Int. J. Chem. Sci.: 10(1), 2012, 539-547
ISSN 0972-768X
www.sadgurupublications.com

IN SILICO QUANTITATIVE STRUCTURE PHARMACOKINETIC RELATIONSHIP MODELING ON ARYL CARBOXYLIC ACID: VOLUME OF DISTRIBUTION AND SERUM PROTEIN BINDING

BHUPINDER SINGH[*], PRIYADEEP, AMAN SINGLA[a] and YASH PAUL[a]

University Institute of Pharmaceutical Sciences, UGC Centre of Advanced Studies,
Panjab University, Sector 14, CHANDIGARH – 160014 INDIA
[a]Lord Shiva College of Pharmacy, SIRSA (Haryana) INDIA

ABSTRACT

An estimate of volume of distribution (V_d) and serum protein binding (% SPB) is of paramount importance in assessing the efficacy of drugs used to treat acute conditions like pain, which can be treated by a single dose. This study was conducted to develop Quantitative Structure Pharmacokinetic Relationship (QSPR) for the prediction of V_d and % SPB in men for congeneric series of twelve arylcarboxylic acid derivatives, using computer assisted Hansch approach. The QSPR correlations were duly analyzed using a battery of apt statistical procedures and validated using leave-one-out (LOO) approach. Analysis of several hundreds of QSPR correlations developed in this study revealed high degree of cross-validated coefficients (Q^2) using LOO method ($p < 0.001$). The overall predictability was found to be high in case of V_d ($R^2 = 0.9846$ $F = 117.51$ $S^2 = 0.0007$, $Q^2 = 0.9617$ $p < 0.001$) as well as % SPB ($R^2 = 0.9764$ $F = 204.31$ $S^2 = 0.0942$, $Q^2 = 0.9684$ $p < 0.001$). Topological and steric parameters were found to primarily ascribe the variation in V_d and % SPB. The results indicated the involvement of dissolution rate limited absorption rather than permeation limited, as hardly any dependence on Log P was observed.

Key words: Quantitative structure pharmacokinetic relationships (QSPR), Volume of distribution (V_d), Serum protein binding, In Silico ADME, Arylcarboxylic acid derivatives.

INTRODUCTION

It is now duly recognized by the pharmaceutical industry that undesirable absorption, distribution, metabolism and excretion of new drug candidates are the cause of many clinical

[*]Author for correspondence; Ph.: +91-172-2534103 (O); 2727575 (R); Fax: 2541142;
E-mail: bsbhoop@yahoo.com

phase drug development failure. Nearly 40% of the drug candidates fail during the clinical trials owing to their poor pharmacokinetic properties. This is an economic disaster as the failed drugs have been in the pipeline for several years with high expenditure of efforts, time and money invested in their development. More recently in silico ADME modelling has been investigated as a tool to optimize selection of the most suitable drug candidate for development. The use of computational models in the prediction of ADME properties has been growing rapidly in drug discovery as they provide immense benefits in throughput and early application of drug design[1,2].

The major aim of in silico QSPR is to enable the drug designer to modify the chemical structure of a pharmacodynamically active drug so that its pharmacokinetic property may be altered without compromising pharmacodynamic potential. An early assessment of ADME properties will help pharmaceutical scientist to select the best drug candidate for development and as well as to reject those with a low plausibility of success. In silico QSPR technique tend to save considerable amount of time, money, animal life and involvement of "normal, healthy and drug-free volunteers" required for conducting the experimental pharmacokinetic studies[3,4].

Volume of distribution (V_d) and serum protein binding (% SPB) values of a drug are vital pharmacokinetic parameters because they are directly related to the bioavailability and can be used in assessing the efficacy of drug used to treat acute conditions like pain, arthritis, which can be treated by single dose. Hence, it is important to predict the V_d and % SPB values of drug leads during drug discovery, so that compounds with acceptable rate of absorption can be identified and those with poor bioavailability can be eliminated. The current study was conducted to investigate *in silico* QSPR amongst various aryl carboxylic acid anti-inflammatory drugs for volume of distribution (V_d) and % SPB. Arylcarboxylic acid derivatives were chosen for QSPR as this category of drugs has extensively been used as antiinflammatory agents in the treatment of acute conditions like pain, arthritis, etc. Moreover, arylcarboxylic acid derivatives consist of significant number of compounds thoroughly investigated for their pharmacokinetic performance particularly V_d (n = 12) Further, the congeners in this class have many common pharmacokinetic characteristics, mechanism and degree of affinity with body tissues.

