Fuzzy Adaptive Least Squares Applied to Structure–Activity and Structure–Toxicity Correlations

Ikuo MORIGUCHI,* Shuichi HIRONO, Yasuo MATSUSHITA, Qian LIU, and Izumi NAKAGOME

School of Pharmaceutical Sciences, Kitasato University, Shirokane, Minato-ku, Tokyo 108, Japan. *Received October 9, 1991*

A 1991 version of fuzzy adaptive least squares (FALS91) has been developed to analyze structure–activity rating data for generation of QSAR (quantitative structure–activity relationship) models. Improvements were made in this version especially to the error-correcting feedback adaptation of discriminant functions to avoid falling into a local optimum. This paper showed effective applications of FALS91 to structure–activity correlation of 29 antihypertensive arylacryloylpiperazine derivatives and structure–toxicity correlation of human acute toxicity of 504 organic chemicals as typical examples of small- and large-scale data sets, respectively. These examples demonstrated the high reliability of FALS91 in both recognition and leave-one-out prediction.

Keywords activity rating; antihypertensive activity; human acute toxicity; fazzy set; membership function; pattern recognition; structure–activity correlation; structure–toxicity correlation

Introduction

Biological activity and toxicity of chemical substances are often described in the form of activity ratings. For correlation of molecular structure with activity rating, we developed the adaptive least-squares (ALS) method,[1-3] which has been successfully used for a decade.[4-6] Activity ratings comprise not only statistical vagueness such as inaccuracy of measurements and individual differences in a living body but also intrinsic vagueness such as subjective criteria for classification. Such indefiniteness can be treated by the concepts of fuzzy variance. According to the technique of the fuzzy set theory,[7] we have introduced a membership function to ALS, to develop the fuzzy ALS (FALS). The membership function is assumed to be a fuzzy degree of membership in classes. A trial product, FALS89, was outlined in a previous paper.[8]

In the 1991 version of FALS (FALS91),[9] improvements were made especially to the error-correcting feedback adaptation of discriminant functions to avoid falling into a local optimum. In this paper, effective applications of FALS91 are shown to quantitative structure–activity relationship (QSAR) studies of 29 antihypertensive arylacryloylpiperazine derivatives and human acute toxicity of 504 organic chemicals as typical examples of large and small sets of data, respectively.

Methods

FALS91 Like ALS, FALS is a nonparametric pattern classifier, and formulates QSAR in a single equation (Eq. 1), irrespective of the number of activity ratings, by an error-correcting feedback adaptation.

$$Z = w_0 + w_1x_1 + w_2x_2 + \cdots + w_px_p \quad (1)$$

In this equation, $x_k = k$th descriptor ($k = 1, 2, \cdots, p$) for structures, w_k ($k = 0, 1, 2, \cdots, p$) = weight coefficient, and Z = discriminant score. For a set of n compounds with p structural descriptors, Eq. 1 can be rewritten as Eq. 2.

$$\boldsymbol{Z} = \boldsymbol{XW} \quad (2)$$

where the discriminant score vector $\boldsymbol{Z}$ consists of Z_i ($i = 1, 2, \cdots, n$), the data matrix $\boldsymbol{X}$ consists of x_{ik} ($i = 1, 2, \cdots, n; k = 0, 1, 2, \cdots, p; x_{i0} = 1$), and the weight vector $\boldsymbol{W}$ consists of w_k ($k = 0, 1, 2, \cdots, p$).

In the m-class discrimination, starting scores, a_j ($j = 1, 2, \cdots, m$), for the members of class j are assumed, and class boundaries, b_j ($j = 1, 2, \cdots, m-1$), are fixed in advance.

$$a_j = 4\left(\sum_{g=1}^{j-1} n_g + \frac{n_j}{2}\right) \Big/ n - 2 \quad (3)$$

$$b_j = (a_j + a_{j+1})/2 \quad (4)$$

In Eq. 3, n_g and n_j are the number of samples in classes g and j, respectively.

