Mathematical Biology and Bioinformatics
2023. V. 18. № 1. P. 72–88. doi: 10.17537/2023.18.72

========== **BIOINFORMATICS** ==========

Cancer Therapeutics: Structure-Based Drug Design of Inhibitors for a Novel Angiogenic Growth Factor

Navaneetha Nambigari*

Department of Chemistry, University College of Science, Saifabad, Osmania University, Hyderabad, Telangana State, India

Abstract. Angiogenesis, the formation of new blood vessels, is a critical and rate-limiting tumor growth step controlled by pro-angiogenic factors and specific inhibitors. Tumor angiogenesis is essential for cancer progression and metastasis. Platelet growth factors (PDGF) and their receptors (PDGFR) are associated with tumor angiogenesis through overexpression of PDGF. Inhibition of PDGF and its signaling pathway is a new approach to the discovery of anticancer therapeutic agents. The present study focuses on the PDGF-C protein in the identification of novel anti-angiogenic compounds. MODELLER 9.10 software allows users to create and refine a 3D homology model of the PDGF-C protein (345 AA length). Secondary structure analysis of the 3D energy model reveals 16 β sheets held together by four cation–π and one π–σ interactions, and three salt bridges. The quality of the model is assessed using the Ramachandran plot (90 percent amino acids in the favorable region) and the ProSA server (*Z*-score = –2.28). Active site residues are identified using Castp, QSite search engine, site map, and protein docking of the protein to its receptor. In addition, virtual screening is performed at the active site using the Glide module of the Schrodinger Suite. Glide score, glide energy and ADME are being measured to discover new benefits of pyrazolone and pyrrolidine-2,3-dione scaffolds as potent PDGF-C antagonists for anti-angiogenic cancer chemotherapy drugs.

Key words: *angiogenesis, PDGF-C antagonists, tumor pathogenesis, metastasis, antiangiogenic, virtual screening.*

INTRODUCTION

The role of angiogenesis in cancer metastasis is an interesting area of research in the previous decade [1]. Angiogenesis is a hallmark cancer process in the development of new blood vessels in tumors. Pro-angiogenic factors (Growth factors) have a significant role in angiogenesis and the development of new vasculature that "nourishes" cancer [2]. These proteins are the vascular endothelial growth factor (VEGF), platelet-derived growth factor (PDGF), fibroblast growth factor (FGF), and the angiopoietin/Tie2 receptor axis, which are significant in molecular pathways in the tumor cell and the focus of the development of anti-angiogenic therapies. Selective upregulation of angiogenic factors led to the resistance to anti-VEGF therapy, and antiangiogenic therapy. Resolving the challenges of drug resistance, adverse effects, and antiangiogenic drugs; new strategies targeting pericytes (PCs) or smooth muscle cells (SMCs) and pro-angiogenic growth factors is essential [3] .

In studies conducted over the last two decades, PDGF is implicated in human tumors, including glioma [4]. platelet-derived growth factors (PDGF) are a pleiotropic family of peptides that stimulate cellular functions such as growth, proliferation, and differentiation by binding to cell surface tyrosine kinase receptors (PDGFR) [5]. PDGF-C plays a significant role in recruiting fibroblasts associated with drug resistance in tumors [6–8], induces

*navaneeta@osmania.ac.in, nitha379@gmail.com

monocyte migration, and up-regulates matrix metalloproteinase (MMP–2 and MMP-9) expression, implying a critical role in pathological angiogenesis [9]. Figure 1 shows the angiogenic activity of PDGF-C, exerted through its effect on fibroblasts [10]. The role of PDGF-C is studied in two angiogenic models: chick chorio allantoic membrane (CAM) and mouse corneal [11]. Furthermore, increased expression and secretion of PDGF-C in Ewing sarcoma cell lines suggested a role for PDGF-C in malignancy [12, 13].

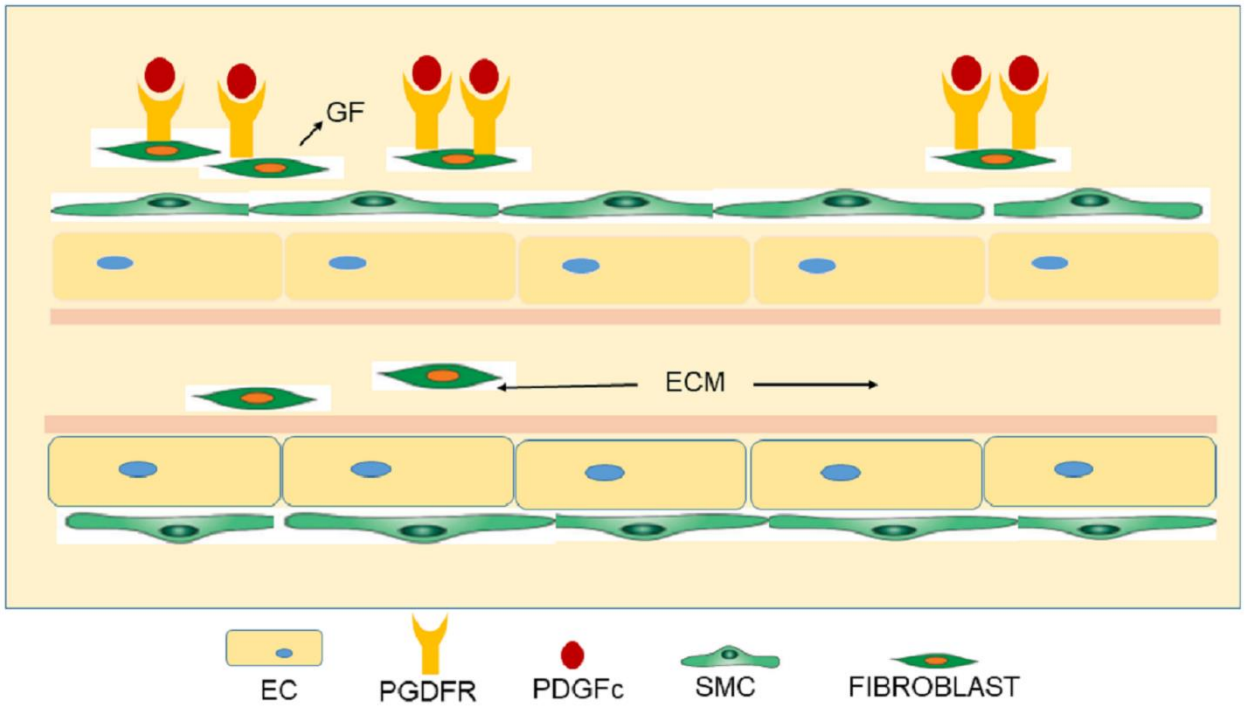

Fig. 1. Angiogenic activity of PDGF-C. Fibroblasts are the most abundant mesenchymal cell component of angiogenic growth factors (GF), extracellular matrix (ECM), and ECM-degrading proteases including matrix metalloproteinases (MMP). Enhanced expression of chemotactic and oncogenic factors, the percentage of fibroblasts elevated during disease processes. PDGF-C stimulates fibroblast proliferation, migration, and recruitment. The angiogenic activity of PDGF-C is exerted through its effect on fibroblast.

The novel PDGF-C ligand is a dormant homodimer with a distinct two entities and expression pattern. Activated PDGF-C stimulates its cognate receptor, -PDGFR [14, 15] via the extracellular signal-regulated kinase cascade (PI3K/AKT Pathway), resulting in overexpression [16]. As a result, inhibiting PDGF-C may create additional therapeutic options for treating neovascular diseases. PDGF C appears to be a promising new biomarker and potential cancer therapeutic target. Recognizing structural features of PDGF-C is therefore beneficial for drug design and development. Since the PDGF-C protein crystal structure is not reported in the Brookhaven PDB database, a rational modelled 3D structure of PDGF - C is generated using homology modelling. Various servers, such as PROCHECK and PRoSA, validate the modeled structure. The natural substrate is docked to the PGDF-C to investigate the putative pharmacological profile of PDGF-C protein. Furthermore, structure-based virtual screening studies yield novel and selective molecules.

