Supporting Information

Structure-based design of DevR inhibitor active against non-replicating *Mycobacterium tuberculosis*

*Rajesh Kumar Gupta[1], Tejender S. Thakur[2], Gautam R. Desiraju[2¶] and Jaya Sivaswami Tyagi[1]**

[1]Department of Biotechnology, All India Institute of Medical Sciences, New Delhi-110026 and [2]School of Chemistry, University of Hyderabad, Hyderabad 500 046, INDIA

Corresponding authors

*Corresponding author: Tel: +91 11 26588491; Fax: +91 11 26588663; Email: jstyagi@gmail.com

[¶]Co-corresponding author: Present address: Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore 560 012, India. Tel: +91 80 22933311; Fax: 91 80 23602306; Email: gautam_desiraju@yahoo.com

Table S1. List of 100 top scored compounds obtained from MOE and GOLD docking results.

Table S2. Segregation of top 200 top scored docked compounds into chemical classes (1-11) of similar structural scaffolds.

Table S3. Hydrogen bond analysis of protein ligand interactions for the top scored docked conformation for the biologically active compounds (**5**, **7**, **8** and **10**).

Table S4. Comparison of secondary structure of the homology model with the X-ray crystal structure crystal structure of DevR (PDB ID: 3C3W).

Table S5. Domain-wise RMSD of Cα atoms of the homology model with the X-ray crystal structure of DevR (PDB ID: 3C3W).

Table S6. Comparison of pocket used for docking of compounds in homology model with pocket present in crystal structure.

Fig. S1. Effect of compound **10** on (a) DevR-regulated promoter activity and (b) viability of *M. tuberculosis* in nutrient starvation model of dormancy.

Fig. S2. Hydrogen bond interaction array predicted between compounds and DevR.

Fig. S3. Ligand-based pharmacophore generated from the biologically active ligands based upon maximum feature alignment criteria.

Fig. S4. Interaction of compound **10** and DevR in protein model (top) and structure (bottom).

Fig. S5. 3D view of inhibitor (compound **10**) docked into the putative binding site of the DevR protein.

Supporting Text:

Domains-wise distribution of amino acids in DevR protein
Sequence Alignment
Homology model generation using MOE- homology module
Identification of the active site
Molecular dynamics
Pharmacophore generation
Docking.

Table S1. List of 200 top scored compounds obtained from MOE (top 100 hits) and GOLD (top 100 hits) docking results.

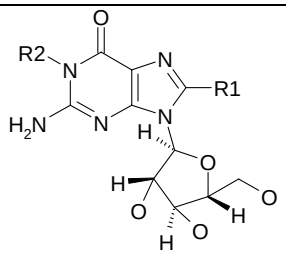
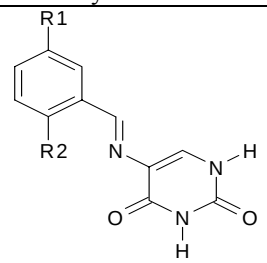
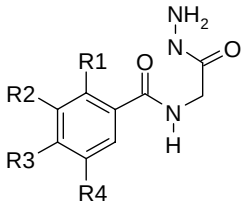
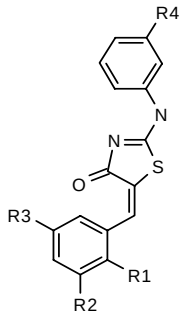
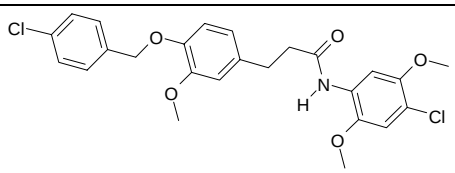
MOE		GOLD	
ZINC ID	**Scoring function (S)**	**ZINC ID**	**Scoring function (Chemscore)**
ZINC02861338	**-6.14 "Compound 10"**	ZINC01143229	26.61
ZINC01899311	-5.87	ZINC00343818	25.25
ZINC00293872	-5.54	ZINC00506272	25.98
ZINC00538490	-5.42	ZINC00289621	27.57
ZINC00360911	-4.87	ZINC01465446	26.52
ZINC01902121	-4.79	ZINC01630640	26.41
ZINC01792763	-4.79	ZINC01630639	25.45
ZINC02842577	-4.77	ZINC00842994	27.49
ZINC00506273	-4.73	ZINC00842995	28.83
ZINC00754747	**-4.64 "Compound 11"**	ZINC01691118	25.19
ZINC00899555	-4.60	**ZINC01142933**	**28.20 "Compound 6"**
ZINC01249774	-4.59	ZINC01818682	26.31
ZINC01185324	-4.57	ZINC00626860	25.67
ZINC00625928	-4.56	ZINC00904577	26.91
ZINC01835467	-4.52	ZINC01419352	27.03
ZINC00346436	-4.51	ZINC02168222	25.03
ZINC01401696	-4.51	ZINC01636696	26.76
ZINC01845183	-4.39	ZINC01041809	26.46
ZINC01723972	-4.33	ZINC02853833	25.92
ZINC00542541	-4.33	ZINC00891483	29.49
ZINC00824775	-4.30	ZINC00878981	26.61
ZINC01147488	-4.30	ZINC00924154	29.01
ZINC02070202	-4.27	ZINC00878969	27.82
ZINC01723965	**-4.25 "Compound 1"**	ZINC01631971	27.10