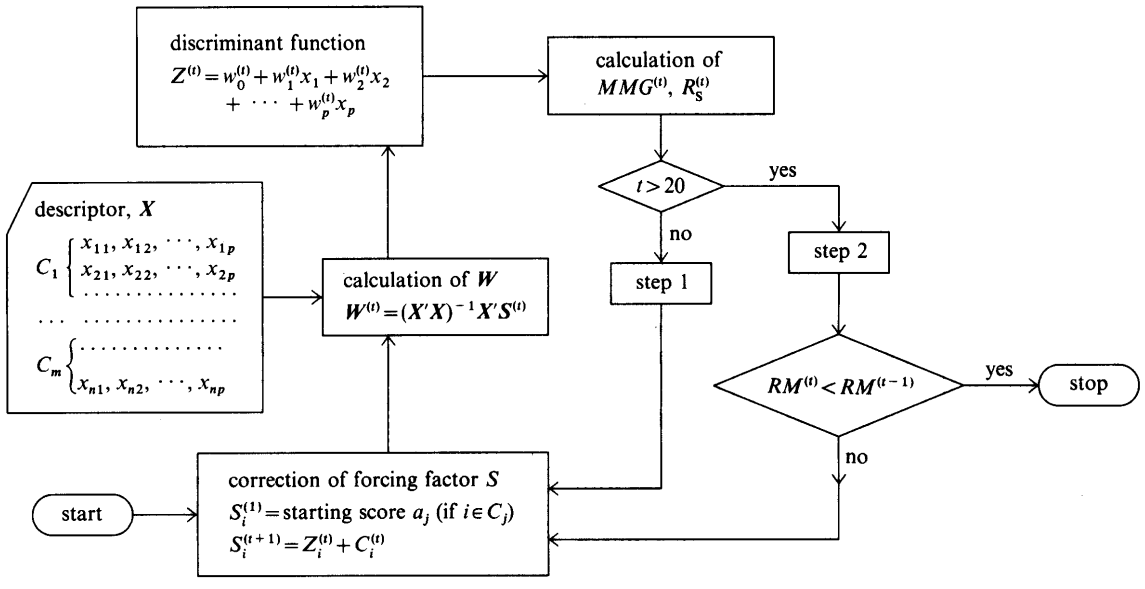


Fig. 1. Process of FALS91 Calculation
$RM^{(t)} = R_S^{(t)}MMG^{(t)}$

Classes are usually numbered in ascending order of potency.

The process of FALS91 calculation is shown in Fig. 1. Forcing factors, S_i ($i=1,2,\cdots,n$), are assumed in place of Z in Eq. 1 (or Eq. 2) for compounds, and W is calculated using the ordinary least squares, as

$$W^{(1)}=(X'X)^{-1}X'S^{(1)} \tag{5}$$

The iterative learning of W is done by error-correcting feedback adaptation of S as follows. The initial forcing factors, $S_i^{(1)}$, are taken to be

$$S_i^{(1)}=a_j \tag{6}$$

where a_j = the starting score for class j to which compound i was observed to belong. The adaptation of S is carried out in two steps: step 1 ($2\leq t\leq 20$) for rough but wide adaptation to avoid falling into a local optimum, and step 2 ($t>20$) for minute correction to obtain the best discriminant function. t is the iteration time.

At iteration 2 and thereafter, the forcing factor $S_i^{(t+1)}$ ($t\geq 1$) is adapted using the correction term $C_i^{(t)}$ as

$$S_i^{(t+1)}=Z_i^{(t)}+C_i^{(t)} \tag{7}$$

In step 1, the following correction term[3] is used for $C_i^{(t)}$.

$$C_i^{(t)}=\begin{cases} 0.1/(\delta_i^{(t)}+0.45)^2+0.1 & Z_i^{(t)}<b_{j-1} \\ 0 & b_{j-1}\leq Z_i^{(t)}\leq b_j \\ -0.1/(\delta_i^{(t)}+0.45)^2+0.1 & Z_i^{(t)}>b_j \end{cases} \tag{8}$$

Here, $\delta_i^{(t)}$ (>0) is the distance between $Z_i^{(t)}$ and the nearer boundary, b_{j-1} or b_j, for compound i which is a member of class j. A weight vector giving the best discrimination within 20 iterations is selected as the starting vector in step 2. The maximum Spearman's rank correlation coefficient with a local minimum apparent variance of errors[2] was used for the criterion of the best discrimination in the step 1 calculation.