The literature data shows that the membrane-proximal homotypic PDGFRα interaction, albeit required for activation, contributes negatively to ligand binding. The structural and biochemical data together offer insights into PDGF-PDGFR signaling, as well as strategies for PDGF-antagonism.

MATERIALS AND METHODS

The comparative modeling method can predict the 3D structure of a protein with the same accuracy as a low-resolution experimentally determined structure [17]. This method identifies one or more known protein structures that are likely to resemble the structure of the query sequence to generate an alignment that maps query sequence residues to template sequence residues. The structural model of the target is built using sequence alignment and template structure. Even though protein structures are more conserved than DNA sequences, detectable sequence similarity usually implies significant structural similarity [18].

Homology modeling. Template identification and generation of 3D structure

The amino acid sequence of PDGF-C (345 AA residues) is retrieved from proteomics server (UniProtKB, ExPASy Swiss-Prot/TrEMBL) with accession number Q9NRA1.PSI-BLAST [19], Jpred3 [20], and Domain Fishing [21], were used to search the sequence database for the homolog of a known structure from the PDB for PDGF-C template proteins. The coordinates of a template (PDB ID: 3KQ4, 3.3 Å resolution) were chosen to design and build the preliminary PDGF-C architecture [22]. ClustalW2 was used to perform pairwise sequence alignment between PDGF-C and the template protein, defining the major blocks of similarity between the query and template sequences using the Gonnet matrix [23–25]. The tertiary structure of PDGF-C is built using MODELLER (Version 9.10), an automated programme that extracts spatial restraints from two sources (homology derived and CHARMm force field derived) [26]. The three-dimensional (3D) model is created by generating ideally satisfying spatial restraints from the alignment and expressing them as probability density functions (pdfs) for the restrained features [27, 28]. Initially, 25 models were built, the model with the lowest modeller objective function being opted for further refinement in structurally variable regions (SVRS). The non-conserved residues are refined using the build loop and scan loop modules of the Swiss-PDB viewer, which account for bad contacts, clashes, and hydrogen bonds [29, 30].

Energy minimization and model validation

The energy minimization was carried out using the Impact refinement module (Impreff) (Impact v 5.0, Schrodinger LLC, New York, NY) at the default cut-off RMSD (Root Mean Score Deviation) of 0.30, with the force field set to the optimized potential for liquid simulations (OPLS 2005) [31, 32]. The 3D model of PDGF-C is validated to assess stereochemical quality using the PROCHECK [33] program of the Structural Analysis and Verification server (SAVES). To analyze the quality of the 3D model via a Ramachandran plot of dihedral angles against amino acid residues [34]. The constructed structure is validated further by measuring sequence and structure compatibility using an energy or scoring function, such as PRoSA's Z-score. The ProSA tool compares the 3D model to experimental protein structures [35].

Active site and protein–protein docking

The prediction of ligand-binding sites is an important step in understanding the molecular recognition mechanism and function of a protein [36]. The three-dimensional structure of proteins provides the required shape and physicochemical texture for binding interactions. The identification of putative binding residues, active site prediction servers such as CASTp [37], Q-Site finder[38], and Sitemap module [39] from Schrodinger LLC, New York, and literary studies are used.

In silico predictions of protein interactions using protein-protein docking studies enable the identification of significant residue-residue contacts involving target interactions. ZDOCK (V 3.1) [40], a Fast Fourier Transform-based docking algorithm, is used to interpret the key

interacting residues of PDGF-C and its receptor PDGFR- (treated as rigid bodies) involved in angiogenesis signaling activation. The amino acid residues not involved in the active site were blocked using ZDOCK 3.1's block.pl. The top 2000 predictions generated at 6° rotational and translational degrees of freedom were thoroughly explored using the scoring function – "pairwise shape complementarity function," and docked complexes generated with creating pl were ranked further based on z-score.

The complexes with the highest z-scores were re-ranked using ZRANK [41]. Another geometry-based molecular docking algorithm, PATCHDOCK (V 1.3), is used to obtain high accuracy interaction predictions based on docking transformations evaluated by a scoring function that yields good molecular shape complementarity [42]. The predicted top 20 solutions were downloaded, analyzed, and the solvent accessible surface area (SASA) calculated using Discovery Studio version 3.5, Accelrys program [43].

Structure-based virtual screening using molecular docking

Computational approaches that 'dock' small molecules with macromolecules and score their complementarity to binding sites are extensively used in hit identification and lead optimization [44]. The primary objectives of drug discovery are to find new chemical entities that are highly capable of binding to the target protein and elicit the desired biological response. Virtual screening necessitates knowledge of the receptor's spatial and energetic requirements [45].

Docking is a crucial conformational sampling technique that is quite often used to predict the binding perspective of small molecule drug candidates to their protein targets [46]. As a result, docking is extremely important in rational drug design. GLIDE (Grid Based Ligand Docking with Energetics) is a computational method for rapidly docking ligands to protein sites and estimating the binding affinities of the docked compounds. It uses a series of hierarchical filters to search for possible ligand locations in the target protein's active-site region [47, 48].

The energy-minimized PDGF-C protein was further considered for structure-based virtual screening using the GLIDE (v 5.6) programme [49]. The scaling factor was 1.0, and the partial charge cut-off was set to 0.25 in Van Der Waal's Radius Scaling. The scoring Grid dimensions are 32 Å × 32 Å × 32 Å so that the ligands bind in that groove [50]. Secondly, a small molecule OTAVA prime screen and NCI cancer ligand dataset were used for docking studies. The 3D coordinates of ligands were generated at a physiological pH of 7.0 ± 2.0 using Ligprep Module (v 4.0) in Maestro v 9.1 (Schrödinger, LLC, 20, New York, NY) using an OPLS_2005 force field [51]. Generation of Ionizers, tautomer, and retention of Specified chiralities to generate low energy conformers was at default parameters.

The ligands with the lowest energy were then subjected to flexible docking against the putative active site of the PDGF-C protein in a sequential flow using the GLIDE module (V 5.6) [52], first in HTVS (High throughput virtual screening) mode to screen a multiple-ligand file for structures that interact favorably with the protein's active site. The top-scoring 10 % of ligand poses evaluated in SP mode (Standard Precision) were further refined in XP mode (Extra Precision) to obtain the best ligand poses for docking simulations [53]. The top-ranked poses are post-docked to optimize bond lengths and angles, as well as torsional angles, and the poses are rescored using the Glide Score.The ADME properties of the hits with the highest Glide score were investigated.

ADME/Pharmacokinetic prediction

Additional research on the computational ADME properties of ligand molecules leads us down an intriguing path that provides information to predict drug-able properties. As a result, drug discovery research is under intense pressure to improve the chances of a drug's success

in clinical trials [54]. Ineffective drug candidates with poor ADME (Absorption, Distribution, Metabolism, and Excretion) properties were eliminated.

Potential ligands with acceptable Glide score and Glide Energy values were further checked for ADME. The new Swiss ADME web tool uses a robust in-house methods like BOILEDEgg, iLOGP, and Bioavailability Radar to adapt quick predictive models for physicochemical properties, pharmacokinetics, drug-like, and drug-chemical friendliness [55].

The best docked ligand molecules are identified as the new potent leads against PGDF-C protein inhibition based on physicochemical properties (Swiss ADME), binding score, RMSD, and visual inspection, were identified as new hits.

RESULTS AND DISCUSSION

The present investigation looks at PDGF-C as a novel target protein for identifying new leads as drug candidates to inhibit pathological angiogenesis. Diverse validation techniques were used for the analysis of the 3D model of target protein. Protein-protein docking confirmed the active site residues. Virtual screening was used to find the best-docked molecule.

