PNAS

Supporting Information for

A molecular mechanism for the 'digital' response of p53 to stress

Jessy Safieh[1], Ariel Chazan[1], Hanna Saleem[1], Pratik Vyas[1,2], Yael Danin-Poleg[1], Dina Ron[1,3], and Tali E. Haran[1,3].

[1] Department of Biology, Technion, Technion City, Haifa, Israel

[2] Present address: Department of Chemical and Biological Physics, Weizmann Institute of Science, Rehovot, Israel

[3] To whom correspondence may be addressed.
Email: bitali@technion.ac.il
Email: dinar@technion.ac.il

This PDF file includes:

> Supporting text
> Figures S1 to S4
> Tables S1 to S5
> Legend for Dataset S1
> SI References

Other supporting materials for this manuscript include the following:

> Dataset S1

Supporting Information Text

SI Materials and Methods

Cyclization kinetics and simulation of cyclization data

DNA constructs for cyclization kinetics were synthesized using the top library/bottom test sequence PCR scheme, described previously (1-4), except that here we used as test sequences DNA sequences containing the full 20-bp long p53 sites plus 5-bp flanking sequences on both sides, specific for each natural p53 site (see Table S3 for sequences). The DNA test sequences for cyclization experiments (bottom PCR templates) were synthesized by Sigma Genosys (Israel), whereas the library DNA sequences (top PCR templates) and the fluorescein- and tetramethylrhodamine (TAMRA)-labeled oligonucleotide primers were synthesized by the Keck foundation Laboratory at Yale University. Cyclization kinetics measurements were carried out as described previously (2-4). PCR reactions (50 µl) contained 1× PrimeSTAR buffer, 0.2 mM dNTP mix, 0.3 µM of each of the primers, 100 nM of the top and bottom templates, and 0.04 units of PrimeSTAR DNA polymerase (Takara, Japan, Cat # R010B). Ligase concentration was varied as a function of phasing length (0.08 U/µl for the in-phase 156L14 and 156L16 constructs and 1.0 U/µl for all other phasing constructs) and total length (from 0.08 U/µl for the 157 constructs, to 0.8 U/µl for the 154–155 and 158–160 constructs and 2.5 U/µl for the 150–153 constructs). We derived quantitative data on the conformational properties of the test sequences by simulating the cyclization data as previously described (2-5), using the simulation program developed by Zhang and Crothers (5). The outcome of the simulations are the bend angle, twist angle, roll and tilt flexibility and twist flexibility per DNA sequence.

Cells, plasmids, and reagents

H1299 human non-small lung carcinoma cells (p53-null cells) were obtained from ATTC (CRL-5803). H1299 cells were grown in RPMI-1640 (Biological Industries, Beth Haemek, Cat # 01-101-1A), and MCF-7 breast cancer cells were grown in DMEM (Biological Industries). Both mediums were supplemented with 10% Fetal Bovine Serum (Biological Industries, Beth Haemek, Cat # 04-127-1A), 2 mM L-glutamine, and 1% Pen-Strep solution (Biological Industries, Beth Haemek, Cat # 31-031-1C). All experiments were performed in these specified mediums. Cells were maintained in a humidified incubator at 37°C and 5% CO_2. Cells were routinely checked for negative mycoplasma infection.

The reporter gene vectors pCLuc Mini- TK 2, encoding *Cypridina* secreted luciferase under the control of a minimal promoter, and pCMV-GLuc, encoding constitutively expressed *Gaussia* secreted luciferase, were purchased from Thermo-Fisher. pC53-SN3 construct (where p53 expression is driven by the CMV promoter) was used for wild-type p53 expression in H1299 cells. p53 expression plasmid was kindly provided by Varda Rotter, Weizmann Institute of Science, Rehovot, Israel.

Generation of reporter plasmids containing p53 REs

All DNA sequences were synthesized by Sigma Genosys (Israel), purified by standard desalting, and are described in Table S5. The sequences corresponding to p53 REs, each containing also 5-bp flanking sequences on both the 5' and the 3' sides, were cloned upstream to the CLuc gene (encoding the *Cypridina* luciferase) in the pCluc Mini-TK plasmids using site-directed mutagenesis with KAPA HiFi™ DNA Polymerase kit (Gamidor Diagnostics), and transformed into competent DH5α cells, according to manufacturer recommendations. Overexpressed plasmids were extracted using QIAprep® Spin Miniprep Kit (Qiagen), according to the manufacturer's protocol. The correct identity of incorporated sequences was verified using Sanger sequencing.

