Computational identification of a transiently open L1/S3 pocket for reactivation of mutant p53

Christopher D. Wassman1,2,*, Roberta Baronio2,3,*, Özlem Demir4,*,†, Brad D. Wallentine5, Chiung-Kuang Chen5, Linda V. Hall2,3, Faezeh Salehi1,2, Da-Wei Lin3, Benjamin P. Chung3, G. Wesley Hatfield2,6,7, A. Richard Chamberlin4,8,9, Hartmut Luecke2,5,9,10,11, Richard H. Lathrop1,2,9,12, Peter Kaiser2,3,9 & Rommie E. Amaro1,4,8,†

The tumour suppressor p53 is the most frequently mutated gene in human cancer. Reactivation of mutant p53 by small molecules is an exciting potential cancer therapy. Although several compounds restore wild-type function to mutant p53, their binding sites and mechanisms of action are elusive. Here computational methods identify a transiently open binding pocket between loop L1 and sheet S3 of the p53 core domain. Mutation of residue Cys124, located at the centre of the pocket, abolishes p53 reactivation of mutant R175H by PRIMA-1, a known reactivation compound. Ensemble-based virtual screening against this newly revealed pocket selects stictic acid as a potential p53 reactivation compound. In human osteosarcoma cells, stictic acid exhibits dose-dependent reactivation of p21 expression for mutant R175H more strongly than does PRIMA-1. These results indicate the L1/S3 pocket as a target for pharmaceutical reactivation of p53 mutants.
Tumour suppressor protein p53 is a transcription factor that has an important role in human cancers. It responds to various cell stress conditions, such as DNA damage, and controls pathways leading to apoptosis (programmed cell death), cellular senescence and cell cycle arrest, among others. Tumour initiation and maintenance depend upon inactivation of p53 pathways, which otherwise would deter uncontrolled cell growth. Consequently, p53 is the most frequently mutated gene in human cancers. Nearly half of all human tumours have mutant p53. Approximately three-quarters of those tumours that have mutant p53 are found to express full-length p53 protein with a single-residue missense mutation in the p53 DNA-binding core domain.

The resulting presence of full-length and single-point-mutated p53 protein in about one-third of all human tumours holds the promise of drugs that reactivate mutant p53 function, deter cancer cell proliferation, and shrink or kill the tumour. Several p53-controlled pathways inhibit cell proliferation, and so are valid pharmaceutical targets. Studies on transgenic mice demonstrate that p53 reactivation enables tumour regression in vivo, even in advanced tumour cases. Reactivation of p53 function has been demonstrated in various p53 cancer mutants by both second-site suppressor mutations and a handful of small molecules, although the precise reactivation mechanism remains elusive.

Here we describe computational analysis of p53 structural models that suggests a transiently open L1/S3 binding pocket around Cys124, Cys135 and Cys141 as a possible target for reactivation of p53 mutants by small molecules. The pocket was identified through genetic mutations, all-atom molecular dynamics (MD) simulations of the p53 core domain, computational solvent mapping of the resulting p53 structural ensemble and ensemble-based virtual screening. Mutation of Cys124 to alanine abolished known R175H reactivation effects of PRIMA-1 in human Saos-2 cancer cells. Ensemble-based virtual screening was used to select 45 compounds with the most favourable predicting binding affinity to the L1/S3 pocket. Of these, stictic acid emerged as a p53 reactivation compound in cultured human osteosarcoma cells, with the ability to stabilize p53.

Results

Computational identification of the L1/S3 binding pocket.

Initial docking of known p53 reactivation compounds (MQ, 3-methylene-2-norbornanone (NB), STIMA-1, MIRA-1, 2, 3 (refs 17–22)) indicated preferential binding to a specific region in the mouse p53 core domain crystal structure (PDB ID 3EXJ), which has high sequence identity (~90%) to human p53. The corresponding region in the human p53 core domain is flanked by loop L1 and sheet S3 (hence, the L1/S3 pocket). The region had been labelled by genetic mutation studies as highly enriched for p53 reactivation mutations (Fig. 1; Phe113, Leu114, His115, Thr123, Leu137 and Asn235), indicating that subtle structural changes in this region can reactivate p53 cancer mutants. Four of those reactivation mutation sites (Phe113, Leu114, His115 and Thr123) plus Cys124 exhibit significant or moderate nuclear magnetic resonance (NMR) chemical shifts in response to binding of the p53-stabilizing peptide CDB3 (ref. 31). Cys124, Cys135 and Cys141 are near the centre of this putative p53 reactivation region, and covalent alkylation of a cysteine thiol group had been suggested for p53 reactivation by PRIMA-1 conversion products such as MQ and other p53 reactivation compounds. In this cysteine triad, Cys124 was the most solvent-accessible, followed by Cys135 and then Cys141, in both the crystal structures and the MD simulations (Table 1). Initial alkylation of Cys124 had been suggested to enable alkylation of Cys141 and then Cys135 (ref. 32).