In step 2, the following correction term[8] is used for $C_i^{(t)}$.

$$C_i^{(t)}=\begin{cases} \alpha\sqrt{(1-MG_i^{(t)})Fl} & Z_i^{(t)}\leq a_j \\ -\alpha\sqrt{(1-MG_i^{(t)})Fl} & Z_i^{(t)}>a_j \end{cases} \tag{9}$$

In Eq. 9, α is a constant usually taken to be 0.5. MG is the membership grade, which is the value of a membership function, $M(Z)$, assumed to give the grade of membership of classes for compounds. MG ranges from 0 to 1, and is taken to be 0.5 at the class boundaries. Fl is the parameter for fuzziness of the boundary between classes, and usually taken to be 0.1. A bell-shaped membership function, $M(Z)$ for class j is assumed as[8]

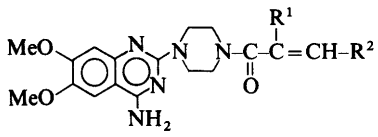
$$M(Z)=\begin{cases} 1/[1+\{(Z-b_{j-1})/Fl-1\}^4] & Z\leq b_{j-1}+Fl \\ 1 & b_{j-1}+Fl<Z\leq b_j-Fl \\ 1/[1+\{(b_j-Z)/Fl-1\}^4] & b_j-Fl<Z \end{cases} \tag{10}$$

Additionally, $M(Z)$ is taken to be 1 when $Z\leq b_1-Fl$ for class 1 and when $Z>b_{m-1}+Fl$ for class m. The iterative least squares calculation is carried out so as to minimize $\Sigma(S_i-Z_i)^2$, or ΣC_i^2 from Eqs. 2, 5, and 7. Therefore, as is obvious from Eq. 9, we can expect to obtain a discriminant function giving maximum ΣMG_i over the set of n compounds. As the criterion for the best discrimination in step 2, the product of mean membership grade (MMG) and Spearman's rank correlation coefficient (R_S) is used; R_S supplements the information about overall accuracy of discrimination.

The results of FALS91 are validated by the leave-one-out prediction.[10] The discriminant function with a scientifically reasonable subset of descriptors giving the best leave-one-out prediction is finally adopted.

As we have outlined, the main points of FALS91 differing from FALS89 are the use of the ALS81 correction term[3] in the step 1 calculation and

TABLE I. Structure and Activity of 29 Antihypertensive Arylacryloylpiperazines



No.	R^1	$R^{2a)}$	ΔRI	$\Sigma\sigma$	I (2-OR)	Obs Rating[b)]	Recog		Leave-one-out	
							Rating[c)]	MG	Rating	MG
1	H	4-iso-PrPh	2.36	−0.15	0.00	1	1	1.000	1	1.000
2	H	2-MeOPh	0.49	−0.27	1.00	1	1	1.000	1	1.000
3	H	2-EtOPh	1.18	−0.24	1.00	1	1	1.000	1	1.000
4	H	3,5-diMeOPh	0.56	0.24	0.00	1	2	0.004	2	0.002
5	H	3-NO_2Ph	0.03	0.71	0.00	1	1	1.000	1	1.000
6	H	3,4-diClPh	2.19	0.60	0.00	1	1	1.000	1	1.000
7	H	3-CF_3Ph	1.17	0.43	0.00	1	1	1.000	1	1.000
8	H	3-MeOPh	0.28	0.12	0.00	1	2	0.000	2	0.000
9	H	4-ClPh	1.01	0.23	0.00	1	1	1.000	1	1.000
10	H	4-BrPh	1.26	0.23	0.00	1	1	1.000	1	1.000
11	H	5-Cl-2-Th	0.89	0.26	0.00	1	1	1.000	1	0.994
12	H	3-MePh	1.01	−0.07	0.00	2	2	1.000	2	1.000
13	H	4-MePh	0.95	−0.17	0.00	2	2	1.000	2	1.000
14	H	4-iso-PrOPh	1.53	−0.45	0.00	2	2	0.998	2	0.734
15	H	2,3,4-triMeOPh	−0.09	−0.42	1.00	2	2	1.000	2	1.000
16	H	5-Me-2-Th	0.69	−0.14	0.00	2	3	0.080	3	0.044
17	H	2-MePh	0.83	−0.17	0.00	3	3	0.548	2	0.136
18	Me	Ph	0.20	0.00	0.00	3	3	1.000	3	1.000
19	H	3-Fu	−1.29	0.04	0.00	3	4	0.000	4	0.000
20	H	5-Me-2-Fu	0.04	0.15	0.00	3	3	1.000	3	0.997
21	Me	2-Th	−0.13	0.03	0.00	3	3	1.000	3	1.000
22	H	3-Me-2-Th	0.48	−0.14	0.00	3	3	1.000	3	1.000
23	H	Ph	0.00	0.00	0.00	3	3	1.000	3	1.000
24	H	2-Th	−0.36	0.03	0.00	3	3	0.999	4	0.036
25	H	4-MeOPh	0.19	−0.27	0.00	4	4	0.925	4	0.515
26	H	4-EtOPh	0.92	−0.24	0.00	4	3	0.000	3	0.000
27	H	2-Fu	−0.84	0.32	0.00	4	3	0.006	3	0.002
28	Me	2-Fu	−0.71	0.32	0.00	4	3	0.002	3	0.001
29	H	3-Th	−0.61	0.04	0.00	4	4	1.000	4	1.000

