Acta Poloniae Pharmaceutica – Drug Research, Vol. 68 No. 5 pp. 677–686, 2011 ISSN 0001-6837
Polish Pharmaceutical Society

A STRUCTURE-ACTIVITY RELATIONSHIP STUDY OF THIAZOLE DERIVATIVES WITH H_1-ANTIHISTAMINE ACTIVITY

ELŻBIETA BRZEZIŃSKA, JUSTYNA STOLARSKA and ANNA SOBAŃSKA

Department of Analytical Chemistry, Medical University of Lodz, 1 Muszyńskiego St., 90-151 Łódz, Poland

Abstract: A structure-activity relationship (QSAR) analysis of 19 thiazole derivatives with H_1-antihistamine activity was carried out. The semi-empirical method AM1 was employed to calculate a set of physicochemical parameters for investigated compounds. The principal component analysis (PCA), discriminant function analysis (DFA) and regression analysis (RA) were employed to reduce dimensionality and investigate which subset of variables is effective for classifying the thiazole derivatives according to their degree of anti-H_1 activity. In PCA the studied compounds were separated into two groups: group A with lower degree of H_1-antihistamine activity and group B with higher activity. The DFA showed that the parameters: α (polarizability), AB (distance between aliphatic and aromatic nitrogen atoms), E_b (binding energy), H_h (hydration energy), e_{HOMO} (HOMO energy), and Q_{Ar} are responsible for separation between compounds exhibiting higher and lower H_1-antihistamine activity. The importance of hydrophobic and steric parameters for thiazole derivatives 1-19 with H_1-antihistamine activity was established *via* RA. On the basis of PCA, DFA and RA methods, a prediction rule for classifying new thiazole derivatives with H_1-antihistamine activity was elaborated.

Keywords: thiazole derivatives, QSAR, discriminant function analysis, principal component analysis, regression analysis, drug design

The quantitative relations between physicochemical and structural properties of chemical compounds and their biological response are the subject of quantitative structure-activity relationship (QSAR) studies. Traditional QSAR studies are often restricted to related or congeneric series of compounds. The physicochemical parameters describe electronic, hydrophobic and steric properties of investigated cases. They can be obtained empirically in laboratory assays. Application of molecular descriptors calculated from structures of these compounds by the use of semi-empirical methods can be useful as well. Both groups of parameters usually successfully correlate with a variety of biological data in QSAR. The results of previous study (1) on thiazole derivatives, investigated in the present work, revealed that the calculated descriptors strongly correlated with the same parameters determined by laboratory methods. Considering that they can be determined for the planned structures, they are important for the design of new drugs and in the establishment of QSAR models. The previous study showed that the calculated hydrophobic, electronic and steric parameters effectively describe the variety of biological activity of investigated compounds. The structural variability of the compounds is connected with their pharmacologic variety (1).

The quantitative analysis of the results of other previous studies (2–5) on thiazole and benzothiazole derivatives (some of thiazoles were used in this study – **1**, **2**, **4–6**, **12–15**, **17**, **19**), which are antagonistic towards H_1-histamine receptor revealed the possibility of applying biochromatographic data and regression analysis (RA) to predict the quantitative effect of activity of these compounds. The study reported also some divergence of the results for the group of compounds with anti-H_1 activity and those which exhibited both anti-H_1 and anti-H_3 activity in biological investigations. In the next study, the application of biochromatographic data and discriminant function analysis (DFA) were used for qualitative discrimination and prediction of direction of potential antihistamine drug activity (6). Similar studies related to quantitative analysis have been reported (7) for antihistamine compounds in which data from RP HPLC and the principal component analysis (PCA) were used. A variety of behavior of

* Corresponding author: e-mail: elzbieta.brzezinska@umed.lodz.pl

investigated compounds in chromatographic environment strongly depends on their structure.

The lack of the crystal structure of the histamine H_1 receptor hinders the development of the new types of antagonists. A theoretical three-dimensional model of human H_1 receptor was recently developed on the basis of other human receptors structures (8). The structural requirements for histamine H_1 binding for antihistamines have been reported (8, 9). Application of thermodynamic analysis of ligand-binding may be a novel approach to dissect agonist- and antagonist-specific H_1-receptor conformations (10).

The present work employs physicochemical parameters of selected anti-H_1 compounds as their structural descriptors and statistical methods (PCA, DFA and RA) in systematic structure-activity relationship analysis. The PCA and DFA methods were employed in order to reduce dimensionality and investigate which subset of variables should be more effective for classification of thiazole derivatives according to their degree of anti-H_1 activity. The PCA can be useful as a tool for initial selection of the parameters which significantly are related with higher anti-H_1 activity. The DFA method can be extremely efficient in new anti-H_1 drug design for classification of new planned structures to groups of potentially higher or lower activity group. The RA method will be used for predicting the quantitative effect of H_1-antihistamine activity of different thiazole derivatives.

