Bioorganic & Medicinal Chemistry 17 (2009) 1654–1662

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

# Application quantum and physico chemical molecular descriptors utilizing principal components to study mode of anticoagulant activity of pyridyl chromen-2-one derivatives

M. S. Bhatia [a,*], K. B. Ingale [a], P. B. Choudhari [a], N. M. Bhatia [a], R. L. Sawant [b]

[a] Department of Pharmaceutical Chemistry, Bharati Vidyapeeth College of Pharmacy, Near Chitranagri, Kolhapur-416013 (MS), India
[b] Department of Pharmaceutical Chemistry, PDVVF's College of Pharmacy, Ahemadnagar – 414111 (MS), India

ARTICLE INFO

*Article history:*
Received 15 October 2008
Revised 20 December 2008
Accepted 23 December 2008
Available online 29 December 2008

*Keywords:*
Prothrombin time
Activated partial thromboplastin time
Chromen-2-one
QSAR

ABSTRACT

Factors II, V, VII and Xa have materialized as a key enzymes for the intervention of blood coagulation cascade and for the development of new anti thrombotic agents. The combined density functional quantum mechanical/molecular mechanical (QM/MM) approach has been used to access inhibition of prothrombin and thrombin production. The biological activities of coumarin derivatives as clotting factor inhibitors was quantitatively analyzed in terms of physicochemical parameters utilizing the principal component analysis. Structural requirements for maximal potency were derived from the results of a quantitative structure activity relationship analysis.



## 1. Introduction

Anticoagulant agents are used to prevent the extension of blood clots in phlebitis or deep vein thrombosis and as a prophylactic agent in patients that have mechanical heart valves (life-long therapy). Fatal damage is often caused by thrombotic events. Hence, numerous efforts have been made to synthesize anti thrombotics, such as antiplatelet, anticoagulant, and thrombolytic agents. Prothrombin and thrombin plays a critical role in thrombosis, since they not only convert fibrinogen to fibrin for clot formation but also strongly induces platelet aggregation. The human factor Xa is a trypsin-like serine protease that serves a critical role in blood coagulation events.[1] The factor Xa combines with factor Va and platelet phospholipid, as a result a prothrombinase complex is formed. This prothrombinase complex, in turn, initiates the splitting of prothrombin to form thrombin thereby setting in motion the final clotting process. Since it was calculated that a molecule of factor Xa could generate 138 molecules of thrombin,[2] inhibition of factor Xa may be more efficient than inactivation of thrombin in interrupting the blood coagulation system. Coumarin ring containing products inhibit coagulation by interfering with the incorporation of vitamin K into vitamin-K dependent clotting factors (Factors II, VII, IX, and X). Coumarin products do not dissolve clots but prevent clots from forming (prophylaxis) and getting larger.[3,4] The main side effect associated with the use of coumarin derivatives is hemorrhaging. The degree of anticoagulation is measured by a blood sample and expressed as the International Normalized Ratio (INR). The INR is the international standard. The therapeutic INR is generally between 2 and 3.5 which correlates with a 30–50% inhibition of vitamin-K dependent clotting factors. Ideally, the patient should have his/her INR checked every 4–6 weeks while on this medication.

The main objective of the present report was to use quantum and other physicochemical molecular descriptors (MDs) to generate predictive linear discriminant analysis (LDA)-assisted QSAR models, enabling the selection of novel coumarin derivatives with good anticoagulant activity. The data sets were generated by synthesizing various coumarin derivatives and testing them for elongation of prothrombin time (PT) and activate partial thrombin time (APTT). Prothrombin time was measured by utilizing rabbit brain thromboplastin reagent and measuring the time required for the formation of clot. In order to determine the Activated partial thromboplastin time brain extract was incubated with Kaolin (Factor XII) and the clotting time in seconds was measured after addition of calcium ions.

The synthetic availability and interesting structural features of chromen-2-one a ubiquitous structural unit in many biologically active compounds prompted us to carry out an extension toward the preparation of 3-(6-aminopyridin-3-yl)-4-methyl-2H-chromen-2-one derivatives (Fig. 1) for synthesis and QSAR studies of 4-methyl-3-(6-{[arylmethylene] amino} pyridine-3-yl)-2H-chro-

* Corresponding author. Tel.: +91 9822172940.
*E-mail address:* manish_bhatia13@yahoo.co.in (M.S. Bhatia).
*URL:* http://drmanishbhatia.com (M.S. Bhatia).

0968-0896/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2008.12.055

men-2-one as anti-coagulant agents. The compounds obtained by substitutions are shown in Table 1.

## 2. Results

### 2.1. Analysis of principal components to detect structure in the relationships between variables (MDs)

To conduct the analysis of information contents and linear independence of the MDs calculated in this study, we performed a factorial analysis.[5–11] Factorial loadings from the principal component analysis, after a varimax normalized rotation of the first twelve factors, are shown in Table 2. Furthermore, the results of the factorial analysis performed here are summarized in Table 3. The twelve principal factors explain approximately 99.5% of the variance. The first factor explains 42.1% of the variance in the MDs studied. The addition of the second factor increases to 55.9% the explained variance, and the addition of the third factor allow 70.7% of the index variance to be accounted for. The other factors are summarized in Table 3. Most of the steric MDs are strongly loaded into factor 1 (F1), such as Molwt, Vol, SMR, Idaverage, MIz, RGYR, Kappa1 and kappa2. Some of these MDs, in addition to hydrophobic information, also have steric and topological information. For instance, distTop is also robustly loaded (loading > 0.70, but with opposite relation) into this factor. Thus, these 'F1-MDs' produce much redundancy and overlapping between them, and their relations are rather complex. This factor F1 can be taken into account as a hydrophobic-steric factor. The second factor (F2) also shows an electronic dimension (DM, LUMO, QMDM) along with hydrophobic dimensions (H-acc) strongly loaded. This result showed that electronic and hydrophobic MDs have a parallel relation. The third factor (F3) appears to be most significant for the other hydrophobic features (H-don, SAaverage). The most of the local charge MDs are strongly loaded in factors F4 (HF), F5 [IP, HOMO], F6 (QMx), 7 [zdipole], F8 (QMy), showing a rather complex relation. Finally, in factor 9, 10, 11 no MDs are strongly loaded, but the highest loadings were found to be of electronic parameters (not given in Table 1) therefore, these dimension are electronic factors. As stated in Material and Methods section, these MDs with high loadings in the same factor are interrelated, while no correlation exists between those MDs having nonzero loadings only in different factors. Thus, we can say that the MDs included in each dimension contain structural information, not contained in any other MD incorporated in others factors.

### 2.2. Molecular diversity study and design of the training and test sets using CA

The quality of a classification (or any QSAR) model depends on the quality of the selected dataset. One of the most critical aspects of constructing the training set is to warrant enough molecular diversity for it. Taking this into account, we selected a database of 64 compounds having a great degree of structural variability. To demonstrate the structural diversity of these datasets, we performed hierarchical CAs of the series.[12–16] The dendrograms are given in Figures 2 and 3, using the Euclidean distance (*X*-axis) and the complete linkage (*Y*-axis), illustrating the results of the k-NNCA developed into active and inactive sets, correspondingly. As can be seen in the binary tree, there is a great number of different subsets, which prove the molecular variability of the selected chemicals in these databases. Because of the difficulty in evaluating the output dendrogram, other kind of CAs is usually performed. Therefore, to split the whole group into two data sets (training and test ones), we perform a k-MCA. The main idea of this procedure consists in making a partition of anticoagulant series of chemicals into several statistically representative classes of compounds.[12–16] Afterward, the selection of the training and prediction sets was performed by taking, in a random way, compounds belonging to each cluster. From these 64 compounds, 52 were chosen at random to form the training set. By using a training data set including maximum possible structural variations makes it possible to explore all structures likely to have the targeted activity. Thus the discovery of lead compounds with good anticoagulant activity, not only with determined action modes but also with novel molecular scaffolds becomes more likely. The remaining series of 12 compounds was prepared as the test sets for the external validation of the models. These compounds were never used in the development of the classification models. The compounds used in training Set and test set are shown in Table 1.

