Homology modeling, active site prediction, and targeting the anti hypertension activity through molecular docking on endothelin – B receptor domain

Daddam Jayasimha Rayalu[1]*, Chandrabose Selvaraj[2], Sanjeev Kumar Singh[2], Ramakrishan Ganeshan[3], Nagapatla Udaya Kumar[1] & Panthangi Seshapani[4]

[1]Department of Bioinformatics, Global Institute of Biotechnology, 3-6-276/2, Himayat Nagar, Hyderabad-29, Andhra Pradesh, India; [2]Computer Aided Drug Design and Molecular Modeling Laboratory, Department of Bioinformatics, Alagappa University, Karaikudi-630003, Tamilnadu, India; [3]Department of Microbiology, P.K.N arts and science college, P.K.N. Nagar, Vidathakulam Road, Tirumangalam -625 706; [4]Department of Microbiology, Sri Venkateswara University, Tirupati - 517 502, Andhra Pradesh, India; Daddam Jayasimha Rayalu – Email: jaimscbiotech2007@gmail.com; *Corresponding author

Abstract:
In cardiovascular system, activation of Endothelin receptors causes vasoconstriction which leads to Pulmonary Arterial Hypertension (PAH). Endothelin receptor antagonism has emerged as an important therapeutic strategy in pulmonary arterial hypertension. Bosentan is intended to affect vasoconstriction, hypertrophic and fibrotic effects by blocking the actions of receptors ET_A and ET_B. In this study we identified the action of Bosentan on endothelin B receptor using docking studies with homology modeled endothelin B receptor. Through the modeled protein, the flexible Docking study was performed with Bosentan and its derivatives with theoretically predicted active sites. The results indicated that amino acid ARG82, ARG84 and HIS197 present in endothelin B receptor are core important for binding activities and these residues are having strong hydrogen bond interactions with Bosentan. We have investigated the Bosentan and its derivatives interactions and scoring parameters using gold docking package. Among the docked compounds, one of the Bosentan derivatives BD_6 shows better interaction than Bosentan with endothelin B receptor. Our results may be helpful for further investigations in both in vivo and in vitro conditions.

Background:
Pulmonary Arterial Hypertension (PAH) is a severe disease of Pulmonary arteries involving vasoconstriction, pulmonary vascular remodeling, and insitu thrombosis [1]. PAH develops the blood flow through the pulmonary arteries and the right side of the heart is under increasing strain to pump blood through to the lungs, leads to breathlessness, chest stiffness and fatigue [2]. Laboratory and clinical investigations have clearly shown that Endothelin receptors are over expressed in several

forms of pulmonary vascular disease and play a significant pathogenetic role in the development and progression of pulmonary vasculopathy [3]. There are two distinct receptors for the endothelin family of peptides, endothelin receptor A (ET_A) and endothelin receptor B (ET_B) and these two receptors have unique locations and binding affinities for the endothelin peptides [4]. The ET_A receptors are expressed on pulmonary vascular smooth muscle cells, whereas ET_B receptors are located on both pulmonary vascular endothelial cells and smooth

muscle cells. ET$_B$ receptors have an important role in the ET-dependent protracted presser effects. Under some circumstances it actually contributes to pulmonary vasoconstriction, through a population of ET$_B$ receptors located on vascular smooth muscle cells [5]. Bosentan is the first oral endothelin receptor antagonist (ERA) that blocks the effect of endothelin receptors. It belongs to a class of highly substituted pyrimidine derivatives, with no chiral centers. It is designated chemically as 4-tert-butyl-N - [6-(2-hydroxy-ethoxy) -5- (2-methoxy-phenoxy) - [2, 2] - bipyrimidin-4-yl] - benzene sulfonamide monohydrate [6]. *In vitro* functional experiments performed on animal and human tissues showed that Bosentan behaves as a competitive antagonist on ET receptors. It is intended to inhibit vasoconstriction, hypertrophic and fibrotic effects by blocking the actions of receptors (ET$_A$ and ET$_B$) which leads to reduction of blood pressure in lungs. Bosentan interact with the binding of ET-1 and other ET peptides to both ET$_A$ and ET$_B$ receptors [7]. The objective of this work is to identify the role the mechanism of Bosentan on ET$_B$ receptor and to model the protein, and to identify the interactions with Bosentan and their derivatives.

