Journal of Ilam University of Medical Sciences

P-ISSN: 1563-4728
E-ISSN: 2588-3135

Quantitative structure-activity relationship study of some angiotensin-converting enzyme inhibitor drugs in the treatment of hypertension based on Monte Carlo optimization method

Shahram Lotfi [1*], Shahin Ahmadi [2], Ali Azimi [3]

[1] Dept of Chemistry, Payame Noor University (PNU), Tehran, Iran
[2] Dept of Pharmaceutical Chemistry, Faculty of Pharmaceutical Chemistry, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran
[3] Dept of Chemistry, Science and Research Branch, Islamic Azad University, Tehran, Iran

Article Info

Article type:
Research article

Article History:
Received: Dec. 28, 2024
Received in revised form: Feb. 02, 2025
Accepted: Mar. 10, 2025
Published Online: May. 26, 2025

**** Correspondence to:***
Shahram Lotfi
Dept of Chemistry, Payame Noor University (PNU), Tehran, Iran

Email:
Sh.lotfi@pnu.ac.ir

A B S T R A C T

Introduction: Hypertension is a severe cardiovascular disease affecting human health, and its control and prevention are crucial for global health. Angiotensin-converting-enzyme inhibitors are primarily used in treating this condition. The aim of this work was to develop quantitative structure-activity relationship (QSAR) models to predict the activity of some chemical compounds as angiotensin I-converting enzyme inhibitors using CORAL software.

Materials & Methods: In this study, quantitative structure-activity relationship to predict the inhibitory activity of the data set containing 255 angiotensin-converting enzyme inhibitor compounds based on the algorithm Monte Carlo was studied. The input file of CORAL software contains the structures of compounds with SMILES symbols along with their inhibitory activity, which were randomly divided into four sets, including active training, passive training, calibration and validation sets. The whole data set is randomly divided into three splits, and an individual model was created for each division.

Results: A hybrid optimal descriptor computed from SMILES and molecular hydrogen-suppressed graphs is employed to construct QSAR models. Four target functions, i.e., TF_0 (WIIC = WCII = 0), TF_1 (WIIC = 0.3 and WCII = 0), TF_2 (WIIC = 0 and WCII = 0.3), and TF_3 (WIIC = WCII = 0.3), are employed to build 12 QSAR models.

Conclusion: The study analyzed experimental data on chemical compounds as inhibitors of angiotensin I-converting enzyme. The QSAR model of split 2 calculated by TF_3 was found to be the best model, and it identified structural attributes responsible for the increase and decrease of inhibitory activity.

Keywords: QSAR models, Angiotensin-converting enzyme inhibitors, Hypertension disease, CORAL software

How to cite this paper: Lotfi Sh, Ahmadi Sh, Azimi A. Quantitative structure-activity relationship study of some angiotensin-converting enzyme inhibitor drugs in the treatment of hypertension based on Monte Carlo optimization method. *Journal of Ilam University of Medical Sciences*. 2025;33(2):26-44.

Introduction

The complications like hypertension, dysrhythmia, heart failure (HF), myocardial infarction, etc. are pivotal factors linked to cardiovascular diseases (CVD), accounting for a substantial portion (32 %) of total mortality in both developed and developing nations (1). High blood pressure, or hypertension, is a condition of sustained increase in blood pressure levels and is the primary risk factor for CVD (2). The renine-angiotensin system is a hormone system that regulates blood pressure and fluid balance through Angiotensin-Converting-Enzyme-I (ACE-I) inhibition, and plays an important role in the pathophysiology of CVDs such as congestive heart failure and hypertension (2). Angiotensin-converting enzyme inhibitors are a group of drugs used in the treatment of cardiac pathologies like hypertension, congestive heart failure, left ventricular dysfunction and myocardial infarction (3). ACE inhibitors act as potent vasodilators by blocking the conversion of angiotensin I to angiotensin II, thereby lowering blood pressure by generating prolonged

© The Author(s)
Publisher: Ilam University of Medical Sciences

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

hypotensive responses (2). Currently, numerous drugs as ACE inhibitors have been synthesized to treat hypertension disease (2). However, experimental methods for drug discovery are expensive and time-consuming (4). In order to save time, reduce costs, and eliminate the need for extensive equipment, the development of in silico methods (e.g., Quantitative Structure-Activity Relationship (QSAR)) to predict the activity/property of drugs prior to the clinical stage would be beneficial (4-5). QSAR, are advanced computational techniques that relate chemical structure to specific properties or biological activity using mathematical models based on molecular descriptors (5). Complications such as hypertension, arrhythmia, HF, and myocardial infarction are critical factors associated with CVDs, contributing to approximately 32% of total mortality in both developed and developing countries (1-4). Hypertension, defined as a sustained elevation in blood pressure, is considered the primary risk factor for CVD (2-5). The renin-angiotensin system plays a key role in regulating blood pressure and fluid balance through the inhibition of ACE-I, and it is closely linked to the pathophysiology of disorders such as hypertension and congestive heart failure (2). ACE inhibitors are a class of pharmaceutical agents widely used in the treatment of various cardiac conditions, including hypertension, congestive heart failure, left ventricular dysfunction, and myocardial infarction (5). These drugs act as potent vasodilators by inhibiting the conversion of angiotensin I to angiotensin II, thereby reducing blood pressure through sustained hypotensive effects (5). Although numerous ACE-I drugs have been synthesized for hypertension management, traditional drug discovery methods remain time-consuming and costly (3,4). The aim of this work was to develop quantitative structure-activity relationship (QSAR) models to predict the activity of some chemical compounds as angiotensin I-converting enzyme inhibitors using CORAL software.

Methods

TThe SMILES notation of 255 agents, along with numerical data on their ACE-I inhibitory activity, was acquired from a study published by Shah et al. The inhibition potential of compounds was defined as half-maximal inhibitory concentration in nM (IC50). The numerical values of activity were converted to a − logIC50 (pIC_{50}). The pIC_{50} was taken as the dependent parameter for constructing QSAR models with the CORAL software. Three splits were made from the dataset and each split was further randomly divided into four sets, i.e., active training, passive training, calibration, and validation sets. A dataset of 255 compounds, including their SMILES representations and ACE-I inhibitory activities (IC50 values in nM), was obtained from a study by Shah et al. The activity data were transformed into a negative logarithmic scale (pIC_{50} = −logIC50) and used as the dependent variable in QSAR model construction via CORAL software. The dataset was divided into three distinct splits, with each split further partitioned into four subsets: active training, passive training, calibration, and validation sets. The hybrid descriptor DCW (T*, N*) was used for model development, combining features derived from SMILES and molecular graphs. Monte Carlo optimization was employed to determine the correlation weights (CW) for the descriptors. Four target functions ($_{TF0}$, TF_1, TF_2, and TF_3) were evaluated, and the statistical performance of models based on each function was compared.

Results

The optimal descriptors from the hybrid model resulting from the combination of SMILES descriptors and hydrogen-free molecular graph were used to build QSAR models. Four objective functions, namely TF_0 (WIIC=WCII = 0), TF_1 (WIIC=0.3), TF_2 (WIIC = 0, WCII = 0.3) and TF_3 (WIIC = WCII = 0.3) were used to build 12 QSAR models. The statistical results of each objective function were compared with each other.

Conclusion

The study evaluated predictive quality of models using the TF_3 target function, with the best performance observed in models developed using this function. Various statistical metrics were used to ensure robustness, including coefficient of determination, IIC, CII, CCC, and cross-validated coefficient. Evaluation criteria like root mean square error, mean absolute error, F-statistic, mean predictive squared correlation coefficient, delta, and Y-randomization tests were also used. The applicability domain of each model was examined using the statistical defect value, denoted as d(A), to define reliable predictions.

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

Structural features from SMILES notation and molecular graph invariants were also analyzed to identify chemical substructures contributing to the enhancement or reduction of pIC_{50} values.

Authors' Contribution

Conceptualization, Methodology, Validation, Formal Analysis, Investigation: SL, SA, Software, Resources, Data Curation: SL, SA, AA, Writing– Original Draft Preparation, Writing– Review & Editing, Visualization, Supervision: SL, AA, Project Administration: SL.

Ethical Statement

No human or animal subjects were used in this study, and therefore, no ethical approval was required. The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

This research received no financial support.

Acknowledgment

The authors thank the all of participants and persons who help in this project.

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

مجله دانشگاه علوم پزشکی ایلام

JIUMS

P-ISSN: ۱۵۶۳-۴۷۲۸

E-ISSN: ۲۵۸۸-۳۱۳۵

مطالعهٔ ارتباط کمی ساختار – فعالیت برخی از داروهای مهارکنندهٔ آنزیم مبدل آنژیوتانسین I در کنترل فشارخون بر اساس روش بهینه‌سازی مونت‌کارلو

شهرام لطفی [۱*]، شهین احمدی [۲]، علی عظیمی [۳]

[۱] گروه شیمی، دانشگاه پیام نور، تهران، ایران

[۲] گروه شیمی، واحد علوم پزشکی تهران، دانشگاه آزاد اسلامی، تهران، ایران

[۳] گروه شیمی، واحدعلوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

اطلاعات مقاله

نوع مقاله: پژوهشی

تاریخ دریافت: ۱۴۰۳/۱۰/۰۸

تاریخ ویرایش: ۱۴۰۳/۱۱/۱۴

تاریخ پذیرش: ۱۴۰۳/۱۲/۲۰

تاریخ انتشار: ۱۴۰۴/۰۳/۰۵

نویسنده مسئول:
شهرام لطفی
گروه شیمی، دانشگاه پیام نور، تهران، ایران

Email:
Sh.lotfi@pnu.ac.ir

چکیده

مقدمه: فشارخون به‌عنوان یکی از جدی‌ترین بیماری‌های قلبی عروقی مطرح است که تأثیر نامطلوبی بر سلامت انسان دارد. کنترل و پیشگیری از این عارضه برای سلامت جهان حیاتی تلقی شده است. مهارکننده‌های آنزیم مبدل آنژیوتانسین گروهی از داروها هستند که عمدتاً در کنترل بیماری‌های قلبی عروقی استفاده می‌شوند. هدف از این کار توسعهٔ مدل‌های ارتباط کمی ساختار-فعالیت برای پیش‌بینی فعالیت برخی ترکیبات شیمیایی، به‌عنوان مهارکنندهٔ آنزیم مبدل آنژیوتانسین I در کنترل فشارخون، بر اساس روش بهینه‌سازی مونت‌کارلو بود.

