**Iranian Journal of Analytical Chemistry**
Volume 8, Issue 1, March 2021 (29-38) Origonal Research Article

**مجله ایرانی شیمی تجزیه**
دوره۸، شماره ۱، فروردین ۱۴۰۰ (۳۸-۲۹)

# Molecular Modeling Studies of Triazolyl Thiophenes As CDK5/P25 inhHibitors Using 3D-QSAR and Molecular Docking

**Zahra Garkani-Nejad [1]*, Abouzar Ghanbari [2]**
1. Chemistry Department, Faculty of Science, Shahid Bahonar University of Kerman, Kerman, Iran
2. MAPNA Group Operation and Maintenance, North Mosaddeq St., Mirdamad Blvd., Tehran, Iran

**Received:** 7 January 2021 **Accepted:** 11 March 2021
**DOI:** 10.30473/ijac.2021.59304.1189

**Abstract**
Three-dimensional quantitative structure-activity relationship (3D-QSAR) techniques are useful methods for ligand-based drug design by correlating physicochemical descriptors from a set of related compounds to their known molecular activity or molecular property values. A novel clubbed triazolyl thiophene series of cdk5/p25 inhibitors were selected to establish 3D-QSAR models using Comparative molecular field analysis (CoMFA) and Comparative molecular similarity indices analysis (CoMSIA) methods. The optimum CoMFA and CoMSIA models obtained, were statistically significant with cross-validated correlation coefficients $r^2_{cv}$ ($q^2$) of 0.539 and 0.558, and conventional correlation coefficients ($r^2$) of 0.980 and 0.967, respectively. A training set containing 88 molecules and a test set containing 24 molecules served to establish the QSAR models. Independent test set validated the external predictive power of both models with predicted correlation coefficients ($r^2_{pred}$) 0.968 and 0.945 for CoMFA and CoMSIA, respectively. Molecular docking was applied to explore the binding mode between the ligand and the receptor. The information obtained from molecular modeling studies may be helpful to design novel CDK5/P25 inhibitors with desired activity.

**Keywords:** 3D-QSAR; Molecular Docking; Cdk5/Pp25 Inhibitors; Triazolyl Thiophene; Alzheimer Disease.

## 1. INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disorder characterized by multiple cognitive deficits and progressive memory impairment in mid- to late-life. Both genetic and environmental factors have been implicated in the development of AD, but it is still unclear how these factors combine and ultimately lead to the neurodegenerative process [1-3]. The neuropathological processes underlying AD include the extracellular deposition of β-amyloid protein in senile plaques, the intracellular formation of neurofibrillary tangles, and loss of cholinergic neurons in areas of the brain associated with learning and memory, executive functioning, behavior and emotional responses [4].

Cyclin dependent kinase5 (CDK5) plays an essential role in the development of the central nervous system. Its deregulation has profound cytotoxic effects and has been implicated in the development of neurodegenerative diseases such as Alzheimer's disease. Cdk5 is a member of a family of proline directed serine/threonine kinases. The serine/threonine kinase cdk5 along with its cofactor p25 has been supposed to hyper phosphorylate tau, leading to the formation of paired helical filaments and deposition of cytotoxic neurofibrillary tangles and thus responsible to neurodegenerative disorders such as Alzheimer's disease and Parkinson's disease [5-10].

QSAR methods have been applied in different scientific studies including chemistry, biology, toxicology and drug discovery to predict and classify biological activities of virtual or newly-synthesized compounds [11-12]. QSAR analysis in computational research is responsible for the generation of models to correlate biological activity and physicochemical properties of a series of molecules. The underlying assumption is that the variations of biological activity within a series can be correlated with changes in measured or computed molecular features of the molecules. A successful 3D-QSAR model not only helps in better understanding of the structure activity relationships of any class of molecules, but also provides researcher with an insight at molecular level about the lead molecules for further developments. Quantitative structure activity relationship QSAR models have another ability, which is providing a deeper knowledge about the mechanism of biological activity.

In the present work, a series of triazolyl thiophene derivatives as cdk5/p25 inhibitors potentially useful for the treatment of Alzheimer's disease, were selected to establish three-dimensional quantitative structure activity relationship (3D-

*Corresponding Author: z_garkani@uk.ac.ir

QSAR) models using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) methods. CoMFA relates the biological activity of a series of molecules with their steric and electrostatic fields. The basic principle of CoMSIA is the same as that of CoMFA, but includes some additional descriptors such as hydrophobicity, hydrogen bond donor and hydrogen bond acceptor.

Both approaches of CoMFA and CoMSIA are based on not only molecular structure, but also the calculation of the interaction energy between the molecule and probe. In these approaches, the structure of the studied molecules should be aligned and placed in a 3D lattice constituted by several thousands of grid points and use a probe (steric, electrostatic, hydrophilic, etc.) to map the surface of the molecule on the basis of the molecule interaction with the probe. The approaches of CoMFA and CoMSIA are almost similar except for molecular similarity, which is computed in the case of CoMSIA. In CoMSIA, the Lennard–Jones and Coulombic potentials used in CoMFA to compute the steric and electrostatic grids are replaced by an exponential functional form derived from alignment algorithm.

## 2. EXPERIMENTAL

### *2.1. Data set*

All of one hundred and twelve compounds of methyl linked cyclo hexyl thiophenes with triazole inhibitors and experimental IC50 values were collected from the literature [13]. The IC50 values were converted to the corresponding log IC50 and used as dependent variable in CoMFA and CoMSIA studies. The log IC50 values provide a homogenous data set for 3D-QSAR study. The 3D-QSAR models were generated using a training set including 88 molecules. A test set including 24 molecules was used to an external validation of the models. The structures of different groups of studied compounds along with different substituents are shown in Tables 1 and 2.

