



مطالعه کیوسار مشتقات کومارین به عنوان عوامل ضد بیماری آلزایمر

شریعه حسینی: دانشیار، گروه شیمی، دانشکده شیمی دارویی، دانشگاه علوم پزشکی آزاد اسلامی، تهران، ایران (* نویسنده مسئول) shariehosseini@yahoo.com

اصغر داوود: استاد، گروه شیمی دارویی، دانشکده علوم دارویی، دانشگاه علوم پزشکی آزاد اسلامی، تهران، ایران

الهام عسکری زاده: استادیار، گروه شیمی کاربردی، دانشکده شیمی دارویی، دانشگاه علوم پزشکی آزاد اسلامی، تهران، ایران

سارا یغمایی: دانشجو، گروه شیمی دارویی، دانشکده علوم دارویی، دانشگاه علوم پزشکی آزاد اسلامی، تهران، ایران

چکیده

کلیدواژه‌ها

کومارین،
فعالیت ضدآلزایمر،
نظریه تابعیت چگالی،
کیوسار

زمینه و هدف: بیماری آلزایمر یکی از شایع‌ترین علل از دست دادن عملکرد ذهنی است که به صورت کلی و با عنوان دمانس شناخته می‌شود. این بیماری تقریباً ۲ درصد را در کشورهای در حال توسعه درگیر می‌کند. با توجه به خاصیت ضد آلزایمری کومارین و مشتقات آن در این تحقیق از مشتقات کومارین استفاده شده است.

روش کار: با استفاده از نرم افزار گوسین و روش محاسباتی DFT/B3LYP و مجموعه پایه 6-31G(d) ۴۴ ترکیب از مشتقات کومارین اِپتیمایز شده و توصیفگرهای کوانتومی بدست آمده است، همچنین با استفاده از نرم افزار دراگون سایر توصیفگرها تعیین شدند. برای بررسی ارتباط این توصیفگرها با فعالیت این دسته از ترکیبات از روش همبستگی خطی MLR استفاده شد.

یافته‌ها: معادله کیوسار با ضریب تعیین و همبستگی $R^2 = 0.9$ و $R = 0.95$ برای این دسته از ترکیبات بدست آمد. همچنین $MSE = 0.15$ ، $RMSE = 0.12$ و $Q^2 = 0.83$ تاییدی بر قابل قبول بودن مدل بدست آمده می‌باشد.

نتیجه‌گیری: معادله بدست آمده نشان می‌دهد فعالیت این ترکیبات با ضرایب منفی $Jhetp$ و $nCbH$ ، $RDF065m$ ارتباط دارد، به این معنی که با افزایش مقادیر این توصیفگرها مقدار فعالیت کاهش می‌یابد. از طرفی فعالیت مهارکنندگی این ترکیبات با ضریب مثبت $nR=Ct$ و SPH بستگی داشته، یعنی با افزایش این مقدار، فعالیت این ترکیبات نیز افزایش می‌یابد، همچنین همبستگی خوبی بین مقادیر فعالیت تجربی و پیش بینی شده بدست آمد.

تعارض منافع: گزارش نشده است.

منبع حمایت‌کننده: حامی مالی ندارد.

شیوه استاندارد به این مقاله:

Hosseini S, Davood A, Askarizadeh E, Yaghmei S. The QSAR Study of Coumarin Derivatives as Anti Alzheimer Agents. Razi J Med Sci. 2024(8 Jun);31.44.

Copyright: ©2024 The Author(s); Published by Iran University of Medical Sciences. This is an open-access article distributed under the terms of the CC BY-NC-SA 4.0 (<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.en>).

*انتشار این مقاله به صورت دسترسی آزاد مطابق با **CC BY-NC-SA 4.0** صورت گرفته است.

The QSAR Study of Coumarin Derivatives as Anti Alzheimer Agents

① **Sharieh Hosseini:** Associate Professor, Department of Chemistry, Faculty of Pharmaceutical Chemistry, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran (* Corresponding Author) shariehosseini@yahoo.com

Asghar Davood: Professor, Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences Tehran Medical sciences, Islamic Azad University, Tehran, Iran

Elham Askarizadeh: Assistant Professor, Department of Chemistry, Faculty of Pharmaceutical Chemistry, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran

Sara Yaghmei: Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran

Abstract

Background & Aims: Alzheimer's disease is one of the most common causes of loss of mental function, which is generally known as dementia (1). Alzheimer's disease affects almost two percent of people in developing countries (2).

The risk of developing the disease increases with age. Alzheimer's disease and its consequences are the Greatest Health Crisis of the 21st Century. Coumarin and its derivatives are one of the most abundant natural compounds in plants. This compound is a heterocyclic oxygen that is known as the parent chemical structure for a class of phytochemicals naturally found in several plant species. These compounds have been used as anticoagulant, antibacterial, anti-osteoporosis and antiasthmatic (6). Quantitative structure- activity relationship (QSAR) is widely used in drug design processes to improve the therapeutic indicators of designed compounds. QSAR models are mathematical equations written based on the relationship between the chemical structure of compounds and their biological activity (7). These equations use different descriptors; these descriptors are divided into different groups such as structural, geometric, spatial, quantum, and the like and are obtained by experimental and computational methods. Due to the time-consuming and costly experimental methods, computational methods are preferred (8). Density functional theory has been one of the most popular methods of solid-state physics since 1970. Though it was not until 1990, when the considered approximations, in theory, were revised, and a better model for exchange interactions was proposed, it was not considered an exact method in quantum chemistry.

Methods: Due to the anti-Alzheimer's property of coumarin and its derivatives, In this research, 44 coumarin derivatives have been selected that were previously synthesized and their experimental activity was measured. All of compounds optimized and their quantum descriptors have been obtained using Gaussian software and computational DFTmethod and basis set 6-31G (d), other descriptors were also determined using Dragon software.

At first, using IBMSPSS20 software and stepwise multiple linear regression method, the samples were randomly divided into training series (33 molecules) and prediction series (11 molecules). 75% of the samples were considered training and 25% of them as a test, and this ratio was repeated three times to get the best correlation. The training series was used to create a suitable model, and the prediction series was used to evaluate the model. The most important part of creating an efficient model is the selection of appropriate descriptors. After calculating different descriptors, they are selected as suitable descriptors to build the model. In this study, the stepwise method (SW) was used to select the most appropriate descriptors. In a stepwise method by examining all descriptors, the selection process continues until a model with a high correlation coefficient is obtained. Using the stepwise method from among 95 descriptors, six descriptors were selected as the most suitable and were modeled by the MLR method. After selecting the most suitable

Keywords

Coumarins,
QSAR,
Anti Alzheimer activity,
Density functional theory

Received: 30/12/2023

Published: 08/06/2024

descriptors, the relationship between the selected descriptors and the activity of the drug compounds was established by a stepwise incremental method. To check the accuracy and validity of the model, internal and cross-validation methods were used. Internal validation is related to the statistical parameters of the training category, one of these parameters is R or the correlation coefficient. An acceptable equation has a correlation coefficient value greater than 0.6. The reported values for R^2 are between 0 and 1. $R^2=0$ means no correlation between the activity and the selected variable, and $R^2=1$ means a very good correlation. Increasing the number of parameters in an equation R^2 cannot show the correlation value correctly. For this reason, in an equation with several parameters, R^2_{adj} is a more accurate scale for expressing correlation. The closer this value is to one, the stronger the linear relationship between the dependent and independent variable. Cross-validation was used to predict the properties of the compounds that were not included in the new model.