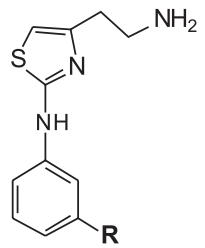
MATERIALS AND METHODS

Compounds

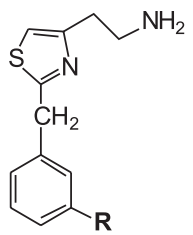
The investigated substances make the set of 19 thiazole-analogs with known biological activity. The statistical analysis applied the values of biological activity (pA_2) of studied compounds determined in the particular biological *in vitro* tests. The synthesis method, analytical data and biological activity of thiazole-analogs **1-19** were described previously (11, 12) (Fig. 1).

Calculation of the theoretical descriptors of molecular properties

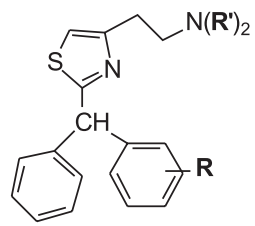
t was suggested that the QSAR analysis, carried out using the molecular descriptors of the compounds in their probably un-ionized or ionized forms (at particular pH), can be most effective (1, 13). Physicochemical parameters of compounds in the group of cases **1-19** were evaluated for un-ionized molecules and, in addition, for their mono-protonated forms as described in the previous study (1). The number of the physicochemical parameters calculated from un-ionized forms of structures was 19. Some of them could not be calculated from ionized forms. The calculation of the percentage of ionization of the investigated compounds in their action place was carried out using an algorithm (14) with pK_a values of the particular compound and pH = 7.4 of biological experiment environment (11, 12). In the case of thiazole derivatives **1-19** with H_1-antihis-



1, 4, 6-7, 11



2, 3, 5, 8-10



12-19

Comp.	R	R'	Comp.	R	R'
1	-CH_3	-	**11**	-Br	-
2	-OCH_3	-	**12**	-4-Br	-H
3	-H	-	**13**	-H	-H
4	-H	-	**14**	-H	-CH_3
5	-F	-	**15**	-4-F	-H
6	-Cl	-	**16**	-4-Cl	-H
7	-F	-	**17**	-3-F	-H
8	-CH_3	-	**18**	-3-Br	-H
9	-Br	-	**19**	-3-Cl	-H
10	-Cl	-			

Figure 1. Structures of the examined compounds **1-19**

Table 1. The molecular descriptors and anti-H_1 activity values, and the group codes for compounds **1–19**.

Comp.	pA_2	E_b [kcal/mol]	A_w [Å^2]	V_w [Å^3]	H_h [kcal/mol]	log P	log D	MR [Å^3]	α [Å^3]	log M_w	ε_{HOMO} [eV]	ε_{LUMO} [eV]	μ [D]	H_f [kcal/mol]	Q_{Ar} [a.u.]	Q [a.u.]	pK_{a1}	pK_{a2}	AB [Å]
1	4.0	–3181.6	262.4	219.8	–9.61	1.30	0.64	70.82	27.11	2.37	–8.714	–0.323	1.24	56.02	–0.141	–0.350	3.79	8.77	4.481
2	4.14	–3396.2	281.5	232.9	–9.30	0.29	0.50	73.32	28.24	2.40	–8.989	–0.226	1.97	11.00	–0.106	–0.350	2.28	8.86	4.330
3	4.16	–3023.9	247.9	208.2	–7.62	0.55	0.54	66.86	25.76	2.34	–9.242	–0.196	1.34	48.62	–0.108	–0.350	2.50	8.91	4.489
4	4.44	–2898.9	240.9	203.0	–10.90	0.83	0.28	65.78	25.28	2.34	–8.761	–0.340	1.10	63.60	–0.141	–0.350	4.35	8.67	4.479
5	4.53	–3035.8	252.3	210.9	–7.34	0.69	0.65	67.07	25.67	2.37	–9.343	–0.309	0.90	3.47	–0.106	–0.303	2.16	8.85	4.490
6	4.61	–2882.6	258.4	217.3	–10.49	1.35	1.22	70.59	27.21	2.40	–8.950	–0.473	0.85	56.81	–0.133	–0.350	3.48	8.76	4.480
7	4.62	–2910.6	246.2	205.6	–10.56	0.97	0.68	66.00	25.19	2.38	–8.964	–0.487	1.11	18.68	–0.134	–0.350	3.48	8.76	4.477
8	4.63	–3306.6	270.5	225.2	–6.47	1.01	1.04	71.90	27.60	2.37	–9.211	–0.181	1.59	41.01	–0.108	–0.350	2.38	8.87	4.318
9	4.65	–2993.8	272.9	230.1	–7.27	1.34	1.36	74.48	28.39	2.47	–9.345	–0.307	0.83	53.39	–0.107	–0.350	2.25	8.85	4.490
10	4.82	–3007.8	265.4	222.6	–7.28	1.06	1.20	71.66	27.69	2.40	–9.323	–0.286	2.26	41.61	–0.105	–0.350	2.16	8.85	4.485
11	5.04	–2868.3	266.3	224.7	–10.48	1.62	1.39	73.41	27.91	2.47	–8.976	–0.487	0.96	68.89	–0.133	–0.350	3.62	8.78	4.479
12	5.87	–4194.8	348.9	303.7	–7.99	3.12	3.12	99.15	38.05	2.57	–9.313	–0.318	2.14	86.05	–0.102	–0.350	1.78	8.76	4.497
13	5.88	–4224.8	324.9	282.2	–8.33	2.33	2.24	91.53	35.42	2.47	–9.232	–0.207	1.20	81.42	–0.102	–0.350	2.05	8.88	4.493
14	5.98	–4768.3	365.1	317.4	–2.23	3.10	2.92	101.60	39.09	2.51	–9.031	–0.185	0.93	88.18	–0.102	–0.268	1.70	8.87	4.505
15	5.99	–4237.1	327.7	284.8	–8.05	2.47	2.38	91.74	35.33	2.49	–9.273	–0.030	2.19	35.98	–0.103	–0.350	1.83	8.77	4.496
16	6.04	–4209.0	341.9	296.5	–8.00	2.84	2.94	96.33	37.35	2.52	–9.285	–0.299	1.99	74.10	–0.102	–0.350	1.78	8.76	4.499
17	6.15	–4236.7	328.5	284.9	–8.06	2.47	2.39	91.74	35.33	2.49	–9.312	–0.294	2.43	36.30	–0.101	–0.350	1.75	8.76	4.675
18	6.16	–4194.7	348.8	303.8	–8.02	3.12	3.11	99.15	38.05	2.57	–9.310	–0.291	2.30	86.21	–0.101	–0.350	1.82	8.77	4.485
19	6.38	–4208.8	341.9	296.6	–8.03	2.84	2.94	96.33	37.35	2.52	–9.302	–0.282	2.18	74.31	–0.101	–0.350	1.75	8.76	4.486

