MOLECULAR PHARMACOLOGY, 22:725–731

A Theorem on *de Novo* Group Contributions

ISTVÁN LUKOVITS

Central Research Institute of Chemistry, Hungarian Academy of Sciences, H-1525 Budapest, Hungary

Received January 8, 1982; Accepted July 7, 1982

SUMMARY

A theorem has been proved stating that the correlation coefficient calculated between two types of pharmacological potencies of series of compounds may be approximated by calculating the weighted correlation coefficient between the corresponding sets of Free-Wilsonian (*de novo*) group contributions. Comparison of activities of series with no common derivatives is also possible. *De novo* substituent constants also depend on other groups, if these emerge at the same substitution site. A simple method has been developed in order to take into account this interdependence. *De novo* substituent constants extracted from pharmacological potencies of phenylethylamine derivatives have been collected. Moderate but significant correlations have been demonstrated between sets of *de novo* substituent constants related to inhibition of uptake$_1$, uptake$_2$, phenyl-*N*-methyltransferase (PNMT), and of the pressor activity of epinephrine as well as for those extracted from potencies of substrates of PNMT. The method was compared with Zahradnik's approach. It was concluded that the existing moderate correlations between *de novo* group contributions are due to similar *nonspecific* drug-receptor interactions or transport processes.

INTRODUCTION

Pharmacological potencies of derivatives of a common "parent" molecule are often compared with activities of the same derivatives measured in a different test. Comparison of a series of sympathomimetic amines for lipolytic activity with effects on heart, bronchioles, and blood pressure led Lands *et al.* (1) to the conclusion that there are two types of *beta*-receptors, *beta*$_1$ and *beta*$_2$. The results of the comparative studies by Timmermans *et al.* (2) are considered later (see Discussion). However, comparative studies are impossible if series of molecules tested in two experimental assays do not possess common derivatrives or if there are only one or two common derivatives. It has been suggested (3) that, instead of pharmacological activities, the *de novo* substituent constants extracted from them should be compared in such cases. *De novo* constants may be calculated by using the Free-Wilson approach (4) or the Fujita-Ban model (5). According to these models, the activity of each molecule in a series of closely related compounds may be approximated by adding the contributions of the substituents and a constant term which is the contribution of the parent structure itself. The group contributions are the *de novo* substituent constants. This can be expressed in a compact mathematical notation:

$$A_i = \sum_{j=1}^{S} X_{ij}a_j + \mu \quad (i = 1, 2, \ldots, M) \qquad (1)$$

where A_i denotes the estimated activity of derivative i, whereas the actual activity of it is $\mathcal{A}_i$ and $\mathcal{A}_i \approx A_i$, a_j denotes the contribution of substituent j, and μ is the constant term. The contribution of a substituent is not equal to the contribution of the same group at a different site of the molecule, and two constants have to be calculated. The total number of substituents will be denoted by S, and M is the number of derivatives. Matrix X has M rows and S columns. The ijth element of it is $X_{ij} = 1$, if derivative i is substituted with group j, and $X_{ij} = 0$ if not. Non-equivalent substitution sites have been assumed throughout this paper.

The aim of the present paper is to perform comparative studies for the pharmacological activities of phenylethylamine derivatives (Fig. 1). *De novo* substituent constants published for these types of activities were collected from the literature. A theorem is proved stating that group contributions calculated by using the Free-Wilson approach may be used to demonstrate correlations between pharmacological responses. A method is derived in order to take into account the mutual interdependence of the contributions of groups belonging to the same site.

In order to state the theorem several notations and concepts must be introduced. Let us assume that there are k sites at which the molecules are modified by replacing their substituents with others. The number of groups appearing at site 1 is s_1, the number of groups at site 2 is s_2, etc., and the number of groups at site k is s_k. A series of derivatives is *complete* if all possible combinations of the existing groups appear, meaning that there are $s_1xs_2x, \ldots, xs_k = M_t$ derivatives altogether. Pharmacological activities of complete series are referred to

0026-895X/82/060725-07$02.00/0

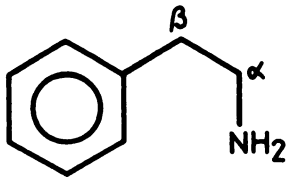

FIG. 1. *The phenylethalamine molecule*

as complete samples. Pharmacological activities of *incomplete* series are referred to as incomplete samples. Let us assume that each derivative in a complete series has been tested in two assays, the respective pharmacological potencies $\mathscr{A}$ and $\mathscr{B}$ were used to extract *de novo* substituent constants a and b and the estimated activities are A and B. The number of times a group j is contained in a series is denoted by m_j.

The following theorem is proved: The correlation coefficient r_{AB}, calculated between the estimated activities A and B, is equal to the weighted correlation coefficient $r_{ab}^{(w)}$ calculated between group contributions a and b, if these were extracted from complete samples. Weighting means that each pair of values a_j and b_j has to be considered m_j times.

Proof of the theorem.[1] Equation 1 can be written in a more condensed form:

$$A_i = c_i^{(A)} + \mu \quad (i = 1, 2, \ldots, M_t) \tag{2}$$

where $\sum_j X_{ij} a_j = c_i^{(A)}$. A similar expression defines $c^{(B)}$.

