2D and 3D QSAR Analysis of Some Valproic Acid Metabolites and Analogues as Anticonvulsant Agents

Tatiana Netzeva,[1,2] Irini Doytchinova,[1] and Roumiana Natcheva[1]

Purpose. To investigate the structural features, responsible for the variations in anticonvulsant activity of a series of twenty six valproic acid (VPA) metabolites and analogues.

Methods. Different approaches for quantitative structure—activity relationship analysis (QSAR) as conventional 2D QSAR analysis and comparative molecular field analysis (3D QSAR) were used. The 2D QSAR was performed with more than twenty structure descriptors as the partition and distribution coefficients, topological, geometrical and electronic descriptors, and indicator variables. The electronic descriptors were calculated for the energetically most stable conformers. For the need of 3D QSAR steric and electrostatic potential maps were generated. Partial least squares (PLS) analysis has been carried out for the statistical evaluation of the models and weighted least squares (WLS) analysis was used for the visualization of the results.

Results. It was established that the two approaches—2D and 3D QSAR, prove the importance of the lipophilicity of the compounds for anticonvulsant activity. The results from both the approaches suggest that a substitution at α-position is essential for a higher activity.

Conclusions. 3D QSAR is useful for describing the steric and electrostatic fields, important for the activity. For predicting the activity of new compounds 2D QSAR tools were proposed.

KEY WORDS: valproic acid; anticonvulsant activity; lipophilicity; quantum mechanics; quantitative structure—activity relationship analysis (QSAR).

INTRODUCTION

Valproic acid (2-propylpentanoic acid, VPA) and its prodrug, primary amide (valpromide, VPD) are well known and widely used antiepileptic drugs (1,2). Their effectiveness is based on several mechanisms of action: an increase of GABA levels in the brain by glutamate decarboxylase activation and GABA transaminase inhibition, an increase of postsynaptic GABA responses, direct membrane effects on the neurons, as well as blockade of the Na^+ and Ca^{++} channels (3,4). The hypotheses recently proposed suggest that the action of valproates is not receptor specific. However, there is some literature data about successful structure-activity and structure-toxicity relationships of valproate (5,6) and valpromide (7,8) derivatives.

The molecular flexibility of VPA derivatives complicates the QSAR of these compounds. The choice of the conformer, for which electronic descriptors are calculated is an important step in QSAR analysis. In this study we accept the energetically most stable ones due to the absence of structural data in the literature on the drug-receptor complex.

The aim of this study is to investigate the anticonvulsant activity of the simple but highly flexible VPA derivatives by means of conventional 2D QSAR analysis and comparative molecular field analysis.

MATERIALS AND METHODS

Structure and Activity

The structure and anticonvulsant activity of 26 VPA derivatives is taken from the literature (9) and shown in Table I. Anticonvulsant activity is determined as threshold for maximal electroconvulsions. All compounds have been injected intraperitoneally in groups of twenty mice at doses of 75–200 mg/kg and the threshold was determined 30 min after injection. The electroshock was delivered via eye electrodes for 0.2 sec. The extension of the hind limbs was taken as the endpoint (predefined effect). The threshold was determined in volts by the "up and down" method of Kimball, Burnett and Doherty (10) and calculated as the voltage inducing the extensor phase in 50% of the mice (EV_{50}). On each experimental day a group of mice has been treated with VPA and its activity has been regarded as a control activity. Anticonvulsant potency relative to VPA (RA), on molar base have been calculated as a ratio between the increase in voltage above the control, caused by the test drug, and increase in voltage induced by VPA. This ratio then has been multiplied by the ratio between the molar dose of VPA and the molar dose of test compound. In our QSAR study RA is used as log RA.