tamine activity, the participation of mono-protonated form (at N atom of the aliphatic amine) predominates (94.89–97.00%). The results of the study (1) on these compounds revealed that the descriptors calculated from both ionized and un-ionized forms of **1-19** can be successfully used in QSAR analysis. Based on them, in the present study the original data set of variables calculated from un-ionized forms **1-19** was used (Table 1). Before applying the stepwise DFA method, each variable was standardized so that they could be compared to each other on the same scale. The dissociation constants (pK_a) and distribution coefficients (log D) values were calculated using a pK ACD/Labs 8.0 and log D ACD/Labs 8.0 software (15).

Other parameters such as: van der Waals surface area (A_w), van der Waals volume (V_w), partition coefficient (log P), molecular refractivity (MR), polarizability (a), molecular weight logarithm (log M_w), distance between aliphatic and aromatic nitrogen atoms (AB), binding energy (E_b), heat of formation (H_f), hydration energy (H_h), HOMO energy (ε_{HOMO}), LUMO energy (ε_{LUMO}), electric charge on N-aliphatic atom (Q), electric charge focused on N-aromatic atom (Q_{Ar}) and dipole moment (μ) were calculated by the HyperChem 7.0 program. All the structures of the studied compounds were geometrically optimized by the use of semi-empirical method AM1 (algorithm Polac-Ribiere, RMS grad = 0.01 kcal/(Å mol) in vacuo). The systematic conformational analysis was not used. The physicochemical parameter values derived from quantum mechanical calculations of the chemical structures are summarized in Table 1.

Values of all descriptors were subjected to mathematical analyses. The principal component analysis, stepwise discriminant function analysis and regression analysis were carried out using the STATISTICA 7.0 software (16). The use of more than one variable in a multivariate regression was justified by an inter-correlation study.

Statistical methods

Principal component analysis (PCA) is a commonly used method for reducing the dimensionality of a data set. The dimensionality of data set is the number of variables that are used to describe each object. It is often found that there are significant correlations between these variables. Under such circumstances, other assays such as: cluster analysis (CA), RA or DFA are often facilitated by using PCA to reduce the number of variables and to eliminate these correlations (17). The principal components (PCs), which explain the variance-covariance structure, are calculated using standard matrix techniques (18) by linear transformation of the original data set of variables into a smaller number of uncorrelated (orthogonal), significant descriptors. The first principal component – PC_1 corresponds to the largest eigenvalue and usually explains the majority of the variation of the data. As the result of the PCA a good classification of the investigated cases can be obtained. The best separation is obtained with several variables out of the original set of data. This suggests that other variables are not so important for classification of these compounds. Considering that the investigated compounds can be separated into groups with higher and lower degree of activity, the PCs determines the direction of the lead compounds modification (19).