The correlation coefficient between variables A and B, r_{AB}, is equal to the correlation coefficient calculated between variables $c^{(A)}$ and $c^{(B)}$, since these are linearly related to A and to B, respectively.

$$r_{AB} = r_{c^{(A)}c^{(B)}} = \frac{\operatorname{cov}(c^{(A)}, c^{(B)})}{s_A s_B} \tag{3}$$

where the symbol $\operatorname{cov}(c^{(A)}, c^{(B)})$ denotes the covariance of variables $c^{(A)}$ and $c^{(B)}$.

$$\operatorname{cov}(c^{(A)}, c^{(B)}) = \frac{\sum_i (c_i^{(A)} - \bar{c}^{(A)})(c_i^{(B)} - \bar{c}^{(B)})}{M_t - 1} \tag{4}$$

and s_A is the standard deviation of variable $c^{(A)}$.

$$s_A = \sqrt{\frac{\sum_i (c_i - \bar{c}^{(A)})^2}{M_t - 1}} \tag{5}$$

s_b can be defined in a similar way. Both $\bar{c}^{(A)}$ and $\bar{c}^{(B)}$ are zero (Eq. 2).

Using the definition given for $c^{(A)}$, Eq. 4 can be written as:

$$\operatorname{cov}(c^{(A)}, c^{(B)}) = \frac{\sum_i \left(\sum_j X_{ij} a_j\right)\left(\sum_{j'} X_{ij'} b_{j'}\right)}{M_t - 1} \tag{6}$$

where the primed index j' indicates that it is not dependent upon index j. Expansion of this expression followed by a summation over i yields

$$\operatorname{cov}(c^{(A)}, c^{(B)}) = \left(\sum_j M_{jj} a_j b_j + \sum_j \sum_{j'} M_{jj'} a_j b_j\right)/M_t - 1 \tag{7}$$

where

$$M_{jj'} = \sum_{i=1}^{M_t} X_{ij} X_{ij'} \tag{8}$$

The diagonal element M_{jj} of matrix $\boldsymbol{M}$ is the number of occurrences, m_j, of substituent j. For complete series,

$$m_j = M_{jj} = \frac{M_t}{s_1} \tag{9}$$

where the site to which group j belongs is denoted arbitrarily. If substituent j' belongs to a different site 2, the number of simultaneous occurrences of substitutents j and j' is $M_{jj'}$. For complete series,

$$M_{jj'} = \frac{M_t}{s_1 s_2} \tag{10}$$

$M_{jj'}$ is constant for all values j' at site 2 and M_{jj} is also constant for all values j at site 1. For complete samples it can be written

$$\sum_j{}' a_j = 0 \tag{11}$$

for each site. Since $M_{jj'}$ is constant within a site, the off-diagonal elements of Eq. 7 vanish. From Eq. 1 it follows that $\bar{a} = 0$ and $\bar{b} = 0$, and the diagonal terms of Eq. 7 may be written as

$$\operatorname{cov}(c^{(A)}, c^{(B)}) = \frac{\sum_j m_j (a_j - \bar{a})(b_j - \bar{b})}{M_t - 1} \tag{12}$$

Analogously to the covariance, the standard deviation s_A can be expressed in terms of the group contributions a_j:

$$s_A = \sqrt{\frac{\sum_j m_j (a_j - \bar{a})^2}{M_t - 1}} \tag{13}$$

s_B may be obtained by replacing a_j with b_j and $\bar{a}$ with $\bar{b}$ in Eq. 13. The correlation coefficient can now be expressed by using Eqs. 3, 12, and 13 as:

$$r_{AB} = \frac{\sum_j m_j (a_j - \bar{a})(b_j - \bar{b})}{\sqrt{\sum_j m_j (a_j - \bar{a})^2 \sum_j m_j (b_j - \bar{b})^2}} \tag{14}$$

The right-hand term of this expression is the correlation coefficient $r_{ab}^{(w)}$ calculated between the variables a_j and b_j considered m_j times ($j = 1, 2, \ldots, S$). *Q.E.D.* In most cases the Free-Wilson approach approximates the experimental activities quite correctly and it can be assumed that $A_i \approx \mathscr{A}_i$ *and* $B_i \approx \mathscr{B}_i$, *and also* $r_{AB} \approx r_{\mathscr{AB}}$. Thus it follows that $r_{\mathscr{AB}} \approx r_{ab}^{(w)}$.

The off-diagonal elements in Eq. 7 do not vanish when incomplete samples are considered. However, it can be shown that the expected value of these terms is zero. For

[1] The symbol $\sum$ denotes that the variable following it has to be added for all values of the lower index; $\sum'$ means that addition is performed only for groups within a site; $\bar{x}$ denotes the mean value of variable x.

this it is enough to show that the expected value of the sum $\sum_{j'} M_{jj'} b_{j'}$ is zero for all values of index j. But this expression is simply a sample drawn from a population for which the expected value of contribution b is zero, and because of that the expected value of $\sum_{j'} M_{jj'} b_{j'}$ is also zero. Therefore, expected values of the off-diagonal terms of Eq. 7 are zero. Thus it follows that for incomplete samples the approximation $r_{\not a \not b} \approx r_{ab}^{(w)}$ is also true.