In the crystal structures of the human p53 core DNA-binding domain (PDB IDs 1TSR-B, 2OCJ), Cys124 is partially occluded and the putative L1/S3 reactivation pocket is not readily accessible to small molecules much larger than water. This occlusion in the crystal structure might seem to preclude alkylation of Cys124 in human p53. Blind docking of known p53 reactivation compounds onto the entire protein surface of the human p53 core domain crystal structure PDB ID 1TSR chain B (1TSR-B) indicated that none of the compounds bound in close proximity to any cysteine, including those in the L1/S3 pocket (Supplementary Fig. S1). However, the L1 loop can undergo a conformational switch to regulate DNA binding by an induced fit mechanism. NMR studies demonstrate that Cys124 is exposed to solvent ~5% of the time. Mass spectrometry results indicate that Cys124 and Cys141 react preferentially with alkylating agents and with glutathione. Alkylation agents cause NMR chemical shifts of several residues within the L1/S3 region, including Leu114, His115 and Thr123, already implicated above by genetic mutations and CDB3-induced NMR chemical shifts.
Wild-type p53 is marginally stable at physiological temperature. The DNA binding domain unfolds with a half-life of around 9 min and has a low melting temperature ($T_m$) of about 36–46 °C depending on experimental conditions. All-atom explicit solvent MD simulations for 30 ns of wild-type p53 and three frequent cancer mutants, R175H, G245S and R273H, revealed a high degree of flexibility for the core domain and particularly the L1/S3 region surrounding Cys124 (Supplementary Movie 1). Clustering the MD ensembles with respect to the L1/S3 pocket conformation further demonstrated the flexibility of the region (Table 1). In the MD ensembles of different p53 variants, average solvent-accessible surface area (SASA) values of the Cys124 side chain were calculated to be in the range of 6–11 Å² (Table 1). In sharp contrast, in the p53 core domain crystal structure 1TSR-B the Cys124 side chain SASA was calculated to be only 0.43 Å².

Docking into the L1/S3 pocket. A transiently open L1/S3 pocket formed around Cys124 due to positional and conformational changes of Ser116, Thr123, Cys124, Thr140 and especially Leu114 (Fig. 1). Three of these residues, Leu114, Thr123 and Cys124, already were implicated above. We determined five distance/dihedral angle criteria (Supplementary Methods) that indicated formation of an open L1/S3 pocket in about 5–8% of each simulation (Table 1). Fpocket, a program to measure the volume of a defined binding pocket (Fig. 1), also identified an open L1/S3 pocket around Cys124 across several p53 variants in 5–30% of the 3,000 frames of each MD trajectory (Table 1). For evaluating Fpocket results, the L1/S3 pocket was considered open if its volume was >150 Å³, a lower bound chosen based on cancer mutant R273H, which exhibited the lowest average L1/S3 pocket volumes. FTMap, a program to identify ligand-binding hot spots in proteins using organic solvent probes, predicted the L1/S3 pocket to be among the most highly occupied by the solvent probes when used to assess druggability of MD-generated conformations (Fig. 1). For wt p53 and mutant R273H MD trajectories, FTMap predicted a druggable L1/S3 pocket to occur, respectively, in 11 and 7 of the 15 most-populated cluster centroids. In contrast, neither Cys124 nor any other cysteine in the original PDB crystal structure 1TSR-B was identified as a druggable site by FTMap.

Several previously identified alkylating p53 reaction compounds, MQ, NB, STIMA-1 (ref. 22) and MIRA-1, 4, 5 (ref. 19), subsequently were docked into the most-populated MD trajectory cluster representative that had an open L1/S3 binding pocket as defined (Table 1, Supplementary Methods). All of these known active compounds presented low-energy predicted docking poses with the reactive methylene oriented in close proximity to the Cys124 sulphydryl group for p53 cancer mutant R175H (Fig. 2). Supplementary Fig. S2 depicts the corresponding ligand interaction diagrams, and Supplementary Table S1 provides the corresponding docking statistics. Reliability of the molecular docking strategy was supported by a positive

### Table 1 | Summary of MD simulation trajectories and open pocket occurrences.

|                  | wt p53 | R175H | G245S | R273H | C124A |
|------------------|--------|-------|-------|-------|-------|
| Total number of clusters in trajectory | 28     | 39    | 74    | 41    | 50    |
| Open pocket as % of total trajectory by geometric criteria | 6.3    | 7.3   | 6.7   | 5.0   | 5.4   |
| Open pocket as % of total trajectory by Fpocket | 27.3   | 29.9  | 22.9  | 7.3   | 35.3  |
| Avg. SASA Å² ± s.d. of Cys124 side chain in MD | 10.93 ± 4.81 | 9.58 ± 4.45 | 6.94 ± 4.71 | 8.38 ± 5.84 | (124Ala value) 4.45 ± 2.99 |
| Cys124 exposure as % of MD frames with SASA > 5 Å² | 90.80 | 86.97 | 60.70 | 64.97 | (124Ala value) 34.20 |
| Avg. SASA Å² ± s.d. of Cys141 side chain in MD | 0.51 ± 0.84 | 1.03 ± 1.27 | 0.56 ± 0.86 | 0.78 ± 1.32 | 1.56 ± 1.54 |
| Cys141 exposure as % of MD frames with SASA > 5 Å² | 0.33   | 1.53  | 0.43  | 2.06  | 3.00  |
| Avg. SASA Å² ± s.d. of Cys135 side chain in MD | 0.17 ± 0.48 | 0.41 ± 0.91 | 0.09 ± 0.27 | 0.09 ± 0.34 | 0.20 ± 0.44 |
| Cys135 exposure as % of MD frames with SASA > 5 Å² | 0.06   | 0.50  | 0.00  | 0.00  | 0.03  |