### 2.3. Development of the discriminant function

Although the number of existing statistical methods of getting classification functions is relatively extensive, we select LDA given the simplicity of the method.[17] The use of LDA in drug design has been extensively reported by different authors. Therefore, LDA was also the technique used in the generation of discriminant functions in the current work. Making use of the LDA technique implemented in the STATISTICA software,[18] the following linear model was obtained, in which the quantum and physicochemical MDs were used as independent variables:

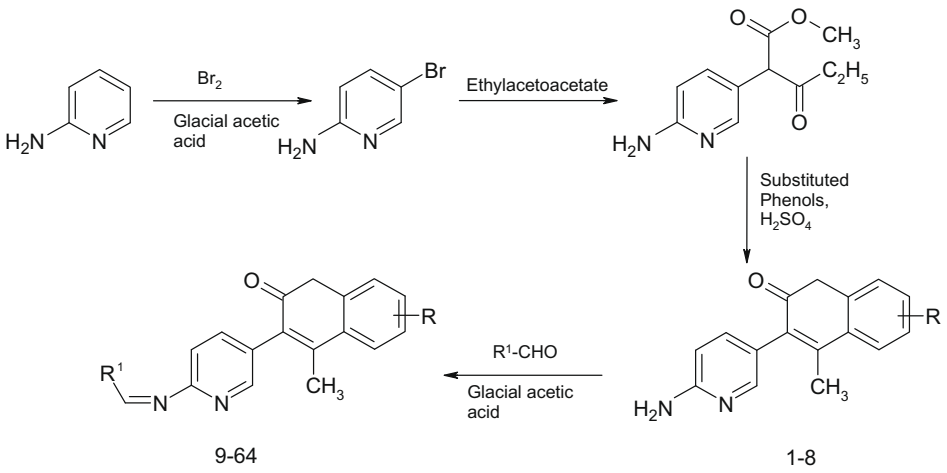


**Figure 1.** Figure showing the synthesis of the pyridyl chromen-2-one derivatives.

**Table 1**
Table showing compounds used in this study for the anticoagulant activity measurements

| Sr. no | R | X | R[1] |
|---|---|---|---|
| 1 | — | H | — |
| 2 | — | $NO_2$ | — |
| 3 | 7-OH | H | — |
| 4 | 7-OH | $NO_2$ | — |
| 5 | 6-$NH_2$ | H | — |
| 6 | 6-$NH_2$ | $NO_2$ | — |
| 7 | 8-$NO_2$ | H | — |
| 8 | 8-$NO_2$ | $NO_2$ | — |
| 9 | — | H | 2-Hydroxyphenyl |
| 10 | — | H | 4-Chlorophenyl |
| 11 | — | H | 4-Methoxyphenyl |
| 12 | — | H | Phenyl |
| 13 | — | H | H |
| 14 | — | H | *NN*-Dimethyl aminophenyl |
| 15 | — | H | 3,4,5-Trimethoxy phenyl |
| 16 | — | $NO_2$ | 2-Hydroxyphenyl |
| 17 | — | $NO_2$ | 4-Chlorophenyl |
| 18 | — | $NO_2$ | 4-Methoxyphenyl |
| 19 | — | $NO_2$ | Phenyl |
| 20 | — | $NO_2$ | H |
| 21 | — | $NO_2$ | *NN*-Dimethyl aminophenyl |
| 22 | — | $NO_2$ | 3,4,5-Trimethoxy phenyl |
| 23 | 7-OH | H | 2-Hydroxyphenyl |
| 24 | 7-OH | H | 4-Chlorophenyl |
| 25 | 7-OH | H | 4-Methoxyphenyl |
| 26 | 7-OH | H | Phenyl |
| 27 | 7-OH | H | H |
| 28 | 7-OH | H | *NN*-Dimethyl aminophenyl |
| 29 | 7-OH | H | 3,4,5-Trimethoxy phenyl |
| 30 | 7-OH | $NO_2$ | 2-Hydroxyphenyl |
| 31 | 7-OH | $NO_2$ | 4-Chlorophenyl |
| 32 | 7-OH | $NO_2$ | 4-Methoxyphenyl |
| 33 | 7-OH | $NO_2$ | Phenyl |
| 34 | 7-OH | $NO_2$ | H |
| 35 | 7-OH | $NO_2$ | *NN*-Dimethyl aminophenyl |
| 36 | 7-OH | $NO_2$ | 3,4,5-Trimethoxy phenyl |
| 37 | 6-$NH_2$ | H | 2-Hydroxyphenyl |
| 38 | 6-$NH_2$ | H | 4-Chlorophenyl |
| 39 | 6-$NH_2$ | H | 4-Methoxyphenyl |
| 40 | 6-$NH_2$ | H | Phenyl |
| 41 | 6-$NH_2$ | H | H |
| 42 | 6-$NH_2$ | H | *NN*-Dimethyl aminophenyl |
| 43 | 6-$NH_2$ | H | 3,4,5-Trimethoxy phenyl |
| 44 | 6-$NH_2$ | $NO_2$ | 2-Hydroxyphenyl |
| 45 | 6-$NH_2$ | $NO_2$ | 4-Chlorophenyl |
| 46 | 6-$NH_2$ | $NO_2$ | 4-Methoxyphenyl |
| 47 | 6-$NH_2$ | $NO_2$ | Phenyl |
| 48 | 6-$NH_2$ | $NO_2$ | H |
| 49 | 6-$NH_2$ | $NO_2$ | *NN*-Dimethyl aminophenyl |
| 50 | 6-$NH_2$ | $NO_2$ | 3,4,5-Trimethoxy phenyl |
| 51 | 8-$NO_2$ | H | 2-Hydroxyphenyl |
| 52 | 8-$NO_2$ | H | 4-Chlorophenyl |
| 53 | 8-$NO_2$ | H | 4-Methoxyphenyl |
| 54 | 8-$NO_2$ | H | Phenyl |
| 55 | 8-$NO_2$ | H | H |
| 56 | 8-$NO_2$ | H | *NN*-Dimethyl aminophenyl |
| 57 | 8-$NO_2$ | H | 3,4,5-Trimethoxy phenyl |
| 58 | 8-$NO_2$ | $NO_2$ | 2-Hydroxyphenyl |
| 59 | 8-$NO_2$ | $NO_2$ | 4-Chlorophenyl |
| 60 | 8-$NO_2$ | $NO_2$ | 4-Methoxyphenyl |
| 61 | 8-$NO_2$ | $NO_2$ | Phenyl |
| 62 | 8-$NO_2$ | $NO_2$ | H |
| 63 | 8-$NO_2$ | $NO_2$ | *NN*-Dimethyl aminophenyl |
| 64 | 8-$NO_2$ | $NO_2$ | 3,4,5-Trimethoxy phenyl |

$$\begin{aligned} pIC_{50}(PT) = {} & 2.4769 + 10.4372\ \text{SAAverage} \\ & - 0.0212\ \text{H-AcceptorCount} \\ & - 0.3111\ \text{IonizationPotential} \\ & + 0.0134\ \text{DipoleMoment} \\ & + 0.0336\ \text{QMDipoleX} \\ & - 0.0012\ \text{Quadrupole1} \qquad (\text{MODEL 1}) \end{aligned}$$