Application

As an instrument for prediction

(i) Estimation of physicochemical properties using subsistent constants,

(ii) Reduction of the number of compounds to be synthesized,

(iii) Faster detection of the most promising compounds, and

(iv) Avoidance of synthesis of compounds with same activity.

As a diagnostic instrument

(i) Information on possible types of interaction forces,

(ii) Information on the nature of receptor, and

(iii) Information on the mechanism of fraction.

Detection of exceptions (outlier)

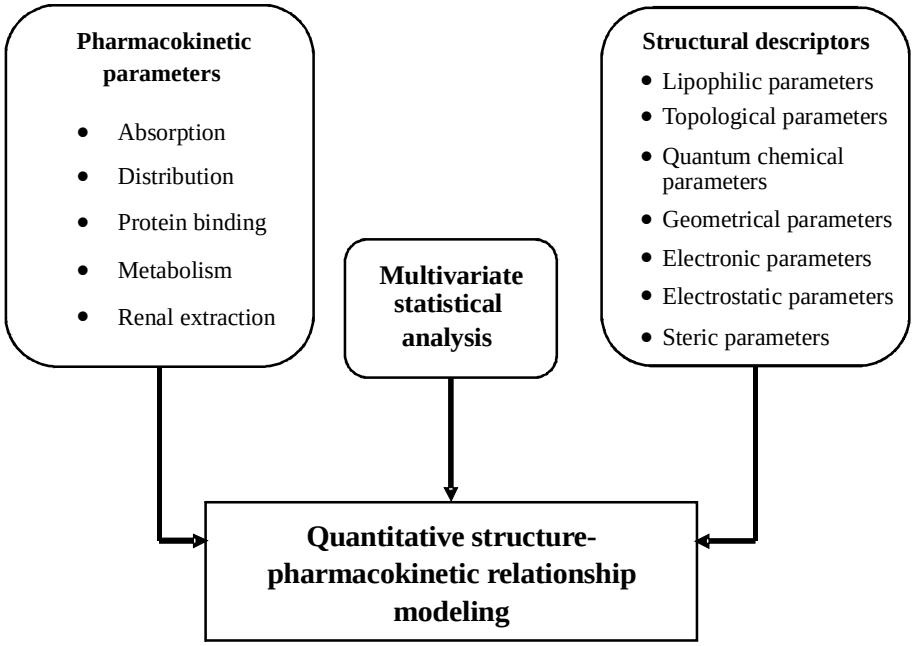

Fig. 1: Quantitative structure pharmacokinetic relationship (QSPR) modeling[5]

Methods

QSPR was conducted amongst arylcarboxylic acid derivatives employing extra-thermodynamic Multi Linear Regression Analysis (MLRA) or Hansch approach. The general steps for developing QSPR model include data set selection, chemical structure entry, 3D structure generation and descriptor calculation, model construction that involves selection of descriptors and validation of testing set using a Pentium dual core (Intel, USA) Desktop (IBM, USA) with 1GB RAM and 160 GB Hard Disk.

Dataset selection

Twelve arylcarboxylic acid derivatives with known human volume of distribution (V_d) and % SPB values were selected from literature[6-8]. In order to ensure that experimental variations in determining V_d and % SPB do not significantly affect the quality of our datasets, only V_d and % SPB values obtained from healthy adult males after oral administration of drug were used for constructing the dataset. The V_d and % SPB value of each of these compounds was also log-transformed (Log V_d) to normalize the data to reduce unequal error variance.

Molecular structure and descriptors

Chemical structures were drawn using suitable templates under Chem draw 7.0 software (Cambridge Soft Corporation, Cambridge, MA) and energy minimization was carried out using Chem 3D pro 3.5 software and the files were saved as MDL *molfiles*. *Molfiles* generated by Chem 3D were exported to DRAGON software, and as many as 1497 diverse descriptors, *viz.* constitutional, geometrical, topological, Whim 3D, electronic, electrostatic etc. were calculated. *Molfiles* were also imported in CODESSA 2.0 software (Semichem, Shawnee, USA) for calculation of more molecular descriptors.

