In Silico designing and optimization of EPO mimetics ligands targeting EPOR binding

A Major Project dissertation submitted

in partial fulfilment of the requirement for the degree of

Master of Technology

In

Bioinformatics

Submitted by

Neeraj Kumar

(DTU/12/MTECH/398)

Delhi Technological University, Delhi, India

Under the supervision of

Dr Vimal Kishor Singh

Department of Biotechnology
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road,
Delhi-110042, INDIA

CERTIFICATE

This is to certify that the M. Tech. Dissertation entitled ***"In Silico* designing and optimization of EPO mimetics ligands targeting EPOR binding."**, submitted by **NEERAJ KUMAR (DTU/12/MTECH/398)** in partial fulfillment of the requirement for the award of the degree of Master of Engineering, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi), is an authentic record of the candidate's own work carried out by her under my guidance.

The information and data enclosed in this dissertation is original and has not been submitted elsewhere for honoring of any other degree.

Date:

Dr Vimal Kishor Singh
(Project Mentor)
Department of Biotechnology
Delhi Technological University
(Formerly Delhi College of Engineering, University of Delhi)

ACKNOWLEDGEMENT

First and foremost, my sincere thanks to Prof B. D. Malhotra, Head, Department of Biotechnology, Delhi Technological University, for giving me an opportunity to study and work in this prestigious institute.

No words are adequate to express my feeling of profound gratitude to my mentor Dr. Vimal Kishor Singh not only for giving me the opportunity to work under him, but also for his guidance, valuable suggestions and persistent encouragement and generosity which inspired me to submit this work in the present form. He has been responsible for smoothing all the rough edges in this investigation by their constructive criticism and deep insight.

NEERAJ
2K12/BIO/16

CONTENTS

LIST OF FIGURES AND TABLES

List of tables

LIST OF ABBREVIATIONS

EPO - Erythropoietin
EPOR - Erythropoietin receptor
ECD - Extra cellular domain
RBC - Red blood cell
JAK - Janus kinase
EMP1 - Erythropoietin mimetic peptide 1
STAT - Signal transducer and activator of transcription
ABT - Antibody
NCBI - National center for biotechnology information
NLM - National library of medicine
PDB - Protein data bank
VEGF - Vascular endothelial growth factor
FLK-1 - Fetal liver kinase 1
MAB - Monoclonal antibody
NESP - Novel erythropiesis stimulating protein
rHuEPO - Recombinant human erythropoietin

In Silico designing and optimization of EPO mimetics ligands targeting EPOR binding

Neeraj Kumar

Delhi Technological University, New Delhi

ABSTRACT

EPOR is cytokine receptor protein; activation of cell surface EPOR by EPO triggers the intracellular phosphorylation cascade response for differentiation and proliferation of progenitor cells into RBC. Due to kidney failure EPOR does not activate which cause the anaemia disease arises due to low RBC count. EPO is glycoprotein hormone which stimulates proliferation and final maturation of RBC precursors in bone marrow. For curing anaemia use of recombinant EPO has significantly improved the curing capacity of patients but it is inconvenient and expensive. Finding small molecules and peptides to alternative EPO, is popular in recent years. Therapeutic targeting binding site of EPOR to EPO mimetic antibody, deriving oligopeptide and dimer analogues of chemical compound may be promising approach for designing new small and more efficient mimetic for EPOR activation. Dimerization of compounds reported to increase the efficiency of compound to activate its target protein. In this study combinatorial library designed in two classes, first chemical compounds and second peptides. First class of library designed by SMND309 analogues and dimers. Second class of library by peptide designed from interaction sites of EPO mimetic human monoclonal antibody ABT007 to EPOR and previously reported mimetic EMP1 and ERB1-7. SMND-309 is a novel derivative of salvianolic acid B which activates EPO receptor first, and then stimulate JAK2/STAT3 pathway and regulate erythropoiesis. Through the screening of combinatorial library of series of analogues of chemical compound SMND309 and peptidic mimetics, efficient mimetics are reported with the results of docking and physicochemical properties. Mimetics are found showing good docking results -9.968 and -5.98 of chemical compound and peptide respectively as compared to reported mimetics SMND309, EMP1 and ERB1-7. Resulting Chemical compound evaluated for binding affinity and found with high binding affinity -78.37 and peptide mimetic found stable and hydrophilic in nature.

Keywords- Erythropoietin, Erythropoietin receptor, Mimetic, Library designing, Docking

Introduction

Anaemia is an abnormal reduction in red blood cells. It occurs from red blood cell degeneration, bleeding or insufficient erythrocytes production. Erythropoiesis is process of making red blood cells. If the body needs more oxygen for instance, the kidney triggers the release of the hormone erythropoietin (EPO), a glycoprotein hormone acts in the bone marrow to growth of erythroid progenitors cell to increase the formation of red blood cells.

For curing anemia disease, in some cases iron is administered through muscular injections or intravenously and transfusions are used to replace blood loss due to injuries and during certain surgeries however, with these techniques side effects and infections are associated. Later recombinant human erythropoietins are available to EPOR activation and erythropiesis regulation, but using recombinant protein is very expensive and laborious process. Hence study required to find out potent mimetic and can be available conveniently. In 1996 Wrighton NC et al reported a 20 amino acids sequence EMP1 which activates EPOR, and in similar fashion to EMP1, ERB1-7 reported which also involved in dimerization of the EPOR and its activation. They reported mimtics based on idea that EPO is a 165 amino acid glycoprotein but its minimal binding region or functional epitope only involved in interaction with its specific receptor. Hence researcher tried to find out short region of amino acid responsible in activation. Therapeutic targeting of binding sites of EPOR to EPO mimetic antibody, deriving oligopeptide and chemical compound analogues may be a promising approach for designing new smaller and more efficient mimetic for EPOR activation.

Known mimetics EMP1, ERB1-7, minimal EMP1 activates EPOR and signalling cascade but these mimetic are reported not to be potent as EPO for erythropoiesis. It is required for finding mimics which can activate EPOR efficiently. Dimerization of 2 EPO mimetic peptides strongly increased their activity for its specific receptor EPOR [Wrighton NC et al 1997 et al 1997]. On this basis, attempt is made to design, combinatorial library with chemical compound SMND309 and peptides derived from interaction site of antibody ABT007 and EPOR and reported mimetic EMP1and ERB1-7.

SMND-309 is a novel derivative of salvianolic acid B which activates EPO receptor first, and then stimulate JAK2/STAT3 pathway and regulate erythropoiesis. Human agonist antibody ABT007 fab fragment binding to EPOR extracellular domain and activates it for the regulation of erythropoiesis.

Here, in this study combinatorial library designed with dimerization of chemical compound SMND309 and the peptides mimetic derived from ABT007 and EPOR complex interaction site and known mimetic EMP1, ERB1-7. Combinatorial library

screened for efficient EPO mimetic which can activate the EPOR and induce the down regulation of RBC formation. This study involves screening of library by docking analysis using Glide docking (Maestro server) for chemical compounds and HEX 6.3 for peptide mimetics. Lead mimetic are further validated by physicochemical properties. Determining the binding affinity of mimetic chemical compound determined by using Prime-GBSA. Peptide mimetic properties molecular weight pI and hydrophobic nature determined by using Peptide protein calculator and Cello predictor.

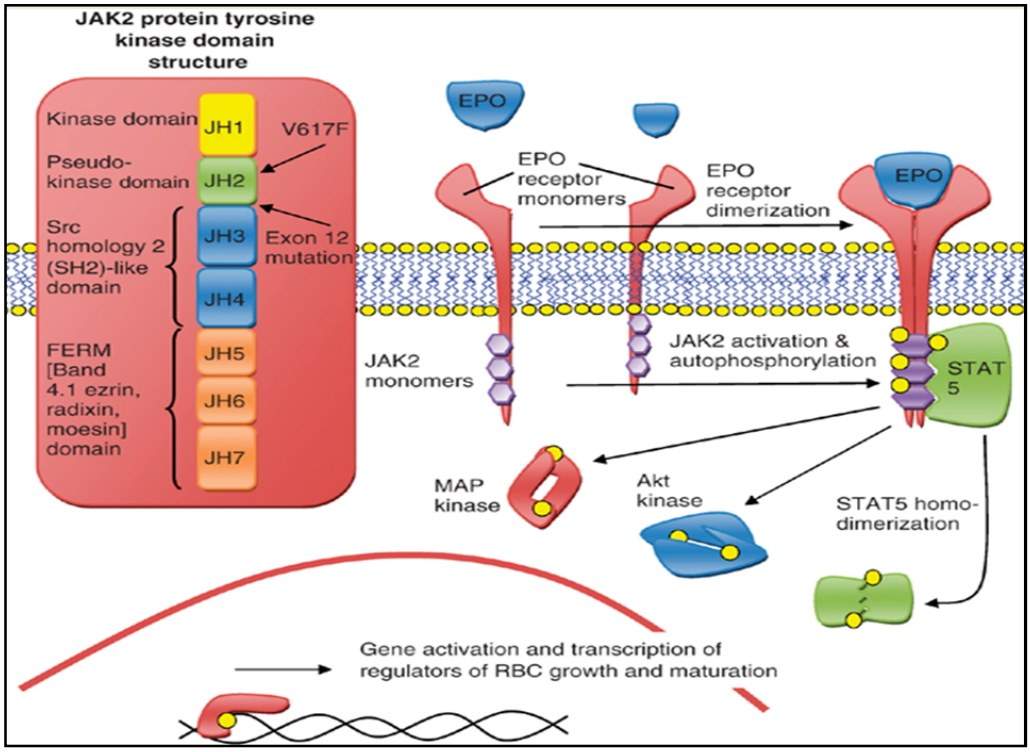


Fig 1. EPO binding to its receptor EPOR and its activation

Review of literature

Severe anaemia that accompanies chronic renal failure was managed by regular blood transfusions, as often as every 2-3 weeks transfusion is required. EPO is a glycoprotein hormone which stimulates proliferation and final maturation of RBC precursors in bone marrow. EPOR is synthesized in predominately in the kidney [Faulds D, sorkin EM et al 1989, Lappin TR, Rich IN et al 1996, Lacombe C, Mayeux P et al 1998, Maiese K, Li K et al, 2005]. EPOR is a member of the cytokine receptor super family and members share similar structural motifs. Cytokine receptors signalling efficiency depends mainly on receptor orientation [Syed R S et al 1998] and mechanism of signal transduction in which activation is believed to be achieved thorough ligand induced homodimerization.

For treating anaemia Androgen therapy had been shown to potentiate erythropoiesis [Alexanian R et al, 1967] but the affect was too weak and having side effects. Regular red cell transfusion to patient is having risk and chances of transmissible risks and infection are associated. Major breakthrough that transformed the therapeutic field of anaemia management came in 1977 with the successful purification of small amounts of human EPO from the urine of patient with aplastic anaemia [Miyake T et al., 1977]. In 1983 the gene for EPO human was then isolated and cloned [Lin FK, Suggs S et al., 1985].

In some reports, intravenously iron is administered through muscular injections. Intravenous iron has the advantage of causing less gastrointestinal passage and obstacles. It may be in the form of iron dextran, sodium ferric gluconate complex in sucrose, or iron sucrose are the various form used introduction in cells. In one report it is defined Ferrlecit or Venofer seem to be at equally effective and also safer than iron dextran.

