HOMOLOGY MODELLING AND DOCKING STUDIES OF FGFR3 PROTEIN TO SEARCH MOST EFFECTIVE DRUG AGAINST ACHONDROPLASIA

Chauhan Nutan[1], Ahmad Faizan[1], Siddiqui M. A.[2*], Amir Asad[2], Singh Priyansha[3], Sahu Sarika[1]

[1] Centre of Bioinformatics, Shobhit University, Meerut
[2] Department of Biotechnology, Meerut Institute of Engineering and Technology, Meerut
[3] Bioinformatics Institute of India, Noida (UP), India

E mail: - asifsiddiqui82@gmail.com

ABSTRACT: The present investigation was carried out with the aim of modeling the 3D structure of FGFR3 protein and predicting the most effective drug using SU5402 and its analogues. FGFR3 protein, responsible for long bone growth, causes Achondroplasia in *H. sapiens* when it becomes mutated. When mutated, the dimmer of FGFR3 stabilizes without interacting with its ligand results in constitutive activation of downstream pathway and inhibits the bone growth. No known structure of FGFR3 was available. The 3D modeling of FGFR3 was done using Robetta server and from the various model predicted, model 5 was selected as the best model after evaluating the models using PROCHECK. . The total number of residues in selected model was found as: 589 (86.5%) residues in most favored region, 84 (12.3%) in additional allowed region, 8 (1.2%) in generously allowed region and 0 (0.0%) in disallowed region of the Ramachandran plot. Three analogues were constructed by using the existing FGFR3 specific inhibitor SU5402. Receptor-analogue interaction study was performed in FlexX3 docking software. Ligand 3 (IUPAC Name: 3-{2-[Z)-(4-hydroxy-2-oxo-1,2-dihydro-3H-indol-3-ylidene) methyl]-4-oxo-4,5-dihydro-1H-pyrrol-3-yl} propanoic acid) was showing the best binding energy (-10.458 KJ/Mol) that can be predicted the most effective inhibitor for FGFR3. It should be noted that these predicted data should be validated using suitable assays for further consideration.

KEYWORDS: FlexX3, FGFR3, *in silico,* ligands, Robetta

INTRODUCTION

Achondroplasia or ACH (OMIM 100800) is an autosomal dominant skeletal dysplasia characterized by rhizomelic short limbs and short stature and is one of the most common causes of dwarfism and predominantly the long bone's (such as those of the upper arms and thighs) disorders (Shiang, *et al*., 1994; Rousseau *et al*., 1994). It affects about 1 in 15,000 to 1 in 40,000 births, and occurs in all races and in both sexes (Bellus *et al*., 2000). The average height of adult male with this disorder is 52 inches (4 feet, 4 inches) and of adult females is 49 inches (4 feet, 1 inch). Characteristic features for clinical diagnosis are a long and narrow trunk, a large head with frontal bossing, hypoplasia of the midface, depressed nasal bridge, and trident hands. As many as 7.5% of infants with achondroplasia die in the first year of life from obstructive apnea or central apnea (Hecht *et al*., 1987). Occasionally, a baby or young child with achondroplasia may die suddenly, often during sleep occurs in 2 to 5 percent of affected babies (Trotter *et al*., 2005), due to compression of the upper end of the spinal cord, which can interfere with breathing.

Achondroplasia was genetically mapped to the short arm of chromosome 4 in 1994 (Francomano *et al*., 1994). Although ACH inherits in an autosomal dominant manner, 80% of cases are due to new, sporadic mutations that involve the gene encoding the protein called Fibroblast Growth Factor Receptor 3 (FGFR3) that is involved in converting cartilage to bone. *FGFR3* is one of four closely related high affinity FGF transmembrane receptors (FGFR1-4). The *FGFR3* is the only gene known to be associated with achondroplasia. All people who have only a single copy of the normal *FGFR3* gene and a single copy of the *FGFR3* gene mutation have achondroplasia. The molecule of F*GFR3* is comprised of 6 domains: 3 immunoglobulin-like extracellular domains (IgI, IgII, and IgIII, each possesses S-S binding) mainly involved in ligand binding, one transmembrane domain (TM), and 2 of paired tyrosine kinase (TK1, TK2) intracellular domains. Most commonly, a point mutation causes the substitution of arginine for glycine (G380R) in the transmembrane region of the receptor that confers a "gain-of-function" mutation that severely limits the bone growth.

