

# MedChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



## Rational design of promiscuous binding modulators of p53 inducing E3(Ub)-ligases (Mdm2 and Pirh2) as anticancer agents: An *in silico* approach†

Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Sarfaraj Niazi<sup>a\*</sup> and Madhusudan Purohit<sup>a\*</sup>

In this paper, for the first time, we report on *in silico* based development of Mdm2 and Pirh2 promiscuous binding small molecule modulators. The methodology involved, capturing the important Mdm2 and Pirh2 interacting residues of p53 TAD and TET respectively, from the literature, protein-protein docking and molecular dynamics simulation studies. The knowledge of important residues delineated was used as the benchmark for the design of Mdm2 and Pirh2 focused ligand libraries; obtained by pharmacophore based screening of 3.9 million compounds deposited in MMSINC<sup>®</sup> database. This was followed by 2D fingerprint similarity based filtering of focused ligand libraries with respect to known reference set of Mdm2 inhibitors, which further confined the chemical space to 608 molecules. These included 365 Mdm2 like small molecule mimetics of p53 TAD and 243 Mdm2 like small molecule mimetics of p53 TET. Docking iterations with respective targets and reverse docking resulted in twelve potential best fit molecules that showed favourable binding interactions with both Mdm2 pDB and Pirh2 CTD. The quality of docking protocol was assessed by using experimentally determined IC<sub>50</sub> values of the known 213 Mdm2 inhibitors and their docking scores with a set of statistical measures which showed good correlations.

### Introduction

Functionally active homotetrameric p53 is a powerful suppressor of tumorigenesis. It serves as an epitome in the present scenario of cancer research due to its unprecedented role as the guardian of genome. p53 serves as the hub for several regulatory pathways involved in apoptosis, DNA repair, senescence, angiogenesis and regulation of cell cycle<sup>1</sup>. Several reports indicated that, nearly all cancer types essentially have either non-functional mutant p53 (~50% of human cancers)<sup>2</sup> or relatively high levels of E3 Ubiquitin (Ub) protein-ligases, which serve as negative regulators of wild type p53<sup>3,4</sup>. In the later case, p53 levels, in concept can be restored by inhibiting its interactions with E3(Ub)-ligases. Hence, the role of p53 induced E3(Ub)-ligases as oncoproteins have been widely acclaimed as rational anticancer therapeutic targets. E3(Ub)-ligases, primarily murine/human double minute 2 (Mdm2/Hdm2) act by facilitating ubiquitylation and proteasomal degradation of p53.

In the recent times, there is a growing interest in the research fraternity for identifying small molecule Mdm2 inhibitors, which

serve as promising anticancer agents and is progressing fast. Nutlins (cis-imidazole analogs) are the first in class, potent inhibitors of Mdm2, identified by screening library of diverse synthetic compounds<sup>5</sup>. Owing to their non-genotoxic effects while maintaining potent inducer of apoptosis, Nutlin class of molecules marks as one of the interesting and promising anticancer agents. Later on, a series of new Mdm2 inhibitors such as, spiro-oxindole<sup>6,7</sup>, benzodiazepinedione analogues<sup>8,9</sup>, etc. were identified which are having similar binding mode as the Nutlins (Supplementary Fig. S1). Apart from Mdm2, similar p53 inducing E3(Ub)-ligases which regulate p53 levels roughly similar to Mdm2 have been deciphered. Pirh2 (p53-induced RING-H2 domain containing protein) is one of the several new p53 inducing E3(Ub)-ligases discovered recently, known to play an apparent secondary role in modulating p53 levels<sup>10,11</sup>. Additionally, COP1, E6-AP, TOPORS, ARF-BP1, synoviolin and E4F1 reportedly also promote ubiquitin mediated degradation of p53<sup>12,13</sup>. However, their detailed mechanisms still remains elusive. With the exception of Mdm2, discoveries of new E3(Ub)-ligases in the light of p53 interactome add more complexity to the p53 degradation pathways<sup>13,14</sup>. One plausible reason for the physiological requirement of such multiple ligases in the mammalian cells is to maintain steady state levels of p53; if one of them turns functionally compromised or deleted<sup>15</sup>. This hypothesis was supported by the conclusions made from earlier studies (review from Brooks et al., 2006<sup>16</sup>), which revealed that inhibitors of Mdm2 alone are not sufficient to rescue p53 from degradation by proteasome. Therefore, there is a strong need to develop promiscuous binding

<sup>a</sup> Department of Pharmaceutical Chemistry, JSS College of Pharmacy, JSS University, Mysuru-570015, India; Phone: +91 (821) 254 8353, Fax: +91 (821) 254 8359

\* To whom correspondence should be addressed:

sarfarazkoilkar@gmail.com, sarfarajniazi@jssuni.edu.in  
mnpurohit04@yahoo.com, madhusudhanpurohit@jssuni.edu.in  
Phone: +91 (821) 254 8353, Fax: +91 (821) 254 8359

† Electronic Supplementary Information (ESI) available: See DOI: 10.1039/x0xx00000x

inhibitors of multiple p53 inducing E3(Ub)-ligases, which may exhibit synergistic effect on p53 stabilization and thereby cancer suppression. The 3D atomic coordinates of two E3(Ub)-ligases, namely Mdm2 and Pirh2 are now available in the RCSB Protein Data Bank (PDB). A pictorial overview of the residues/regions elucidated by several experimental groups<sup>10,14,17–20</sup> for Mdm2, Pirh2 and p53 interactions is shown in the Supplementary Fig. S2.

Structure based virtual screening (SBVS) has gained a prominent spot in the recent scenario particularly in early stage of drug development process wherein, the large databases of natural and/or synthetic product libraries are screened *in silico* against a defined region of therapeutically important target.<sup>21–24</sup> Exploration of “druggable” hot spot regions in the novel biomolecular targets, and rapid virtual screening of large chemical databases using 2D-fingerprint and/or 3D- pharmacophore similarity based methods, often in concert with molecular docking simulations would leverage drug development process in a cost-effective manner.<sup>25–28</sup> In the present study, an attempt was made to design small molecule modulators which show promiscuous binding to multiple p53 E3(Ub)-ligases, in particular Mdm2 and Pirh2 using advanced computational tools.

## Experimental

### Modelling full length structure of Pirh2

The solution state NMR structures of three distinct domains of Pirh2<sup>14</sup>, viz., N-terminal domain (NTD)/p53 Binding Domain (pBD), RING-H2 domain and C-terminal domain (CTD) (UniProtKB sequence ID: Q96PM5; RCSB PDB IDs: 2K2C, 2JRJ and 2K2D) were used to model full length structure of Pirh2. Build Homology Models (MODELER v.9.4) module of Accelrys Discovery studio 3.5 (Acc. DS 3.5) with default settings was used for modeling. The modeled Pirh2 structure was refined using Prime v3.1 of Schrödinger LLC<sup>29</sup> with the implementation of OPLS 2005 force field. This was followed by CHARMM minimization using 400 steps of steepest decent method and Generalized Born as implicit solvent model. The optimized full length structure of Pirh2 was used for further *in silico* studies.

