Faculty of Pharmacy, Meijo University, Nagoya, Japan

Free-Wilson analysis of the antibacterial activity of fluoronaphthyridines against various microbes. A new application of indicator variables

Y. TERADA and K. NANYA

The quantitative structure-activity relationship of the antibacterial activity of eighteen fluoronaphthyridines against nine species of bacteria was analyzed by the Free-Wilson method. A new indicator variable, which designates the species of bacteria, was useful in predicting the antibacterial activity against the various microbes.

1. Introduction

The Free-Wilson analysis is an important method used in the quantitative structure-activity relationship (QSAR) studies [1]. The Free-Wilson method uses indicator variables, which take the value of one or zero, depending on the presence or absence of a substituent at a certain position. The activity of a compound is calculated as the sum of the contribution of a reference compound and substituents [2, 3].

Bouzard et al. reported the antibacterial activity of fluoronaphthyridine derivatives against various microbes [4]. From the literature, we selected compounds **1**–**18** suitable for the Free-Wilson analysis and their antibacterial activity against nine species of bacteria.

The Free-Wilson analysis and the leave-one-out prediction of the eighteen compounds were performed for every species of bacteria. The predictions of these calculations were generally satisfying, however, the statistics were not sufficient. We then combined all the activity data for the various bacterial species and applied new indicator variables, which designate the species of the bacteria. The Free-Wilson analysis was performed using these indicator variables together with the usual indicator variables. Instead of the leave-one-out prediction, we employed the leave-nine-out prediction. In the leave-nine-out prediction, we removed all data concerning a compound from the training set and predicted the activity using the regression coefficients. While the accuracy of the prediction was slightly improved, the significance of the statistics was significantly improved.

The aim of the present paper is the evaluation of the new indicator variables for calculating the activity of the different bacterial species. This kind of indicator variable can be applied in various aspects and an improvement in the statistical reliability is expected.

2. Investigations, results and discussion

2.1 Individual analysis of the bacterial species

The antibacterial activity of the eighteen compounds toward individual bacterial species was analyzed by the Free-Wilson method. The resulting regression coefficients and the statistical parameters are shown in Table 1. In this table, a blank means a variable deleted by the F test and values with an asterisk are not significant at the 95% confidence level. The contributions of the substituents except for 2,5-diazabicyclo[2.2.1]heptyl ([dbHp]) were not always significant. This means that the variation in the activity is small and the frequency is not high enough to reveal the contribution of the substituents.

The antibacterial activity was predicted by the leave-one-out procedure for the individual bacterial species and the results are shown in Table 2. The average of the standard deviations for the individual bacterial species was 0.604 (n = 18, k = 5). The overall standard deviation estimated by all the data was 0.530 (assumed n = 162, k = 13).

2.2. Overall analysis of various bacterial species

The antibacterial activities for the nine bacterial species were combined and were analyzed by the variables designating the structure and the bacterial sensitivity. The activ-

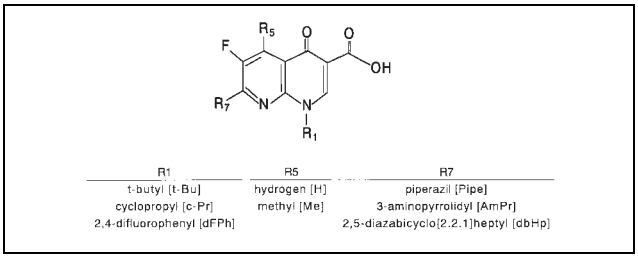

Fig. 1: Molecular structure of the studied compounds

Table 1: Group contributions and statistical parameters of the Free-Wilson analysis of the individual bacterial species