مواد و روش‌ها: در این مطالعه، ارتباط کمی ساختار-فعالیت برای پیش‌بینی فعالیت مهارکنندگی مجموعه داده، شامل ۲۵۵ ترکیب شیمیایی مهارکنندهٔ آنزیم مبدل آنژیوتانسین I، بر اساس الگوریتم مونت‌کارلو مطالعه شد. فایل ورودی نرم‌افزار CORAL شامل ساختارهای ترکیبات به نماد SMILES به همراه فعالیت مهارکنندگی آن‌ها است که به‌طور تصادفی به چهار گروه شامل آموزش فعال، آموزش غیرفعال، کالیبراسیون و اعتبارسنجی تقسیم گردیدند. این تقسیم‌بندی سه بار به‌صورت تصادفی تکرار شد و برای هر تقسیم‌بندی یک مدل جداگانه ایجاد گردید.

یافته‌های پژوهش: توصیفگرهای بهینه از مدل هیبریدی حاصل از ترکیب میان توصیفگرهای SMILES و گراف مولکولی بدون هیدروژن برای ساخت مدل‌های QSAR به‌کار برده شد. چهار تابع هدف یعنی TF_0 (WIIC=WCII=0)، (WIIC=0.3) TF_1، (WIIC=0, WCII=0.3) TF_2 و TF_3 (WIIC=WCII=0.3) برای ساخت ۱۲ مدل QSAR استفاده گردید. نتایج آماری هر تابع هدف با یکدیگر مقایسه شد.

بحث و نتیجه‌گیری: در این مطالعه، مدل‌سازی داده‌های تجربی برخی از ترکیبات شیمیایی به‌عنوان مهارکنندهٔ آنزیم مبدل آنژیوتانسین I انجام گردید. کیفیت آماری مدل ارتباط کمی ساختار-فعالیت ایجادشده با استفاده از تابع هدف TF_3 برای شکاف ۲ بهتر از مدل‌های دیگر بود؛ بنابراین، به‌عنوان بهترین مدل در نظر گرفته شد. با استفاده از مدل به‌دست‌آمده، ویژگی‌های ساختاری مسئول افزایش و کاهش فعالیت مهارکنندگی شناسایی گردید.

واژه‌های کلیدی: مدل QSAR، مهارکنندهٔ آنزیم مبدل آنژیوتانسین I، بیماری فشارخون، نرم‌افزار کورال

استناد: لطفی شهرام، احمدی شهین، عظیمی علی. مطالعهٔ ارتباط کمی ساختار- فعالیت برخی از داروهای مهارکنندهٔ آنزیم مبدل آنژیوتانسین I در کنترل فشارخون بر اساس روش بهینه‌سازی مونت‌کارلو. *مجله دانشگاه علوم پزشکی ایلام*، خرداد ۱۴۰۴؛ (۲)۳۳ : ۴۴-۲۶.

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

CC BY NC

حق‌مؤلف © نویسندگان

مقدمه

بیماری‌های قلبی عروقی (CVD) به‌عنوان بزرگ‌ترین علت مرگ‌ومیر در سراسر جهان شناخته‌شده است. فشارخون بالا یا پرفشاری خون وضعیت افزایش مداوم سطح فشارخون است و عامل خطر اولیه برای CVD است. در حال حاضر، انواع مختلفی از داروها مانند مهارکننده‌های رنین، مسدودکننده‌های گیرندهٔ آنژیوتانسین II (ARBs) و مهارکننده‌های آنزیم مبدل آنژیوتانسین (ACEIs) برای کنترل فشارخون بالا وجود دارد. آنژیوتانسینوژن یک پروتئین سنتزشده در کبد است که در خون آزاد می‌شود و به‌وسیلهٔ هورمون رنین به آنژیوتانسین I تبدیل می‌گردد؛ سپس آنژیوتانسین I توسط آنزیم مبدل آنژیوتانسین (ACE) در ریه، به آنژیوتانسین II تبدیل می‌شود. عملکرد نهایی هورمون آنژیوتانسین II از طریق انقباض عروقی، افزایش‌دهندهٔ فشارخون است (۲، ۱)؛ بنابراین، ACE دارای یک عملکرد اساسی در تنظیم فشارخون در نظر گرفته می‌شود و موادی با اثر بازدارنده بر فعالیت ACE را می‌توان به‌عنوان کنترل بالقوهٔ فشارخون بالا در نظر گرفت.

تاکنون تعداد بسیاری از مهارکننده‌های ACE قوی و خاص به‌عنوان داروهای خوراکی سنتز شده‌اند که در کنترل فشارخون بالا استفاده می‌گردند. در سال ۱۹۷۷، کوشمن و همکاران (۳) برای اولین بار کاپتوپریل را به‌عنوان مهارکنندهٔ ACE سنتز کردند. ازآنجاکه این دارو برای استفاده در سال ۱۹۸۱ تأیید شد، نقش مهارکننده‌های ACE در کنترل بالینی فشارخون بالا به‌طورکلی توسط جامعهٔ پزشکی شناخته‌شده است. در حال حاضر، داروهای سنتزی مانند کاپتوپریل، انالاپریل، لیزینوپریل و سایر مهارکننده‌های ACE برای کنترل فشارخون بالا استفاده می‌گردد که اثر بالایی بر کاهش فشارخون دارند (۴). بااین‌حال، مصرف برخی از این داروها با عوارض جانبی مانند سرفهٔ خشک، واکنش‌های اگزمایی، افت فشارخون، اختلالات چشایی و آنژیوادم رودهٔ کوچک همراه است (۷–۵). بروز تقریباً همهٔ این آثار جانبی را می‌توان با طراحی ACE I با میل ترکیبی بیشتر به محل فعال ACE، مقادیر IC50 کمتر و سمیت کمتر کاهش داد؛ بنابراین، مطالعاتی که بر مهارکنندهٔ ACE تمرکز دارند، به‌سوی توسعهٔ داروهایی با عوارض جانبی کمتر برای ایجاد راهبردی جایگزین حرکت می‌کنند (۸).

فرایند توسعهٔ دارو اغلب به علت محدودیت‌های مختلف در زمان و منابع، با تأخیر مواجه می‌شود. از سویی، رویکرد تجربی برای بررسی رابطهٔ میان فعالیت‌ها و ساختارهای مولکولی زمان‌بر و پرزحمت است. مطالعات شیمی انفورماتیک به‌طور فزاینده‌ای برای رسیدگی به این چالش استفاده می‌گردد. شیمی انفورماتیک با استفاده از روش‌های محاسباتی طیف وسیعی از کاربردها، ازجمله کشف ترکیبات دارویی جدید، شناسایی اهداف دارویی و بهینه‌سازی خواص دارویی یا سنتیک دارویی ترکیبات شیمیایی موجود با فعالیت‌های بیولوژیکی را ارائه می‌کند. در میان ابزارهای مختلف در شیمی انفورماتیک، رابطهٔ کمی ساختار–فعالیت (QSAR) به‌عنوان رویکردی برجسته و به‌طور گسترده پذیرفته‌شده است. امروزه شیمی دارویی بر روش‌های محاسباتی متعددی تمرکز دارد که از نظر زمان و هزینه مقرون‌به‌صرفه هستند (۱۱–۹). هریک از روش‌های نظری ویژگی‌های خاص خود را دارند. روش مدل‌سازی QSAR نقش مهمی در طراحی و توسعهٔ دارو دارد (۱۳، ۱۲). مدل QSAR یک معادلهٔ ریاضی است که ساختار شیمیایی مولکول‌های مطالعه را به خواص فیزیکی، شیمیایی یا فعالیت دارویی آن‌ها از طریق توصیفگرهای شیمیایی تولیدشده از ساختار مولکولی مرتبط می‌کند (۱۴).

مدل‌های ایجادشده با استفاده از QSAR می‌تواند در تخمین اثربخشی داروهای مطالعه‌شده از سوی محققان و متخصصان شیمی دارویی، حتی پیش از سنتز آن‌ها مؤثر باشد؛ بنابراین، با صرف هزینه و زمان کمتر و پیش از انجام آزمایش‌های برون‌تن یا درون‌تن می‌توان آثار بیولوژیک احتمالی آن‌ها را پیش‌بینی کرد. با استفاده از روش‌های محاسباتی، الگوهای ساختاری مهارکننده‌ها را می‌توان بهتر شناسایی کرد و بینش بهتری دربارهٔ ساختار اثر دارو ارائه داد (۱۵)؛ بنابراین، با استفاده از این روش، تعداد فراوانی از

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

ترکیبات ازجمله آن‌هایی را که هنوز سنتز نشده‌اند، می‌توان بررسی نمود تا ساختارهایی با خصوصیات مدنظر برای سنتز و آزمایش انتخاب گردند.