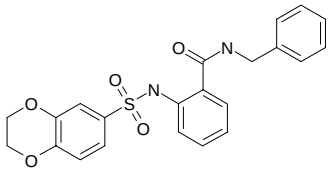
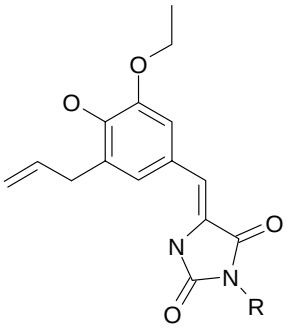
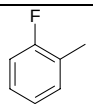
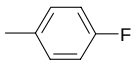
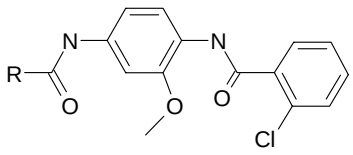
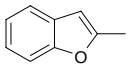
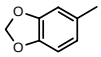
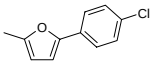
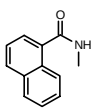
ZINC01407745	-4.24	ZINC01660285	27.49
ZINC00865559	-4.21	ZINC01606657	25.01
ZINC00360910	-4.17	ZINC01559281	28.77
ZINC00316445	-4.12	ZINC02776457	25.21
ZINC00906858	-4.12	ZINC01340002	25.16
ZINC01603988	-4.12	ZINC00871048	26.94
ZINC01863772	-4.11	ZINC00871049	25.48
ZINC01407251	-4.10	ZINC01618680	30.29
ZINC00289621	-4.11	ZINC00671644	25.50
ZINC01131848	-4.10	ZINC01700730	25.02
ZINC00824777	-4.09	ZINC01611795	30.03
ZINC01844672	-4.05	ZINC01055132	27.94
ZINC01183999	**-4.04 "Compound 4"**	ZINC01472675	26.17
ZINC00747852	-4.01	ZINC00726366	28.74
ZINC00671644	-4.01	ZINC01472680	25.77
ZINC01500864	-4.00	**ZINC00438332**	**26.06 "Compound 2"**
ZINC01649591	-3.99	ZINC01571909	25.07
ZINC00564244	-3.96	ZINC00709396	25.96
ZINC00005704	-3.96	ZINC00709395	26.43
ZINC01655170	-3.90	ZINC01335587	28.42
ZINC02827369	-3.90	ZINC00927401	26.36
ZINC02718848	-3.89	ZINC01092459	25.99
ZINC02020930	-3.89	ZINC01461512	25.46
ZINC01591462	-3.87	ZINC00564244	26.43
ZINC01122423	-3.86	ZINC01084332	25.67
ZINC01652970	-3.86	ZINC01109183	25.22
ZINC01582058	-3.85	ZINC02756570	25.40
ZINC01286314	-3.85	ZINC00793821	25.23
ZINC01433736	-3.85	ZINC01723937	27.35

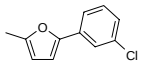
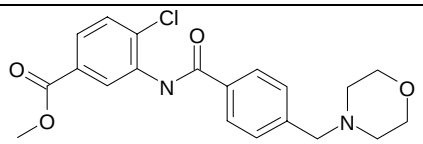
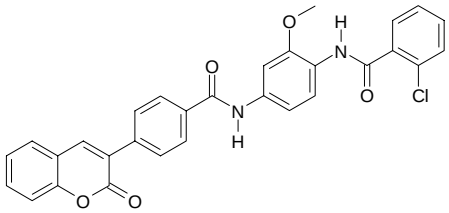
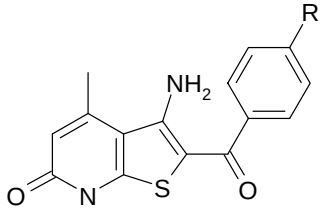
ZINC01691787	-3.85	ZINC01694055	28.98
ZINC00625676	-3.85	ZINC01434177	25.58
ZINC00868304	-3.84	ZINC01128276	27.47
ZINC01401045	-3.84	ZINC01668707	26.19
ZINC01768587	**-3.83 "Compound 5"**	ZINC00832247	27.92
ZINC02609664	-3.81	ZINC00832246	29.89
ZINC01438091	-3.81	ZINC00832245	28.11
ZINC00726366	-3.80	ZINC00832244	25.07
ZINC00349493	-3.77	ZINC01468958	27.55
ZINC00495623	-3.75	ZINC00445428	25.12
ZINC00826908	-3.74	ZINC01884490	25.04
ZINC00982253	**-3.74 "Compound 7"**	**ZINC00255951**	**26.09 "Compound 3"**
ZINC01886425	-3.72	ZINC01658312	26.16
ZINC01460765	-3.69	ZINC01140861	25.75
ZINC02830752	-3.69	**ZINC00982253**	**26.62 "Compound 7"**
ZINC01233313	-3.68	ZINC00663488	26.32
ZINC02668812	-3.67	ZINC00293872	25.15
ZINC01759984	-3.65	ZINC02178457	26.09
ZINC00438332	**-3.62 "Compound 2"**	ZINC01458622	26.29
ZINC00708958	-3.62	ZINC01217738	25.69
ZINC01577568	-3.61	ZINC01445799	29.83
ZINC02607350	-3.60	ZINC01148951	25.86
ZINC02607708	-3.59	ZINC00813644	26.77
ZINC02615405	-3.59	ZINC00813647	26.09
ZINC01703684	-3.58	**ZINC01108279**	**29.01 "Compound 8"**
ZINC02825177	-3.56	ZINC00809586	27.36
ZINC01308891	-3.56	ZINC00807667	27.44
ZINC02718932	-3.55	ZINC02839838	26.82
ZINC02614884	-3.55	ZINC01141113	25.75

ZINC01460837	-3.52	ZINC01324393	27.39
ZINC01764559	-3.51	ZINC01141029	25.70
ZINC00626157	-3.51	ZINC02829782	28.62
ZINC01281228	-3.49	ZINC01131066	28.43
ZINC00950315	-3.49	ZINC02696029	27.05
ZINC01724573	-3.48	**ZINC02859156**	**27.22 "Compound 9"**
ZINC01604383	-3.47	ZINC02838996	26.54
ZINC01710602	-3.46	ZINC01903294	25.24
ZINC02607140	-3.44	**ZINC01768587**	**29.00 "Compound 5"**
ZINC01774035	-3.42	ZINC00825225	25.74
ZINC02608686	-3.42	ZINC01159049	26.82
ZINC02822841	-3.42	ZINC00349493	27.59
ZINC00936925	-3.42	ZINC01033431	28.11
ZINC00917403	-3.41	ZINC02451344	26.02
ZINC01375405	-3.39	ZINC01110122	25.97
ZINC03766872	-3.39	ZINC00631056	27.43
ZINC00708956	-3.37	ZINC01859436	28.40
ZINC02607136	-3.37	ZINC01324557	26.00
ZINC02610387	-3.37	ZINC01603988	25.89
ZINC00626142	-3.37	ZINC01077343	25.09
ZINC00420110	-3.36	ZINC02455297	26.08
ZINC00793821	-3.34	ZINC00683574	25.97
ZINC02607714	-3.34	ZINC01500864	29.16
ZINC00502611	-3.34	**ZINC01183999**	**27.10 "Compound 4"**

Table S2. Segregation of top scored docked compounds into chemical classes (1-11) of similar structural scaffolds**.**

Class	Molecular scaffold	Compound	Substituent				MOE score	GOLD score
			R1	R2	R3	R4		
1	 9-furanyl -Purin-6-one	**1**	Br	H	-	-	-4.25	NA
		1.1	I	H	-	-	-4.33	NA
		1.2	OH	NH_2	-	-	-3.58	NA
2	 5-N-substituted pyrimidine-di-one	**2**	NO_2	OH	-	-	-3.62	26.06
		2.1	Cl	OH	-	-	-3.77	NA
3	 N-substituted tri methoxy benzamide	**3**	H	OCH_3	OCH_3	OCH_3	NA	26.08
		3.1	H	H	H	F	-5.54	25.15
		3.2	Cl	H	H	Cl	NA	26.18
		3.3	CH_3	H	H	Br	NA	26.08
4	 2-amino, 4-oxo, 1, 3 thiazole derivative	**4**	OH	OCH_3	Br	COOH	-4.04	27.10
		4.1	OH	H	NO_2	OCH_3	-3.42	NA
5	 3-{4-[(4-chlorobenzyl)oxy]-3-methoxyphenyl}-N-(4-chloro-2,5-dimethoxyphenyl) acrylamide	**5**	-	-	-	-	-3.83	29.00