FGFR3 is activated and dimerize by FGF 1,2,4,8 and 9 (Chellaiah *et al*., 1994; Hecht *et al*., 1995; Kanai *et al*., 1997) and leads to autophosphorylation of tyrosine residues in the cytoplasmic domain that stimulates intrinsic tyrosine kinase activity. The phosphorylated tyrosines serve as binding sites for cellular substrates (Mohammadi *et al*., 1996). The pathways downstream of the receptor are not well defined. However, the MAP (mitogen-activated protein) kinase, STAT1 (member of a family of proteins that carry signals from activated receptors to the transcription machinery of a cell, i.e., **S**ignal **T**ransduction and **A**ctivation of **T**ranscription proteins) and PLCγ (phospholipase C) cascades have been implicated. The ultimate effect of FGFR3 signalling on bone growth is inhibitory. Gly380Arg transmembrane mutation (Webster and Donoghue, 1996) confers ligand-independent stabilization of FGFR3 dimers that results in constitutive activation of the downstream pathways so the inhibition of bones growth.

Until recently, we had no choice in improving the short stature of achondroplasia except by surgical lengthening. Although growth hormone therapy for FGFR3-related dwarfisms, including ACH and HCH, has been tried (Tanaka *et al*., 1998) but with no long term effects (Horton *et al*., 2007). The other trial is the inhibition of FGFR3 kinase using its selective inhibitors. Several works have been done on inhibition of FGFR3 using FGFR3 specific or protein tyrosine kinase specific inhibitors in treatment of multiple myelomas, bladder carcinomas and others caused due to mutation in FGFR3. Earlier studies shows that NDGA, CHIR-258, U5402, PD173074, 2-Arylidenedihydroindole-3-ones, PKC412, have been proved good for regulating multiple myelomas and bladder carcinomas in cell cultures or in mice model (Grand *et al*.,2004; Xin *et al*., 2006; Meyer *et al*., 2008; Chen *et al*., 2005; Gerby *et al*., 2007).

In the present work, we model our protein using Robetta server followed by inhibitor design using ACD/ChemSketch 12.0. **SU5402 (**IUPAC name**:** 3-[(3-(2-carboxyethyl)-4-methylpyrrol-2-yl)methylene]-2-indolinone), FGFR3 specific inhibitor, was used to construct the different analogues.

MATERIALS AND METHODS

Protein modelling and evaluation

The amino acid sequence of FGFR3 was obtained from the sequence database of National Center of Biotechnology Information (http://www.ncbi.nlm.nih.gov/) (Accession number: AAA52450, Gene ID: 182569). It was ascertained that the three-dimensional structure of the protein was not available in Protein Data Bank (PDB), so the FASTA format of the protein was submitted to Robetta server (http://robetta.bakerlab.org/) (David *et al*., 2001), an Internet service that provides automated structure prediction and analysis tools that can be used to infer protein structural information from genomic data, and obtained modelled 3-D structure, that, afterwards, evaluated by Structure Analysis and Verification Server (http://nihserver.mbi.ucla.edu/SAVES/), run by a research group at UCLA.

Ligand construction and docking

Possible ligands were designed by using ACD/ChemSketch 12.0 freeware provided by ACD/LABS software and predicted their drug–likeness and molecular properties by molSoft LLC. For finding possible binding pockets for ligands in the modelled protein metaPocket (http://metapocket.eml.org/) was used. It is an online service that uses four different methods LIGSITE, PASS, Q-SiteFinder and SURFNET to identify pocket sites (Huang, 2009). All the ligands then docked against the modelled protein with the help of FlexX3 software. It is a fast, flexible docking method that uses an incremental construction algorithm to place ligands into an active site of protein.