**Table 1.** The structure of different groups of methyl linked cyclo hexyl thiophenes with triazole studied in this work

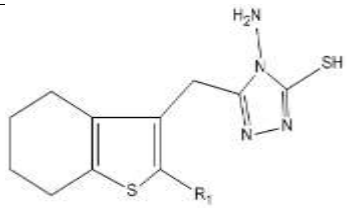
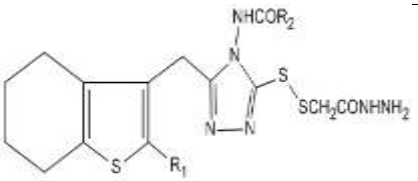
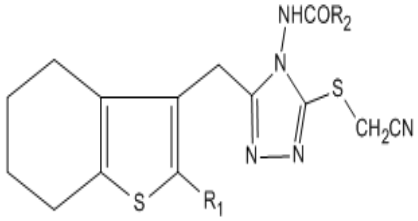
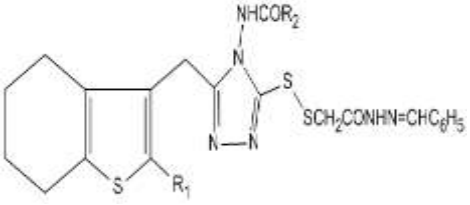
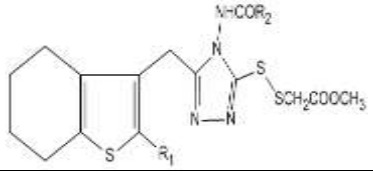
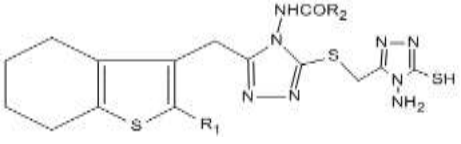
| No. | Structure | No. | Structure |
|---|---|---|---|
| 1 |  | 5 |  |
| 2 |  | 6 |  |
| 3 |  | 7 |  |
| 4 |  | 8 |  |

**Table 2.** Data set with different substituents along with experimental and predicted log IC50 values using CoMFA and CoMSIA models

