J Young Pharm, 2018; 10(2): 178-182
A multifaceted peer reviewed journal in the field of Pharmacy
www.jyoungpharm.org | www.phcog.net

Original Article

Molecular Modeling and Docking Assessment of Thymic Stromal Lymphopoietin for the Development of Natural Anti Allergic Drugs

Rahma Muhammad Adamu*, B. K. Malik
Department of Biotechnology, school of Engineering and Technology, Sharda University, Greater Noida, Uttar Pradesh, INDIA.

ABSTRACT

Objective: To model a 3D-structure of Thymic Stromal Lymphopoietin (TSLP) and docking of small molecules for development of novel natural anti-allergic drug. **Method:** Crystal structure of the *Mus musculus* cytokine receptor (Protein Data Bank ID: 4NN5) was used as template for comparative modeling in MODELLER (v 9.16). Loop refinement of the modeled structure was performed in MODELLER and further verified and analyzed using Ramachandaran plot and ERRAT. Molecules from ZINC Natural product database were randomly docked into the predicted model using Autodock 4.2 program. Molecules with lower binding affinity were filtered based on molecular descriptors of Wlogp < 5 and TPSA < 140 Å² and further assessed for pharmacokinetics and bioavailability property using Boiled egg model, and later checked for mutagenic, tumorigenic, reproductive effect and irritant toxicity analysis. **Results:** The modeled 3-D structure of TSLP has shown a good quality model with 91.2% at the most favored region by Ramachandran statistics and 86.4 overall qualities of ERRAT evaluations. Six molecules were selected as top hits by passing the filtering criteria and their binding interaction of hydrophilic, hydrophobic and pi-cation with the target. They have good pharmacokinetic properties with high gastro-intestinal absorption, the toxicity analysis shows all the six molecules to have non-mutagenic, non-tumorigenic and non-reproductive effect. **Conclusion:** This work predict the 3D structure of TSLP using comparative modelling methods. Six ligands were identified with high binding energy by molecular docking approaches and optimized to have good gastrointestinal absorption and non-toxic. They can be good oral drugs against allergic diseases.

Key words: Allergy, TSLP, Modeling, Natural, Drugs, Pharmacokinetics.

Key Message: Investigation of novel natural molecules that form hydrophobic and pi-cation interactions with TSLP predicted to be good oral anti-allergic drugs.

Correspondence

Dr. Rahma Muhammad Adamu, PhD Scholar in Computational Biology Lab, Department of Biotechnology, School of Engineering and Technology, Sharda University Greater Noida, Uttar Pradesh, INDIA.

Phone: +917428568519

Email: turanjoji84@gmail.com

DOI: 10.5530/jyp.2018.10.40

INTRODUCTION

Among the significant health problems of the world today, allergic diseases are not to be excluded and it affect both people in developed and developing countries. Allergic diseases are caused by different allergens and resulting in allergic diseases including asthma, rhinitis, food allergies, and atopic dermatitis. The incidence of allergic disease is increasing worldwide with a great epidemic proportion in association to environmental exposure to allergens and with modern lifestyles. TSLP is a master switch at the interface between the environmental allergens and the pulmonary allergic immunologic responses, and plays a central role in polarizing dendritic cells (DCs) by enhancing OX40L expression, which induces the differentiation of naive T cells into Th2 cells.[1-3]

Thymic stromal lymphopoietin (TSLP) is related to IL-7 and is an epithelial-derived cytokine significantly elevated in those with asthma and allergic diseases. TSLP is expressed by epithelial cells of the skin, gut and lung and primes resident dendritic cells to promote Th2 cytokine production by their subsequently engaged effector T cells. Allergens stimulate the production of TSLP in inflamed tissue and their receptors primarily expressed by dendritic cells are expressed on mast cells and also promotes allergic responses. TSLP is also produced at barrier surfaces, fibroblasts, mast cells, and keratinocytes.[4,5] Expression of TSLP when proteases interact with PAR2 contribute to allergic inflammation,[5] this indicate that TSLP might have a function in early activation of the innate host defense response at the initial site of exposure in the epithelium, leading to Th2 polarization.[5]

