QSAR Based Design, Synthesis & Evaluation of Anticancer & Analgesic Activity of Novel Benzimidazole Derivatives

Shubham Namdev[1], Priyal Jain[1], Jitendra Banweer[1], Sukhwant Singh[1], Swati Pandey[1]

SIRT-Pharmacy, Sage University Bhopal

Abstract

A set of benzotriazine derivatives were tested for their $GABA_A$/BzR complex binding affinity and minimum inhibitory concentrations were determined for all the compounds. Quantitative structure activity relationship (QSAR) analysis was applied to 28 of the above mentioned derivatives using a combination of various physicochemical, steric, electronic, and structural molecular descriptors. Different QSAR models revealed that CDE, PC & DDE parameters have significant impact on receptor binding affinity so as anticonvulsant activity of the benzotriazine derivatives. The low residual activity and adjacent r^2 values (r^2_{adj}) observed indicated the predictive ability of the developed QSAR models. A series of new 3-acylpyrazolo [5, 1-c] [1, 2, 4] benzotriazine 5-oxide were synthesized on the basis of QSAR findings, screened for their anticonvulsant activity in MES and PTZ models and were compared with standard drugs phenytoin sodium and sodium valproate.

Introduction

Epilepsy, a ubiquitous disease characterized by recurrent seizures, inflicts more than 60 million people worldwide according to epidemiological studies. The available drug therapy, however, cause notable adverse side effects such as drowsiness, ataxia, gastrointestinal disturbance, hepatotoxicity and megaloblastic anaemia 8

Gamma-aminobutyric acid (GABA), the major inhibitory neurotransmitter in the mammalian brain, targets the ionotropic GABAA and GABAC receptors and the metabotropic GABAB receptors. Of these, it is the GABAA receptor family which has been the most widely studied since these receptors is the molecular target for the action of a number of clinically important drugs, including benzodiazepines (BZd), barbiturates, and anaesthetics. 10

In order to obtain drugs which are more selective than BZd, the development of $GABA_A$ receptor subtype selective ligands or functionally selective ligands is the current strategy of medicinal chemistry researchers. 10

The quantitative structure-activity relationship (QSAR) research field provides medicinal chemists with the ability to predict drug activity by mathematical equations which construct a relationship between the chemical structure and the biological activity [1, 2]. These mathematical equations are in the form of $y = Xb + e$ that describe a set of predictor variables (X) with a predicted variable (y) by means of a regression vector (b) [3]. After the earlier QSAR studies by Hansch, who showed a correlation between biological activity and octanol-water partition coefficient [2], it is now assumed that the sum of substituent effects on the steric, electronic and hydrophobic interaction of compounds with their receptor determines their biological activity [4-6]. The first step in constructing the QSAR models is finding one or more molecular descriptors that represent variation in the structural property of

the molecules by a number [7]. Nowadays, a wide range of descriptors are being used in QSAR studies which can be classified into different categories according to the Karelson approach including; constitutional, geometrical, topological, quantum, chemical and so on [8]. There are different variable selection methods available including; multiple linear regression (MLR), genetic algorithm (GA), principal component or factor analysis (PCA/FA) and so on. The mathematical relationships between molecular descriptors and activity are used to find the parameters affecting the biological activity and/or estimate the property of other molecules. 10

In our previous work, a QSAR study of new 3-acylpyrazolo [5, 1-c] [1, 2, 4] benzotriazine 5-oxide derivatives for BZR ligand affinity has been performed and a series of derivatives of 2-substituted-

5-nitrobenzimidazole were synthesized based on concept of bioisosteric replacement which was first found to have anticonvulsant activities, among which 2-(4-Chloro-phenyl)-5-nitro-1*H*-benzimidazole showed the strongest activity with an ED50 value of 30 mg/kg in the maximal electroshock (MES) and Pentylenetetrazole induced seizure model (*sc*PTZ) test. 10

Analyzing the relationship of the structure of series of benzotriazine derivatives which shows different GABA/BZd binding affinity, result findings of QSAR study and anticonvulsant activity of designed compounds, it was found that ring possessing substitutions with partition coefficient, charge-dipole and dipole-dipole

energy as principle properties, enhanced the hydrophobic ability of target compounds, thus make them more permeable to the blood–brain barrier and enhance anticonvulsant activity.

