iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** V **Month of publication:** May 2020

DOI: http://doi.org/10.22214/ijraset.2020.5134

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

Ligands Based Drug Design for Covid 19 - A Multi-Faceted Approach using Ligand Design, Molecular Docking and Binding Probability Calculation

Mukundan Satyanarayanan[1], Dr. V. Radhika[2]
[1] *Concepts 2 Consumer, Chennai, India*
[2]*Department of Mathematics, Easwari Engineering College, Chennai, India*

Abstract: ***Covid 19 has been the most devastating pandemic of the recent years, affecting 3 million people in about 210 Countries. Entire World is working on inventing a drug for this pandemic. As a major breakthrough, the crystal structure of Covid 19 main protease $3CL^{Pro}$ or M^{Pro} , which plays a major role in mediating the replication and transcription of the virus , was derived by Jin, Z., Du, X., Xu, Y. et al., paving way for the drug design. The crystal structure of M^{Pro} with a computer aided design inhibitor N (6LU7) has been used as a potential target for drug design in this work.***

Three possible binding sites were identified for $3CL^{Pro}$ or M^{Pro} using DEEPSITE, a protein binding pocket predictor. Complimentary Ligand shapes were generated for the SARS CoV2 main protease M^{Pro} , making use of LIGANN, a structure based de novo drug design tool. They are purely structure based designs and do not have any previous history of synthesis or usage.

Molecular docking of the new ligands with the target protein 6LU7 was done using iGEMdock. The binding free energy values were calculated. 10 best ligand designs , for each binding site, based on lowest free energy requirement for have been selected for further study.

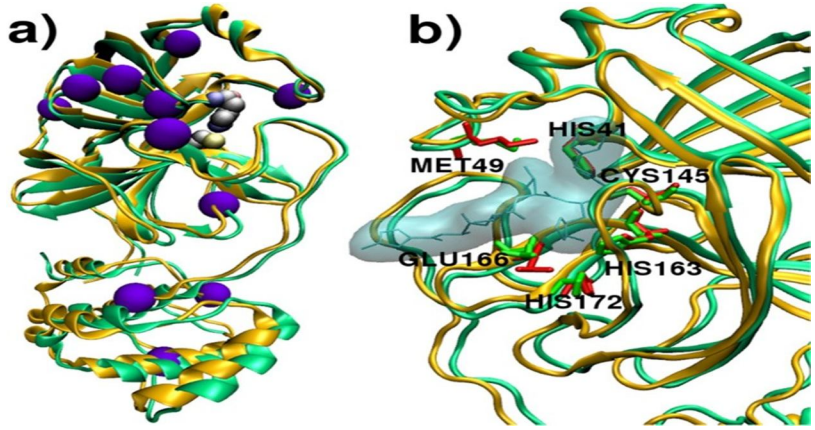
The binding probability for the 10 ligands were calculated using BINDSCOPE, a structure based protein ligand binding predictor. The identical structures for these ligands were identified using Drug bank database. The results were verified with Tanimoto Coefficient calculation.

Based on various parameters like free binding energy, binding probability, structural identity and Tanimoto coefficient, top 5 ligand structures have been selected as potential leads for drug discovery.

Keywords: Covid -19, SARS CoV2, Drug design, Natural Ligands, M^{Pro} , 3CL Main Protease, 6LU7, binding free energy, binding probability, Tanimoto coefficient

I. INTRODUCTION

Covid 19 has been the most devastating pandemic of the recent years, affecting 210 Countries so far. Close to 3 million people have been affected till date. The death toll is close to 0.2 million. This disease is caused due to a novel corona virus named SARS CoV2. Entire World is working on inventing a drug for this pandemic. No drug has proceeded till the commercial production stage, though a few drugs have entered into the clinical trial stage.



ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

II. SARS Co v2 INFECTION MECHANISM

Majority of the Coronaviridae genome encodes pp1a and through ribosomal frameshifting during translation, pp1ab. These polyproteins are cleaved and transformed into mature non-structural proteins (NSPs) by 2 proteases $3CL^{Pro}$ (3CL like Protease) and $PL^{Pro.}$ (Papain like protease) , encoded by open reading frame 1(3). The NSPs in turn play a fundamental role in the transcription/replication during infection [1]

$3CL^{Pro}$ or M^{Pro} is a catalytically active dimer. Cleavage by M^{Pro} occurs at the Glutamine residue in the P1 position of the substrate through the protease CYS-HIS dyad. Cysteine thiol functions as the nucleophile in the proteolytic process.

