In silico Modeling and Toxicity Profiling of a Set of Quinoline Derivatives as c-MET Inhibitors in the treatment of Human Tumors

İnsan Tümörlerinin Tedavisinde c-MET İnhibitörü Olarak Kullanılan Bir Dizi Kinolin Türevinin In silico Modellemesi ve Toksisite Profili

ABSTRACT

Objectives: 4-(2-fluorophenoxy) quinoline derivatives constitute one of the chemical classes of hepatocyte growth factor receptor (c-MET) inhibitors, a promising treatment against various human tumors. There are three aims of the present study: (1) To develop a robust and validated quantitative structure-activity relationship model to predict the c-Met kinase inhibition; (2) to examine the toxicity profiles of these compounds; (3) to design new quinoline derivatives and apply the developed model on these compounds to observe its pertinence.

Materials and Methods: A multiple linear regression method was used to develop the model with calculated descriptors. State-of-the-art internal and external validation parameters were calculated. The in silico toxicity profiles including structural alerts and the lowest observed adverse effect level (LOAEL) values were evaluated using online tools. New derivatives were designed and tested on the developed model.

Results: A series of 4-(2-fluorophenoxy) quinoline derivatives was linearly modeled and vigorously validated to predict the molecules’ c-MET kinase inhibition potential. Statistical metrics of the developed model showed that it was robust and able to make successful predictions for this chemical class. The mass, electronegativity, partial charges, and the structure of the molecules had an effect on the activity. Moreover, the toxicity profiles of the studied compounds were found to be adequate.

Conclusion: Five of the synthesized compounds were observed to be auspicious for the toxicity/activity ratio. The developed model is useful in the virtual screening and in the design of new anti-tumor compounds.

Key words: Toxicity, LOAEL, anti-tumor, c-MET, QSAR

ÖZ

Amaç: 4-(2-florofenoksi) kinolin türevleri, insan tümörlerinin tedavisinde kullanılan hepatosit büyüme faktörü reseptörü (c-MET) inhibitörleri kimyasal sınıfından biridir. Bu çalısmının üç amacı vardır: (1) c-MET kinaz inhibisyonunu tahmin etmek için sağlıklı ve doğrulanmış bir kantitatif yapı-aktivite iliskisi modeli geliştirmek; (2) üzerinde çalışılan bileşiklerin toksisite profililerini incelemek; (3) yeni kinolin türevleri tasarlamak ve geliştirilen modeli bu bileşikler üzerinde uygulayarak uygunlukunu incelemektir.

Gereç ve Yöntemler: Bu çalışmada hesaplanmış tanımlayıcılarla modeli geliştirmek için çoklu doğrusal regresyon yöntemi kullanılmıştır. Güncel iç ve dış doğrulama parametreleri hesaplanmıştır. Böylece in silico toksisite profileri yapısal uyarılar ve gözlemlenen en düşük advers etki düzeyi değerleri çevrimiçi araçları kullanılarak değerlendirilmiştir. Yeni bileşikler tasarlanmış ve geliştirilen model üzerinde test edilmiştir.

Bulgular: Bir dizi 4-(2-florofenoksi) kinolin türevi doğruluk olarak modellenmiş ve moleküllerin c-MET kinaz inhibisyon potansiyeline tahmin etmek için doğrulanmıştır. Geliştirilen modelin istatistiksel metrikleri, sağlam olduğunu ve bu kimyasal sınıf için başarılı tahminler yapabilğini göstermiştir. Kütle, elektronegatiflik, kısmi yükler ve moleküllerin yapısının aktive üzerinde etkili olduğu görülmüştür. Ek olarak, inceленen bileşiklerin toksisite profilinin kabul edilebilir olduğu görülmüştür.

Sonuç: Sentezlenen bileşiklerden beşinin risk-yarar profili incelendiğinde, etkisinin risklerinden daha fazla olduğu görülmüştür. Geliştirilen model, sanal tarama ve yeni anti-tümör bileşiklerinin tasarlanmasına yararlı olduğu belirlenmiştir.