Methodology:

The sequence information of Endothelin-B receptor-*Homosapiens* (Accession No: P24530) is retrieved from Swiss Prot (http://www.swissprot.org) and submitted to SBASE server (http:// hydra icgeb.trieste.it/sbase) for domain prediction. The predicted domain (118-386) is search for similarity (Protein blast) and the related protein structure was identified by the BLAST program against PDB. The macromolecular structure (PDB ID 3C9M [8]), which shows maximum identity with high score and less e-value designated as template. The 3D model was generated by using the academic version of MODELER9v7 (http//:www.salilab.org/modeler), based on the information obtained from sequence alignment [9]. The 3D structure obtained from modeler were further refined through molecular dynamics and equilibration methods using NAMD 2.5 with CHARMM27 force field and the system contains TIP3P water model. The energy of the structure was minimized with 10,000 steps with cutoff of 12 Å for vdW interactions. The equilibrated system was simulated with 1 picoseconds (PS); to hold back the original conformation, for that (NPT) constant pressure and 310K constant temperatures has been choosed [10]. Finally, the structure having the least energy with low RMSD was used for further studies. The final structure obtained, was analyzed by Ramachandran's map drawn using PROCHECK v.3.0, and environment profile using ERRAT graph (Structure Evaluation server). The model satisfying all the parameters after evaluation was considered for the further process [9]. Theoretical Active site prediction of Endothelin-B receptor was predicted using CASTP server. CASTP identifies and measures pockets and pocket mouth openings, as well as cavities. Here we input the modeled protein for predicting the ligand binding sites and the CASTP server predicts the amino acids crucial for binding interactions [11]. The Bosentan library was generated manually by addition of polar groups in the required position and 30 new compounds were generated as derivatives of Bosentan **Table 1 (see supplementary material)**. The Docking interactions are done with GOLD (Genetic Optimization of Ligand Docking), based on genetic algorithm which is used for docking of Bosentan and their derivative compounds [12]. The interaction

of these derivatives with the active site residues are thoroughly studied using molecular mechanics calculations. The parameters used for GA were population size 100, selection pressure 1.1, and operator parameters for crossover, mutation and migration were set to 100, 100 and 10 respectively. Default cutoff values of 3.0 X (dH-X) for hydrogen bonds and 6.0 X for vdW were employed. During docking, the default algorithm speed was selected and the ligand binding site in the Endothelial B receptor was defined within a 10 A° radius with the centroid as CE atom of ASP81. The number of poses for each inhibitor was set 100, and early termination was allowed if the top three bound conformations of a ligand were within 1.5X RMSD [13]. After docking, the individual binding poses of each ligand were observed and their interactions with the protein were studied.
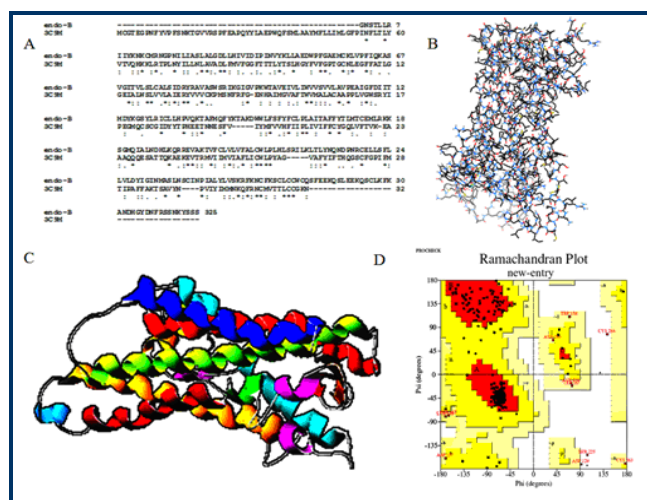

