

Supplementary information for:

**Computational strategy for intrinsically disordered protein
ligand design leads to the discovery of p53 transactivation
domain I binding compounds that activate the p53 pathway**

*Hao Ruan[†], Chen Yu[†], Xiaogang Niu^{§,‡}, Weilin Zhang[†], Hanzhong Liu[⊥], Limin Chen^{//},
Ruoyao Xiong[†], Qi Sun[†], Changwen Jin^{§,‡}, Ying Liu^{†,⊥,*}, Luhua Lai^{†,⊥,/**}*

[†]BNLMS, State Key Laboratory for Structural Chemistry of Unstable and Stable Species, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China.

[§]College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China.

[‡]Beijing Nuclear Magnetic Resonance Center, Peking University, Beijing 100871, China.

[⊥]Center for Quantitative Biology, Academy of Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China.

^{//}Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China.

*Corresponding authors:

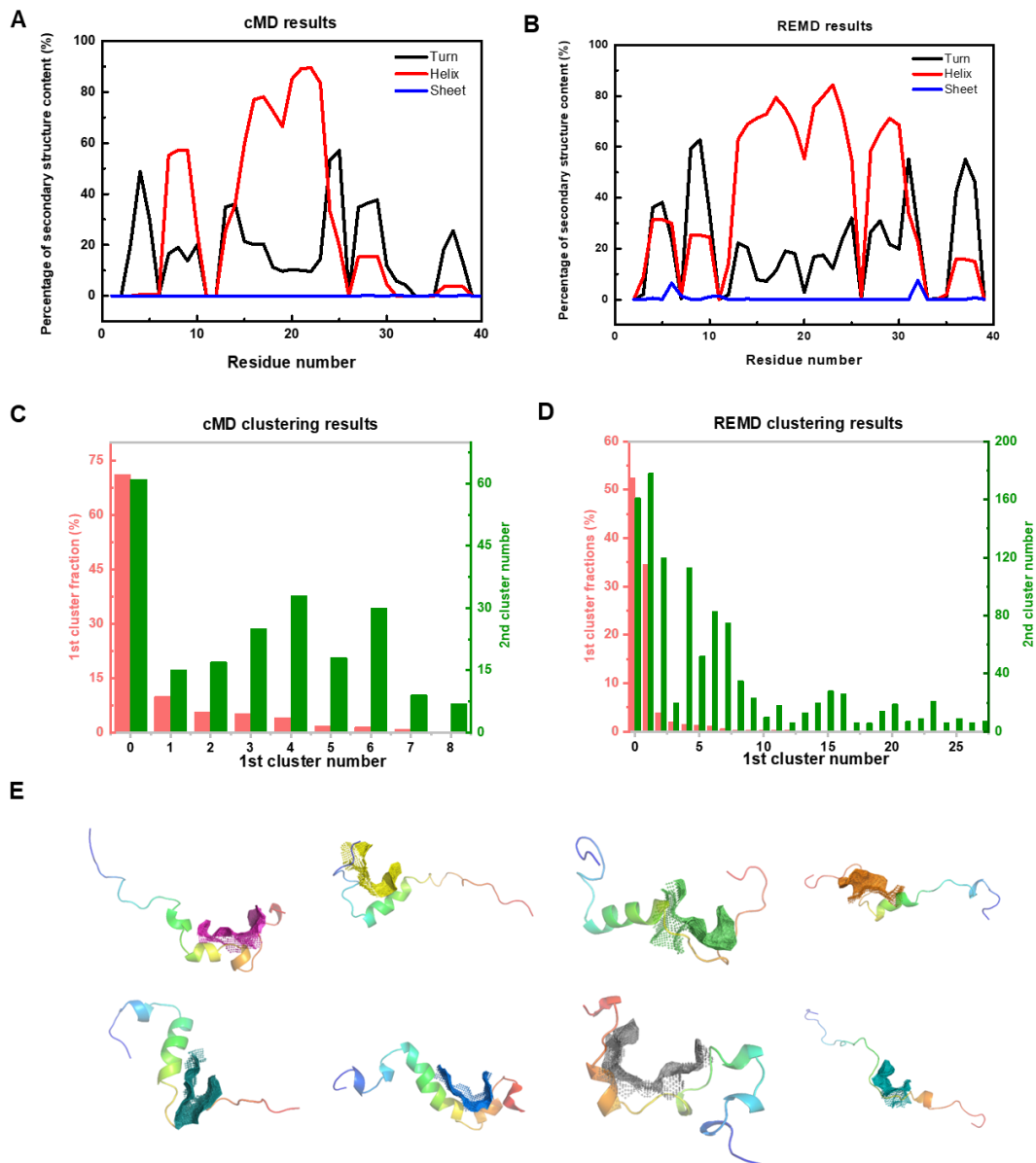
Ying Liu, Email: liuying@pku.edu.cn; Tel: 861062751490

Luhua Lai, Email: lhlai@pku.edu.cn; Tel: 861062757486

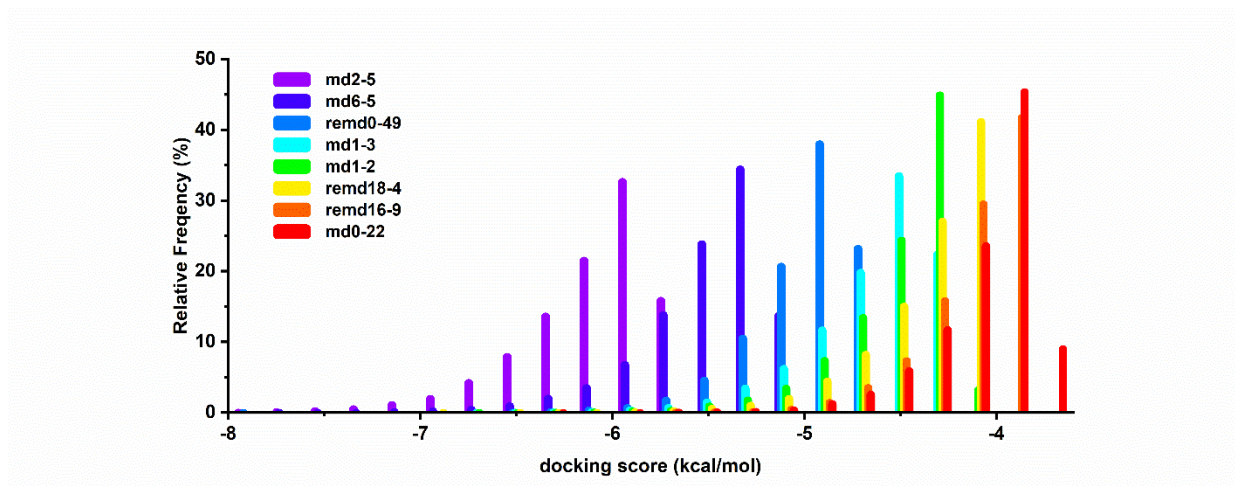
This SI includes:

Supplementary Figs. S1 to S15

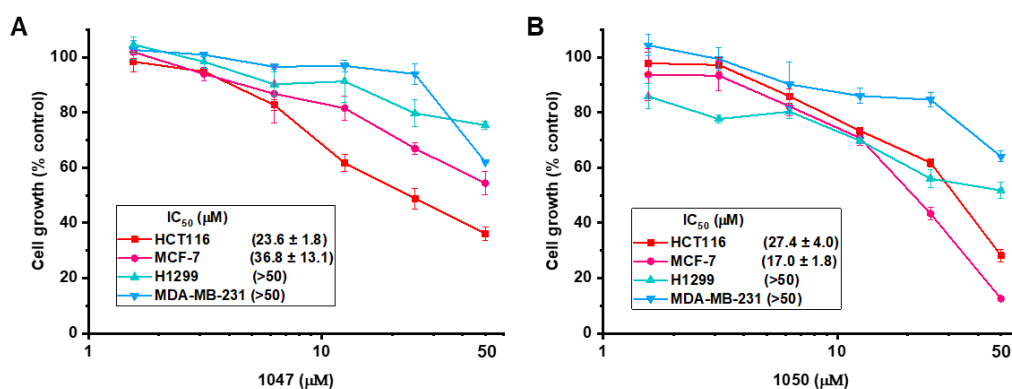
Supplementary Tables S1 to S7



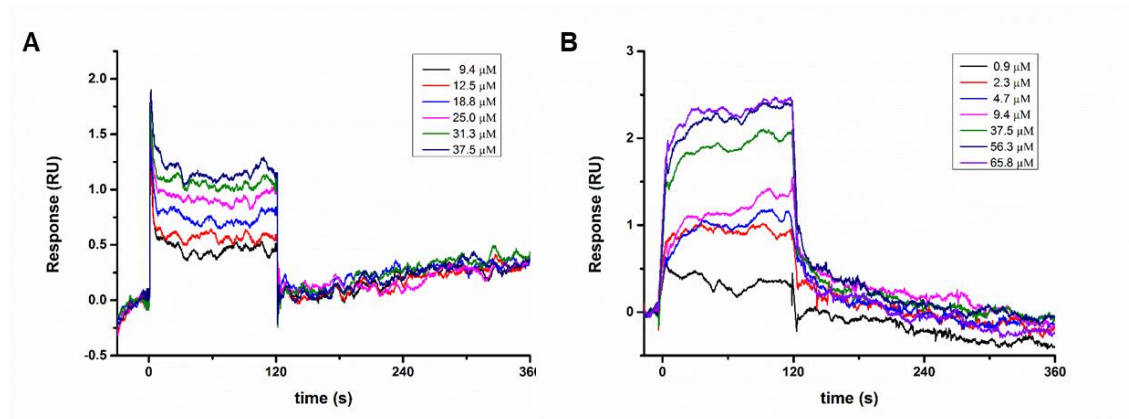
Supplementary Figure S1. Secondary structure and pockets detection analysis of p53 TAD1 trajectories. (A, B) Secondary structure content of p53 TAD1 trajectories from classical all-atom MD simulations with initial structure from the p53 TAD1-MDMX complex, and from the replica exchange molecular dynamics (REMD) using implicit water model with random coil initial structure. (C, D) Results of stepwise clustering based on p53 TAD1 functional sites. The first round of local clustering was based on p53 functional site of residues 19-26, while the second round conducted for each cluster from the primary clustering results was based on the full-length TAD1. (E) Eight potential druggable pockets selected in consideration of functional site location, population percent and CAVITY druggability score.



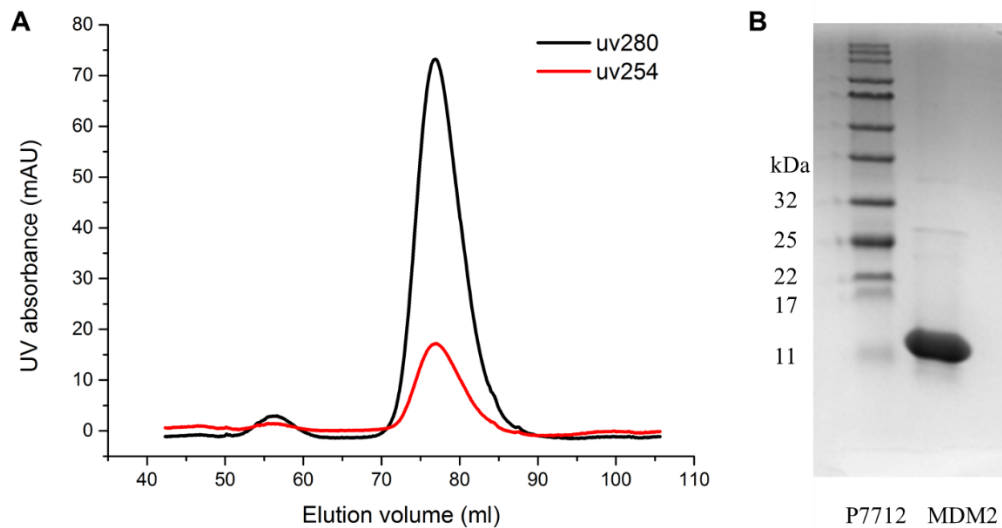
Supplementary Figure S2. Distribution of the docking scores (from the SP mode of Glide) of the eight representative conformations against the screening compound library. Top 10% docking scores of each conformation are presented to acquire better display.