6	 N-benzyl-2-[(2,3-dihydro-1,4-benzodioxin-6-yl sulfonyl) amino] benzamide	**6**	-	-	-	-	NA	28.20
7	 5-(3-allyl-5-ethoxy-4-hydroxybenzylidene)-3-Substituted-2,4-imidazolidinedione	**7**		-	-	-	-3.74	26.70
		7.1	H	-	-	-	-4.12	NA
		7.2		-	-	-	-3.69	26.43
		7.3		-	-	-	NA	26.09
8	 N-{4-[(2-chlorobenzoyl)amino]-3-methoxyphenyl}-1-substituted-2-carboxamide	**8**		-	-	-	-	29.01
		8.1		-	-	-	-	27.39
		8.2		-	-	-	-	27.35
		8.3		-	-	-	-	26.82
		8.4		-	-	-	-	26.77
		8.5		-	-	-	-	26.54
		8.6		-	-	-	-	26.09
		8.7		-	-	-	-	25.86
		8.8		-	-	-	-	25.75

		8.9		-	-	-	-	25.75
		8.10		-	-	-	-	25.70
9	 Methyl 4-chloro-3-{[4-(4 morpholinylmethyl) benzoyl] amino} benzoate	**9**	-	-	-	-	-	27.22
10	 2-chloro-N-(2-methoxy-4-{[4-(2-oxo-2H-chromen-3-yl) benzoyl]amino}phenyl) benzamide	**10**	-	-	-	-	-6.14	NA
11	 3-amino-2-(4-substituted benzoyl)-4-methylthieno pyridin-6-one	**11**	Br	-	-	-	-4.64	NA
		11.1	F	-	-	-	-4.11	27.57

NA: Not available in the top 100 hits

Table S3. Hydrogen bond analysis of protein ligand interactions in the top scored docked conformation for the biologically active compounds.

Ligand	*X−H···A*	*Donor (X)*	*Acceptor (A)*	$d_{H\cdots A}$ *(Å)*	$D_{X\cdots A}$ *(Å)*	$\theta_{X-H\cdots A}$ *(°)*
5	N−H···Cl	ARG77 (NE)	LIG210 (Cl)	1.409	2.271	138.0
	C−H···N	LIG210 (C)	ARG77 (NH2)	2.242	3.253	153.8
	C−H···O	LIG210 (C)	SER123 (O)	2.273	3.103	131.3
	C−H···O	LEU124 (CD2)	LIG210 (O)	2.969	3.888	141.9
	C−H···O	ALA130 (CB)	LIG210 (O)	1.776	2.669	135.4
	C−H···O	MET134 (CG)	LIG210 (O)	2.457	3.359	138.8
	C−H···S	LIG210 (C)	MET134 (SD)	1.778	2.647	132.7
	N−H···Cl	ARG197 (NH1)	LIG210 (Cl)	1.321	2.278	153.5
7	C−H···O	ARG77 (CG)	LIG210 (O)	2.566	3.656	175.3
	C−H···O	LIG210 (C)	ALA98 (O)	1.997	2.841	131.2
	O−H···O	SER99 (OG)	LIG210 (O)	2.030	2.903	148.3
	C−H···O	LEU124 (CD1)	LIG210 (O)	2.586	3.677	177.1
	C−H···O	LEU125 (CB)	LIG210 (O)	2.953	3.808	135.1
	C−H···O	LIG210 (C)	LEU125 (O)	2.581	3.411	132.4
	C−H···O	LIG210 (C)	ASN127 (OD1)	2.870	3.887	155.6
	C−H···N	LIG210 (C)	ASN127 (ND2)	1.415	2.435	152.8
	C−H···O	LIG210 (C)	GLY164 (O)	1.573	2.533	143.9
	C−H···O	THR198 (CG2)	LIG210 (O)	1.493	2.421	138.3
	C−H···O	THR198 (CG2)	LIG210 (O)	2.154	2.992	131.5
	C−H···O	VAL202 (CG2)	LIG210 (O)	2.619	3.689	166.1
	C−H···N	VAL202 (CG2)	LIG210 (N)	2.413	3.270	134.0
8	O−H···Cl	SER99 (OG)	LIG210 (Cl)	2.687	3.649	170.0
	C−H···O	LIG210 (C)	SER99 (OG)	2.738	3.740	152.7
	C−H···N	LIG210 (C)	LEU125 (N)	2.244	3.268	155.0
	C−H···O	ALA130 (CB)	LIG210 (O)	2.330	3.276	143.7
	C−H···O	LIG210 (C)	LEU161 (O)	2.688	3.582	139.0
	C−H···O	LIG210 (C)	SER162 (O)	2.735	3.706	148.6
	C−H···O	GLY164 (CA)	LIG210 (O)	1.735	2.652	138.0
	C-H...O	LIG210 (C)	LEU165 (O)	2.480	3.456	148.5
	C−H···O	LIG210 (C)	ARG197 (NH1)	1.633	2.490	131.2
	C−H···Cl	THR198 (CG2)	LIG210 (Cl)	2.890	3.959	165.8
	C−H···O	ALA201 (CB)	LIG210 (O)	1.681	2.639	143.2

Ligand	*X−H···A*	*Donor (X)*	*Acceptor (A)*	$d_{H\cdots A}$ (Å)	$D_{X\cdots A}$ (Å)	$\theta_{X-H\cdots A}$ (°)
10	C−H···O	LIG210 (C)	SER99 (OG)	2.184	3.028	132.6
	C−H···O	ALA130 (CB)	LIG210 (O)	2.594	3.583	149.9
	C−H···O	MET134 (CG)	LIG210 (O)	2.283	3.187	138.7
	C−H···O	GLY164 (CA)	LIG210 (O)	1.656	2.675	152.6
	C−H···O	ALA201 (CB)	LIG210 (O)	2.175	3.252	167.8
	C−H···N	THR205 (CG2)	LIG210 (N)	1.731	2.644	137.6
	C−H···O	LIG210 (C)	THR205 (OG1)	1.697	2.552	131.8

Table S4. Comparison of secondary structure of the homology model with the X-ray crystal structure crystal of DevR (PDB ID: 3C3W)

