DNA dynamics dictates p53 functional outcome

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

Faculty of Biology, Technion, Technion City, Haifa, Israel

[1] present address: Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel

* To whom correspondence should be addressed.
Email address: bitali@technion.ac.il
Phone: 972-4-8293767
Fax: 972-4-8225153

ABSTRACT

The tumor suppressor protein p53 is situated in the midst of a complex cellular network that is activated in response to cellular stress. Activated p53 functions mainly as a transcription factor, regulating the expression of numerous genes involve in various cellular pathways critical for preventing cancer, and in pathways unrelated to cancer surveillance. An unresolved question in the field is how p53 is able to parse its myriad functions in response to the severity of the stress signal and consequently to coordinate the functional outcome in a timely manner. We have previously shown that DNA torsional flexibility distinguishes between different p53 response elements (REs). Here we show across the genome that p53 target genes belonging to pathways acting early in the stress response (e.g., DNA damage response and innate immunity) have REs that are significantly more flexible than REs of genes involved in pathways that need to be more strictly regulated, or that their functional outcome occurs later in the response to stress (e.g., intrinsic apoptosis and p53 negative regulation). We validated these statistical findings by several complementary experimental approaches, *in vitro* and in cells, for six p53 REs belonging to pathways that operate at different times post p53 induction. Our results clearly demonstrate that the flexibility of p53 REs contributes significantly to the temporal expression of p53 target genes and thereby to life versus death decisions in the p53 system.

INTRODUCTION

Gene regulation depends on the correct spatiotemporal arrangements of transcription factors and their specific DNA response elements (REs). It is now commonly accepted that recognition of DNA by regulatory proteins occurs not only by direct hydrogen bonds from the protein side chains to the DNA bases, termed direct, or base, readout (1), but is complemented by direct hydrogen bonds to the DNA backbone, termed indirect readout (2), and by recognition of structural and dynamical (flexibility) properties of DNA REs, termed shape readout (3). The ability of DNA to be deformed upon interaction with proteins ("deformability") is sequence-dependent (4,5).

The tumor suppressor protein p53 is part of a complex cellular network that is activated in response to various cellular stress signals. Activated p53 functions mainly as a transcription factor, regulating the expression of numerous genes, involved not only in various cellular pathways critical for preventing cancer, but also in pathways unrelated to cancer surveillance (6,7). p53 is currently viewed more as the "guardian of homeostasis", rather than the "guardian of the genome" (8). It is now appreciated that p53 respond both to acute DNA damage and other life and death decisions, as well as to more "routine" activities, such as energy metabolism (9-12) and embryonic development (13). In most of its activities p53 has both positive as well as negative outcomes, such as apoptosis and anti-apoptosis (14-16).

A major question in the field is how p53 "chooses" the appropriate response to the acute or routine stress signals. Broadly, two alternative models were proposed (17,18). *The selective binding model* focusses on inherent differences in p53 REs as a central driving force for discrimination between p53 target genes. *The selective context model*, suggests that p53 is not the crucial force behind target gene selectivity, but that discrimination is achieved by other events at the cellular and genomic level. Within the selective binding model, readout by base, indirect and shape were all shown to contribute to the interaction of p53 with its target promoters (19-27). In particular, we have previously demonstrated, using cyclization kinetics of DNA minicircles experiments, that DNA torsional (twist) flexibility changes significantly between p53 REs that differ in the uncontracted CWWG (W = A or T) motif at the center of each p53 half-site (20). The CATG motif, that has higher torsional flexibility, was shown to be bound with high affinity and low cooperativity by p53, whereas REs with the more rigid CTAG and CAAG motifs bound p53 with lower affinity and high cooperativity. Moreover, we have shown that when p53 levels are low, the torsional flexibility of p53 REs positively correlates with transactivation levels (21). On the other hand, when p53 levels are highly induced, transactivation levels positively correlated with p53 binding affinity (21). Using binding

stability measurements (*in vitro* kinetic off rates), we established that high torsional flexibility leads to more kinetically stable p53/RE complexes (21). Thus, at basal p53 levels, transactivation is correlated with kinetic stability of p53/RE complexes.

The torsionally flexible CATG motif is often observed in binding sites of cell-cycle arrest genes, whereas the rigid CAAG and CTAG motifs are more abundant in apoptosis related genes (21,28,29). Here, we hypothesize that across the genome p53-dependent genes with flexible REs will belong to outcome categories in which there is a need for an early and immediate response upon stress. Rigid REs on the other hand, will occur in functional categories where the response to stress is irreversible and likely to be more strictly regulated, or that it should occur at late stages in the response to stress. We show that this is indeed the case for all known p53 dependent genes that are activated by direct binding of p53 to spacer-less REs. Moreover, we prove this hypothesis using several complementary experimental approaches, *in vitro* and in cells, for six p53 REs belonging to different outcome categories that function at different times after p53 induction. Thus, the DNA torsional flexibility of p53 RE is important for singling out the time-wise appropriate target site among the numerous p53 dependent promoters.

Materials and Methods

Cells, plasmids and reagents

H1299 human non-small lung carcinoma cells (p53-null cells) were obtained from ATTC (CRL-5803). Cells were thawed and propagated in RPMI-1640 with 10% Foetal Bovine Serum (FBS), 2 mM L-glutamine, and 1% Pen-Strep solution (Biological Industries). All experiments were performed in this medium. Cells were maintained in a humidified incubator at 37 °C and 5% CO_2. Cells were routinely checked for mycoplasma infection and tested negative.

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

Generation of reporter plasmids containing p53 - REs

All DNA sequences were synthesized by Sigma Genosys (Israel) and purified by standard desalting. The sequences corresponding to p53 REs were incorporated in 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. DNA Primers used for incorporation are found in Table S6. Overexpressed plasmids were extracted using QIAprep® Spin Miniprep Kit (Qiagen), according to manufacturer's protocol. The identity of incorporated sequence were verified using Sanger sequencing.

Reporter gene assays

H1299 cells were seeded in 24-well plates at a density of $5X10^4$ and incubated for 24 h. Then cells were co-transfected with 800 ng of either pCLuc Mini-TK 2 harboring p53 RE or a corresponding empty vector along with 50 ng of pCMV-GLuc and 0 to 75 ng (0, 1, 2, 5, 10, 25, 75 ng) of p53 expression vector (pC53-SN3). DNA concentration was adjusted using pcDNA 3.1 plasmid. Transfection was performed with jetPEI (Polyplus) reagent according to the manufacturer's recommendation. Transfection efficiency was monitored by transfecting separate wells with 0.1 μg eGFP along with Gluc and pcDNA plasmids. An aliquot of growth medium (20 μl) was taken at various time points post transfection and luminescence was measured using Pierce™ Cypridina Luciferase Flash Assay Kit and Pierce™ Gaussia Luciferase Flash Assay Kit (both from Thermo Fisher) according to the manufacturer's protocol. Signal was detected by CLARIOstar 96 well plate reader (BMG LABTECH). Four to six independent experiments were carried out, each containing two technical replicates.

Analysis of transactivation from p53 REs

To correct for transfection variability, all results from Cypridina Luciferase luminescence were normalized to Gaussia Luciferase luminescence. Next, the results were normalized in two steps. First, the results obtained with pCLuc Mini-TK 2 harboring p53RE ($Cluc^{+RE}$) were divided by the results obtained with the corresponding empty vector ($Cluc^{-RE}$), Second, the results obtained with p53 (1-75ng of pC53-SN3) were divided by the results obtained without p53 (0 ng of pC53-SN3). Fold increase in transactivation is thus given by the equation:

Fold increase in transactivation= $(Cluc^{+RE}/Cluc^{-RE})^{+p53}/(Cluc+RE/Cluc-RE)^{-p53}$

p53 protein quantitation and detection

Cells were lysed in RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% Na deoxychlate, 0.1% SDS, 25 mM Tris-Cl, pH 7.5) containing 1/100 (Vol/Vol) Protease inhibitor cocktail. Equal amount of protein from the different samples were resolved on 10% SDS PAGE, and transferred to

PVDF membrane. The blots were subsequently incubated with primary antibodies against p53 (DO-7, Cell Signaling) and β-Actin (13E5, Cell Signaling). Membranes were then reacted with a secondary antibody (Anti-mouse IgG, and Anti-rabbit IgG) conjugated to Horseradish Peroxidase (HRP), developed using Western Bright™ enhanced chemiluminescence reagent (Advansta), air-dried, and exposed using Fusion Pulse (Vilber) detection system.