Homology modelling of PGDF-C

The ExPASy server was used to retrieve the FASTA sequence of human PDGF-C (Accession: Q9NRA1). The target protein is of 345 residues in length, molecular weight of 39,029 Da, cytogenetic location is 4q32.1 with genomic coordinates (GRCh38) 4:156 760 453–156 971 798.

The template search was carried out on servers, such as NCBI-Blast, JPred, and Domain fishing; the results of the predicted template are shown in Table 1. Position-Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST), a tool used to identify a template, a homologous amino acid sequence with the PGDF-C protein sequence [56]. The BLAST program employs a stochastic method to identify the template with the highest similarity to the target protein [57]. The low *E*-value of 3KQ4-B indicates a strong biological relationship with the PGDF-C amino acid sequence.

Table 1. The template search results for the PDGF-C protein

S. No.	Name of the protein database search server	Parameters (s) considered for template selection	*E*- value	PDB code
1	BLAST	Sequence position specificity.	7×10^{-10}	3KQ4-B
2	Jpred3	Secondary structure prediction, solvent accessibility and coiled-coil region prediction	4×10^{-9}	3KQ4-B
3	Domain fishing	Protein fold recognition	5×10^{-48}	3KQ4-B

The Jpred 3 server tool identified template proteins of homologous secondary structural elements (-helices, -sheets, and loops) with solvent accessibility, and coiled-coil region prediction for proteins in the RCSPDB using the JNet algorithm. The JPred server also identified 3KQ4-B as a PGDF-C template [58].

Domain fishing uses a sequence to split into domains, and the search template from PFAM plus PDB and SCOP is ranked based on sequence identity (percent ID), coverage, and resolution with an *E*-value = 5×10^{-48}, and the predicted template structure is 3KQ4-B [21].

The percent identity of bases that are identical to the reference sequence and query sequence can still be a true hit. While looking for homology between conserved regions, the *E*-value is critical [59]. The template has a query coverage of 33 %, a percent identity of

37.07 %, and is chosen as a template protein (retrieved from the RSC protein data bank) based on the lowest *E*-score to build a reliable model for the PGDF-C protein.

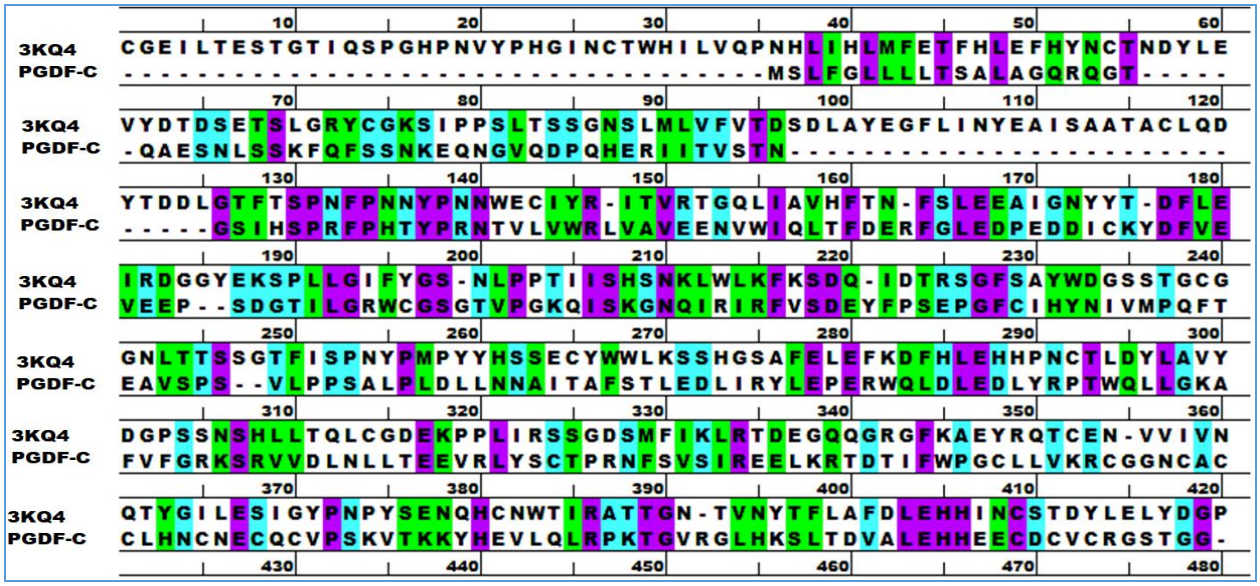

Fig. 2. Sequence alignment (pairwise) of PDGF-C with the template protein 3KQ4-B by using Clustal-W2. The conserved residues are indicated with purple, strongly similar residues with green, weakly similar residues with cyan, and white for diversity.

The prerequisite for creating a reliable 3D structure is to reliably align PGDF-C with its phylogenetically related protein sequence, 3KQ4-B. The pairwise alignment of the PGDF-C protein sequence with the template protein sequence using CLUSTAL-W software is shown in Figure 2, with a 40.57 % similarity between the target and template sequences.

Structural analysis of the protein

The generated 3D model of the protein (PGDF-C) secondary structure is shown in Figure 3.

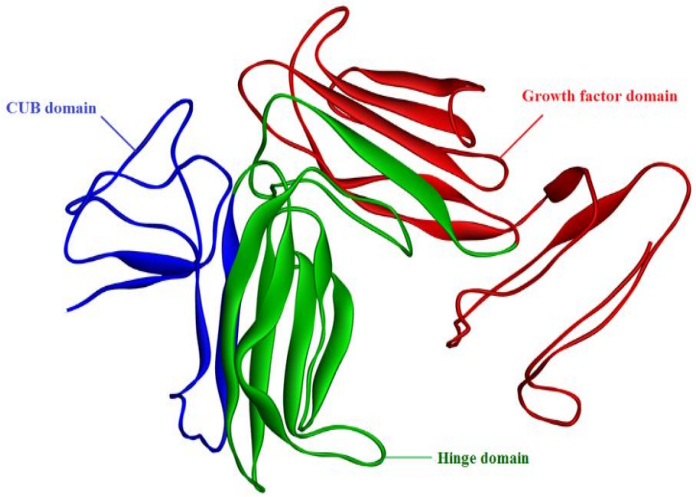

Fig. 3. 3D-structure of PDGF-C protein showing the 3 domains: Cub domain (violet color), hinge domain (green color), growth factor domain (maroon color).

Two long, highly twisted antiparallel pairs of β-strands in an antiparallel side-by-side mode characterize these growth factor domains. They all have eight highly conserved

cysteines (I-VIII).The three loops connecting the monomeric antiparallel ββ-strands are referred to as loops 1, 2, and 3; due to the head-to-tail arrangement, loop 2 of one monomer will be close to loops 1 and 3. Table 2 contains information on the 16 β-sheets.

Table 2. Outline of residues representing the secondary structure of PDGF-C protein

B sheets	Aminoacids*
Sheet 1	FGL
Sheet 2	VQD
Sheet 3	IHS
Sheet 4	TVLVWRL
Sheet 5	IQL
Sheet 6	YDFVEVEE
Sheet 7	TILGRW
Sheet 8	QIS
Sheet 9	IRIRFVSD
Sheet 10	FCI
Sheet 11	AITAFST
Sheet 12	YLEPERW
Sheet 13	WQLLGKA
Sheet 14	DLN
Sheet 15	FSVSIREE
Sheet 16	WPGCLLV

*single letter codes are used for amino acid residues

Protein tertiary structure (3D) is critical for their function as well as their interaction with drugs [39]. Among the PGDF homologs, the binding domain is the most divergent in both sequence and structure, this growth factor site is important for ligand binding for most CUB domain-containing proteins. The 3D model of PGDF-C contains 16 ß-sheets that are required for the separation of the growth factor domain from the CUB domain.

Salt bridges and π-π interactions preserve the amino acids in the PGDF-C protein structure together. The model includes four π-cation interactions and one π–σ interaction, as well as three salt bridges (Table 3), which increase the protein's stability.

