IJPSR (2019), Volume 10, Issue 6 **(Research Article)**

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 09 October 2018; received in revised form, 19 December 2018; accepted, 30 December 2018; published 01 June 2019

SELECTING PROTEIN STRUCTURE/S FOR DOCKING-BASED VIRTUAL SCREENING: A CASE STUDY ON TYPE II INHIBITORS OF VEGFR-2 KINASE

H. R. Bhojwani and U. J. Joshi *

Department of Pharmaceutical Chemistry, Prin. K. M. Kundnani College of Pharmacy, Cuffe Parade, Mumbai - 400005, Maharashtra, India.

Keywords:

VEGFR-2 kinase, Crystal Structures, DFG-in/out confirmation, Type II inhibitors, Performance Indices, DUD Set Validation, Enrichment Studies

Correspondence to Author:
Dr. Urmila Joshi

Head of Department, Department of Pharmaceutical Chemistry, Prin. K. M. Kundnani College of Pharmacy, Cuffe Parade, Mumbai - 400005, Maharashtra, India.

E-mail: urmila.joshi1365@gmail.com

ABSTRACT: In this study, 36 crystal structures available with type I-V inhibitors of VEGFR-2 kinase in the RCSB PDB were classified into DFG-in/-out conformation using visual analysis and KLIFS database. The focus was on Type II inhibitors as most kinase inhibitors belong to this category. Therefore, the crystal structures with DFG-out confirmation with a type II inhibitor were selected depending on the resolution and r-free value. 11 selected crystal structures were subjected to self-docking studies and interaction analysis, leading to the elimination of one crystal structure *viz*. PDB id 3U6J. 10 crystal structures were subjected to cross-docking analysis. No crystal structures were eliminated at this stage as 50% ligands were docked accurately at RMSD cut off ≤ 2Å. These structures were further evaluated for screening performance by calculation of five performance indicating terms. A rank order was established by performance terms. The next stage of selection was the calculation of enrichment factor and assessment of the number of chemical classes retrieved after docking of the DUD set along with actives. Considering the EF values and the rank order of performance terms; 5 crystal structures were eliminated. Lastly, advanced enrichment parameters such as ROC, AUC, RIE, the average number of outranked decoys, and BEDROC were calculated for the remaining 5 structures. After considering all the stages of evaluation, 4ASE was identified as the most suitable crystal structure.

INTRODUCTION: Docking - based virtual screening (DBVS) is a method of choice for identification of chemically diverse hits when the 3 dimensions (3D) structures of the target are available [1]. The success of the docking-based virtual screening is sensitive to the choice of the 3D structure of the target [2].

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.10(6).2998-11

The article can be accessed online on www.ijpsr.com

DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.10(6).2998-11

When a limited number of crystal structures were available for a particular target, selection of a high-resolution crystal structure was a method of choice for the selection of crystal structure [3].

However, recent years have seen an explosion in the availability of the crystal structures of many "druggable" targets [4]. It is a well-known fact that the crystal structure of any target in a complex with a bound ligand represents a confirmation of the target, optimally adapted to accommodate that particular ligand. The flexibility of the target protein allows the target to adopt a different conformation in the presence of a chemically diverse ligand [5, 6].

DBVS in such cases is more likely to identify more hits belonging to the chemical class of co-crystallized ligand. Docking is a computationally intensive procedure; multiple runs of docking-based virtual screening using multiple crystal structures are difficult. Selection of appropriate crystal structure for virtual screening assumes importance in such a situation.

The vascular endothelial growth factor receptor type 2 (VEGFR-2), a tyrosine kinase linked receptor; represents one such case. It plays an important role in normal physiological processes such as cell proliferation, differentiation, migration, and angiogenesis, which makes VEGFR-2 an attractive target for cancer [7]. VEGFR-2 exists in two conformations, active and inactive depending on the orientation of the DFG - motif of the activation segment. In the 'DFG-in' or active conformation, the aspartate (Asp1046) residue orients toward the ATP binding cleft, and the phenylalanine (Phe1047) is buried in a hydrophobic pocket adjacent to the ATP site. In the 'DFG-out' or inactive conformation, the flipping of the DFG motif causes the phenylalanine side chain to occupy the ATP binding cleft, thus uncovering the hydrophobic pocket [7-10].

Different types of inhibitors occupy different sites and bind to a different conformation of VEGFR-2. Type I inhibitors bind to the DFG-in/active as well as DFG-out/inactive conformation of the receptor, whereas type II inhibitors bind to the inactive conformation only, thereby occupying the hydrophobic pocket in addition to ATP-binding site. Presently, most of the inhibitors of kinases belong to the type II category, and they impart selectivity as well [11]. There are other types of inhibitors, such as type $I_{1/2}$, which bind to an inactive conformation without occupying the hydrophobic pocket that is characteristic of Type II inhibitors [12]. Type III as well as IV inhibitors, do not compete with ATP and bind solely to allosteric pockets [13, 14]. Type V inhibitors bind to ATP-binding site and neighboring region like the type II inhibitor, however, they lock the kinase in the DFG-in form as reported [14].

There are several reports in the literature on the use of docking or docking-based virtual screening for the discovery of VEGFR-2 inhibitors [15-18].

None of the studies, take the confirmation of the DFG motif of the activation segment into consideration. This conformational analysis is important as the binding modes for different types of inhibitors differ and could lead to inaccurate results. The present study focuses on a systematic approach for selection of crystal structure/s for docking-based virtual screening of VEGFR-2 kinase inhibitors using Glide software with a special emphasis on the conformational complexities that are involved.

MATERIALS AND METHODS:

Classification and Selection of Crystal Structures: 36 crystal structures (resolution 1.5 - 2.95Å) of VEGFR-2 kinase domain bound to inhibitors have been deposited in the RCSB Protein Data Bank (PDB) and listed in the KLIFS database [19]. A series of sequential filters were applied for selecting the relevant crystal structures for the study. The first filter was selection of crystal structures with resolution < 2.5 Å and R-free value <0.25 [20, 21]. The second filter was the selection of only those crystal structures present in DFG-out conformation. These were identified on the basis of visual analysis and literature report. They were in complex with a Type II inhibitor as per literature. The sequential filtering procedure gave 11 crystal structures which were used for further study.

Detailed Analysis of Ligand Binding Mode using KLIFS Database: In addition to the structural information, the KLIFS database also reports the ligand binding mode. Ligand binding mode includes the details on the pockets and sub-pockets occupied. Therefore, pocket and sub-pocket occupancy were studied for the 11 selected crystal structures using the KLIFS database [19].

Receptor Preparation: 11 crystal structures (PDB id 1YWN, 2OH4, 3BE2, 3EWH, 3U6J, 3VHE, 3VNT, 3VO3, 3WZE, 4ASD, and 4ASE) were prepared using the protein preparation workflow in Maestro 10.2 [24]. Hydrogen atoms were added to the protein structures consistent with a pH of 7.4. Missing residues and loops were added. Since none of the crystal structures had water-mediated protein-ligand interactions, all the water molecules were deleted. Hydrogen atoms were also added to the co-crystallized ligand molecules followed by the generation of energetically accessible ionization

and tautomeric states using Epik 3.2 [25]. The protein assignment program in Maestro was used to set terminal rotamer states for Asn, Gln automatically, and His as well as tautomeric and protonation states of His to optimize the hydrogen-bonding network in the complex. Besides, hydroxyl and thiol torsions in Cys, Ser, and Tyr were optimized. The optimized protein-ligand complexes were subjected to a restrained minimization using the Imperf tool available in Protein Preparation Wizard of Maestro. An RMSD cutoff value of 0.3 Å was used.

