INTERNATIONAL RESEARCH JOURNAL OF PHARMACY
www.irjponline.com ISSN 2230 – 8407
Research Article

FINDING NEW INHIBITORS FOR EML4-ALK FUSION PROTEIN: A COMPUTATIONAL APPROACH

Kesavan Sabitha* and Ramshankar Vijayalakshmi
Department of Molecular Oncology, Cancer Institute (WIA), Adyar, Chennai, India

Article Received on: 01/02/12 Revised on: 06/03/12 Approved for publication: 22/03/12

*Dr K Sabitha, Email: rk_sabitha@gmail.com

ABSTRACT
The fusion between echinoderm microtubule-associated protein-like 4 (EML4) and anaplastic lymphoma kinase (ALK) has recently been identified in a subset of non-small cell lung cancers (NSCLC). PF-02341066 (crizotinib) is an orally bioavailable ALK inhibitor currently under clinical development. PF-02341066 in EML4-ALK NSCLC was designed for patients not eligible for the phase III trial or patients randomized to chemotherapy who subsequently developed progressive disease. PF-02341066 is a dual inhibitor of mesenchymal epithelial transition growth factor (c-met) and anaplastic lymphoma kinase translocation gene and caused tumour shrinkage in 52% of patients in a phase I study. However, some studies also show denovo mutations within the kinase domain of EML4-ALK that confers resistance to multiple ALK inhibitors. Hence development of new inhibitors with better binding affinities towards the EML4-ALK is the need of the hour for subsequent clinical validation.
Computational (virtual) screening of drug-like compounds against the protein targets like EML4-ALK, might help to identify specific lead inhibitors more efficiently. The Protein- Ligand interaction plays a significant role in structure based drug designing.
In the current study, we have considered EML4-ALK, a fusion protein involved in NSCLCs as a receptor and NCI subset Ligands as drugs. The receptor was docked to the NCI database of drugs and a docking score was calculated using GLIDE docking software. Based on the docking score, we choose the best drugs and analyzed its ADME properties using Qikprop tool. The results of this analysis show some novel compounds that can be further evaluated as EML4-ALK inhibitors in experimental NSCLC cell lines. The study further supports the application of computer-aided techniques to the discovery of novel and specific drug for EML4-ALK fusion protein.
Key Words: EML4-Alk, Glide, QikProp, I-Tasser, ADME, Docking

INTRODUCTION

EML4-ALK fusion gene seems to be unique to NSCLC and its multiple variants have been identified in lung cancer [1]. EML4–ALK is a fusion type protein tyrosine kinase that is present in 4-5% of NSCLCs and is generated due to a small inversion within the short arm of Chromosome 2 (2p21 and 2p23) separated by 12 Mb, they are normally present in opposite orientations[2]. EML4-ALK fusion protein is formed due to disruption at a point ~ 3.6 kb upstream of exon 13 of EML4 and is inverted to connect to a position 297 bp upstream of exon 21 of ALK, yielding EML4-ALK variant 1. Although the fusions contain variable truncation points of EML4 occurring at exons 2, 6, 13, 14, 15 and 18 and 20, the ALK fusion in all of them starts at a portion encoded by exon 20 of the kinase gene [3]. To date, all the EML4-ALK fusions tested biologically show gain of function properties [4,5].
Current study shows binding affinity of active inhibitors for the EML4-ALK, downloaded from ligand database to characterize interaction features using GLIDE (Schrodinger) and to interpret inhibition mechanism of small molecules for this fusion protein involved in lung cancer. The compounds which showed good binding affinity and specificity were analyzed for their absorption, distribution, metabolism and excretion profile (ADME). The aim of the present study was to investigate the inhibitory activity of the compounds using molecular docking studies and to analyze the ADME properties of these compounds for drug like candidates by using the Qikprop (a Schrodinger software) to identify specific drug molecules for EML4-ALK.

MATERIALS & METHODS

Modelling of EML4-ALK using I-TASSER

The amino acid sequence of EML4-ALK from genbank (BAF73611.1) was selected as target. I-TASSER, an online server for protein structure and function predictions, identifies the global or local threading alignments using either MUSTER –single threading server [6 -8] or LOMETS – meta-threading server [9]. This was done to thread the query sequence of EML 4-ALK through a representative PDB structure library. The structures of lowest energy were selected and then refined by a fragment-guided molecular dynamic procedure, with the purpose of optimizing the hydrogen –binding network and removing steric clashes.