a) Ph: phenyl; Fu: furyl; Th: thienyl. *b*) Ratings (*ED*: dose for blood pressure lowering of more than 15%): 1, $ED>10$ mg/kg; 2, $ED=10$; 3, $ED=3$; 4, $ED=1$. *c*) Class boundaries: b_1, 0.175; b_2, 0.515; b_3, 0.860.

the use of a single *Fl* value common to all the boundaries. Compared with the FALS89 correction term,[8] the ALS81 correction term gives rougher but wider searching to avoid falling into a local optimum, although it gives no convergence in step 1. The use of a single *Fl* value assures the obtaining of maximum ΣMG_i in the step 2 calculation.

Calculation All calculations were carried out on a Sony NWS-830 computer and a Kobe Steel KTR-B08 transputer attached to an Epson PC-286VF microcomputer.

Results and Discussion

Structure–Activity Correlation of Antihypertensive Arylacryloylpiperazine Derivatives The first example of the application of FALS91 studied 29 arylacryloylpiperazine derivatives with antihypertensive activity, which have been well studied by ALS,[11] functional-link net (FUNCLINK),[12] and neural networks.[12,13] For comparison, we analyzed the same set of Sekiya's data[11] using FALS91. The structural and activity data are listed in Table I.

The FALS91 analysis gave Eq. 11 as the best QSAR model. The results of ALS81,[11] FUNCLINK,[12] and neural networks[12] are also shown here.

FALS91

$$Z = -1.088\Delta RI - 2.423\Sigma\sigma - 1.858I(\text{2-OR}) + 0.806 \quad (11)$$
$$(CI=0.91)\quad (CI=0.68)\quad (CI=0.57)$$
$$n=29(4\text{gr}),\ \text{recog: } n_{\text{mis}}=7(0),\ R_S=0.914,\ MMG=0.743$$
$$\text{leave-one-out: } n_{\text{mis}}=9(0),\ R_S=0.886,\ MMG=0.671$$

ALS81

$$L = -1.018\Delta RI - 2.357\Sigma\sigma - 1.838I(\text{2-OR}) + 0.769 \quad (12)$$
$$n=29(4\text{gr}),\ \text{recog: } n_{\text{mis}}=7(0),\ R_S=0.914$$
$$\text{leave-one-out: } n_{\text{mis}}=11(0),\ R_S=0.859$$

FUNCLINK

$$Iq = -8.035\Delta RI + 2.882\cos(\pi\Sigma\sigma) - 4.244I(\text{2-OR}) + 2.882 \quad (13)$$
$$n=29(4\text{gr}),\ \text{recog: } n_{\text{mis}}=7(0),\ R_S=0.914$$
$$\text{leave-one-out: } n_{\text{mis}}=8(0),\ R_S=0.898$$

neural network ($N_{\text{node}}=10$, $\alpha=0.7$, $\eta=0.9$)[12]

$$n=29(4\text{gr}),\ \text{recog: } n_{\text{mis}}=0(0),\ R_S=1.000 \quad (14)$$
$$\text{leave-one-out: } n_{\text{mis}}=13(5),\ R_S=0.642$$

In these equations, ΔRI is the difference of the high performance liquid chromatography (HPLC) retention index from that for compound 23, $\Sigma\sigma$ is the sum of the Hammett constants for aromatic substituents in R^2, and I(2-OR) is the indicator variable for the presence of 2-alkoxyphenyl structure in R^2, CI is the contribution index[2], n is the number of compounds, n_{mis} is the number misclassified, and the figure in parentheses after that is the number misclassified by two ratings. The squared correlation matrix for the three parameters included in Eq. 11 is listed in Table II. There appears no possibility of chance correlation from this matrix.