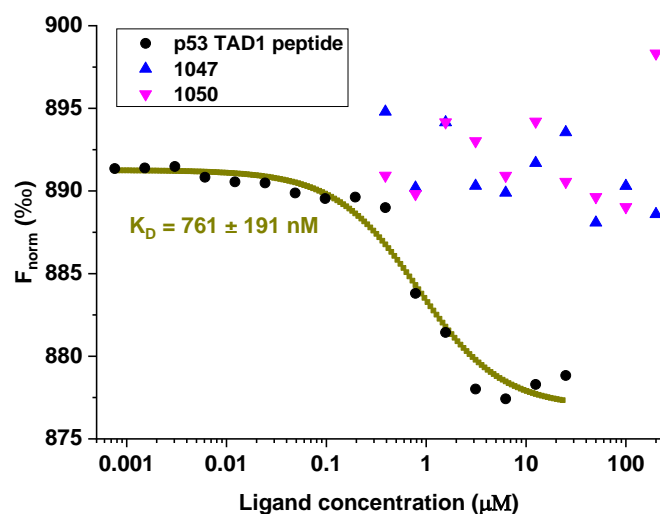
Supplementary Figure S3. Cell growth inhibition activity of 1047 (A) and 1050 (B) in different genetic backgrounds (MCF-7/p53 wt, HCT116/p53 wt, H1299/p53 null, MDA-MB-231/p53 mt).



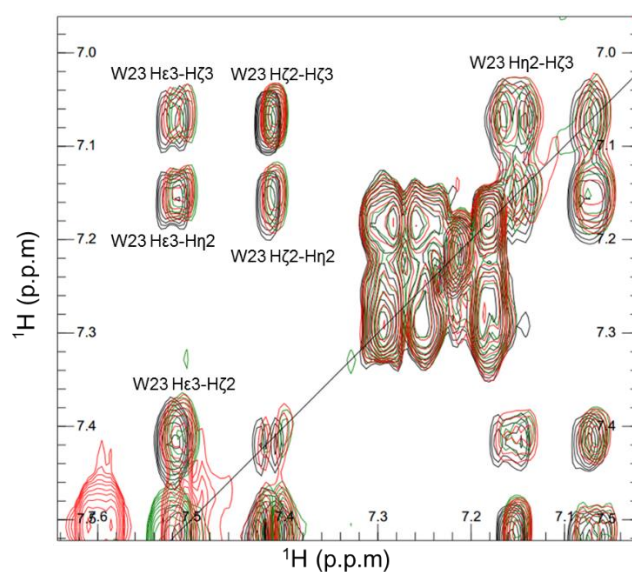
Supplementary Figure S4. Surface plasmon resonance (SPR) dose-response curves of binding between p53 TAD1 and 1047 (A) and 1050 (B), respectively.



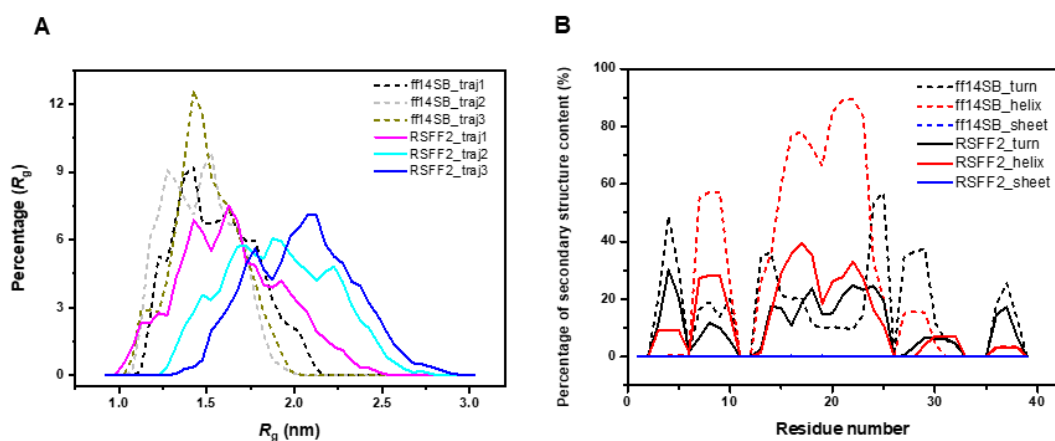
Supplementary Figure S5. Purification and characterization of MDM2. (A) Elution profile of the final Sephacryl S-200 gel-filtration column. MDM2 was eluted as a single peak. Absorbance at 254 nm is shown in red and at 280 nm is shown in black. **(B)** SDS-PAGE analysis of MDM2 and molecular mass markers run under reducing conditions.



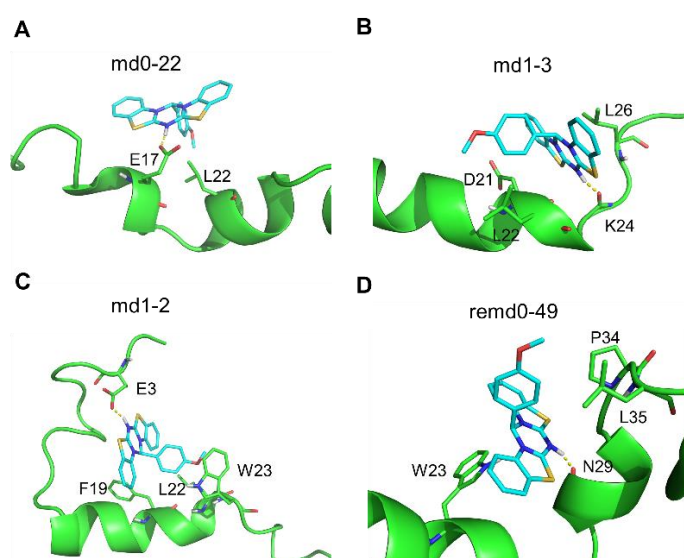
Supplementary Figure S6. The binding measurements between MDM2 and p53 TAD1 (black), 1047 (blue) and 1050 (magenta), respectively. In the MST experiment we have kept the concentration of NHS-labeled MDM2 constant, while the concentration of the non-labeled ligands was varied.



