IOCVS 2025 Conference
The 1st International Online Conference on Veterinary Sciences: Veterinary Virology in the New Era
03–05 December 2025 | Online
MDPI

Development of a Quantitative Structure–Retention Relationship (QSRR) Model for Predicting Veterinary Drug Retention Time in Food Matrices

Hamada A. A. Noreldeen [1]*, Abdel-Rahman Mustafa Hamed [2], Yassein Abdel-Maksoud Ahmed Osman [1]

[1] National Institute of Oceanography and Fisheries, NIOF, Cairo, 4262110, Egypt. [2] Faculty of Science, Chemistry, Faculty of Science, Al-Azhar University, Assuit, 71111, Egypt.

Abstract

Illegal additives, such as veterinary drugs, in food matrixes pose continuous risks to human health, necessitating new screening methods. We established a robust Quantitative Structure-Retention Relationships (QSRR) model based on Multiple Linear Regression (MLR) to predict the LC retention time of 95 veterinary drugs. The model showed a strong correlation and high reliability across four validation methods, providing a cost-effective tool for rapid illegal additive screening in food safety.

Introduction

Food safety monitoring requires fast screening methods, but reference standards for all potential illegal additives are often unavailable. QSRR is a chemo-informatic technique that relates a compound's structure-derived molecular descriptors to its chromatographic behavior. This method can significantly improve prediction accuracy and enhance the identification of unknown compounds in complex samples. The objective was to develop a simple, highly predictive QSRR model to predict the retention time of three classes of veterinary drugs in food matrices

Hypothesis

The core objective was to establish a stable and reliable Multiple Linear Regression (MLR)-based QSRR model that can accurately predict the LC retention time of diverse veterinary drugs in complex food matrices.

Materials and Methods

The study used 95 compounds (62 training, 30 test, 3 real samples) from three veterinary drug classes, extracted from fish tissue and analyzed via UHPLC-LTQ Orbitrap XL MS. Molecular descriptors were calculated using the free software tools ACD and TEST, followed by optimization to select the ten most significant descriptors (e.g., ACDlogP, BEHP1). A Multiple Linear Regression (MLR) model was built on the training set using the R-package "leaps" for variable selection. Model validation included internal (k-fold, leave-one-out cross-validation) and external checks using the test set and complex fish real samples

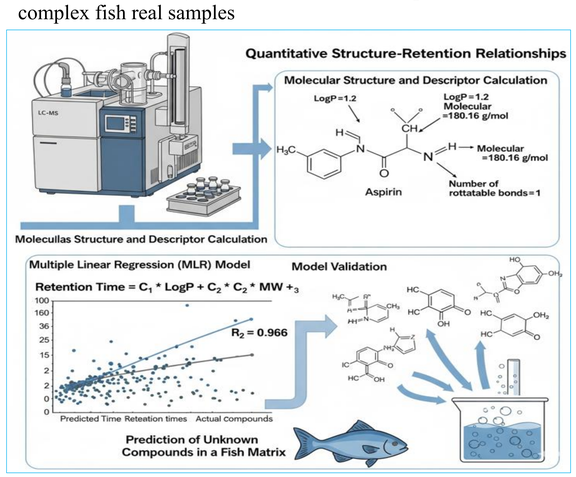

Figure 1. Schematic of QSRR model establishment and validation

Results & Discussion

The final MLR model provided an excellent fit for the training data (), demonstrating a strong linear relationship between the selected descriptors and retention time. Hydrophobicity (ACDlogP and ALOGP) was identified as the primary factor influencing chromatographic retention behavior among the ten selected descriptors. Crucially, the external validation results across all sets confirmed the model's strong predictive capability in complex fish tissue matrices. This straightforward QSRR method is a powerful, cost-effective tool for improving the detection and risk assessment of illegal additives in food safety control.

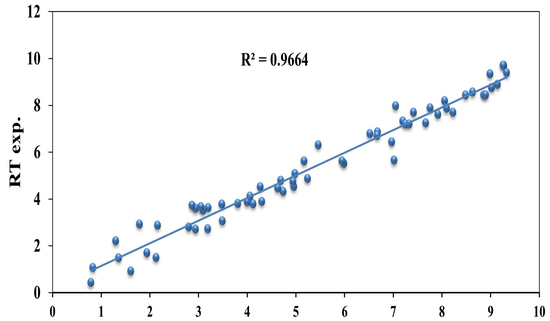

Figure 2. Experimental retention times versus predicted gradient retention times generated by MLR of 62 data points in the training sets.

The following equation (Eq. 1) describes the QSRR models with their statistical and internal validation parameters.

$RT_{pred.}$ = -35.693 + 0.848*ACDlogP + 0.132*ib + 5.348*BEHp1 + 5.779*BEHp2 + 1.532*GATS1m - 1.182*GATS2m + 0.788*ALOGP - 0.195*ALOGP2 - 0.556*Hy - 1.889*Ui

R^2 = 0.966, R^2_{Adj} = 0.959, p-value: < 2.2e-16

MAE = 0.455, RMSE =0.622.

MAEcv = 0.341, RMSEcv = 0.586, K=10

MAEloocv = 0.344, RMSEloocv = 0.587, K=10.