Anahtar kelimeler: Toksisite, LOAEL, anti-tümör, c-MET, QSAR

*Correspondence: gulcin.tugcu@yeditepe.edu.tr, Phone: +90 216 578 00 00, ORCID-ID: orcid.org/0000-0002-9750-6563
Received: 05.02.2021, Accepted: 15.03.2021
© Turk J Pharm Sci, Published by Galenos Publishing House.
INTRODUCTION

Cancer is a worldwide health problem. Hepatocyte growth factor receptor (c-MET) kinase inhibition has been a novel treatment for various human cancers. In addition to being effective in treating diseases, the approved drugs are expected to be safe at an acceptable level. In other words, the risk-benefit profile of these drugs was considered to be greater than their risks. Among these compounds, crizotinib and cabozantinib are two c-MET inhibitors that meet these criteria and are used in the treatment of various cancers. There are other molecules that have been added to the list of approved drugs. However, studies on promising c-MET tyrosine kinase inhibitors such as JNJ-38877605 were suspended in phase 1 due to the high risk of toxicity. Therefore, before starting clinical trials, it is greatly important to explore the safety profile of a drug candidate in the preclinical phase as early as possible so that studies for safer drugs continue expeditiously. Quantitative structure-activity-relationship (QSAR) modeling can be utilized to reduce costs, save time, and understand mechanisms of active substances even before a chemical is synthesized.

The safety evaluation of drugs is determined as the ratio of the safe amount of the drug exposed to the amount of drug that is effective. All preclinical data on the efficacy and safety endpoints need to be used as early as possible to understand the preliminary therapeutic index (TI). This index is calculated by dividing the highest amount of the drug that does not cause any toxicity to the amount that yields the aimed effect. The derivation of TI ratios from translational studies helps in proper decision-making for a drug in the next stage of drug development.

This study developed a QSAR model to estimate the c-MET kinase inhibition of a set of compounds. The compounds' toxicity profiles are then explored in silico, and several compounds were determined to be both effective and safe. The model, which revealed satisfying internal and external validation statistics, was then applied to a set of newly designed compounds.

MATERIALS AND METHODS

Data set

A data set of 32 4-(2-fluorophenoxy) quinoline derivatives as c-MET kinase inhibitors was obtained from the literature. The negative logarithm of the activity value (p c-MET) was used as the dependent variable in the model equation. Table 1 lists the structures and biological activities (IC_{50}) of the studied compounds.

Molecular descriptors and descriptor selection

Compounds were drawn in Spartan v.18 (Wavefunction Inc., Irvine, CA) and corresponding geometries were optimized with the semi-empirical PM7 using MOPAC 6. The Chemopy v.3.2, PaDEL v.2.21, and alvaDesc 1.0.20 (www.alvascience.com/alvades) descriptors were calculated using geometry-optimized molecule files. Descriptor selection was performed via genetic algorithm (GA) using QSARINS v.2,2,4 following the removal of the constant and near-constant value descriptors. A search with GA within the software follows the tournament selection method.

Model development and validation

The data set was split into training and test sets for an external validation procedure by systematically selecting compounds for the test set from the data set sorted according to activity. In this way, the activities were distributed homogenously and the training set was set as large as possible. Additionally, training and test sets are congruent so that the test set resides in the applicability domain. The training set consisted of 80% of the whole data set.

Selected significant descriptors were used as independent variables in MLR models to predict the biological activity. The delta K limit was set at 0.05 to eliminate models with collinear descriptors. The maximum number of descriptors in the model has been limited to at least five compounds per descriptor (Topliss Ratio) to avoid overfitting in the model. Linear models were developed using the ordinary least square method, which is carried out in QSARINS v.2,2,4 with the selected descriptors as the independent variables.

Internal and external validation procedures were performed vigorously using widely known parameters in the literature. The coefficient of determination (R^2), leave-one-out cross-validation (Q_{loo}^2), standard error (SE), root mean squared error (RMSE), and Fischer statistics (F) as model fitting and internal validation criteria were listed. The reliability of the developed model was tested by the randomization procedure (Y-scrambling). Activity values in the training set were rearranged and new correlation coefficients were calculated for the randomized models. The average correlation coefficient of the new models was expected to be significantly low corresponding to the model correlation coefficient, indicating the absence of a chance correlation.