Ligand Preparation: Two sets of VEGFR-2 ligands were used in this study: (1) Co-crystallized ligands from the VEGFR-2 PDB complexes listed above were used to access the ability of each receptor to dock the ligands accurately in cross-docking studies and (2) a collection of 42 VEGFR-2 kinase inhibitors selected from the literature [26-36] with IC_{50} values ranging from 0.035 nM to 930 nM belonging to diverse chemical classes were used for the performance indices calculation and database enrichment studies. The 42 VEGFR-2 kinase inhibitors were chosen in a manner such that most of the chemical classes of the type II inhibitors could be included. The DUD decoy set of 2906 molecules with ‘‘drug-like’’ characteristics and a molecular weight less than or equal to 500 was used for the database enrichment studies [37]. All ligands were prepared using LigPrep 3.4 [38] by generating low energy ionization and tautomeric states with a pH of 7.4. In the case of the co-crystallized ligands and any analogs in set 2, chiralities were retained from the input structure. All molecules subjected to LigPrep 3.4 were energy minimized using the OPLS 2005 force field.

Generation of Receptor Grid and Molecular Docking: Energy grids for docking studies were computed for each of the prepared protein-ligand complexes using default settings in the Receptor Grid Generation Tab in Glide 6.7 [39, 40]. The centroid of the co-crystallized ligand was used to define the center of each grid box. Default values were accepted for van der Waals scaling a partial input charges were used. Extra precision docking was used for docking accuracy, and standard precision docking runs for performance indices and enrichment studies, with default settings for all other parameters and no constraints or similarity scoring were applied [39, 41].

Docking Accuracy and Interaction-Based Analysis: To assess docking accuracy, all possible self and cross-docking studies were carried out with the extra precision (XP) scoring function. The final docked conformation of the inhibitor was aligned to the original conformation, and root means square deviation (RMSD) was calculated. After performing the self-docking studies, the interaction-based analysis was carried out wherein the hydrogen bonding and the other (π-π and π-cation as well as hydrophobic) interactions produced by the co-crystallized ligand after the self-docking studies were compared with the interactions reported in RCSB for that particular ligand. The crystal structures which did not reproduce hydrogen-bonding interactions were eliminated during self-docking studies, and remaining were subjected to cross-docking [42].

Calculation of Screening Performance Index and Enrichment Studies: Docking of the set 2 ligands (42 Type II VEGFR-2 kinase inhibitors mentioned earlier in the Ligand Preparation section) for the calculation of performance indices was performed on each of the 9 prepared PDB structures using the standard precision (SP) mode of Glide with default parameters into the shortlisted crystal structures [39]. Following this, the docking performance was evaluated by calculating five parameters referred to as term 1-5 using the mathematical equations as reported in the literature [43]. These five terms include most favorable docking energies, comparison of average docking energies of all actives with the docking energy of the best binder, number of actives giving favorable docking energies, number of actives successfully docked and the screening performance index (SPI).

Docking using the standard precision (SP) mode of Glide 6.7 was performed on each of the shortlisted PDB structures. Enrichment studies involved the calculation of classical and advanced enrichment parameters after the screening of the enriched decoy set. Standard procedures reported in the literature [44, 45] were used for the calculation of enrichment factor (at 1, 2, 5, 10, and 20% of the ranked dataset), receiver-operating characteristic (ROC) curve, area under the curve, average number of out-ranking decoys, robust initial enhancement (RIE), and Boltzmann-enhanced receiver operating curve (BEDROC) [44, 45].

RESULTS AND DISCUSSION:

Classification and Selection of Crystal Structures: The details of 36 crystal structures have been given in **Table 1** and **2**. Three filters were used for the selection of the crystal structures.

TABLE 1: A LIST OF CRYSTAL STRUCTURES OF VEGFR-2 KINASE USED IN THIS STUDY WITH THEIR CORRESPONDING RESOLUTION, R-FREE VALUE, CHAIN, DFG-CONFORMATION AND THE TYPE OF INHIBITOR (I, I1/2, III, AND V) HAS BEEN PROVIDED

S. no.	PDB ID	Resolution (Å)	Chain	R-value Free	Type of Inhibitor	DFG-conformation	
						as per KLIFS	as per a visual analysis
1	1Y6A	2.10	A	0.217	I	Out-like	Out
2	1Y6B	2.10	A	0.236	I	Out-like	Out
3	2P2H	1.95	A	0.229	I	In	In
4	2XIR	1.50	A	0.238	$I_{1/2}$	Out	Out
5	3B8R	2.70	A/B	0.259	$I_{1/2}$	In	In
6	3C7Q	2.10	A	0.279	I	Out	Out
7	3CJF	2.15	A	0.258	I	Out	Out
8	3CJG	2.25	A	0.253	I	In	In
9	3VHK	2.49	A	0.262	III	Out	Out
10	3VID	2.30	A	0.281	I	Out	Out
11	3WZD	1.57	A	0.212	V	In	In
12	4AG8	1.95	A	0.241	$I_{1/2}$	Out	Out
13	4AGC	2.00	A	0.251	$I_{1/2}$	Out-Like	Out
14	4AGD	2.81	A	0.307	I	Out-Like	Out

TABLE 2: A LIST OF CRYSTAL STRUCTURES OF VEGFR-2 KINASE USED IN THIS STUDY IS GIVEN WITH THEIR CORRESPONDING RESOLUTION, R-FREE VALUE, CHAIN, DFG-CONFORMATION WITH TYPE II INHIBITORS HAS BEEN PROVIDED

S. no.	PDB ID	Resolution (Å)	Chain	R-value Free	Type of Inhibitor	DFG-conformation	
						as per KLIFS	as per the visual analysis
1	1YWN	1.71	A	0.230	II	Not Mentioned	Out
2	2OH4	2.05	A	0.231	II	Out	Out
3	2P2I	2.40	A/B	0.266	II	Out	Out
4	2QU5	2.95	A	0.251	II	Out	Out
5	2QU6	2.10	A/B	0.272	II	Out	Out
6	2RL5	2.65	A	0.233	II	Out	Out
7	3B8Q	2.75	A/B	0.276	II	Out	Out
8	3BE2	1.75	A	0.226	II	Out	Out
9	3CP9	2.50	A/B	0.263	II	Out	Out
10	3CPB	2.70	A/B	0.286	II	Out	Out
11	3CPC	2.40	A/B	0.272	II	Out	Out
12	3DTW	2.90	A/B	0.279	II	Out	Out
13	3EFL	2.20	A/B	0.264	II	Out	Out
14	3EWH	1.60	A	0.234	II	Out	Out
15	3U6J	2.15	A	0.230	II	Out	Out
16	3VHE	1.55	A	0.209	II	Out	Out
17	3VNT	1.64	A	0.192	II	Out	Out
18	3VO3	1.52	A	0.182	II	Out	Out
19	3WZE	1.90	A	0.224	II	Out	Out
20	4ASD	2.03	A	0.230	II	Out	Out
21	4ASE	1.83	A	0.231	II	Out	Out
22	5EW3	2.50	A/B	0.254	II	Out	Out

The first filter used was the resolution of the crystal structure for selecting good quality structures only. The limit of filter value was set to < 2.5 Å which resulted in the elimination of eight crystal structures (2QU5, 2RL5, 3B8Q, 3B8R, 3CP9, 3CPB, 3DTW, and 4AGD). The next criterion used was R-free value <0.25, which led to the elimination of 13 crystal structures. The third criterion used for selecting crystal structures was the activation loop representing a DFG-out

confirmation with a Type II inhibitor. This procedure resulted in the elimination of 4 crystal structures (PDB ID: 1Y6A, 1Y6B, 2XIR, and 3WZD), as these crystal structures either do not represent DFG-out confirmation or do not possess a Type II inhibitor. 11 crystal structures with PDB IDs 1YWN, 2OH4, 3BE2, 3EWH, 3U6J, 3VHE, 3VNT, 3VO3, 3WZE, 4ASD, and 4ASE were selected for further study.