METHODS

Adjustment of group contributions. In practical cases the series to be compared are *not* identical. There might be several molecules which are contained in only one of the samples. Thus the sets of substituents to be compared might not be identical. Several groups appearing in one series may not occur in the second. *De novo* substituent constants are essentially regression coefficients and are not independent of one another. Therefore, only identical sets of group contributions are suitable for comparison. Derivatives bearing noncommon substituents should be omitted from the series. This procedure would be very time-consuming, since the Free-Wilson analysis must be repeated for each pair of samples. On the other hand, omission of any derivative reduces the reliability of the calculated *de novo* substituent constants. In order to take into account the effects of omission of a derivative, a method for correcting the group contributions is proposed here.

A series of five derivatives is considered. The results obtained may be generalized easily. All molecules are derivatives of a common parent structure which is modified at a single substitution site. The substituents are denoted by the letters *T*, *U*, *V*, *W*, and *Z*. The respective pharmacological potencies are A_T, A_U, A_V, A_W, and A_Z, and the mean activity is $\bar{A}$. The individual potencies can be expressed in terms of the group contributions a_T, a_U, a_V, a_W, and a_Z:

$$A_T = a_T + \bar{A} \tag{15}$$

$$A_U = a_U + \bar{A} \tag{16}$$

$$A_V = a_V + \bar{A} \tag{17}$$

$$A_W = a_W + \bar{A} \tag{18}$$

$$A_Z = a_Z + \bar{A} \tag{19}$$

The group contributions satisfy the symmetry equation (Eq. 11):

$$a_T + a_U + a_V + a_W + a_Z = 0 \tag{20}$$

The contributions are obtained by solving the system of Eqs. 15–19.

Now suppose that for some reason the derivative bearing group *Z* has to be omitted from the sample. The mean value of the remaining data (A_T, A_U, A_V, and A_W) will be denoted by $\bar{A}'$. The activities can be expressed in terms of the group contributions a'_T, a'_U, a'_V, and a'_W:

$$A_T = a'_T + \bar{A}' \tag{21}$$

$$A_U = a'_U + \bar{A}' \tag{22}$$

$$A_V = a'_V + \bar{A}' \tag{23}$$

$$A_W = a'_W + \bar{A}' \tag{24}$$

where the *de novo* substituent constants a'_T, a'_U, a'_V, and a'_W satisfy the symmetry restriction (Eq. 11):

$$a'_T + a'_U + a'_V + a'_W = 0 \tag{25}$$

Comparison of Eqs. 20 and 25 shows that the primed constants cannot be equal to the unprimed *de novo* constants. Solving Eq. 15 for a_T and Eq. 21 for a'_T, after division and rearranging a'_T can be expressed in terms of a_T:

$$a'_T = a_T \frac{A_T - \bar{A}'}{A_T - \bar{A}} \tag{26}$$

Similar expressions can be derived for a'_U, a'_V, and a'_W. The same rule can be applied if more groups are discarded.

In order to generalize this result, a complete series bearing group *T* at a specific site is considered. The mean activity of the derivatives is denoted by $\bar{A}_T$. Replacement of group *T* by group *U* yields a second series which is complete, and the mean activity of it is $\bar{A}_U$. The mean activities $\bar{A}_V$, $\bar{A}_W$, and $\bar{A}_Z$ are defined in a similar way, whereas $\bar{A}_t$ denotes the over-all mean activity. $\bar{A}'_t$ denotes the mean activity of four complete samples remaining after discarding all data on *Z* derivatives. Replacing values $\bar{A}_T$, $\bar{A}_U$, $\bar{A}_V$, $\bar{A}_W$, and $\bar{A}_Z$ in Eqs. 15–19 and in Eqs. 21–24, the arguments leading to Eq. 26 can be used again and one may write:

$$a'_T = a_T \frac{\bar{A}_T - \bar{A}'_t}{\bar{A}_T - \bar{A}_t} \tag{27}$$

Similar expressions can be derived for the other groups. This result is equivalent to the outcome of the derivation starting from direct expansion of determinants defining the group contributions.[2]

De novo substituent constants related to different sites are independent of one another. This is easily shown. Instead of matrix *X* (Eq. 1), a contracted matrix *Y* is considered to calculate the group contributions *a* (6, 7). Matrix *Y* ensures that the symmetry equations are automatically fufilled within the Free-Wilson approach. It can be easily proved by working out a numerical example, that the correlation between columns of matrix *Y* is zero for complete series, if the columns refer to different sites.

Data sets. Group contributions extracted from pharmacological activity data of PEA[3] derivatives are listed in Table 1. All values are dimensionless figures since they refer to activities expressed in $-\log C$ units, where *C* denotes the concentration needed to evoke a definite biological response. The activities of substrates of PNMT are the rates of formation of *N*-methylated PEA derivatives. The group contributions a_{PS} related to this activity were calculated from logarithms of the reaction rates (5). The values a_{PI} originate from inhibition of PNMT (3), whereas values a_{BE} and a_{BN} denote the contributions extracted from blocking potencies of *N*,*N*-dimethyl-2-

[2] I. Lukovits, unpublished results.

[3] The abbreviations used are: PEA, phenylethylamine; PNMT, phenyl-*N*-methyltransferase.

