QSARs on Bactericidal Activity of 3-carboxy-4-quinolones

LASZLO TARKO[1]*, LUCIA PINTILIE[2], CĂTĂLINA NEGUÞ[2], CORNELIU ONISCU[3] , MIRON TEODOR CĂPROIU[1]

[1] Romanian Academy, Center of Organic Chemistry "C.D.Nenitzescu", 202B Spl. Independentei, 060023, Bucharest, Romania

[2] National Institute for Chemical-Pharmaceutical Research and Development, 112 Vitan Av., 074373, Bucharest, Romania

[3] Technical University "Gh. Asachi", 71 A D. Mangeron Av.,700050, Iasi, Romania

This paper presents results of three QSAR (Quantitative Structure Activity Relationship) studies realized with the PRECLAV computer program. The database we used contains initially 100 derivatives of 3-carboxy-4-quinolone. The dependent property is bactericidal activity against Staphylococcus aureus, Escherichia coli and Pseudomonas aeruginosa. A specific criterion identifies the outlier molecules in the calibration set. Two molecules are identified as "possible outliers for lead hopping". After the elimination of outliers, we obtained: N = 77 / 86 / 84, s = 0.2904 / 0.3583 / 0.2993, r^2 = 0.8850 / 0.7943 / 0.8645, F = 91.1 / 37.6 / 82.9 and r^2_{CV} = 0.8415 / 0.7337 / 0.8415. The bactericidal activity against the three studied bacteria was favored by the presence of saturated C substituted (hetero)cycles, by the presence of certain groups (-F, unconjugated -NH/-NH_2) and by a non-balanced molecular shape. The bactericidal activity was disfavored by the presence of certain chemical groups (-NO_2, -C_6H_4, -CO-) and of the triazole cycle. The lipophilic/hydrophilic feature of quinolones has little impact upon bactericidal activity.

Keywords: QSAR, bactericidal activity, quinolones, PRECLAV

The first bactericidal substances, either produced by microorganisms or synthesized in the lab, were discovered soon after Ehrlich (1913) introduced to chemotherapy the concept of selective toxicity. The first chemotherapeutic drugs were the sulfamides, soon followed by a variety of biosynthetic antibiotics such as β-lactams, macrolides, aminoglycosides and tetracyclines, isolated from various cultures of fungi, bacteria, and actinomicetes. The biosynthetic and semi-synthetic antibiotics have dominated the drug industry, while chemosynthetic antibiotics have played, with some exceptions, only a minor role.

J.R. Price and his collaborators were the first to test a quinolone [1]. They found lacking any bactericidal activity. G.Y. Lesher [53] and his colleagues discovered nalidixic acid, the first therapeutic quinolone. Nalidixic acid has a mild activity against some gram-negative microorganisms. Later on, several other similar quinolones were synthesized, some having a narrow anti-bactericidal spectrum (especially against enterobacteria), some exhibiting rapid elimination, and some with low tissue absorption. Those features allow their use as urinary antiseptics only.

After 1980, second-generation quinolones were produced having stronger bactericidal properties and a wider activity spectrum, which allows their use against localized infections. Koga *et al.*[2] introduced Norfloxacin, the first quinolone that includes fluorine in position 6 and piperazin-1-yl in position 7. Norfloxacin is 500 times more active than the previously synthesized compounds. Moreover, it is effective against both gram-positive and gram-negative microorganisms, including *P. aeruginosa*, which is particularly difficult to control. Today number of synthesized therapeutic quinolones is large [1-25].

Norfloxacin and other similar quinolones were designed by traditional QSAR analysis [2, 49-52] using regression analysis. The purpose of the QSAR studies presented here was to identify the molecular features with the highest impact (favorable or unfavorable) upon the bactericidal activity of 3-carboxy-4-quinolones.

Methods and formulae

In the present QSAR studies the dependent property is the bactericidal activity A defined as

$$A = \log (200 / MIC)$$

The starting point for the computation is the database (100 derivatives of the 3-carboxy-4-quinolone) shown in fig. 2, table 1 and table 2. The values of MIC (minimum inhibitory concentration, μg/mL), for *S. aureus, E. coli* and *P. aeruginosa,* are quoted from the literature (table 2, last column). For MIC evaluation quoted authors use standard twofold serial dilution method, using agar media [54]. They use standard collection bacteria strains (89 cases) or non-standard bacteria strains (11 cases, table 2, last column and footnotes).

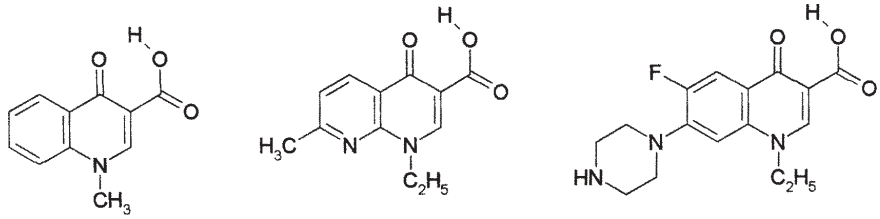

Fig. 1. Some reference quinolones

* *email: ltarko@cco.ro*

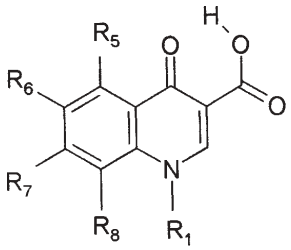

Fig. 2. Structure of analyzed quinolones

In table 1 and table 2, the molecules are ordered according to the ascending sum value of their activities $A_{Staphilococcus} + A_{Escherichia} + A_{Pseudomonas}$.

The molecules (not zwitterions) have been virtually constructed using the molecular mechanics program, PCMODEL [26]. The geometry of the minimum energy conformer was obtained by the MMX force field. Then the quantum mechanics program MOPAC [27], using the keyword string "am1 pulay gnorm=0.01 shift=50 geo-ok camp-king bonds vectors", optimized the geometry more rigorously.