Ligand-Binding Mode Analysis: The ligand binding modes of the 11 shortlisted crystal structures were assessed using the KLIFS database, as shown in **Table 3**.

TABLE 3: GIVES THE DETAILS OF THE POCKET AND SUB-POCKET OCCUPANCY FOR 11 CRYSTAL STRUCTURES AS PER THE KLIFS DATABASE

S. no.	PDB ID	Pocket	Sub-pocket
1	1YWN	front, gate, back	AP, BP-IB, BP-II-out
2	2OH4	front, gate, back	AP, BP-I-B, BP-II-out, BP-III
3	3BE2	front, gate, back	AP, BP-I-A, BP-I-B, BP-II-out, BP-III, BP-V
4	3EWH	front, gate, back	AP, BP-I-A, BP-I-B, BP-II-out, BP-III
5	3U6J	front, gate, back	AP, BP-I-A, BP-I-B, BP-II-out, BP-III
6	3VHE	front, gate, back	AP, BP-I-B, BP-II-out, BP-III
7	3VNT	front, gate, back	AP, BP-I-B, BP-II-out, BP-III
8	3VO3	front, gate, back	AP, BP-I-B, BP-II-out
9	3WZE	front, gate, back	AP, BP-I-B, BP-II-out, BP-III
10	4ASD	front, gate, back	AP, BP-I-B, BP-II-out, BP-III
11	4ASE	front, gate, back	AP, BP-I-B, BP-II-out

As per the KLIFS classification of the binding site for kinases, a kinase inhibitor can occupy front cleft, gate area, and back cleft. These regions are further divided into the sub-pockets. The front cleft consists of adenine pocket (AP) and front pocket-I (FP-I). The front pocket-II (FP-II) occupies both the front cleft and part of the gate region. The gate area connects the front cleft with the back cleft. Back pocket I-B (BP-I-B) is located in the center of the gate area, whereas Back pocket I-A (BP-I-A) is located at the top of the gate area.

The back pockets IA and IB (BP-IA and BP-IB) are accessible to ligands binding to both in both DFG-in and -out conformations of the receptor. Type II ligands, binding to the DFG-out conformation of the receptor, can access some of the additional sub-pockets in the gate region and back cleft. This is due to flipping of the phenylalanine residue of DFG-motif. These include a large pocket back pocket –II out (BP-II out), back pocket III (BP-III) located in the basement of BP-II-out and two more back pockets *viz*. BP-IV and BP-V. This representation from the KLIFS database on the sub-pocket occupancy has been provided in the literature [19]. These observations highlight the importance of considering the conformation of DFG-motif as well as pocket occupancy in available crystal structures when selecting them for virtual screening experiments for a particular type of the inhibitor. An insight into the ligand binding mode for PDB ID 2OH4 shows that the ligand occupies AP in the front cleft along with interactions with the hinge region. BP-IB, BP-II out, and BP-II, I were occupied by the ligand in the gate area and the back cleft. The interactions for 2OH4 ligand were found to be with Cys919 (hinge region) as well as Glu885 and Asp1046 (BP-II out). Similar ligand binding mode was observed in the case of 3VHE, 3VNT, 3WZE, and 4ASD. In crystal structures with PDB id 3EWH and 3U6J, the ligands were found to occupy the hinge region, AP, BP-I-A, BP-I-B, BP-II-out, and BP-III. 3BE2 ligand occupies the hinge region, AP, BP-I-A, BP-I-B, BP-II-out, BP-III, and BP-V. 1YWN, 4ASE, and 3VO3 ligands were found to occupy BP-IB and BP-II out. All interactions, as observed for each co-crystallized ligand have been mentioned in **Table 4**. This difference in the binding mode for the type II inhibitors was an important factor in the study since the shortlisted crystal structures covered an entire range of possibilities for the type II ligand binding.

Self-Docking, Interaction-Based, and Cross-Docking Analysis: RMSD between a generated docking pose and the co-crystallized ligand conformation represents the most established benchmark for the ability of docking to predict the protein-bound ligand conformation [47]. **Table 5**

summarizes the results of self-docking of all the cognate ligands into their native protein structures. 10 of 11 cognate ligands had an RMSD ≤ 1Å while 3BE2 had RMSD of 1.52. It is not only important that the native ligands are docked accurately, but also they must have a similar binding-pattern reproduced as they possess in their native crystal structure. For this purpose, an interaction-based analysis was conducted wherein the hydrogen bonds, hydrophobic interactions, π-π, and π-cation interactions formed by the ligand with the receptor in the crystal structure, were studied. This assessment was based on the fact [42] that correctly docked ligands will reproduce correct hydrogen-bonding and the other interactions with the protein. If the hydrogen-bonding interactions were represented accurately, but the other interactions differed moderately, the docking pose was considered 'nearly correct' and accepted. However, if the hydrogen bonding interactions were not reproduced, then that pose is incorrect. 11 crystal structures were subjected to interaction-based analysis. The result for interaction-based analysis has been given in **Table 5**. 1YWN, 2OH4, 3BE2, 3VHE, 3VO3, and 4ASE; reproduced a correct pose. 3EWH, 3WZE, 3VNT, and 4ASD could reproduce a nearly correct pose. 3U6J did not reproduce a correct or a nearly correct pose. It was observed that the hydrogen bonding interactions with Cys919 and Asp 1046 were not reproduced for 3U6J ligand post the self-docking study although it had an acceptable RMSD value of 0.26. Taking this into consideration, 3U6J was eliminated from the study at this stage.

TABLE 4: INTERACTIONS AS OBSERVED IN THE 11 SELECTED CRYSTAL STRUCTURES DEPOSITED IN THE PROTEIN DATA BANK

S. no.	PDB ID	Hydrogen Bonding				Other interactions	
		Glu 917	Cys 919	Glu 885	Asp 1046	Hydrophobic	Pi-Pi and Pi-cation
1	1YWN	*	*	*	*	Val 916, Phe 918, Ile 890, Ile 1017, Leu 840	
2	2OH4		*	*	*	Leu1035, Phe1047, Val916, Ala866, Asp1046	
3	3BE2		*	*	*	Thr916, Phe918, Leu840, Val848, Ala866, Lys868, Leu889, Asp1046, Phe1047	pi-pi (Phe1047)
4	3EWH		*	*		Thr916, Leu840, Val848, Ala866, Lys868, Leu889, Asp1046, Phe1047	pi-cation (Lys868) pi-pi (Phe1047)
5	3U6J		*		*	Ile892, Leu1019, Phe918, Leu840, Ala866, Leu889, Asp1046, Phe1047	pi-pi (Phe1047)
6	3VHE		*	*	*	Phe918,Val848,Cys919, Leu840, Ala866, Leu889, Asp1046, Phe1047, Cys1045	
7	3VNT		*	*	*	Val916, Ala866 Leu889, Lys868 Cys1045, Leu1035 Asp1046	pi-cation (Lys868)
8	3VO3		*	*	*	Ala866, Val916, Lys868, Leu1035	
9	3WZE		*	*	*	Val848, Val916, Leu1035, Ala866, Cys1045, Phe1047	
10	4ASD		*	*	*	Val848, Phe918, Leu1035, Ala866, Cys1045, Lys868 Leu840, Phe1047	
11	4ASE		*	*	*	Val848, Phe918, Leu1035, Ala866, Cys1045, Lys868 Leu840, Phe1047	

* Indicates hydrogen bonding interaction was observed for a co-crystallized ligand with that particular amino acid residue.

