5. I. I. Chervin, V. S. Nosova, V. F. Rudchenko, V. I. Shevchenko, and R. G. Kostyanovskii, Khim. Geterotsikl. Soedin., No. 3, 396 (1988).
6. V. F. Rudchenko, V. I. Shevchenko, and R. G. Kostyanovskii, Izv. Akad. Nauk SSSR, Ser. Khim., No. 3, 598 (1986).
7. B. J. R. Nicolaus, G. Pagani, and E. Testa, Helv, Chim. Acta, 45, 358 (1962).
8. L. A. Simonyan, N. P. Gambaryan, P. V. Petrovskii, and I. L. Knunyants, Izv. Akad. Nauk SSSR, Ser. Khim., No. 2, 370 (1968).
9. J. Leitich, Angew. Chem. Int. Ed. Engl., 15, 372 (1976).
10. I. L. Charlton, C. C. Liao, and P. de Mayo, J. Am. Chem. Soc., 93, 2463 (1971).
11. V. F. Rudchenko, V. G. Shtamburg, A. P. Pleshkova, Sh. S. Nasibov, I. I. Chervin, and R. G. Kostyanovskii, Izv. Akad. Nauk SSSR, Ser. Khim., No. 11, 2539 (1981).

NEW METHOD FOR PREDICTING THE ANTIBACTERIAL PROPERTIES OF SEMISYNTHETIC PENICILLINS.

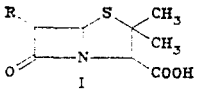
1. LOGICAL STRUCTURAL ANALYSIS

G. A. Veinberg, A. M. Kats, L. S. Gitlina, V. E. Golender, A. B. Rozenblit, and É. Lukevits

UDC 615.334:519.25:547.789.6'718

A new method for predicting the antibacterial properties of semisynthetic penicillins based on statistical evaluation of the merit of including functional groups, heteroatoms, and cyclic systems in fixed positions on the side chain of antibiotics is proposed.

The necessity of finding new medicinal semisynthetic penicillins is explained by the ability of pathogenic microorganisms to develop resistance to regularly used antibiotics. The methodology of identifying new effective compounds in a series of structural analogs of penicillin by antibacterial screening in vitro has remained unchanged for almost 30 years and is based mostly on the identification of empirical relations between the structure of the R group which is varied and the antimicrobial properties of the antibiotic I.



I

A characteristic feature of such an approach is the absence of any limitation on structural modification of the side chain. It is not surprising that the quantity of semisynthetic penicillins, I, synthesized in the 1960-1980 period is estimated at not less than 20 thousand and the structure of the R substituent in them is characterized by unusual variety.

However, recently the intensity of screening new compounds has markedly decreased, aided somewhat by an impasse in planning the synthesis of structural analogs of antibiotics. This important stage of the study until now was primarily based on intuitive evaluation of the usefulness of including in the side chain of I aliphatic, aromatic, and heterocyclic systems which have not been used or modification of those already used in practice (ampicillin, amoxycillin, etc.). Such a method of identifying promising substances is accompanied by the synthesis of an unjustifiably large quantity of their weakly active derivatives.

The main reason for the low productivity of this approach is the absence of an effective predictive method which allows a quantitative evaluation of the promise of including

Institute of Organic Synthesis, Academy of Sciences of the Latvian SSR, Riga 226006.
Translated from Khimiya Geterotsiklicheskikh Soedinenii, No. 3, pp. 396-403, March, 1989. Original article submitted February 18, 1988.

 0009-3122/89/2503-0332 $12.50

new structures in the side chain, using for this the accumulated information on antibacterial properties of semisynthetic penicillins which have already been synthesized. Numerous attempts to compare quantitatively the activity of antibiotics with their physicochemical parameters (lipophilicity, electronic effects of substituents, electron density on separate atoms, etc.) have not led to identification of a universal rule covering all structural types of semisynthetic penicillins [1-5].