RESULTS AND DISCUSSION

Achondroplasia in humans is most frequently due to constitutive activation mutations within the transmembrane domain of FGFR3 (Wang *et al*., 1999). Spontaneous inactivating or genetically engineered null mutations of the FGFR3 gene result in appendicular abnormalities, enhanced proliferation of growth plate chondrocytes, and reduced cortical bone thickness (Eswarakumar and Schlessinger, 2007; Vanek *et al*., 1986). Expression of FGFR3 has been proposed to be the essential link to coordinate growth plate cell proliferation with the osteogenesis and mineralization necessary to structurally support the elongating skeleton (Eswarakumar and Schlessinger, 2007). Disruption of this coordination could jeopardize the integrity of the skeleton in terrestrial species. The pathology in animals with gain of function mutations likewise exhibit dysregulation between cellular proliferation at the growth plate and ossification (Brodie and Deng, 2003).

Until recently, we had no choice in improving the short stature of achondroplasia except by surgical lengthening. Although growth hormone (GH) therapy for FGFR3-related dwarfisms, including ACH and HCH, has been tried (Tanaka *et al*., 1998; Ramaswami *et al*., 1999). Earlier studies also reported that GH treatment for skeletal dysplasias produces a positive effect for a short time (Yamanaka *et al*., 2003). The other trial is the inhibition of FGFR3 kinase using its selective inhibitors. Several works have been done on inhibition of FGFR3 using FGFR3 specific or protein tyrosin kinase specific inhibitors in treatment of multiple myelomas, bladder carcinomas and other caused due to mutation in FGFR3. Earlier studies shows that NDGA, CHIR-258, SU5402, PD173074, have been proved good for regulating multiple myelomas in cell cultures or in mice model (Grand *et al*., 2004; Xin *et al*., 2006; Meyer *et al*., 2008).

Other works have been done using 2-Arylidenedihydroindole-3-ones, PKC412 in bladder carcinoma cell lines and hematopoetic malignancies in vitro (Chen *et al*., 2005; Gerby *et al*., 2007). More work on the treatment of achondroplasia through FGFR3 inhibition has not been done, however, one work suggested that the over activity of FGFR3 was inhibited by PPADS that reduced the tyrosine phosphorylation of FGF receptor type 3 triggered by fibroblast growth factor 9 (FGF9) (50% reduction) on achondroplasic chondrocytes (Guzmán-Aránguez *et al*., 2008). These approaches have been proven to have a blocking effect on FGFR3 activation in vitro. Despite progress in characterization of the FGFR3, many aspects remain unclear and are poorly understood.

Therefore, the FASTA format of full sequence of FGFR3 protein was obtained from NCBI which consists 806 aa. Robetta server was used for homology modelling that models protein in fully automated manner (David *et al*., 2004). This internet service parses the input sequence into domains and builds models for domains with sequence homology to proteins of known structure using comparative modelling, and models for domains lacking such homology using Rosetta do novo structure prediction method. A total of 5 rough models were generated and domain predictions and molecular coordinates of the models spanning the full length query were given as results.

The generated rough models were subjected to SAVS for testing their internal consistency and reliability. For this, Phi and Phi torsion angles of predicted models were evaluated by PROCHECK analysis (Laskowski *et al*., 1996) and model 5 (Figure. 1) was selected for further studies. The total number of residues in selected model was found as: 589 (86.5%) residues in most favored region, 84 (12.3%) in additional allowed region, 8 (1.2%) in generously allowed region and 0 (0.0%) in disallowed region of the Ramachandran plot (Figure. 2).