Results: The final evaluation of the model is done by external validation, which indicates how much predictive ability the QSAR model has for the compounds outside the model. For the QSAR equation $R^2=0.89665$ and $R=0.94691$ were obtained. Also, $MSE=0.51397$ and $RMSEP=0.81297$ and $Q^2=0.83216$ are a confirmation of the acceptability of the obtained model. The obtained equation shows that the activity of these compounds is related to the negative coefficients of $RDF065m$, $nCbH$, $Jhetp$, which means that by increasing the values of these descriptors, the amount of activity decreases. On the other hand, the activity of these compounds depends on the positive coefficient of $nR=Ct$ and SPH that means by increasing this value, the activity of these compounds also increases. By placing the effective descriptors, the anti-Alzheimer activity values of these compounds were predicted. The comparison of these values with the obtained experimental activities shows an acceptable correlation between the values of experimental and predicted activity. To investigate how much the increase or decrease of the selected descriptors affects the desired activity, a graph of the average effect of the descriptors has been drawn. The y-axis represents the standardized coefficients, and the x-axis represents the selected descriptor using the SW-MLR method. According to this chart, the $Jhetp$ has the most negative, and the $nR=Ct$ descriptor has the most positive effect. In this chart, the descriptors drawn above the chart mean that the amount of activity increases with increasing the value of the descriptor, and for those drawn at the bottom of the chart, the amount of activity decreases as they increase. The evaluation of the domain of application shows that there is no systematic error in the obtained model, and it also shows the efficiency of the model for predicting behavior in experimental methods.

Conclusion: According to the examination of the results presented above, it can be generally concluded that according to the predicted activity values in Table 3, compounds 36 and 37 have the most activity and compounds 21 and 15 have the least anti-Alzheimer activity. The obtained model helps to design new drugs with more activity in the future based on the factors affecting the activity obtained from the QSAR equation. Since computational methods are faster and more accessible than experimental methods, it can be less predicted the necessary structures to design effective drugs.

Conflicts of interest: None

Funding: None

Cite this article as:

Hosseini S, Davood A, Askarizadeh E, Yaghmei S. The QSAR Study of Coumarin Derivatives as Anti Alzheimer Agents. *Razi J Med Sci.* 2024(8 Jun);31.44.

Copyright: ©2024 The Author(s); Published by Iran University of Medical Sciences. This is an open-access article distributed under the terms of the CC BY-NC-SA 4.0 (<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.en>).

***This work is published under CC BY-NC-SA 4.0 licence.**

مقدمه

بیماری آلزایمر (Alzheimer disease) یک نمونه اختلال پیشرونده‌ی عصبی است که به صورت زوال عقل، اختلال شناختی و از دست دادن حافظه نقش بازی میکند (۱). در سال‌های اخیر مونوآمین اکسیدازها (MAOs) آنزیم‌های مهم در سوخت و ساز نوروترنسمیتر مونوآمین و استرس اکسیداتیو توجه زیادی را برای نقش بالقوه خود در درمان آلزایمر دریافت کرده اند. MAOها دی نوکلئوتید آدنین فلاوینی هستند که حاوی آنزیم‌های مسئول دآمینا سیون اکسیداتیو از آمین‌های مختلف بیوژنیک و غذایی می‌باشند (Monoamine Oxidase type B Inhibitor) MAO_B یکی از ایزوفرم‌های آنزیمی است که به دآمینا سیون بتا فنتیل آمین و بنزیل آمین سرعت می‌بخشد. مهارکننده‌های MAO_B برای درمان بیماری‌های عصبی استفاده می‌شود. مطالعات نشان داده‌اند که بیان MAO_B در بافت‌های عصبی چهار برابر در سالمندی افزایش می‌یابد و فعالیت آن در مغز بیماران مبتلا به آلزایمر میزان مصرف انتقال دهنده‌های عصبی و آسیب نورونی را افزایش می‌دهد (۲، ۳). در همین حال، شواهد نشان داده است که MAO_B می‌تواند به میزان قابل توجهی حافظه یادگیری در مدل‌های حیوانی را با توجه به مهار استرس اکسیداتیو در آلزایمر بهبود بخشد. به طور کلی این شواهد نشان می‌دهد که MAO_B پتانسیل زیادی برای درمان آلزایمر دارد. عدم ثبات و توازن یون فلزی در مغز یکی از مهمترین نشانه‌های آلزایمر است که مربوط به شکل‌گیری پلاک‌های آمیلوئید و پیچ خوردگی نوروفیبریلاری (neurofibrillary) می‌شود (۱). بنابراین ایجاد کمپلکس از یون‌های فلزی به عنوان یک گزینه درمانی بالقوه برای درمان آلزایمر با تعدیل سطح بیومتال‌ها در مغز ارائه شده است. از نظر اهمیت MAO_B و بیومتال‌ها و استرس اکسیداتیو در درمان آلزایمر، به طراحی (MTDLs) با مهار MAO_B، شلات آهن و توانایی آنتی اکسیدان در سال‌های اخیر توجه زیادی شده است. بسیاری از ترکیبات توانایی کارآمد شلات یون آهن و فعالیت مطلوب مهار MAO_B را نشان داده‌اند. یکی از این ترکیبات کومارین است، محققان نشان داده‌اند که کومارین یکی از ممتازترین داربست‌های MAO_B