The statistical computations were conducted using an improved version of the PRECLAV program [28-32, 55, 56]. The output files created by MOPAC for each analyzed molecule are input files for PRECLAV and they contain the values of some descriptors. Using the data from the files generated by MOPAC, PRECLAV has computed some descriptors and has performed the statistical analysis.

We have used PRECLAV descriptors [30, 31, 32] ("whole molecule" descriptors and some parabolic functions of these descriptors, "virtual fragmentation" descriptors, "grid" descriptors) and DRAGON descriptors [33].

The "significant" descriptors are those which are sufficiently correlated with the dependent property ($r^2 > 4/N$, N is number of molecules in calibration set). The quality criteria for the descriptors, the method for grouping the significant descriptors in sets, the quality criteria for the calculated QSARs and the criteria for ending the computations have been presented in previous works [28, 30-32, 55, 56].

Table 1

CHEMICAL GROUPS IN ANALYZED QUINOLONES

Index	R_1	R_5	R_6	R_7	R_8
1	2,4 –di-CH_3- C_6H_3	H	F	4-CH_3-piperazin-1-yl	H
2	cyclopropyl	NH_2	NH_2	4-CH_3-piperazin-1-yl	H
3	4-NO_2-C_6H_4	H	F	4-CH_3-piperazin-1-yl	H
4	C_2H_5	H	F	1,2,4-triazol-4-yl	H
5	3-F-C_6H_4	H	F	4-CH_3-piperazin-1-yl	H
6	C_2H_5	H	F	pyrazol-1-yl	H
7	C_2H_5	H	F	3-(hydroxymethyl)-piperazin-1-yl	H
8	C_2H_5	H	F	4-(4-nitrobenzyl)-piperazin-1-yl	H
9	4-Br-C_6H_4	H	F	4-CH_3-piperazin-1-yl	H
10	*trans*-2-CH_3-cyclobutyl	H	F	piperazin-1-yl	H
11	C_2H_5	H	NO_2	piperazin-1-yl	H
12	C_2H_5	H	F	CH_3	H
13	4-NH_2-C_6H_4	H	F	4-CH_3-piperazin-1-yl	H
14	C_2H_5	H	F	$S(CH_2)_2NH_2$	H
15	C_2H_5	H	F	piperidin-1-yl	H
16	cyclopropyl	H	F	piperazin-1-yl	H
17	cyclopropyl	C_2H_5	F	piperazin-1-yl	H
18	C_2H_5	H	CN	piperazin-1-yl	H
19	4-CH_3-C_6H_4	H	F	4-CH_3-piperazin-1-yl	H
20	cyclopropyl	SH	F	4-CH_3-piperazin-1-yl	F
21	C_2H_5	H	F	pyrrol-1-yl	H
22	C_2H_5	H	F	$N(CH_3)_2$	H
23	C_2H_5	H	F	pyrrolid-1-yl	H
24	cyclopropyl	H	NH_2	4-CH_3-piperazin-1-yl	F
25	4-F-C_6H_4	H	NH_2	4-CH_3-piperazin-1-yl	H
26	2-F-C_6H_4	H	F	4-CH_3-piperazin-1-yl	H
27	CH_3	H	F	piperazin-1-yl	H
28	3,4-OCH_2O-C_6H_3	H	F	4-CH_3-piperazin-1-yl	H
29	cyclopropyl	H	NH_2	4-CH_3-piperazin-1-yl	H
30	C_2H_5	H	F	morpholin-1-yl	H
31	4-F-C_6H_4	H	F	piperidin-1-yl	H
32	CH_2OH	H	F	piperazin-1-yl	H
33	C_2H_5	H	F	4-(4-NH_2-benzyl)-piperazin-1-yl	H