Ramachandran plot (from PROCHECK server) measures the angular distribution of ψ/φ torsion angles of the protein model's backbone residues. The Figure 4 reveals that 242 (80.4 %) of the residues have φ/ψ angles in the favored regions, 52 (17.5 %) in the additional allowed region, 7 (2.3 %) in the generously allowed region, and none (0 %) in the disallowed region. The stereo-chemical quality of any protein structure of a good quality model is expected to have over 90 % in the most favored regions [A, B, L].

Table 3. The hydrophobic interactions and salt bridges in the protein*

Bond	π–cation interactions	π–σ interactions	salt bridges (A^0)
Residue name and number	Phe 270 - Arg 207 Phe 32 - Lys 36 Phe 30 - Arg 76 Tyr 248 - Lys225	Tyr 202 - Cys 250	Asp 102 - Lys 105 (3.439) Asp 198 - Lys 278 (3.988) Arg 201 - Glu 307 (2.671)

*three letter codes are used for amino acid

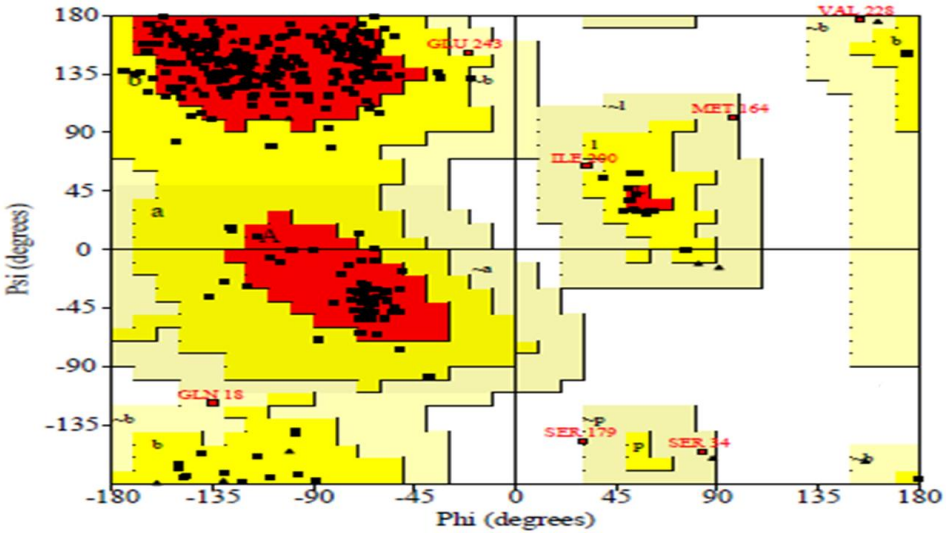

Fig.4. Ramachandran plot of PDGF-C protein. The favored region is in red, the additionally allowed region in the yellow, the generously allowed region in light yellow, and the disallowed region in white.

The Figure S1 (see Supplementary Material) depicts the Verify3D compatibility of the PDGF-C protein model (3D) with its own amino acid sequence (1D). The Verify3D server ascertained whether an atomic PDGF-C model (3D) was compatible with its amino acid sequence (1D). For each of the 345 residues, the scores of a sliding 21-residue window (from –10 to +10) are added and plotted. The average 3D-1D score of 51.73 % of the residues is greater than 0.2.

The ProSA server calculates the energy required for protein folding architecture as a function of the amino acid sequence. The low *Z*-score indicates a high overall model quality and compares the deviation of the ProSA server to calculate the energy required for protein folding architecture as a function of the amino acid sequence. ProSA-Web *Z*-score determined by X-ray crystallography (light blue) and NMR spectroscopy for all proteins in the PDB (dark blue). The black spot in Figure 5,a corresponds to the PDGF-C protein and has the *Z*-score value of –2.28.

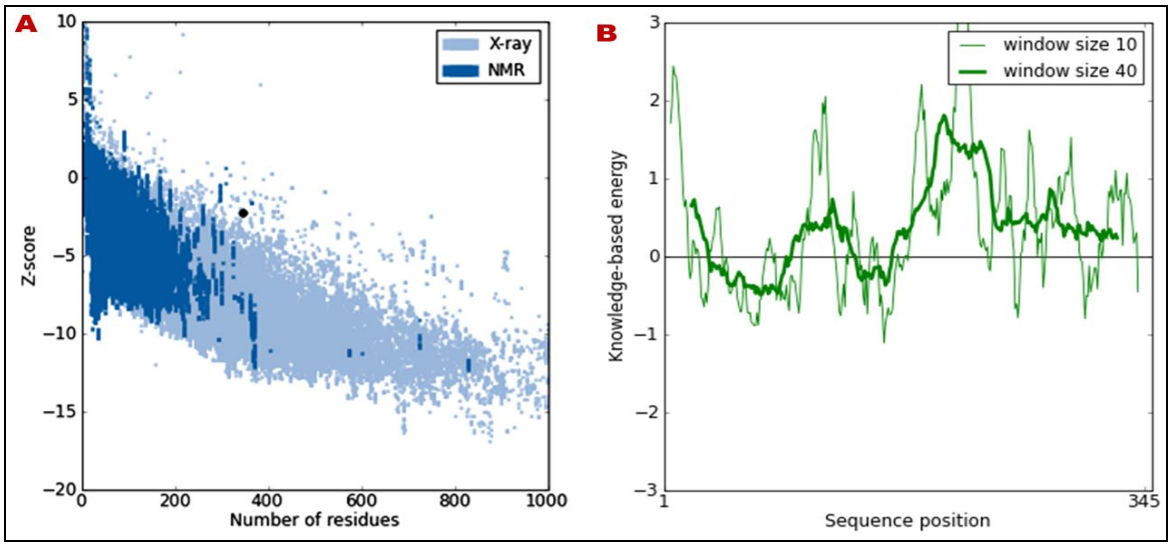

Fig. 5. ProSA profile of PDGF-C protein. **a**) The overall model quality of PDGF-C by ProSA plot. **b**) Energy profile of residue energies as a function of amino acid sequence position.

The low *Z*-score indicates good overall model quality and compares the structure's total energy deviation from an energy distribution derived from native conformations [60]. Overall

folding energies of the protein residues are highly negative, with the folding energy for the model protein in the range of native conformations having a *Z*-score of –2.28. This value is very close to the template value (–6.45), indicating that the obtained model is reliable and closely matches the experimentally determined structures. The C-terminal domain contains a large number of residues with an energy distribution that is entirely below the zero baselines, which is consistent with protein parameters. The ProSA plot (Fig. 5,b) shows that the modelled PGDF-C and the native structure are well correlated in almost all parts of the sequence.

In recent years, inhibiting PDGF-PDGFR signaling has become an attractive pursuit in anticancer therapy, and combined inhibition of PDGF and VEGF has emerged as a promising strategy for suppressing angiogenesis in tumor progression [61, 62]. Several strategies have been utilized to block PDGF/PDGFR signaling at the extracellular level: neutralizing antibodies for PDGF ligands and receptors, aptamers, N-terminal processing-deficient PDGFs, and soluble receptors without the kinase domain [63]. The development of such therapies would greatly benefit from a detailed structural model of PDGF/PDGFR interaction.

All members of the PGDF family share a growth factor core domain containing a conserved set of cysteine residues. The core domain is necessary and sufficient for receptor binding and activation. Activation in the extracellular space requires dissociation of the growth factor domain from the CUB domain. Plasmin and tissue plasminogen activator (tPA) has been demonstrated to proteolytically remove the CUB domain in PDGF-C, rendering it biologically active [64]. TPA needs to interact with both the CUB domain and the core domain in order for cleavage and activation of PDGF-C to occur, which likely explains this specificity. The Figure 6 show the conserved CUB domain of PDGF-C, implicated in protein-protein and protein-carbohydrate interactions and may regulate the extracellular distribution of latent PDGF-C [65].