پیش از مطالعات تجربی روی مهارکننده‌های ACE، باید یک روش جایگزین در نظر گرفته شود که بتواند در آن، رابطهٔ میان فعالیت‌های زیستی و ساختارهای شیمیایی استفاده گردد. برای بررسی سازوکار میان ساختار و فعالیت زیستی ACE I، پرابهاکار و گوپتا (۱۶) اولین مطالعهٔ QSAR را روی مهارکننده‌های ACE انجام داده‌اند.

نرم‌افزار (http://www.insilico.eu/coral) CORAL نرم‌افزاری کاربردی برای ایجاد مدل‌های QSAR، به‌منظور پیش‌بینی نقاط پایانی مختلف است (۲۰-۱۷). در این نرم‌افزار، از نماد SMILES از ساختار مولکولی ترکیبات برای ایجاد مدل‌های QSAR استفاده می‌شود؛ همچنین از نماد SMILES برای تعیین وزن همبستگی (DCW) توصیفگرها با کمک روش مونت‌کارلو استفاده می‌گردد. ضریب همبستگی میان توصیفگر بهینه (وزن‌های همبستگی) و نقطهٔ پایانی توسط این نرم‌افزار محاسبه می‌شود و متعاقباً از آن برای تولید تابع هدف استفاده می‌گردد. گزارش‌های پیشین نشان داده‌اند که شاخص شدت همبستگی (CII) و شاخص ایدئال همبستگی (IIC) برای افزایش قابلیت پیش‌بینی مدل‌های QSAR به‌کار گرفته شده‌اند (۲۳-۲۱). هنگامی‌که CII و IIC باهم به‌کار می‌روند، کیفیت بهینه‌سازی مونت‌کارلو را در ساخت مدل QSAR افزایش می‌دهند (۲۴).

هدف از این کار توسعهٔ مدل‌های قابل‌اعتماد QSAR برای پیش‌بینی فعالیت 255 (IC50) ترکیب شیمیایی، به‌عنوان مهارکنندهٔ آنزیم مبدل آنژیوتانسین (ACE I) I، با استفاده از نرم‌افزار CORAL مبتنی بر بهینه‌سازی مونت‌کارلو است. در این مطالعه، از روش تعادل همبستگی برای تولید مدل‌های QSAR استفاده می‌شود. در این روش، داده‌ها به چهار گروه آموزش فعال، آموزش غیرفعال، کالیبراسیون و اعتبارسنجی تقسیم می‌گردند، درحالی‌که در روش‌های دیگر، داده‌ها به سه گروه آموزش، کالیبراسیون و اعتبارسنجی تقسیم می‌شوند. قابلیت اطمینان و قدرت پیش‌بینی مدل‌های پیشنهادی با استفاده از چهار زیرمجموعه (آموزش فعال، آموزش غیرفعال، کالیبراسیون و اعتبارسنجی) از داده‌ها، به‌منظور سه تقسیم تصادفی برای هر تابع هدف (TF_0، TF_1، TF_2 و TF_3) به‌دقت ارزیابی شد.

مواد و روش‌ها

داده‌های استفاده‌شده در این تحقیق برای ایجاد مدل QSAR که شامل ۲۵۵ ترکیب شیمیایی متنوع از ترکیبات هتروسیکلیک با فعالیت مهارکنندگی (IC50/ میانهٔ غلظت بازدارندگی) در برابر آنزیم تبدیل‌کنندهٔ آنژیوتانسین I، به همراه نماد SMILES ترکیبات، از مقالهٔ شاه و همکاران (۲۵) جمع‌آوری گردید. دامنهٔ فعالیت بازدارنده‌ها در واحد منفی لگاریتم در محدودهٔ ۳/۵۵۳ تا ۹/۸۵۸ بود.

برای مدل‌سازی رابطهٔ کمی ساختار-فعالیت از تابع pIC_{50} (-logIC50) این ترکیبات، به‌عنوان متغیر وابسته استفاده شد. داده‌ها به‌صورت غیرانتخابی به چهار گروه آموزش فعال، آموزش غیرفعال، کالیبراسیون و اعتبارسنجی تقسیم گردید و به هرکدام از گروه‌های آموزش فعال و آموزش غیرفعال تقریباً ۲۷ درصد، کالیبراسیون ۲۴ درصد و آزمون ۲۲ درصد از داده‌ها اختصاص داده شد. جدول شمارهٔ ۱ درصد توزیع یکسانی ترکیبات دستهٔ آموزش فعال، آموزش غیرفعال، کالیبراسیون و اعتبارسنجی برای شکاف ۱ تا ۳ با سایر شکاف‌ها را نشان می‌دهد. این درصدها بر اساس رابطهٔ زیر محاسبه‌شده است:

$$Identity\ (\%) = \frac{\square_{i,}}{0.5\times(i+N_i)} \times 100$$

در اینجا $N_{i,j}$ تعداد ترکیباتی که برای هر دو شکاف i^{th} و j^{th}در یک دسته توزیع شده‌اند (یعنی دسته می‌تواند آموزش فعال، آموزش غیرفعال، کالیبراسیون یا اعتبارسنجی باشد)، N_i تعداد ترکیباتی است ک برای شکاف i^{th} در یک دسته توزیع می‌شوند و N_j تعداد ترکیباتی است که در دسته برای شکاف j^{th} توزیع می‌گردند.

هرکدام از دسته‌های آموزش فعال، آموزش غیرفعال، کالیبراسیون و اعتبارسنجی وظیفه‌ای دارند. دستهٔ آموزش مسئول ساخت مدل QSAR با استفاده از محاسبهٔ CWها

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

(وزن‌های همبستگی) است که بیشترین ضریب همبستگی را در میان توصیفگرها و pIC_{50} دارد؛ دستهٔ آموزش غیرفعال بررسی می‌کند که آیا مدل ایجادشده برای ترکیباتی که در مجموعهٔ آموزشی فعال نیستند، مناسب است یا خیر؛ مجموعهٔ کالیبراسیون برای شناسایی شروع پیش‌برازش مدل QSAR استفاده می‌شود و مجموعهٔ اعتبارسنجی برای ارزیابی نهایی پتانسیل پیش‌بینی مدل‌های توسعه‌یافته استفاده می‌گردد (۲۷، ۲۶، ۲۲). این کار سه بار به‌صورت تصادفی برای هر تابع هدف (TF_0، TF_1، TF_2 و TF_3) انجام‌شده است.

جدول شمارهٔ ۱. درصد یکسانی ترکیبات دستهٔ آموزش فعال، آموزش غیرفعال، کالیبراسیون و اعتبارسنجی هر شکافت با سایر شکافت‌ها

Split	Set	Split1	Split2	Split3
1	Active Training	100	22.0	18.6
	Passive training	100	26.8	25.2
	Calibration	100	28.3	19.0
	Validation	100	27.9	18.8
2	Active Training		100	15.6
	Passive training		100	25.0
	Calibration		100	20.5
	Validation		100	18.9
3	Active Training			100
	Passive training			100
	Calibration			100
	Validation			100

انتخاب توصیفگرهای مولکولی مناسب

برای تحویل دوز تجویزشدهٔ پرتودرمانی 300 (cGy) به بیمار در حالت استفاده از MLC، برای میدان AP یک شیلد جزیره‌ای در چند مرحله با استفاده از زاویهٔ پرتو و زاویهٔ کولیماتور مشابه روش بلوک سروبندی طراحی شد. هر یک از میدان‌های قدامی و خلفی به سه مرحله تقسیم گردید.

$$Y = C_0 + C_1 \times {}^{\text{Hybrid}}\text{DCW}\left(T^*, N^*\right) \qquad (۲)$$

DCW تابعی از ویژگی‌های مولکولی (به‌عنوان‌مثال، ویژگی‌های ساختاری) است که از نماد SMILES استخراج می‌شود. C0 و C1 ضرایب رگرسیونی هستند که با روش حداقل مربعات محاسبه می‌گردد. T* و N* شاخص‌های بهینه‌سازی مونت‌کارلو هستند که مقدار آستانه و تعداد دوره‌های بهینه‌سازی وزن‌های همبستگی را نشان می‌دهند.

در این مطالعه، مدل QSAR برای پیش‌بینی نقطهٔ پایانی با استفاده از توصیفگر بهینهٔ ترکیبی بر اساس وزن‌های همبستگی ویژگی‌های SMILES و گراف مولکولی بدون هیدروژن (HSG) طراحی‌شده است و از طریق معادله زیر محاسبه می‌شود:

$${}^{\text{Hybrid}}\text{DCW}(T^*, N^*) = {}^{\text{SMILES}}\text{DCW}(T^*, N^*) + {}^{\text{HSG}}\text{DCW}(T^*, N^*) \qquad (۳)$$

رابطهٔ ریاضی محاسبهٔ وزن‌های همبستگی هرکدام از توصیفگرهای SMILES و HSG در اطلاعات تکمیلی مقاله به‌ترتیب در معادله‌های شمارهٔ ۱ و ۲ آورده شده است.