Figure 1: Modeling of Endothelin B receptor with sequence aligned **(A)** with 3C9M morphed **(B)** to build 3D model structure **(C)** and the model structure Ramachandran's plots **(D)**

Results:

Homology Modeling of Endothelin – B receptor Domain

A high level of sequence identity should guarantee a more accurate alignment between the target sequence and template structure, and here 3C9M (Chain A, Structure of a mutant bovine Rhodopsin in hexagonal crystal form protein) with Endothelin-B receptor have high sequence identity and the identity of the reference protein with the Endothelin B receptor domain is 36%. Structurally conserved regions (SCRs) for the model and the template were determined by multiple sequence alignment (A) superimposition of template and model structure (B), 3D structure of modeled protein (C) and Ramachandran's plot (D) are illustrated in **Figure 1**. In this study, the model is made up of residues 118 - 386 because the active domain region was identified in between these residues by the SBASE server. Coordinates from the reference protein (3C9M) to the SCRs, structurally variable regions (SVRs), N-termini and C-termini were assigned to the target sequence based on the satisfaction of spatial restraints (C).

Superimposition of 3C9M with Endothelin- B receptor domain

This model was refined by molecular dynamics method and the final stable structure of the Endothelial B receptor is

superimposed with C$^\alpha$ trace of template. The weighted root mean square deviation of C$^\alpha$ trace between the template and final refined models was 1.81 Å. This final refined model was used for the identification of active site and for docking of the substrate with the protein. This protein has the structural alignment with 13 helices and 2 sheets like a covered barrel shape. The overall scores indicates acceptable protein environment. After the refinement process, validation of the model was carried out using Ramachandran plot calculations computed with the PROCHECK program. Altogether 86.7% of the residues of Endothelin B receptor were in favored and allowed regions. The overall PROCHECK G-factor of Endothelin-B receptors was – 0.16 and VERIFY 3D environment profile was good. The compatibility score above zero in the VERIFY-3D program showed except ten, all residues are reasonable which makes us to believe that the structure of the Endothelin-B receptor protein is reliable.

Active site Identification of Endothelin B receptor

The possible binding sites of Endothelin B receptor was searched based on the structural comparison of template and the model build with CASTP server. Since, Endothelin B and the 3C9M are well conserved in both sequence and structure; their biological function should be identical. The predicted results are interesting and its shows the amino acid position 17-22, 78-98 and 197-208 are predicted to conserved with the active site. Here we consider the experimental binding sites of 3C9M include some of the residues as mentioned above. Thus in this study ARG 82, ARG 84 and HIS 197 are chosen as the more favorable sites to dock the substrate.

Docking of inhibitors within the active site of Endothelin B receptor

In the presence of inhibitory results of Bosentan, new structural features and fictionalization requirements were proposed for the basic Bosentan scaffold that could increase affinity with Endothelin B receptor and consequently improves the anti hyper tense property [14]. Among these requirements, the modifications of the group in the structure by a more polar group were expected to increase the activity. These interesting results prompted us to prepare Bosentan analogues. Here we developed different analogues based on the structure of Bosentan by replacing with more polar groups than already existing groups, with little change in the properties of analogues represented in **Table 1 (see supplementary material).** These were used for docking studies to identify better drug derivative. Docking of inhibitors given in **Table 1 (see supplementary material)** with Endothelin B was performed using GOLD 3.0.1, which is based on genetic algorithm (CCDC). The structure of Bosentan used for designing derivatives. The docking procedure includes several steps. First, the protein-ligand complex is generated using the GOLD package without constraints between the ligand and the specific amino acids of the pocket. The algorithm exhaustively searches the entire rotational and translational space of the ligand with respect to the receptors. The flexibility of the ligand is given by dihedral angle variations. The various solutions evaluated by a score, which is equivalent to the absolute value of the total energy of the ligand in the protein environment. Thus docking with the program GOLD version 3.0.1 was employed to locate the appropriate binding orientation and conformation of compounds with Endothelin B receptor.