Multivariate statistical analyses

Attempts were made to correlate various descriptors with the V_d and % SPB values. The initial regression analysis was carried out using heuristic analysis followed by best MLRA (RGMS) options of CODESSA software. All the descriptors were checked to ensure that value of each descriptor was available for each structure and there is a significant variation in these values. Descriptors for which values were not available for every structure in the data in question were discarded. Thereafter, the one and multiple parameter correlation equations for each descriptor were calculated.

Pharmacokinetic data of V_d and % SPB parameter available for twelve arylcarboxylic acid derivatives were analyzed, limiting the ratio of descriptors: drug to 1 : 4. As a final result, the heuristic method yields a list of the best ten correlations each with the highest r^2 and F-values. Many such attempts were carried out to obtain significant correlations for arylcarboxylic acid derivatives. A set of important descriptors found to significantly ascribe the variation of V_d and % SPB, was constructed. Further, a search for the multi-parameter regression with the maximum predicting ability was performed. A number of sets of descriptors were thus made and MLRA performed with V_d and % SPB. Regression plots of each correlation; thus, attempted were examined. Residual plots were also studied for absence of randomization and distinct patterns to eliminate chance correlations.

Validation of testing set

The predictability of the final models was tested by LOO method. Briefly, the descriptors of one compound are removed, the model is redefined and the target properties of the removed compound are predicted. This process is repeated until all target properties have been predicted once for each drug. A value of cross-validated R^2, commonly called Q^2, is then computed analogous to the conventional R^2 according to equation (1):

$$Q^2 = 1 - \frac{\sum\left(y_{pred} - y_{obs}\right)^2}{\sum\left(y_{obs} - y_{mean}\right)^2} \quad \ldots(1)$$

A model with good predictive performance has a Q^2 value close to 1, models that do not predict better than merely chance alone can have negative values.

The F-values were computed according to Equation (2):

$$F = \frac{S_1^2}{S_2^2} \quad \ldots(2)$$

Where, S_1 is variance between samples and S_2 variance within samples.

The values of computed F-ratio were compared with the critical values tabulated in statistical texts and levels of significance discerned. The correlations found to be statistically significant were compiled from CODESSA software.

RESULTS AND DISCUSSION

Variable QSPR results were obtained following application of multivariate statistical analysis on arylcarboxylic acid derivatives. The prominent descriptors explaining variation in V_d encompasses, melting point, VDW, mass, Kier Hall indices, information contents and structured information contents. The electronic parameters like TMSA, WNSA, DPSA minimum partial charges, etc. and geometric parameters like XYS yielded minor contributions towards improvement in relationships. Thus overall diffusional (rather than permeational) interactions seem to play pivotal role in attributing V_d. The small magnitude of Q^2, however may tend to lessen the significance of QSPRs on $V_{d.}$. Albeit the logarithmic values tend to marginally improve the correlation degree (r^2), but the Q^2 values tend to be very significantly higher connoting QSPRs using log V_d yield far more significant results. Volume of distribution is a negative indicator of the rate of absorption of a drug. Good QSPR correlations were obtained for the value of V_d for arylcarboxylic acid.

Table 1: Significant linear and logarithmic QSPR polynomial equations along with the statistical parameters for a series of aryl carboxylic acid derivatives using volume of distribution as the pharmacokinetic parameters

Eq. No.	Equation	R^2	Q^2	S	F-value	P <
1	$_{Vd}$ = -4.6821 + 0.4308*$_{VD}$W	0.4647	0.3469	3.425	16.25	0.01
2	$_{Vd}$ = -43.326-316.8* Min PCO	0.4402	0.2641	5.984	12.26	0.01
3	$_{Vd}$ = -21.479-0.5876*$_{VD}$W + 0.02415*Ql	0.9215	0.8942	0.426	24.34	.005
4	$_{Vd}$ = -15.8427 + 0.3157*PLN-0.6124*Bl	0.8672	0.8127	0.975	22.37	.005
5	$_{Vd}$ = 8.8124-0.9031*$_{VD}$W + 0.5124* Ql-0.3199*AEC	0.9846	0.9617	0.001	117.51	.001
6	$_{Vd}$ = -5.35124-0.7621*$_{VD}$W + 0.2157* KFI-0.3199*AEC	0.9527	0.9017	0.001	113.24	.001
7	Log_{Vd} = -1.805 + 0.1732*KHI2	0.6278	0.5421	1.548	19.82	.01
8	Log_{Vd} = -0.8126 + 0.0059*IC2	0.5410	0.5017	1.648	17.48	.01
9	Log_{Vd} = -3.9457 + 0.6128 *KSI3 + 5.3169*XY/XZ	0.9405	0.9127	0.025	79.25	.005
10	Log_{Vd} = -5.6124 + 0.04215 *ECC + 0.3215*AEC	0.9348	0.9084	0.049	64.75	.005
11	Log_{Vd} = -6.3127 + 0.02145 *ZX + 0.01578*TMSA-0.1348*IC2	0.9957	0.9864	0.001	120.35	.005
12	Log_{Vd} = -3.2182 + 0.6127*$_{VD}$W + 0.3512 *CIC2 + 2.9427*RNGC	0.9924	0.9543	0.001	119.84	.005