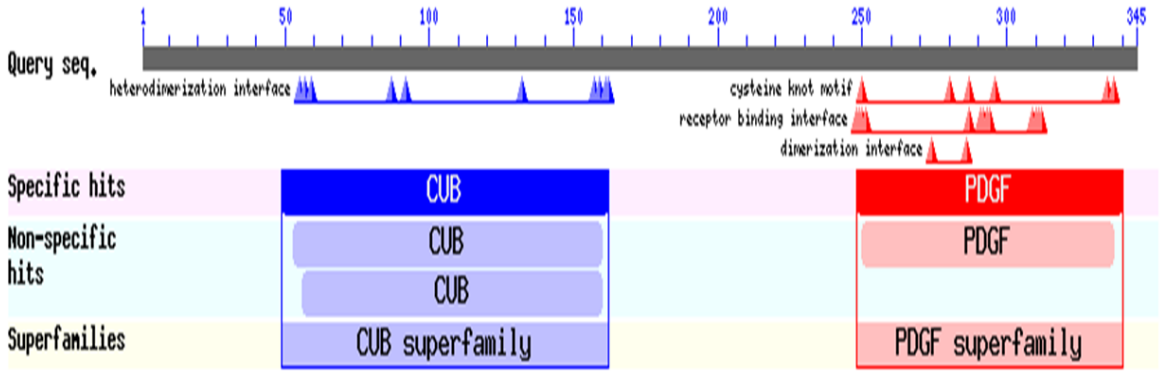

Fig.6. Conserved domains in the PDGF-C protein sequence. PDGF-C sequence representation N-terminal CUB Domain (46–63 residues) shown in blue and C-terminal growth factor domain (235–345 residues) shown in red.

The region of PGDF-C protein binding to its receptor was taken as a reference to determine the putative active site. The active site residues and pocket volumes were identified using CASTp site map and Qsite finder servers as shown in Table 4.

A pictorial representation of the site is shown in Figure 7. The CASTp analysis shows a binding cavity which possess the following residues: 288, 289, 291, 292, 305, 307, 309, 311–314, 319, 320, 323, 337, 339,342, 344, 345; The Q site analysis identified a cavity the protein with 294-296, 300, 302, 303, 305, 306, 307 amino acid residues and site map identified 284-288, 290, 292-294, 317-321 amino acid being present. The amino acid residues in the region of Tyr304 – Gly 307, Leu311 – Lys314, His331 – Cys 335 and Arg 340 to Gly 345 of PDGF-

C protein have putative binding interactions with these PGDFR. At the active site region, a grid with dimensions of 24 $Å^3$ was created for virtual screening studies.

A library of compounds (from the OTAVA prime screen and NCI cancer ligand databases, respectively) was subjected to energy minimization using the Schrödinger Software LigPrep module in Maestro 9.1. Virtual screening studies were conducted against the PDGF-C protein using a library of molecular structures obtained after ligand preparation. 10 % structures were screened in each docking stage, namely HTVS, SP, and XP dockings. As a result of the virtual screening, a total of 22 docked complexes were obtained. In the virtual screening study, the Glide module of the Schrödinger suite employs the default Monte Carlo simulation search method to generate the best candidates [48].

Table 4. Active site residues are determined by using various servers and their corresponding volume of cavities in Å

Name of the server	The volume of cavity (Å)	No. of amino acid
CASTp	2494.9	288, 289, 291, 292, 305, 307, 309, 311–314, 319, 320, 323, 337, 339, 342, 344, 345
Q-site Finder	137	294–296, 300, 302, 303, 305, 306, 307
Site map	175.3	284–288, 290, 292–294, 317–321

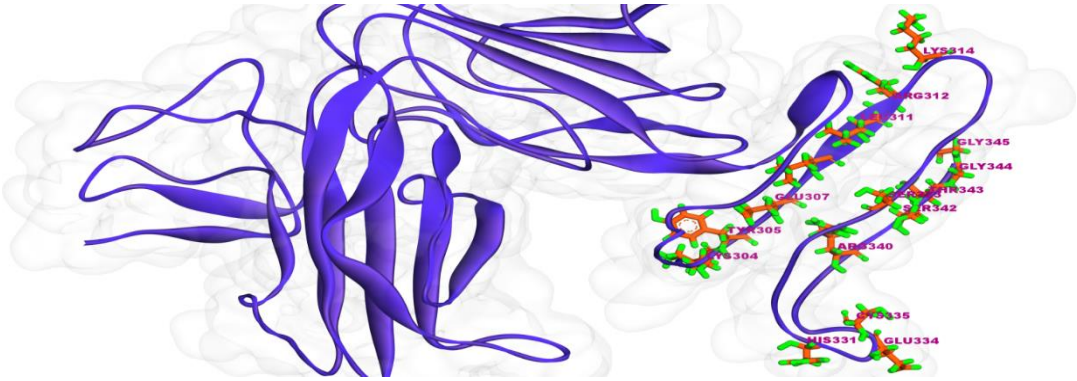

Fig. 7. Active site of PDGF-C protein Putative active site residues (Tyr304 – Gly 307, Leu311 – Lys314, His331 – Cys 335 and Arg 340 to Gly 345) of PDGF-C protein shown in ball and stick (orange, green) and protein is represented in the violet color ribbon.

The binding interactions, specifically the H-bond distances, (< 2.5Å) and π-cations interactions, were investigated for all docked complexes, and ligands were prioritized based on the glide score and energy. Table S1 contains data on the total output (22 docked) complexes. The docking scores of the ligands range from 10.96 to 9.01 kcal/mol. Accelrys DS visualizer 3.5 is used to display the docked structures and H-bonds. The best four compounds with the highest binding energies and significant affinities with the PGDF-C target protein are shown in Figure S2.

All the ligands embedded within the target protein's active site form hydrogen bonds with the same position as the target protein's established active site. At the default condition of glide, the ligand molecules M1 to M8 are docked flexibly to the PGDF-C protein, which is held rigid in the docking process. Table 5 displays the best-docked compound data for a sample of 8 prioritized docked complexes from a total output of 22 docked complexes with permissible ADME.

Based on the Glide score, Glide Energy, and a review of the ligand structures in Table 5, it is clear that the ligands share common pharmacophore properties, which may be attributed to their bioisosteric character. The pyrazolone, pyrrolidine-2,3-dione, and pyrimidine acetamide moieties are present in the majority of the ligands.

Table 5. Predicted ADME of the docked molecules*

Molecule No. (ID)	Glide Energy	HBA	HBD	log P	MW	TPSA (Å)	Lead likeness
M1 (Ligand-82)	–57.54	7	2	2.09	454.47	128.39	2
M2 (Ligand -67)	–54.69	6	3	2.94	429.47	83.06	1
M3 (Ligand-134)	–57.67	4	3	1.39	311.74	120.38	0
M4 (Ligand-133)	–57.90	4	3	1.39	311.74	120.38	0
M5 (Ligand-85)	–55.18	5	2	3.81	472.51	99.05	3
M6 (Ligand-86)	–55.13	6	2	4.11	495.59	89.74	3
M7 (Ligand-132)	–45.65	5	2	1.84	363.43	118.61	2
M8 (Ligand-161)	–48.68	4	2	3.73	407.49	108.56	2

*The permissible ADME values are as follows: molecular weight (MW) < 500, hydrogen bond donor (HBD) < 5, hydrogen bond acceptor (HBA) <10, lipophilicity (log P) < 5.6, total polar surface area (TPSA) <=120 Å.

The pyrrolidine-2,3-dione (M5 and M6) and pyrazolone (M2) scaffolds are shown in this study to act as selective pharmacophores for binding with the PGDF-C. The NH group of the pyrazolone moiety in M2 and the pyrrolidine-2,3-dione moiety in M5 and M6 ligand molecules binds to the amino acid residues Ala 285, Arg253, Leu311 and Ser 323 of the PGDF-C active site. Competitive binding of ligand molecules at the active site can prevent the formation of the PDGF-C-PDGFR complex.