Protein Preparation

EML4-ALK model obtained by I-TASSAR server was subjected to Protein Preparation Wizard workflow implemented in the Schrodinger package. Hydrogens were added and subsequently minimization of the structure was done using OPLS_2005 force field. EML4-ALK model was optimized and His residues were protonated. Hydroxyl groups of Asn and Gln residues are also optimized.

Ligand preparation

Drug-likeness NCI subset Ligands were downloaded from Ligand database (http://bicpu.edu.in/ligandinfo.htm). The LigPreP module of Maestro (Schrödinger) was used to correct for bond orders. All ligands were subjected to full minimization with OPLS (optimized potential for liquid simulations) force field to correct its bond length and bond angles. Using 'ionizer' option all ligands were assigned by appropriate ionization state. These ligands were used for docking studies with the modeled EML4-ALK.

Molecular Docking

A molecular docking study was carried out using Glide, [10] considering the ligands as flexible but treating the receptor as a rigid structure. A docking grid was centered on EML4-ALK, in its junctional region (amino acids 485-505), with a dimension sufficient to accommodate ligands with a length equal to $15A^0$ .Glide docking algorithm helped perform a series of hierarchical searches for locations of possible ligand affinity within the binding site of EML4-ALK. The stability and binding affinity of ligands docked with receptor (EML4-ALK) was due to hydrogen bond and van der Waals interaction. The glide score, glide energy value, H-bonds and

Vander Waals contacts (good, bad and ugly) to the EML4-ALK were visualized and its binding sites were analyzed.

ADME predictions by QikProp 3.2

QikProp is a quick, accurate, easy-to-use absorption, distribution, metabolism, and excretion (ADME) prediction program designed by Professor William L. Jorgensen [11, 12].

ADME properties for the compound were obtained using QikProp.

RESULTS

Computational modeling of the structure of EML4-ALK using I-TASSER was done. I-TASSER is a hierarchical protein structure modeling approach based on the multiple threading alignments and an iterative implementation of the Threading ASSEmbly Refinement (TASSER) program. Using I-TASSER server, the target sequence was modeled and was used for further studies. Thus modelled EML4-ALK structure was further corrected by protein preparation wizard of Maestro. Optimal protonation states for histidine residues were determined and potentially transposed heavy atoms in arginine, glutamine, and histidine side chains were corrected. A restrained minimization that allowed hydrogen atoms to be freely minimized was done, while allowing for sufficient heavy-atom movement to relax strained bonds and angles. Figure 1 shows the ribbon representation of the modelled EML4-ALK.

A grid generated around junction region of EML4-ALK consisting of amino acids from 485 – 497, specific to EML4-ALK had the prepared ligands docked into this region using standard precision Glide algorithm. The candidate ligands with high docking score and low binding energy are listed in Table 1. The ligands were ranked based on the glide score. The ligand 1 formed hydrogen bonds with Gly 492, Tyr 498 , Gln509 and His 502 and formed covalent interaction with Pro489, with a docking score of -7.89, its binding energy was found to be low (Figure 2).

Deamino DPN was another ligand with very low glide energy but its docking score was less when compared to other compounds. Compounds ranked within top 68 were taken for further studies such as ADME properties. To provide initial information about ADME properties of these compounds, Qikprop module (Schrodinger 2010 software) was used.

We analyzed 44 physically significant descriptors and pharmaceutically relevant properties of docked compounds. Among these properties, molecular weight, H bond donors, H bond acceptors, logP (octonol/water) , logP MDCK , log Kp(skin permeability), humoral absorption and their position was according to Lipkinski's rule of 5. In accordance with Lipkinski's rule 5, the molecular weights of compounds should be < 500 Daltons with <5 hydrogen bond donors, <10 hydrogen bond acceptors and logP of < 5. Out of 67 compounds only 22 compounds are found to be well within acceptance range of the Lipinski's rule for drug like molecules. Compound 1 showed a high docking score with low energy but its ADME properties did not show drug like characteristics. For compounds numbered 4, 8, 14, 17- 20, 23, 24, 31, 35 – 37, 52, 55 – 58, 63, 65, 67 the partition coefficient (QPlogPo/w), water solubility (QPlogs), cell permeability (QPPCaco) and human oral absorption were found to be within the acceptable range defined for human use. (Table 2).