	[c-Pr]	[dFPh]	[Me]	[AmPr]	[dbHp]	const.	r[a]	s[b]	F[c]
Efa	−0.230*		0.262*	0.647	1.151	5.745	0.899	0.287	13.664
Sau	−0.330*		0.443	0.585	0.933	7.253	0.806	0.407	6.011
Eco					0.534	7.847	0.520	0.439	5.925
Ecl	0.668	0.568		0.419*	0.737	6.857	0.760	0.424	4.450
Pmi				0.515	0.808	6.750	0.671	0.404	6.156
Mmo	0.554				0.402	6.922	0.780	0.283	11.676
Sma	0.321*				0.476	6.821	0.571	0.427	3.691
Pae	0.401	0.301	−0.334	0.301	0.401	6.033	0.848	0.224	6.160
Kpn					0.534	7.573	0.501	0.461	5.360

[a] Correlation coefficient. [b] Standard deviation. [c] Variance ratio. * Not significant at the 95% confidence level.

ities predicted by the leave-nine-out procedure are shown in Table 2 and the standard deviation calculated using these data was 0.507 (n = 162, k = 13). Since the standard deviation is a measure of the fit of the data, depending only on the unexplained variance and the degrees of freedom [1], the prediction was slightly improved by the overall analysis. The group contributions and the statistic parameters are shown in Table 3. In this case, the statistic parameters were significantly improved. The indicator variable of *Pseudomonas aeruginosa* ([*Pae*]) was never significant, therefore, the sensitivity of this bacteria is almost equal to *Enterococcus faecalis* ([*Efa*]). It can be seen

Table 2: The observed and predicted antibacterial activity of 1,5,7-trisubstituted fluoronaphthyridines against the nine bacterial species