Molecular modeling, conformational analysis, quantum mechanical optimization and electronic descriptor calculation, as well as steric and electrostatic field generation for 3D QSAR and statistical evaluation of the models were performed with the Chem-X computer program (11). Conformational analysis of VPA and its derivatives was carried out by use of molecular mechanics (MM2 force field) with a 0.0001 energy gradient convergence criterion. The method of Gasteiger (12) was used for charge computation. Conformational analysis was performed as systematic search using the default torsion increment of 30° (12 incremental steps) for all rotatable bonds. The conformers of the global minimum underwent AM1 (13) optimization with keyword PRECISE. AM1 is a part of MOPAC (Ver. 6.49), implemented in Chem-X.

2D QSAR Analysis

For the needs of 2D QSAR analysis various structure descriptors were generated. The n-octanol-water partition coefficients (log P) were calculated by the unique additive-constitutive ACD/Log P algorithm, which is sensitive to unsaturations. It is based on well characterized log P contributions of separate atoms, structural fragments, and intramolecular interactions between different fragments. These contributions have been derived from a database of 3,600 structures for which one or more experimental log P values have been reported in the literature. A correction for ionization distribution coefficient for pH = 7.4 (log D) was done. Log P and pK_a values were obtained with ACD/Labs computer program (14).

[1] Department of Chemistry, Faculty of Pharmacy, Medical University-Sofia, Sofia 1000, Bulgaria.

[2] To whom correspondence should be addressed. (e-mail: tnetzeva@mbox.pharmfac.acad.bg)

Table I. Compounds considered in this study and their relative anticonvulsant activities

$$R_2-\overset{\overset{\displaystyle R_1}{|}}{\underset{\underset{\displaystyle R_3}{|}}{C}}_4-\overset{\overset{\displaystyle O_1}{\|}}{C}_3-OH_2$$

No	Compound	RA
1	$R_1 = H$, $R_2 = R_3 = CH_2CH_2CH_3$ (VPA)	1.00
	VPA-metabolites	
2	$R_1 = -H$, $R_2 = -CH_2CH_2CH_3$, $R_3 = -CH(OH)CH_2CH_3$	0.13
3	$R_1 = -H$, $R_2 = -CH_2CH_2CH_3$, $R_3 = -CH_2CH(OH)CH_3$	0.26
4	$R_1 = -H$, $R_2 = -CH_2CH_2CH_3$, $R_3 = -C(O)CH_2CH_3$	0.17
5	$R_1 = -H$, $R_2 = -CH_2CH_2CH_3$, $R_3 = -CH_2CH=CH_2$	0.87
6	$R_1 = -H$, $R_2 = -CH_2CH_2CH_3$, $R_3 = -CH=CH_2CH_3$	0.54
7	$R_1 = -H$, $R_2 = -CH_2CH_2CH_3$, $R_3 = =CHCH_2CH_3$ (cis-isomer)	0.56
8	$R_1 = -H$, $R_2 = -CH_2CH_2CH_3$, $R_3 = =CHCH_2CH_3$ (trans-isomer)	0.84
9	$R_1 = -H$, $R_2 = -CH_2CH=CH_2$, $R_3 = =CHCH_2CH_3$	0.46
10	$R_1 = -H$, $R_2 = R_3 = -CH_2CH=CH_2$	0.29
	VPA-analogues	
11	$R_1 = -H$, $R_2 = R_3 = -CH_2CH_3$	0.22
12	$R_1 = -H$, $R_2 = -CH_2CH_2CH_3$, $R_3 = -CH_3$	0.18
13	$R_1 = -H$, $R_2 = -CH_2CH_2CH_3$, $R_3 = -CH_2CH_3$	0.25
14	$R_1 = -H$, $R_2 = -CH_2CH_2CH_2CH_3$, $R_3 = -CH_3$	0.40
15	$R_1 = -H$, $R_2 = -CH_2CH_2CH_2CH_3$, $R_3 = -CH_2CH_2CH_3$	0.96
16	$R_1 = -H$, $R_2 = R_3 = -CH_2CH_2CH_2CH_3$	1.60
17	$R_1 = -H$, $R_2 = R_3 = -CH_2CH_2CH_2CH_2CH_3$	1.40
18	$R_1 = -H$, $R_2 = R_3 = -CH_2CH_2CH_2CH_2CH_2CH_3$	3.00
19	$CH_3-CH_2-CH_2$ $CH-CH_2-COOH$ $CH_3-CH_2-CH_2$	3.70
20	$R_1 = -H$, $R_2 = -CH_2CH_2CH_3$, $R_3 = -CH(CH_3)CH_2CH_3$	0.82
21	$R_1 = -H$, $R_2 = R_3 = -CH(CH_3)_2$	0.58
22	$R_1 = R_2 = -CH_3$, $R_3 = -CH_2CH_2CH_3$	1.60
23	$R_1 = -CH_3$, $R_2 = -CH_2CH_3$, $R_3 = -CH_2CH_2CH_3$	2.70
24	(cyclohexyl)—COO	0.11
25	(cycloheptyl)—COO	0.21
26	(cyclohexyl)—COOH, CH	0.88