### Docking p53 tetramerization (TET) domain to the CTD of full length Pirh2 structure

Since the structure of Pirh2 complexed with p53 is not available in the RCSB PDB, docking simulations were performed specifically for TET domain of p53 and CTD of full length Pirh2. Crystal structure of p53 TET domain (PDB ID: 1C26)<sup>30</sup> was retrieved from RCSB PDB database and prepared for docking using protein preparation wizard of Acc. DS 3.5. The prepared p53 TET domain was docked to the modelled full length structure of Pirh2 using HADDOCK web server protocol v2.1<sup>31,32</sup>. To set the docking constraints, the p53 TET interacting residues of Pirh2 CTD (Sheng *et al.* 2008) were defined as active residues, while all other solvent-accessible residues were defined as passive residues. To drive the docking simulations, the information of user defined docking constraints were converted into ambiguous interaction restraints by HADDOCK. The protocol follows three stages of docking namely, (a) rigid-body energy minimization, (b) semi-flexible refinement and (c) final model

refinement in explicit solvent. A maximum of 200 water refined models obtained in the final stage of docking run were clustered using pair-wise main-chain RMSD (Root Mean Square Deviation) cut-off of 7.5 and minimum cluster size of 4 as criteria. Several clusters were generated. These clusters were ranked based on HADDOCK scores calculated on the basis of weighted intermolecular energy terms. The two best refined models each from highest ranked cluster and the largest cluster among total clusters obtained from docking run were used as HADDOCK representative structures.

### Molecular Dynamic (MD) simulations of Mdm2 pBD-p53 Trans Activation Domain (TAD) and HADDOCK representative models of Pirh2 CTD-p53 TET complexes

MD simulations were performed until 6 ns for two HADDOCK generated representative Pirh2 CTD-p53 TET docked complexes, viz., or1, or2, obtained each from cluster 1 and cluster 2 respectively, and also for Mdm2 pBD-p53 TAD co-crystal structure (PDB ID: 1YCR<sup>17</sup>). Simulations were performed using GROMACS version 4.5.4<sup>33</sup> with the implementation of the CHARMM27 forcefield<sup>34</sup>. The complexes were solvated using Simple Point Charge (SPC) water molecules in an octahedron box with periodic boundary conditions. The positively charged system was neutralized by adding chloride ions as counter ions. Long range electrostatics was calculated using Particle Mesh Ewald summation with 10 Å cutoff for coulombic interactions<sup>35</sup>. Energy minimization was performed using steepest descent algorithm with a tolerance of 1000 KJ mol<sup>-1</sup> nm<sup>-1</sup>. The atomic positions of the complexes were restrained and the system was equilibrated in NVT (isochoric-isothermal) ensemble for 100 ps at 300K temperature. This was further extended to another 200 ps in NPT (isobaric-isothermal) ensemble or until the pressure was maintained at 1.0 bar. After equilibrating the system, the position restraints were released and subjected to 2 ns (1 million steps) of production run using a time step of 2 fs. The trajectories were saved every 800 steps and analyzed using GROMACS tools. Cluster analysis was performed for all MD trajectories (3 – 6 ns, using Gromos method<sup>36</sup> with backbone atoms RMSD cutoff of either 0.1 or an arbitrary number was chosen such that the total clusters obtained should be less than 100. Middle structures which belonged to the largest clusters were chosen as the representative of each MD trajectory for further *in silico* studies. All the 3D structures were rendered using PyMol (<http://www.pymol.org/>).

### Rational design of promiscuous Mdm2 and Pirh2 binding ligands

#### Screening based on 3D pharmacophore models of p53 residues interacting with Mdm2 and Pirh2

The 3D structures of E3(Ub)-ligase interacting residues of p53 such as TAD (F19, W23 and L26) and TET (L330, M340, F341 and L344) belonged to Mdm2\_p53\_5562ps and Pirh2\_p53\_or2\_5950ps complexes (obtained from cluster analysis of MD frames) respectively (See Fig. 5a and Fig. 5c in the results and discussion section), were used as the templates for pepMMsMIMIC, a web based 3D peptidomimetic compound virtual screening tool<sup>37</sup>. The pepMMsMIMIC web server was earlier used for the design of non

peptide small molecule hit candidates for hFSHR-ECD<sup>38</sup>. The pharmacophore models generated based on input residues of each p53 peptide stretch were used to screen against library of 17 million conformers obtained from 3.9 million commercially available chemical structures present in MMsINC<sup>®</sup> database<sup>39</sup>.

The pepMMsMIMIC web server employs five types of scoring methods to improve the selection of ligands *viz.*, 1) fingerprint based filtering of shape similarity 2) based only on shape similarity and 3) based only on the pharmacophoric similarity 4) shape based filtering of pharmacophoric similarity 5) hybrid search (60% pharmacophoric, 40% shape). Each scoring method was designed to deliver top 200 small molecule peptide mimetics along with similarity scores. Thus Mdm2 and Pirh2 focused libraries each contained 1000 small molecule peptidomimetics were retrieved using 3D atomic coordinates of p53-TAD and p53-TET residues.

#### Fingerprint based similarity searching using 2D similarity coefficients

In general, compounds which share maximum structural and physicochemical similarities are more likely to exhibit similar pharmacological properties. Therefore, to reduce chemical space disparity with the known Mdm2 inhibitors, the focused Mdm2 and Pirh2 ligand libraries retrieved from 3D screening protocol were filtered to obtain sub-libraries that share maximum fingerprints or structural similarities with the reference Mdm2 inhibitors. A list of 32 2D structures of such reference inhibitors are shown in the Supplementary Fig. S3.

Using four 2D similarity coefficients *viz.*, Tanimoto, Dice, Cosine and Target (with fingerprints: FPFC<sub>6</sub>) implemented in Acc. DS 3.5 and a similarity cut-off score of 0.3, molecules in the focused libraries were further screened.

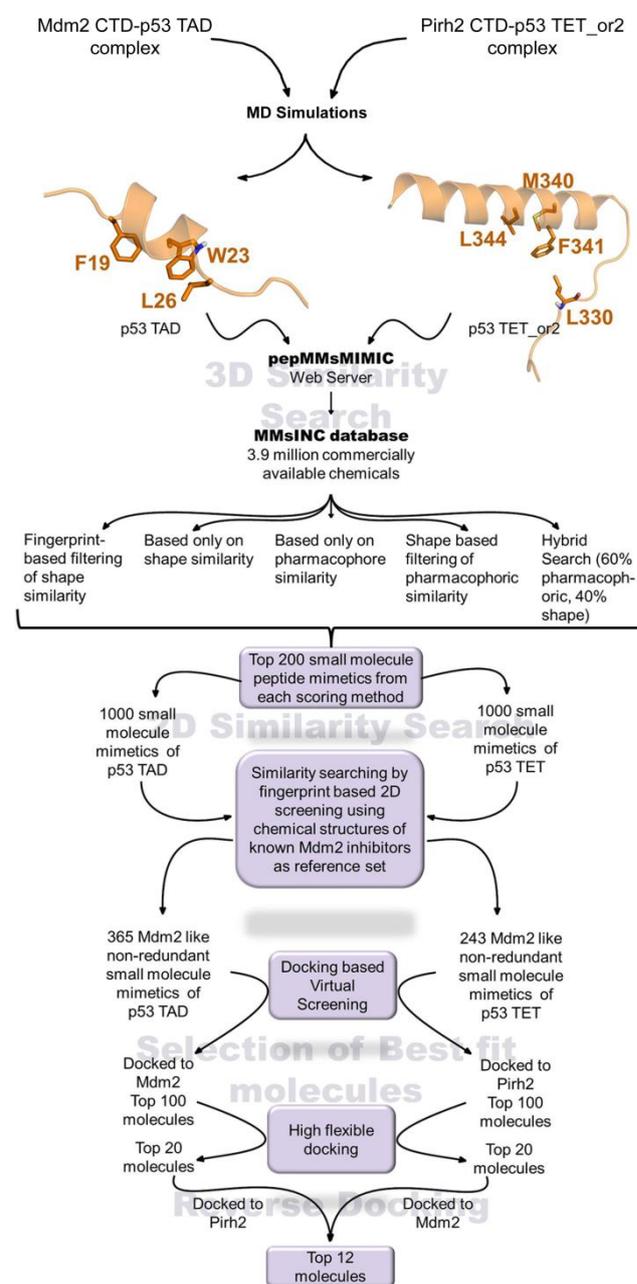
#### Ligand preparation for docking

The Mdm2 and Pirh2 focused sub-libraries retrieved from fingerprint based screening protocol were prepared for docking using Prepare Ligands protocol of Acc. DS. 3.5. The non-redundant ligands from each focused library were discerned, separated and prepared at pH 6.5–8.5 to generate their possible ionisation states.