Discriminant function analysis (DFA) is a multivariate technique useful to determine which variables discriminate the investigated cases between two or more naturally occurring or *a priori* defined groups. On that basis, the cases can be separated from distinct populations. Next, the classification functions are calculated as models (equations) useful to allocate new cases into previously defined populations (19–21).

The QSAR analysis by regression analysis (RA) of H_1-antihistamine activity and physicochemical properties of **1-19** thiazole derivatives was made. The correlation between biological activity data (11, 12) (pA_2 values in Table 1) and the calculated molecular descriptors of the examined compounds (data in Table 1) were investigated by univariate and multivariate regression analysis method. The general purpose of multiple regressions is to analyze the relationship between several independent variables (physicochemical parameters of examined compounds) and a dependent variable (anti-H_1 activity of compounds). The regression analysis was carried out using the STATISTICA 7.0 program (16). The use of more than one variable in a multivariate equation was justified by inter-correlation study.

RESULTS AND DISCUSSION

Principal component analysis (PCA)

The PCA was carried out with the application of original parameters calculated from un-ionized structures of compounds **1-19** (data in Table 1). After several attempts to obtain the good classification of the compounds, the best separation was achieved with six variables. The first three principal components explained 97.02% of total variance in the data as follows: PC_1 = 68.71%, PC_2 = 18.85%

Table 2. The loading vectors values of PC_1–PC_3 and correlations between PCs and variables for the compounds **1**–**19**.

Variable	Loading vectors			Correlation PC/variable		
	PC_1	PC_2	PC_3	PC_1	PC_2	PC_3
E_b	–0.4533	0.3017	0.0943	–0.9203	0.3208	0.0710
H_h	0.3435	0.3113	–0.8277	0.6974	0.3310	–0.6237
log P	0.4137	–0.4525	0.1540	0.8400	–0.4811	0.1161
α	0.4568	–0.3380	0.0951	0.9275	–0.3594	0.0716
ε_{HOMO}	–0.3405	–0.5572	–0.5116	–0.6913	–0.5925	–0.3855
Q_{Ar}	0.4251	0.4274	0.1075	0.8632	0.4545	0.0810

Table 3. Classification functions for the group A and B.

Variable	A	B
α	–19.772	27,186
AB	–7.173	9,863
E_b	124.519	–171,213
DH_{hyd}	0.865	–28,689
ε_{HOMO}	43.110	–59,276
Q_{arom}	39.767	–54,680
constant	–53.943	–101,817

and PC_3 = 9,46%. The plot of the score vectors for the first two PCs ($PC_1 \times PC_2$) is shown in Figure 2. This projection conserves 87.56% of the total variance of the original data.

The scatterplot shows that the studied compounds are separated into two groups (A and B). Group A contains compounds **1-11** with lower degree of H_1-antihistamine activity ($pA_2 \leq 5.04$), and group B includes compounds **12-19** with higher activity ($pA_2 \geq 5.87$). It is evident in PCA assay that PC_1 alone is responsible for the separation of compounds with higher and lower activity. The less active compounds **1-11** are on the left hand side of the scatterplot and they are connected with negative values of PC_1. The more active compounds **12-19** are connected with positive PC_1 values, on the right hand side of the plot. The loading vectors of the first three PCs were calculated. Table 2 shows the loading vectors values of PC_1–PC_3 and correlations between PCs and variables.

The PC_1 can be expressed by loading vectors *via* the following equation:

$PC_1 = -0.453\ [E_b] + 0.343\ [H_h] + 0.414\ [\log P] + 0.457\ [a] - 0.341\ [e_{HOMO}] + 0.425\ [Q_{Ar}]$

More active thiazole derivatives can be obtained when the PC_1 assumes positive values. So, it is evident that the active compounds can be obtained when they have higher values for the variables: log P and α; low value of H_h; less negative charge on aromatic nitrogen atom – Q_{Ar}; combined with lower values of binding energy and more negative value of HOMO energy. Such characteristics can be useful for the new thiazole anti-H_1 drugs design process.

Discriminant Function Analysis (DFA)

The DFA was performed by means of STATISTICA 7.0 (16). Within the current QSAR study, physicochemical parameters derived from the semi-empirical calculations (Table 1) were used as grouping variables. Before applying the DFA method, each variable was auto scaled so that they could be compared to each other on the same scale. The DFA was performed on 19 cases of thiazole derivatives (**1-19**). The compounds studied were initially divided into two groups of activity. The group codes (A and B) were assigned to them *a priori*. Affiliation of the particular compounds with two groups of lower – A (compounds **1-11**) and higher – B (compounds **12-19**) anti-H_1 activity was established based on pA_2 values (11, 12) and results of PCA described above. In the performed analysis the forward stepwise method was applied. In stepwise DFA a model of discrimination was built step-by-step. Grouping variables (physicochemical parameters) were successively introduced to the analysis. Specifically, at each step the program reviews all variables and evaluates which one will contribute most to the discrimination between groups. That variable is then included in the model, and the program proceeds to the next step. The stepwise procedure is guided by the respective F to enter and F to remove values. The F value for a variable indicates its statistical significance in the discrimination between groups. It is the measure of the extent to which a variable makes a unique contribution to the prediction of group mem-

Table 4. Classification matrix.