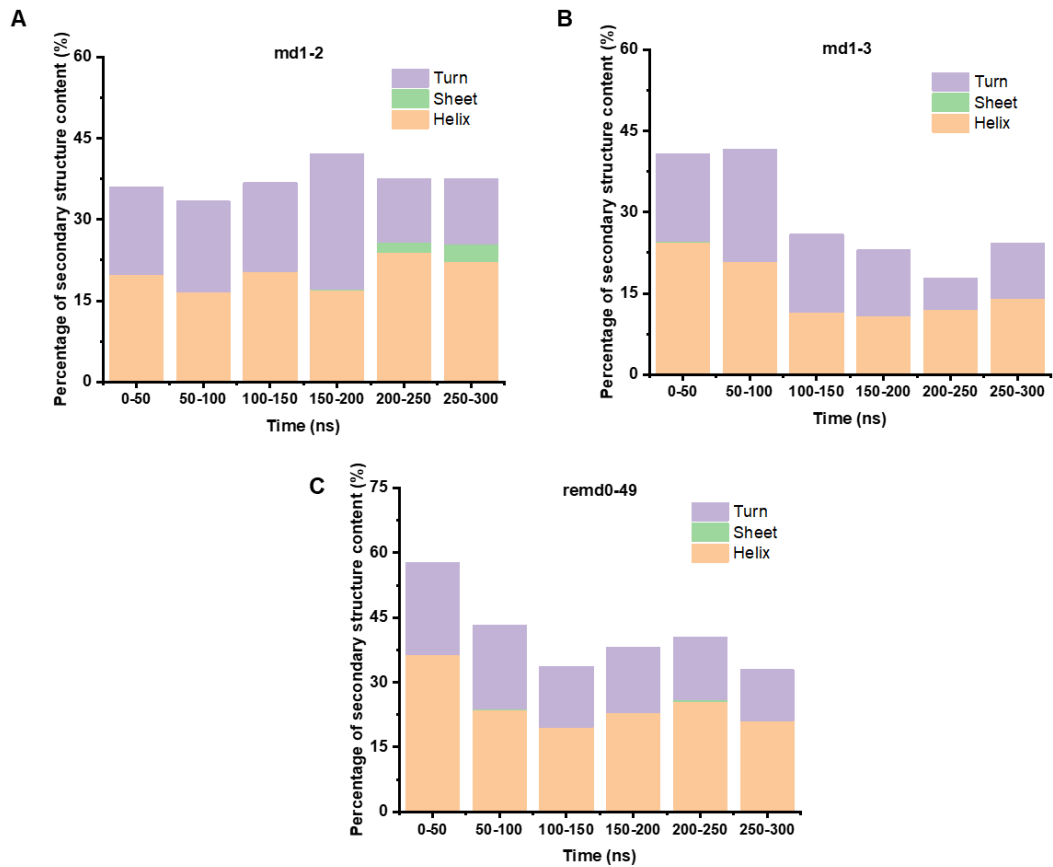
Supplementary Figure S7. Overlay TOCSY spectra of free shuffled p53 TAD1 peptide (black) and shuffled p53 TAD1 peptide-p53 TAD1 ligands complex (1047 red, 1050 green). The shuffled p53 TAD1 peptide sequence: MSPPSEDNWE SKSQPQELVP PVDSLQLESL PFTLNEPLA.



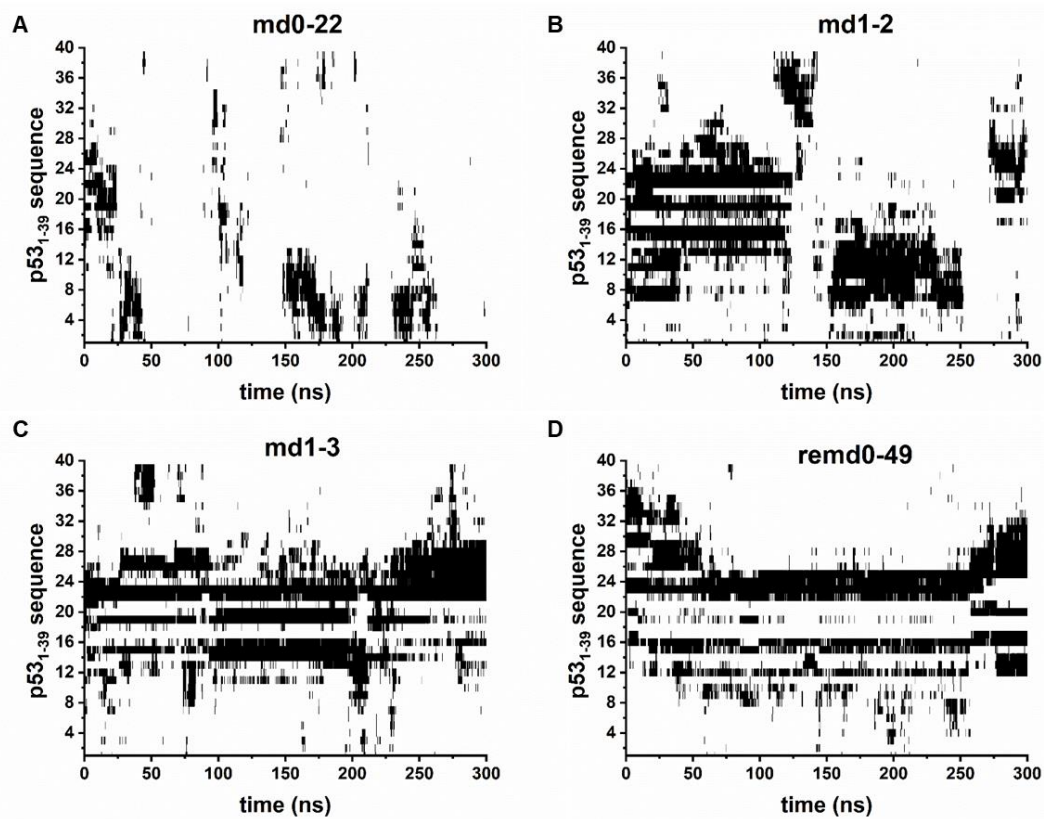
Supplementary Figure S8. FF14SB and RSFF2 force fields comparison. (A) Density estimate of the radius of gyration $p(R_g)$ according to MD simulations trajectories. **(B)** Secondary structure content of p53 TAD1 with different force fields.