#### Docking based virtual screening and high flexible docking of small molecule peptide mimetics

Two steps docking protocol (virtual screening and selection of best fit molecules) was carried out independently for the prepared Mdm2 and Pirh2 focused libraries. GOLD v5.2<sup>40,41</sup> was used for aforementioned docking exercise. E3(Ub)-ligases were prepared for docking by detaching p53 peptides from both complexes *viz.* Mdm2\_p53\_5562ps and Pirh2\_p53\_or2\_5950ps which are extracted from MD trajectory. No further protein preparation (adding polar and non-polar hydrogen atoms, including those necessary to define the correct ionization and tautomeric states of residues) was needed, since the structures were found to be optimized. The binding sites in the Mdm2 and Pirh2 were defined by picking C $\beta$  of I99 and C $\gamma$  of H52 respectively, as the centroid atoms and the binding site radius was set to 15 Å, so as to encompass all the important residues. Docking accuracy was set to 30% for virtual screening, while default settings were used for selection of best fit molecules. Top ranked

solutions obtained from each genetic algorithm run were shortlisted based on highest number of interactions with important residues and GoldScore\_Fitness. An overview of methodology followed for rational design of Mdm2 and Pirh2 promiscuous binding ligands has been illustrated as flowchart in the Fig. 1.



**Figure 1.** Flow chart illustrating the methodology followed for rational design of promiscuous binding Mdm2 and Pirh2 hits using *in silico* approach.

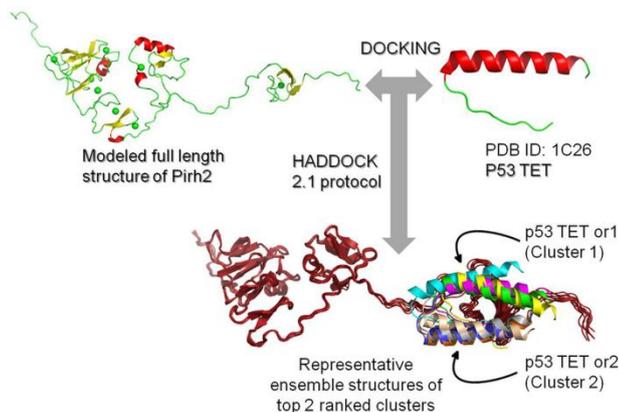
## Results and discussion

Reports from biochemical and protein-protein interaction experiments conducted earlier revealed strong binding affinity of Pirh2 CTD towards p53 TET<sup>14</sup>. Therefore, *in silico* structure based studies were carried out for CTD of modeled full length Pirh2

structure and p53 TET, in order to understand their binding modes. As the co-crystal structure of p53 TET-Pirh2 CTD complex is not available in the RCSB PDB, docking studies were performed for CTD of modeled Pirh2 and p53 TET using HADDOCK 2.1 web server. Post docking, HADDOCK clustered 169 Water Refined Models (WRMs) among a maximum of 200 WRMs in 8 clusters, which represents 84.5 % of the total WRMs generated. These clusters were ranked based on HADDOCK scores calculated on the basis of weighted intermolecular energy terms. Of the 8 clusters generated, cluster 1 had highest cluster size (66) whereas, cluster 2 had lowest mean HADDOCK score (-119.5 +/- 14.3). Results of other parameters evaluated for cluster 1 and cluster 2 are shown in Table 1. Examination of superimposed representative models of top 2 ranked clusters revealed two possible binding orientations for p53 TET viz., or1 (cluster 1) and or2 (cluster 2) (Fig. 2). To evaluate the stability of these representative models, along with Mdm2-p53 complex (PDB ID: 1YCR), MD simulations were carried out until 6 ns.

**Table 1.** Parameter values calculated for HADDOCK generated top two ranked clusters of Pirh2-p53 TET docked complexes

Parameters	Cluster 1	Cluster 2
HADDOCK score	-82.9 +/- 5.0	-119.5 +/- 14.3
Cluster size	66	38
RMSD from the overall lowest-energy structure	4.4 +/- 0.2	1.2 +/- 0.7
Van der Waals energy	-78.9 +/- 8.9	-92.0 +/- 3.8
Electrostatic energy	-387.9 +/- 106.4	-401.3 +/- 59.5
Desolvation energy	6.8 +/- 5.4	-9.0 +/- 8.1
Restraints violation energy	668.0 +/- 75.13	617.8 +/- 22.28
Buried Surface Area	2513.0 +/- 108.3	2638.0 +/- 91.4
Z-Score	-0.9	-2.1

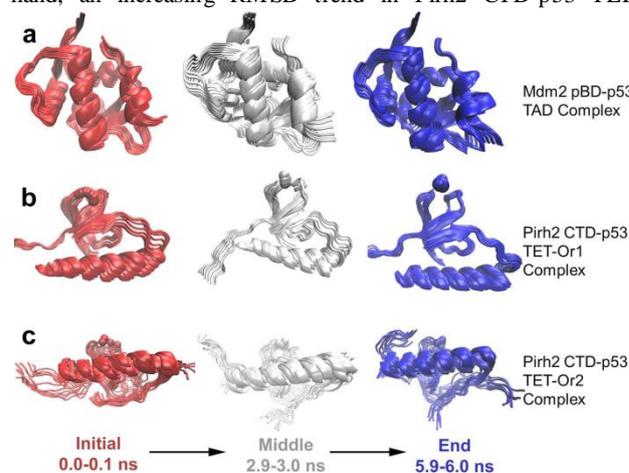


**Figure 2.** Illustration of two possible binding modes or orientations of p53 TET viz., or1 (cluster 1) and or2 (cluster 2) predicted by HADDOCK 2.1 protocol.

Multiple frames were extracted from the beginning, middle and at the end of each MD trajectory. Structural analysis of these multiple frames displayed at different time intervals indicated relatively less main chain fluctuations in Mdm2 pBD-p53 TAD (Fig. 3a) and Pirh2 CTD-p53 TET or1 (Fig. 3b) complexes as compared with Pirh2

CTD-p53 TET or2 complex. Higher fluctuations in the terminal loop regions was observed in Pirh2 CTD-p53 TET or2 complex (Fig. 3c). However, middle region of Pirh2 CTD and p53 TET or2 proteins exhibited relatively lower fluctuations which are evident from RMSF plots (Fig. 4f and Fig. 4h).