Secondary Structure	Homology Model	Crystal structure
α-sheet-1	3-7	2-7
α-helix-1	11-22	11-22
α-sheet-2	28-29	27-33
α-helix-2	36-41	36-46
α-sheet-3	49-53	50-57
α-helix-3	62-70	62-72
α-sheet-4	77-82	77-80
α-helix-4	87-95	87-97
α-sheet-5	99-103	-
α-helix-5	108-119	101-124
α-helix-6	127-140	127-143
α-helix-7	152-163	152-162
α-helix-8	167-174	167-174
α-helix-9	178-191	178-191
α-helix-10	197-208	199-209

Table S5. Domain-wise RMSD (Å) of Cα atoms of the homology model structures with the X-ray crystal structure of DevR (PDB ID: 3C3W)

Domain	Homology model (Å)
N-terminal (1-97 aa)	5.50
Linker (98-149 aa)	14.50
C-terminal (150-210 aa)	8.38
Full protein (1-210 aa)	18.64

Table S6. Comparison of Pocket 1 used for docking of compounds in DevR homology model with pocket present in crystal structure

Pocket	Pocket features			Active site residues	Remarks
	Size*	Hydrophobicity score#	Side chain contribution†		
Crystal structure	165	21	122	55, 56, 57, 58, 59, 60, 61, 82, 161, 165, 166, 167, 170, 182, 185, 186, 189, 194, 195, 197, 198, 199, 200	-Large and hydrophobic in size -Pocket is at the interface of N-and C-terminal domains -Suitable for ligand binding
Homology model	188	27	143	77, 94, 97, 98, 99, 118, 123, 124, 125, 127, 130, 133, 134, 137, 161, 162, 164, 165, 197, 198, 201, 202, 205	-Large and hydrophobic in size -Pocket is at the interface of N- and C-terminal domains. -Suitable for ligand binding

*Pocket size is defined as the total number of contact atoms in the pocket.
#Hydrophobicity score is defined as the total number of hydrophobic contact atoms in pocket.
†Side chain contribution is defined as the total number of side chain contact atoms in the pocket

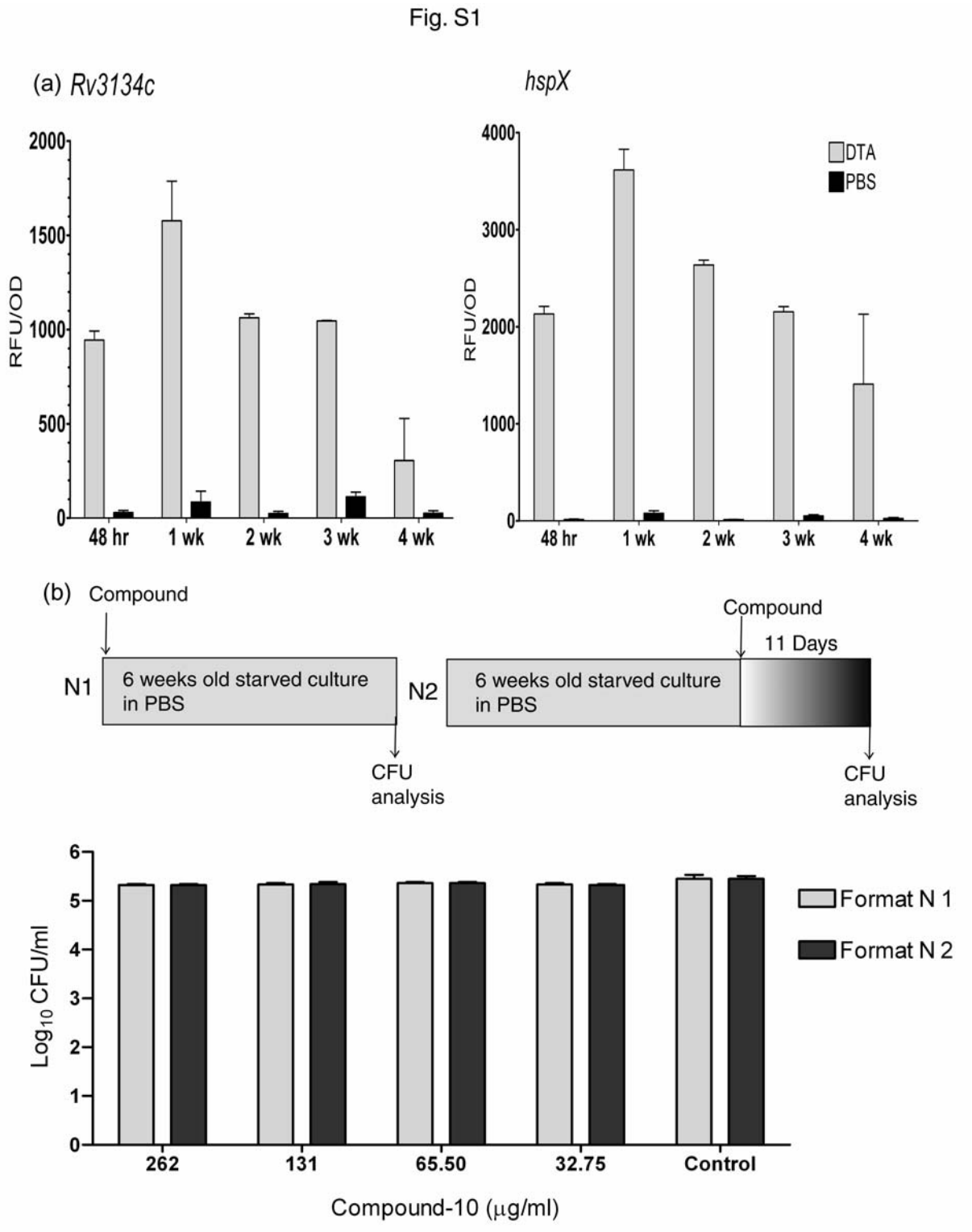


Fig. S1. Effect of compound **10** on (a) DevR-regulated promoter activity and (b) viability of *M. tuberculosis* in nutrient starvation model of dormancy. PBS, phosphate buffered saline; DTA, Dubos Tween Albumin culture medium; Control, no compound.

Fig. S2

LIGAND	ARG77	ALA98	SER99	SER123	LEU124	LEU125	ASP127	ALA130	LEU133	MET134	LEU161	SER162	GLY164	LEU165	ARG197	THR198	ALA201	VAL202	THR205
5	■			■	■										■				
7	■	■	■		■	■							■			■		■	
8						■		■			■	■	■	■	■	■	■		
10			■					■		■			■				■		■

Fig. S2. Hydrogen bond interaction array predicted between compounds and DevR. G164, A201 and T205 residues interact across the α10 DevR dimer interface and T205 is critical for this interaction

Fig. S3. Ligand-based pharmacophore generated from the biologically active ligands based upon maximum feature alignment criteria.