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 (20,27,30,31), except that here we used as test sequences DNA sequences containing the full 20-bp long p53 sites plus 5-bp flanking sequences on either side, specific for each natural p53 site (See Table S5 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 (20,27,31). 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). 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 (20,27,31,32).

Functional annotation of p53 genes and clustering into pathways

We previously assembled a set of 250 p53 REs without spacers that contain sites validated for binding in cellular context and for p53-dependent gene expression (27) that belong to 235 genes (some genes contain multiple REs). We located each gene in the Genecards suite (https://www.genecards.org/) by their Entrez ID, gene symbol and gene name. We then looked for existing functional annotations in UniProtKB/Swiss-Prot (https://www.uniprot.org), extracted from the Gene Ontology (GO) project, derived from the 'Biological process' and 'Molecular function' categories, but only when a relevant publication was available. We then looked for more recent publications for each gene product. All our annotations are thus

supported by individual publications, given in Table S1 by their Pubmed ID (PMID) codes. For each gene we looked for its specific function, the biological pathway or process involved, and the functional outcome at the end of the biological cascade initiated by the specific gene.

Other computational and statistical tests

Information content per DNA binding site (Rseq) was calculated separately for each p53 outcome category, using the method of Schneider *et al.* (33), with the correction for small sample size suggested in this study. Consensus sequence patterns ("logos") were created using Weblogo 3.7.4., based on (34,35). Testing the strength of relationships between variables was carried out by Spearman's rank correlation as a non-parametric measure of correlation. In these tests we used ties whenever the values were within ±SEM of each other. In testing for statistically significant differences between functional groups clustered by their main p53 dependent functional outcome, as well as for differences in transactivation level of the studied REs, we used one-way analysis of variance (ANOVA). We used the Steel-Dwass method as a post-hoc test, to determine the groups that are significantly differed from each other by their deformability, because this test is adjusted for multiple comparisons, and the Tukey HSD (α = 0.05) as a post-hoc test, for the differences between transactivation levels. All analyzes were conducted using JMP®, Version 15 (SAS Institute Inc., Cary, NC, 2019). The values for the deformability (V(B) in units of $°^3Å^3$) of each DNA base-pair step were calculated as previously described (27), using the values of (36).

RESULTS AND DISCUSSION

Classification of p53 REs by their functional categories yield distinct structural signature

We have previously compiled a list of 250 p53 REs, validated for binding in cellular context and for p53-dependent gene expression (27). These REs belong to 235 genes that are solely activating, do not contain spacer sequences between the two half sites, and are found within the promoter or the first intron/exon region (27). To classify these p53 genes by their cellular function, we looked for their specific function, the biological pathway (or process) initialized by them, and the functional outcome at the end of their biological pathway (Table S1). Of the 235 genes, functions were assigned to 219 genes, where some of the genes turned out to have more than one end outcome. Still, some important functional categories were too small for a meaningful statistical analysis. Hence, we added four more genes whose p53 REs reside in the second intron of the gene. Thus, we ended up with 223 genes and 240 p53 dependent REs. To assess the flexibility properties of these REs we used a parameter termed "deformability" that denotes the volume of conformational space within common energy contours (36,37). We previously observed that this rigorous measure of helical flexibility, derived from crystal structures of protein-DNA complexes, has a strong and significant correlation with twist fluctuations (indicating twist flexibility) of p53 REs, measured experimentally using cyclization kinetics assays (27). In analyzing the clustering of functional categories by their deformability values (Table S1), we took only one RE per gene, also in cases where p53 promoters contain multiple REs, because transactivation produces just a single mRNA molecule and a single outcome, even when multiple p53 tetramers are bound to the promoter. p53 clusters (i.e., p53 REs that contain more than two half sites) are known to confer additional binding affinity and increase transactivation levels from the nearby gene (29,38). We chose to include only the most flexible full-site RE of any given cluster, based on our previous study, where we showed that there is a weak but significant correlation between RE conservation and deformability (27). We therefore hypothesize that the more conserved full-site RE (i.e., RE with sequence that is closer to the consensus sequence) is the primary RE of a cluster of REs, and hence will be bound first. Three genes contained two partially overlapping REs consisting of three abutting half sites, where the middle half site is common to both REs (LIF, PRKAB1 and RRAD genes). In these three cases we used the calculated deformability of the complete 30-bp RE, because of ambiguity in assigning the primary RE (Table S1).

A box plot of the deformability of p53 full site REs versus functional outcome of p53 activation, shows substantial differences in mean deformability values between functional categories (Figure 1). The data used for Figure 1 is for functional categories having at least

eight genes in each group, to minimize noise and scatter from small groups that do not have sufficient number of genes to form a cohesive category. The complete list of deformability values and their mean per functional category is shown in Table S2. Results show that genes belonging to p53 positive regulation category (mean V(B) = 2.34 ± $°^3Å^3$) and genes belonging to survival and proliferation categories - angiogenesis (mean V(B) = 2.36 ± 0.06$°^3Å^3$), autophagy (mean V(B) = 2.37 ± 0.08$°^3Å^3$), cell-cycle arrest (mean V(B) = 2.36 ± 0.03$°^3Å^3$), development (mean V(B) = 2.54 ± 0.03$°^3Å^3$), DNA damage response (mean V(B) = 2.48 ± 0.03$°^3Å^3$), energy metabolism (mean V(B) = 2.48 ± 0.04$°^3Å^3$), and innate immunity (mean V(B) = 2.49 ± 0.05$°^3Å^3$), show higher than average deformability values. Remarkably, the last four groups, those that are needed for establishing a fast and timely response to cellular needs, have the highest deformability. On the other hand, genes belonging to death-related outcome (intrinsic apoptosis, mean V(B) = 1.96 ± 0.03$°^3Å^3$), to p53 negative regulation (mean V(B) = 1.96 ± 0.06$°^3Å^3$), and to a lesser extent, to cell adhesion (including both cell-cell and cell-ECM adhesion genes, mean V(B) = 2.13 ± 0.03$°^3Å^3$,) and to cytoskeleton organization (mean V(B) = 2.07 ± 0.04$°^3Å^3$), have very low mean deformability values. Other categories included in Figure 1 and Table S2 have average deformability values.

The "survival" category (Figure 1 and Table S2) comprise p53-dependent genes that inhibit death, such as anti-apoptosis, anti-necrotic and anti-pyroptosis genes, but nonetheless its deformability is slightly below average V(B) values (mean V(B) = 2.19 ± 0.04$°^3Å^3$). Note that the extrinsic (death receptor) apoptosis category, though it is death-related, has a slightly above average mean deformability (mean V(B) = 2.28 ± 0.03$°^3Å^3$). This is because the extrinsic apoptosis pathway is initiated by signals originating from outside the cell, such as natural killer lymphocytes, or cytotoxic T lymphocytes, and thus are intimately connected with the innate and adaptive immune systems (39). Interestingly, the REs of the autophagy category have deformability values that are either very high or very low, as this group consists of two sub-groups that originate from two different linkages. Autophagy genes that have REs with high V(B) values (2.59 ± 0.07$°^3Å^3$) are linked to autophagy that comes together with the activation of innate immunity genes (SPATA18, TNFAIP8, TRIM5, and TRIM22), or DNA damage response genes (AEN), or alternatively, connected with autophagy machinery genes (ATG4A, PLEKHF1). Autophagy genes having REs with low V(B) values (2.11 ± 0.08$°^3Å^3$) are linked with autophagy that come together with apoptosis-related events (Table S2).

To assess the statistical significance of the differences in deformability between functional categories, we carried out one-way analysis of variance on functional categories that contain at least eight genes each (178 genes, 18 categories, 223 functions). The F ratio obtained

from the analysis was significant (F = 15.67, P = <0.0001). Thus, we can reject the null hypothesis that the groups are sampled from populations with identical distributions. A post-hoc test (Steel-Dwass method), established that there was a significant difference in mean deformability between many functional groups (35 relationships out of 153 possible ones, Table S3), the largest being between development and cell adhesion (Z = 4.998, P <0.0001). Of note to p53 functions as a tumor suppressor protein, are the differences between intrinsic apoptosis and cell-cycle arrest (Z = -3.88, P = 0.013), intrinsic apoptosis and DNA damage response (Z = -4.00, P = 0.008), as well as between intrinsic and extrinsic apoptosis (Z = -3.82, P = 0.016). DNA repair, on the other hand, is one of the groups with high variations in deformability within the group, and thus its average does not differ significantly from intrinsic apoptosis (Z = -3.24, P = 0.11) or other categories. This is due mainly to the XPC RE. XPC functions also as a negative p53 regulator (40). The difference between innate immunity (mean V(B) = 2.49 ± 0.05°^{3}Å^3, 8 genes) and adaptive immunity (mean V(B) = 2.08 ± 0.03°^{3}Å^3, 4 genes) should also be noted.

