Jordan Journal of Pharmaceutical Sciences, Volume 14, No. 2, 2021

2D-QSAR modeling of novel pleconaril derivatives (isoxazole-based molecules) as antiviral inhibitors against Coxsackievirus B3 (CVB3)

Youness Moukhliss [1], Khalil El Khatabi [1], Yassin Koubi [1], Hamid Maghat [1*], Abdelouahid Sbai [1], Mohammed Bouachrine [2] and Tahar Lakhlifi [1]

[1] Molecular Chemistry and Natural Substances Laboratory, Faculty of Science, Moulay Ismail University of Meknes, Morocco.
[2]EST Khenifra, Sultan Moulay Sliman University, Benimellal, Morocco.
*Hamid Maghat Tel: 00212614571353 ; E-mail: h.maghat@umi.ac.ma

ABSTRACT

Because of their acute pathologies, the search for antiviral drugs for coxakievirus B3 (CVB3) is becoming an urgent and unavoidable necessity. In the present study, a series of Pleconaril derivatives (anti-CVB3 molecules) are subjected to 2D-QSAR study, in which the objective is the construction of a predictive model of new anti-CVB3 candidates that are more active than the ones studied. Two models are obtained by multiple linear regression ($R^2 = 0.893$; $Q^2_{CV} = 0.837$; $R^2_{test} = 0.778$) and multiple non-linear regression ($R^2 = 0.918$; $Q^2_{CV} = 0.784$; $R^2_{test} = 0.734$) showing very satisfactory results. Based on the obtained optimal QSAR models, novel Pleconaril derivatives are designed as new CVB3 inhibitors, showing remarkably improved inhibitory activity compared to the existing system. These results might be useful for advanced research in future experimental work.
Keywords: 2D-QSAR, MLR, MNLR, Coxakievirus B3, Pleconaril, Isoxazole, Antiviral.

INTRODUCTION

The heterocyclic compounds constitute a very important rate of organic compounds of natural or synthetic origins. These compounds have a wide range of biological activity [1-6].

The derivations of the isoxazole form a particular and interesting class of heterocyclic compounds with five links, due to their availability and their biological and pharmaceutical activities which are important and efficient such as the anti-inflammatory, analgesic, antiviral, antioxidant, anticancer and antimicrobial activities [7-12].

Coxakievirus B3 (CVB3) is a class of viruses belonging to the Picornaviridae family, one of the oldest and most diverse families of viruses and more specifically to the group of enteroviruses [13]. These are genera that have been detected in both humans and animals [13].

Received on 29/5/2020 and Accepted for Publication on 17/4/2021.

Like most viruses, CVB3 is transmitted by two main routes: aerosol and fecal-oral routes [14-15] . They target all age groups and precisely new infants and children under 15 years of age [14-16]. As soon as reach their target, CVB3s cause infections of the central nervous system, respiratory infections, mucocutaneous, muscular, and digestive disorders [17-18].

These infections occur after a set of processes that make up the replication cycle that begins with the entry of the virus after attaching itself to specific receptors. A process is then triggered to release viral RNA into the cytoplasm of the cells. After the protein expression and replication is followed by the assembly, new virions are then released and transmitted through the bloodstream to reach different destinations [15].

CVB3s are small viruses (diameter about 30nm), non-enveloped, their capsid has icosahedral symmetry. It consists of 60 copies of four proteins VP1, VP2, VP3, and VP4. The VP1, VP2, and VP3 proteins are located at the

outer surface of the capsid while VP4 is located at the inner surface [19]. Several studies have confirmed that the VP1 protein has been involved in the process of binding to the receptors of the cells [20,21].

Inside VP1 and just below the bottom of the canyon, there is a hydrophobic pocket that contains fatty acid [22]. This pocket contributes to the stability of the virus and also represents the binding site for antiviral molecules. These molecules displace the lipid presented in the hydrophobic pouch, then blocking the decapsidation process [23] so expression and replication of new viruses.

The main objective of our study is to model the inhibitory effect of a series of 28 Pleconaril derived compounds (isoxazole-based molecules) against CVB3, using statistical tools Principal Component Analysis (PCA), Multiple Linear Regression (MLR), Multiple Non-Linear Regression (MNLR). The current study includes four main steps: the selection of the dataset, generation of molecular descriptors, construction and validation of predictive models and finally proposing new candidate compounds that are more active than the studied ones.

MATERIAL AND METHODS

1- Database

In the present study, we selected 28 Pleconaril derivatives (Table 1) with activity values reported in the literature [24]. The activity expressed by IC_{50} is defined as the concentration necessary to decrease the initial rate of effect of CVB3 with a percentage of inhibition 50%. For modelling purposes, the activity was expressed in pIC_{50} ($pIC_{50}= -\log_{10}(IC_{50})$). To build and validate our QSAR model, the data set was divided into two sets: 23 molecules constituted the training set and the remaining 5 molecules constituted the test set. The division of the data set was performed by random selection.

2- Data processing

A wide variety of molecular descriptors were computed using Gaussian 09W, Chemoffice2016, MarvinSketch and ACD/ChemSketch 2019.13 to predict the correlation between these descriptors and the activity of the molecules studied (Tables 2 and 3) [25-28].

The electronic descriptors were obtained after optimization of the studied molecules using the 6-31G(d,p) [29] base of the Lee-Yang-Parr three-parameter Becke function (B3LYP) [30], which is a kind of density functional theory DFT method [31].The calculation of the descriptors started by drawing the molecules in GaussView 5.0 [32], then opening these structures in Gaussian 09W, and then executing the optimization (the calculations).

3- Selection of descriptors

The QSAR model is supposed to be simple and understandable, the descriptors chosen must be meaningful and interpretable and it must have the fewest parameters to explain the activity [33]. In order to reduce their number of descriptors, the highly correlated descriptors are removed based on the results of the Principal Component Analysis (PCA). A random procedure was used for the selection of the compounds of the learning set and the test set.

4- Principal Component Analysis (PCA)

PCA is an essentially descriptive statistical method, which aims at extracting the maximum amount of information contained in the compounds from the dataset [34] . The results of the PCA analysis are used to select the MLR input data. Thus, initially, we eliminated all descriptors with small (non-significant, $r \leq 0.3$) correlations with the dependent variable (pIC_{50}). To reduce redundancy in our data matrix, the descriptors that are highly correlated ($r \geq 0.9$) and have a low value of the correlation coefficient concerning the dependent variable were excluded (Table 4).

5- Multiple Linear Regression (MLR)

It is one of the most transparent modeling methods whose the prediction is easily achievable [35]. The principal of MLR is based on the fact that the dependent variable Y (Activity/Property) depends linearly on the different variables (the descriptors), according to the relation:

$$Y = a_i + \sum_{i=0}^{n} a_i X_i$$

Y: the dependent variable; X_i: the independent variables; n: the number of independent variables; a_0: the constant of the model equation; a_i: the coefficients of descriptors in the model equation.

6- Multiple non-Linear Regression (MNLR)

Despite its transparency, linear regression has certain disadvantages. It is deficient in the detection of non-linear dependence (exponential, logarithmic, polynomial,). To make up for this deficiency, a non-linear regression (NLR) is performed.

In our work we used the second-degree polynomial model, based on the descriptors proposed by the linear model, as shown in the following equation:

$$Y = a_0 + \sum_{i=1}^{n} (a_i X_i + b_i X_i^2)$$

Y: is the dependent variable; X_i: the independent variables; n: the number of explanatory variables; a_0: the constant of the model equation; a_i and b_i: the coefficients of descriptors in the model equation.

7- Model Validation

The stability and robustness of the model must be verified before using it for prediction [36]. There are two types of validation: internal and external validation.

7.1- Internal validation

Internal validation techniques include cross-validation, which consists of extracting one or more compounds (LOO) from the training set and reconstructing the model to calculate the activity (or property) of these compounds [37]. The correlation coefficient between the predicted and observed activities of the extracted compounds must be greater than 0.5 $(Q^2_{CV} > 0,5)$ [38].

Another internal validation technique; Y-Randomization is performed to assert that the model is not due to chance. It consists of randomly disorganizing the activities/properties N times (e.g. 100) and the columns of the descriptors remain unchanged [39]. This gives N models with specific statistical characteristics. These N models must have low performance [39].

In general, these internal validation techniques allow the evaluation of the robustness and the stability of the QSAR/QSPR model parameters concerning the molecules of the training set. However, they do not demonstrate in any way the predictive power of the models, which is why an external validation of the model is required [38,40].