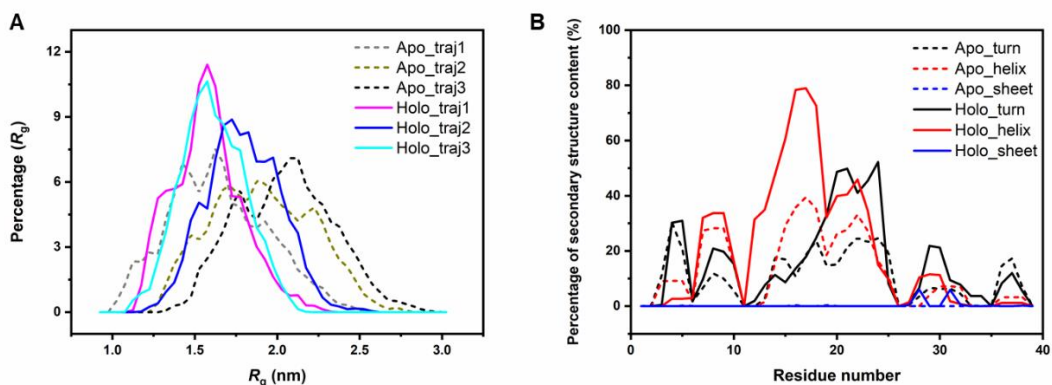
Supplementary Figure S9. Binding models of p53 TAD1 with 1047 generated by Glide. p53 TAD1 is represented in green and 1047 is represented in cyan. The conformation number is consistent with Figure 8S and Table S4.



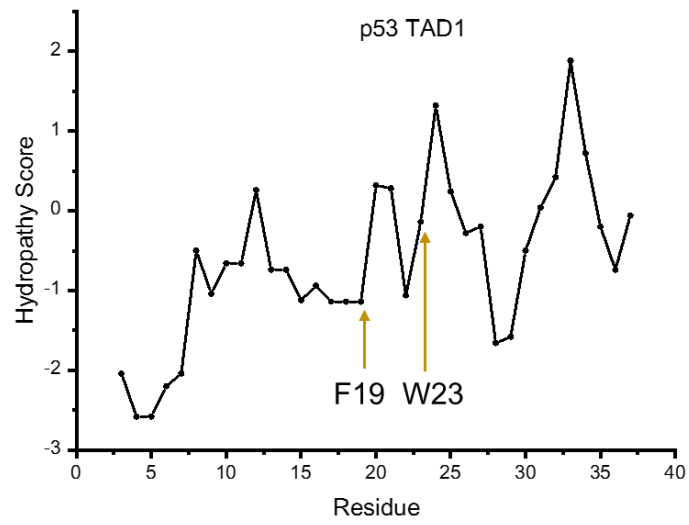
Supplementary Figure S10. Time-evolution of secondary structure content with a time interval of 50 ns starting from 0 ns from p53-TAD1 and 1047 complex simulations. In late stages of simulations, secondary structure content tends to be stable, indicating convergence of simulations.



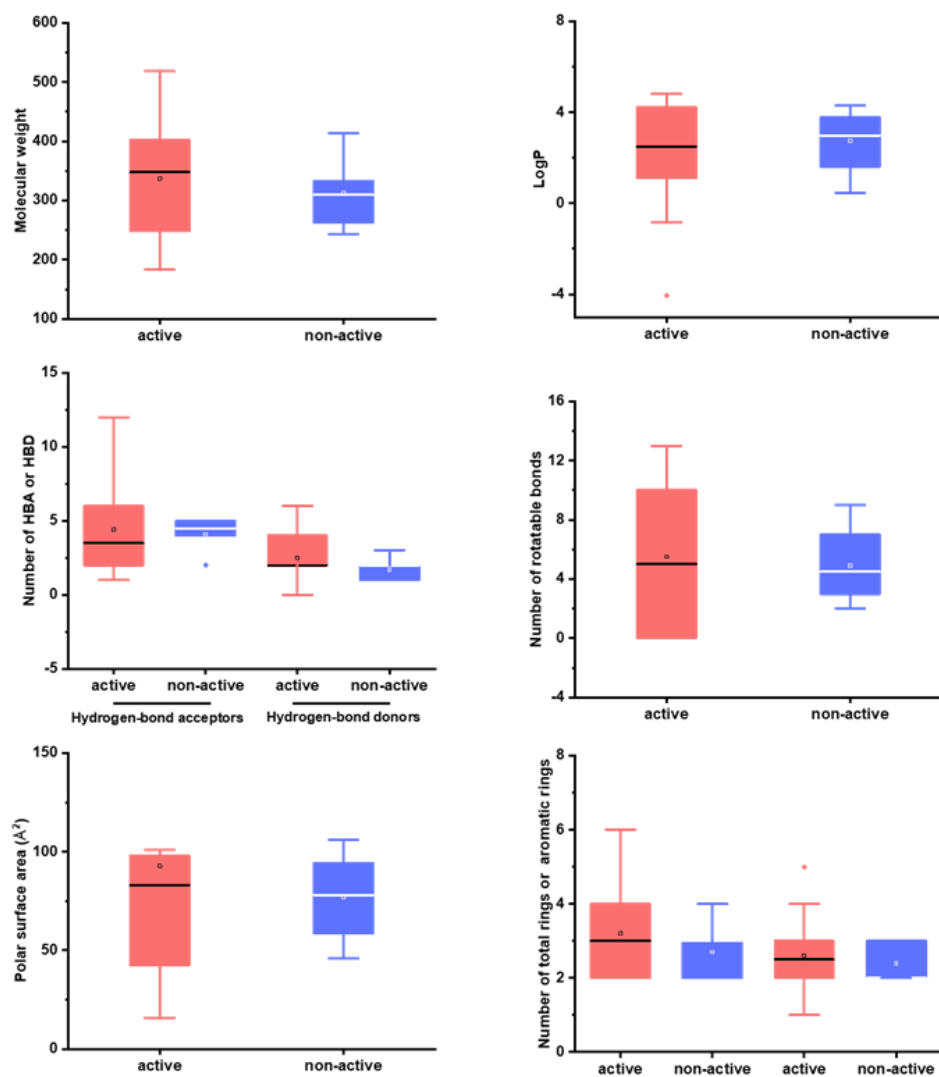
Supplementary Figure S11. Interactions between 1047 and p53 TAD1 in MD simulations. The four starting complex structures were built by molecular docking and therefore there were four 300-nanosecond simulations with different initial structures. Contact interactions were defined as the distance between hydrogen atoms in **1047** and different protein residues $< 4.5 \text{ \AA}$.