Compd.		1[a]	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	obsd.[c]	6.00	5.70	6.30	6.30	7.22	6.89	5.10	6.00	6.00	6.89	6.60	6.89	6.00	6.00	6.30	6.89	6.89	7.22
Efa[b]	pred.[d]	5.46	6.11	6.42	6.77	6.59	7.25	5.83	5.67	6.40	6.19	6.88	7.14	5.57	6.12	6.55	6.49	7.01	7.14
	pred.[e]	5.55	5.94	5.95	6.37	6.54	6.63	6.25	6.16	6.68	6.33	6.79	6.77	5.97	6.17	6.37	6.60	6.87	6.71
	obsd.[c]	7.22	7.52	8.52	7.70	8.40	8.40	6.60	7.52	6.89	8.70	7.80	8.40	7.52	7.80	7.80	8.10	8.10	8.70
Sau	pred.[d]	7.26	7.75	7.38	8.48	8.12	8.71	7.30	7.29	7.97	7.58	8.16	8.25	7.16	7.66	7.85	8.34	8.22	8.61
	pred.[e]	7.08	7.42	7.42	7.86	8.05	8.11	7.79	7.67	8.24	7.81	8.31	8.27	7.48	7.66	7.88	8.10	8.39	8.20
	obsd.[c]	7.70	7.22	8.10	7.52	7.70	8.40	7.52	8.10	7.22	8.70	8.70	8.70	7.80	7.80	8.40	8.10	8.10	8.70
Eco	pred.[d]	7.86	8.08	7.82	8.29	8.52	8.38	7.88	7.61	8.49	7.77	8.32	8.32	7.85	7.85	7.63	7.82	8.44	8.32
	pred.[e]	7.22	7.60	7.61	8.04	8.25	8.27	7.90	7.80	8.38	7.97	8.42	8.41	7.63	7.82	8.01	8.27	8.55	8.36
	obsd.[c]	7.22	6.60	7.70	6.89	7.52	7.52	7.22	7.52	7.52	8.70	7.80	8.70	7.52	7.52	7.80	7.52	7.80	8.70
Ecl	pred.[d]	7.50	7.56	6.91	7.85	8.10	8.10	7.86	7.53	8.11	7.50	8.44	7.87	7.39	7.39	7.86	7.97	8.30	7.96
	pred.[e]	6.85	7.24	7.24	7.68	7.87	7.93	7.53	7.44	7.97	7.58	8.08	8.02	7.25	7.45	7.65	7.91	8.17	7.97
	obsd.[c]	6.89	6.30	7.22	6.60	6.89	7.22	6.60	6.89	7.22	7.80	7.52	8.10	7.22	6.60	7.52	7.22	7.52	8.10
Pmi	pred.[d]	6.72	6.84	7.27	7.40	7.69	7.63	7.03	6.72	7.27	7.29	7.56	7.58	6.81	7.05	7.21	7.52	7.81	7.59
	pred.[e]	6.38	6.77	6.78	7.21	7.42	7.46	7.07	6.99	7.50	7.14	7.61	7.56	6.78	7.01	7.18	7.44	7.70	7.51
	obsd.[c]	7.22	6.60	7.22	6.89	6.89	7.22	7.22	7.22	7.22	7.80	8.40	7.80	6.89	6.89	7.22	6.89	7.22	7.52
Mmo	pred.[d]	6.73	6.96	6.88	6.93	7.43	7.35	7.54	7.54	7.71	7.40	7.68	7.91	6.93	6.93	6.88	6.93	7.35	7.27
	pred.[e]	6.41	6.80	6.83	7.24	7.47	7.51	7.09	7.02	7.55	7.19	7.61	7.63	6.85	7.05	7.25	7.51	7.77	7.60
	obsd.[c]	7.22	6.60	7.22	6.89	6.60	6.89	6.60	6.89	6.60	7.52	7.80	8.40	6.89	6.60	7.22	6.89	7.22	7.52
Sma	pred.[d]	6.57	6.96	6.77	6.81	7.56	7.51	7.27	7.20	7.53	6.87	7.33	7.21	6.81	6.96	6.77	6.81	7.44	7.24
	pred.[e]	6.25	6.64	6.67	7.08	7.32	7.37	6.96	6.88	7.42	7.05	7.48	7.44	6.69	6.90	7.09	7.35	7.61	7.44
	obsd.[c]	6.30	5.40	6.60	5.70	6.60	6.00	6.30	6.30	6.60	6.60	6.60	6.60	6.30	6.00	6.60	6.30	6.60	6.60
Pae	pred.[d]	5.90	5.85	6.20	6.15	6.35	6.15	6.50	6.00	6.80	6.00	6.95	6.11	6.35	6.00	6.65	6.30	6.80	6.00
	pred.[e]	5.55	5.94	5.95	6.37	6.54	6.63	6.25	6.16	6.68	6.33	7.71	6.77	5.97	6.17	6.37	6.60	6.87	6.71
	obsd.[c]	7.52	7.22	7.70	7.22	7.52	7.52	7.22	7.70	7.22	8.70	8.40	8.40	7.52	7.52	8.10	7.22	8.10	8.70
Kpn	pred.[d]	7.58	7.60	7.56	7.60	8.22	8.22	7.89	7.56	8.14	7.54	8.05	8.05	7.58	7.58	7.40	7.60	8.11	7.85
	pred.[e]	6.94	7.31	7.34	7.76	7.97	8.03	7.63	7.53	8.09	7.68	8.15	8.14	7.35	7.55	7.74	8.03	8.26	8.07

[a] The compound number and structure are shown in Fig. 1 and Table 4
[b] Abbreviation of the bacterial species are indicated in Table 5
[c] Antibacterial activity was taken from the reference [4] and converted into log (1/MIC)
[d] The leave-one-out prediction by the individual bacterial analysis
[e] The leave-nine-out prediction by the overall analysis

Table 3: Group contributions and statistical parameters of the Free-Wilson analysis of the overall 162 data