بهینه‌سازی وزن‌ها به روش مونت‌کارلو

به‌منظور ایجاد مدل‌های QSAR، چهار تابع هدف TF_0، TF_1، TF_2 و TF_3 به‌صورت زیر، برای بهینه‌سازی مونت‌کارلو بررسی گردید و نتایج آماری آن‌ها مقایسه شد:

$$TF_0 = R_{ATRN} + R_{PTRN} - |R_{ATRN} - R_{PTRN}| \times dr_{weight} \qquad (۴)$$

$$TF_1 = TF_0 + IIC_{CAL} \times \text{Weight For IIC } (IIC_{weight}) \qquad (۵)$$

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

$$TF_2 = TF_0 + CII_{CAL} \times \text{Weight For CII } (CII_{weight}) \quad (۶)$$

$$TF_3 = TF_0 + IIC_{CAL} \times IIC_{weight} + CII_{CAL} \times CII_{weight} \quad (۷)$$

RATRN و RPTRN ضرایب همبستگی میان نقاط پایانی pIC_{50} مشاهده‌شده و پیش‌بینی‌شده برای مجموعه‌های آموزش فعال و آموزش غیرفعال هستند. مقدار عددی وزن‌های IIC و CII ثابت هستند. در این تحقیق، به‌ترتیب IICweight برابر ۰/۳ و CIIweight نیز ۰/۳ در نظر گرفته‌شده است. IICCAL(شاخص ایدئال بودن همبستگی) و CIICAL (شاخص شدت همبستگی) با استفاده از داده‌های مجموعه کالیبراسیون محاسبه می‌شوند. این شاخص‌ها برای بهبود کیفیت بهینه‌سازی مونت‌کارلو برای ایجاد مدل QSAR به‌کار می‌روند. روابط ریاضی محاسبهٔ IIC و CII در اطلاعات تکمیلی مقاله در معادله‌های ۳ تا ۷ آورده شده است.

دامنهٔ کاربرد مدل

بر اساس اصل سوم سازمان همکاری اقتصادی و توسعه (OECD)، دامنهٔ کاربرد (AD) معیار مهمی است که باید در هر مدل QSAR ایجادشده گنجانده گردد. AD برای تعیین اینکه آیا مدل‌های طراحی‌شدهٔ QSAR پیش‌بینی‌های درست یا اشتباه دارند، استفاده می‌شود؛ بنابراین، اگر مولکول در AD مدل ایجادشده قرار گیرد، پیش‌بینی نقطهٔ پایانی با استفاده از مدل قابل‌تصور است و اگر مولکول خارج از AD باشد، پیش‌بینی صحیح نقطهٔ پایانی غیرممکن می‌گردد. در نرم‌افزار CORAL، AD بر اساس نقص SMILES است که کیفیت ویژگی‌های مولکولی استخراج‌شده از ماده را تعیین می‌کند. در مدل CORAL، AD با توجه به توزیع ویژگی‌های ساختاری (Ak) استخراج‌شده از SMILES یا HSG در مجموعه‌های آموزش فعال و کالیبراسیون تعریف می‌شود. رابطهٔ ریاضی محاسبهٔ نقص آماری مرکب (Defect compound) برای یک توصیفگر (Ak) در اطلاعات تکمیلی مقاله در معادلهٔ ۸ آورده شده است.

نقص آماری مرکب SMILES را می‌توان به‌عنوان مجموع عیوب آماری همهٔ ویژگی‌ها محاسبه کرد.

$$\text{Defect}_{\text{compound}} = \sum_{k=1}^{NA} \text{Defect}_{A_K} \quad (۸)$$

NA تعداد توصیفگرهای SMILES فعال برای یک ترکیب است.

در نرم‌افزار CORAL، اگر نامساوی زیر برقرار باشد، یک ترکیب در محدودهٔ دامنهٔ کاربرد مدل قرار می‌گیرد، در غیر این صورت، به‌عنوان یک دادهٔ پرت شناخته می‌شود:

$$\text{Defect}_{\text{compound}} < 2 \times \overline{\text{Defect}}_T \quad (۹)$$

که $\overline{\text{Defect}}_T$ میانگین نقص آماری برای دستهٔ آموزش فعال است.

یافته‌های پژوهش

در این مطالعه، برای ارائهٔ نتایج آماری قابل‌اعتماد، سه مدل CORAL برای هر نوع تابع هدف (TF_0، TF_1، TF_2 و TF_3) و درمجموع، ۱۲ مدل QSAR با استفاده از توصیفگرهای بهینهٔ ترکیبی استخراج‌شده از ویژگی‌های SMILES و HSG ساخته شد. توصیفگرهای بهینه با استفاده از وزن‌های همبستگی (DCW) از طریق بهینه‌سازی مونت‌کارلو برای هر تابع هدف محاسبه گردیدند. ویژگی‌های آماری مدل‌های CORAL محاسبه‌شده با TF_0، TF_1، TF_2 و TF_3 در اطلاعات تکمیلی مقاله در جدول شمارهٔ ۱ نشان داده‌شده است. از روش توازن همبستگی برای ایجاد مدل‌های QSAR با تابع هدف (TF_0 (WIIC=WCII=0 استفاده شد. برای ایجاد مدل‌های QSAR با استفاده از تابع هدف TF_1، تنها IIC (WIIC=0.3 و WCII=0) به‌کار رفت و نتیجه نشان داد که افزودن IIC به TF_0، کیفیت آماری R2 و RMSE را برای

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

مجموعهٔ کالیبراسیون و اعتبارسنجی بهبود می‌بخشد.

برای ساخت مدل‌های مبتنی بر بهینه‌سازی TF_2، تنها از CII (WIIC=0 & WCII=0.3) استفاده گردید و نتایج تأیید می‌کند که افزودن CII به TF_0 باعث افزایش مقدار عددی R2 برای همهٔ زیرمجموعه‌ها (آموزش فعال، آموزش غیرفعال، مجموعه‌های کالیبراسیون و اعتبارسنجی) شد. درنهایت، در مدل ساخته‌شده با تابع هدف TF_3، IIC و CII (WIIC=0.3 و WCII=0.3) به‌طور همزمان استفاده گردیدند. در گزارش‌های پیشین آمده است که تولید تابع هدف با گنجاندن IIC و CII به‌طور همزمان می‌تواند کارآمدتر از استفاده از IIC یا CII به‌تنهایی باشد (۲۴). درنهایت، مدل‌های بهبودیافته با تابع هدف TF_3 از نظر آماری بر همهٔ مدل‌های دیگر برتری داشتند؛ بنابراین، آن‌ها را به مدل‌های قابل‌اعتماد برای پیش‌بینی احتمالی آینده تبدیل کرد. ویژگی‌های آماری مدل‌های محاسبه‌شده با TF_3 در جدول شمارهٔ ۲ آورده شده است.

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

جدول شمارهٔ ۲. مشخصات آماری برای مدل‌های QSAR مربوط به اثر مهارکنندگی آنزیم مبدل آنژیوتانسین I ترکیبات مطالعه‌شده بر اساس تابع هدف TF_3

Split	Set	n	R^2	CCC	IIC	CII	Q^2	Q^2F_1	Q^2F_2	Q^2F_3	RMSE	MAE	F	$\overline{r_m^2}$	Δr_m^2	Y-Test	Rm^{*2}(test)	5-fold
۱	ActivTRN	68	0.8494	0.9159	0.9121	0.9121	0.8354				0.686	0.540	360					0.8347
	Pass TRN	69	0.8451	0.9175	0.8575	0.9012	0.8363				0.650	0.515	363					0.8376
	Calib	63	0.7113	0.8399	0.8434	0.8487	0.6942	0.6840	0.6639	0.7695	0.819	0.658	150	0.6009	0.0837		0.5590	0.6975
	Valid	55	0.7783	0.8755	0.8203	0.8721	0.7593				0.7528	0.6007	186	0.6885	0.0295	0.0183	0.6737	0.7637
۲	ActivTRN	68	0.7792	0.8759	0.8323	0.8716	0.7658				0.731	0.591	233					0.7595
	Pass TRN	69	0.8298	0.9003	0.5151	0.9048	0.8200				0.712	0.573	327					0.8177
	Calib	64	0.7581	0.8626	0.8707	0.8691	0.7398	0.7350	0.7292	0.7765	0.759	0.618	194	0.6621	0.0050		0.6595	0.7398
	Valid	54	0.8553	0.9۲29	0.7990	0.8983	0.8۵22				0.6006	0.4528	230	0.7386	0.0182	0.0294	0.7295	0.8024
۳	ActivTRN	70	0.8371	0.9113	0.8641	0.9066	0.8275				0.643	0.510	394					0.8277
	Pass TRN	71	0.8523	0.9095	0.8624	0.9125	0.8443				0.717	0.551	398					0.8461
	Calib	63	0.7154	0.8361	0.8458	0.8350	0.6931	0.7006	0.6941	0.6058	0.754	0.624	153	0.6058	0.0798		0.6458	0.6936
	Valid	51	0.7647	0.8661	0.8354	0.8481	0.7469				0.7577	0.6029	159	0.6677	0.1408	0.0231	0.7381	0.7433

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

مدل‌های ایجادشده با استفاده از TF_0، TF_1 و TF_2 در اطلاعات تکمیلی مقاله آمده است و مدل‌های ساخته‌شده برای pIC_{50} به‌دست‌آمده با تابع هدف TF_3 توسط معادلات زیر ارائه می‌شوند:

شکاف ۱ pIC_{50}=-0.18839 (± 0.04813) +0.08860 (± 0.00058) * DCW(۷ ،۳۵)

شکاف ۲ pIC_{50}=0.47962 (± 0.05711) +0.07845 (± 0.00063) * DCW(۷ ،۳۵)

شکاف ۳ pIC_{50}=-0.34947 (± 0.05090) +0.07023 (± 0.00046) * DCW(۷ ،۳۵)

نتایج آماری مدل‌ها (جدول شمارهٔ ۲) نشان داد که همهٔ مدل‌های به‌دست‌آمده با استفاده از TF_3 به‌اندازه کافی برای پیش‌بینی pIC_{50} خوب بودند و با الزامات مؤلفه‌های اعتبارسنجی مختلف موافقت دارند (۲۸). مقادیر pIC_{50} مشاهده‌شده و محاسبه‌شده و همچنین دامنهٔ کاربرد ترکیبات استفاده‌شده در این کار در جداول S1 در فایل Excel آمده است.