In this respect, it was expedient to study the possibility of solving this problem using the automated TOPLOG system developed at the IOS AN Latv. SSR, which is a package of programs for logical analysis on a computer of the relationships between structures of compounds and their biological activity [6]. A characteristic feature of such an approach is the role of a function which reflects the specific structural features of the penicillin side chain and not the physicochemical parameters, but the combination of atoms bonded to each other which form the structure and determine the corresponding biological properties of the compounds. The merit of the approach is the possibility of studying the structure-activity relationship based on the most available information, the main part of which is data on the structure and activity of hundreds of compounds in vitro relative to standard test-microorganisms which have been revealed in reviews, scientific journals, and patent literature.

Development of the logical structural analysis of the semisynthetic penicillins using the TOPLOG system included the following steps: 1) formation of the data base; 2) division of the compounds into activity classes corresponding to given criteria; 3) choice and statistical evaluation of indicators; and 4) analysis of indicators using the computer and their use for directed synthesis of compounds with given properties.

FORMATION OF THE DATA BASE FOR STRUCTURE-ACTIVITY ANALYSIS

The comparative data on the structure and antibacterial activity in vitro of 1467 compounds revealed between 1961-1983 in 174 specialized reviews, articles, and patents comprised the data base for carrying out the logical structural analysis of semisynthetic penicillins by the TOPLOG system.

From the data base, 265 antibiotics representing basic structural types of semisynthetic penicillins were chosen for arranging the volume of analyzed material in the memory of the computer NR 1000.

DIVISION OF THE COMPOUNDS INTO ANTIBIOTIC CLASSES

Execution of the TOPLOG analysis guaranteeing the highest effect in a series of compounds of similar biological activity should include the threshold activity values (attainment and surpassing of which is the goal of structural modification). The threshold activity values for semisynthetic penicillins which were established in correspondence with the requirements representative for this type of antibacterial preparation [7, 8] are given in Table 1 together with the analogous indicators for medicinal semisynthetic penicillins.

The compounds chosen for analysis were separated according to threshold values into conditionally active (175) and inactive (90). The active ones are those whose minimum effective concentration (MEC) relative to one of the test-microorganisms in Table 1 would be equal to or lower than the corresponding threshold value. The inactive category includes antibiotics the MEC of which is higher than all the threshold values given in Table 1. By this approach, the carbenicillin is conditionally inactive because of the level of antibacterial activity which is less than the threshold indicators (Table 1).

DESCRIPTION OF THE STRUCTURE OF PENICILLINS

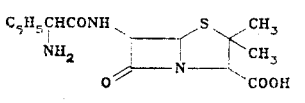
The method of describing the molecular structure of chemical compounds is the most important part of the logical structure analysis carried out on the computer. A variant matrix of binding of descriptive centers (MBDC) [6] was used in the TOPLOG system for this purpose. The potentially active structural elements (heteroatoms, functional groups, aromatic and cyclic systems, etc.) are understood to be descriptive centers (DC) and are designated by the corresponding code. Penicillanic acid was separated as an independent descriptive center for simplifying the analysis since it occurs in the composition of all compounds comprising the data base in an unchanged form.

TABLE 1. Threshold Values of the MEC and Level of Antibacterial Activity of Some Semisynthetic Penicillins

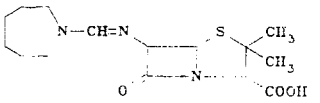
Bacterial type	Threshold values of the MEC, μg/ml	Level of antibacterial activity, μg/ml			
		ampicillin	carbenicillin	piperacillin	mecillinam
S. aureus (not producing β-lactamase)	0,25	0,05	1,5	0,8	5,0
S. aureus (producing β-lactamase)	5,0	250	25	100	100
E. coli	2,5	3	6	2	0,4
Ps. aeruginosa	5,0	500	50	3	100
P. vulgaris	1,0	500	12	6	100
K. pneumoniae	5,0*	250	250	12	2

A characteristic feature of the matrix codification was that it considered not only all structural elements which form the molecular structure but also their relative distribution due to the fixation of the chemical bonds separating them.

The scheme below shows examples of the matrix description of the ampicillin and mecillinam preparations.