continuare

34	1-CH_3-cyclobutyl	H	F	piperazin-1-yl	H
35	allyl	H	F	piperazin-1-yl	H
36	C_2H_5	H	F	pyrrolidin-1-yl	H
37	C_2H_5	H	F	4-OH-piperidin-1-yl	H
38	2,2-di-CH_3-butyl	H	F	piperazin-1-yl	H
39	cyclopropyl	H	F	3-ethyl-amino-methyl-azetidin-1-yl	H
40	isopropyl	H	F	piperazin-1-yl	H
41	C_2H_5	H	F	3-fluoromethyl-piperazin-1-yl	
42	cyclopropyl	H	F	3-dimethylamino-3-methyl-azetidin-1-yl	H
43	4-F-C_6H_4	H	F	homopiperazin-1-yl	H
44	4-F-C_6H_4	H	F	4-OH-piperidin-1-yl	H
45	*tert*-butyl	H	NH_2	4-CH_3-piperazin-1-yl	H
46	1,3-di-F-2,2-di-CH_3-propyl	H	F	piperazin-1-yl	H
47	C_2H_5	H	F	thiazolidin-1-yl	H
48	vinyl	H	F	piperazin-1-yl	H
49	phenyl	H	F	piperazin-1-yl	H
50	4-F-C_6H_4	H	F	pyrrolidin-1-yl	H
51	4-F-C_6H_4	H	F	morpholin-1-yl	H
52	C_2H_5	H	F	thiomorpholin-1-yl	H
53	cyclopropyl	H	NH_2	thiomorpholin-1-yl	F
54	cyclopropyl	H	F	*trans*-2-methyl-3-ethylaminomethyl-azetidin-1-yl	H
55	cyclobutyl	H	F	piperazin-1-yl	H
56	cyclopropyl	H	NH_2	3-aminomethyl-morpholin-1-yl	H
57	4-F-C_6H_4	H	F	3-OH-pyrrolidin-1-yl	H
58	4-F-C_6H_4	H	F	4-NH_2-piperazin-1-yl	H
59	cyclopropyl	H	F	*cis*-2-ethyl-3-amino-azetidin-1-yl	H
60	2-allyl	H	F	piperazin-1-yl	H
61	2,4-diF-C_6H_3	H	F	4-CH_3-piperazin-1-yl	H
62	cyclopropyl	H	F	3-methylamino-3-methyl-azetidin-1-yl	H
63	cyclopropyl	H	F	*cis*-2-methyl-3-amino-3-methyl-azetidin-1-yl	H
64	4-F-C_6H_4	H	F	3-amino-3-methyl-azetidin-1-yl	H
65	cyclopropyl	H	F	3-amino-pyrrolidin-1-yl	NH_2
66	cyclopropyl	F	F	4-CH_3-piperazin-1-yl	F
67	cyclopropyl	$NH(CH_2)_2OH$	F	4-CH_3-piperazin-1-yl	F
68	cyclopropyl	H	F	3-fluoromethyl-piperazin-1-yl	H
69	cyclopropyl	H	F	3-dimethylamino-azetidin-1-yl	H
70	cyclopropyl	H	F	1-amino-cyclopent-3-en-3-yl	H
71	cyclopropyl	OH	F	4-CH_3-piperazin-1-yl	F
72	*tert*-butyl	H	F	piperazin-1-yl	H
73	cyclopropyl	H	F	homopiperazin-1-yl	H
74	4-F-C_6H_4	H	F	piperazin-1-yl	H
75	cyclopropyl	H	F	3-methylamino-azetidin-1-yl	H
76	cyclopropyl	H	F	*cis*-2-methyl-3-amino-azetidin-1-yl	H
77	cyclopropyl	H	F	3-amino-pyrrolidin-1-yl	H
78	cyclopropyl	H	F	3-aminomethyl-3-methyl-pyrrolidin-1-yl	H
79	cyclopropyl	H	F	3-methylamino-pyrrolidin-1-yl	H
80	4-OH-C_6H_4	H	F	4-CH_3-piperazin-1-yl	H
81	cyclopropyl	H	F	3-amino-azetidin-1-yl	H
82	cyclopropyl	H	F	*trans*-2-methyl-3-amino-azetidin-1-yl	H
83	4-F-C_6H_4	H	F	3-methyl-3-amino-pyrrolidin-1-yl	H
84	cyclopropyl	H	F	6,8-diazabicyclo[3,2,2]non-6-yl	H
85	fluoro-*tert*-butyl	H	F	piperazin-1-yl	H

continuare

86	cyclopropyl	H	F	piperazin-1-yl	H
87	cyclopropyl	NH_2	F	piperazin-1-yl	Cl
88	cyclopropyl	NH_2	F	4-CH_3-piperazin-1-yl	F
89	cyclopropyl	Cl	F	piperazin-1-yl	F
90	cyclopropyl	H	F	3-methyl-piperazin-1-yl	H
91	cyclopropyl	NH_2	F	piperazin-1-yl	H
92	cyclopropyl	H	F	3-aminomethyl-pyrrolidin-1-yl	H
93	cyclopropyl	H	NH_2	piperazin-1-yl	H
94	cyclopropyl	H	F	3-amino-pyrrolidin-1-yl	OMe
95	cyclopropyl	CH_3	F	piperazin-1-yl	F
96	cyclopropyl	CH_3	F	piperazin-1-yl	Cl
97	cyclopropyl	NH_2	F	piperazin-1-yl	F
98	cyclopropyl	CH_3	F	piperazin-1-yl	H
99	cyclopropyl	H	F	3-amino-pyrrolidin-1-yl	F
100	cyclopropyl	H	F	3-amino-pyrrolidin-1-yl	Cl

Table 2

BACTERICIDAL ACTIVITY OF ANALYZED QUINOLONES

Index in Table 1	Bactericidal activity against			Reference for MIC value
	S. aureus	*E. coli*	*P. aeruginosa*	
1	0.301	0.301	0.301	[3]
2	0.194	1.097	0.194	[6]
3	0.495	1.398	0.194	[4]
4	0.903	0.903	0.903	[12]
5	1.204	1.509	0.602	[3]
6	1.505	1.204	0.903	[12]
7	1.398	2.000	0.495	[7]
8	2.108	1.505	0.301	[2]
9	1.810	1.509	0.602	[3]
10	1.699	2.000	0.796	[15]
11	0.903	2.409	1.204	[2]
12	1.505	2.710	0.602	[2]
13	2.000	2.301	0.796	[4]
14	0.903	2.398	1.810	[1]
15	2.409	2.108	0.602	[2]
16	2.903	0.495	2.000	[6]
17	1.502	2.699	1.204	[19*]
18	1.204	2.710	1.505	[2]
19	2.108	2.108	1.204	[3]
20	1.805	2.710	1.204	[13**]
21	2.710	1.805	1.204	[12]
22	2.409	2.710	0.602	[2]
23	2.699	2.097	1.204	[1]
24	1.699	2.301	2.000	[6]
25	1.398	3.222	1.398	[6]
26	2.108	2.409	1.509	[3]
27	1.505	2.710	2.108	[2]
28	2.426	2.409	1.509	[3]
29	2.903	2.000	1.699	[6]
30	2.409	3.000	1.204	[2]
31	3.000	2.108	1.509	[3]
32	2.108	2.710	1.805	[2]
33	2.710	2.710	1.204	[2]