As a major breakthrough, the crystal structure of SARS CoV2 main protease $3CL^{Pro}$ (3CL like Protease) or M^{Pro} , which plays a major role in mediating the replication and transcription of the virus , was derived by Jin, Z., Du, X., Xu, Y. *et al.* [2], paving way for the drug design, considering this enzyme as a potential target molecule. The crystal structure of M^{Pro} with a computer aided design inhibitor N (6LU7) has been used as a potential target for drug design in this work.

III. RESEARCH METHODOLOGY

The possible binding sites of 3CLPro or M^{Pro} were identified using Deep Site protein binding site predictor [3]. The binding sites identified are mentioned in Table 1

Binding site	scores	centres
0	0.996279	[-34.87799835 16.13400078 52.64800072]
1	0.996786	[-16.87799835 34.13400078 58.64800072]
2	0.99318	[-10.87799835 20.13400078 68.64800072]

Complimentary Ligand shapes were generated for the Covid 19 main protease M^{Pro} , for binding site all the three binding sites making use of LIGANN [4], a structure based de novo drug design tool, based on generative neural-networks. Site 2 (-10.88, 20.13, 68.64) is considered as the most appropriate site for binding (presence of CIS-HYS diad) based on the study conducted by Marina Macchiagodena, et al (1).

(a) SARS-CoV2 (orange) and SARS-CoV (green) main proteases. Violet spheres correspond to the alpha carbons of the 12 differing residues. Grey spheres indicate the CYS-HIS dyad (b) binding pocket with the main residues in bond representation (green and red for SARS-CoV2 and SARS-CoV, respectively). The shaded region marks the binding site for the substrate.

The tool generated 88, 94 & 93 ligand models for binding sites 0,1,& 2 respectively. They are purely structure based designs and are not from the list of ligands which are already being used or have been discovered previously. In spite of this, a few of them closely resemble some of the ligands which have already been discovered.

Molecular docking of the new ligands with the target protein 6LU7 was done using the graphical automatic drug design system for docking, iGEMdock [5], version 2.1. The application suggested 3 binding centres for the ligands. A validation trial was done for all the three centres using the first 5 ligands of the binding site 0 (-10.88, 20.13, 68.64). Binding centre PJE depicted the lowest docking fitness score for all the 5 ligands. The validation results are shown in Table 2. This was selected as the binding centre to measure the docking fitness score for all the 3 sets of ligands. The binding site radius was fixed at 10.0 Å. The following parameters were used for testing. Population Size: 150, Generations: 70, No of solutions: 2

The docking fitness was measured for all the 3 sets of new ligands. Out of this, the top 10 ligands, with lowest docking fitness score were selected for analysis from each of the binding site. The docking fitness score for all the three sets are shown in Table 3, 4 & 5.

The binding probabilities for these selected ligands with the SARS CO V2 main protease M^{Pro} were calculated using the tool BINDSCOPE [6]. BINDSCOPE is a structure based protein ligand binding predictor. It calculates the binding probability based on the binding pocket and ligand pose. The binding probability was measured at the binding site suggested by DeepSite i.e. 0 (-34.88, 16.13, 52.65), 1(16.88, 34.13, 58.64) & 2 (-10.88, 20.13, 68.64). The results are shown in Table 3, 4 &5.

New Ligands showing lowest binding fitness scores and highest binding probabilities were identified for all the three binding sites. They can function as potential ligand models for drug discovery. The structures of these ligands were checked for similarity using Drug bank database in Open Babel [7] platform. The results were verified with Tanimoto Coefficient calculation [8]. The results are shown in Table 3, 4 & 5.

The research work finally lists out 5 new ligands based binding fitness score, binding probability, molecular similarity and Tanimoto coefficient. The details are listed in Table 6.