است. مدل‌سازی مولکولی با فراهم آوردن یک دید مولکولی از ساختار و دینامیک مولکولی مواد، ما را در درک بهتر ارتباط بین ساختار مولکولی و خواص نهایی کمک می‌کند (۴). آزمایشات معمول اطلاعاتی راجع به پدیده‌ها و فرایندهای مولکولی برای ما فراهم نمی‌آورند و همچنین شبیه‌سازی‌ها و آزمایش‌ها می‌توانند به کمک یکدیگر موجب درک مفاهیم جدیدی در خواص مواد شوند. از طرف دیگر مدل‌سازی مولکولی امکان پیش‌بینی خواص موادی را فراهم می‌آورد، که هنوز سنتز نشده‌اند یا اندازه‌گیری خواص آنها با مشکلاتی همراه است یا اینکه اطلاعات چندانی در مورد آنها وجود ندارد (۵). طراحی، توسعه، و تجاری کردن یک دارو، فرآیندی پیچیده، وقت‌گیر و هزینه‌بر است. به دلایل مذکور، هر وسیله یا روشی که سرعت و بازده هر مرحله از فرآیند کشف دارو را افزایش دهد مطلوب است. در علم شیمی دارویی، فعالیت بیولوژیکی مانند ویژگی‌های فیزیکی-شیمیایی و خواص شیمیایی ترکیبات آلی به ساختار مولکولی آنها وابسته است، طراحی دارو به کمک کامپیوتر بر اساس روش‌های کمی ساختار-فعالیت یکی از ابزارهایی است که می‌تواند سرعت و بازده فرآیند کشف دارو را افزایش دهد (۵، ۶). رابطه کمی ساختار-فعالیت (Quantitative structure-activity relationship) یا همان کیوسار (QSAR) به طور گسترده‌ای در پروسه‌های طراحی دارو به منظور بهبود شاخص‌های درمانی ترکیبات طراحی شده، استفاده می‌شود. مدل‌سازی QSAR یک پیش‌بینی و تشخیص است که برای یافتن ارتباط بین ساختارهای شیمیایی و فعالیت بیولوژیکی بر اساس معادلات ریاضی و روابط آماری استفاده می‌شود (۷) و در تلاش برای برآورد نیاز و تمایل شیمیدانان دارویی برای پیش‌بینی پاسخ بیولوژیکی ترکیبات تازه مهندسی شده یا توسعه نیافته تکامل یافته است (۸).

هدف از این تحقیق، مطالعه‌ی رابطه کمی ساختار-فعالیت ۴۴ ترکیب از مشتقات کومارین با استفاده از توصیف‌گرهای مختلف و روش رگرسیون خطی چندگانه گام به گام توسط نرم افزار SPSS به منظور پیش‌بینی فعالیت مهارکنندگی آنها در برابر آنزیم MAO_B می‌باشد.

روش کار

انتخاب داده ها: اولین قدم در مدلسازی QSAR انتخاب و جمع آوری مجموعه‌ای از داده‌ها شامل مولکول‌هایی با مکانیسم مهار کنندگی مشابه و فعالیت مشخص می‌باشد.

در انتخاب دیتا ست باید توجه کرد که فعالیت ترکیبات مورد نظر به دقت اندازه گیری شده باشد. در این پژوهش ۴۴ مشتق جدید کورکومین با فعالیت بازدارندگی آنزیم MAO_B از مقالات معتبر انتخاب شد (۱، ۹، ۱۰). فعالیت مهار کنندگی ترکیبات فوق با IC50 نمایش داده می‌شود. IC50 غلظتی از ترکیبات است که باعث اثر بازدارندگی ۵۰ درصد از آنزیم MAO_B می‌شود که در واقع معیاری از فعالیت ترکیبات می‌باشد. اما در این مطالعه از معیار pIC50 استفاده شده است که این معیار با فرمول منفی لگاریتم IC50 محاسبه می‌شود. علت استفاده از pIC50، پراکندگی مقادیر IC50 ترکیبات می‌باشد. برای رسم ساختارها از نرم افزار گوس ویو ۰۶ و بارعایت ساختار فضایی استفاده شد (۱۱).

پس از رسم ساختارها، با استفاده از نرم افزار گوسین ۰۹ (۱۱) کلیه ساختارهای بهینه و در مرحله بعد فرکانس آنها نیز محاسبه شد. بهینه سازی باروش نظریه تابعیت چگالی (DFT) Density Functional Theory و تابع ترکیبی B3LYP و مجموعه پایه 6-31G(d) با انجام شد. منظور از بهینه کردن ترکیبات بدست آوردن پایدارترین حالت ساختار است. استفاده از دستور فرکانس، علاوه بر کمک به محاسبه توصیفگرهای ترمودینامیکی، برای اطمینان از صحت بهینه سازی ترکیبات و عدم تشکیل حالت گذار می‌باشد. داده‌ها بصورت تصادفی به دو دسته آموزش و آزمون تقسیم شده است. وظیفه دسته آموزش تهیه مدل کیوسار با استفاده از روش رگرسیون خطی چندگانه گام به گام و دسته آزمون پیش بینی و اعتبارسنجی خارجی می‌باشد. این کار ۳ بار بصورت تصادفی انجام شده است تا بهترین همبستگی بدست آید.

انتخاب توصیفگرهای مولکولی مناسب: مهمترین بخش در ایجاد یک مدل کار آمد، انتخاب توصیف

کننده‌های مناسب است. در این مطالعه توصیفگرهای مورد نظر با استفاده از خروجی گو سین و نرم افزار دراگون بدست آمده و پس از بررسی همبستگی بین متغیرها ۴۵۱ توصیفگر به منظور بدست آمدن معادله کیوسار وارد محاسبات شدند. توصیفگرهای کوانتومی که از خروجی گوسین بدست آمده شامل انرژی کل گرمایی، انرژی الکترونی، آنتالپی، انتروپی، انرژی آزاد گیبس، ممان دو قطبی، انرژی بالاترین اربیتال مولکولی اشغال شده (HOMO)، انرژی پایین ترین اربیتال اشغال نشده (LUMO)، گپ انرژی، قطبش پذیری، الکترون‌گاتیوی (χ)، سختی (η)، الکترون دوستی (ω) و.. (۱۳) توصیفگرهای نرم افزار دراگون شامل انواع توصیفگرهای یک بعدی، دو بعدی و سه بعدی در نظر گرفته شدند.

مدلسازی: پس از انتخاب مناسب ترین توصیفگرها توسط روش افزایش مرحله‌ای برقراری ارتباط مناسب بین توصیفگرها و متغیر و فعالیت ترکیبات دارویی صورت گرفت. مدلسازی بر روی گروه آموزش به روش رگرسیون خطی چندگانه گام به گام انجام شد (۱۲).

این روش به کمک نرم افزار اس پی اس بر روی گروه آموزش پیاده شد و از میان ۴۵۱ توصیفگر موجود، آنهایی که بیشترین میزان R^2 را داشتند انتخاب شدند. به این ترتیب چندین معادله به دست آمد که با توجه به تعداد مولکول‌ها و تعداد توصیفگرهای انتخابی و همچنین پایین بودن میزان خطای آن‌ها انتخاب مدل مناسب صورت گرفت. برای بررسی دقت و صحت مدل از روش اعتبارسنجی داخلی و متقاطع استفاده شد. اعتبارسنجی داخلی مربوط به پارامترهای آماری دسته آموزش است که یکی از این پارامترها R یا ضریب همبستگی می‌باشد.