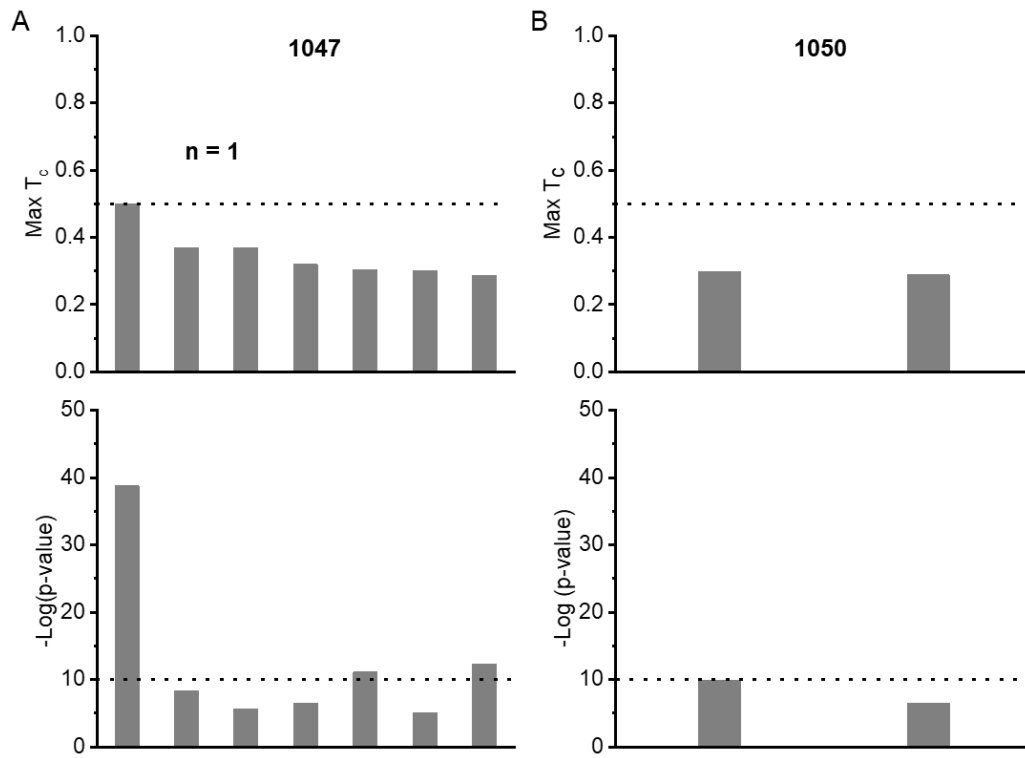
Supplementary Figure S12. MD simulations trajectories analysis of apo and holo p53 TAD1 ensemble. (A) Density estimate of the radius of gyration $p(R_g)$ with or without p53 TAD1 ligands according to MD simulation trajectories. (B) Secondary structure content of apo and holo p53 TAD1 ensemble upon binding to **1047**.



Supplementary Figure S13. Hydrophobicity analysis of p53 TAD1 according to the scale of Kyte & Doolittle with 5 residues window size.



Supplementary Figure 14. Physicochemical property distributions of the active compounds and the negative control compounds.



Supplementary Figure S15. SEA analysis for off-target probabilities of 1047 (A) and 1050 (B). Targets with both $\text{Max } T_c > 0.5$ and $p\text{-value} < 10^{-10}$ are considered as possible targets.

Supplementary Tables

Supplementary Table S1. Initial screening results by SPR and cell viability results by MTT using MCF-7 cell-line.

Compounds ID	Response/RU (50.0 μ M)	Response/RU (100 μ M)	IC ₅₀ (μ M, MCF-7)
AG-690/33033005	0.5	0.8	> 100
AN-943/13651024	0.5	0.8	> 75
AE-473/30516010	0.5	1.2	> 75
AF-399/42100324	0.9	1.4	> 100
AM-814/41092115	0.5	1.1	> 75
AM-814/41092143	1.0	1.4	> 75
AJ-644/41040360 (1047)	2.4	4.3	36.8 \pm 13.1
AO-476/40829160	0.5	1.1	> 100
AN-465/43384132	0.6	0.9	> 100
AN-465/43384049	0.6	1.4	> 100

Supplementary Table S2. Cell growth inhibition evaluation results for **1047** analogs.

Compounds ID	IC ₅₀ (μ M, HCT116, p53 wt)	IC ₅₀ (μ M, MCF-7, p53 wt)	IC ₅₀ (μ M, H1299, p53 null)
AM-944/40947971	6.8 \pm 0.6	3.0 \pm 0.3	6.3 \pm 3.1
AQ-405/42300135 (1050)	27.4 \pm 4.0	17.0 \pm 1.8	>50
AT-761/43494449	44.2 \pm 27.8	50.0 \pm 10.0	45.0 \pm 6.3
AT-761/43494450	21.9 \pm 2.4	27.0 \pm 6.6	36.0 \pm 4.6
AT-761/43502416	18.4 \pm 8.0	13.0 \pm 5.4	27.5 \pm 1.3

Supplementary Table S3. Assignments of TOCSY spectra. Only residue W23 was identified to show CSPs and average $\Delta F1$ CSPs were calculated for five cross peaks.

Sample	F1 Peak/ppm					Average $\Delta F1$ /ppm
	H ζ 2-H ζ 3	H ζ 2-H η 2	H ϵ 3-H ζ 2	H ϵ 3-H ζ 3	H ϵ 3-H η 2	
WT	7.412	7.412	7.495	7.495	7.496	–
WT + 1047	7.403	7.401	7.483	7.483	7.483	0.011
WT + 1050	7.400	7.399	7.478	7.478	7.478	0.015
Shuffle	7.417	7.416	7.516	7.520	7.520	–
Shuffle + 1047	7.412	7.411	7.513	7.514	7.513	0.005
Shuffle + 1050	7.409	7.408	7.510	7.510	7.505	0.008

Supplementary Table S4. Docking scores and ranking of the ten active compounds and the ten negative control compounds with the eight conformations generated by the SP mode of Glide. Compounds that are among the top 1% (shown in red) in at least three different conformations were considered as potential p53 TAD1 binding compounds for further experimental study. Ten compounds that were among the top 1% (shown in red) in only one conformation were chose as negative controls.

