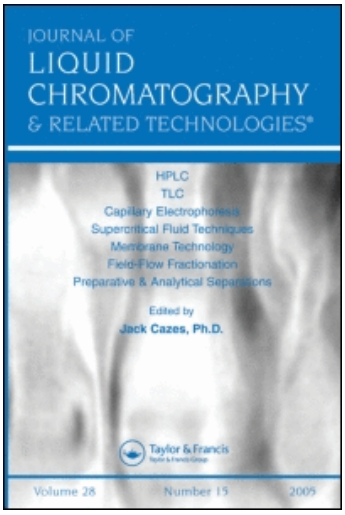



**Journal of Liquid Chromatography & Related Technologies**
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

# EVALUATION OF LIPOPHILICITY OF PIPERAZINE DERIVATIVES BY THIN LAYER CHROMATOGRAPHY AND PRINCIPAL COMPONENT ANALYSIS

C. Sârbu[a]; Barbara Malawska[b]
[a] Analytical Chemistry Department, Faculty of Chemistry and Chemical Engineering, "Babeş-Bolyai" University, Cluj-Napoca, Romania [b] Department of Pharmaceutical Chemistry, Collegium Medicum of Jagiellonian University, Kraków, Poland

Online publication date: 21 August 2000

**To cite this Article** Sârbu, C. and Malawska, Barbara(2000) 'EVALUATION OF LIPOPHILICITY OF PIPERAZINE DERIVATIVES BY THIN LAYER CHROMATOGRAPHY AND PRINCIPAL COMPONENT ANALYSIS', Journal of Liquid Chromatography & Related Technologies, 23: 14, 2143 — 2154
**To link to this Article: DOI:** 10.1081/JLC-100100477
**URL:** http://dx.doi.org/10.1081/JLC-100100477

J. LIQ. CHROM. & REL. TECHNOL., 23(14), 2143–2154 (2000)

# EVALUATION OF LIPOPHILICITY OF PIPERAZINE DERIVATIVES BY THIN LAYER CHROMATOGRAPHY AND PRINCIPAL COMPONENT ANALYSIS

C. Sârbu,[1,]* Barbara Malawska[2]

[1]Analytical Chemistry Department
Faculty of Chemistry and Chemical Engineering
"Babeş-Bolyai" University
Arany Janos 11
3400 Cluj-Napoca, Romania
[2]Department of Pharmaceutical Chemistry
Collegium Medicum of Jagiellonian University
Medyczna 9
30-688 Kraków, Poland



## ABSTRACT

The lipophilic character of a series of active anticonvulsant N-substituted amides of α-piperazine-γ-hydroxybutyric was studied. The classical $R_{Mo}$- values were measured by means of reversed thin layer chromatography using a mixture of methanol, TRIS buffer, and acetic acid as the solvent system and compared with the factors scores obtained by principal component analysis based also on the TLC retention data. The significant correlation between the $R_{Mo}$- values and slopes (the specific hydrophobic surface areas) usually indicated that this group of N-substituted amides could be considered as a homologous series of compounds independent of their structural heterogeneity. It is emphasized, once again, that this procedure can not be a rational and objective way for congeneric studies because there is always a high correlation between slope and intercept. The reliability of the factor scores values as lipophilicity indices is shown by their high correlation with the classical $R_{Mo}$-values. In addition, the "lipophilic-

ity chart" described by the first two components will have the effect of separating compounds from each other most effectively from the congeneric aspect point of view.

Finally, a better correlation was observed between scores corresponding to the first principal component and the partition coefficients (log P) of the amides, calculated by using the Prolog P module of the Pallas system.

## INTRODUCTION

In the last decades QSAR and QSRR (quantitative structure-retention relationship) models have been developed and applied for the prediction of a broad range of physicochemical properties and biological activities.[1-4] The primary goal in developing QSARs and QSRRs is to identify structurally based numerical descriptors that can be mathematically correlated to a property of interest. To obtain a significant correlation, it is crucial that appropriate descriptors be employed, whether they are theoretical, empirical, or derived from readily available experimental characteristics of the structures. Many descriptors reflect simple molecular properties and, thus, can provide insight into the physicochemical nature of the activity/property under consideration.[4-8]



Lipophilicity is one of the molecular parameters frequently used in QSAR studies, because biological activity of a molecule can generally be correlated with its ability to penetrate the different hydrophobic barriers.

