Exploration of potential ligands against cancer-causing transcription factor E2F3

Muzammal Hussain[1†], Aqeel Javeed[1†], Muhammad Ashraf[1], Samerene Siddique[2], Amjad Riaz[3] and Muhammad Mahmood Mukhtar[4*]

[1]Department of Pharmacology & Toxicology, [3]Department of Theriogenology, and [4]Department of Microbiology, University of Veterinary and Animal Sciences, Lahore, Pakistan

[2]School of Human Sciences, London Metropolitan University, London, UK

Abstract: The transcription factor-based therapeutic approaches are the mainstay of current anticancer drug design options to develop highly selective agents with novel modes of action. In this paper, a homology model of DNA-binding domain of transcription factor E2F3 was generated according to X-ray structure of E2F4. As a first step of our proposed project aspired towards exploration of highly selective potential E2F3 ligands, we performed structure-based virtual screening of ZINC 3D chemical database by using Dock Blaster server. Then 31 compounds, selected by filtration step, were docked against the prominent DNA binding site residues of E2F3 model. Two of them have shown a promising interaction with respect to binding poses. The aim is to propose new active ligands against neoplasias characterized by overexpression of E2F3 transcription factor.

Keywords: E2F3, Homology modeling, Rational drug design, Virtual screening, transcription factor.

INTRODUCTION

Despite all of the discernible progress in cancer chemotherapy, a great deal of solicitude is there regarding indiscrimination between neoplastic and normal cells. This great concern has engendered unfeigned interest for developing target-based therapies (Sams-Dodd, 2005). One novel approach is to target transcription factors and other signal transduction kinase proteins, thus, divulging highly selective novel agents for cancer intervention (Karamouzis *et al.*, 2002; Nebert, 2002). While extensive amount of cancer drug discovery research is being zealously projected on various transcription-factor families (Libermann *et al.*, 2006), E2F family still needs exploitation for this purpose (Tsantoulis *et al.*, 2005).

E2F is a family of transcription factors that in coordination with DP and pocket proteins (pRB, p107, p130) regulate the cell cycle (Attwooll *et al.*, 2004). Transcription factor E2F3 protein, a member of E2F family, is encoded by the E2F3 gene in humans (Lees *et al.*, 1993). Under normal physiological conditions, after being transcribed by MYC (Adams *et al.*, 2000), E2F3 conspire to promote G1/S-phase transition of cell cycle together with its DP-dimer protein. This protein binds specifically to retinoblastoma protein (pRB) in a cell-cycle dependent manner and has a central role in linking cell cycle proteins, such as cyclins, cyclin-dependent kinases (CDKs) and pRB, to the expression of a variety of genes involved in cell cycle progression and cellular proliferation (Humbert *et al.*, 2000; Leone *et al.*, 1998). E2F3 modulates different genes that play a crucial role in a variety of cellular processes such as cell cycle transition, DNA synthesis, transcription and signal transduction. When overexpressed, especially under pathological conditions of cancer, E2F3 can bear a classical oncogenic axis as demonstrated in 67% of human prostate cancer (Foster *et al.*, 2004), in 55-70% of squamous cell carcinomas and 79% of adenocarcinomas of the lung (Cooper *et al.*, 2006), and in human bladder and ovarian cancer ((Feber *et al.*, 2004; He *et al.*, 2008; Oeggerli *et al.*, 2004). Furthermore, many studies have suggested E2F3 as an important target for therapeutic intervention in cancer ((Eiring *et al.*, 2008; He *et al*., 2008; Parisi *et al.*, 2007). In this study, we propose that small molecules that would block the DNA-binding site of E2F3 in a promising manner might have a significant benefit in preventing the oncogenic overexpression of genes regulated by E2F3. This approach can be extremely useful in the treatment of cancers in which overexpressed E2F3 plays a crucial role in the oncogenic axis.

Generally, structure-based virtual screening is reliable enough to identify drug candidates from a large chemical space of compound libraries, especially for projects based in academic laboratories, by reducing the costs of time and infrastructure needed for experimental screening (Ghosh *et al.*, 2006; Shoichet, 2004). It also has revolutionized the way drugs are designed by illustrating receptor-ligand interactions. As far as E2F transcription family is concerned, recently, Ma *et al.* (2008) used the computational approach based on virtual screening to study E2F-inhibitor recognition, thus, setting an example for its further application as an effective target in cancer

**Corresponding author:* e-mail: mahmood.mukhtar@gmail.com

†*Equally contributed to this work*

drug discovery a (Ma *et al.*, 2008) Here, we report homology modeling and an exploratory docking-based virtual screening study against DNA-binding domain of transcription factor E2F3. The ZINC (Irwin *et al.*, 2005) database was searched through a web-based service for identification of small molecules that might inhibit the interaction of E2F3 with DNA.

METHODOLOGY

Homology modeling

The experimental 3D structure of transcription factor E2F3 has not been solved up to now, so a homology model of its DAN-binding domain was generated by adopting the following appropriate steps.

Sequence retrieval and choice of template

The amino acid (aa) sequence of human transcription factor E2F3 (Accession code NP_001940) was retrieved from NCBI data base (Pruitt *et al.*, 2009), and submitted in the 'Domain Annotation' session of the SWISS-MODEL workspace for structural and functional domain interpretation (Arnold *et al.*, 2006; Bordoli *et al.*, 2009).