Structurally, TSLP resembles IL-7, consisting of a four-helix bundle cytokine. Despite poor amino acid sequence identity between murine and human TSLP they have similarity at the functional level.[6,7] Therefore the murine TSLP with only 40% identity with human TSLP was used as a template in this study to build the 3-Dimensional structure of human TSLP to develop drug that can inhibit it function because of its important in the initiation of allergic inflammation. Allergic diseases has become complicated disease involving several cells and increase rate of recurrence it is therefore necessary to develop new drugs that can potentially treat these diseases. Hence, drugs discovered from virtual screening of large natural database using computational methods by targeting this potential therapeutic target (Thymic stromal lymphopoietin) will help in the treatment of this disease at minimum time, low cost and free side effects. This will definitely reduce the rate of recurrence of this disease because of the promiscuous effectiveness of natural compounds. This study focuses on determining the 3D structure of TSLP and inhibiting it function using natural small molecules by molecular docking methods for the development of anti-allergic drugs.

MATERIALS AND METHODS

Sequence Retrieval, Template Selection and Secondary Structure Prediction

The amino sequence of TSLP (159 residues) was retrieved and saved in FASTA format for comparative modeling from National Center for Biotechnology Information (NCBI). The sequence (NP_149024.1) was then subjected to protein-protein Basic Local Alignment Search Tool (BLASTp) against the Protein Data Bank for search of suitable template. Chrystal structure of chain A Cytokine receptor from *Mus musculus* (mouse) with PDB ID 4NN5, sequence Identity of 40%, E-value of 3e-23, Query coverage of 77% and resolution of 1.9 Å was selected as a template. The amino sequence of the target sequence (TSLP) was analyzed by Self-optimized Prediction Method with Alignment (SOPMA).[8] further analysis of the template (4NN5) structure was performed in UCSF Chimera.[9]

Homology Modeling and Evaluation of Models

During the course of this study the human 3-Dimensional structure of Thymic Stromal Lymphopoetin (TSLP) has not been determined by either X-ray crystallography or Nuclear Magnetic Resonance methods. To understand the binding of small molecules to TSLP for drug design, it is therefore apparent to determine its 3-Dimensional structure thereby we apply homology modeling approaches to determine the 3D structure. The alignment of TSLP and the template sequences was generated using the **Align2d** command (it takes into account structural information from the template when constructing an alignment) in MODELLER 9v16.[10] once a target-template alignment is constructed, using **auto model** class MODELLER calculates a 3D model of the target automatically. Five similar models of TSLP were generated based on the template structure (4NN5) and the alignment file. Among the five 3D structures developed, the one with low Discrete Optimized Protein Energy (DOPE) was chosen and evaluated using PROCHECK [11] and ERRAT [12] online servers. Observed loops from superimposition (TSLP model vs. 4NN5) and unconstructed loops comparative to the non-corresponding structures of the template were refined by running the loop modeling script in MODELLER. The generated loop optimized model was further evaluated using PROCHECK and ERRAT.

Molecular Docking based Virtual Screening

Virtual screening based on molecular docking is one of the most widely used methods of structure based drug design. This approach provides molecular information about protein-ligand interactions and thus is important for lead discovery and optimization. To identify novel hits natural compounds we therefore screened freely accessible large database; ZINC Natural Product Database. The docking processes were performed in Autodock 4.2.[13] Polar hydrogens were added and Lamarckian algorithm was applied. A grid of 60, 60, and 60 points in X (4.348), Y (7.451), and Z (-5.352) directions defined by the Autodock program was built with a grid spacing of 0.375Å. All other parameters were remained as default. The ones with the lowest binding energy were retained as hits compounds and further subjected to pharmacokinetics and toxicity evaluations.