Figure 1: **Predicted 3D structure of FGFR3 Protein**

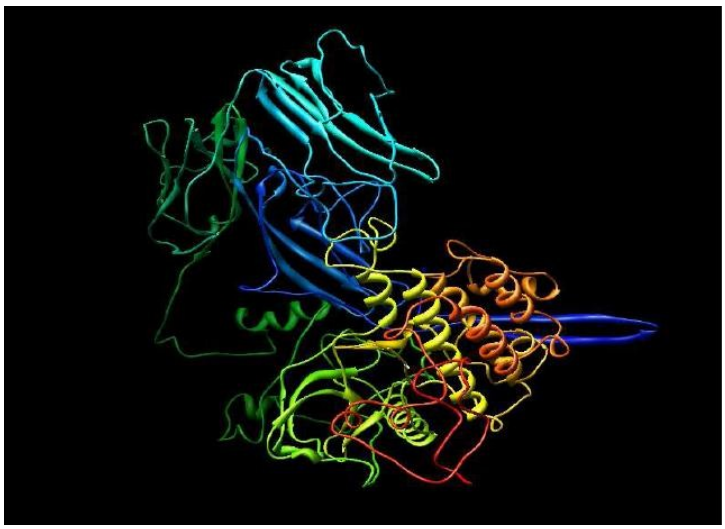
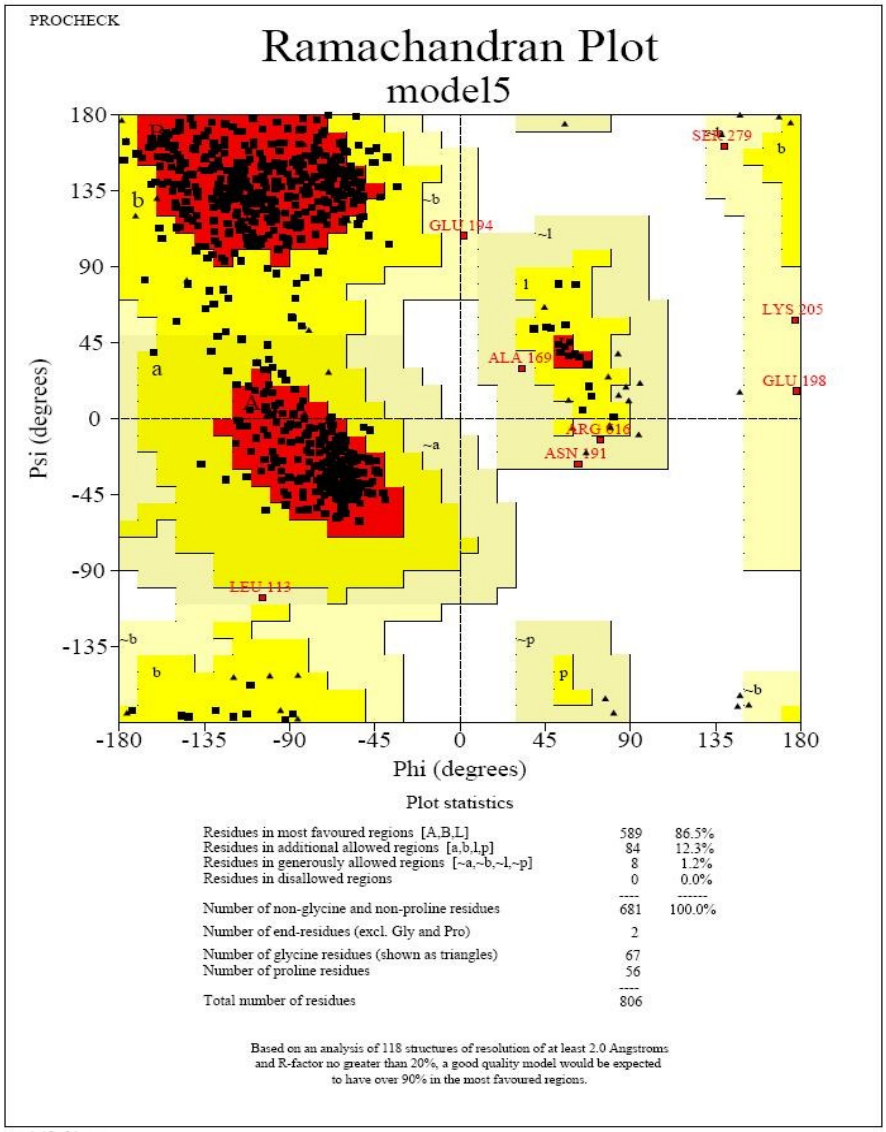


Figure 2: Ramachandran Plot analysis of predicted model structure of FGFR3



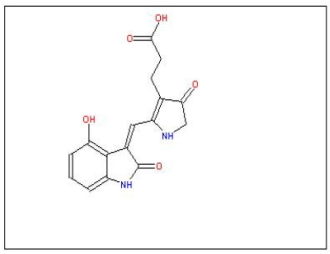
For finding ligand binding pockets metaPocket server was used. Total 10 binding pockets were predicted from which first binding pocket (Table. 1), which was used in docking study with FlexX3 (Rarey *et al*., 1996). The residues taken as active site by FlexX 3 are shown highlighted. SU5402 (Paterson *et al*., 2004) (a FGFR3 specific inhibitor) was used as template for *in silico* designing of ligands in ACD/ChemSketch 12.0 software. Total 3 ligands were designed. All the proposed ligands were submitted to MolSoft L.L.C. for the prediction of their molecular properties and drug-likeness. Ligand 3 shows best drug likeness model score (Figure. 3).