| No. | Comp. | R1 | R2 | R3 | Log IC50 | | |
|---|---|---|---|---|---|---|---|
| | | | | | Exp. | CoMFA | CoMSIA |
| 1 | 1a | $NHCOCH_2Cl$ | - | - | 1.63 | 1.562 | 1.721 |
| 2 | 1b | $NHCOCH_3$ | - | - | 2.60 | 2.604 | 2.559 |
| 3 | 1c* | $NHCOC_6H_5$ | - | - | 2.80 | 2.857 | 2.900 |
| 4 | 1d | $NHCH_2CH_2COOH$ | - | - | 2.43 | 2.435 | 2.413 |
| 5 | 2a | $NHCOCH_2Cl$ | -4-$ClC_6H_4$ | - | 2.54 | 2.472 | 2.473 |
| 6 | 2b* | $NHCOCH_3$ | -4-$ClC_6H_4$ | - | 2.66 | 2.666 | 2.680 |
| 7 | 2c | $NHCOC_6H_5$ | -4-$ClC_6H_4$ | - | 2.26 | 2.244 | 2.227 |
| 8 | 2d | $NHCH_2CH_2COOH$ | -4-$ClC_6H_4$ | - | 2.66 | 2.654 | 2.667 |
| 9 | 2e* | $NHCOCH_2Cl$ | -$C_6H_5$ | - | 2.88 | 2.900 | 2.866 |
| 10 | 2f | $NHCOCH_3$ | -$C_6H_5$ | - | 2.76 | 2.676 | 2.731 |
| 11 | 2g | $NHCOC_6H_5$ | -$C_6H_5$ | - | 2.95 | 2.958 | 2.932 |
| 12 | 2h | $NHCH_2CH_2COOH$ | -$C_6H_5$ | - | 2.83 | 2.879 | 2.787 |
| 13 | 2i | $NHCOCH_2Cl$ | -$CH_3$ | - | 2.84 | 2.857 | 2.748 |
| 14 | 2j | $NHCOCH_3$ | -$CH_3$ | - | 2.77 | 2.805 | 2.675 |
| 15 | 2k | $NHCOC_6H_5$ | -$CH_3$ | - | 2.79 | 2.802 | 2.799 |
| 16 | 2l | $NHCH_2CH_2COOH$ | -$CH_3$ | - | 2.64 | 2.641 | 2.621 |
| 17 | 3a | $NHCOCH_2Cl$ | -4-$ClC_6H_4$ | - | 1.76 | 1.888 | 2.093 |
| 18 | 3b | $NHCOCH_3$ | -4-$ClC_6H_4$ | - | 2.62 | 2.491 | 2.275 |
| 19 | 3c | $NHCOC_6H_5$ | -4-$ClC_6H_4$ | - | 2.74 | 2.700 | 2.674 |
| 20 | 3d | $NHCH_2CH_2COOH$ | -4-$ClC_6H_4$ | - | 2.62 | 2.631 | 2.615 |
| 21 | 3e* | $NHCOCH_2Cl$ | -$C_6H_5$ | - | 2.82 | 2.710 | 2.708 |
| 22 | 3f* | $NHCOCH_3$ | -$C_6H_5$ | - | 2.60 | 2.667 | 2.687 |
| 23 | 3g | $NHCOC_6H_5$ | -$C_6H_5$ | - | 2.69 | 2.700 | 2.735 |
| 24 | 3h | $NHCH_2CH_2COOH$ | -$C_6H_5$ | - | 3.69 | 3.688 | 3.735 |
| 25 | 3i* | $NHCOCH_2Cl$ | -$CH_3$ | - | 1.78 | 2.042 | 2.136 |
| 26 | 3j | $NHCOCH_3$ | -$CH_3$ | - | 2.66 | 2.519 | 2.417 |
| 27 | 3k | NHCOC6H5 | -$CH_3$ | - | 2.75 | 2.720 | 2.797 |
| 28 | 3l* | $NHCH_2CH_2COOH$ | -$CH_3$ | - | 2.81 | 2.742 | 2.809 |
| 29 | 4a | $NHCOCH_2Cl$ | -4-$ClC_6H_4$ | - | 2.65 | 2.644 | 2.706 |
| 30 | 4b | $NHCOCH_3$ | -4-$ClC_6H_4$ | - | 2.59 | 2.523 | 2.642 |
| 31 | 4c | $NHCOC_6H_5$ | -4-$ClC_6H_4$ | - | 2.47 | 2.409 | 2.445 |
| 32 | 4d | $NHCH_2CH_2COOH$ | -4-$ClC_6H_4$ | - | 2.73 | 2.725 | 2.719 |
| 33 | 4e | $NHCOCH_2Cl$ | -$C_6H_5$ | - | 2.83 | 2.775 | 2.766 |
| 34 | 4f | $NHCOCH_3$ | -$C_6H_5$ | - | 2.76 | 2.724 | 2.76 |
| 35 | 4g | $NHCOC_6H_5$ | -$C_6H_5$ | - | 2.55 | 2.561 | 2.497 |
| 36 | 4h* | $NHCH_2CH_2COOH$ | -$C_6H_5$ | - | 2.81 | 2.834 | 2.843 |
| 37 | 4i | $NHCOCH_2Cl$ | -$CH_3$ | - | 2.84 | 2.804 | 2.768 |
| 38 | 4j | $NHCOCH_3$ | -$CH_3$ | - | 2.59 | 2.746 | 2.764 |
| 39 | 4k | $NHCOC_6H_5$ | -$CH_3$ | - | 2.75 | 2.759 | 2.757 |
| 40 | 4l | $NHCH_2CH_2COOH$ | -$CH_3$ | - | 2.78 | 2.831 | 2.917 |
| 41 | 5a | $NHCOCH_2Cl$ | -4-$ClC_6H_4$ | - | 1.59 | 1.563 | 1.507 |
| 42 | 5b* | $NHCOCH_3$ | -4-$ClC_6H_4$ | - | 1.81 | 1.814 | 1.803 |
| 43 | 5c* | $NHCOC_6H_5$ | -4-$ClC_6H_4$ | - | 2.37 | 2.419 | 2.311 |
| 44 | 5d | $NHCH_2CH_2COOH$ | -4-$ClC_6H_4$ | - | 2.30 | 2.266 | 2.257 |
| 45 | 5e | $NHCOCH_2Cl$ | -$C_6H_5$ | - | 1.57 | 1.747 | 1.852 |
| 46 | 5f | $NHCOCH_3$ | -$C_6H_5$ | - | 2.28 | 2.050 | 1.937 |
| 47 | 5g | $NHCOC_6H_5$ | -$C_6H_5$ | - | 2.42 | 2.418 | 2.483 |
| 48 | 5h | $NHCH_2CH_2COOH$ | -$C_6H_5$ | - | 2.66 | 2.622 | 2.665 |
| 49 | 5i | $NHCOCH_2Cl$ | -$CH_3$ | - | 1.79 | 1.884 | 1.854 |
| 50 | 5j | $NHCOCH_3$ | -$CH_3$ | - | 2.40 | 2.422 | 2.415 |
| 51 | 5k | $NHCOC_6H_5$ | -$CH_3$ | - | 2.26 | 2.203 | 2.319 |
| 52 | 5l | $NHCH_2CH_2COOH$ | -$CH_3$ | - | 2.97 | 2.972 | 2.949 |
| 53 | 6a* | $NHCOCH_2Cl$ | -4$ClC_6H_4$ | -$CH_3$ | 2.67 | 2.694 | 2.743 |
| 54 | 6aa | $NHCOC_6H_5$ | -$CH_3$ | -$CH_3$ | 2.81 | 2.833 | 2.819 |
| 55 | 6ab | $NHCH_2CH_2COOH$ | -$CH_3$ | -$CH_3$ | 2.56 | 2.551 | 2.577 |
| 56 | 6ac* | $NHCOCH_2Cl$ | -$CH_3$ | -$C_6H_5$ | 2.56 | 2.522 | 2.540 |
| 57 | 6ad | $NHCOCH_3$ | -$CH_3$ | -$C_6H_5$ | 2.64 | 2.680 | 2.634 |
| 58 | 6ae | $NHCOC_6H_5$ | -$CH_3$ | -$C_6H_5$ | 2.61 | 2.628 | 2.608 |
| 59 | 6af * | $NHCH_2CH_2COOH$ | -$CH_3$ | -$C_6H_5$ | 2.68 | 2.747 | 2.649 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 60 | 6ag | $NHCOCH_2Cl$ | $-CH_3$ | $CH_2Cl$ | 2.66 | 2.663 | 2.662 |
| 61 | 6ah* | $NHCOCH_3$ | $-CH_3$ | $CH_2Cl$ | 2.68 | 2.617 | 2.673 |
| 62 | 6ai | $NHCOC_6H_5$ | $-CH_3$ | $CH_2Cl$ | 2.55 | 2.566 | 2.532 |
| 63 | 6aj | $NHCH_2CH_2COOH$ | $-CH_3$ | $CH_2Cl$ | 2.42 | 2.428 | 2.404 |
| 64 | 6b | $NHCOCH_3$ | $-4-ClC_6H_4$ | $-CH_3$ | 2.81 | 2.791 | 2.771 |
| 65 | 6c | $NHCOC_6H_5$ | $-4-ClC_6H_4$ | $-CH_3$ | 2.50 | 2.473 | 2.513 |
| 66 | 6d | $NHCH_2CH_2COOH$ | $-4-ClC_6H_4$ | $-CH_3$ | 2.55 | 2.498 | 2.569 |
| 67 | 6e* | $NHCOCH_2Cl$ | $-4-ClC_6H_4$ | $-C_6H_5$ | 2.79 | 2.798 | 2.770 |
| 68 | 6f | $NHCOCH_3$ | $-4-ClC_6H_4$ | $-C_6H_5$ | 2.72 | 2.628 | 2.712 |
| 69 | 6g | $NHCOC_6H_5$ | $-4-ClC_6H_4$ | $-C_6H_5$ | 2.68 | 2.695 | 2.668 |
| 70 | 6h* | $NHCH_2CH_2COOH$ | $-4-ClC_6H_4$ | $-C_6H_5$ | 2.43 | 2.454 | 2.399 |
| 71 | 6i | $NHCOCH_2Cl$ | $-4ClC_6H_4$ | $CH_2Cl$ | 2.36 | 2.337 | 2.403 |
| 72 | 6j | $NHCOCH_3$ | $-4-ClC_6H_4$ | $CH_2Cl$ | 2.41 | 2.389 | 2.413 |
| 73 | 6k* | $NHCOC_6H_5$ | $-4-ClC_6H_4$ | $CH_2Cl$ | 2.53 | 2.545 | 2.500 |
| 74 | 6l | $NHCH_2CH_2COOH$ | $-4-ClC_6H_4$ | $CH_2Cl$ | 2.27 | 2.286 | 2.230 |
| 75 | 6m | $NHCOCH_2Cl$ | $-C_6H_5$ | $-CH_3$ | 2.54 | 2.476 | 2.568 |
| 76 | 6n | $NHCOCH_3$ | $-C_6H_5$ | $-CH_3$ | 2.74 | 2.707 | 2.777 |
| 77 | 6o | $NHCOC_6H_5$ | $-C_6H_5$ | $-CH_3$ | 2.75 | 2.764 | 2.752 |
| 78 | 6p* | $NHCH_2CH_2COOH$ | $-C_6H_5$ | $-CH_3$ | 2.67 | 2.667 | 2.669 |
| 79 | 6q | $NHCOCH_2Cl$ | $-C_6H_5$ | $-C_6H_5$ | 2.67 | 2.643 | 2.616 |
| 80 | 6r | $NHCOCH_3$ | $-C_6H_5$ | $-C_6H_5$ | 2.62 | 2.587 | 2.701 |
| 81 | 6s* | $NHCOC_6H_5$ | $-C_6H_5$ | $-C_6H_5$ | 2.67 | 2.712 | 2.696 |
| 82 | 6t | $NHCH_2CH_2COOH$ | $-C_6H_5$ | $-C_6H_5$ | 2.64 | 2.654 | 2.637 |
| 83 | 6u | $NHCOCH_2Cl$ | $-C_6H_5$ | $CH_2Cl$ | 2.73 | 2.760 | 2.792 |
| 84 | 6v | $NHCOCH_3$ | $-C_6H_5$ | $CH_2Cl$ | 2.81 | 2.804 | 2.812 |
| 85 | 6w | $NHCOC_6H_5$ | $-C_6H_5$ | $CH_2Cl$ | 2.74 | 2.738 | 2.718 |
| 86 | 6x | $NHCH_2CH_2COOH$ | $-C_6H_5$ | $CH_2Cl$ | 2.56 | 2.568 | 2.566 |
| 87 | 6y | $NHCOCH_2Cl$ | $-CH_3$ | $-CH_3$ | 2.66 | 2.672 | 2.63 |
| 88 | 6z | $NHCOCH_3$ | $-CH_3$ | $-CH_3$ | 2.51 | 2.507 | 2.512 |
| 89 | 7a | $NHCOCH_2Cl$ | $-4-ClC_6H_4$ | - | 1.54 | 1.643 | 1.547 |
| 90 | 7b | $NHCOCH_3$ | $-4-ClC_6H_4$ | - | 2.33 | 2.306 | 2.357 |
| 91 | 7c | $NHCOC_6H_5$ | $-4-ClC_6H_4$ | - | 2.34 | 2.406 | 2.346 |
| 92 | 7d* | $NHCH_2CH_2COOH$ | $-4-ClC_6H_4$ | - | 2.41 | 2.402 | 2.429 |
| 93 | 7e | $NHCOCH_2Cl$ | $-C_6H_5$ | - | 1.51 | 1.460 | 1.519 |
| 94 | 7f * | $NHCOCH_3$ | $-C_6H_5$ | - | 2.59 | 2.608 | 2.573 |
| 95 | 7g* | $NHCOC_6H_5$ | $-C_6H_5$ | - | 2.64 | 2.692 | 2.664 |
| 96 | 7h | $NHCH_2CH_2COOH$ | $-C_6H_5$ | - | 2.43 | 2.387 | 2.427 |
| 97 | 7i | $NHCOCH_2Cl$ | $-CH_3$ | - | 2.53 | 2.584 | 2.476 |
| 98 | 7j | $NHCOCH_3$ | $-CH_3$ | - | 2.66 | 2.697 | 2.603 |
| 99 | 7k* | $NHCOC_6H_5$ | $-CH_3$ | - | 2.67 | 2.715 | 2.657 |
| 100 | 7l | $NHCH_2CH_2COOH$ | $-CH_3$ | - | 2.73 | 2.745 | 2.734 |
| 101 | 8a | $NHCOCH_2Cl$ | $-4-ClC_6H_4$ | - | 3.46 | 3.530 | 3.588 |
| 102 | 8b | $NHCOCH_3$ | $-4-ClC_6H_4$ | - | 3.80 | 3.624 | 3.685 |
| 103 | 8c | $NHCOC_6H_5$ | $-4-ClC_6H_4$ | - | 3.52 | 3.469 | 3.527 |
| 104 | 8d | $NHCH_2CH_2COOH$ | $-4-ClC_6H_4$ | - | 3.77 | 3.778 | 3.719 |
| 105 | 8e | $NHCOCH_2Cl$ | $-C_6H_5$ | - | 3.80 | 3.756 | 3.736 |
| 106 | 8f | $NHCOCH_3$ | $-C_6H_5$ | - | 3.81 | 3.808 | 3.797 |
| 107 | 8g* | $NHCOC_6H_5$ | $-C_6H_5$ | - | 3.74 | 3.771 | 3.826 |
| 108 | 8h | $NHCH_2CH_2COOH$ | $-C_6H_5$ | - | 3.70 | 3.738 | 3.703 |
| 109 | 8i | $NHCOCH_2Cl$ | $-CH_3$ | - | 3.66 | 3.717 | 3.624 |
| 110 | 8j | $NHCOCH_3$ | $-CH_3$ | - | 3.70 | 3.795 | 3.760 |
| 111 | 8k | $NHCOC_6H_5$ | $-CH_3$ | - | 3.79 | 3.787 | 3.762 |
| 112 | 8l | $NHCH_2CH_2COOH$ | $-CH_3$ | - | 3.79 | 3.792 | 3.809 |