continuare

34	2.602	2.903	1.398	[15]
35	1.805	3.000	2.108	[2]
36	3.000	2.710	1.204	[2]
37	2.710	2.710	1.505	[2]
38	2.903	2.602	1.699	[15]
39	2.602	2.903	1.699	[9]
40	2.301	2.602	2.301	[15]
41	3.222	3.222	0.796	[7]
42	2.903	2.903	1.699	[9]
43	2.403	3.000	2.108	[3]
44	3.301	2.710	1.509	[3]
45	2.000	3.523	2.000	[6]
46	2.301	3.187	2.301	[14]
47	3.000	3.000	1.810	[1]
48	1.805	3.301	2.710	[2]
49	3.000	2.409	2.409	[3]
50	3.301	2.409	2.108	[3]
51	3.301	2.710	1.810	[3]
52	4.523	2.097	1.204	[1]
53	3.222	2.903	1.699	[6]
54	2.903	3.222	1.699	[9]
55	2.602	3.187	2.301	[15]
56	3.000	3.000	2.108	[5]
57	3.301	3.000	1.810	[3]
58	2.710	3.000	2.409	[1]
59	2.903	3.222	2.000	[10]
60	2.602	3.187	2.602	[15]
61	3.301	3.000	2.108	[3]
62	2.903	3.523	2.000	[9]
63	2.903	3.523	2.000	[9]
64	3.222	3.222	2.000	[5]
65	3.523	2.577	2.577	[24]
66	2.710	3.602	2.409	[13**]
67	2.710	3.602	2.409	[13**]
68	3.222	3.523	2.000	[7]
69	3.222	3.523	2.000	[9]
70	3.301	3.301	2.398	[17]
71	3.000	3.602	2.409	[13**]
72	2.903	3.523	2.602	[14]
73	2.903	3.824	2.301	[7]
74	3.000	3.602	2.710	[3]
75	2.903	3.824	2.602	[9]
76	3.222	3.523	2.602	[10]
77	3.155	3.347	3.000	[17]
78	4.187	3.000	2.398	[16]
79	3.903	3.301	2.398	[16]
80	3.602	3.301	2.710	[3]
81	2.903	3.824	2.903	[9]
82	3.222	3.824	2.903	[10]
83	3.602	4.000	2.409	[18]
84	3.903	3.523	2.699	[8]
85	3.187	4.097	2.903	[14]
86	3.187	3.824	3.187	[1]
87	3.903	3.903	2.398	[20*]

continuare

88	3.301	3.903	3.000	[5*; 13**]
89	3.301	4.204	2.710	[13**]
90	3.222	4.125	2.903	[7]
91	3.903	4.187	2.398	[20*]
92	4.523	3.602	2.398	[16]
93	3.222	3.824	3.523	[6]
94	3.824	3.456	3.456	[23; 25]
95	3.903	3.903	3.000	[19]
96	3.903	3.903	3.000	[19]
97	3.602	4.187	3.301	[5*; 13**; 20*]
98	3.903	4.187	3.000	[19*]
99	3.824	3.699	3.699	[23; 24; 25]
100	3.699	3.824	3.824	[11; 24]

* against non-standard strains UC 76 *S. aureus*; Vogel *E. coli*; UI-18 *P. aeruginosa*
** against non-standard strains 209P JC-1 *S. aureus*; NIIHJ JC-2 *E. coli*; 12 *P. aeruginosa*

The descriptors included in QUSAR (equation useful in prediction) are "predictors".

The computed QSAR are multilinear. Here "A" denotes biological activities.

$$A = c_0 + \Sigma c_k \cdot p_k \quad (1)$$

Ordinary Least Square Method computes weighting factors c_k of predictors p_k. The PRECLAV program does not compute errors related to regression coefficients.

The outlier molecules are molecules for which the QSAR resulted from computations offers only a poor estimation of the bactericidal activity, although for the rest of the molecules in the calibration set the estimates have been good. The presence of outlier molecules lowers the predictive quality of the whole calibration set and often determines the inclusion into the final equation of a different set of predictors. To identify these molecules, PRECLAV uses an improved version of specific criterion, called COIN (*Combined Outlier INdex)*.

$$COIN = f \cdot \Delta_{value} \cdot \Delta_{rank} / \Sigma | A_{obs}| \quad (2)$$

where:

$\Delta_{value} = |A_{obs} - A_{calc}|_{value}$

$\Delta_{rank} = |A_{obs} - A_{calc}|_{rank}$

A_{obs} is the observed value of bactericidal activity

A_{calc} is the computed value of bactericidal activity

Difference Δ_{value} compares the calculated and the observed value of the dependent property. Difference Δ_{rank} compares the rank of the molecule in the set ordered by the calculated or observed values. For factor *f* the program uses the value *f* = 12. This parameterization was obtained empirically after the analysis of a large number of structurally diverse databases.

Some molecules in the calibration set may give high values of Δ_{value}, but low values of Δ_{rank} or vice versa. Here only the molecules with COIN > 1 were considered as outliers and eliminated from further computations.

The biochemical active (A_{obs} value is large enough, here A_{obs} > 3.5) high outliers (COIN > 4) may be good starting points for lead hopping [34, 35] because these molecules are both active and different from the other calibration set molecules.

After removing the outliers, the program repeats the statistical analysis. By computing a new QSAR, the program identifies other outliers. The repetitive step-by-step process of outlier identification / elimination is interrupted when the number of discovered outliers becomes null. During this repetitive process, the value of COIN gradually decreases. This does not happen to the values of other classical mathematical functions used to identify outliers.

After the outliers have been eliminated, the program calculates the relative utility U of various predictors in the final QSAR, using equation (3).

$$U = (R^2 - r^2) / (1 - r^2) \quad (3)$$

where:

R^2 is the square of the Pearson correlation between the observed and calculated values of activity (values calculated using an equation with *k* predictors);

r^2 is the square of the Pearson correlation between the observed and calculated values of activity (values calculated using an equation with *k-1* predictors, that is the equation that does not contain the analyzed predictor).

After computing the value of U for all predictors, these values are normalized according to the highest U (the highest value becomes 1000). The predictors with a high value for U (U>500) may be considered very useful in estimating the activity because they correlate very well with activity and do not correlate with other predictors. Each "useful" predictor offers plenty of information about why activity varies from molecule to molecule. Moreover, each "useful" predictor offers a different kind of information from the other predictors.

To obtain reliable enough conclusions from computed QSARs we use:

- the physical significance of predictors, based on PRECLAV/DRAGON documentation [32, 36-44] (for MlogP [45]);
- the mathematical sign of predictors in QSAR;
- the computed value of utility U;
- the result of "virtual fragmentation" analysis.