**Table 2**
Factor loadings obtained from the principal component analysis of the quantum and physicochemical MDs computed in this work

| MDs[a] | F1 | F2 | F3 | F4 | F5 | F6 | F7 | F8 | F9 | F10 | F11 | F12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Molwt | 0.98 | | | | | | | | | | | |
| Vol | 0.97 | | | | | | | | | | | |
| H-acc | 0.63 | 0.71 | | | | | | | | | | |
| H-don | | | 0.70 | | | | | | | | | |
| *S*log*P* | 0.80 | | | | | | | | | | | |
| SMR | 0.97 | | | | | | | | | | | |
| c3c | 0.85 | | | | | | | | | | | |
| cV3c | 0.87 | | | | | | | | | | | |
| DM | | −0.70 | | | | | | | | | | |
| Q1 | | | | | | | | | | | | |
| Q2 | | | | | | | | | | | | |
| IDAverage | 0.91 | | | | | | | | | | | |
| IP | | | | | −0.52 | | | | | | | |
| HOMO | | | | | 0.52 | | | | | | | |
| LUMO | | −0.78 | | | | | | | | | | |
| QMx | | | | | | −0.62 | | | | | | |
| QMy | | | | | | | | 0.60 | | | | |
| QMz | | | | | | | | | | | | |
| SKaverage | 0.67 | | | | | | | | | | | |
| SAaverage | | | −0.73 | | | | | | | | | |
| MIx | | | | | | | | | | | | |
| MIy | 0.88 | | | | | | | | | | | |
| MIz | 0.89 | | | | | | | | | | | |
| Xdipole | | | | | | | | | | | | |
| Ydipole | −0.66 | | | | | | | | | | | |
| Zdipole | | | | | | | 0.63 | | | | | |
| RGYR | 0.93 | | | | | | | | | | | |
| HF | | | | −0.67 | | | | | | | | |
| DistTop | −0.72 | | | | | | | | | | | |
| K1 | 0.97 | | | | | | | | | | | |
| K2 | 0.97 | | | | | | | | | | | |
| QMDM | | 0.79 | | | | | | | | | | |

[a] Factor loadings between −0.5 and 0.5 (0.5 < factor loading > −0.5) were deleted to clarify the relationships between MDs and identify their factor structure.

$N = 52$, $r^2 = 0.82$, $q^2 = 0.77$, $\lambda = 0.96$, $F = 41.96$, pred $r^2 = 0.78$.

$$\begin{aligned} pIC_{50}\ (\text{APTT}) = {} & 0.0538 + 0.0013\ \text{Heat of Formation} \\ & - 0.0306\ \text{H-AcceptorCount} \\ & - 0.0215\ \text{XcompDipole} \\ & + 0.0305\ \text{QMDipoleX} \\ & - 0.0299\ \text{QMDipoleY} \\ & + 4.0633\ \text{SAAverage} \qquad (\text{MODEL 2}) \end{aligned}$$

$N = 52$, $r^2 = 0.73$, $q^2 = 0.65$, $\lambda = 0.87$, $F = 32.25$, pred $r^2 = 0.81$.
Where $N$ is the number of compounds, $r^2$ is the correlation coefficient, $q^2$ is the, $\lambda$ is Wilks' statistics, $F$ is the Fisher ratio, pred $r^2$ is the $r^2$ for external test set.. The observed and predicted biological activities along with the residuals are shown in Tables 4 and 5. The correlation plot for model 1 and 2 are shown in Figures 4 and 5, respectively.

As we said earlier, the factors obtained by factorial analysis capture all of the 'essence' of the original MDs, because they are linear combinations of the original items, and these consecutive factors are uncorrelated to each other. Taken into account this procedure (the 12 previously extracted factors), the above stated QSAR models were developed.

### 2.4. Internal and external validations

Validation of QSAR equations is a very important aspect in ligand-based drug design. Together with the quantification of the goodness-of-fit, it is quite important to obtain a measure of the model predictive capacity and stability. The most popular validation criterion to explore the robustness of a predictive model is through the analysis of the influence of each one of individual ob-

**Table 3**
Results of the factor analysis by using the principal component method for 23 quantum and physicochemical MDs for 64 compounds

| Factor | % Total variance[a] | % Cumulative variance[a] |
|---|---|---|
| F1 | 42.1 | 42.1 |
| F2 | 14.9 | 55.9 |
| F3 | 11.0 | 70.8 |
| F4 | 6.1 | 76.8 |
| F5 | 4.7 | 81.5 |
| F6 | 3.9 | 85.4 |
| F7 | 3.6 | 88.9 |
| F8 | 2.9 | 91.8 |
| F9 | 2.5 | 94.2 |
| F10 | 2.1 | 96.4 |
| F11 | 1.7 | 98.1 |
| F12 | 1.4 | 99.5 |

[a] Calculated on Systat 11.

jects that configure the final equation. This procedure is known as CV or internal validation by leave-one-out (LOO). In recent years, the LOO press statistics have been used as a means of indicating predictive ability, and many authors consider high values for these statistics as an indicator, or even as the ultimate proof, of the high predictive power of a QSAR model.

## 3. Discussion

### 3.1. Interpretation of MODEL 1

The model for anticoagulant activity by estimation of prothrombin time [MODEL 1], using Principal Component Regression explains the coumarin-clotting factor interactions. It is mostly due to the biological activities expressed by the SAAverage, a mesure of hydrophobic function, electronic MDs, the Dipole Moment and quantum mechanic Dipole moment along *X -axis,* which showed positive contribution. The positive contribution of SAAverage implies that the binding affinity increases with the rise in hydrophobic features of compounds up to reach a critical value, after which the affinity will decrease. The electronic parameters like DM and QMx indicate that electronic interactions are important in formation of thrombin. The generated model shows that biological activity is negatively contributed with MDs like H-Acceptor count, Ionizationpotential and Quadrapole1. The presence of Hydrogen bond acceptor moiety in the structure will reduce the anticoagulant effect. Thus the addition of hydrogen bond donor groups like hydroxyl, amino will increase the activity.

### 3.2. Interpretation of MODEL 2

On the other hand, MODEL 2 shows that the Activated partial thromboplastin time increases with the increase in the Heat of Formation, SAAverage and Quntaum mechanic dipole along *X*-axis. The positive correlation with SAAAverage indicates that the hydrophobic interactions are involved in the inhibition of thrombus formation. Involvement of electronic interactions contributed by Quntaum mechanic dipole along *X*-axis are also equivalently important for the inhibition of Clotting. Similar to the MODEL 1 H-acceptor count is negatively contributing to the activity. In addition, the x component of Dipole and Quntaum mechanic dipole along Y-axis also showed a negative contribution to anticoagulant activity by APTT.