Stress induction

MCF-7 cells were grown in 35 mm plates at a density of $4.5X10^5$ cells. Cells were allowed to reach 80% confluency in a fresh medium before treatment. Exponentially growing cells

were exposed to either 2 μM or 10 μM Camptothecin (CPT) and harvested at the indicated time points post incubation with CPT.

p53 protein quantitation and detection
All cells were lysed in RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% Na deoxycholate, 0.1% SDS, 25 mM Tris-Cl, pH 7.5) containing 1/100 (Vol/Vol) Protease inhibitor cocktail (Sigma, Cat # P8340). An equal amount of whole cell lysate proteins from the different samples were resolved on 10% SDS PAGE, and transferred to PVDF membrane. The blots were subsequently incubated at 4°C with primary antibodies against p53 (DO-7 mouse mAb, Cell Signaling, Cat # 48818) and β actin (13E5 rabbit mAb, Cell Signaling, Cat # 4970), sequentially. Membranes were then reacted with a secondary antibody (Anti-mouse IgG, and Anti-rabbit IgG conjugated to Horseradish Peroxidase (both from Cell Signaling, Cat # 7076 and # 7074, respectively), developed using Western Bright^{TM} enhanced chemiluminescence reagent (Advansta), air-dried, and exposed using Fusion Pulse (Vilber) detection system, followed by quantitation using Cliqs software (version 1.1, TotalLab Ltd., UK). For p53 serial-dilution curve, full-length recombinant p53 protein was resolved on gels together with the transfection samples.

Information content analysis
Information content per DNA binding site (Iseq, ref. 6, 7) was calculated separately for the whole set of REs as well as for the flexible, the rigid functional outcome categories and the group with intermediate flexibility, on the mononucleotide level, using the following equations:

(1) $Iseq\ (mono) = \sum_{l=1}^{s} Iseq_l(mono) = \sum_{l=1}^{s} \sum_{B=0}^{3} f_{lb} * \ln\frac{f_{lb}}{P_B} = \sum_{l=1}^{s} \sum_{B=0}^{3} f_{lb} * \ln(4 * f_{lb})$

Where $Iseq_l$ (mono) is the mononucleotide information content per position, f_{lB} is the frequency of occurrence of base B in position l in the population of all possible binding sites, and P_B of base B is the frequency in the whole genome, which is usually taken to be 0.25 for each base. To correct for small sample size, we replaced equation (1) with the following equation, shown to be the best estimate for the sequence information when f_{lB} is unknown precisely (7):

(2) $Iseq\ (mono) = \sum_{l=1}^{s} Iseq_l(mono) = \sum_{l=1}^{s} \sum_{B=0}^{3} \left(\frac{n_{lb}+1}{N+4}\right) * \ln\left(4 * \frac{n_{lb}+1.5}{N+4.5}\right)$

Where N is the sample size, and n_{lb} is the number of occurrences of base-pair B ($B = 0$, I, 2, 3) at position l. The expected variance in assigning f_{lB} values was calculated from (6):