A gapped BLAST (Altschul *et al.*, 1997) query was carried out against SWISS-MODEL Template Library (SMTL) (Arnold *et al*., 2006) by using 'Template Identification' tool option of the SWISS-Model workspace. Then, an appropriate template was selected from the result hit list based on analysis of template coverage with respect to appropriate target domain.

Modeling by using SWISS-MODEL workspace

On the basis of target and the selected template similarity result by Blast search, the 'Automated mode' of SWISS-MODEL workspace was used for homology modeling. The amino acid sequence of E2F3 was submitted as FASTA format and the selected template was specified by its PDB identifier and chain ID after assuring that this ID is present in SMTL. After the completion of automated modeling step, the project file returned by server containing the sequence alignment of the template and modeled structure was analyzed as a 'DeepView Project' in order to check for any improvements.

Quality evaluation of the predicted model

The quality of returned model was verified with ANOLEA (Melo *et al.*, 1997), GROMOS (van-Gunsteren *et al.*, 1996), Verify3D (Eisenberg *et al.*, 1997), and PROCHECK (Laskowski *et al.*, 1993) programs provided under the 'Structure Assessment' tool option of the SWISS-MODEL workspace. Moreover, the modeled structure of the target protein was compared with the experimental structure of the template by loading them in the DeepView (Swiss-Pdb Viewer) (Guex *et al.*, 1997), and superposing by Cα-atoms. Finally, the average root mean square deviation (RMSD) value was noted to decide the quality of the generated model.

DNA-Binding site mapping

The present study was established on knowledge-based *in silico* analysis of the DNA binding site residues with respect to interactions with DNA bases, and then, to use that information for virtual screening and docking studies. So, on the basis of literature review, the conserved residues of the DNA-binding domain of template were mapped out in the model by sequence and structural alignment in DeepView.

Virtual screening by using Dock Blaster Server

Dock Blaster is a web-based virtual screening service (http://blaster.docking.org) that uses ZINC (Irwin-Shoichet, 2005) as a dockable database. The modeled 3D structure of target was submitted in Dock Blaster along with DNA binding site residues as a separate pdb file. The top 'hits' predicted by the server were further filtered by selecting only those molecules having molecular weight between 300 to 350, neutral charge, hydrogen bond donors 0/1, hydrogen bond acceptors 3/5, rotatable bonds 2/7 and a value of xLogP between -2 and 4. This helped to obtain a small library of 31 compounds (not shown) that were downloaded from server list in pdb format for further docking evaluations against the model.

Automated molecular docking studies

The concept of multi-step docking applications with energy barriers, introduced by Tuccinardi *et al.* (2007), was adopted for auspicating the interaction mode of each compound against the active DNA-binding site of E2F3 model. The program AutoDock 4.2 (Morris *et al.*, 2009) was used for computing the potential interaction of each compound with E2F3 model and the strength of association and binding affinity was predicted by using scoring functions like hydrogen bonding, binding energy and binding pose. A brief description of the docking studies criteria along with filtration and subsequent docking steps is shown in table 1.

Table 1: General outline of the docking steps by AutoDock as part of virtual screening procedure

Step performed	Compounds selected
Filtration[a] of compounds searched by Dock Blaster	31
10 runs of docking Compounds showing maximum H-bonding interaction with binding site residues and having binding energy less than -3.5 Kcal/mol.	6
30 runs of docking Compounds showing maximum H-bonding interaction and having energy less than -4.0 Kcal/mol.	3
Visual inspection with respect to binding mode against the binding site residues	2
50 runs of docking	2

[a] Net charge = 0, xLogP = -2/4, rotatable bonds = 2/7, Molecular weight = 300-350, H-bond donors = 0/1, H-bond acceptors = 3/5.

AutoDock Tools were used for creating PDBQT files for E2F3 model and each of the filtered compounds by adding polar hydrogen atoms and assigning partial atomic charges (Kollman for the modeled protein and Gasteiger for the ligands). Meanwhile, all the torsion angles of the ligands to be explored during the docking process were defined. A 3D grid of 40, 40, and 40 points in the x, y, and z directions with a grid spacing of 0.375Å was created by the AutoGrid algorithm (Morris *et al.*, 1998) and centered on target site residues. Whereas, a distance-dependent function of the dielectric constant was used to evaluate the energy maps between the potential inhibitors and the modeled protein.

Initially, all the compounds selected by filteration step were subjected to 10 runs of the AutoDock calculations, by using the Lamarckian genetic algorithm (LGA) (Morris *et al*., 1998), while all other parameters were used by default values. After cluster analysis, using an RMS tolerance of 1.0 Å (Tuccinardi *et al*., 2007), the compounds with at least one cluster showing maximum number of H-bond interaction specifically with the target residues with an average of estimated binding energy greater than -3.5 Kcal/mol were further subjected to 30 runs of docking studies. Then, the compounds with at least one cluster with an average of estimated binding energy greater than -4.0 Kcal/mol were visually inspected for their binding poses in order to check whether they are covering the target site residues promisingly. Furthermore, the compounds with best binding poses were selected and subjected to 50 runs of docking studies in order to check for any improvement in binding energy. Finally, the compound with best binding energy and binding pose was also docked against the template (E2F4) by 50 run AutoDock search.