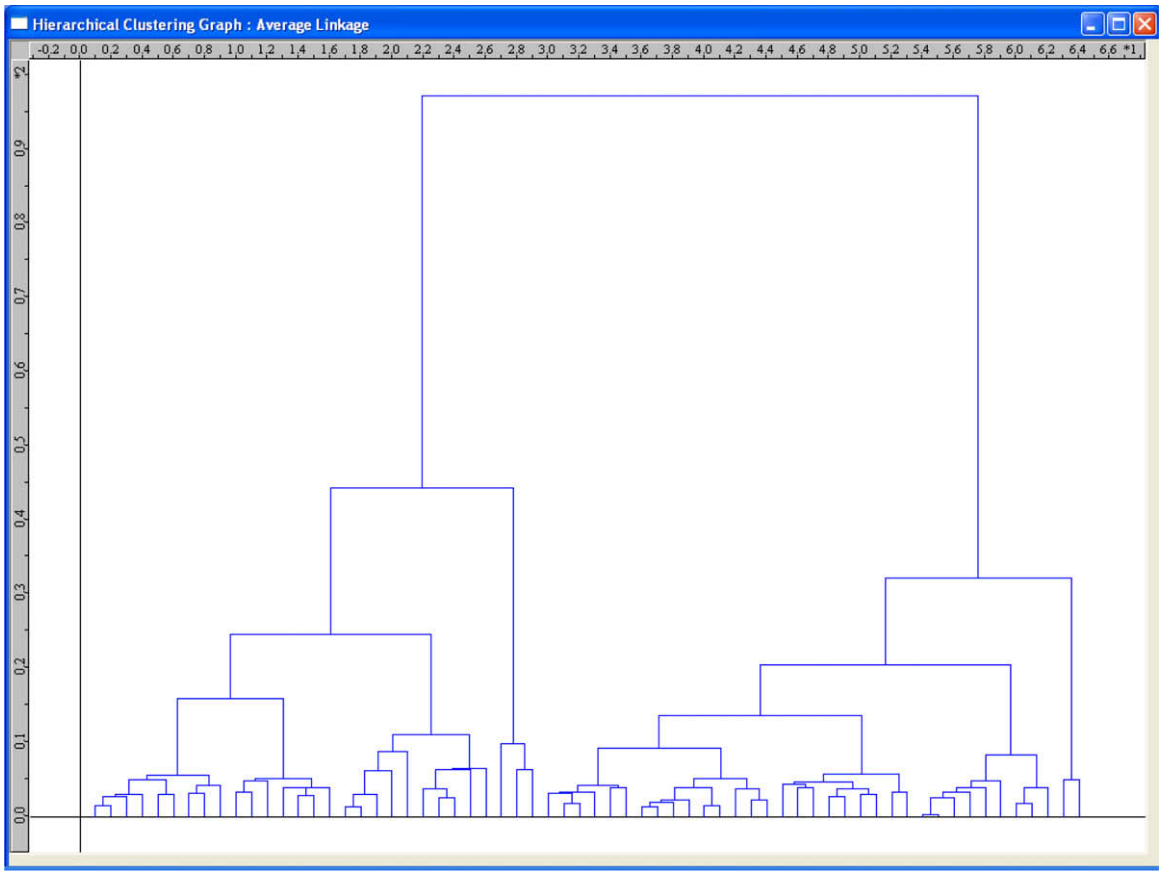


**Figure 2.** A dendrogram illustrating the results for the hierarchical k-NNCA of the set of compounds for prothrombin time in the training and prediction sets of the present work.

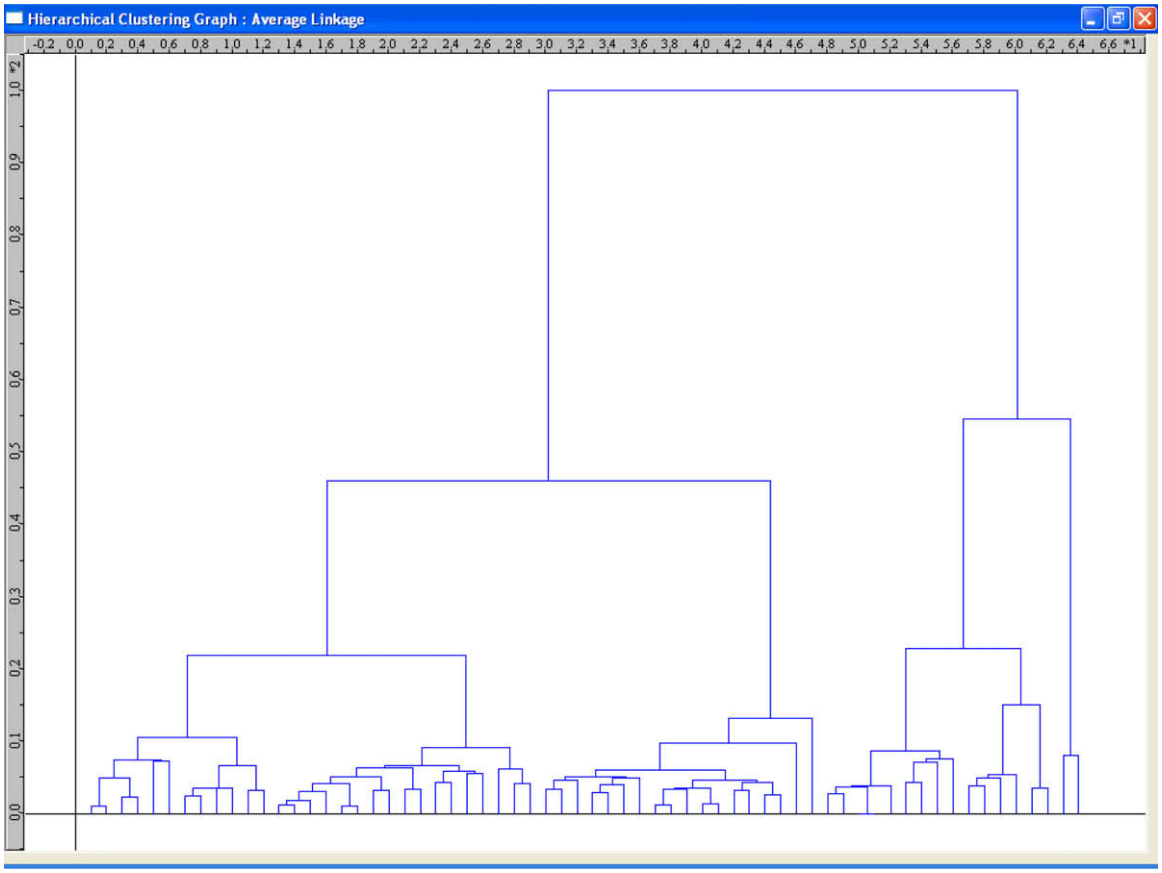


**Figure 3.** A dendrogram illustrating the results for the hierarchical k-NNCA of the set of compounds for activated partial thromboplastin time in the training and prediction sets of the present work.