اعتبارسنجی مدل: اعتبارسنجی مدل‌های توسعه‌یافته در ارزیابی قابلیت اطمینان و استحکام مدل‌های QSAR مهم است. اعتبارسنجی مدل را می‌توان با استفاده از: الف. اعتبارسنجی متقابل (Q2) یا اعتبارسنجی داخلی (R2)، ب. اعتبارسنجی خارجی و ج. تصادفی‌سازی داده‌ها یا درهم‌آمیختگی Y بررسی کرد. معادلهٔ ریاضی این معیارهای آماری در مقالات پیشین ارائه گردیده (۳۱-۲۹) و همچنین در جدول شمارهٔ ۳ آورده شده است.

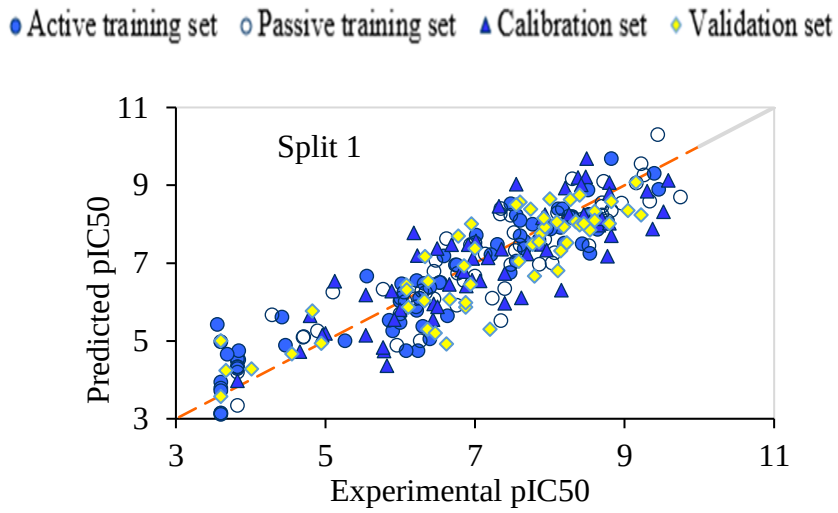
جدول شماره ۳. معادلهٔ ریاضی شاخص‌های اعتبارسنجی مدل‌های QSAR

نوع اعتبارسنجی	معیار پتانسیل پیشگویی
داخلی	$R^2 = 1 - \frac{\sum(Y_{obs} - Y_{prd})^2}{\sum(Y_{obs} - \bar{Y})^2}$
	$Q^2 = 1 - \frac{\sum(Y_{prd} - Y_{obs})^2}{\sum(Y_{obs} - \bar{Y}_{train})^2}$
خارجی	$Q^2_{F1} = 1 - \frac{\sum(Y_{per(test)} - Y_{obs(test)})^2}{\sum(Y_{obs(test)} - \bar{Y}_{train})^2}$
	$Q^2_{F2} = 1 - \frac{\sum(Y_{prd(test)} - Y_{obs(test)})^2}{\sum(Y_{obs(test)} - \bar{Y}_{ext})^2}$
	$Q^2_{F3} = 1 - \frac{\sum(Y_{prd(test)} - Y_{obs(test)})^2/n_{ext}}{\sum(Y_{obs(test)} - \bar{Y}_{train})^2/n_{train}}$
	$R^2_m = R^2 \times \left(1 - \sqrt{R^2 - R^2_0}\right)$
	$CCC = \frac{2\sum(X - \bar{X})(Y - \bar{Y})}{\sum(X - \bar{X})^2 + \sum(Y - \bar{Y})^2 + n((\bar{X} - \bar{Y})^2}$
	$MAE = \frac{1}{n} \times \sum\lvert Y_{obs} - Y_{prd}\rvert$
آزمون تصادفی‌سازی	$C_{R^2_p} = R\sqrt{(R^2 - R^2_r)}$

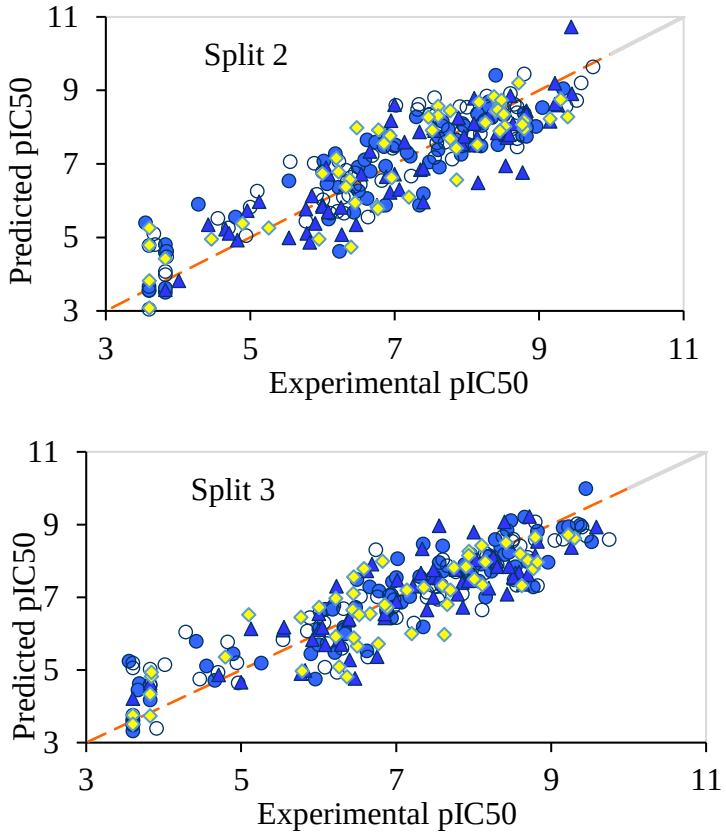
[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

توانایی پیش‌بینی مدل QSAR در صورتی قابل‌قبول است که مقدار عددی Q2 و R2 بیشتر از ۰/۷ باشد (۳۲). $Q2$ به‌عنوان یک معیار آماری، برای ارزیابی دقت پیش‌بینی عمل می‌کند. مقدار Q^2 بالا نشان می‌دهد که مدل می‌تواند داده‌های جدید را به‌دقت پیش‌بینی کند، درحالی‌که مقدار $Q2$ پایین نشان می‌دهد که مدل بسیار خاص برای داده‌های آموزشی است و نمی‌تواند به‌طور مؤثر به داده‌های جدید تعمیم داده شود (۳۳). دربارهٔ اعتبارسنجی خارجی، اگر CCC>۰.۸۵، r2m>0.5 و Δr2m<0.2 باشد، مدل را می‌توان به‌عنوان مدلی با قابلیت پیش‌بینی خوب بیان کرد (۳۴). Q_F1^2 توانایی مدل برای پیش‌بینی فعالیت ترکیبات موجود در مجموعهٔ اعتبارسنجی را اندازه‌گیری می‌کند. Q_F2^2 همبستگی میان مقادیر تجربی و مقادیر پیش‌بینی‌شده را در مجموعهٔ اعتبارسنجی اندازه‌گیری می‌نماید. Q_F3^2 توانایی مدل را برای پیش‌بینی فعالیت ترکیباتی اندازه‌گیری می‌کند که در مجموعه دادهٔ استفاده‌شده برای ساخت مدل گنجانده نشده‌اند؛ همچنین دربارهٔ آزمون تصادفی‌سازی Y، پس از چندین بار تکرار، مدل‌های تصادفی جدید ایجاد گردید و مقادیر R2 زیر ۰/۱ یافت شد که نشان‌دهندهٔ این است که همبستگی میان pIC_{50} و توصیفگرهای مولکولی بر اساس همبستگی تصادفی نیست. علاوه بر این، برای سه مدل، cR_p^2 بیشتر از ۰/۷ به‌دست آمد که همبستگی غیرشانسی مدل‌های توسعه‌یافته را تأیید کرد (۳۵). همان‌طور که جدول شمارهٔ ۲ نشان می‌دهد، مقادیر عددی ویژگی‌های آماری محاسبه‌شده برای همهٔ مدل‌های طراحی‌شده در محدوده‌های قابل قبولی قرار داشتند که در منابع گزارش‌شده است (۳۷، ۳۶)؛ بنابراین، مدل‌های QSAR ایجادشده از نظر آماری قابل‌اعتماد هستند و پتانسیل پیش‌بینی خوبی دارند.

بر اساس نتایج آماری مدل‌های ایجادشده، مدل مربوط به شکافت ۲ که بالاترین شاخص‌های آماری برای دستهٔ اعتبارسنجی را داشت، به‌عنوان بهترین مدل انتخاب گردید و می‌توان pIC_{50} مربوط به ترکیباتی که pIC_{50} آن‌ها گزارش نشده است، با استفاده از این مدل پیشگویی و گزارش نمود. شکل شمارهٔ ۱ نمایش تصویری داده‌های مشاهده‌شدهٔ pIC_{50} در مقابل pIC_{50} محاسبه‌شده برای سه مدل به‌دست‌آمده با استفاده از تابع هدف TF_3 را نشان می‌دهد.

