SUPPLEMENTARY INFORMATION

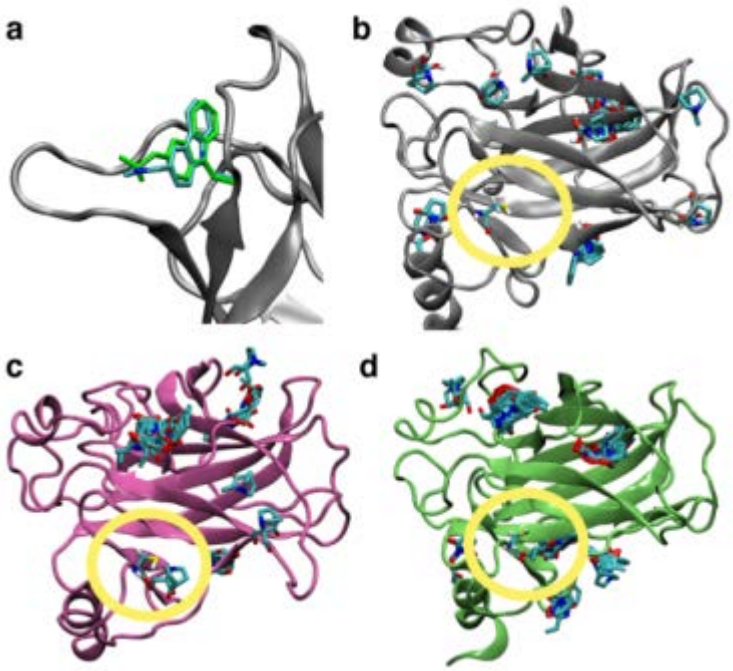
for:

Computational Identification of a Transiently Open L1/S3 Pocket for p53 Cancer Mutant Reactivation

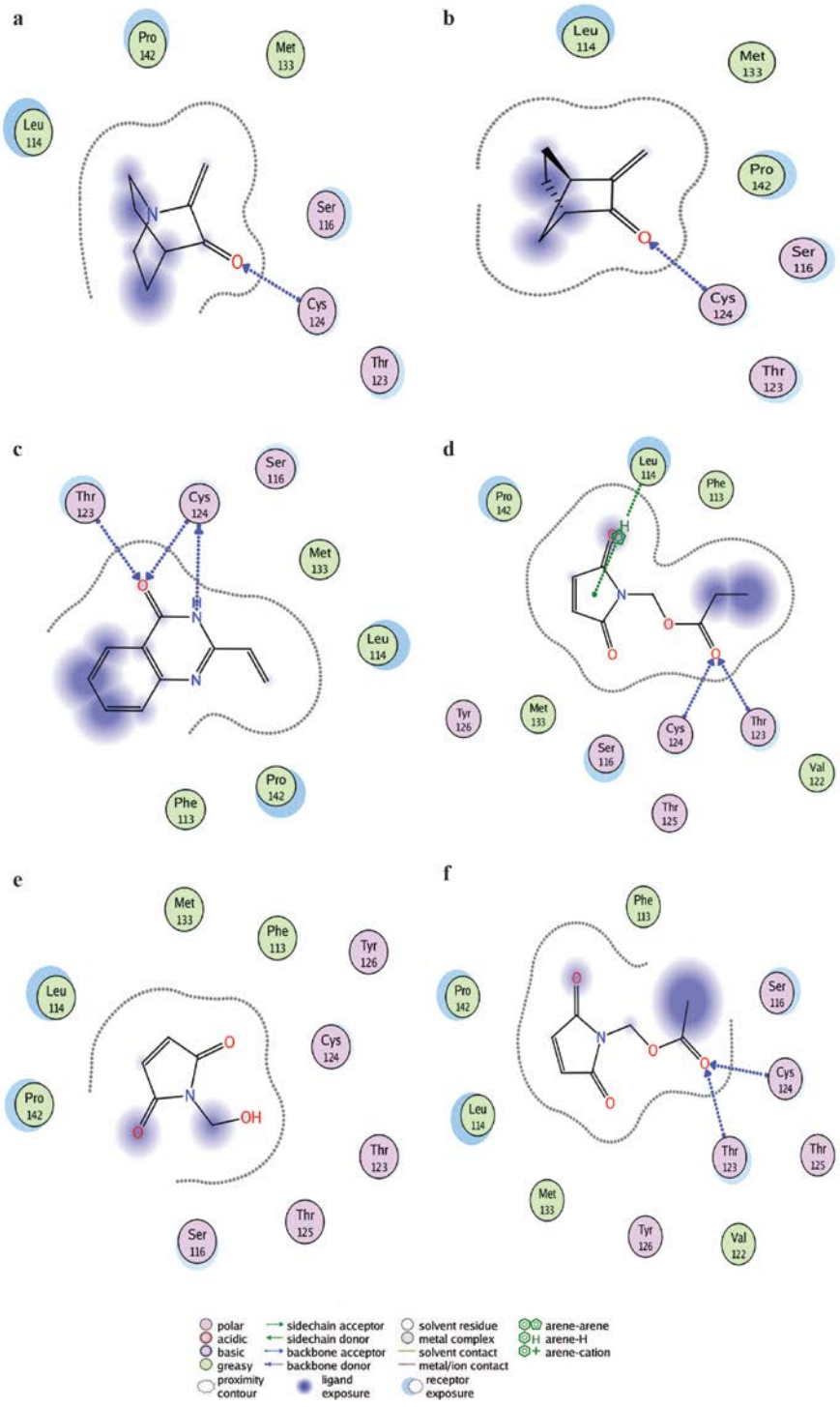
Christopher D. Wassman, Roberta Baronio, Özlem Demir, Brad D. Wallentine, Chiung-Kuang Chen, Linda V. Hall, Faezeh Salehi, Da-Wei Lin, Benjamin P. Chung, G. Wesley Hatfield, A. Richard Chamberlin, Hartmut Luecke, Richard H. Lathrop, Peter Kaiser, Rommie E. Amaro

