-II) isozyme inhibitory activity, Int. J. Mol. Sci. 8 (2007) 145–155, https://doi.org/10.3390/i8020145.

[33] A.K. Jain, R. Veerasamy, A. Vaidya, V. Mourya, R.K. Agrawal, QSAR analysis of some novel sulfonamides incorporating 1,3,5-triazine derivatives as carbonic anhydrase inhibitors, Med. Chem. Res. 19 (2010) 1191–1202, https://doi.org/10.1007/s00044-009-9262-0.

[34] S. Singh, A QSAR study on novel series of carbonic anhydrase inhibitors hCA IX—tumor-associated (hypoxia), Med. Chem. 8 (2012) 656–672, https://doi.org/10.2174/157340612801216391.

[35] R. Guha, E. Willighagen, A survey of quantitative descriptions of molecular structure, Curr. Top. Med. Chem. 12 (2012) 1946–1956, https://doi.org/10.2174/156802612804910278.

[36] A. Bender, R.C. Glen, A discussion of measures of enrichment in virtual screening: comparing the information content of descriptors with increasing levels of sophistication, J. Chem. Inf. Model. 45 (2005) 1369–1375, https://doi.org/10.1021/ci0500177.

[37] S. Gosav, M. Praisler, D.O. Dorohoi, G. Popa, Structure–activity correlations for illicit amphetamines using ANN and constitutional descriptors, Talanta 70 (2006) 922–928, https://doi.org/10.1016/j.talanta.2006.05.054.

[38] D.C. Young, Computational Drug Design: A Guide for Computational and Medicinal Chemists, Wiley, Hoboken, N.J., 2009.

[39] R.S. kord Abadi, A. Alizadehdakhel, S.T. Paskiabei, A DFT and QSAR study of several sulfonamide derivatives in gas and solvent, J. Korean Chem. Soc. 60 (2016) 225–234, https://doi.org/10.5012/jkcs.2016.60.4.225.

[40] V. Pandey, Modeling of carbonic anhydrase (Ii) inhibitory activities of sulphonilamide schiff bases by artificial neural network trained with different numerical techniques, Int. J. Pharm. Pharm. Sci. 10 (2018) 202–207, https://doi.org/10.22159/ijpps.2018v10i1.22775.

[41] A. Maleki, H. Daraei, L. Alaei, A. Faraji, Comparison of QSAR models based on combinations of genetic algorithm, stepwise multiple linear regression, and artificial neural network methods to predict Kd of some derivatives of aromatic sulfonamides as carbonic anhydrase II inhibitors, Russ. J. Bioorgan. Chem. 40 (2014) 61–75, https://doi.org/10.1134/S106816201306006X.

[42] H. Chauhan, J. Bernick, D. Prasad, V. Masand, The role of artificial neural networks on target validation in drug discovery and development, in: M. Puri, Y. Pathak, V. K. Sutariya, S. Tipparaju, W. Moreno (Eds.), Artificial Neural Network For Drug Design, Delivery And Disposition, Academic Press, USA, 2016, pp. 15–27. https://doi.org/10.1016/B978-0-12-801559-9.00002-8.

[43] T. Pivetta, F. Isaia, F. Trudu, A. Pani, M. Manca, D. Perra, et al., Development and validation of a general approach to predict and quantify the synergism of anti-cancer drugs using experimental design and artificial neural networks, Talanta 115 (2013) 84–93, https://doi.org/10.1016/j.talanta.2013.04.031.

[44] M.B. Mikulová, D. Kružlicová, D. Pecher, C.T. Supuran, P. Mikuš, Synthetic strategies and computational inhibition activity study for triazinyl-substituted benzenesulfonamide conjugates with polar and hydrophobic amino acids as inhibitors of carbonic anhydrases, Int. J. Mol. Sci. 21 (2020) 3661, https://doi.org/10.3390/ijms21103661.

[45] S. Agatonovic-Kustrin, R. Beresford, Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research, J. Pharm. Biomed. Anal. 22 (2000) 717–727, https://doi.org/10.1016/S0731-7085(99)00272-1.

[46] F.R. Burden, Using artificial neural networks to predict biological activity from simple molecular structural considerations, Quant. Struct.-Act. Relat. 15 (1996) 7–11, https://doi.org/10.1002/qsar.19960150103.

[47] F. Amato, A. López, E.M. Peña-Méndez, P. Vaňhara, A. Hampl, J. Havel, Artificial neural networks in medical diagnosis, J. Appl. Biomed. 11 (2013) 47–58, https://doi.org/10.2478/v10136-012-0031-x.

[48] A. Guerra, P. Gonzalez-Naranjo, N.E. Campillo, H. Cerecetto, M. Gonzalez, J. A. Paez, Artificial neural networks based on CODES descriptors in pharmacology: identification of novel trypanocidal drugs against chagas disease, Curr. Comput. Aided Drug Des. 9 (2013) 130–140, https://doi.org/10.2174/1573409911309010012.

[49] A. Ion, S. Gosav, M. Praisler, Screening for NBOMe hallucinogens based on artificial neural networks and structural descriptors, in: 2019 E-Health and Bioengineering Conference (EHB), Iasi, Romania, 2019, pp. 1–4. https://doi.org/10.1109/EHB47216.2019.8970048.

[50] M.H. Keshavarz, A.R. Akbarzadeh, A simple approach for assessment of toxicity of nitroaromatic compounds without using complex descriptors and computer codes, SAR and QSAR, Environ. Res. 30 (2019) 347–361, https://doi.org/10.1080/1062936X.2019.1595135.

[51] T. Gaudin, P. Rotureau, G. Fayet, Mixture descriptors toward the development of quantitative structure-property relationship models for the flash points of organic mixtures, Ind. Eng. Chem. Res. 54 (2015) 6596–6604, https://doi.org/10.1021/acs.iecr.5b01457.

[52] G.N. Srivastava, A.S. Malwe, A.K. Sharma, V. Shastri, K. Hibare, V.K. Sharma, Molib: A machine learning based classification tool for the prediction of biofilm inhibitory molecules, Genomics 112 (2020) 2823–2832, https://doi.org/10.1016/j.ygeno.2020.03.020.

[53] T. Wang, L. Tang, F. Luan, M.N.D.S. Cordeiro, Prediction of the toxicity of binary mixtures by QSAR approach using the hypothetical descriptors, Int. J. Mol. Sci. 19 (2018), https://doi.org/10.3390/ijms19113423.

[54] ACD/ChemSketch 2015 Pack2 (File Version C20H41, Build 78694, 08 Jun 2015) from ACD/Labs, Copyright@1994-2015, Advanced Chemistry Development, Inc.

[55] Trajan 4.0 software package, Trajan Software, Trajan neural network simulator, Release 4.0 D. Durham, UK: Trajan Software Ltd., 1999.

[56] E. Havránková, J. Csöllei, P. Pazdera, New approach for the one-pot synthesis of 1,3,5-triazine derivatives: application of Cu(I) supported on a weakly acidic cation-exchanger resin in a comparative study, Molecules 24 (2019), https://doi.org/10.3390/molecules24193586.