Previous research has confirmed that the potential ligands for SARS Co V2 main protease M^{Pro} are all aromatic moieties with rotatable bonds of pseudo linear type [1]. The structural details of the potential 5 ligands have also been verified. The details are provided in the Table 7. The molecular structure of these 5 new ligands and their similar ligands are shown in Table 8.

These molecules can be further studied. Their structure can be slightly modified to improve the above parameters, so that, they can function as potential drugs for Covid 19. These new ligands can efficiently bind with the SARS Co V2 main protease M^{Pro}, inhibiting the formation of NSPs and thereby the replication and transcription during infection.

IV CONCLUSION

Covid 19 has been the most devastating pandemic of the recent years, affecting 210 Countries so far. Close to 3 million people have been affected till date. This disease is caused due to a novel corona virus named SARS CoV2. The crystal structure of SARS CoV2 main protease $3CL^{Pro}$ (3CL like Protease) or M^{Pro}, which plays a major role in mediating the replication and transcription of the virus , was derived by Jin, Z., Du, X., Xu, Y. *et al.* (2), paving way for the drug design, considering this enzyme as a potential target molecule.

The possible binding sites of $3CL^{Pro}$ or M^{Pro} were identified using Deep Site protein binding site predictor (3). Complimentary Ligand shapes were generated for the Covid 19 main protease M^{Pro}, for binding site all the three binding sites making use of LIGANN, a structure based de novo drug design tool, based on generative neural-networks.

Molecular docking of the new ligands with the target protein 6LU7 was done using the graphical automatic drug design system for docking, iGEMdock, version 2.1. The docking fitness was measured for all the 3 sets of new ligands. Out of this, the top 10 ligands, with lowest docking fitness score were selected for analysis from each of the binding site. The binding probabilities for these selected ligands with the SARS CO V2 main protease M^{Pro} were calculated using the tool BINDSCOPE.

New Ligands showing lowest binding fitness scores and highest binding probabilities were identified for all the three binding sites. The structures of these ligands were checked for similarity using Drug bank database in Open Babel platform. The results were verified with Tanimoto Coefficient calculation.

New Ligand molecules showing some resemblance to existing drug molecules like Arlasetone, Ajmaline, Rufinamide, Nilotinib and Deferaserox can be taken up for further drug research. The molecular structure of these 5 new ligands and their similar ligands are shown in Table 8.

REFERENCES

[1] Marina Macchiagodena, Marco Pagliai, and Piero Procacci, Identification of potential binders of the main protease 3CLpro of the COVID-19 via structure-based ligand design and molecular modelling, Chem Phys Lett. 2020 Jul; 750: 137489, Published online 2020 Apr 18. doi: 10.1016/j.cplett.2020.137489

[2] Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., Jiang, R., Yang, X., You, T., Liu, X., Yang, X., Bai, F., Liu, H., Liu, X., Guddat, L.W., Xu, W., Xiao, G., Qin, C., Shi, Z., Jiang, H., Rao, Z., Yang, H, Structure of Mpro from COVID-19 virus and discovery of its inhibitors in PubMed, 2020 Apr 9. Doi: 10.1038/s41586-020-2223-y. [Epub ahead of print]

[3] Jiménez, Doerr, Martinez-Rosell, Rose, De Fabritiis, DeepSite: Protein binding site predictor using 3D-convolutional neural networks, *Bioinformatics*, Volume 33, Issue 19, 01 October 2017, Pages 3036–3042.

[4] Skalic, M., Jiménez Luna, J., Sabbadin, D., & De Fabritiis, Shape-Based Generative Modeling for de-novo Drug Design. Journal of chemical information and modelling, 2019, Volume 59, Issue 3, Pages 1205-1214

[5] Hsu, K., Chen, Y., Lin, S. et al. iGEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis. BMC Bioinformatics 12, S33 (2011). https://doi.org/10.1186/1471-2105-12-S1-S33

[6] Jiménez Luna, J., Skalic, M., Martinez-Rosell, G., & De Fabritiis, G, K DEEP : Protein-ligand absolute binding affinity prediction via 3D-convolutional neural networks, Journal of Chemical Information and Modeling, 2018, Volume 58, Issue 2, Pages 287-296.