Figure 3 shows the structural details of Compound 4 that formed hydrogen bonds with amino acids such as Gly 492, Gly 494, His 502 and Asn 579 of EML4-ALK fusion protein (Figure 2 A). Docking score of compound 4 was around -7.085256 and its glide energy was - 44.121472. The docking score of compound 8 was -6.919415 and it formed hydrogen bonds with Gln 503 and Ile 622. Further , ADME analysis of these compounds showed optimal values for permeability through MDCK cells (QPlog MDCK), QikProp predicted log IC50 value for blockage of K+ channels (QPlogHERG), QikProp predicted gut-blood barrier (QPPCaco) and violations of the Lipinski's rule of five (LROF) were within acceptable range.

Based on overall analysis we conclude that the compounds 4, 8, 14, 17 -20, 23, 24, 31, 35 – 37, 53, 56 – 59, 64, 66, 68 are the most potent compounds that can be used as a specific drug for EML4-alk fusion protein. ADME properties of these compounds were found under acceptable range. So, these drugs have to be further validated in cancer cell lines harbouring this fusion gene and evaluated as a potential inhibitor to help identify therapeutic targets of EML4-ALK fusion protein by various experimental studies.

DISCUSSION

Computational Biology and bioinformatics have the potential not only for speeding up the drug discovery process, reducing the costs, but also of changing the way drugs are designed. *In silico* drug design (DD) helps to facilitate and speedup the drug designing process, which involves variety of methods to identify novel compounds. One such method involves docking of the drug molecule with the receptor (target). Docking studies procedures basically aim to identify the correct conformation of ligands in the binding pocket of a protein and to predict the affinity between the ligand and the protein.

EML4 –ALK is a fusion type protein tyrosine kinase present in 4-5% of NSCLCs but forms a distinct entity for targeted therapy for these patients. Recent evidences also show several new acquired new mutations in patients harbouring this fusion protein, hence newer inhibitors are also necessary apart from the established Crizotinib, which has shown promise in this category of patients.

In the current study, we have shown compounds downloaded from ligand database after characterization of the interaction features, having binding affinity for the EML4-ALK, and have attempted to interpret inhibition mechanism of small molecules for this targetable fusion protein involved in lung cancer.

The purpose of scoring procedure is the identification of the correct binding pose by its lowest energy value and the ranking of protein ligand complexes according to their binding affinities. Glide scoring function can be enumerated by the displacement of waters by the ligand from hydrophobic regions of the protein active site, protein –ligand hydrogen bonding interactions as well as other strong electrostatic interactions such as salt bridges, desolvation effects, entropic effects due to the restriction on binding of the motion of flexible protein or ligand groups and also interaction of the ligand with metal ions. Such chemical entities are likely to be able to enter higher phases of the drug development process. In vitro approaches are now widely used to investigate the ADME properties of new chemical entities and, more recently, computational (in silico) modelling has been investigated as a tool to optimize selection of the most suitable candidates for drug development. ADME properties of top 68 docked molecules obtained by Glide docking were checked, using Qikprop module of Schrodinger.

It helps in analyzing the pharmacokinetics and pharmacodynamics properties of molecules by accessing the

drug like properties. In addition to predicting drug like properties, QikProp provides ranges for comparing a particular molecule's properties with those of 95% of known drugs. QikProp also flags 30 types of reactive functional groups that may cause false positives in high-throughput screening (HTS) assays .

Our results establish that the compounds 4, 8, 14, 17 -20, 23, 24, 31, 35 – 37, 53, 56 – 59, 64, 66, 68 can develop as a promising lead in the design of EML4-ALK inhibitors. Because of large number compounds are available, a combination of selecting candidates by docking and biochemical testing might prove suitable for finding highly potent EML4-ALK inhibitors.

CONCLUSION

Crystal structure of EML4-ALK should shed light on identity of new class of inhibitors but currently computational tools and insilico approaches for drug design are invaluable to provide a sense of direction and utility of new compounds that can be identified. This approach will pave way for development of next generation ALK inhibitors to be tried, evaluated, validated and used in targeted therapy for patients harbouring EML4-ALK fusion protein.