* Test set

*2.2. 3D-QSAR modeling and structure alignment*

All the molecular modeling calculations and database alignment were performed using the molecular modeling package SYBYL [14]. Partial atomic charges calculated using the Gasteiger–Hückel method were assigned to each atom and the energy minimization of each molecule was performed using Powell method and Tripos standard force field with a distance-dependent dielectric function. The minimization was terminated when the energy gradient convergence criterion 0.05 kcal/mol was reached or when the 2000 step minimization cycle limit was exceeded. Structural alignment was considered as one of the most sensitive parameters in CoMFA and CoMSIA study, since much experience shows that the resulting CoMFA and CoMSIA model is often sensitive to the particular alignment [15-18]. In this study, the lowest energy conformation of compound 7e with the most potent activity was

selected as the template structure for the molecular alignment (Fig. 1).

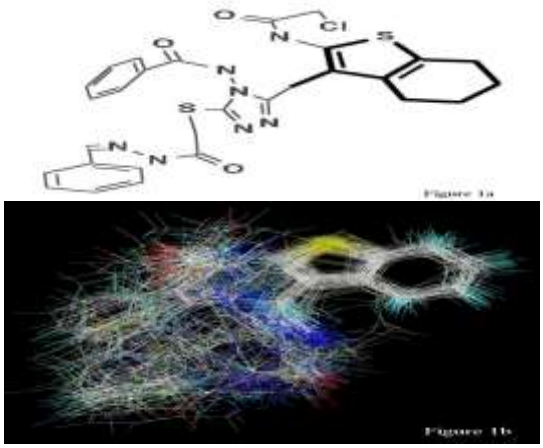


**Fig. 1.** Superposition of compounds in the training and test sets using the common substructure-based alignment rules (a) Compound 7e used as the template molecule for database alignment. (b) Database alignment all the compounds in the dataset are shown.

*2.2.1. CoMFA and CoMSIA studies*

Generally, to achieve better understanding, the contribution of electrostatic field, steric and hydrophobic by a set of data, and also to build predictive 3D-QSAR models, CoMFA and CoMSIA studies were performed based on the molecular alignment as described. The Lennard–Jones potentials and Coulombic terms, which represent steric and electrostatic fields for CoMFA, respectively; whiles CoMSIA calculates similarity indices in the space surrounding each of the molecules in the data set.