A wide range of topological indices was used for describing the chemical structure: Wiener index (WIN), the connectivity index of Randic (RAN), Balaban J index (BJ), Hosoya Z index (HZ), and Kier and Hall index (KH). Calculation of all the topological indices was carried out with OASIS program for Windows (15).

The molecular volume (MV) and molecular surface area (MSA), as well as the electrostatic potential (EP) on MSA were calculated within Chem-X computer program.

Electronic descriptors as dipole moment (DIP), volume polarisability (VP), energy of the highest occupied and the lowest unoccupied molecular orbitals (HOMO and LUMO) and the partial charges of separate atoms (QA_X) were computed by AM1 routines.

The quantitative linear models were obtained by a stepwise procedure. Statistical significance of the regression equations was tested on the basis of conventional regression coefficient (R^2), Fisher's ratio (F), and the standard error of estimate (s). The standard errors of the coefficients of independent variables are given in brackets. The best equation generated is tested by cross-validation (leave-one-out) procedure (16).

3D QSAR Analysis

In the 3D QSAR analysis (17) probably the most crucial step is molecular alignment. For this purpose rigid and flexible fitting of the studied compounds was used and the results were compared. The flexyfit method, used by us, is particularly

appropriate for structures with a large number of degrees of freedom, such as VPA derivatives. Each structure is modified by the rotation of non-rigid bonds to try to fit it to the query, without taking into account internal energy changes. While in the rigid fitting one molecule is presented by only one conformer, in the flexible fitting a large number of conformers for one and the same molecule were generated. In our field comparison with flexible fitted structures only one conformer for a molecule was used—the one that fits the model best.

The molecules of the studied series contain a common invariant structural fragment (the atoms from the carboxylic group and the vicinal carbon atom—O_1, O_2, C_3, and C_4) which was taken as the basic query for molecular alignment. One interesting question is in what manner the query influences the statistical characteristics of the 3D QSAR models. Since the query should be sufficiently general so that all the structures can be accepted, we started our investigation with the minimum number of atoms (four) and enlarged the query to six atoms.

The steric (S) and electrostatic (E) fields around the molecules were calculated on a $10 \times 10 \times 6$ A grid with a grid spacing of 1.0 A and the proton probe with $+1.0$ charge. A value of 1.0 was assigned to the dielecric constant. PLS analysis (18) was used to derive a linear relationship between the anticonvulsant activity and independent variables. Cross-validation (leave-one-out procedure) was applied to measure the predictive capabilities of the models. Automatic calculation of sample numbers when the structures are omitted only once was allowed. The statistical significance of the models was estimated by the squared conventional (R^2_{conv}) and predictive (R^2_{pred}) correlation coefficients, where:

$$R^2_{pred} = 1 - \frac{PRESS}{SS}$$

Here, PRESS is the sum of the squared differences between the observed and predicted activity and SS is the sum of squared deviations of each activity value from their mean.