REFERENCES

1) Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 2007;448 (7153):561-6.

2) Martelli MP, Sozzi G, Hernandez L, Pettirossi V, Navarro A, Conte D, et al. EML4-ALK rearrangement in non-small cell lung cancer and non-tumor lung tissues. Am J Pathol 2009;174 (2):661-70.

3) Sasaki T, Rodig SJ, Chirieac LR, Jänne PA**.** The biology and treatment of EML4-ALK non-small cell lung cancer. Eur J Cancer. 2010 Jul;46(10):1773-80. Epub 2010 Apr 24.

4) Shaw AT, Yeap BY, Mino-Kenudson M, Digumarthy SR, Costa DB, Heist RS,et al. Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. J Clin Oncol. 2009 Sep 10;27(26):4247-53. Epub 2009 Aug 10.

5) Li Y, Ye X, Liu J, Zha J, Pei L. Evaluation of EML4-ALK fusion proteins in non-small cell lung cancer using small molecule inhibitors. Neoplasia. 2011 Jan;13(1):1-11.

6) Ambrish Roy, Alper Kucukural, Yang Zhang. I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols 2009;5: 725-738.

7) Ambrish Roy, Dong Xu, Jonathan Poisson, Yang Zhang .A Protocol for Computer-Based Protein Structure and Function Prediction. Journal of Visualized Experiments 2011; 57: e3259.

8) Yang Zhang. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 2008; 9: 40.

9) S. Wu, Y. Zhang. *LOMETS:* A local meta-threading-server for protein structure prediction. Nucleic Acids Research 2007;35: 3375-3382.

10) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L **.** "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening," J. Med. Chem. 2004; *47*, 1750–1759.

11) Jorgensen, W.L.; Duffy, E.M. . "Prediction of drug solubility from structure," Adv. Drug Deliv. Rev 2002; 54, 355-366

12) Jorgensen, W. L.; Duffy, E. M., "Prediction of Drug Solubility from Monte Carlo Simulations," .Bioorg. Med. Chem. Lett., 2000 ; 10, 1155-8.

Table 1 : Docking Score of Compounds

S No	WEIGHT	Name	Docking Score	Glide energy
1	738.78	2,3-dihydroxy-4,6-bis((2-hydroxybenzylidene) amino) cyclohexyl 2,6-dideoxy-2,6-bis((2-hydroxybenzylidene)amino)hexopyranoside	-7.891495	-65.1727
2	570.51	N-(2-((4-(((2-amino-4-hydroxy-6-pteridinyl)methyl)amino) benzoyl)amino)-4-carboxybutanoyl)glutamic acid	-7.599904	-54.3829
3	312.28	4-(((2-amino-4-hydroxy-6-pteridinyl)methyl)amino)benzoic acid	-7.318057	-40.7673
4	458.59	11,17-dihydroxy-3,20-dioxopregn-4-en-21-yl cyclopentanecarboxylate	-7.085256	-44.1215
5	327.34	N-(2-(1,3-dimethyl-2,6-dioxo-2,3,6,9-tetrahydro-1H-purin-8-yl) ethyl)benzamide	-7.0482	-42.8381
6	270.29	8-benzyl-1,3-dimethyl-3,9-dihydro-1H-purine-2,6-dione	-7.000363	-39.4029
7	678.59	1,3,4,6-tetra-O-acetylhex-2-ulofuranosyl 2,3,4,6-tetra-O-acetyl hexopyranoside	-6.965	-63.9081
8	178.19	9-ethyl-6-hydrazino-9H-purine	-6.919415	-31.3074
9	722.64	1,2,3,5,6-penta-O-acetyl-4-O-(2,3,4,6-tetra-O-acetylhexopyranosyl) hexitol	-6.904637	-63.1708
10	354.36	1,4-di(1H-benzimidazol-2-yl)-1,2,3,4-butanetetrol	-6.841174	-51.8025
11	521.48	2-benzyl-6-(((2-((3-benzyl-5-chloro-2-hydroxybenzyl)amino)ethyl) amino)methyl)-4-chlorophenol	-6.71109	-53.272
12	327.21	1H-imidazol-4-ylmethanol compound with 2,4,6-tris(hydroxy(oxido) amino)phenol (1:1)	-6.690823	-32.9516
13	348.27	2,4,6-tris(hydroxy(oxido)amino)phenol compound with 6,7-dihydro-5H-cyclopenta[b]pyridine (1:1)	-6.690823	-32.9516
14	230.22	1-naphthyl aminocarbonylcarbamate	-6.654698	-35.4709
15	183.12	2-hydroxy-5-(hydroxy(oxido)amino)benzoic acid	-6.61657	-25.4203
16	363.22	Sodium GMP	-6.602558	-43.7396
17	460.47	N,N'-bis((1,3-dimethyl-2,6-dioxo-2,3,6,9-tetrahydro-1H-purin-8-yl)methyl)thiourea	-6.601636	-55.7831
18	456.53	11,17-dihydroxy-3,20-dioxopregn-4-en-21-yl 2-furoate	-6.598462	-43.7102
19	544.65	2-allyl-N~1~,N~3~-bis(4-amino-2-methyl-6-quinolinyl)-2-benzylmalonamide	-6.564677	-57.2027
20	203.19	(8-quinolinyloxy)acetic acid	-6.55234	-29.1265
21	232.71	4-(3-chlorobenzyl)-2-methylphenol	-6.551623	-29.2932
22	294.31	Aminonucleoside	-6.513542	-39.8795
23	313.31	N-((1,3-dimethyl-2,6-dioxo-2,3,6,9-tetrahydro-1H-purin-8-yl)methyl)benzamide	-6.469549	-38.6459
24	224.25	3-tert-butyl-2-hydroxy-5-methoxybenzoic acid	-6.4613	-26.9308
25	344.4	N-(4-(dimethylamino)-2-methylphenyl)-3,4,5-trimethoxybenzamide	-6.448657	-38.8859