MD, molecular dynamics; SASA, solvent-accessible surface area.
Biological validation of the L1/S3 binding pocket hypothesis. We examined the role of Cys124 in the L1/S3 binding pocket by site-directed mutagenesis of Cys124 to alanine (C124A). If Cys124 were a critical residue for p53 reactivation, then removal of its thiol group by this otherwise innocuous mutation would abolish the PRIMA-1 reactivation effect. For example, Cys124 might be a key attachment residue for PRIMA-1 conversion products such as MQ, it might initiate a Cys124-Cys141 alkylation cascade, or it might form a disulfide bond that stabilized p53 in an active conformation induced by compound binding. The literature suggests that a reducing environment leads to correct p53 folding and disulfide bond formation may lead to aggregation. There is no a priori reason to rule out any of these hypotheses.

p53 null human osteosarcoma Saos-2 cells were engineered to have doxycycline-inducible expression of either wild-type p53 or various p53 mutants. Single clones with similar expression levels of the different p53 variants were selected for direct comparison of small-molecule effects. As expected, expression of wild-type p53 in Saos-2 cells halted cell proliferation, whereas expression of mutant R175H had no effect on cell proliferation (Fig. 3). Mutation of Cys124 to alanine had little effect on otherwise wild-type p53 (C124A) nor did it restore activity of R175H (R175H_C124A), indicating that the C124A mutation is largely neutral for general p53 function. Consistent with previous results, addition of PRIMA-1 to cells expressing inactive mutant R175H led to p53 reactivation and subsequent inhibition of cell cycle progression. Importantly, PRIMA-1 reactivation of p53 function was blocked in the double mutant R175H_C124A (Fig. 3), demonstrating the importance of Cys124 in pharmacological reactivation of mutant R175H by PRIMA-1. These experimental results supported the computational identification of the L1/S3 binding pocket surrounding Cys124 as significant for p53 reactivation.
Experimental support for stictic acid and p53 reactivation.

Stictic acid partially restored p53 activity in human Saos-2 cells expressing cancer mutant R175H (Fig. 5), as demonstrated by induction of the cell cycle inhibitor p21. p21 is a canonical gene target of active p53 and a sensitive and selective output for p53 activity. Induction of p21 by stictic acid was dose-responsive, as shown in Fig. 5. It was also dependent upon p53, because it had only a minor effect on p21 expression in cells lacking p53 (Fig. 5, Supplementary Fig. S4).

p53 primarily functions as a transcriptional activator. To test more directly the effect of stictic acid on restoring transcriptional activator function of p53 mutants we generated stable cell lines harbouring reporter constructs driven by two different p53-activator function of p53 mutants we generated stable cell lines more directly the effect of stictic acid on restoring transcriptional activity. Induction of p21 reporter by reactivation compounds. PRIMA-1 and MQ stabilized each of the p53 variants by ~≤ 1° for concentrations ranging from 10 to 80 μM, 2.5× to 20× stoichiometric ratio of compound to protein (Fig. 7). The greatest thermal stabilization was observed with stictic acid, up to 4.60 ± 0.39 ºC at 80 μM in mutant R175H. All three of these compounds demonstrated a relative dose dependence in their ability to stabilize p53. None of these compounds had stabilizing effects on three control proteins tested, annexin A2, S100A4 and S100A14, suggesting that their interactions are specific to p53.

Discussion

Despite continuous progress in cancer research and care, millions of people die annually from this disease. Therapeutic reactivation of p53 could prevent or delay many of these deaths. The p53 Y220C reactivation pocket already has been used for rational drug design and discovery in that specific mutation. Here we report computer-guided identification of a transiently open L1/S3 specific manner as expected (Fig. 6a). Activation of p21 reporter transcription was more pronounced and confirmed that stictic acid is a p53 mutant reactivation compound comparable to previously identified compounds.

Differential scanning fluorimetry (DSF) provides the ability to detect molecular interactions between a small molecule and a protein based on changes in the melting temperature (Tm) of the protein. These Tm changes are taken to reflect changes in protein stability upon binding of the small molecule. Variants of the p53 core domain (wild type, G245S and R175H) were tested in DSF assays with a range of concentrations of PRIMA-1, MQ and stictic acid (Fig. 7). The Tm values in the absence of compounds for wild type, G245S and R175H, at 4 μM, were 36.54 ± 0.11, 33.51 ± 0.09 and 27.89 ± 0.11 ºC, respectively. PRIMA-1 and MQ stabilized each of the p53 variants by ~≤ 1° for concentrations ranging from 10 to 80 μM, 2.5× to 20× stoichiometric ratio of compound to protein (Fig. 7). The greatest thermal stabilization was observed with stictic acid, up to 4.60 ± 0.39 ºC at 80 μM in mutant R175H. All three of these compounds demonstrated a relative dose dependence in their ability to stabilize p53. None of these compounds had stabilizing effects on three control proteins tested, annexin A2, S100A4 and S100A14, suggesting that their interactions are specific to p53.
Figure 7 | Thermal stabilization of p53 mutants by reactivation compounds. Selected concentrations of PRIMA-1 (blue, square), MQ (green, circle) and stictic acid (red, triangle) were added to 4 µM of wild type, G245S and R175H core domain protein. Thermal stabilization was measured using DSF. Data points are plotted as the mean of sextuplet measurements, with corresponding 95% confidence intervals. Confidence interval bars not visible are smaller than the symbol used to plot the data point. PRIMA-1 and MQ stabilized each p53 variant about 1 °C or less over this range of concentrations. Stictic acid resulted in the greatest stabilization for both p53 mutants studied.