Results and discussions

The Pearson correlation r^2 and Spearman correlation ρ^2 coefficients of the activity values from table 2 are:

0.4098; 0.4695 for the pair *S. aureus – E. coli*

0.4385; 0.4941 for the pair *S. aureus – P. aeruginosa*

0.6052; 0.7210 for the pair *E. coli – P. aeruginosa*

Considering the number of these values (n = 100), the correlation can be considered "very large", according to usual statistical criteria. Consequently, even before any QSAR computations, we can guess, roughly speaking, that many of the molecular features determining the

Table 3
IDENTIFIED COIN OUTLIER MOLECULES (INDEX OF TABLE 1)

	Bacterium		
	S. aureus	*E. coli*	*P. aeruginosa*
step #1	**2, 5, 10, 17, 20, 24, 29, 33, 34, 41, 46, 47, 52, 61, 69, 80, 93**	**7, 17, 24, 73, 80, 93**	**7, 10, 20, 34, 41, 65, 67, 80, 84, 85, 93**
step #2	**1, 8, 22, 51, 55, 63**	**16, 45, 85, 91, 98**	**17, 56**
step #3	-	**10, 25, 90**	**6, 13, 68**
Possible outliers for lead hopping	**52**	**93**	**93**

bactericidal activity upon the three microorganisms are the same, especially in the case of the *E. coli – P. aeruginosa* pair.

In order to obtain the QSAR equations we have computed almost 2500 descriptors, and we have tried hundreds of thousands of combinations between them. Even so, the QSAR equations based on table 1, obtained before the outlier elimination, have a poor predictive quality. For *S. aureus* we identified only 88 significant descriptors and a type (1) equation was obtained with k = 5, s = 0.6246, r^2 = 0.5249, F = 21.0, r^2_{CV} = 0.4519. For *E. coli* we identified 126 significant descriptors and a type (1) equation was obtained with k = 10, s = 0.4932, r^2 = 0.6528, F = 16.9, r^2_{CV} = 0.5611. For *P. aeruginosa* we identified 168 significant descriptors and a type (1) equation was obtained with k = 7, s = 0.4894, r^2 = 0.6550, F = 25.2, r^2_{CV} = 0.5725.

The outlier identification procedure led to the results presented in table 3.

There are twenty-three *S. aureus* outliers, fourteen *E. coli* outliers and sixteen *P. aeruginosa* outliers in table 3. There are eleven quinolones in table 2, used against non-standard bacteria strains. Only two *S. aureus* outliers, only three *E. coli* outliers and only three *P. aeruginosa* outliers are used against non-standard bacteria strains. Consequently, the outlier feature of some quinolones cannot be explained by using a non-standard bacteria strain for MIC evaluation.

We know that all quinolones do the same thing – they inhibit the action of DNA-gyrase – although the details about how exactly this inhibition is achieved are not yet elucidated. A molecule might be an outlier due to the different manner in which the hydrogen bonds to the bacterial DNA are formed.

The molecules **52** and **93** - possible outliers for lead hopping - are very active (A_{obs} > 3.5) and high outliers (COIN > 4). The molecule **52** is the sole molecule that includes R_6 = F and R_7 = thiomorpholin-1-yl. The molecule **93** is the sole molecule that includes R_6 = NH_2 and R_7 = piperazin-1-yl.

Once the outliers were eliminated, a calibration set of 77 molecules was utilized for *S. aureus*, of 86 molecules for *E. coli*, and of 84 molecules for *P. aeruginosa*. Eliminating fewer outliers (e.g. only those identified during step #1) leads to QSAR equations with poorer predictive quality, and the information offered by those equations is less reliable.

QSAR #1 for activity against *S. aureus*

Calibration set: 77 molecules (table 1 molecules without 23 outliers of Table 3)

Number of significant descriptors: 190

The type (1) QSAR for prediction:

c_0 = -1.1251

c_1 = -1.1734; p_1 – Mor09v (3D-MoRSE descriptor [36, 37]) (U = 878)

c_2 = 0.3824; p_2 – parabolic function of "HOMO energy weighted molecular volume" (U = 285)

c_3 = 0.5317; p_3 – parabolic function of "Maximum free valence of F atoms" (U = 512)

c_4 = 0.0624; p_4 – RDF070u (RDF descriptor [38]) (U = 523)

c_5 = - 0.1502; p_5 – number of H atoms in R of R-Ar [39] (U = 1000)

c_6 = 0.9291; p_6 – L3u (WHIM descriptor [40, 41]) (U = 231)

Standard error of values: 0.2904

Standard error of ranks: 6.7746

Pearson square correlation r^2*:* 0.8850

Fisher F function: 91.1

Kendall rank correlation K = 0.8353

Pearson cross-validated correlation r^2_{CV}*:* 0.8415

Figure 3 presents the scatter-plot of observed/computed values of bactericidal activity related to *S. aureus.*

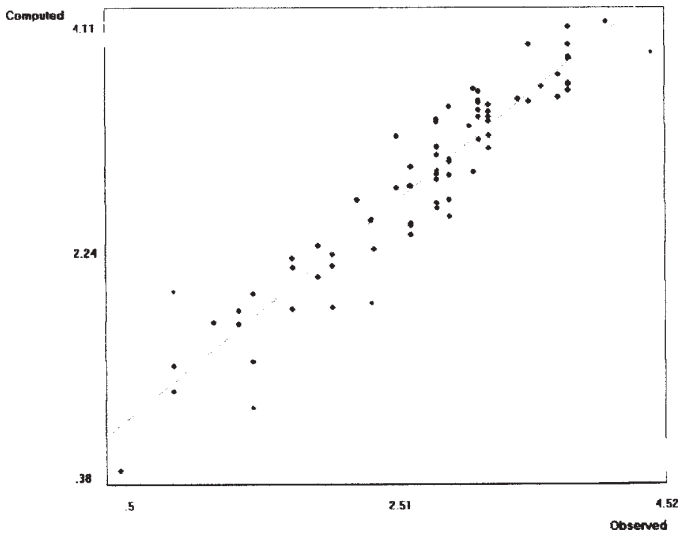

Fig. 3. Scatter-plot related to *S. aureus QSAR*

The value of predictors QSAR #1 may be received from authors.