Equation 11 with negative coefficients for all the descriptors indicates that less hydrophobic compounds with electron-donating aromatic substituents in R^2 are favorable to activity, and the presence of 2-alkoxyphenyl structure in R^2 reduces the potency. The discrimination results were quite reliable: $R_S=0.914$ in the recognition, and $R_S=0.886$ in the leave-one-out prediction. The results of each compound are listed in Table I.

Table II. Squared Correlation Matrix of Descriptors in Eq. 11

	ΔRI	$\Sigma\sigma$	I (2-OR)
ΔRI	1.000		
$\Sigma\sigma$	0.007	1.000	
I (2-OR)	0.000	0.177	1.000

Equation 11 derived using FALS91 is very similar to Eq. 12 obtained by ALS81 which has been successfully used for a decade. Thus it can be said that the results of FALS91 are as reliable as those of ALS81. Moreover, Eq. 11 is statistically superior to Eq. 12 in the leave-one-out prediction. FUNCLINK further improves the leave-one-out prediction by use of non-linear descriptor, $\cos(\pi\Sigma\sigma)$ (Eq. 13); the validity of the use of non-linear descriptors is not always assured, however. Using the neural network (Eq. 14), all the compounds were correctly classified in the recognition. In contrast to this prominent recognition ability, the leave-one-out prediction is terribly poor. In addition to this, the neural networks do not construct any clear QSAR models. This data set from the Sekiya group[11] is very interesting, since it distinctly reveals the characteristics of the four methods.

Structure–Toxicity Correlation of Human Acute Toxicity of 504 Organic Chemicals The second example of the application of FALS91 concerns structure–toxicity correlation for predicting acute human toxicity of miscellaneous organic chemicals. Predicting human toxicity by computer has been an extremely important subject, since human toxicity cannot be experimentally measured. Toxicity involves various combinations of hazardous effects on multiple biological receptors, so toxicity ratings are often

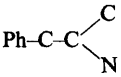
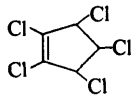
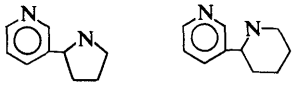
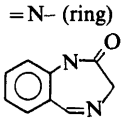
Table III. Typical Structures in the Three Ratings

Rating 1 (273 compds)	Rating 2 (150 compds)	Rating 3 (81 compds)
Fatty acids	Barbiturates	Nitrophenols
Carbohydrates	Phenols	Cyanides
Alcohols	Anilines	Organic phosphorous compounds
Carboxylic esters	Organic phosphorous compounds	Cardiac glycosides
Benzodiazepines		

Table IV. Results of FALS91 Analysis of 504 Organic Chemicals

Number of descriptors	Recognition				Leave-one-out prediction			
	N_{mis}	R_S	MMG	$R_S \cdot MMG$	N_{mis}	R_S	MMG	$R_S \cdot MMG$
34	64 (0)	0.863	0.853	0.736	91 (0)	0.814	0.812	0.661
35	64 (0)	0.862	0.852	0.734	93 (0)	0.810	0.817	0.662
36	62 (0)	0.868	0.856	0.743	87 (0)	0.823	0.822	0.676
37	62 (0)	0.868	0.853	0.740	87 (0)	0.826	0.814	0.672
38	60 (0)	0.874	0.857	0.749	83 (0)	0.837	0.818	0.685
39	60 (0)	0.874	0.857	0.749	86 (0)	0.829	0.817	0.677
40	58 (0)	0.881	0.859	0.757	87 (0)	0.825	0.818	0.674
41	58 (0)	0.877	0.858	0.753	90 (0)	0.820	0.810	0.664
42	57 (0)	0.878	0.862	0.757	90 (0)	0.821	0.813	0.668
43	57 (0)	0.880	0.862	0.758	88 (0)	0.828	0.813	0.673
44	56 (0)	0.882	0.862	0.760	88 (0)	0.824	0.812	0.669
45	55 (0)	0.883	0.862	0.761	89 (0)	0.819	0.816	0.668
46	55 (0)	0.882	0.870	0.767	86 (0)	0.818	0.826	0.676
47	54 (0)	0.884	0.868	0.768	88 (0)	0.813	0.821	0.667
48	53 (0)	0.887	0.871	0.772	87 (0)	0.813	0.822	0.668
49	52 (0)	0.888	0.872	0.774	96 (0)	0.801	0.815	0.652

TABLE V. Discriminant Function (Eq. 15) for Human Acute Toxicity of Organic Chemicals