Functional outcome categories have a unique sequence signature

Sequence logos are a convenient way to depict base-pair preferences and information content of DNA binding sites (33-35). We used sequence logos to assess DNA binding preferences of p53-dependent functional groups. Here again, we analyzed the data by taking one RE per gene, except for the three sites containing three abutting half sites, mentioned above, where we included both sites in the analysis (i.e., once where the common middle half site is the 5' half-site of RE1, and once where it is the 3' half-site of RE2). Figure 2 shows representative patterns for selected functional groups, created by Weblogo (34,35), displaying dramatic changes between them, in all positions (see Figure S1 for the pattern of all other groups), and especially relative to the degenerate pattern of all 250 REs together, taken from our previous study (27). To quantitate the patterns observed in the sequence logos, we calculated Rseq for each pattern (Table S4). Rseq is defined as the total decrease in uncertainty (or the decrease in maximal entropy) per DNA binding site and is a measure for the information content of DNA binding sites (33). One cannot easily compare the information content of groups differing in their number of sequences, especially when only limited sequence data exists, because limited data results in underestimation of the entropy per DNA site. However, a correction to Rseq calculations, for errors resulting from small sample size (i.e., small number of DNA sequences per group), is known to minimize this bias (33) and is also the default option in Weblogo. This limitation does not apply however to the use of the ratio of information content of partial regions

of the site, normalized by the information content of the full site. When we carry out such normalization using the ratio of Rseq of CWWG regions to that of the whole site, within a specific outcome category, we can divide the logos to those in which the information content is high mainly in the two CWWG motifs, versus those in which it is high mainly in all other regions except CWWG (Figure 2). The CWWG motifs are uncontacted by p53, and hence it is where shape readout dominates (20). In the present study these are the sites with particularly high information content, in one or both half sites of the following categories: development (71% of total Rseq is in the CWWG regions), DNA damage response (74%), energy metabolism (73%), innate immunity (70%), and to some extend cell-cycle arrest (69%). These are the outcome categories where we observed high average deformability of their REs. This may mean that high information content in shape readout positions comes together with high deformability, and both occur in categories that are known to lead to immediate response upon stress. Indeed, Rseq of the two CWWG regions together is rank correlated to the deformability of full-sites REs ($\rho = 0.53$, $P = 0.02$ for 18 categories having at least eight genes each). Other sites, such as survival (anti-death category), and membrane trafficking, seem to have moderately high information content in the two CWWG groups (67% each), but in the case of survival, it is high also in non-CWWG regions, and thus survival has a high overall information content (16.3 ± 0.8 bits). Membrane trafficking (and to some extent also cell-cycle arrest) may suffer from the lack of sufficient number of genes in this category (eight genes only each). The opposite pattern is observed in the sites with high information content in all positions except CWWG, such as p53 negative regulation and extrinsic apoptosis (48% of the total Rseq is in CWWG regions). Nonetheless, these two categories have high overall information content (16.3 ± 1 and 19.0 ± 0.9, respectively). In p53 negative regulation the 9th and 12th position are more conserved than those of the highly conserved C and G at the 4th,7th, and 17th positions. This may mean that in these two categories the driving force for p53-DNA interaction is mainly the recognition of specific protein-DNA bases contacts, as these positions are known to be directly contacted by p53 in p53-DNA complexes, (though some indirect readout is observed there as well, 19,22). Last, there are functional groups in which the information content is low at all positions, such as p53 positive regulation, which is highly variable at all but the G at the 7th position, which may mean that the functional outcome of this category have diverse origin and sources.

Experimental validation of the flexibility of six natural p53 REs by cyclization kinetics assays

We used deformability to calculate the flexibility properties of our curated p53 activating REs dataset, above and before (27), because this parameter has a strong and significant correlation to experimental torsional flexibility of p53 sites, measured by cyclization kinetics of DNA minicircles (27). Here, we present the experimental data on six well-known p53 REs, belonging to several functional outcome categories, that were included in the data used to assess this correlation. For the current study, we used a single full-site p53 binding site as the functional unit for interaction with p53. The 20-bp full binding sites were embedded within the 30-bp test sequences together with five base pairs flanking the full-site from each side. The flanking sequences are the REs natural flanking sequences (Table S5). This is because we have previously shown that sequences flanking p53 REs modulate p53 binding and transactivation (27). However, the global structural characteristics of the sites, presented in Table 1 and Figure S2, are those of the central 20-bp p53 sites without flanks. The chosen REs belong to genes that function at different cellular pathways, such as DNA damage response and DNA repair (RRM2B), cell cycle arrest (p21-5', CCNG1), intrinsic apoptosis (TP53AIP1, PMAIP1), and mTOR inhibition, as well as p53 negative regulation (DDIT4). These REs span the entire range of deformability values of p53 REs (Table S1). The DDIT4 site, chosen for all experiments carried out in this study, is the less deformable site of the two p53 REs in the promoter of the DDIT4 gene. This site was chosen because it is the least deformable p53 site in our entire p53 activating p53 dataset. Table 1 shows that, as in our previous studies on p53 binding sites (20,27,31), the only significant change in global structural parameters of these sites is in the twist (torsional) flexibility of the DNA binding sites. More importantly, CCNG1 is the most flexible site from the set of six sites chosen for experimental study, followed by RRM2B and p21-5'. At the bottom of the list are TP53AIP1 and DDIT4. This validates our hypothesis that p53 REs belonging to DNA damage response and cell cycle arrest are more flexible than those belonging to intrinsic apoptosis and negative p53 regulation, as well as further validating the method used in assessing the dynamical characteristics of the RE set across the genome by the deformability parameter.

Reporter gene assays as a function of p53 levels

To validate our hypothesis in cellular context we carried out luciferase reporter gene assays in human non-small lung carcinoma cell line null for p53 (H1299), expressing ectopic p53 along with reporter plasmids, each carrying a distinct RE. First, we ascertained that p53 ectopic

expression is linearly correlated with the amount of transfected p53 expression plasmid pC53-SN3. This was done by transfecting cells with increasing amounts of p53 expression vector and a reporter construct harboring the p21-5' RE and immunoblotting the cell lysate (Figure S3A). Densitometry analysis of the observed p53 band, normalized to the β-Actin band, confirmed the linear relationship between p53 protein level and pC53-SN3 amount (R = 0.998, Figure S3B). We then carried out reporter gene assays as a function of p53 levels for the six well-known p53 REs described above. We chose to analyze transactivation by reporter gene assays 48 h after transfection, to enable observing a significant signal also in cells transfected with the lowest p53 plasmid amount (1 ng p53 plasmid). As can be observed from Figure 3, in cells expressing low levels of p53, transactivation from flexible REs (RRM2B, p21-5' and CCNG1) was higher compared to transactivation from the more rigid REs (PMAIP1, TP53AIP1, and DDIT4). ANOVA analysis revealed that transactivation levels of the REs are significantly different from each other (F = 6.78, P = 4 X10^{-3} for cells transfected with 1 ng p53 plasmid and F = 21.59, P = 8.16 X 10^{-7} for 2 ng p53 plasmid). Post hoc analysis (Tukey HSD) revealed that in cells transfected with 1 ng p53 plasmid, transactivation level from the p21-5' RE was significantly higher than that from PMAIP1, TP53AIP1 and DDIT4 REs. At 2ng p53 plasmid, transactivation level from the RRM2B RE, or the CCNG1 RE, is significantly higher than that from PMAIP1, TP53AIP1 and DDIT4 REs, whereas transactivation level from p21-5' RE is significantly higher only to those from the TP53AIP1 and DDIT4 REs. At the highest amount of p53 (75 ng of pC53-SN3) all REs except DDIT4 exhibit similar levels of transactivation (Figure S4). The transactivation from cells transfected with 2 ng p53 plasmid is strongly and significantly correlated with the experimental twist fluctuation of the six studies REs from cyclization kinetics assays (ρ = 0.986, P = 3.1 X 10^{-4}), as we expected based on our earlier study (21). Moreover, we calculated the difference between transactivation from the three most flexible REs (corresponding to DNA damage response, cell cycle arrest and DNA repair categories) and the three most rigid REs (corresponding to intrinsic apoptosis and p53 negative regulation categories) and found that the two groups differ significantly from each other at both p53 levels (t = 3.67, P = 0.021 and t = 3.88, P = 0.018 at 1 ng and 2 ng p53 plasmid, respectively).