In this regard, to generate the 3D-QSAR models, for current alignments, the CoMFA and CoMSIA methods were used. In CoMFA, Steric and electrostatic interactions were calculated using the Tripos force field with a distance-dependent dielectric constant at all interactions in a regularly spaced (2Å) grid taking a $sp^3$carbon atom as steric probe and a (+1) charge as electrostatic probe. All compounds were aligned based on the lowest energy conformer of the reference compound by the atom fit method. With standard options for scaling of variables, the regression analysis was carried out using the full cross-validated partial least squares (PLS) method. The final model, non cross–validated conventional analysis, was developed with the optimum number of components to yield a non cross–validated $r^2$ value.

The CoMSIA procedure is similar to the CoMFA procedure. In this approach, five different similarity fields are calculated: steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor. The CoMSIA similarity index descriptors were derived using the same lattice box as that used in CoMFA calculations. In general, similarity indices, $A_{F,K}$ between the compounds of interest were computed using Equation (1).

$$A^{q}_{F,K(j)} = - \Sigma W_{probe,k} W_{ik} e^{-a r^2_{iq}} \qquad (1)$$

Where q represents a grid point, i is the summation index over all atoms of the molecule j under computation, $W_{ik}$ is the actual value of the physicochemical property k of atom i and $W_{probe,k}$ is the value of the probe atom [19, 20].

*2.3. Partial Least Square (PLS) Analysis*

The Partial Least Squares (PLS) regression technique was used to construct a linear correlation between the 3D-field (independent variables) and the biological activity values (dependent variables). Then the optimum number of components was employed to construct 3D-QSAR models by non-cross-validations to obtain the conventional correlation coefficient $r^2$ and standard deviation (SE) and significant factor F. The analysis procedure was performed by combing the bioactivity values (log IC50) and the corresponding field descriptor variation. Furthermore, the predictive cross-validated coefficient of $q^2$ (or $r^2_{cv}$) and correlation coefficient $r^2_{pred}$ was calculated.

*2.4. Molecular Docking studies*

Molecular docking was carried out to understand the detailed binding model for the active site receptor with its ligands. In order to determine the appropriate binding conformations of studied compounds and check the main factors affecting the activity from the 3D-QSAR models, docking study for active compound 7e was performed. The X-ray crystal structure of CDK5/P25 was retrieved from the Protein Data Bank (PDB). At the beginning of the docking, all the water and ligands were removed and the random hydrogen atoms were added. Then the receptor structure was minimized in 2000 cycles with Powell method in SYBYL [14].

## 3. RESULT AND DISCUSSION

*3.1. CoMFA and CoMSIA analysis*

The CoMFA and CoMSIA analyses were performed corresponding to the strategies mentioned earlier. Parameters demonstrate that both QSAR models obtained are of high degree of confidence and strong predictive ability. The training set was used for model building and the test set used for an external validation of the model. 88 Compounds out of the total 112 derivatives were used as the training set and 24 compounds were used as the test set. The test set compounds were selected randomly in such a way

that the structural diversity and wide range of activity in the data set were included. All analyses were performed with steric and electrostatic fields calculated at each grid point simultaneously. PLS analysis was carried out for the training set and the test set. The experimental and predicted activities of the data set using CoMFA model are shown in Table 2. Figure 2a shows the relationship between the predicted and experimental log IC50 values from the CoMFA model.

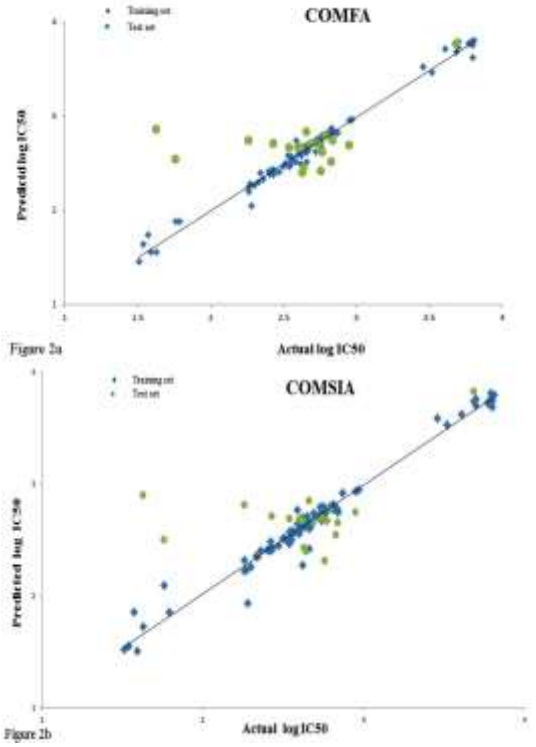



**Fig. 2**. The correlation plots of predicted versus actual log IC50 values using the training and test sets based on (a) CoMFA model and (b) CoMSIA model.

The statistical results of CoMFA and CoMSIA using PLS analysis are presented in Table 3. The CoMFA model gave a cross -validated correlation coefficient $q^2$= 0.539 with an optimal number of principal components of 4 and a non-cross-validated correlation coefficient $r^2$=0.980 with an F value of 744.215 and an estimated standard error (SE) equal to 0.065. The external predictive $r^2_{pred}$ was 0.968 for the CoMFA (Table 3).

**Table 3.** Statistical results observed for the CoMFA and CoMSIA models

| Parameters[a] | CoMFA | CoMSIA |
|---|---|---|
| $q^2$ | 0.539 | 0.558 |
| SE | 0.065 | 0.091 |
| $r^2$ | 0.980 | 0.967 |
| F-values | 744.215 | 377.790 |
| NC | 4 | 5 |
| Fraction | | |
| S | 0.45 | 0.153 |
| E | 0.55 | 0.294 |
| H | -- | 0.271 |
| D | -- | 0.281 |
| A | -- | -- |
| $r^2_{pred}$ | 0.968 | 0.945 |

[a] **$q^2$**, cross-validated correlation coefficient; **SE**, non-cross-validated standard error of estimate; **NC**, number of components; **$r^2$**, non-cross-validated square correlation coefficient; **F**, Fisher's F-value; **$r^2_{pred}$**, predictive $r^2$ in the test set; **S**, steric field; **E**, electrostatic field; **H**, hydrophobic field; **D**, H-bond donor field; and **A**, H-bond acceptor field.

In CoMSIA analysis, the five different descriptor fields, that is, the hydrophobic (H), hydrogen bond donor (D) and acceptor (A), steric (S) and electrostatic (E) fields, are not totally independent from each other. Table 4 summarizes the representative results of different CoMSIA descriptor combinations.