RESULTS

Molecular Geometry

The conformational analysis shows a free rotation around all σ-bonds in the studied molecules without energetical barriers higher than 20 kcal/mol. The 2D plot of O_1-C_3-C_4-C_5 torsion angle against energy ($E = E_{(\varphi)} - E_{min}$) in kcal/mol for VPA molecule is shown in Fig. 1. The torsion angle values obtained for the most stable conformers after MM2 and AM1 calculations are presented in Table II.

The free rotation suggests that the differences in anticonvulsant activity between the compounds cannot be explained with any distinct molecular conformation since all conformers are easily interchangeable, and an equilibrium mixture should be present. For further calculations the energetically most stable conformers were used.

2D QSAR

The calculated values of log P, pK_a and log D are shown in Table III. The experimental log P values are taken from the literature (19). The calculated log P data correlate well with measured log P ($n = 6$, $R^2 = 0.990$, $F = 403.13$, $s = 0.094$).
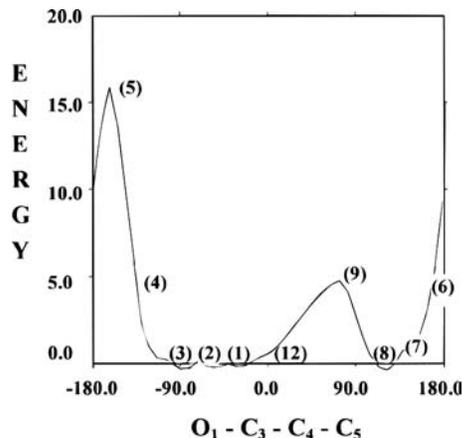
Fig. 1. A 2D plot of O_1-C_3-C_4-C_5 angle against energy for VPA molecule.

There is a strong correlation ($n = 26$, $R^2 = 0.955$, $F = 510.90$, $s = 0.239$) between log P and log D for all the compounds. Compound 4 remains an outlier in this regression with studentised residual equal to 8.06, probably because of its ability to form an internal hydrogen bond (a H-bond length of 1.784 A was established after MM2 optimization and 2.126 A after AM1 optimization). The model with log D is listed below:

$$\log RA = 0.250(0.048) \log D - 0.217(0.059) \quad (1)$$

$$n = 26, R^2 = 0.528, F = 26.80, s = 0.301$$

An attempt to improve the model with log D by including steric and topological indices was done. The main difficulty in this case was the high rate of intercorrelation between these indices and lipophilicity parameters. It was established that only Balaban J index (the values are presented in Table III) does not correlate with log P and log D and the equation with log D obtained is:

$$\log RA = 0.249(0.044) \log D + 0.301(0.121) BJ - 1.253(0.418)$$

$$n = 26, R^2 = 0.628, F = 19.44, s = 0.272 \quad (2)$$

The analysis of the residuals, calculated according to Eq. (2) shows that the compounds 22, 23 and 26 with a quaternary carbon atom next to the carboxylic group have large residuals. To discriminate these compounds an indicator variable Me (Table III, Me = 0 for compounds with tertiary carbon at α-position and Me = 1 for those with quaternary carbon at the same position) was included in the equation above:

Table II. Torsion angles obtained by MM2 and AM1 optimization of VPA molecule
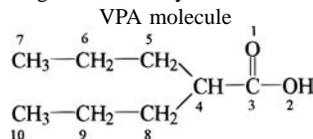
Torsion	MM2 (deg)	AM1 (deg)
O_1-C_3-C_4-C_5	−41.8	−38.2
C_3-C_4-C_5-C_6	−75.4	−72.1
C_4-C_5-C_6-C_7	−70.6	−78.6
C_3-C_4-C_8-C_9	−82.6	−67.8
C_4-C_8-C_9-C_{10}	166.7	171.5