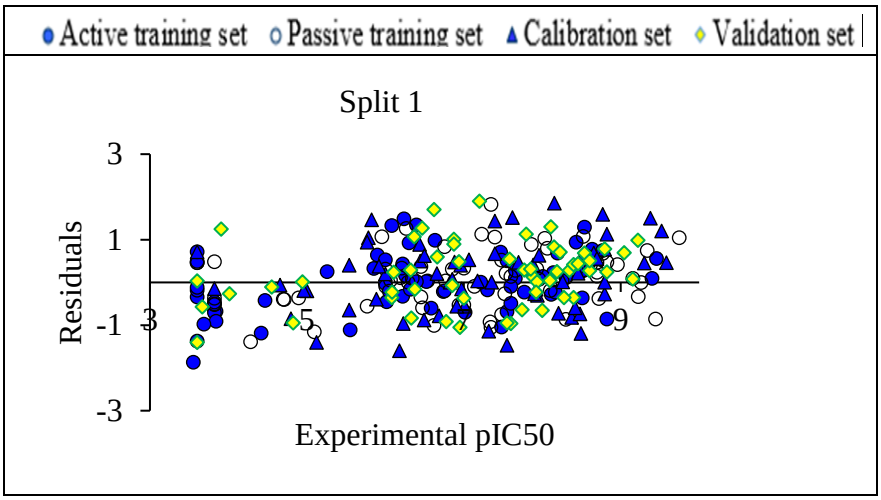


[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

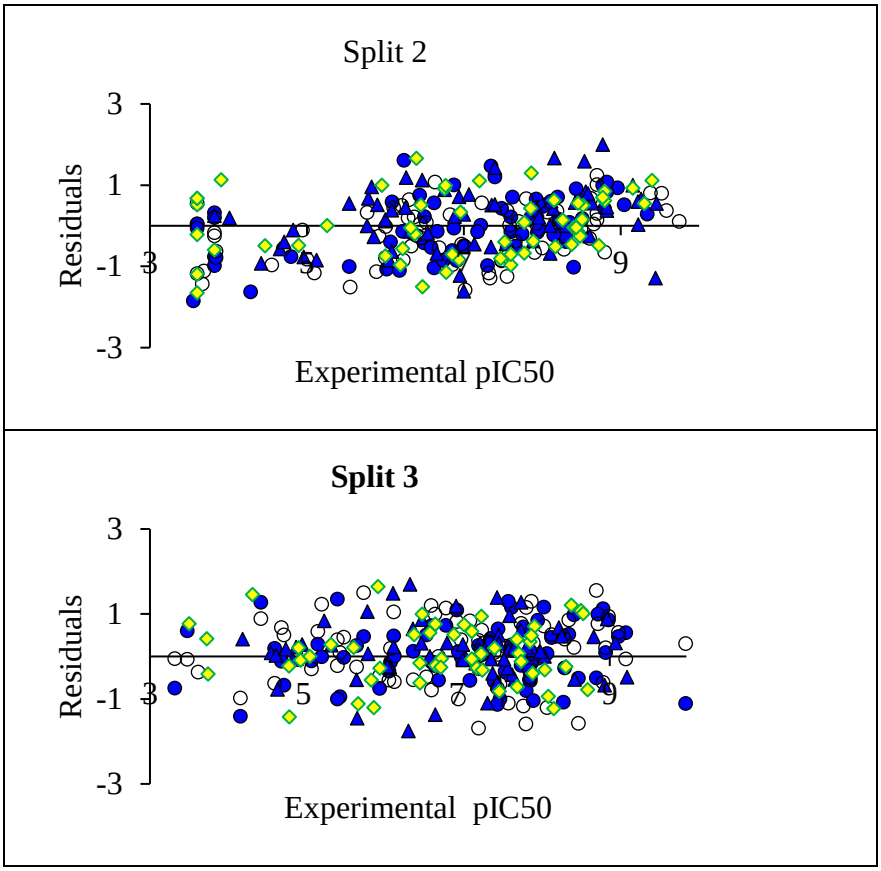


شکل شمارهٔ ۱. نمایش همبستگی میان pIC_{50} مشاهده‌شده و محاسبه‌شده برای مدل‌های ۱ تا ۳ به‌دست‌آمده با استفاده از تابع هدف TF_3

همان‌طور که در شکل شمارهٔ ۱ دیده می‌شود، همبستگی خوبی میان داده‌های مشاهده‌شده و محاسبه‌شده در مدل‌های ایجاد گردیده وجود دارد. این نتایج تأیید می‌کنند که مدل‌های QSAR ایجادشده قابل‌اطمینان هستند و به‌خوبی برازش شده‌اند؛ همچنین نمودارهای باقیمانده‌های pIC_{50} در مقابل مقادیر pIC_{50} پیش‌بینی‌شده برای همهٔ مدل‌های QSAR در شکل شمارهٔ ۲ نشان داده‌شده است. پراکندگی باقیمانده‌های pIC_{50} در نزدیکی خط افقی در مرکز صفر دیده می‌شود که نشان می‌دهد، همهٔ مدل‌های QSAR ایجادشده به‌خوبی برازش شده‌اند.



[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]



شکل شمارهٔ ۲. نمایش تصویری از باقیمانده‌های pIC_{50} در مقابل pIC_{50} تجربی برای همهٔ مدل‌های QSAR ساخته‌شده با استفاده از TF_3

تفسیر مدل QSAR

تفسیر مکانیکی مدل‌ها به درک تأثیر توصیفگرها در نقطهٔ پایانی پیش‌بینی‌شده کمک می‌کند؛ یعنی ویژگی‌های مولکولی افزاینده و کاهندهٔ فعالیت را بیان می‌نماید. در مدل‌سازی QSAR توسط برنامهٔ CORAL، ویژگی‌های ساختاری (Ak) که از سه یا بیشتر از سه اجرای مستقل، از بهینه‌سازی مونت‌کارلو به‌دست می‌آید، برای دستیابی به تفسیر مکانیکی مدل‌ها به‌کار گرفته می‌شوند.

CW برای هر یک از ویژگی‌های SMILES در اجراهای مختلف یک مدل احتمالاً مثبت، منفی یا در یک اجرا مثبت و در اجرای دیگر منفی باشد. اگر ضریب یک توصیفگر در همهٔ اجراها مثبت یا منفی باشد، به‌ترتیب توصیفگر افزاینده و کاهندهٔ فعالیت برای نقطهٔ پایانی محسوب می‌شوند و اگر این ضریب در بعضی اجراها مثبت و در بعضی از اجراها منفی گردد، این توصیفگر تعریف‌نشده به‌شمار می‌رود. جدول شمارهٔ ۴ فهرستی از ویژگی‌های ساختاری را به‌عنوان توصیفگرهای افزاینده یا کاهندهٔ pIC_{50} نشان می‌دهد که از سه اجرای بهینه‌سازی مونت‌کارلو با T* و N* بهینه به همراه تفسیر توصیفگرها به‌دست‌آمده‌اند.

جدول شماره ٤. لیست توصیفگرهای افزاینده و کاهندهٔ pIC_{50}، وزن‌های همبستگی (CWs) در هر اجرا، تعداد ویژگی‌های ساختاری در هر زیرمجموعه به همراه تفسیر توصیفگرها

توصیفگر	وزن همبستگی اجرای ۱	وزن همبستگی اجرای ۲	وزن همبستگی اجرای ۳	N_{CAL}	N_{PT}	N_{AT}	نقص آماری	توضیحات
توصیفگرهای افزاینده								
(F..S)..Y.Y.	2.3862	2.2359	3.7417	68	69	63	0	
	4.7228	4.4229	6.1107	68	69	64	0	

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

حضور همزمان فلوئور و گوگرد	0	63	71	70	3.6381	2.3001	6.3006	
حضور همزمان فلوئور و برم	0	63	69	68	1.8739	5.0025	3.6902	(F.Br)..Y.Y.
	0	64	69	68	7.3434	5.4024	4.5281	
	0	63	71	70	2.0443	3.1968	6.6502	
حضور همزمان برم و گوگرد	0	63	69	68	2.0166	2.4536	6.4085	(Br.S)..Y.Y.
	0	64	69	68	2.6154	3.9911	2.8547	
	0	63	71	70	1.2587	4.1869	5.0175	
حضور همزمان کلر و گوگرد	0	63	69	68	6.6962	3.2312	3.4818	(Cl.S)..Y.Y.
	0	64	69	68	2.9465	4.9324	4.1836	
	0	63	71	70	2.41	1.9537	6.0964	
حضور همزمان کلر و برم	0	63	69	68	0.5254	1.1799	4.3565	(ClBr)..Y.Y.
	0	64	69	68	4.3095	4.3605	4.6398	
	0	63	71	70	6.0322	2.0605	5.5458	
حضور همزمان نیتروژن و گوگرد	0	63	69	68	7.399	2.2569	4.2778	(N..S)..Y.Y.
	0	64	69	68	5.392	6.6499	5.9261	
	0	63	71	70	1.0076	1.7201	5.3479	
حضور همزمان گوگرد و اکسیژن	0	63	69	68	6.5413	3.2167	4.6124	(O..S)..Y.Y.
	0	64	69	68	3.9696	3.9109	6.1881	
	0	63	71	70	4.0276	3.219	7.5101	
حضور نداشتن باند سه‌گانه	0	63	69	68	2.4352	6.7974	6.2725	<#>..0000...
	0	64	69	68	4.3706	3.9011	4.1614	
	0	63	71	70	2.0824	2.2465	6.2491	
حضور نداشتن اتم فاوئر	0.0002	58	65	63	2.833	3.1318	1.9364	<F>..0000...
	0.0004	62	65	63	3.4436	2.6853	3.1184	
	0.0002	60	68	66	6.4975	7.8153	9.9936	
توصیفگرهای کاهنده								
حضور کربن آلیفاتیک	0	63	69	68	-0.7383	-0.1922	-0.2693	C.....
	0	64	69	68	-0.4816	-0.0939	-0.1677	
	0	63	71	70	-0.3825	-0.546	-0.6653	
حضور باند دوگانه همراه با شاخه‌دار شدن	0.0001	63	69	67	-0.4122	-0.1498	-0.7411	=...(.....
	0.0003	64	67	68	-0.6682	-0.163	-0.3104	
	0.0002	62	71	69	-0.7469	-0.5216	-0.1256	
توالی کربن آلیفاتیک، شاخه‌دار شدن و باند دوگانه	0.0001	63	68	67	-0.315	-0.8821	-0.0008	C...(...=...
	0.0003	63	67	68	-0.135	-0.3218	-0.1937	
	0.0002	62	71	69	-0.3028	-0.4217	-0.4337	