Table 1: **The first binding pocket predicted by metaPocket used in docking study.**

ATOM	4502	PRO	304	ATOM	4586	LEU	309
ATOM	4509	PRO	304	ATOM	5710	LEU	385
ATOM	4520	TYR	305	ATOM	5711	LEU	385
ATOM	4521	TYR	305	ATOM	5716	LEU	385
ATOM	4523	TYR	305	ATOM	5717	LEU	385
ATOM	4524	TYR	305	ATOM	5718	LEU	385
ATOM	4525	TYR	305	ATOM	5719	LEU	385
ATOM	4526	TYR	305	ATOM	5720	PHE	386
ATOM	4527	TYR	305	ATOM	5721	PHE	386
ATOM	4528	TYR	305	ATOM	5722	PHE	386
ATOM	4529	TYR	305	ATOM	5723	PHE	386
ATOM	4530	TYR	305	ATOM	5724	PHE	386
ATOM	4531	TYR	305	ATOM	5725	PHE	386
ATOM	4532	TYR	305	ATOM	5726	PHE	386
ATOM	4533	TYR	305	ATOM	5734	PHE	386
ATOM	4534	TYR	305	ATOM	5740	ILE	387
ATOM	4535	VAL	306	ATOM	5741	ILE	387
ATOM	4536	VAL	306	ATOM	5742	ILE	387
ATOM	4537	VAL	306	ATOM	5743	ILE	387
ATOM	4551	THR	307	ATOM	5744	ILE	387
ATOM	4552	THR	307	ATOM	5745	ILE	387
ATOM	4553	THR	307	ATOM	5746	ILE	387
ATOM	4554	THR	307	ATOM	5747	ILE	387
ATOM	4555	THR	307	ATOM	5748	ILE	387
ATOM	4556	THR	307	ATOM	5749	ILE	387
ATOM	4557	THR	307	ATOM	5750	ILE	387
ATOM	4560	THR	307	ATOM	5752	ILE	387
ATOM	4561	THR	307	ATOM	5753	ILE	387
ATOM	4562	THR	307	ATOM	5754	ILE	387
ATOM	4565	VAL	308	ATOM	5755	ILE	387

At last, all the ligands were docked against the modelled protein using FlexX3. This software works on the principle of incremental construction algorithm. The inputs for FlexX3 were ligand and the receptor protein (the binding pocket) in mol2 format. FlexX3 results shows that the binding energy for SU5402 was -4.794 KJ/MOL with only one hydrogen bond involved in the interaction with THR 307 (Figure. 4). Among the all ligands only ligand 3 was showing the lowest docking energy -10.458 KJ/MOL with two hydrogen bonds and three hydrophobic interactions involved in binding between the ligand and the active site residues VAL (308), TYR (305) (Figure. 5).

Figure 3: **Molecular properties and drug-likeness of Ligand 3**



Molecular formula: C16 H14 N2 O5
Molecular weight: 314.09
Number of HBA: 5
Number of HBD: 4
MolLogP : 0.36
MolLogS : -3.32 (in Log(moles/L)) 151.02 (in mg/L)
MolPSA : 96.27 A^2
MolVol : 332.86 A^3
Number of stereo centers: 0