So we thought that the presence of a bulkier and electronegative substitution in the ring was essential structure for the anticonvulsant activity. In this paper, we designed and prepared a series of 3-acylpyrazolo [5, 1-c] [1, 2, 4] benzotriazine 5-oxide derivatives. The hypothesis was that a benzotriazine ring with a bulkier acyl group may have higher affinity for the receptor and enhance their anticonvulsant activity. The new compounds were evaluated as anticonvulsant agents in experimental epilepsy models, i.e., maximal electroshock (MES) induced seizure in mice.

In view of the above and in continuation of our studies on the agonistic activities of benzotriazine derivatives, as well as on correlation of molecular properties with activity [4-8,28-35], the objective of this investigation was to study the usefulness of QSAR in the prediction of the anticonvulsant activity of 28 benzotriazine derivatives. Multiple linear regression (MLR) models have been developed as a mathematical equation which can relate chemical structure to the inhibitory activity. 10

Results and Discussion

In the first step of present study, different substituted benzotriazines were evaluated for GABAA/BzR complex binding affinity activity. Structures and binding data for new, synthesized compounds as inhibition constant (K_i) values are summarized in Table 1.

Table 1: The structure and BZR ligand affinity of new 3-acylpyrazolo [5, 1-c] [1, 2, 4] benzotriazine 5-oxide
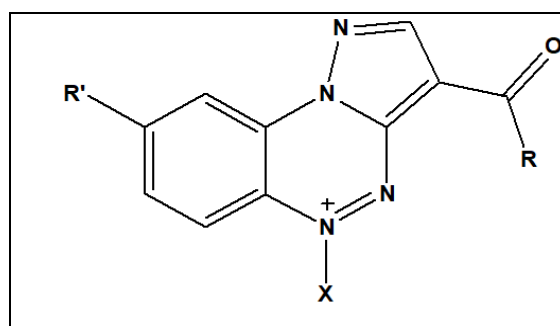
Compound No	R	R'	X	K_i (nM)	log $1/CK_i$
1.	H	Cl	O	64	-1.80618
2.	Me	Cl	O	551	-2.74115
3.	Me	OEt	O	902	-2.95521
4.	Me	OPh	O	53	-1.72428
5.	Ph	Cl	O	3494	-3.54332
6.	*o*-OMe-Ph	Cl	O	4200	-3.62325
7.	*p*-OMe-Ph	Cl	O	1240	-3.09342
8.	*p*-OMe-Ph	OEt	O	919	-2.96332
9.	3-Py	Cl	O	2400	-3.38021
10.	3-Py	OEt	O	2000	-3.30103
11.	1,3-Benzodioxol-5-yl	Cl	O	2900	-3.4624
12.	2-Furyl	Cl	O	72	-1.85733
13.	2-Furyl	OEt	O	28	-1.44716
14.	2-Furyl	OPh	O	7	-0.8451
15.	2-Furyl	OCH2PH	O	1.4	-0.14613
16.	2-Furyl	OCH2C=Ch	O	3.5	-0.54407
17.	2-Thienyl	Cl	O	62	-1.79239
18.	2-Thienyl	OEt	O	45	-1.65321
19.	2-Thienyl	OPh	O	8.9	-0.94939
20.	2-Thienyl	OCH2Ph	O	2.7	-0.43136
21.	3-Thienyl	Cl	O	121	-2.08279
22.	2-Pyrrolyl	Cl	O	4	-0.60206
23.	2-Pyrrolyl	OEt	O	1836	-3.26387
24.	2-Pyrrolyl	OPh	O	276	-2.44091
25.	1-Methylpyrrol-2-yl	Cl	O	5202	-3.71617
26.	1-Methylpyrrol-2-yl	OEt	O	819	-2.91328
27.	1-Methylpyrrol-2-yl	OPh	O	1300	-3.11394
28.	1-Methylpyrrol-2-yl	Cl	O	380	-2.57978

The screening results reveal that all the compounds exhibited appreciable *in vitro* binding affinity towards $GABA_A$/BzR receptor complex. In the second step, we focused our efforts on developing the QSAR models of compounds **1 – 28** as anticonvulsants. A set of benzotriazines was used for MLR model generation. The reference drugs were not included in model generation as they belong to a different structural series. Binding affinity data determined as nM were first transformed to the negative logarithms (log1/c nM), (Table 1) which was used as a dependent variable in the QSAR study. Different physicochemical, steric, electronic, and thermodynamic descriptors were used as independent variables and were correlated with $GABA_A$/BzR complex binding affinity, which finally responsible for anticonvulsant activity of tested compounds.