(3) $\sigma_{lB}^2 = <f_{lB}^2> - <f_{lB}>^2 = \frac{(n_{lB}+1)(N-n_{lB}+1)}{(N+4)^2(N+5)}$
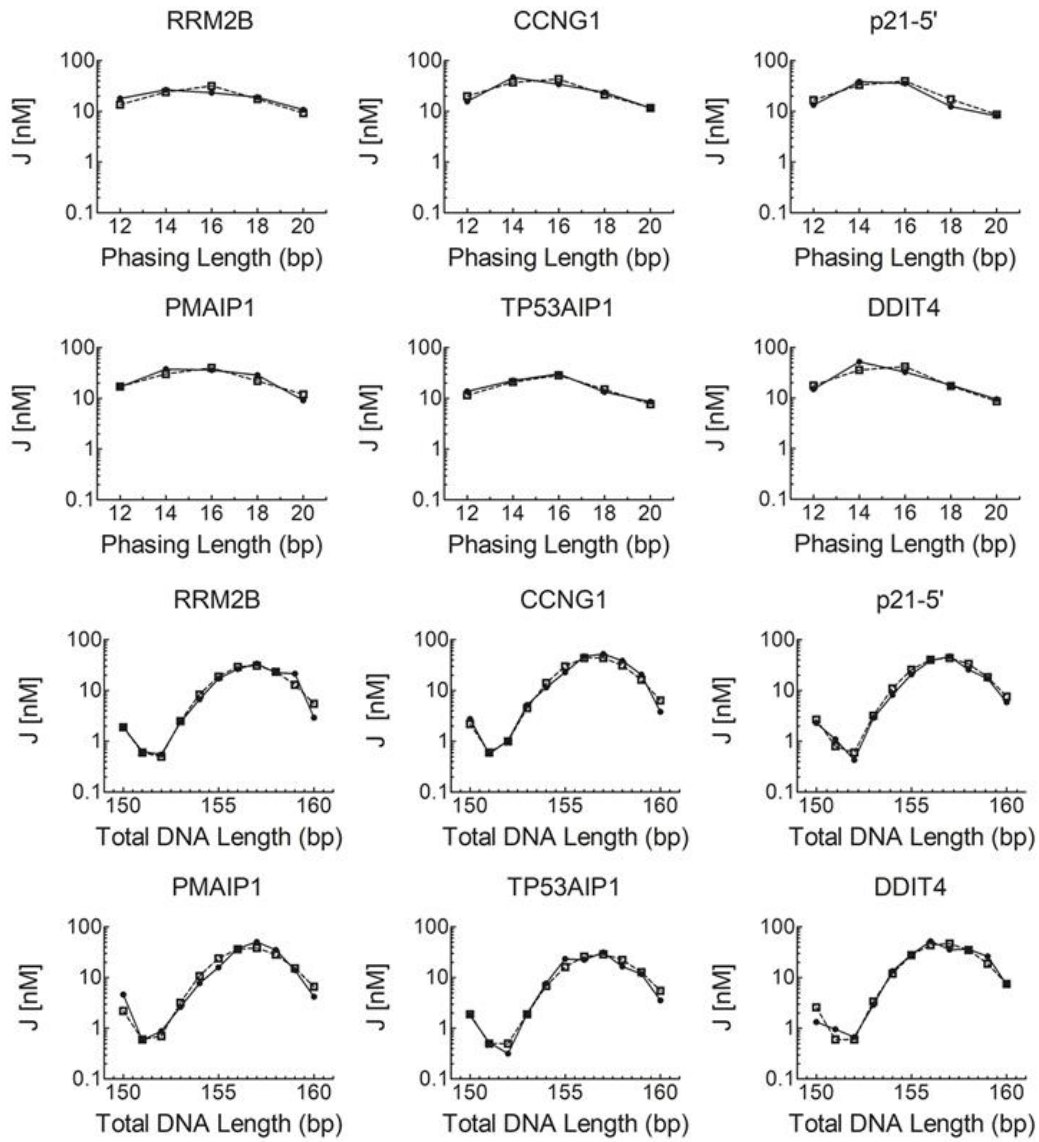
Fig. S1. Cyclization kinetics of p53 REs. The J-factors for the DNA constructs as a function of either the phasing length (top two rows), or the total DNA length (bottom two rows). The solid lines are the experimental curves and the dashed lines are the curves from simulating the data.
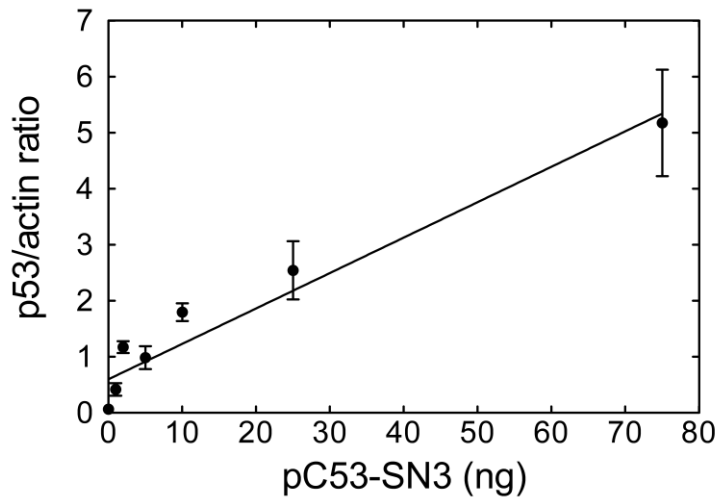
Fig. S2. Validation of ectopic p53 expression level in H1299 cells. Correlation between pC53-SN3 plasmid amount at the transfection against p53 protein level relative to β-actin, derived from the gel shown in Figure 4D. p53 protein level was estimated by normalizing the band intensity of p53 protein to the band intensity of β-Actin from the same sample. There is a linear relationship between the amounts of p53 expressing plasmid used in the transfection and the observed p53 protein levels (r = 0.972, *P* <0.0001). The results are averages of four independent experiments.