**Table 4.** The regression summary analysis of CoMSIA models.

| Parameters[a] | S | E | H | D | SD | SED | SEHD | SEHDA |
|---|---|---|---|---|---|---|---|---|
| $q^2$ | 0.514 | 0.359 | 0.491 | 0.415 | 0.519 | 0.529 | 0.558 | 0.527 |
| NC | 6 | 5 | 5 | 4 | 6 | 6 | 5 | 5 |
| SE | 0.207 | 0.106 | 0.123 | 0.115 | 0.106 | 0.094 | 0.091 | 0.92 |
| $r^2$ | 0.826 | 0.955 | 0.939 | 0.947 | 0.955 | 0.965 | 0.967 | 0.966 |
| F-Value | 83.131 | 275.619 | 199.659 | 231.406 | 276.014 | 353.608 | 377.790 | 366.367 |
| Fraction | | | | | | | | |
| S | -- | -- | -- | -- | 0.407 | 0.223 | 0.153 | 0.125 |
| E | -- | -- | -- | -- | -- | 0.403 | 0.294 | 0.247 |
| H | -- | -- | -- | -- | -- | -- | 0.271 | 0.225 |
| D | -- | -- | -- | -- | 0.593 | 0.374 | 0.281 | 0.243 |
| A | -- | -- | -- | -- | -- | -- | -- | 0.160 |
| $r^2_{pred}$ | 0.622 | 0.87 | 0.914 | 0.838 | 0.894 | 0.940 | 0.945 | 0.941 |

[a]**$q^2$**, cross-validated correlation coefficient; **SE**, non-cross-validated standard error of estimate; **NC**, number of components; **$r^2$**, non-cross-validated square correlation coefficient; **F**, Fisher's F-value; **$r^2_{pred}$**, predictive $r^2$ in the test set; **S**, steric field **E**, electrostatic field; **H**, hydrophobic field; **D**, H-bond donor field; and **A**, H-bond acceptor field.

The CoMSIA model with a combination of steric, electrostatic, hydrophobic and hydrogen bond donor fields yields the highest cross-validated $q^2$= 0.558 with 5 components, non-cross-validated $r^2$= 0.967, F value of 377.790, and $r^2_{pred}$ of 0.945. The experimental and predicted activities of the data set using CoMSIA model are shown in Table 2. Figure 2b shows the relationship between the predicted and experimental log IC50 values from the CoMSIA model.

*3.2. Contour Analysis*

The contour maps were used to display the fields around the molecules, and to rationalize where changes in each field probably affect the activity of the molecule. Contour maps are plotted as the percentages of the contribution of CoMFA or CoMSIA equation. They show regions where variations of steric, electrostatic, hydrophilic, hydrogen-bond donor or acceptor nature in the structural features of the different molecules lead to an increase or decrease in the activity. In the contour maps, each colored contour represents particular properties such as green contours for regions of high steric tolerance (80% contribution), yellow for low steric tolerance (20% contribution), red contours for regions of decreased electrostatic tolerance for positive charge (20% contribution), blue for regions for decreased electrostatic tolerance for negative charge (80% contribution), yellow contours represent hydrophobically favored regions (80% contribution) and white contours for hydrophobically disfavored regions (20% contribution).To aid in visualization, the most active compound 7e is shown as template molecule with the contour maps. In case of CoMFA, the green contours denote favorable steric interactions and the yellow contours show the region where the steric group was not favored. Figure 3 illustrates the contours of the steric fields, showing in green and in yellow the favored and unfavored bulky groups, respectively. In CoMFA maps, sterically favorable (green) contours are observed (Figure 3a) appearing adjacent around the R1 ($NHCOCH_2Cl$) group and indole ring of compound 7e as well as the presence of a bulky substitution at this position should improve biological activity. Conversely, the sterically unfavorable regions (20% contribution) indicated by a yellow contour spotted around the ($-SH_2CONHN=CHC_6H_5$) of compound 7e reveal that less bulky substituents are not favorable in that region (Figure 3a). The electrostatic effects of the substituents were analyzed by the presence of blue and red color electrostatic contour map. The CoMFA electrostatic map (Figure 3b) displays red contours where the partial negative charge is associated with increased activity (80% contribution) while the blue contours indicate regions where the electropositive properties of molecules cause an increase in the activity.

In CoMSIA (SEHD) contour map for the 7e most active compound is shown in Figure 4. The CoMSIA contour maps consist of steric, electrostatic, and hydrophobic and hydrogen bond donor (SEHD) fields. Figure 4a where the sterically favored regions are shown in green and disfavored regions are shown in yellow, respectively. Figure 4b depicts the electrostatic contour maps obtained from the CoMSIA model, where blue contours represent the favorable electropositive regions and red contours account for the favorable electronegative regions, respectively. Figure 4c shows the CoMSIA hydrophobic contour map, where the yellow (hydrophobic favorable) and white (hydrophobic unfavorable) contours represent 80% and 20% contributions, respectively.

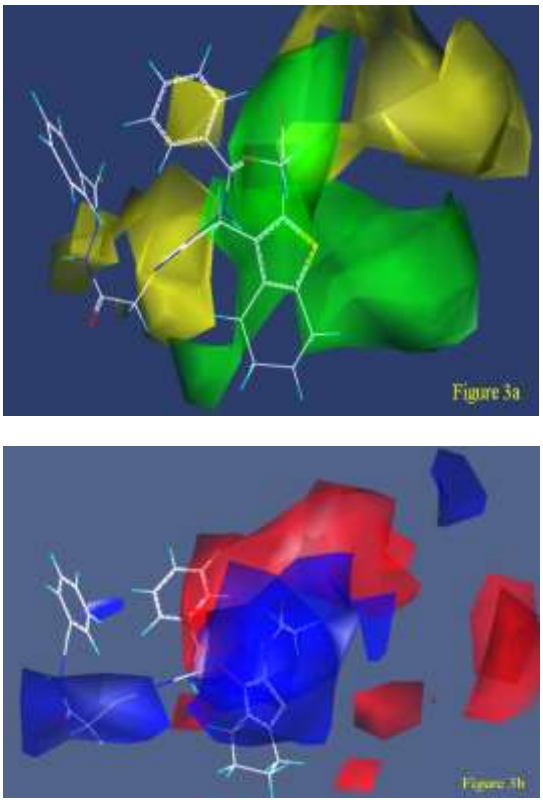



**Fig. 3.** CoMFA steric and electrostatic maps. (a) Sterically favorable and disfavorable areas are shown in green and yellow, respectively and (b) electrostatic field contours shown in red (electronegative substituents favored) and blue (electropositive substituents favored) colors.

Figure 4d shows the CoMSIA hydrogen-bond donor plot, the cyan contours indicate regions where hydrogen bond donor substituents on the ligands are favored and the purple contours represent areas where hydrogen bond donor substituents, are disfavored. Cyan contour plot located around the R1 and R2 positions and indole ring shows that hydrogen bond donor groups exhibit the negative effect on the biological activity.