Pharmacokinetics and Toxicity analysis of the Hit Compounds

The SMILES of the set of compounds with lowest binding energy were submitted to SwissADME [14] online server for study of their pharmacokinetics and drug-likeness properties. Molecules were filtered out based on 2 molecular properties which are Water Partition Coefficient (WlogP) values to be less than 5 which predict low level of toxicity, non-specific binding and possible oral administration,[15] Topological Polar Surface Area (TPSA) less than 140 Å^2 indicating a high possibility of complete absorption.[16] Based on these criteria ligands that form good interaction with the protein were selected as hits. Further the hits ligands were analyzed for Bioavailability property using Boiled Egg analysis [17] and toxicity analysis using Osiris Datawarrior.[18] Protein-ligand interactions in 2D diagrams, bond distance and acceptor angle were analyzed in Schrodinger suite.[19] The atoms involved in the protein, and hit ligands, were discovered using PyMOL [20] software.

RESULTS

Secondary Structure Prediction

The prediction result has shown that the secondary structure of TSLP has been dominated with alpha helix 44.65% (71 residues) followed by random coils 30.19% (48 residues), extended strand 20.13% (32 residues) and Beta turn 9.23% (12 residues). Further analysis of the template structure (Figure 3A) in UCSF chimera indicates non corresponding structures in some residues of the crystal structure of the chain A of 4NN5 (cytokine receptor from *Mus musculus*).

Homology Modeling and Evaluations

The basic modeling of the TSLP resulted in five models and the one with lowest DOPE score of -13572.77344 and Molpdf of 1952.70386 was chosen. The model was further evaluated using Ramachandran plot (Figure 1) with 86.4% in Most favored regions [A,B,L] and 49.219 % of ERRAT results, (Figure 2) from the results important loops at residues number 108 to 126, 38 to 45 and 57 to 68 were observed. The loops were further optimized by running the loop modeling script in MODELLER 9.16 and the resulted optimized model with Molpdf of 2971.45166 was also evaluated and significant change in the Ramachandran plot statistics and ERRAT outcome ware observed (Table 1). The structural superimposition Cα trace of predicted model (TSLP) and template (4NN5) has shown RMSD of 3.977 (Figure 3C). Further, the results of the match-alignment of predicted TSLP 3D structure and chain A of template 4NN5 in UCSF Chimera has shown the missing residues and the loops as observed in Figure 4.

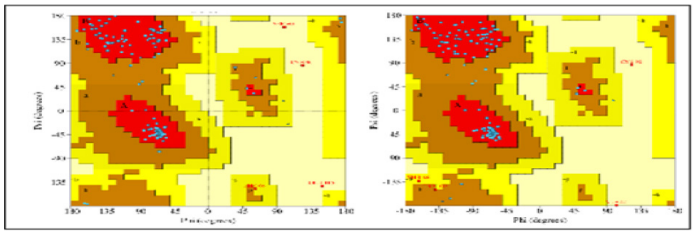

Figure 1: Ramachandran plot of (a) Basic model and (b) optimized loops model of predicted structure of TSLP.

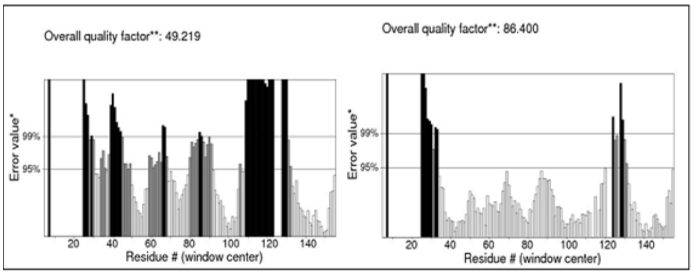

Figure 2: ERRAT graph of (a) Basic model and (b) optimized loops model of predicted 3D structure of TSLP.