F1-Acceptor (Red); cut off radius = 1.2 Å
F2- Donor (Blue); cut off radius = 1.5 Å
F3- Aromatic (light green); cut off radius = 1.2 Å
F4- Hydrophobic (Dark green); cut off radius = 1.8 Å
F5- Acceptor (Red); cut off radius = 1.3 Å

Distance (Å)	F1	F2	F3	F4	F5
F1	-	4.15	4.09	9.68	7.58
F2	4.15	-	3.90	10.03	6.65
F3	4.09	3.90	-	13.04	10.06
F4	9.68	10.03	13.04	-	4.03
F5	7.58	6.65	10.06	4.03	-

Fig. S4

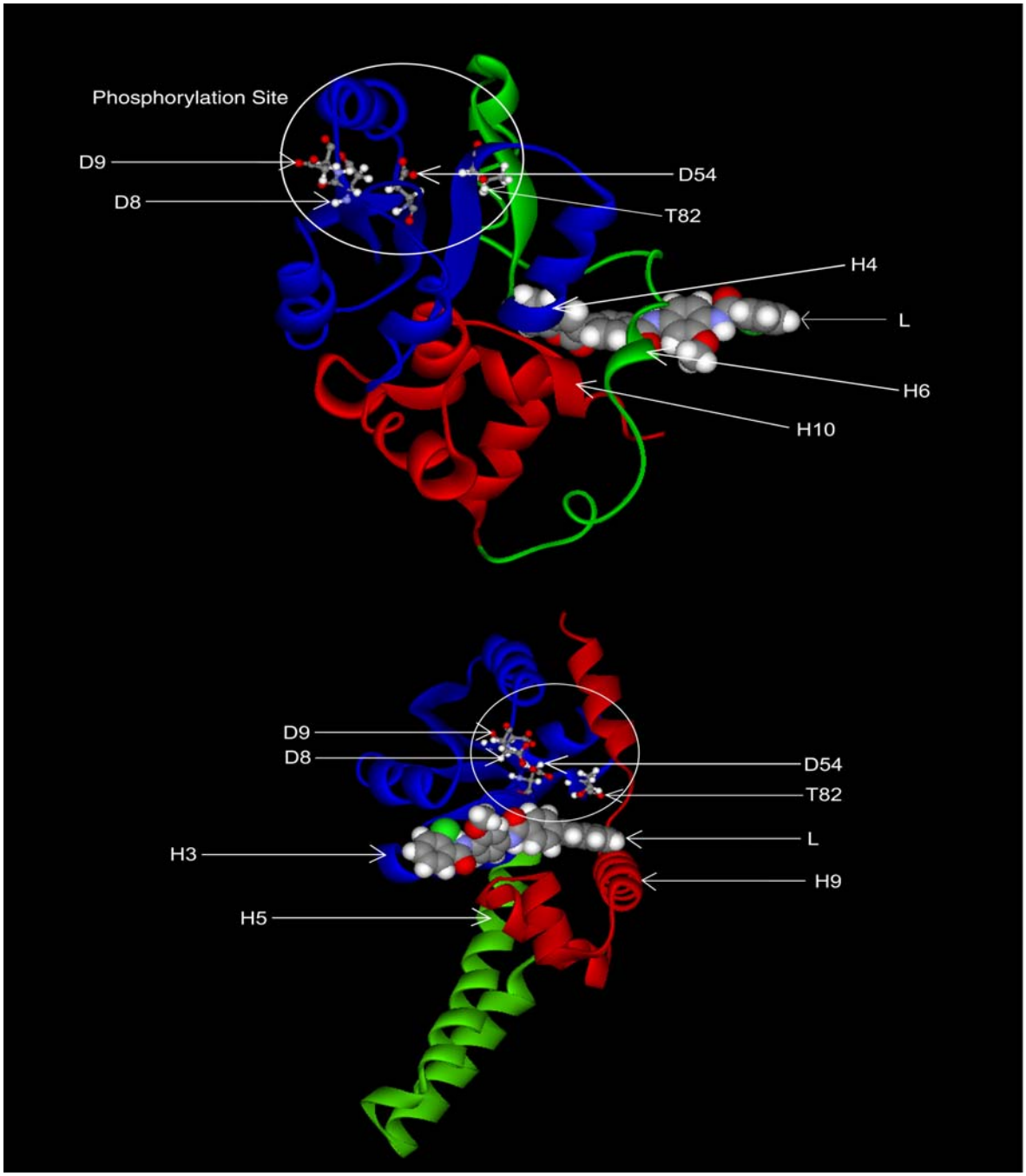


Fig. S4. Interaction of compound **10** (L) and DevR in protein model (top) and structure (bottom). N-terminal domain, linker and C-terminal domain are shown in blue, green and red colors respectively. α-helices are represented as H.

