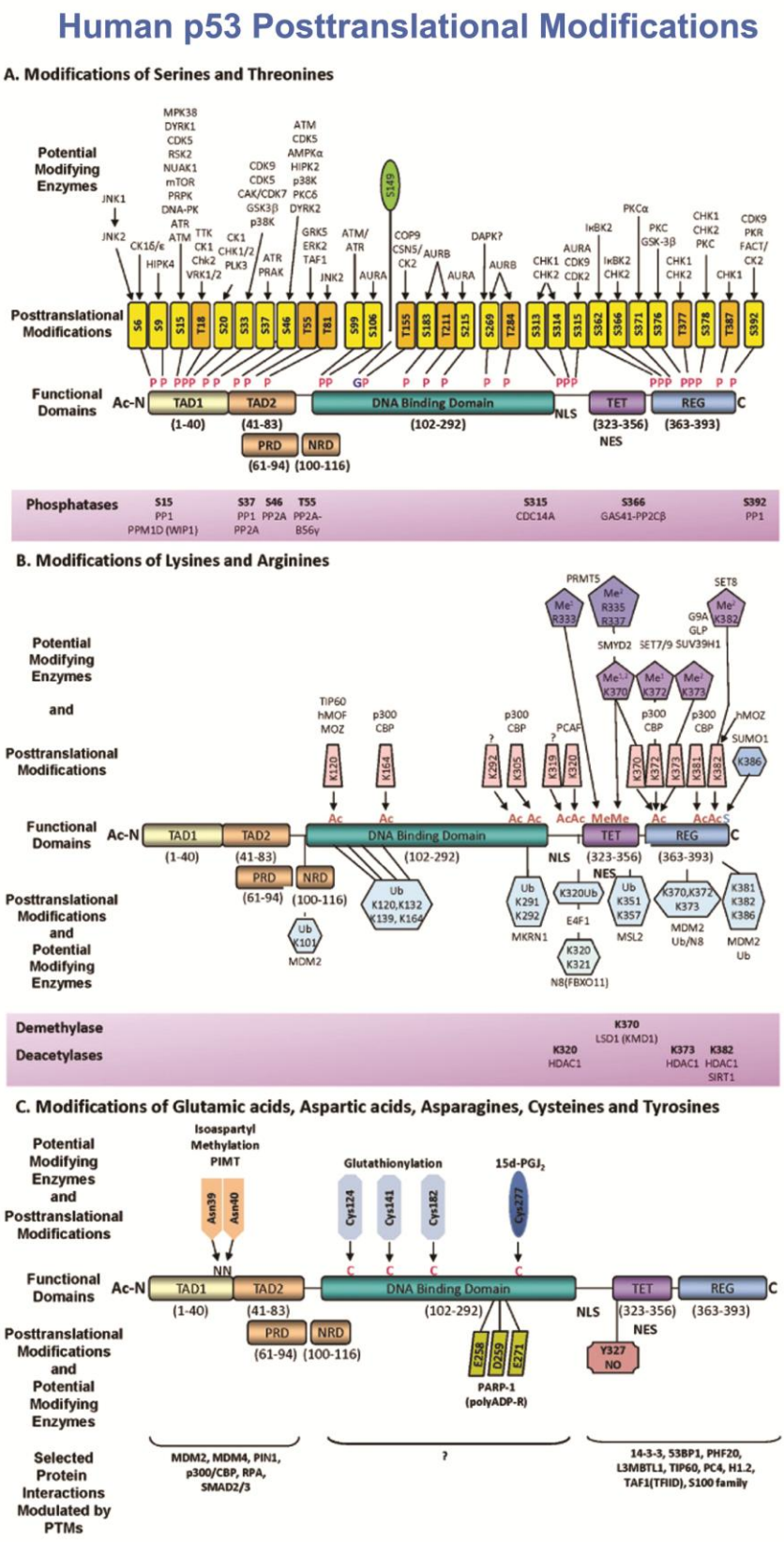
Supp. Figure S1



Supp. Figure S1. Protein domains, posttranslational modification sites and enzymes that can modify WT human p53. The 393 amino acid human p53 polypeptide is represented schematically with postulated functional regions and domains indicated. Residues ~1-40 (TAD1) and 41-83 (TAD2) comprise independent tandem transactivation domains; residues ~61-94 represent a proline-rich domain (PRD); residues 33-80 are poorly conserved. Residues 100-116 constitute an N-terminal repression domain that is required for repressing basal p53 activity in some cell types. Residues ~102-292 contain the central, sequence-specific, DNA binding core region; residues 305-321 contain the primary bi-partite nuclear localization signal (NLS); residues 323-356 comprise the tetramerization domain (TET) which contains a nuclear export signal within residues 339-350; residues 363-393 (REG) negatively regulate DNA binding by the central core to consensus recognition sites in oligonucleotides and interact in a sequence-independent manner with single- and double-stranded nucleic acids but contribute positively to chromatin binding and transactivation in vivo.