Lipophilic character has been defined in many ways, the best known of which is probably the partition coefficient, P, which represents the tendency of a molecule to partition itself between organic and aqueous phases:

$$P = C_o/C_w \quad (1)$$

where $C_o$ and $C_w$ represents concentrations in the organic and aqueous phases, respectively.

Usually the n-octanol-water partition coefficient (log P) estimated by a direct equilibration method[9] or by calculation according to different mathematical models is used for inter-correlation with lipophilicity.[10]

The $R_M$- values obtained from various types of reversed-phase thin layer chromatography are the most widely used chromatographic alternatives to the shake flask method. The use of RPTLC is based on the assumed linear relationship between the molecular parameter (2) and log P.

$$R_M = \log (1/R_F - 1) \quad (2)$$

The advantages of TLC methods are the very small amounts of sample needed for the estimation and the less strict requirement of purity, because the impurities separate during the chromatographic process. They are rapid and relatively simply, low cost, and easy to perform. In addition, we have to stress the dynamic aspect of the chromatographic process and the wide choice of stationary phases and developing solvents.

The $R_M$- value (related to the molecular lipophilicity), determined by using RPTLC, generally depends linearly on the concentration of the organic component of the mobile phase:

$$R_M = R_{Mo} + bC \quad (3)$$

where $R_M$- values were calculated using equation (2).

At the same time, the relationship between the intercept $R_{Mo}$ and the slope b, in the TLC equation, (3), has also been studied by many researchers.[11-16] In each case a high correlation between the two regression parameters was observed. As a consequence, different explanations were given for slope. For example, the slope has been regarded as a characteristic of the specific hydrophobic surface area of the compound[14,15] or was explained in terms of a "displacement model".[16] However, the most important aspect of the high linear correlation was considered as a possibility of finding congeneric classes within large groups of compounds.



Recently, we addressed the question regarding whether a high linear correlation between the intercept and slope can be an objective criterion to find the real (natural) congeneric classes within the chemical compounds series. This was because we demonstrated that by using fuzzy regression[17] there always is a high correlation between slope and intercept. This very interesting and useful conclusion was strongly supported in recent papers.[18,19] Taking into account these results and other disadvantages of the regression method applied in this scope (logarithmic scale, extrapolation) we proposed the use of scores obtained applying principal component analysis as a support for the lipophilicity scale and the "lipophilicity chart" as a new concept.[20-22]

The purpose of this paper is to investigate the feasibility of the scores, obtained by principal component analysis using RPTLC retention data, as a measure of lipophilicity in correlation with partition coefficient (log P) of some anticonvulsant N-substituted amides of α-piperazine-γ-hydroxybutyric acid. In addition, the scatterplots of the scores on the plane described by the first two components appear to be very useful having the effect of separating compounds, one from each other, most effectively, obtaining in this way the "congeneric lipophilicity chart" of the series.

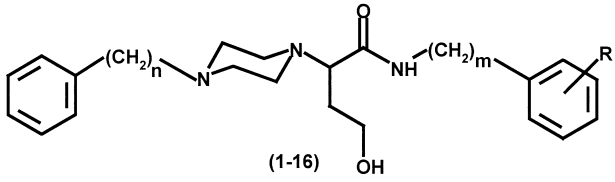
## EXPERIMENTAL

### Thin-Layer Chromatography

Chromatography was performed in a normal developing chamber on 20 cm x 20 cm precoated RP-18 $F_{254}s$ plates from Merck (Darmstadt, Germany). Mixtures of methanol, TRIS buffer (pH 7.4) and acetic acid (5.7%), with methanol content between 25-65% (v/v) in steps of 5% were used as mobile phase, as the studied compounds differed considerably in their retention.[23]

Solutions of each compound presented in Figure 1 (5 mg $mL^{-1}$) were prepared in methanol and 10 μL of each solution was spotted separately to the origin of the plates. After being developed, the dried plates were observed under UV lamp. The $R_f$ - values were expressed as the mean of four determinations.