RESULTS

In this study, we prepared the structural model of DNA-binding domain E2F3 based on its homology and then performed virtual screening to find the best small molecules that could block the DNA-binding site of E2F3 in a promising manner. These findings can be extremely useful in the treatment of cancers in which overexpressed E2F3 plays a crucial role in the oncogenic axis.

Identification of DNA-binding motif of E2F3

The InterPro scan ((Zdobnov *et al.*, 2001) for sequence annotation identified a functional domain (sequence range 178-243) containing a fold related to the winged helix DNA-binding motif, a characteristic of DNA-binding domains of E2F as well as DP family of transcription factors, which shows a sequence-specific DNA-binding activity (Zheng *et al.*, 1999). This was found to be in accordance with secondary structure prediction results by PsiPred (Jones, 1999), showing alpha-helices and extended beta-sheet portions in the domain residue range (data not shown).

Construction of the structural model of DNA-binding domain of E2F3

Our BLAST search for identification of a suitable structure template yielded a highly significant match with crystal structure of human transcription factor E2F4 having an X-Ray crystallographic resolution of 2.6 Å (ExPDB ID 1cf7A). This was the only template predicted in our target residue range of DNA-binding domain. The BLAST alignment showed 68% identity, 83% similarity, and the template was found to be extending almost the entire length of the target DNA-binding domain residues of E2F3 with only one gap (fig. 1a).

The homology model of DNA-binding domain of E2F3 generated by SWISS-MODEL workspace is shown in Fig. 1b. The analysis of the model in DeepView found it to be authentic one as there were no any residues making clashes with each other or backbone, no chain breaks and no reconstructed amino acids. Similar to that of the template, the model comprises typical winged-helix fold containing α-helices and β-sheet portions. The structural alignment of the Cα-traces of the model and template (fig.1c), and the backbone conformational analysis by superposition of the model over template in DeepView revealed no significant structural divergence. The location of the single gap, as detected in BLAST alignment, was further examined with respect to structural context and was found to be adjusted in a coil region without causing any disturbance in a helical or beta-sheet portion. The root mean square deviation (RMSD) value between equivalent Cα atoms was 0.42Å, which further confirms a very small fluctuation between the generated model and the experimental structure of template, thus, reflecting the presence of strong restraints in most regions of the model.

Evaluation of the structural model of DNA-binding domain of E2F3

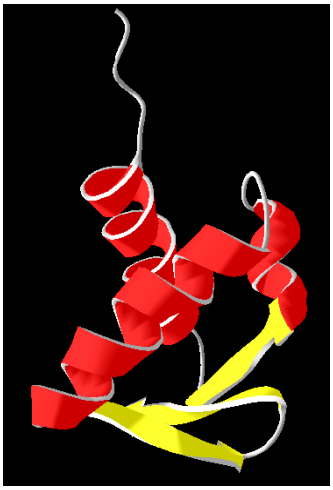
The evaluation of model by ANOLEA, GROMOS and Verify3D at SWISS-MODEL workspace did not detect any major problems. The regions with positive (red) values in ANOLEA and GROMOS plots were very small, while all the residues showed optimal values in Verify3D plot (data not shown) that signal towards a good quality. Similarly, the comparison of stereochemical evaluation of the generated model and template by PROCHECK (table 2) showed that Phi (φ) and Psi (ψ) torsion angle values of 86.4% residues of the generated model occupy the most favored (core) regions of the Ramachandran plot and only 1.7% of them invade the disallowed, whilst the figures for the template (1cf7A) were found to be 77% and 1.68%, respectively. The scattered plots of Phi (φ) and Psi (ψ) torsion angles in Ramachandran plot provide judgment about the stereochemical quality of a structured model; the structure is more correct if a lesser proportion of residues lies in disallowed regions (Pal *et al.*, 2002). As the proportion of residues in disallowed region was almost same for model and template (fig. 1d), and the

Table 2: PROCHECK validation parameters obtained for the model and template from Swiss-Model workspace.

Parameters	E2F3(Model)	1cf7A (Template)
Ramachandran plot values		
aa in core regions	86.4%	77%
aa in additional allowed regions	11.9%	18%
aa in generously allowed regions	0.0%	3.3%
aa in disallowed regions	1.7%	1.68%
PROCHECK		
Overall G factor	0.03	0.11
Bad contacts	1	5

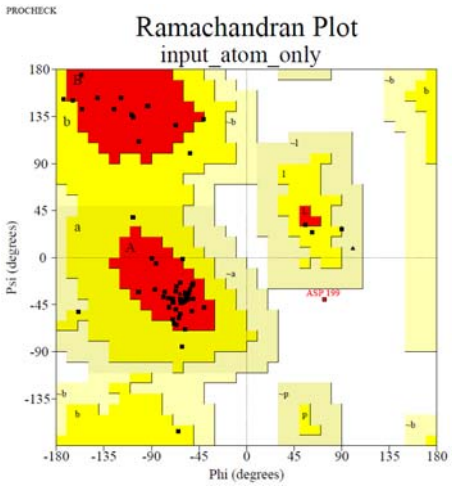
```
TARGET 177  TRYDTSLGLLTKKFIQLLSQSPDGVLDLNKAAEVLKV-QKRRIYDITNVLEGIHLIKKKSKNNVQW-
1cf7A  16   SRHEKSLGLLTTKFVSLLQEAKDGVLDLKLAADTLAVRQKRRIYDITNVLEGIGLIEKKSKNSIQWK
            .* . ****** **. ** .. ******. **. * * *************** ** *****..**
```