7.2- External validation

The external validation consists of predicting the activity/property of the compounds constituting the test set, this validation is characterized by the parameters R^2_{test}, R_{test}, and $Q^2_{cv(test)}$. Several recent studies have shown that these parameters are insufficient to verify the predictive power of the model [41,42]. Therefore other parameters should be calculated [42].

- Metrics r^2_m:

To better indicate the predictability of the model, the metrics introduced by Roy *et al* [43] can be exploited by r^2_m which is defined by the following equations:

$$\bar{r}^2_m = \frac{(r^2_m + r'^2_m)}{2}$$

$$\Delta r^2_m = |r^2_m - r'^2_m|$$

Where $r^2_m = r^2 \times (1 - \sqrt{(r^2 - r^2_0)})$ and

$$r'^2_m = r^2 \times (1 - \sqrt{(r^2 - r'^2_0)})$$

Tropsha and Golbraikh criteria:

Golbraikh and Tropsha [41] proposed a set of parameters to determine the external prediction of the QSAR model (Table 7).

8- Applicability domain

This part of the analysis is explicitly requested [44]. The domain of applicability (DA) defines the zone in which a compound can be predicted with confidence. Indeed, the QSAR/QSPR built model is not intended to be used outside this domain [45].

There are several methods for determining the scope of application of a QSPR/QSAR model, including the "leverage" method. This method is based on the variation of the standardized residuals of the dependent variable according to the leverages. If a compound has a leverage that exceeds the threshold h*=3(k+1)/N (where k is the number of descriptors and N the number of molecules constitutes the learning set), this compound is considered as an influential compound on the model developed [46].

Another recently developed approach [47] gave comparable results to the "leverage" method. The algorithm of this new approach is as follows:

-All the descriptors appearing in the developed model (for training and test set compounds) are standardized using the following formula:

$$S_{ki} = \frac{|X_{ki} - \bar{X}_i|}{\sigma_{ki}}$$

k: total number of compounds (Training set and test set).

i: total number of descriptors.

X_{ki}: Standardized descriptor i for compound k.

$\bar{X}_i$: Mean value of the descriptor Xi for the training set compounds only.

σ_{ki}: standard deviation of the descriptor Xi for the training set compounds only.

The above-calculation should run for all descriptor values present in the model.

- Thereafter, one needs to compute the S_{ik} values. If $S_{ik(max)} < 3$, then that compound is not an X-outlier (if in the training set) or is with in applicability domain (if in the test set).

- If $S_{ik(max)} > 3$, then one should compute $S_{ik(min)}$. If $S_{ik(min)} > 3$, then the compound is an X-outlier (if in the training set) or is not with in applicability domain (if in the test set).

- If $S_{ik(max)} > 3$ and $S_{ik(min)} \leq 3$, then one should compute $S_{new(k)}$ from the following equation:

$$S_{new(k)} = \bar{S}_k + 0.28 \times \sigma_{S_k}$$

$S_{new(k)}$: S_{new} value for the compound k.

$\bar{S}_k$: Mean of $S_{i(k)}$ values of the compound k.

σ_{S_k}: Standard deviation of $S_{i(k)}$ values of the compound k.

RESULTS AND DISCUSSION

A series of 28 molecules (Table 1) derived from pleconaril were studied to determine a quantitative relationship between the structure of these molecules and their antiviral activity (pIC_{50}) against coxsackievirus B3 (CVB3).

1- Multiple linear regression (MLR)

The resulting model is a linear combination of three descriptors: n (Refractive Index), E_T (Total Energy), and μ (Chemical potential).

The model equation is as follows:

$$pIC_{50} = 16,826 - 11,468.n + 1,728.10^{-4} E_T - 3,722.\mu \quad (1)$$

The activity values (pIC_{50}) calculated by equation (1) of the MLR model are given in Table 9.

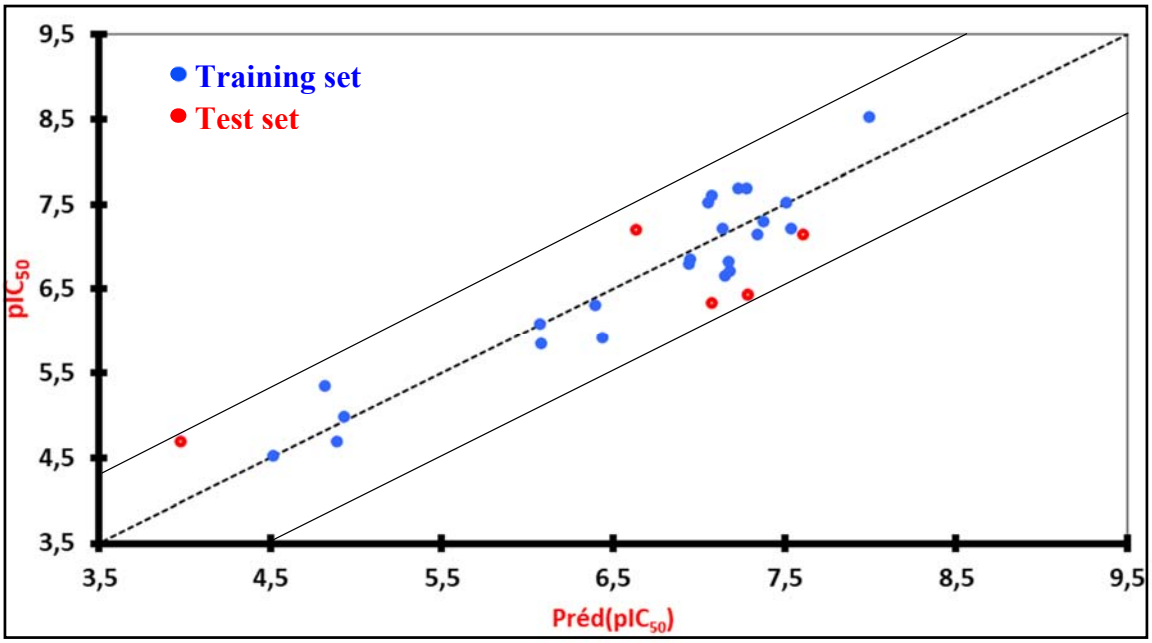


Figure 1: Graphical representation of calculated and observed activity for MLR

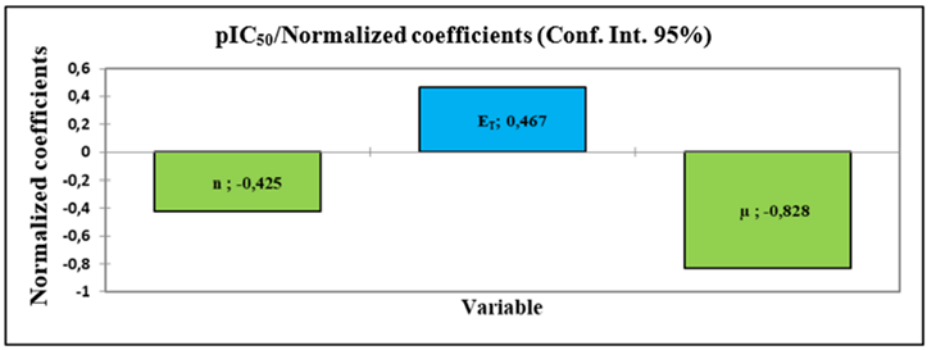


Figure 2: The relative contributions of the three descriptors for MLR

The relative contributions of the three descriptors are shown in Figure 2.

The Refractive index and chemical potential contributed negatively while total energy contributes positively.

The contribution of the refractive index and total energy were comparable, while the contribution of chemical potential was significant.

The value of the variance inflation factor (VIF) is calculated according to the equation:

$$VIF = \frac{1}{1 - r^2}$$

Where r^2 is the coefficient for determining the multiple regression equation between the model descriptors. If the VIF is equal to 1, no intercorrelation exists for each variable and if it is between 1 and 5, the constructed model is acceptable and

if it is greater than 10, the associated model is unstable and verification is required.

The VIF value of each descriptor used in the model (Table 5) is less than 5 and close to 1, which is an indication of the slight correlation of each descriptor to the other two.

The value of t for a descriptor is related to its statistical significance. High absolute values of t indicate that each regression coefficient is significantly larger than the associated standard deviation.

The probability P of t of a descriptor gives its statistical significance when is combined with other descriptors in the model (i.e., provides information on the interactions between descriptors). Descriptors with probability values of t below 0.05 are considered to be statically significant for a given model, i.e., their influence on the dependent variable (the response) is not due to chance.

The values of the probabilities of t for the three descriptors of the developed model (Table 6) were all lower than 0.001, indicating that the chosen descriptors are highly significant.