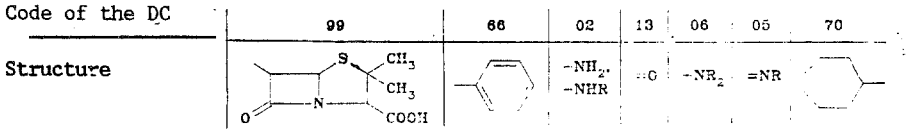


Ampicillin



Mecillinam

Descriptive centers contained in ampicillin and mecillinam



MBDC of ampicillin

	1	2	3	4	5
1	66	2	3	3	4
2		02	3	3	4
3			13	2	3
4				02	1
5					99

MBDC of mecillinam

	1	2	3	4
1	70	−1	2	3
2		06	2	3
3			05	1
4				99

Diagonal elements of the matrix are codes of the DC contained in penicillin; off-diagonal elements are the distance between DC expressed as the number of chemical bonds; inclusion of a nitrogen heteroatom (DC 06) in a saturated seven-membered ring (DC 70) is denoted by −1.

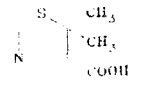
CHOICE AND STATISTICAL EVALUATION OF ACTIVITY INDICATORS

The choice of activity indicators of semisynthetic penicillins was based on an algorithm which located general fragments for the majority of structures represented by the matrices. The structural fragment is considered an activity indicator if the probability of activity of the new compound, P, containing this fragment exceeds a few threshold values:

$$P = (l_1 + l)/(l_1 + l_2 + 2) \geqslant t,$$

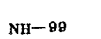
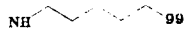
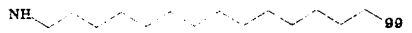
where l_1 is the number of active compounds of the data base which have the indicator which is being evaluated; l_2 is the number of inactive compounds of the data base which have the indicator examined. The threshold values t are determined using successive approximations [6].

TABLE 2. Some Statistical Data on DC Observed in Side Chains of Semisynthetic Penicillins (SP)

No. SP	DC		No. of compounds containing the DC			Probability of DC appearance in active compounds (P)
	code	structure	total	active	inactive	
1	99		265	175	90	0,66
2	02	$-NH_2$, $-NHR$	263	173	90	0,66
3	13	$=O$	257	168	89	0,65
4	66		242	159	83	0,66
5	41	$-CH_3$	95	65	30	0,68
6	64		7	5	2	0,66
7	62		4	3	1	0,66
8	03	$-NR_2$	64	39	25	0,61
9	32	F	8	5	3	0,60
10	05	$=NR$	62	52	10	0,83
11	46	$C=C$	60	48	12	0,79
12	11	$-OH$	54	48	6	0,87
13	67		33	29	4	0,86
14	31	Cl, Br, I	22	20	2	0,87
15	65		9	9	0	0,91
16	53		3	3	0	0,91
17	12	$-OR$	59	43	16	0,73
18	22	$-SR$	20	15	5	0,73
19	24	S^{+6}	15	11	4	0,71
20	56		20	12	8	0,59
21	04	$=NH$	21	11	10	0,52
22	54		4	2	2	0,50
23	50		4	2	2	0,50
24	60		4	1	3	0,33
25	55		8	2	6	0,30
26	52		5	1	4	0,28
27	30		5	1	4	0,28

A complete catalog of DC which are contained in the side chains, a list of active and inactive compounds which contain actual DC, and an evaluation of the probability of DC appearance in conditionally active and inactive compounds were revealed as a result of the structural–statistical analysis of the biological activity of penicillins using the TOPLOG system.

TABLE 3. Structural Indicators Consisting of Two DC

Structural indicator			No. of compds. containing the indicator	
substructure	number of bonds between DC	probability of appearance in active compds. (P)	active	inactive
	1	0,66	173	90
	6	0,55	14	11
	10	0,29	5	14
	13	0,17	0	4
	15	0,80	7	1
	4	0,33	1	3
	7	0,80	7	1
	8	0,92	35	2
	5	0,71	4	1
	8	0,93	38	2
	10	0,67	3	1

It was found that the structural diversity of 265 analogs of penicillin is provided for by variation of 40 DC. The number of their appearance in the analyzed antibiotics is changed from 1 to 265. The most extensive are amino, imino, and hydroxy groups, and aromatic and heterocyclic systems. Aldehyde and cyclobutyl groups are observed in the data base less than three times.