1a



1b

1c



1d

Fig. 1: Sequence and structural comparison of model and template. (a) The improved sequence alignment of the modeled structure (E2F3) and template (1cf7A) after analysis as project file in DeepView. In the sequences, an asterisk (*) indicates an identical or conserved residue and a stop (.) indicates similar substitutions. (b) 3-D homology model of DNA-binding domain of E2F3 in "Ribbon" model representation by DeepView. The structure consists of typical winged-helix fold similar to that of template, and is colored by "secondary structure schematics" (Red-α helices, Yellow-β sheets, Gray-coil). (c) Structural alignment of the model (E2F3) and template (E2F4; 1cf7A) by Cα trace. The model is colored cyan while template is colored yellow. The gap indicates less structural convergence at this position and complies with that of sequence alignment results. (d) Ramachandran plot for the model obtained by PROCHECK analysis at Swiss-Model workspace.

overall G-factor (Morris A. L. *et al.*, 1992) values of the model (0.03Å) and template (0.11 Å) were not making a much difference, it can be said that quality of our generated model is good.

In silico analysis of the crystal structure of E2F4 (PDB ID 1cf7) in DeepView revealed that two arginine residues of a highly conserved Arg56-Arg57-Xxx-Tyr59-Asp60 (RRXYD) sequence in α3 helix make critical contacts in the major groove of DNA. The side chain of each of the two arginines donates a pair of hydrogen bonds to the edges of guanine bases on opposite strands of DNA. The comparison of model and template by superposition revealed that highly conserved motif RRXYD in E2F3 model constitutes Arg216-Arg217-Xxx-Tyr219-Asp220. Since the two arginines of RRXYD motif (fig. 4) make critical contacts in the major groove of DNA; they were the main target of this study.

Virtual screening for the potential binding compounds

The information about scoring function values obtained during the first and second steps of docking-based virtual

Table 3: First and second step docking results

Comp. No.	10 Run Docking			30 Run Docking		
	Binding energy Kcal/mol	No. of H-bonds formed	[a]Residues involved in H-bonding	Binding energy Kcal/mol	No. of H-bonds formed	[a]Residues involved in H-bonding
1	-3.78	2	R 216, R 217			
2	-4.27	4	Q 214, K 215, R 216, R 217			
3	-3.69	3	R 216, R 216, R 217	-4.03	3	R 216, R 216, R 217
4	-2.52	2	R 216, R 217			
5	-3.77	1	R 217			
6	-1.92	2	R 216, R 217			
7	-3.57	2	R 216, R 217			
8	-3.77	3	R 216, R 217, R 217	-3.85	3	R 216, R 217, R 217
9	-3.97	1	R 216			
10	-3.07	3	R 216, R 217, R 217			
11	-3.80	3	R 216, R 217, R 217	-4.06	3	R 216, R 217, R 217
12	-3.80	2	R 216, R 217			
13	-3.20	2	R 216, R 217			
14	-3.06	2	R 216, R 217			
15	-3.98	2	R 216, R 217			
16	-3.54	2	R 216, R 217			
17	-2.9	2	R 216, R 217			
18	-2.84	2	R 216, R 217			
19	-4.03	3	R 216, R 217, R 217	-4.17	3	R 216, R 217, R 217
20	-2.92	2	R 216, R 217			
21	-4.12	4	Q 214, K 215, R 216, R 217			
22	-3.54	2	R 216, R 217			
23	-3.59	4	Q 214, K 215, R 216, N 223			
24	-3.16	2	R 216, R 217			
25	-3.62	3	R 216, R 217, R 217	-3.69	3	R 216, R 217, R 217
26	-3.5	2	R 216, R 217			
27	- 3.54	3	R 216, R 217, R 217	-3.80	3	R 216, R 217, R 217
28	-3.33	2	R 216, R 217			
29	-3.84	1	R 217			
30	-2.87	2	R 216, R 217			
31	-3.34	2	R 216, R 217			

[a] R = Arginine, Q = Glutamine, K = Lysine, N = Asparagine.

screening stuides is presented in table 3. After the first step of docking procedure (10 runs), all of the compounds established effective interaction with a wide range of binding energy (-1.92 Kcal/mol to -4.27 Kcal/mol) and hydrogen bonding (1 to 4). Moreover, almost all of the compounds were showing hydrogen bond interaction with the same arginine residues of DNA binding site that are involved in interaction with DNA bases, nevertheless, some of them also established additional interaction with the adjacent residues like Gln-214, Lys-215 and Asn-223. Being the essentials of drug-receptor interactions, multiple hydrogen bonds must be adequately established

between the drug and the receptor active site in order to confer proper stability because a single hydrogen bond, being relatively weak, is considered insufficient to endorse a stable interaction (Sahu *et al.*, 2008). Accordingly, as per criteria of first step of docking study (hydrogen bond interaction particularly with two binding site arginines = maximum, binding energy = -3.5 Kcal/mol, binding pose = cover the two arginines in a promising manner), only six compounds 3, 8, 11, 19, 25, 27 were selected as each of them was making three hydrogen bonds with the two arginines and the remaining were eliminated.