SUPPLEMENTARY INFORMATION IN THIS DOCUMENT:

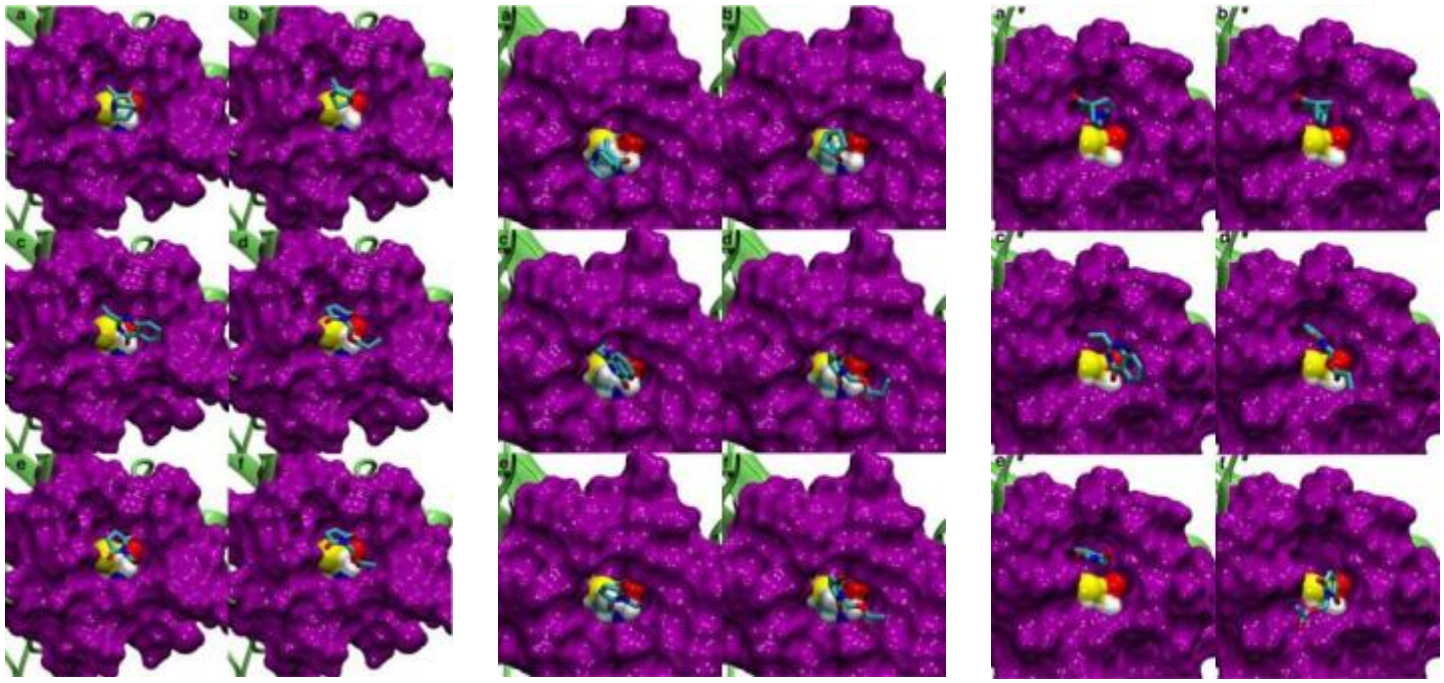
SUPPLEMENTARY FIGURES



Supplementary Figure S1. Positive control and blind docking of PRIMA-1 related compounds.
(a) Positive control using AutoDock Vina and PhiKan083 blind docking into the 2VUK crystal structure**. (b)** Blind docking of PRIMA-1 related compounds into the 1TSR-B crystal structure reveals no binding events near C124. **(c, d)** Blind docking of these compounds reveals favourable binding near C124 in the MD-extracted L1/S3 open pocket. **(c)** wild type. **(d)** R175H. For both (c) and (d) no known active compounds were found to dock favourably to any other cysteine except Cys124.