ساده‌ترین راه برای بررسی میزان همبستگی دو یا چند متغیر محاسبه ضریب همبستگی است. ضریب تعیین یا R^2 یک شاخص برای بیان دقت خط رگرسیون بدست آمده می‌باشد و نشان دهنده نسبت تغییرات متغیر وابسته توسط متغیر مستقل است. یک معادله قابل قبول مقدار ضریب همبستگی بیشتر از ۰/۶ را دارد.

یافته‌ها

مدل کیوساز: بین توصیف کننده‌ها و فعالیت‌های بازدارندگی مشتقات کومارین برای سری آموزش با استفاده از روش همبستگی خطی چندگانه رابطه زیر به عنوان بهترین مدل خطی بدست آمد:

$$pIC_{50} = 12.601 - 5.473 * Jhetp - 0.245 * RDF065m + 1.028 * nR = Ct - 0.389 * nCbH + 5.573 * SPH \quad (1)$$

طبق معادله به دست آمده فعالیت مهارکنندگی ترکیبات مورد بررسی به ۵ توصیفگر وابسته است. معادله بدست آمده نشان می‌دهد فعالیت این ترکیبات با ضریب منفی $nCbH$ ، $Jhetp$ ، $RDF065m$ ارتباط دارد، به این معنی که با افزایش مقادیر این توصیف‌گرها مقدار فعالیت مهارکنندگی کاهش می‌یابد. از طرفی فعالیت این ترکیبات با ضرایب مثبت $nR=Ct$ ، SPH بستگی داشته، یعنی با افزایش آنها، فعالیت این ترکیبات نیز افزایش می‌یابد. در جدول ۱ پارامترهای آماری برای معادله کیوساز به دست آمده نشان داده شده است.

باتوجه به جدول ۲، پارامترهای آماری R^2_{test} ، R^2 ، $RMSE$ ، MSE ، Q^2 برای انتخاب بهترین مدل استفاده شدند. در مدل بدست آمده $R^2=0/9$ و با توجه به بالا بودن مقدار F ، مدل بدست آمده یک مدل قابل قبول می‌باشد. همچنین پارامترهای $RMSE$ و Q^2 و R^2_{test} جهت نشان دادن اعتبار مدل بدست آمده و همچنین قدرت پیش بینی مدل محاسبه شدند. مقدار $Q^2=0/83$ که بالاتر از $0/5$ است نشان می‌دهد مدل بدست آمده از اعتبار خوبی برخوردار است. برای قدرت پیش بینی مدل از ترکیبات گروه آزمون استفاده شد و میزان R^2 این ترکیبات $0/59$ محاسبه شد که نشان دهنده توانایی خوب مدل در پیش بینی قدرت اثر ترکیبات است. جدول ۲ متغیرهای مؤثر در معادله

هر چقدر این مقدار به یک نزدیکتر باشد رابطه‌ی خطی بین متغیر وابسته و مستقل قوی‌تر است (۱۳). روش اعتبار سنجی متقاطع عمدتاً به عنوان یک روش اعتبار سنجی درونی محسوب می‌شود. این روش یک مدل جدید را با استفاده از دسته‌های کاهش یافته از دیتاها تحلیل می‌کند و برای پیشگویی خواص ترکیباتی که وارد مدل جدید نشده‌اند مورد استفاده قرار می‌گیرد. روند تا جایی ادامه می‌یابد که همه ترکیبات محذوف یک بار پیشگویی شوند. در این مرحله از تحقیق، روش Leave one out به کار گرفته شد. در این روش میزان اعتبار و دقت مدل با استفاده از Q^2_{LOO} ، مشخص می‌شود. در این روش میزان اعتبار و دقت مدل با استفاده از Q^2_{LOO} ، مشخص می‌شود ارزیابی نهایی مدل به وسیله اعتبار سنجی خارجی صورت می‌گیرد که مشخص می‌کند مدل کیوساز به دست آمده برای ترکیبات خارج از مدل تا چه حد توانایی پیش‌گویی دارد. این مرحله بر روی ۱۱ ترکیب موجود در دسته آزمون که طی مرحله جداسازی به صورت تصادفی انتخاب شده‌اند انجام گرفته است. به این ترتیب میزان فعالیت ترکیبات بدون انجام فرایندهای آزمایشگاهی تخمین زده می‌شود که این مقادیر با میزان فعالیت‌های به دست آمده از محاسبات آزمایشگاهی مقایسه می‌گردند. R^2_{Test} باید مقادیری بیش از $0/6$ را به خود اختصاص دهد تا نشان دهنده قدرت کافی و مطلوب مدل کیوساز به دست آمده باشد یکی دیگر از پارامترها برای محاسبه قدرت پیش‌گویی مدل بدست آمده در اعتبار سنجی خارجی، محاسبه (Root mean square error of prediction) است.

جدول ۱- پارامترهای آماری محاسبه شده برای مدل بدست آمده

Train		Test			Q ² LOO	
R	R ²	F	SEE	R ²	RMSEP	
0/95	0/90	۳۶/۸۵	0/۵۱	0/۷۰	0/۸۱	0/۸۳