Furthermore, parameters such as main-chain (N, C $\alpha$ , C) Root Mean Square Deviation (RMSD), Rg (Radius of gyration), distance, H bonds and residue-wise average Root Mean Square Fluctuation (RMSF) were plotted to verify the stability of complexes with respect to time (Fig. 4a-h). A relatively stable trend in RMSD, Rg, distance and number of H bonds was observed in case of Mdm2-p53 complex as compared with Pirh2 CTD-p53 TET or1 and Pirh2 CTD-p53 TET or2 complexes. The Pirh2 CTD-p53 TET or1 complex was seen fluctuated with a maximum RMSD of 0.5 nm. On the other hand, an increasing RMSD trend in Pirh2 CTD-p53 TET or2



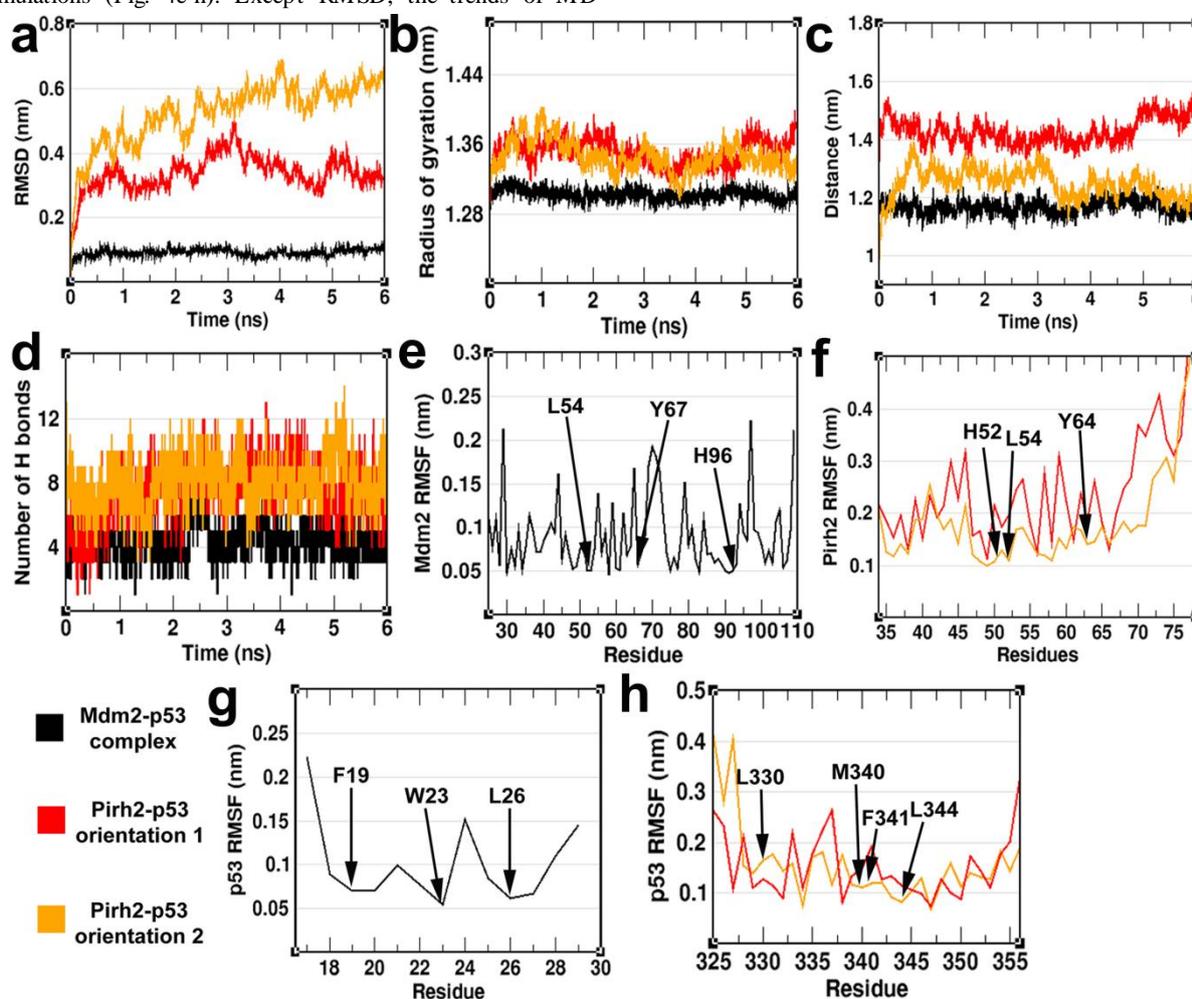
**Figure 3.** Graphical representation of the multiple complexes (a) Mdm2 pBD-p53 TAD, (b) Pirh2 CTD-p53 TET or1 and (c) Pirh2 CTD-p53 TET or2 complexes) obtained during MD simulations over time. Multiple frames were rendered using VMD 1.9.2 from the beginning of the trajectory in red (0.0-0.1 ns), at the middle in white (2.9-3.0 ns), and at the end in blue (5.9-6.0 ns).

trajectory was observed till the end of the simulation time, which was due to the fluctuating terminal loop regions (Fig. 4a). Interestingly Rg which is a measure of structural compactness, was found to be relatively low in case of Pirh2 CTD-p53 TET or2 indicating p53 TET or2 maintained higher number of contacts with Pirh2 CTD as compared with Pirh2 CTD-p53 TET or1 (Fig. 4b). Additionally, the distance between Pirh2 CTD and p53 TET or2 in case of Pirh2 CTD-p53 TET or2 was found on par compared with the distance between Mdm2 and p53 after 3.5 ns till the end of the simulation (Fig. 4c).

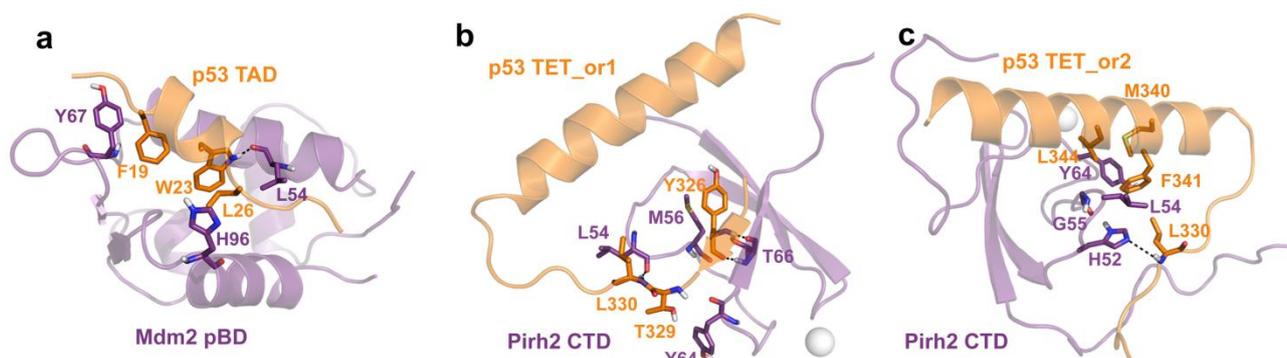
Moreover, on an average, Pirh2 CTD formed eight H bond contacts with p53 TET or2 whereas, six H bond contacts were formed with the p53 TET or1 during the course of MD simulation. The average number of H bond contacts between Mdm2 and p53 during the course of MD simulation was found to be four which remained intact until the end of the simulation time (Fig. 4d). Residue-wise average RMSF plots indicated that the labeled interface residues (Fig. 4e-h) showed least fluctuations and maintained regular noncovalent

contacts throughout the simulations. As opposed to the RMSD trends, the average residue wise RMSF of Pirh2 and p53 during simulations was found to be lower in case of Pirh2 CTD-p53 TET or2 as compared with or1 complex. Clustering Mdm2-p53 MD trajectory frames using RMSD cutoff of 0.10 generated 79 clusters, of which the largest cluster had 268 structures. While clustering Pirh2\_p53\_or1 trajectory generated 59 clusters using a RMSD cutoff of 0.17 and the largest cluster had 202 structures. Furthermore, clustering Pirh2\_p53\_or2 trajectory generated 85 clusters using RMSD cutoff of 0.15 and the largest cluster had 128 structures. Middle structures belonged to the largest clusters viz., Mdm2\_p53\_5562ps, Pirh2\_p53\_or1\_5916ps and Pirh2\_p53\_or2\_5950ps were chosen as the representatives of each MD trajectories for further *in silico* studies. Analysis of interacting interface residues (Mdm2 pBD: L54, Y67, H96; p53 TAD: F19, W23, L26; Pirh2 CTD: H52, L54, Y64; p53 TET: L330, M340, F341, L344) of the representative MD structures (Fig. 5a-c) and their corresponding RMSF plots revealed minimal fluctuations during MD simulations (Fig. 4e-h). Except RMSD, the trends of MD