Table 7 contains the various calculated parameters and the criteria that must be met for the model to be acceptable.

The obtained results have shown a good correlation between the three selected descriptors and the studied activity, which is characterized by satisfactory statistical parameters. The good quality of the fit, the robustness, and the predictive power of the model has been confirmed by the high values of R^2, R^2_{adj}, $Q^2_{CV(LOO)}$, F, and low error values (MSE, RMSE).

The results obtained were very satisfactory and reflect the reliability of our model.

2- Multiple non-linear regression (MNLR)

We also used multiple nonlinear regressions to make up for the nonlinearity that is missing in multiple linear regression (MLR) and thus improve its results.

The method was applied to the same learning set and descriptors used in multiple linear regression.

The model equation obtained is as follows:

$$pIC_{50} = 195,37 - 180,41.n + 8,03.10^{-1}.E_T + 12,85.\mu + 54,52.n^2 + 6,46.10^{-2}.E_T^2 + 2,10.\mu^2 \quad (2)$$

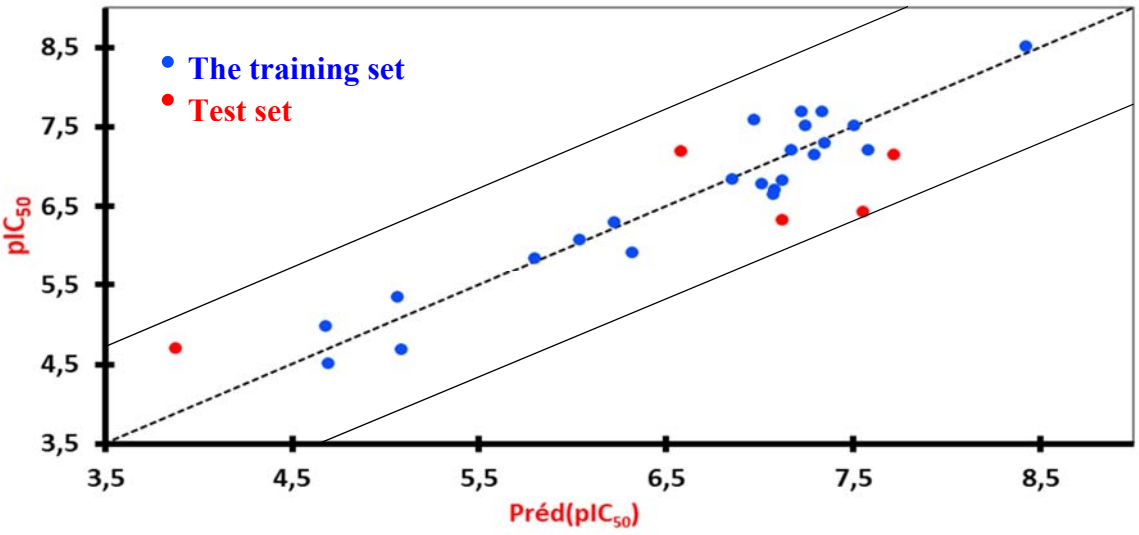


Figure 3: The correlation between observed and calculated activities for MNLR

The activity values (pIC_{50}) calculated by equation (2) of the MNLR model are given in Table 9.

The following table contains the calculated parameters for the model.

The overall obtained results are very satisfactory and reflected the reliability of the model obtained.

The good fit, robustness, and predictive power of the model have been confirmed by the high values of R^2, R^2_{adj}, $Q^2_{CV(LOO)}$, F, and low error values (MSE, RMSE).

3- Randomization test

To avoid chance correlations and validate the MLR and MNLR models built, randomization testing was applied, and 100 models were developed for both MLR and MNLR.

Figures 4 and 5 show the low values of $Q^2_{CVLOO(Rand)}$ and R^2_{Rand} obtained for the two models MLR and MNLR. The results confirm that the obtained models were not due to chance correlation.

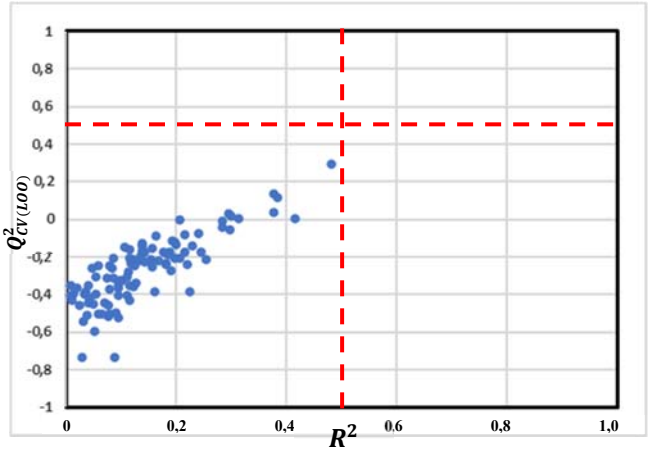


Figure 4:Y- Randomisation plot for MLR

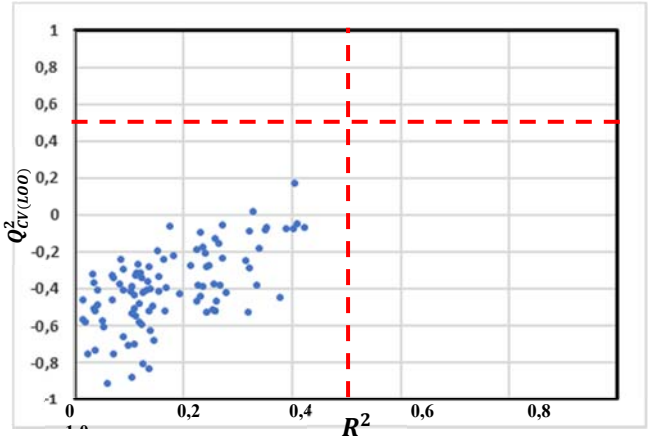


Figure 5:Y- Randomisation plot for MNLR

Figures 4 and 5 show the low values of $Q^2_{CVLOO(Rand)}$ and R^2_{Rand} obtained for the two models MLR and MNLR. The results confirm that the obtained models were not due to chance correlation.

4- Applicability domain

Both Figures 6 and 7 showed the standardized prediction errors as a function of the lever values (hi) and the S_{new} values of the molecules in the two training and test sets.

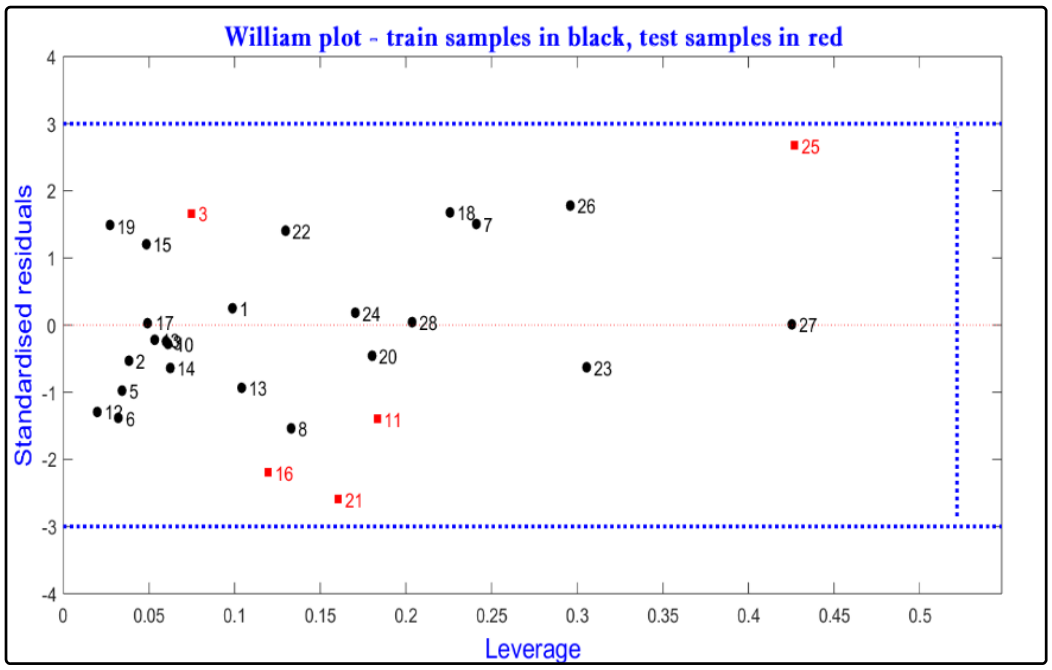


Figure 6: Williams plot to evaluate the applicability domain of MLR model (h*=0.522 and residual limits=±3).