		Predicted classification		
		Percentage of correctly classified	A p = 0.57895	B p = 0.42105
Observed classification	A	100.00	11	0
	B	100.00	0	8
	Total	100.00	11	8

bership (22). The introduction of sufficient grouping of variables to the model and obtaining maximum probability of *a priori* classification, discriminant function (root) discriminating activity groups was calculated. In the two-group case, only one discriminant function could be determined. The last phase of the analysis of the compounds was to determine classification functions for both activity groups (see equations below). These functions can be used for classification of cases. Each function allows us to compute classification scores of each case for each group. The case is classified as belonging to the group for which it has the highest classification score.

After six subsequent steps of analysis and introduction of six grouping variables, discriminant function (root) was achieved. Next, classification functions for each level of compounds activity were calculated. Two subsequent classification functions grouping compounds of low and higher H_1-antihistamine activity range are listed in Table 3.

Six significant variables were extracted such as**:** a, AB, E_b, H_h, e_{HOMO}, Q_{Ar}**.** Comparing the results using DFA and PCA methodologies, we can see also that a, E_b, H_h, ε_{HOMO}, Q_{Ar} are key properties for explaining the H_1-antihistamine activity of the thiazole derivatives **1-19** but also the properties log P and AB are important for design of new thiazoles demonstrating antihistamine activity. The discrimination function for groups A and B is given below:

$A = -19.77\ \alpha - 7.17\ AB + 124.52\ E_b + 20.86\ H_h + 43.11\ e_{HOMO} + 39.77\ Q_{Ar} - 53.94$

$B = 27.19\ \alpha + 9.86\ AB - 171.21\ E_b - 28.69\ H_h - 59.28\ e_{HOMO} - 54.68\ Q_{Ar} - 101.82$

The quality of the discriminant models was determined on the basis of *a posteriori* probability (cases in the model). The investigation of classification matrix showed that 100% of cases were correctly classified according to the *a priori* assigned group codes (Table 4).

The separation of the two groups is good. When H_1-antihistamine activity of new thiazole derivatives is investigated, there are the following allocation rules derived from the DFA results: to calculate the values of six variables obtained here *via* the DFA method (a, AB, E_b, H_h, e_{HOMO}, Q_{Ar}) for the new thiazole derivative; to substitute these values in the two discrimination functions obtained here; to check which discrimination function presents the higher value. If the higher value is related to the discrimination function of group A, the thiazole derivative is probably active.

Regression Analysis (RA)

The application of the biochromatographic data in QSAR analysis of some thiazole derivatives with H_1-antihistamine activity was described in the previous papers (2–5). Additionally, the lipophilicity data of solutes were applied as independent variables in the regression analysis. On the basis of described results it was found that log P is a crucial indicator of the H_1-antihistamine effect of thiazole derivatives. An increase in the log P value favors higher biological activity of the tested compounds. Numerous significant multivariate relationships of the antihistamine effect involved log P values (2–5). The present RA started with inter-correlation study of the used independent variables. Only the uncorrelated physicochemical data can be used in the multivariate relationships (see Table 5).

Next, the systematic analysis was performed. As result, over 80 statistically significant relationships were determined. The obtained relationships of H_1-antihistamine effect and molecular descriptors values explained 77–95% of the variance. The best univariate and multivariate relationships are shown in Table 6.

It is evident in QSAR investigation that the best correlations obtained by the RA method for thiazole derivatives **1-19** with H_1-antihistamine activi-

Table 5. Inter-correlation among parameters used in the multiple regression analysis.