Table III. The values of some calculated descriptors

	log P					
	calcd	expt[a]	pKa	log D	BJ	Me
1	2.72	2.75	4.82	0.14	3.50	0
2	0.80		4.43	−2.17	3.87	0
3	1.06		4.65	−1.69	3.69	0
4	1.42		3.52	−2.46	3.87	0
5	2.40		4.70	−0.30	3.50	0
6	2.48		4.51	−0.41	3.50	0
7	2.95		5.00	0.55	3.50	0
8	2.95		5.00	0.55	3.50	0
9	1.72		4.82	−0.86	3.50	0
10	2.22		4.49	−0.69	3.50	0
11	1.66	1.68	4.80	−0.94	3.35	0
12	1.66	1.80	4.82	−0.92	3.17	0
13	2.19		4.82	−0.39	3.41	0
14	2.72	2.64	4.82	0.14	3.40	0
15	3.25	3.01	4.82	0.67	3.51	0
16	3.78	3.55	4.82	1.20	3.55	0
17	4.84		4.90	2.34	3.55	0
18	5.91		4.90	3.41	3.54	0
19	3.25		4.80	0.65	3.50	0
20	3.60		4.83	1.03	3.85	0
21	2.35		4.74	−0.31	3.98	0
22	2.01		4.86	−0.53	3.70	1
23	2.54		4.86	−0.01	3.94	1
24	1.77		4.91	−0.72	2.23	0
25	2.34		4.82	−0.24	2.28	0
26	2.32		4.95	−0.13	2.48	1

[a] The experimental log P values are taken from ref. 19.

$$\log RA = 0.255(0.033)\log D + 0.321(0.090)\,BJ$$
$$+\ 0.550(0.125)\,Me - 1.385(0.313)$$
$$n = 26,\ R^2 = 0.803,\ R^2_{cv} = 0.762,\ F = 29.80,\ s = 0.203 \quad (3)$$

The plot of observed against predicted by Eq. (3) log RA is shown on Fig. 2.

The electronic parameters did not improve the 2D QSAR models.
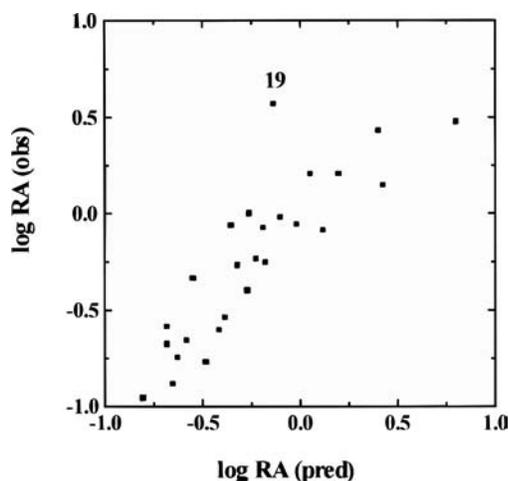
Fig. 2. Plot of observed against predicted by Equation (3) log RA.

3D QSAR

The influence of different alignment techniques (rigid and flexible fitting) and queries on the quality of the 3D QSAR models was investigated. The results are shown in Table IV. Generally, the models with electrostatic fields are better than those with steric ones. The combination between the steric and electrostatic fields results in models, which have medial statistical characteristics to these with steric and electrostatic fields alone. The best model is with log D and electrostatic field values, obtained by use of flexible fit technique and six-atom query.

Weighted least squares (WLS) analysis (20) was used for providing a visual representation of the regions of space around a VPA molecule that correlate well with the anticonvulsant activity. The coefficient maps, which store the correlation coefficients of each remaining point, are shown in Fig. 3.

The localization of the electrostatic potential extremum values was investigated and compared with WLS coefficient map. The localization of the first two minimum and maximum values is shown in Fig. 4 over the structure of VPA.

DISCUSSION

2D QSAR

It was established that the lipophilicity of VPA derivatives is important for anticonvulsant activity. It is higher for the more lipophilic compounds. The positive value of the coefficient in front of BJ index indicates that elongation and branching in

Table IV. Statistical characteristics of the 3D QSAR models obtained for different alignment techniques and queries (n = 26).