| series **A** | series **B** | |
|---|---|---|
| n = 0, m =1, m* = 2 | n = 1, m = 1, m* = 2 | R |
| **1** | **9** | H |
| **2** | **10** | H* |
| **3** | **11** | 2-Cl |
| **4** | **12** | 4-Cl |
| **5** | **13** | 4-F |
| **6** | **14** | 4-$CH_3$ |
| **7** | **15** | 4-$OCH_3$ |
| **8** | **16** | 3,4-$(OCH_3)_2$ |

a) For compounds **2** and **10** $m^* = 2$

**Figure 1**. Molecular structure of the N-substituted amides of α-piperazine-γ-hydroxybutyric acid.

Calculations of log P values for the investigated compounds were performed by means of a computer program using Prolog P module of the PALLAS system.[24]

## Principal Component Analysis

Principal component analysis[20-22,25] was performed on the retention data matrix by the use of a computer program.[26] It displays objects (N-substituted amides of α-piperazine-γ-hydroxybutyric acid) in a reduced space by finding a direction (first principal component) that best preserves the scatter of the observations ($R_F$- values) in the multidimensional space described by the solvent systems. PCA gives both coordinates (scores) of the studied compounds and the loading of variables (solvents) on the principal components.

The results obtained from the initial chromatographic data using covariance matrix (without autoscaling) can be usually presented in three panels, although typically there are only two. The first panel shows the table of data statistics; the second is the table of components and the third panel displays the eigenvectors associated with each of the components.

Table 1 lists the eigenvalues of the covariance matrix, ordered from largest to smallest, the third column of this table shows the difference between each eigenvalues and the next smaller eigenvalue and the fourth column shows the proportion. These results suggest a significant three component model, which explained 99.29% of the total variance (information), considering only the



**Table 1**

**The Eigenvalues and the Ratios of the Variance Explained by the Six Components Using Covariance Matrix**

| Component | Eigenvalue | Difference | Proportion (%) | Cumulative (%) |
|---|---|---|---|---|
| 1 | 0.029535 | 0.028640 | 95.01 | 95.05 |
| 2 | 0.000895 | 0.000472 | 2.88 | 97.93 |
| 3 | 0.000423 | 0.000296 | 1.36 | 99.29 |
| 4 | 0.000127 | 0.000058 | 0.40 | 99.69 |
| 5 | 0.000069 | 0.000032 | 0.21 | 99.90 |
| 6 | 0.000037 | | 0.11 | 100.00 |

eigenvalues higher than one. The first component explains 95.01% of the total variance, the second 2.88% and the third only 1.36%; the subsequent eigenvalues are just sampling noise.

When the significance of the component model retained was tested applying the Bartlett's statistics,[26] testing the hypothesis that (p-k) eigenvalues in variance-covariance matrix are equal, a model with three components was also selected.

## RESULTS AND DISCUSSION

The results of regression analysis using equation (3) are compiled in Table 2. The statistics obtained (see also Table 2) illustrate that the linear equation fits, in a very good way, the experimental data, the linear model explaining approximately 95-99% of the total variance (see $R^2$ values) in the majority of cases. As usual good correlation has also been found between the $R_{Mo}$- and b-values of equation (3) as is shown by the following linear relationship:



**Table 2**

**Regression Data and Scores on the First Three Principal Components for the N-Substituted Amides of α-Piperazine-γ-Hydroxybutyric Acid Studied in this Paper**