Gold Score fitness function:
Gold Score performs a force field based scoring function and is made up of four components namely, hydrogen bond energy (H-bond); vander Waals energy (External vdW); vander Waals energy (internal vdW) and Intermolecular hydrogen bond (internal H-bond) energy. The external vdw score is multiplied by a factor of 1.375 when the total fitness score is computed. This is an empirical correction to encourage protein-ligand hydrophobic contact. The fitness function has been optimized for the prediction of ligand binding positions [15].

GoldScore Fitness = S (hb_ext) + 1.3750*S (vdw_ext) + S (hb_int) + 1.0000*S (int)
S (int) = S (vdw_int) + S (tors)

Where S (hb_ext) is the protein-ligand hydrogen bond score, S (vdw_ext) is the protein-ligand van der Waals score, S (hb_int) is the score from intramolecular hydrogen bond in the ligand and S (vdw_int) is the score from intramolecular strain in the ligand.
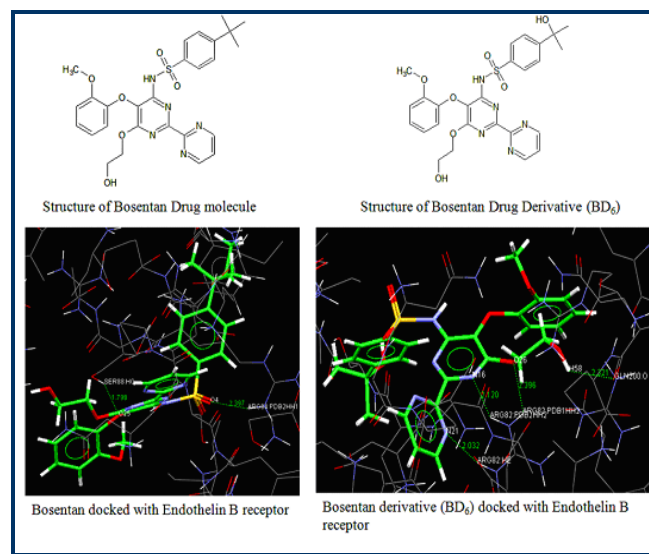

Structure of Bosentan Drug molecule

Structure of Bosentan Drug Derivative (BD$_6$)

Bosentan docked with Endothelin B receptor

Bosentan derivative (BD$_6$) docked with Endothelin B receptor

Figure 2: Comparisons of interactions between Bosentan and its derivatives against Endothelin B receptor

Among the 29 inhibitors BD$_6$ derivate of Bosentan showed high affinity than others towards Endothelin B. The structure of the BD$_6$ was shown in the **Figure 2.** Bosentan docked results shows interaction of Endothelin B receptor with 2 hydrogen bonds. The active residues in the Endothelin B receptor ARG84 binds to O4 of Bosentan with one hydrogen bond. The other hydrogen bond is formed between SER88 and O25 of Bosentan. BD$_6$ derivative forms four hydrogen bonds with Endothelin B receptor with the active residues. It is observed that BD$_6$ derivative was located at center of active site and is stabilized by hydrogen bonding interactions and these interactions are represented in **Table 2.** Bosentan total Gold Score fitness function is 30.03 Kcal/mol. This shows that Bosentan has good inhibitory activity on Endothelin B receptor. Gold Score fitness function of BD$_6$: Total Gold Score fitness function is 45.21 Kcal/mol. Here we have chosen Gold Fitness function than Chem score fitness function as Gold score fitness function is marginally better than ChemScore fitness function. Even though, some other compounds in this library are having better

scoring the BD_6 shows better interactions and those interactions in Bosentan binding region **Table 3 (see supplementary material)**. So, we have demonstrated that, among the 29 derivatives designed the binding capacity is highest for BD_6 and can be further used for further investigations.