TABLE 1

Group contributions of substituents of PEA derivatives and mean activities related to individual substituents[a]

Group	Blocking of the pressor action of epinephrine norepinephrine[b]				Inhibition of PNMT[c]		PNMT substrate activity[d]		Inhibition of			
									Uptake$_1$[e,i]		Uptake$_2$[e,i]	
	a_{BE}	$\bar{A}_X$	a_{BN}	$\bar{A}_X$	a_{PI}	$\bar{A}_X$	a_{PS}	$\bar{A}_X$	a_{U1}	$\bar{A}_X$	a_{U2}	$\bar{A}_2$
H_o					0.014 (44)	3.19						
F_o					0.003 (3)	3.24						
Cl_o					0.253 (4)	3.55						
Me_o					0.016 (3)	2.98						
OMe_o					−0.847 (2)	1.58						
H_m	−0.252 (6)	8.62	−0.291 (6)	8.59	−0.122 (14)	3.07	−0.043 (5)	1.11	0.035 (14)	1.92	0.048 (5)	2.02
F_m	−0.553 (1)	7.52	−0.543 (1)	7.49	0.541 (3)	3.69						
Cl_m	−0.045 (4)	8.56	−0.033 (4)	8.56	1.070 (3)	4.49						
Br_m	0.182 (5)	8.93	0.185 (5)	8.93								
I_m	0.327 (1)	8.40	0.367 (1)	8.40								
OH_m					−0.157 (2)	3.04	−0.065 (7)	1.35	−0.351 (13)	2.34	0.569 (4)	1.71
Me_m	0.202 (5)	8.95	0.227 (5)	8.96	−0.161 (3)	2.95						
OMe_m					−0.774 (3)	1.91	0.671 (1)	2.00	1.358 (3)	−0.03	−0.838 (3)	3.01
H_p	−0.623 (6)	8.05	−0.652 (6)	8.02	−0.051 (13)	3.08	0.319 (5)	1.22	0.113 (12)	2.04	−0.331 (3)	1.91
F_p	−0.283 (4)	8.44	−0.279 (4)	8.43	0.043 (3)	3.31						
Cl_p	0.144 (4)	8.86	0.156 (4)	8.86	0.710 (3)	4.24						
Br_p	0.397 (4)	9.12	0.396 (4)	9.10								
I_p	0.806 (1)	9.25	0.896 (1)	9.29								
OH_p					0.084 (3)	3.41	−0.296 (7)	1.27	−0.237 (16)	1.99	0.041 (7)	2.33
Me_p	0.633 (3)	9.37	0.642 (3)	9.37	−0.094 (3)	2.95						
OMe_p					−0.522 (3)	2.03	0.479 (1)	2.00	1.218 (2)	0.39	0.353 (2)	2.00
H_α[f]							0.024 (12)[g,h]		0.078 (24)[g,h]		−0.031 (12)[g]	
Me_α[f]							−0.178 (2)[g,h]		−0.282 (7)[g,h]		0.713 (1)[h]	
$H(R_\alpha)$							−0.061 (12)[h]		−0.028 (26)[h]			
$Me(R_\alpha)$							0.737 (1)[h]		0.159 (5)[h]			
$H(S_\alpha)$							0.116 (11)[h]		0.204 (22)[h]			
$Me(S_\alpha)$							−0.635 (2)[h]		−0.516 (9)[h]			
H_β[f]							−0.814 (7)[g,h]		−0.108 (22)[g,h]		0.143 (10)[h]	
OH_β[f]							0.950 (6)[g,h]		0.298 (8)[g,h]		−0.543 (6)[h]	
$H(R_\beta)$							−1.097 (6)[h]		−0.052 (20)[h]			
$OH(R_\beta)$							0.804 (8)[h]		0.097 (11)[h]			
$H(S_\beta)$							−0.631 (9)[h]		−0.153 (25)[h]			
$OH(S_\beta)$							1.192 (5)[h]		0.682 (6)[h]			
NH							0.093 (11)		−0.153 (21)		0.121 (9)	
NMe							−0.509 (2)		0.357 (9)		−0.363 (3)	

[a] The values $\bar{A}_X$ are given for the ring substituents only. Numbers of occurrences are shown in parentheses.
[b] The mean activities are given in −log ED_{50} (M/kg) units (9).
[c] The mean activities are given in −log I_{50} (M) units (3).
[d] The mean activities are given as logarithms of relative reaction rates (5).
[e] The mean activities are given in log I_{50} units with respect to PEA (11).
[f] These contributions are related to racemic mixtures.
[g] These contributions are weighted means of the respective values given for antipodes.
[h] The numbers of occurrences given in refs. 5 and 11 are rounded off in this paper.
[i] The actual group contributions a_{U1} and a_{U2} (11) were multiplied by −1 for the sake of compatibility.

bromophenylethylamines (3, 8) against the pressor action of epinephrine and norepinephrine, respectively. Values of a_{PS}, a_{PI}, and a_{BE} were calculated by using the Fujita-Ban (5) approach and had to be transformed[4] into the Free-Wilson type contributions listed in Table 1.