External validation parameters, representing the predictive ability of the model, were listed as R^2, RMSE, Q_{loo}^2, Q_{loo}^2, (0.70), Q_{loo}^2, (0.70), Q_{loo}^2, (0.70), CCC (0.85), and r_{m}^2 (0.65) of the test set with the recommended lower limits given in parentheses. Furthermore, conditions described by Golbraikh and Tropsha were applied to the test set predictions:

I. (R^2 - R_{o}^2)/R^2 < 0.1 and 0.85 ≤ k ≤ 1.15
II. |R_{o}^2 - R_{o}^2| < 0.3,

where R^2 corresponds to the predicted vs. observed, R_{o}^2 refers to the observed vs. predicted, k and k’ are slopes, and R_{o}^2 and R_{o}^2 are coefficients of determination passing through the origin.

Applicability domain

The applicability domain of the model was defined via the Williams plot. Leverage (hat) values lie on the x-axis, while standardized residuals are on the y-axis in the graph. Response outliers were diagnosed by the limit of ±2.5 standardized
Table 1. Molecular structures and activities of studied compounds

| Compound | R1    | R2              | IC₅₀ (nm) |
|----------|-------|-----------------|-----------|
| 10a      | CH₃   | Cyclopentyl     | 52.86     |
| 10b      | CH₃   | 2-thiophenyl    | 37.46     |
| 10c      | CH₃   | Naphtyl         | 48.36     |
| 10d      | CH₃   | Phenyl          | 22.74     |
| 10e      | propyl| Phenyl          | 19.45     |
| 10f      | cyclohexyl| Phenyl      | 30.38     |
| 10g      | phenyl| Phenyl          | 54.42     |
| 10h      | 'But  | Phenyl          | 8.35      |
| 10i      | 'But  | 4-methylphenyl- | 12.35     |
| 10j      | 'But  | 4-methoxyphenyl-| 21.48     |
| 10k      | 'But  | 3,4,5-trimethoxyphenyl-| 35.64     |
| 10l      | 'But  | 'But            | 18.45     |
| 10m      | 'But  | 4-fluorophenyl- | 2.49      |
| 10n      | 'But  | 3-fluorophenyl- | 4.72      |
| 10o      | 'But  | 2-fluorophenyl- | 3.13      |
| 10p      | 'But  | 4-chlorophenyl- | 7.43      |
| 10q      | 'But  | 3-chlorophenyl- | 7.86      |
| 10r      | 'But  | 2-chlorophenyl- | 8.12      |
| 10s      | 'But  | 4-bromophenyl- | 18.44     |
| 10t      | 'But  | 4-CF₃-phenyl-   | 36.62     |
| 10u      | 'But  | 2,3-dichlorophenyl-| 18.66     |
| 10v      | 'But  | 2-thiophenyl   | 32.74     |
| 10w      | 'But  | 2-furanyl      | 42.68     |
| 10x      | 'But  | Cyclopentyl     | 66.54     |
| 10y      | 'But  | Naphtyl        | 46.87     |

| Compound | R1    | R2            | R3            | IC₅₀ (nm) |
|----------|-------|---------------|---------------|-----------|
| 11a      | 'But  | 4-fluorophenyl-| Butyl         | 10.46     |
| 11b      | 'But  | 4-fluorophenyl-| 'But'         | 18.46     |
| 11c      | 'But  | 4-fluorophenyl-| Cyclohexyl    | 20.54     |
| 11d      | 'But  | 4-fluorophenyl-| Phenyl        | 30.92     |
| 11e      | 'But  | 4-fluorophenyl-| 3,4-dimethoxyphenyl-| 24.45     |
| 11f      | 'But  | 4-fluorophenyl-| 4-fluorophenyl-| 36.46     |
| 11g      | 'But  | 4-fluorophenyl-| H             | 91.12     |
residuals, while the structural outlier threshold was set at the critical hat (h*). The critical hat value was calculated as 3p/n, where p is the number of descriptors plus one and n is the number of molecules in the training set.