Conclusion:
In this work, we have constructed a 3D model of Endothelin B receptor of *Homosapiens,* and obtained a refined model after energy minimization. The validation of protein structure through SAVS server and NAMD validates the protein is stable enough. The stable structure is further used for docking with Bosentan and its derivatives. Docking results indicate that conserved amino-acid residues in Endothelin B receptor play an important role in maintaining a functional conformation and are directly involved in donor substrate binding. The interaction between the domain and the inhibitors proposed in this study are useful for understanding the potential mechanism of domain and the inhibitor binding. The hydrogen bonds play important role for the structure and function of biological molecule in this study, and we found that ARG 82, ARG 84 and HIS 197 of Endothelin B are important for strong hydrogen bonding interaction with these inhibitors. To the best of our knowledge ARG 82, ARG 84 is conserved in the domain and may be important for structural integrity or maintaining the hydrophobicity of the inhibitor-binding pocket. Among 29 derivatives used for docking, BD_6 derivative showed best binding energy than other compounds. The total fitness function is high for BD_6 when compared to other derivatives including Bosentan. According to our investigations from the docking results BD_6 ($C_{26}H_{27}N_5O_7S$) can act as better Endothelin antagonist. Future investigation of other theoretical studies and experimental studies may confirm that BD_6 as a potent inhibitor.

Reference
[1] Honore JC *et al*. *Clin Sci*. 2002 **103**: 280S [PMID: 12193104]
[2] Mutlu B *et al*. *Anadolu Kardiyol Derg*. 2010 **1:** 43 [PMID: 20819766]
[3] Raza T *et al*. *Heart Reviews*. 2007 **8**: 90
[4] Benigni A. *Curr Opin Nephrol Hypertens*. 1995 **4**: 349 [PMID: 7552102]
[5] Maegele M *et al*. *Neurol Res.* 2011 **33**: 119 [PMID: 21801586]
[6] Strange G *et al*. *Expert Rev Pharmacoecon Outcomes Res*. 2011 **11**: 253 [PMID: 21671693]
[7] Boniface S *et al*. *Rev Mal Respir*. 2011 **28**: e94 [PMID: 22099418]
[8] Stenkamp RE. *Acta Crystallogr D Biol Crystallogr*. 2008 **64**: 902 [PMID: 18645239]
[9] Sundaram S *et al*, *Bioinformation*. 2010 **5**: 177 [PMID: 21364783]
[10] Wang Y *et al*. *Comput Sci Discov*. 2011 4: 015002 [PMID: 21686063]
[11] Kaistha SD & Sinha R, *Bioinformation*. 2009 **3**: 240 [PMID: 19255642]
[12] Verdonk ML *et al*. *Proteins*. 2003 **52**: 609 [PMID: 12910460]
[13] Annamala MK *et al*. *Bioinformation*. 2007 **1**: 339. [PMID: 17597917]
[14] Goirand F *et al*. *J Cardiovasc Pharmacol*. 2003 **41**: 117 [PMID: 12500029]
[15] John S *et al*. *BMC Bioinformatics*. 2011 12 [PMID: 21342558]

Edited by P Kangueane

BIOINFORMATION

Supplementary material:

Table 1: Library of Bosentan

S.NO	DRUG	Molecular formula	Molecular weight	Molar refractivity cm^3	Index of refraction	Density g /cm^3	Polaris ability 10^{-24} cm^3
1	BOSTN	$C_{27}H_{29}N_5O_6S$	551.61	143.69 ± 0.4	1.607± 0.02	1.325± 0.06	56.96±0.5
2	BD$_1$	$C_{28}H_{31}N_5O_6S$	565.64	148.33 ± 0.4	1.601±0.02	1.307±0.06	58.80±0.5
3	BD$_2$	$C_{28}H_{31}N_5O_6S$	565.64	148.33± 0.4	1.601±0.02	1.307±0.06	58.80±0.5
4	BD$_3$	$C_{28}H_{31}N_5O_5S$	549.64	146.79 ± 0.4	1.589±0.02	1.263±0.06	58.19±0.5
5	BD$_4$	$C_{28}H_{31}N_5O_6S$	565.64	148.33 ± 0.4	1.601±0.02	1.301±0.06	58.80±0.5
6	BD$_5$	$C_{26}H_{27}N_5O_7S$	553.58	140.60 ± 0.4	1.626±0.02	1.394±0.06	55.74±0.5
7	BD$_6$	$C_{26}H_{27}N_5O_7S$	553.58	140.60± 0.4	1.626±0.02	1.394±0.06	55.74±0.5
8	BD$_7$	$C_{26}H_{27}N_5O_7S$	553.58	140.60 ± 0.4	1.626±0.02	1.394±0.06	55.74±0.5
9	BD$_8$	$C_{26}H_{27}N_5O_7S$	553.58	140.60 ± 0.4	1.626±0.02	1.394±0.06	55.74±0.5
10	BD$_9$	$C_{29}H_{33}N_5O_5S$	563.55	151.42 ± 0.4	1.585±0.02	1.248±0.06	60.02±0.5
11	BD$_{10}$	$C_{26}H_{27}N_5O_7S$	553.58	140.60 ± 0.4	1.626±0.02	1.394±0.06	55.74±0.5
12	BD$_{11}$	$C_{26}H_{27}N_5O_6S$	537.58	139.03 ± 0.4	1.613 ±0.02	1.345±0.06	55.11±0.5
13	BD$_{12}$	$C_{26}H_{27}N_5O_6S$	537.58	139.03±0.4	1.613 ±0.02	1.345 ±0.06	55.11±0.5
14	BD$_{13}$	$C_{27}H_{29}N_5O_5S$	535.61	142.16± 0.4	1.594 ±0.02	1.279 ±0.06	56.35 ±0.5
15	BD$_{14}$	$C_{28}H_{30}N_5O_5S$	548.63	NA	NA	NA	NA
16	BD$_{15}$	$C_{27}H_{28}N_5O_6S$	550.60	NA	NA	NA	NA
17	BD$_{16}$	$C_{30}H_{34}N_5O_6S$	592.68	NA	NA	NA	NA
18	BD$_{17}$	$C_{30}H_{34}N_5O_6S$	592.68	NA	NA	NA	NA
19	BD$_{18}$	$C_{31}H_{36}N_5O_5S$	590.71	NA	NA	NA	NA
20	BD$_{19}$	$C_{30}H_{34}N_5O_6S$	592.68	NA	NA	NA	NA
21	BD$_{20}$	$C_{28}H_{30}N_5O_7S$	580.63	NA	NA	NA	NA
22	BD$_{21}$	$C_{28}H_{30}N_5O_7S$	580.63	NA	NA	NA	NA
23	BD$_{22}$	$C_{28}H_{31}N_5O_7S$	581.64	149.86 ± 0.4	1.613±0.02	1.352 ±0.06	59.41 ±0.5
24	BD$_{23}$	$C_{28}H_{31}N_5O_6S$	565.64	148.33 ± 0.4	1.601±0.02	1.307 ±0.06	58.79 ±0.5
25	BD$_{24}$	$C_{28}H_{31}N_5O_6S$	565.64	148.30 ± 0.4	1.601±0.02	1.308 ±0.06	58.79 ±0.5
26	BD$_{25}$	$C_{28}H_{31}N_5O_6S$	565.64	148.30 ± 0.4	1.601±0.02	1.308 ±0.06	58.79 ±0.5
27	BD$_{26}$	$C_{29}H_{33}N_5O_5S$	563.66	151.42 ± 0.4	1.585 ±0.02	1.585 ±0.02	60.02 ±0.5
28	BD$_{27}$	$C_{28}H_{31}N_5O_6S$	565.64	148.12 ± 0.4	1.616 ±0.02	1.335 ±0.06	58.72 ±0.5
29	BD$_{28}$	$C_{28}H_{31}N_5O_7S$	581.64	149.86 ± 0.4	1.613 ±0.02	1.352 ±0.06	59.41 ±0.5
30	BD$_{29}$	$C_{28}H_{31}N_5O_7S$	581.64	149.86 ± 0.4	1.613 ±0.02	1.352 ±0.06	59.41 ±0.5