TABLE 5: INDICATES THE RMSD VALUES FOR COGNATE LIGAND OF EACH CRYSTAL STRUCTURE OBTAINED AFTER SELF-DOCKING STUDIES AND DETAILS OF THE INTERACTION-BASED ANALYSIS FOR 11 SELECTED CRYSTAL STRUCTURES

S. no.	PDB ID	Self-docking RMSD	H-bonding	Other interactions
1	1YWN	0.24	Reproduced	Reproduced
2	2OH4	0.21	Reproduced	Reproduced
3	3BE2	1.52	Reproduced	Reproduced
4	3EWH	0.22	Reproduced	Not Reproduced
5	3U6J	0.26	Not Reproduced	Reproduced
6	3VHE	0.23	Reproduced	Reproduced
7	3VNT	0.29	Reproduced	Not Reproduced
8	3VO3	0.18	Reproduced	Reproduced
9	3WZE	0.09	Reproduced	Not Reproduced
10	4ASD	0.29	Reproduced	Not Reproduced
11	4ASE	0.11	Reproduced	Reproduced

The structures that have the best ability to dock non-native ligands with a lower RMSD are possibly are more successful in virtual screening [47]. This was investigated by cross-docking studies for type II inhibitors of the VEGFR-2 kinase. The remaining 10 crystal structures were subjected to cross-docking studies. Each of the non-native ligands was docked into each of the protein structure and RMSD calculated. The calculated RMSD values for cross-docking studies are as shown in **Table 6**. The average RMSD values were found to be 0.52, 0.92, and 0.94 for 3EWH, 1YWN, and 4ASE; respectively. Other crystal structures had average RMSD values higher than 1. 3VHE had the highest median RMSD value while 3EWH had the lowest. For shortlisting the crystal structures at this stage, the cut-off value for RMSD was set to value of ≤ 2Å. 3EWH could cross-dock all the ligands accurately at cutoff ≤ 2Å. Furthermore, all the crystal structures could dock at least 5 out of 9 non-native ligands at cutoff ≤ 2Å.

Additionally, the performance was evaluated at RMSD cutoff ≤ 3Å, 4Å, and 5Å in addition to above mentioned one and the results have been given in **Table 6**. 3VHE was the only crystal structure to cross-dock at RMSD ≤ 5Å while 2OH4, 3VNT, 3VO3, and 3WZE cross-docked 1 ligand each at RMSD ≤ 4Å. Since RMSD ≤ 2Å is most widely used cut off and as all the crystal structures could cross-dock more than 50% (*viz*. 5 out of 9) ligands accurately at this cut off; none of them was eliminated at this stage. Furthermore, similar to self-docking, the interactions after the cross-docking studies were analyzed.

A close look at the RMSD values for all the ligands revealed that 3BE2 and 3EWH ligands had higher values in most cross-docking screens. As mentioned earlier, only 3BE2 and 3EWH ligands occupied IA BP- and BP-IA and BP-V, respectively, in addition to the hinge region, AP, FP, BP-IB, BP-II out and BP-III occupied in most other cases. It can be presumed that 3BE2 and 3EWH ligands result in induced fit effects such that they occupy these additional pockets. Thus, the use of other crystal structures can lead to binding of these ligands in a pose that is different than their native poses. Therefore, this could result in high RMSD values for these two ligands in the cross-docking screens. To further verify that the docked pose for 3BE2 and 3EWH differs from the native one, an inspection of the interactions made by these ligands in the non-native crystal structures was undertaken. It was observed that in most of the cases, all the parent hydrogen bond interactions characteristic for the type II binding were not reproduced when docked into non-native structures. Additionally, it was observed that except for 3BE2 ligand, either partial or complete hydrogen bond interactions were reproduced in all crystal structures.

TABLE 6: RMSD VALUES OF THE 9 CO-CRYSTALLIZED LIGANDS IN THE CROSS-DOCKING STUDIES PERFORMED FOR 10 CRYSTAL STRUCTURES OF VEGFR-2 KINASE

	PDB ID									
Ligand	**1YWN**	**2OH4**	**3EWH**	**3BE2**	**3VHE**	**3VNT**	**3VO3**	**3WZE**	**4ASD**	**4ASE**
1YWN		2.19	0.47	0.57	2.30	1.74	ND	2.40	1.46	1.79
2OH4	0.34		0.37	1.22	1.92	0.69	0.50	0.52	0.51	0.75
3EWH	1.33	2.34		1.33	2.55	2.76	2.64	2.30	2.52	2.51
3BE2	0.53	3.59	1.08		4.15	3.85	3.67	3.83	2.63	1.94
3VHE	0.89	0.37	0.35	2.05		0.40	0.36	0.33	0.45	0.40
3VNT	2.40	0.76	0.93	0.76	2.68		0.74	0.69	0.57	0.40
3VO3	1.23	0.37	0.27	0.21	0.49	0.13		0.36	0.24	0.29
3WZE	0.26	0.27	0.21	1.43	0.79	0.19	0.81		0.18	0.12
4ASD	0.51	0.33	0.19	1.43	0.72	0.34	0.44	0.67		0.23
4ASE	0.80	0.85	0.83	1.32	0.94	0.83	0.89	0.24	0.83	
Minimum RMSD	0.26	0.27	0.19	0.57	0.49	0.13	0.36	0.24	0.18	0.12
Maximum RMSD	2.40	3.59	1.08	1.43	4.15	3.85	3.67	3.83	2.63	2.51
Average	0.92	1.23	0.52	1.15	1.84	1.21	1.26	1.26	1.04	0.94
Cross-docking RMSD										
Cross-docked at < 2Å	8	6	9	8	5	7	6	6	7	8
Cross-docked at <3Å	1	2	0	1	3	1	1	2	2	1
Cross-docked at < 4Å	0	1	0	0	0	1	1	1	0	0
Cross-docked at < 5Å	0	0	0	0	1	0	0	0	0	0

ND: Not Docked

Calculation of Screening Performance Index: Screening performance index calculations are based on the docking results of actives only and were carried out for all the 10 crystal structures, and results have been mentioned in **Table 7**. Calculation of term-1 was done intending to identifying the structure that binds ligands with the most favorable docking energies [43]. These structures are proposed to be more likely to select other true actives and to reject false positives.

In simple terms, term 1 measured how favorable the best docking energy within a receptor structure was in comparison to the best docking energies to all structures. The results indicated that 3WZE docked an active with the most favorable docking energy (-15.397 kcal/mol) and a term 1 value of 1 in comparison to the 9 other crystal structures. It can also be seen that 1YWN had the least value for term 1 and it could dock an active with most favorable docking energy of -11.351 kcal/mol.

TABLE 7: TERM 1, 2, 3, AND 5 VALUES OBTAINED AFTER CALCULATION FOR 10 CRYSTAL STRUCTURES

PBD ID	Lowest Docked Energy (kcal/mol)	Average Docking Energy (kcal/mol)	Term 1	Term 2	Term 3	Term 4	Term 5	Average Rank
1YWN	-11.350	-9.759	0.737 (10)	0.859 (2)	0.542 (1)	NC	0.261 (10)	5.75
2OH4	-14.096	-10.721	0.915 (4)	0.760 (7)	0.254 (7)	NC	0.619 (3)	5.25
3BE2	-12.339	-9.713	0.801 (9)	0.787 (5)	0.523 (2)	NC	0.380 (7)	5.75
3EWH	-12.447	-10.215	0.808 (8)	0.820 (3)	0.386 (3)	NC	0.357 (9)	5.75
3VHE	-12.557	-11.189	0.815 (7)	0.891 (1)	0.362 (4)	NC	0.714 (1)	3.25
3VNT	-13.565	-10.773	0.881 (5)	0.794 (4)	0.295 (5)	NC	0.547 (4)	4.5
3VO3	-13.401	-9.995	0.870 (6)	0.745 (8)	0.245 (8)	NC	0.380 (7)	7.25
3WZE	-15.397	-10.639	1 (1)	0.691 (10)	0.185 (10)	NC	0.404 (5)	6.5
4ASD	-14.334	-10.322	0.930 (3)	0.720 (9)	0.217 (9)	NC	0.404 (5)	6.5
4ASE	-14.360	-11.275	0.932 (2)	0.785 (6)	0.276 (6)	NC	0.642 (2)	4