Developing a QSAR model requires a diverse set of data, and, thereby a large number of descriptors have to be considered. Descriptors are numerical values that encode different structural features of the molecules. Selection of a set of appropriate descriptors from a large number of them requires a method, which is able to discriminate between the parameters. Pearson's correlation matrix has been performed on all descriptors by using SYSTAT 12.0 Statistical Software. The analysis of the matrix revealed six descriptors for the development of MLR model. The values of descriptors selected for MLR model are presented in Table 2. Linear models were then formed by a stepwise addition of terms. A deletion process was then employed, whereby each variable in the model was held out in turn and using the remaining parameters models was generated. Each descriptor was chosen as input for the statistical

software package and then the stepwise addition method implemented in the software was used for choosing the descriptors contributing to the binding affinity of benzotriazine derivatives.

Table 2: Values of molecular descriptors used in the regression analysis.

Compound	HE	CDE	DDE	SBE	TRE	PC
1	-13.774	6.530	0.384	-0.972	5.140	3.896
2	-13.699	6.484	0.214	-1.010	4.772	2.223
3	-13.455	4.170	-0.100	-0.979	4.640	-0.483
4	-12.598	4.540	4.125	-1.170	0.166	1.156
5	-12.925	6.438	0.366	-1.017	7.381	1.376
6	-12.164	5.601	-3.819	-1.016	9.921	2.324
7	-12.089	6.937	0.534	-1.008	7.560	1.323
8	-11.950	4.632	0.216	-0.969	7.704	1.059
9	-13.158	7.514	1.107	-1.022	15.074	2.838
10	-13.009	5.216	0.784	-0.986	15.272	-0.205
11	-11.807	8.060	3.377	-1.204	6.379	0.863
12	-12.674	0.825	0.837	-1.634	10.954	0.112
13	-12.532	-1.471	0.595	-1.569	10.963	0.291
14	-12.499	4.979	3.059	-1.745	4.979	4.931
15	-12.237	-1.393	1.309	-1.568	5.280	3.530
16	-12.514	-0.815	2.524	-1.530	11.081	0.362
17	-12.798	2.959	2.791	-1.163	14.362	1.035
18	-12.583	2.612	2.501	-1.099	14.359	0.771
19	-12.526	0.770	2.628	-1.308	9.506	2.145
20	-11.780	-0.112	3.283	-1.060	8.524	3.010
21	-12.630	8.834	2.024	-0.962	17.058	1.035
22	-12.130	6.686	2.213	-1.922	9.195	0.114
23	-11.976	4.385	1.916	-1.874	9.063	-0.510
24	-11.929	4.780	2.110	-2.090	4.184	1.493
25	-12.040	6.088	1.296	-1.990	8.500	0.467
26	-11.928	3.865	0.994	-2.167	11.710	0.203
27	-11.824	4.161	1.119	-2.211	1.757	1.846
28	-12.094	6.073	0.484	-2.024	10.126	0.467

It was considered that two variables having correlation greater than 0.5 were not included in the final equation, so that by introducing different groups on the pharmacophore each variable could be regulated independently to enhance the biological activity of the designed compound. For the same purpose Pearson correlation matrix has been derived between selected descriptors which were shown in table 3.

Table 3: Pearson correlation matrix between selected descriptors

	BA	HE	CDE	DDE	SBE	TRE
BA	1.000					
HE	-0.138	1.000				
CDE	-0.619	-0.099	1.000			
DDE	0.433	0.133	-0.139	1.000		
SBE	0.071	-0.491	0.132	-0.073	1.000	
TRE	0.097	-0.007	-0.055	-0.061	0.127	1.000

Now for the purpose of validation of the final model, the data set was randomly divided into training and test group. Compounds of training set were used for internal validation and compounds of test set were used for external validation. In this study we removed compound 3, 9, 10, 14 and 16 as test compound and remaining were used as training set. Compounds of training set were used for the development of the equation.

The specifications for the best-selected MLR models are shown in Table 4. The di-parametric model indicated the importance of charge-dipole energy (CDE) and partition coefficient (PC) in contribution to $GABA_A$ receptor binding affinity which ultimately responsible for anticonvulsant activity (model 1, Table 4). Addition of dipole-dipole energy (DDE) as an additional parameter to model 1 significantly increased the linear regression coefficient (R^2) from 0.812 to 0.923 (model 2, Table 4). Similarly, the addition of a another parameter HOMO energy (HE) also increased the regression coefficient, but a MLR method only can be used when a relatively small number of molecular descriptors are used (at least five to six times smaller than the total number of compounds). In this case (for 19 compounds), only three descriptors can be used to develop a good QSAR model in order to avoid a high chance of spurious correlations. Finally by removing a compound as an outlier we reached to squared R value of 0.954 with only three descriptors explaining the biological activity of the series. In this approach, only the QSAR models 1 and 2 can be considered.