| Compound | $R_{MO}$ | b | r | $R^2$ | PCl | PC2 | PC3 | log P |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.896 | -0.0342 | 0.9929 | 0.9859 | 0.540 | 0.197 | 0.062 | 1.44 |
| 2 | 2.089 | -0.0355 | 0.9865 | 0.9732 | 0.390 | 0.183 | 0.047 | 1.89 |
| 3 | 2.257 | -0.361 | 0.9866 | 0.9734 | 0.358 | 0.179 | 0.048 | 2.15 |
| 4 | 2.509 | -0.396 | 0.9915 | 0.9831 | 0.286 | 0.151 | 0.042 | 2.16 |
| 5 | 1.846 | -0.0308 | 0.9798 | 0.9600 | 0.499 | 0.212 | 0.064 | 1.59 |
| 6 | 1.918 | -0.0324 | 0.9779 | 0.9563 | 0.494 | 0.219 | 0.063 | 1.87 |
| 7 | 1.857 | -0.0313 | 0.9789 | 0.9582 | 0.506 | 0.212 | 0.064 | 1.39 |
| 8 | 1.613 | -0.0281 | 0.9741 | 0.9489 | 0.621 | 0.234 | 0.056 | 1.37 |
| 9 | 1.005 | -0.0205 | 0.9691 | 0.9392 | 0.857 | 0.130 | 0.012 | 1.16 |
| 10 | 1.225 | -0.0210 | 0.9584 | 0.9185 | 0.672 | 0.129 | 0.012 | 1.61 |
| 11 | 1.497 | -0.0252 | 0.9766 | 0.9537 | 0.611 | 0.211 | 0.055 | 1.87 |
| 12 | 1.747 | -0.0294 | 0.9830 | 0.9663 | 0.533 | 0.208 | 0.046 | 1.88 |
| 13 | 1.209 | -0.0222 | 0.9826 | 0.9655 | 0.780 | 0.200 | 0.057 | 1.31 |
| 14 | 1.485 | -0.0257 | 0.9807 | 0.9618 | 0.627 | 0.199 | 0.069 | 1.59 |
| 15 | 1.210 | -0.0224 | 0.9815 | 0.9633 | 0.786 | 0.201 | 0.072 | 1.11 |
| 16 | 1.085 | -0.0216 | 0.9759 | 0.9524 | 0.907 | 0.212 | 0.047 | 1.06 |

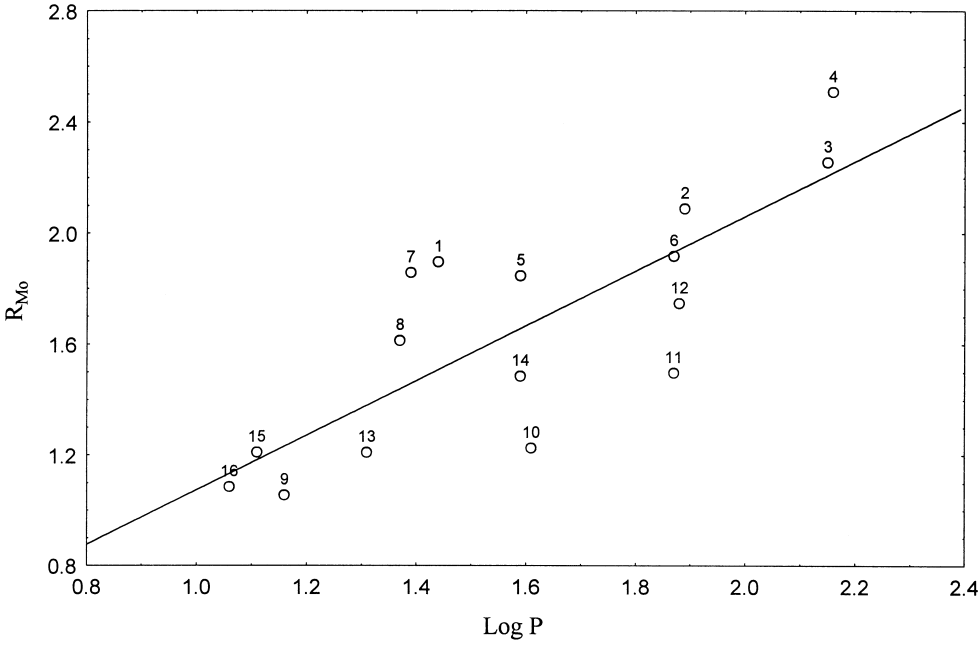



**Figure 2**. Relationship between $R_{Mo}$ and log P ($R_{Mo}$ = 0.085 + 0.987logP; r = 0.8018).

$$R_{Mo} = 0.385 - 71.992\ b;\ r = 0.9882. \quad (4)$$

This finding indicates that the intercept, $R_{Mo}$, (lipophilicity) and slope, b, (specific hydrophobic surface area) for the majority of these compounds are highly correlated and, in that case, they might form a homologous series of compounds as has been suggested by some authors.[11-15] Moreover, a high correlation was obtained between $R_{Mo}$- values and the scores of the same anticonvulsant compounds on the first principal component, as is described by the linear equation (5).