Fig. S5

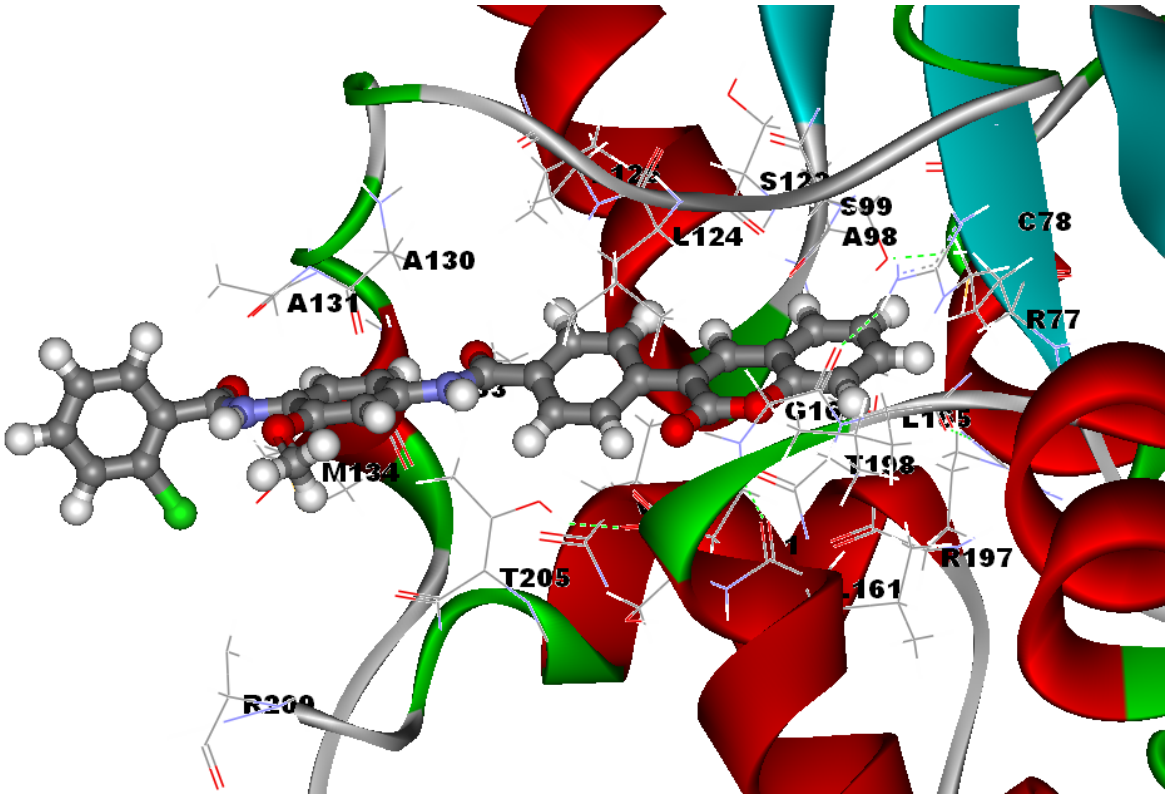


Fig. S5. 3D view of inhibitor (compound **10**) docked into the putative binding site in DevR protein.

Supporting Text

Domains-wise distribution of amino acids in DevR protein.

N-Terminal region (1-95 aa)

VVKVFLVDDHEVVRRGLVDLLGADPELDVVGEAGSVAEAMARVPAARPDVAVLDVRL
PDGNGIELCRDLLSRMPDLRCLILTSYTSDEAMLDAIL

Linker Region (96-144 aa)

AGASGYVVKDIKGMELARAVKDVGAGRSLLDNRAAAALMAKLRGAAEKQ

C-Terminal region (144-217 aa)

DPLSGLTDQERTLLGLLSEGLTNKQIADRMFLAEKTVKNYVSRLLAKLGMERRTQAAVF
ATELKRSRPPGDGP

Sequence Alignment

MOE-Homology search parameters

Tuple size = 1
Scoring matrix = blosum62
Gap start = -12
Gap extend = -2
E-value = 10
E-value accept = 1e-12
Z-iterations = 100
Z-cutoff = 6

Query sequence “TWO COMPONENT TRANSCRIPTIONAL REGULATORY PROTEIN DEVR (1-217 Amino acids)”.

VVKVFLVDDHEVVRRGLVDLLGADPELDVVGEAGSVAEAMARVPAARPDVAVLDVRL
PDGNGIELCRDLLSRMPDLRCLILTSYTSDEAMLDAILAGASGYVVKDIKGMELARAVK
DVGAGRSLLDNRAAAALMAKLRGAAEKQDPLSGLTDQERTLLGLLSEGLTNKQIADRM
FLAEKTVKNYVSRLLAKLGMERRTQAAVFATELKRSRPPGDGP

Sequence alignment details

	PDB ID	*Description*	*E-Value*	*%Identity*	*%Similarity*	*Total AA*	*Overlap AA*
1	1A04	Signal Transduction	9.5e-30	32.3%	70.6%	215	201
2	1ZLJ	Transcription	2.6e-27	95.0%	97.5%	78	74
3	1YIO	DNA Binding	1.2e-26	33.0%	56.9%	208	197
4	1DZ3	Response regulator	3.7e-14	24.2%	64.2%	130	120
5	1A20	Response regulator	4.2e-10	24.0%	56.9%	256	167
6	1F4V	Response regulator	0.00083	28.6%	61.3%	128	119

Template protein used: - E. coli response regulator NarL protein (PDB ID: 1A04)

Sequence alignment

```
                  10        20        30        40        50
DevR       MVKVFLVDDHEVVRRGLVDLLGADPELDVVGEAGSVAEAMARVPAARPDVAVLDVR
             ..:.::: ..: :. .:..  :.. ::::::.  ...  . .  ::. .::..
1A04    SNQEPATILLIDDHPMLRTGVKQLISMAPDITVVGEASNGEQGIELAESLDPDLILLDLN
             10        20        30        40        50        60

         60        70        80        90       100       110
DevR    LPDGNGIELCRDLLSRMPDLRCLILTSYTSDEAMLDAILAGASGYVVKDIKGMELARAVK
        .:  ::.:    : ..  . : ....  . .: .. :.  ::.::.::::.  .: .:.
1A04    MPGMNGLETLDKLREKSLSGRIVVFSVSNHEEDVVTALKRGADGYLLKDMEPEDLLKALH
             70        80        90       100       110       120

        120       130         140       150       160       170
DevR    DVGAGRSLLDNRAAAALMAKLRG--AAEKQDPLSGLTDQERTLLGLLSEGLTNKQIADRM
        ...::. .:..  . .: :.::.  :. ..: .. :: .:: .: :..:: ::.:: :.
1A04    QAAAGEMVLSEALTPVLAASLRANRATTERD-VNQLTPRERDILKLIAQGLPNKMIARRL
             130       140       150       160       170

          180       190       200       210
DevR    FLAEKTVKNYVSRLLAKLGMERRTQAAVFATELKRSRPPGDGP
         ..:.::: .:...: :. .. :..:::.
1A04    DITESTVKVHVKHMLKKMKLKSRVEAAVWVHQERIF
     180       190       200       210
```

(i) Homology model generation using MOE- homology module

Force Field = Amber 99

Method used = Best intermediated method

No of models generated = 10

Minimization = Medium (with RMS = 10)

(ii) Identification of the active site

Method used = MOE a-site finder

Probe radius 1 = 1.4 (Angstrom) corresponds to hydrophilic probe

Probe radius 2 = 1.8 (Angstrom) corresponds to hydrophobic probe

Isolated Donor/Acceptor = 3 (Angstrom)

Connection Distance = 2.5 (Angstrom) maximum radius used for clustering spheres to define a pocket

Minimum site size = 3 lower cutoff in no. of alpha sphere

Cutoff radius = 2(Angstrom) cutoff pocket size

Top four probable pockets identified in the homology model

Pocket 1

Size = 188 (Total no of contact atoms in the pocket)

Hyd = 27 (Total no of hydrophobic contact atoms in the pocket)

Side = 143 (Total no of side chain contact atoms in the pocket)

Residues = ARG77, ILE94, GLY97, ALA98, SER99, VAL118, SER123, LEU124, LEU125, ASN127, ALA130, LEU133, MET134, LEU137, LEU161, SER162, GLY164, LEU165, ARG197, THR198, ALA201, VAL202, THR205

Pocket 2

Size = 154 (Total no of contact atoms in the pocket)

Hyd = 26 (Total no of hydrophobic contact atoms in the pocket)

Side = 116 (Total no of side chain contact atoms in the pocket)