Posttranslational modification sites (P, phosphorylation; Ac, acetylation; G, glycosylation; Me^1, Me^2, mono- or di-methylation; Ub, ubiquitylation; N8, neddylation; SUMO1, sumoylation; isoaspartyl methylation; glutathionylation; NO, nitrosylation; ADP-R, ADP-ribosylation; 15d-PGJ_2, 15-deoxy-$\Delta^{12,14}$-prostaglandin J_2 conjugation) are indicated together with enzymes that can accomplish the modifications in vitro.

Enzyme/Protein Abbreviations: 14-3-3, eluting from DEAE-cellulose in the 14th fraction from bovine brain homogenate and found on positions 3.3 after subsequent electrophoresis (YWHA, Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein); 53BP1, p53 binding protein 1; AMPK, Protein kinase AMP-activated; ATM, Ataxia telangiectasia mutated; ATR, Ataxia telangiectasia-related; AurA, Aurora kinase A (STK15); AurB, Aurora kinase B; CBP, CREB binding protein (KAT3A); CDC14A, Cell division cycle 14A (phosphatase); CDK5/7/9, Cyclin-dependent kinase 5/7/9 (CAK/CDK7); CHK1/2, Checkpoint kinase 1/2; CK1/2, Casein kinase 1/2; COP9/CSN5, Constitutive photomorphogenic (COP) *arabidopsis* (homolog) associated subunit 5; DAPK, Death-associated protein kinase; DNA-PK, DNA-activated protein kinase; DYRK1/2, Dual-specificity tyrosine phosphorylation regulated kinase 1/2; E4F1, E4F transcription factor 1 (Represses transcription from adenovirus E4 promoter in the absence of E1A); ERK2, Extracellular signal-regulated kinase 2 (p44 MAPK); FACT/CK2, Facilitates chromatin transcription complex with CK2; FBXO11, F-box protein 11; G9A

(KMT1C), SET domain containing protein G9A; GAS41-PP2Cβ, YEATS-domain containing common subunit of TIP60-SRCAP complexes complexed with PP2Cβ; GLP (KMT1D), G9A-like protein; GRK5, G-protein coupled receptor kinase 5; GSK3ß, Glycogen synthase kinase beta; H1.2, Histone cluster 1, H1c; HIPK2/4, Homeodomain-interacting protein kinase 2/4; HDAC1, Histone demethylase 1; hMOF (KAT8), ortholog of *Drosophila* Males Absent on the First; IκBK2, Inhibitor of kappa light chain enhancer in B-cells kinase 2; JNK1/2, Jun N-terminal kinase 1/2; L3MBTL1, Lethal (3) malignant brain tumor-like protein 1; LSD1 (KMD1), Lysine-specific demethylase 1; MDM2/4, Mouse double minute 2/4; MKRN1, Makorin ring finger protein 1; hMOZ (KAT6A), Monocytic leukemia zinc finger protein; MPK38, Murine serine-threonine kinase 38; MSL2, Male-specific lethal 2 homolog (*Drosophila*); mTOR, Mechanistic target of Rapamycin; p38K, Mitogen-activated protein kinase p38 alpha; NUAK1, Novel (nua) kinase family 1; p38K, Mitogen-activated protein kinase 1; p300 (KAT3B), E1A binding protein p300; PARP-1, Poly (ADP ribose) polymerase 1; PCAF; p300/CBB-associated factor (KAT2B); PC4, Positive coactivator 4 (SUB1 homolog (*S. cerevisiae*)); PCAF (KAT2B), p300/CBP-associated factor; PIMT, Protein-L-isoaspartate (D-aspartate) O-methyltransferase; PHF20, PHD finger protein 20; PIN1, Peptidylproline Cis/Trans isomerase 1; PIRH1, Ring finger and CHY zinc finger domain containing 1 (p53-induced protein with a RING H2); PKCδ/ε□□□, Protein kinase C delta/epsilon; PKR, Protein kinase interferon-inducible, double-stranded RNA activator; PP1/2a, Protein phosphatase 1/2a; PPM1D, Protein phosphatase Mg2+/Mn2+-dependent (WIP1, wild-type induced p53); PRAK, p38-regulated-activated protein kinase; PRMT5, Protein arginine methyltransferase 5; PRPK, TP53-regulating kinase; RPA, Replication protein A; RSK, Ribosomal S6 kinase; S100, S100 calcium-binding family of proteins; SET7/9 (KMT7), SET domain containing (lysine methyltransferase) 7; SET8 (KMT5A), SET domain containing protein 8; SIRT1, Silent mating type information regulation 2 homolog 1 (*S. cerevisiae*); SMYD2 (KMT3C), SET and MYND domain containing 2; SMAD2/3, Sma- and Mad-related protein 2; SUMO1, SMT3 suppressor of Mif two 3 homolog 1 (*S. cerevisiae*); SUV39H1 (KMT1A), Suppressor of variegation 3-9 homolog 1 (*Drosophila*); TAF1, TATA-box binding protein associated factor 1; TIP60 (KAT5), HIV Tat interactive protein; TTK, Phosphotyrosine picked threonine kinase; VRK1/2, Vaccinia-related kinase 1/2.