جدول ۲- مقایسه مقادیر PIC50 تجربی و PIC50 محاسبه شده به همراه توصیفگرهای مناسب

Name	Status	Jhetp	RDF065m	nR=Ct	nCbH	SPH	PIC50 تجربی	PIC50 پیش بینی شده
۱	train	۴۵۵.۱	۶۵۳.۷	۱	۳	۷۶۴.۰	۰۳.۷	۸۸.۶
۲	train	۱۱۶.۱	۸۸۱.۱۰	۱	۸	۸۲۹.۰	۴۷.۶	۳۶.۶
۳	train	۱۲۹.۱	۹۶۶.۸	۱	۸	۸۴۱.۰	۸۹.۶	۸۳.۶
۴	train	۱۱۵.۱	۳۸۸.۹	۱	۸	۸۴۸.۰	۵۳.۶	۸۴.۶
۶	train	۱۱۲.۱	۶۶۷.۱۱	۱	۷	۸۴.۰	۸۹.۶	۶۴.۶
۷	train	۱۱.۱	۷۷۴.۱۰	۱	۷	۸۰۸.۰	۵۷.۶	۶۹.۶
۱۰	train	۱۰۶.۱	۷۰۵.۱۰	۱	۷	۷۹۲.۰	۷۸.۶	۶۴.۶
۱۲	train	-۹۷.۱	۵۲۱.۱۱	۱	۶	۷۹۴.۰	۶۵.۶	۸۹.۶
۱۴	train	۸۴۸.۱	-۸۷.۱۱	۱	۶	۹۶۲.۰	۶۸.۳	۸۳.۳
۱۵	train	۸۲۶.۱	۲۱۱.۱۲	۱	۶	۹۶۲.۰	۱۸.۴	۶۷.۳
۱۶	train	۷۱۵.۱	۹۴۹.۱۴	۱	۵	۹۲۷.۰	۱۹.۴	۸۰.۳
۱۷	train	۷۵۲.۱	۶۷۸.۱۱	۱	۷	۹۶۲.۰	۲۹.۳	۸۲.۳
۱۹	train	۷۹۶.۱	-۸۳.۱۲	۱	۶	۹۵۶.۰	۶۳.۳	۸۳.۳
۲۰	train	۸۸۳.۱	-۲۷.۹	۱	۶	۹۵۸.۰	۵۶.۴	۱۲.۴
۲۱	train	۹۴۴.۱	۸۲۲.۸	۱	۵	۸۳۳.۰	۴۸.۳	۵۳.۳
۲۲	train	-۱۶.۲	۹۰.۱۲	۰	۴	۱	۹۲.۴	۵۸.۵
۲۳	train	۱۹۸.۱	۸۷۲.۳	۰	۸	۹۳۲.۰	۲۶.۷	۱۸.۷
۲۵	train	۳۷۲.۱	-۶۲.۴	۰	۸	۷۱۹.۰	۶۷.۵	۹۹.۴
۲۷	train	۲۷۷.۱	۸۴۹.۴	۰	۷	۷۵۶.۰	۶۹.۵	۹۱.۵
۲۸	train	۲۴۶.۱	۸۲۱.۳	۱	۸	۷۶۳.۰	۷۴.۷	-۱.۷
۳۰	train	۲۶۹.۱	۲۶۴.۶	۱	۸	۷۵۸.۰	۸۶.۵	۲۶.۶
۳۱	train	۲۲۶.۱	۶۵۵.۳	۰	۸	۷۴۹.۰	۸.۵	-۰.۶
۳۲	train	۲۹۳.۱	۲۶۶.۵	۲	۸	۷۹۲.۰	۳۶.۸	۵۹.۷
۳۳	train	۴۵.۱	۷۸۷.۳	۲	۸	۸۲۶.۰	۷۹.۶	۲۸.۷
۳۴	train	۲۳۴.۱	۹۸۸.۵	۲	۸	۸۵۶.۰	۲۵.۸	-۰.۸
۳۵	train	۳۲.۱	۴۵۱.۶	۲	۸	۷۸۴.۰	۴۹.۶	۱۱.۷
۳۶	train	۱۵۲.۱	۴۶۲.۴	۲	۸	۸۰۲.۰	۴۶.۸	۶۲.۸
۳۸	train	۴۸۵.۱	۴۷۸.۵	۲	۸	۸۲۷.۰	۷۲.۶	۶۸.۶
۳۹	train	۵۳۴.۱	۵۴۱.۸	۲	۸	۷۸۸.۰	۲۸.۵	۴۵.۵
۴۰	train	۴۷۵.۱	۶۹۴.۱۳	۲	۷	۸۳۵.۰	۷۷.۴	۱۶.۵
۴۱	train	۷۵.۱	۲۰۷.۲	۲	۸	۹۸۹.۰	۵۵.۷	۹۴.۶
۴۲	train	۳۴۱.۱	۱۴۴.۱۰	۲	۷	۷۹۴.۰	۵۵.۷	۵۳.۶
۴۴	train	۲۹۳.۱	۱۶.۹	۲	۸	۷۹۱.۰	۴۸.۵	۶۳.۶
۵	test	۱۲۶.۱	۹۲.۱۳	۱	۷	۸۲۷.۰	۳۷.۶	۹۴.۵
۸	test	-۹۶.۱	۹۱۶.۱۰	۱	۷	۸۴۳.۰	۷۱.۶	۹۳.۶
۹	test	-۹۲.۱	۶۷۱.۱۲	۱	۷	۷۶۹.۰	۵۲.۶	۱۱.۶
۱۱	test	-۹۱.۱	۵۵۴.۱۱	۱	۷	۸۴۲.۰	۸۳.۶	۸۰.۶
۱۳	test	۸۶۵.۱	۷۴.۸	۱	۷	۹۵۵.۰	۱۳.۴	۸۸.۳
۱۸	test	۷۳۹.۱	۱۳۷.۱۳	۱	۶	۹۶۲.۰	۲۵.۳	۹۲.۳
۲۴	test	۲۳۳.۱	۲۴۶.۴	۰	۸	۹۶۸.۰	-۰.۷	۱۵.۷
۲۶	test	۲۵۱.۱	۲۳.۸	۰	۷	۹۷۳.۰	۱۷.۵	۴۴.۶
۲۹	test	۲۱.۱	۴۴۳.۶	۱	۷	۶۵۹.۰	۸۸.۷	۳۸.۶
۳۷	test	۱۷۷.۱	۱۹۵.۶	۲	۸	۸۲۱.۰	۴۶.۸	۱۶.۸
۴۳	test	۳۷۹.۱	۹۴۹.۶	۲	۷	۷۸۵.۰	۵۱.۵	-۰.۷

بررسی میزان همبستگی بین مقادیر پیش بینی شده و مقادیر تجربی نمودار همبستگی آن‌ها رسم شده است شکل ۱ نشان دهنده همبستگی بین مقادیر تجربی و پیش بینی شده می‌باشد که همبستگی قابل قبولی را با $R^2 = ۰/۸۹۷۶$ نشان می‌دهد.

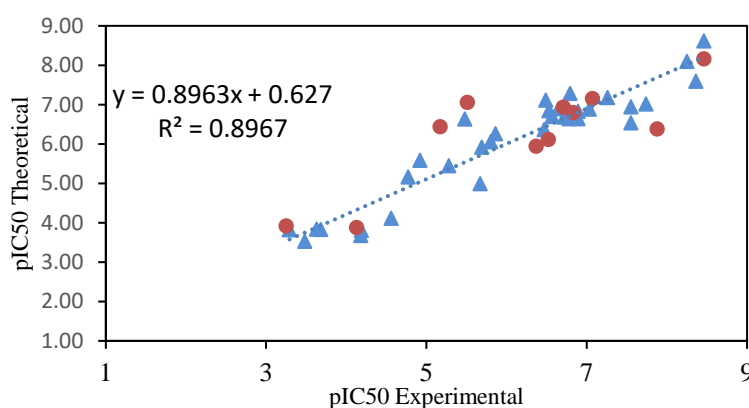
کیوسار بدست آمده، فعالیت تجربی و فعالیت پیش بینی شده را نشان می‌دهد.