The lowest correlation with activity is calculated for predictor p_6 (r^2 = 0.0792). The highest intercorrelation between predictors is calculated for the pair p_4, p_6 (r^2 = 0.3913).

The "useful" predictors (U > 500) for describing the bactericidal activity of the analyzed quinolones are p_1, p_2, p_4, and p_5. The physical significance of those "useful" predictors, and the sign of the coefficients in the QSAR equation, suggests that:

-the presence of the R-Ar fragment is detrimental to bactericidal activity upon *S. aureus;*

-there is an optimal number of fluorine atom grafted on aromatic cycle: the presence of two atoms is better than the presence of one or three such atoms;

-non-balanced molecules (i.e. substitutes R_7 and R_8 have large volumes and polarizabilities while substitutes R_1, R_5 and R_6 have small volumes and polarizabilities) favours the biochemical activity.

QSAR #2 for activity against *E. coli*

Calibration set: 86 molecules (table 1 molecules without 14 outliers of table 3)

Number of significant descriptors: 169

The type (1) QSAR for prediction:

c_0 = 1.9509

c_1 = 0.7658; p_1 – Number of NH single bonds (U = 1000)

c_2 = 0.5844; p_2 – number of N atoms attached to Ar [39] (U = 500)

c_3 = 3.1290; p_3 – R5m (GETAWAY descriptor [42, 43]) (U = 598)

c_4 = -1.1685; p_4 – product Percent of fluorine · Maximum charge of F atoms (U = 778)

c_5 = -15.3316; p_5 – R2e+ (GETAWAY descriptor [42, 43]) (U = 468)

c_6 = - 0.471; p_6 – parabolic function of "Percent of hydrogen" (U = 750)

c_7 = -1.3603; p_7 – R6e (GETAWAY descriptor [42, 43]) (U = 261)

c_8 = 0.0937; p_8 – Average distance between oxygen atoms (U = 750)

Standard error of values: 0.3583

Standard error of ranks: 10.3572

Pearson square correlation r^2: 0.7943

Fisher F function: 37.6

Kendall rank correlation K = 0.7609

Pearson cross-validated correlation r^2_{CV}: 0.7337

Figure 4 presents the scatter-plot of observed/computed values of bactericidal activity related to *E. coli.*

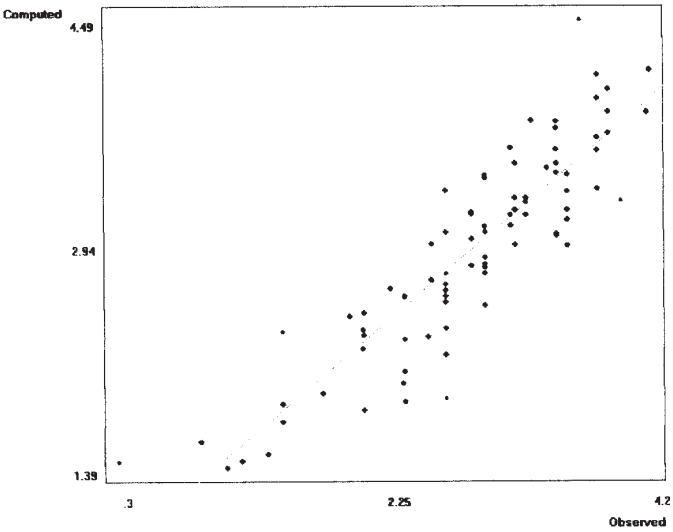

Fig. 4. Scatter-plot related to *E. coli QSAR*

The value of predictors QSAR #2 may be received from authors.

The lowest correlation with activity is calculated for predictor p_8 (r^2 = 0.0448). The highest intercorrelation between predictors is calculated for the pair p_3, p_7 (r^2 = 0.3574).

The "useful" predictors for describing the bactericidal activity of the analyzed quinolones are p_1, p_3, p_4, p_6 and p_8. The physical significance of those "useful" predictors, and the sign of the coefficients in the QSAR equation, suggests that:

-the biochemical activity upon *E. coli* is favored by :

i)the presence of halogen atoms grafted on the quinolone cycles, especially fluorine;

ii)the presence of unconjugated groups NH or NH_2;

iii)a large distance between the groups OH and COOH;

iv)small volume and polarizability of substitutes R_1, R_6 and R_7 (especially R_1 and R_7);

- the biochemical activity upon *E. coli* is disfavored by the presence of "keto" and "amide" groups.

QSAR #3 for activity against *P. aeruginosa*

Calibration set: 84 molecules (table 1 molecules without 16 outliers of table 3)

Number of significant descriptors: 253

The type (1) QSAR for prediction:

c_0 = -2.5379

c_1 = 0.6526; p_1 – Mor15e (3D-MoRSE descriptor [36, 37]) (U = 806)

c_2 = 0.6352; p_2 – parabolic function of "percent of NH single or weak bonds" (U = 1000)

c_3 = 0.6375; p_3 – parabolic function of "Sum of [net charge · exposed atomic surface] products (charge < 0)" (U = 963)

c_4 = -5.0864; p_4 – MATS2p (2D-autocorrelation descriptor [44]) (U = 482)

c_5 = 0.2128; p_5 – Moriguchi LogP [45] (U = 176)

c_6 = 1.5249; p_6 – HATS7e (GETAWAY descriptor [42, 43]) (U = 486)

Standard error of values: 0.2993

Standard error of ranks: 7.2528

Pearson square correlation r^2: 0.8645

Fisher F function: 82.9

Kendall rank correlation K = 0.8365

Pearson cross-validated correlation r^2_{CV}: 0.8415

Figure 5 presents the scatter-plot of observed/computed values of bactericidal activity related to *P. aeruginosa.*

The value of predictors QSAR #3 may be received from authors.