Numbers in brackets indicate the rank obtained for that term value. NC: Not Calculated

Term-2 and Term-3 used the average docking energies and the average deviation of the docking energies from the most favorable one to evaluate whether compounds other than the one with the most favorable docking energy also possessed favorable docking energies - the more compounds with docking energies close to the most favorable one the better [43]. Comparison of average docking energies of all actives with the docking energy of the best binder (Term-2) indicated that if the average docking energy of all the actives to a crystal structure is closer to the docking energy of the best binder, more actives are bound with favorable energies to this structure, thereby suggesting that this structure could pick out actives more readily [43]. It was observed that the term 2 value was the maximum for 3VHE for which the lowest docking energy/most favorable energy and the average docking energies of 42 actives were -12.557 kcal/mol and -11.189 kcal/mol, respectively. Likewise, the term 2 values for other crystal structures were found to be as follows: 1YWN (0.859), 3EWH (0.82), 3VNT (0.794), 3BE2 (0.787), 4ASE (0.785), 2OH4 (0.76), 3VO3 (0.745), 4ASD (0.72), and 3WZE (0.691).

A value nearer to 1 indicates the lesser difference between the lowest docking energy and the average energy for the 42 active compounds. Term-3 is qualitatively similar to Term-2 but quantitatively different. It is based on the rationale that if more actives give favorable docking energies as the best binder, this structure is more likely to pick out actives [43]. 3BE2 had the highest value for term 3, while 3WZE had the least (0.185). These values indicate that 3BE2 is likely to pick out more actives and is followed by 3EWH, 3VHE, 3VNT, 4ASE, 2OH4, 3VO3, and 4ASD. Term-4 penalized structures to which fewer compounds could be docked successfully [43]. Term 4 was not calculated as all 42 actives were successfully docked into 10 selected crystal structures, and therefore, none of the crystal structures could be penalized.

The calculation of the number of actives giving favorable docking energies (term-5) was based on the assumption that if many actives can dock to a structure with docking energies more favorable than the overall average docking energies to all structures, this structure might be more likely to pick out many actives in virtual screening [43]. The overall average docking energy considering 10 structures were found to be -10.460 kcal/mol. During the calculations, we observed that 3VHE could dock 30 out of 42 actives with favorable energies greater than the overall average docking

energy (-10.460 kcal/mol). Likewise, the number of actives (out of 42) that could be docked with better energies by other crystal structures in comparison to overall average docking energy were as follows: 4ASE (27), 2OH4 (26), 3VNT (23), 4ASD (17), 3WZE (18), 3VO3 (16), 3BE2 (16), 3EWH (15), and 1YWN (11). Thus, the screening performance index *viz.* term 5 was the maximum for 3VHE (0.71), and the other crystal structures followed in the same order.

Furthermore, the crystal structures were ranked for their values obtained for each term so that it could be understood which crystal structure had better performance when considering all the performance indices. The ranking order observed based on the average ranks was 3VHE, 4ASE, 3VNT, 2OH4, 1YWN, 3BE2, 3EWH, 3WZE, 4ASD, and 3VO3, as shown in **Table 7**.

The performance indices calculations indicated the ability of the crystal structures to identify the actives. However, in the virtual screening process, the ability of a crystal structure to reject the false positives is equally important. Given this, we decided to subject each of these crystal structures to docking-based virtual screening using a decoy set enriched with the same 42 actives. In the next step, 10 crystal structures were evaluated for their comparative ability to discriminate between active and inactive during docking protocol. This was done by use of a decoy set, which is a set of compounds that are presumed to be inactive. The largest publicly accessible database of decoys is the Directory of Useful Decoys, which was used in the present study along with 42 actives.

Enrichment Studies:

Calculation of Standard and Advanced Enrichment Parameters: Enrichment studies provide useful insights into the screening efficiency of a crystal structure. They quantify the number of active compounds found in the hit list, concerning the fraction of inactive. The success of virtual screening is correlated with its ability to rank the active compounds at high positions of the hit list since only the first fraction of a hit list will be screened experimentally.

Enrichment Factor: The first enrichment parameter calculated during the enrichment studies was the enrichment factor at 1, 2, 5, 10, and 20% of the ranked data set [44, 45]. The theoretical maximum enrichment factors at 1, 2, 5, 10, and 20% that could be obtained were 70.19, 49.97, 20.05, 9.99, and 5.00, respectively. The values of enrichment factors for 10 PDB structures are as shown in **Table 8**.

The results of the calculation indicated that six PDB structures *viz.* 1YWN, 2OH4, 3BE2, 3EWH, 3VHE, 3VNT, and 4ASE could produce EF1% greater than 50. The remaining three crystal structures produced EF1% values of 38.72 and 41.15. The values of enrichment factors at 2, 5, 10, and 20% of the ranked dataset also showed similar trends with 3VO3, 3WZE, and 4ASD remaining at the bottom.

TABLE 8: ENRICHMENT FACTOR VALUES CALCULATED AT VARIOUS PERCENTAGES OF THE RANKED DATASET FOR 10 SELECTED CRYSTAL STRUCTURES

PDB	EF1%	EF2%	EF5%	EF10%	EF20%	Number of Chemical Classes in 1% of the dataset
1YWN	65.35	32.12	14.32	7.37	3.92	8
2OH4	65.35	35.69	16.23	8.33	4.28	8
3BE2	55.67	27.36	13.85	7.61	4.40	7
3EWH	55.67	35.69	16.23	8.56	4.40	7
3VHE	62.93	35.69	14.80	7.61	4.04	8
3VNT	55.67	27.36	12.89	7.12	3.69	9
3VO3	38.72	19.03	7.64	5.23	2.74	6
3WZE	41.15	22.60	10.50	5.47	3.57	6
4ASD	41.15	21.41	10.50	5.47	3.09	8
4ASE	60.51	34.50	15.28	8.80	4.64	9

Since, chemical diversity is an important aspect in identifying hits in virtual screening, the number of chemical classes to which the actives belonged was identified for 1% of the ranked data set. It was seen that 3VNT and 4ASE identified actives belonging to 9 chemical classes. 1YWN, 2OH4, 3VHE, and 4ASD could identify actives belonging to 8 chemical classes while 3EWH and 3BE2 identified 7 chemical classes. The active compounds belonging to 6 chemical classes were

retrieved by 3VO3 and 3WZE. Therefore, by results of lower ranking in both screening performance and values of enrichment factors; 3BE2, 3EWH, 3VO3, 3WZE, and 4ASD were eliminated from the study.

ROC Curves: ROC curves allow a visual comparison of the ability of the crystal structures to discriminate the actives and decoys reflected in the form of sensitivity and specificity pairs. **Fig. 1** represents the ROC curve for 1YWN, 2OH4, 3VHE, 3VNT, and 4ASE. Usually, the ROC curve representing ideal distributions will be the one where there is no overlap between the scores of active molecules and decoys. The ideal ROC curve continues as a horizontal straight line to the upper-right corner where all actives and all decoys are retrieved, which corresponds to sensitivity = 1 and specificity = 0.

In contrast to that, the ROC curve for a set of actives and decoys with randomly distributed scores tends towards the Se = 1-Sp line asymptotically with an increasing number of actives and decoys. This represents a random performance and is reflected as a diagonal. The ROC curves for all 5 crystal structures in this study indicated that they performed better than random screening. ROC curve for 4ASE depicts that the curve starts from the origin and closely follows the y-axis till the point where Se = 0.8 and after which it begins to drift upwards and right with subsequent retrieval of actives and decoys.