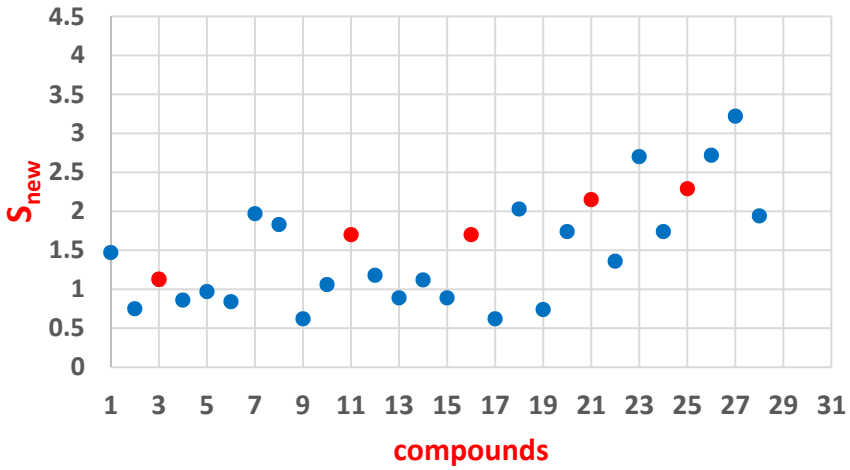


Figure 7: The S_{new} values of the test training and test sets

All the tailings are found to be in the range of (±3SD) (horizontal lines) and all compounds had levers $h_i < h^* = 0.522$ (Figure 6) and had a S_{new} of less than 3 (Figure 7), so there are no outliers for both training and testing sets, meaning that the model has a good predictive capability.

PERFORMANCE COMPARISON OF TWO MODELS MLR AND MNLR

A comparison of the results obtained by the two models was made (Figure 8).

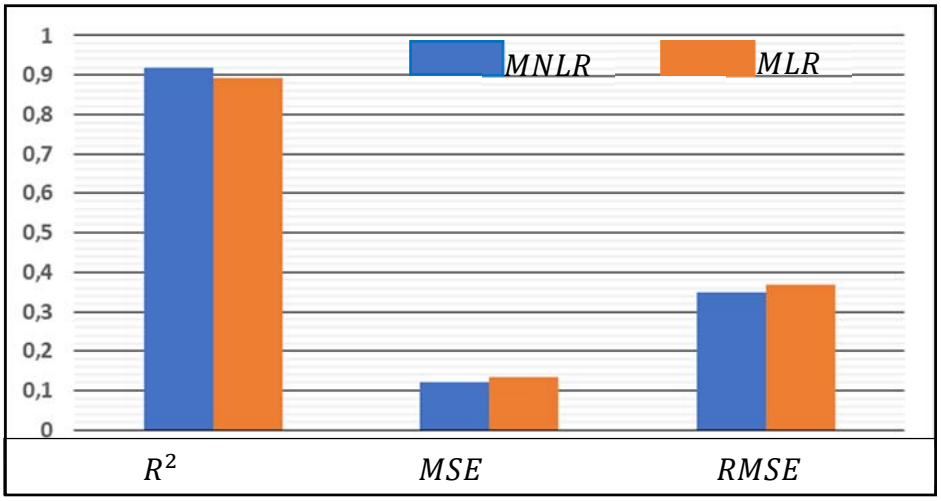


Figure 8: Performance comparison of the two models MLR and MNLR

The two established linear (MLR) and non-linear (MNLR) models generally performed well. Based on this comparison, it can be said that the non-linear model gave slightly better results than the linear model, this implied that the studied activity has certain non-linear characteristics and the introduction of MNLR improved the quality of prediction.

NEWLY DESIGNED COMPOUNDS

The main objective of the construction of QSAR models is the prediction of the activity of new candidate molecules that are more active than the studied molecules. This will be simple if we could understand the effect of each descriptor on the activity and the right choice of substitutes that vary these descriptors in the right direction.

According to the equation (1) and taking into account the sign of the values of the descriptors we could say that the activity increases with:

- The decrease in the refractive index (n)
- The increase in total energy (E_T)
- The decrease in chemical potential (μ)

The greater the absolute value of the t-test, the greater the influence of the descriptor. According to the obtained results (Table 7), the chemical potential μ has a stronger influence than the other two descriptors (n and E_T). The fact that the chemical potential is proportional to the energies of the HOMO and LUMO boundary orbitals ($\mu=(E_{HOMO}+E_{LUMO})/2$) allows us to conclude that a decrease of μ is equivalent to a decrease of the energy of the boundary orbitals, thus increasing the electrophile character of the molecule. Finally, to decrease the chemical potential, in other words, to increase the activity, it is necessary to introduce electron acceptor groups (deactivating groups such as NO_2, CN, CHO,...) which makes the molecule react as an electrophile.

Taking into account the above results, we introduced new substituents and then we calculated the activities of the proposed molecules. The structures of the proposed molecules and the pIC_{50} values theoretically predicted by

the MLR model as well as the corresponding levers are shown in the following table.

The proposed compounds **1p-4p** (Table10) showed higher activities than the most active compound 18 in the existing system (Table 1) and the corresponding lever values were found to be less than h*=0.522.

These results confirm that these compounds might be synthesized and evaluated as anti-CVB3 agents.

CONCLUSION

In this study, two QSAR models (MLR and MNLR) were constructed linking the inhibitory activity of a set of pleconaril-derived compounds (anti-CVB3) to three descriptors: refractive index (n), total energy (E_T) and chemical potential (μ). The set of the obtained results (statistical parameters, internal validation, external validation, and the field of applicability) showed that the two used models seem to be useful, although the advantage is to the non-linear model (MNLR). The strength of this study is the design of new candidate compounds that are more active than the studied compounds and that might be effective as antiviral agents against CVB3.

REFERENCES

1. Koubi Y., Moukhliss Y., Sbai A., El Masaoudy Y., Maghat H., EL-Mernissi R., Aziz Ajana M., Bouachrine M. and Lakhlifi T. Antimicrobial evaluation against Escherichia coli (MTCC 1652) by using 1,4-disubstituted 1,2,3-triazole and derivatives: QSAR study. Arabian Journal of Chemical and Environmental Research. 2019; 06: 57–69.
2. El Masaoudy Y., Aanouz I., Moukhliss Y., Koubi Y., Maghat H., Lakhlifi T. and Bouachrine M. 2D-QSAR study of the antimicrobial activity of a series of 5-(substituted benzaldehyde) thiazolidine-2,4-dione derivatives against Staphylococcus aureus by Multiple Linear Regression method. J. Mater. Environ. Sci. 2020, 11: 1914-1927.
3. Unni J., Mohammed A. A., Ashish D. W., Sameer K. V., Krishnan R. and Susobhan M.. Synthesis, Biological Evaluation and Molecular Modeling Studies of novel 2-[(2,4-dioxo-1,3-thiazolidin-3-yl)acetyl]-N-arylhydrazinecarbothioamides as Antibacterial Agents Targeting Alanine Racemase Enzyme. Jordan Journal of Pharmaceutical Sciences. 2020 ; 13, No. 3.
4. Santosh K., Manish B. and Prafulla Ch. Investigation of QSAR kNN-MFA on a Series of Substituted Chromen-2-one Derivatives as FXa Inhibitors. Jordan Journal of Pharmaceutical Sciences. 2018; 11, No. 3.
5. Qosay A. A., Nizar Al., Katreen B., Mohammad H., Ghazi Al. and Ammar A. Design, Synthesis and Biological Evaluation of Potential Novel Zinc Binders Targeting Human Glyoxalase-I; A Validated Target for Cancer Treatment. Jordan Journal of Pharmaceutical Sciences. 2018; 11, No. 1.
6. HAJJI H., Koubi Y., Moukhliss Y., EL-Mernissi R., Lakhlifi T, Aziz A. M. and Bouachrine M. Modeling and statistical study of series of derivatives 5- (1H-Indol-5-yl) -1, 3, 4-Thiadiazol-2-amines as potent inhibitors of PIM-1 kinase. RHAZES: Green and Applied Chemistry. 2020; 10: 88-102.
7. Caliskan B., Sinoplu E., Ibis K., Akhan G. E., Cetin A. R. and Banoglu E. Synthesis and cellular bioactivities of novel isoxazole derivatives incorporating an aryl piperazine moiety as anticancer agents. J.Enzym. Inhib. Med. Chem. 2018; 33: 1352–1361.
8. Fernandez J., Chicharro J., Bueno J.M. and Lorenzo M. Isoxazole mediated synthesis of 4-(1H)pyridones: Improved preparation of antimalarial candidate GSK932121. Chem. Commun. 2016; 52: 10190–10192.
9. Amgad G., Habeeb P. N., Praveen R. and Knaus, E. E. Design and Synthesis of 4,5-Diphenyl-4-isoxazolines: Novel Inhibitors of Cyclooxygenase-2 with Analgesic and Antiinflammatory Activity. J. Med. Chem. 2001; 44, 18: 2921–2927
10. Daidone G., Raffa D., Maggio B., Plescia F., Cutuli V.M. C., Mangano N.G. and Caruso A. Synthesis and

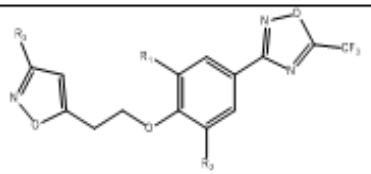
Antiproliferative Activity of Novel 3-(Indazol-3-yl)-quinazolin-4(3H)-one and 3-(Indazol-3-yl)-benzotriazin-4(3H)-one Derivatives. Arch. Pharm.1999; 50: 332.