reactivation pocket in p53 that may present a more broadly applicable druggable region for such cancer therapies. The coordinates of the human p53 core domain structures with the open binding pocket conformations shown in Fig. 4 appear as Supplementary Data S1, and also have been deposited in the Protein Model DataBase (PMDB) (http://mi.caspur.it/PMDB/; PMDB ID (a) PM0078199 for wt p53; (b) PM0078200 for R175H; (c) PM0078201 for G245S; and (d) PM0078202 for R273H) and in SourceForge (http://sourceforge.net/projects/p53cancerrescue/, in folder ‘Coordinates of p53 core domain’). We anticipate that the L1/S3 reactivation region reported here will be useful to other researchers carrying out structure-based drug discovery and development of p53 reactivation drug leads.

A notable advantage of the L1/S3 pocket is that it occurs across several p53 variants, as seen in Fig. 4 and Supplementary Figs S3 and S6, giving hope that p53 reactivation compounds can be developed to target several p53 cancer mutants simultaneously. In contrast, for example, known compounds that target the p53 Y220C reactivation pocket may be specific to the Y220C mutant, because specific structural details of the Y to C mutation were used in their design or discovery. None of the R175H, G245S or R273H mutations lie in the L1/S3 pocket or have direct effects on its conformation. Its generalizability is illustrated here by the result that virtual screening against one mutant, R273H, produced a compound that resulted in biological validation against a different mutant, R175H, and induced DSF temperature shifts in all p53 variants studied.

Protein conformational flexibility is well-recognized as an important factor in ligand binding. A few computational drug discovery approaches attempt to accommodate protein-induced fit on ligand binding by allowing side chains to move to improve the fit of the ligand into the binding pocket. Very few such approaches allow the protein backbone to move to accommodate the ligand. Nevertheless, the results above provide a cautionary tale: a ligand-binding pocket of medical importance may be hidden and difficult to recognize in structural databases, yet other virtual structures extracted from its conformational ensemble may support successful structure-based drug discovery. With appropriate protein conformations in hand for computational analysis, we screened only 45 compounds through biological assays to find a novel p53 reactivation compound, stictic acid.

Our relaxed complex scheme clusters are based on conformations of all residues surrounding the L1/S3 pocket, not on specific conformations of particular residues. No cluster centroid produced a suitable docking pose for stictic acid docked to G245S (Fig. 4d), due to the position and rotamer state of the crucial Leu114 residue. A suitable Leu114 conformation occurred in only three or four frames of the entire 3,000-frame 30-ns trajectory (Methods, RMSD-based clustering). The sampling in MD is limited, so the fact that any suitable pocket-opening events were seen at all is indication that such pockets exist at thermal equilibrium for this mutant. Indeed, three occurrences in 30 ns is equivalent to 10^8 occurrences per second, whereas three occurrences in 3,000 frames is equivalent to a probability of 10^-3, and so the event cannot be considered rare. Instead, we see here a limitation of our current computational methods for modelling complex molecular processes.

As seen in Supplementary Figs S7 and S9, stictic acid remained docked after 60 ns of stable MD simulation. The rotamer state of Leu114 was critical to opening up a binding site large enough to accommodate large molecules such as stictic acid. In the initial docked pose of stictic acid (Supplementary Fig. S9a), there is only one direct H-bond interaction between stictic acid and the Cys124 backbone amine group, in addition to various hydrophobic interactions including Pro142 and Leu114 (Supplementary Fig. S9c). The hydrophobic interactions with Pro142 and Leu114 were conserved throughout the simulation. The final pose (Supplementary Fig. S9b) provides an elegant illustration of induced fit upon binding. Stictic acid moved deeper into the pocket, achieved a very nice shape complementarity, substantially expanded the initial network of favourable direct and water-mediated interactions, and shifted the ligand and receptor exposure patterns (Supplementary Fig. S9d).

PRIMA-1 has conversion products that are more active than it is itself, and this may be the case for stictic acid as well. The existence of a highly active and bioavailable stictic acid decomposition product was suggested by initial experiments carried out with a stictic acid solution that had a long storage time. Such hypothetical conversion products were not present in the freshly ordered newly dissolved pure compound, from which the results presented here were derived.