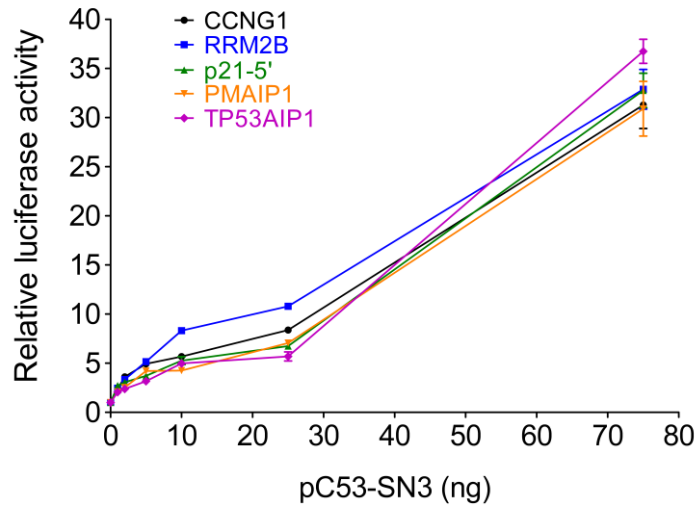
Fig. S3. Luciferase activity from p53 REs using 75 ng p53 expression vector. Shown are luciferase reporter gene assays in human non-small lung carcinoma cell line (H1299) ectopically expressing p53 along with reporter plasmids. The figure shows fold increase in transactivation level 48 h post-transfection, using 75 ng p53 expression vector. Luminescence values were normalized to the transfection efficiency of co-transfected, constitutively expressed Gaussia luciferase (Gluc). Results were further normalized first to the empty Cypridina secreted luciferase (CLuc) vector, and then to results obtained without p53. Error bars represent the mean ± SD of four independent experiments, each containing two technical replicas.
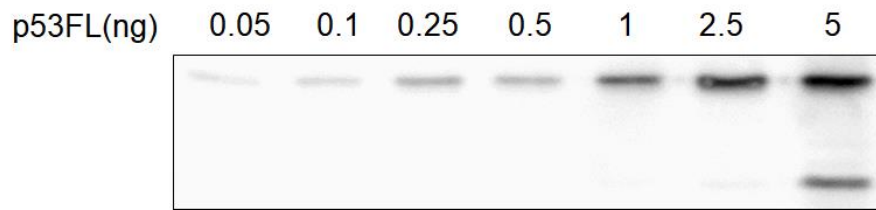
Fig. S4. Serial dilution of purified full-length p53 recombinant protein (5 sec exposure). The lower band in the 5-ng lane is a degradation product, which was included in the quantification.