$$R_{Mo} = 3.061 - 2.375PC1;\ r = -0.9731 \quad (5)$$

It is very interesting to stress, also, that the correlation between scores and the log P values (0.8672) is higher than the correlation between $R_{Mo}$ values and log P values (0.8019) (see Figure 2 and 3). On the basis of these findings and as can been seen from data provided in Table 2, the scores on the first prin-

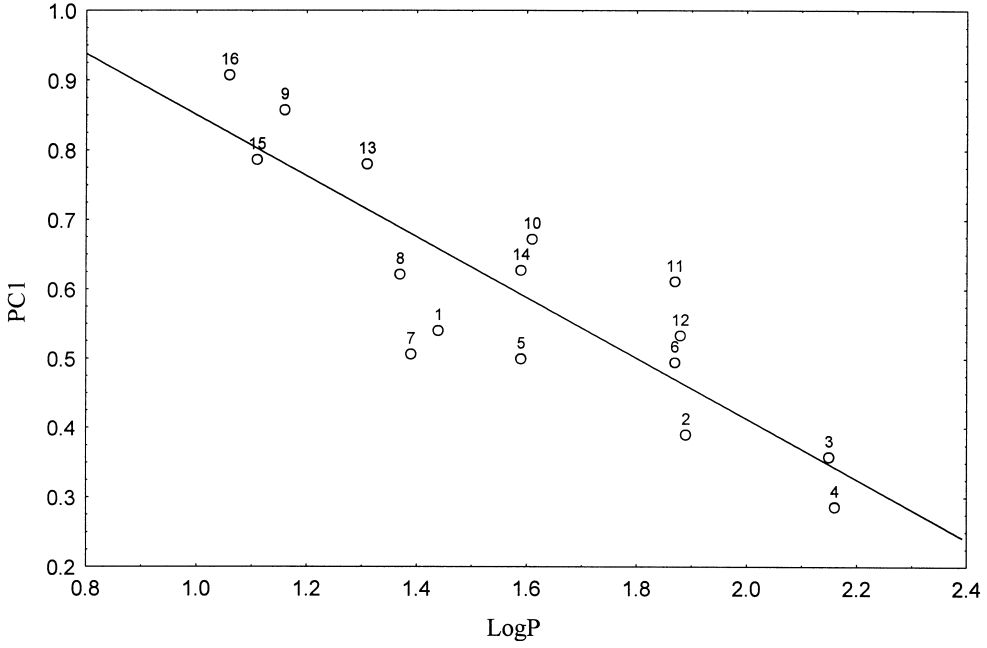



**Figure 3**. Relationship between PC1 and log P ( PC1 = 1.288 – 0.438LogP; r = -0.8672).



cipal component can replace, efficiently, the $R_{Mo}$ values in the estimation experiments of the lipophilicity of these compounds.

In addition, as is shown in Figure 4, scores plots are very useful as a display tool for examining the relationships between compounds, looking for trends, groupings, or outliers. Hence, graphing scores onto the plane described by PC1 and PC2 we obtain what could be named the "congeneric lipophilicity chart". It appears, clearly, that the compounds studied in this paper form practically three different linear congeneric classes, in very good agreement with the selected eigenvectors when the Bartlett test was applied, and also the consideration[23] of two main groups (1-8) and (9-16), respectively. The position of outliers of the compounds (9) and (10) to the second linear cluster and also of the compound (1) to the first linear cluster reveals its relatively different chemical structure compared with the rest of the drugs.

It is easy to observe, from Figure 1, that the structural difference between the compound (9) and (10)—a methylene group—is the same with that between