It is unclear whether or not stictic acid requires covalent linkage to p53 for reactivation according to the general cysteine alkylation mechanism proposed by Lambert et al. for several other p53 reactivation compounds. Stictic acid may, in principle, bind covalently to nucleophilic protein side chains by acylating or alkylating nucleophilic residues such as cysteine. Likely reaction sites and modes of reactivity (Supplementary Fig. S10) include the direct addition of a nucleophile (a cysteine thiol group of the protein, as represented in the figure) to either the seven-membered ring lactone carbonyl group (A) or to the carbonyl group of the five-membered ring lactone (B), resulting in
acetylation of the thiol and covalent attachment to the protein as the respective thioester. Alkylation is another possible mechanism, occurring via a more circular mechanism (C) in which dehydration generates a reactive quinone methide-type species that can alkylate nucleophiles, as illustrated schematically for a p53 cysteine side chain thiol group. An alternative dehydration pathway (loss of a proton on the para-methyl group, not shown) could produce a different methide that would result in alkylating at the corresponding benzylic position. There is no a priori reason to rule out any of these pathways.

In DSF experiments, preheated PRIMA-1 (to induce formation of conversion products26) induced thermal shifts ($T_m$ increases) that were slightly higher than that for MQ. Though slight, the differences were statistically significant at higher concentrations, which is consistent with the previous evidence that MQ is not the most active conversion product of PRIMA-1 (ref. 26). These two known reactivation compounds showed significantly lower stabilizing effects on the p53 mutants tested than did stictic acid. Stictic acid is a much larger molecule than MQ. It has the potential to make more interactions within the binding pocket (compare Figs 2 with 4), perhaps giving a better fit to PRIMA-1 or MQ for stabilizing p53 cancer mutants. We cannot rule out other effects, for example, the possibility that the low thermal shifts for PRIMA-1 may be due to its incomplete conversion into more active products, or that the low thermal shifts for MQ may be due to its highly reactive nature leading to rapid decomposition on contact with water.

Stictic acid displayed promising results in reactivating severely destabilized mutant R175H in human cancer cells. Reactivation was evident from dose-dependent and p53-dependent induction of p21 expression. Cell death was not activated by stictic acid, which may be due to insufficient reactivation of cancer mutant R175H or may indicate pathway-selective restoration of p53 activity. Such probes with pathway discriminatory activity will become valuable tools for the study of p53 biology. PUMA reporter assays demonstrated comparable p53 reactivation by stictic acid and PRIMA-1 (Fig. 6), yet PRIMA-1 is a potent inducer of apoptosis in cells with mutant p53. It is possible that additional, p53-independent cytotoxic functions of PRIMA-1 or one of its degradation products26 act synergistically with p53 reactivation to induce apoptosis. Such additive effects may not be induced by stictic acid.

Mutant R175H showed a greater degree of thermal stabilization in DSF by stictic acid than did wild type or G245S. Perhaps, this is because R175H is initially highly destabilized (a Zn region mutant, 30% folded at $37^\circ C$) while wild type and G245S (a DNA region mutant, 95% folded at $37^\circ C$) are not. These results indicate that stabilization effects and reactivation outcome depend on the p53 mutation, and that thermal stabilization of the p53 core domain may reactivate some p53-controlled pathways and not others in a way that depends on the ligand.

DSF in vitro experiments indicated direct binding of stictic acid to all p53 variants tested. Induction of p21 expression by stictic acid showed dependence on p53-mediated pathways, because no such induction was observed in the Saos-2 parent cell lines lacking p53. Taken together, these in vitro and in cellulo results indicate that stictic acid acts directly on p53.

Recently, two compounds from the thiosemicarbazone family (for example, NSC-319725) were identified as allele-selective reactivation compounds for p53 cancer mutant R175H25. Whether these compounds directly bind to p53 and target the L1/S3 reactivation pocket identified in this study is currently unclear, but will be interesting questions to address in future experiments.

Our results demonstrate the value of integrated strategies combining genetics and computation for identification of druggable protein regions, and highlight the feasibility of using MD simulations to expose druggable protein regions that otherwise may not be evident.

### Methods

**MD simulations.** Wild-type p53 MD simulations were performed using crystal structure 1TSR-B33 for the initial coordinates. Models for mutants R273H, R175H, G245S and C124A were prepared by virtual mutation of wild-type p53 using the AMBER10 suite49. The crystallographic water molecules were retained and each system was solvated in a TIP3P water box50. Zinc-coordinating residues were modeled using the cationic dummy atom model51. Protonation states of titratable residues were determined using the Whatif Web Interface51. Histidine, asparagine and glutamine side chain orientations were corrected as needed using the Molprobity web server52. Each system was neutralized by adding chloride ions as needed. The structure files of each system consisted of about 27,260 atoms, which were prepared using the AmberFF9SB force field53.

Each of the p53 variant systems was first relaxed using 36,000 steps of minimization followed by four consecutive restrained 250 ps MD simulations. Thereafter, 30 ns unrestrained MD simulation was performed at 310 K and 1 atm for each system using the NAMD2.7 suite54. Details of the minimization and MD simulation protocol can be found in Demir et al.55.

To gauge the stability of the stictic acid in the binding pocket, a 60 ns MD simulation was performed. The best-scoring docked pose of the molecule in the R175H L1/S3 open-pocket conformation was used as the starting geometry (Fig. 4b, Supplementary Table S1). Stictic acid parameters were generated with the Antechamber module of Amber10 using the Generalized Amber Force Field55,56 with RESP HF 6-31G* charges.