No.	[c-Pr]	[dFPh]	[Me]	[AmPr]	[dbHp]	[*Sau*]	[*Eco*]	[*Ecl*]	[*Pmi*]	[*Mmo*]	[*Sma*]	const.	r[a]	s[b]	F[c]	**1**
	0.425	0.383	0.163	0.481	0.814	1.530	1.666	1.300	0.829	0.862	0.700	1.385	5.551	0.886	0.373	42.510
2	0.285	0.243	0.131	0.341	0.674	1.477	1.658	1.302	0.828	0.863	0.701	1.367	5.808	0.868	0.387	35.570
3	0.419	0.378	0.160	0.298	0.720	1.471	1.660	1.290	0.827	0.880	0.718	1.392	5.654	0.883	0.376	41.220
4	0.258	0.217	0.149	0.459	0.720	1.493	1.667	1.311	0.873	0.873	0.711	1.394	5.761	0.879	0.380	39.487
5	0.292	0.250		0.387	0.759	1.505	1.710	1.328	0.874	0.926	0.781	1.430	5.783	0.882	0.378	44.930
6	0.296	0.254	0.124*	0.387	0.755	1.478	1.642	1.300	0.827	0.879	0.737	1.402	5.751	0.876	0.384	38.546
7	0.406	0.289		0.311	0.645	1.540	1.649	1.274	0.820	0.835	0.710	1.376	5.847	0.881	0.370	44.361
8	0.321	0.289	0.095*	0.396	0.730	1.512	1.642	1.283	0.829	0.862	0.720	1.374	5.741	0.871	0.393	36.657
9	0.442	0.289		0.498	0.720	1.558	1.702	1.292	0.819	0.871	0.745	1.411	5.738	0.895	0.357	51.664
10	0.213	0.289		0.269	0.720	1.478	1.642	1.249	0.811	0.863	0.717	1.351	5.846	0.883	0.362	45.418
11	0.324	0.289	0.105*	0.387	0.714	1.523	1.633	1.293	0.819	0.819	0.693	1.360	5.753	0.874	0.384	37.567
12	0.276	0.289		0.387	0.665	1.496	1.642	1.249	0.793	0.863	0.666	1.368	5.829	0.868	0.383	39.218
13	0.331	0.257	0.122*	0.418	0.752	1.512	1.660	1.283	0.810	0.882	0.720	1.385	5.709	0.873	0.392	37.268
14	0.331	0.309	0.114*	0.367	0.701	1.487	1.651	1.274	0.837	0.873	0.728	1.376	5.750	0.870	0.392	36.216
15	0.331	0.263	0.119*	0.360	0.720	1.514	1.642	1.284	0.810	0.880	0.718	1.369	5.745	0.872	0.391	37.075
16	0.331	0.339	0.134	0.436	0.720	1.504	1.668	1.309	0.836	0.908	0.746	1.429	5.691	0.880	0.383	39.897
17	0.331	0.338		0.387	0.769	1.513	1.677	1.302	0.827	0.897	0.735	1.386	5.765	0.873	0.390	41.204
18	0.331	0.230		0.387	0.661	1.488	1.652	1.258	0.803	0.889	0.727	1.360	5.819	0.868	0.383	39.202

[a] Correlation coefficient. [b] Standard deviation. [c] Variance ratio. * Not significant at the 95% confidence level.

Table 4: Structural matrix

	R_1			R_5		R_7		
Compd.	t-Bu	c-Pr	dFPh	H	Me	Pipe	AmPr	dbHp
1	1			1		1		
2	1				1	1		
3	1			1			1	
4	1				1		1	
5	1			1				1
6	1				1			1
7		1		1		1		
8		1			1	1		
9		1		1			1	
10		1			1		1	
11		1		1				1
12		1			1			1
13			1	1		1		
14			1		1	1		
15			1	1			1	
16			1		1		1	
17			1	1				1
18			1		1			1

from the Table that substitution of the C-5 hydrogen by a methyl group has almost no effect on the activity. Different from the individual bacterial analysis, the inclination of the group contributions and the bacterial susceptibility can be clearly seen from Table 3. Judging from the magnitude of the coefficient of the indicator variables, a more pronounced effect on biological activity is exhibited by the bacterial sensitivity than by the structural variation.

In the present work, an indicator variable designating the bacterial species was introduced into the Free-Wilson analysis. This variable is based on a comparable idea with the usual structural indicator variable. Because the bacterial susceptibility is an expression of drug-receptor interactions, we can consider that the parameter designating the bacterial species has the same dimension as the structural parameter.