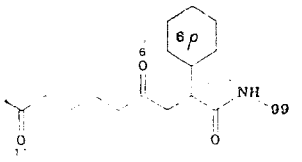
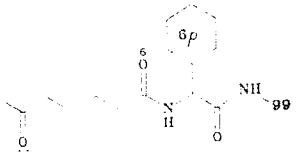
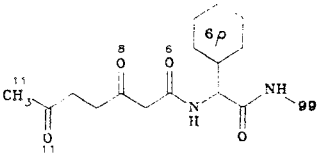
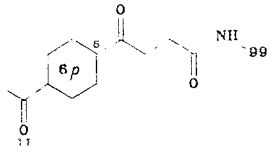
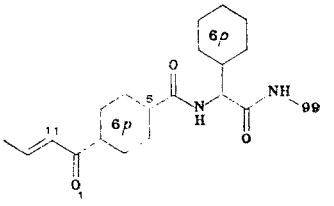
The quantitative characteristics for DC observed in not less than three compounds are given in Table 2. The indicators of this table give a descriptive representation of the structural material used for construction of the side chain of the semisynthetic penicillins and also allows qualitative evaluation of the merit of including the DC in the molecule. For this, the probability of appearance of an actual DC in the active compounds to the probability of appearance DC 99 (coding for penicillanic acid) which is contained in all antibiotics and characterizes the ratio of active and inactive compounds in the data base which is analyzed should be compared. Such comparison allows the DC given in Table 2 to be divided into three groups.

The first group (Nos. 2-9) includes DC coding for the amino group, carbonyl and terminal methyl groups, fluorine atoms, and six-membered cyclic systems containing 2, 4, and 6 p-electrons. The majority of these structural elements are characterized by a large number of representatives in the antibiotics. The probability of their appearance in the active antibiotics changes in the limits P = 0.60-0.68 and is comparable to the analogous indicator for DC 99. Consequently, being basic construction materials, they of themselves do not substantially affect the strengthening or weakening of the antibacterial properties of the antibiotics.

The second group (Nos. 10-19) contains DC coding for a nitrogen atom in a heterocyclic or aldimino group, hydroxy, ester and thioester groups, hexavalent sulfur, halogens (except fluorine), and 3, 5, and 7 p-electron five- and six-membered cyclic systems. The probability of their appearance in active penicillins occurs in the range P = 0.71-0.91 and exceeds the analogous indicator for DC 99, which indicates a definite improvement of the antibacterial properties of the antibiotics with their help.

The third group (Nos. 20-27) contains DC coding for unsubstituted aldimino groups, saturated three-, five-, and six-membered rings, and also five-membered rings containing 2, 4, 5, and 6 p-electrons. Judging from the values P = 0.28-0.59, their introduction into the side chain adversely affects the activity of the penicillins.

TABLE 4. Structural Indicators Containing DC 02, 13, 41, 46, and 66

Structural indicator		Number of compounds containing the indicator	
substructure*	probability of appearance in active compds. (P)	active	inactive
	0,93	13	0
	0,93	12	0
	0,80	3	0
	0,87	6	0
	0,86	5	0

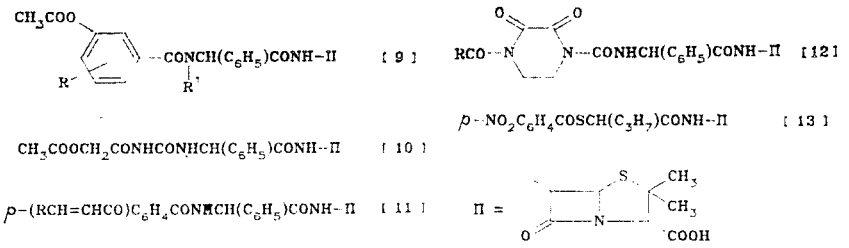
*The number in the formula indicates the number of bonds from DC 99