Table 4: Best MLR models for the prediction of GABAA/BzR complex binding affinity

No.	Descriptor used	MLR equation	N	R^2	SE	F
E1	CDE & PC	BA= -1.698 (±0.242) -0.300 (±0.040) CDE+ 0.360 (±0.098) PC	20	0.812	0.494	36.681
E2	CDE, PC & DDE	BA=-2.176 (±0.188)-0.273 (±0.027) CDE +0.220 (±0.046) DDE +0.415 (±0.066) PC	20	0.923	0.325	64.037
E3	CDE, PC & DDE	BA= -2.160(±0.149)-0.271 (±0.021) CDE +0.219 (0.036) DDE +0.429 (±0.052) PC	19	0.954	0.257	103.223

It is well known that there are three important components in any QSAR study: development of models, validation of models and utility of developed models. Validation is a crucial aspect of any QSAR analysis [36]. The statistical quality of the resulting models, as depicted in Table 4, is determined by r^2, SE, and F [37-39]. The F-value presented in Table 4 is found statistically significant at 99% level since all the calculated F values are higher as compared to tabulated values.

For the testing the validity of the predictive power of selected MLR models the LOO technique was used. The developed models were validated by the calculation of following statistical parameters: PRESS, S_{PRESS} and r^2_{adj} (Table 5). These parameters were calculated from the following equations

$$PRESS = \sum (Y_{obs}-Y_{cal})^2$$

$$S_{PRESS} = \sqrt{PRESS} / n$$

$$r^2_{adj} = 1- r^2 (n-1/n-p-1)$$

Where, *Y*obs, *Y*calc and *Y*mean are observed, calculated and mean values; *n*, number of compounds; *p*, number of independent parameters.

PRESS is an acronym for prediction sum of squares. It is used to validate a regression model with regards to predictability. To calculate PRESS, each observation is individually omitted. The remaining n – 1 observations are used to calculate a regression and estimate the value of the omitted observation. This is done n times, once for each observation. The difference between the actual Y value, y_{obs}, and the predicted Y, y_{calc}, is called the prediction error. The sum of the squared prediction errors is the PRESS value. The smaller PRESS is, the better the predictability of the model. These values are in terms of the dependent variable, y.

| Model No.	PRESS	S_{PRESS}	r^2_{adj}
1	4.1411	0.4550	0.790
2	1.6923	0.2908	0.909
3	1.7329	0.3020	0.945

In many cases r^2_{adj} is taken as a proof of the high predictive ability of QSAR models. A high value of these statistical characteristic (> 0.5) is considered as a proof of the high predictive ability of the model, although recent reports have proven the opposite [40]. It is proven that the only way to estimate the true predictive power of a model is to test it on a sufficiently large collection of compounds from an external test set. The test set must include no less than five compounds, whose activities and structures must cover the range of activities and structures of compounds from the training set. This application is necessary for obtaining trustful statistics for comparison between the observed and predicted activities for these compounds. A reliable model should be also characterized by a high correlation coefficient between the predicted and observed activities of compounds from a test set of molecules that was not used to develop the models.

Conclusion

Experimental

The synthetic pathway for preparation of the benzimidazoles is shown in Scheme 1. Compounds PN1, PN2, PN3 and PN4 were synthesized by condensation of corresponding aldehydes i.e. 4-Chlorobenzaldehyde, 4-Fluorobenzaldehyde, pyridine-2-carbaldehyde and 4-Methylbenzaldehyde; appropriate 4-Nitro substituted o-phenylenediamine, and sodium metabisulfite in absolute ethanol (95%).
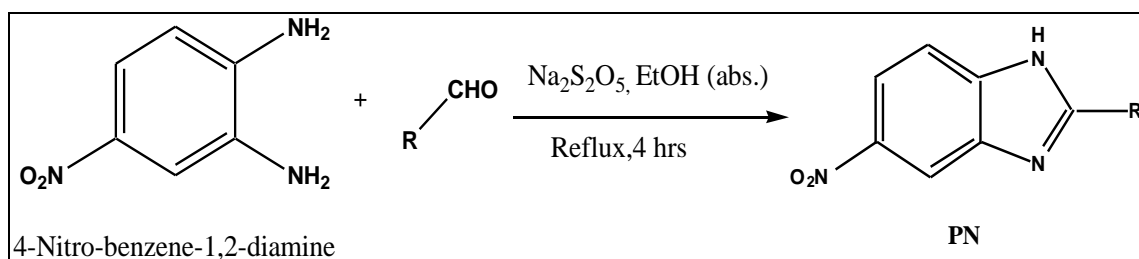
General procedure for Synthesis of 2-substituted-5-nitrobenzimidazole derivatives