Exploring toxicity profile of c-MET inhibitors

The toxicity profile of c-MET inhibitors was predicted using online tools and software. As a first step, structural alerts were sought for pan-assay interference compounds (PAINS)19 and Brenk et al.21 using SwissADME.22 These alerts filter unwanted substructural features and screen compounds for lead-likeness. One of the parameters used in the safety evaluation is the lowest observed adverse effect level (LOAEL), which is defined as the lowest dose in a study at which adverse effects were observed, although it can lead to a significant overestimation of the TI.3 LOAEL values (obtained from oral rat chronic toxicity tests) were predicted via pkCSM.23

The TI can be used as the ratio of the highest amount of the drug that does not cause any toxicity to the amount that yields/creates the desired effect in a translational research environment. This study used in vitro efficacy and in silico safety endpoint to compare safety margin interpretations. The greater the difference between the effective dose of a drug candidate molecule and its toxic dose, the safer that drug is. Consequently, a ratio was calculated as LOAEL to IC_{50} to evaluate the toxicity over the activity. This study defined a threshold level of 1 for this ratio.

Design of new c-MET inhibitors

New compounds were designed to explore potential anti-tumor agents. Inspired from papers of Parikh and Ghate24 and Liu et al.25 useful molecular fragments were selected. For instance, pyrazol and oxolane functional groups were introduced to the designed molecules. Additionally, the single methoxy group on the quinoline backbone was proved on some compounds. New compounds were designed to explore potential anti-tumor agents. Inspired from papers of Parikh and Ghate24 and Liu et al.25 useful molecular fragments were selected. For instance, pyrazol and oxolane functional groups were introduced to the designed molecules. Additionally, the single methoxy group on the quinoline backbone was proved on some compounds.

RESULTS AND DISCUSSION

The training and test set selection was performed using systematic division. Activity values were sorted in decreasing order. The most and the least active compounds were assigned into the training set to provide a broader applicability domain. Finally, 80% of the data set was allocated into the training set. A total of 5182 descriptors were calculated using Chemopy (633), Padel (1444), and alvaDesc (3105) programs for all compounds. Selected descriptors were used as independent variables in the model developed on the training set. A GA-based search with parameters of 500 iterations, a population of 50 models, and a mutation rate of 20 was employed to find significant descriptors for the model. The four-descriptor MLR model (Eqn. 1) below was found to be predictive with an acceptable goodness-of-fit value. SEs of the coefficients were given in the parentheses (p<0.05).

\[
\text{c-MET = 30.428 (±8.794) + 0.087 (±0.018) PEOEVSA2 + 0.034 (±0.016) AATSC4m -5.751 (±1.094) SpMin8_Bh(e) -25.005 (±8.584) VR2}\_\text{RG (1)}
\]

The reliability of the model was investigated using a randomization test, which was run for 2000 iterations for the developed model. The average R^2 was 0.159. Additionally, the distinct distribution of R^2 and Q^2 values belong to randomized models, and the developed model (Figure SI) revealed that the generated model is robust and the descriptors are not selected by chance.

The four-descriptor linear model has fulfilled the internal and external validation criteria listed in the materials and methods section. Figure 1 depicts the predicted versus observed activities, which are scattered around the y = x line. Table SI1 gives the predicted values by the model, leverages, and standardized residuals for the studied compounds. Table SI2 gives the calculated descriptor values, while Table 2 lists the standardized coefficients of the descriptors in the model, representing their relative importance. Partial charges, surface area contributions, and the Sanderson electronegativity were seen to be the most influential properties of the compounds. The result indicated that partial charges, surface area, mass, electronegativity, and the structure of the molecules contributed to the activity (Table 2). Table 2 lists the standardized coefficients of the descriptors in the model, representing their relative importance. Partial charges, surface area contributions, and the Sanderson electronegativity were seen to be the most influential properties of the compounds. Table SI2 gives the calculated descriptor values, while Table

Figure 1. Predicted versus experimental activity values of training and test set compounds
The correlation matrix of the descriptors of the model, showing that the descriptors were not intercorrelated.

The Williams plot was depicted to show both the structurally distant compounds to the others (leverages) and the standardized residuals that represent the model’s prediction accuracy. Horizontal reference lines are at ±2σ and the vertical reference line is at the critical hat value (h* = 0.877). The model did not have any outliers in terms of leverages and standardized residuals. Hence, all compounds are located within the AD of the model. Supplementary information Table S1 lists the leverages and standardized residuals for the studied compounds.