[7] O'Boyle, N.M., Banck, M., James, C.A. et al. Open Babel: An open chemical toolbox. J Cheminform 3, 33 (2011). https://doi.org/10.1186/1758-2946-3-33

[8] Bajusz, D., Rácz, A. & Héberger, K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations?. J Cheminform 7, 20 (2015). https://doi.org/10.1186/s13321-015-0069-3

TABLES

Table 2: Validation of best docking site for the ligand 6LU7 (Docking site 0: =-34.88, 16.13, 52.64)

Protein	6LU7	Binding site	0	
S.No	Ligand SMILES	Binding Site centre		
		02J	PJE	010
1	CN(C)c1ccccc1CNc1cc(-c2cccs2)n[nH]1	-50.23	-67.52	-61.2
2	CSc1ccccc1CC(=O)NC(c1nc2ccccc2o1)c1c(C)noc1C	-55.37	-66.79	-56.54
3	CSCC(C(=O)O)N(C(=O)c1nn(-c2ccccc2)c2c1CCC2)c1ccccc1	-48.85	-75.18	-59.48
4	CCC(C)C(C(=O)NS(=O)(=O)C(C)C(F)(F)F)c1ccccc1	-67.34	-94.76	-70.21
5	Cn1cnc2cccc(-c3nc(-c4ccsc4)no3)c21	-63.40	-80.78	-67.41

Table 3: Docking Fitness Score, Binding Probability, Molecular Similarity and Tanimoto coefficient for the top 10 new ligands for the binding site 0(-34.88, 16.13, 52.64)

S.No	SMILES	Binding Affinity	Fitness Score	Similarity	Tanimoto coefficient
22	CN(Cc1ccccc1N1CCCC1=O)C(=O)C1(S(=O)(=O)c2ccccc2)CCCC1	0.0377	-96.88	Irbesatran	0.26
24	Cc1cccc(-c2nc(-c3ccsc3)no2)c1F	0.9283	-103.34	Rufinamide	0.17
25	CC(c1ccccc1)n1cccc1C=NNc1cccc(F)c1	0.0833	-99.12	Antipyrine	0.29
32	Cc1cccc(NCc2nc(-c3ccccc3)c[nH]2)c1Cl	0.3861	-102.96	Thiabendazole	0.22
37	Cc1cccc(NC(=S)Nc2ccccc2-n2ccnc2)c1	0.9359	-102.18	Azathioprine	0.2
42	CN(Cc1nc(N)c2ccccc2n1)Cc1ccccc1F	0.1185	-99.74	Bretylium	0.28
45	Cc1cccc(C(C)C)c1NC(=O)COC(=O)c1cc2ccccc2[nH]1	0.4852	-102.9	Nitazoxanide	0.16
54	Cn1c(-c2cnc3ccccc3n2)nc2c(F)cccc21	0.6428	-106.95	Nilotinib	0.31
77	Cc1cc(Oc2ccc(-c3ccccc3)nn2)ccc1F	0.478	-100.48	Papaverine	0.27
84	O=Cc1sc(-c2ccccc2)nc1-c1ccncc1	0.8465	-97.14	Sulfamethaxazole	0.16

Table 4: Docking Fitness Score, Binding Probability, Molecular Similarity and Tanimoto coefficient for the top 10 new ligands for the binding site 1(-16.88, 34.13, 68.64)

S.No	SMILES	Binding Probability	Binding Free Energy	Similar structure	Tanimoto Coeffcient
1	O=C(c1coc(-c2ccccc2)n1)N1CCC(CO)c2ccccc21	0.0007	-93.08	Alosetron	0.3
3	C#CCOC(=O)c1cc(=O)n(-c2ccc(C)cc2)[nH]1	0.4166	-95.25	Conivapton	0.2
5	Fc1ccccc1-c1nnc(Sc2cccc(OC3CCCC3)c2)o1	0.106	-105.01	Rabeprazole	0.2
8	O=C(O)c1cn(-c2ccc(Br)cc2)nn1	0.8213	-102.48	Deferasirox	0.27
9	CCCC(=O)OCc1n[nH]c(-c2ccc(OC)cc2)n1	0.2641	-90.3	Oxaprozin	0.23
22	COc1cccc2oc(=O)n(C)c12	0.9831	-94.85	Tolnaftate	0.21
24	Cc1ccccc1N1CC(C(=O)O)C(F)(F)C1	0.9937	-90.34	Ajmaline	0.38
34	CC(C)NC(=O)COC(=O)c1ccc(CCC#N)cc1	0.0403	-96.8	Benzonatate	0.44
44	CCC(=O)OCc1nnc(-c2cc(C)cc(C)c2)[nH]1	0.5155	-91.39	Oxaprozin	0.24
88	CN(C)C(=O)c1cccc(-c2noc(Cc3ccccc3F)n2)c1	0.3889	-94.3	Estazolam	0.21