The contributions a_{BN} were calculated from the original data of Graham and Karrar (9). The observed and estimated potencies are listed in Table 2. R denotes the multiple correlation coefficient, s is the standard error of the estimate, and F is the result of Fisher's F-test. Comparison with the theoretical (10) F-value ($F_{10,11,p=0.005}$ =

[4] The Fujita-Ban contributions [$a^{(FB)}$] may be expressed (8) in terms of the Free-Wilsonian ones [$a^{(FW)}$] as follows:

$$a^{(FB)} = a^{(FW)} - a_H^{(FW)},$$

where $a_H^{(FW)}$ denotes the contribution of atom H. For each site the respective value of $a_H^{(FW)}$ has to be considered. On the other hand, values of $a_H^{(FW)}$ can be expressed in terms of Fujita-Ban contributions using the following formula:

$$a_H^{(FW)} = -\frac{\sum_j m_j a_j^{(FB)}}{M}$$

for each site. From values $a^{(FB)}$ and $a_H^{(FW)}$, the other Free-Wilsonian contributions [$a^{(FW)}$] can also be expressed.

TABLE 2

Experimental and estimated blocking potencies of N,N-dimethyl-2-bromophenylethylamine derivatives against the pressor activity of norepinephrine and epinephrine (in $-\log ED_{50}$ M/kg units)[a]

Substituents		Norepinephrine		Epinephrine	
m	*p*	Observed	Estimated[b]	Observed	Estimated[c]
H	H	7.25	7.74	7.46	7.82
H	F	8.13	8.11	8.16	8.16
H	Cl	8.68	8.55	8.68	8.59
H	Br	8.89	8.79	8.89	8.84
H	I	9.29	9.29	9.25	9.25
H	Me	9.28	9.04	9.30	9.08
F	H	7.49	7.49	7.52	7.52
Cl	H	8.15	8.00	8.16	8.03
Br	H	8.30	8.22	8.30	8.26
I	H	8.40	8.40	8.40	8.40
Me	H	8.52	8.26	8.46	8.28
Cl	F	8.19	8.37	8.19	8.37
Br	F	8.57	8.59	8.57	8.60
Me	F	8.82	8.63	8.82	8.63
Cl	Cl	8.89	8.81	8.89	8.80
Br	Cl	8.92	9.03	8.92	9.02
Me	Cl	8.96	9.07	8.96	9.04
Cl	Br	9.00	9.05	9.00	9.05
Br	Br	9.30	9.27	9.35	9.28
Me	Br	9.22	9.31	9.22	9.30
Me	Me	9.30	9.55	9.30	9.53
Br	Me	9.52	9.51	9.52	9.51

[a] Data from ref. 9.
[b] Data from refs. 3 and 8.
[c] $R = 0.959$, $s = 0.23$, $F = 12.45$.

5.85) indicates that the multiple linear regression equation is significant at the $p \leq 0.005$ level.

Inhibitory activities against norepinephrine uptake$_1$ and uptake$_2$ (11) were expressed in log C units; the signs of the corresponding group contributions a_{U1} and a_{U2} were changed for the sake of compatibility.

Numerical calculations. Linear regression analysis has been used to demonstrate correlations between sets of *de novo* substituent constants. The pairs of contributions a_j and b_j were considered m_j times. In cases when the number of occurrence of group j is different in the two series, the smaller figure has been considered. The sum of these weighting factors is the total number of data points denoted by n. For each substituent j the mean activity $\bar{A}_j$ is also listed in Table 1. The over-all mean $\bar{A}$ is the arithmetical mean of values $\bar{A}_j$. The same definition is valid also for values $\bar{A}'$.

RESULTS

The correlation coefficient calculated between blocking potencies against the pressor activity of epinephrine and of norepinephrine (Table 2) is $r = 0.997$, whereas the correlation coefficient calculated between the estimated activities is $r = 0.999$. For the extracted group contributions the following relationship could be demonstrated:

$$a_{BE} = 0.95(\pm 0.01)a_{BN} + 0.00 \quad \begin{matrix} n & s & r & F \\ 44 & 0.02 & 0.999 & 18801 \end{matrix} \qquad (28)$$

where the number in parentheses denotes the 95% confidence interval of the regression coefficient. The calculated intercept is 0 to 5 places of decimals. Equation 28 is significant at the $p \leq 0.005$ level, since the theoretical F-value given for 40 df (10) is $F_{1,40,p=0.005} = 8.83$. The correlation coefficient demonstrated between the group contributions is in accordance with the correlation coefficients calculated between the estimated activities and between the experimental data. Because of the high intercorrelation between values a_{BN} and a_{BE}, the latter is considered below.

The set of substituents associated with values of a_{BE} is not equivalent with the set of substituents associated with values of a_{PI}. The groups Br_m, I_m, Br_p, and I_p had to be omitted from the 2-bromophenylethylamine series whereas the groups OH_m, OMe_m, OH_p, and OMe_p were discarded in the series used to extract values of a_{PI}. Ortho substituents do not influence the contributions of groups at sites m and p, thus they were not discarded. The corrected values of a_{BE} are: $H_m = -0.421$, $F_m = -0.505$, $Cl_m = -0.106$, $Me_m = -0.239$, $H_p = -0.492$, $F_p = -0.168$, $Cl_p = 1.715$, and $Me_p = 0.835$. The corrected values of a_{PI} are: $H_m = -0.474$, $F_m = 0.153$, $Cl_m = 0.774$, $Me_m = -0.402$, $H_p = -0.174$, $F_p = -0.024$, $Cl_p = 0.561$, and $Me_p = -0.190$. The sum of the contributions is not zero now, because incomplete samples have been considered. The following equation could be derived for the corrected values of a_{BE} and a_{PI}:

$$a_{BE} = 0.75(\pm 0.62)a_{PI} + 0.10 \quad \begin{matrix} n & s & r & F \\ 28 & 0.66 & 0.443 & 6.34 \end{matrix} \qquad (29)$$

This equation is significant at the $p \leq 0.05$ level ($F_{1,26,p=0.05} = 4.22$). The correction is essential, since no significant correlation could be derived between the primary values of a_{BE} and a_{PI} in Table 1 ($r^{(w)} = 0.113$). Comparison of Fujita-Ban type group contributions did not indicate any relationship between blocking potencies and inhibition of PNMT (3).