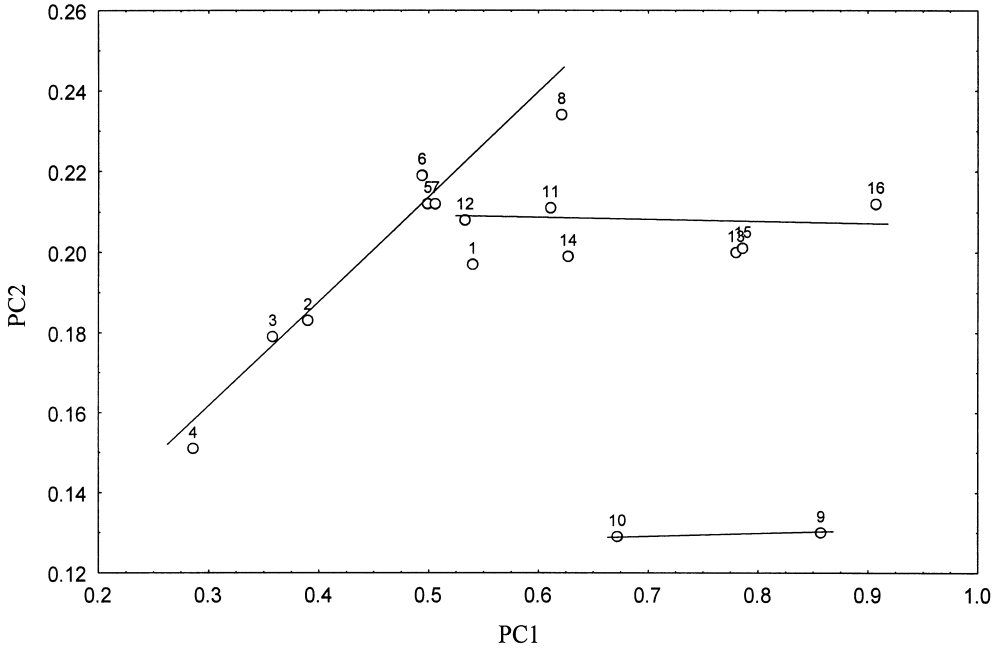



**Figure 4**. Congeneric lipophilicity chart using scores corresponding to PC1 and PC2.



compounds (1) and (2). However, a more interesting point arises from a comparison of the compounds (5), (7) and (13), (15), respectively. The compounds of each pair seem to be equivalent from the lipophilicity point of view estimated by RPTLC method. The relatively similar hydrogen-bond basicity of the fluorine atom and $OCH_3$ group could be the most probable explanation for this phenomenon. The three dimensional graphic display of scores described by PC1, PC2, and PC3, respectively, shown in Figure 5 also supports the same conclusions. The position of each compound within the graphs is also in good agreement with the effect of substituent in the phenyl ring and the presence of aliphatic methylene group.

The presence of the methyl and chloro substituents in the phenyl ring increased the lipophilicity of compounds (3), (4), and (6) in series A and compounds (11), (12) and (14) in series B. Different lipophilicities were obtained, also, for isomers with *ortho* chloro (3), (11) and *para* chloro (4), (12) substituents in the phenyl ring.

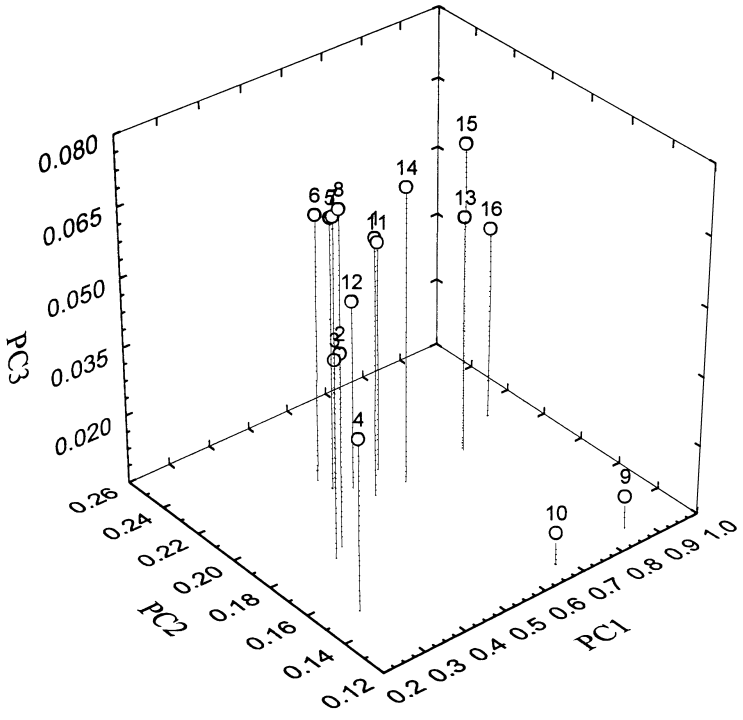



**Figure 5**. Tridimensional graphically display of scores corresponding to PC1, PC2, and PC3, respectively.



## CONCLUSIONS

The lipophilic character of some bioactive anticonvulsant N-substituted amides of α-piperazine-γ-hydroxybutyric acid was studied by means of reversed thin layer chromatography using a mixture of methanol, TRIS buffer, and acetic acid as the solvent system.