BIOINFORMATION

Table 2: Scoring difference between Bosentan and its derivatives against Endothelin B receptor

Ligands	Gold score	Interacting sites	Bosentan atom	Bond length(Å)	Bond angles (Deg)
Bosentan	30.03 Kcal/mol	ARG84 (HH1)	O4	2.397	105.05
		SER 88 (HG)	O25	1.798	89.62
Bosentan D_6	45.21 Kcal/mol	ARG82 (HHE)	N16	2.120	91.48
		ARG82 (HE)	N21	2.032	80.81
		ARG84 (HH2)	O26	2.396	82.38
		GLN200 (O)	H58	2.321	63.86

Table 3: Fitness scoring parameters from GOLD docking with known and unknown compounds

S.NO	DRUG	S(hb_ext)	S(vdW_ext)	S(hb_int)	S(int)	Total Fitness
1	BOSTN	2.51	37.90	0.00	-24.59	30.03
2	BD1	2.33	36.69	0.00	-26.86	25.92
3	BD2	2.87	33.58	0.00	-14.14	34.91
4	BD3	2.49	37.27	0.00	-11.77	41.97
5	BD4	2.19	40.20	0.00	-15.38	42.09
6	BD5	2.92	40.55	0.00	-134.81	-76.13
7	BD6	4.35	44.70	0.00	-20.60	45.21
8	BD7	4.63	35.47	0.00	-15.87	37.54
9	BD8	4.50	39.47	0.00	-15.04	43.73
10	BD9	2.00	39.83	0.00	-32.29	24.47
11	BD10	0.78	38.22	0.00	-24.47	28.87
12	BD11	1.38	38.59	0.00	-13.16	41.29
13	BD12	0.55	42.66	0.00	-14.99	44.21
14	BD13	0.48	40.74	0.00	-13.06	43.43
15	BD14	2.04	37.98	0.00	-18.60	35.66
16	BD15	1.43	42.38	0.00	-21.09	38.61
17	BD16	0.54	36.63	0.00	-24.93	25.98
18	BD17	0.45	36.63	0.00	-24.42	26.40
19	BD18	2.62	36.66	0.00	-32.76	20.26
20	BD19	2.72	35.50	0.00	-28.27	23.25
21	BD20	0.58	36.32	0.00	-31.81	18.71
22	BD21	1.63	37.91	0.00	-26.21	27.55
23	BD22	3.84	36.63	0.00	-30.18	24.02
24	BD23	1.59	37.06	0.00	-21.68	30.86
25	BD24	1.28	36.95	0.00	-34.15	17.94
26	BD25	1.34	48.86	0.00	-37.25	20.27
27	BD26	2.13	37.67	0.00	-31.81	22.11
28	BD27	0.46	37.99	0.00	-31.68	21.02
29	BD28	1.10	35.58	0.00	-16.93	33.10
30	BD29	0.98	38.74	0.00	-27.93	26.32

Where: S (hb_ext) = Protein-ligand hydrogen bond score; S (vdw_ext) = Protein-ligand van der Waals score; S (hb_int) = Score from intramolecular hydrogen bond in the ligand; S (int) = Score from intramolecular strain in the ligand.