The values of a_{PI} had to be transformed again in order to compare them with values of a_{U1}, a_{U2}, and a_{PS}. The effects of removal of groups F_m, Cl_m, Me_m, F_p, Cl_p, and Me_p have been taken into account (Eq. 27), and the corrected values of a_{PI} are now: $H_m = 0.388$, $OH_m = 0.369$, $OMe_m = -0.460$, $H_p = 0.131$, $OH_p = 0.201$, and $OMe_p = -0.370$. Contributions a_{U1}, a_{U2}, and a_{PS} may be compared directly with one another and with values of a_{PI}.

There is only one compound, (±)-amphetamine, tested for its inhibitory activity on PNMT as well as for its activity as substrate of PNMT (3, 5). Nevertheless, the two actions could be compared by demonstrating relationship between the *de novo* substituent constants a_{PI} and a_{PS}:

$$a_{PS} = -0.85(\pm 0.42)a_{PI} + 0.25 \quad \begin{matrix} n & s & r & F \\ 17 & 0.20 & -0.742 & 18.41 \end{matrix} \qquad (30)$$

The equation is significant at the $p \leq 0.005$ level ($F_{1,15,p=0.005} = 10.8$).

The series tested for inhibition of norepinephrine uptake$_1$ (11) has three derivatives in common with the series tested for inhibition of PNMT (3, 8). These derivatives are (±)-methyldopamine, *p*-OH-amphetamine, and (±)-amphetamine, and the respective potencies are −2.79, −2.79, and −2.38: and 3.39, 3.12, and 2.89. The first

set of numbers indicates logarithms of relative values (11) (see Methods), whereas the second set expresses pI_{50} values (molar). The correlation coefficient calculated between these potencies is $r = -0.899$. For the extracted contributions the following result could be obtained:

$$a_{U1} = -1.58(\pm 0.26)a_{PI} + 0.47 \qquad (31)$$

n	s	r	F
36	0.20	−0.904	152.81

The equation is significant at the $p \leq 0.005$ level ($F_{1,30,p=0.005} = 9.18$).

(±)-Amphetamine is the only substance tested both for inhibition of uptake$_2$ and for inhibition of PNMT (3, 11). These actions may be compared by performing a regression analysis for the extracted sets of *de novo* substituent constants a_{U2} and a_{PI}:

$$a_{U2} = 0.72(\pm 0.56)a_{PI} - 0.13 \qquad (32)$$

n	s	r	F
18	0.37	0.559	7.26

The equation is significant at the $p \leq 0.05$ level ($F_{1,16,p=0.05} = 4.49$).

The series of PEA derivatives serving as substrates of PNMT (5) and the series investigated for inhibition of norepinephrine uptake$_1$ (11) have six substances in common: (±)-octopamine, (−)-norepinephrine, (+)-norepinephrine, (−)-epinephrine, (±)*p*-hydroxynorephedrine, and PEA. The respective values of logarithms of the relative potencies are 1.26, 1.32, 1.18, 0.60, 1.26, and 0.00 (5), and −1.93, −2.61, −1.89, −2.04, −2.41, and −2.00 (11). The correlation coefficient calculated between these values ($r = 0.390$) is not significant, even at the $p = 0.05$ level. The group-extracted contributions (Table 1) might be related to some extent:

$$a_{PS} = 1.00(\pm 0.30)a_{U1} + 0.02 \qquad (33)$$

n	s	r	F
93	0.45	0.567	43.22

The equation is significant at the $p \leq 0.005$ level ($F_{1,60,p=0.005} = 8.49$).

The regression equation derived for the contributions a_{PS} and a_{U2} is

$$a_{PS} = -0.66(\pm 0.36)a_{U2} + 0.04 \qquad (34)$$

n	s	r	F
55	0.38	−0.450	13.58

It is significant at the $p \leq 0.005$ level. The only common compound occurring in both series used to extract these values is PEA.