A. Posttranslational modifications to serines and threonines; **B.** posttranslational modifications to lysines and arginines; **C.** posttranslational modifications to glutamic acids, aspartic acids,

cysteines and tyrosines. The figure is updated and modified from recent reviews (Anderson and Appella, 2009 and Meek and Anderson, 2009); references for most modifications can be found at PhosphoSitePlus (http://www.phosphosite.org/homeAction.do). A handful of additional residues are reported to be modified (e.g. R110Me[1], H115Me, S149p, T150p, R209Me[1], R213Me[1], K291Ac, T312p), but evidence for their occurrence *in vivo* currently either is weak or non-existent.

Supp. Table S1. Cancer Mutants in Known p53 PTM Sites

Residue	Domain	Modification	Enzyme	Signal	Vertebrate Residue Conservation	p73 p63	Mutant Residue	Number Mutations at this Residue	Mouse Mutant Model	Cell Line
Ser6	TAD1	P	JNK2	DNA Damage	0.63	NO	L	1	-	-
Ser9	TAD1	P	HIPK4, CK1	DNA Damage	0.51	NO	-	0	-	-
Ser15	TAD1, NES1	P	ATM, ATR, NUAK1	DNA Damage	0.98	NO	R	1	S18A	-
Thr18	TAD1, NES1	P	CK2, VRK1/2	DNA Damage	0.84	NO	-	0	-	-
Ser20	TAD1, NES1	P	CHK1/2, CK1, PLK3	DNA Damage	0.74	NO	-	0	S23A	-
Asn29	TAD1	carboxyl M	PIMT	?	0.65	NO	-	0	-	-
Asn30	TAD1	carboxyl M	PIMT	?	0.60	NO	S	1	-	YES
Ser33	TAD1	P	p38K, CDK5, CDK9	DNA Damage	0.58	NO	T, F	3	-	-
Ser37	TAD1	P	ATR, PRAK	DNA Damage	0.53	NO	T, P	2	-	YES
Ser46	TAD2	P	ATM, p38K, CDK5, DYRK2	DNA Damage	0.53	NO	P, F	6	Hupki S46A	-
Thr55	TAD2	P	TAF1, ERK2	Unstressed	0.23	NO	I	1	-	-
Thr81	TAD2, PRD	P	JNK2	DNA Damage	0.35	NO	I	2	-	-
Ser99	PRD	P	ATM, ATR, DNA-PK	DNA Damage	0.79	YES	P, F,	4	-	-
Lys101	NRD	Ub	MDM2	Unstressed	0.65	NO	R, N	4	-	-
Ser106	DBD	P	AURA	?	0.40	NO	R, G	8	-	YES
Lys120	DBD	Ac	TIP60, hMOF	DNA Damage	1	YES	Q, E, M, N, R	16	K117R 3KR	-
Cys124	DBD	GTN	?	?	0.84	NO	S, R, G, Y, W	6	-	YES