Some side effects different depending on the way of administration of iron and some of side effects includes pain at the site of injection and pain in the vein, metallic taste and flushing [Bennett CL et al 2008].

During certain strategies for curing of diseases transfusions are used to replace blood loss due to any injuries. Blood transfusion used to treat severely anaemic patients with sickle cell disease, thalassemia, myelodysplastic syndromes, or other type anaemia. In some disease condition patient required continuously blood transfusion. Frequent blood transfusions can results in iron overloaded but If untreated, iron overload can lead to heart and liver damage.

Genetically engineered as recombinant human erythropoietin (rHuEPO) which is genetically engineered known as epoetin alfa. Novel erythropoiesis stimulating protein (NESP), also called darbepoetin alfa, stays longer in the blood than epoetin alfa hence requires fewer injections.

Chronic anaemia disease arises due to lower levels of erythropoietin. In order to avoid receiving blood transfusions Injections of recombinant erythropoietin can help in the increasing of the number of erythrocytes. Erythropoietin is used to treat anaemia. This erythropoietin drug can cause side effects like including blood clots [Rizzo JD et al 2007].

Use of recombinant EPO for cure of anaemia has significantly improved the curing capacity of patients but despite the enormous success, therapy is inconvenient and expensive. Hence it is required to find out the efficient and target specific EPO mimetic. Here we take attempt to design mimetics.

The EPO mimetic peptide is small size and production in large numbers relatively easy, but their activity is sill very low that of native EPO [Livnah O, Stura EA et al 1996]. New methods for ligand discovery are based on combinatorial procedures for assembling large numbers of compounds to produce diverse molecules for binding molecular targets with higher specificity and efficiency [Gallop M A et al 1994, Gordon E M et al 1985]. Overcoming the issue, later it was found that dimerization of 2 EPO mimetic peptides strongly increased the activity [Wrighton NC et al 1997, Johnson DL et al 1997]. In the consequent year it is reported that dimerization of EPO proteins results in enhanced erythopoietic activity in vitro and in vivo [Bruno Dalle et al 2001]. Simultaneously with chemical compounds it is reported that designing dimers of small molecules using different linkers enhance the binding efficiency and specificity [Joel Goldberg et al 2001]. We have formed combinatorial library by series of analogue of SMND309 and peptides derived from interaction sites of monoclonal antibody ABT007 and reported mimetic EMP1, minimal EMP1 peptide, ERB1-7.

SMND-309 is a novel derivative of salvianolic acid B. SMND-309 works by controlling the effects of ischemia and reperfusion injury on brain by lowering the infract volume, increasing the survival of neurons, improving neurological function and promotes angiogenesis by increasing the levels of erythropoietin (EPO), erythropoietin receptor (EPOR), phosphorylated JAK2, phosphorylated STAT3, VEGF and VEGF receptor 2 (Flk-1) in the brain [Zhu H et al 2013].

Recently it is found that using HUVEC (EPO receptor) and AG-490 define that SMND-309 first activate EPO receptor then stimulate JAK2/STAT3 signalling pathway, which up-regulates the expression of VEGF. SMND-309 has strong angiogenic activity on HUVEC, which results in the up-regulation of VEGF through EPOR receptor JAK/STAT3 signalling pathways [Du G et al 2013].Which indicates that SMND309 have role in EPOR activation The more specific chemical compound is advantageously as it will bind to the target receptor avoiding any other risks and can pass through the membranes and can be easily excreted out due its small size.

Human agonist antibody ABT007 fab fragment binding to EPOR extracellular domain and activates it to regulate erythropoiesis However, these antibodies activate the EPOR are less efficient so it is required of potent mimetic which mimic the role of EPO [Zhihong Liu et al 2007]. Unique binding site docking with monoclonal antibody determination a possible approach to design mimetic which can mimic the EPO protein role and which in future can innovate to mimetic with better efficient binding to the receptor

EPO is a 165 amino acid glycoprotein but its minimal binding region or functional epitope only involved in interaction with its specific receptor hence scientist tried to find mimetic of small size which can mimic the role of the EPO. Emp1 is 20 amino acids EPO mimetic [Wrighton NC et al 1996]. Consequent study on peptide EMP1 finds that the minimal peptide sequence of 13 amino acids can trigger the activation of EPOR and cascade signalling [Johnson DL et al 1998].

ERB1-7 cyclic peptide of 18 amino acids belongs to EPO mimetic family [McConnell SJ et al 1998]. Recognition of short linear peptide sequences by receptor proteins is important to many essential cellular processes such as signalling, regulation and the formation of protein networks [Neduva V et al 2005]. These peptide mimetics advantageous for activation or EPOR however, not potent. Hence study is required to find out mimetics which can efficiently activate and regulate erythropoiesis.

METHODOLOGY

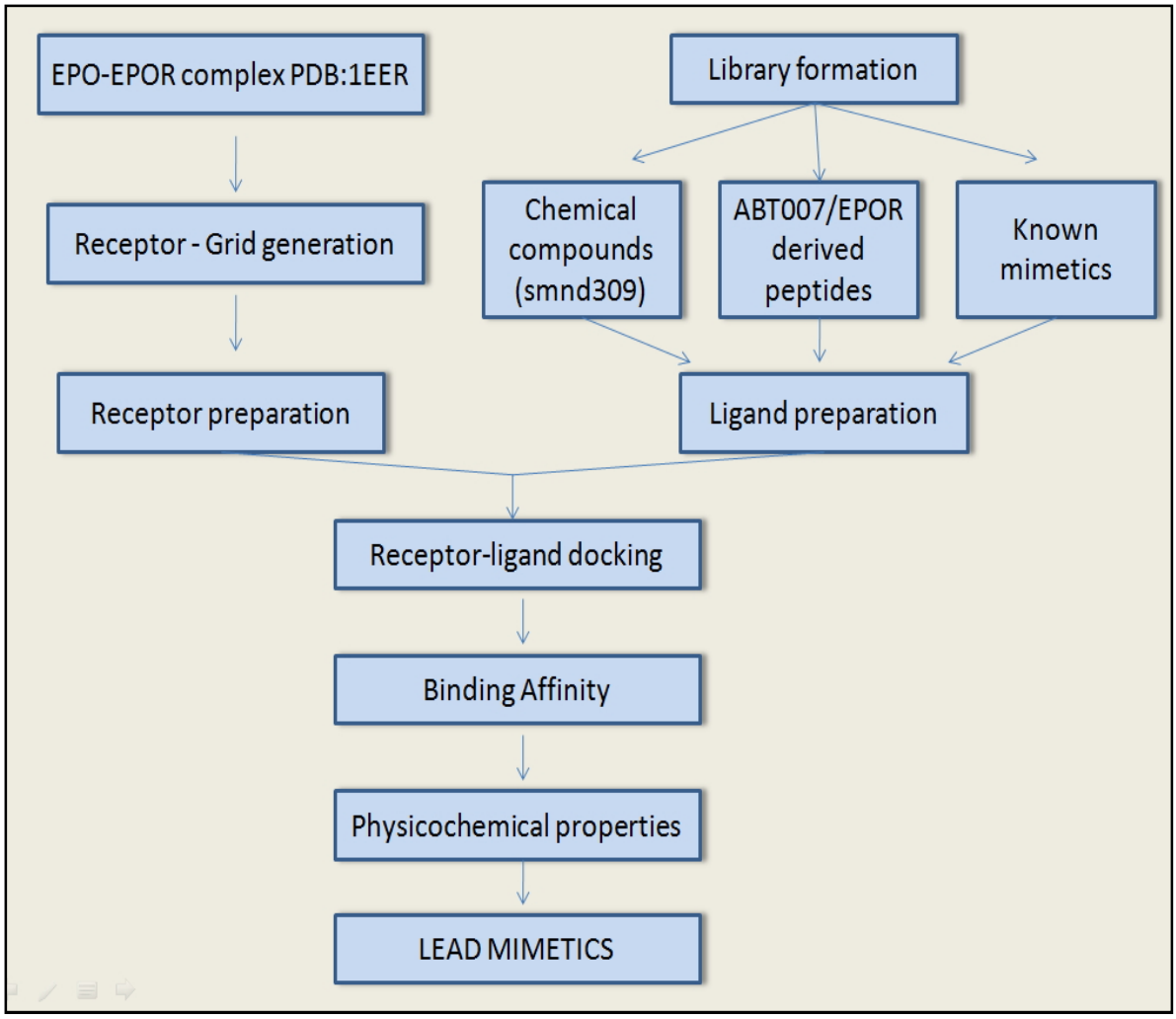
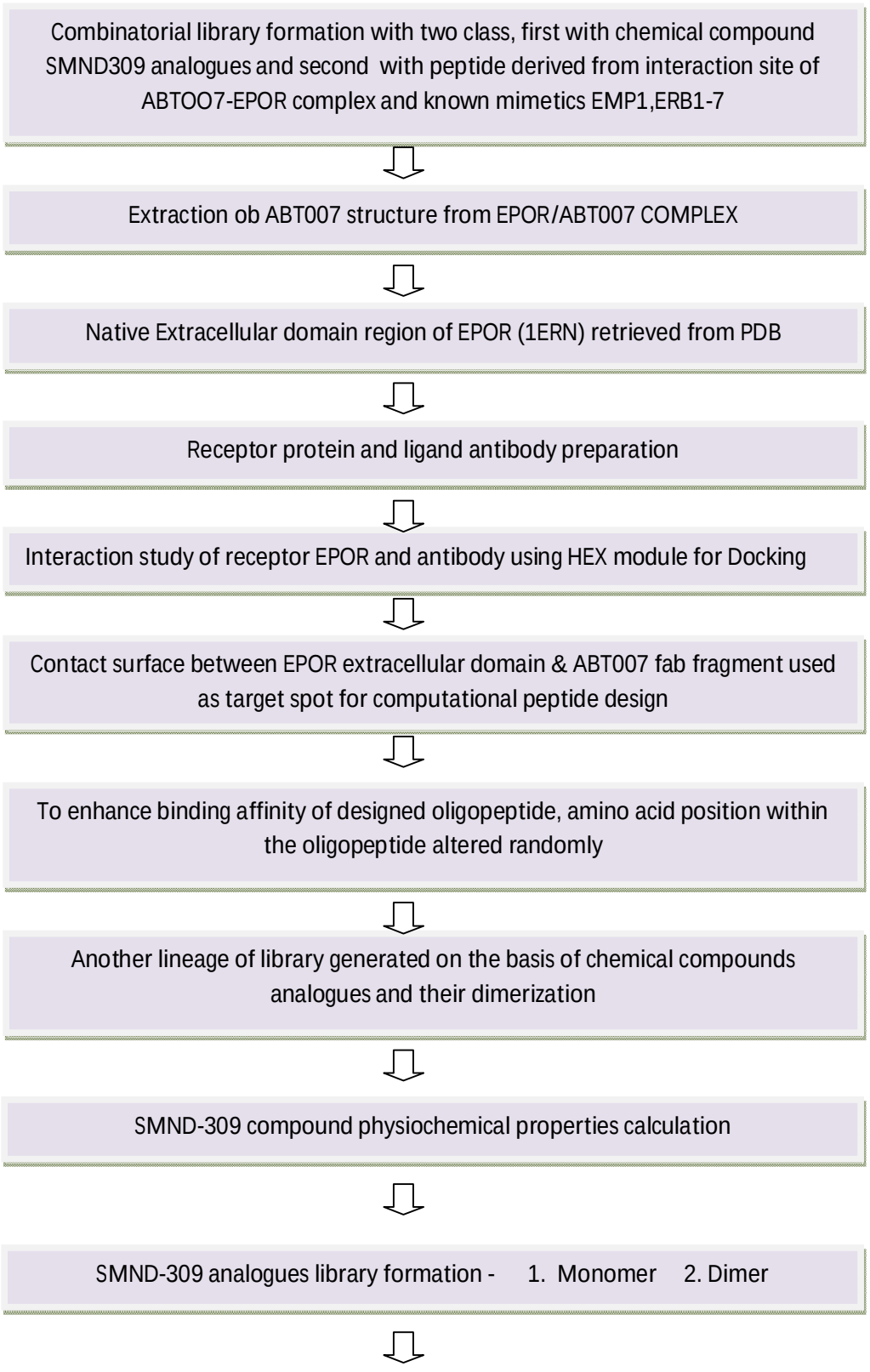
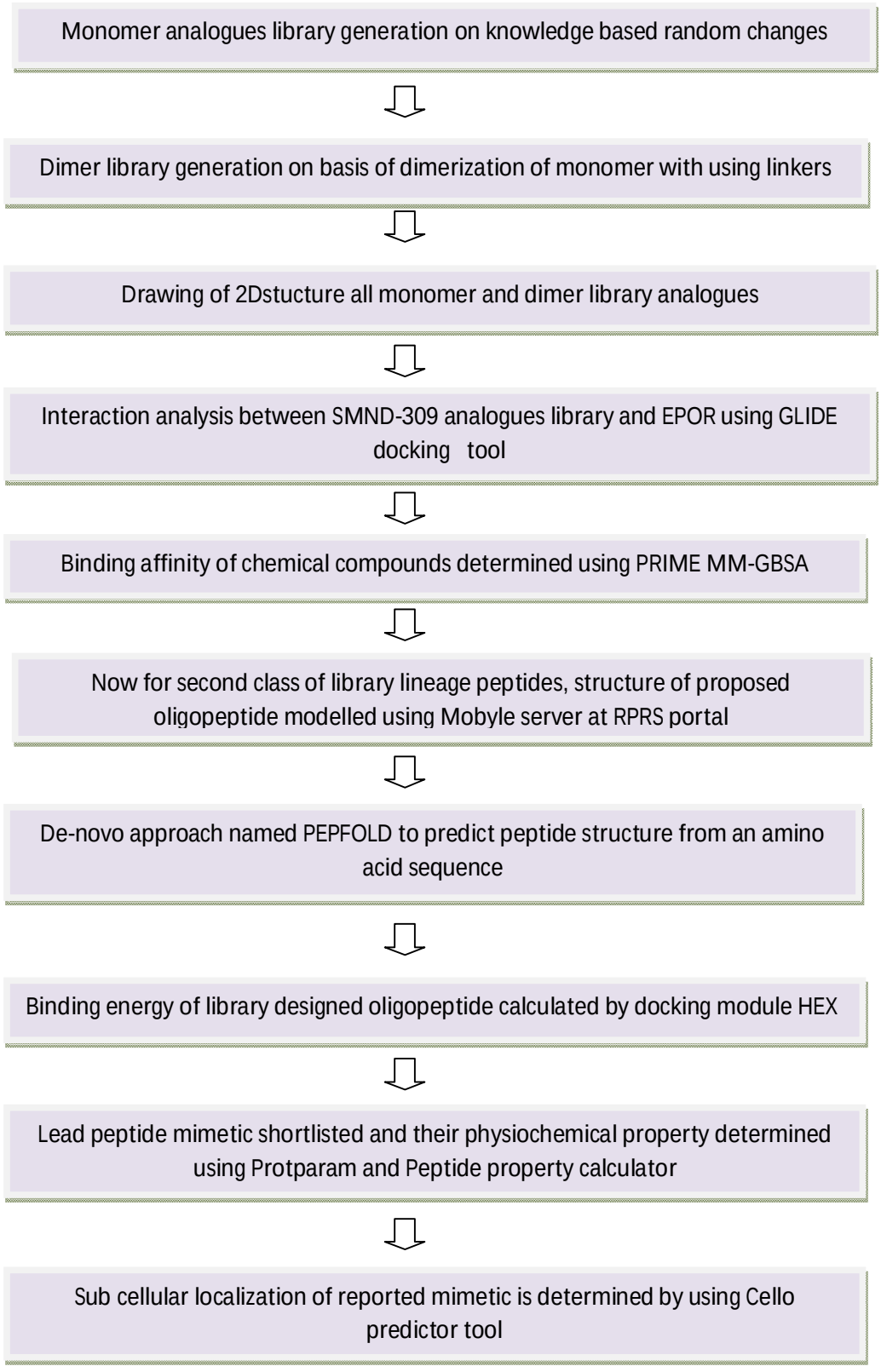