There are 10 compounds for which data on inhibition of uptake$_1$ and of uptake$_2$ are available. These are (±)-metanephrine, 3,4-dimethoxyphenylethylamine, (±)-epinephrine, PEA, metatyramine, (±)-amphetamine, *p*-methoxyphenylethylamine, (±)-norepinephrine, and dopamine (11). The respective potencies are −0.42, 0.26, −1.89, −2.00, −2.33, −2.39, −2.38, −1.04, −2.22, and −2.81 for uptake$_1$ (A_{U1}) and −3.41, −2.37, −2.16, −2.00, −1.90, −1.88, −1.83, −1.57, −1.48, and −1.28 for uptake$_2$ (A_{U2}). The regression of A_{U2} on A_{U1} is

$$A_{U2} = -0.40(\pm 0.36)A_{U1} + 2.68 \qquad (35)$$

n	s	r	F
10	0.47	−0.668	6.44

Equation 35 is significant at the $p \leq 0.05$ level ($F_{1,8,p=0.05} = 5.32$). Comparison of the extracted group contributions a_{U2} and a_{U1} yields a very similar slope and correlation coefficient, being significant at the $p = 0.005$ level:

$$a_{U2} = -0.48(\pm 0.16)a_{U1} - 0.03 \qquad (36)$$

n	s	r	F
62	0.25	−0.624	38.36

DISCUSSION

The results obtained are summarized in Fig. 2. With the exception of the correlation between a_{BN} and a_{BE} (Eq. 28), most correlation coefficients calculated are rather low, indicating that the respective regression equations cannot be used for predictions. However, the significance levels associated with the regression equations indicate that there is at least a moderate relationship between the various sets of *de novo* group contributions extracted from the pharmacological potencies of PEA derivatives. According to our theorem, similar moderate relationships may be expected between the activities of the corresponding series of derivatives. It is assumed hereafter that correlations between contributions indicate that correlations between pharmacological activities exist. Equations 35 and 36 support this assumption, and the results obtained for values a_{PS} and a_{U1} (Eq. 33) and for a_{U1} and a_{P1} (Eq. 31) do not contradict this assumption, although the number of pharmacological data is not enough to make the moderate correlations significant.

For the interpretation of these results the basic assumptions of the approach of Hansch and Fujita (12) are used. According to this theory, biological activities of derivatives of a common "parent" structure may be described in terms of substituent constants. It has been proved that a Hansch-Fujita approach employing linear terms of the substituent constants is equivalent to the Free-Wilson model (13, 14). Thus the group contributions listed in Table 1 could be decomposed into hydrophobic, electronic, and steric factors (5). Such decomposition was not made in this paper. We considered the *de novo* group contributions as an "idealized" Hansch-type substituent constant incorporating all chemical and physical characteristics necessary for evoking pharmacological response. This variable is the optimal one, since the Free-Wilson approach gives the upper limit of multiple correlation coefficient which may be obtained with any additive linear Hansch model (8).

Correlations demonstrated between these *de novo* substituent constants indicate that some of the physical and chemical properties of the respective series necessary to

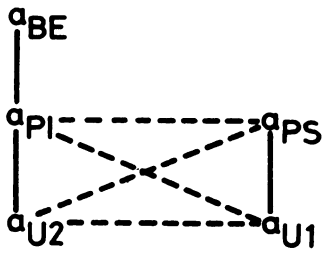

FIG. 2. *Scheme of relationships between de novo substituent constants*

——, Significant positive correlation coefficients; - - -, significant negative correlation coefficients.

evoke pharmacological response are similar. The similarity of group contributions has been observed by Zahradnik (15) in series of compounds with *nonspecific* toxic action. The relative toxicities of compounds of the type *R-X* could be described in terms of the contribution of the aliphatic substituent *R*, denoted by β:

$$\log \frac{\tau_i}{\tau_{Et}} = \beta \qquad (37)$$

where τ_i denotes the molar concentration of derivative *i* necessary to ellicit a definite biological response and τ_{Et} is a similar quantity corresponding to the reference compound, usually the ethyl derivative. Although Eq. 37 can be used for only a definite reference system, it can be extended to different systems by replacing β in Eq. 37 with $\alpha\beta$, α denoting a constant characteristic for the system.

The symbol *X* refers to various functional groups treated by Zahradnik (e.g., —OH, —COO^-, —NH_3^+, —$NHCSS^-$). Formally, Zahradnik's model is identical with the Fujita-Ban approach (5), since both express the relative potencies in terms of group contributions; the latter one is used for multiple substitution sites. In Zahradnik's model the group contributions obtained for the different systems (Eq. 37) are linearly related if nonspecific actions are considered, and it may be assumed that this is valid for Fujita-Ban group contributions, too. It was shown that the constants β are related to the substituent constants π (16) and were not related to Hammett constants (17). In analogy with Zahradnik's results it is assumed that the correlations demonstrated between *de novo* group contributions in this paper are due to similar nonspecific processes occurring during drug-receptor interaction. Many correlations are moderate, and the main part of the drug-receptor interactions must be specific. Nothing can be said about the mechanism of drug-receptor interaction at the molecular level. In this paper a process termed "nonspecific" means that it may take place in at least two types of receptors (18).

The pharmacological activities considered in this paper, with one exception (substrate activity against PNMT), are inhibitory potencies. It has been demonstrated by Timmermans *et al.* (2) that "... the binding of affinity *in vitro* of antagonists for *alpha*$_1$-adrenoceptors corresponds with their functional antagonism of *alpha*$_1$-adrenoceptors *in vivo* and *in vitro*." It is assumed that a similar correspondence is valid for the inhibitors considered in this paper, although transport processes may alter the picture.

The correlation coefficient is negative between group contributions a_{PS} and a_{PI}, indicating that factors favorable to one action are unfavorable to the second (Eq. 30). Similarly, the correlation coefficient between a_{PS} and a_{U2} (Fig. 2) is also negative (Eq. 34), and in accordance with this result the correlation coefficient between a_{PI} and a_{U2} is positive (Eq. 31).