26	767.54	Coenzyme A	-6.426437	-58.2554
27	246.26	4,4'-dimethyl[1,1'-biphenyl]-2,2',5,5'-tetrol	-6.396548	-30.3469
28	213.22	2-(methylsulfonyl)-9H-purin-6-amine	-6.389219	-33.1945
29	265.27	3-(1,3,7-trimethyl-2,6-dioxo-2,3,6,7-tetrahydro-1H-purin-8-yl)propanamide	-6.37938	-34.7123
30	215.23	3-((methylsulfonyl)amino)benzoic acid	-6.376619	-26.8443
31	290.27	5a,11a-dihydro-5,6,11,12-naphthacenetetrone	-6.370437	-34.2925
32	168.19	3H-naphtho[1,2-d]imidazole	-6.348643	-26.0695
33	308.36	4-amino-N-(5-(2-hydroxyethyl)-4-methyl-2-pyrimidinyl)benzenesulfonamide	-6.339011	-36.8598
34	276.33	8-cyclohexyl-1,3,7-trimethyl-3,7-dihydro-1H-purine-2,6-dione	-6.329084	-32.4666
35	372.5	11-hydroxy-3-oxopregna-4,17-dien-21-yl acetate	-6.325232	-36.6815
36	329.44	4-((4-amino-3-methylphenyl)(4-imino-3-methyl-2,5-cyclohexadien-1-ylidene)methyl)-2-methylphenylamine	-6.321081	-37.2819
37	279.36	10-(1H-imidazol-4-ylmethyl)-10H-phenothiazine	-6.315267	-35.2222
38	194.19	1,3,8-trimethyl-3,9-dihydro-1H-purine-2,6-dione	-6.304098	-26.2965
39	323.39	2-(4-morpholinyl)-1,4-diphenyl-1,4-butanedione	-6.303717	-32.4667
40	439.42	3-((4-aminophenyl)diazenyl)-4,5-dihydroxy-2,7-naphthalene disulfonic acid	-6.297331	-42.7914
41	719.59	3,6-dichloro-2-(3-((4-ethoxyphenyl)imino)-6-(4-ethoxy-2-sulfoanilino)-3H-xanthen-9-yl)benzoic acid	-6.28756	-51.3989
42	286.37	8-hydroxy-12a-methyl-3,4,4a,4b,5,6,10b,11,12,12a-decahydro-2H-naphtho[2,1-f]chromen-2-one	-6.284349	-29.9826
43	180.16	1,3-dimethyl-3,9-dihydro-1H-purine-2,6-dione	-6.272798	-26.4863
44	222.24	8-isopropyl-1,3-dimethyl-3,9-dihydro-1H-purine-2,6-dione	-6.270316	-28.8652
45	241.25	N-(3-methoxyphenyl)-9H-purin-6-amine	-6.267433	-37.6168
46	546.01	11,17-dihydroxy-3,20-dioxopregn-4-en-21-yl 2-chloro-4-(hydroxy(oxido)amino)benzoate	-6.25706	-43.3715
47	270.24	1,5-diamino-4,8-dihydroxyanthra-9,10-quinone	-6.243334	-34.4153
48	300.31	8-(benzyloxy)-1,3,7-trimethyl-3,7-dihydro-1H-purine-2,6-dione	-6.232216	-36.0595
49	296.32	8-cinnamyl-1,3-dimethyl-3,9-dihydro-1H-purine-2,6-dione	-6.224808	-40.4242
50	413.3	1,3-bis((7-chloro-4-quinolinyl)amino)-2-propanol	-6.224204	-45.0416
51	162.19	5,8-dihydro-1,4-naphthalenediol	-6.214844	-23.8579
52	276.27	4-phenoxathiincarboxylic acid 10,10-dioxide	-6.212579	-32.6681
53	230.26	2-(2,2-dimethylpropanoyl)-1H-indene-1,3(2H)-dione	-6.208426	-27.1288
54	234.32	9-cyclohexyl-9H-purine-6-thiol	-6.208398	-32.0361
55	210.23	3-tert-butyl-2,5-dihydroxybenzoic acid	-6.206497	-26.9308
56	400.39	Theophylline , 8,8'-trimethylenedi-	-6.201116	-45.6937
57	181.22	2-(methylthio)-9H-purin-6-ylamine	-6.19682	-29.2162
58	178.19	N~2~,N~2~-dimethyl-9H-purine-2,6-diamine	-6.183883	-27.1963
59	440.54	(4-((4-(hexanoylamino)phenyl)sulfonyl)anilino)methanesulfonic acid	-6.18351	-42.3094
60	160.17	1,5-naphthalenediol	-6.181651	-22.6973
61	251.26	2,3,4,5,6-pentahydroxycyclohexanone thiosemicarbazone	-6.181139	-41.3279
62	312.38	2-(2-phenoxyethyl)-1-benzothien-3-yl acetate	-6.176936	-35.5759
63	259.21	2-(2,6-dioxotetrahydro-2H-pyran-3-yl)-1H-isoindole-1,3(2H)-dione	-6.172012	-32.1171
64	232.71	4-(2-chlorobenzyl)-2-methylphenol	-6.170044	-25.1865
65	198.17	3-hydroxy-4,5-dimethoxybenzoic acid	-6.16998	-28.8365
66	566.3	URIDINE-DYPHOPHATE-GLUCOSE	-6.16629	-25.1865
67	266.25	3-(1,3,7-trimethyl-2,6-dioxo-2,3,6,7-tetrahydro-1H-purin-8-yl)propanoic acid	-6.164191	-56.1676