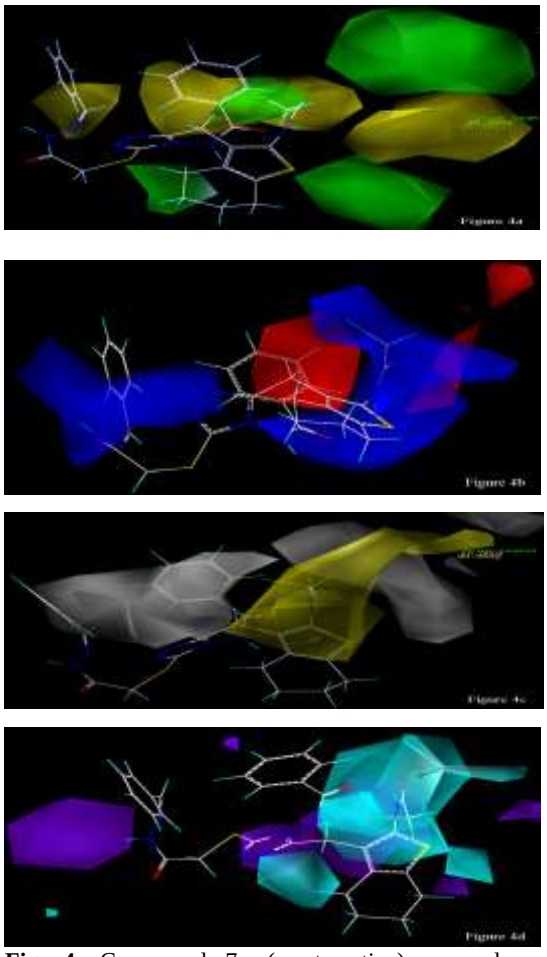


**Fig. 4.** Compound 7e (most active) mapped on CoMSIA: (a) Steric coefficient contour map. Green contours refer to sterically favored regions; yellow contours indicate disfavored areas, (b) Electrostatic coefficient contour map. Red contours refer to regions where negatively charged substituents are disfavored; blue contours indicate regions where negatively charged substituents are favored, (c) Hydrophobic coefficient contour map. Yellow contours refer to regions where hydrophobic substituents are favored; white contours indicate regions where hydrophilic substituents are disfavored, (d) hydrogen bond donor contour maps. Cyan contours refer to regions where H-bond donor substituents are favored; purple contours indicate regions where H-bond donor substituents are disfavored.

*3.3. Docking Results*

Docking studies were applied to investigate the binding mode between CDK5/P25 with 7e molecule. In Figure 5a, the hydrogen bonding (dashed lines) interactions between the reference compound 7e with the highest inhibitory activity was shown and the key residues with two amino acids (SER229 and LYS177) and total of four hydrogen bonds were formed. Indole ring performed two interactions with SER229 and R1 position and two interactions with LYS177.

In Figure 5b, the MOLCAD Multi-Channel cavity depth potential surface structure of the binding site within the compound 7e is displayed. In Figure 5b, hydrogen bonding of the R1 position and Indole ring of compound 7e is observed in a cyan area. In each panel, compound 7e is shown as ball and stick representation; hydrogen bonds are shown as dashed yellow lines; residues are shown as line representation.

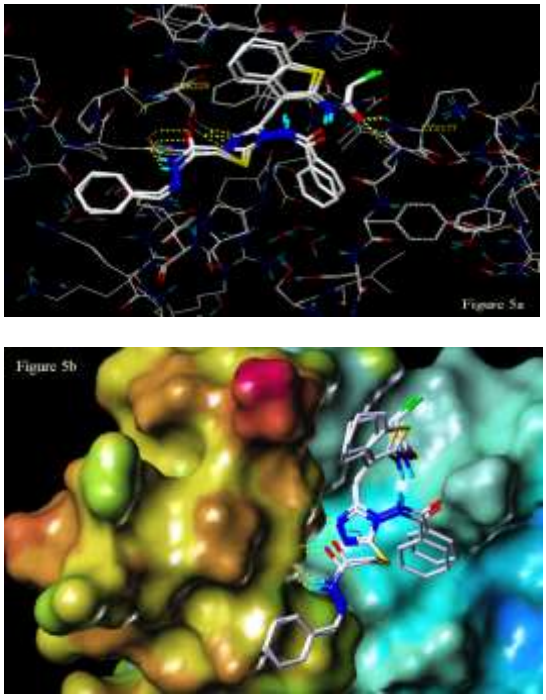



**Fig.5 (a, b).** Structure of compound 7e docked into CDK5/P25.

## 4. CONCLUSION

In this paper, the ligand- and receptor-based 3D-QSAR studies of 112 of triazolyl thiophene derivatives as a series of cdk5/p25 inhibitors performed using CoMFA and CoMSIA tools. From the resultant models, the high $q^2$ ($r^2_{cv}$) and $r^2_{pred}$ values proved that the 3D-QSAR models developed in this work are statistically reliable and predictable. On the basis of the CoMFA and CoMSIA model contour maps, significant regions for steric, electrostatic, hydrophobic, H-bond interactions were identified to enhance the bioactivity. The obtained results can be used as a guideline to design new potent cdk5/p25 inhibitors.

**REFERENCES**


[1] M.W. Bondi, E.C. Edmonds and D.P. Salmon, Alzheimer's Disease: Past, Present, and Future, *J. Int. Neuropsychol. Soc.* 23 (2017) 818-831.

[2] R.U. Haque and A.I. Levey, Alzheimer's disease: A clinical perspective and future nonhuman primate research opportunities, *PNAS* 116 (2019) 26224-26229.

[3] D.J. Selkoe, Alzheimer 's disease: genes, proteins and therapy, *Physiol. Rev.* 81 (2001) 741–766.

[4] J. Park, J. Seo , J. Won , H. Yeo and Y. Ahn, Abnormal Mitochondria in a Non-human Primate Model of MPTP-induced Parkinson's Disease: Drp1 and CDK5/p25 Signaling. *Exp Neurobiol* 28 (2019) 414–424.

[5] K. Pozo and J.A. Bibb, The Emerging Role of Cdk5 in Cancer. *Trends Cancer* 10 (2016) 606–618.

[6] J. Seo, O. Kritskiy, L.A. Watson, S.J. Barker, D. Dey, W.K. Raja, Y.T. Lin, T. Ko, S. Cho, J. Penney, M.C. Silva, S.D. Sheridan, D. Lucente and J.F. Gusella, Inhibition of p25/Cdk5 Attenuates Tauopathy in Mouse and iPSC Models of Frontotemporal Dementia. *J. Neurosci.* 37 (2017) 9917–9924.