ANALYSIS OF THE PREDICTIVE PROPERTIES OF THE STRUCTURAL INDICATORS CHARACTERIZING THE SIDE CHAIN OF SEMISYNTHETIC PENICILLINS

A description of the molecular structures of antibiotics in matrix form accounting for the number of chemical bonds separating the DC allows the automated search of general fragments of the side chain of 265 compounds and reveals more than 2300 types of structural indicators consisting of two and more DC. For further analysis, only those which include DC 99, which is observed in not less than three compounds, were chosen.

Specific examples of structural indicators consisting of two DC allow evaluation of the merit of including amino, imino, and hydroxy groups in the N-acyl fragment of penicillin (Table 3). The data of Table 3 visually demonstrate the larger informativeness of statistical evaluation of the appearance of the indicator in active compounds accounting for the number of chemical bonds separating the structural element from the penicillin nucleus (DC 99), by comparison to the average indicators given in Table 2. Thus, the probability of appearance of amino and imino groups in active compounds depending on the number of bonds changes from 0.17 to 0.80, and without accounting for this factor, in agreement with the data of Table 2, DC 02 is characterized by P = 0.66. Analogous differentiation is followed for structural indicators containing the aldimino and hydroxy groups.

Examples of more complicated indicators yielding structural-statistical evaluation for some DC placed in fixed positions of the N-acyl fragment are given in Table 4. The high frequency of appearance of the indicators in active compounds indicates the promise of using them as structural bases for modification of antibiotics since they give information on the optimal distribution in the side chain for amide groups, carbonyl, and 6p-electron six-membered rings, etc. These indicators appear as a result of computational analysis of the five types of antibiotics



Thus, the study carried out allows a more efficacious variation of the placement in the side chain of original aliphatic, aromatic, and heterocyclic systems which have not been used before for this purpose, and also to construct new antibiotics combining the structural fragments of various indicators. Examples of practical use of this method for prediction of biological activity of new semisynthetic penicillins will be presented in the following article.

LITERATURE CITED

1. C. Hansch and E. W. Deutsch, J. Med. Chem., 8, 706 (1965).
2. B. Wetzel, E. Weitin, W. Reuter, R. Maier, and U. Lechner, Artzneim.-Forsch., 35, (1A), 343 (1985).
3. J. U. Indelicato, A. Dinner, L. R. Peters, and W. L. Wilham, J. Med. Chem., 20, 961 (1977).
4. B. Coene, A. Schanck, J. M. Dereppe, and M. V. Meerssche, J. Med. Chem., 27, 694 (1984).
5. D. B. Boyd, J. Med. Chem., 27, 63 (1984).
6. A. B. Rozenblit and V. E. Golender, Logical Combination Methods in Construction of Medicines [in Russian], Zinatne, Riga (1983).
7. H. Ferres, M. J. Basker, and P. J. O'Hanlon, J. Antibiot., 27, 922 (1974).
8. J. R. E. Hoover, Handbook of Experimental Pharmacology, Pt. 2, 67, 131 (1983).
9. N. Oi, B. Aoki, T. Shinozaki, K. Moro, I. Matsunaka, T. Noto, T. Nebashi, J. Harada, and H. Enso, U.S. Pat. No. 4229348, Chem. Abstr., 92, 198392 (1980).
10. K. G. Metzger and G. Schmidt, U.S. Pat. No. 3954730, Chem. Abstr., 83, 97277 (1975).
11. Y. Fujimoto and Y. Kishi, U.S. Pat. No. 4219477, Chem. Abstr., 91, 140857 (1979).
12. I. Saikawa, S. Takano, C. Yoshida, O. Takashima, K. Momonoi, S. Kuroda, M. Komatsu, T. Yasuda, and Y. Kodama, DE. Pat. No. 2519400, Chem. Abstr., 85, 33052 (1976).
13. A. Lutz and K. Volger, GB Pat. No. 1348198, Chem. Abstr., 79, 92208 (1973).