Query	Descriptors	R^2_{pred}	R^2_{conv}	R^2_{pred}	R^2_{conv}
		Rigid fitting		Flexible fitting	
	S	−0.564	0.495	−0.793	0.599
	E	−0.166	0.740	−0.171	0.628
	S + E	−0.190	0.602	−0.320	0.612
	log D + S	−0.486	0.507	−0.458	0.623
	log D + E	−0.108	0.752	−0.107	0.657
	Log D + S + E	−0.175	0.608	−0.308	0.634
	S	−0.765	0.507	0.253	0.803
	E	−0.404	0.625	−0.063	0.759
	E + S	−0.458	0.543	0.101	0.760
	log D + S	−0.673	0.529	0.304	0.855
	log D + E	−0.284	0.655	0.011	0.778
	Log D + S + E	−0.407	0.569	0.107	0.814
	S	−0.476	0.630	0.103	0.810
	E	−0.755	0.644	0.709	0.937
	E + S	−0.564	0.635	0.328	0.859
	log D + S	−0.414	0.642	0.236	0.843
	log D + E	−0.649	0.664	0.718	0.939
	Log D + S + E	−0.486	0.650	0.485	0.879
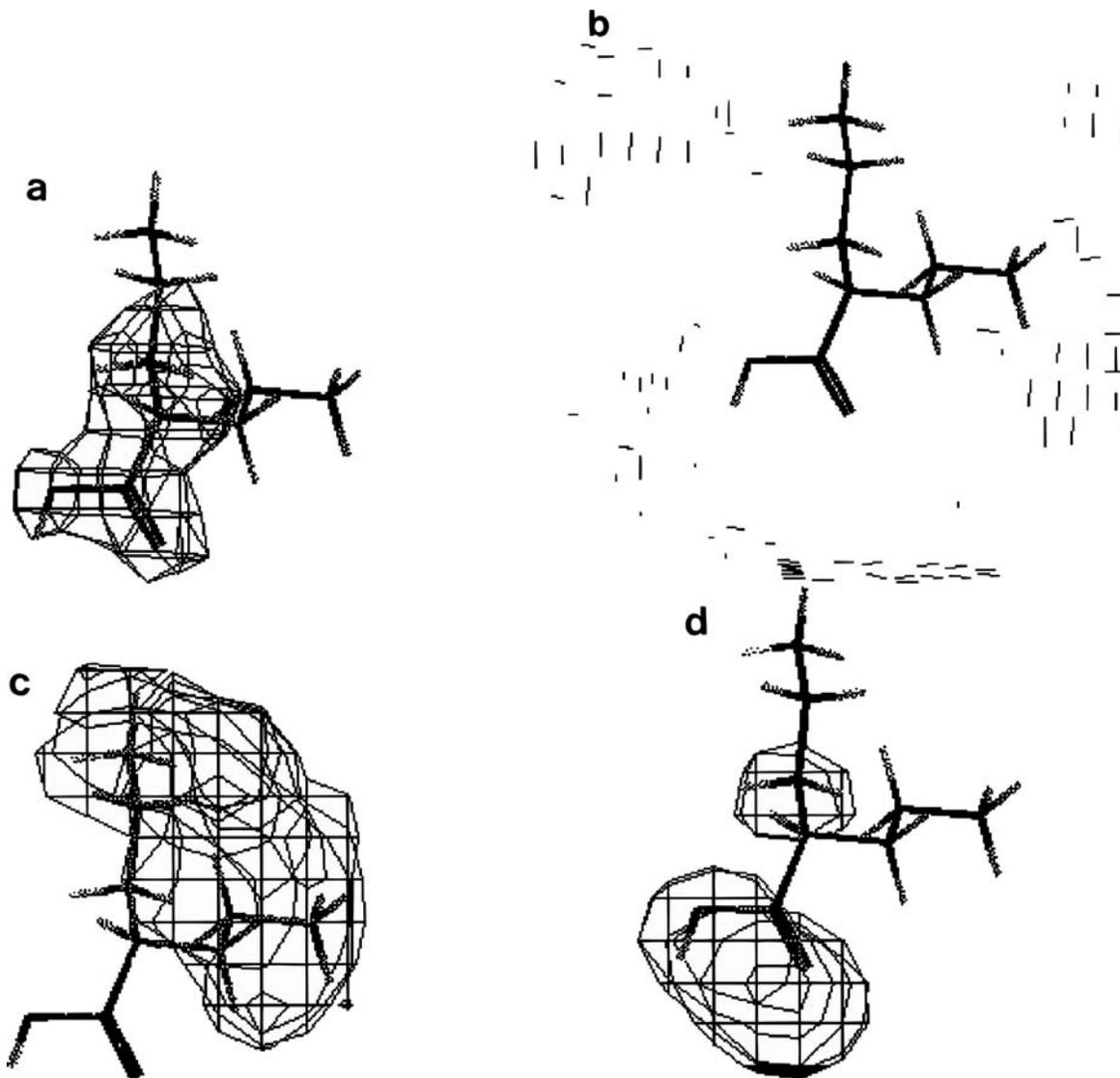
Fig. 3. WLS coefficient maps. Steric maps: positive (a) and negative (b) coefficients describe regions of space whose occupancy respectively increases or decreases the activity. Electrostatic maps: positive coefficients (c) describe regions where an electron deficiency increases the log RA, negative coefficients (d) describe regions where an electron density increases the log RA. The maps are displayed over the structure of VPA.
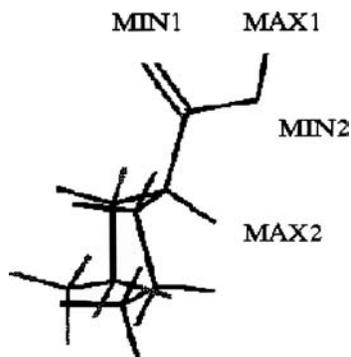
Fig. 4. Localisation of EP minimum and maximum values for the VPA molecule.