11. Zaibo Y., Pei L. and Xiuhai G.. Novel Pyrazole-Hydrazone Derivatives Containing an Isoxazole Moiety: Design, Synthesis, and Antiviral Activity. Molecules. 2018 Jul; 23(7): 1798.
12. Afzal S. , Richie R. B., Kishor P. , Srinath N. , Venkata K. and Shahanaaz S. Antimicrobial, Antioxidant, and Anticancer Activities of Some Novel Isoxazole Ring Containing Chalcone and Dihydropyrazole Derivatives. Molecules .2020 ; 25 : 1047.
13. Tapparel C., Siegrist F., Petty T.J., and Kaiser L. Picornavirus and Enterovirus Diversity With Associated Human Diseases. Infect Genet Evol. 2013; 14: 282-293.
14. Andréoletti L., Lesay M., Deschildre A., and al. Differential detection of rhinoviruses and enteroviruses RNA sequences associated with classical immunofluorescence assay detection of respiratory virus antigens in nasopharyngeal swabs from infants with bronchiolitis. J Med Virol 2000 ; 61 : 341-6.
15. American Academy of Pediatrics. Enterovirus (nonpoliovirus) infections. In : Pickering LK, ed. Red book : report of the committee on infectious diseases. Elk Grove Village, IL : American Academy of Pediatrics, 2003 : 269-70.
16. Andréoletti L., Renois F., Jacques J. and Nicolas L. Human enteroviruses and respiratory infections. Med Sci (Paris) 2009; 25: 921–930.
17. Thibaut H.J., Lacroix C., De Palma A.M., Franco D., Decramer M. and Neyts J. Toward antiviral therapy/prophylaxis for rhinovirus-induced exacerbations of chronic obstructive pulmonary disease: challenges, opportunities, and strategies, Rev. Med. Virol. 2016; 26 : 21-33.
18. Palacios G. and Oberste MS. Enteroviruses as agents of emerging infectious diseases. J Neurovirol 2005 ; 11 : 424-33.
19. K Muckelbauer J., Kremer M., Minor I., Diana G., J Dutko F., Groarke J., and G Rossmann M. The structure of coxsackievirus B3 at 3.5 A° resolution.1995; 3:653-667.
20. Makarov V.A., Braun H., Richter M., Riabova O.B., Kirchmair J., Kazakova E.S., Seidel N., Wutzler P. and Schmidtke M. Pyrazolopyrimidines: Potent Inhibitors Targeting the Capsid of Rhino- and Enteroviruses. Chem Med Chem. 2015 ;10 :1629-1634.
21. Braun H., Kirchmair J., Williamson M.J., Makarov V.A., Riabova O.B., Glen R.C., Sauerbrei A. and Schmidtke M. Molecular mechanism of a specific capsid binder resistance caused by mutations outside the binding pocket. Antiviral Research. 2015 ;123 :138-145.
22. Liu Y., Hill M.G., Klose T., Chen Z., Watters K., Bochkov Y.A., Jiang W., Palmenberg A.C. and Rossmann M.G. Atomic structure of a rhinovirus C, a virus species linked to severe-childhood asthma. Proc. Natl. Acad. Sci. U.S.A. 2016; 113:8997- 9002.
23. Thibaut HJ., De Palma A.M. and Neyts J. Combating enterovirus replication: State-of-the-art on antiviral research. Biochemical Pharmacology. 2012; 83: 185-192.
24. Egorova A., Kazakova E., Jahn B., Ekins S., Makarov V. and Schmidtke M. Novel pleconaril derivatives: Influence of substituents in the isoxazole. European Journal of Medicinal Chemistry. 2020; 188:112007.
25. Frisch MJ., Trucks GW., Schlegel HB., Scuseria GE., Robb MA., Cheeseman JR. and al. Gaussian 09, Revision D.01. Gaussian, Inc., Wallingford CT, 2009.
26. Perkin Elmer Informatics. Chem Bio Office 2016.
27. MarvinSketch 20.2, ChemAxon Ltd.2020.
28. Inc., Toronto, ON, Canada. ACD LABS 10 Advanced Chemistry Development 2015.
29. Hariharan P. C. and Pople J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. Theoretica Chimica Acta.1973 ; 28: 213–222.
30. Lee C., Yang W. and Robert G. Parr. Development of the Colle-Salvetti Correlation-Energy Formula into a

Functional of the Electron Density. Physical Review. 1988 ; 37: 785–789.

31. Cramer C.J. Essentials of Computational Chemistry Theories and Models. John Wiley and Sons Ltd, 2004; 2nd edition, Chapter 8, pp 249-284.
32. Dennington R., Keith T. and Millam J. GaussView, version 5.0.9. Semichem Inc: Shawnee Mission, KS, 2009.
33. Crawley M. J. Statistics: An Introduction Using R, Wiley Chichester, UK, 2005; p 389.
34. Jolliffe I.T., Principal Component Analysis, New-York, NY: Springer, 2002 ; 2nd edition, Chapter 3, pp 29-61.
35. Roy K., Kar S. and Narayan Das R. Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment. Academic Press Boston. 2015, 1st edition, Chapter 6, p191.
36. Shao J. Linear model selection by cross-validation. Journal of the American statistical association. 1993; 88:486-494.
37. Zhang L., Zhu H., Oprea T.I., Golbraikh A. and Tropsha A. QSAR Modeling of the Blood-Brain Barrier Permeability for Diverse Organic Compounds. Pharm Res.2008 ; 25 :1902–1914.
38. Tropsha A., Gramatica P. and Gombar V.K. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. QSAR & Combinatorial Science. 2003 ; 22 :69-77.
39. He L. and Jurs P.C., Assessing the reliability of a QSAR models predictions. J Mol Graph Model. 2005 ;23 : 503-523.
40. Roy P.P., Paul S., Mitra I. and Roy K.. On Two Novel Parameters for Validation of predictive QSAR models. Molecules. 2009; 14:1660-1701.
41. Golbraikh A. and Tropsha A. Beware of q2 . J Mol Graph Model. 2002 ; 20 :269–276.
42. T. M. Martin, P. Harten, D. M. Young, E.N. Muratov, A. Golbraikh, H. Zhu, and Tropsha. Does Rational Selection of Training and Test Sets Improve the Outcome of QSAR Modeling. J Chem Inf Model. 2012 ; 52 :2570-2578.
43. Roy K., Chakraborty P., Mitra I., Ojha P. K., Kar S. and Das R. N. Some case studies on application of r^2_m metrics for judging Quality of Quantitative Structure-Activity Relationship Predictions: Emphasis on Scaling of Response Data. J Comput Chem. 2013 ; 34 :1071-1082.
44. Tunkel J., Mayo K., Austin C., Hickerson A. and Howard P. Environ. Sci. Techn.2005 ; 39 :2188- 2199.
45. Netzeva T.I., Worth A.P., Aldenberg T., Benigni R., Cronin M.T.D., Gramatica P., Jaworska J.S., Kahn S., Klopman G., Marchant C.A., Myatt G., Nikolova-Jeliazkova N., Patlewicz G.Y., Perkins R., Roberts D.W., Schultz T.W., Stanton D.T., Van De Sandt J.J.M., Tong W., Veith G., and Yang C. Current status of methods for defining the applicability domain of (quantitative) structure–activity-relationships. Alternatives to Laboratory Animals. 2005; 33:155-173.
46. Gramatica P. Principles of QSAR models validation: internal and external. QSAR Comb. Sci. 2007 ; 26 : 694-701.
47. Roy K., Supratik K. and Pravin A. On a simple approach for determining applicability domain of QSAR models. Chemometrics Intell Lab Syst. 2015 ; 145: 22-29.