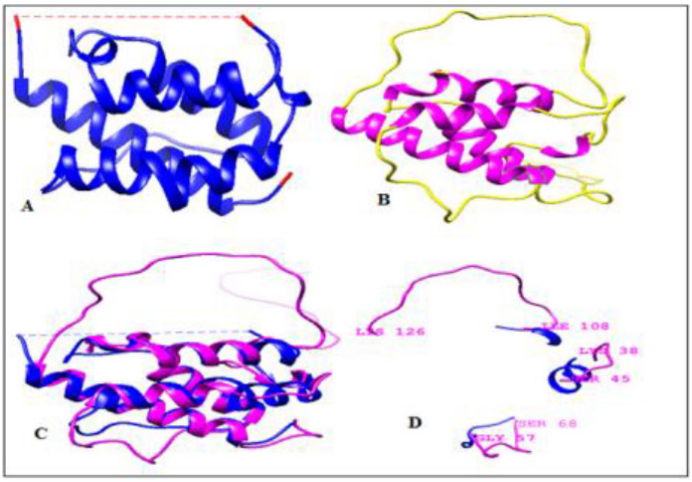

Figure 3: (A) 3D structure of the template PDB ID 4NN5, (B) predicted 3D structure of TSLP constructed using MODELLER 9.16, Helices are in Magenta and loops are in yellow (C) Superimposition of the template (4NN5) and the modeled structure of TSLP with RMSD of 3.977 Å (D) The modeled loops (magenta) from ILE108-lys126 in comparison with the missing structure of the template, modeled loops (magenta) from LYS38-SER45 in comparison with the helices (blue) from the template and the optimized loop (magenta) from GLY57-SER68 in comparison with loops (blue) from the template.

Table 1: Evaluations of the Basic and Loop optimized Models.

PROCHECK	Basic Model	Loop Optimized Model
Most favoured regions [A,B,L]	86.4%	91.2%
Additional regions allowed [a,b,l,p]	8.8%	6.1%
Generously allowed regions [~a,~b,~l,~p]	0.7%	1.4%
Disallowed regions [XX]	2.0%	1.4%
ERRAT (overall quality)	49.219%	86.4%

Molecular Docking and Pharmacokinetic properties of the hits ligands

Several molecules bind to the predicted 3D structure of TSLP, among which 58 molecules with lowest binding affinity were filtered out and evaluated for pharmacokinetic and physicochemical properties using SwissADME server. Six molecules (Table 2) were selected as top hits by passing the filtering criteria of Wlogp < 5 and TPSA < 140 Å² and were further studied for their bonding interaction (Table 3) with the target protein. Boiled egg model is proposed as an accurate predictive model that works by computing the lipophilicity and polarity of small molecules.[21] The Boiled egg analysis of the six molecules (Figure 5) has shown all to be highly absorbable in the gastrointestinal tract.

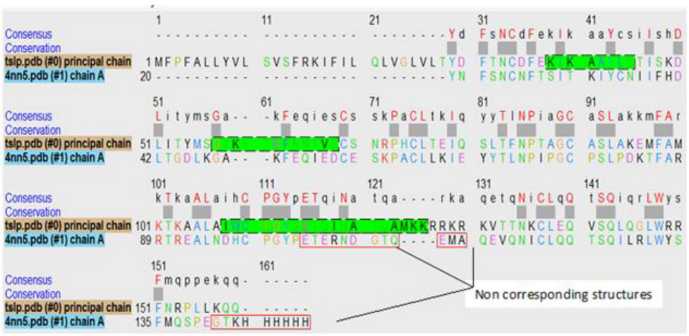

Figure 4: Match-Alignment of predicted TSLP model and chain A of template 4NN5. The dots indicate the missing residues, the red letters indicate the conserved residues in the human predicted 3D-structure of TSLP and that of mouse crystal structure (4NN5). The green highlighted residues show the optimized loops that were further modeled.

Protein Ligand Interactions

The protein ligand interactions is specific because different side chains formed different bonds. ZINC19370008, ZINC19309311, ZINC00718292 and ZINC15959260 form hydrogen bonds with amides (Asn85 Gln80) residues of TSLP, they are polar due to Oxygen and Nitrogen drawing electrons towards them and tend to be better in hydrogen bonding. The aliphatic (Leu106) and the aromatic (Tyr29 & Trp109) amino acids are essentially non-polar and therefore interact mainly via hrdrophobic interactions and vander waals forces, therefore the interaction between ZINC20760321 and ZINC98368471 with TSLP residues is hydrophobic interaction.