To a Solution of 1eq of 4-nitro o-phenylenediamine and 1 eq of corresponding aldehyde in ethanol was added 4 eq of $Na_2S_2O_5$ and the resulting mixture was reflux for 4 h. After reaction mixture was cooled to room temperature, diethyl ether was added and the crude product was filtered off. The crude product was suspended in mixture of ethanol–diethyl ether several times until the powder was obtained analytically pure.

Reference:

1. C. Hansch, D. Hoekman, H. Gao, *Chem. Rev.* **1996**, *96*, 1045-1076.
2. C. Hansch, P.P. Maloney, T. Fujita, R.M. Muir, *Nature* **1962**, *194*, 178-180.
3. B. Hemmateenejad, *Chemom. Intell. Lab. Syst.* **2005**, *75*, 231-245.
4. T. Fujita, J. Iwasa, C. Hansch, *J. Am. Chem. Soc.* **1964**, *86*, 5175-5180.
5. C. Hansch, *Acc. Chem. Res.* **1968**, *2*, 232-239.
6. C. Hansch, J.M. Clayton, *J. Pharm. Sci.* **1973**, *62*, 1-21.
7. K. S. Agatonovic, I.G. Tucker, M. Zecevic, L.J. Ziva-novic, *Anal. Chem. Acta* **2000**, *418*, 181-195.
8. Mohajeri, B. Hemmateenejad, A. Mehdipour, R. Miri, *J. Mol. Graph. Model* **2008**, *26*, 1057-1065.
9. U. Rudolph, H. Mohler, *Annu. Rev. Pharmacol. Toxicol.* **2004**, *44*, 475–498.
10. J. R. Atack, *Expert Opin. Invest. Drugs* **2005**, *14*, 601–618.

11. H. V. Morris, G. R. Dawson, D. S. Reynolds, J. R. Atack, D. N. Stephens, *Eur. J. Neurosci.* **2006**, *23*, 2495–2504.

12. W. Sieghart, *Adv. Pharmacol.* **2006**, *54*, 231–263.

13. S. Basile, A. S. Lippa, P. Skolnick, *Drug Discovery Today: Ther. Strateg.* **2006**, *3*, 475–481.

14. P. J. Whiting, *Curr. Opin. Pharmacol.* **2006**, *6*, 24–29.

15. M. M. Savic, D. I. Obradovic, N. D. Ugresic, D. R. Bokonjic, *Neural Plast.* **2005**, *12*, 289–298.

16. R. Dias, W. F. A. Sheppard, R. L. Fradley, E. M. Garrett, J. L. Stanley, S. J. Tye, S. Goodacre, R. J. Lincoln, S. M. Cook, R. Conley, D. Hallett, A. C. Humphries, S. A. Thompson, K. A. Wafford, L. J. Street, J. L. Castro, P. J. Whiting, T. W. Rosahl, J. R. Atack, R. M. McKernan, G. R. Dawson, D. S. Reynolds, *J. Neurosci.* **2005**, *25*, 10682–10688.

17. G. R. Dawson, K. Maubach, N. Collinson, M. Cobain, B. J. Everitt, A. M. MacLeod, H. I. Choudhury, L. M. McDonald, G. V. Pillai, W. Rycroft, A. J. Smith, F. Sternfeld, D. Tattersall, K. A. Wafford, D. S. Reynolds, G. Seabrook, J. R. Atack, *JPET* **2006**, *316*, 1335–1345.

18. K. Maubach, *Drugs Future* **2006**, *31*, 151–162.

19. P. F. McKay, K. L. Foster, D. Mason, R. Cummings, M. Garcia, L. S. Williams, C. Grey, J. M. Cook, H. L. June, *Psycopharmacology* **2004**, *172*, 455–462.