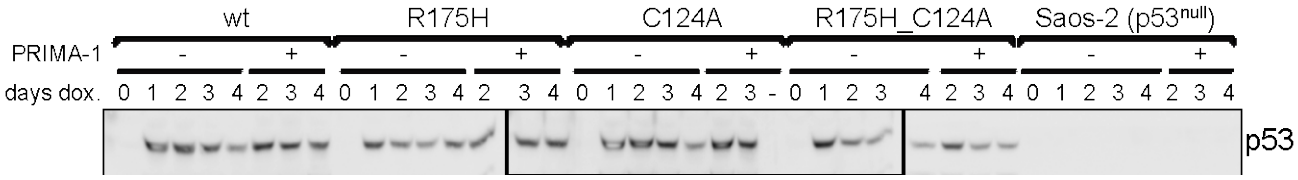


Supplementary Figure S2. Ligand interaction diagrams of the complexes in Fig. 2.
(a) MQ (from PRIMA-1) [26], **(b)** NB [21], **(c)** STIMA-1 [22], **(d)** MIRA-1 [19], **(e)** MIRA-2 [19], **(f)** MIRA-3 [19]. All compounds are docked with mutant R175H in the poses of Fig. 2.

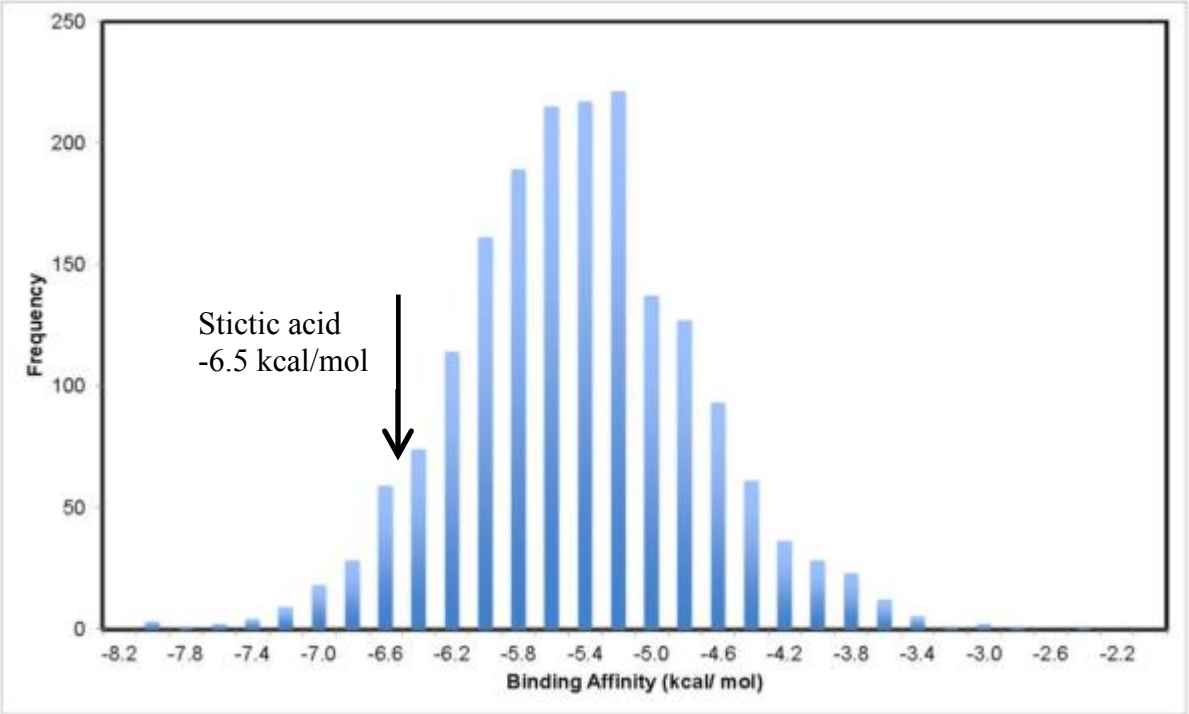


(A) **(B)** **(C)**

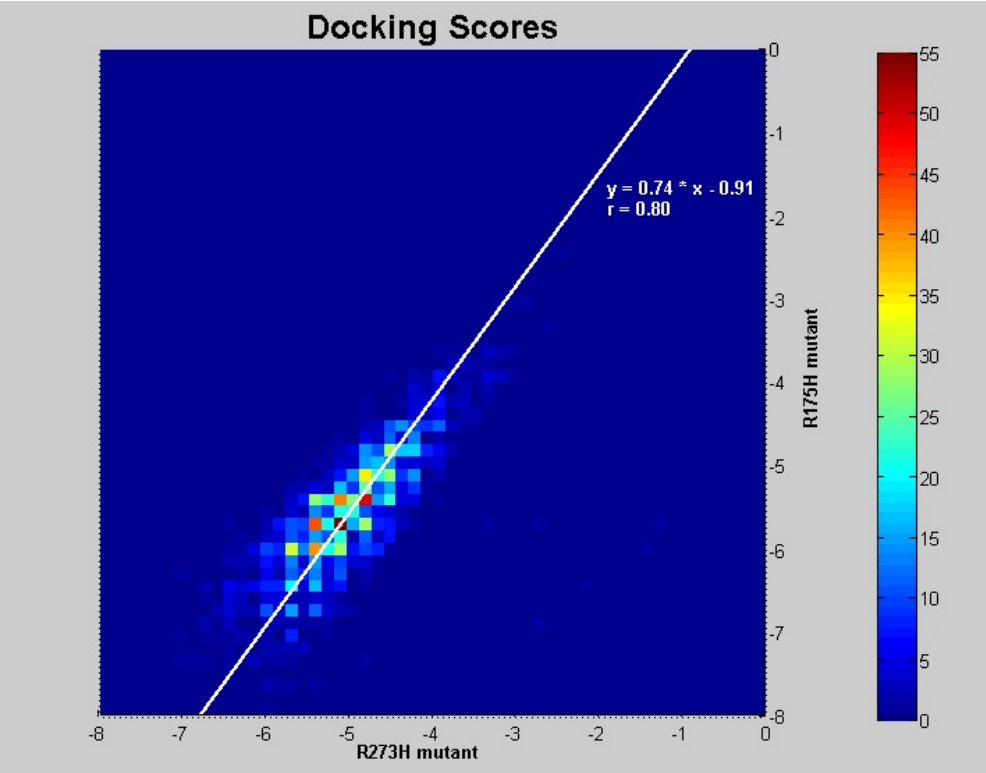
Supplementary Figure S3. Known p53 reactivation compounds docked into open L1/S3 pocket.
Variants **(A)** wt p53, **(B)** G245S, and **(C)** R273H across compounds **(a)** MQ (from PRIMA-1), **(b)** NB, **(c)** STIMA-1, **(d)** MIRA-1, **(e)** MIRA-2, and **(f)** MIRA-3. Known p53 reaction compounds could dock favourably into the most-populated MD trajectory cluster centroid that had an open L1/S3 pocket. In each case, the binding pose that is depicted had the smallest distance between the reactive methylene of the small molecule and the Cys124 sulphydryl group.



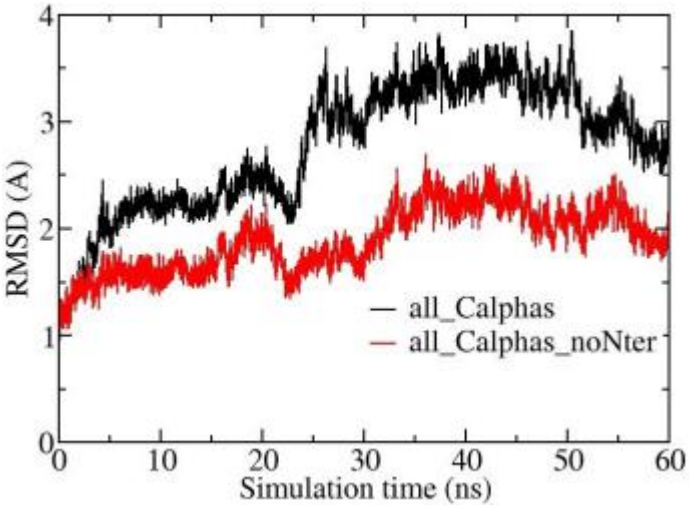
Supplementary Figure S4. Expression levels of wild type and mutant p53 in Saos-2 lines.