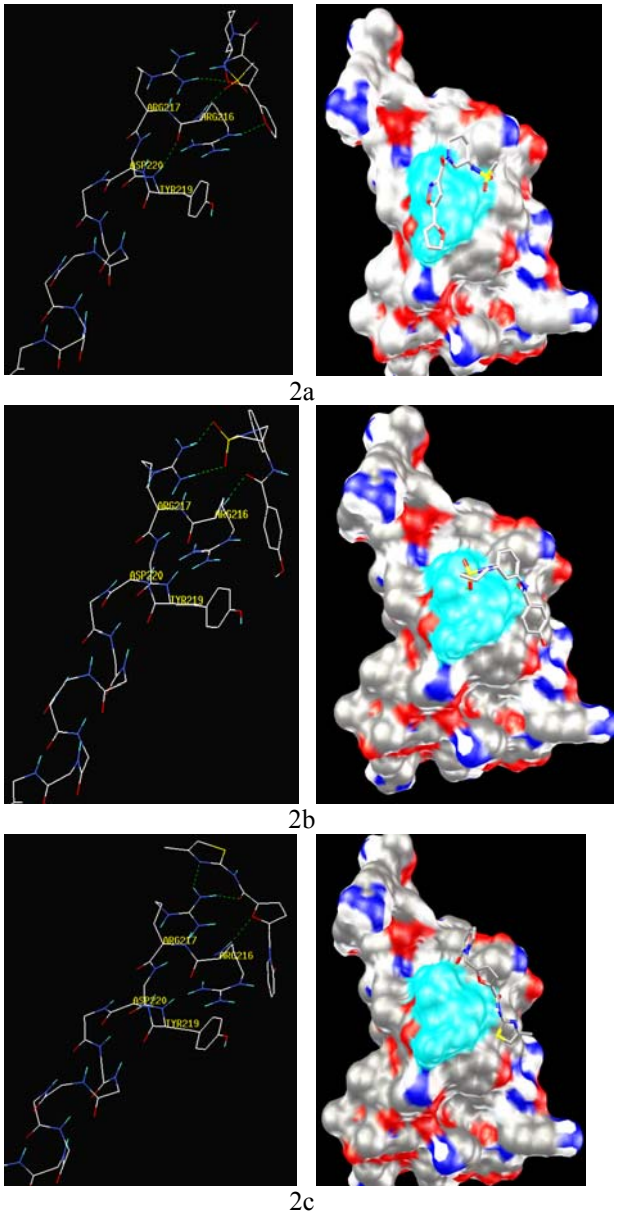
In the second step (30 runs docking), though all the six compounds demonstrated improvement in binding energy, compounds 8, 25, 27 were also eliminated as they could not satisfy the energy barrier limit (-4.0 Kcal/mol) for this step. However, the hydrogen bonding interaction was found complying with the first step without any improvement in number (table 3).

The visual inspection of the remaining three compounds 3, 11 and 19 (secondary structures depited in fig. 3) in terms of binding pose was very intriguing as illustrated in fig. 2. The data about their hydrogen bonding interaction with the target site residues of the model is shown in table 4. According to our docking experiments, oxygen atoms of furan and oxazole rings in compound 3 coordinated a hydrogen bond with -HE and -HH22 in the guanidine moieties of Arg-216 and Arg-217 respectively, while one of the oxygen atoms of methanesulfonamide moiety accepted H-bond from backbone nitrogen (NH) of Arg-216 (fig. 2a). All the bond lengths were within optimal range of hydrogen bonding (table 4). Overall, it seemed to make a better interaction as it was covering the binding site in a promising manner.

Table 4: Summary of amino acids and atoms involved in hydrogen bonding between E2F3 model and compounds 3, 11 and 19.

H-bonds formed Amino acid: atom…… atom: compound	Bond distance (Å)	Bond energy (Kcal/mol)
Arg 216: NH….O29 :3	1.988	-5.186
Arg 216: HE….O37 :3	1.791	-5.783
Arg 217: HH22….O5 :3	2.088	-4.339
Arg 216: NH….O17 :11	1.833	-7.391
Arg 217:HH12….O36 :11	2.142	-2.322
Arg217: HH22….O35 :11	1.882	-6.261
Arg 216: NH….O6: 19	2.19	-2.194
Arg 217:HH11….N22 :19	1.906	-0.403
Arg 217:HH12….O12 :19	2.145	-2.124

Likewise, compound 11 also appeared to be assuring the active site as it has covered one of the two active site residues by accepting two hydrogen bonds from side chain –NH1 (HH11) and –NH2 (HH22) of Arg-217 via the oxygen atoms of dioxidothiazolidin ring. In general, it could not provide a front cover to Arg-216 regardless of interaction with backbone nitrogen through carbonyl oxygen of methoxybenzamide (fig. 2b); however, the bond lengths were found satisfying the optimal limits (table 4), further ensuring it a good interaction.