مقایسه فعالیت تجربی و فعالیت پیش بینی شده: با جایگذاری مقادیر هر یک از ۵ متغیر موجود در معادله، میزان فعالیت پیش بینی شده بدست آمده سپس برای

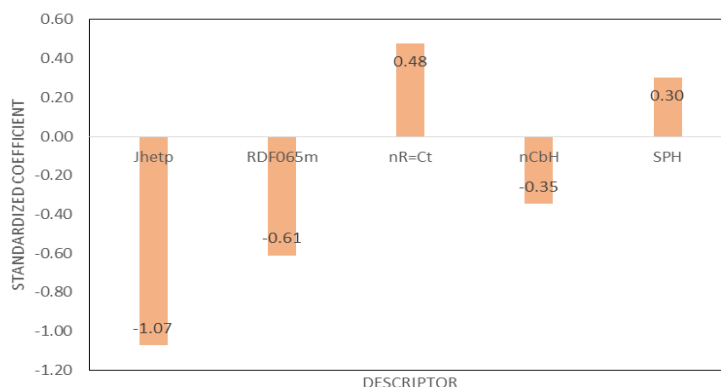
شوند. تفسیر مکانیکی مدل‌های QSAR به درک تأثیر توصیفگرها در پیش‌بینی فعالیت یا نقطه پایانی کمک می‌کند. دامنه‌ی کاربرد (Applicability Domain) برای اعتبارسنجی مدل‌های توسعه یافته QSAR توصیه می‌شود. تجزیه و تحلیل دامنه کاربرد به ما اجازه می‌دهد تا بررسی کنیم که آیا مدل QSAR ساخته شده می‌تواند برای هر مجموعه‌ای از مولکول‌ها اعمال شود یا خیر. پیش‌بینی یک نقطه پایانی مدل شده با استفاده از QSAR تنها زمانی مناسب است که مولکول پیش‌بینی شده در محدوده دامنه کاربرد توسعه یافته QSAR قرار گیرد. بنابراین، دامنه کاربرد را می‌توان به عنوان اطلاعات فضای فیزیکوشیمیایی، بیولوژیکی یا ساختاری بیان کرد که بر اساس آن مجموعه آموزشی مدل ساخته می‌شود و مدل برای پیش‌بینی مولکول‌های جدید در حوزه‌های خاص اعمال می‌شود (۱۴).

مقایسه اثر توصیفگرها: برای بررسی این که افزایش یا کاهش توصیفگرهای انتخابی، چقدر بر روی فعالیت مورد نظر تاثیرگذار است نمودار اثر متوسط توصیفگرها (شکل ۲) رسم شده است. محور yها بیانگر ضرایب استاندارد شده (β) و محور xها نیز نمایش دهنده‌ی توصیفگر انتخاب شده با استفاده از روش SW-MLR است. بر اساس این نمودار، توصیفگر Jhetp بیشترین تاثیر منفی و nR=Ct بیشترین تاثیر مثبت را دارد. در این نمودار توصیفگرهایی که بالای نمودار رسم شده‌اند به این معنی است که با افزایش مقدار توصیفگر، مقدار فعالیت افزایش یافته و برای آنها که پایین نمودار رسم شده‌اند با افزایش آنها، مقدار فعالیت کاهش می‌یابد.

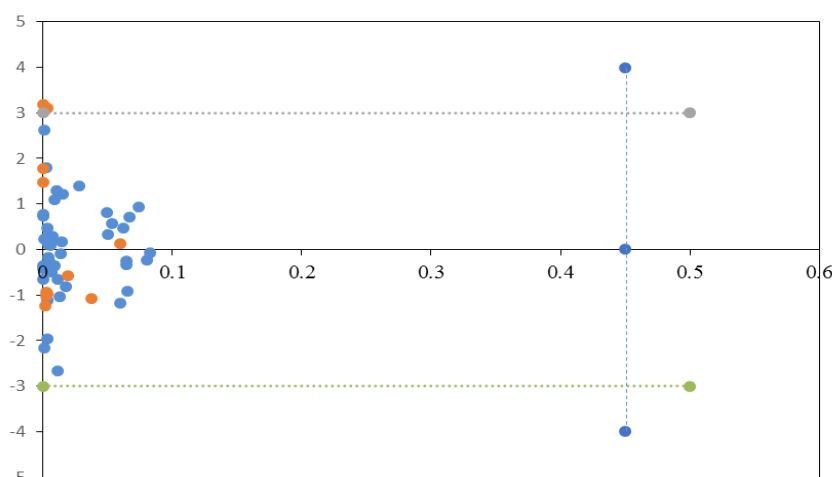
ارزیابی دامنه کاربرد مدل: پارامترها یا توصیفگرهای مورد استفاده در توسعه مدل‌های QSAR باید آنالیز



شکل ۱- مقادیر فعالیت پیش‌بینی شده در مقابل تجربی (دسته‌ی آموزش و دسته‌ی تست)



شکل ۲- نمودار اثر توصیفگر



شکل ۳- نمودار ویلیامز پلات

به دست نیامده است (۱۶). این تست ده بار انجام شد. مقدار R^2 و Q^2 محاسبه شده برای ترکیبات کمتر از ۰/۵ و مقدار cRp2، ۰/۸۸ به دست آمد. پارامترهای آماری این تست نشان دهنده این است که مدل ما بر اساس شانس و به صورت تصادفی به دست نیامده است. **بررسی اثر توصیفگرها (Thetp):** این توصیفگر از دسته توصیفگرهای توپولوژیکی و دویبعدی دراگون می باشد. این توصیف کننده شاخص بالابان مبنی بر وزن ماتریکس بر اساس قطبیت پذیری می باشد. شاخص بالابان با فرمول زیر محاسبه میشود.

$$J = \frac{B}{C+1} \sum_b (\sigma_i \sigma_j)_b^{-1/2} \quad (2)$$

در این فرمول c تعداد حلقه ها B عدد نمودار مولکولی b همه ی اتصالات مولکول $\sigma_i \sigma_j$ فاصله ی مربوط به اتم i و اتم j را نشان می دهد (۱۷). ضریب این توصیفگر در مدل به دست آمده منفی است یعنی با افزایش آن فعالیت مهارکنندگی ترکیبات کاهش می یابد. کمترین میزان این پارامتر برای ترکیبات ۸،۱۰ و ۱۱ و بیشترین آن برای ترکیب ۲۲ می باشد. در نتیجه به نظر می رسد با افزایش تعداد حلقه ها این پارامتر کاهش می یابد و استخلاف های دارای حلقه می توانند سبب کاهش این توصیف گر و افزایش فعالیت مهارکنندگی شوند. **RDF065** این تابع یکی از کمیت های مهم برای

برای ارزیابی دامنه کاربرد مدل از تکنیک خطای باقی مانده به روش لوریچ استفاده شده است. در این روش، نمودار ویلیامز پلات برای نشان دادن نتایج رسم می شود. در این نمودار محور X ها میزان لوریچ ترکیبات است و محور Y ها مقادیر باقی مانده استاندارد شده است. به اختلاف بین میزان پیش بینی شده و تجربی فعالیت مقدار باقی مانده گفته می شود. پراکندگی یکنواخت نقاط، حول محور افقی که بی سانگر خطای باقیمانده صفر است، نشان دهنده ی آن است که خطای سیستماتیکی در مدل به دست آمده وجود ندارد و همچنین کارآمد بودن مدل برای پیش بینی رفتار در روش های تجربی را نشان می دهد (۱۵) (شکل ۳).