20. Costanzo, G. Guerrini, G. Ciciani, F. Bruni, S. Selleri, B. Costa, C. Martini, A. Lucacchini, A. P. Malmberg, A. Ipponi, *J. Med. Chem.* **1999**, *42*, 2218–2226.

21. Costanzo, G. Guerrini, G. Ciciani, F. Bruni, S. Selleri, B. Costa, C. Martini, A. Lucacchini, A. P. Malmberg, A. Ipponi, *Med. Chem. Res.* **1999**, *9*, 322–339.

22. Costanzo, G. Guerrini, G. Ciciani, F. Bruni, C. Costagli, S. Selleri, B. Costa, C. Martini, P. Malmberg-Aiello, *Med. Chem. Res.* **2001**, *10*, 366–377.

23. Costanzo, G. Guerrini, G. Ciciani, F. Bruni, C. Costagli, S. Selleri, B. Costa, C. Martini, P. Malmberg-Aiello, *J. Med. Chem.* **2002**, *45*, 5710–5720.

24. S. K. Sullivan, R. E. Petroski, G. Verge, R. S. Gross, A. C. Foster, D. E. Grigoriadis, *J. Pharmacol. Exp. Ther.* **2004**, *311*, 537–546.

25. R. E. Petroski, J. E. Pomeroy, R. Das, H. Bowman, W. Yang, A. P. Chen, A. C. Foster, *J. Pharmacol. Exp. Ther.* **2006**, *317*, 369–377.

26. S. Lippa, P. Czobor, J. Stark, B. Beer, E. Kostakis, M. Gravielle, S. Bandypadhyay, S. J. Russek, T. T. Gibbs, D. H. Farb, P. Skolnick, *PNAS* **2005**, *102*, 7380–7385.

27. S.O. Podunavac-Kuzmanović, S.L. Markov, D.J. Barna, *J. Theor. Comp. Chem.* **2007**, *6*, 687-698.

28. S.O. Podunavac-Kuzmanović, D.D. Cvetković, *CICEQ* **2007**, *13*, 68-71.

29. N.U. Perišić-Janjić, S.O. Podunavac-Kuzmanović, *J. Planar. Chromatogr.* **2008**, *21*, 135-141.

30. S.O. Podunavac-Kuzmanović, S.L. Markov, D.J. Barna, *J. Serb. Chem. Soc.* **2008**, *73*, 967-978.

31. J.G. Topliss, R.P. Edwards, *J. Med. Chem.* **1979**, *22*, 1238-1244.

32. G.W. Snedecor, W.G. Cochran in *Statistical Methods*, Oxford and IBH, New Delhi, India, **1967**, 381.

33. S. Chaltterjee, A.S. Hadi, B. Price in *Regression Analysis by Examples*, Wiley VCH, New York, USA, **2000**.

34. M.V. Diudea in *QSPR/QSAR Studies for Molecular Descriptors*, Nova Science, Huntingdon, New York, USA, **2000**.

35. Golbraikh, J. Tropsha, *J. Mol. Graph. Mod.* **2002**, *20*, 269-276.

36. G. Guerrini, G. Ciciani, G. Cambi, F. Bruni, S. Selleri, F. Melani, M. Montali, C. Martini, C. Ghelardini, M. Norcini, A. Costanzoa, *Bioorganic & Medicinal Chemistry* **2008**, *16*, 4471–4489.

37. J. L. Falcó, M. Piqué, M. González, I. Buira, E. Méndez, J. Terencio, C. Pérez, M. Príncep, A. Palomer, A. Guglietta, *European Journal of Medicinal Chemistry* **2006**, *41*, 985–990.

38. Kamal, P. P. Kumar, K. Sreekanth, B. N. Seshadri, P. Ramulu, *Bioorganic & Medicinal Chemistry Letters* **2008**, *18*, 2594–2598.

39. D. C. White, T. D. Greenwood, A. L. Downey, J. R. Bloomquistc, J. F. Wolfea,

Bioorganic & Medicinal Chemistry **2004, 12**, 5711–5717.

40. R. J. Krall, J. K. Penry, B. G. White, H. J. Kupferberg, E. A. Swinyard, *Epilepsia* 1978, *9*, 409.

41. R. J. Poter, J. J. Cereghino, G. D. Gladding, B. J. Hessie, H. J. Kupferberg, B. Scoville, *Clev. Clin. Q.* **1984**, *51*, 293.