The lowest correlation with activity is calculated for predictor p_6 (r^2 = 0.0459). The highest intercorrelation between predictors is calculated for the pair p_2, p_5 (r^2 = 0.1930).

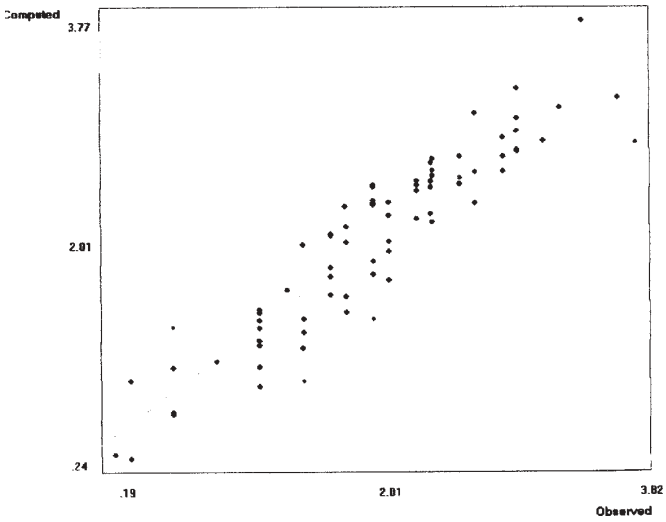

Fig. 5. Scatter-plot related to *P. aeruginosa QSAR*

The "useful" predictors for describing the bactericidal activity of the analyzed quinolones are p_1, p_2, and p_3. The physical significance of those "useful" predictors, and the sign of the coefficients in the QSAR equation, suggests that:

- from the point of view of activity, there is an optimal proportion of NH bonds;

- the presence of cycles favours the biochemical activity

- the smaller (especially for R_1) and the more numerous, the better.

We observe that the QSAR equations do not contain any "grid" predictors. Those descriptors have "lost" the mathematical competition, which was "won" by other kinds of descriptors. The same thing can be said about the descriptors measuring the lipophilic/hydrophilic feature (e.g. MlogP). Other authors noted as well the small influence of logP upon quinolones activity [46].

A small volume and low polarizability of substitute R_1 are favoring activity. This is also true for substitute R_7, but only for *E. coli*. We deduce that the structure of R_7 has a large influence (either favorable or unfavorable) upon the bactericidal activity of quinolones, which was also noted by other authors [47].

It is possible that the favorable influence of unconjugated groups NH/NH_2 is due to the appearance of intramolecular hydrogen bonds. Some authors noted the importance of zwitterions [48].

The analysis of virtual fragmentation descriptors, specific to PRECLAV, offers additional information. The value of those descriptors is weight percent of different molecular fragments and it is well correlated to the bactericidal activity. However, if the number of null values is large (i.e. a certain fragment is absent from many molecules), those descriptors are not utilized in computation of the QSAR equations. Here the sign of correlation *r* and value of square correlation r^2 of those descriptors to the bactericidal activity suggest that the presence of fragment CH (which reflects here the presence of a saturated cycle substituted at C atom) has a favorable influence ($r > 0$, $r^2 > 0.25$) upon the bactericidal activity against the three studied microorganisms. On the other hand, the presence of fragment NO_2 has a very negative influence ($r < 0$, $r^2 = 0.1558$) upon the activity against *S. aureus*, while the presence of fragments C_6H_4 or of triazole cycle has a very negative influence ($r < 0$, $r^2 > 0.12$) upon the activity against *E. coli* and *P. aeruginosa*.

The computed values of bactericidal activity, using QSAR #1, #2 and #3 may be received from authors.

Conclusions

The bactericidal activity against the three studied bacteria is favorably influenced by:

-the presence of saturated (hetero)cycles substituted at C atom;

-the presence of groups -NH/-NH_2- (unconjugated) and/or the presence of two fluorine atoms;

-a non-balanced molecular shape (i.e. substitutes R_7 and R_8 should have large volumes and polarizabilities while substitutes R_1, R_5 and R_6 should have small volumes and polarizabilities);

The bactericidal activity is unfavorably influenced by the presence of substitutes -NO_2, -C_6H_4, -CO-, and by the presence of triazole cycle.

The lipophilic/hydrophilic feature has a small influence on the bactericidal activity.

References

1.CHU, D.T., FERNANDES, P.B., Advances in Drug Research, **21,** 1991, p. 39

2.KOGA, H., ITOH A., MURAYAMA, S., J. Med. Chem., **23,** 1980, p.1358

3. CHU, D.T.W., PRABHAVATHI, B.F., J. Med. Chem., **28,** 1985, p. 1558

4.RADL, S., ZIKAN, V., Coll. Chem. Czech. Commun, **54,** 1989, p. 2181

5.NARAYANAN, S.R., KELLER, T.B., J. Med. Chem., **31,** 1988, p. 503

6.CECCHETTI, V., CLEMENTI, S., CRUCIANI, G., FRAVOLINI, A., PAGELLA, P. G., SAVINO, A., TABARRINI, O., J. Med. Chem., **38,** 1995, p. 973