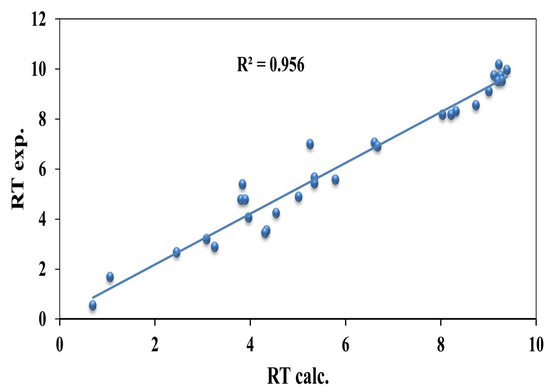

Figure 3. Experimental retention times versus predicted gradient retention times generated by MLR of 30 data points in the test sets.

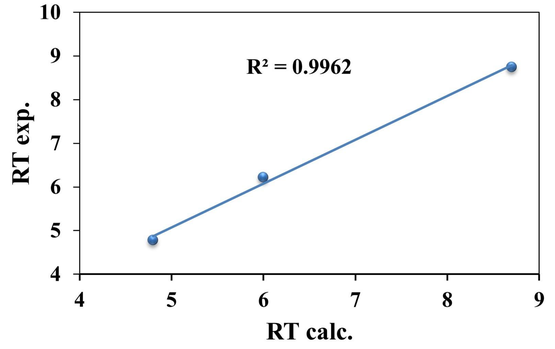

Figure 4. Experimental retention times versus predicted gradient retention times generated by MLR of 3 data points in the real sample.

Table 1. Symbols and definition of molecular descriptors in final QSRR models.

Symbol	Class	Definition
ACDlogP	Molecular properties	Octanol-water partitioning coefficients.
ALOGP	Molecular properties	Ghose-Crippen octanol water coefficient.
ALOGP2	Molecular properties	Ghose-Crippen octanol water coefficient squared.
Hy	Molecular properties	Hydrophilic index.
Ui	Molecular properties	Unsaturation index.
ib	Information Indices	Information bond index.
BEHp1	Burden Eigenvalue Descriptors	Highest eigenvalue n. 1 of Burden matrix/ weighted by atomic polarizabilities.
BEHp2	Burden Eigenvalue Descriptors	Highest eigenvalue n. 2 of Burden matrix/ weighted by atomic polarizabilities.
GATS1m	2D Autocorrelation Descriptors	Geary autocorrelation-lag 1/weighted by atomic masses.
GATS2m	2D Autocorrelation Descriptors	Geary autocorrelation-lag 2/weighted by atomic masses.

Table 2. Experimental gradient retention times (RT exp.) and predicted gradient retention times (RT pred.) for the test set.

NO.	Name	RT exp.	RT pred.	NO.	Name	RT exp.	RT pred.
Sulfonamides				16	Prochlorperazine Dimaleate	6.61	7.06
1	Sulfabenzamide	5.35	5.65	17	Colterol	1.06	1.67
2	Sulfamethoxypyridazine	3.97	4.05	18	Ractopamine	3.84	5.38
3	Sulfamoxole	3.81	4.76	**Cortical hormones**			
4	Sulfamethazine	3.90	4.77	19	Amcinonide	9.22	10.18
5	N4-Phthalylsulfathiazole	5.26	6.99	20	Halcinonide	9.12	9.75
6	Sulfadimethoxine	5.79	5.58	21	Clobetasone 17-butyrate	9.39	9.95
7	Sulfamethoxazole	4.55	4.24	22	Alclometasone-17,21-dipropionate	9.26	9.73
8	Diaveridine	3.26	2.88	23	Fluticasone propionate	9.20	9.51
Beta-receptor agonist				24	Mometasone furoate	9.29	9.52
9	Brombuterol	4.35	3.54	25	Prednisone	6.68	6.90
10	Formoterol	4.31	3.45	26	Hydrocortisone 21-acetate	8.04	8.16
11	Mapenterol	5.35	5.43	27	Diflorasone diacetate	9.01	9.10
12	Clorprenaline	3.09	3.19	28	Cortisone 21-acetate	8.32	8.29
13	Bambuterol	5.02	4.88	29	Prednisolone 21-acetate	8.22	8.17
14	Isoproterenol	0.70	0.53	30	Betamethasone 21-Acetate	8.74	8.56
15	Cimbuterol	2.46	2.66				

Future Studies

This validated QSRR model can be immediately integrated into laboratory pipelines for the preliminary identification and risk warning of unknown additives. Future work will focus on extending this predictive approach to other compound classes or different food matrices to broaden its application scope.

References

- M. Talebi, et al., 2015, J. Chromatogr. A, 1424, 69-76.
- S.H. Park, et al., 2017, J. Chromatogr. A, 1523, 173-182.
- E. Tyteca, et al., 2017, J. Chromatogr. A, 1486, 50-58.
- A.K. Ghose, G.M. Crippen, 1986, J. Comput. Chem., 7, 565-577.
- A.K. Ghose, A. Pritchett, G.M. Crippen, 1988, J. Comput. Chem., 9, 80-92.