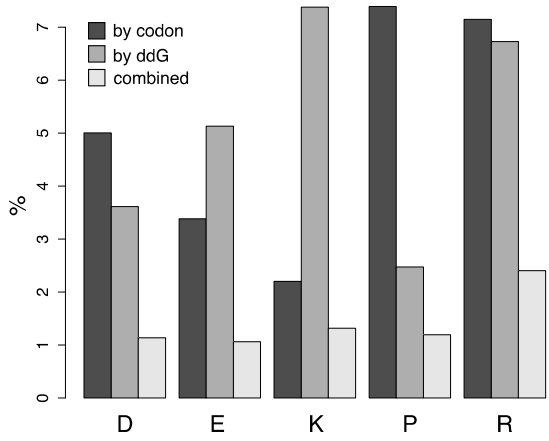
Cell Reports, Volume *31*

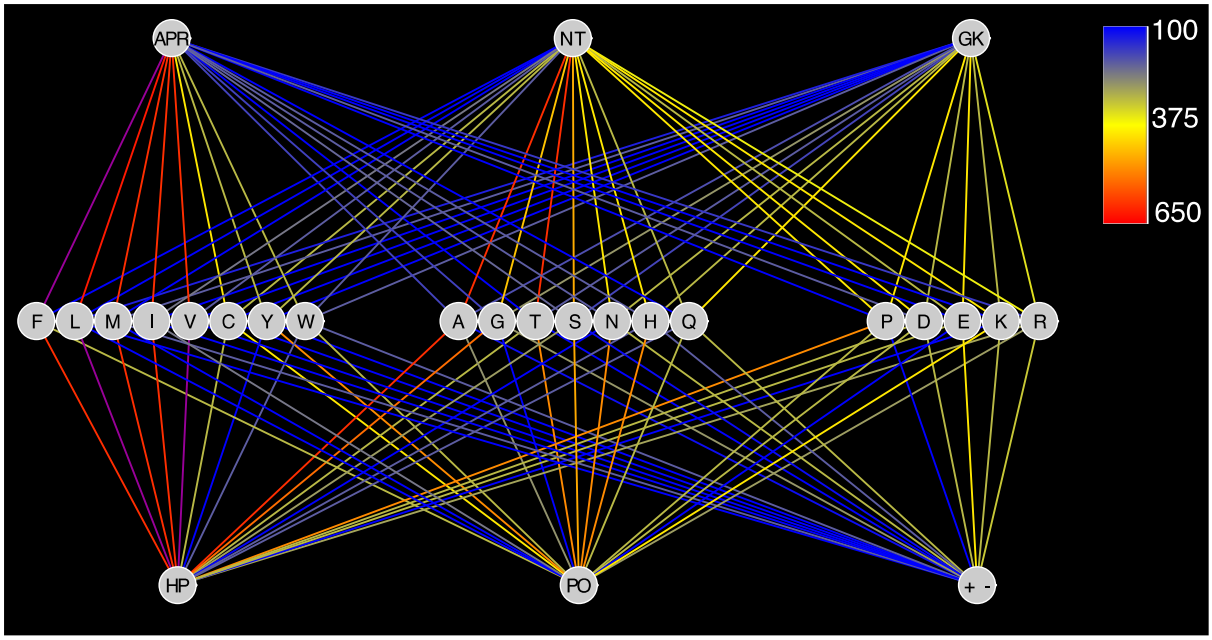
Supplemental Information

Thermodynamic and Evolutionary Coupling between the Native and Amyloid State of Globular Proteins