the hydrocarbon chains of VPA derivatives result in an enhancement of anticonvulsant activity. Unfortunately, this elongation and branching enhance also the toxicity of the compounds (9). The substitution of the proton with a methyl group at the α-position also leads to an increase in activity. Eq. (3) has goods statistical characteristics, especially according to the cross-validated regression coefficient, and can be used for prediction of activity of new compounds. Compound 19 remains an outlier (Fig. 2). As it is more active than predicted, it can be taking into consideration for a lead compound of new series of more active VPA derivatives.

3D QSAR

It was found that both the alignment technique and the type of query influence the quality and the predictive ability

of the models. The R^2_{pred} values are more sensitive to the alignment technique than to the number of atoms in the query. It was shown in Table IV that all the R^2_{pred} values for the models with rigid fitted molecules are negative, while a good half of these values in the models obtained after flexible fitting, are positive. The type of the query has greater influence in the models, in which the flexible fitting for molecular alignment is used. It was established that the greater number of atoms is included in the query, the higher R^2_{pred} values are obtained. Inclusion of the lipophilicity term log D improves statistical characteristics in all the 3D QSAR models.

The separation of steric and electrostatic fields leads to four non-correlated grid contours (Fig. 3). The comparison between the WLS electrostatic coefficient maps and the Fig. 4, representing the localization of EP extremum values shows that the high electron density, responsible for the appearance of the negative potential with extremum values MIN1 and MIN2 is favorable for the anticonvulsant activity. It seems that the electron deficiency in the vicinity of the carboxylic group is not desirable for the activity.