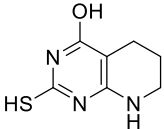
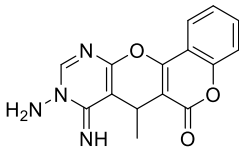
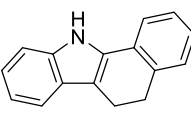
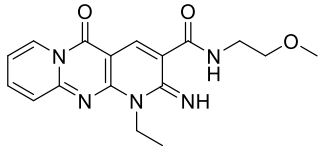
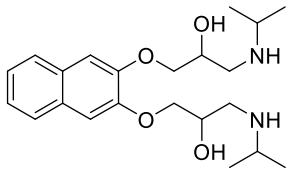
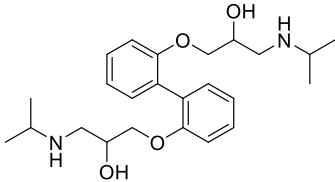
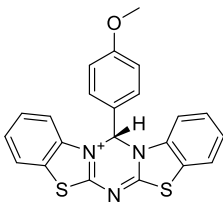
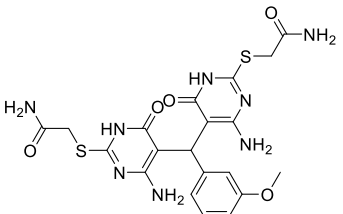
	SPECS ID	Conformation ID/Docking score (kcal/mol)/Ranking							
		md0-22	md1-2	md1-3	md2-5	md6-5	remd0-49	remd16-9	remd18-4
The ten active compounds	AG-690/33033005	-4.21	-4.38	-5.58	-6.48	-6.16	-4.90	-4.64	-4.29
		2.1%	5.6%	0.15%	1.3%	0.44%	5.5%	0.46%	4.5%
	AN-943/13651024	-4.98	-5.11	-4.50	-5.31	-6.15	-5.42	-4.44	-5.01
		0.08%	0.47%	5.9%	23%	0.45%	0.71%	1.1%	0.40%
	AE-473/30516010	-4.08	-3.86	-5.85	-6.84	-4.24	-5.55	-4.28	-4.43
		3.5%	23.7%	0.05%	0.35%	43%	0.39%	2.1%	2.9%
	AF-399/42100324	-3.69	-5.09	-4.29	-5.49	-5.42	-5.35	-3.86	-4.85
		8.5%	0.5%	10%	16%	49%	0.98%	9.5%	0.42%
	AM-814/41092115	-4.09	-6.08	-5.41	-5.31	-6.08	-4.69	-4.11	-4.31
		13.2%	0.01%	0.28%	23%	0.57%	11%	3.9%	2.6%
	AM-814/41092143	-4.96	-5.08	-5.18	-5.20	-6.50	-3.94	-4.69	-4.74
		0.08%	0.52%	0.69%	28%	0.12%	56%	0.35%	1.1%
	AJ-644/41040360	-4.73	-5.38	-6.30	-3.88	-1.07	-5.50	-4.30	-4.24
		0.25%	0.17%	0.01%	95%	99%	0.48%	1.9%	5.3%
	AO-476/40829160	-4.03	-5.82	0	-4.91	-4.99	-5.90	-4.23	-5.18
		4.0%	0.03%	99%	45%	13%	0.09%	2.5%	0.2%
	AN-465/43384132	-4.57	-5.35	-4.33	-5.04	-5.33	-4.02	-5.10	-4.55
		0.50%	0.19%	9.3%	37%	6.3%	51%	0.03%	2.0%
	AN-465/43384049	-4.58	-4.24	-1.89	-5.38	-5.11	-5.53	-4.49	-4.08
0.47%		8.7%	98%	21%	11%	0.41%	0.08%	8.5%	
Average docking score	-4.76	-5.42	-5.67	-6.84	-6.22	-5.57	-4.87	-5.01	
The ten negative control	AT-057/43315013	-5.51	-4.5	-4.07	-4.86	-5.81	-3.57	-3.14	-3.95
		0.007%	3.9%	18%	48%	1.4%	83%	54%	12%
	AK-968/12383032	-3.12	-5.87	0.00	-5.66	-4.46	0.00	-3.71	-3.18
		51%	0.02%	99%	12%	32%	99%	15%	56%
	AJ-292/41701750	-3.12	-3.70	-5.90	-5.53	-4.17	-3.36	-3.71	-3.17
		51%	33.3%	0.04%	15%	47%	92%	15%	57%
AC-907/34131002	-3.95	-3.71	-3.61	-7.60	-3.29	-4.80	-3.06	-2.71	
	5.4%	32%	26%	0.02%	89%	7.7%	61%	85%	

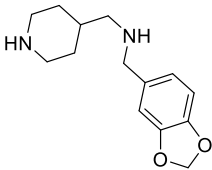
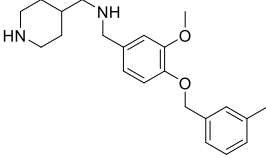
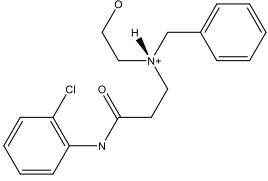
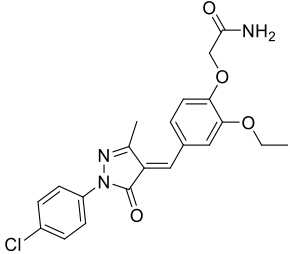
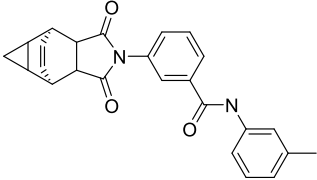
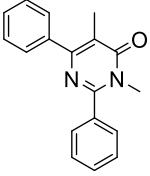
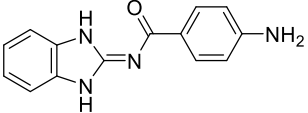
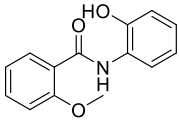
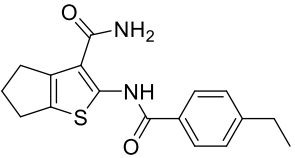
The ten negative control	AT-057/43469664	-4.10	-3.42	-4.83	-7.73	-4.52	-4.26	-2.99	-3.56
		3.1%	53%	2.2%	0.01%	29%	3.2%	66%	31%
	AF-399/33193056	-3.49	-3.99	-4.83	-4.96	-7.10	-5.10	-4.21	-4.30
		23%	17%	2.2%	42%	0.01%	2.7%	2.7%	4.3%
	AP-970/43481810	-3.37	-4.04	-3.45	-5.02	-7.05	-4.23	-3.57	-4.74
		31%	15%	54%	38%	0.01%	34%	22%	1.1%
	AK-968/11286189	-3.52	-2.12	-3.15	-5.62	-3.79	-6.41	-3.69	-3.67
		21%	98%	74%	12%	67%	0.01%	16%	25%
	AO-080/43378259	-3.22	0.00	0.00	-6.00	-4.01	-3.94	-5.24	-3.16
		43%	99%	99%	5%	55%	57%	0.02%	57%
AG-670/36669027	-3.94	-3.96	0.00	0.00	-4.56	0.00	0.00	-5.85	
	5.8%	18%	99%	99%	42%	99%	99%	0.01%	

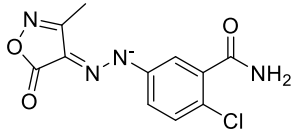
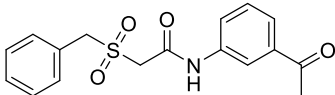
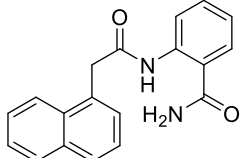
Supplementary Table S5. SPR response values of the ten negative control compounds.