The Figure S3 depicts the Bioavailability Radar plots for the M1–M8 demonstrating rapid assessment of drug-likeness. The bioavailability plot considers six physicochemical properties: lipophilicity, size, polarity, solubility, flexibility, and saturation. A physicochemical range on each axis defines the adapted descriptors and is depicted as a pink area into which the molecule's radar plot had to fall entirely to be considered drug-like [66]. The pink area represents the optimal range for each properties (Lipophilicity: XLOGP3 between −0.7 and +5.0, size: MW between 150 and 500 g/mol, polarity: TPSA between 20 and 120 Å2, solubility: log S not higher than 6, saturation: fraction of carbons in the sp3 hybridization not less than 0.25, and flexibility: no more than 9 rotatable bonds). The M1–M8 are predicted to be orally bioavailable, but M1, M3 and M4 are too polar and none of the compound molecules crosses the BBB.

The Figure 8 illustrates a BOILED-Egg predictive model for gastrointestinal and BBB prediction. The TPSA vs. WlogP plot shows a white region indicating the physicochemical space of molecules with the highest probability of being absorbed by the gastrointestinal tract, and the yellow region (yolk) is the physicochemical space of molecules with the highest probability of permeating to the brain. P-glycoprotein (PGA) substrates are susceptible to pharmacokinetic changes caused by drug interactions with PGA inhibitors or inducers; key

mechanism underlying decreased intracellular drug accumulation in various cancers is PGA overexpression [67].

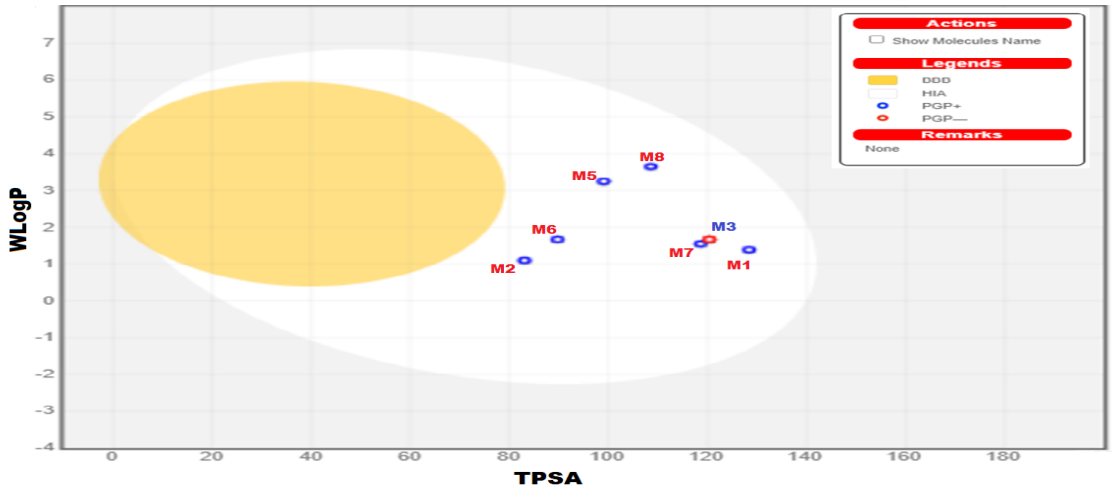

Fig. 8. BOILED-Egg predictive model. The BOILED-Egg model plot based on TPSA vs WLogP for the gastrointestinal and BBB prediction.

Except for M3, the molecules M1 - M8 are PGA+, indicating that they are PGA-inducing drug-like molecules. The Jorgensen Lipinski law (as most anti-cancer agents demonstrate cardiovascular toxicity and obey the Lipinski rule of five, Jorgensen rule of three) and drug-like properties (Table 5) are mentioned [68]. The study of pharmacokinetic properties reveals molecules with permissible ADME properties and lower scaffold toxicity. A critical stage of the drug development process is ADMET assessment before a pre-clinical trial. Polar surface area values (120 Å) are within an acceptable range for all known ligands and have strong overall ADME properties. The results show that the M2, M5, and M6 molecules have synthetic viability compared to other molecules and can thus be considered as novel leads for the design of new PGDF-C protein inhibitors and scaffolds for further development of new lead against pathogenic angiogenesis and cancer.

CONCLUSIONS

The 3D structure of the PDGF-C protein generated using 3KQ4-B as a template is comparable to the X-ray resolved protein structure. Protein-protein docking of PDGF-C with its natural receptor (PDGFR) confirmed the binding residues in the active site region. Virtual screening studies show that the lead molecules have putative binding interactions with the amino acid residues Asn292, Leu309, Leu311, Lys322, Gly341, Thr434, Ser332, and Ser342 of the PDGF - C protein. The studies revealed that pyrazolone and pyrrolidine-2,3-dione scaffolds (Glide Energy = –10.56, –10.12, and –10.04 kcal/mol) with acceptable ADME will inhibit PGDF-C protein by blocking the active site residues and act as leads in the design of inhibitors against pathological angiogenesis in cancer.

The author NNT is thankful to The Head, Department of Chemistry and the Principal, University College of Science, Saifabad, Osmania University, Hyderabad and DST FIST, New Delhi for the Computational Lab facilities to carry out this work.

REFERENCES

1. Harris A.L. Angiogenesis as a New Target for Cancer Control. *Eur. J. Cancer, Suppl.* 2003. V. 1. No. 2. P. 1–12. doi: 10.1016/S1359-6349(03)00007-7
2. Gavalas N., Liontos M., Trachana S.P., Bagratuni T., Arapinis C., Liacos C.,

Dimopoulos M., Bamias A., Angiogenesis-Related Pathways in the Pathogenesis of Ovarian Cancer. *Int. J. Mol. Sci.* 2013. V. 14. No. 8. P. 15885–15909. doi: 10.3390/ijms140815885