CONCLUSIONS

Finally, it may be assumed that the two approaches—2D and 3D QSAR analysis, complement each other. The 3D QSAR analysis is more descriptive and lays down the directions for rational design. For better results in 3D QSAR, flexible fitting to a large number of atoms, forming the query, is proposed, when molecules with high conformational flexibility are investigated. However, for predicting the activity of new VPA derivatives we prefer the conventional 2D QSAR analysis, because the parameters, describing the activity in the best model are independent from the method for optimization and the choice of conformation.

REFERENCES

1. G. Carraz, H. Beriel, Luu-Duc, and S. Lebreton. Aprochess dans la pharmacodinamique biochimique de la structure N-dipropyla-cetique. *Therapie* **20**:419–426 (1965).
2. V. Rogiers, A. Callaerts, A. Vercruisse, M. Arkrawi, E. Shephard, and I. Phillips. Effects of valproate on xenobiotic biotransformation in rat liver. In vivo and in vitro experiments. *Pharm. Weekbl. Sci.* **14**:127–131 (1992).
3. D. Johnston. Valproic acid: update on its mechanisms of action. *Epilepsia* **25**:S1-4 (1984).
4. R. Davis, D. Peters, and D. McTavish. Valproic acid. A reappraisal of its pharmacological properties and clinical efficacy in epilepsy. *Drugs* **47**:332–372 (1994).
5. D. Dawson. Additive incidence of developmental malformation for Xenopus embryos exposed to a mixture of ten aliphatic carboxylic acids. *Teratology* **44**:531–536 (1991).
6. G. Rekatas, E. Tani, V. Demopoulos, and P. Kouronakis. Synthesis of GABA-valproic acid derivatives and evaluation of their anticonvulsant and antioxidant activity. *Arch. Pharm. (Weinheim)* **329**:393–398 (1996).
7. A. Haj-Yehia and M. Bialer. Structure-pharmacokinetic relationships in a series of valpromide derivatives with antiepileptic activity. *Pharm. Res.* **6**:683–689 (1989).
8. A. Haj-Yehia and M. Bialer. Structure-pharmacokinetic relationships in a series of short fatty acid amides that possess anticonvulsant activity. *J. Pharm. Sci.* **79**:719–724 (1990).
9. W. Loscher and H. Nau. Pharmacological evaluation of varios metabolites and analogues of valproic acid. *Neuropharmacology* **24**:427–435 (1985).
10. A. Kimball, W. Burnett, and D. Doherty. Chemical protection against ionizing radiation. *Radiat. Res.* **7**:1–12 (1957).
11. Chem-X. Chemical Design Ltd., Chipping Norton, Oxfordshire Ox7 5SR UK (1998).
12. J. Gasteiger and M. Marsili. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron* **36**:3219–3288 (1980).
13. M. Dewar, E. Zoebisch, E. Healy, and J. Stewart. AM1: A new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **107**:3902–3909 (1985).
14. ACD/Labs. Advanced Chemistry Development Inc., Toronto, Canada (1995).
15. O. Mekenyan, St. Karabunarliev, J. Ivanov, and D. Dimitrov. A new development of the OASIS computer system for modeling molecular properties. *Computers Chem.* **18**:173–187 (1994).
16. S. Wold. Cross-validatory estimation of the number of components in factor and principal components models. *Technometrics* **20**:397–405 (1978).
17. R. Cramer III, D. Patterson, and D. Bunce. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **110**:5959–5967 (1988).
18. W. Dunn III, S. Wold, U. Edlund, S. Hellberg, and J. Gasteiger. Multivariate structure-activity relationships between data from a battery of biological tests and an ensemble of chemical descriptors: The PLS method. *Quant. Struct. Act. Relat.* **3**:131–137 (1984).
19. J. Sangster. Octanol—water partition coefficients of simple organic compounds. *J. Chem. Phys. Ref. Data* **18**:1111–1229 (1989).
20. B. Hudson, D. Livingstone, and E. Rahr. Pattern recognition display methods for the analysis of computed molecular properties. *J. Comp. Aided Mol. Des.* **3**:55–65 (1989).