Fig 2. Flow chart of methodology


Combinatorial library formation with two class, first with chemical compound SMND309 analogues and second with peptide derived from interaction site of ABTOO7-EPOR complex and known mimetics EMP1,ERB1-7
Extraction ob ABT007 structure from EPOR/ABT007 COMPLEX
Native Extracellular domain region of EPOR (1ERN) retrieved from PDB
Receptor protein and ligand antibody preparation
Interaction study of receptor EPOR and antibody using HEX module for Docking
Contact surface between EPOR extracellular domain & ABT007 fab fragment used as target spot for computational peptide design
To enhance binding affinity of designed oligopeptide, amino acid position within the oligopeptide altered randomly
Another lineage of library generated on the basis of chemical compounds analogues and their dimerization
SMND-309 compound physiochemical properties calculation
SMND-309 analogues library formation - 1. Monomer 2. Dimer


Monomer analogues library generation on knowledge based random changes
Dimer library generation on basis of dimerization of monomer with using linkers
Drawing of 2Dstucture all monomer and dimer library analogues
Interaction analysis between SMND-309 analogues library and EPOR using GLIDE docking tool
Binding affinity of chemical compounds determined using PRIME MM-GBSA
Now for second class of library lineage peptides, structure of proposed oligopeptide modelled using Mobyle server at RPRS portal
De-novo approach named PEPFOLD to predict peptide structure from an amino acid sequence
Binding energy of library designed oligopeptide calculated by docking module HEX
Lead peptide mimetic shortlisted and their physiochemical property determined using Protparam and Peptide property calculator
Sub cellular localization of reported mimetic is determined by using Cello predictor tool

1. Library formation

Combinatorial library is formed based on the dimerization of compounds and sequences. In this study library is formed in two classes one is by analogues of SMND309 compound and second predicted peptide sequences on the basis of interaction sites of antibody ABT007-EPOR, and dimerization of known mimetics. SMND309 compound with C18H14O8 formula is a soluble compound in saline water (hydrophilic) in nature. Hydrophobic amino acids play important role in EPOR activation and so increasing hydrophobicity in chemical compound more hydrophobic compound added to SMND309. For dimerization of chemical compounds we have used different linkers. In combination to linker some random changes are done to designed whole library.

Antibody ABT007 binds to EPOR and activates it. Interaction site between EPOR and mimetic antibody ABT007 determined by docking is used for library formation. Light chain and heavy chain residues interacting with the receptor active site with high affinity were selected for the designing of new mimetic oligopeptide. Random changes of amino acid at different location are done for making the various combinations of oligopeptide library.

EMP1, ERP, ERB1-7 are known mimetic which activates EPOR phosphorylation. But it is reported that these mimetic peptides are not efficient in the activation. Amino terminal dimerization of peptide mimetic can increase specificity and binding to the target receptor hence we generate dimer of these peptides and random changes for designing the library.

2. Receptor protein preparation-

EPO binds to EPOR on extra cellular domain region (ECD). The 3D structure of EPOR retrieved from RSCB protein data bank (PDB ID 1ERN). Binding site region of EPOR protein is selected. Native EPOR protein was prepared for docking analysis using the Maestro server. EPOR protein chains were identified for finding any duplication of any chain, if present one of the chains will be removed. Water molecules and any other unnecessary ligand molecule removed and hydrogen was added to the protein molecule then protein structure is selected for energy minimization and followed by optimization for docking analysis.

3. ABTOO7 antibody preparation-

Human antibody ABT007 mimics the role of EPO for activation and dimerization of EPOR. 3D structure of the antibody was retrieved ands saved in PDB format. For docking analysis, antibody was prepared by extracting water and unwanted attached molecules. PDB file of antibody is imported to Maestro workspace and prepared.

4. Interaction analysis of EPOR and ABT007-

The binding site of EPOR and ABT007 antibody is determined using HEX 6.3 docking module. ABT007 PDB file is retrieved from its complex 2JIX by splitting the chains and then export the antibody in PDB format. Prepared receptor protein and antibody imported in GLIDE and allowed to run. Resulting PDB file exported to Maestro server and the binding site of EPOR and determined and which binding sites were used to predict peptide sequences.

5a. Ligand preparation of chemical compounds-

Library class one lineage of dimer chemical compounds prepared for the docking. SMND309 have been reported for activation of EPOR and JAK–STAT signalling pathways, which is important for RBC formation. SMND309 structure retrieved from NCBI. SMND309 is reported to activate the EPOR cascade signalling. 2D Structure of analogues and dimer molecules ligands are drawn using the Chemdraw. Using Chemdraw tool and ligands were saved in '.mol' format. Library chemical compounds are prepared by removing water or any other unwanted binding molecules. Ligand preparation of library is done by using Ligprep module of Meastro server. First imported to the ligand molecule on workspace and project saved and allowed to run.

5b. Ligand preparation of peptide sequences

Library class two lineage peptides first the 3D structure was predicted by Pepfold peptide structure prediction server. Different peptides imported in workspace of Pepfold and saving the project name and sequence and format allowed to run. Pepfold returns in PDB file of target peptides. PDB file of ligands were prepared using Prepwizard removing water molecules and unwanted ligands.

6a.Interaction analysis of EPOR and combinatorial library of chemical compounds-

Combinatorial library was studied for interaction analysis for ligand and receptor EPOR. For first class lineage of library (chemical compounds) used the GLIDE docking module for the interactive study of erythropoietin receptor and designed mimetic library. GLIDE is commercial available molecular docking software which is used to analyze molecule interaction between the protein molecules and ligand molecules [R.A. Friesner et al., 2004]. This required receptor protein and ligand molecule first prepared to remove duplication of chains and any other ligands. For docking chemical molecules saved in.mol file is imported into the workspace of GLIDE. Ligand chemical compound before is require to prepare before molecule interaction analysis. LigPrep module is used to prepare chemical compound is prepared. Next for covering whole binding interface groove of receptor protein grid generated and then receptor ligand allowed to run. The

output of the binding energy/efficiency between ligand and receptor protein retrieved as Glide score, which gives binding energy value and interaction site.

6b. Interaction analysis of EPOR and combinatorial library of peptides-

Secondly, for second class of library (peptide sequences) binding energy is analyzed by using the docking tool HEX 6.3. This required prepared receptor and ligand in PDB format. Receptor and ligand PDBs were imported in the workspace of HEX and grid for domain of interaction are selected and allowed to run. HEX results in a number of best PDB of interacting receptor and ligand and log files. This gives the binding energy Etotal score.