بحث و نتیجه‌گیری

توسعه/پیش‌بینی فعالیت دارو فرایند پیچیده و پرهزینه‌ای است که هدف آن تولید داروهای ایمن و مؤثر است. این فرایند با چالش‌های فراوانی مواجه است و بسیاری از کاندیداهای دارویی به علت مشکلاتی مانند حلالیت و پایداری، به مرحلهٔ تجاری نمی‌رسند. بهینه‌سازی این فرایند برای افزایش شانس موفقیت اهمیت بسیاری دارد. طراحی/پیش‌بینی فعالیت دارو به کمک کامپیوتر به‌عنوان ابزاری قدرتمند و کارآمد، نقش مهمی در تسریع و بهبود فرایند طراحی و توسعهٔ داروها ایفا می‌کند و به کاهش زمان و هزینه‌های مربوطه کمک می‌نماید. در این پژوهش، مدل‌های QSAR با هدف پیش‌بینی فعالیت ۲۵۵ ترکیب شیمیایی، به‌عنوان مهارکنندهٔ آنزیم مبدل آنژیوتانسین I در درمان بیماری فشارخون، با استفاده از روش مونت‌کارلو و بر اساس معیارهای

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

شاخص ایدئال همبستگی (IIC) و شاخص شدت همبستگی (CII) ساخته‌شده و از توصیفگرهای هیبریدی بهینه که ترکیبی از SMILES و HSG است، در ساخت مدل‌های QSAR توسط نرم‌افزار CORAL استفاده گردیده است.

۱۲ مدل QSAR از ۳ تقسیم با استفاده از چهار تابع هدف (TF_0، TF_1، TF_2 و TF_3) ایجاد شد. نتایج آماری مدل‌های QSAR ایجادشده با هر تابع هدف با یکدیگر مقایسه گردید. مدل‌های ایجادشده توسط روش‌های الف. اعتبارسنجی متقابل (اعتبارسنجی داخلی/ CCC، IIC، R2، Q2 و F)، ب. اعتبارسنجی خارجی (Q_F1^2〗 ، Q_F2^2 〖Q، 〖Cr〗 _p^2 ، s، MAE، (r_m^2)¯ و 〖Δr 〖3^2_ m^2_ 〖() و ج. تصادفی‌سازی داده‌ها ارزیابی شدند که همهٔ مدل‌های ساخته‌شده قابل‌اطمینان، پیش‌بینی پذیر و قابل تکرار بودند. مدل‌های طراحی‌شده با تابع هدف TF_3پتانسیل پیش‌بینی بهتری نسبت به مدل‌های طراحی‌شده با سایر توابع هدف داشتند. مدل‌های QSAR پیشنهادی ضریب همبستگی ((R2 در محدودهٔ ۰/۸۵۳۵-۰/۷۶۴۷ برای دستهٔ آزمون قرار داشتند که رضایت‌بخش بود. همهٔ مدل‌ها از نظر آماری رضایت‌بخش و قوی بودند و با معیارهای آماری پیشنهادی موافقت داشتند.

کارایی مدل QSAR توسعه‌یافته در این مطالعه با مدل موجود پیشنهادشدهٔ شاه و همکاران (۳۸) مقایسه گردید. آنان از رویکرد الگوریتم ژنتیک-رگرسیون خطی چندگانه (GA-MLR) برای ساخت مدل استفاده کردند و معیارهای آماری R2=0.70-0.75، Q2LOO=0.67-0.73، Q2LMO=0.66-0.72 و CCC=0.70-0.78 را برای مدل خود گزارش نمودند و بیان کردند، این نتایج نشان‌دهندهٔ یک مدل QSAR با پیش‌بینی بسیار قوی و از نظر آماری معنی‌دار است. مقایسهٔ ویژگی‌های آماری میان مدل‌های QSAR ساخته‌شده و مدل QSAR پیشین نشان می‌دهد که مدل‌های ساخته‌شده توسط CORAL نتایج برتری را ارائه می‌دهد. مدل ایجادشده برای شکاف ۲ توسط تابع هدف TF_3 که معیارهای آماری دارد و کارایی بهتری نسبت به سایر شکاف‌ها داشت؛ بنابراین، به‌عنوان مدل برتر انتخاب شد.

مقالات پیشین ثابت کردند که معیارهای r_m^2 برای توصیف پتانسیل پیش‌بینی مدل‌ها، معتبرتر از معیارهای معمولی مانند R2 هستند (۳۹). اگر مقدار عددی r_(m 0.5<r_m^2 و Δ^2<0.2) برای هر دو مجموعه آموزشی و اعتبارسنجی مدل باشد قابل‌قبول است. مدل‌های QSAR توسعه‌یافته در این پژوهش، همه این معیارها را برآورده می‌کنند اما در کار قبلی معیارهای r_m^2 گزارش نشده است. مدل گزارش‌شده قبلی با استفاده از یک تقسیم توسعه‌یافته، اما در این مطالعه، از ۱۲ تقسیم برای توسعه مدل‌های QSAR با استفاده چهار مجموعه (آموزش فعال، آموزش غیرفعال، مجموعه کالیبراسیون و اعتبارسنجی) استفاده‌شده است. در مطالعه حاضر، ۱۲مدل با استفاده از چهار تابع هدف TF_0، TF_3، F1،F2 سه شده است. روش گزارش‌شده توسط شاه و همکاران مدل‌سازی با استفاده از بیش از یک توصیفگر ساخته‌شده است، اما در این مطالعه تنها یک توصیفگر، DCW، برای ساخت مدل QSAR استفاده می‌شود. ازاین‌رو، مدل‌های QSAR فعلی قوی‌تر هستند و قابلیت پیش‌بینی بهتری دارند.

در این مطالعه، ۱۲ مدل QSAR برای پیش‌بینی pIC_{50} تعداد ۲۵۵ ترکیب شیمیایی با اثر مهارکنندگی آنزیم مبدل آنژیوتانسین I با استفاده از روش مونت‌کارلو توسعه داده شد. برای طراحی مدل‌های QSAR از الگوریتم تعادل همبستگی استفاده شد. برای بررسی اهمیت روش آماری مدل‌های ساخته‌شده چهار تابع هدف، یعنی TF_0 (بدون IIC یا CII)، TF_1 (با IIC به‌تنهایی)، TF_2 (با CII به‌تنهایی) و TF_3 (با IIC و CII، به‌طور همزمان) بکار برده شد. استفاده همزمان از IIC و CII (TF_3) پتانسیل پیش‌بینی مدل QSAR را بهبود بخشید. همه مدل‌های پیشنهادی، پیش‌بینی رضایت‌بخشی را برای تخمین pIC_{50} ارائه کردند، اما بهترین پتانسیل پیش‌بینی‌شده با TF_3 برای مدل به‌دست‌آمده از شکاف ۲ محاسبه شد، بنابراین به‌عنوان بهترین مدل مشخص شد. برای ارزیابی پایایی و قابلیت پیش‌بینی تمامی مدل‌های ایجادشده، از شاخص‌های آماری مختلفی مانند R2، IIC، CCC، Q2، F، Q_F1^2، Q_F2^2، Q_F3^2، 〖Cr〗 _p^2 ، s، MAE، (r_m^2)¯ و 〖Δr〗

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

_m^2 و تصادفی‌سازی داده‌ها یا درهم‌آمیختگی Y استفاده شد. دامنه کاربرد مدل برای شناسایی نقاط پرت در مدل‌های QSAR تولیدشده استفاده شد. ویژگی‌های ساختاری به‌عنوان محرک‌های افزایش/کاهش pIC_{50} نیز شناسای شدند.

سپاس گزاری

از همه‌ی اشخاصی که در این پژوهش همکاری داشتند، تشکر و قدردانی می‌شود.

تعارض منافع

نویسندگان اعلام می‌نمایند که هیچگونه تعارض منافعی وجود ندارد.

کد اخلاق

با توجه به استفاده نکردن از نمونه انسانی –حیوانی، نیاز به تهیه کد اخلاق نبود.

حمایت مالی

این تحقیق هیچ کمک مالی خاصی از سازمان ها ی تأمین مالی در بخش های عمومی، تجاری یا غیر انتفاعی در یافت نکرده است.

مشارکت نویسندگان

شهرام لطفی: مدیریت پروژه، نگارش پیش‌نویس اولیه، اعتبارسنجی

شهین احمدی: نگارش بررسی و ویرا یش، نرم‌افزار، تجز یه و تحلیل داده ها

علی عظیمی: نرم افزار، مدل سازی، گردآوری داده‌ها.