The compounds were inspected in silico according to their structural alerts and LOAELs to explore the toxicity profile of the set of c-MET inhibitors as potential anti-tumor agents. No alert was fired for PAINS and Brenk screens in the SwissADME software. LOAEL values that were predicted via pkCSM were used to evaluate the safety of drugs. LOAEL predictions varied between 0.237 and 75.683 mg/kg/day. 10m, 11a, 11b, 11c, and 11g with the highest toxic dose/effective dose ratio were seen to be safe at the effective dose as expected from a drug (Table SI3) (Figure 2).

As described in the materials and methods section in detail, 20 quinoline derivatives were designed. The developed model was deployed on this set of new compounds (Table SI4) to explore possible inhibitors. Among 20 compounds, 16 of them were located in the AD and showed promising results as drug candidates (Figure 3). Compounds D06 and D07 were structurally distant from the modeled compounds. D02 and D13 had extreme prediction values, resulting in falling outside the AD of the model.

### Table 2. Descriptors used in the model along with their standardized coefficients

| Descriptor | Software | Meaning | Standardized coefficient |
|------------|----------|---------|-------------------------|
| PEOEVSA2   | ChemoPy  | MOE-type descriptor using partial charges and surface area contributions | 0.492 |
| AATSC4m    | PaDEL    | Autocorrelation descriptor Average-centered Broto-Moreau autocorrelation - lag 4 (weighted by mass) | 0.297 |
| SpMin8_Bh(e) | alvaDesc | Smallest eigenvalue n.8 of the Burden matrix weighted by the Sanderson electronegativity | -0.610 |
| VR2_RG     | alvaDesc | 3D matrix-based descriptor Normalized Randic-like eigenvector-based index from the reciprocal squared geometrical matrix | -0.350 |

*ChemoPy: Chemoinformatics in Python, PaDEL: PaDEL-Descriptor, alvaDesc: Alvascience Descriptor*

### Table 3. Correlation matrix of descriptors used in the model

|        | PEOEVSA2 | AATSC4m | SpMin8_Bh(e) | VR2_RG |
|--------|----------|---------|--------------|--------|
| PEOEVSA2 | 1        | -       | -            | -      |
| AATSC4m | 0.14     | 1       | -            | -      |
| SpMin8_Bh(e) | -0.17 | 0.52    | 1            | -      |
| VR2_RG | 0.02     | -0.61   | -0.25        | 1      |

CONCLUSION

A QSAR model was developed with four descriptors to predict the c-MET inhibitory activity. The model revealed that the mass, electronegativity, partial charges, and the structure of the molecules have an influence on the c-MET inhibition. The model was validated internally and externally to prove its robustness and predictivity. Compounds were predicted for their availability as drug candidates and toxicity profile. It can be concluded that the compounds are opportune to be drug candidates and five of them ensued promising as both active and relatively safe. Additionally, the developed model was applied to 20 newly designed compounds. The model showed promising results on the new compounds. The model can be used in the early design process of c-MET inhibitors as anti-tumor agents.
ACKNOWLEDGMENTS

The authors would like to thank Prof. Paola Gramatica for providing QSARINS program.

Conflict of interest: No conflict of interest was declared by the authors. The authors are solely responsible for the content and writing of this paper.