7. Physiochemical and biological activity prediction-

Lead mimetic sorted out with better binding efficiency, then known mimetics. Selected mimics are further analyzed for physicochemical properties. Chemical compound evaluated for the binding affinity using PRIME–MM-GBSA module. Binding affinity indicates the efficient binding of ligand to receptor. Shortlisted peptide mimics sequence biological properties and physiochemical properties determined by using Peptide protein calculator, Protparam and Cello server for determine peptide position localization in humans.

RESULTS

1. Library formation-

1a. Peptide mimetic predication

Peptide sequences were predicted on the basis of binding site of ABT007 antibody and EPOR. First protein is prepared and then antibody prepared and both allowed docking and binding sites were identified and random sequences from the binding sites were predicted.

Receptor Preparation:

EPOR 3D structure Analysis- EPO-EPOR complex PDB 1EER was imported to Maestro and it was found that EPOR binding to EPO with two different sites. EPOR is having two similar chains A and B chain binding to EPO. Receptor ligand binding critically depends on the orientation of interaction. Both chain A and B of EPOR have similar amino acid sequences so for docking analysis duplicated chain B chain is removed from the receptor protein. Protein is prepared using Prepwizard by importing receptor file and minimizing and optimizing the receptor protein.

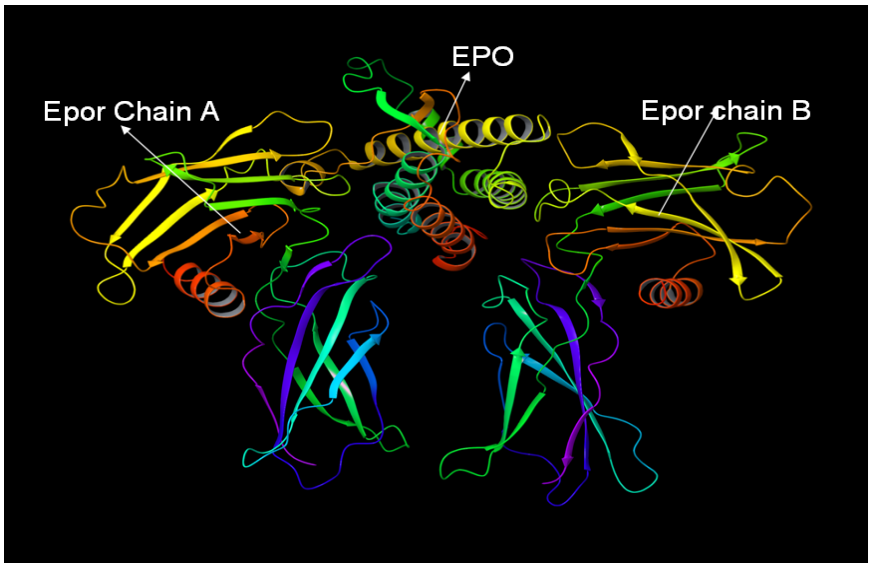


Fig 3. 3D sturucture of EPO-EPOR complex

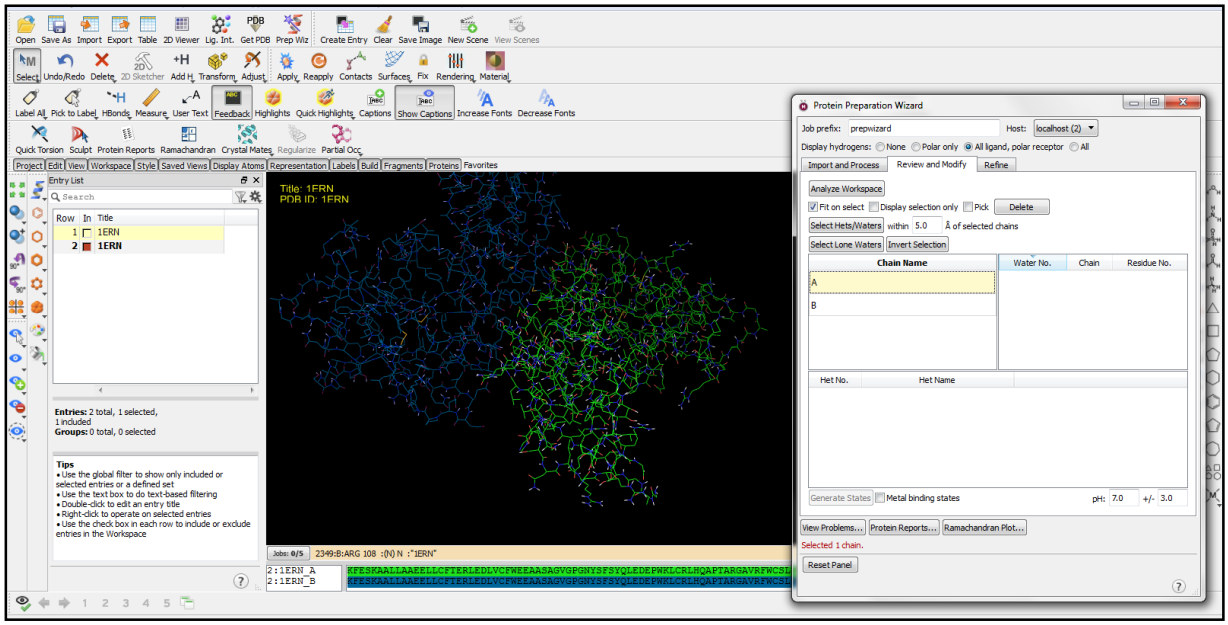


Fig. 4 Maestro interface showing EPOR chain A and chian B

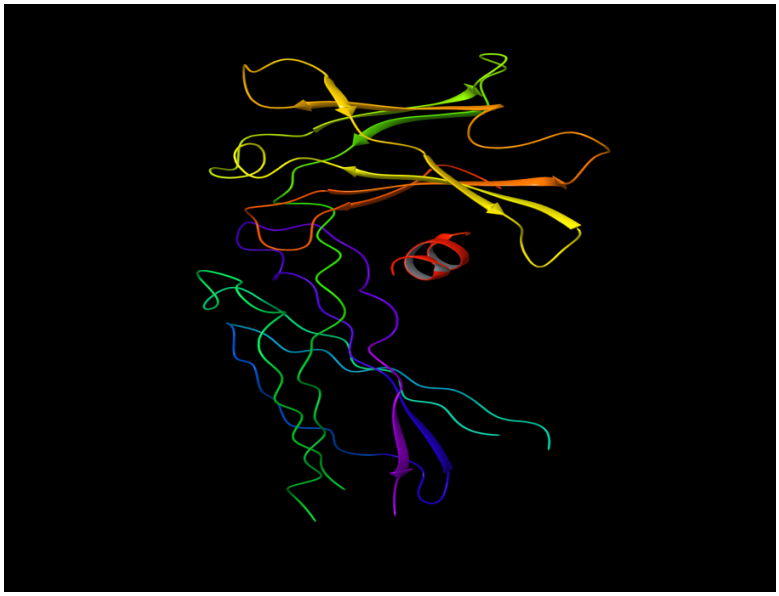


Fig. 5- EPOR 3D structure,representing single chain A, after removing duplicate chan B and water molecules

Binding Site Identification:

PDB 1EER complex between EPO and EPOR retrieved (Livnah O et. al. 1999). The complex is analyzed in Maestro work space for identifying the interacting residues responsible for binding. EPOR interact with EPO with two different sites 1 and 2. First site of EPOR hydrophobic in nature mainly due to Phe93 which are responsible for non polar interaction, its side chain also consisted of hydrophilic amino acids also involved in the interaction with the ligand.

Fig.6- intercting EPO-EPOR and showing binding sites

Identified residue which are interacting were - Leu59, Glu60, Asp61, Thr90, Ser91, Ser92, Phe93, Val94, Pro95, Leu96, Ileu113, His114, Ileu115, Asn116, Ser152, His153, Glu202, Pro203, Ser204, Phe205.

Interaction analysis of EPOR and ABT007-

Prepared receptor protein EPOR and ABT007 antibody were docked using HEX server. PDB files of receptor and ligand imported and interaction domain and number out coming result 10 were selected. Best resulting PDB of ligand and receptor analyzed to determining binding sites. Light chain find to interacting to EPOR with Arg25, Arg26, Asn29, Glu31, Ala32, Glu33. Heavy chain find to interacting with Tyr27, Asn29, Ser34, Tyr48, Tyr90, Leu91. Random combination of these amino acids made and peptide sequences predicted.

2. Ligand preparations-

2a. Ligand preparation of chemical compounds

First chemical compounds were drawn on ChemDraw and then compounds file was saved in '.mol' format. Ligand file was imported to Maestro and using Ligprep, ligand is prepared for further analysis**.**