Table1: Chemical structures and corresponding activities



N°	R_1	R_2	R_3	pIC_{50}
1	Me	Me	MeOCO	7.222
2	Me	Me	EtOCO	7.155
3[b]	H	Me	EtOCO	7.208
4	H	MeO	EtOCO	6.310
5	H	NO_2	EtOCO	6.830
6	H	F	EtOCO	6.664
7	H	CF_3	EtOCO	7.523
8	MeO	NO_2	EtOCO	5.925
9	Me	Me	$(Me)_2CHOCO$	6.721
10	Me	Me	Ph	6.854
11[b]	Me	Me	NH_2CO	7.155
12	Me	Me	MeNHCO	7.222
13	Me	Me	EtNHCO	7.301
14	Me	Me	$PhCH_2NHCO$	5.857
15	Me	Me	Me_2NCO	7.699
16[b]	H	H	Me_2NCO	6.339
17	H	Me	Me_2NCO	7.523
18[a]	H	NO_2	Me_2NCO	8.523
19	H	F	Me_2NCO	7.602
20	H	CF_3	Me_2NCO	6.796
21[b]	MeO	NO_2	Me_2NCO	6.438
22	H	CF_3	MeNHCO	7.699
23	H	NH_2	EtOCO	4.699
24	H	Pyrrolyl	EtOCO	4.991
25[b]	H	$NCHNMe_2$	EtOCO	4.700
26	H	NH_2	$(Me)_2NCO$	5.350
27	H	NHCOMe	$(Me)_2NCO$	4.521
28	H	Pyrrolyl	$(Me)_2NCO$	6.086

((a): more active (b): test set)

Table 2: Software used to calculate the selected descriptors and models construction and validation

• Gaussian 09W	Highest Occupied Molecular Orbital Energy $E_{HOMO}(eV)$; Lowest Unoccupied Molecular Orbital Energy $E_{LUMO}(eV)$; Hardness η(eV)=((E$_{LUMO}$-E$_{HOMO}$))/2; Dipole moment μ_D (Debye) ; Chimical Potential μ(eV) =(E$_{LUMO}$+E$_{HOMO}$)/2;Electrophilicity Index ω(eV) =χ^2/2η; Total Energy E$_T$(eV); Energy Gap between $EHOMO$ and $ELUMO$ values $EGap(eV)$; Activation Enrgy Ea(eV) ; Wavelength of Absorption Maximum λ_{max} (nm) ; Oscillation force $f_{(so)}$.
• ChemSketchACD Labs	Molecular Weight MW (cm^3); Molar Refractivity MR (cm^3) ; Molecular Volume MV (cm^3) ; Parachor Pc (cm^3); Index of Refraction (n); Surface Tensionγ (dyne/cm); Density D (g/cm^3) ; Polarisabilty αe (cm^3) .
• MarvinSketch 20.2	Partition Coefficient log P ; Polar Surface Area PSA (A°)^2.
• Chemoffice 2016	Solubility logS ; Winner Index IW.
• XLSTAT 2016.02.284	Models construction and validation.
• MATLAB 2018a	

Table 3: Calculated values of the different descriptors

N°	IC$_{50}$	pIC$_{50}$	MW	MR	MV	Pc	n	γ	D	αe	logP	PSA	logS	IW	E$_{HOMO}$	E$_{LUMO}$	E$_T$	ΔE	μ_D	E$_a$	λ_{max}	f$_{(SO)}$	η	μ	ω
1	0,06	7,222	425,359	96,55	321,6	816,7	1,512	41,5	1,322	38,27	4,88	100,48	-5,5396	3015,00	-6,327	-1,662	-41780,131	4,665	3,217	4,433	279,680	0,0002	2,333	-3,995	3,420
2	0,07	7,155	439,385	101,18	338,1	856,5	1,510	41,1	1,299	40,11	5,23	100,48	-5,8805	3342,00	-6,576	-1,709	-43919,92	4,867	2,714	1,528	2098,150	0,0039	2,434	-4,143	3,526
3	0,062	7,208	425,359	96,36	321,8	818,9	1,510	41,8	1,321	38,20	4,73	100,48	-5,6817	3150,00	-6,2	-1,608	-42850,122	4,592	3,439	4,281	289,640	0,0003	2,296	-3,904	3,319
4	0,49	6,310	441,538	98,21	329,6	837,9	1,507	41,7	1,339	38,93	4,04	109,71	-5,3006	3368,00	-6,198	-1,652	-44896,296	4,546	3,743	1,469	843,760	0,0006	2,273	-3,925	3,389
5	0,148	6,830	470,356	102,47	334,6	876,6	1,524	47,0	1,405	40,62	4,12	143,62	-5,1920	3868,00	-7,287	-2,674	-47344,426	4,613	4,330	1,723	719,440	0,0004	2,307	-4,981	5,377
6	0,217	6,664	429,322	91,52	309,8	788,3	1,502	41,9	1,385	36,28	4,38	100,48	-5,7121	3150,00	-6,396	-1,789	-44480,186	4,607	3,340	1,558	795,720	0,0007	2,304	-4,093	3,635
7	0,03	7,523	479,330	96,51	339,1	838,4	1,481	37,3	1,413	38,26	5,11	100,48	-6,4989	3810,00	-6,689	-1,917	-50951,229	4,772	3,911	1,823	680,300	0,0001	2,386	-4,303	3,880
8	1,189	5,925	486,355	104,75	341,4	893,4	1,525	46,8	1,424	41,52	3,99	152,85	-5,3009	4025,00	-6,821	-2,541	-50460,765	4,28	4,867	1,608	771,030	0,0004	2,140	-4,681	5,120

N°	IC_{50}	pIC_{50}	MW	MR	MV	Pc	n	γ	D	αe	logP	PSA	logS	IW	E_{HOMO}	E_{LUMO}	E_T	ΔE	μ_D	E_a	λ_{max}	$f_{(SO)}$	η	μ	ω
9	0,19	6,721	453,412	105,77	355,0	893,7	1,507	40,1	1,277	41,93	5,64	100,48	-6,2521	3671,00	-6,573	-1,707	-44989,91	4,866	2,866	0,848	1462,810	0,0018	2,433	-4,140	3,522
10	0,14	6,854	443,418	109,37	349,0	885,2	1,539	41,3	1,270	43,35	6,88	74,18	-7,3703	3640,00	-6,473	-1,69	-42936,741	4,783	4,159	1,237	1002,290	0,0000	2,392	-4,082	3,483
11	0,07	7,155	410,347	93,70	302,4	786,3	1,531	45,6	1,356	37,14	4,00	117,27	-4,9466	2718,00	-6,588	-1,722	-41239,937	4,866	2,242	1,526	812,470	0,0005	2,433	-4,155	3,548
12	0,06	7,222	424,374	98,40	324,1	825,0	1,519	41,9	1,309	39,00	4,24	103,28	-5,2015	3015,00	-6,581	-1,716	-42309,679	4,865	2,414	1,494	829,650	0,0000	2,433	-4,149	3,538
13	0,05	7,301	438,400	103,03	340,6	864,8	1,516	41,5	1,287	40,84	4,59	103,28	-5,5346	3342,00	-6,578	-1,714	-43379,616	4,864	2,552	1,486	834,510	0,0000	2,432	-4,146	3,534
14	1,39	5,857	500,470	122,88	384,8	997,1	1,551	45,0	1,300	48,71	5,96	103,28	-7,0567	5274,00	-6,575	-1,719	-48597,094	4,856	2,213	1,355	915,160	0,0000	2,428	-4,147	3,542
15	0,02	7,699	438,400	103,16	339,0	862,5	1,520	41,8	1,293	40,89	4,47	94,49	-5,3490	3314,00	-6,562	-1,701	-43379,152	4,861	2,924	1,223	1014,140	0,0003	2,431	-4,132	3,511
16	0,458	6,339	410,347	93,51	306,4	787,2	1,522	43,5	1,338	37,07	3,47	94,49	-4,9510	2936,00	-6,314	-1,654	-41239,364	4,66	3,233	1,535	808,000	0,0007	2,330	-3,984	3,406
17	0,03	7,523	440,373	100,19	330,4	843,9	1,518	42,5	1,332	39,71	3,28	103,72	-4,7691	3340,00	-6,853	-1,615	-44355,61	5,238	2,711	1,408	880,370	0,0005	2,619	-4,234	3,422
18	0,003	8,523	455,345	100,05	318,3	842,7	1,541	49,1	1,430	39,66	3,43	137,63	-4,9564	3559,00	-6,656	-2,44	-46803,733	4,216	4,419	1,596	776,670	0,0006	2,108	-4,548	4,906
19	0,025	7,602	428,338	93,50	310,7	794,4	1,513	42,7	1,378	37,06	3,62	94,49	-5,1805	3123,00	-6,383	-1,781	-43939,42	4,602	2,977	1,439	861,570	0,0006	2,301	-4,082	3,621
20	0,16	6,796	478,345	98,49	340,0	844,4	1,491	38,0	1,406	39,04	4,35	94,49	-5,9674	3780,00	-6,65	-1,908	-50410,465	4,742	3,321	1,570	789,670	0,0006	2,371	-4,279	3,861
21	0,365	6,438	485,371	106,73	342,3	899,4	1,535	47,6	1,417	42,31	3,23	146,86	-4,7693	3994,00	-6,566	-2,402	-49919,902	4,164	5,242	1,521	815,010	0,0000	2,082	-4,484	4,829
22	0,02	7,699	464,319	93,73	325,0	806,9	1,488	37,9	1,428	37,15	4,12	103,28	-5,8205	3454,00	-6,681	-1,912	-49340,772	4,769	3,374	1,623	763,730	0,0000	2,385	-4,297	3,871
23	20,02	4,699	426,347	95,77	307,8	807,0	1,534	47,2	1,384	37,96	3,45	126,50	-5,1453	3150,00	-5,532	-1,523	-43286,52	4,009	3,807	1,517	817,280	0,0005	2,005	-3,528	3,104