REFERENCES

1. WHO, 2018. Available from: https://www.who.int/health-topics/cancer#tab=tab_1
2. Liu J, Gong Y, Shi J, Hao X, Wang Y, Zhou Y, Hou Y, Liu Y, Ding S, Chen Y. Design, synthesis and biological evaluation of novel N-[4-(2-fluorophenoxy) pyridin-2-yl] cyclopropanecarboxamide derivatives as potential c-Met kinase inhibitors. Eur J Med Chem. 2020;194:112244.
3. Muller P, Milton M. The determination and interpretation of the therapeutic index in drug development. Nat Rev Drug Discov. 2012;11:751-761.
4. Mo HN, Liu P. Targeting MET in cancer therapy. Chronic Dis Transl Med. 2017;3:148-153.
5. Puccini A, Marin-Ramos NI, Bergamo F, Schirripa M, Lonardi S, Lenz HJ, Loupakis F, Battaglin F. Safety and tolerability of c-MET inhibitors in cancer. Drug Saf. 2019;42:211-233.
6. Wang Q, Yang S, Wang K, Sun SY. MET inhibitors for targeted therapy of EGFR TKI-resistant lung cancer. J Hematol Oncol. 2019;12:63.
7. Loikema MP, Bohets HH, Arkenau HT, Lampo A, Barale E, de Jonge MJA, van Doorn L, Hellemans P, de Bono JS, Eskenys FALM. The c-Met tyrosine kinase inhibitor JNJ-38877605 causes renal toxicity through species-specific insoluble metabolite formation. Clin Cancer Res. 2015;21:2297-2304.
8. Bass AS, Cartwright ME, Mahon C, Morrison R, Snyder R, McNamara P, Bradley P, Zhou YY, Hunter J. Exploratory drug safety: a discovery strategy to reduce attrition in development. J Pharmacol Tox Met. 2009;60:69-78.
9. Nan X, Li HJ, Fang SB, Li QY, Wu YC. Structure-based discovery of novel 4-(2-fluorophenoxo) quinoline derivatives as c-Met inhibitors using isocyanide-involved multicomponent reactions. Eur J Med Chem. 2020;193:112241.
10. Stewart JJ. MOPAC: a semiempirical molecular orbital program. J Comput Aid Mol Des. 1990;4:1-103.
11. Dong J, Cao DS, Miao HY, Liu S, Deng BC, Yun YH, Wang NN, Lu AP, Zeng WB, Chen AF. ChemDes: an integrated web-based platform for molecular descriptor and fingerprint computation. J Cheminform. 2015;7:60.
12. Yap CW. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. J Comput Chem. 2011;32:1466-1474.
13. Gramatica P, Chirico N, Papa E, Kovarich S, Cassani S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. J Comput Chem. 2013;34:2121-2132.
14. Gramatica P, Cassani S, Chirico N. QSARINS-Chem: Insubria datasets and new QSAR/QSPR models for environmental pollutants in QSARINS. J Comput Chem. 2014;35:1036-1044.
15. Haupt RL, Haupt SE. Practical Genetic Algorithms (2nd ed). New Jersey: Wiley-Interscience; 2004:103-104.
16. Todeschini R, Consonni V. Molecular descriptors for chemoinformatics: volume I: alphabetical listing. Weinheim; John Wiley & Sons; 2009:647.
17. Lee PH, Cucurull-Sanchez L, Wu J, Du VJ, Lee, PH, Cucurull-Sanchez L, Lu J, Du VJ. Development of in silico models for human liver microsomal stability. J Comput Aid Mol Des. 2007;21:665-673.
18. Gramatica P, Sangion A. A historical excursus on the statistical validation parameters for QSAR models: a clarification concerning metrics and terminology. J Chem Inf Model. 2016;56:1127-1131.
19. Golbraikh A, Tropsha A. Beware of q2! J Mol Graphics Modell. 2002;20:269-276.
20. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem. 2010;53:2719-2740.
21. Brenk R, Schipani A, James D, Krasowski A, Gilbert IH, Frearson J, Wyatt PG. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. ChemMedChem. 2008;3:435-444.
22. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:1-13.
23. Pires DE, Blundell TL, Ascher DB. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. J Med Chem. 2015;58:4066-4072.
24. Parikh PK, Ghate MD. Recent advances in the discovery of small molecule c-Met Kinase inhibitors. Eur J Med Chem. 2018;143:1103-1138.
25. Liu L, Norman MH, Lee M, Xi N, Siegmund A, Boezio AA, Booker S, Choquette D, D’Angelo ND, Germain J, Yang K, Yang Y, Zhang Y, Bellon SF, Whittington DA, Harmange JC, Dominguez C, Kim TS, Dussault I. Structure-based design of novel class II c-Met inhibitors: 2. SAR and kinase selectivity profiles of the pyrazolone series. J Med Chem. 2012;55:1868-1897.

Click for Supplementary access link: https://cms.galenos.com.tr/SolvePark/Uploads/Files/f51106ab41b841c9b0063ece0c95030e.pdf