Lys132	DBD	Ub	MDM2?	Unstressed?	1	YES	Q, E, W, L, T, R, M, N	205	-	YES
Lys139	DBD	Ub	MDM2?	Unstressed?	1	YES	Q, E, T, R, N	30	-	YES
Ser149	DBD	OGN P?	?	?	0.74	NO	T, P, A, F	21	-	-
Thr155	DBD	P	COP9/CSN5	Unstressed	0.53	NO	P, A, S, N, I, M	102	-	YES
Lys164	DBD	Ac Ub	p300 PIRH2	?	1	YES	Q, E, T, R, M, N	49	K161R 3KR	YES
Cys182	DBD	GTN	?	?	0.63	NO	S, R, Y	18	-	-
Ser183	DBD	P	AURB	?	0.70	NO	P, L	9	-	-
Thr211	DBD	P	AURB	?	0.98	YES	P, A, S, N, I	35	-	YES
Ser215	DBD	P	AURA	?	1	YES	R, G, C, N, T, I, K	110	-	YES
Glu258	DBD	ADR	PARP1	?	1	YES	K, Q, A, G, V, D	129	-	YES
Asp259	DBD	ADR	PARP1	?	0.72	NO	N, H, Y, P, S, A, G, V, E	97	-	YES
Ser269	DBD	P	AURB, DAPK	?	0.79	NO	R, G, C, N, T, I	26	-	-
Glu271	DBD	ADR	PARP1	?	1	YES	K, Q, P, R, A, G, V, D	67	-	YES
Cys277	DBD	ALK	15d-PGJ2	?	1	YES	R, G, Y, S, F, W	100	-	YES
Thr284	DBD	P	AURB, DAPK?	?	0.95	NO	P, A, K, I	19	-	YES
Lys291	DBD	Ub	MKRN1	Unstressed	0.93	NO	Q, E, T, R, M, N	23	-	-
Lys292	DBD	Ub	MKRN1, MDM2, PIRH2	Unstressed	0.77	NO	Q, E, G, T, R, I, N	24	-	-
Lys305	LNK	Ac Ub	p300 MDM2, PIRH2	DNA Damage Unstressed	0.74	YES	E, T, R, M, N	16	-	-
Ser313	LNK	P	CHK1/2	DNA Damage	0.70	NO	R, C, N, I	5	-	-
Ser314	LNK	P	CHK1/2	DNA Damage	0.65	NO	F	1	-	-
Ser315	LNK	P	AURA, CDK9	DNA Damage	0.79	NO	P, C, F	4	Hupki S315A S312A	-

Lys319	LNK, NLS1	Ac Ub	PCAF MDM2	DNA Damage Unstressed	0.88	YES	**E , R, N**	7	-	-
Lys320	LNK, NLS1	Ub, N8	E4F1, FBX011, MDM2	Unstressed?	0.72	NO	**N**	6	K317R	YES
Lys321	LNK, NLS1	N8	FBX011, MDM2	Unstressed?	0.74	NO	**R**	1	-	-
Tyr327	TET	NO	?	NO	0.70	NO	**H, D, S, C**	4	-	-
Arg333	TET	Me	PRMT5	?	0.91	YES	-	0	-	-
Arg335	TET	Me2	PRMT5	?	0.84	YES	**G, H, L**	3	-	-
Arg337	TET	Me2	PRMT5	?	0.95	NO	**C, H, P, L**	37	-	YES
Lys351	TET, NES2	Ub	MLS2, MDM2	Unstressed?	0.72	NO	**N**	1	-	-
Lys357	REG	Ac Ub	hMOF MDM2, PIRH2	DNA Damage	0.58	NO	-	0	-	-
Ser362	REG	P	IκB2	DNA Damage	0.65	NO	-	0	-	-
Ser366	REG	P	IκB2, CHK2	DNA Damage	0.74	NO	**A**	2	-	-
Lys370	REG	Ac Me1,2 N8 Ub	p300, CBP SMYD2 MDM2 MDM2, PIRH2	DNA Damage Unstressed	0.81	NO	**Q**	1	K367R K6R, 7KR	-
Ser371	REG	P	PKCα	?	0.58	NO	-	0	-	-
Lys372	REG	Ac Me1,2 N8 Ub	p300, CBP SET7, SET9 MDM2 MDM2, PIRH2	DNA Damage Unstressed	0.70	NO	-	0	K369R K6R, 7KR	-
Lys373	REG	Ac Me1,2 N8	p300, CBP G9A, GLP MDM2	DNA Damage Unstressed	0.72	NO	-	0	K370R K6R, 7KR	-