p53^{null} Saos-2 cells harbouring inducible p53 expression modules (wt, R175H, C124A, and R175H_C124A) were cultured with doxycycline (dox.) to express p53. Time point "0" shows repressed p53 before doxycycline induction. Pre-heated PRIMA-1 (50 μM final concentration) was added to cells as indicated 8 h after doxycycline. Equal amounts of total protein lysates were analysed by immunoblotting with anti-p53 antibodies. Samples were loaded on three different gels as indicated in the arranged panels.



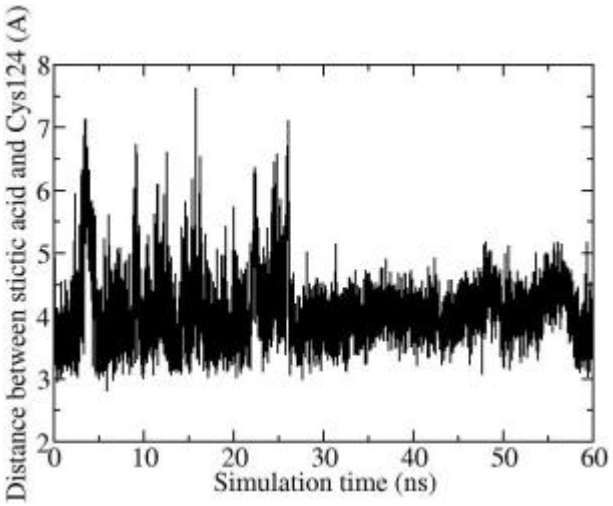
Supplementary Figure S5. Histogram of best docking scores for NCI/DTP compounds against R273H.
The R273H open L1/S3 pocket (cluster 13 centroid) was used as the docking target. The dataset of compounds (1,842 total) was a combination of the entire NCI Diversity Set II and an ADME-filtered subset of the NCI Mechanistic Set, the NCI Natural Products Set, the NCI Combo Set, and the NCI Oncology Approved Set. The arrow indicates the position of stictic acid (-6.5 kcal/mol). Methods are described in the text.