ارزیابی مدل ارائه شده با آزمون Y تصادفی: در این روش متغیر وابسته به صورت تصادفی انتخاب می شود و مدل بر اساس توصیفگرهایی که در مدل اصلی به دست آمده ساخته می شود و برای شرایط جدید R^2 ، Q^2 محاسبه می شود. مقدار قابل قبول R^2 و Q^2 رای مدل تصادفی کمتر از ۰/۵ است. از طرف دیگر با محاسبه اختلاف بین ضریب تعیین مدل تصادفی و ضریب تعیین مدل اصلی (R^2)، مقدار cRp2 محاسبه می شود. مقدار بالاتر از ۰/۵ برای cRp2 نشان دهنده این است که مدل به صورت تصادفی و بر اساس شانس

جدول ۳- آزمون Y تصادفی

MODEL TYPE	R	R ²	Q ² LOO
Original	۹۴۶۹۱۶۴.۰	۸۹۶۶۵۰۷.۰	۸۳۲۱۵۸۹.۰
۱Random	۳۵۱۷۵۷۲.۰	۱۲۸۷۰۶۷.۰	۳۰۰۹۹۹۳.۰
۲Random	۳۷۹۲۶۱۳.۰	۱۴۳۸۳۹۱.۰	۳۰۱۰۰۸۰.۰
۳Random	۴۶۱۸۰۷۵.۰	۲۱۳۲۶۶۲.۰	۲۷۴۲۵۸۴.۰
۴Random	۳۳۲۹۴۰۴.۰	۱۱۰۸۴۹۳.۰	۳۶۹۷۵۸۳.۰
۵Random	۲۹۵۴۳۴۴.۰	۰۸۷۲۸۱۵.۰	۳۸۰۹۰۱۲.۰
۶Random	۴۳۵۵۶۴۰.۰	۱۸۹۷۱۶۰.۰	۵۲۴۸۳۰۱.۰
۷Random	۴۴۰۸۵۶۶.۰	۱۹۴۲۵۴۵.۰	۲۱۶۶۳۳۷.۰
۸Random	۳۸۱۸۵۸۱.۰	۱۴۵۸۱۵۶.۰	۳۳۸۹۶۱۶.۰
۹Random	۹۴۶۹۱۶۴.۰	۸۹۶۶۵۰۷.۰	۸۳۲۱۵۸۹.۰
۱۰Random	۳۵۱۷۵۷۲.۰	۱۲۸۷۰۶۷.۰	۳۰۰۹۹۹۳.۰

سیستم‌های مسطح و غیر مسطح و نوع اتم‌ها فراهم می‌کند. RDF065 تابع توزیع شعاعی در فاصله 065 و برا ساس جرم اتمی محاسبه شده است (۱۷). ضریب RDF065m در معادله کیوسار منفی است که نشان می‌دهد با افزایش مقدار این متغیر فعالیت این ترکیبات کاهش می‌یابد. از مقایسه مقادیر این پارامتر در جدول ۳ در هر کدام از این سه دسته ترکیبات می‌توان نتیجه گرفت که هرچه جرم مولکولی ساختار مورد نظر بیشتر باشد مقدار این توصیفگر بیشتر بوده و فعالیت بهاری کمتر است. بنابراین برای افزایش قدرت مهارکنندگی باتوجه به تاثیر این توصیفگر باید به دنبال مشتقاتی بود که دارای جرم مولکولی کمتری باشند. باتوجه به مقادیر بدست آمده در جدول ۳ می‌توان مشاهده نمود که به عنوان مثال ترکیب ۱۴ بیشترین مقدار RDF065 دارای فعالیت به نسبت کمتری بوده و ترکیباتی مثل ترکیب ۲۳ یا ۳۸ که فعالیت بالاتری دارند دارای مقادیر کوچکی از این پارامتر می‌باشند.

$nR=Ct$ این توصیفگر مربوط به توصیفگرهای یک بعدی نرم افزار دراگون می‌باشد. تفاوت ساختار شیمیایی ترکیبات مورد مطالعه را حساب می‌کند. اعداد مربوط به این توصیفگر بیانگر تعداد اتم‌های کربن سوم آلیفاتیک می‌باشد (۱۸). ضریب این توصیفگر در معادله بدست آمده مثبت است. همانطور که از نتایج بدست آمده در جدول ۳ پیداست هر چه تعداد اتم‌های کربن

تعیین ساختار سیستم، تابع توزیع شعاعی RDF است. که ساختار موضعی و نظم اتم‌ها را نسبت به یکدیگر مشخص می‌کند و به دما و چگالی سیستم وابسته است. می‌توان گفت RDF یا توزیع شعاعی مربوط به یک دسته از اتم‌ها، معادل توزیع احتمال یافتن یک اتم در فضای کروی به شعاع R است. رابطه زیر نحوه محاسبه تابع RDF را نشان می‌دهد:

$$g(R) = f \sum_{i=1}^{N-1} \sum_{j>i}^N A_i A_j e^{-B(R-r_{ij})^2}$$

که f یک فاکتور مقیاس، و N تعداد اتم‌های مولکول است. همچنین r_{ij} فاصله i بین دو اتم نو A و A یک ویژگی اتمی وزن دار نشده (u جرم اتمی، m حجم و اندروالس، v الکترونگاتیویته، e قطبش پذیری p است. یک فاکتور تسهیل کننده است که توزیع احتمال فاصله بین اتمی را مشخص می‌کند و می‌توان از آن به فاکتور دما برای تعریف جنبش اتمی تعبیر کرد $g(R)$ در نقاط گسسته‌ای با فواصل معین محاسبه و با مجموعه‌ای از کدهای RDF با ویژگی‌های مختلف اتمی می‌توان ساختار سه بعدی یک مولکول را به طور واضح توصیف کرد. این توصیفگرها اهمیت توزیع اتم‌ها در مولکول روی اثر بازدارندگی آن را نشان می‌دهند. همچنین این توصیفگرها علاوه بر فاصله اتمی، اطلاعات با ارزشی در مورد فاصله پیوند، انواع حلقه،

آمده از معادله کیوسار سنتز و طراحی نمود. از آن جایی که روش‌های محاسباتی سریعتر و در دسترس‌تر نسبت به روش‌های تجربی می‌باشند، می‌توان با صرف زمان کمتری به پیش‌بینی ساختارهای لازم برای طراحی داروهای موثر پرداخت.