		Ub	MDM2, PIRH2							
Ser376	REG	P	GSK3β, PKC	?	0.70	NO	T, A	2	-	-
Thr377	REG	P	CHK1/2	DNA Damage	0.58	NO	-	0	-	-
Ser378	REG	P	CHK1/2, PKC, p38K	?	0.84	NO	-	0	-	-
Lys381	REG	Ac	p300, CBP	DNA Damage	1	NO	-	0	K378R K6R, 7KR	-
		Ub	MDM2, PIRH2	Unstressed						
Lys382	REG	Ac	p300, CBP, hMOZ	DNA Damage	0.88	YES	-	0	K379R K6R, 7KR	-
		Me2	SET8, PR-SET7	DNA Damage						
		Ub	MDM2, PIRH2	Unstressed						
Lys386	REG	Ac	p300, CBP	DNA Damage	0.88	NO	-	0	K383R K6R, 7KR	-
		SUMO	PIAS, PIASxβ	?						
		Ub	MDM2, PIRH2, COP1	Unstressed						
Thr387	REG	P	CHK1	DNA Damage	0.26	NO	-	0	-	-
Ser392	REG	P	CDK9, PKR, FACT/CK2	DNA Damage	1	NO	L	1	S389A	-

This supplemental table is identical in content to Table 1 in the published text except that the cells are colored to match the TP53 domain structure shown in Supp. Figure S1 above, and mutant residue function (column 8) is indicated by a color as described below.

Table information was abstracted largely from the p53 IARC TP53 Database R16 http://p53.iarc.fr/ (Petitjean et al., 2007) which describes the properties of 29,575 somatic mutations and 635 germline mutations. The human TP53 sequence is taken from SwissProt # P04637-1 and is known as the 'canonical' sequence or 'p53alpha'; residue 1 is the initiating methionine. The mouse TP53 protein sequence is derived from RefSeq entry NP_035770; Residue 1 corresponds to the first methionine codon which is located three residues 5' of the first

methionine codon in the human cDNA. Residue: Bold indicates a more frequently mutated residue (>90/29,575). Domain: TAD1, Transactivation Domain 1; TAD2, Transactivation Domain 2; PRD, Proline Rich Domain; NRD, N-terminal Repression Domain; DBD, Sequence-specific DNA Binding Domain; LNK, Linker Domain; TET, Tetramerization Domain; REG, Regulatory Domain, non-specific DNA binding domain; NES, Nuclear Export Signal; NLS, Nuclear Import Signal; gray shading indicates the different TP53 domains. Modification: P, phosphorylation; Ac, acetylation; ALK, alkylation; Ub, ubiquitylation; Me1, monomethylation; Me2, dimethylation; Su, sumoylation; N8, neddylation; ADR, ADP-ribosylation; OGN, O-glycosylation; GTN, Glutathionylation; NO, Nitration. Residue conservation data was obtained from TP53 Mut Assessor (Leroy et al., 2013). Vertebrate: Conservation of the residue in vertebrate TP53: Forty-two vertebrate TP53 sequences were aligned using CLUSTAL. For each position from 1 to 393, conservation of each residue of the human TP53 has been ranked from 0 (not found in any other TP53 to 1 (residue conserved in all vertebrate TP53s). Only full-length TP53 sequences were used for this analysis. TP73/ TP63: Conservation of the residue in human TP73 and TP63. The human TP53 sequence was aligned to human TP63 and TP73 to identify residues conserved in the three proteins. Yes: residue conserved; No: residue not conserved. Missense Mutants: Blue, transactivation competent; Purple, partial loss of transactivation; Green, supertransactivator; Red, loss of transactivation; Black, not evaluated. Enzyme/Protein Abbreviations: see Supp. Figure S1. Mouse Mutant Model: Mutant models were taken from the literature (see text); the mouse p53 residue is given together with the amino acid (single letter code) encoded by the mutant. Mutation of multiple residues is indicated as per the literature. Cell Line: The availability of a mutant human cell line is indicated as per the IARC TP53 database.

Note added during review: A recent proteomics study has identified a large number of additional TP53 posttranslational modifications that occur under some circumstances in human fibroblasts and COS-1 cells (Dehart et al., 2013).

Supp. References

Anderson CW, Appella E. 2009. Signaling to the p53 tumor suppressor through pathways activated by genotoxic and non-genotoxic stresses. In: Bradshaw RA, Dennis EA, editors. Handbook of Cell Signaling: Elsevier. p 2185-2204.

Dehart CJ, Chahal JS, Flint SJ, Perlman DH. 2013. Extensive post-translational modification of active and inactivated forms of endogenous p53. Mol Cell Proteomics (in press).

Meek DW, Anderson CW. 2009. Posttranslational modification of p53: cooperative integrators of function. Cold Spring Harb Perspect Biol 1(6):a000950.