DISCUSSION

The non-structural residues of the template 4NN5 correspond to the loop of the TSLP from residues 108 to 126 that were further modeled and form a good TSLP 3D-structure suitable for further docking studies. Two molecules (ZINC19370008 and ZINC19309311) cross the blood brain barrier, ZINC98368471 and ZINC00718292 are non-substrate to P-gp. It is of great importance in drug design to screened compounds that are non-substrate to P-gp at early stage to avoid drug-drug interactions. The physicochemical properties such as solubility and lipophilicity play a significant role of whether a drug can progress to be a successful drug candidate[21] therefore determination of these parameters at early stage of drug discovery is necessary. Understanding the protein ligand interactions is very important because many drugs bind selectively to target proteins in the body. All the six molecules bind (Table 3) to the predicted model of TSLP with good binding anergy and form protein ligand interactions (Figure 6). The molecule ZINC98368471 form a pi-cation interactions with the protein which is a strong, non-covalent

Table 2: Calculated physiochemical and pharmacological properties of the hits ligands.

Ligands	WlogP	TPSA (Å²)	GI absorption	BBB permeant	P-gp substrate	CYP450 inhibitors
ZINC19370008	2.06	58.12	High	yes	Yes	CYP2C9
ZINC98368471	1.00	65.27	High	No	No	CYP2D6
ZINC19309311	2.05	47.12	High	Yes	Yes	CYP1A2, CYP2C9
ZINC00718292	1.76	86.46	High	No	No	CYP2C9
ZINC15959260	-0.30	118.64	High	No	Yes	CYP3A4
ZINC20760321	-0.33	121.60	High	No	Yes	CYP3A4

wlogP; water partition coefficient, TPSA; topological polar surface area, GI: gastrointestinal, BBB; blood brain barrier, LogKp; skin permeation coefficient.

Table 3: Binding energy and protein ligand interaction analysis of the hits molecules.

Ligands	Binding energy (kcal/mol)	Inhibition constant(K_i)	Interacting residues		Bond distance	Acceptor Angle
			Protein	Ligand		
ZINC19370008	-9.98	48.63 nM	GLN80CD-OE1	H_7N_1	1.87	138.83
			ASN85CG-OD1	$H_{22}O_{21}$	1.94	147.28
ZINC98368471	-9.55	99.70 nM	TRP109 cation-π	NH^+	5.98	20.27
			TRP109 cation-π	NH^+	6.17	24.69
			TYR29 cation- π	NH^+	4.38	20.33
ZINC19309311	-9.43	123.11 nM	GLN80CD-OE1	H_7N_1	1.89	136.51
ZINC00718292	-8.66	447.42 nM	GLN80CD-OE1	$H_{26}O_{25}$	2.02	139.537
ZINC15959260	-8.27	859.98 nM	GLN80CD-OE1	$H_{22}N_8$	1.89	169.833
ZINC20760321	-7.99	1.40 uM	LEU106C-O	$H_{27}N_{26}$	1.89	144.34

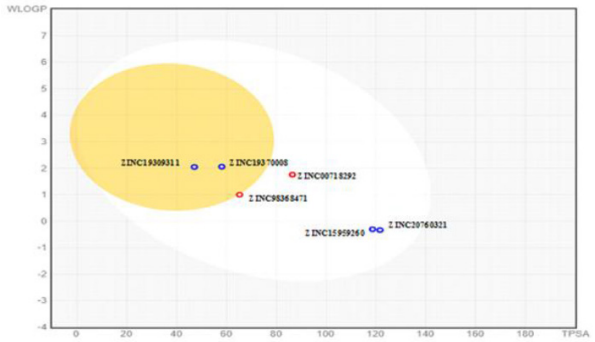

Figure 5: Boiled egg plot, water partition coefficient (WlogP) vs. Topological polar surface area (TPSA) of the hits Legends.