## 4. Conclusion

The elongation of prothrombin time due to the addition of the pyridyl chromon-2-one derivatives indicates that the synthesized derivatives have anticoagulant properties. The mechanistic data obtained from the elongation of the prothrombin time indicates that the activity may be due to inhibition of clotting factors II, V, VII and X. While the elongation in activated partial prothrombin time may be due to inhibition of clotting factors V, VIII, IX, X and XIII. The initial and maximum effect of anticoagulants is based on the half-lives of each of these factors. Factor VII has a half-life of 6 h so the effect of coumarin on this factor will increase the bleeding to a certain degree within 6 h, however Factor II and X exhibit half-lives of 48–72 h, so the maximum effect of coumarin is not seen until three days after initiation or dose change. The coumarin types of drugs act by inhibition of factor. Thus inspection of actual interacting site and the interactions at the active site of the clotting factors can lead to discovery potent anticoagulant agents.

## 5. Experimental

### 5.1. Materials and methods

The IR (KBr) spectra of the synthesized compounds were recorded on a Jasco- FTIR 4100 instrument. The $^1$H NMR spectra of the compounds were recorded on 400 MHz Varian NMR, using DMSO-$d_6$ as solvent. This series was subjected to QSAR studies using Vlife Molecular Design Suite (MDS) 3.0 (obtained from Vlife sciences, Pune) running on P-IV processor. The quantum and physicochemical properties of each compound were specified using more than 500 descriptors, which delineate liphophilic, conformational, electronic, spatial, structural, thermodynamic and quantum mechanical information. For the calculation of quantum descriptors MOPAC simulations were used with the PM3 Hamiltonian.

### 5.2. Step-I: synthesis of ethyl-2-(6-aminopyridin-3-yl)-3-oxobutanoate

A mixture of 2 g of potassium hydroxide, 3 g of potassium carbonate and 3 mL of ethyl acetoacetate (0.01 mol) was irradiated under microwave irradiation for 2 min at 455 W. After the formation of potassium salt of ethyl acetoacetate the solution of 4g (0.01 mol) 2-amino-5-bromopyridine in 10 mL DMSO was added, and the mixture was stirred vigorously under microwave irradiation for 40 min at 700 W. The mixture was cooled and the DMSO was removed under vacuum. The product was extracted in three 25 mL portions of chloroform all the fractions were combined and chloroform was evaporated to obtain the product which was recrystallized from methanol.

### 5.3. Step-II: synthesis of 3-(6-aminopiridin-3-yl)-4-methyl-2H-chromen-2-one[19–22]

In 500 mL three-necked round bottom flask, with sealed magnetic stirrer, a thermometer reaching to bottom placed 30 mL of concentrated sulfuric acid. The flask was surrounded with ice bath and when temperature falls below 10 °C added a mixture of 1 g (2 mol) of substituted phenol and 3.2 g (2 mol) of ethyl-2-(6-aminopyridin-3-yl)-3-oxobutanoate the mixture was stirred, and the temperature kept below 10 °C by means of ice and salt. After

**Table 4**
Table showing the results of the observed and predicted biological activity for prothrombin time (PT)

| Sr. no | Observed $PIC_{50}$ PT | Predicted $PIC_{50}$ PT[a] | Residuals[a] |
|---|---|---|---|
| 1 | 0.461 | 0.448 | 0.012 |
| 2[b] | 0.028 | 0.126 | −0.098 |
| 3 | 0.250 | 0.263 | −0.013 |
| 4 | 0.244 | 0.072 | 0.171 |
| 5 | 0.473 | 0.498 | −0.025 |
| 6 | 0.455 | 0.385 | 0.069 |
| 7 | 0.326 | 0.221 | 0.104 |
| 8 | 0.161 | 0.031 | 0.129 |
| 9 | 0.857 | 0.709 | 0.147 |
| 10 | 0.782 | 0.690 | 0.091 |
| 11 | 0.792 | 0.728 | 0.063 |
| 12 | 0.657 | 0.728 | −0.071 |
| 13 | 0.392 | 0.476 | −0.084 |
| 14 | 0.666 | 0.638 | 0.027 |
| 15 | 0.590 | 0.516 | 0.073 |
| 16[b] | 0.384 | 0.436 | −0.052 |
| 17 | 0.554 | 0.538 | 0.015 |
| 18 | 0.311 | 0.368 | −0.057 |
| 19 | 0.447 | 0.493 | −0.046 |
| 20[b] | 0.364 | 0.390 | −0.026 |
| 21 | 0.783 | 0.742 | 0.040 |
| 22 | 0.526 | 0.503 | 0.022 |
| 23[b] | 0.559 | 0.563 | −0.004 |
| 24 | 0.785 | 0.771 | 0.013 |
| 25[b] | 0.812 | 0.632 | 0.179 |
| 26 | 0.682 | 0.684 | −0.002 |
| 27 | 0.594 | 0.526 | 0.067 |
| 28 | 0.751 | 0.800 | −0.049 |
| 29 | 0.350 | 0.436 | −0.086 |
| 30 | 0.459 | 0.423 | 0.035 |
| 31 | 0.808 | 0.573 | 0.234 |
| 32 | 0.210 | 0.393 | −0.183 |
| 33 | 0.326 | 0.497 | −0.171 |
| 34 | 0.442 | 0.373 | 0.068 |
| 35[b] | 0.759 | 0.590 | 0.168 |
| 36 | 0.252 | 0.197 | 0.054 |
| 37 | 0.610 | 0.680 | −0.070 |
| 38 | 0.774 | 0.802 | −0.028 |
| 39 | 0.658 | 0.706 | −0.048 |
| 40 | 0.791 | 0.791 | 0.000 |
| 41 | 0.615 | 0.641 | −0.026 |
| 42 | 0.780 | 0.751 | 0.028 |
| 43 | 0.493 | 0.561 | −0.068 |
| 44 | 0.494 | 0.330 | 0.163 |
| 45 | 0.724 | 0.736 | −0.012 |
| 46 | 0.366 | 0.372 | −0.006 |
| 47 | 0.398 | 0.473 | −0.075 |
| 48[b] | 0.310 | 0.360 | −0.050 |
| 49 | 0.612 | 0.566 | 0.045 |
| 50 | 0.309 | 0.283 | 0.025 |
| 51 | 0.573 | 0.563 | 0.009 |
| 52[b] | 0.684 | 0.671 | 0.012 |
| 53 | 0.461 | 0.574 | −0.113 |
| 54 | 0.693 | 0.629 | 0.063 |
| 55 | 0.407 | 0.408 | −0.001 |
| 56 | 0.831 | 0.900 | −0.069 |
| 57[b] | 0.402 | 0.563 | −0.161 |
| 58[b] | 0.350 | 0.377 | −0.027 |
| 59 | 0.656 | 0.602 | 0.053 |
| 60 | 0.182 | 0.327 | −0.145 |
| 61[b] | 0.211 | 0.365 | −0.154 |
| 62 | 0.226 | 0.139 | 0.086 |
| 63[b] | 0.705 | 0.585 | 0.119 |
| 64 | 0.290 | 0.232 | 0.057 |

[a] Calculated on Vlife Sciences Molecular design Suite 3.0.
[b] Indicates molecules of test set.