	pA_2	E_b	A_w	V_w	ΔH_{hyd}	log P	log D	MR	α	log M_w	ε_{HOMO}	ε_{LUMO}	$\varepsilon_{HOMO}-\varepsilon_{LUMO}$	μ	ΔH_f	Q_{arom}	Q_{alif}	pK_{a2}	AB
pA_2	1.00	–0.88	0.91	0.93	0.36	0.94	0.95	0.93	0.93	0.88	–0.52	0.27	–0.51	0.53	0.56	0.60	0.12	–0.22	0.48
E_b	–0.88	1.00	–0.97	–0.97	–0.56	–0.88	–0.89	–0.96	–0.96	–0.78	0.40	–0.53	0.52	–0.52	–0.51	–0.67	–0.30	0.01	–0.32
A_w	0.91	–0.97	1.00	1.00	0.52	0.93	0.96	1.00	1.00	0.90	–0.45	0.41	–0.51	0.54	0.59	0.68	0.24	–0.06	0.31
V_w	0.93	–0.97	1.00	1.00	0.51	0.95	0.96	1.00	1.00	0.90	–0.45	0.40	–0.51	0.53	0.61	0.67	0.23	–0.08	0.34
ΔH_{hyd}	0.36	–0.56	0.52	0.51	1.00	0.38	0.44	0.48	0.48	0.31	–0.46	0.59	–0.59	0.08	0.22	0.68	0.72	0.58	0.04
log P	0.94	–0.88	0.93	0.95	0.38	1.00	0.98	0.96	0.95	0.92	–0.37	0.20	–0.36	0.43	0.72	0.47	0.16	–0.27	0.46
log D	0.95	–0.89	0.96	0.96	0.44	0.98	1.00	0.97	0.97	0.95	–0.52	0.25	–0.50	0.51	0.66	0.62	0.15	–0.15	0.39
MR	0.93	–0.96	1.00	1.00	0.48	0.96	0.97	1.00	1.00	0.91	–0.45	0.37	–0.50	0.53	0.64	0.65	0.20	–0.11	0.35
α	0.93	–0.96	1.00	1.00	0.48	0.95	0.97	1.00	1.00	0.90	–0.45	0.38	–0.50	0.54	0.64	0.65	0.20	–0.11	0.35
log M_w	0.88	–0.78	0.90	0.90	0.31	0.92	0.95	0.91	0.90	1.00	–0.50	0.13	–0.43	0.48	0.62	0.56	0.07	–0.20	0.37
ε_{HOMO}	–0.52	0.40	–0.45	–0.45	–0.46	–0.37	–0.52	–0.45	–0.45	–0.50	1.00	–0.39	0.92	–0.49	–0.02	–0.87	0.00	–0.36	–0.23
ε_{LUMO}	0.27	–0.53	0.41	0.40	0.59	0.20	0.25	0.37	0.38	0.13	–0.39	1.00	–0.72	0.38	–0.01	0.65	0.18	0.44	–0.14
ε_{HOMO}-ε_{LUMO}	–0.51	0.52	–0.51	–0.51	–0.59	–0.36	–0.50	–0.50	–0.50	–0.43	0.92	–0.72	1.00	–0.53	–0.01	–0.93	–0.07	–0.45	–0.12
μ	0.53	–0.52	0.54	0.53	0.08	0.43	0.51	0.53	0.54	0.48	–0.49	0.38	–0.53	1.00	0.05	0.57	–0.36	–0.18	0.15
ΔH_f	0.56	–0.51	0.59	0.61	0.22	0.72	0.66	0.64	0.64	0.62	–0.02	–0.01	–0.01	0.05	1.00	0.11	0.04	–0.20	0.21
Q_{arom}	0.60	–0.67	0.68	0.67	0.68	0.47	0.62	0.65	0.65	0.56	–0.87	0.65	–0.93	0.57	0.11	1.00	0.21	0.48	0.10
Q_{alif}	0.12	–0.30	0.24	0.23	0.72	0.16	0.15	0.20	0.20	0.07	0.00	0.18	–0.07	–0.36	0.04	0.21	1.00	0.33	0.09
pK_{a2}	–0.22	0.01	–0.06	–0.08	0.58	–0.27	–0.15	–0.11	–0.11	–0.20	–0.36	0.44	–0.45	–0.18	–0.20	0.48	0.33	1.00	–0.32
AB	0.48	–0.32	0.31	0.34	0.04	0.46	0.39	0.35	0.35	0.37	–0.23	–0.14	–0.12	0.15	0.21	0.10	0.09	–0.32	1.00

Table 6. The relationships between $pA_2(H_1)$ values and molecular descriptors calculated for the compounds **1**–**19**.