3. Bergers G., Hanahan D. Modes of Resistance to Anti-Angiogenic Therapy. *Nat. Rev. Cancer*. 2008. V. 8. No. 8. P. 592–603. doi: 10.1038/nrc2442
4. Plate K.H., BreierG., Farrell C.L., Risau W. Platelet-Derived Growth Factor Receptor-Beta Is Induced Duringtumor Development and Upregulated during Tumor Progressionin Endothelial Cells in Human Gliomas. *Lab. Invest*. 1992. V. 67. P. 529–534.
5. Yu J.H., Ustach C., ChoiKim H.R. Platelet-Derived Growth Factor Signaling and Human Cancer. *BMB Rep.* 2003. V. 36. No. 1. P. 49–59. doi: 10.5483/BMBRep.2003.36.1.049
6. Anderberg C., Li H., Fredriksson L., Andrae J., Betsholtz C., Li X., Eriksson U., Pietras K. Paracrine Signaling by Platelet-Derived Growth Factor-CC Promotes Tumor Growth by Recruitment of Cancer-Associated Fibroblasts. *Cancer Res.* 2009. V. 69. No. 1. P. 369–378. doi: 10.1158/0008-5472.CAN-08-2724
7. Crawford Y., Kasman I., Yu L., Zhong C., Wu X., Modrusan Z., Kaminker J., Ferrara N. PDGF-C Mediates the Angiogenic and Tumorigenic Properties of Fibroblasts Associated with Tumors Refractory to Anti-VEGF Treatment. *Cancer Cell*. 2009. V. 15. No. 1. P. 21–34. doi: 10.1016/j.ccr.2008.12.004
8. Reigstad L. J., Varhaug J. E.; Lillehaug, J. R. Structural and Functional Specificities of PDGF-C and PDGF-D, the Novel Members of the Platelet-Derived Growth Factors Family. *FEBS J.* 2005. V. 272. No. 22. P. 5723–5741. doi: 10.1111/j.1742-4658.2005.04989.x
9. Wågsäter D., Zhu C., Björck H. M., Eriksson P. Effects of PDGF-C and PDGF-D on Monocyte Migration and MMP-2 and MMP-9 Expression. *Atherosclerosis* 2009. V. 202. No. 2. P. 415–423. doi: 10.1016/j.atherosclerosis.2008.04.050
10. Li X., Kumar A., ZhangF., Lee C., Li Y. Tang Z., Arjunan P. VEGF-Independent Angiogenic Pathways Induced by PDGF-C. *Oncotarget.* 2010. V. 1. No. 4. P. 309–314. doi: 10.18632/oncotarget.141
11. Risau W., Drexler H., Mironov V., Smits A., Siegbahn A., Funa K., Heldin C.-H. Platelet-Derived Growth Factor Is Angiogenic *In Vivo*. *Growth Factors.* 1992. V. 7. No. 4. P. 261–266. doi: 10.3109/08977199209046408
12. Zwerner J.P., May W.A. Dominant Negative PDGF-C Inhibits Growth of Ewing Family Tumor Cell Lines. *Oncogene.* 2002. V. 21. No. 24. P. 3847–3854. doi: 10.1038/sj.onc.1205486
13. Zwerner J.P,. May W.A. PDGF-C Is an EWS/FLI Induced Transforming Growth Factor in Ewing Family Tumors. *Oncogene.* 2001. No. 20. No. 5. P. 626-633. doi: 10.1038/sj.onc.1204133
14. Gilbertson D.G., Duff M.E., West J.W., Kelly J.D., Sheppard P.O., Hofstrand P.D., Gao Z., Shoemaker K., Bukowski T.R., Moore M., Feldhaus A.L., Humes J.M. Palmer T.E., Hart C.E. Platelet-Derived Growth Factor C (PDGF-C), a Novel Growth Factor That Binds to PDGF α and β Receptor. *J. Biol. Chem.* 2001. V. 276. No. 29. P. 27406–27414. doi: 10.1074/jbc.M101056200
15. Li X., Pontén A., Aase K., Karlsson L. Abramsson A., Uutela M., Bäckström G., Hellström M., Boström H., Li H., Soriano P., Betsholtz C., Heldin C.-H., Alitalo K., Östman A. Eriksson U. PDGF-C Is a New Protease-Activated Ligand for the PDGF α-Receptor. *Nat. Cell Biol.* 2000. V. 2. No. 5. P. 302–309. doi: 10.1038/35010579
16. Raica M., Cimpean A.M. Platelet-Derived Growth Factor (PDGF)/PDGF Receptors

(PDGFR) Axis as Target for Antitumor and Antiangiogenic Therapy. *Pharmaceuticals.* 2010. V. 3. No. 3. P. 572–599. doi: 10.3390/ph3030572

17. Chothia C., Lesk A.M. The Relation between the Divergence of Sequence and Structure in Proteins. *EMBO J.* 1986. V. 5. No. 4. P. 823–826. doi: 10.1002/j.1460-2075.1986.tb04288.x
18. Martí-Renom M.A., Stuart A.C., Fiser A., Sánchez R., Melo F., Šali A. Comparative Protein Structure Modeling of Genes and Genomes. *Annu. Rev. Biophys. Biomol. Struct.* 2000. V. 29. No. 1. P. 291–325. doi: 10.1146/annurev.biophys.29.1.291
19. Altschul S. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* 1997. V. 25. No. 17. P. 3389–3402. doi: 10.1093/nar/25.17.3389
20. Cole C,. Barber J.D., Barton G.J. The Jpred 3 Secondary Structure Prediction Server. *Nucleic Acids Res.* 2008. 36. P. W197–W201. doi: 10.1093/nar/gkn238
21. Contreras-Moreira B., Bates P.A. Domain Fishing: A First Step in Protein Comparative Modelling. *Bioinformatics* 2002. V. 18. No. 8. P. 1141–1142. doi: 10.1093/bioinformatics/18.8.1141
22. Berman H.M. The Protein Data Bank. *Nucleic Acids Res.* 2000. V. 28. No. 1. P. 235–242. doi: 10.1093/nar/28.1.235
23. Gonnet G.H., Cohen M.A., Benner S.A. Exhaustive Matching of the Entire Protein Sequence Database. *Science.* 1992. V. 256. No. 5062. P. 1443–1445. doi: 10.1126/science.1604319
24. Thompson J.D., Higgins D.G., Gibson T.J. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* 1994. V. 22. No. 22. P. P. 4673–4680. doi: 10.1093/nar/22.22.4673
25. Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G. Clustal W and Clustal X Version 2.0. *Bioinformatics.* 2007. V. 23. No. 21. P. 2947–2948. doi: 10.1093/bioinformatics/btm404
26. Brooks B.R., Bruccoleri R.E., Olafson B.D., States D.J., Swaminathan S.K.M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* 1983. V. 4. P. 187–217. doi: 10.1002/jcc.540040211
27. Šali A., Potterton L., Yuan F., van Vlijmen H., Karplus M. Evaluation of Comparative Protein Modeling by MODELLER. *Proteins Struct. Funct. Genet.* 1995. V. 23. No. 3. P. 318–326. doi: 10.1002/prot.340230306
28. Šali A., Blundell T.L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* 1993. V. 234. No. 3. P. 779–815. doi: 10.1006/jmbi.1993.1626
29. Guex N., Peitsch M.C. SWISS-MODEL and the Swiss-Pdb Viewer: An Environment for Comparative Protein Modeling. *Electrophoresis.* 1997. V. 18. No. 15. P. 2714–2723. doi: 10.1002/elps.1150181505
30. Fiser A., Do R.K.G., Šali A. Modeling of Loops in Protein Structures. *Protein Sci.* 2000. V. 9. No. 9. P. 1753–1773. doi: 10.1110/ps.9.9.1753
31. Jorgensen W.L., Tirado-Rives J. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* 1988. V. 110. No. 6. P. 1657–1666. doi: 10.1021/ja00214a001
32. Jorgensen W.L. Maxwell D.S., Tirado-Rives J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids.