Table S1. End functional-outcome groups with at least five genes each (related to Figure 1)

Group	Functional outcomes included in the group
Cell adhesion/anti-cell-motility	Increase in cell-matrix adhesion; increase in cell-cell adhesion; anti-metastasis; anti-migration; anti-invasion; anti-angiogenesis
Cell growth	Cell division machinery; mitotic cell cycle; mitotic spindle organization; regulation of cytokinesis; symmetric cell division
Cell signaling	Signal transduction by various pathways, when the signaling event is followed by multiple possible functional outcome results
Cellular stress response	Response to severe stress resulting from: ER stress, metabolic stress, energetic stress, hypoxia, heat stress, cold stress, osmotic stress, oxidative stress, and oxidative DNA damage. Unfolded protein response; stress granules formation; cellular stress response; response to severe stress
Cytoskeleton/cell motility	Cytoskeleton organization; actin polymerization; ECM degradation; muscle contraction; neuronal migration; movement along microtubules; directional migration; apical-basal polarity; cytoskeleton re-organization; axon guidance; migration; invasion; wound healing; EMT; platelet aggregation; angiogenesis; decrease in cell adhesion
Development	Embryonic development; morphogenesis; neurogenesis; synaptogenesis; muscle development
DNA repair	Damage repair; translesion DNA synthesis; synthesis of dNTPs
Early DNA damage response (DDR) steps	DNA damage response activation; cell-cycle arrest (G1/S and G2/M); p53 positive regulation; other events closely following DNA damage; immediate early genes for DDR
Energy metabolism	Energy metabolism; glucose metabolism; oxidative phosphorylation; ATP synthesis
Extrinsic apoptosis	Death receptor pathway
Innate immunity	Innate immune response; inflammation; pyroptosis
Intrinsic apoptosis	Intrinsic apoptosis machinery and effectors
Ion homeostasis	Ion channels; ion transport; transport of small molecules
Lipid metabolism	Lipid metabolism; lipid droplets; cholesterol metabolism; steroids metabolism; steroidogenesis
Membrane trafficking	Endocytosis; exocytosis; vesicle transport; secretory pathways; Golgi; membrane transport; vesicular lipid transport
p53 negative regulation	Targeting p53 for destruction; p53 down regulation after damage repair
Stem cells and differentiation	Stem cell biology; stem cells self-renewal; pluripotency maintenance and regulation; stem cells differentiation
Survival	Anti-apoptosis; anti-necrosis; anti-pyroptosis; anti other kinds of deaths; cell survival

Table S2. Significant differences between deformability values (V(B)°^{3}Å^3) grouped by functional outcome of gene activation by p53

First outcome group	to second outcome group	Score mean difference [a]	Std error difference [b]	*Z* [c]	*P*-Value
Development	Cellular stress response	146.3	23.9	6.109	<.0001
Development	Cytoskeleton/cell motility	125.6	20.7	6.058	<.0001
Development	Intrinsic apoptosis	-134.4	24.6	-5.462	<.0001
Development	p53 negative regulation	-145.5	29.3	-4.961	0.0001
Early DDR steps	Cellular stress response	125.9	29.3	4.293	0.0027
Early DDR steps	Cytoskeleton/cell motility	105.2	26.8	3.930	0.013
Early DDR steps	Intrinsic apoptosis	-114.0	29.9	-3.817	0.0207
Early DDR steps	p53 negative regulation	-125.1	33.9	-3.694	0.0338
Energy homeostasis	Cellular stress response	147.6	27.7	5.336	<.0001
Energy homeostasis	Cytoskeleton/cell motility	126.9	24.9	5.090	<.0001
Energy homeostasis	Intrinsic apoptosis	-135.6	28.2	-4.806	0.0002
Energy homeostasis	p53 negative regulation	-146.8	32.4	-4.526	0.0009
Innate immunity	Cellular stress response	121.7	21.3	5.702	<.0001
Innate immunity	Cytoskeleton/cell motility	101.0	17.7	5.718	<.0001
Innate immunity	Intrinsic apoptosis	-109.8	22.1	-4.973	0.0001
Innate immunity	p53 negative regulation	-120.9	27.2	-4.437	0.0014
Lipid homeostasis	Cellular stress response	103.6	25.4	4.077	0.007
Lipid homeostasis	Cytoskeleton/cell motility	82.9	22.4	3.700	0.033
Cell signaling	Cellular stress response	-86.3	20.7	-4.160	0.0049
Cell signaling	Cytoskeleton/cell motility	-65.6	16.9	-3.874	0.0164

[a] Score mean difference: The mean of the rank score of the deformability values in the first outcome group minus the mean of the rank scores of the deformability values in the second outcome group.

[b] Std error difference. The standard error of the score mean difference.

[c] Z: The standardized test statistic.