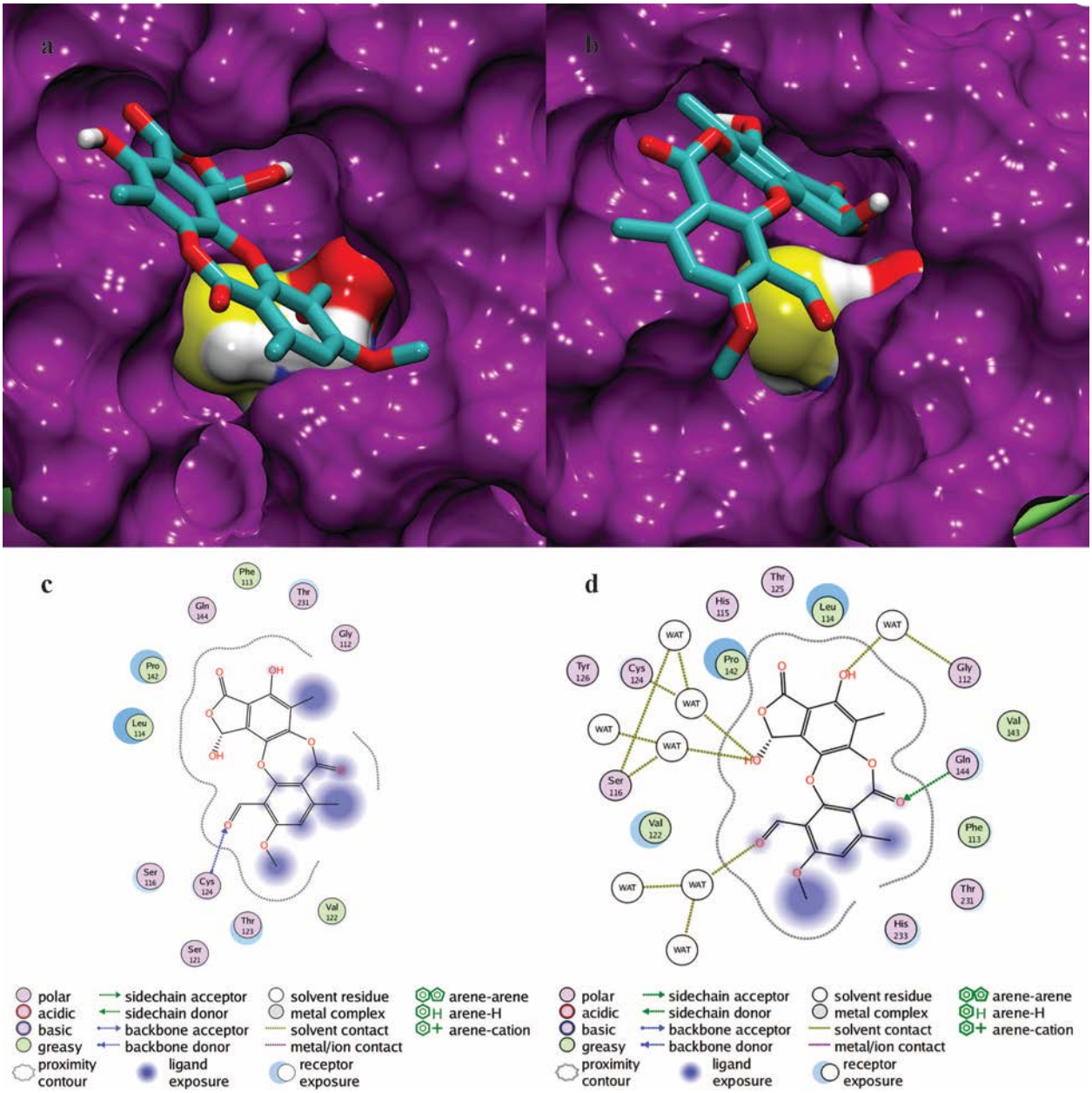
Supplementary Figure S6. Docking scores for NCI/DTP compounds against mutants R273H and R175H.
The graph displays Autodock Vina docking scores for the 1,842 NCI/DPT compounds shown in Supplementary Fig. S5. Each compound corresponds to a single point on the graph. Coordinates correspond to the docking scores against mutants R273H (x-axis) and R175H (y-axis). The mutant R273H conformation was the same as in Supplementary Fig. S5. The mutant R175H conformation was chosen as the R175H cluster centroid that had the smallest L1/S3 pocket RMSD relative to the R273H conformation. The correlation coefficient was 0.80.