J. Am. Chem. Soc. 1996. V. 118. No. 45. P. 11225–11236. doi: 10.1021/ja9621760

33. Laskowski R.A., MacArthur M.W., Moss D.S. Thornton J.M. PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* 1993. V. 26. No. 2. P. 283–291. doi: 10.1107/S0021889892009944
34. Ramachandran G.N., Ramakrishnan C. Sasisekharan V. Stereochemistry of Polypeptide Chain Configurations. *J. Mol. Biol.* 1963. V. 7. No. 1. P. 95–99. doi: 10.1016/S0022-2836(63)80023-6
35. Wiederstein M., Sippl M.J. ProSA-Web: Interactive Web Service for the Recognition of Errors in Three-Dimensional Structures of Proteins. *Nucleic Acids Res.* 2007. V. 35. P. W407–W410. doi: 10.1093/nar/gkm290
36. Laurie A.T. Methods for the Prediction of Protein-Ligand Binding Sites for Structure-Based Drug Design and Virtual Ligand Screening. *Curr. Protein Pept. Sci.* 2006. V. 7. P. 395–406. doi: 10.2174/138920306778559386
37. Dundas J., Ouyang Z., Tseng J., Binkowski A., Turpaz Y., Liang J. CASTp: Computed Atlas of Surface Topography of Proteins with Structural and Topographical Mapping of Functionally Annotated Residues. *Nucleic Acids Res.* 2006. V. 34. P. W116–W118. doi: 10.1093/nar/gkl282
38. Laurie A.T.R., Jackson R.M. Q-SiteFinder: An Energy-Based Method for the Prediction of Protein-Ligand Binding Sites. *Bioinformatics.* 2005. V. 21. No. 9. P. 1908–1916. doi: 10.1093/bioinformatics/bti315
39. Halgren T.A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* 2009. V. 49. No. 2. P. 377–389. doi: 10.1021/ci800324m
40. Goverdhan L., Revanth B., Mahendar D., Manan B., Sarita Rajender P. Identification and optimisation of novel selective inhibitors against human regulator of G protein signalling 2 (RGS2) protein for type 2 diabetes mellitus: an in silico approach. *Int. J. Comput. Biol. Drug Des.* 2021. V. 14. No. 3. P. 166–189. doi: 10.1504/IJCBDD.2021.10040273
41. Navaneetha N., Kiran Kumar M., Vasavi M., Bhargavi K., Sarita Rajender P., Ramasree D., Uma V. Angiogenesis: An Insilico Approach to Angiogenic Phenotype. *J. Pharm. Res.* 2012. V. 5. No. 1. P. 583–588.
42. Schneidman-Duhovny D., Inbar Y., Nussinov R., Wolfson H.J. PatchDock and SymmDock: Servers for Rigid and Symmetric Docking. *Nucleic Acids Res.* 2005. V. 33. P. W363–W367. doi: 10.1093/nar/gki481
43. *Accelrys Discovery Studio Visualiser v 3.5.0.12158*. San Diego: Accelrys Software Inc., 2012.
44. Ghosh S., Nie A., An J., Huang Z. Structure-Based Virtual Screening of Chemical Libraries for Drug Discovery. *Curr. Opin. Chem. Biol.* 2006. V. 10. No. 3. P. 194–202. doi: 10.1016/j.cbpa.2006.04.002
45. Klebe G. Virtual Ligand Screening: Strategies, Perspectives and Limitations. *Drug Discov. Today.* 2006. V. 11. No. 13–14. P. 580–594. doi: 10.1016/j.drudis.2006.05.012
46. Lengauer T., Rarey M. Computational Methods for Biomolecular Docking. *Curr. Opin. Struct. Biol.* 1996. V. 6. No. 3. P. 402–406. doi: 10.1016/S0959-440X(96)80061-3
47. Lanka G., Bathula R., Bhargavi M., Potlapall S.R. Homology modeling and molecular docking studies for the identification of novel potential therapeutics against human PHD3 as a drug target for type 2 diabetes mellitus. *Journal of Drug Delivery and Therapeutics*. 2019. V. 9. No. 4. P. 265–273.
48. Friesner R.A. Banks J.L. Murphy R.B., Halgren T.A., Klicic J.J., Mainz D.T., Repasky M.P., Knoll E.H., Shelley M., Perry J.K. et al. Glide: A New Approach for

Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* 2004. V. 47. No. 7. P. 1739–1749. doi: 10.1021/jm0306430

49. Kawatkar S., Wang H., Czerminski R., Joseph-McCarthy D. Virtual Fragment Screening: An Exploration of Various Docking and Scoring Protocols for Fragments Using Glide. *J. Comput. Aided. Mol. Des.* 2009. V. 23. No. 8. P. 527–539. doi: 10.1007/s10822-009-9281-4
50. Podvinec M., Lim S.P., Schmidt T., Scarsi M., Wen D., Sonntag L.-S., Sanschagrin P., Shenkin P.S., Schwede T. Novel Inhibitors of Dengue Virus Methyltransferase: Discovery by in Vitro-Driven Virtual Screening on a Desktop Computer Grid. *J. Med. Chem.* 2010. V. 53. No. 4. P. 1483–1495. doi: 10.1021/jm900776m
51. *LigPrep, Version 3.3*. New York, NY: Schrödinger, LLC, 2010.
52. *GLIDE, Version 5.6*. New York, NY: Schrödinger, LLC, 2010.
53. Friesner R.A., Murphy R.B., Repasky M.P., Frye L.L., Greenwood J.R., Halgren T.A., Sanschagrin P.C., Mainz D.T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *J. Med. Chem.* 2006. V. 49. No. 21. P. 6177–6196. doi: 10.1021/jm051256o
54. Rodrigues A.D. Preclinical Drug Metabolism in the Age of High-Throughput Screening: An Industrial Perspec. *Pharm. Res.* 1997. V. 14. No. 11. P. 1504–1510. doi: 10.1023/A:1012105713585
55. Daina A., Zoete V. A BOILED-Egg To Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *Chem.Med.Chem.* 2016. V. 11. No. 11. P. 1117–1121. doi: 10.1002/cmdc.201600182
56. Johnson M., Zaretskaya I., Raytselis Y., Merezhuk Y., McGinnis S., Madden T.L. NCBI BLAST: A Better Web Interface. *Nucleic Acids Res.* 2008. V. 36. P. W5–W9. doi: 10.1093/nar/gkn201
57. Karlin S., Altschul S.F. Methods for Assessing the Statistical Significance of Molecular Sequence Features by Using General Scoring Schemes. *Proc. Natl. Acad. Sci.* 1990. V. 87. No. 6. P. 2264–2268. doi: 10.1073/pnas.87.6.2264
58. Drozdetskiy A., Cole C., Procter J., Barton G.J. JPred4: A Protein Secondary Structure Prediction Server. *Nucleic Acids Res.* 2015. V. 43. No. W1. P. W389–W394. doi: 10.1093/nar/gkv332
59. Kerfeld C.A., Scott K.M. Using BLAST to Teach “E-Value-Tionary” Concepts. *PLoS Biol.* 2011. V. 9. No. 2. P. e1001014. doi: 10.1371/journal.pbio.1001014
60. Wiederstein M., Sippl M.J. ProSA-Web: Interactive Web Service for the Recognition of Errors in Three-Dimensional Structures of Proteins. *Nucleic Acids Res.* 2007. V. 35. P. W407–W410. doi: 10.1093/nar/gkm290
61. Bergers G., Song S., Meyer-Morse N., Bergsland E., Hanahan D. Benefits of Targeting Both Pericytes and Endothelial Cells in the Tumor Vasculature with Kinase Inhibitors. *J. Clin. Invest.* 2003. V. 111. No. 9. P. 1287–1295. doi: 10.1172/JCI17929
62. Östman A. PDGF Receptors-Mediators of Autocrine Tumor Growth and Regulators of Tumor Vasculature and Stroma. *Cytokine Growth Factor Rev.* 2004. V. 15. No. 4. P. 275–286. doi: 10.1016/j.cytogfr.2004.03.002
63. Erber R., Thurnher A., Katsen A.D., Groth G., Kerger H., Hammes H., Menger M.D., Ullrich A., Vajkoczy P. Combined Inhibition of VEGF- and PDGF-signaling Enforces Tumor Vessel Regression by Interfering with Pericyte-mediated Endothelial Cell Survival Mechanisms. *FASEB J.* 2004. V. 18. No. 2. P. 338–340. doi: 10.1096/fj.03-0271fje
64. Andrae J., Gallini R., Betsholtz C. Role of Platelet-Derived Growth Factors in

Physiology and Medicine. *Genes Dev.* 2008. V. 22. No. 10. P. 1276–1312. doi: 10.1101/gad.1653708

65. Fredriksson L., Li H., Fieber C., Li X., Eriksson U. Tissue Plasminogen Activator Is a Potent Activator of PDGF-CC. *EMBO J.* 2004. V. 23. No. 19. P. 3793–3802. doi: 10.1038/sj.emboj.7600397
66. Ritchie T.J., Ertl P., Lewis R. The Graphical Representation of ADME-Related Molecule Properties for Medicinal Chemists. *Drug Discov. Today.* 2011. V. 16. No. 1–2. P. 65–72. doi: 10.1016/j.drudis.2010.11.002
67. Breier A., Gibalova L., Seres M., Barancik M., Sulova Z. New Insight into P-Glycoprotein as a Drug Target. *Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry - Anti-Cancer Agents).* 2013. V. 13. No. 1. P. 159–170. doi: 10.2174/187152013804487380
68. Park S.J., Baars H., Mersmann S., Buschmann H., Baron J.M., Amann P.M., Czaja K., Hollert H., Bluhm K., Redelstein R., Bolm C. *N* -Cyano Sulfoximines: COX Inhibition, Anticancer Activity, Cellular Toxicity, and Mutagenicity. *Chem. Med. Chem.* 2013. V. 8. No. 2. P. 217–220. doi: 10.1002/cmdc.201200403

Received 19.12.2022.
Revised 07.03.2023.
Published 26.03.2023.