Table S3. Best-fit parameters from cyclization kinetics measurements for the sequences studied here and the corresponding calculated deformability

Functional outcome	Gene symbol	Test sequence *	Bend angle (°) [†,‡]	Twist angle (°) [‡]	Roll and tilt flexibility (°) [‡,§]	Twist flexibility (°) [‡,¶]	Deformability V(B) (°³Å³) [#]
Cell growth	CCNG1	tgtcc**TGACATGCCCAGGCATGTCT**tccaa	-0.72 (0.09)	34.52 (0.08)	4.54 (0.09)	5.77 (0.09)	2.326
DNA Repair	RRM2B	tgtga**GCACAAGCCCAGGCTAGTCC**gaggc	-1.49 (0.10)	34.14 (0.10)	4.20 (0.09)	5.65 (0.09)	2.495
Early DDR steps	p21-5'	gtcag**GAACATGTCCCAACATGTTG**agctc	-0.06 (0.08)	34.04 (0.09)	4.45 (0.08)	5.34 (0.08)	2.263
Intrinsic apoptosis	PMAIP1	gtcgg**GAGCGTGTCCGGGCAGGTCG**cgctc	0.83 (0.12)	34.30 (0.15)	4.33 (0.11)	5.07 (0.12)	2.042
Intrinsic apoptosis	TP53AIP1	cctcc**TCTCTTGCCCGGGCTTGTCG**agatg	-1.02 (0.11)	33.96 (0.11)	4.20 (0.12)	4.91 (0.10)	1.926
Range of values Observed [‖]			−10.0 to 2.4	33.96 to 34.85	4.20 to 6.40	3.9 to 5.80	

[*] Capital bold letters are sequences of natural p53 REs. Lower letters are the natural flanking sequences. The presented parameters are those of the 20-bp REs only.

[†] Bending is by roll and its center is located at the fourth step of all sequences, i.e. at the C-A, C-G, or C-T step.

[‡] Numbers in parenthesis are the simulation errors, calculated as described in ref. (5).

[§] Roll and tilt flexibility are the thermal fluctuations of the roll and tilt angle between adjacent base pairs (1, 5).

[¶] Twist flexibility is the thermal fluctuation of the twist angle between adjacent bases (1, 5).

[#] The deformability values are those from Dataset S1, and were calculated as previously described (4), using the values of (8).

[‖] These values denote the range of observed values for these parameters in other DNA sequences, determined by the same method used here (2-4, 9).

1

Table S4: ANOVA analysis and Tukey HSD Post-hoc test results of p53-dependent reporter gene assays

A. Transactivation 48 h after transfection with various levels of p53 plasmid

p53 plasmid level		1 ng	2 ng	5 ng	10 ng	25 ng	75 ng
ANOVA (F, P value)		4.95, 0.004	22.85, <0.0001	16.4, <0.0001	54.9, <0.0001	520, <0.0001	1.13, 0.38
Comparison between		**P values ***					
CCNG1	RRM2B	ns	ns	ns	***	***	ns
CCNG1	p21-5'	ns	*	**	ns	**	ns
CCNG1	PMAIP1	ns	***	ns	***	ns	ns
CCNG1	TP53AIP1	ns	***	***	ns	***	ns
RRM2B	p21-5'	ns	ns	***	***	***	ns
RRM2B	PMAIP1	ns	***	*	***	***	ns
RRM2B	TP53AIP1	ns	***	***	***	***	ns
p21-5'	PMAIP1	*	*	ns	*	ns	ns
p21-5'	TP53AIP1	**	**	ns	ns	ns	ns
PMAIP1	TP53AIP1	ns	ns	*	ns	*	ns

B. Transactivation with 25 ng p53 plasmid as a function of time

Time		12 h	18 h	24 h
ANOVA (F, P value)		4.34, 0.008	4.36, 0.008	79, <0.0001
Comparison between		**P values ***		
CCNG1	RRM2B	ns	ns	***
CCNG1	p21-5'	ns	ns	ns
CCNG1	PMAIP1	*	ns	**
CCNG1	TP53AIP1	*	ns	***
RRM2B	p21-5'	ns	ns	***
RRM2B	PMAIP1	ns	*	***
RRM2B	TP53AIP1	ns	**	***
p21-5'	PMAIP1	ns	ns	***
p21-5'	TP53AIP1	ns	ns	***
PMAIP1	TP53AIP1	ns	ns	**