Reporter gene assays as a function of time

The higher transactivation level observed 48 h post transfection from flexible REs relative to rigid REs could be due to an earlier start of transactivation, or alternatively due to faster transactivation rates. To distinguish between these possibilities we examined transactivation

levels, at 25 ng p53 plasmid (to enable observation of a significant signal also at shorter times), at additional time points (12, 18, and 24 h post transfection, Figure 4). Two factorial analysis revealed that there is a significant difference between the transactivation level of the six REs at the four time points post transfection. ANOVA analysis at 12 h post transfection revealed the small but significant difference between the transactivation level of the six REs (F = 3.16, P = 0.03). A post-hoc analysis indicated that transactivation from CCNG1 is significantly different from DDIT4, and that this is the only significant difference among the six studied REs (Tukey HSD P = 0.027). At 18 h post infection, post-hoc analysis indicated that transactivation from RRM2B is significantly different from DDIT4 (F = 4.84, P = 0.006, Tukey HSD P = 0.0018). Only at 24 h post infection transactivation from most of the studied REs were significantly different from each other (F = 35.75, P = 9.75 X 10^{-9}), and similar to the differences at 48 h. At 24 h TP53AIP1 and DDIT4 were both significantly different from CCNG1 RRM2B and p21-5', whereas PMAIP1 was different only from RRM2B and p21-5'. Spearman rank correlation between p53 dependent transactivation level from the six studied REs 24 h post transfection and their experimental twist fluctuations (indicating twist flexibility), established a positive strong and significant correlation between these two parameters (ρ = 0.96, P = 0.0029). The difference between transactivation from the three most flexible REs and the three most rigid REs differ significantly from each other at all the time points, however, the most significant differences were observed at 24 h post transfection (t = 3.31, P = 0.03).

mRNA level from endogenic p53-dependent genes are correlated with p53 REs deformability and with their functional outcome

In the gene reporter system used in this study, p53 REs transactivation level is monitored from a transiently transfected plasmid. To be able to mimic a more natural environment where transcription is affected by multiple other factors (e.g., chromatin organization and the presence of other transcription factors), we searched in the literature for data on mRNA levels of cellular genes in response to endogenous p53. Studies by the group of Lahav and colleagues have introduced the notion that different stress levels evoke different dynamic patterns of p53 expression (41-46). Low stress level induces a pulsatile pattern of expression (the level of p53 expression exhibit a series of pulses with fixed amplitude and frequency), whereas more severe stress induces sustained pattern of expression (where only one continuous peak of p53 expression is observed). Moreover, Purvis *et al*. (46) demonstrated that the sustained pattern can be mimicked by gradual addition of a specific inhibitor of MDM2, Nutlin-3 (47). In the study by Purvis *et al.* (46) the mRNA level of a large set of p53-dependent genes was measured

by qPCR, every hour up to 24 h post stress induction, for both pulsatile and sustained p53-expression pattern. To concur with our hypothesis, we took from this study only genes that p53 directly activates (27). We also deleted REs that contain long (> 9 bp) spacer sequences between the two half-sites, because p53 REs that harbor long spacer sequences can bind concurrently also in the hemi-specific mode of interaction (31). We then looked whether there is a correlation between mRNA level 1h post stress induction and deformability of the REs involved. Figure 5 and Figure S5 show that a strong and significant positive correlation exist between mRNA level 1 h post stress induction and deformability of the REs, for both the pulsatile ($\rho = 0.88$, $P = 8.4 \times 10^{-5}$) and the sustained mode of p53 expression ($\rho = 0.85$, $P = 2.6 \times 10^{-4}$). No correlation was observed at times later than 1 h. It is not surprising that the same correlation is observed for both pulsatile and sustained mode of p53 expression, since at 1 h they are still very similar. It is however, interesting to note that not only is there a correlation between mRNA level and deformability of the individual REs, but that they cluster by their functional outcome. Highest deformability and mRNA level is observed for the two cell-cycle arrest genes (CDKN1A and GADD45A), then two REs that function both in autophagy as well as intrinsic apoptosis (BAX and PUMA), followed by REs to genes that function only in intrinsic apoptosis (PMAIP1 and TP53AIP1), followed by three REs to genes that function in p53 negative regulation (MDM2, PPM1D and XPC). These last two categories have the most rigid REs of all spacer-less p53 REs. The most rigid REs, and least amount of mRNA, are those for genes in which the two half sites are interrupted by spacer sequences, DDB1 (1-bp spacer, 48), and CASP6 (4-bp spacers, 49), with CASP6 being worse than DDB1, in agreement with our previous observation on binding of p53 to REs with spacers (31). Two additional REs appear in Figure 4, BCL6 belonging to adaptive immunity outcome, which we have not analyzed statistically because this group currently contain only four genes, and SESN1, belonging to oxidative stress outcome. The average deformability of this category is $2.13 \pm 0.071.94^{\circ 3}Å^3$, but SESN1 is an outlying REs within this category, much lower than average ($V(B) = 1.94^{\circ 3}Å^3$).

GENERAL DISCUSSSION

Temporal expression of p53 target genes govern cell-fate decisions. An important question in the field is how p53 discriminates between the myriad p53-dependent target genes. Whereas the importance of differential binding affinities of p53 to its REs (28), the dependence of cell-fate decisions on p53 levels (50-52), and on the elapsed time post p53 induction (51) are widely acknowledged, the mechanistic basis for p53-dependent temporal gene expression is currently unknown. In our previous studies we showed experimentally that the torsional flexibility of the

central CWWG motif is the major structural change between p53 REs (20), and that at high p53 levels transactivation from a given RE is determined by binding affinity, whereas at basal p53 levels is it regulated by p53 binding kinetics, governed by DNA flexibility of the RE (21). These mechanistic observations support the selective binding model, as a major contributor to target gene selectivity in the p53 system, but by themselves do not explain the unique manner by which p53 can temporally coordinate its myriad target genes, to counteract an acute stress signal, or a routine cellular demand. Using our previous observations on the strong correlation between calculated deformability and experimentally measured twist flexibility, from cyclization kinetics assays (27), we show here that DNA flexibility is a strong classifier of p53 functional outcome. Genes that are part of outcome categories in which a response is needed immediately, time-wise, have REs which are capable of transactivating the genes at a faster rate and at low levels of p53, because they are flexible. Thus, our results obtained *in vitro* and in cells clearly indicate that the information for the selective binding of p53 to DNA is intrinsically encoded within the base sequence of the p53 REs. Hence, one should conclude that more than "p53 is smart" (18), the DNA double helix where the REs reside is "smart"- the REs base sequences are information rich by virtue of their structural properties.

The dependence of the functional outcome of genes activation by p53 on RE flexibility can be illustrated by the cascade that begins with damage to DNA. Looking broadly on the events starting with damage to DNA, it is known that initial sensing of DNA damage is an immediate event, preceding cell-cycle arrest (CCA) (53,54). Thus, DNA damage response (DDR) should, according to our predictions, belong to a category with very flexible REs, as shown in this study (Figure 1 and Table S2). The deformability of DDR REs is however not statistically different as a group from that of CCA-related REs, mainly because these two groups share few genes (Table S2). Following CCA p53 regulates the attempts at repairing the damage to DNA, unless the damage is too severe to be repaired, which depends on the accumulated level of DNA damage (55). Intrinsic apoptosis should come after DNA repair attempts, and only if the repair process failed (56,57). This last step, being an irreversible step, is partly regulated by the high rigidity of the REs of apoptosis-related genes. Thus, the qualitative deformability series DDR>CCA>DNA repair>Intrinsic apoptosis follows the order of events in the cells upon DNA damage. Autophagy and differentiation are additional functional outcomes that can be activated by p53, as an alternative mechanism to eliminate damaged cells (58). Apoptosis-related autophagy (upper six REs in the autophagy category in Table S2) and differentiation genes have REs with deformability values in the range between those of DNA repair and intrinsic apoptosis. One should of course consider the spread of deformability values

within each functional category, and likewise in the temporal order of events that takes place within each functional outcome group. For example, RRM2B supplies nucleotides for the DNA repair process (59), thus it functions early in DNA repair. Indeed, it is the most deformable RE in the DNA repair category (Table S2), and one of the most torsionally flexible sequence, as determined from the cyclization kinetics measurements. However, within the whole category of DNA repair genes, it is an upper outlier (Figure 1 and Table S2).