**Table 5**
Table showing the results of the observed and predicted biological activity for activated partial thromboplastin time (APTT)

| Sr. No | Observed $PIC_{50}$ APTT | Predicted $PIC_{50}$ APTT[a] | Residuals[a] |
|---|---|---|---|
| 1 | 0.439 | 0.348 | 0.090 |
| 2[b] | −0.098 | −0.026 | −0.071 |
| 3 | 0.076 | 0.138 | −0.062 |
| 4 | 0.070 | −0.099 | 0.169 |
| 5 | 0.392 | 0.287 | 0.104 |
| 6 | 0.154 | 0.190 | −0.037 |
| 7 | 0.186 | 0.085 | 0.101 |
| 8 | −0.060 | −0.083 | 0.023 |
| 9 | 0.532 | 0.430 | 0.101 |
| 10 | 0.537 | 0.448 | 0.088 |
| 11 | 0.489 | 0.471 | 0.017 |
| 12 | 0.433 | 0.541 | −0.108 |
| 13 | 0.291 | 0.356 | −0.065 |
| 14 | 0.521 | 0.473 | 0.047 |
| 15 | 0.289 | 0.247 | 0.041 |
| 16[b] | 0.344 | 0.308 | 0.035 |
| 17 | 0.403 | 0.380 | 0.022 |
| 18 | 0.163 | 0.186 | −0.023 |
| 19 | 0.493 | 0.534 | −0.041 |
| 20[b] | 0.370 | 0.402 | −0.032 |
| 21 | 0.304 | 0.463 | −0.159 |
| 22 | 0.025 | 0.171 | −0.146 |
| 23[b] | 0.280 | 0.293 | −0.013 |
| 24 | 0.434 | 0.423 | 0.010 |
| 25[b] | 0.446 | 0.301 | 0.144 |
| 26 | 0.347 | 0.396 | −0.0496 |
| 27 | 0.311 | 0.305 | 0.005 |
| 28 | 0.401 | 0.293 | 0.107 |
| 29 | 0.049 | 0.102 | −0.053 |
| 30 | 0.158 | 0.153 | 0.004 |
| 31 | 0.507 | 0.449 | 0.057 |
| 32 | 0.190 | 0.197 | −0.007 |
| 33 | 0.202 | 0.236 | −0.034 |
| 34 | 0.141 | 0.221 | −0.080 |
| 35[b] | 0.318 | 0.291 | 0.026 |
| 36 | −0.049 | −0.076 | 0.027 |
| 37 | 0.275 | 0.364 | −0.089 |
| 38 | 0.533 | 0.506 | 0.026 |
| 39 | 0.385 | 0.419 | −0.034 |
| 40 | 0.490 | 0.487 | 0.002 |
| 41 | 0.396 | 0.368 | 0.027 |
| 42 | 0.400 | 0.332 | 0.067 |
| 43 | 0.239 | 0.229 | 0.009 |
| 44 | 0.193 | 0.084 | 0.108 |
| 45 | 0.523 | 0.392 | 0.130 |
| 46 | 0.265 | 0.281 | −0.016 |
| 47 | 0.097 | 0.194 | −0.097 |
| 48[b] | 0.159 | 0.249 | −0.090 |
| 49 | 0.483 | 0.431 | 0.051 |
| 50 | 0.099 | 0.160 | −0.061 |
| 51 | 0.158 | 0.100 | 0.057 |
| 52[b] | 0.263 | 0.301 | −0.039 |
| 53 | 0.308 | 0.217 | 0.090 |
| 54 | 0.292 | 0.338 | −0.046 |
| 55 | 0.306 | 0.231 | 0.074 |
| 56 | 0.380 | 0.329 | 0.050 |
| 57[b] | 0.101 | 0.100 | 0.001 |
| 58[b] | 0.249 | 0.320 | −0.071 |
| 59 | 0.455 | 0.485 | −0.030 |
| 60 | 0.051 | 0.121 | −0.070 |
| 61[b] | 0.110 | 0.176 | −0.066 |
| 62 | 0.486 | 0.334 | 0.151 |
| 63[b] | 0.054 | 0.040 | 0.013 |
| 64 | 0.073 | 0.129 | −0.056 |

[a] Calculated on Vlife Sciences Molecular design Suite 3.0.
[b] Indicates molecules of test set.

all the mixture had been added (about 2 h) the reaction mixture was stirred for 24 h without further cooling. The reaction mixture was poured with vigorous stirring into ice and water mixture. The precipitate was then collected on filter paper and washed with 50 mL portions of cold water. The crude product was dissolved in 50 mL of 5% NaOH solution and the solution was filtered, and substituted coumarin was reprecipited from the filtrate by slow addition of dilute (1:10) sulfuric acid until solution until the solution was neutral to litmus. During neutralization the mixture must be well stirred. The product was collected on Buchner funnel and washed with four 50 mL portions of cold water and dried. IR (KBr), 1725–1720(C=O); 1150–1200(C–O–C); 1360–1310(C–N);

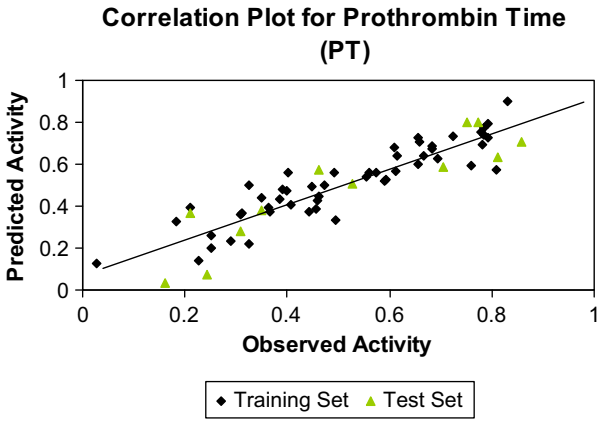


**Figure 4.** Figure showing correlation plot between observed and predicted activity for prothrombin time (PT).

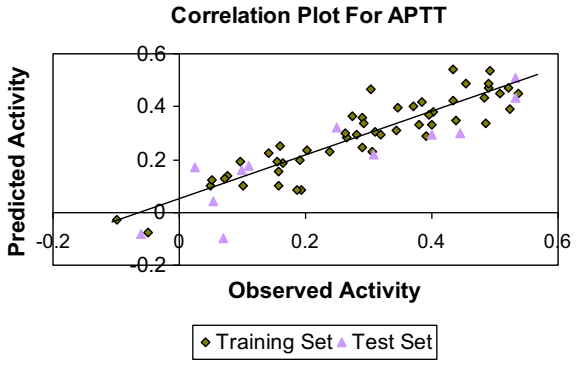


**Figure 5.** Figure showing correlation plot between observed and predicted activity for activated partial thromboplastin time (APTT).

3398–3381(N–H); 1675–1610(C=N); NMR (DMSO) $\delta$ : 0.94 (s, 3H, $CH_3$), 2.48 (s, 2H, $NH_2$), 6.91 (s, 1H, 3-pyridine), 7.14, (s, 1H, 4-pyridine), 7.35–8.25 (m, H, aromatic), 8.91 (s, 1H, 6-pyridine).

### 5.4. Step-III: synthesis of 4-methy-3-(6-[phenyl methylene] amino pyridine-3-yl)-2H chromen-2-one[23–25]

In flask containing 15 mL of glacial acetic acid 1.33 g (0.005 mol) of 3-(6-aminopiridin-3-yl)-4-methyl-2H-chromen-2-one was added, to this mixture 0.53 g (0.005 mol) of aldehyde was added. And the reaction mixture was treated with microwave irradiation of 10 min at 700 W. The mixture after cooling was poured in cold water the imines precipitate out. IR (KBr), 1725–1720(C=O); 1150–1200(C–O–C); 1360–1310(C–N); 1675–1610(C=N); NMR (DMSO) $\delta$ : 0.92 (s, 3H, $CH_3$), 6.99 (s, 1H, 3-pyridine), 7.15, (s, 1H, 4-pyridine), 7.40–8.25 (m, H, aromatic), 8.88 (s, 1H, 6-pyridine), 9.20 (s, 1H, CH=N).