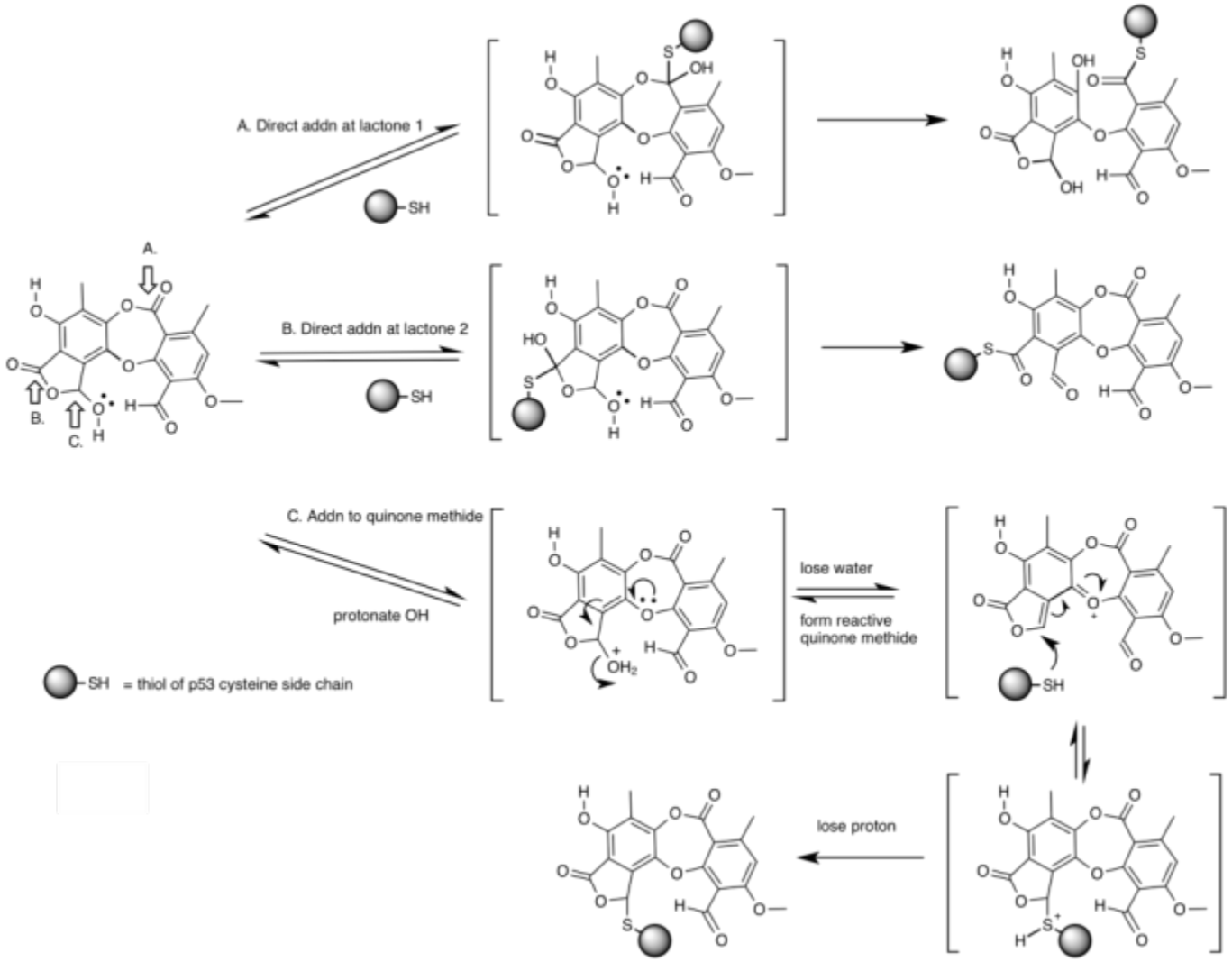
Supplementary Figure S7. Stability of the p53 core domain over 60 ns of MD simulation.
The graph corresponds to the complex of Supplementary Fig. S9. **(black; top trace)** RMSD in Angstroms of all C-alphas, after superimposing all frames with respect to all C-alphas. The N-terminal flexible loop moved substantially, while the rest of the protein structure remained mainly intact. **(red; bottom trace)** RMSD of all C-alphas except those of the 12 N-terminal residues (residues 96 to 108). Both metrics show that the core domain remained stable over the entire MD simulation.



Supplementary Figure S8. Persistence of stictic acid in the L1/S3 pocket over 60 ns of MD simulation.
The graph corresponds to the complex of Supplementary Fig. S9. It shows the distance in Angstroms between the sulphur atom of Cys124 and the oxygen atom of the 5-membered heterocyclic ring of stictic acid during the MD simulation. The graph shows that the complex remained stable over the entire MD simulation. The variance decreased after about 25 ns upon accommodating stictic acid deeper in the binding pocket.