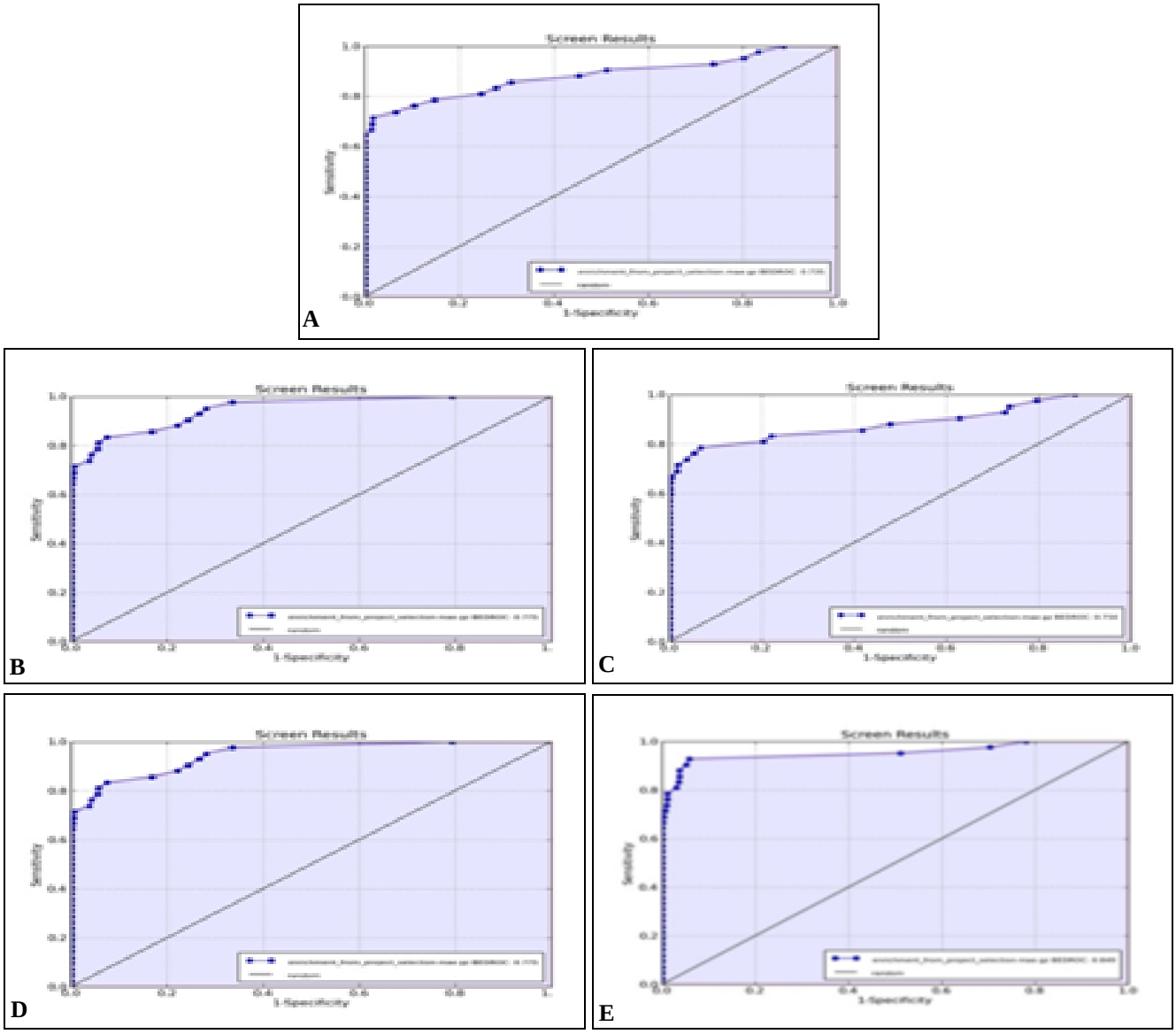

FIG. 1: REPRESENTS ROC CURVES OBTAINED FOR 5 SELECTED CRYSTAL STRUCTURES WHERE (A) 1YWN, (B) 2OH4, (C) 3VHE, (D) 3VNT, AND (E) 4ASE

Following this, when the Se > 0.9 at a particular point, only decoys were retrieved, and hence the curve moved only rightwards. The remaining actives were retrieved after this. When the last activity was identified, the Se = 1 and 1-Sp was slightly lesser than 0.8. Likewise, the ROC curves were analyzed for the other 4 crystal structures.

However, it was difficult to differentiate between the crystal structures using ROC curves only.

The area under the ROC Curve: In the ROC context, the area under the ROC curve (AUC) measures the performance numerically and can provide insights for quantitative comparison. A general guide for classifying the accuracy of screening is as follows: $0.9 \leq AUC \leq 1$ is excellent; $0.80 \leq AUC < 0.9$ is good; $0.70 \leq AUC < 0.8$ is fair; $0.50 \leq AUC < 0.7$ is poor; and $AUC < 0.5$ is a failure [44, 45]. AUC values have been mentioned in **Table 9.**

TABLE 9: GIVES THE ADVANCED ENRICHMENT PARAMETER VALUES FOR 5 SELECTED CRYSTAL STRUCTURES

Enrichment Parameters	1YWN	2OH4	3VHE	3VNT	4ASE
AUC	0.87 (good)	0.94 (excellent)	0.87 (good)	0.85 (good)	0.95 (excellent)
Ave. Number of outranking decoys	375	179	366	444	158
RIE	12.79	13.49	13.05	10.95	14.77
BEDROC(alpha=20.0, alpha*Ra=0.2849)	0.735	0.775	0.750	0.629	0.849

An Average Number of Outranking Decoys: The rank of each action is adjusted by the number of outranking actives. The number of outranking decoys is then defined as the adjusted rank of that active minus one. The number of outranking decoys is calculated for each docked active and averaged. The average number of outranking decoys was calculated in the same manner for the 5 crystal structures. It was seen that this value was the least for 4ASE. 4ASE was followed by 2OH4, 3VHE, 1YWN, and 3VNT, as shown in **Table 9**.

From the results, it was evident that 4ASE docked fewer decoys higher than the actives when compared with other 4 crystal structures. However, the classical enrichment parameters such as enrichment factors, ROC and AUC suffer from the problem of "early recognition".

In other words, these parameters do not distinguish high ranked active molecules from actives ranked at the end of a rank-ordered list.

In other words, two crystal structures that differ in the ability to rank the highest scored active molecules at the beginning of such an ordered list, but show the same enrichment for active molecules, would be assessed to perform equal and therefore, it was decided to compare them based on advanced enrichment parameters. Therefore, it was important to compare these 5 crystal structures from advanced enrichment descriptors such as Robust Initial Enhancement (RIE) and Boltzmann-Enhanced Discrimination of ROC (BEDROC).

Robust Initial Enhancement: RIE is an advanced enrichment metric which quantitatively indicates the ability of a ranking method to achieve a distribution of actives better than a method performing randomly. The RIE values for selected crystal structures which have been given in **Table 9**. An RIE value of greater than 1 indicates performance better than random. In our study, all four crystal structures produced RIE greater than 1.Since, the aim was to identify the best crystal structure to carry out virtual screening for type II VEGFR-2 inhibitors, and a crystal structure that performs exceptionally well will be of importance. 4ASE had the highest RIE of 14.77, followed by 2OH4, 3VHE, 1YWN, and 3VNT with 13.49, 13.05, 12.79, and 10.95, respectively.

The Boltzmann-Enhanced Discrimination of ROC (BEDROC): This metric is related to α-values. The α-value is a value that contributes to θ% of the total score at z% of the rank. The α-value of 20 indicates that 80% of the maximum contribution comes from the first 8% of the list, thereby ensuring the measurement of early recognition. If we keep the maximum contribution value to be constant at 80%, then the α-value of 20 will indicate this contribution coming from 8% of the list, and for comparing any two structures an important criterion is $\alpha Ra << 1$ [47].