2a

2b

2c

Fig. 2: Illustration of docked and binding modes of compounds against the targeted binding site residues of E2F3 DNA-binding domain model. The model is in "Wireframe" and "Molecular surface" representations and binding site residues are labeled in "Wireframe" and colored light blue in "Molecular surface", while the compounds are in "Wireframe" and "Stick and Balls" representation. The color scheme is CPK (Red-Oxygen, Blue-Nitrogen, Cyan-Hydrogen, Gray-Carbon, and Yellow-Sulfur); (a**)** E2F3 complexed with compound 3 (ZINC ID: 29198731); (b) E2F3 complexed with

compound 11 (ZINC ID: 5339121); (c) E2F3 complexed with compound 19 (ZINC ID: 29414722). The hydrogen bonds established by compounds can be seen as green dotted lines.

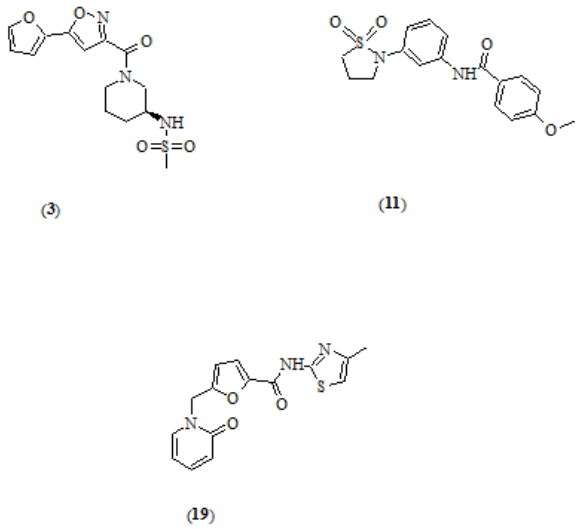
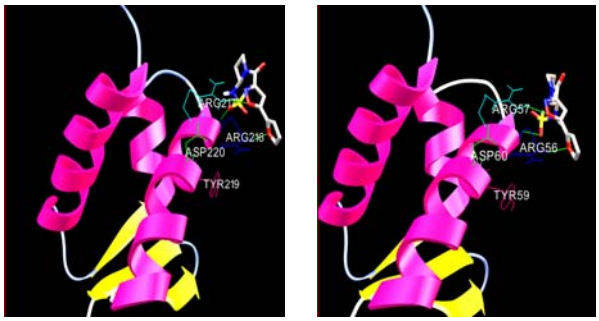


Fig. 3: Chemical structures of compounds 3, 11 and 19 which are potential DNA-binding site ligands for E2F3.

In a succinct comparison with compounds 3 and 11, compound 19 could not establish a desirable interaction in terms of binding mode. Even though it established three hydrogen bonds with optimal bond lengths; two with side chain guanidine moiety of Arg-217 and one with backbone nitrogen of Arg-216 (table 4), it had gone awry (fig. 2c), and was not expected to meet our criteria of analysis. Therefore, it could not be promoted for third step docking studies.

In the last step of docking studies (third step: 50 runs), compound 3 showed improvement in its binding energy from -4.03 Kcal/mol to -4.10 Kcal/mol, but the values for compound 11 remained the same as that in step 2. The inhibitory constant (Ki) values for compounds 3 and 11 were 0.988 mM and 1.06 mM, respectively. As the binding energy difference of compound 3 (-4.10 Kcal/mol) and 11 (-4.06 Kcal/mol) is very small and both are making equal number of hydrogen bonds (3) with the target site residues of E2F3 model, both can be considered for further *in vitro* evaluation.

Furthermore, 50 run docking studies of compound 3 with the template (E2F4) showed that binding energy, inhibitory constant and hydrogen bonding interaction values with the template (-4.21 Kcal/mol, 0.817mM and 3 respectively) do not make a much difference as that of the model (-4.10 Kcal/mol, 0.988 mM, 3). It, further, corroborates the better quality of generated homology model and results of this study (fig. 4).



Model (E2F3) Template (E2F4)

Fig. 4: Representation of the docked structure for compound 3 with model (E2F3) and template (E2F4). The model and template are represented by "Ribbon" model colored by "secondary structure schematics" while the ligand is in "Ball and Stick" colored by CPK. The residues of the conserved motif RRXYD are exposed in both model and template and are colored individually for distinction (Blue and Cyan for the two consecutive arginines respectively, Red-Tyr and Green-Asp). The three hydrogen bonds established by the ligand with two arginine residues can be clearly seen in green color.