### 5.5. Determination of prothrombin time (PT) by Quick's method[26]

- Pipette 0.1 mL of plasma in small test tube. Add 0.1 mL of brain thromboplastin and mix.
- Wait for 2 min and add pre warmed calcium chloride solution at 37 °C, mix and start the stop watch.
- Hold the tube in front a source of light and keep tilting the test tube gently. At first appearance of fibrin clot stop the watch immediately and record the time. Record this time as control reading.
- Pipette 0.1 mL of plasma in small test tube; add 0.1 mL of 100 mg/mL test compound solution. Incubate for 5 min and repeat the procedure to record the elongation in mean prothrombin time.
- The $IC_{50}$ values were calculated from the dose response curve.

### 5.6. Elongation of activated partial thromboplastin time (APTT)[27]

- Pipette out 0.1 mL of brain extract, 0.1 mL of kaolin reagent and 0.2 mL of plasma in a test tube. Incubate for 1 min at 37 °C.
- Add 0.1 mL of calcium chloride and start the stop watch.
- After 20 s observe the formation of clot by tilting the test tube. As soon as the clot is observed note the time as control reading.
- Pipette out 0.1 mL of brain extract, 0.1 mL of kaolin reagent and 0.2 mL of plasma in small test tube; add 0.1 mL of 100 mg/mL test compound solution. Incubate for 5 min and repeat the procedure to record the elongation in mean activate partial thromboplastin time.
- The $IC_{50}$ values were calculated from the dose response curve.

### 5.7. QSAR studies

#### 5.7.1. Data Set

To obtain mathematical expressions able to discriminate between good (high) and moderate-low activity compounds, the chemical information contained in a great number of compounds, with and without the desired biological profile, must be statistically processed. In the present study, the anti-coagulant activity was used to establish a classical linear-classification-based QSAR equation. Taking into account that the most critical aspect, in the construction of a training dataset, is the molecular diversity of the included compounds, we had selected compounds having as much structural variability as possible. The anti-coagulant activity was evaluated on basis of elongation of prothrombin time and thrombin time of each compound in the database was evaluated in vitro. The builder module of the Vlife MDS 3.0 program (Vlife Sciences Technologies Pvt. Ltd) was used to generate molecular models of 64, 3-[2-amino(pyridine-3-yl)]chromen-2-one derivatives. These molecular models were built using the template of the most active molecule from the data set, that is, 3(6-amino-5-nitro pyridine-3-yl) 4-methyl-*2H*-chromen-2-one. They were then energy-minimized using the Merck Molecular Force Field (MMFF) until the root mean square (rms) gradient reached value 0.001 kcal $mol^{-1}$ Å. The chemical information contained in a great number of compounds, with and without the desired biological profile, must be statistically processed. In the present study, the $IC_{50}$ was used to establish a classical linear-classification based QSAR equation.

#### 5.7.2. Molecular descriptors for QSAR analysis

A large number of MDs is usually used in QSAR methods. The specific biological action of drugs is frequently described by hydrophobic, electronic, and steric properties. The hydrophobic properties express the ability of a molecule to be transported through the organism to interact with biological membranes and to be bound to the receptor by van der Waals forces. We considered as hydrophobic descriptor the decimal logarithm of the octanol–water partition coefficient (*slogP*). Both electronic and steric properties characterize the pharmacodynamic properties in the ligand–receptor interaction. They define the ability of the drug to join to the receptor. The electronic descriptors calculated by quantum mechanical procedures for the 2D QSAR Studies were: molecular weight (molwt), Volume (Vol), H-bond acceptor count (H-acc), H-

bond doner count (H-don), Chi3cluster (c3c), chiV3Cluster (cV3c), dipole moment (DM), QMDipoleX (QMDx), QMDipoleY (QMDy), QMDipoleZ (QMDz), Information based descriptor (Idaverage), Quadrupole1 (Q1), Quadrupole2 (Q2), Heat of Formation (HF), highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO), ionization potential (IP), molar refractivity (SMR), partition coefficient (*slogP*), Moment of inertia along *x-axis* (MIx) Moment of inertia along y axis (MIy), Moment of inertia along *z-axis* (MIz) average hydrophobicity by kellog method (SKaverave), average hydrophobicity by Audry method (SAaverage), Xcomponent of Dipole (Xdipole), Ycomponent of Dipole (Ydipole), Zcomponent of Dipole (Zdipole), radius of gyration (RGYR), Distanc Topological (DistTop), Kappa1 (K1), Kappa2 (K2), quantum mechanics dipole moment (QMDM).

### 5.7.3. Analysis of principal components

To conduct the comparison of the MDs computed in this work, we performed a factorial analysis, using the principal components method. The theoretical aspects of this statistical technique have been extensively exposed in the literature including many chemical applications.[5–11] The main uses of factorial analytical techniques are: (1) to reduce the number of variables, and (2) to detect structure in the relationships between variables, namely, to classify variables[10,11] In this approach, factorial loadings (or 'new' variables) are obtained from original variables of Molecular Descriptors. Thus, these factors capture all the 'essence' of these MDs, because they are linear combinations of the original items. Because each consecutive factor is defined to maximize that variability not captured by the preceding factor, consecutive factors are independent of each other. Put in another way, consecutive factors are uncorrelated or orthogonal to each other. The first obtained factor is generally more highly correlated with the variables than the other factors. This is to be expected, because these factors are successively extracted and will account for less and less overall variance. The factor analysis was carried out using 'varimax normalized' as rotational strategy to obtain the factorial loadings from the principal component analysis. The goal of this rotational procedure is to obtain a clearer pattern of the loadings, that is, factors that be somehow clearly marked by high loadings for some variables and low loadings for others. The 'varimax normalized rotation' is the method that is most commonly used as 'varimax' rotation.[18] This rotational strategy is aimed at maximizing the variances of the squared normalized factorial loadings (row factorial loadings divided by squared roots of the respective communalities), across the variables for each factor. This strategy makes the structure of the factorial pattern as simple as possible, permitting a clearer interpretation of the factors without loss of orthogonality between them. Finally, some of the most important conclusions, which could can be drawn from a factor analysis that will be of large usefulness in the present article are the following (1) variables with a high loading in the same factor are interrelated and will be the more, the higher the loadings, (2) no correlation exists between variables having nonzero loadings only in different factors. These are the principal ideas that permit the interpretation of the factorial structure, obtained using the factorial analysis as a classification method, and (3) only variables with high loadings in different factors may be combined in a regression equation to eliminate collinearities.