[7] Y.L. Zheng, N.D. Amin, Y.F. Hu, P. Rudrabhatla, V. Shukla, J. Kanungo, S. Kesavapany, P. Grant, W. Albers and H.C. Pant, A 24-residue peptide (p5), derived from p35, the Cdk5 neuronal activator, specifically inhibits Cdk5-p25 hyperactivity and tau hyperphosphorylation. *J. Biol. Chem.* 285 (2010) 34202–34212.

[8] M. Kolarova, F. García-Sierra, A. Bartos, J. Ricny and D. Ripova, Structure and pathology of tau protein in Alzheimer disease, *Int. J. Alzheimers Dis.* 2012 (2012) 731526.

[9] J. Herzog, S.M. Ehrlich, L. Pfitzer, J. Liebl, T. Frohlich, G.J. Arnold, W. Mikulits, C. Haider, A.M. Vollmar and S. Zahler, Cyclin-dependent kinase 5 stabilizes hypoxia-inducible factor-1alpha: a novel approach for inhibiting angiogenesis in hepatocellular carcinoma. *Oncotarget.* 7 (2016) 27108-27121.

*[10]* J.H. Christopher, A.S. Mark and B.C. Christopher, Discovery and SAR of 2-aminothiazole inhibitors of CDK5/p25 as a potential treatment for Alzheimer's disease, *Bioorg. Med. Chem. Lett.* 14 (2004) 5521-5525.

[11] M. Tadayon and Z. Garkani-Nejad, *In silico* study combining QSAR, docking and molecular dynamics simulation on 2,4-disubstituted pyridopyrimidine derivatives, *J. Recept. Signal Trans.* 39:2 (2019) 167-174.

[12] H. Safarizadeh and Z. Garkani-Nejad, Molecular docking, molecular dynamics simulations and QSAR studies on some of 2-arylethenylquinoline derivatives for inhibition of Alzheimer's amyloid-beta aggregation: Insight into mechanism of interactions and parameters for design of new inhibitors, *J. Mol. Graph. Model.* 87 (2019) 129-143.

[13] M. Shiradkar, J. Thomas, V. Kanase and R. Dighe, Studying Synergism of Methyl Linked Cyclohexyl Thiophenes with Triazole: Synthesis and Their Cdk5/P25 Inhibition Activity, *Eur. J. Med. Chem.* 46 (2011) 2066-2074.

[14] SYBYL package, Tripos Inc.: St. Louis, MO, USA, (2006), Available online: http://www.tripos.com.

[15] D.M. Chhatbar, U.J. Chaube, V.K. Vyas and H.G. Bhatt, CoMFA, CoMSIA, Topomer CoMFA, HQSAR, molecular docking and molecular dynamics simulations study of triazine morpholino derivatives as mTOR inhibitors for the treatment of breast cancer, *Comp. Bio. Chem.* 80 (2019) 351–363.

[16] A. Khaldan, K.E. khatabi, R. El-Mernissi, A. Ghaleb, R. Hmamouchi, A. Sbai, M. Bouachrine and T. Lakhlifi, 3D-QSAR Modeling and Molecular Docking Studies of novel triazoles-quinine derivatives as antimalarial agents, *J. Mater. Environ. Sci.* 11 (2020) 429-443.

[17] J. Sun, S. Cai, N. Yan and H. Mei, Docking and 3D-QSAR studies of influenza neuraminidase inhibitors using three-dimensional holographic vector of atomic interaction field analysis, *Eur. J. Med. Chem.* 45 (2010) 1008–1014.

[18] J. Zhu, Q. Yu, Y. Cai, Y. Chen, H. Liu, W. Liang and J. Jin, Theoretical Exploring Selective-Binding Mechanisms of JAK3 by 3D-QSAR, Molecular Dynamics Simulation and Free Energy Calculation. *Front. Mol. Biosci.* 7 (2020) 1-12.

[19] Z.Q. Yang and P.H. Sun, 3D-QSAR Study of Potent Inhibitors of Phosphodiesterase-4 Using a CoMFA Approach, *Int. J. Mol. Sci.* 8 (2007) 714–722.

[20] Q.L. Song, P.H. Sun and W.M. Chen, Exploring 3D-QSAR for Ketolide Derivatives as Antibacterial Agents Using CoMFA and CoMSIA, *Lett. Drug Des. Discov.* 7 (2010) 149–159.

# مطالعه مدل‌سازی مولکولی تری آزولیل تیوفن‌ها به عنوان بازدارنده‌های CDK5/P25 با استفاده از روش‌های 3D-QSAR و داکینگ مولکولی

**زهرا گرکانی نژاد[1,*]، ابوذر قنبری[2]**

۱. گروه شیمی، دانشکده علوم، دانشگاه شهید باهنر کرمان، کرمان، ایران

۲. گروه عملیات و نگهداری مپنا، خیابان مصدق شمالی، بلوار میرداماد، تهران، ایران

**تاریخ دریافت: ۱۸ دیماه ۱۳۹۹ تاریخ پذیرش: ۲۱ اسفند ۱۳۹۹**

**چکیده**

روش های ارتباط کمی ساختار–فعالیت سه بعدی(3D-QSAR) برای طراحی دارو بر پایه لیگاند بسیار مفید می باشند. یک سری جدید از تری آزولیل تیوفن ها به عنوان بازدارنده های CDK5/P25 انتخاب شده اند و با استفاده از روش‌های 3D-QSAR (CoMFA) و (CoMSIA) مدل‌سازی انجام شده است. برای مدل های بهینه (CoMFA) و (CoMSIA) به ترتیب ضرایب همبستگی ارزیابی متقاطع ۰,۵۳۹ (q2) r2cv و ۰,۵۹۸ و ضرایب همبستگی (r2) ۰,۹۸۰ و ۰,۹۶۷ بدست آمده است. از یک سری آموزشی شامل ۸۸ مولکول و یک سری پیش بینی شامل ۲۴ مولکول برای بدست آوردن مدل ها استفاده شده است. ضرایب همبستگی مدل ها برای سری پیش بینی (r2pred) به ترتیب ۰,۹۶۸ و ۰,۹۴۵ بدست آمده است. از داکینگ مولکولی برای بررسی اتصال لیگاند و گیرنده استفاده شده است. نتایج حاصل از داکینگ مولکولی می تواند در طراحی بازدارنده های جدید مفید باشد.

**واژه‌های کلیدی**

3D-QSAR؛ داکینگ مولکولی؛ بازدارنده های CDK5/P25؛ تری آزولیل تیوفن؛ بیماری آلزایمر.