Table 5: Docking Fitness Score, Binding Probability, Molecular Similarity and Tanimoto coefficient for the top 10 new ligands for the binding site 2(-16.88, 34.13, 68.64)

S.No	SMILES	Binding Probability	Binding free energy	Similar Molecule	Tanimoto Coefficient
14	Cc1cccc(CNc2nc(N)nc(C(F)(F)F)n2)c1	0.5852	-97.54	Edrophonium	0.13
20	C#CCn1cc(-c2nc(N)nc(C)c2Br)nn1	0.06	-100.42	Rufinamide	0.24
51	Cc1cc(N(C)C)ccc1C(=O)Nc1nc2ccccc2[nH]1	0.0001	-86.49	Omeprazole	0.25
55	COCC(=O)N1CCc2c([nH]c3c(F)cccc23)C1	0.9808	-89.76	Alosetron	0.52
58	Cc1cccc(C)c1OCc1nc(C(N)=O)no1	0.0371	-86.92	Nitazoxanide	0.22
61	Cc1cccc(OCCNc2ccc(O)nn2)c1	0.0004	-86.29	Erlotinib	0.35
63	c1cc(-c2cc(CN3CC4CCC(C3)N4)no2)no1	0.0001	-87.09	Alizapride	0.23
90	CCC(C)(C)c1cc(Nc2cc(C)on2)[nH]n1	0.0001	-86.53	Nitrofural	0.12
91	Cc1[nH]nc(NS(=O)(=O)c2ccc3ccccc3c2)c1C(N)=O	0.6863	-91.11	Sulfamoxole	0.35
93	CC(=NNc1ccc(OC(F)F)cn1)c1cccc(O)c1	0.0576	-93.9	Riluzole	0.32

Table 6: Top 5 ligands suitable for drug design based on docking fitness score, binding probability and Tanimoto coefficient.

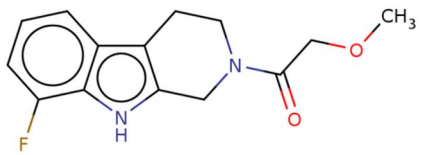
S.No	Binding site	SMILES	Binding Probability	Fitness Score	Similar Molecule	Tanimotto Coefficient
1	2	COCC(=O)N1CCc2c([nH]c3c(F)cccc23)C1	0.9808	-89.76	Alosetron	0.52
2	1	Cc1ccccc1N1CC(C(=O)O)C(F)(F)C1	0.9937	-90.34	Ajmaline	0.38
3	1	O=C(O)c1cn(-c2ccc(Br)cc2)nn1	0.8213	-102.48	Deferasirox	0.27
4	0	Cn1c(-c2cnc3ccccc3n2)nc2c(F)cccc21	0.6428	-106.95	Nilotinib	0.31
5	0	Cc1cccc(-c2nc(-c3ccsc3)no2)c1F	0.9283	-103.34	Rufinamide	0.17

Table 7: Molecular weight, rotatable bonds and aromatic rings in new ligands selected for drug design

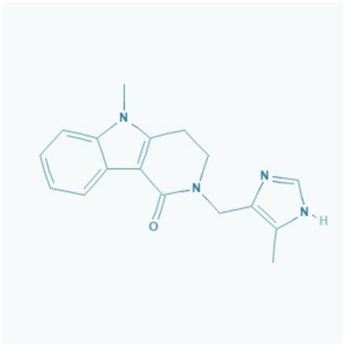
S.No	SMILES	Aromatic rings	Molecular Weight
1	COCC(=O)N1CCc2c([nH]c3c(F)cccc23)C1	3	262.28
2	Cc1ccccc1N1CC(C(=O)O)C(F)(F)C1	2	241.24
3	O=C(O)c1cn(-c2ccc(Br)cc2)nn1	2	268.07
4	Cn1c(-c2cnc3ccccc3n2)nc2c(F)cccc21	4	278.29
5	Cc1cccc(-c2nc(-c3ccsc3)no2)c1F	3	260.29