### 5.7.4. Cluster analysis

Cluster analysis (CA) encompasses a number of different classification algorithms, and it permits to organize the observed data into meaningful structures. Many CA algorithms have been invented, and they belong to two categories: hierarchical clustering and partitional (nonhierarchical) clustering. Hierarchical clustering rearranges objects in a binary tree-structure (joining clustering), and these methods are implemented in either an agglomerative (bottom–up) or divisive (top–down) procedure. On the other hand, the partitional clustering assumes that the objects have nonhierarchical characters.[12–16] Most popular partitional cluster algorithms are k-mean cluster algorithms (k-MCAs) and another by Jarvis and Patrick (also known as k-nearest neighbor cluster algorithm; k-NNCA) algorithms. The k-mean clustering algorithms use an interchange (or switching) method to divide *n* data points into k groups (clusters) so that the sum of distances/dissimilarities between the objects within the same cluster is minimized. The k-mean approach requires that k (the number of clusters) is known before clustering. The Jarvis–Patrick method requires that the user specify the number of nearest neighbors, as well as the number of neighbors in common to merge two objects. The Jarvis–Patrick method is a deterministic algorithm; it does not require iterations for computations.[12–16] To design the training and test series, as well as to demonstrate the structural diversity of the present database, we carried out one of these kinds of cluster analyses (k-NNCA) for anti-coagulant series. The Vlife graph package from the QSAR Module was used to develop these CA's. In this study, we used the 'average linkage' metric as the method to merge objects into clusters. The average linkage distance between two clusters is defined as the average (Euclidean squared arithmetic mean) distance between pairs of objects, one in each cluster. Average linkage tends to join those clusters with small variances and produces new clusters with roughly the same variance.

### 5.7.5. LDA and classification-based QSAR model

The discriminant functions were obtained using LDA,[17] as implemented in Vlife model building Wizard, the default parameters of this program were used in the development of the model. Forward stepwise was fixed as the strategy for variable selection. The principle of maximal parsimony (Occam's razor) was taken into account as the strategy for model selection. In its original form, Occam's razor states that 'Entities should not be multiplied beyond necessity'. In this case, simplicity is loosely equated with the number of parameters in the model. If we understand the predictive mistake to be the error rate for unseen examples, the Occam's razor can be stated for the selection of QSAR models as ('QSAR Occam's Razor'): Given two QSAR models with the same predictive error, the simplest one should be preferred because simplicity is desirable in itself.[28] Relation to this, we select that model with the highest statistical signification, but having as few parameters as possible.

### 5.7.6. Validation of the obtained model

The statistical robustness and predictive power of the obtained model were assessed using a prediction (test) set.[29] In addition; a leave-group-out (LGO) cross-validation (CV) strategy was carried out. Eliminate a compound in the training set and predict its biological activity on the basis of the k-NN principle, that is, as the weighted average activity of k most similar molecules. The similarities are evaluated as Euclidean distances between compounds using only the subset of descriptors that corresponds to the current model. Repeat until every compound in the training set has been eliminated and its activity predicted once. In this way, every observation was predicted once (in its group of left-out observations).

## Acknowledgement

The authors are thankful to Dr. H.N. More, Principal, Bharati Vidyapeeth College of Pharmacy, Kolhapur for providing facilities to carry out the research work.

## References and notes


1. Coleman, R. W.; Marder, V. J.; Salzman, E. W.; Hirsh, J.. In Coleman, R. W., Marder, V. J., Salzman, E. W., Hirsh, J., Eds., 3rd ed.; J.B. Lippincott: Philadelphin, 1994; pp 3–18.
2. Elodi, S.; Varadi, K. *Thromb. Res.* **1979**, *15*, 617.
3. Sturzebecher, J.; Markwardt, F.; Walsmann, P. *Thromb. Res.* **1976**, *9*, 637.
4. Rameshwar, N.; Krishna, K.; Ashok Kumar, B.; Parthasarathy, T. *Bioorg. Med. Chem.* **2006**, *14*, 319.
5. Marrero Ponce, Y. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2010.
6. Marrero-Ponce, Y. *Bioorg. Med. Chem.* **2004**, *12*, 6351.
7. Cramer, R. D., III *J. Am. Chem. Soc.* **1980**, *102*, 1849.
8. Cramer, R. D., III *J. Am. Chem. Soc.* **1980**, *102*, 1837.
9. Needham, D. E.; Wei, I.-C.; Seybold, P. G. *J. Am. Chem. Soc.* **1988**, *110*, 4186.
10. Malinowski, E. R.; Howery, D. G. In *Factor Analysis in Chemistry*; Wiley: Interscience, New York, 1980.
11. Estrada, E.; Rodrýguez, L. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1037.
12. Xu, J.; Hagler, A. *Molecules* **2002**, *7*, 566.
13. Mc Farland, J. W.; Gans, D. J. In *Chemometric Methods in Molecular Design*; Waterbeemd, H., Ed.; VCH Publishers: Weinheim, Germany, 1995; pp 295–307.
14. Johnson, R. A.; Wichern, D. W. In *Applied Multivariate Statistical Analysis*; Prentice-Hall: Upper Saddle River, NJ, 1988.
15. Marrero-Ponce, Y.; Meneses-Marcel, A.; Machado-Tugores, Y.; Montero Pereira, D.; Escario, J. A.; Nogal-Ruiz, J. J.; Ochoa, C.; Arán, V. J.; Martýnez-Fernández, A. R.; Garcýa Sánchez, R. N.; Montero-Torres, A.; Torrens, F. *Curr. Drug Disc. Technol.* **2005**, *2*, 245.
16. Montero-Torres, A.; Celeste Vega, M.; Marrero-Ponce, Y.; Rolón, M.; Gómez-Barrio, A.; Escario, J. A.; Arán, V. J.; Martýnez-Fernández, R.; Meneses-Marcel, A. *Bioorg. Med. Chem.* **2005**, *13*, 6264.
17. Van de Waterbeemd, H. In *Chemometric Methods in Molecular Design*; van de Waterbeemd, H., Ed.; VCH Publishers: Weinheim, Germany, 1995; pp 265–288.
18. STATISTICA, Statsoft, Inc. 1999. http://www.statsoft.com.
19. Horning, E. C.. In *Organic Synthesis Collective*; John Wiley & sons, Inc.: USA, 1955; Vol. III, pp 281–283.
20. Rajendra, E.; Ramesh, P.; Ramu, K. *Ind. J. Chem.* **2004**, *43B*, 2650.
21. Wang, L.; Xia, J.; Ian, H.; Qiun, C.; Yun, M. S. *Ind. J. Chem.* **2003**, *42B*, 2097.
22. Jin, T. S.; Gua, T.; Shauangli, T.; Yin, Y.; Liu, H. *Ind. J. Chem.* **2003**, *42B*, 2612.
23. Delia, T. J.; Wilcox, T. M.; Ottoman, R. R. *J. Heterocycl. Chem.* **1979**, *16*, 1647.
24. Morrison, R. T.; Boyd, R. E. In *Organic Chemistry*, 6th ed.; New Delhi, Prentice-Hall of India Private Ltd: New Delhi, 1996; pp 657–675.
25. Loudon, G. M. In *Organic Chemistry*, 4th ed.; Oxford University press: New York, 2002; pp 874–880.
26. Godkar, P. B.; Godkar, P. B. In *Textbook of Medical Laboratory Technology*, 2nd ed.; Bhalani publishing house: Mumbai, India, 2005; pp 840–841.
27. Godkar, P. B.; Godkar, P. B. In *Textbook of Medical Laboratory Technology*, 2nd ed.; Bhalani publishing house: Mumbai, India, 2005; p 843.
28. Estrada, E.; Patlewicz, G. *Croat. Chim. Acta* **2004**, *77*, 203.
29. Wold, S.; Erikson, L. In *Chemometric Methods in Molecular Design*; van de Waterbeemd, H., Ed.; VCH Publishers: Weinheim, Germany, 1995; p 309.