Residues = CYS66, ARG67, LEU70,ILE94, LEU95,ALA96, GLY97, LEU133, LYS136, LEU137, ARG138, GLY139, ALA140, GLU142, LYS143, GLN144, MET194, ARG196, THR198, GLN199, VAL202, PHE203, GLU206

Pocket 3

Size = 92 (Total no of contact atoms in the pocket)

Hyd = 19 (Total no of hydrophobic contact atoms in the pocket)

Side = 79 (Total no of side chain contact atoms in the pocket)

Residues = LEU53, VAL55, GLY62, ILE63, VAL65, CYS66, ILE80, THR82, TYR84, SER86, ALA89, ASP92, ALA93, ALA96, TYA101

Pocket 4

Size = 118 (Total no of contact atoms in the pocket)

Hyd = 14 (Total no of hydrophobic contact atoms in the pocket)

Side = 94 (Total no of side chain contact atoms in the pocket)

Residues = MET1, VAL2, LYS3, GLU26, ASP49, GLY119, ALA120, GLY121, THR166, ASN167, LYS168, GLB169, ASP172, ARG173, GLU178, LYS182

Remarks- Pocket no 3 and 4 were on the surface of the protein hence discarded. Pocket 2 was having very narrow opening unsuitable for the ligand entry thus discarded for the docking studies.

Molecular dynamics

Solvent addition was done by using water soak module

Soak mode = Sphere

Layer width =5

Dynamics parameters

Ensemble used = NVT

Temperature = 300K, Initial temperature = 300 K

Pressure = 101 kPa

MD duration = 200 ps

Time step = 0.002 ps

Sample period = 0.5 ps

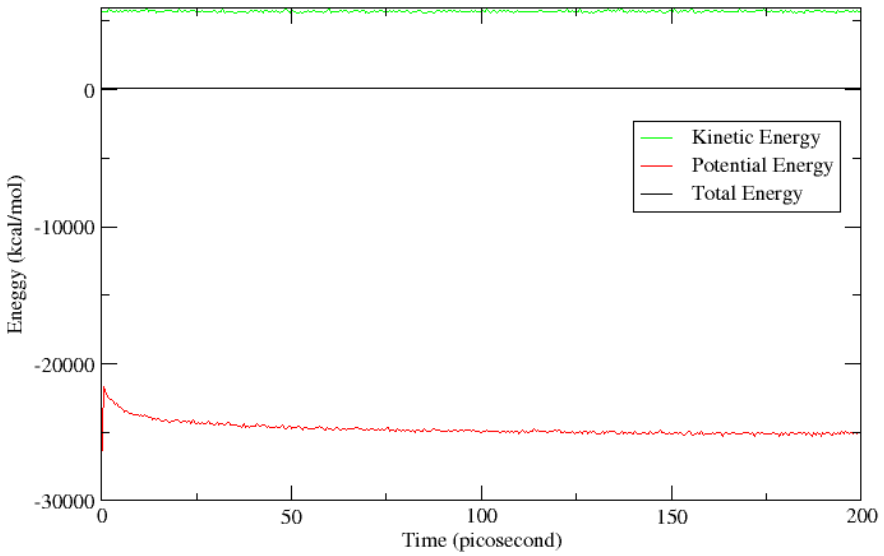
Temperature response = 1 ps; Pressure response = 0.5 ps

Constraints used

Bond length =light atoms

Water = none

Constraint tolerance = 1e-09 ps



Potential energy, kinetic energy and total energy (kcal/mol) vs. time (ps) plot for 200ps MD run.

Pharmacophore generation

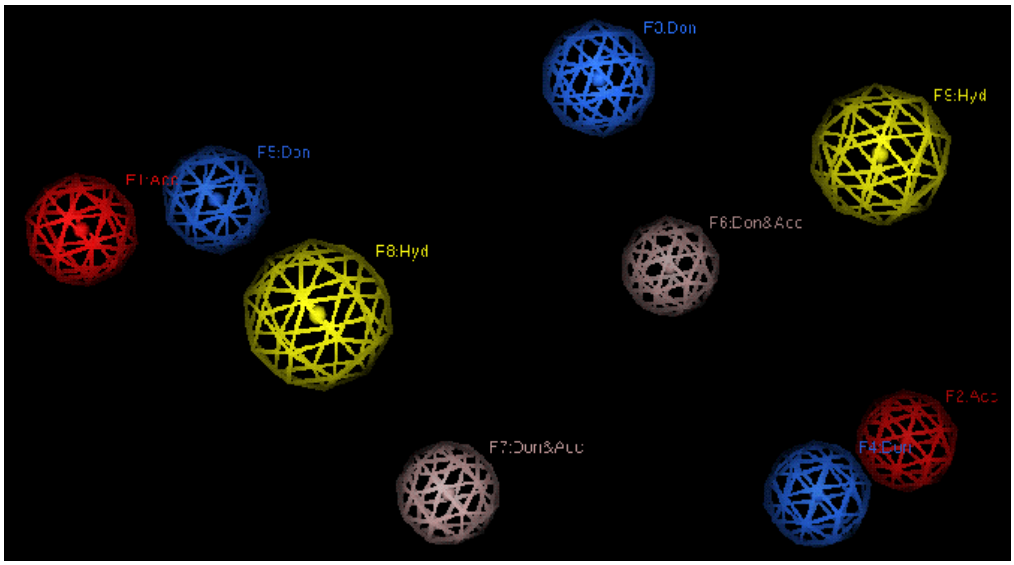
Sampling of structure for water based pharmacophore generation was done at time steps (step size = 0.5 ps):- 259, 272, 301, 341, 345, 378, 394 and 400. Pixel map showing retention of protein water interactions during the MD run for water acting as hydrogen bond acceptor.