Eq. no.	pA_2 =	R[a]	R^2	F[b]	S[c]	n[d]
1.	a – b E_b	0.88	0.77	58.178	0.40272	19
2.	a + b A_w	0.91	0.83	85.884	0.34426	19
3.	a + b logP	0.94	0.88	121.51	0.29670	19
4.	a + b logD	0.95	0.91	171.32	0.25445	19
5.	a + b A_w + c AB	0.94	0.88	57.376	0.30537	19
6.	a + b A_w – c Q_{Ar}	0.93	0.87	55.187	0.31062	19
7.	a + b V_w – c Q_{Ar}	0.94	0.89	66.618	0.28584	19
8.	a + b V_w + c AB	0.94	0.89	63.015	0.29300	19
9.	a + b logP – c e_{HOMO}	0.96	0.91	84.324	0.25697	19
10.	a + b logP – c e_{HOMO} – e_{LUMO}	0.95	0.91	81.33	0.26124	19
11.	**a + b logP – c Q_{Ar}**	**0.96**	**0.92**	**97.759**	**0.24009**	**19**
12.	a + b logD – c Q_{Ar}	0.96	0.92	97.613	0.24026	19
13.	a + b logD – c pK_a	0.96	0.92	86.464	0.25404	19
14.	a + b logD + c AB	0.96	0.92	94.660	0.24369	19
15.	a + b A_w – c H_h – d e_{HOMO} – e_{LUMO}	0.94	0.88	37.897	0.30781	19
16.	a + b logP – c H_h – d e_{HOMO} – e_{LUMO}	0.96	0.92	58.618	0.25276	19
17.	a + b logP + c H_h – d Q_{Ar}	0.96	0.92	61.106	0.24796	19
18.	a + b logP – c e_{HOMO} + d AB	0.96	0.92	53.863	0.26277	19
19.	a + b logP + c e_{LUMO} – d Q_{Ar}	0.97	0.94	73.088	0.22814	19
20.	**a + b logP + c m – d Q_{Ar}**	**0.97**	**0.95**	**86.800**	**0.21041**	**19**
21.	a + b logD – c Q_{Ar} + d AB	0.97	0.94	76.578	0.22321	19
22.	a + b MR – c Q_{Ar} + d AB	0.97	0.94	70.529	0.23197	19
23.	a + b a – c Q_{Ar} + d AB	0.97	0.94	72.488	0.22902	19

[a]Correlation coefficient. [b]Value of the F-test of significance. [c]Standard error of estimate. [d]Number of compounds used to derive the regression equation. Significance level for all equations $p < 0.00000$.

ty underline important role of hydrophobic and steric parameters for this kind of activity. The same parameters (A_W; V_w; log P; log D; MR; a) with strong factor loadings built the most significant factor, which was obtained in factor analysis of the investigated compounds **1-19** (1). It is also evident that some of the electronic parameters are very important. These parameters were linked to higher and lower compounds activity *via* DFA (a, AB, H_h, e_{HOMO}, Q_{Ar}) and PCA (a, log P, H_h, ε_{HOMO}, Q_{Ar}) methods.

All calculated significant relationships can be applied to predict the pharmacological activity of new drug candidates. The best of these relationships can be expressed by the equation ***20*** (Table 6), which explains 95% of the total variance:

$pA_2 = 5.16(\pm 0.05) + 0.71(\pm 0.06)$ log P $+ 0.133(\pm 0.06)$ m $- 0.177(\pm 0.05)\ Q_{Ar}$

It was concluded that the high lipophilicity and dipole moment combined with less negative charge on aromatic nitrogen atom are the properties of active H_1-antihistamine thiazole derivatives.

It is clearly seen that equation ***20*** may have predictive value for the design of new thiazole derivatives as the H_1-antihistamine drugs (Table 7 and Fig. 3). The correlation of calculated $pA_2(H_1)$ values of the tested compounds predicted by the use of the above mentioned equation *versus* their $pA_2(H_1)$ values obtained from biological tests was significant ($R^2 = 0.95$).

However, the range of pA_2 data of the examined compounds obtained from biological tests clustered around two sets (compounds **1-11** have pA_2 values between 4.00 and 5.04; compounds **12-19** have pA_2 values between 5.87 and 6.38). For the two-point data distribution the possibility of coincidence in the model presented in the figure cannot be eliminated.

Table 7. The obtained and predicted $pA_2(H_1)$ values of the examined compounds **1–19**.

Compound No.	pA_2 obtained	pA_2 predicted
1	4.00	4.04
2	4.14	4.21
3	4.16	4.26
4	4.44	4.51
5	4.53	4.26
6	4.61	4.82
7	4.62	4.60
8	4.63	4.66
9	4.65	4.72
10	4.82	4.84
11	5.04	5.04
12	5.87	6.30
13	5.88	5.52
14	5.98	6.02
15	5.99	5.84
16	6.04	6.07
17	6.15	5.89
18	6.16	6.34
19	6.38	6.12

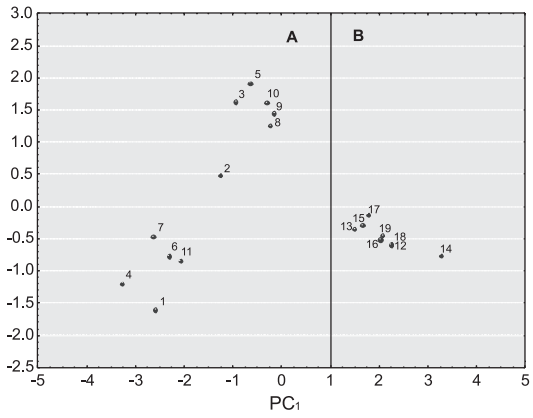


Figure 2. PCA score (PC_1 and PC_2) for the thiazole derivatives **1–19** with H_1-antihistamine activity. Group A represents compounds with lower activity and group B compounds with higher activity