Another illustration of the relationship between flexibility of REs and functional outcome is depicted in genes that up- and down-regulate p53. Here, genes that positively regulate p53 levels and stability at the start of a stress event have REs with higher torsional flexibility relative to those that down-regulate p53. Genes that down-regulate p53 possess more rigid REs, as these are transcribed towards the end of stress events when p53 levels should be decreased. Additional example, is the relationship between innate immunity – a fast and early event, with genes harboring REs that are very flexible, versus adaptive immunity, which is not an immediate event, and its genes have relatively rigid REs (Table S2).

p53 sequence logo is known to be degenerate, as depicted in Figure 2D, which is based on 250 validated activating REs that do not have spacers sequences between the half-sites (27). Here we show that when we cluster p53 REs by their functional outcome, the pattern becomes more information rich. Degeneracy in consensus sequence may be critical for gene regulation, as it allows for variable temporal and p53 level-dependent response (60). Indeed, we show here that p53 consensus sequence pattern is a mixture of different functional outcomes that have different sequence/structure requirements. Some genes lead to early response and some to late ones, some genes are activated at basal p53 levels and some at highly induced levels. Putting them together obscure the fine details of the pattern.

ACKNOWLEDGEMENT

We thank Prof. Varda Rotter (Weizmann Institute of Science) for the p53 expression plasmid, pC53-SN3.

FUNDING

This work was supported by the Israel Science Foundation (Grant #1517/14 to T.E.H.)

Conflict of interest statement. None declared

REFERENCES

1. Seeman, N.C., Rosenberg, J.M. and Rich, A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. USA*, **73**, 804-808.
2. Otwinowski, Z., Schevitz, R.W., Zhang, R.G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F. and Sigler, P.B. (1988) Crystal structure of *trp* repressor/operator complex at atomic resolution [published erratum appears in Nature 1988 Oct 27;335(6193):837]. *Nature*, **335**, 321-329.
3. Rohs, R., Jin, X., West, S.M., Joshi, R., Honig, B. and Mann, R.S. (2010) Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.*, **79**, 233-269.
4. Luisi, B. (1995) In Lilley, D. M. J. (ed.), *DNA-Protein: Structural interactions*. IRL Press, Oxford, pp. 1-48.
5. Travers, A.A. (1995) In Lilley, D. M. J. (ed.), *DNA-Protein: Structural interactions*. IRL Press, Oxford, pp. 49-75.
6. Vousden, K.H. and Prives, C. (2009) Blinded by the Light: The Growing Complexity of p53. *Cell*, **137**, 413-431.
7. Kastenhuber, E.R. and Lowe, S.W. (2017) Putting p53 in Context. *Cell*, **170**, 1062-1078.
8. Aylon, Y. and Oren, M. (2016) The Paradox of p53: What, How, and Why? *Cold Spring Harb. Perspect. Med.*, **6**, a026328.
9. Gottlieb, E. and Vousden, K.H. (2010) p53 regulation of metabolic pathways. *Cold Spring Harb. Perspect. Biol.*, **2**, a001040.
10. Liang, Y., Liu, J. and Feng, Z. (2013) The regulation of cellular metabolism by tumor suppressor p53. *Cell Biosci.*, **3**, 9.
11. Lacroix, M., Riscal, R., Arena, G., Linares, L.K. and Le Cam, L. (2020) Metabolic functions of the tumor suppressor p53: Implications in normal physiology, metabolic disorders, and cancer. *Mol. Metab.*, **33**, 2-22.
12. Lahalle, A., Lacroix, M., De Blasio, C., Cisse, M.Y., Linares, L.K. and Le Cam, L. (2021) The p53 Pathway and Metabolism: The Tree That Hides the Forest. *Cancers (Basel)*, **13**, 133.
13. Koifman, G., Aloni-Grinstein, R. and Rotter, V. (2019) p53 balances between tissue hierarchy and anarchy. *J. Mol. Cell. Biol.*, **11**, 553-563.
14. Portt, L., Norman, G., Clapp, C., Greenwood, M. and Greenwood, M.T. (2011) Anti-apoptosis and cell survival: a review. *Biochim. Biophys. Acta*, **1813**, 238-259.
15. Sheikh, M.S., Huang, Y., Fernandez-Salas, E.A., El-Deiry, W.S., Friess, H., Amundson, S., Yin, J., Meltzer, S.J., Holbrook, N.J. and Fornace, A.J., Jr. (1999) The antiapoptotic decoy receptor TRID/TRAIL-R3 is a p53-regulated DNA damage-inducible gene that is overexpressed in primary tumors of the gastrointestinal tract. *Oncogene*, **18**, 4153-4159.
16. Pan, G., Ni, J., Yu, G., Wei, Y.F. and Dixit, V.M. (1998) TRUNDD, a new member of the TRAIL receptor family that antagonizes TRAIL signalling. *FEBS Lett.*, **424**, 41-45.
17. Espinosa, J.M. (2008) Mechanisms of regulatory diversity within the p53 transcriptional network. *Oncogene*, **27**, 4013-4023.
18. Vousden, K.H. (2000) p53: death star. *Cell*, **103**, 691-694.
19. Kitayner, M., Rozenberg, H., Kessler, N., Rabinovich, D., Shaulov, L., Haran, T.E. and Shakked, Z. (2006) Structural basis of DNA recognition by p53 tetramers. *Mol. Cell*, **22** 741-753.
20. Beno, I., Rosenthal, K., Levitine, M., Shaulov, L. and Haran, T.E. (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.

21. Jordan, J.J., Menendez, D., Sharav, J., Beno, I., Rosenthal, K., Resnick, M.A. and Haran, T.E. (2012) Low-level p53 expression changes transactivation rules and reveals superactivating sequences. *Proc. Natl. Acad. Sci. USA*, **109**, 14387-14392.
22. Kitayner, M., Rozenberg, H., Rohs, R., Suad, O., Rabinovich, D., Honig, B. and Shakked, Z. (2010) Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs. *Nat. Struct. Mol. Biol.*, **17**, 423-429.
23. Chen, Y., Dey, R. and Chen, L. (2010) Crystal structure of the p53 core domain bound to a full consensus site as a self-assembled tetramer. *Structure*, **18**, 246-256.
24. Chen, Y., Zhang, X., Dantas Machado, A.C., Ding, Y., Chen, Z., Qin, P.Z., Rohs, R. and Chen, L. (2013) Structure of p53 binding to the BAX response element reveals DNA unwinding and compression to accommodate base-pair insertion. *Nucleic Acids Res.*, **41**, 8368-8376.
25. Zhang, X., Dantas Machado, A.C., Ding, Y., Chen, Y., Lu, Y., Duan, Y., Tham, K., Chen, L., Rohs, R. and Z., Q.P. (2014) Conformations of p53 Response Elements in Solution Deduced Using Site-Directed Spin Labeling and Monte Carlo Sampling. *Nucleic Acids Res.*, **42**, 2789–2797.
26. Golovenko, D., Brauning, B., Vyas, P., Haran, T.E., Rozenberg, H. and Shakked, Z. (2018) New Insights into the Role of DNA Shape on Its Recognition by p53 Proteins. *Structure*, **26**, 1-14.
27. Senitzki, A., Safieh, J., Sarma, V., Golovenko, D., Danin-Poleg, Y., Inga, A. and Haran, T.E. (2021) The complex architecture of p53 binding sites. *Nucleic Acids Res.*, **49**, 1364–1382.
28. Weinberg, R.L., Veprintsev, D.B., Bycroft, M. and Fersht, A.R. (2005) Comparative binding of p53 to its promoter and DNA recognition elements. *J. Mol. Biol.*, **348**, 589-596.
29. Riley, T., Sontag, E., Chen, P. and Levine, A. (2008) Transcriptional control of human p53-regulated genes. *Nat. Rev. Mol. Cell Biol.*, **9**, 402-412.
30. Zhang, Y. and Crothers, D.M. (2003) High-throughput approach for detection of DNA bending and flexibility based on cyclization. *Proc. Natl. Acad. Sci. USA*, **100**, 3161-3166.
31. Vyas, P., Beno, I., Xi, Z., Stein, Y., Golovenko, D., Kessler, N., Rotter, V., Shakked, Z. and Haran, T.E. (2017) Diverse p53/DNA binding modes expand the repertoire of p53 response elements. *Proc. Natl. Acad. Sci. USA*, **114**, 10624-10629.
32. Zhang, Y. and Crothers, D.M. (2003) Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys. J.*, **84**, 136-153.
33. Schneider, T.D., Stormo, G.D., Gold, L. and Ehrenfeucht, A. (1986) Information content of binding sites on nucleotide sequences. *J. Mol. Biol.*, **188**, 415-431.
34. Schneider, T.D. and Stephens, R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.*, **18**, 6097-6100.
35. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188-1190.
36. Balasubramanian, S., Xu, F. and Olson, W.K. (2009) DNA sequence-directed organization of chromatin: structure-based computational analysis of nucleosome-binding sequences. *Biophys. J.*, **96**, 2245-2260.
37. Olson, W.K., Gorin, A.A., Lu, X.J., Hock, L.M. and Zhurkin, V.B. (1998) DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA*, **95**, 11163-11168.
38. Bourdon, J.C., Deguin-Chambon, V., Lelong, J.C., Dessen, P., May, P., Debuire, B. and May, E. (1997) Further characterisation of the p53 responsive element--identification of new candidate genes for trans-activation by p53. *Oncogene*, **14**, 85-94.