In the present approach it is necessary to assume that all compounds act by the same mechanism and that the activity data are congeneric regardless of the bacterial species. Using the present indicator variable, it is possible to estimate the activity of a bacteria from the data of different bacteria. Moreover, this approach should be effective for various biological situations.

3. Experimental

3.1. Free-Wilson procedure

The minimum inhibitory concentrations (MIC, μg/ml) of the eighteen fluoronaphthyridines against the nine species of bacteria were taken from the literature [4]. These compounds have three, two and three different substituents in the three positions of the parent skeleton (N-1, C-5 and C-7, respectively as shown in Fig. 1). The other substituents and bacterial species in the literature were ignored, because their frequency was insufficient and caused blanks in the data matrix. The antibacterial activities transformed into log (1/MIC) are listed in Table 2 and the indicator variables for the substituents are shown in Table 4. The Free-Wilson method [2] in the Fujita-Ban modification [3] was used to calculate the substituent effect or the substituent and the bacterial sensitivity effects. In the regression calculations, variables were selected by the stepwise method.

The precision of the Free-Wilson method was examined using the cross-validation procedure. The accuracies of the leave-one-out and the leave-nine-out predictions were evaluated by the standard deviation (s) given by the following equation:

$$s = \sqrt{\frac{\sum (y_{obs} - y_{pred})^2}{n - k - 1}} \quad (1)$$

where n is the number of compounds and k is the number of variables.

All calculations were carried out on a personal computer NEC VALUESTAR NX, using the add-in software of Microsoft Excel [5].

Table 5: Bacterial species and their indicator variables

Enterococcus faecalis (Efa)	1							
Staphylococcus aureus (Sau)		1						
Escherichia coli (Eco)			1					
Enterobacter cloacae (Ecl)				1				
Proteus mirabilis (Pmi)					1			
Morganella morganii (Mmo)						1		
Serratia marcescens (Sma)							1	
Pseudomonas aeruginosa (Pae)								1
Klebsiella pneumoniae (Kpn)								1

3.2. Individual analysis of the bacterial species

The antibacterial activity of the eighteen compounds was analyzed for every bacterial species in Table 2. As the Fujita-Ban modification was applied, the indicator variables related to compound **1** ([t-Bu], [H] and [Pipe]) were eliminated and analyzed by the remaining five parameters.

3.3. Overall analysis of various bacterial species

All the MIC data against the nine bacterial species were combined and the leave-nine-out prediction was performed on the 162 activity data (Table 2). In this calculation, the nine activity data concerning a compound were predicted as a test set by the remaining 153 training data. The species of bacteria was represented by indicator variables as shown in Table 5. The two kinds of variables (Tables 4 and 5) were linked together, and a 16-bit dependent variable set was then formed. The activity of compound **1** against *Enterococcus faecalis* was treated as the reference; therefore, the related four variables ([t-Bu], [H], [Pipe] and [*Efa*]) were eliminated from the calculations.

References

1 Kubinyi, H.; in Ramsden, C.A. (Ed.): Comprehensive Medicinal Chemistry, Vol. 4, p589, Pergamon Press, Oxford 1990
2 Free, S.M.; Wilson, J. W.: J. Med. Chem., **7**, 395 (1964)
3 Fujita, T.; Ban, T.: J. Med. Chem., **14**, 148 (1971)
4 Bouzard, D.; Di Cesare, P.; Essiz, M.; Jacquet, J. P.; Leddoussal, B.; Remuzon, P.; Kessler, R. E.; Fung-Tonc, J.: J. Med. Chem., **35**, 518 (1992)
5 The Excel Statistics 97 Win Program, SSRI Co., Ltd., Tokyo

Received April 27, 1999
Accepted August 11, 1999

Dr. Yukimasa Terada
150 Yagotoyama, Tenpaku
Nagoya 468-8503
Japan
yterada@meijo-u.ac.jp