N°	IC_{50}	pIC_{50}	MW	MR	MV	Pc	n	γ	D	αe	logP	PSA	logS	IW	E_{HOMO}	E_{LUMO}	E_T	ΔE	μ_D	E_a	λ_{max}	$f_{(SO)}$	η	μ	ω
24	10,22	4,991	476,405	113,22	333,6	868,2	1,594	45,8	1,420	44,88	5,14	105,41	-6,8028	4091,00	-5,911	-1,927	-47466,405	3,984	2,383	1,069	1159,660	0,002	1,992	-3,919	3,855
25	19,95	4, 700	481,425	113,85	349,0	889,1	1,565	42,0	1,370	45,13	4,08	116,08	-5,7915	4181,00	-5,653	-1,538	-47968,862	4,115	4,914	1,291	960,690	0,0009	2,058	-3,596	3,142
26	4,47	5,350	425,362	97,74	308,7	813,1	1,545	48,0	1,377	38,75	2,69	120,51	-4,6137	3123,00	-5,521	-1,515	-42745,752	4,006	4,448	1,111	1116,150	0,0000	2,003	-3,518	3,089
27	30,14	4,521	467,399	107,78	339,5	891,0	1,647	47,3	1,376	42,72	2,80	123,59	-4,8301	3838,00	-6,072	-1,817	-46899,782	4,255	3,225	1,045	1186,740	0,0002	2,128	-3,945	3,657
28	0,82	6,086	475,421	115,61	337,0	875,6	1,601	45,5	1,410	45,83	4,39	99,42	-6,2713	4060,00	-6,662	-1,782	-46925,713	4,88	3,200	1,063	1166,800	0,0020	2,440	-4,222	3,6527

Table 4: Matrix of correlation

	pIC_{50}	MW	MR	MV	Pc	n	γ	D	αe	logP	PSA	logS	IW	E_{HOMO}	E_{LUMO}	E_T	ΔE	μ_D	Ea	λmax	$f_{(SO)}$	η	μ	ω
pIC_{50}	1																							
MW	-0,276	1																						
MR	**-0,451**	**0,700**	1																					
MV	-0,156	**0,751**	**0,816**	1																				
Pc	-0,299	**0,769**	**0,905**	**0,940**	1																			
n	**-0,657**	0,287	**0,629**	0,180	**0,397**	1																		
γ	**-0,410**	0,069	0,259	-0,16	0,182	**0,622**	1																	
D	-0,168	0,429	-0,10	-0,27	-0,163	0,160	0,334	1																
αe	-0,451	0,700	1,000	0,816	0,905	0,629	0,259	-0,10	1															
logP	0,175	0,182	0,378	0,564	0,395	-0,19	-0,48	-0,50	0,378	1														
PSA	-0,252	0,323	0,056	-0,06	0,176	0,216	0,709	0,565	0,056	-0,56	1													
logS	0,052	-0,41	-0,52	-0,60	-0,457	-0,012	0,427	0,219	-0,524	-0,90	0,540	1												
IW	-0,368	0,922	0,857	0,849	0,897	0,387	0,150	0,185	0,857	0,322	0,195	-0,53	1											
E_{HOMO}	-0,666	-0,23	0,014	-0,26	-0,194	0,356	0,199	-0,00	0,014	-0,23	-0,05	0,080	-0,13	1										
E_{LUMO}	-0,192	**-0,508**	-0,08	-0,12	-0,253	0,005	-0,38	-0,58	-0,086	0,143	-0,69	-0,13	-0,32	0,585	1									
E_T	0,146	**-0,927**	**-0,39**	**-0,526**	**-0,508**	-0,073	0,033	**-0,633**	**-0,391**	-0,01	**-0,380**	0,268	**-0,743**	0,272	**0,584**	1								
ΔE	**0,624**	-0,184	-0,09	0,200	-0,001	**-0,423**	**-0,592**	**-0,529**	-0,09	**0,411**	**-0,575**	-0,22	-0,13	**-0,662**	0,222	0,214	1							
μ_D	-0,193	0,314	0,013	-0,01	0,082	0,005	0,298	**0,472**	0,013	-0,29	**0,575**	0,219	0,162	0,126	**-0,471**	**-0,400**	**-0,587**	1						
Ea	0,285	-0,27	-0,33	-0,22	-0,280	-0,316	-0,187	-0,10	-0,33	0,107	-0,06	0,056	-0,29	-0,007	0,059	0,232	0,063	0,027	1					
λ_{max}	-0,214	0,109	0,334	0,276	0,283	0,276	0,032	-0,21	0,334	0,156	-0,11	-0,17	0,163	0,066	0,146	0,004	0,056	-0,21	**0,630**	1				
$f_{(SO)}$	-0,091	0,034	0,183	0,071	0,052	0,132	-0,05	-0,06	0,183	0,228	-0,14	-0,23	0,047	0,011	0,088	0,033	0,068	-0,25	-0,19	**0,760**	1			
η	**0,624**	-0,184	-0,09	0,200	-0,001	**-0,423**	**-0,592**	**-0,529**	-0,09	**0,411**	**-0,575**	-0,22	-0,13	**-0,662**	0,222	0,214	**1,000**	**-0,587**	0,063	0,056	0,068	1		
μ	**-0,516**	**0,398**	-0,03	-0,22	0,246	0,228	-0,06	-0,287	-0,033	-0,07	-0,37	-0,01	-0,24	**0,919**	**0,858**	**0,457**	-0,311	-0,15	0,024	0,113	0,050	-0,311	1	
ω	0,222	**0,496**	0,086	0,134	0,260	-0,024	0,372	**0,551**	0,086	-0,13	**0,683**	0,143	0,319	**-0,619**	**-0,998**	**-0,568**	-0,179	**0,456**	0,053	-0,14	-0,087	-0,179	**-0,879**	1

Table 5: Multicollinearity statistics

	n	E_T	μ
Tolerance	0.906	0.739	0.688
VIF	1.104	1.353	1.454

Table 6: Caracteristics of the MLR model parameters

Source	Value	Standard Error	t	Pr > $\|t\|$
Constant	16.8258	3.845	4.376	0.000
n	-11.4683	2.126	-5.394	< 0.001
E_T	0.0002	0.000	5.349	< 0.001
μ	-3.7219	0.406	-9.156	< 0.001

Table 7: Comparison of model parameters (MLR) with Golbraikh and Tropsha criteria