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

References

1. Jao CL, Huang SL, Hsu KU. Angiotensin I-converting enzyme inhibitory peptides: Inhibition mode, bioavailability, and antihypertensive effects. BioMedicine. 2012; 2: 130-6. doi: 10.1016/j.biomed.2012.06.005.
2. Amaya, J.A.G., et al., In-silico design of new enalapril analogs (ACE inhibitors) using QSAR and molecular docking models. Inform Med Unlocked. 2020; 19: 100336. doi:10.1016/j.imu.2020.100336.
3. Cushman, D.W., et al., Design of potent competitive inhibitors of angiotensin-converting enzyme. Carboxyalkanoyl and mercaptoalkanoyl amino acids. Biochemistry. 1977; 16: 5484-91. doi:10.1021/bi00644a014.
4. Lin K, Zhang L, Han X, Meng Z, Zhang J, Wu Y, Cheng D. Novel angiotensin I-converting enzyme inhibitory peptides from protease hydrolysates of Qula casein: Quantitative structure-activity relationship modeling and molecular docking study. J Agric Food Chem. 2017; 32: 266-77. doi:10.1016/j.jff.2017.03.008.
5. Yılmaz, İ., Angiotensin-converting enzyme inhibitors induce cough. Turk Thorac J. 2019; 20: 36-42. doi: 10.5152/TurkThoracJ.2018.18014.
6. Vena GA, Cassano N, Coco V, De Simone C. Eczematous reactions due to angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers. Immunopharmacol Immunotoxicol. 2013; 35: 447-50. doi: 10.3109/08923973.2013.797992.
7. Murray BA, FitzGerald RJ. Angiotensin converting enzyme inhibitory peptides derived from food proteins: biochemistry, bioactivity and production. Curr Pharm Des. 2007; 13: 773-91. doi: 10.2174/138161207780363068.
8. Wu Q, Cai QF, Tao ZP, et al. Purification and characterization of a novel angiotensin I-converting enzyme inhibitory peptide derived from abalone (Haliotis discus hannai Ino) gonads. Eur Food Res Technol. 2015; 240: 137-45. doi: 10.1016/j.peptides.2011.11.006.
9. Lorenz S, et al. Toward application and implementation of in silico tools and workflows within benign by design approaches. ACS Sustain Chem Eng. 2021; 9: 12461-75. doi:10.1021/acssuschemeng.1c03070.
10. Zhao P, Peng Y, Xu X, Wang Z, Wu Z, Li W, Tang Y, Liu G. In silico prediction of mitochondrial toxicity of chemicals using machine learning methods. J Appl Toxicol. 2021; 41: 1518-26. doi: 10.1002/jat.4141.
11. Azimi A, Ahmadi Sh, Kumar A, Qomi M, Almasirad A, et al. SMILES-Based QSAR and Molecular Docking Study of Oseltamivir Derivatives as Influenza Inhibitors. PACs. 2022;43: 1-21. doi:10.1080/10406638.2022.2067194.
12. Ahmadi S, Ghanbari H, Lotfi S, Azimi N. Predictive QSAR modeling for the antioxidant activity of natural compounds derivatives based on Monte Carlo method. Mol Divers. 2021; 25: 87-97. doi:10.1007/s11030-019-10026-9.
13. Ahmadi S, Khani R, Moghaddas M. Prediction of anti-cancer activity of 1, 8-naphthyridin derivatives by using of genetic algorithm-stepwise multiple linear regression.Med Sci. 2018;28:181-94. doi:10.29252/iau.28.3.181.
14. Lotfi S, Ahmadi S, Zohrabi P. QSAR modeling of toxicities of ionic liquids toward Staphylococcus aureus using SMILES and graph invariants. Struct Chem. 2020; 31: 2257-70. doi: 10.1007/s11224-020-01568-y
15. Ahmadi S, Moradi Z, Kumar A, Almasirad A. SMILES-based QSAR and molecular docking study of xanthone derivatives as α-glucosidase inhibitors. J Recept Signal Transduct Res. 2021; 42:361-72. 1-12. doi:10.1080/10799893.2021.1957932.
16. rabhakar YS, Gupta SP. Structure-Activity relationship study on angiotensin-converting enzyme inhibitors--investigation of hydrophobic interaction in inhibition mechanism. Indian J Biochem Biophys. 1985, 22: 318-20.
17. Javidfar M, Ahmadi S. QSAR modelling of larvicidal phytocompounds against Aedes aegypti using index of ideality of correlation. SAR QSAR Environ Res. 2020; 31: 717-39. doi:10.1080/1062936X.2020.1806922.
18. Deokar H, Deokar M, Wang W, Zhang R, Buolamwini JK. QSAR studies of new pyrido [3, 4-b] indole derivatives as inhibitors of colon and pancreatic cancer cell proliferation. Med Chem Res. 2018; 27: 2466-81. doi: 10.1007/s00044-018-2250-5.
19. Lotfi S, Ahmadi S, Kumar P. Correction: Ecotoxicological prediction of organic chemicals toward Pseudokirchneriella subcapitata by Monte Carlo approach. RSC Adv. 2022; 12: 34567. doi:10.1039/D2RA03936B.
20. Toropov AA, Toropova AP, Roncaglioni A, Benfenati E. The system of self-consistent models for pesticide toxicity to Daphnia magna. Toxicol Mech Methods. 2023;33: 578-83. doi:10.1080/15376516.2023.2197487.
21. Toropova AP, Toropov AA, Roncaglioni A, Benfenati E. The index of ideality of correlation improves the predictive potential of models of the antioxidant activity of tripeptides from frog skin (Litoria rubella).

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]

Comput Biol Med. 2021; 133: 104370. doi:10.1016/j.compbiomed.2021.104370.
22. Lotfi S, Ahmadi Sh, Kumar P. A hybrid descriptor based QSPR model to predict the thermal decomposition temperature of imidazolium ionic liquids using Monte Carlo approach. J Mol Liq. 2021; 338: 116465. doi:10.1016/j.molliq.2021.116465.
23. Kumar P, Kumar A. In silico enhancement of azo dye adsorption affinity for cellulose fibre through mechanistic interpretation under guidance of QSPR models using Monte Carlo method with index of ideality correlation. SAR QSAR Environ Res. 2020; 31: 697-715. doi:10.1080/1062936X.2020.1806105.
24. Toropov AA, Toropova AP, Achary PGR, Raškova M, Raška I. The searching for agents for Alzheimer's disease treatment via the system of self-consistent models. Toxicology Toxicol Mech Methods. 2022; 32: 549-557. doi: 10.1080/15376516.2022.2053918.
25. Shah S, Chaple D, Masand VH, Zaki MEA, Al-Hussain SA, Shah A,et al. In silico study to recognize novel angiotensin-converting-enzyme-I inhibitors by 2D-QSAR and constraint-based molecular simulations. J Biomol Struct Dyn. 2023; 43:2211-30. doi: 10.1080/07391102.2023.220326.
26. Duhan M, Sindhu J, Kumar P, Devi M, Singh R, Kumar R, et al. Quantitative structure activity relationship studies of novel hydrazone derivatives as α-amylase inhibitors with index of ideality of correlation. J Biomol Struct Dyn. 2020; 40: 4933-53. doi: 10.1080/07391102.2020.1863861.
27. Achary PGR, Toropova AP, Toropov AA. Combinations of graph invariants and attributes of simplified molecular input-line entry system (SMILES) to build up models for sweetness. Food Res Int. 2019; 122: 40-46. doi:10.1016/j.foodres.2019.03.067.
28. Roy PP, Roy K. QSAR studies of CYP2D6 inhibitor aryloxypropanolamines using 2D and 3D descriptors. Chem Biol Drug Des. 2009; 73: 442-55. doi: 10.1111/j.1747-0285.2009.00791.x.
29. Lotfi S, Ahmadi S, Kumar P. The Monte Carlo approach to model and predict the melting point of imidazolium ionic liquids using hybrid optimal descriptors. RSC Adv. 2021; 11: 33849-57. doi:10.1039/D1RA06861J.
30. Ahmadi S, Ketabi S, Qomi M. CO 2 uptake prediction of metal–organic frameworks using quasi-SMILES and Monte Carlo optimization. New J Chem. 2022; 46: 8827-37. doi:10.1039/D2NJ00596D.
31. Kumar P, Kumar A. CORAL: QSAR models of CB1 cannabinoid receptor inhibitors based on local and global SMILES attributes with the index of ideality of correlation and the correlation contradiction index. Chemometr Intell Lab Sys. 2020; 200: 103982. doi:10.1016/j.chemolab.2020.103982.
32. Toropova AP, Toropov AA, Veselinović AM, Veselinović JB, Benfenati E, Leszczynska D, et al. Nano-QSAR: Model of mutagenicity of fullerene as a mathematical function of different conditions. Ecotoxicol Environ Saf. 2016; 124: 32-36. doi:10.1016/j.ecoenv.2015.09.038.
33. Tabti K, et al. HQSAR, CoMFA, CoMSIA docking studies and simulation MD on quinazolines/quinolines derivatives for DENV virus inhibitory activity. Chem Afr. 2022; 5: 1937-58. doi:10.1007/s42250-022-00484-4.
34. Chirico N, Gramatica P. Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection. J Chem Inf Model. 2012; 52: 2044-58. doi:10.1021/ci300084j.
35. Rücker C, Rücker G, Meringer M. Y-randomization–a useful tool in QSAR validation, or folklore. J Chem Inf Model. 2007; 47: 2345-57. doi:10.1021/ci700157b.
36. Yordanova D, W. Schultz T, Kuseva C, Tankova K, Ivanova H. Automated and standardized workflows in the OECD QSAR Toolbox. ComputToxicol. 2019; 10: 89-104. doi:10.1016/j.comtox.2019.01.006.
37. Gatidou G, Vazaiou N, Thomaidis NS, Stasinakis AS. Biodegradability assessment of food additives using OECD 301F respirometric test. Chemosphere. 2020; 241: 125071. doi:10.1016/j.chemosphere.2019.125071.
38. Shah S, Chaple D, Masand VH, Zaki MEA, Al-Hussain SA, Shah A. In silico study to recognize novel angiotensin-converting-enzyme-I inhibitors by 2D-QSAR and constraint-based molecular simulations. J Biomol Struct Dyn. 2024; 42: 2211-30. doi: 10.1080/07391102.2023.220326.
39. Roy K, Chakraborty P, Mitra I, Ojha PK, Kar S, Das RN. Some case studies on application of "rm2" metrics for judging quality of quantitative structure–activity relationship predictions: emphasis on scaling of response data. J Comput Chem. 2013; 34: 1071-82. doi.org/10.1002/jcc.23231.

[Downloaded from sjimu.medilam.ac.ir on 2026-11-09]