No.[a]		Descriptor	Coef.	*CI*
1	B	Partially arom. polycycles	1.381	0.37
2	B	α,β-Unsat. lactones	3.100	0.34
3	B	>N–COO–X (X: phenyl, hetero ring, –N=)	1.628	0.29
4	B		1.551	0.23
5	B	Aliph. cyanides	1.394	0.20
6	A	Benzene rings, quinone rings	0.307	0.20
7	B	–P–O–X, –P–S–X (X: P, C=C, Ph, hetero ring, –C–C–N<, –C–C–S–)	1.121	0.18
8	B		1.525	0.18
9	B	Ar–X–Ar (X: >C=O, >N–, >CH–O–, >CHCO–)	0.892	0.17
10	B	Aziridines	1.477	0.16
11	B	Aliph. tertiary amines	0.531	0.16
12	A[b]	Proximity effect	0.050	0.13
13	A	Heteroarom. rings with 1 hetero atom	0.572	0.13
14	B	Tropines	1.589	0.12
15	B		1.452	0.11
16	B	$-PX_n-$ (X: 2 or more kinds of atoms out of F, N, O, S)	0.413	0.10
17	B	Nitrophenols	0.983	0.10
18	B	–P(=O)–X (X: F, N)	0.791	0.08
19	C[b]	Intramolecular hydrogen bonds	0.589	0.06
20	C[b]	Rings except benzene and fuzed rings with benzene	0.088	0.04
21	B	Arom. –CN	0.886	0.04
22	A	log *P* calcd.	0.014	0.02
23	B	X–C–Y (X: CN, halogen, NO_n; Y: –C=O, –O–)	0.111	0.01
24	A	C(sp^2, ring)	−0.056	0.26
25	B		−1.872	0.20
26	A	Aliph. –OH	−0.161	0.19
27	B	Ureas, thioureas	−0.741	0.15
28	A	–COO–, –CSO– (acids, esters, thioesters)	−0.255	0.14
29	A	=N– (ring)	−0.173	0.10
30	B		−0.936	0.10
31	B	–S–, $-SO_n-$, $-SO_2N<$	−0.370	0.10
32	B	$-SO_3H$	−0.364	0.09
33	B[b]	Amphoteric substructures	−1.870	0.08
34	A[b]	$(CX)^{0.6}$ (*CX*: sum of weighted numbers of carbon and halogen atoms)	−0.055	0.07
35	C[b]	Alkanes, cycloalkanes, alkenes and cycloalkenes	−0.930	0.07
36	A	C(sp, sp^2; non-ring)	−0.049	0.05
37	B	Aliph. secondary amines	−0.115	0.02
38	B	Aliph. ketones, aldehydes	−0.082	0.02
Constant			0.134	

a) A: numerical variable; B: seminumerical variable (taken to be 0, 1, or 2 for the absence, the presence of singular number, or the presence of plural number, respectively); C: dummy variable. *b*) Ref. 17.

used for the expression of toxicity levels. The following rating definitions based on a probable lethal dose were used in this FALS91 study: rating 1 (non- or slightly toxic), above 0.5 g/kg; rating 2 (toxic), 0.05—0.5 g/kg; rating 3 (severely toxic), less than 0.05 g/kg.

The data were collected mainly from Gosselin's compilation[14] which contains toxicological information about acute poisonings arising through misuse of consumer products. In addition, some estimated data of medicines[15] and general organic compounds[16] were used. The data set used for FALS91 analysis included 71 heteroaromatic compounds, 203 compounds bearing aromatic hydrocarbon or quinone ring(s), and 230 other miscellaneous organic compounds. The typical structures included in the three toxicity ratings are shown in Table III. As a matter of fact, structural and pharmacological features for each class are not particularly clear. For example, riboflavin (vitamin B_2) was assigned to rating 1 whereas menadione (vitamin K_3) to rating 2, sulfisoxazole was assigned to rating 1 whereas sulfamerazine to rating 2, and metharbital was assigned to rating 2 whereas amobarbital to rating 3.

For the FALS91 analysis, most descriptors investigated were those for molecular fragments and substructures, since the set of compounds included diverse structures and functionalities. The descriptors were divided into numerical and seminumerical parameters according to their effects on toxicity. Numerical parameters included physicochemical properties of compounds, and numbers of specified structural fragments. Seminumerical parameters were also used for the number of specified substructures, but in this case, they were taken to be 1 or 2 denoting the presence of singular number and plural number, respectively. The values of log *P* (*P*: octanol/water partition coefficient) were calculated using our simple method.[17] Several dummy variables were also investigated for the presence or absence of specified structures.