	Parameter	Model score	Threshold
Fitting criteria	$R^2 = 1 - \frac{\sum(Y_{obs} - Y_{calc})^2}{\sum(Y_{obs} - \overline{Y}_{obs})^2}$	0.893	> 0.6
	$R^2_{adj} = \frac{(N-1).R^2 - k}{N-k-1}$	0.876	> 0.6
	$MSE = \frac{\sum(Y_{obs} - Y_{calc})^2}{N}$	0.135	A low value
	$RMSE = \sqrt{\frac{\sum(Y_{obs} - Y_{calc})^2}{N}}$	0.367	A low value
	$F = \frac{\sum(Y_{calc} - \overline{Y}_{calc})^2.N - k - 1}{\sum(Y_{obs} - Y_{calc})^2.k}$	52.972	A high value
Internal validation	$Q^2_{CV} = 1 - \frac{\sum(Y_{calc} - Y_{obs})^2}{\sum(Y_{obs} - \overline{Y}_{obs})^2}$	0.837	> 0.5
	R_{Rand} (Average of the 100 $R_{Rand(i)}$)	0.149	$< R$
	R^2_{Rand} (Average of the 100 $R^2_{Rand(i)}$)	0.360	$< R^2$
	$Q^2_{CV\,LMO(Rand)}$ (Average of the 100 $Q^2_{CV\,LMO(Rand)(i)}$)	-0.254	$< Q^2_{CV}$
	$cR^2_p = R.\sqrt{(R^2 - R^2_{Rand})}$	0.828	> 0.5
	$\bar{r}^2_{m(CV\,LMO)} = \frac{\left\|r^2_m + r'^2_m\right\|}{2}$	0.774	> 0.5
	$\Delta r^2_{m(CV\,LMO)} = \left\|r^2_m - r'^2_m\right\|$	0.066	< 0.2
External validation	$R^2_{test} = 1 - \frac{\sum(Y_{calc(test)} - Y_{obs(test)})^2}{\sum(Y_{obs(test)} - \overline{Y}_{obs(train)})^2}$	0.778	> 0.5

	Parameter	Model score	Threshold
Fitting criteria	$R^2 = 1 - \frac{\sum(Y_{obs} - Y_{calc})^2}{\sum(Y_{obs} - \overline{Y}_{obs})^2}$	0.893	> 0.6
	$R^2_{adj} = \frac{(N-1).R^2 - k}{N-k-1}$	0.876	> 0.6
	$MSE = \frac{\sum(Y_{obs} - Y_{calc})^2}{N}$	0.135	A low value
	$RMSE = \sqrt{\frac{\sum(Y_{obs} - Y_{calc})^2}{N}}$	0.367	A low value
	$F = \frac{\sum(Y_{calc} - \overline{Y}_{calc})^2.N - k - 1}{\sum(Y_{obs} - Y_{calc})^2.k}$	52.972	A high value
	$\overline{r}^2_{m(\text{test})} = \frac{\left\|r^2_m + r'^2_m\right\|}{2}$	0.594	> 0.5
	$\Delta r^2_{m(\text{test})} = \left\|r^2_m - r'^2_m\right\|$	0.193	< 0.2
	$\Delta r^2_{0(\text{test})} = \left\|r^2_0 - r'^2_0\right\|$	0.239	< 0.3
	$\frac{(r^2 - r^2_0)}{r^2}$	0.345	< 0.1
	$\frac{(r^2 - r'^2_0)}{r^2}$	0.038	< 0.1
	$K = \frac{\sum(Y_{obs}.Y_{calc})^2}{\sum(Y_{calc})^2}$	0.963	$0.85 \leq K \leq 1.15$
	$K' = \frac{\sum(Y_{obs}.Y_{calc})^2}{\sum(Y_{obs})^2}$	1.028	$0.85 \leq K' \leq 1.15$

Table 8: Statistical parameters of the MNLR model

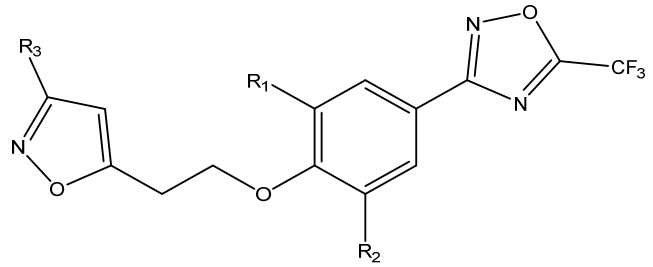
Model parameters							Internal and external validation							
N	R	R^2	R^2_{adj}	MSE	RMSE	F	$Q^2_{CV\,(LOO)}$	$Q^2_{CV\,LMO\,(Rand)}$	R_{Rand}	R^2_{Rand}	cR^2_p	R^2_{test}	K	K'
23	0.958	0.918	0.905	0.122	0.349	67.511	0.784	0.397	0.397	0.176	0.836	0.737	0.875	1.097

Table 9: Comparison of the observed values with those calculated by MLR and MNLR

N°	$pIC_{50(observed)}$	$pIC_{50(predicted)}$		N°	$pIC_{50(observed)}$	$pIC_{50(predicted)}$	
		MLR	MNLR			MLR	MNLR
1	7.222	7.137	7.166	**15**	7.699	7.279	7.220
2	7.155	7.341	7.292	**16**[b]	6.339	7.075	7.117
3[b]	7.208	6.636	6.581	**17**	7.523	7.513	7.501
4	6.310	6.395	6.221	**18**[a]	8.523	7.995	8.423
5	6.830	7.173	7.119	**19**	7.602	7.076	6.970
6	6.664	7.150	7.069	**20**	6.796	6.944	7.013
7	7.523	7.054	7.240	**21**[b]	6.438	7.287	7.551
8	5.925	6.437	6.314	**22**	7.699	7.230	7.330
9	6.721	7.179	7.082	**23**	4.699	4.886	5.085
10	6.854	6.951	6.849	**24**	4.991	4.931	4.681
11[b]	7.155	7.608	7.719	**25**[b]	4.700	3.975	3.873
12	7.222	7.538	7.579	**26**	5.350	4.816	5.061
13	7.301	7.377	7.343	**27**	4.521	4.518	4.689
14	5.857	6.078	5.794	**28**	6.086	6.072	6.033

((a): more active (b): test set)

Table 10: Values of descriptors, calculated anti- CVB3 activity pIC_{50} and leverages (h) for the new compounds



N°	R_1	R_2	R_3	n	E_T (eV)	μ (eV)	pIC_{50}	h
18	H	NO_2	$(Me)_2NCO$	1.541	-46803.733	-4.548	8.523	0.2260
1p	NO_2	NO_2	$(Me)_2NCO$	1.558	-52367.980	-5.033	8.644	0.2103
2p	NO_2	NO_2	NO_2	1.575	-51202.249	-5.637	10.898	0.4300
3p	CN	CN	CN	1.578	-42038.746	-5.265	11.063	0.4398
4p	CHO	CHO	CHO	1.581	-43759.955	-4.956	9.580	0.1939

دراسة مشتقات جديدة للبْلِكُونَارِيلْ (مركبات ايزوجزازول) المضاد للفيروسات كُوكْسَاكِي ب 3 باعتماد المقاربة "العلاقة الكمية نشاط – بنية ثنائية البعد"

. يونس مخلص [1]، خليل الخطابي [1]، ياسين كوبي [1]، حميد مغاط [1*]، عبد الواحد السباعي [1]، محمد بوعشرين [2]، الطاهر لخليفي [1]

1 مختبر الأنواع الطبيعية والجزيئات الكيميائية، كلية العلوم، جامعة مولاي إسماعيل، مكناس، المغرب.

2 المدرسة العليا للتكنولوجيا خنيفرة، جامعة السلطان مولاي سليمان، بني ملال، المغرب.

ملخص

نظرا لأعراضها الحادة، أصبح البحث عن دواء مضاد للفِيرُوسَات كُوكْسَاكِي ب 3 (CVB3) ضرورة ملحة ومستعجلة. 28 مركب مشتقة من البْلِكُونَارِيلْ (مضاد CVB3) كانت موضوع الدراسة باعتماد المقاربة " العلاقة الكمية نشاط – بنية ثنائية البعد" من أجل بناء نموذج يمكننا من التنبؤ بمركبات جديدة أكثر نشاطا من المركبات المدروسة. بتوظيف الانحدار الخطي المتعدد والانحدار غير الخطي المتعدد تم الحصول على نموذجين بنتائج جيدة. استثمار هذين النموذجين مكننا من اقتراح مركبات نشاطها أكبر مقارنة مع المركبات المدروسة. تعتبر هذه النتائج مرجعا مهما لأعمال تجريبية مقبلة.

الكلمات الدالة: العلاقة الكمية نشاط – بنية ثنائية البعد، الانحدار الخطي المتعدد، الانحدار غير الخطي المتعدد، فيروس كُوكْسَاكِي ب 3، البْلِكُونَارِيلْ، إيزُوكْسَازُولْ، مضاد فيروسات.

تاريخ استلام البحث 2020/5/29 وتاريخ قبوله للنشر 2021/4/17.