39. Ricci, M.S. and El-Deiry, W.S. (2007) In Gewirtz, D. A., Holt, S. E. and Grant, S. (eds.), *Apoptosis, Senescence, and Cancer. Cancer Drug Discovery and Development series*. Humana Press.
40. Krzeszinski, J.Y., Choe, V., Shao, J., Bao, X., Cheng, H., Luo, S., Huo, K. and Rao, H. (2014) XPC promotes MDM2-mediated degradation of the p53 tumor suppressor. *Mol. Biol. Cell*, **25**, 213-221.
41. Hafner, A., Stewart-Ornstein, J., Purvis, J.E., Forrester, W.C., Bulyk, M.L. and Lahav, G. (2017) p53 pulses lead to distinct patterns of gene expression albeit similar DNA-binding dynamics. *Nat. Struct. Mo.l Biol.*, **24**, 840-847.
42. Batchelor, E., Loewer, A. and Lahav, G. (2009) The ups and downs of p53: understanding protein dynamics in single cells. *Nat. Rev. Cancer*, **9**, 371-377.
43. Batchelor, E., Mock, C.S., Bhan, I., Loewer, A. and Lahav, G. (2008) Recurrent initiation: a mechanism for triggering p53 pulses in response to DNA damage. *Mol. Cell*, **30**, 277-289.
44. Loewer, A., Batchelor, E., Gaglia, G. and Lahav, G. (2010) Basal dynamics of p53 reveal transcriptionally attenuated pulses in cycling cells. *Cell*, **142**, 89-100.
45. Batchelor, E., Loewer, A., Mock, C. and Lahav, G. (2011) Stimulus-dependent dynamics of p53 in single cells. *Mol. Syst. Biol.*, **7**, 488.
46. Purvis, J.E., Karhohs, K.W., Mock, C., Batchelor, E., Loewer, A. and Lahav, G. (2012) p53 dynamics control cell fate. *Science*, **336**, 1440-1444.
47. Vassilev, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C. *et al.* (2004) In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science*, **303**, 844-848.
48. Nguyen, T.T., Grimm, S.A., Bushel, P.R., Li, J., Li, Y., Bennett, B.D., Lavender, C.A., Ward, J.M., Fargo, D.C., Anderson, C.W. *et al.* (2018) Revealing a human p53 universe. *Nucleic Acids Res.*, **46**, 8153–8167.
49. Yang, C., Kaushal, V., Haun, R.S., Seth, R., Shah, S.V. and Kaushal, G.P. (2008) Transcriptional activation of caspase-6 and -7 genes by cisplatin-induced p53 and its functional significance in cisplatin nephrotoxicity. *Cell Death Differ.*, **15**, 530-544.
50. Chen, X., Ko, L.J., Jayaraman, L. and Prives, C. (1996) p53 levels, functional domains, and DNA damage determine the extent of the apoptotic response of tumor cells. *Genes Dev*, **10**, 2438-2451.
51. Zhao, R., Gish, K., Murphy, M., Yin, Y., Notterman, D., Hoffman, W.H., Tom, E., Mack, D.H. and Levine, A.J. (2000) Analysis of p53-regulated gene expression patterns using oligonucleotide arrays. *Genes Dev.*, **14**, 981-993.
52. Kracikova, M., Akiri, G., George, A., Sachidanandam, R. and Aaronson, S.A. (2013) A threshold mechanism mediates p53 cell fate decision between growth arrest and apoptosis. *Cell Death Differ.*, **20**, 576-588.
53. Niida, H. and Nakanishi, M. (2006) DNA damage checkpoints in mammals. *Mutagenesis*, **21**, 3-9.
54. Surova, O. and Zhivotovsky, B. (2013) Various modes of cell death induced by DNA damage. *Oncogene*, **32**, 3789-3797.
55. Offer, H., Erez, N., Zurer, I., Tang, X., Milyavsky, M., Goldfinger, N. and Rotter, V. (2002) The onset of p53-dependent DNA repair or apoptosis is determined by the level of accumulated damaged DNA. *Carcinogenesis*, **23**, 1025-1032.
56. Elmore, S. (2007) Apoptosis: a review of programmed cell death. *Toxicol. Pathol.*, **35**, 495-516.
57. Nowsheen, S. and Yang, E.S. (2012) The intersection between DNA damage response and cell death pathways. *Exp. Oncol.*, **34**, 243-254.

58. Helton, E.S. and Chen, X. (2007) p53 modulation of the DNA damage response. *J. Cell. Biochem.*, **100**, 883-896.
59. Yamaguchi, T., Matsuda, K., Sagiya, Y., Iwadate, M., Fujino, M.A., Nakamura, Y. and Arakawa, H. (2001) p53R2-dependent pathway for DNA synthesis in a p53-regulated cell cycle checkpoint. *Cancer Res.*, **61**, 8256-8262.
60. Hoh, J., Jin, S., Parrado, T., Edington, J., Levine, A.J. and Ott, J. (2002) The p53MH algorithm and its application in detecting p53-responsive genes. *Proc. Natl. Acad. Sci. USA*, **99**, 8467-8472.
61. Zhang, Y., Xi, Z., Hegde, R.S., Shakked, Z. and Crothers, D.M. (2004) Predicting indirect readout effects in protein-DNA interactions. *Proc. Natl. Acad. Sci. USA*, **101**, 8337-8341.

Table 1: Best-fit parameters from cyclization kinetics measurements for the sequences studied here

Functional outcome	Gene symbol	Sequence	Bend angle (°) [a,b]	Twist angle (°) [b]	Roll and Tilt fluctuation (°) [b]	Twist fluctuation (°) [b]
DNA damage response & DNA Repair	RRM2B	GAACATGTCCCAACATGTTG	-1.24 (0.11)	34.16 (0.11)	4.31 (0.12)	5.60 (0.12)
Cell cycle	CCNG1	TGACATGCCCAGGCATGTCT	-0.72 (0.09)	34.52 (0.08)	4.54 (0.09)	5.77 (0.09)
Cell cycle	p21-5’	GCACAAGCCCAGGCTAGTCC	-0.06 (0.08)	34.04 (0.09)	4.45 (0.08)	5.34 (0.08)
Apoptosis	PMAIP1	GAGCGTGTCCGGGCAGGTCG	-1.24 (0.15)	34.29 (0.15)	4.5 (0.14)	5.13 (0.15)
Apoptosis	TP53AIP1	TCTCTTGCCCGGGCTTGTCG	-1.02 (0.11)	33.96 (0.11)	4.20 (0.12)	4.91 (0.10)
mTOR inhibition & p53 negative regulation	DDIT4	AAACAAGTCTTTCCTTGATC	0.15 (0.13)	34.17 (012)	4.45 (0.12)	4.67 (0.13)
Range of values observed [c]			−10.0 to 2.4	34.04 to 34.85	4.30 to 6.40	3.9 to 5.80

[a] 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.
[b] Numbers in parenthesis are the simulation errors, calculated as described in ref. (32).
[c] Roll and tilt fluctuations are the thermal fluctuations of the roll and tilt angle between adjacent base pairs (30,32).
[d] Twist fluctuations are the thermal fluctuations of the twist angle between adjacent bases (30,32).
[e] These values denote the range of observed values for these parameters in other DNA sequences, determined by the same method and the same constructs used here (20,27,31,61).