Table 2 : **ADME properties of docked compounds**

S No	WEIGHT	NAME	donorHB	accptHB	QPlogPC16	QPlogPoct	QPlogPw
1	458.59	11,17-dihydroxy-3,20-dioxopregn-4-en-21-yl cyclopentanecarboxylate	2	8.45	13.279	23.382	12.834
2	178.19	9-ethyl-6-hydrazino-9H-purine	3	5	6.45	12.557	11.102
3	230.22	1-naphthyl aminocarbonylcarbamate	2	3	8.473	13.127	12.866
4	460.47	N,N'-bis((1,3-dimethyl-2,6-dioxo-2,3,6,9-tetrahydro-1H-purin-8-yl)methyl)thiourea	3	14	13.557	28.068	20.605
5	456.53	11,17-dihydroxy-3,20-dioxopregn-4-en-21-yl 2-furoate	2	8.95	13.835	23.532	14.418
6	544.65	2-allyl-N~1~,N~3~-bis(4-amino-2-methyl-6-quinolinyl)-2-benzylmalonamide	3	7	18.627	28.156	14.86
7	203.19	(8-quinolinyloxy)acetic acid	1	3.75	7.27	10.611	7.677
8	313.31	N-((1,3-dimethyl-2,6-dioxo-2,3,6,9-tetrahydro-1H-purin-8-yl)methyl)benzamide	1	9	10.736	18.212	13.547
9	224.25	3-tert-butyl-2-hydroxy-5-methoxybenzoic acid	1	2.5	6.835	9.731	5.004
10	290.27	5a,11a-dihydro-5,6,11,12-naphthacenetetrone	0	5.75	9.624	13.884	9.637
11	276.33	8-cyclohexyl-1,3,7-trimethyl-3,7-dihydro-1H-purine-2,6-dione	0	5	7.421	12.613	6.518
12	372.5	11-hydroxy-3-oxopregna-4,17-dien-21-yl acetate	1	5.7	11.036	18.175	8.997
13	329.44	4-((4-amino-3-methylphenyl)(4-imino-3-methyl-2,5-cyclohexadien-1-ylidene)methyl)-2-methylphenylamine	4	3	12.443	20.268	11.442
14	276.27	4-phenoxathiincarboxylic acid 10,10-dioxide	1	6.5	8.995	14.281	11.553
15	210.23	3-tert-butyl-2,5-dihydroxybenzoic acid	2	2.5	7.006	11.005	6.866
16	400.39	Theophylline , 8,8'-trimethylenedi-	1	11.5	10.805	21.203	14.557
17	181.22	2-(methylthio)-9H-purin-6-ylamine	3	4	6.209	12.057	10.128
18	178.19	N~2~,N~2~-dimethyl-9H-purine-2,6-diamine	3	4.5	6.151	12.228	10.594
19	259.21	2-(2,6-dioxotetrahydro-2H-pyran-3-yl)-1H-isoindole-1,3(2H)-dione	0	7.5	7.957	14.009	10.615