7. ZIEGLER, C.B., BITHA, P., J. Med. Chem.,, **33,** 1990, p. 142

8. PETERSEN, U., GROHE, K., US Patent 5,153,204,1992.

9. FRIGOLA, J., PERES, J., J. Med. Chem., **36,** 1993, p. 801

10.FRIGOLA, J., TORRENS, A., J. Med. Chem.,,, **37,** 1994, p. 4105

11.KAZIKO, A., TSUYOSHI, K., J. Med. Chem., **36,** 1993, p. 1356

12.TOSHIO, U.M., J. Med. Chem., **30,** 1987, p. 2163

13.MYAMOTO, A., J. Med. Chem.,, **33,** 1990, p. 1645

14.REMUZON, P., BOUZARD, D., J. Med. Chem.,, **34,** 1991, p. 29

15.BOUZARD, D., DI CESARE, P., J. Med. Chem.,, **32,** 1989, p. 537

16.DOMAGALA, J.M., J. Med. Chem., **36,** 1993, p. 871

17.LABORDE, L., J. Med. Chem., ,**36,** 1993, p. 1964

18.ROSEN, T., J. Med. Chem., **31,** 1988, p. 1598

19.HAGEN, E.S., DOMAGALA, J.M., J. Med. Chem.,, **34,** 1991, p. 1155

20.DOMAGALA, J.M., BRIDGES, A.J., CULBERTSON, L.G., J. Med. Chem.,, **34,** 1991, p. 1142

21.SANCHEZ, J.P., BRIDGES, A.J., J. Med. Chem.,, **35,** 1992, p. 361

22.SUTO, M.J., J. Med. Chem., **35,** 1992, p. 4745

23.HIDETO, F., Bioorg. & Med. Chem. Lett., 1998, p. 2833.

24.SANCHEZ, J.P., DOMAGALA, J.M., HAGEN, S.E., J. Med. Chem.,, **31,** 1988, p. 983

25.SANCHEZ, J.P., GOGLIOTTI, R.D., DOMAGALA, J.M., J. Med. Chem.,, **38,** 1995, p. 4478

26.PCModel v.9.0 program is available from GILBERT, K., GAJEWSKI, J.J., Serena Software, Box 3076, Bloomington, IN, USA

27.MOPAC93 program is available from STEWART, J.J.P., 15210 Paddington Circle, Colorado Springs, CO 80921, E-mail: MrMOPAC@OpenMOPAC.net

28. TARKO, L. Rev. Chim. (Bucureºti), **56,** 2005, p. 639

29. PRECLAV program is available from TARKO, L., Center of Organic Chemistry (CCO)-Bucharest, Romanian Academy, E-mail: ltarko@cco.ro

30. TARKO, L. Rev. Chim. (Bucureºti), **55,** nr. 7, 2004, p. 539

31. TARKO, L. Arkivoc, 2004, xiv, p. 74

32. TARKO, L., LUPESCU, I., GROPOSILA-CONSTANTINESCU, D., Arkivoc, 2005, x, p. 254

33. DRAGON v. 5.4 program is available from TODESCHINI, R., Talete srl., via V. Pisani,13-20124, Milano, Italy, website http://www.talete.mi.it

34. CRAMER, R. D., JILEK, R. J., GUESSREGEN, S., CLARK, S. J., WENDT, B., CLARK, R. D., J. Med. Chem., **47,** 2004, p. 6777

35. SAEH, J. C., LYNE, P. D., TAKASAKI, B. K., COSGROVE, D.A., J. Chem. Inf. Comput. Sci., **45,** 2005, p. 1122

36. SCHUUR, J., SELZER, P., GASTEIGER, J., J. Chem. Inf. Comput. Sci., **36,** 1996, p. 334

37. SCHUUR, J., GASTEIGER, J., Anal. Chem., **83,** 1997, p. 2398

38. HEMMER, M.C., STEINHAUER, V., GASTEIGER, J., Vibrat. Spect., **19,** 1999, p. 151

39. VISVANADHAN, V.N., GHOSE, A.K., REVANKAR, G.R., ROBINS, R.K., J. Chem. Inf. Comput. Sci., **29,** 1989, p. 163

40. GRAMATICA, P., CONSONNI, V., TODESCHINI, R., Chemosphere, **38,** 1999, p. 1371

41. GRAMATICA, P., CORRADI, M., CONSONNI, V., Chemosphere, **41,** 2000, p. 763

42. CONSONNI, V., TODESCHINI, R., PAVAN, M., J. Chem. Inf. Comput. Sci. **42,** 2002, p. 682

43. CONSONNI, V., TODESCHINI, R., PAVAN, M., GRAMATICA, P., J. Chem. Inf. Comput. Sci., **42,** 2002, p. 693

44. BROTO, P., MOREAU, G., VANDICKE, C., Eur. J. Med. Chem., **19,** 1984, p. 79

45. MORIGUCHI, I., HIRONO, S., LIU, Q., NAKAGAME, I., MATSUSHITA Y., Chem. Pharm. Bull., **40,** 1992, p. 127

46. NIET, M. J., ALOVERO, F.L., MANZO, R.H., MAZZIERI, M.R., Eur. J. Med. Chem., **40,** 2005, p. 361

47. ALOVERO, F.L., PAN X. S., MORRIS J. E., MANZO R.H., FISHER, L.M.,

Antimicrob. Agents Chemother., **44,** 2000, p. 320
48. TAKAHASHI, H., HAYAKAWA, I., AKIMOTO, T., Yakushigaku Zasshi, **38,** 2003, p. 181
49. CLARK, D. E., Expert Opinion on Drug Discovery, **1,** 2006, p. 103
50. LANCE, R. P., Clinical Infectious Diseases, **33,** 2001, p. S180
51. XU, J.L., GAO, S.L., ZHANG, X. F., XIE, Y. P., HUANG, X. Y., XIE, X. G., Chinese J. Struc. Chem., **26,** 2007, p. 91
52. CECCHETTI, V., TABARRINI, O., SABATINI, S., MIAO, H., FILIPPONI, E., FRAVOLINI, A., Bioorg. Med. Chem., **7,** 1999, p. 2465
53. LESHER, G.Y., FROELICH, E. J., *GRUETT,* M. D., BAILEY, J. H., BRUNDAGE, P. R., J. Med. Chem.,, ***5,*** 1962, p. 1063
54. OHGOSHI, M., Chemotherapy, **22,** 1974, p. 1126
55. TARKO, L., LUPESCU, I., GROPOSILA-CONSTANTINESCU, D., Arkivoc, 2006, xiii, p. 22
56. BETERINGHE, A., FILIP, P., TARKO, L., Arkivoc, 2005, x, p. 45

Manuscript received: 22.10.2007