Figures:

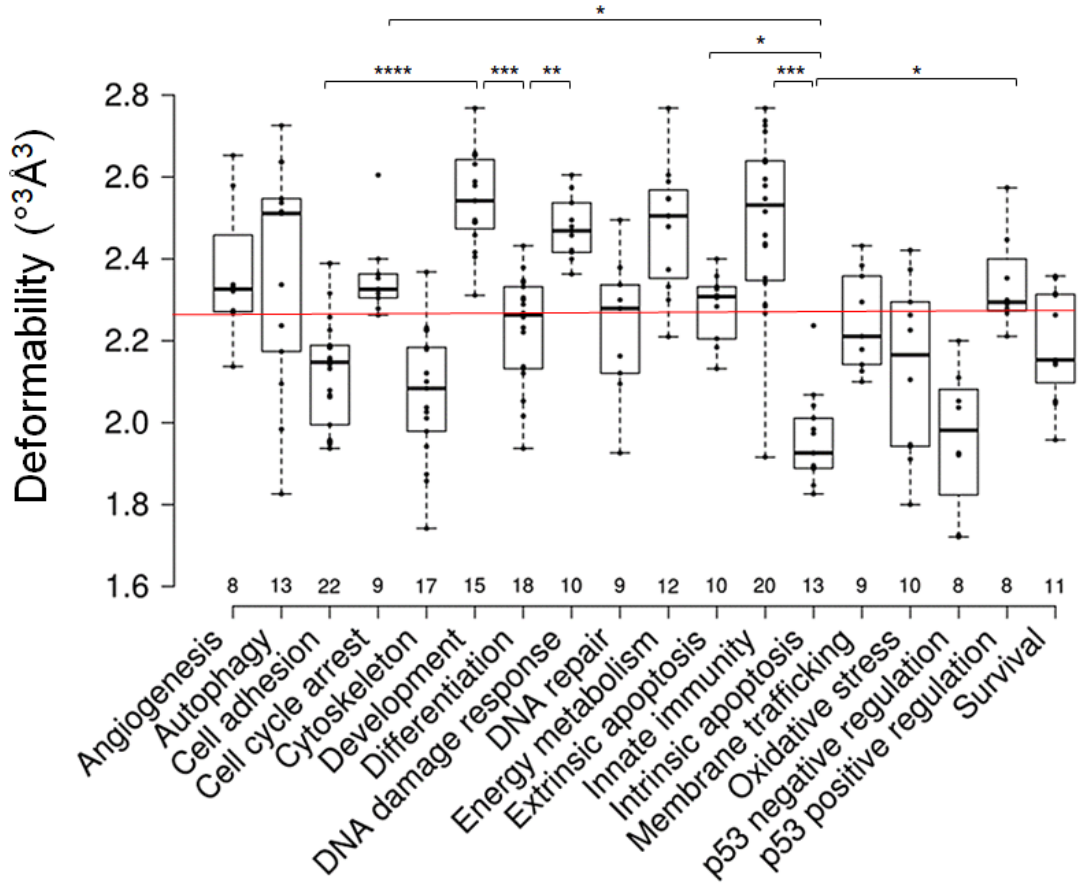

Figure 1: p53 REs belonging to different functional categories are clustered by their deformability. Shown are functional categories with at least 8 genes each. Center lines show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles. All data points are shown as dots. The red line marks the average deformability of the entire set used in the analysis. Numbers above the x axis mark the number of REs in each functional category. P values are marked as follows: P < 0.03 *, P < 0.01 **, P < 0.001 ***, P < 0.0001 ****. Only representative differences are marked by stars, to avoid clattering the pattern. For full list of significant differences see Table S3.

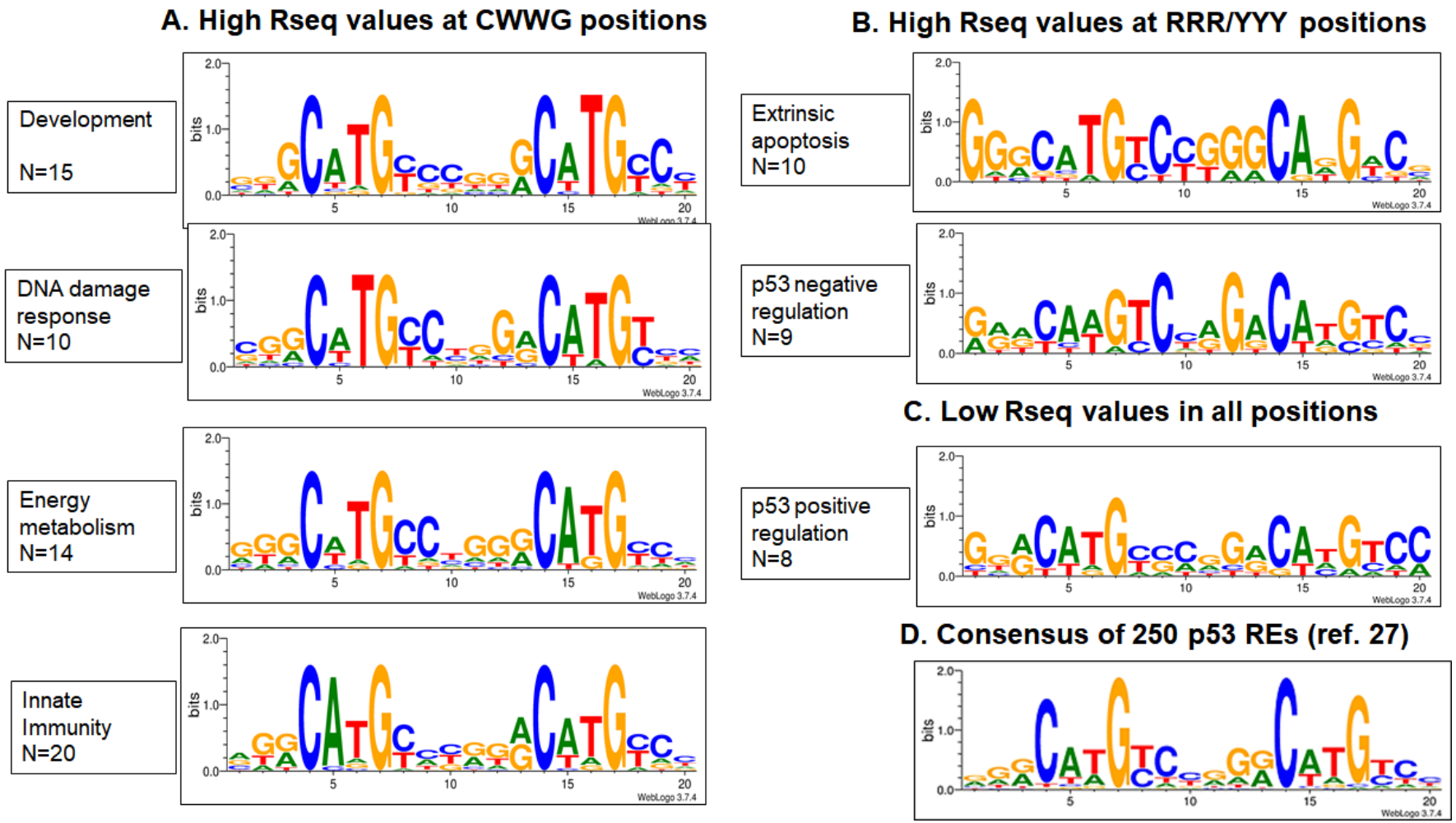

Figure 2: Consensus p53-binding motifs by p53-dependent functional outcome categories. (A) Functional outcome categories with high Rseq values mainly at the two CWWG motifs. (B) Functional outcome categories with high Rseq values mainly at the RRR/YYY motifs. (C) Lowest Rseq values category throughout the entire site. (D) Consensus motif of the entire set of 250 activating p53 REs without spacers. Sequence motifs were generated by weblogo (33-35).