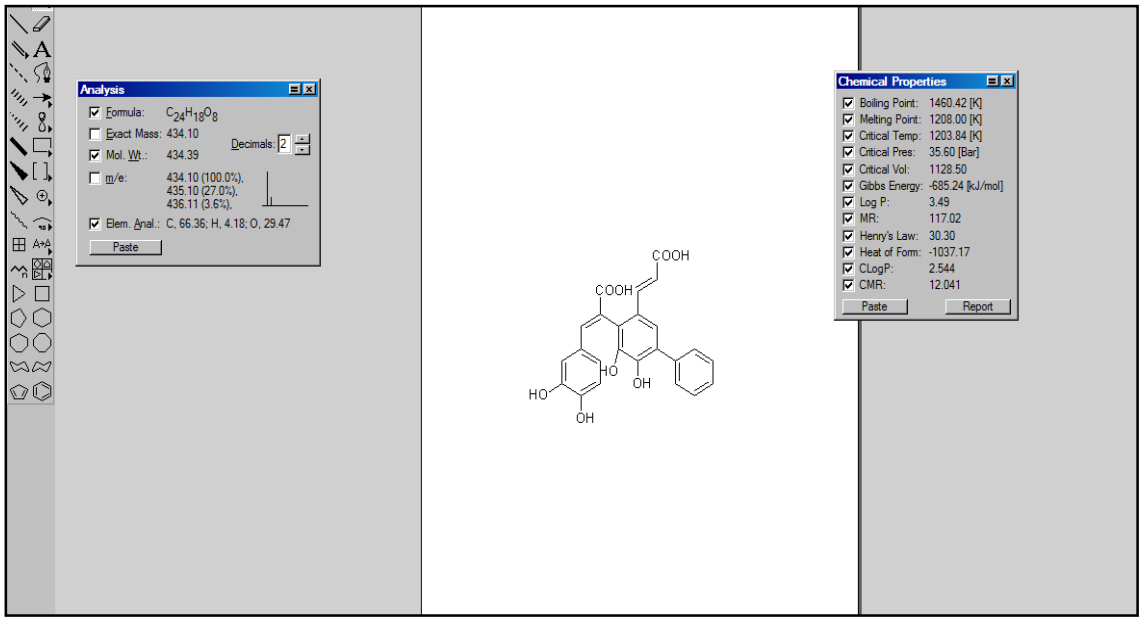


Fig. 7- Ligand drawn on work spcae of Chemsketch

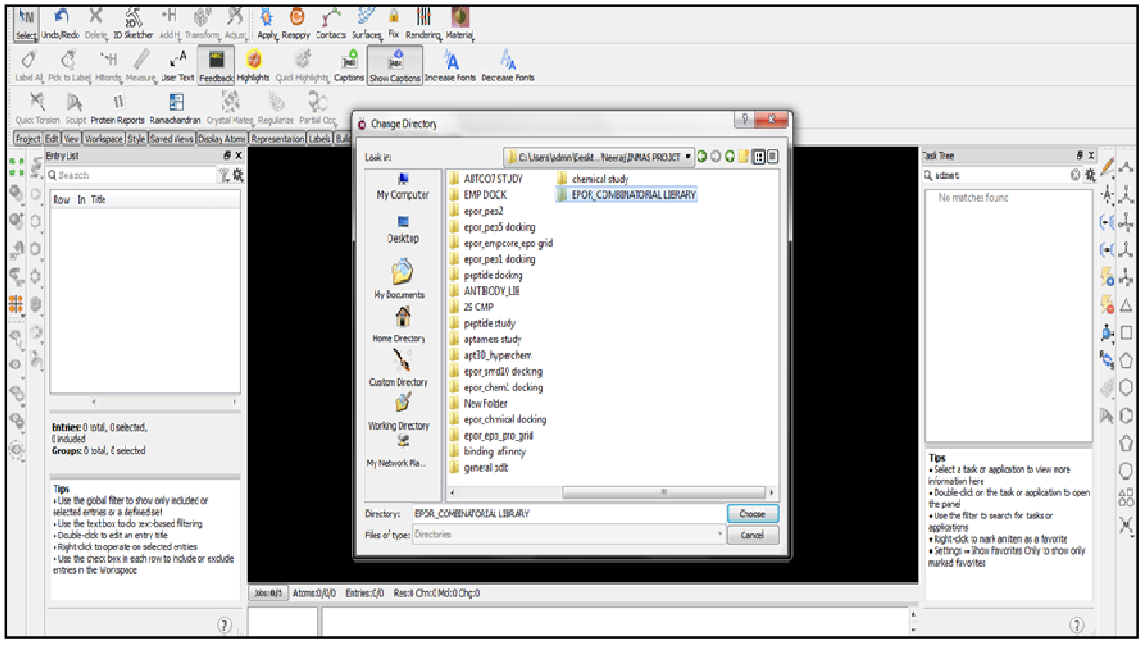


Fig.8- Change directory panel of Maestro for saving the project

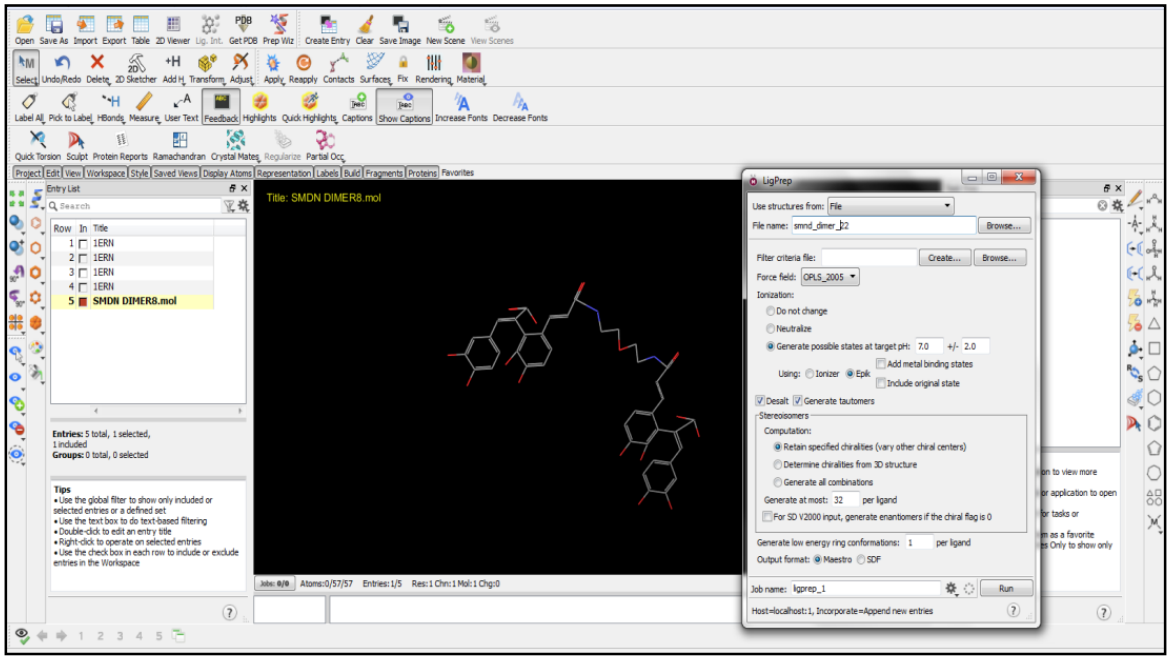
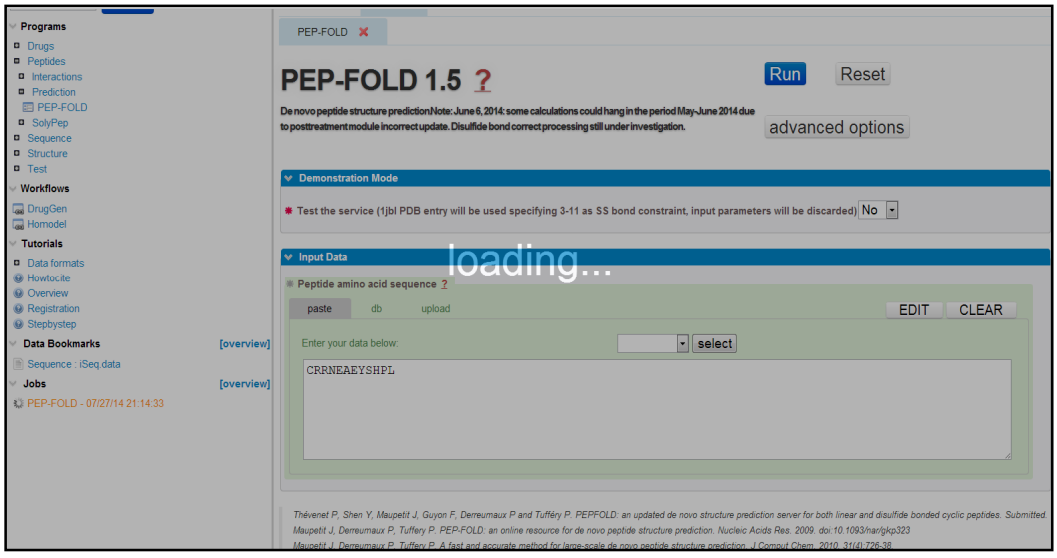


Fig.9- Ligprep panel for ligand preparation

2b. Ligand preparation of peptide sequences

3D structures of peptides are predicted by Pepfold. It is de novo peptide structure prediction tool. Peptide sequences were paste on the workspace and allowed to run which gives 3D structure in PDB format. PDB files imported to Maestro and Prepwizard used to prepare structures.



.

Fig.10- Web server Interface of Pepfold: To genertae 3D strcture of peptides

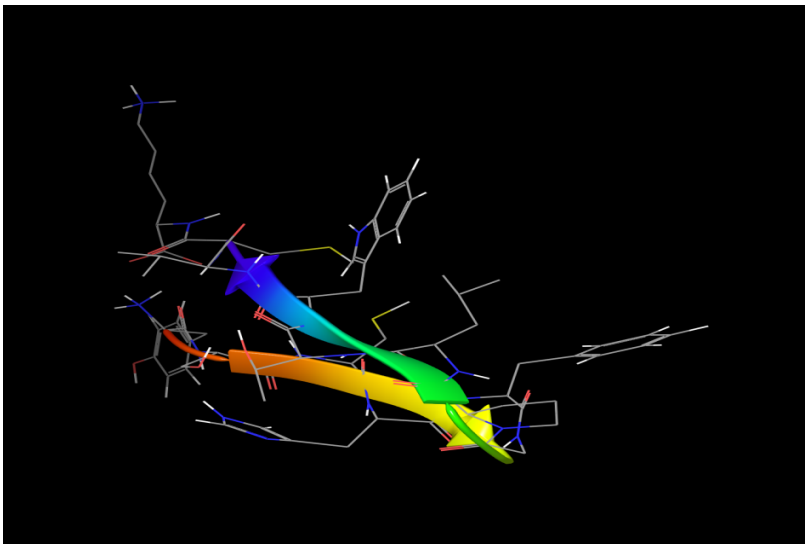


Fig.11- 3D stucture for EMP1 retrived from Pepfold

3a.Interaction analysis of EPOR and combinatorial library of chemical compounds- Chemical analogues were studied for interacting with the EPOR using Glide module. Receptor protein 1ERN and ligand imported to the workspace. Grid used to select interacting domain of protein browse to panel and allow to run. GLIDE returns in Glide score, lipophilic fraction, H-bond and expos penalty. Glide score is based on binding energy value of ligand to receptor. Lipophilic fraction indicated the hydrophobicity of the chemical compounds and Expos penalty is obstacle which arises during interaction of ligand with receptor.

Chemical compounds	Glide Score	Expos penalty	Activity	H-bond	Electrostatic rewards	Lypophillic term & fraction
SMND309	-4.080	0.2	-4.1	-1.9	-0.6	-1.4
Monomer-1	-5.113	0.4	-5.1	-2.0	-0.8	-1.8
Monomer-2	-3.686	0.2	-3.7	-1.8	-0.5	-1.5
Monomer-3	-4.867	0.4	-4.9	-2.4	-0.8	-2.3
Monomer-4	-5.063	0.1	-5.1	-3.4	-0.6	-1.2
Dimer-1	-5.418	0.4	-5.4	-4.1	-0.9	-1.3
Dimer-2	-6.089	0.5	-6.0	-3.7	-0.8	-1.6
Dimer-3	-7.551	0.6	-7.5	-4.0	-1.2	-1.9
Dimer-4	-6.413	0.5	-6.4	-4.3	-0.5	-2.9
Dimer-5	-5.081	0.5	-5.1	-1.6	-1.9	-2.1
Dimer-6	-5.607	0.4	-5.6	-2.4	-1.4	-2.2
Dimer-7	-7.066	1.4	-7.1	-4.5	-1.3	-2.9
Dimer-8	-5.879	1.3	-5.9	-4.4	-1.2	-1.8
Dimer-9	-6.727	0.6	-6.7	-2.3	-1.6	-3.4
Dimer-10	-5.957	0.8	-6.0	-3.1	-1.3	-2.3
Dimer-11	-7.973	2.0	-8.0	-5.4	-3.5	-1.2
Dimer-12	-6.508	0.2	-6.5	-1.9	-1.7	-3.0
Dimer-13	-7.166	1.3	-7.2	-5.7	-0.9	-2.0
Dimer-14	-8.445	2.4	-8.4	-5.2	-2.0	-2.7
Dimer-15	-8.621	1.7	-8.6	-4.8	-1.9	-3.6
Dimer-16	-6.709	3.2	-6.8	-5.8	-2.0	-2.3
Dimer-17	-9.125	1.6	-9.2	-5.9	-1.4	-3.6
Dimer-18	-9.525	0.7	-9.5	-4.0	-3.2	-3.0
Dimer-19	**-9.968**	**2.0**	**-10.0**	**-4.5**	**-3.5**	**-4.1**
Dimer-20	-8.53	1.4	-8.5	-4.1	-2.0	-3.9
Dimer-21	-7.885	1.6	-7.9	-4.0	-2.0	-3.3
Dimer-22	-9.051	1.9	-9.1	-6.4	-2.0	-2.6
Dimer-23	-6.707	1.4	-6.7	-4.3	-1.1	-2.7
Dimer-24	-4.784	1.9	-4.8	-3.2	-1.2	-2.3

Table 01- Compounds (17, 18, 19, and 22) have shown better binding efficiency then SMND309.