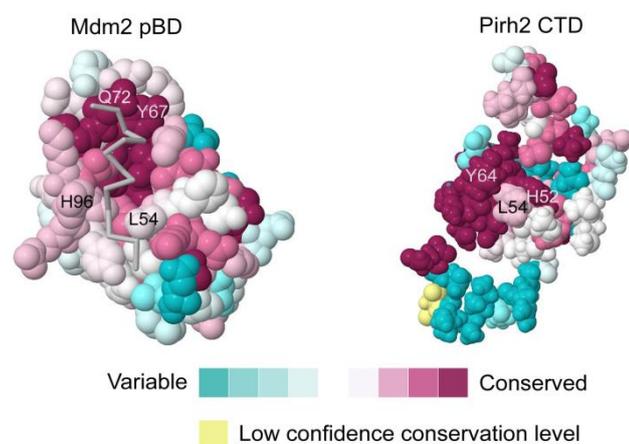
parameters such as, distance, Rg, H Bonds and RMSF of Pirh2 CTD-p53 TET or2 complex were superior as compared with Pirh2 CTD-p53 TET or1. Therefore, the representative 3D structure of the former complex i.e. Pirh2\_p53\_or2\_5950ps along with Mdm2\_p53\_5562ps were further utilized for ligand screening exercise. The Mdm2 pBD and Pirh2 CTD residues were further analysed for evolutionary conservation based on the phylogenetic relations between the non-redundant homologous sequences of UniRef90 database.<sup>42</sup> As expected, the p53 interacting interface residues viz., L54, Y67, H96 and H52, L54, Y64 of Mdm2 pBD and Pirh2 CTD respectively, are highly conserved which further fortify their functional importance in p53 binding (Fig. 6). Furthermore, the Pirh2 CTD residues identified to be crucial for p53 ubiquitylation based on previously reported *in vitro* biochemical studies<sup>14</sup>, the evolutionary conservation of select Pirh2 residues reported herein will further refine our understanding about their functional importance, which can be utilised for target specific ligand identification.



**Figure 4.** Analysis of MD parameters such as, (a) main chain RMSD (b) Rg (c) distance (d) H bonds (e) residue wise average RMSF plot of Mdm2 (f) residue wise average RMSF plots of Pirh2 in the Pirh2 CTD-p53 TET or1 and Pirh2 CTD-p53 TET or2 complexes (g) residue wise average RMSF plot of Mdm2 binding p53 TET domain and (h) residue wise average RMSF plots of p53 in the Pirh2 CTD-p53 TET or1 and Pirh2 CTD-p53 TET or2 complexes.



**Figure 5.** Analysis of protein-protein interactions in the representative middle structures viz., (a) Mdm2 pBD-p53 TAD\_5562ps, (b) Pirh2 pBD-p53 TET\_or1\_5916ps and (c) Pirh2 pBD-p53 TET\_or2\_5950ps complexes extracted from the largest clusters obtained from clustering of MD trajectory frames. The labelled Pirh2 CTD residues are numbered as per NMR structure, RCSB PDB ID: 2K2D.



**Figure 6.** Sequence conservation analysis for 3D structures of Mdm2 pBD and Pirh2 CTD and their p53 binding sites. The amino acid residues are colour coded according to the sequence conservation. The residues important for p53 binding which also showed high degree of conservation are labelled.

### Rational design of promiscuous binding ligands for Mdm2 and Pirh2

*Mdm2 and Pirh2 focused library designed on 3D pharmacophore models of p53 TAD and TET residues*

Five scoring methods (described in Methods) implemented in pepMMsMIMIC web server were used to create Mdm2 and Pirh2 focused libraries. Top 2000 molecules (200 molecules per each scoring method) from 3.9 million commercially available chemical structures present in MMsINC<sup>®</sup> database were retrieved using pepMMsMIMIC web server based on the pharmacophore models generated for p53 TAD and TET residues.

*Screening based on 2D fingerprint similarity with reference Mdm2 inhibitors*

In order to refine the above focused libraries, four 2D similarity coefficients (see methods) and a similarity cut-off score of 0.3 were used to screen structurally similar ligands against 32 experimentally tested reference Mdm2 inhibitors. Summary of number of most similar molecules retrieved from 3D pharmacophore and 2D similarity search protocols are shown in Table 2 and Table 3. The most similar ligands thus obtained were further curated to remove redundant structures. The total size of the focused libraries was therefore reduced from 2000 to 608 non-redundant structures (Mdm2 focused - 365 Mdm2 like small molecule mimetics of p53 TAD and Pirh2 focused -243 Mdm2 like small molecule mimetics of p53 TET).

**Table 2.** Summary of the scores of most similar ligands with respect to Mdm2 pBD binding p53 residues and number of Mdm2 inhibitor like molecules retrieved using 2D similarity coefficients

Sl. No.	Scoring functions implemented in pepMMsMIMIC web server	Molecules with 3D pharmacophore similarity scores (range) with respect to with p53 TAD residues (Trp23, Phe19, Leu26)	Number of molecules retrieved using minimum 2D similarity coefficient cut-off score of 0.3 with respect to reference Mdm2 ligands			
			Tanimoto	Dice	Cosine	Target
1	Fingerprint-based filtering of shape similarity	0.86-0.77	4	29	29	52
2	Based only on shape similarity	0.89-0.83	0	51	53	89
3	Based only on pharmacophore similarity	1.00	2	132	133	107
4	Shape based filtering of pharmacophoric similarity	1.00-0.39	2	91	91	106
5	Hybrid Search (60% pharmacophoric, 40% shape)	0.88-0.73	2	131	132	103

**Table 3.** Summary of the scores of most similar ligands with respect to Pirh2 CTD binding p53 residues and number of Mdm2 inhibitor like molecules retrieved using 2D similarity coefficients

Sl. No.	Scoring functions implemented in pepMMsMIMIC web server	Molecules with 3D pharmacophore similarity scores (range) with respect to p53 TET residues (Leu330, Phe341, Leu344, Met 340)	Number of molecules retrieved using minimum 2D similarity coefficient cut-off score of 0.3 with respect to reference Mdm2 ligands			
			Tanimoto*	Dice	Cosine	Target
1	Fingerprint-based filtering of shape similarity	0.84-0.74	16	39	40	68
2	Based only on shape similarity	0.89-0.82	20	37	39	70
3	Based only on pharmacophore similarity	0.78-0.35	3	12	13	47
4	Shape based filtering of pharmacophoric similarity	0.78-0.30	4	14	14	59
5	Hybrid Search (60% pharmacophoric, 40% shape)	0.70-0.44	4	16	16	71

\* In this case, the minimum Tanimoto coefficient cutoff score of 0.2 was used.