The group contributions a_{U1} extracted from inhibitory activities of PEA derivatives against norepinephrine uptake$_1$ deserve special interest, since the correlation coefficients calculated between them and other sets of *de novo* substituent constants related to inhibitory potencies (a_{U2} and a_{PI}) are negative (Eqs. 31 and 36). The correlation coefficient demonstrated between variables a_{U1} and a_{PS} is again positive (Eq. 33). This "anomalous" feature of uptake$_1$ is not yet clear, although Burgen and Iversen (19) found that structural requirements for the inhibition of uptake$_1$ and uptake$_2$ must be strikingly different.

The relationship between a_{BE} and a_{PI} (Eq. 29) is more complex, since antagonism of the pressor activity of epinephrine and inhibition of enzyme PNMT are independent processes (20). At the present time we cannot explain this result, but it is possible that similar transport processes may be accounted for.

ACKNOWLEDGMENTS

The author is indebted to Prof. A. S. Koch (Budapest) for reading and correcting the manuscript. This work was partially done during the author's stay in Yugoslavia. Thanks are due to Prof. D. Hadži for kind hospitality.

REFERENCES

1. Lands, A. M., A. Arnold, J. P. McAuliff, F. P. Luduena, and T. G. Brown, Jr. Differentiation of receptor systems activated by sympathomimetic amines. *Nature (Lond).* **214**:597–598 (1967).
2. Timmermans, P. B. M. W. M., F. Karamat-Ali, H. Y. Kwa, A. M. C. Schoop, F. P. Slothorst-Grisdijk, and P. A. van Zwieten. Identical antagonist selectivity of central and peripheral *alpha*$_1$-adrenoceptors. *Mol. Pharmacol.* **20**:295–301 (1981).
3. Lukovits, I. Comparative studies on drug-receptor binding of phenylethylamine derivatives using calculated group contribution constants. *Int. J. Quant. Chem.* **20**:429–438 (1981).
4. Free, S. M., and J. M. Wilson. A mathematical contribution to structure-activity studies. *J. Med. Chem.* **7**:395–399 (1964).
5. Fujita, T., and T. Ban. Structure-activity study of phenethylamines as substrates of biosynthetic enzymes of sympathetic transmitters. *J. Med. Chem.* **14**:148–152 (1971).
6. Purcell, W. P., G. E. Bass, and J. M. Clayton. *Strategy of Drug Design: A Guide to Biological Activity.* Wiley, New York, 65–80 (1973).
7. Craig, P. N. Comparison of the Hansch and Free-Wilson approaches to structure-activity correlation. *Adv. Chem. Ser.* **114**:115–129 (1972).
8. Kubinyi, H., and H. O. Kehrahahn. Quantitative structure-activity relationships. 1. The modified Free-Wilson approach. *J. Med. Chem.* **19**:578–586 (1976).
9. Graham, J. D. P., and M. A. Karrar. Structure-action relations in N,N-dimethyl-2-halogenophenethylamines. *J. Med. Chem.* **6**:103–107 (1963).
10. Snedecor, G. W., and W. G. Cochran. *Statistical Methods.* Iowa State University Press, Ames, Iowa, 560–566 (1972).
11. Ban, T., and T. Fujita. Mathematical approach to structure-activity study of sympathomimetic amines. *J. Med. Chem.* **12**:353–356 (1969).
12. Hansch, C., and T. Fujita, δ-σ-π analysis: a method for the correlation of biological activity and chemical structure. *J. Am. Chem. Soc.* **86**:1616–1626 (1964).
13. Cammarata, A. Interrelationship of the regression models used for structure-activity analysis. *J. Med. Chem.* **15**:573–577 (1972).
14. Schaad, L. J., B. A. Hess, Jr., W. P. Purcell, A. Cammarata, R. Franke, and H. Kubinyi. Compatibility of the Free-Wilson and Hansch quantitative structure-activity relations. *J. Med. Chem.* **24**:900–901 (1981).
15. Zahradnik, R. Influence of the structure of aliphatic substituents on the magnitude of the biological effect of substances. *Arch. Int. Pharmacodyn.* **135**:311–329 (1962).
16. Kopecký, J., and K. Boček. A correlation between constants used in structure-activity relationships. *Experientia* **23**:125 (1967).
17. Zahradnik, R. Correlation of the biological activity of organic compounds by means of the linear free energy relationships. *Experientia* **18**:534–536 (1962).
18. Ariëns, E. J., and A. M. Simonis. Receptors and receptor mechanisms, in *Beta-Adrenoceptor Blocking Agents* (P. R. Saxena, and R. P. Forsyth, eds.). North-Holland Publishing Company, Amsterdam, 3 (1976).
19. Burgen, A. S. V., and L. L. Iversen. The inhibition of noradrenaline uptake by sympathomimetic amines in the rat isolated heart. *Br. J. Pharmacol.* **25**:34–49 (1965).
20. Patil, P. N., D. D. Miller, and U. Trendelenburg. Molecular geometry and adrenergic drug activity. *Pharmacol. Rev.* **26**:323–392 (1975).

Send reprint requests to: Dr. István Lukovits, Central Research Institute of Chemistry, Hungarian Academy of Sciences, H-1525 Budapest, P.O. B. 17, Hungary.