Highest probability of being absorbed by the gastrointestinal tract

Highest probability to permeate to the brain

Non-substrate to P-gp

Substrate to P-gp

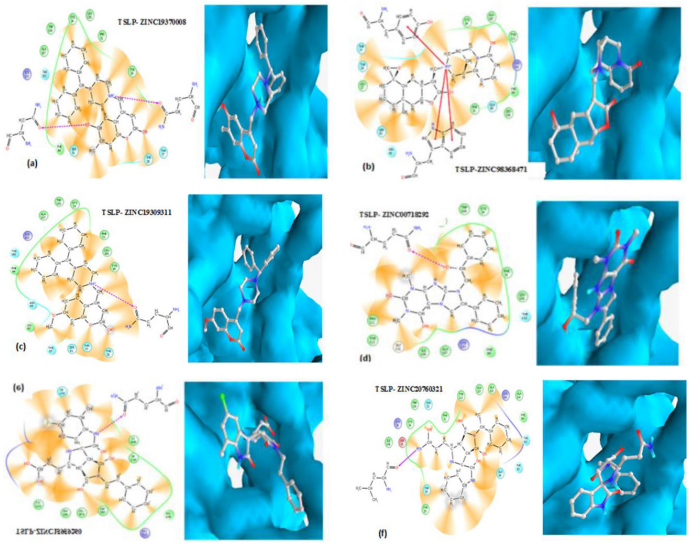

Figure 6: (a-f) showing the protein ligand interaction of docked natural molecules to predicted 3D structure of TSLP. Legends: H-bond (backbone) Hydrogen-bond (side chain) Pi-cation.

Table 4: Toxicity analysis and drug-likeness of the hits molecules.

Ligands	Mutagenic	Tumorigenic	Reproductive effect	Irritant
ZINC19370008	none	none	none	None
ZINC98368471	none	none	none	Low
ZINC19309311	none	none	none	Low
ZINC00718292	none	none	none	none
ZINC15959260	none	none	none	none
ZINC20760321	none	none	none	Low

LLE, Liphophilic Ligand Efficiency

binding force that is used throughout nature. Tyrosine (Tyr), and Tryptophan (Trp) are generally hydrophobic and can contribute hydrogen bonds, the pi-cation interactions between amino acids contribute significantly to stabilizing protein secondary structure. In addition, drug-receptor interactions across a wide array of systems use pi-cation interactions.[22] Therefore from the six molecules that bind to the predicted TSLP, ZINC98368471 with -9.55 binding affinity has shown to be the lead because it has good interaction with the protein and good pharmacokinetics properties. It has high gastrointestinal absorption, CYP2D6 inhibition and non-substrate to P-glycoprotein. Although from the toxicity analysis ZINC19309311 has low irritant character but it is predicted to be non-mutagenic, non-tumorigenic and has zero reproductive effect (Table 4).

CONCLUSION

This work predict the 3D structure of Thymic Stromal Lymphopoietin; an important therapeutic target for allergy using comparative modelling methods. Six ligands were identified with good binding energy by structure based drug design and molecular docking approaches and later optimized to have good gastrointestinal absorption and non-toxic. Taken together, this study demonstrate that these six ligands can further be tested under *in-vivo* and *in-vitro* condition for prediction of new drugs against allergic diseases.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENT

The Author acknowledge the conducive study area provided by the Deparment of Biotechnology Sharda University Greater Noida, India and also support from colleague Mustafa Alhaji Isa.

ABBREVIATIONS

TSLP: Thymic stromal lymphopoietin; **DC's:** dendritic cells; **PAR2:** Protease Activated Receptor Type 2; **NCBI:** National Center for Biotechnology Information; **DOPE:** Discrete Optimized Protein Energy; **SOPMA:** Self-optimized Prediction Method with Alignment; **RMSD:** Root Mean Square Deviation; **WlogP:** Water Partition Coefficient; **TPSA:** Topological Polar Surface; **P-gp:** P-glycoprotein; **GI:** Gastrointestinal; **BBB:** Blood Brain Barrier; **LogKp:** Skin permeation coefficient.