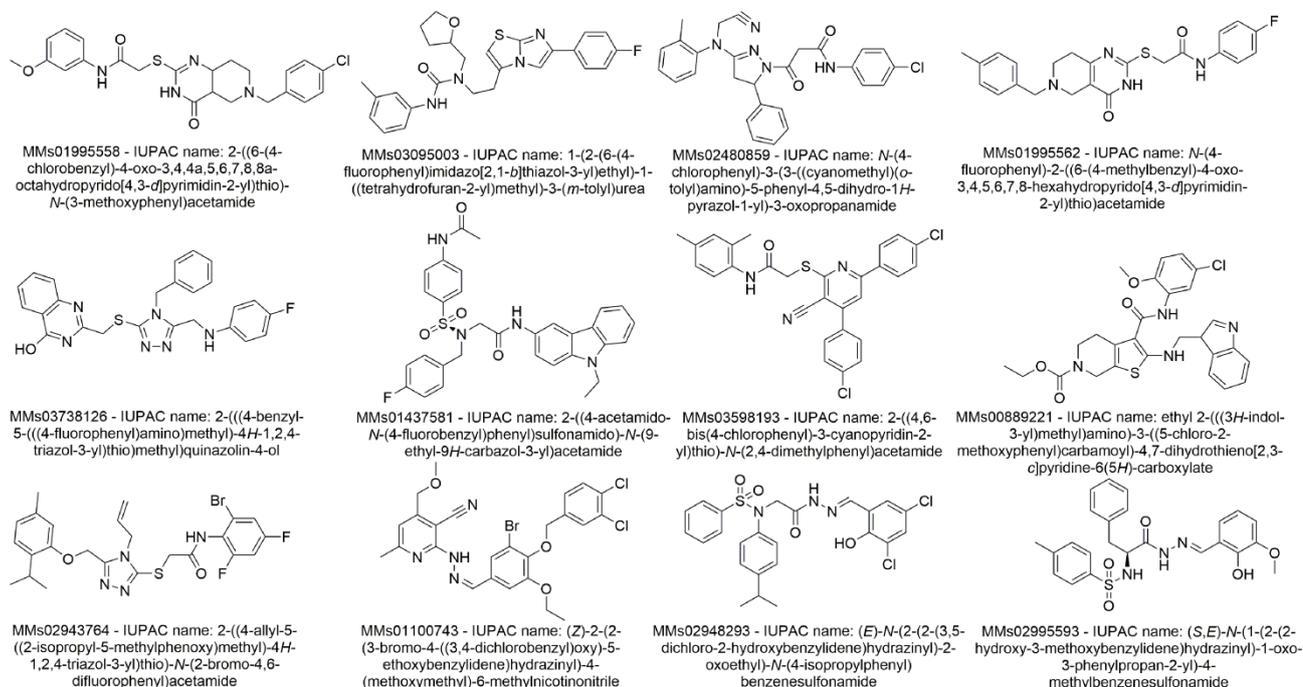
Best fit molecules identified based on docking based virtual screening, high flexible docking and reverse docking.

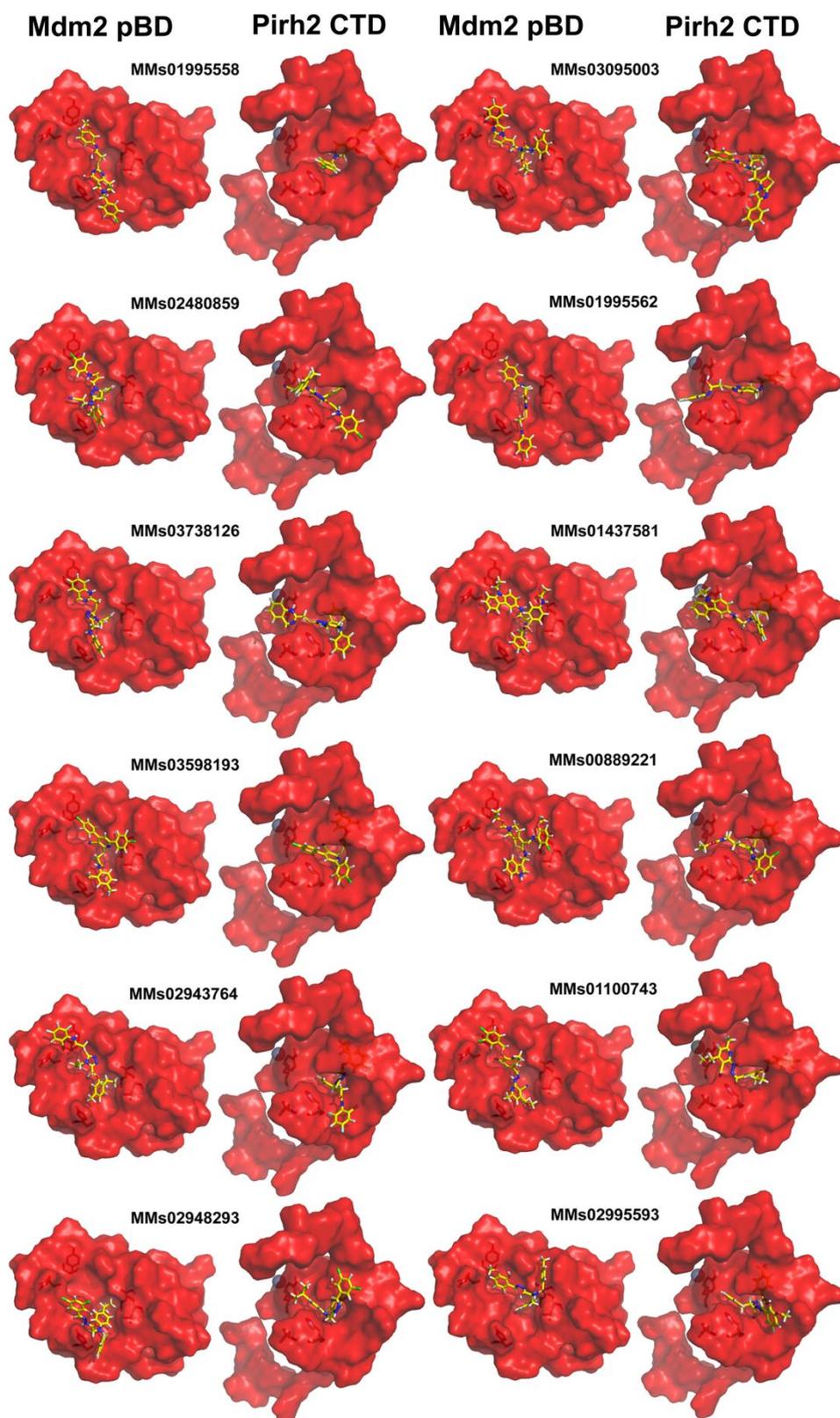
Docking simulations was carried out for the selection of best fit molecules that could bind to both Mdm2 pBD and Pirh2 CTD. Two steps docking protocol, initially for screening and later for selection of best fit molecules was carried out. The non-redundant Mdm2 and Pirh2 focused libraries obtained from 3D and 2D virtual screening were docked to respective E3(Ub)-ligases. The 100 best molecules were selected from each docking exercise (Mdm2 and Pirh2 separately, total 200 best molecules) based on GoldScore\_Fitness cutoff of 50 as the criteria. The above docking exercise was repeated using high flexible mode (default) of Gold docking algorithm for the above shortlisted molecules. Based on highest number of nonbonding interactions with the p53 interacting residues of Mdm2

pBD and Pirh2 CTD along with minimum steric clashes, 12 molecules (20 molecules each for Mdm2 and Pirh2) were further shortlisted from each docking output.