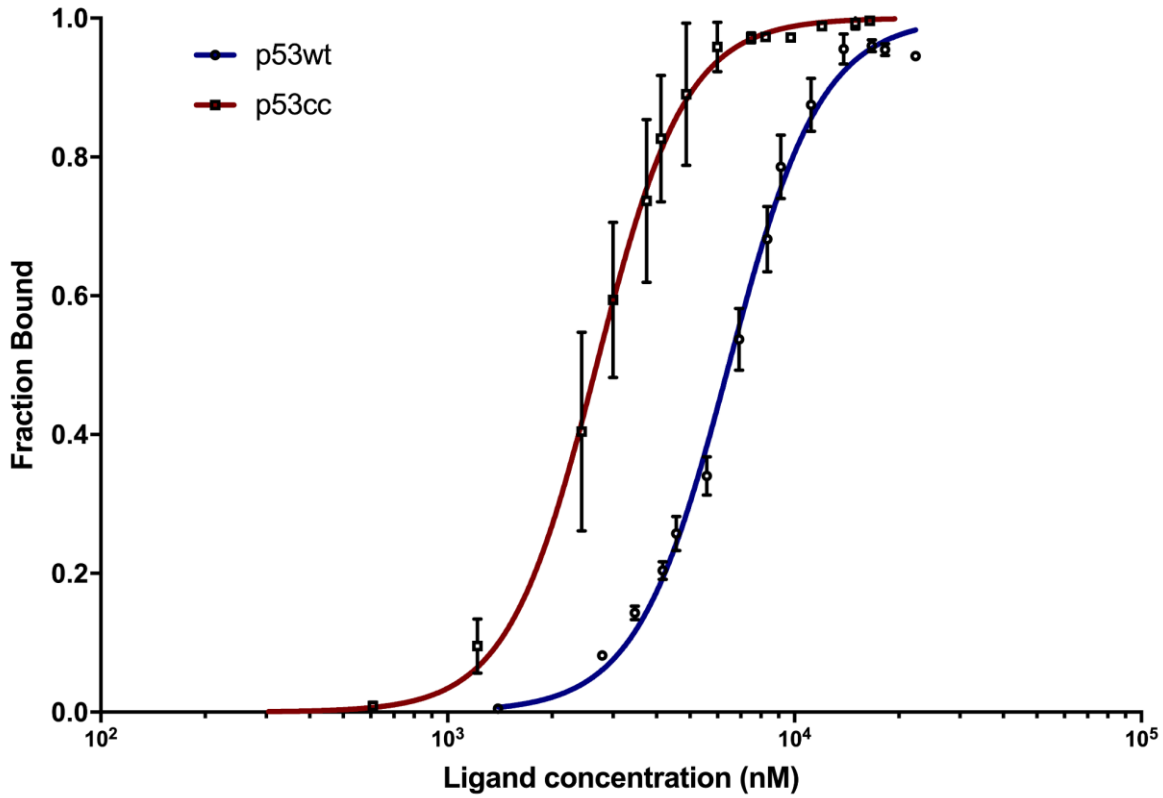
Tobias Langenberg, Rodrigo Gallardo, Rob van der Kant, Nikolaos Louros, Emiel Michiels, Ramon Duran-Romaña, Bert Houben, Rafaela Cassio, Hannah Wilkinson, Teresa Garcia, Chris Ulens, Joost Van Durme, Frederic Rousseau, and Joost Schymkowitz



Supplementary Figure 1. Percent chance of placing a gatekeeper residue in an APR based on codon usage or stability (by ddG), or both combined. Related to Figure 3.



Supplementary Figure 2. Network mapping of single nucleotide based transitional propensities related to physicochemical stability (bottom half) and aggregation proclivity (top half) of the 20 amino acid residue side chains. Related to Figure 3. Top row nodes categorise mutational transitions as aggregation-prone, neutral or gatekeepers, whereas the bottom row is represented by nodes for transitions to hydrophobic, polar or charged states. Initial residue states are shown in single letter coding as the central node row of the graph. Network edges are colour-coded based on strength of transitional propensity (red to blue gradient). Single nucleotide mutations rarely switch to codons encoding for residues with altered physicochemical properties and this conservation is even stronger regarding the amyloid propensity of the transitional residues. This codon bias conservation highlights the coupling between aggregation propensity and globular stability of amino acid side chains.



Supplementary Figure 3. DNA-binding curves of p53cc and wt measured with microscale thermophoresis (MST). Related to Figure 5. Raw fluorescence is normalized to the fraction of bound target shown in the vertical axis. Similar affinities were calculated for both p53cc and wt constructs (2.69 μM ± 0.67 μM and 6.47 μM ± 0.70 μM, respectively). Measurements were carried out in triplicates and plotted using the Hill equation.

Supplementary Tables

Supplementary Table 1. Tm values used. Related to Figure 1.

Species	Average Tm value (degree Celsius)
E. coli	54
T. thermophiles	79
S. cerevisiae	51
H. sapiens (Hela cells)	56

Supplementary Table 2. p53 homologs with low APR strength. Related to Figure 4.