Supplementary Figure S9. Stictic acid in the docked pose of Fig. 4(b) and after 60 ns stable MD simulation. **(a)** Stictic acid with mutant R175H in the initial docked pose of Fig. 4(b). The perspective is slightly modified here to provide better visibility for panel b. **(b)** Stictic acid after MD simulation for an additional 60 ns as described in the text. Over the MD simulation, stictic acid moved deeper into the pocket, achieving a very nice shape complementarity. **(c)** Ligand interaction diagram corresponding to (a). **(d)** Ligand interaction diagram corresponding to (b). The network of favourable interactions has expanded and the exposure pattern has shifted.



Supplementary Figure S10. Pathways by which stictic acid could form a covalent bond to a cysteine side chain. Direct attack of the thiol on either lactone carbonyl group (pathways A and B) would produce the respective thioester adducts shown via standard two step trans-acylations. Another more circuitous but also well-precedented pathway for covalent C-S bond is via addition of the thiol to a quinone methide intermediate formed (either before or after binding) by the elimination of water from the cyclic hemiacetal, as shown for pathway C.

SUPPLEMENTARY TABLES

Compound	Binding Score (kcal/mol)	z-score	Mean	Std.Dev.
MQ	-3.8	-1.71	-3.51	0.17
	-3.6	-0.53		
	-3.6	-0.53		
*	-3.6	-0.53		
	-3.6	-0.53		
	-3.4	0.65		
	-3.4	0.65		
	-3.3	1.24		
	-3.3	1.24		
NB	-4.1	-1.45	-3.81	0.20
	-4.0	-0.95		
	-4.0	-0.95		
	-3.9	-0.45		
	-3.8	0.05		
	-3.7	0.55		
*	-3.7	0.55		
	-3.6	1.05		
	-3.5	1.55		
STIMA1	-5.8	-1.64	-5.26	0.33
	-5.6	-1.03		
	-5.5	-0.73		
	-5.3	-0.12		
	-5.2	0.18		
	-5.2	0.18		
*	-5.0	0.79		
	-4.9	1.09		
	-4.8	1.39		
MIRA1 *	-4.8	-1.13	-4.62	0.16
	-4.8	-1.13		
	-4.8	-1.13		
	-4.6	0.13		
	-4.6	0.13		
	-4.6	0.13		
	-4.6	0.13		
	-4.5	0.75		
	-4.3	2.00		
MIRA2	-4.4	-1.78	-4.08	0.18
*	-4.2	-0.67		
	-4.2	-0.67		
	-4.1	-0.11		
	-4.1	-0.11		
	-4.0	0.44		
	-4.0	0.44		
	-3.9	1.00		
	-3.8	1.56		
MIRA3 *	-4.8	-1.21	-4.57	0.19
	-4.8	-1.21		
	-4.7	-0.68		

	-4.6	-0.16		
	-4.6	-0.16		
	-4.5	0.37		
	-4.5	0.37		
	-4.3	1.42		
	-4.3	1.42		
Stictic acid *	-7.0	-1.40	-6.41	0.42
	-6.8	-0.93		
	-6.7	-0.69		
	-6.6	-0.45		
	-6.4	0.02		
	-6.4	0.02		
	-6.1	0.74		
	-6.0	0.98		
	-5.0	1.69		

Supplementary Table S1. Statistics for the docking of known reactivation compounds.
The predicted binding affinities (kcal/mol), z-scores, mean, and standard deviation (std. dev.) are shown for the nine docked poses of each compound shown in Fig. 2 and Fig 4(b) against mutant R175H in an open L1/S3 pocket conformation; see Methods section for details. An asterisk indicates the binding pose chosen for depiction, based on the smallest distance between the reactive methylene and the Cys124 sulphydryl group.

NSC Number	Predicted binding affinity
NSC-80313	-6.42
NSC-84100_b	-6.41
NSC-84100_a	-6.40
NSC-121868_b	-6.18
NSC-121868_a	-6.16
NSC-332670	-6.16
NSC-37641_a	-6.13
NSC-37641_b	-6.13
NSC-332186	-6.00
NSC-24951_b	-5.92
NSC-90737	-5.89
NSC-24951_a	-5.88
NSC-359472	-5.84
NSC-102742	-5.81
NSC-94600_b	-5.79
NSC-45545	-5.78
NSC-86467_a	-5.77
NSC-94600_a	-5.77
NSC-371765	-5.75
NSC-86467_b	-5.70
NSC-13316_a	-5.66
NSC-87838	-5.65
NSC-117197_b	-5.62
NSC-117197_a	-5.61
NSC-13316_b	-5.57
NSC-69359_a	-5.55
NSC-211340	-5.53
NSC-69359_b	-5.53
NSC-95909	-5.51
NSC-10211	-5.46
NSC-659434	-5.28
NSC-34875_a	-5.26
NSC-34875_b	-5.26
NSC-146770	-5.11

Supplementary Table S2. Selected NCI Diversity Set II compounds and predicted binding affinities.
The predicted binding affinities (kcal/mol) for each of the selected NCI Diversity Set II compounds were calculated using a population-weighted average of binding affinities to the most-populated 30 cluster centroids of the R273H MD trajectory. Suffixes “a” and “b” correspond to enantiomers of the same compound.