Type	Donor	Donor atom	Residue No.	Acceptor	Acceptor atom	Residue No.	259	272	301	341	345	378	394	400
C-H...O	ARG	CG	77	HOH	O	755								
N-H...O	ARG	NE	77	HOH	O	755								
N-H...O	ARG	NH2	77	HOH	O	755								
C-H...O	SER	CA	123	HOH	O	292								
C-H...O	LEU	CB	124	HOH	O	296								
C-H...O	LEU	CB	124	HOH	O	315								
C-H...O	LEU	CB	124	HOH	O	318								
C-H...O	LEU	CB	124	HOH	O	321								
C-H...O	LEU	CD1	124	HOH	O	313								
C-H...O	LEU	CD1	124	HOH	O	740								
C-H...O	LEU	CD2	124	HOH	O	293								
C-H...O	LEU	CD2	124	HOH	O	315								
C-H...O	LEU	CD2	124	HOH	O	740								
C-H...O	LEU	CG	124	HOH	O	292								
C-H...O	LEU	CG	124	HOH	O	296								
N-H...O	LEU	N	124	HOH	O	292								
N-H...O	LEU	N	124	HOH	O	296								
C-H...O	LEU	CA	125	HOH	O	211								
C-H...O	LEU	CB	125	HOH	O	318								
C-H...O	LEU	CD1	125	HOH	O	215								
C-H...O	LEU	CD1	125	HOH	O	295								
C-H...O	LEU	CD2	125	HOH	O	215								
C-H...O	LEU	CD2	125	HOH	O	295								
C-H...O	LEU	CD2	125	HOH	O	318								
N-H...O	LEU	N	125	HOH	O	318								
N-H...O	ASN	ND2	127	HOH	O	162								
N-H...O	ASN	ND2	127	HOH	O	311								
N-H...O	ASN	ND2	127	HOH	O	316								
N-H...O	ASN	ND2	127	HOH	O	321								
C-H...O	ALA	CB	130	HOH	O	162								
C-H...O	ALA	CB	130	HOH	O	308								
C-H...O	ALA	CB	130	HOH	O	309								
C-H...O	ALA	CB	130	HOH	O	310								
C-H...O	ALA	CB	130	HOH	O	316								
C-H...O	LEU	CD1	133	HOH	O	222								
C-H...O	LEU	CD1	133	HOH	O	309								
C-H...O	MET	CE	134	HOH	O	309								
C-H...O	GLY	CA	164	HOH	O	293								
C-H...O	ARG	CD	197	HOH	O	297								
C-H...O	ARG	CD	197	HOH	O	755								
C-H...O	ARG	CG	197	HOH	O	297								
N-H...O	ARG	NH1	197	HOH	O	755								
C-H...O	THR	CB	198	HOH	O	215								
C-H...O	THR	CG2	198	HOH	O	215								
C-H...O	THR	CG2	198	HOH	O	297								
C-H...O	ALA	CB	201	HOH	O	297								
C-H...O	ALA	CB	201	HOH	O	315								
C-H...O	VAL	CA	202	HOH	O	222								
C-H...O	VAL	CG1	202	HOH	O	222								
C-H...O	VAL	CG2	202	HOH	O	215								
C-H...O	VAL	CG2	202	HOH	O	295								
C-H...O	VAL	CG2	202	HOH	O	315								
C-H...O	THR	CB	205	HOH	O	222								
C-H...O	THR	CG2	205	HOH	O	222								
C-H...O	THR	CG2	205	HOH	O	310								
C-H...O	THR	CG2	205	HOH	O	320								
C-H...O	THR	CG2	205	HOH	O	740								
C-H...O	THR	CG2	205	HOH	O	765								
O-H...O	THR	OG1	205	ALA	O	201								

Pixel map showing retention of protein water interactions during the MD run for water acting as hydrogen bond donor.

Type	Donor	Donor atom	Residue No.	Acceptor	Acceptor atom	Residue No.	259	272	301	341	345	378	394	400
O-H...O	HOH	O	215	ALA	O	98								
O-H...O	HOH	O	211	SER	O	99								
O-H...O	HOH	O	292	ARG	O	122								
O-H...O	HOH	O	293	SER	O	123								
O-H...O	HOH	O	335	SER	OG	123								
O-H...N	HOH	O	292	LEU	N	124								
O-H...O	HOH	O	211	LEU	O	124								
O-H...O	HOH	O	335	LEU	O	124								
O-H...N	HOH	O	321	LEU	N	125								
O-H...O	HOH	O	321	LEU	O	125								
O-H...O	HOH	O	321	ASN	OD1	127								
O-H...O	HOH	O	309	ALA	O	130								
O-H...O	HOH	O	740	SER	O	162								
O-H...O	HOH	O	755	GLY	O	164								
O-H...O	HOH	O	297	ARG	O	197								
O-H...O	HOH	O	315	ALA	O	201								
O-H...O	HOH	O	740	ALA	O	201								
O-H...O	HOH	O	223	VAL	O	202								
O-H...O	HOH	O	765	THR	O	205								
O-H...O	HOH	O	740	THR	OG1	205								

Color Scheme represents: - Yellow = water based acceptor pharmacophoric features; Blue = water based donor pharmacophoric features; Green= water based donor/acceptor pharmacophoric features.



F1- Acceptor (Red); cut of radius = 0.8 Å; F2- Acceptor (Red); cut of radius = 0.8 Å
F3- Donor (Blue); cut of radius = 0.8 Å; F4- Donor (Blue); cut of radius = 0.8 Å
F5- Donor (Blue); cut of radius = 0.8 Å; F6- Donor/Acc (Blue); cut of radius = 0.8 Å
F7- Donor/Acc (Blue); cut of radius = 0.8 Å; F8- Hydrophobic (Yellow); cut of radius = 1.0 Å;
F9- Hydrophobic (Yellow); cut of radius = 1.0 Å

Distance (Å)	F1	F2	F3	F4	F5	F6	F7	F8	F9
F1	-	13.5	7.78	11.58	2.54	10.12	7.21	4.79	11.47
F2	13.5	-	9.19	2.97	11.58	4.71	7.37	11.51	6.27
F3	7.78	9.19	-	7.57	6.66	7.1	7.81	5.22	4.31
F4	11.58	2.97	7.57	-	10.03	5.17	5.7	8.99	5.44
F5	2.54	11.58	6.66	10.03	-	7.81	5.80	5.34	9.92
F6	10.12	4.71	7.1	5.17	7.81	-	5.37	9.56	6.21
F7	7.21	7.37	7.81	5.7	5.80	5.37	-	6.79	8.68
F8	4.79	11.51	5.22	8.99	5.34	9.56	6.79	-	8.50
F9	11.47	6.27	4.31	5.44	9.92	6.21	8.68	8.50	-

Docking

Pre-processing of raw Zinc database was done using molecular fingerprint methodologies in MOE in order to remove repeated molecular entries. Further drug-like filters derived from the original "Lipinski rules of five" e. g. logP less than 7; molecular weight between 100-600 Da; number of hydrogen bond donors less than 5; number of hydrogen bond acceptors less than 10; number of rotatable bonds allowed less than 15 were also used during the initial screening of the database.

Active site residues

ARG77, ILE94, GLY97, ALA98, SER99, VAL118, SER123, LEU124, LEU125, ASN127, ALA130, LEU133, MET134, LEU137, LEU161, SER162, GLY164, LEU165, ARG197, THR198, ALA201, VAL202, THR205

Gold docking

Number of independent genetic algorithm runs =5
Pressure = 1.1
Number of genetic algorithm operations =1000
Number of set of Islands used =3
Population size=100
Default operator weights used
Crossover = 95
Mutation = 95
Migration = 10

MOE docking

Methodology = Alpha triangle
Max iterations = 500000
Min Iterations = 300000
Time out (sec) = 600
Scoring function used = London dG