Table 8: Molecular structure and Structure of similar molecule for the 5 new ligands

Ligand 1:



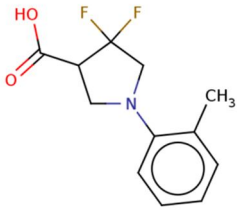
Alosetron

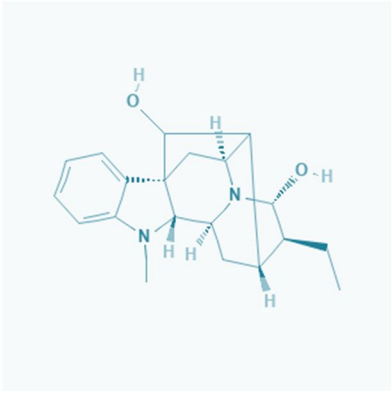


Ligand Name: 1-(8-fluoro-1, 3, 4, 9-tetrahydropyrido [3, 4-b] indol-2-yl)-2-methoxy-ethanone

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

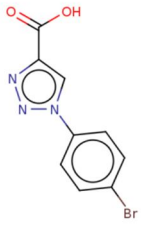
Ligand 2 Ajmaline

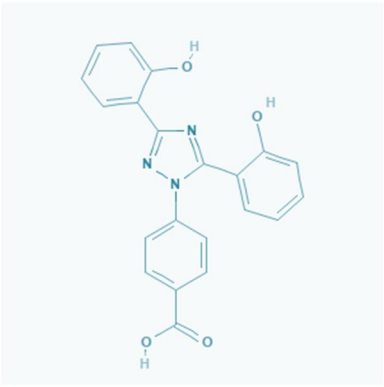




Ligand Name: 4, 4-difluoro-1-(o-tolyl) pyrrolidine-3-carboxylic acid

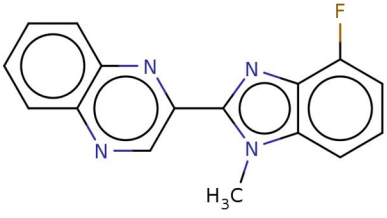
Ligand 3: Deferasirox

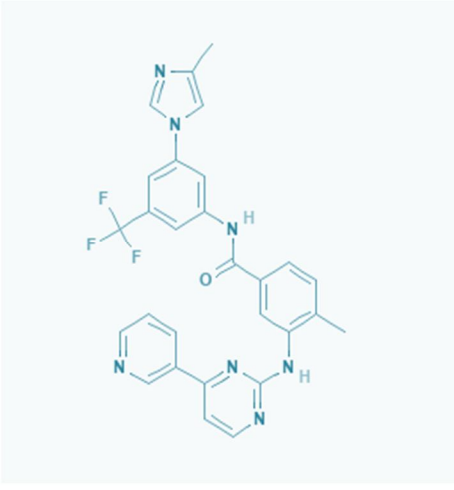




Ligand Name: 1-(4-bromophenyl) triazole-4-carboxylic acid

Ligand 4: Nilotinib

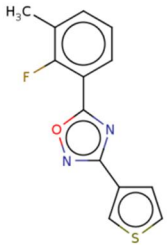




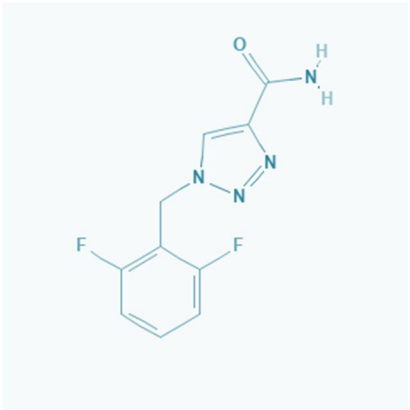
Ligand Name: 2-(4-fluoro-1-methyl-benzimidazol-2-yl)quinoxaline