Drug-likeness model score: -0.76

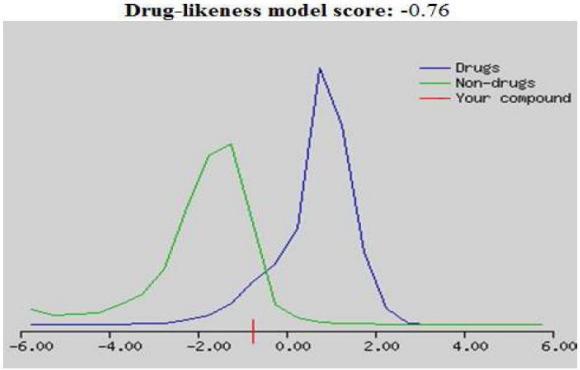
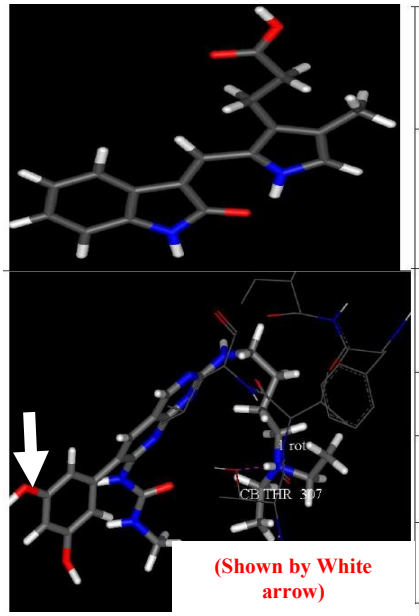


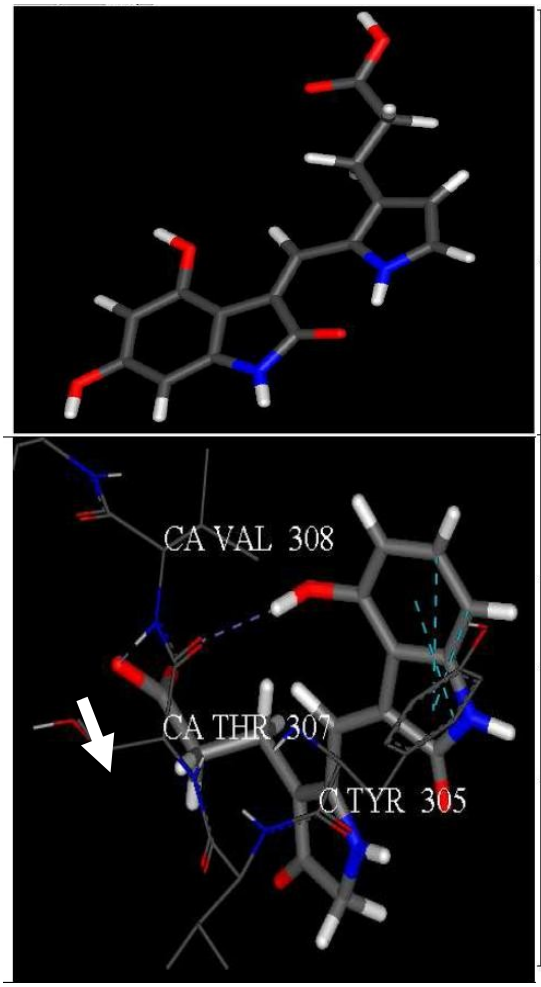
Figure 4: **Docking Result of SU5402**



(Shown by White arrow)

IUPAC NAME: 3-{4-methyl-2-[(*Z*)-(2-oxo-1,2-dihydro-3*H*-indol-3-ylidene)methyl]-1*H*-pyrrol-3-yl}propanoic acid
SMILE NOTATION: O=C(O)CCc1c(enc1/C=C3/c2ccccc2NC3=O)C
DOCKING ENERGY: -4.794 Kj/mol
HYDROGEN BONDS: ONE
HYDROPHOBIC INTERACTIONS: NO
AMINO ACIDS: THR 307

Figure 5: **Docking result of Ligand 3**



IUPAC NAME: 3-{2-[(*Z*)-(4-hydroxy-2-oxo-1,2-dihydro-3*H*-indol-3-ylidene)methyl]-4-oxo-4,5-dihydro-1*H*-pyrrol-3-yl}propanoic acid
SMILE NOTATION: O=C(O)CCC=1C(=O)CNC=1/C=C3/c2c(cccc2O)NC3=O
DOCKING ENERGY: -10.458 KJ/MOL
HYDROGEN BONDS: TWO
HYDROPHOBIC INTERACTIONS: THREE
AMINO ACIDS: VAL 308, TYR 305
(Shown by White arrow)

CONCLUSION

Analyzing the diseases and designing drugs with respect to the gene and protein level helps to find the underlying causes of the diseases, and to improve their rate of cure. To draw a link between genes and their relation to disease outcomes and drug discovery, bioinformatics is a promising method. Most techniques of bioinformatics are used in the first two phases of drug discovery to extract interesting information and find important genes and/or proteins for speeding the process of drug discovery, enhancing the accuracy of analysis and reducing the cost.