20	198.17	3-hydroxy-4,5-dimethoxybenzoic acid	2	4.25	6.298	10.573	8.416
21	266.25	3-(1,3,7-trimethyl-2,6-dioxo-2,3,6,7-tetrahydro-1H-purin-8-yl)propanoic acid	1	7	7.521	13.423	9.855

Sno	QP logS	CIQP logS	QPlog HERG	QPP Caco	QPlog BB	QPP MDCK	QPlog Kp	QPlog Khsa	Human Oral Absorption	Percent Human Oral Absorption	Rule Of Five
1	-5.867	-5.208	-4.718	219.924	-1.564	96.255	-4.06	0.543	3	88.603	0
2	-0.296	-0.604	-4.47	102.817	-0.423	46.815	-5.565	-0.65	2	60.419	0
3	-1.4	-2.569	-3.167	135.234	-0.802	145.045	-3.344	-0.409	2	71.299	0
4	-4.575	-4.046	-5.032	48.639	-2.174	33.259	-5.559	-0.558	2	45.631	1
5	-5.129	-5.297	-5.282	146.92	-1.702	62.239	-3.857	0.281	3	82.121	0
6	-7.695	-8.879	-7.203	423.389	-1.629	195.389	-1.683	1.221	1	70.34	3
7	-2.001	-2.495	-2.463	192.41	-0.578	105.955	-2.528	-0.532	3	78.944	0
8	-3.826	-2.885	-5.772	166.761	-1.573	71.372	-3.955	-0.472	3	79.873	0
9	-3.039	-3.103	-1.641	175.945	-0.689	96.19	-3.208	-0.115	3	82.347	0
10	-2.495	-3.584	-5.12	359.637	-0.818	163.792	-3.062	-0.472	3	81.04	0
11	-3.376	-3.133	-3.52	1639.62	-0.145	844.233	-3	-0.078	3	100	0
12	-5.778	-4.686	-4.438	337.386	-1.117	152.866	-3.975	0.68	3	92.624	0
13	-5.176	-5.379	-5.629	369.251	-1.296	168.529	-2.86	0.521	3	94.542	0
14	-2.296	-3.071	-2.729	58.883	-0.918	29.464	-3.594	-0.637	3	65.543	0
15	-2.375	-2.738	-1.617	54.982	-1.113	27.36	-4.14	-0.344	3	68.948	0
16	-3.216	-3.755	-4.343	175.858	-1.46	75.589	-4.481	-0.624	2	57.735	1
17	-1.751	-1.925	-3.547	388.805	-0.659	322.372	-3.752	-0.607	3	75.234	0
18	-1.726	-1.765	-3.664	475.457	-0.716	221.486	-3.578	-0.578	3	76.531	0
19	-0.9	-1.804	-3.978	174.057	-0.962	74.753	-4.325	-1.176	2	64.327	0
20	-1.531	-2.002	-1.5	85.057	-0.925	43.846	-3.753	-0.703	3	67.299	0
21	-2.279	-2.459	-1.702	54.662	-1.121	27.188	-4.419	-0.696	3	63.599	0