C. Transactivation with 2 ng p53 plasmid as a function of time

Time		12 h	18 h	24 h
ANOVA (F, P value)		152, <0.0001	903, <0.0001	8110, <0.0001
Comparison between		**P values ***		
CCNG1	RRM2B	***	***	***
CCNG1	p21-5'	***	***	***
CCNG1	PMAIP1	***	***	***
CCNG1	TP53AIP1	***	***	***
RRM2B	p21-5'	*	***	***
RRM2B	PMAIP1	**	***	***
RRM2B	TP53AIP1	***	***	***
p21-5'	PMAIP1	ns	ns	***
p21-5'	TP53AIP1	***	***	***
PMAIP1	TP53AIP1	***	***	***

[*] p values are marked as follows: $p > 0.05$ ns, $p \leq 0.05$ *, $p \leq 0.01$ **, $p \leq 0.001$ ***

Table S5. Primers used for incorporation of p53 REs and flanking sequences into pCluc Mini-TK plasmids

Gene symbol	Primer direction	Sequence (5'->3') *
RRM2B	Forward	**AGGCATGTCT**tccaaGGATCCTTCGCATATTAAGGTGAC
	Reverse	**GGGCATGTCA**ggacaGAATTCCAAGATCTCCCGATCCG
CCNG1	Forward	**AGGCTAGTCC**gaggcGGATCCTTCGCATATTAAGGTGAC
	Reverse	**GGGCTTGTGC**tcacaGAATTCCAAGATCTCCCGATCCG
p21-5'	Forward	**CAACATGTTG**agctcGGATCCTTCGCATATTAAGGTGAC
	Reverse	**GGACATGTTC**ctgacGAATTCCAAGATCTCCCGATCCG
PMAIP1	Forward	**GGGCAGGTCG**cgctcGGATCCTTCGCATATTAAGGTGAC
	Reverse	**GGACACGCTC**ccgacGAATTCCAAGATCTCCCGATCCG
TP53AIP1	Forward	**GGGCTTGTCG**agatgGGATCCTTCGCATATTAAGGTGAC
	Reverse	**GGGCAAGAGA**ggaggGAATTCCAAGATCTCCCGATCCG

* Capital bold letters are the sequences of natural p53 REs. Lower letters are the natural flanking sequences. Underlined capital letters are the plasmid-specific annealing sequences.

Dataset S1 (separate file). p53 genes used in the analysis. The Excel file contains the details on the studied genes, their REs, their specific function and the pathways initiated by the specific function, as well as the final functional outcome of their activation, as defined in the main text. Also given are the deformability values (V(B) $°^3Å^3$) of the REs and the Pubmed ID of the references supporting the data on each gene. The list includes REs that directly activates p53, do not include spacer sequences between the two half sites, and are found in the promoter on the first intron/exon of the nearby gene.

SI References
1. Zhang Y & Crothers DM (2003) High-throughput approach for detection of DNA bending and flexibility based on cyclization. *Proc. Natl. Acad. Sci. USA* 100:3161-3166.
2. Beno I, Rosenthal K, Levitine M, Shaulov L, & Haran TE (2011) Sequence-dependent cooperative binding of p53 to DNA targets and its relationship to the structural properties of the DNA targets. *Nucleic Acids Res.* 39:1919-1932.
3. Vyas P*, et al.* (2017) Diverse p53/DNA binding modes expand the repertoire of p53 response elements. *Proc. Natl. Acad. Sci. USA* 114:10624-10629.
4. Senitzki A*, et al.* (2021) The complex architecture of p53 binding sites. *Nucleic Acids Res.* 49:1364–1382.
5. Zhang Y & Crothers DM (2003) Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys. J.* 84:136-153.
6. Berg OG & von Hippel PH (1987) Selection of DNA binding sites by regulatory proteins. Statistical- mechanical theory and application to operators and promoters. *J. Mol. Biol.* 193:723-750.
7. Berg OG & von Hippel PH (1988) Selection of DNA binding sites by regulatory proteins. II. The binding specificity of cyclic AMP receptor protein to recognition sites. *J. Mol. Biol.* 200:709-723.
8. Balasubramanian S, Xu F, & Olson WK (2009) DNA sequence-directed organization of chromatin: structure-based computational analysis of nucleosome-binding sequences. *Biophys. J.* 96:2245-2260.
9. Zhang Y, Xi Z, Hegde RS, Shakked Z, & Crothers DM (2004) Predicting indirect readout effects in protein-DNA interactions. *Proc. Natl. Acad. Sci. USA* 101:8337-8341.