The significant correlation between the $R_{Mo}$- values and b-slopes (specific hydrophobic surface areas) indicate that this group of N-substituted amides could be considered as a homologous series of compounds independent of their structural heterogeneity as far as it was considered. The reliability of the factor scores values as lipophilicity indexes is shown by their high correlation with the classical $R_{Mo}$-values.

In addition, the "lipophilicity chart" described by the first two components had the effect of separating compounds from each other most effectively from the congeneric aspect point of view. Finally, a better correlation was observed between scores corresponding to the first principal component and the partition coefficients (log P) of the amides calculated by using the Prolog P module of the Pallas system.

## REFERENCES

1. S. D. Brown, T. B. Blank, S. T. Sum, L. G. Weyer, Anal. Chem., **66**, 337R-340R (1994).

2. S. D. Brown, S. T. Sum, F. Despagne, Anal. Chem., **68**, 42R-44R (1996).

3. B. K. Lavine, Anal. Chem., **70**, 216R-217R (1998).

4. Q. S. Wang, L. Zhang, J. Liq. Chrom. & Rel. Technol., **22**, 1-14 (1999).

5. M. Karelson, V. S. Lobanov, A. R. Katritzky, Chem. Rev., **96**, 1027-1043 (1996).

6. R. F. Rekker, R. Mannhold, **Calculation of Drug Lipophilicity**, Weinheim, 1992.

7. M. V. Diudea, O. Ivanciuc, **Molecular Topology**, Comprex Publishing House, Cluj, 1995.

8. A. Chiriac, D. Ciubotariu, Z. Simon, **QSAR-Quantitative Relationships**, Mirton Publishing House, Timişoara, 1996.

9. C. Hansch, S. M. Anderson, J. Org. Chem., **32**, 2583-2586 (1967).

10. A. Leo, C. Hansch, D. Elkins, Chem. Rev., **71**, 525-616 (1971).

11. C. Horváth, W. Melander, I. Molnar, J. Chromatogr., **125**, 129-156 (1976).

12. T. Cserháti, Chromatographia, **18**, 18-20 (1984).

13. T. Cserháti, Chromatographia, **18**, 318-322 (1984).

14. T. Cserháti, J. Biochem. Biophys. Methods, **27**, 133-142 (1993).

15. G. L. Biagi, A. M. Barbaro, A. Sapone, M. Recanatini, J. Chromatogr. A, **662**, 341-361 (1994).

16. F. Murakami, J. Chromatogr., **178**, 393-399 (1979).

17. H. Pop, C. Sârbu, Anal. Chem., **68**, 771-778 (1996).

18. A. G. Gonzáles, M. A. Herrador, A. G. Asuero, Talanta, **48**, 729-736 (1999).

19. P. D. Lark, Anal. Chem., **26**, 1712-1725 (1954).

20. S. Todor, C. Sârbu, Rev. Chim. (Bucharest), **49**, 177-182 (1998).

21. C. Sârbu, S. Todor, J. Planar Chromatogr., **11**, 123-126 (1998).

22. C. Sârbu, S. Todor, J. Chromatogr., **822**, 263-269 (1998).

23. B. Malawska, J. Planar Chromatogr., **11**, 137-140 (1998).

24. Pallas; (version for Windows 1.2) distributed by Compu Drug Chemistry Ltd. (1995).

25. S. Wold, K. Esbensen, P. Geladi, Chemom. Intell. Lab. Syst., **2**, 37-52 (1987).

26. A. Mackiewicz, W. Ratajczak, Comput. Geosci., **19**, 303-342 (1993).



Received September 10, 1999 Author's Revisions March 23, 2000
Accepted December 27, 1999 Manuscript 5183