For each p53 variant studied, the solvent-accessible surface area of the Cys124 side chain over the MD trajectory was calculated using a customized visual molecular dynamics27 (VMD) script.

The coordinates of the human p53 core domain structures with the open-binding pocket conformations shown in Fig. 4 appear as Supplementary Data S1, and also have been deposited in the PMDB 47 (http://mi.caspar.it/PMDB/; PMDB ID (a) PM0078199 for wt p53; (b) PM0078200 for R175H; (c) PM0078201 for G245S; and (d) PM0078202 for R273H) and in SourceForge (http://sourceforge.net/projects/p3ccancerrescue/), in folder 'Coordinates of p53 core domain'.

### RMSD-based clustering.

Protein snapshots extracted from the simulations were clustered with respect to all heavy atoms of the L1/S3 site. For each variant, atomic coordinates were extracted at 10 ps intervals over their 30 ns MD trajectory, resulting in 3,000 simulated structures per trajectory. These 3,000 structures were superimposed with respect to all C$_\alpha$ atoms to remove overall translation and rotation. They were then clustered using the gromos algorithm58 in the GROMACS 4.0.5 program59 at an RMSD cutoff value of 1.65 Å, based on atomic coordinates of all atoms of those residues that had at least one atom within 10 Å of Cys124 in the minimized and equilibrated form of each p53 variant. In each cluster, the structure with the smallest average RMSD to all other structures in the cluster was designated as the ‘centroid’ or ‘representative’ structure of the cluster.

Cluster centroids exhibiting open L1/S3 pockets were identified for each p53 variant using the geometric criteria (Supplementary Methods). These open centroids were used for docking stictic acid and the known p53 reactivating small molecules. In the case of G245S, the known small molecules were successfully docked into the cluster centroid with an identified open L1/S3 pocket. However, it was not possible to identify a cluster centroid for G245S that produced a binding pose for stictic acid (a relatively much larger molecule) in which a possible alkylation reaction with the Cys124 side chain seemed reasonable. A closer look at the G245S MD trajectory revealed that the position and rotamer state of Leu114 was critical for L1/S3 pocket opening. Leu114, already implicated above, exhibited a favourable conformation in only three or four frames of the entire 3,000-frame trajectory. Thus, a snapshot with an open L1/S3 pocket was extracted manually from the G245S MD trajectory and subsequently used to dock stictic acid.

### In silico docking.

Initial docking studies were performed using AutoDock 4.2.2 (SMILES strings for all small molecules considered explicitly in this paper are given in Supplementary Table S4). All subsequent docking studies used AutoDock Vina60 with an exhaustiveness value of 100 and a $20 \times 20 \times 20$ Å search space centred on the sulphur atom of Cys124. Nine different docking poses were calculated for each compound in combinations of a compound and a p53 variant. For initial compound scoring, population-weighted average binding affinities of the most-populated 15 cluster centroids from the R273H MD trajectory were calculated using the best binding affinity found by AutoDock Vina for each compound. The 48 compounds selected were re-scored using the population-weighted average binding affinities of the most-populated 30 cluster centroids. Ligand interaction diagrams were prepared using the Molecular Operating Environment program (Chemical Computing Group Inc., Montreal, Canada).
Computational fragment mapping using FTMap. The FTMap program44 utilizes 16 probe molecules to identify and rank the most energetically favourable binding pockets (PF binding). For this purpose, FTMap first performs fast and body docking of probe molecules with a fast Fourier transform method, then minimizes and rescres the docked poses. Next, it clusters them for each probe using a simple greedy algorithm, and finally locates the consensus sites of binding pockets identified by all probes.

Cell lines. Soas-2 cells were obtained from the American Type Culture Collection (Manassas, VA, USA). They were maintained at 37 °C in DMEM high glucose supplemented with 10% FBS, penicillin G sodium (units ml^−1) and streptomycin (100 μg ml^−1).

Stable Soas-2 cell lines harbouring doxycycline-inducible p53 wild type or mutants were established by a lentiviral-based strategy. Human p53 genes were cloned as SacI/PstI fragments into pEN_TmiRc3 and the lentivirus vector pSlik B-containing media. To isolate single p53-expressing clones, cells were induced plates using Lipofectamine 2000 (Invitrogen, Carlsbad, CA). Supernatants were performed on mutant R175H was identical except that the starting temperature and DSF was used to various concentrations and added to diluted protein for overnight incubation at 4 °C. An equivalent volume of DMSO was used as a control and also incubated overnight at 4 °C. For western blot analysis, each protein and compound condition was tested as replicates of 6 with 20 μl per well of a 96-well plate. DSF was performed on wild-type p53 and mutant Ga245S using the CFX96 Touch RT–PCR system (Bio-Rad) by increasing the temperature of the plate 0.5 °C per min from 25 to 52.5 °C. DSF performed on mutant R175H was identical except that the starting temperature was 25 °C. Fluorescence measurements were taken once every minute during all experiments. T_m values for each well, mean T_m values, and 95% confidence intervals were calculated with the program Prism 5 (GraphPad Software, La Jolla, CA) by fitting the raw fluorescence data to a Boltzmann sigmoidal curve.

References

1. Olivier, M. et al. The IARC TP53 database: new online mutation analysis and recommendations to users. Hum. Mutat. 19, 607–614 (2002).
2. Soussi, T. & Beroud, C. Assessing TP53 status in human tumours to evaluate clinical outcome. Nat. Rev. Cancer 1, 233–240 (2001).
3. Joerger, A. C. & Fersht, A. R. The tumor suppressor p53: from structures to mechanisms and therapeutic potential. Oncogene 26, 2243–2254 (2007).
4. Martins, C. P., Brown-Swigart, L. & Evan, G. I. Modeling the therapeutic efficacy of p53 restoration in tumours. Cell 127, 1323–1334 (2006).
5. Selivanova, G. & Wiman, K. G. Reactivation of mutant p53: molecular mechanisms and therapeutic potential. Oncogene 26, 2243–2254 (2007).
6. Sharpless, N. E. & DePinho, R. A. Cancer biology: gone but not forgotten. Nature 445, 606–607 (2007).
7. Ventura, A. et al. Restoration of p53 function leads to tumour regression in vivo. Nature 445, 661–665 (2007).
8. Wiman, K. G. Restoration of wild-type p53 function in human tumors: strategies for efficient cancer therapy. Adv. Cancer Res. 97, 321–338 (2007).
9. Xue, W. et al. Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. Nature 445, 656–660 (2007).
10. Baroni, T. E. et al. A global suppressor motif for p53 cancer mutants. Proc. Natl Acad. Sci. USA 101, 4930–4935 (2004).
11. Baranov, R. et al. All-code scanning identifies p53 cancer rescue mutations. Nucleic Acids Res. 38, 7079–7088 (2010).
12. Brachmann, R. K., Yu, K., Eby, Y., Patil, N. P. & Boeke, J. D. Genetic selection of intragenic suppressor mutations that reverse the effect of common p53 cancer mutations. EMBO J. 17, 1847–1859 (1998).
13. Danziger, S. A. et al. Predicting positive p53 cancer rescue regions using Most Informative Positive (MIP) active learning. PLoS Comput. Biol. 5, e1000498 (2009).
14. Danziger, S. A. et al. Functional classification of mutation sequence spaces: the example of p53 cancer rescue mutants. IEEE ACM Transact. Comput. Biol. Bioinform. IEEE ACM 3, 114–125 (2006).
15. Danziger, S. A., Zeng, J., Wang, Y., Brachmann, R. K. & Lathrop, R. H. Choosing where to look next in a mutation sequence space: active Learning of informative p53 cancer rescue mutants. Bioinformatics 23, i104–i114 (2007).
16. Wieczorek, A. M., Waterman, J. L., Waterman, M. J. & Halazonetis, T. D. Structure-based rescue of common tumor-derived p53 mutations. Nat. Med. 2, 1143–1146 (1996).
17. Bykov, V. J., Issaeva, N., Selivanova, G. & Wiman, K. G. Mutant p53-dependent growth suppression distinguishes PRIMA-1 from known anticancer drugs: a statistical analysis of information in the National Cancer Institute database. Carcinogenesis 23, 2011–2018 (2002).
18. Bykov, V. J. et al. Restoration of the tumor suppressor function to mutant p53 by a low-molecular-weight compound. Nat. Med. 8, 282–288 (2002).
19. Bykov, V. J. et al. Reactivation of mutant p53 and induction of apoptosis in human tumor cells by maleimide analogs. J. Biol. Chem. 280, 30384–30391 (2005).
20. North, S., Pluquet, O., Maurici, D., El-Ghissassi, F. & Hainaut, P. Restoration of wild-type conformation and activity of a temperature-sensitive mutant of p53 (p53V272M)) by the cytosine analog 5-aza-2′-deoxycytidine in the esophageal carcinoma cell line TE-1. Mol. Carcinog. 33, 181–188 (2002).
21. Reddy, N. L., Hill, J., Ye, L., Fernandez, P. R. & Stout, D. M. Identification and structure-activity relationship studies of 3-methylen-2-norbornanone as
potent anti-proliferative agents presumably working through p53 mediated apoptosis. Bioorg. Med. Chem. Lett. 14, 5645–5649 (2004).

22. Zagar, I. et al. p53 targets the low molecular weight compound STIMA-1. Mol. Oncol. 2, 70–80 (2008).

23. Boeckler, F. M. et al. Targeted rescue of a destabilized mutant of p53 by an in silico screened drug. Proc. Natl Acad. Sci. USA 105, 10360–10365 (2008).

24. Wüllen, R. et al. Halogen-Enriched Fragment Libraries as Leads for Drug Rescue of Mutant p53. J. Am. Chem. Soc. 134, 6810–6818 (2012).

25. Xu, Y., Vazquez, A., Levine, A. J. & Carpio, D. R. Allele-Specific p53 Mutant Reactivation. Cancer Cell 21, 614–625 (2012).

26. Lambert, J. M. et al. PRIMA-1 reactivates mutant p53 by covalent binding to the core domain. Cancer Cell 15, 376–388 (2009).

27. Wiman, K. G. Pharmacological reactivation of mutant p53: from protein structure to the cancer patient. Oncogene 29, 4245–4252 (2010).