References

1. Zhang C, Yang K, Yu S, Su J, Yuan S, Han J, et al. Design, synthesis and biological evaluation of hydroxypyridinone-coumarin hybrids as multimodal monoamine oxidase B inhibitors and iron chelates against Alzheimer's disease. *Eur J Med Chem.* 2019;180:367-82.
2. Youdim MB, Edmondson D, Tipton KF. The therapeutic potential of monoamine oxidase inhibitors. *Nat Rev Neurosci.* 2006;7(4):295-309.
3. Shih J, Chen K, Ridd M. Monoamine oxidase: from genes to behavior. *Ann Rev Neurosci.* 1999;22(1):197-217.
4. Berger EA, Murphy PM, Farber JM. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Ann Rev Immunol.* 1999;17(1):657-700.
5. Fahey JL, Taylor JM, Detels R, Hofmann B, Melmed R, Nishanian P, et al. The prognostic value of cellular and serologic markers in infection with human immunodeficiency virus type 1. *N Eng J Med.* 1990;322(3):166-72.
6. Nikolaev DM, Manathunga M, Orozco-Gonzalez Y, Shtyrov AA, Guerrero Martinez YO, Gozem S, et al. Free Energy Computation for an Isomerizing Chromophore in a Molecular Cavity via the Average Solvent Electrostatic Configuration Model: Applications in Rhodopsin and Rhodopsin-Mimicking Systems. *J Chem Theory Comput.* 2021;17(9):5885-95.
7. Roy K, Kar S, Das RN. Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment: Academic press; 2015.
8. Kesar S, Paliwal SK, Mishra P, Chauhan M. Quantitative structure-activity relationship analysis of selective rho kinase inhibitors as neuro-regenerator agent. 2019.
9. Hu YH, Yang J, Zhang Y, Liu KC, Liu T, Sun J, et al. Synthesis and biological evaluation of 3-(4-aminophenyl)-coumarin derivatives as potential anti-Alzheimer's disease agents. *J Enz Inhibit Med Chem.* 2019;34(1):1083-92.
10. Gnerre C, Catto M, Leonetti F, Weber P, Carrupt PA, Altomare C, et al. Inhibition of

سوم آلیفاتیک بیشتر باشند میزان فعالیت ضدآلزایمری ترکیبات نیز بیشتر خواهد بود.

nCbH این توصیفگر همانند توصیفگر قبلی جز توصیفگرهای یک بعدی نرم افزار دراگون است. این توصیفگر قادر به فشرده سازی اطلاعات در سایت‌های واکنشی احتمالی در حلقه‌های آروماتیک است و نشان دهنده‌ی تعداد بنزن جایگزین نشده اتم کربن می‌باشد (۱۹). با توجه به ضریب منفی این توصیفگر در معادله کاهش تعداد این توصیفگر منجر به افزایش فعالیت مهارکنندگی می‌شود.

SPH این توصیفگر مربوط به توصیفگرهای سه بعدی نرم افزار دراگون می‌باشد و بیانگر شکل هندسی ساختار است. SPH یک توصیف کننده‌ی نامتقارن و نامساوی بودن است که به عنوان تابعی از مقادیر ویژه‌ی ماتریکس کوواریانس محاسبه می‌شود که مقادیر آن از صفر تا یک برای مولکول‌های تخت تا کاملاً کروی را شامل می‌شود (۲۰). ضریب این توصیفگر در معادله مثبت است یعنی با افزایش آن فعالیت بیشتر می‌شود. از بین ۴۴ ترکیب بررسی شده ساختار ۲۲ بیشترین مقدار این توصیفگر یعنی عدد یک را به خود اختصاص داده است که نشان دهنده‌ی متقارن بودن این ساختار است.

نتیجه‌گیری

بر اساس نتایج مشاهده شده در این ارزیابی، بین ساختار و فعالیت مشتقات کومارین ارتباط معنادار آماری وجود دارد، با توجه به مدل بدست آمده توصیفگرهای یک بعدی ($nR=Ct$, $nCbH$), دو بعدی $Jhetp$ و سه بعدی ($RDF065m$, SPH) از نرم افزار دراگون می‌توانند بر فعالیت این ترکیبات موثر باشند. با توجه به توضیحات و بررسی نتایجی که در بالا ارائه شد می‌توان به طور کلی نتیجه گرفت که با توجه به مقادیر فعالیت‌های پیش‌بینی شده در جدول ۳، ترکیب ۳۶ و ۳۷ بیشترین فعالیت و ترکیب ۲۱ و ۱۵ کمترین فعالیت ضدآلزایمری را دارند. مدل بدست آمده کمک می‌کند تا بتوان در آینده داروهای جدید با فعالیت بیشتر را بر اساس عوامل موثر بر فعالیت بدست

monoamine oxidases by functionalized coumarin derivatives: biological activities, QSARs, and 3D-QSARs. *J Med Chem.* 2000;43(25):4747-58.

11. Frisch M, Trucks G, Schlegel H, Scuseria G, Robb M, Cheeseman J, et al. Gaussian 09, Revision D. 01, Gaussian, Inc., Wallingford CT. See also: URL: <http://www.gaussian.com>. 2009.

12. Chtita S, Ghamali M, Ousaa A, Aouidate A, Belhassan A, Taourati AI, et al. QSAR study of anti-Human African Trypanosomiasis activity for 2-phenylimidazopyridines derivatives using DFT and Lipinski's descriptors. *Heliyon.* 2019;5(3).

13. Gramatica P, Chirico N, Papa E, Cassani S, Kovarich S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. Wiley Online Library; 2013.

14. Aminae H, Hosseini S, Davood A, Askarizadeh E. QSAR study of indole derivatives as active agents against *Candida albicans*: a DFT calculation. *Journal of Receptors and Signal Transduction.* 2022;42(6):614-22.

15. Wang YN, Chen J, Li X, Wang B, Cai X, Huang L. Predicting rate constants of hydroxyl radical reactions with organic pollutants: algorithm, validation, applicability domain, and mechanistic interpretation. *Atmospher Environ.* 2009;43(5):1131-5.

16. Rücker C, Rücker G, Meringer M. y-Randomization and its variants in QSPR/QSAR. *J Chem Inform Model.* 2007;47(6):2345-57.

17. Helguera AM, Combes RD, Gonzalez MP, Cordeiro M. Applications of 2D descriptors in drug design: a DRAGON tale. *Curr Topics Med Chem.* 2008;8(18):1628-55.

18. Valenzuela L, del Valle JM, de la Fuente J. Modelling of carotenoids solubility in supercritical carbon dioxide using quantitative structure-property relationships. III Iberoamerican Conference on Supercritical Fluids Cartagena de Indias (Colombia), 2013.

19. Roy PP, Kovarich S, Gramatica P. QSAR model reproducibility and applicability: A case study of rate constants of hydroxyl radical reaction models applied to polybrominated diphenyl ethers and (benzo-) triazoles. *J Comput Chem.* 2011;32(11):2386-96.

20. Molnar M, Brahmabhatt H, Rastija V, Pavić V, Komar M, Karnaš M, et al. Environmentally friendly approach to Knoevenagel condensation of rhodanine in choline chloride: Urea deep eutectic solvent and QSAR studies on their antioxidant activity. *Molecules.* 2018;23(8):1897.