DISCUSSION

Homology models have been found worthful for docking applications to search compound libraries for novel ligands (Hillisch *et al.*, 2004), although high target-template similarity is of prime importance to produce better results (Krovat *et al.*, 2005). In this study, homology modeling was used to generate three dimensional structure of DNA-binding domain of E2F3 transcription factor on the basis of E2F4 (template). The major emphasis was on exploring the potential inhibitors against E2F3 by using a docking-based virtual screening procedure via a server and software combination. We screened ZINC chemical database, containing 13 million compounds, with a Dock Blaster server ((Irwin *et al.*, 2009) search and the returned 5,000 compounds were further subjected to a filtration step on the basis of binding site analysis, preferably, selecting compounds having at least 3-5 hydrogen bond acceptors and minimum hydrogen bond donors (0-1) in their structure. Multi-step docking of a small library of selected compounds against the binding site residues of generated model screened two compounds (3 & 11) showing reasonable interaction in terms of docked energy and binding mode. The exploration of potential interactions between these novel agents and our generated model represents the first but very exciting step of our intended project towards the development of novel anti-cancer drugs against many forms of cancer characterized by transcription factor E2F3 overexpression. The next phase of this project is aimed at performing *in vitro* binding assays so that we could be able to take this first step towards validation of these potential interactions as

promising ones for the development of anti-cancer agents with a novel mode of action.

In principle, the driving influence in computer-aided drug discovery process, particularly the structure-based, has always been to grasp the nature of protein-protein and protein-ligand interactions to develop novel pharmaco-logical agents within a short period of time. It is a general conviction that prediction of perfect posing (conformation and orientation) of a ligand within a binding site can be of extreme importance to revolutionize the development of more selective and potent inhibitors against a specific target. High-throughput structure-based virtual screening through implementation of docking calculations provides an opportunity in pharmaceutical research to identify high efficiency ligands on the basis of posing and scoring functions (Kitchen *et al*., 2004). As such, multi-step docking and scoring techniques, though usually rely on approximations, are increasingly being coupled in virtual screening procedures for specific therapeutic targets, thus, contributing to useful proceedings in structure-based inhibitor discovery. The combined implementation of compound filtering/ranking and well-suiting scoring functions in concert with computational docking serves an important tool to screen small number of candidate compounds from chemical databases possessing fairly large and diverse chemical entities. The main objective of the homology model-based virtual screening procedure adopted in this study was to test the reliability of our generated model and the solidity of our docking screening procedure for searching compounds that show better interaction in terms of hydrogen bonding, minimum binding energy and the best binding pose, and the results obtained validated the system.

Stated simply, this approach can be highly useful to narrow down the search process to select a small number of compounds with best interaction modes. The outcomes of this study will encourage the researchers to use Dock Blaster server in order to perform structure-based virtual screening procedures within short time spans. To all intents and purposes, this study may give an impregnable perception to researchers for developing novel anti-cancer drugs against types of neoplasias with overexpressed E2F3.

REFERENCES

Adams MR, Sears R, Nuckolls F, Leone G and Nevins JR (2000). Complex transcriptional regulatory mechanisms control expression of the E2F3 locus. *Mol. Cell. Biol.,* **20**: 3633-3639.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*., **25**: 3389-3402.

Arnold K, Bordoli L, Kopp J and Schwede T (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics,* **22**: 195-201.

Attwooll C, Lazzerini Denchi E and Helin K (2004). The E2F family: Specific functions and overlapping interests. *EMBO J*., **23**: 4709-4716.

Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J and Schwede T (2009). Protein structure homology modeling using SWISS-MODEL workspace. *Nat. Protoc*., **4**: 1-13.

Cooper CS, Nicholson AG, Foster C, Dodson A, Edwards S, Fletcher A, Roe T, Clark J, Joshi A, Norman A, Feber A, Lin D, Gao Y, Shipley J and Cheng SJ (2006). Nuclear overexpression of the E2F3 transcription factor in human lung cancer. *Lung Cancer,* **54**: 155-162.

Eiring AM, Neviani P, Santhanam R, Oaks JJ, Chang JS, Notari M, Willis W, Gambacorti-Passerini C, Volinia S, Marcucci G, Caligiuri MA, Leone GW and Perrotti D (2008). Identification of novel posttranscriptional targets of the BCR/ABL oncoprotein by ribonomics: requirement of E2F3 for BCR/ABL leukemogenesis. *Blood,* **111**: 816-828.

Eisenberg D, Luthy R and Bowie JU (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Method Enzymol.,* **277**: 396-404.

Feber A, Clark J, Goodwin G, Dodson AR, Smith PH, Fletcher A, Edwards S, Flohr P, Falconer A, Roe T, Kovacs G, Dennis N, Fisher C, Wooster R, Huddart R, Foster CS and Cooper CS (2004). Amplification and overexpression of E2F3 in human bladder cancer. *Oncogene*. **23**: 1627-1630.

Foster CS, Falconer A, Dodson AR, Norman AR, Dennis N, Fletcher A, Southgate C, Dowe A, Dearnaley D, Jhavar S, Eeles R, Feber A and Cooper CS (2004). Transcription factor E2F3 overexpressed in prostate cancer independently predicts clinical outcome. *Oncogene*. **23**: 5871-5879.

Ghosh S, Nie A, An J and Huang Z (2006). Structure-based virtual screening of chemical libraries for drug discovery. *Curr. Opin. Chem. Biol.,* **10**: 194-202.

Guex N and Peitsch MC (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis,* **18**: 2714-2723.

He L, Yang H, Ma Y, Pledger WJ, Cress WD and Cheng JQ (2008). Identification of Aurora-A as a direct target of E2F3 during G2/M cell cycle progression. *J. Biol. Chem.,* **283**: 31012-31020.

Hillisch A, Pineda LF and Hilgenfeld R (2004). Utility of homology models in the drug discovery process. *Drug Discov. Today,* **9**: 659-669.