Compounds ID	Response/RU (50.0 μ M)	Response/RU (100 μ M)
AT-057/43315013	0.3	0.9
AK-968/12383032	0.2	1.3
AJ-292/41701750	0.0	0.7
AC-907/34131002	0.2	1.5
AT-057/43469664	0.0	1.1
AF-399/33193056	0.3	0.6
AP-970/43481810	-0.1	0.6
AK-968/11286189	0.4	1.0
AO-080/43378259	-0.1	2.3
AG-670/36669027	0.1	1.0

Supplementary Table S6. Chemical structures of the ten active compounds and the ten negative control compounds.

Specs ID	Compound ID	Structure	
AG-690/33033005	A1		
AN-943/13651024	A2		
AE-473/30516010	A3		
AF-399/42100324	A4		
The ten active compounds	AM-814/41092115	A5	
	AM-814/41092143	A6	
	AJ-644/41040360	A7	
	AO-476/40829160	A8	

The ten active compounds	AN-465/43384132	A9	
	AN-465/43384049	A10	
The ten negative control	AT-057/43315013	N1	
	AK-968/12383032	N2	
	AJ-292/41701750	N3	
	AC-907/34131002	N4	
	AT-057/43469664	N5	
	AF-399/33193056	N6	
	AP-970/43481810	N7	

The ten negative control	AK-968/11286189	N8	
	AO-080/43378259	N9	
	AG-.2670/36669027	N10	

Supplementary Table S7. The Tanimoto similarity metric of the active and the negative control compounds (taking a value between 0 and 1, with 1 corresponding to identical structures).

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10
A1	1.00	0.04	0.03	0.05	0.07	0.16	0.04	0.05	0.14	0.18	0.05	0.02	0.03	0.02	0.07	0.06	0.05	0.04	0.02	0.03
A2	0.04	1.00	0.02	0.02	0.04	0.03	0.03	0.09	0.03	0.03	0.01	0.02	0.01	0.02	0.08	0.04	0.12	0.12	0.02	0.04
A3	0.03	0.02	1.00	0.02	0.02	0.05	0.03	0.02	0.09	0.04	0.01	0.02	0.00	0.01	0.02	0.02	0.03	0.01	0.01	0.02
A4	0.05	0.02	0.02	1.00	0.05	0.04	0.02	0.03	0.06	0.04	0.02	0.03	0.02	0.01	0.03	0.05	0.02	0.02	0.02	0.10
A5	0.07	0.04	0.02	0.05	1.00	0.07	0.06	0.05	0.05	0.07	0.05	0.02	0.01	0.02	0.06	0.06	0.03	0.03	0.02	0.03
A6	0.16	0.03	0.05	0.04	0.07	1.00	0.06	0.05	0.09	0.17	0.03	0.03	0.02	0.02	0.07	0.13	0.05	0.04	0.03	0.04
A7	0.04	0.03	0.03	0.02	0.06	0.06	1.00	0.04	0.05	0.07	0.02	0.02	0.03	0.01	0.02	0.02	0.03	0.03	0.01	0.02
A8	0.05	0.09	0.02	0.03	0.05	0.05	0.04	1.00	0.06	0.06	0.01	0.02	0.01	0.01	0.05	0.04	0.06	0.06	0.01	0.03
A9	0.14	0.03	0.09	0.06	0.05	0.09	0.05	0.06	1.00	0.11	0.02	0.03	0.01	0.01	0.05	0.04	0.05	0.03	0.01	0.04
A10	0.18	0.03	0.04	0.04	0.07	0.17	0.07	0.06	0.11	1.00	0.06	0.02	0.02	0.01	0.06	0.07	0.04	0.03	0.03	0.03
N1	0.05	0.01	0.01	0.02	0.05	0.03	0.02	0.01	0.02	0.06	1.00	0.01	0.01	0.01	0.03	0.03	0.02	0.02	0.03	0.01
N2	0.02	0.02	0.02	0.03	0.02	0.03	0.02	0.02	0.03	0.02	0.01	1.00	0.02	0.01	0.02	0.02	0.02	0.02	0.03	0.03
N3	0.03	0.01	0.00	0.02	0.01	0.02	0.03	0.01	0.01	0.02	0.01	0.02	1.00	0.00	0.02	0.02	0.02	0.02	0.01	0.02
N4	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.00	1.00	0.02	0.02	0.02	0.01	0.01	0.02
N5	0.07	0.08	0.02	0.03	0.06	0.07	0.02	0.05	0.05	0.06	0.03	0.02	0.02	0.02	1.00	0.24	0.12	0.13	0.03	0.04
N6	0.06	0.04	0.02	0.05	0.06	0.13	0.02	0.04	0.04	0.07	0.03	0.02	0.02	0.02	0.24	1.00	0.07	0.07	0.03	0.03
N7	0.05	0.12	0.03	0.02	0.03	0.05	0.03	0.06	0.05	0.04	0.02	0.02	0.02	0.02	0.12	0.07	1.00	0.46	0.02	0.05
N8	0.04	0.12	0.01	0.02	0.03	0.04	0.03	0.06	0.03	0.03	0.02	0.02	0.02	0.01	0.13	0.07	0.46	1.00	0.02	0.03
N9	0.02	0.02	0.01	0.02	0.02	0.03	0.01	0.01	0.01	0.03	0.03	0.03	0.01	0.01	0.03	0.03	0.02	0.02	1.00	0.04
N10	0.03	0.04	0.02	0.10	0.03	0.04	0.02	0.03	0.04	0.03	0.01	0.03	0.02	0.02	0.04	0.03	0.05	0.03	0.04	1.00