Dimer19

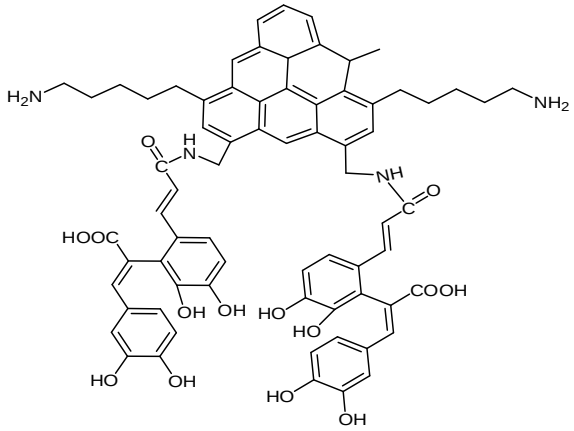


Fig.12- Chemical structure of Dimer 19

Ligand interaction sites are identified using 2D viewer application

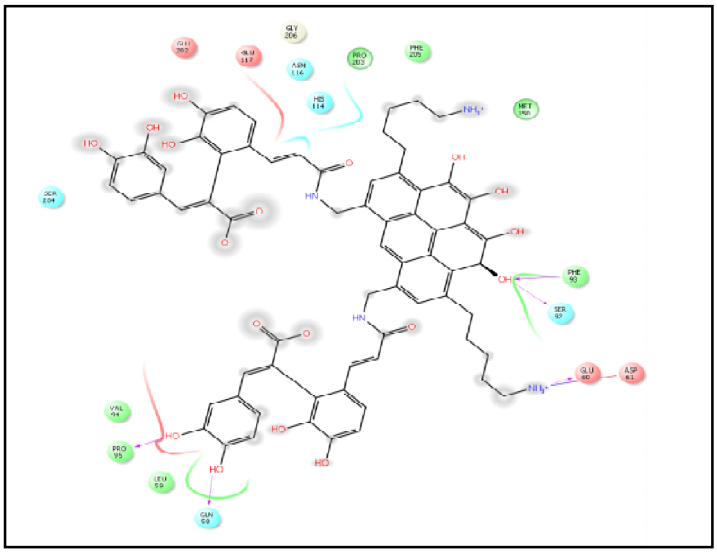


Fig.13- 2D view of interaction sites Dimer 17

This graph representing the plot between Glide score and lipophilic fraction, here this study have shown with increasing in the glide score lipophilic fraction values is increasing. Lipophilic fraction represents the hydrophobic molecules of compound which have importance compound stability and movements through membrane.

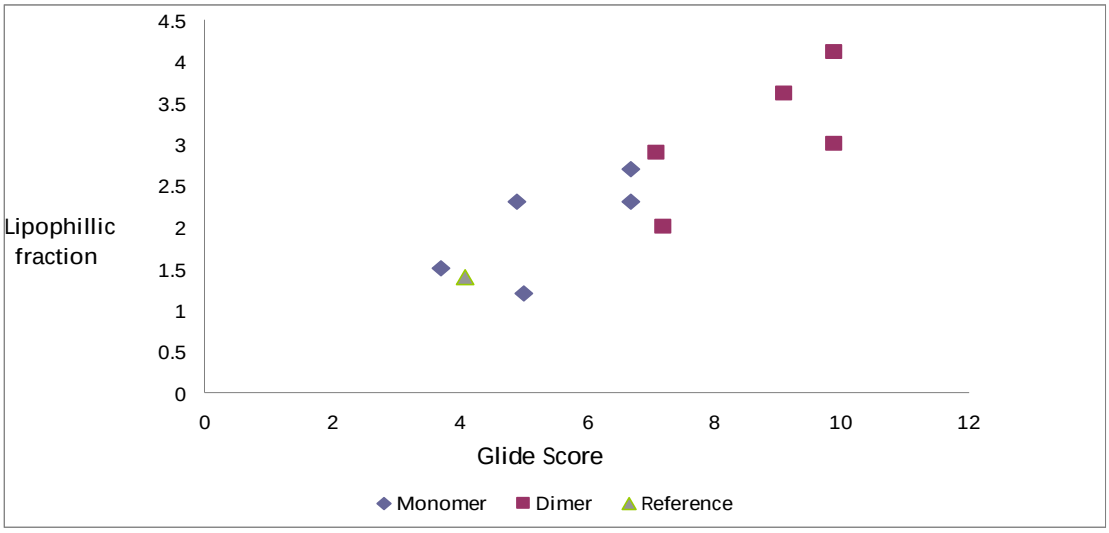


Fig. 14- Plot between Glide score and lipophilic fraction for monomer, dimer and SMND309.

3b. Interaction analysis of EPOR and combinatorial library of peptides-

Combinatorial library class 2 lineage peptides interaction with EPOR analyzed using HEX software. Initially receptor and ligand are prepared for removing water and unwanted lignads then PDBs of receptor and ligand peptides were uploaded to the software. Parameters of interaction analysis calculation device number of results and domain of interaction is selected and then run the programme. Hex returns the Etotal in log file and best PDBs for interaction of receptor and ligand. Here in the table results of peptides library are mentioned**.**

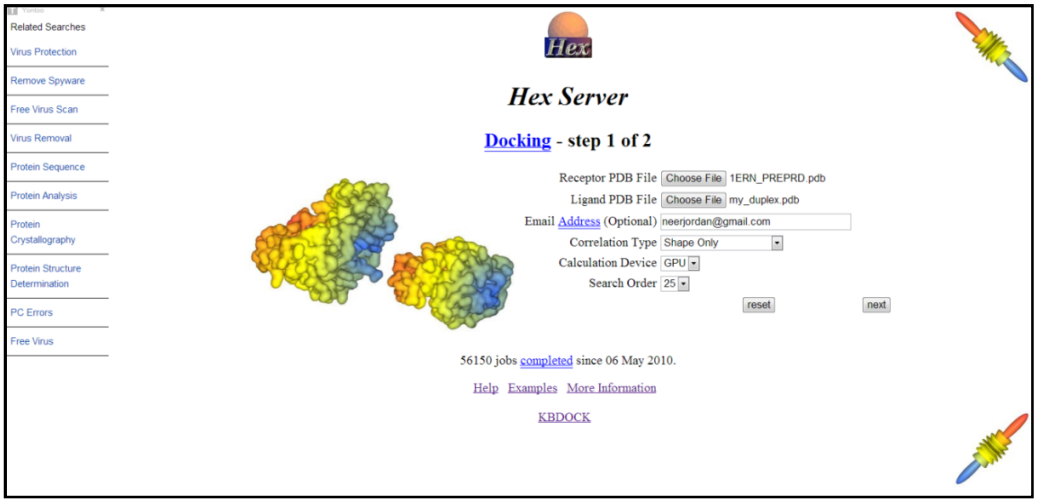


Fig.15- HEX 6.3 home page

Sr. no.	Oligopeptides	Etotal
	Known mimetic	
1	EMP1-GGTYSCHFGPLTWVCKPQGG	-459.4
2	ERB7- DREGCRRGWVGQCKAWFN	-426.6
3	YSCHFGPLTW	-402.5
4	YSCHFGPLTWVCK	-437.6
	Unknown mimetic	
	Light chain	
6	CRRNEAE	-336.4
7	CCRRNEAEC	-390.5
8	RRNEYYAYS	-397.6
9	RREASAHY	-368.7
10	RNEASHYC	-394.9
11	CRRNEASHY	-397.5
12	CRNEAESHY	-427.8
	Heavy chain	
13	CYNSYHYLC	-355.1
14	CCYNSAHYLC	-372.4
15	YYNASYHLL	-359.3
16	CYYNNASHL	-371.7
	Random mimetics	
17	CYSCHLCY	-351.3

18	CCYSCAHL	-351.2
19	CCYSCNNASHL	-386.0
20	CRRNAECLTW	-383.0
21	CRRNEAERRNC	-412.0
22	RREAERRSHL	-425.5
23	CRREAESSYC	-365.5
25	**CRNEAESHYCYYNNASHL**	**-479.7**
26	CYYNNASHLRREAERRSHL	-447.6
27	CRNEAESHYCCYNSAHYLC	-460.9
28	ILVGTLIVLIPVLIVLVFLYWQ	-455.6

Table.02- Etotal score of EPOR and peptide library using HEX.

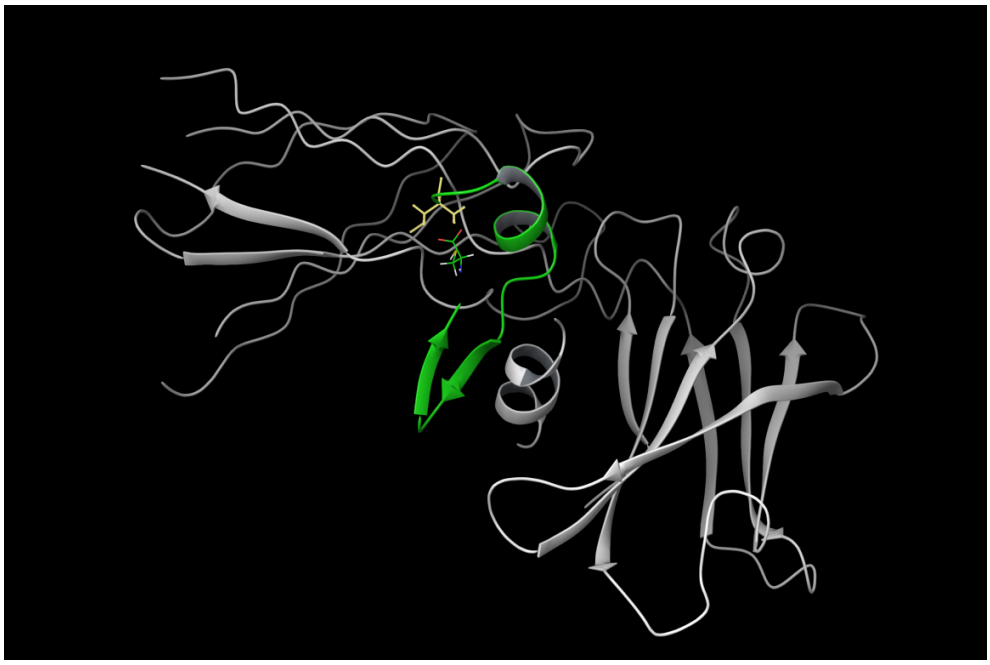


Fig.16- Peptide 25 interacting to its target with EPOR

4. Physiochemical and biological activity prediction-

Binding affinity of chemical compouds

Short listed chemical compound dimer 19, on the basis of Glide score is further evaulated for binding efficieny to its receptor were determined using PRIME-MM-GBSA.