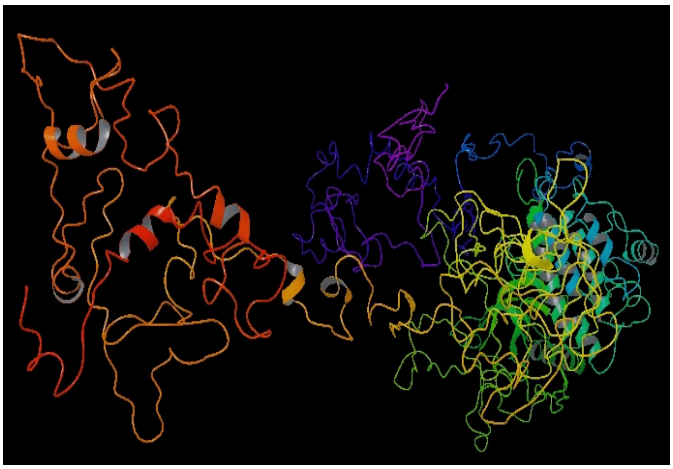

Figure 1 : Ribbon representation of EML4-Alk

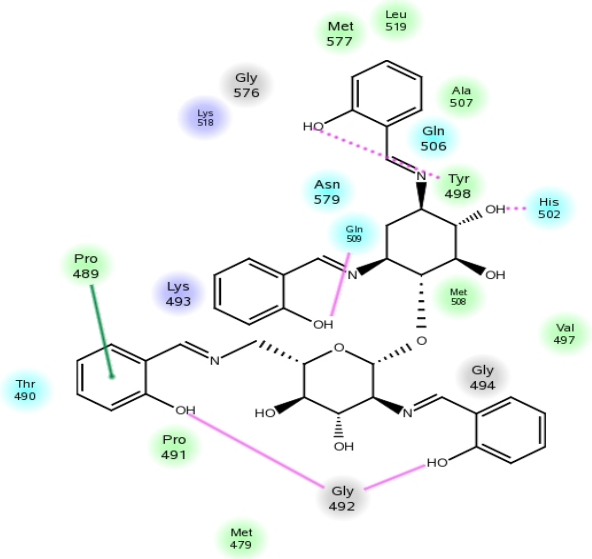

Figure 2 : Binding of Compound 1 to EML4-ALK

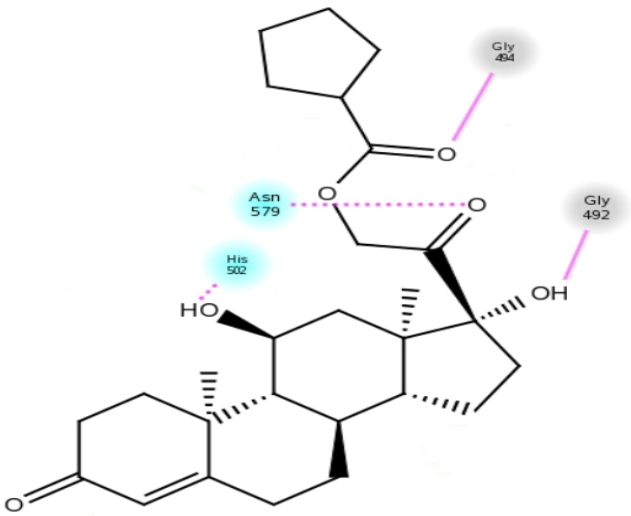

Figure 3 Binding of Compound 4 to EML4-ALK fusion protein

Source of support: Nil, Conflict of interest: None Declared