REFERENCES

1. Liu YJ, Soumelis V, Watanabe N, Ito T, Wang YH, De Waal Malefyt R, *et al*. TSLP: an epithelial cell cytokine that regulates T cell differentiation by conditioning dendritic cell maturation. Annu. Rev Immunol. 2007;25:193-219.
2. Rochman Y, Leonard WJ. Thymic stromal lymphopoietin: a new cytokine in asthma. Curr Opin Pharmacol. 2008;8(3):249-54.
3. You H, Li R, Wei C, Chen S, Mao L, Zhang Z, *et al*. Thymic Stromal Lymphopoietin Neutralization Inhibits the Immune Adjuvant Effect of Di-(2-Ethylhexyl) Phthalate in Balb/c Mouse Asthma Model. PloS one. 2016;11(7):e0159479.
4. Reche PA, Soumelis V, Gorman DM, Clifford T, Liu MR, Travis M, *et al*. Human thymic stromal lymphopoietin preferentially stimulates myeloid cells. J Immunol. 2001;167(1):336-43.
5. Ziegler SF, Artis D. Sensing the outside world: TSLP regulates barrier immunity. Nat Immunol. 2010;11(4):289-93.
6. Saenz SA, Taylor BC, Artis D. Welcome to the neighborhood: epithelial cell-derived cytokines license innate and adaptive immune responses at mucosal sites. Immunol Rev. 2008;226(1):172-90.
7. Liu YJ, Soumelis V, Watanabe N, Ito T, Wang YH, de Waal Malefyt R, *et al*. TSLP: an epithelial cell cytokine that regulates T cell differentiation by conditioning dendritic cell maturation. Annu Rev Immunol. 2007;25:193-219.
8. Geourjon C, Deleage G. SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics. 1995;11(6):681-4.
9. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, *et al*. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605-12.
10. Šali A, Potterton L, Yuan F, van Vlijmen H, Karplus M. Evaluation of comparative protein modeling by MODELLER. Proteins: Structure, Function, and Bioinformatics. 1995;23(3):318-26.
11. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. J Appl. Crystallogr. 1993;26(2):283-91.
12. Colovos C, Yeates TO. ERRAT: An empirical atom-based method for validating protein structures. Protein Sci. 1993;2:1511-9.
13. Norgan AP, Coffman PK, Kocher JP, Katzmann DJ, Sosa CP. Multilevel parallelization of Auto Dock 42. J Cheminform. 2011;3(1):12.
14. Daina A, Michielin O, Zoete V. Swiss ADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci. Rep. 2017;7:42717.
15. Patel SK, Prasanth Kumar S, Jasrai YT, Pandya HA, Rawal RM. Computational analysis of naturally occurring marine compounds (NOMC) targeting gap junctions and cell adhesive molecules for the identification of anticancer drug targets. J Compt. Intel. Bioinf. 2011;4:165-71.
16. Palm K, Stenberg P, Luthman K, Artursson 1 P. Polar molecular surface properties predict the intestinal absorption of drugs in humans. Pharm Res. 1997; 14(5):568-71.
17. Daina A, Zoete V. A BOILED-Egg to Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. Chem Med Chem. 2016;11(11):1117-21.
18. Sander T, Freyss J, Von Korff M, Rufener C. Data Warrior: an open-source program for chemistry aware data visualization and analysis. Journal of chemical information and modeling. 2015;55(2):460-73.
19. Schrödinger L. "Schrödinger Suite 2016." Maestro version 11. 2016.
20. DeLano WL. Pymol: An open-source molecular graphics tool. CCP4 Newsletter on Protein Crystallography. 2002;40:82-92.
21. Kwong E, editor. Oral Formulation Roadmap from Early Drug Discovery to Development. John Wiley and Sons. 2017.
22. Dougherty DA. Cation-π interactions involving aromatic amino acids. J Nutr. 2007;137(6):1504S-8S.

Article History: Submission Date : 26-10-2017 ; Revised Date : 09-12-2017; Acceptance Date : 08-01-2018.
Cite this article: Adamu RM, Malik BK. Molecular Modeling and Docking Assessment of Thymic Stromal Lymphopoietin for the Development of Natural Anti Allergic Drugs. J Young Pharm. 2018;10(2):178-82.