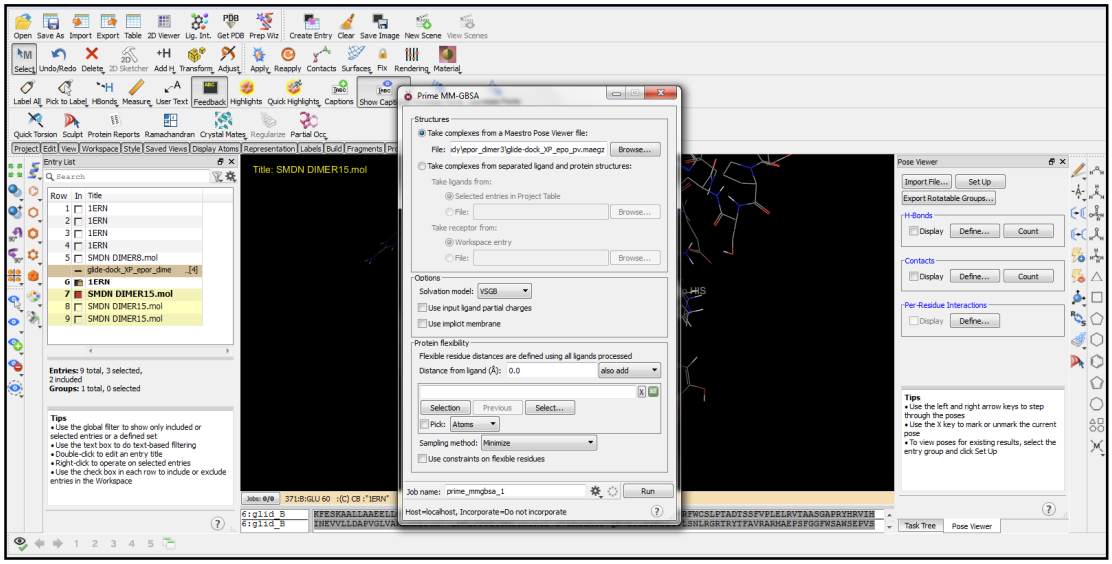


Fig.17- Prime MM- GBSA panel

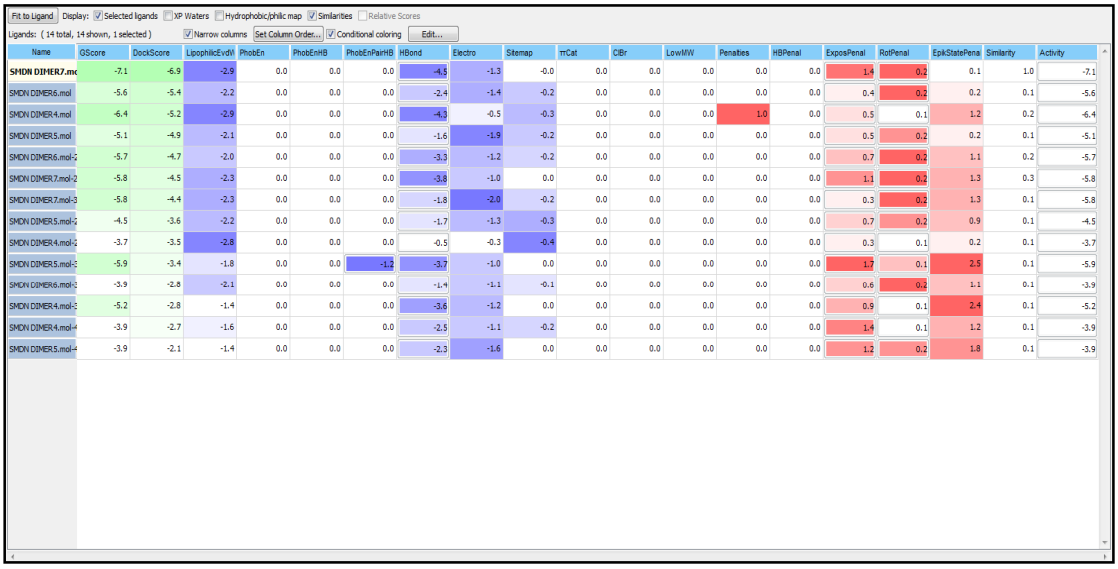


Fig.18- Resulting table of Prime MM- GBSA for chemical compond

CHEMICAL COMPOUND	MMGBSA dG Bind	CHEMICAL COMPOUND	MMGBSA dG Bind
SMND309	-23.245	Dimer 13	-67.344
Dimer 01	-42.369	Dimer 14	-71.302
Dimer 02	-38.124	Dimer 15	-73.132
Dimer 03	-54.063	Dimer 16	-41.165
Dimer 04	-33.509	Dimer 17	-54.278
Dimer 05	-43.663	Dimer 18	-54.409
Dimer 06	-43.024	Dimer 19	-78.374
Dimer 07	-54.756	Dimer 20	-56.169
Dimer 08	-53.163	Dimer 21	-58.760
Dimer 09	-59.576	Dimer 22	-69.215
Dimer 10	-43.106	Dimer 23	-39.964
Dimer 11	-47.214	Dimer 25	-33.824
Dimer 12	-51.840		

Table.03- Binding affinity analysis of chemical compounds library

Physicochemical proprety of peptide

Peptide mimetics 25, found on basis of good binding energy value was further evaluted for physicochemical properties determined using Protein peptide calculator. Which results in finding the molucluer wieght, isoelectric pH, hydrophillic nature and hydrphobicity.

Peptide Sequence:
CRNEAESHYCYYNNASHL

Modifications:
No modifications.

Chemical Formula:
$C_{91}H_{128}N_{28}O_{31}S_2$

Molecular Weight:
2174.30

Isoelectric Point:
6.48

Net Charge

pH

Hydrophilicity Analysis:

Peptide	Charge	Attribute
CRNEAESHYCYYNNASHL	1	basic

Note:

- Red: acidic residues, like D E and C-terminal -COOH
- Blue: basic residues, like R K H and N-terminal -NH2
- Green: hydrophobic uncharged residues, like F I L M V W A and P
- Black: other residues, like G S T C N Q and P
- **Z:** Unrecognized codes are replaced of 'Z'.

Scheme for you choosing suitable solvent:

- for basic peptide, initially try to dissolve the peptide in water; if the peptide does not dissove, try 10% and higher solutions of acetic acid; if it still does not dissove, add TFA(<50ul) to solubilize the peptide and dilute to 1ml with deionized water.

- When input the custom peptide sequence data, please only use the code specified in the **Amino Aicd Code Table**. You can use single-letter-code or multiple-letter-code. If you use mulitiple-letter-code, please use {} to surround each multiple-letter-code.
  e.g. AC{GLY}{ILE}{d-GLY}.
- Select the proper modification for custom peptide synthesis.
- For disulfide bridge, please provide the positions with the format like **5-12, 15-24** (one bridge between positions 5 and 12 cysteine, and another bridge between positions 15 and 24).

Guidelines for solubilizing your peptide:

1. Peptides that are shorter than 5 residues are generally soluble in aqueous media, except in extreme cases where all the residues are very hydrophobic.
2. Hydrophilic peptides containing >25% charged residues and <25% hydrophobic residues also generally dissolve in aqueous media, provided that the charged residues are fairly distributed throughout the sequence. Peptides are generally purified with 0.1% TFA/water ane 0.1% TFA/ACN solvent system.
3. Hydrophobic peptides containing 50% to 75% hydrophobic residues may be insoluble or only partially soluble in aqueous solutions, even if the sequence contains 25% charged resideus. It is best to first dissolve these peptides in a minimal amount of stronger solvents such as DMF, acetonitrile, isopropyl alcohol, ethanol, acetic acid, 4-8M GdnHcl or urea, DMSO (if the sequence does not contain C,W or M), and other similar organic solvents, and then slowly add the solution to a stirred aqueous buffer solution. If the resulting peptide solution begins to show turbidity, you might have reached the solubility limit and it will be futile to proceed. Again, it is important to remember that the initial solvent of choice should be compatible

Fig.19- peptide 25 parameters identification using protein peptide calculator

PARAMETERS	VALUE
Molecular weight	2174.30
Isoelectric point	6.48
Hydrophiilicity	28%
Hydrophobicty	21%

Table.04- Results of peptide 25 using protein peptide calculator

Cello predictor

Subcellular localization of shortlisted peptides is detemined by Cello predictor. it detrmines the localalization of peptides after secratory patheway.

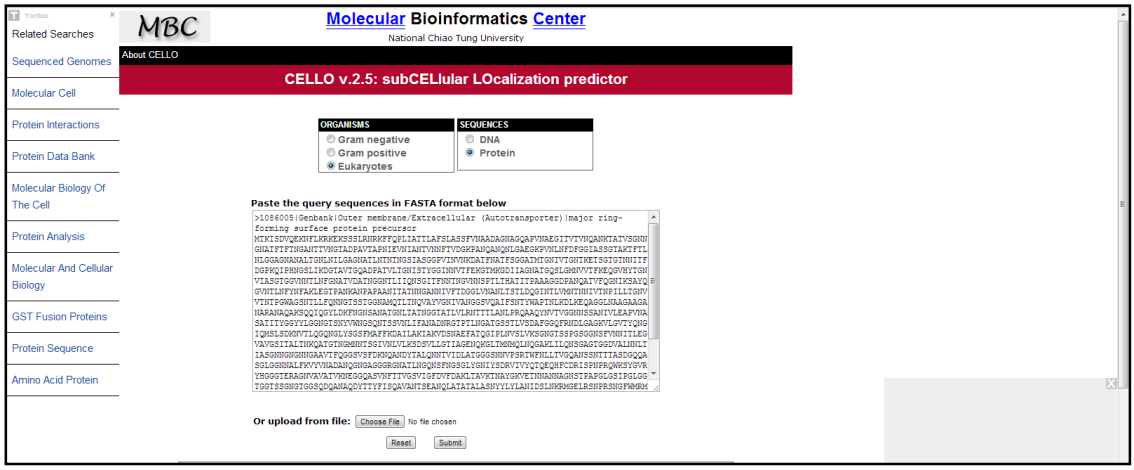


Fig.20- Cello home page

CELLO Prediction-

Nuclear	1.678 *	Peroxisomal	0.029
Mitochondrial	1.269 *	ER	0.025
Chloroplast	0.676	Vacuole	0.018
Extracellular	0.606	Cytoskeletal	0.018
Cytoplasmic	0.509	Lysosomal	0.017
Plasmamembrane	0.142	Golgi	0.013

Table.05- Cello predictor results for peptide 25

Discussion

In this study, Dimerization technique used to design the efficient mimics for EPO. Combinatorial library of dimers designed by series of analogue of SMND309 compound which activates the EPOR and peptide sequence derived from binding sites of antibody ABT007-EPOR and dimmer of known mimetics EMP1, ERB1-7 and minimal peptide sequence involved in EPOR activation. Library tested for interaction analysis with EPOR extra cellular domain region for binding and regulation of erythropoiesis. Library class one lineage consisted of monomer compound with adding hydrophobic chains and dimer of chemical compound with using different linkers. For an interaction study first receptor EPOR extra cellular domain region (PDB 1ERN) was prepared by removing water and any other unnecessary ligand binding. EPO-EPOR active site is analyzed using PDB structure (1EER) of the complex in maestro server. EPOR found to interact with two regions to EPO forming the receptor dimer. Receptor binding sites posses hydrophobic amino acids mainly Phe93 responsible for hydrophobic interaction with the ligand molecules. EPOR hydrophobic region is sounded by hydrophilic amino acid which shows polar amino acids also play important role in the interactions. EPOR site side chains also showing effective binding of hydrophilic amino acids with Arg. EPO Grid is generated and then docking for binding energy analysis had been done using Glide module. Chemical library is then analyzed for docking analysis to find out the interaction with lesser binding energy. Among them chemical compound 19 and have shown better binding energy value comparable to previously reported mimetic peptides. This study have found (-9.968) about double binding efficiency, then SMND309 compound (-4.081) binding energy which is reported to activating the EPOR for erythropoiesis regulation. Peptide mimetic 25 have Etotal value -479.7 which better then Etotal value of known mimetic EMP1, ERB1-7, minimal peptide sequence -459.4, -426.6 and -437.6 respectively. Lead chemical compound Peptidic mimetic further analyzed for physicochemical properties and stability. Binding affinity of chemical compound mimetic 19 found -78.374 using Prime-GBSA. Peptide mimetic is found 28% hydrophilic in nature which shows it is stable in water and hydrophobicity is 21% which is important for crossing the membranes and required for globular structure of proteins stability. Lead peptide hydrophilic and hydrophobic nature may have important role in binding to EPOR as we had seen polar and non polar both interaction involved in binding of EPO-EPOR complex. Cello predictor results shows that peptide localization by secretary pathway is to nuclear and mitochondrial where it can bind to its receptor and activate the signalling for erythrocytes formation. Biological activity and physiochemical properties showing its importance in future treatment of patient with anaemia disease. These dimer mimetics in future posses important role which can be used for ex vivo erythrocytes generation and in many other applications in field of regenerative medicine.