Table 9 gives the BEDROC values for crystal structures under consideration. Taking into consideration, the criterion for αRa, BEDROC value was considered at α = 20 where αRa =

0.2849. At this stated criterion, it was found that BEDROC value was the maximum for 4ASE followed by 2OH4, 3VHE, 1YWN, and 3VNT for 8% of the ranked data set. On the basis of results obtained for self- and cross-docking studies, interaction analysis, performance indices as well as classical and advanced enrichment metrics, it could be inferred that 4ASE performed consistently and was the most appropriate crystal structure for virtual screening of type II VEGFR-2 inhibitors.

Comparative Analysis with Previous Studies: 1Y6A, 2P2H, 2QU5, 2RL5, 3C7Q, 3CJG, and 3CJF that were used in the study conducted by Planes as and co-workers [17] have not been included in our study after applying the selection criteria. PDB 1YWN was selected by them as the working structure after taking into account the crystallographic resolution and the analysis of the docking results obtained for each PDB, to use for VEGFR-2 docking-based virtual screening. Likewise, 1YWN was identified as one of the top three crystal structures by Zhang and co-workers [18]. In the present study, 1YWN performed fairly but was not identified as the best one taking the conformational state and type II inhibitors into consideration. The conformational analysis, along with the type of inhibitor, is an important consideration in the case of kinases, as mentioned earlier. However, the lack of conformational analysis (DFG-in/out) is fairly evident in the results given by Zhang and co-workers [18]. Amongst the 31 crystal structures included in their study, 2P2H, 3CJG, and 3B8R are present in DFG-in conformation. The count is given for ligands cross-docked at RMSD ≤ 2Å for 2P2H, 3CJG, and 3B8R are 5, 4, and 8; respectively.

Additionally, there are structures present in DFG-out form but in complex with other types of inhibitors. The number of ligands cross-docked at RMSD≤ 2Å in each of these cases is as follows: 1Y6A (4), 1Y6B (4), 3C7Q (2), 3CJF (4), 3VHK (6), 3VID (4), 4AGC (5) and 4AGD (1). This indicates the incapability of these cavities to accommodate chemically diverse type II ligands. These results highlight the importance of considering the conformations in the beginning, not only to have the correct data being taken ahead but also to avoid intensive computational procedures. It is essential to mention that the conformational consideration, along with the type of inhibitors, will significantly change statistical RMSD results for all the structures in the study reported by Zhang and co-workers. In comparison to Zhang's study, 3B8Q was eliminated at the primary stage due to its resolution while 3EWH was eliminated at a later stage owing to low performance in comparison to other structures as evident from enrichment factor and number of chemical classes retrieved. Lastly, in the present study, an evenhanded chance was given to all the structures by considering them at each stage and eventually eliminating the poor performers after the enrichment factor and chemical class considerations.

CONCLUSION: A methodical study of arriving at the most appropriate crystal structure of VEGFR-2 which can be used in docking-based virtual screening to identify Type II inhibitors was undertaken. The methods serially employed classification of crystal structures, ligand binding mode analysis, self-docking, interaction-based analysis, cross-docking, docking of known actives followed by calculation of screening performance index, docking of enriched decoy set and calculation of enrichment parameters; both classical and advanced. Taking the results of all the studies into consideration, it is proposed that 4ASE is the most promiscuous structure that can be used for the docking-based virtual screening studies of type II inhibitors. Lastly, this work is a more accurate representation of a systematic approach that can be applied for selection of crystal structure for virtual screening from limitless data available in a protein data bank for the difficult targets (along with a particular type of inhibitor which is particularly, applicable to kinases).

ACKNOWLEDGEMENT: The authors acknowledge the Department of Biotechnology (Project File No. BT /PR14373 /MED/30/ 530/2010) for sanction of grant

CONFLICT OF INTEREST: Both authors have none to declare.

REFERENCES:

1. Tuccinardi T: Docking-based virtual screening: recent developments. Comb Chem High Throughput Screening 2009; 12(3): 303-14.
2. Erickson JA, Jalaie M, Robertson DH, Lewis RA and Vieth M: Lessons in molecular recognition: the effects of