In present work, the model was designed for FGFR3 protein and docked with the designed ligands. All the three proposed ligands were showing good binding energy than SU5402, the original inhibitor of FGFR3. Among these ligands, ligand 3 (3-{2-[(*Z*)-(4- hydroxy-2-oxo-1, 2-dihydro-3*H*-indol-3-ylidene) methyl]-4-oxo-4, 5-dihydro-1 *H* pyrrol- 3-yl} propanoic acid) showed the best docking energy (-10.458 KJ/MOL) with 2 H-bonds. This ligand can be used as a drug against FGFR3 protein in future, but it should be noted that these predicted data should be validated using suitable assays for further consideration in future studies.

REFERENCES

Bellus G.A., Speiser P.W., Weaver C.A., Garber A.T., C.R. Bryke *et al*., (2000). Distinct missense mutations of the FGFR3 Lys650 codon modulate receptor kinase activation and the severity of the skeletal dysplasia phenotype. *Am. J. Hum. Genet*., **67:** 1411–1421.

Brodie S.G., Deng C.X., (2003). Mouse models orthologous to FGFR3-related skeletal dysplasias. *Pediatr Pathol Mol Med*, **22**:87-103.

Chellaiah A.T., McEwen D.G., Werner S., Xu J., Ornitz D.M., (1994). Fibroblast growth factor receptor (FGFR) 3. Alternative splicing in immunoglobulin-like domain III creates a receptor highly specific for acidic FGF/FGF-1. J. Biol. Chem., **269**: 11620- 11627.

Chen J., Lee B.H., Williams I.R., Kutok J.L., Mitsiades C.S. *et al*., (2005). FGFR3 as a therapeutic target of the small molecule inhibitor PKC412 in hematopoietic malignancies. *Oncogene*., 2**4:** 8259-67.

David E.K., Dylan C., Baker D., (2004). Prediction and analysis using the Robetta server, *Nucleic Acids Research*, **32:** 526–531.

David E.K., Chivian D., David B., (2001). Protein structure method for binding affinity estimation. *J. computer-aided molecular design*, **15:** 426-429.

Eswarakumar V.P., Schlessinger J., (2007). Skeletal overgrowth is mediated by deficiency in a specific isoform of fibroblast growth factor receptor 3. *Proc Natl Acad Sci USA*, **104**:3937-42.

Francomano C.A., de Luna R.I.O., Hefferson T.W., Bellus G.A. and Turner C.E. *et al*., (1994). Localization of the achondroplasia gene to the distal 2.5 Mb of human chromosome 4p. *Hum. Molec. Genet*., **3**: 787-792.

Gerby B., Boumendjel A., Blanc M., Bringuier P.P., Champelovier P. *et al*., (2007). 2-Arylidenedihydroindole-3-ones: design, synthesis, and biological activity on bladder carcinoma cell lines. *Bioorg. Med. Chem. Lett*., **17:** 208-13.

Grand E.K., Chase A.J., Heath C., Rahemtulla A., (2004). Targeting FGFR3 in multiple myeloma: inhibition of t(4;14)-positive cells by SU5402 and PD173074. Cross NC. *Leukemia.* **18:** 962-6.

Guzmán-Aránguez A., Almudena C., Avner Y., Jesús P., (2008). Effect of PPADS on achondroplasic chondrocytes: Inhibition of FGFreceptor type 3 over-activity. *European Journal of Pharmacology*, **584**: 72–77

Hecht D., Zimmerman N., Bedford M., Avivi A., Yayon A., (1995). Identification of fibroblast growth factor 9 (FGF9) as a high affinity, heparin dependent ligand for FGR receptors 3 and 2 but not for FGF receptors 1 and 4. *Growth Factors*, **12**: 223- 233.