The results of FALS91 analysis using 34 to 49 descriptors are summarized in Table IV. In the recognition, a 49-descriptor equation gave the best result with 89.7% correct discrimination. However, the best leave-one-out prediction was obtained with the 38-descriptor equation (Eq. 15) shown in Table V.

In Table V, descriptors with positive coefficients and those with negative coefficients are respectively listed in the order of the contribution index which indicates the degree of contribution of descriptors to discrimination. Partially aromatic polycyclic structures, α,β-unsaturated lactones, special carbamates, *etc.* are probable to enhance acute toxicity, whereas sp^2 ring carbons, 1,2-methylenedioxybenzene structures, aliphatic alcohols, *etc.* probably contribute to lowering toxicity. Those coefficients, however, cannot be used to make inferences about the contribution of each substructure; they are valid only when used in the context of this multidimensional model. The maximum squared correlation between the descriptors was 0.485 (between descriptors 6 and 24); this may indicate no serious statistical problem in the derivation of Eq. 15. The correct classification into three ratings by Eq. 15 was 88.1% in the recognition and 83.5% in the leave-one-out prediction; a reasonably accurate structure–toxicity model could be generated for the estimation of human acute toxicity using FALS91.

Conclusion

The reliability of FALS91 shown in the two examples of the application is quite good in spite of the diversity of

structures and vagueness of potencies. In recent years, QSAR's in toxicity for large-scale sets of data have been studied and their use attempted by regulatory agencies and industry to screen compounds for possible health and environmental hazards. The computerized pattern classifier, FALS91, will become a useful tool for structure–activity and structure–toxicity correlation studies.

References

1) I. Moriguchi and K. Komatsu, *Chem. Pharm. Bull.*, **25**, 2800 (1977).
2) I. Moriguchi, K. Komatsu, and Y. Matsushita, *J. Med. Chem.*, **23**, 20 (1980).
3) I. Moriguchi and K. Komatsu, Abstracts of Papers, 8th Symposium on Structure–Activity Relationships, Tokyo, Oct. 1981, p. 55.
4) S. L. Rose and P. C. Jurs, *J. Med. Chem.*, **25**, 769 (1982).
5) P. C. Jurs, T. R. Stouch, M. Czerwinski, and J. N. Narvaez, *J. Chem. Inf. Comput. Sci.*, **25**, 296 (1985).
6) V. K. Gombar, E. P. Jaeger, and P. C. Jurs, *Quant. Struct.-Act. Relat.*, **7**, 225 (1988).
7) V. Novak, "Fuzzy Sets and Their Applications," Adam Hilger, Bristol, 1989, Chapter 9.
8) I. Moriguchi, S. Hirono, Q. Liu, Y. Matsushita, and T. Nakagawa, *Chem. Pharm. Bull.*, **38**, 3373 (1990).
9) I. Moriguchi, Q. Liu, S. Hirono, Y. Matsushita, and I. Nakagome, Abstracts of Papers, 19th Symposium on Structure–Activity Relationships, Kawaguchi, Nov. 1991, p. 244.
10) J. W. Tukey, *Ann. Math. Stat.*, **29**, 614 (1958).
11) T. Sekiya, "Structure–Activity Relationship and Drug Design," ed. by T. Fujita, Kagaku Dojin, Kyoto, 1986, p. 129.
12) Q. Liu, S. Hirono, and I. Moriguchi, *Quant. Struct.-Act. Relat.*, in press.
13) T. Aoyama, Y. Suzuki, and H. Ichikawa, *J. Med. Chem.*, **33**, 905 (1990).
14) R. E. Gosselin, R. P. Smith, and H. C. Hodge (eds.), "Clinical Toxicology of Commercial Products," 5th ed., Williams and Wilkins, Baltimore, 1984.
15) Pharmacopoeia of Japan, 12th ed., 1991.
16) D. V. Sweet (ed.), "Registry of Toxic Effects of Chemical Substances," 1985—86 Edition, U.S. Government Printing Office, Washington, 1987.
17) I. Moriguchi, S. Hirono, Q. Liu, I. Nakagome, and Y. Matsushita, *Chem. Pharm. Bull.*, **40**, 127 (1992).