ligand and protein flexibility on molecular docking accuracy. J Med Chem 2004; 47(1): 45-55.
3. Anderson AC. The process of structure-based drug design: Chem Biol 2003; 10(9): 787-97.
4. Andricopulo AD, Salum LB and Abraham DJ: Structure-based drug design strategies in medicinal chemistry. Curr Top Med Chem 2009; 9(9): 771-90.
5. Rueda M, Bottegoni G and Abagyan R: Recipes for the selection of experimental protein conformations for virtual screening. J Chem Inf Model 2009; 50(1): 186-93.
6. Teague SJ: Implications of protein flexibility for drug discovery. Nat Rev Drug Discovery 2003; 2(7): 527-41.
7. Neufeld G, Cohen T, Gengrinovitch S and Poltorak Z: Vascular endothelial growth factor (VEGF) and its receptors. FASEB J 1999; 13(1): 9-22.
8. Roskoski Jr R: VEGF receptor protein–tyrosine kinases: Structure and regulation. Biochem Biophys Res Comm 2008; 375(3): 287-91.
9. Holmes K, Roberts OL, Thomas AM and Cross MJ: Vascular endothelial growth factor receptor-2: structure, function, intracellular signaling and therapeutic inhibition. Cell Signal 2007; 19(10): 2003-12.
10. Treiber DK and Shah NP: the ins and outs of kinase DFG motifs. Chem Biol 2013; 20(6): 745-46.
11. Muller S, Chaikuad A, Gray NS and Knapp S: The ins and outs of selective kinase inhibitor development. Nature Chemical Biology 2015; 11(11): 818-21.
12. Huang L, Huang Z, Bai Z, Xie R, Sun L and Lin K: Development and strategies of VEGFR-2/KDR inhibitors. Future Med Chem 2012; 4(14): 1839-52.
13. Wu P, Nielsen TE and Clausen MH: Small-molecule kinase inhibitors: an analysis of FDA-approved drugs. Drug Discovery Today 2016; 21(1): 5-10.
14. Okamoto K, Ikemori-Kawada M, Jestel A, Von Koonig K, Funahashi Y, Matsushima T, Tsuruoka A, Inoue A and Matsui J: Distinct binding mode of multikinase inhibitor lenvatinib revealed by biochemical characterization. ACS Medicinal Chemistry Letters 2014; 6(1): 89-94.
15. Ai G, Tian C, Deng D, Fida G, Chen H, Ma Y, Ding L and Gu Y: A combination of 2D similarity search, pharmacophore, and molecular docking techniques for the identification of vascular endothelial growth factor receptor-2 inhibitors. Anti-cancer Drugs 2015; 26(4): 399-09.
16. Kar RK, Suryadevara P, Sahoo BR, Sahoo GC, Dikhit MR and Das P: Exploring novel KDR inhibitors based on pharmaco-informatics methodology. SAR QSAR Environ Res 2013; 24(3): 215-34.
17. Planesas JM, Claramunt RM, Teixidó J, Borrell JI and Perez-Nueno VI: Improving VEGFR-2 docking-based screening by pharmacophore post filtering and similarity search postprocessing. J Chem Inf Model 2011; 51(4): 777-87.
18. Zhang Y, Yang S, Jiao Y, Liu H, Yuan H, Lu S, Ran T, Yao S, Ke Z, Xu J and Xiong X: An integrated virtual screening approach for VEGFR-2 inhibitors. J Chem Inf Model 2013; 53(12): 3163-77.
19. Vijayan RS, He P, Modi V, Duong-Ly KC, Ma H, Peterson JR, Dunbrack Jr RL and Levy RM: Conformational analysis of the DFG-out kinase motif and biochemical profiling of structurally validated type II inhibitors. J Med Chem 2014; 58(1): 466-79.
20. Van Linden OP, Kooistra AJ, Leurs R, De Esch IJ and De Graaf C: KLIFS: a knowledge-based structural database to navigate kinase-ligand interaction space. J Med Chem 2013; 57(2): 249-77.
21. Wlodawer A, Minor W, Dauter Z and Jaskolski M: Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. The FEBS Journal 2008; 275(1): 1-21.
22. Kleywegt GJ: Validation of protein crystal structures. Acta Crystallogr Sect D: Biol Crystallogr 2000; 249-65.
23. Lintnerova L, Garcia-Caballero M, Gregan F, Melichercik M, Quesada AR, Dobias J, Lac J, Salisova M and Bohac A: A development of chimeric VEGFR2 TK inhibitor based on two ligand conformers from PDB: 1Y6A complex–Medicinal chemistry consequences of a TKs analysis. Eur J Med Chem 2014; 72: 146-59.
24. Maestro, version 10.2, Schrodinger, LLC, New York, NY, 2015.
25. Epik, version 3.2, Schrödinger, LLC, New York, NY, 2015.
26. Dai Y, Guo Y, Frey RR, Ji Z, Curtin ML, Ahmed AA, Albert DH, Arnold L, Arries SS, Barlozzari T and Bauch JL: Thienopyrimidine ureas as novel and potent multitargeted receptor tyrosine kinase inhibitors. J Med Chem 2005; 48(19): 6066-83.
27. Dai Y, Hartandi K, Ji Z, Ahmed AA, Albert DH, Bauch JL, Bouska JJ, Bousquet PF, Cunha GA, Glaser KB and Harris CM: , Discovery of N-(4-(3-Amino-1 H-indazol-4-yl) phenyl)-N ‘-(2-fluoro-5-methyl phenyl) urea (ABT-869), a 3-Aminoindazole-Based Orally Active Multitargeted Receptor Tyrosine Kinase Inhibitor. J Med Chem 2007; 50: 1584-97.
28. Hasegawa M, Nishigaki N, Washio Y, Kano K, Harris PA, Sato H, Mori I, West RI, Shibahara M, Toyoda H and Wang L: Discovery of novel benzimidazoles as potent inhibitors of TIE-2 and VEGFR-2 tyrosine kinase receptors. J Med Chem 2007; 50(18): 4453-70.
29. Frey RR, Curtin ML, Albert DH, Glaser KB, Pease LJ, Soni NB, Bouska JJ, Reuter D, Stewart KD, Marcotte P and Bukofzer G: 7-Aminopyrazolo 1, 5-a pyrimidines as potent multitargeted receptor tyrosine kinase inhibitors. J Med Chem 2008; 51(13): 3777-87.
30. Kubo K, Shimizu T, Ohyama SI, Murooka H, Iwai A, Nakamura K, Hasegawa K, Kobayashi Y, Takahashi N, Takahashi K and Kato S: Novel potent orally active selective VEGFR-2 tyrosine kinase inhibitors: synthesis, structure-activity relationships, and antitumor activities of n-phenyl-n ‘-{4-(4-quinolyloxy) phenyl} ureas. J Med Chem 2005; 48(5): 1359-66.
31. Wang C, Gao H, Dong J, Zhang Y, Su P, Shi Y and Zhang J: Biphenyl derivatives incorporating urea unit as novel VEGFR-2 inhibitors: Design, synthesis and biological evaluation. Bioorg Med Chem 2014; 22(1): 277-84.
32. Curtin ML, Frey RR, Heyman HR, Sarris KA, Steinman DH, Holmes JH, Bousquet PF, Cunha GA, Moskey MD, Ahmed AA and Pease LJ: Isoindolinone ureas: a novel class of KDR kinase inhibitors. Bioorg Med Chem Lett 2004; 14(17): 4505-09.
33. Matsumoto S, Miyamoto N, Hirayama T, Oki H, Okada K, Tawada M, Iwata H, Nakamura K, Yamasaki S, Miki H and Hori A: Structure-based design, synthesis, and evaluation of imidazo 1, 2-b pyridazine and imidazo 1, 2-a pyridine derivatives as novel dual c-Met and VEGFR2 kinase inhibitors. Bioor Med Chem 2013; 21(24): 7686-98.
34. Liu L, Siegmund A, Xi N, Kaplan-Lefko P, Rex K, Chen A, Lin J, Moriguchi J, Berry L, Huang L and Teffera Y: Discovery of a potent, selective, and orally bioavailable c-Met inhibitor: 1-(2-Hydroxy-2-methyl propyl)-N-(5-(7-methoxyquinolin-4-yloxy) pyridine-2-yl)-5-methyl-3-oxo-2-phenyl-2, 3-dihydro-1 H-pyrazole-4-carboxamide (AMG 458). J Med Chem 2008; 51(13): 3688-91.
35. Liu L, Norman MH, Lee M, Xi N, Siegmund A, Boezio AA, Booker S, Choquette D, D’Angelo ND, Germain J

and Yang K: Structure-based design of novel class II c-Met inhibitors: 2. SAR and kinase selectivity profiles of the pyrazolone series. J Med Che 2012; 55(5): 1868-97.
36. Musumeci F, Radi M, Brullo C and Schenone S: Vascular endothelial growth factor (VEGF) receptors: drugs and new inhibitors. J Med Chem 2012; 55(24): 10797-22.
37. Lagarde N, Zagury JF and Montes M: Benchmarking data sets for the evaluation of virtual ligand screening methods: review and perspectives. Journal of chemical information and modeling. 2015; 55(7): 1297-07.
38. LigPrep, version 3.4, Schrodinger, LLC, New York, NY, 2015.
39. Glide, version 6.7, Schrödinger, LLC, New York, NY, 2015.
40. Joshi AK, Gadhwal MA and Joshi UJ: Identification of potential novel EGFR inhibitors using a combination of pharmacophore and docking methods. International Journal of Pharmacy and Pharmaceutical Sciences. 2015; 7(6): 77-91.
41. Bhojwani HR and Joshi UJ: Pharmacophore and Docking Guided Virtual Screening Study for Discovery of Type I Inhibitors of VEGFR-2 Kinase. Current Computer-Aided Drug Design 2017; 13(3): 186-07.
42. Kroemer RT, Vulpetti A, McDonald JJ, Rohrer DC, Trosset JY, Giordanetto F, Cotesta S, McMartin C, Kihlen M and Stouten PF: Assessment of docking poses: interactions-based accuracy classification (IBAC) versus crystal structure deviations. J Chem Inf Comput Sci 2004; 44(3): 871-81.
43. Huang Z and Wong CF: Inexpensive method for selecting receptor structures for virtual screening. J Chem Inf Model 2015; 56(1): 21-34.
44. Kirchmair J, Markt P, Distinto S, Wolber G and Langer T: Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection-What can we learn from earlier mistakes? J Comp Aided Mol Des 2008; 22(3-4): 213-28.
45. Braga RC and Andrade CH: Assessing the performance of 3D pharmacophore models in virtual screening: how good are they? Curr Top Med Chem 2013; 13(9): 1127-38.
46. Ramezani M and Shamsara J: A cross-docking study on matrix metalloproteinase family. Anti-Inflammatory Anti-Allergy Agents Med Chem 2015; 14(3): 164-71.
47. Truchon JF and Bayly CI: Evaluating virtual screening methods: good and bad metrics for the "early recognition" problem. J Chem Inf Model 2007; 47(2): 488-08.

How to cite this article:
Bhojwani HR and Joshi UJ: Selecting protein structure/s for docking-based virtual screening: a case study on type II inhibitors of VEGFR-2 kinase. Int J Pharm Sci & Res 2019; 10(6): 2998-11. doi: 10.13040/IJPSR.0975-8232.10(6).2998-11.