Conclusion

EPO mimetics chemical compound and small peptide sequence which can activate the EPOR cytokine protein and consequent signalling for eythropoiesis have been designed using computational analysis. SMND309 reported to activate EPOR in absence of EPO. which seems it have potential role in erythrocytes generation and blood formation in future however SMND309 is not bind so efficiently. Similarly EMP1, ERB1-7 peptides were known to mimic EPO but they are less efficient then natural endogenous hormone; hence study required to find potent mimetics. Here in this study, combinatorial library of chemical compound and peptide by forming dimer designed. study find out dimer compounds with better binding efficiency and determined the physiochemical properties which seem it is stable and effective compound. In future these compounds posses important role for erythrocytes generation and anaemic patient.

REFERENCES

Alexanian, R; Vaughn, WK; Ruchelman, MW; (1967) Erythropoietin excretion in man following androgens. J Lab Clin Med.;70(5):777–785.

Bazan, JF; (1990) Structural design and molecular evaluation of a cytokine receptor superfamily. Proc Natl Acad Sci USA; 87:6934–6938.

Dalle, B; Henri, A; Fessard, PR; Bettan, M; Scherman, D; Beuzard, Y; and Emmanuel Payen, E; (2001). Dimeric erythropoietin fusion protein with enhanced erythropoietic activity in vitro and in vivo, Blood: 97 (12).

Du, G; Zhu, H; Yu, P; Wang, H; He, J; Ye, L; Fu, F; Zhang, J; Tian, J; (2013). SMND-309 promotes angiogenesis in human umbilical vein endothelial cells through activating erythropoietin receptor/STAT3/VEGF pathways Eur J Pharmacol. 700(1-3):173-80.

Faulds, D; Sorkin, EM; Epoetin (1989) (recombinant human erythropoietin). A review of its pharmacodynamic and pharmacokinetic properties and therapeutic potential in anemia and the stimulation of erythropoiesis. Drugs. ;38:863–899.

Gallop, MA; Barrett,RW; Dower, WJ; Fodor, SPA; Gordon, EM; (1994). Applications of combinatorial technologies to drug discovery. Background and peptide combinatorial libraries" Journal of Medicinal Chemistry. 37: 1233-51.

Goldberg, J; Jin, Q; Ambroise, Y; Satoh, S; Desharnais, J; Capps, K; Boger, DL; (2002).Erythropoietin Mimetics Derived from Solution Phase Combinatorial Libraries, J. Am. Chem. Soc.124 (4), 544–555.

Johnson, DL; Farrell, FX; Barbone, FP; (1997). Amino terminal dimerization of an erythropoietin mimetic peptide results in increased erythropoietic activity. Chem Biol; 4: 939–950.

Lacombe, C; Mayeux, P; (1998) Haematologica, 83,724-732.

Lappin, TR; Rich, IN;(1996) Erythropoietin--the first 90 years. Clin Lab Haematol Sep; 18(3):137-45.

Lin, FK; Suggs, S; Lin, CH; Browne, JK;(1985) Smalling R, Egrie JC, Chen KK, Fox GM, Martin F, Stabinsky Z, Cloning and expression of the human erythropoietin gene. Proc Natl Acad Sci U S A. Nov;82(22):7580-4.

Maiese K, Li F, and Chong ZZ (2005) New avenues of exploration for erythropoietin. JAMA 293: 90–95.

McConnell, SJ; Dinh, T; Le, MH; (1998). Isolation of Erythropoietin Receptor Agonist Peptides Using Evolved Phage Libraries. Biol Chem; 379: 1279–1286.

Miyake, T; Kung, CK; Goldwasser, E; (1977) Purification of human erythropoietin , The Journal of Biological Chemistry,252, 5558-5564.

Neduva, V; Russell, RB; (2005). Linear motifs evolutionary interaction switches FEBS Lett, 579 3342-3345.

Liu, Z; Stoll, VS; Devries, PJ; Jakob, CG; Xie, N; Simmer, RL; Lacy, SE; Egan, DA; Harlan, JE; Lesniewski, RR; Reilly, EB; (2007). A potent erythropoietin-mimicking human antibody interacts through a novel binding site. Blood. 110(7):2408-13.

Livnah, O; Stura,EA; Johnson, DL; Middleton, SA; Mulcahy, LS; Wrighton, NC; Dower, WJ; Jolliffe, LK; Wilson, IA; (1996).Functional Mimicry of a Protein Hormone by a Peptide Agonist: The EPOReceptor Complex at 2.8 Å. 273. 5274. 464-471.

Syed, RS; Reid, SW; Li, C; Cheetham, JC; Aoki, KH; Liu, B; Zhan, H; Osslund, TD; Chirino, AJ; Zhang, J; Finer-Moore, J; Elliott, S; Sitney, K; Katz, BA; Matthews, DJ; Wendoloski, JJ; Egrie, J; Stroud, RM;(1998) Efficiency of signalling through cytokine receptors depends critically on receptor orientation. Nature. 395(6701):511-6.

Wrighton, NC; Balasubramanian, P; Barbone, FP; (1997). Increased potency of an erythropoietin peptide mimetic through covalent dimerization. Nature Biotechnol; 15: 1261–1265.

Wrighton, NC; Farrell, FX; Chang, R; Kashyap, AK; Barbone, FP; Mulcahy, LS; Johnson, DL; Barrett, RW; Jolliffe, LK; Dower, WJ; (1996). Small Peptides as Potent Mimetics of the Protein Hormone Erythropoietin. Science : Vol. 273 no. 5274 pp. 458-463.

Zhu, H; Zou, L; Tian, J; Du, G; Gao, Y; (2013). SMND-309, a novel derivative of salvianolic acid B, protects rat brains ischemia and reperfusion injury by targeting the JAK2/STAT3 pathway.Eur J Pharmacol. 714(1-3):23-31.

Appendix

Software used

Maestro server

Maestro is Schrödinger application which involves combination of various molecular modelling tools. It can be run on graphical interface or command bas on windows and Linux depends on the tools.

Glide

Glide means grid base ligand docking with energtics. It searches the interaction efficiency between the one or two ligand with receptors.

Ligprep

Ligprep is collection of tool for preparing the ligand drug like molecule for interaction analysis. This provides the low energy model of target ligand with appropriate chiralities.

Prepwizard

Prepwizard tool is used to preparation of receptor proteins. This tool removes the water molecules and unwanted ligands attached to the receptor protein. It differentiates the different chains of the protein and minimization and optimization of the protein.

XP visualiser

XP visualiser is application of Maestro for visualizing various physicochemical properties of interacting ligand and receptor. This application considers the Lipophilic fraction, activity, H-bond and Expos penalty.

Prime MM-GBSA

Prime MM-GBSA is designed to determine the binding affinity of lignd molecules with the receptor. It is compatible with the ligand sand receptors which are prepared first for removing water and other unwanted ions**.**

Pepfold

Pepfold tool denovo 3D structure prediction tool for peptide sequence. This gives PDB structure prediction in PDB format. It is tool of Mobyler server include the Pepfold another various peptide sequences analysis tools.

HEX

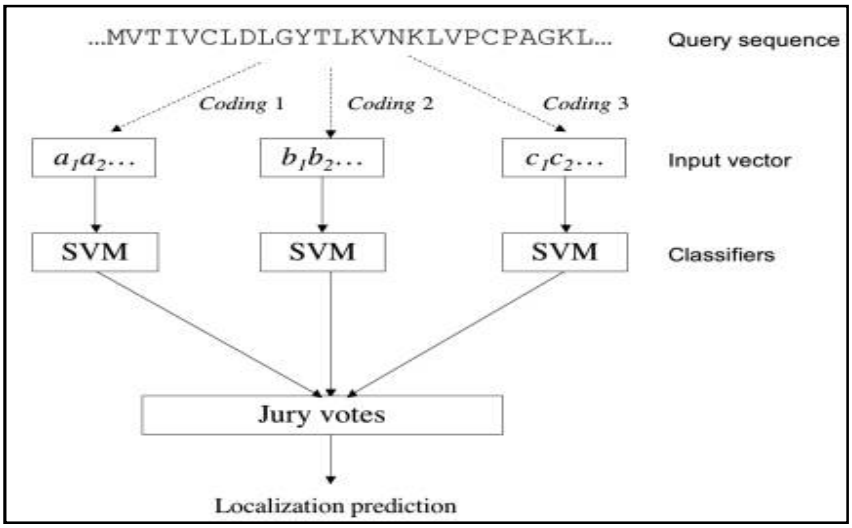
HEX server is protein docking analysis module. Receptor and ligand only in PDB format can be imported in the module. Receptor and ligands imported to the tool first prepared for removing water and unwanted molecules then interaction analysis executed. HEX results in PDB of ligand and receptor protein and with the log files reporting the Etotal values for interaction. Default parameters and interaction domain have been set which can be changed according to the need.

Protein peptide calculator

Protein peptide calculator is designed to identify various properties of the target protein sequences. It determines the pI, molecular weight, stability prediction, hydrophilicity and hydrophobicity of the protein.

Cello

This is prediction tool for determining the sub cellular localization of target protein in the cell. It is support vector machine system which identifies physicochemical property and then predicts the location of target protein. Architecture is shown below-



Pymol

Pymol is freely available tool for the 3D view of PDB structure of proteins. It supports various known format for 3D structures. This tool discriminates the different chains and domain of protein structure for sharp view.

Aliphatic index

The aliphatic index of a protein is the relative volume occupied by aliphatic amino acids side chains (alanine, valine, isoleucine, and leucine). It can be considered as a positive factor for the increasing the thermo stability of globular proteins.

Gravy

The Gravy (grand average of hydropathy) determines sum of hydrophobic values the amino acids for a target peptide or protein. It shows results as hydropathy divided by the residues number in the sequence.