28. Amaro, R. E., Baron, R. & McCammon, J. A. Improved relaxed complex scheme for receptor flexibility in computer-aided drug design. J. Comput. Aided Mol. Des. 22, 693–705 (2008).

29. Cheng, L. S. et al. Ensemble-based virtual screening reveals potential novel antiviral compounds for avian influenza neuraminidase. J. Med. Chem. 51, 3878–3881 (2008).

30. Malecka, K. A., Ho, W. C. & Marmorsein, R. Crystal structure of a p53 core tetramer bound to DNA. Oncogene 28, 325–333 (2009).

31. Friedler, A. et al. A peptide that binds and stabilizes p53 core domain: chaperone strategy for rescue of oncogenic mutants. Proc. Natl Acad. Sci. USA 99, 937–942 (2002).

32. Kaar, J. L. et al. Stabilization of mutant p53 via alkylation of cysteines and effects on DNA binding. Proc. Natl Acad. Sci. USA 109, 2267–2278 (2010).

33. Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. Science 265, 346–355 (1994).

34. Wang, Y., Rosengarth, A. & Luecke, H. Structure of the human p53 core domain in the absence of DNA. Acta crystallographica. Sect. D Biol. Crystallogr. 63, 276–281 (2007).

35. Eminazadah, S., Tropia, L. & Halazonetis, T. D. Crystal structure of a monodomain human p53 tetramer bound to the natural CDKN1A (p21) response element. Mol. Cancer Res. 9, 1493–1499 (2011).

36. Petz, T. J. et al. An induced fit mechanism regulates p53 DNA binding kinetics to identify specific sequence specificity. EMBO J. 30, 2167–2176 (2011).

37. Canadillas, J. M. et al. Solution structure of p53 core domain: structural basis for its instability. Proc. Natl Acad. Sci. USA 103, 2109–2114 (2006).

38. Veu, C. S., Nitue, S. K., Doneau, C. E., Pattabiraman, N. & Srivenugopal, K. S. Human p53 is inhibited by glutathionylation of cysteines present in the proximal DNA-binding domain during oxidative stress. Biochemistry 46, 7765–7780 (2007).

39. Bullock, A. N. et al. Thermodynamic stability of wild-type and mutant p53 core domain. Proc. Natl Acad. Sci. USA 94, 14338–14342 (1997).

40. Bullock, A. N., Hencel, J. & Fersht, A. R. Quantitative analysis of residual folding and DNA binding in mutant p53 core domain: definition of mutant states for rescue in cancer therapy. Oncogene 19, 1245–1256 (2000).

41. Natan, E. et al. Interaction of the p53 DNA-binding domain with its n-terminal extension modulates the stability of the p53 tetramer. J. Mol. Biol. 409, 358–368 (2011).

42. Demir, O. et al. Ensemble-based computational approach discriminates functional activity of p53 cancer and rescue mutants. PLoS Comput. Biol. 7, e1002318 (2011).

43. Le Guilloux, V., Schmidtke, P. & Tuffery, P. Fpocket: an open source platform for ligand pocket detection. J. Comput. Aided Mol. Des. 22, 693–705 (2008).

44. Brenek, R. et al. Fragment-based identification of druggable ‘hot spots’ of proteins using Fourier domain correlation techniques. Bioinformatics 25, 621–627 (2009).

45. Koh, A. M. et al. Using targeted transgenic reporter mice to study promoter-specific p53 transcriptional activity. Proc. Natl Acad. Sci. USA 109, 1685–1690 (2012).

46. Vedadi, M. et al. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proc. Natl Acad. Sci. USA 103, 15835–15840 (2006).

47. Canfield, T., De Meo, P. D., Cozzetto, D., Calamari, G. & Tramontano, A. The PMDB Protein Model Database. Nucleic Acids Res. 34, D306–D309 (2006).

48. Case, D. A. et al. AMBER 10 (University of California, 2008).

49. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926–935 (1983).

50. Pang, Y.-P. Novel zinc protein molecular dynamics simulations: steps toward antiangiogenesis for cancer treatment. J. Mol. Model 5, 196–202 (1999).

51. Rodriguez, R., Chinea, G., Lopez, N., Pons, T. & Vriend, G. Homology modeling, model and software evaluation: three related resources. Bioinformatics 14, 523–528 (1998).

52. Chen, V. B. et al. MolProbit: all-atom structure validation for macromolecular crystallography. Acta crystallographica. Sect. D Biol. Crystallogr. 66, 12–21 (2010).

53. Hornak, V. et al. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 65, 712–725 (2006).

54. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. J. Comput. Chem. 26, 1781–1802 (2005).

55. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. J. Mol. Graph. Model 25, 247–260 (2006).

56. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. J. Comput. Chem. 25, 1157–1174 (2004).

57. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. J. Mol. Graph. 14, 33–28 (1996).

58. Baron, R. & McCormann, J. A. Dynamics, hydration, and motional averaging of a loop-gated artificial protein cavity: the W191G mutant of cytochrome c peroxidase in water as revealed by molecular dynamics simulations. Biochemistry 46, 10629–10642 (2007).

59. Hess, B., Kutzner, C., Spuel, D. v. d. & Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J. Chem. Theory Comput. 4, 435–447 (2008).

60. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461 (2010).