Thus logarithmic, transformation furnishes better correlations, LOO correlations, residuals and the overall statistical significance. The residual plot widely shows that V_d values tend to be clustered. However in log V_d plot the cluster tends to be partially dispersed, thus improving the regulations of residuals. Fig. 2 shows the linear and residuals plots between the calculated and predicted value of V_d using multi-parameters QSPR studies for a series of 12 arylcarboxylic acid derivatives.

Serum protein binding (Table 2) was found to depend upon the topological parameters like ASIC, AIC, CIC, etc. and steric parameters like RMW. Thus, % SPB does not seem to have any dependence on lipophilic and electronic parameters ruling out the

hydrophobic and ionic bonding. Constitutional parameters like NH, RNH, RNC, No, RNO etc. are also found to be important. Hydrogen and Vander Waals interactions are likely to play a stellar role in governing serum protein binding.

Table 2: Significant linear and logarithmic QSPR polynomial equations along with the statistical parameters for a series of aryl carboxylic acid derivatives using % serum protein binding as the pharmacokinetic parameters

Eq. No.	Equation	R^2	Q^2	S	F-value	P <
1	% SPB = -82.656+2.2457*XYS	0.5187	0.4988	5.948	21.54	0.005
2	% SPB = 85.836-27.977*ACIC2	0.4976	0.4857	7.247	11.57	0.010
3	% SPB = -147.518 + 12.596*BIN + 0.2015*SPH	0.9470	0.9287	1.943	132.75	0.001
4	% SPB = -60.5124 + 1.2615*XYS-0.1059*LCI	0.9528	0.9456	1.984	78.22	0.005
5	% SPB = -125.84 + 31.517*BIN + 2.0294* ACIC2-0.2157*GI	0.9764	0.9684	0.094	204.31	0.001
6	% SPB = -102.84-12.945*GI + 21.94* KSI-0.3218*LCI	0.9658	0.9268	1.102	184.24	0.001
7	Log % SPB = -0.5862 + 0.4646*AIC2	0.4929	0.4591	7.948	10.26	0.010
8	Log % SPB = -0.8691 + 0.6386*AIC1	0.4860	0.4262	8.249	9.84	0.050
9	Log % SPB = -1.234 + 0.6378*XYS + 0.3157*SPH	0.9084	0.8423	2.165	59.86	0.005
10	Log % SPB = 2.9457 + 0.2648*AIC2 + 0.1237*SPH	0.8857	0.8216	3.249	52.14	0.005
11	Log % SPB = -3.8461 + 0.1021*KSI3 + 0.0686*FNSA1 + 3.3852*HI	0.9619	0.9215	1.439	178.52	0.001
12	Log % SPB = -4.5127 + 0.6124*RI + 0.1027* PI-0.5127*DPSA2	0.9584	0.9042	1.525	157.67	0.001

Logarithmic transformation does not improve the degree of correlations.

Fig. 3 shows the linear and residuals plots between the calculated and predicted value of % SPB using multi-parameters QSPR studies for a series of 12 arylcarboxylic acids.

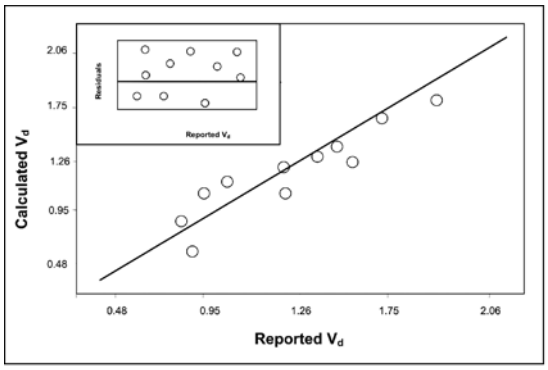

Fig. 2: Plot between the predicted and reported values of V_d for QSPR of arylcarboxylic acid derivatives (The inset shows residual plot)