Humbert PO, Verona R, Trimarchi JM, Rogers C, Dandapani S and Lees JA (2000). E2f3 is critical for normal cellular proliferation. *Gene Dev.,* **14**: 690-703.

Irwin JJ and Shoichet BK (2005). ZINC – A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.,* **45**: 177-182.

Irwin JJ, Shoichet BK, Mysinger MM, Huang N, Colizzi F, Wassam P and Cao Y (2009). Automated docking screens: A feasibility study. *J. Med. Chem.,* **52**: 5712-5720.

Jones DT (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.,* **292**: 195-202.

Karamouzis MV, Gorgoulis VG and Papavassiliou AG (2002). Transcription factors and neoplasia: Vistas in novel drug design. *Clin. Cancer Res*., **8**: 949-961.

Kitchen DB, Decornez H, Furr JR and Bajorath J (2004). Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Rev. Drug Discov*., **3**: 935-949

Krovat EM, Steindl T and Langer T (2005). Recent Advances in Docking and Scoring. *Curr. Comput-Aid. Drug Des.,* **1**: 93-102.

Laskowski RA, MacArthur MW, Moss DS and Thornton JM (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, **26**: 283-291.

Lees JA, Saito M, Vidal M, Valentine M, Look T, Harlow E, Dyson N and Helin K (1993). The retinoblastoma protein binds to a family of E2F transcription factors. *Mol. Cell. Biol.,* **13**: 7813-7825.

Leone G, DeGregori J, Yan Z, Jakoi L, Ishida S, Williams RS and Nevins JR (1998). E2F3 activity is regulated during the cell cycle and is required for the induction of S phase. *Gene Dev.,* **12**: 2120-2130.

Libermann TA and Zerbini LF (2006). Targeting transcription factors for cancer gene therapy. *Curr. Gene Ther.,* **6**: 17-33.

Ma Y, Kurtyka CA, Boyapalle S, Sung SS, Lawrence H, Guida W and Cress WD (2008). A small-molecule E2F inhibitor blocks growth in a melanoma culture model. *Cancer Res.,* **68**: 6292-6299.

Melo F, Devos D, Depiereux E and Feytmans E (1997). ANOLEA: a www server to assess protein structures. *Proceedings.* International Conference on Intelligent Systems for Molecular Biology, *ISMB*. **5**: 187-190.

Morris AL, MacArthur MW, Hutchinson EG and Thornton JM (1992). Stereochemical quality of protein structure coordinates. *Proteins,* **12**: 345-364.

Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK and Olson AJ (1998). Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem*., **19**:1639-1662.

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS and Olson AJ (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.,* **30**: 2785-2791.

Nebert DW (2002). Transcription factors and cancer: an overview. *Toxicology,* **181-182**: 131-141.

Oeggerli M, Tomovska S, Schraml P, Calvano-Forte D, Schafroth S, Simon R, Gasser T, Mihatsch MJ and Sauter G (2004). E2F3 amplification and over-expression is associated with invasive tumor growth and rapid tumor cell proliferation in urinary bladder cancer. *Oncogene,* **23**: 5616-5623.

Pal D and Chakrabarti P (2002). On residues in the disallowed region of the Ramachandran map. *Biopolymers,* **63**: 195-206.

Parisi T, Yuan TL, Faust AM, Caron AM, Bronson R and Lees JA (2007). Selective requirements for E2f3 in the development and tumorigenicity of Rb-deficient chimeric tissues. *Mol. Cell. Biol.,* **27**: 2283-2293.

Pruitt KD, Tatusova T, Klimke W and Maglott DR (2009). NCBI Reference Sequences: Current status, policy and new initiatives. *Nucleic Acids Res.,* **37**: D32-36.

Sahu VK, Khan AKR, Singh RK and Singh PP (2008). Hydrophobic, Polar and Hydrogen Bonding Based Drug-Receptor Interaction of Tetrahydroimidazo-benzodiazepinones. *Am. J. Immunol.,* **4**: 33-42.

Sams-Dodd F (2005). Target-based drug discovery: Is something wrong? *Drug Discov. Today,* **10**: 139-147.

Shoichet BK (2004). Virtual screening of chemical libraries. *Nature,* **432**: 862-865.

Tsantoulis PK and Gorgoulis VG (2005). Involvement of E2F transcription factor family in cancer. *Eur. J. Cancer*, **41**: 2403-2414.

Tuccinardi T, Cascio MG, Marzo VD, Manera C, Ortore G, Saccomanni G and Martinelli A (2007). Structure-Based Virtual Screening: Identification of Novel CB2 Receptor Ligands. *Lett. Drug Des. Discov.,* **4**: 15-19.

van-Gunsteren WF, Billeter SR, Eising AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP and Tironi IG (1996). The GROMOS96 Manual and User Guide. Vol. VdF Hochschulverlag ETHZ, Zürich.

Zdobnov EM and Apweiler R (2001). InterProScan – An integration platform for the signature-recognition methods in InterPro. *Bioinformatics,* **17**: 847-848.

Zheng N, Fraenkel E, Pabo CO and Pavletich NP (1999). Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F-DP. *Gene Dev.,* **13**: 666-674.