In the final stage of *in silico* work flow, reverse docking was carried out i.e. shortlisted "Mdm2 pBD docked ligands" were docked to Pirh2 CTD and vice-versa, using high flexible mode of GOLD docking algorithm. Finally, the 12 best fit molecules were selected based on highest GoldScore\_Fitness and interactions with important p53 binding residues of Mdm2 pBD and Pirh2 CTD.

The 3D pharmacophore shape similarity scores, 2D fingerprint similarity scores and GoldScores of the identified Mdm2 and Pirh2 promiscuous binding best fit ligands are summarised in the Supplementary Table S1. The 2D chemical structures and the pose of the twelve best fit molecules binding to both Mdm2 pBD and Pirh2 CTD are showed in the Fig. 7 and Fig. 8.

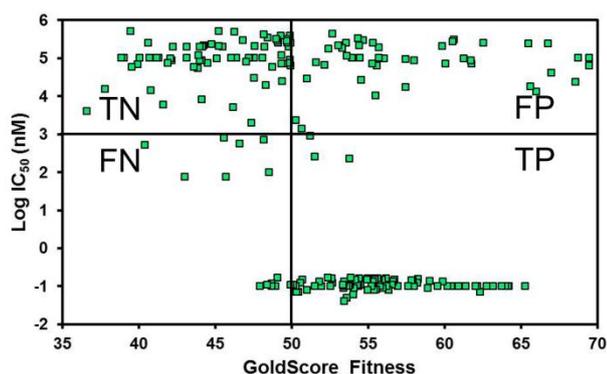
**Figure 7.** 2D structures and chemical names of identified best fit molecules for Mdm2 pBD and Pirh2 CTD.



**Figure 8.** Poses of twelve best fit ligands that bind to p53 TAD and TET interacting regions of Mdm2 pBD and Pirh2 CTD, respectively, predicted by GOLD docking algorithm. The best fit ligands were shown as yellow carbon sticks whereas the Mdm2 pBD and Pirh2 CTD were shown as surface representation. Interacting residues were rendered as transparent sticks.

### Validation of docking protocol

The docking protocol was validated using 213 compounds whose  $IC_{50}$  values were experimentally determined for their potential to bind to Mdm2 (Supplementary Table S2). Of the 213 compounds, 95 had  $IC_{50}$  values  $<1,000$  nM, and hence were considered actives (true positives) and the remaining 118 compounds whose  $IC_{50}$  values were  $>1,000$  nM, thus considered inactive (true negatives). The 3D structures of the above compounds were obtained from BindingDB database<sup>43</sup> and docked to Mdm2 pBD using GOLD. Using an arbitrary GoldScore\_Fitness cut-off of 50, the docking algorithm correctly predicted 80 of the 95 actives as True Positives (TP) and 74 of the 118 inactives as True Negatives (TN). 44 of the inactives and 15 of the actives were wrongly predicted as False Positives (FP) and False Negatives (FN), respectively (Fig. 9). The above data was used to calculate (a) sensitivity, (b) specificity, (c) precision and (d) accuracy of the docking protocol. The sensitivity, specificity, precision and accuracy of the docking protocol were found to be  $\sim 0.8$ ,  $0.6$ ,  $0.6$  and  $0.7$  respectively. Furthermore, Matthews Correlation Coefficient (MCC) and Cohen's kappa ( $\kappa$ ) index were calculated to assess the prediction quality of GOLD docking algorithm. The MCC and  $\kappa$  were found to be  $\sim 0.5$ . The results of all the above statistical measures indicated a good correlation between experimental  $IC_{50}$  values and Goldscore\_Fitness, and therefore the quality of docking protocol.



**Figure 9.** Plot of experimentally determined  $\log IC_{50}$  values of 213 compounds versus their GoldScore\_Fitness predicted for Mdm2 pBD binding.

### Conclusions

In the current study, for the first time, an attempt was made to design Mdm2 and Pirh2 promiscuous binding small molecule ligands using the knowledge of important Mdm2 and Pirh2 interacting p53 residues obtained from the biochemical findings reported in the literature, protein-protein docking and MD simulation studies. Focused ligand libraries consisting of 608 molecules with confined chemical space were designed based on the 3D pharmacophore shape similarity to the important residues of p53 TAD and TET and 2D similarity to reference set of Mdm2 inhibitors. These included 365 Mdm2 like small molecule mimetics of p53 TAD and 243 Mdm2 like small molecule mimetics of p53 TET. Docking

simulations with respective targets and reverse docking using GOLD, resulted in twelve potential best fit molecules that showed favourable binding interactions with both Mdm2 pDB and Pirh2 CTD. The GOLD docking algorithm was validated using experimentally determined  $IC_{50}$  values of 213 known compounds and their docking scores. The results of statistical measures calculated *viz.*, sensitivity, specificity, precision, accuracy of the docking protocol; additionally, MCC and  $\kappa$  index indicated a good correlation between experimental  $IC_{50}$  values and Goldscore\_Fitness. In conclusion, the *in silico* designed promiscuous E3(Ub)-ligase binding hit candidates would be further complemented with *in vitro* experimental testing to corroborate their binding promiscuity towards E3(Ub)-ligases and to establish the proof-of-concept which is under way in our laboratory. It has been reported that the persistent use of Mdm2 inhibitors alone may result in tumor resistance<sup>4</sup>, therefore these potential promiscuous binding hit candidates may be highly desirable and could find applications in the treatment of cancer.

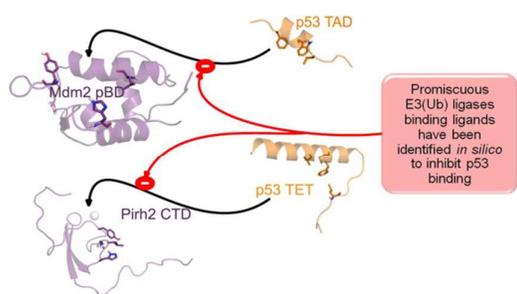
### Acknowledgements

We thank the Principal, JSS College of Pharmacy, JSS University Mysuru for providing necessary facilities. Sarfaraj Niazi thanks Indian Council of Medical Research (ICMR), Govt. of India, New Delhi for the award of Senior Research Fellowship (SRF), No. BIC/11(16)/2014, IRIS ID No. 2014-23020.