NSC Number	Predicted binding affinity
NSC-640974	-6.15
NSC-643148	-6.08
NSC-641228	-6.00
NSC-634568	-5.99
NSC-658494	-5.99
NSC-40341	-5.94
NSC-401005	-5.91
NSC-92339*	-5.89
NSC-641240	-5.86
NSC-641253	-5.81
NSC-32192	-5.77
NSC-642649	-5.72
NSC-623051	-5.66
NSC-106408	-5.64
NSC-668270	-5.60
NSC-87511	-5.60
NSC-204985	-5.59
NSC-132791	-5.55
NSC-653010*	-5.54
NSC-635438	-5.46
NSC-653016*	-5.43
NSC-407010	-5.42
NSC-635326	-5.32
NSC-266535	-5.26

Supplementary Table S3. Selected compounds from four NCI databases and predicted binding affinities.
The binding affinities (kcal/mol) for each of the selected compounds from 4 NCI databases were calculated using a population-weighted average of the binding affinities of the most-populated 30 cluster centroids of R273H MD trajectory. An asterisk marks compounds that were not assayed because not available.

Compound	SMILES String
Stictic Acid (NSC-87511)	C3=C(C(=C2OC1=C4C(=C(O)C(=C1OC(=O)C2=C3C)C)C(=O)OC4O)C=O)OC
PRIMA-1	OCC1(CO)N2CCC(CC2)C1=O
MQ	C[C@H]1C(=O)C2CCN1CC2
MIRA-1	C(CC)(=O)OCN1C(C=CC1=O)=O
MIRA-2	C1CC(=O)N(C1=O)CO
MIRA-3	C1CC(=O)N(C1=O)COC(=O)C
NB	C[C@H]1C(=O)[C@H]2CC[C@@H]1C2
NSC-319725	CN(C)C(=S)NN=C(C)C1=CC=CC=N1
STIMA-1	C(=C)C2=NC1=CC=CC=C1C(=N2)O

Supplementary Table S4. SMILES strings for all small molecules considered explicitly in this paper.

SUPPLEMENTARY METHODS

Criteria that distinguish C124 open from closed L1/S3 binding pockets.

The criteria consist of (1) an amino acid composition criterion and (2) a distance geometry criterion.

(1) The amino acid composition criterion is that the binding pocket is lined by residues Val122, Thr123, Pro142, Leu114, Ser121, Cys124, Thr231, Cys141, Ser116, Gln144, and Phe113.

(2) The distance geometry criterion consists of five constraints. The five constraints, as stated in words, are that:

1. Either the distance between Pro142's beta carbon and Leu114's delta carbon 1 atom (Leu114.CD1) or the distance between Pro142's beta carbon and Leu114's delta carbon 2 atom (Leu114.CD2) should be greater than 5.6 Å.
2. Both the distance between Pro142's beta carbon and Leu114's delta carbon 1 atom (Leu114.CD1) and the distance between Pro142's beta carbon and Leu114's delta carbon 2 atom (Leu114.CD2) should be greater than 4.9 Å.
3. Either the distance between Cys124's alpha carbon and Leu114's delta carbon 1 atom (Leu114.CD1) or the distance between Cys124's alpha carbon and Leu114's delta carbon 2 atom (Leu114.CD2) should be greater than 10.8 Å.
4. Both the distance between Cys124's alpha carbon and Leu114's delta carbon 1 atom (Leu114.CD1) and the distance between Cys124's alpha carbon and Leu114's delta carbon 2 atom (Leu114.CD2) should be greater than 9.0 Å.
5. The dihedral angle of Leu114's N atom, Leu114's alpha C, Leu114's beta C atom, and Leu114's gamma C atom should either be less than 40 degrees, or larger than 80 degrees.

The same five distance geometry constraints, as stated in an equivalent mathematical expression, are that:

[(distance (Pro142.CB, Leu114.CD1) > 5.6 Å) or (distance (Pro142.CB, Leu114.CD2) > 5.6 Å)] and
[(distance (Pro142.CB, Leu114.CD1) > 4.9 Å) and (distance (Pro142.CB, Leu114.CD2) > 4.9 Å)] and
[(distance (Cys124.CA, Leu114.CD1) > 10.8 Å) or (distance (Cys124.CA, Leu114.CD2) > 10.8 Å)] and
[(distance (Cys124.CA, Leu114.CD1) > 9.0 Å) and (distance (Cys124.CA, Leu114.CDs) > 9.0 Å)] and
[(dihedral angle (Leu114.N, Leu114.CA, Leu114.CB, Leu114.CG) > 80 degrees) or
(dihedral angle (Leu114.N, Leu114.CA, Leu114.CB, Leu114.CG) < 40 degrees)].