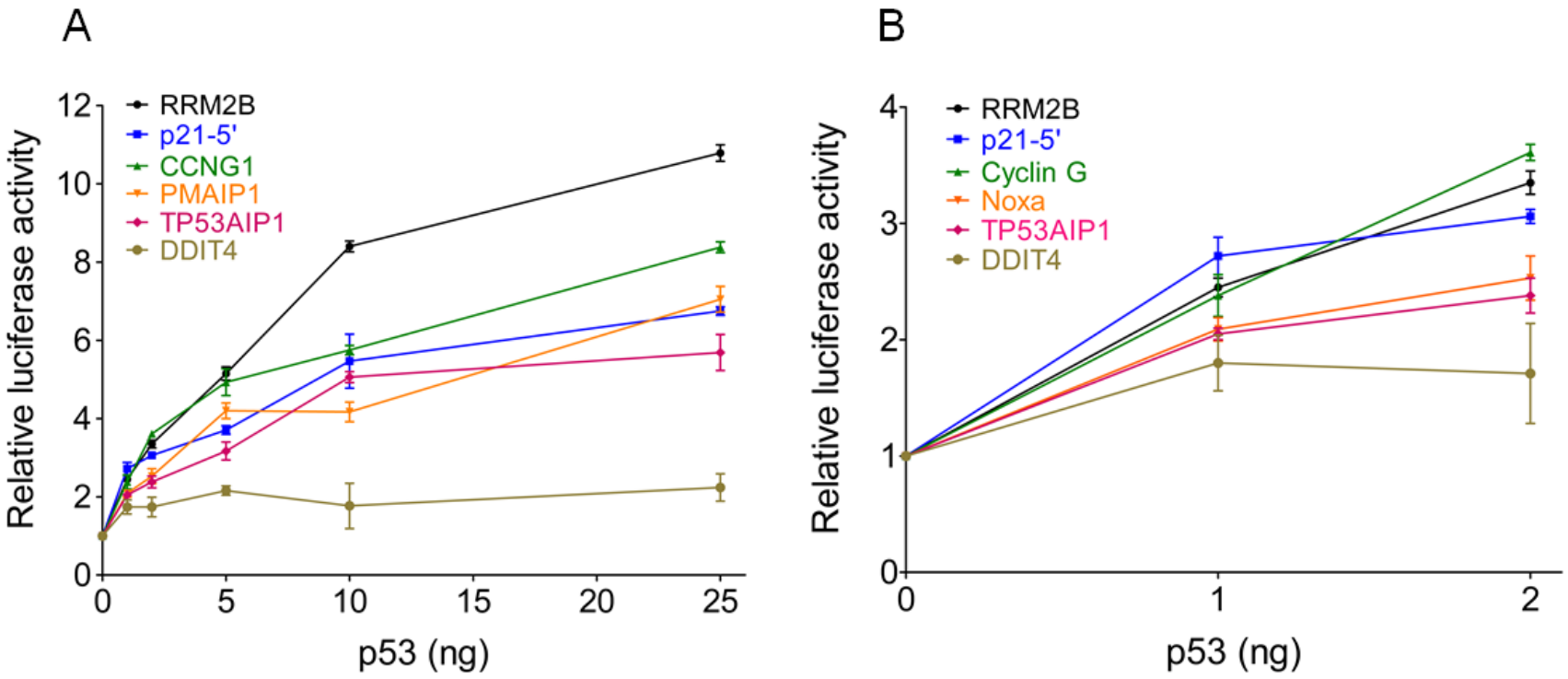

Figure 3: Transactivation from p53 REs as a function of different p53 concentrations. Fold increase in transactivation level measured from six different p53 REs as a function of six different p53 expression levels (1 ng, 2 ng, 5 ng, 10 ng, 25 ng, and 75 ng pC53-SN3 plasmid in the transfection). Luminescence was measured 48 h post transfection. Luminescence values were normalized to the transfection efficiency of co-transfected, constitutively expressed Gaussia luciferase. Results were normalized first to the empty cLuc vector and then to results obtained without p53. Error bars represent mean ± SD of five independent experiments, each containing 2 technical replicas. (A) shows the range of p53 levels from 1 to 25ng, (B) is an inset focusing on the 1-2 ng levels. See Figure S4 for the whole range of p53 levels, including 75 ng.

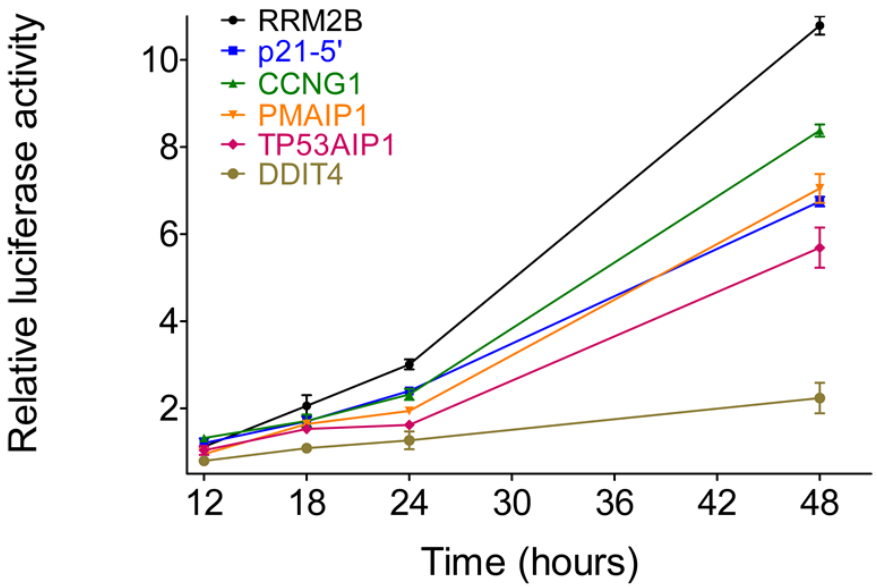

Figure 4: Transactivation from p53 REs as a function of time. Fold increase in transactivation level measured from six different p53 REs as a function of three time points (12 h, 18 h, and 24 h post transfection), using 25 ng pC53-SN3 plasmid. Error bars represent mean ± SD of five independent experiments, each containing 2 technical replicas. Luminescence analysis details are as in Figure 3.

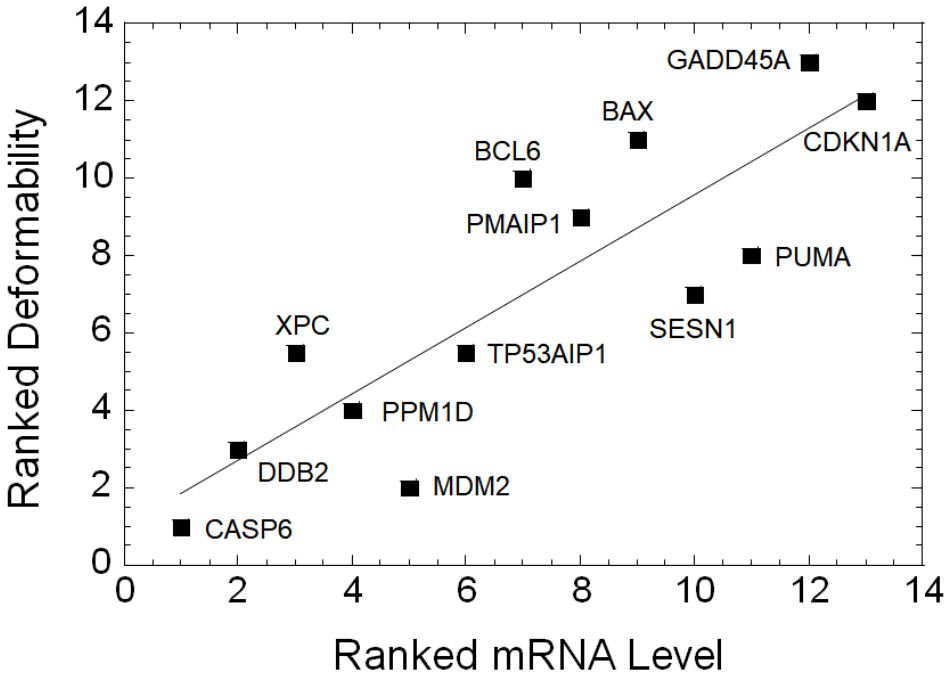

Figure 5: mRNA level from endogenous p53-dependent genes is correlated with the calculated deformability of their REs ($\rho = 0.88$, $P = 8.4 \times 10^{-5}$). The mRNA level data is from Purvis *et al.* (46), and is the pulsatile expression pattern 1 h post stress induction. See text for details.