### References

- 1 C. C. Harris, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **107**, 1659–1660.
- 2 S. Haupt, M. Berger, Z. Goldberg and Y. Haupt, *J. Cell Sci.*, 2003, **116**, 4077–4085.
- 3 B. Vogelstein, S. Sur and C. Prives, *Nat Educ*, 2010, **3**, 6.
- 4 S. Shangary and S. Wang, *Clin. Cancer Res.*, 2008, **14**, 5318–5324.
- 5 L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi and E. a Liu, *Science*, 2004, **303**, 844–848.
- 6 S. Shangary and S. Wang, *Annu. Rev. Pharmacol. Toxicol.*, 2009, **49**, 223–241.
- 7 K. Ding, Y. Lu, Z. Nikolovska-Coleska, G. Wang, S. Qiu, S. Shangary, W. Gao, D. Qin, J. Stuckey, K. Krajewski, P. P. Roller and S. Wang, *J. Med. Chem.*, 2006, **49**, 3432–3435.
- 8 H. K. Koblish, S. Zhao, C. F. Franks, R. R. Donatelli, R. M. Tominovich, L. V LaFrance, K. a Leonard, J. M. Gushue, D. J. Parks, R. R. Calvo, K. L. Milkiewicz, J. J. Marugán, P. Raboisson, M. D. Cummings, B. L. Grasberger, D. J. Johnson, T. Lu, C. J. Molloy and A. C. Maroney, *Mol. Cancer Ther.*, 2006, **5**, 160–169.
- 9 B. L. Grasberger, T. Lu, C. Schubert, D. J. Parks, T. E. Carver, H. K. Koblish, M. D. Cummings, L. V LaFrance, K. L. Milkiewicz, R. R. Calvo, D. Maguire, J. Lattanze, C.

- F. Franks, S. Zhao, K. Ramachandren, G. R. Bylebyl, M. Zhang, C. L. Manthey, E. C. Petrella, M. W. Pantoliano, I. C. Deckman, J. C. Spurlino, A. C. Maroney, B. E. Tomczuk, C. J. Molloy and R. F. Bone, *J. Med. Chem.*, 2005, **48**, 909–912.
- 10 R. P. Leng, Y. Lin, W. Ma, H. Wu, B. Lemmers, S. Chung, J. M. Parant, G. Lozano, R. Hakem and S. Benchimol, *Cell*, 2003, **112**, 779–791.
- 11 J. Shloush, J. E. Vlassov, I. Engson, S. Duan, V. Saridakis, S. Dhe-Paganon, B. Raught, Y. Sheng and C. H. Arrowsmith, *J. Biol. Chem.*, 2011, **286**, 4796–4808.
- 12 D. Dornan, I. Wertz, H. Shimizu, D. Arnott, G. D. Frantz, P. Dowd, K. O'Rourke, H. Koeppe and V. M. Dixit, *Nature*, 2004, **429**, 86–92.
- 13 E. Tai and S. Benchimol, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 11431–11432.
- 14 Y. Sheng, R. C. Laister, A. Lemak, B. Wu, E. Tai, S. Duan, J. Lukin, M. Sunnerhagen, S. Srisailam, M. Karra, S. Benchimol and C. H. Arrowsmith, *Nat. Struct. Mol. Biol.*, 2008, **15**, 1334–1342.
- 15 D. Hoeller and I. Dikic, *Nature*, 2009, **458**, 438–444.
- 16 C. L. Brooks and W. Gu, *Mol. Cell*, 2006, **21**, 307–315.
- 17 P. H. Kussie, S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A. J. Levine and N. P. Pavletich, *Science*, 1996, **274**, 948–953.
- 18 H. Shimizu, L. R. Burch, A. J. Smith, D. Dornan, M. Wallace, K. L. Ball and T. R. Hupp, *J. Biol. Chem.*, 2002, **277**, 28446–28458.
- 19 M. Wallace, E. Worrall, S. Pettersson, T. R. Hupp and K. L. Ball, *Mol. Cell*, 2006, **23**, 251–263.
- 20 G. W. Yu, S. Rudiger, D. Veprintsev, S. Freund, M. R. Fernandez-Fernandez and A. R. Fersht, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 1227–1232.
- 21 Z. Zhong, L. J. Liu, Z.-Q. Dong, L. Lu, M. Wang, C.-H. Leung, D.-L. Ma and Y. Wang, *Chem. Commun.*, 2015, **51**, 11178–11181.
- 22 D. Ma, D. S. H. Chan, G. Wei, H. Zhong, H. Yang, L. To, L. T. Leung, E. a Gullen, P. Chiu, Y.-C. Cheng and C.-H. Leung, *Chem. Commun. (Camb)*, 2014, **50**, 13885–13888.
- 23 L. J. Liu, K. H. Leung, D. S. H. Chan, Y. T. Wang, D. L. Ma and C. H. Leung, *Cell Death Dis.*, 2014, **5**, e1293.
- 24 D. S. H. Chan, H. M. Lee, F. Yang, C. M. Che, C. C. L. Wong, R. Abagyan, C. H. Leung and D. L. Ma, *Angew. Chemie*, 2010, **122**, 2922–2926.
- 25 D. L. Ma, D. S. H. Chan and C. H. Leung, *Chem. Soc. Rev.*, 2013, **42**, 2130–2141.
- T. Cheng, Q. Li, Z. Zhou, Y. Wang and S. H. Bryant, *AAPS J.*, 2012, **14**, 133–141.
- P. Willett, *Drug Discov. Today*, 2006, **11**, 1046–1053.
- D. B. Kitchen, H. Decornez, J. R. Furr and J. Bajorath, *Nat. Rev. Drug Discov.*, 2004, **3**, 935–949.
- N. Prime version 3.1, Schrödinger, LLC, New York, 2013.
- P. D. Jeffrey, S. Gorina and N. P. Pavletich, *Science*, 1995, **267**, 1498–502.
- T. A. Wassenaar, M. Dijk, N. Loureiro-Ferreira, G. Schot, S. J. Vries, C. Schmitz, J. Zwan, R. Boelens, A. Giachetti, L. Ferella, A. Rosato, I. Bertini, T. Herrmann, H. R. A. Jonker, A. Bagaria, V. Jaravine, P. Güntert, H. Schwalbe, W. F. Vranken, J. F. Doreleijers, G. Vriend, G. W. Vuister, D. Franke, A. Kikhney, D. I. Svergun, R. H. Fogh, J. Ionides, E. D. Laue, C. Spronk, S. Jurkša, M. Verlato, S. Badoer, S. Dal Pra, M. Mazzucato, E. Frizziero and A. M. J. J. Bonvin, *J. Grid Comput.*, 2012, **10**, 743–767.
- S. J. de Vries, M. van Dijk and A. M. J. J. Bonvin, *Nat. Protoc.*, 2010, **5**, 883–897.
- B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
- P. Bjelkmar, P. Larsson, M. A. Cuendet, B. Hess and E. Lindahl, *J. Chem. Theory Comput.*, 2010, **6**, 459–466.
- T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089.
- X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angew. Chemie Int. Ed.*, 1999, **38**, 236–240.
- M. Floris, J. Masciocchi, M. Fanton and S. Moro, *Nucleic Acids Res.*, 2011, **39**, W261–9.
- A. Sonawani, S. Niazi and S. Idicula-Thomas, *PLoS One*, 2013, **8**, e64475.
- J. Masciocchi, G. Frau, M. Fanton, M. Sturlese, M. Floris, L. Pireddu, P. Palla, F. Cedrati, P. Rodriguez-Tomé and S. Moro, *Nucleic Acids Res.*, 2009, **37**, D284–290.
- G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, *J. Mol. Biol.*, 1997, **267**, 727–748.
- GOLD, version 5.0 User Guide. CCDC Software Ltd Cambridge, 2010.
- H. Ashkenazy, E. Erez, E. Martz, T. Pupko and N. Ben-Tal, *Nucleic Acids Res.*, 2010, **38**, W529–533.
- BindingDB database, <https://www.bindingdb.org/bind/index.jsp>, (accessed May 2015).

Graphical Abstract

Twelve promiscuous binding p53 inducing E3(Ub) ligases (Mdm2 and Pirh2) hit candidates have been identified by structure based virtual screening