Hecht J.T., Francomano C.A., Horton W.A., Annegers J.F., (1987). Mortality in achondroplasia. *Am. J. Hum. Genet*., **41**: 454-467.

Horton W.A., Hall J.G., Hecht J.T., (2007). Achondroplasia. *Lancet*, **370:** 162-172.

Huang B., (2009). MetaPocket: A meta approach to improve protein ligand binding site prediction. *Omics,* **13:** 325-330**.**

Kanai M., Goke M., Tsunekawa S., Podolsky D.K., (1997). Signal transduction pathway of human fibroblast growth factor receptor 3. *J. Biol. Chem*., **272**: 6621- 6628.

Laskowski R.A., MacArthur M.W., Moss D.S., Thornton J.M., (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*., **8**: 477-486.

Meyer A.N., McAndrew C.W., Donoghue D.J., (2008). Nordihydroguaiaretic acid inhibits an activated fibroblast growth factor receptor 3 mutant and blocks downstream signaling in multiple myeloma cells. *Cancer Res*., **68:** 7362-7362.

Mohammadi M., Dikic I., Sorokin A., Burgess W.H., Jaye M., Schlessinger J., (1996). Identification of six novel autophosphorylation sites on fibroblast growth factor receptor 1 and elucidation of their importance in receptor activation and signal transduction. *Molec. Cell. Biol*., **16**: 977- 989.

Paterson J.L., Zhihua L., Wen X. , Khan E. M., Chang H., Pollett J.B. , Trudel S., Stewart A., (2004). Preclinical studies of fibroblast growth factor receptor 3 as a therapeutic target in multiple myeloma. *Br. J. Haematol.*, 124 (5):**595– 603.**

Ramaswami U., Hindmarsh P.C., Brook C.G., (1999). Growth hormone therapy in hypochondroplasia. *Acta Paediatr Suppl*, **88(428):** 116–117.

Rarey M., Kramer B., Lengauer T., Klebe G., (1996). A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol*., **261**:470–489.

Rousseau F., Bonaventure J., Legeai-Mallet L., Pelet A., Rozet J.M. *et al*., (1994). Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. *Nat. (Lond)*, **371**: 252–254.

Shiang R., Thompson L.M., Zhu Y.Z., Church D.M., Fielder T.J. *et al*., (1994). Mutations in the transmembrane domain of FGFR3 cause the most commongenetic form of dwarfism, achondroplasia. Cell, **78**: 335–342.

Tanaka H., Kubo T., Yamate T., Ono T., Kanzaki S., Seino Y., (1998). Effect of growth hormone therapy in children with achondroplasia: growth pattern, hypothalamicpituitary function, and genotype. *Eur. J. Endocrinol*., **138:** 275–280.

Trotter T.L., Hall J.G., (2005). Health supervision for children with achondroplasia. *Pediatrics*, **116**: 771-783.

Vanek J, Alstad A., Berg I., Misek A., Moore B., Limesand W., (1986). Spider syndrome in lambs: A clinical and postmortem analysis. *Vet Med*, **81**:663-668.

Wang Y., Spatz M.K., Kannan K., Hayk H., Avivi A., Gorivodsky M., Pines M., Yayon A., Lonai P., Givol D., (1999). A mouse model for achondroplasia produced by targeting fibroblast growth factor receptor 3. *Proc Natl Acad Sci USA*, **96**: 4455-60.

Webster M.K., Donoghue D.J., (1996). Constitutive activation of fibroblast growth factor receptor 3 by the transmembrane domain point mutation found in achondroplasia. *EMBO. J.,* **15** : 520-527.

Xin X., Abrams T.J., Hollenbach P.W., Rendahl K.G., Tang Y. *et al*., (2006). CHIR-258 is efficacious in a newly developed fibroblast growth factor receptor 3-expressing orthotopic multiple myeloma model in mice. *Clin. Cancer Res*., **12:** 4908-15.

Yamanaka Y., Kanazawa H., Tanaka H., Inoue M., Namba N., Seino Y., (2003). Efficacy of growth hormone therapy for patients with skeletal dysplasia. *J Bone Miner Metab* **21(5):** 307–310.

****