CONCLUSION

The dimensionality of physicochemical parameters was reduced by the PCA and DFA methods, and the subset of variables more effective for classification the thiazole derivatives according to their degree of anti-H_1 activity were determined. The PCA method can be useful as an efficient tool for initial selection of the parameters which significantly enhance anti-H_1 activity. The analysis determined the direction of the lead compounds modification. The results of DFA method showed that α, AB, E_b, H_h, e_{HOMO} and Q_{Ar} parameters are key properties for explaining the H_1-antihistamine activity of thiazole derivatives **1-19** but log P is also important for design of new thiazoles exhibiting antihistamine activity. The determined discrimination function for groups A and B can be an efficient tool in further investigations. Good univariate and multivariate relationships obtained by the use of RA method can be used for predicting the quantitative effect of H_1-antihistamine activity of different thiazole derivatives. These relationships involved the parameters determined *via* DFA and PCA methods.

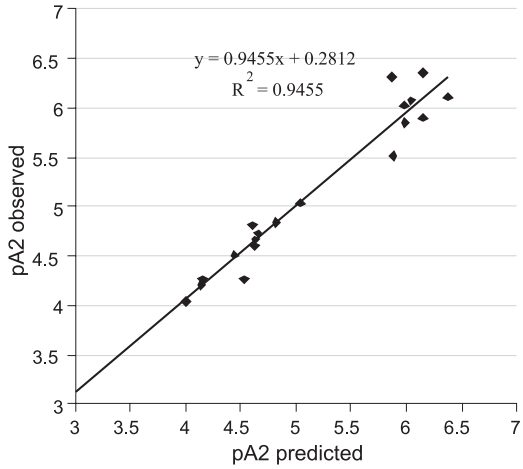


Figure 3. Correlation of calculated $pA_2(H_1)$ values of the tested compounds predicted by equation ***20*** (Table 6) *versus* their $pA_2(H_1)$ values obtained from the biological tests

Acknowledgment

This study was supported by the Medical University of Łódź, Poland, Research Program No. 503-3016-3.

REFERENCES

1. Brzezińska E., Stolarska J.: Acta Pol. Pharm. Drug Res. 66, 25 (2009).
2. Brzezińska E., Kośka G., Walczyński K.: J. Chromatogr. A 1007, 145 (2003).

3. Brzezińska E., Kośka G., Klimczak A.: J. Chromatogr. A 1007, 157 (2003).
4. Brzezińska E., Kośka G.: J. Planar Chromatogr. 16, 451 (2003).
5. Brzezińska E., Kośka G.: J. Planar Chromatogr. 17, 40 (2004).
6. Brzezińska E., Kośka G.: Biomed. Chromatogr. 20, 1004 (2006).
7. Gami-Yilinkou R., Nasal A., Kaliszan R.: J. Chromatogr. 633, 57 (1993).
8. Wang X., Yang Q., Li M., Yin D., You Q.: J. Mol. Model. 16, 1529 (2010)
9. Aslanian R., Piwinski J.J., Zhu X., Priestley T., Sorota S., Du X-Y., Hang X-S. et al.: Bioorg. Med. Chem. Lett. 19, 5043 (2009)
10. Wittmann H.-J., Seifert R., Strasser A.: Mol. Pharmacol. 76, 25 (2009)
11. Walczyński K., Timmerman H., Zuiderveld O.P., Zhang M.Q., Glinka R.: Farmaco 54, 533 (1999) .
12. Walczyński K., Guryn R., Zuiderveld O.P., Zhang M.Q., Timmerman H.: Farmaco 55, 569 (2000).
13. Scherrer R.A., Leo A.J.: Mol. Inf. 29, 687 (2010)
14. Hulanicki A.: Reactions of acids and bases in analytical chemistry (Polish), 3rd edn., PWN, Warszawa 1992.
15. ACD/Labs™ Log D Suite 8.0, pKa dB 7.0, Advanced Chemistry Development Inc.
16. STATISTICA 7.1 – version for Windows, StatSoft, Inc. 2005
17. Leach A.R.: Molecular modeling. Principles and applications. Addison Wesley Longman Ltd., Edinburgh 1998.
18. Chatfield C., Collins A.J.: Introduction to multivariate analysis, Chapman and Hall, London 1980.
19. Alves C.N., Pinheiro J.C., Camargo A.J., Ferreira M.M.C., da Silva A.B.F.: J. Mol. Struct. (THEOCHEM) 530, 39 (2000).
20. Johnson R.A., Wichern D.W.: Applied Multivariate Statistical Analysis, Prentice-Hall, Englewood Cliffs, NJ 1992.
21. Marida K.V., Kent J.T., Bibbly J.M.: Multivariate Analysis. Academic Press, New York 1979.
22. Casabán-Ros E., Antün-Fos G.M., Gálvez J., Duart M.J., Garcia-Doménech R.: QSAR 18, 35 (1999).

Received: 3, 01. 2011