APR or equivalent	TANGO score	UniProtKB ID	homology	species	% identity human p53 DBD
IRVVFRLI	0	A9UZX3_MONBE	p53-like	*Monosiga brevicollis* (Choanoflagellate)	34
MMQLHFTL	0	A0A0N5E6D9_TRIMR	p63	*Trichuris muris* (Mouse whipworm)	32
VYVDFSL	0	A0A0N4VCR4_ENTVE	p53-like	*Enterobius vermicularis* (Human pinworm)	31
TELEFTL	0	A0A195C1G0_9HYME	p63	*Cyphomyrmex costatus* (fungus breeding ant)	29
TELEFTL	0	A0A151XH99_9HYME	p63	*Trachymyrmex zeteki* (fungus breeding ant)	29
TELEFTL	0	F4X756_ACREC	p63	*Acromyrmex echinatior* (Panamanian leafcutter ant)	27
AIFTL	1,1	V5GSV4_ANOGL	p53	*Anoplophora glabripennis* (Asian longhorn beetle)	32
LIFNL	1,2	A0A154PH19_9HYME	p53	*Dufourea novaeangliae* (solitary bee)	31
IHTVF	1,8	A0A0L0G1F4_9EUKA	p53-like	*Sphaeroforma arctica* (*Opisthokonta, Protozoa*)	30
IIFHL	1,8	A8CSR7_NEMVE	p53-like	*Nematostella vectensis* (Starlet sea anemone)	35
LNAVF	2,0	Q4H300_CIOIN	p53-like	*Ciona intestinalis* (transparent sea squirt)	40
AIQLSITL	2,4	A0A0V0ZFM8_9BILA	p53	*Trichinella patagoniensis* (Nematode)	33
AIQLSVTL	2,5	E5SRY1_TRISP	p53	*Trichinella spiralis* (Nematode)	32
AIQLSVTL	2,5	A0A0V1GZU5_9BILA	p53	*Trichinella zimbabwensis* (Nematode)	31
TVFSL	4,3	A0A0A9W1L4_LYGHE	p53	*Lygus hesperus* (Western plant bug)	30
LNICFAL	4,4	F1L1F1_ASCSU	p53	*Ascaris suum* (Pig roundworm)	31
“ILTIITL”	**76**	**-**	**p53**	***Craniata***	**66**

Supplementary Table 3. p53 APR TANGO score and standard deviation per clade. Related to Figure 4.

Clade	N	TANGO score (SD)
Craniata	136	76 (13)
Mollusca	12	76 (9)
Diptera	57	53 (22)
Nematoda	25	36 (30)
Ixodida	8	35 (22)
Coleoptera	7	21 (21)

Supplementary Table 4. Selected mutations and codon usage. Related to Figure 4.

Mutation	WT codon	Mutant codon
L252K	CTC	AAG
T256R	ACA	AGA
A138G	GCC	GGC
R267L	CGG	CTG
N268D	AAC	GAC

Supplementary Table 5. Data collection and refinement statistics. Related to Figure 4.

	P53 charged core
Data collection	
Beamline	Proxima1 (Soleil, France)
Date of collection	10-Feb-2017
Wavelength (Å)	0.9786
Space group	$C222_1$
Cell dimensions	
a, *b*, *c* (Å)	68.42, 84.96, 84.85
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	42.48-1.67 (1.73-1.67)*
R_{sym} *(%)*	8.2 (72.4)
R_{meas} *(%)*	8.6 (78.0)
$I / \sigma I$	17.2 (2.6)
$CC_{1/2}$ (%)	99.9 (83.4)
Completeness (%)	98.2 (95.1)
Redundancy	7.8 (7.2)
No. total reflections	222350 (19402)
No. unique reflections	28518 (2713)
Refinement	
Resolution (Å)	1.67
R_{work} / R_{free}	17.8 / 19.4
No. atoms	1866
Protein	1602
Ligand/ion	13
Water	251
Average *B*-factor	15.32
B-factor Protein	12.60
B-factor Ligand/ion	30.26
B-factor Water	31.89
R.m.s. deviations	
Bond lengths (Å)	0.024
Bond angles (°)	2.1
Ramachandran outliers (%)	0.0
Ramachandran favored (%)	99.5
Ramachandran allowed (%)	0.5

*Values in parentheses are for highest-resolution shell.