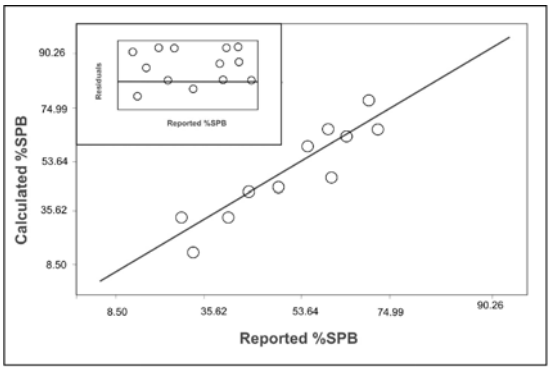

Fig. 3: Plot between the predicted and reported values of % SPB for QSPR of arylcarboxylic acid derivatives (The inset shows residual plot)

CONCLUSIONS

Analysis of several hundreds of QSPR correlations and consequent profiles in the current investigations on arylcarboxylic acid derivatives revealed that:

The quantitative relationships for various pharmacokinetic parameters were highly predictable in most cases ($p < 0.001$).

V_d and % SPB was found to be primarily a function of steric and topological parameters. The geometrical parameters like XYS, MIA also yielded minor contributions towards improvement in relationships. Thus, overall, the diffusional interactions seem to play a pivotal role in attributing V_d rather than the permeational ones, as hardly any dependence upon lipophlic parameters were observed. The overall predictability of V_d was found to be high ($R^2 = 0.9846$, $S = 0.0007$, $Q^2 = 0.9617$, $p < 0.001$). Logarithmic transformation tends to improve the correlations marginally ($R^2 = 0.9924$, $S = 0.0002$, $Q^2 = 0.9543$, $p < 0.005$).

Pharmacokinetic performance of a drug is known to be not merely a function of its physicochemical nature, but of the biological system(s) too like somatic, psychological, pathological environmental, nutritional, genetic, hereditary and diurnal status of the human subjects. This causes a great deal of plausible variation in pharmacokinetic profiles amongst the volunteers/patients undergoing study. The literature values of the pharmacokinetic

parameters, taken up in the present investigations, pertain to diverse subject populations hailing from different age groups, genders, races, nutritional and physical attributes, etc. studied in different geographical regions under different weather conditions. Considering these potentially high inter-subject and intra-subject variations amongst the pharmacokinetic parameters, the currently established relationships assume much higher credibility. It seems highly probable that the *in silico* approaches will evolve farther for ADME prognosis, as did the *in vitro* methods during the last decade. Past experience with the latter could be helpful in avoiding repetition of similar errors and taking the necessary steps to ensure effective implementation of the former.

ACKNOWLEDGEMENTS

The authors are thankful to Rishi Kapil and Babita Garg for providing requisite inputs for the current research work.

REFERENCES

1. Yash Paul, A. S. Dhake, M. Parle and B. Singh, Int. J. Chem. Sci., **7(3)**, 1613-1623 (2009).
2. B. Singh, A. S. Dhake, D. Sethi and Yash Paul, The Pharm. Rev., **8**, 93-100 (2007).
3. Yash Paul, A. S. Dhake, M. Parle and B. Singh, Res. J. Pharm. Tech., **1(2)**, 106-111 (2008).
4. Van de H. Waterbeemed, E. Gifford, Nature Reviews. Drug Disco., 2(3), 192-204 (2003).
5. M. Grover, B. Singh, M. Bakshi and S. Singh, Pharm. Sci. Tech. Today, **3(1)**, 28 (2000).
6. L. Shargel, S. Wu-Pong and Y. U. Andrew, Applied Biopharmaceutics and Pharmacokinetics, Ed. 5th, McGraw Hill, Columbus,OH, USA (2005) p. 259, p. 460, pp. 864-66.
7. D. C. Hooper and J. S. Wolfson, Quinolone Antimicrobial Agents, 2nd (Ed.) American Society for Microbiology, Washington, (1993) pp. 195-223.
8. A. Lubasch, I. Keller, K. Borner, P. Koeppe and H. Lode, Antimicrob. Agents Chemother, **44**, 2600-2603 (2000).

Revised : 22.10.2011 *Accepted : 23.10.2011*