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

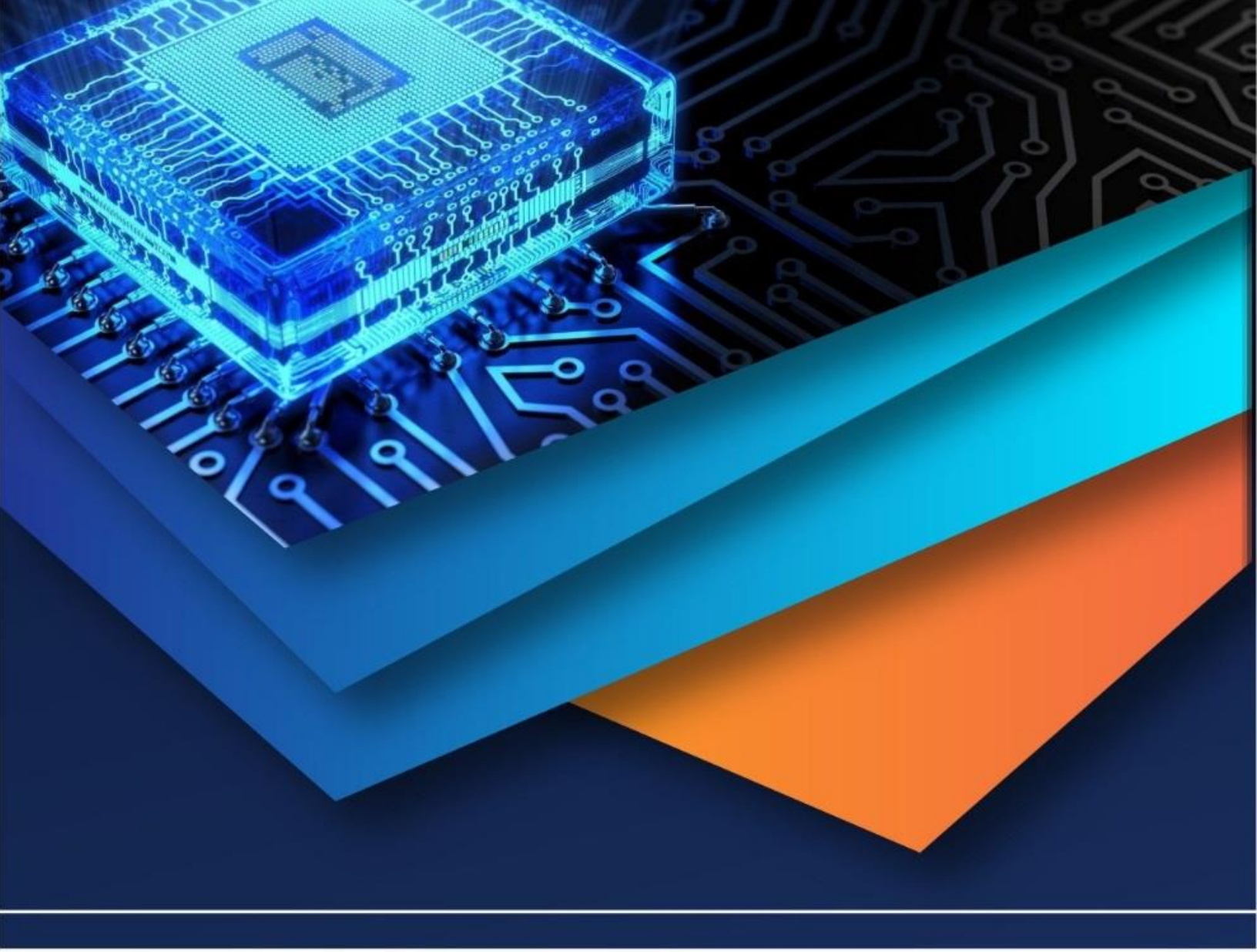
Ligand 5



Rufinamide



Ligand name: 5-(2-fluoro-3-methyl-phenyl)-3-(3-thienyl)-1,2,4-oxadiazole



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY