Non-hydrolyzable acetyllysine analogs to study protein acetylation in vitro and in cells

Simon Maria Kienle[1,3,#,*], Matthias Sigg[1,#], Tobias Schneider[2,3], Katrin Stuber[1,2,3], Jan Lehmann[1,3], Jasmin Jansen[1,3], Florian Stengel[1,3], Andreas Marx[2,3], Michael Kovermann[2,3], Martin Scheffner[1,3,*]

[1]Department of Biology, University of Konstanz, Konstanz, Germany
[2]Department of Chemistry, University of Konstanz, Konstanz, Germany
[3]Konstanz Research School Chemical Biology, University of Konstanz, Konstanz, Germany

[#]equal contribution

*Corresponding authors: simon.kienle@uni-konstanz.de , martin.scheffner@uni-konstanz.de

ABSTRACT

Lysine acetylation plays a prominent regulatory role in eukaryotic cells. Yet, determining the functional consequences of acetylation for a given protein represents a considerable challenge. For instance, lysine residues are subject to various posttranslational modifications, rendering interpretation of mutational studies difficult. The genetic code expansion technology enables site-specific incorporation of acetyllysine (AcK) into proteins, but the applicability of AcK is limited, as within cells, the acetyl group is removed by deacetylases. Here, we show that site-specific incorporation of the non-hydrolyzable AcK analog ketolysine (KeK) into ubiquitin closely resembles the structural and functional effects of AcK incorporation. Furthermore, AcK and KeK can be efficiently incorporated into the tumor suppressor p53 in cells. However, whereas AcK becomes deacetylated, KeK remains stable. Accordingly, incorporation of KeK, but not AcK, affects p53-mediated transcription. Thus, we propose that KeK is the AcK surrogate of choice for studying acetylation of a given protein in cells.

INTRODUCTION

The functions and properties of eukaryotic proteins are controlled by various posttranslational modifications (PTMs) that in most cases are reversible[1]. This allows a cell to rapidly adapt its proteome during time-regulated processes such as the cell cycle or differentiation and to respond to environmental changes. The continuous advancement and refinement of mass spectrometry (MS)-based approaches has been instrumental in identifying and cataloging different PTMs, e.g. phosphorylation, acetylation, and modification of proteins by ubiquitin (Ub) and Ub-like proteins, for individual proteins as well as on a proteome-wide scale[2,3]. Yet, while such studies provide invaluable insights into the plasticity of proteomes, how a given PTM affects the properties of a protein of interest (POI) remains unclear in many cases[4].

A general limitation in studying the effect of PTMs on the biochemical and structural properties of a POI is the generation of sufficient amounts of respectively modified proteins[5]. While in principle this can be achieved by incubation of the POI with the respective modifying

44 enzymes, the efficiency of such reactions is often limited and/or not position-specific (i.e.
45 residues in addition to, or other than the one of interest are modified), resulting in
46 heterogeneous populations of the POI. Moreover, evaluation or validation of the results
47 obtained in vitro may be even more challenging in cells. Common approaches to study the
48 effect of a PTM in cells include the use of POI variants, in which the respective residue has
49 been mutated (e.g. in case of lysine acetylation, replacement of lysine by arginine or
50 glutamine), and modulation of the activity of the respective modifying enzyme[6]. Such
51 approaches have obvious limitations. Targeting the modifying enzyme does in most, if not all
52 cases not only affect the modification status of the POI, but also that of other proteins.
53 Concerning replacement strategies, a main consequence of lysine acetylation, for instance,
54 is neutralization of the positive charge of the side chain of lysine[7]. To provide evidence for
55 the importance of a positive charge at the respective position of a POI, lysine is commonly
56 replaced by arginine. However, lysine is not only subject to acetylation, but also to other
57 PTMs including modification by methyl groups, Ub, and Ub-like proteins[8]. Thus, it is difficult,
58 or even not possible to distinguish if a respective effect is due to interference with acetylation
59 or, for example, with ubiquitination. As surrogate for acetylated lysine (acetyllysine or briefly,
60 AcK), glutamine is generally employed, which mimics elimination of the positive charge of the
61 lysine side chain. However, at least in some cases, the functional consequences of lysine
62 acetylation are not solely explained by charge neutralization[9-11]. Attachment of an acetyl
63 group results in extension of the length of the side chain, which depending on the position in
64 a POI, may induce structural changes, an effect that cannot be mimicked by the side chain of
65 glutamine (see Figure 1a), as we have for instance shown using Ub as model system.[11]
66 The genetic code expansion technology represents an attractive possibility to
67 incorporate AcK at the desired position of a POI, e.g. by the so-called amber codon
68 suppression (ACS) method[5,12-14]. Yet, in applications in eukaryotic cells the use of K-to-Q
69 variants may seem to be superior to ACS-mediated incorporation of AcK, as the latter is
70 technically more complex and more importantly, the acetyl group can be readily removed by
71 the action of deacetylases. To overcome this potential limitation, non-hydrolyzable AcK

72 analogs, such as trifluoroacetyllysine (TFAcK)[15] and 2-amino-8-oxononanoic acid, or briefly
73 ketolysine (KeK)[16], should prove helpful (Figure 1a). To address this proposition, we used
74 Ub[11] and the tumor suppressor p53[17-19] as model proteins in vitro and in cells, respectively.
75 Our results obtained with Ub in vitro clearly indicate that unlike Q, TFAcK and KeK represent
76 highly suitable AcK surrogates. Furthermore, we provide evidence that upon incorporation
77 into p53, AcK, and in part TFAcK, are efficiently deacetylated in cells, while KeK is refractory
78 to deacetylation. Notably, this difference in deacetylation susceptibility correlates with the
79 transcriptional activity of p53 variants harboring either AcK, TFAcK or KeK. In conclusion, our
80 data indicate that KeK is the AcK surrogate of choice for studying protein acetylation in cells.
81

82 **RESULTS**

83 **Incorporation of AcK, TFAcK, or KeK has similar effects on Ub structure and function**

84 Besides its well-established role as a covalent modifier of itself and other proteins, Ub is also
85 subject to PTMs including acetylation[20-22]. By employing the ACS technology, we have
86 recently shown that acetylation of K11 affects the structure, and in consequence functional
87 aspects of Ub[11]. Notably, these effects of acetylation are not caused solely by charge
88 neutralization, but also involve steric effects presumably imposed by the length of the side
89 chain (Figure 1a). Indeed, replacement of K11 by the presumed AcK surrogate Q shows
90 different effects on the structure of Ub, and the respective Ub variant is likewise functionally
91 different, i.e. it is more efficiently used by the E3 ligase HDM2 for autoubiquitination than
92 K11-acetylated Ub[11] (see also Figure 1b). Together with results from other groups[9,10], our
93 studies with Ub therefore indicate that data obtained with protein variants, in which lysine
94 residues were replaced by glutamine, have to be interpreted with caution, especially in the
95 absence of structural data.

96 AcK is likely subjected to deacetylation in cells. Therefore, we resorted to the AcK
97 analogs TFAcK[23] and KeK[16], as their side chains are similar in length to AcK (Figure 1a), but
98 are either less (TFAcK)[15] or not (KeK) susceptible to deacetylation (see below). TFAcK or
99 KeK were incorporated into Ub at position 11 (Ub 11X, where X represents the respective

100 amino acid) by ACS in bacteria (Supplementary Figure 1), and the respective Ub variants
101 were first tested for their ability to be used by HDM2 in autoubiquitination experiments
102 (Figure 1b). In agreement with our previous results[11], Ub 11AcK was less efficiently used by
103 HDM2 than unmodified Ub and Ub 11R (as judged by the levels of free Ub at 60 min), while
104 Ub 11Q was more efficiently used than Ub 11AcK (Figure 1b). Importantly, the efficiency of
105 HDM2 autoubiquitination in the presence of Ub 11KeK or Ub 11TFAcK was similar to that in
106 the presence of Ub 11AcK (Figure 1b).

107 The results obtained in HDM2 autoubiquitination assays support the notion that the
108 effects of K11 acetylation are mediated by a combination of charge neutralization and
109 structural changes induced by extension of the lysine side chain[11]. To corroborate this
110 assumption and to show that incorporation of AcK, TFAcK or KeK induces similar structural
111 changes, we performed high-resolution NMR spectroscopy using isotopically labeled variants
112 of Ub 11AcK, 11TFAcK, and 11KeK. For all three Ub variants, two-dimensional ^{1}H-^{15}N
113 heteronuclear single quantum coherence (HSQC) NMR spectra were acquired
114 (Supplementary Figure 2), and weighted chemical shift perturbations (CSPs) were calculated
115 for backbone amide resonances with reference to unmodified Ub (Figure 2). By this,
116 alterations in the chemical environment of corresponding amide proton and nitrogen spins as
117 consequence of the conformational changes induced by the respective modification at
118 position 11 of Ub can be monitored on a residue-by-residue basis. As expected by the length
119 and composition of their side chains, Ub 11AcK, 11TFAcK, and 11KeK show highly similar
120 CSP patterns, including characteristic perturbations at the C-terminal end of the central α-
121 helix of Ub[11]. In fact, only slight variations are apparent between the Ub variants in terms of
122 amplitude and the exact position of the affected amino acids in the primary sequence.

123 Taken together, the above results indicate that both TFAcK and KeK properly mimic
124 the effects of acetylation of K11 on Ub structure and that the actual side chain length at
125 position 11 is the determining factor for the acetylation-induced structural changes resulting
126 in the inefficient usage of the respective Ub variants by HDM2.
127

128 **Ub 11AcK, Ub 11TFAcK, and Ub 11KeK share similar interactomes**

129 To further prove that TFAcK and KeK are functionally equivalent to AcK, we performed
130 affinity enrichment coupled to high-resolution mass spectrometry (AE-MS) experiments with
131 whole cell extracts to compare the protein-protein interaction properties ("interactomes") of
132 Ub 11AcK/TFAcK/KeK. As controls, we used Ub 11A, Ub 11R, Ub 11Q, and unmodified Ub
133 (Ub 11K). In brief, we equipped all Ub variants with an N-terminal Strep tag II for
134 immobilization on Strep-Tactin® beads. The immobilized Ub variants or Strep beads alone
135 were incubated with whole cell extracts derived from HEK293T cells, and affinity-enriched
136 proteins were identified by label-free quantitative MS. Significantly enriched proteins were
137 determined by ANOVA statistics (S0 = 2, FDR = 0.005) (Figure 3a).

138 Hierarchical clustering resulted in the identification of 338 significant interactors that
139 bind to one or more Ub variants (Figure 3b). Besides proteins that show no preference for
140 any of the Ub variants (Cluster 4), we found 23 proteins that preferentially interact with Ub
141 11AcK, Ub 11TFAcK, and Ub 11KeK (Cluster 1) and 13 proteins that showed a strong
142 decrease in binding to these Ub variants (Cluster 5). Similarly, proteins were identified that
143 preferentially interact with Ub 11R and Ub 11K (Cluster 2, 3), suggesting that in this case the
144 positive charge at position 11 is an important determinant for binding. To validate the AE-MS
145 results, we analyzed some of the interactions by AE followed by western blot analysis
146 (Supplementary Figures 3a, 3b) and by using bacterially expressed proteins (NDP52[11],
147 UCHL3[11]; Supplementary Figure 3c). Reassuringly, we found a high correlation between the
148 interaction patterns obtained by the different methods. For instance, NDP52 showed
149 preferred binding to Ub 11TFAcK, Ub 11KeK, and Ub 11AcK, independent of the method
150 used. Similarly, UCHL3 showed the strongest interaction with Ub 11K and 11R and the
151 weakest interaction with Ub 11Q and Ub 11A.

152 The above results clearly indicate that the identity of the side chain at position 11
153 affects the protein interaction properties of Ub and that Ub variants with similar side chains
154 have similar interaction properties (Figure 3b). To further support this notion, we performed a
155 Pearson correlation analysis, which indicates the degree of similarity or dissimilarity between

156 two different data sets (Figure 3c). Considering all 338 significant interactors, this revealed a
157 high correlation between Ub 11R and Ub 11K and between Ub 11A and Ub 11Q, which
158 reflects the significance of a positive charge at this position. Yet, there is no significant
159 correlation of Ub 11A/11Q with Ub 11AcK, Ub 11TFAcK, and Ub 11KeK, while there is a high
160 correlation between the latter three Ub variants in general and between Ub 11AcK und Ub
161 11KeK in particular (Figure 3c; see also Supplementary Figure 3d). This again indicates that
162 the structural change induced by elongation of the side chain of K11 by acetylation has a
163 major impact on the protein interaction properties of Ub and that this structural and, in
164 consequence, functional change is most reliably mimicked by TFAcK and KeK and not or
165 only partially by Q.
166 Taken together, the results obtained show that Ub variants harboring TFAcK, KeK or
167 AcK are highly similar in their propensity to interact with other proteins.
168
169 **Incorporation of AcK and its non-hydrolyzable analogs into p53 in human cells**
170 From the results obtained with Ub, we reasoned that KeK and TFAcK are suitable AcK
171 surrogates for studies in cells. Since Ub is a highly abundant protein in cells rendering
172 analysis of ectopically expressed Ub rather difficult, we switched to the tumor suppressor p53
173 to test this assumption. p53 is subjected to various PTMs affecting its stability and/or its
174 transcriptional activities resulting in different biological outcomes (e.g. cell cycle arrest,
175 apoptosis, senescence) (summarized in[24-27]). Most importantly for our purpose, the
176 acetylation status of distinct K residues of p53, including K residues in its DNA binding
177 domain (DBD), contributes to determining the selectivity or preference of p53 for certain
178 target genes and thereby the eventual transcriptional program induced by p53[17-19].
179 To determine whether KeK and TFAcK are suitable AcK surrogates in cells, we
180 generated expression constructs for human p53, in which the lysine codons at position 120
181 and/or 164 within the DBD were replaced by the amber stop codon TAG (p53-120 and p53-
182 164, respectively; note that K120 directly contacts DNA, while K164 does not). We chose
183 these positions, since besides their physiological relevance[17-19] (see Discussion),

184 incorporation of unnatural or non-canonical amino acids by ACS is rarely quantitative,
185 resulting in a mixture of full-length and truncated versions of the POI. The more N-terminal
186 the truncation occurs, the easier it is to distinguish between the full-length and the truncated
187 protein, e.g. by western blot analysis. Furthermore, since in our case, the truncations occur in
188 the DBD, the truncated p53 versions are not of functional relevance - they can neither bind to
189 DNA nor oligomerize with full-length p53 - and, thus, do not need to be considered in the
190 analysis.

191 The p53 expression constructs were transfected together with an expression construct
192 encoding the CUA-tRNA as well as an aminoacyl-tRNA synthetase for AcK (AcK-RS)[28,29] and
193 its analogs TFAcK and KeK (for further details, see Methods) in the absence or presence of
194 increasing concentrations of AcK, TFAcK or KeK. As cell line, we used H1299 cells, since
195 they do not express endogenous p53[30]. As shown in Figure 4, the three non-canonical amino
196 acids were incorporated, resulting in the expression of full-length p53, with AcK and KeK
197 being more efficiently used or incorporated into p53 than TFAcK (Figures 4a-c). Furthermore,
198 the system is not limited to the incorporation of AcK or its analogs at a single position, but
199 can also be used to generate p53 variants harboring two acetylated lysine residues
200 (Supplementary Figure 4), expanding the analytic potential of the system to multiple or at
201 least double acetylated p53 variants.

202

203 **Incorporation of AcK, TFAcK or KeK has different effects on the transcriptional**
204 **transactivation properties of p53**

205 To test the functionality of the p53-120 variants, we determined their ability to transactivate
206 the transcription of a reporter construct containing the binding site or response element (RE)
207 for p53 of the *p21* gene[31,32]. As control, we used expression constructs encoding wild-type
208 (wt) p53 and the tumor-derived mutant p53-273H (substitution of R273 by H), which has lost
209 the sequence-specific DNA binding properties (ref[33], and references therein). While the
210 120AcK and 120TFAcK variants were capable of inducing expression of the reporter
211 construct, though with somewhat reduced efficiency as compared to wt p53, the 120KeK

212 variant was not, or only poorly active (Figure 5a). Moreover, the differences in the
213 transactivational potential of the different p53 variants were also reflected in the levels of
214 endogenous p21, indicating that the differences are not due to an artifact of the p21 reporter
215 construct (Figure 5b). Similar results were obtained with the p53-164 variants
216 (Supplementary Figures 5a, 5c).

217 At first glance, the finding that the KeK-containing p53 variants are significantly less
218 active than the AcK and TFAcK variants seems to be surprising. However, it may be
219 explained by the notion that in contrast to KeK, AcK and TFAcK are deacetylated in cells
220 (ref[15], and references therein) and, thus, the respective variants regain wt-like activity. To
221 prove this possibility, we purified the different p53 variants from transfected cells and
222 subjected them to Parallel Reaction Monitoring (PRM)-MS analysis[34]. The data obtained
223 indeed suggest that the AcK and TFAcK variants are deacetylated, though to different
224 extents, while the KeK variant is refractory to deacetylation as expected (Figure 5c;
225 Supplementary Figure 5d). The "partial resistance" of TFAcK is most likely due to the notion
226 that Zn-dependent deacetylases can remove the fluorinated acetyl group[15,23], while NAD-
227 dependent deacetylases cannot do so (see Discussion).

228 To exclude that the data obtained with KeK are the result of unwanted side effects, we
229 performed several control experiments. Firstly, we looked into the possibility that the mere
230 presence of KeK in the cell culture medium interferes with the transactivation activity of wt
231 p53. Yet, a negative effect was not observed (Supplementary Figure 5e). Secondly, p53-
232 101KeK was similarly active as wt p53, p53-101AcK, and p53-101TFAcK, showing that
233 incorporation of KeK per se does not interfere with p53 function (Supplementary Figures 5b,
234 5c). Thirdly, we previously reported that in vitro, p53-120KeK binds to the p21 response
235 element (p21-RE) with an efficiency similar to unmodified p53[35], indicating that incorporation
236 of KeK does not disrupt the conformational integrity of the DBD of p53.

237 Finally, we asked whether or not Q can serve as substitute for AcK, when studying p53
238 function(s) in cells. This showed that unlike p53-120KeK, p53-120Q behaved like wt p53 in
239 the p21 reporter assay (Figure 5d), suggesting that in case of K120 acetylation, Q is not a

240 suitable surrogate for AcK. In conclusion, our results obtained with Ub and p53 indicate that
241 KeK is the most suitable AcK analog to study the effect of lysine acetylation of a POI in vitro
242 and in cells.
243

244 **DISCUSSION**

245 Lysine acetylation is a prevalent PTM affecting many intracellular proteins[7,36-39]. Yet,
246 determining the functional consequences as well as the physiological importance of
247 acetylation for a POI remains a serious challenge. An obvious consequence of acetylation is
248 neutralization of the positive charge of lysine's side chain, the functional significance of which
249 can be determined by mutational analysis by replacing the respective lysine residue by
250 arginine or an uncharged residue such as glutamine or alanine. An alternative, but not
251 mutually exclusive possibility is that the eventual effect of acetylation is mediated by the
252 extension of the length of lysine's side chain and/or the actual presence of the acetyl group.
253 While this cannot be adequately addressed by conventional mutational analysis, the genetic
254 code expansion technology in general and ACS in particular provide excellent tools to do so.
255 Our comparative analysis of AcK and its analogs TFAcK and KeK and their effects on distinct
256 properties of Ub and p53 strongly supports this proposition.

257 In support of the notion that the effects of acetylation are not necessarily explained by
258 simple charge neutralization, we recently reported that acetylation of Ub at K11 results in
259 structural and, in consequence, functional changes that cannot be mimicked by substitution
260 of K11 by Q[11]. Here, we extended the functional comparison of Ub 11AcK and Ub 11Q by
261 determining their respective interactomes in vitro (Figure 3). By comparison with unmodified
262 Ub, Ub 11R, and Ub 11A, this clearly showed that even though the charge status at position
263 11 affects the Ub interactome (Figure 3c), acetylation of K11 has additional or different
264 effects on the interaction properties of Ub. Moreover, we show that incorporation of TFAcK
265 and KeK at position 11 affects the structure of Ub in a manner similar, but not identical to
266 AcK incorporation (Figure 2). These slight differences, together with intrinsic variations in
267 proteome-wide affinity enrichments, are reflected in highly similar, but not identical

268 interactomes of Ub 11AcK, Ub 11TFAcK, and Ub 11KeK (Figure 3, Supplemental Figure 3).

269 To determine the applicability of ACS-mediated incorporation of AcK and its analogs in
270 eukaryotic cells, we switched to p53, since acetylation of distinct lysine residues in its DBD
271 affects the transcriptional properties of p53[17-19] and since p53 activity can be readily
272 monitored in cells by transcriptional reporter assays. Our results show that AcK, TFAcK, and
273 KeK can be efficiently incorporated into p53 at different positions, with incorporation of
274 TFAcK being somewhat less efficient than those of the others for unknown reasons (note
275 that a similar observation was made for the efficiency of AcK and TFAcK incorporation by
276 others[15]). However, only p53-120KeK was resistant to the action of deacetylases in cells,
277 while p53-120AcK was quantitatively deacetylated and p53-120TFAcK was partially
278 deacetylated (Figure 5c). The "partial resistance" of TFAcK is likely due to the notion that Zn-
279 dependent deacetylases can remove the fluorinated acetyl group, while NAD-dependent
280 deacetylases cannot[15,23,40]. From a chemical perspective, TFAcK is more susceptible to
281 nucleophilic attacks than AcK, as the trifluoroacetyl group is a better leaving group than the
282 acetyl group[41]. However, while Zn-dependent deacetylases such as KDACs release the
283 acetyl group via nucleophilic attack of the C atom of the carbonyl group by an activated water
284 molecule, NAD-dependent deacetylases including sirtuins use NAD as electrophile for
285 nucleophilic attack of the carbonyl oxygen atom[40]. Thus, due to the electron withdrawing
286 effect of the three fluorine atoms, the nucleophilicity of the carbonyl oxygen is significantly
287 decreased[41]. Along this line, it has recently been shown that HDAC2 is able to deacetylate
288 FOXO4 variants harboring TFAcK or AcK at position 189, while Sirt2 can only deacetylate
289 the AcK variant[15]. In contrast to TFAcK, KeK is resistant against the attack of Zn-dependent
290 deacetylases as well[16,40]. An intermediate step in HDAC-catalyzed deacetylation is the
291 nucleophilic attack of the carbonyl C atom by a water molecule. To do so, a base,
292 presumably a histidine residue, accepts a proton from the water molecule and transfers it to
293 the nitrogen atom of the amide bond, which facilitates the final cleavage of the amide bond[40].
294 In contrast to the nitrogen, the carbon atom of KeK (Figure 1a) can most probably not
295 function as an acceptor for the proton. Thus, while in the case of KeK the nucleophilic attack

296 may still occur, though with reduced efficiency, cleavage of the C-C bond is highly unlikely. In
297 any case, this difference in the deacetylation susceptibility of AcK, TFAcK, and KeK may
298 represent an attractive possibility to dissect the contribution of the different deacetylase
299 families to the deacetylation of a POI.

300 The difference in acetylation status of the different p53-120 variants is also reflected in
301 different abilities to transactivate a p21-reporter construct as well as to induce expression of
302 endogenous p21 (Figure 5). p53-120AcK is similarly active as wt p53, p53-120TFAcK is
303 somewhat less active, and p53-120KeK does neither detectably transactivate the reporter
304 construct nor induce endogenous p21 expression. In addition and in line with published
305 data[42], p53-120R (substitution of K by R) efficiently transactivated the reporter construct.
306 Moreover, the 120Q variant behaved like wt p53 in the reporter assay, suggesting that in this
307 case, charge neutralization does not suffice to mimic the effect of acetylation. Our data with
308 p53-120KeK are in line with ChIP data indicating that within cells, p53 acetylated at K120
309 does not accumulate at the *p21* gene[43] and with the notion that acetylation of K120 does not
310 result in the induction of expression of the *p21* gene (ref[18,19], and references therein). Yet,
311 similar to our data with p53-120KeK[35], it has previously been shown that in vitro, p53-120AcK
312 generated by the ACS method binds to the p21-RE with an affinity similar to unmodified
313 p53[32]. How can these apparent contradictory data be reconciled? In a simplified view, cell
314 cycle arrest target genes such as *p21* contain "high affinity" REs for p53, while pro-apoptotic
315 target genes contain "low affinity" REs (ref[42], and references therein). Furthermore, there is
316 data to indicate that p53 contacts high affinity REs and low affinity REs in a distinct manner
317 and that binding of p53-120AcK does not affect the conformation of high affinity REs, but
318 changes the conformation of low affinity REs[32,42,44]. Moreover, there is evidence that the
319 conformation of p53-120AcK bound to high affinity REs is different to the one bound to low
320 affinity REs[32]. Such conformational changes that depend on both p53 acetylation and binding
321 of p53 to DNA in a sequence-specific manner may affect the protein-protein interaction
322 properties of p53 and thereby its ability to differentially transactivate p53 target genes.

323 Whereas K120 acetylation has been associated with the "apoptotic program" of p53 by

324 inducing expression of pro-apoptotic genes such as *BAX* and *PUMA*, K164 acetylation has
325 been involved in the "cell cycle program" by inducing the expression of cell cycle arrest
326 genes including *p21*[17-19,43,45]. Yet, p53-120KeK or p53-164KeK were inactive in our reporter
327 assays, irrespective of whether a p21 reporter construct or a PUMA reporter construct was
328 used (Figure 5, Supplementary Figures 5 and 6). As discussed above, all of our data indicate
329 that KeK structurally and functionally resembles AcK and, thus, its incorporation should affect
330 p53 activity in a manner similar to lysine acetylation. Yet, we cannot exclude that proteins
331 exist that bind to KeK with lower affinity than to AcK. For instance, bromodomains are known
332 binders of AcK-containing proteins[46]. Thus, there may be bromodomain-containing proteins
333 or other proteins that specifically bind to p53-120AcK, and such interactions may be critically
334 involved in p53-mediated transactivation of the p21 and/or PUMA reporter constructs.
335 However, this does not appear to be the case, since p53-120R efficiently induced expression
336 of both the p21 reporter construct and the PUMA reporter construct. Irrespectively, a likely
337 conclusion of the data obtained with p53-120KeK (p53-TFAcK), p53-120Q, and p53-120R is
338 that a positive charge at position 120 is not critical for p53-mediated transcription of the p21
339 reporter construct and, thus, the negative effect of K120 acetylation cannot be explained by
340 charge neutralization. For the PUMA reporter construct, charge neutralization at position 120
341 contributes to the negative effect of KeK incorporation, but additional properties of KeK (and
342 thus, AcK) appear to be involved. Taken together, our results obtained with the different p53
343 variants indicate that a comparative analysis of POI variants, in which a distinct K residue is
344 replaced by R, Q, or AcK surrogates, provides critical information about if and how
345 acetylation affects known properties/functions of a POI and that KeK represents the AcK
346 mimic of choice for such studies.
347 A possibility to explain the apparent discrepancy between our data and published data
348 is that our reporter constructs miss additional sequence elements of the respective
349 endogenous genes. At least in case of p21, this seems unlikely, since endogenous p21
350 expression was also not induced by p53-164KeK (Supplementary Figure 5c). Furthermore,
351 many of the published data have been obtained by mutational analysis, for instance by

352 replacing respective lysine residues by arginine (summarized in refs[16-18]). Thus, some of the
353 effects ascribed to lysine acetylation may in part be due to other lysine modifications such as
354 mono-ubiquitination[47,48] or lactylation[49]. Other possibilities would be that mono-acetylation of
355 p53 or acetylation in general is not sufficient for the induction of certain target genes. In other
356 words, it seems likely that a combination of acetylation with additional PTMs such as
357 phosphorylation is required to observe the respective transcriptional effects, a scenario that
358 may not be achieved in our experimental set-up. For instance, it could be envisioned that
359 such phosphorylation events occur only under certain stress conditions. An alternative, but
360 not mutual exclusive possibility is that to be functional, such PTMs have to occur in a timely
361 ordered, i.e. consecutive manner. Along this line, prior lysine acetylation, in particular
362 incorporation of an AcK mimic that cannot be hydrolyzed as in our case, may prevent
363 phosphorylation of p53 at residues required for transcriptional transactivation. To address
364 such possibilities, the continuous development and refinement of the genetic code expansion
365 technology to allow the combinatorial incorporation of different non-canonical amino acids
366 resembling non-reversible, yet faithful analogs of respective PTMs should prove highly
367 valuable.
368
369 **METHODS**
370 **Synthesis of 2-amino-8-oxononanoic acid (Ketolysine; KeK)**
371 KeK was synthesized as described previously[16], except for the final purification step. Instead
372 of using a Dowex column, KeK was purified twice via MPLC and then lyophilized. For MPLC
373 purification, a PrepChrom C-700 (Büchi) with SuperVarioFlash® D26 RP18 30 g or D40
374 RP18 90 g (Götec-Labortechnik GmbH) were used. After sample application, the column was
375 washed with H_2O for 20 min and then a linear gradient (0%-100% acetonitrile in 20 min) of
376 acetonitrile in H_2O was applied with a flow rate of 20 ml/min. Runs were monitored by UV
377 detection at 254 nm. Fractions containing KeK were pooled and lyophilized.
378 To record NMR spectra of KeK, a Bruker Avance III 400 (400 MHz) was used. Spectra
379 were measured in the presence of D_2O (the deuterium resonance signal is used as internal

380 reference frequency for other nuclei) at room temperature. The program MestReNova 14.2
381 was used for spectrum evaluation. Shifts are indicated in ppm (σ-scale) and coupling
382 constants in Hertz (Hz).

383 ^{1}H NMR (400 MHz, D_2O): δ [ppm] = 3.88 (t, J = 6.1 Hz, 1H, H-1), 2.62 (t, J = 7.3 Hz,
384 2H, H-6) 2.25 (s, 3H, CH3), 2.06-1.82 (m, 2H, H-2), 1.61 (p, J = 7.4 Hz, 2H, H-5), 1.53-1.31
385 (m, 4H, H-3 & H-4).

386 ^{13}C NMR (101 MHz, D_2O): δ [ppm] = 217.56 (CO), 174.68 (CO2H), 54.65 (C-1), 43.13
387 (C-6), 30.20 (C-2), 29.31 (CH3), 27.88 (C-4), 24.08 (C-3), 22.98 (C-5).

388

389 **Plasmids**

390 Codon-optimized cDNA encoding Ub with a C-terminal His_6-tag was inserted into the pGEX-
391 2TK backbone replacing the GST cDNA to yield pKS Ub-His[50]. cDNAs of Ub variants, in
392 which the amber stop codon (TAG), a glutamine, an alanine or arginine codon replaced the
393 respective lysine codon, were generated via site-directed mutagenesis.

394 The cDNA encoding the aminoacyl-tRNA synthetase for AcK (AcK-RS) derived from
395 *Methanomethylophilus alvus* was kindly provided by Dr. K. Lang (ETH Zürich) and was
396 inserted into the pRSFDuet vector (Merck Millipore) for bacterial expression. Plasmids
397 encoding tRNA and AcK-RS derived from *Methanosarcina mazei* for incorporation of
398 AcK/TFAcK/KeK in H1299 cells were kindly provided by Dr. Simon Elsässer (Karolinska
399 Institute, Stockholm)[28,29].

400 cDNAs encoding wild-type p53 or the tumor-derived mutant 273H with and without C-
401 terminal StrepII tag was inserted into pcDNA3. cDNAs for p53 variants, in which the TAG
402 stop codon, a glutamine or an arginine codon replaced the respective lysine codon, were
403 generated via site-directed mutagenesis.

404 Reporter constructs encoding luciferase were generated by introducing p53-responsive
405 elements (p21, PUMA)[32] into the pGL4 luciferase reporter system.

406

407

408 **Bacterial expression and protein purification**

409 Human His_6-tagged UBA1[50,51], N-terminal His_6-tagged UCHL3[50], and C-terminal His_6-tagged
410 UbcH5b[50,52] were expressed and purified as described.

411 Expression of the various GST fusion proteins was performed in *E. coli* BL21 (DE3)
412 Rosetta. GST-SUMO fusion proteins of HA-tagged HDM2 and NDP52 were expressed and
413 purified as described[11].

414

415 **Bacterial generation of modified Ub and unmodified Ub**

416 For generation of Ub site-specifically modified at position 11 (with and without an N-terminal
417 Strep II tag), the respective pKS vector was co-transformed with pRSFDuet harboring the
418 AcK-RS/t-RNA pair for amber codon suppression into *E. coli* B834 (DE3) cells. Cells were
419 grown at 37 °C in 100 ml LB medium containing 100 µg/l carbenicillin/ 34 µg/l kanamycin
420 until OD_{600} = 1 was reached. Then, cultures were diluted to an OD_{600} = 0.1 with minimal
421 media devoid of phenylalanine and further incubated. At an OD_{600} = 0.4, either 10 mM AcK
422 (abcr) or 20 mM TFAcK (abcr) or 10 mM KeK (racemate) and 20 mM nicotinamide (Sigma-
423 Aldrich) were added. Gene expression was induced at OD_{600} = 0.7 by the addition of 1 mM
424 IPTG (final conc.). After incubation at 25 °C for 16 h, cells were harvested by centrifugation,
425 resuspended in lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 30 mM imidazole, 1 %
426 Triton X-100, 1 mg/ml aprotinin/leupeptin, 1 mg/ml Pefabloc) and lysed by sonication. The
427 lysate was cleared, and modified Ub-His was purified using a 5 ml HisTrap HP (Äkta pure25,
428 GE Healthcare) with a step from 30 mM to 242 mM imidazole in 50 mM Tris-HCl pH 7.4, 150
429 mM NaCl for elution. Protein-containing fractions were pooled, UCHL3 was added to a final
430 concentration of 75 µg/ml to remove the C-terminal His-tag, and the samples were incubated
431 for 4 h. After cleavage, modified Ub variants were dialyzed against 25 mM acetate pH 4.5
432 overnight at 4 °C. Ub samples were further purified via a 1 ml HiTrap SP HP (GE Healthcare)
433 with a linear gradient from 25 mM acetate pH 4.5 to 25 mM acetate 4.5, 1 M NaCl. Modified
434 Ub containing fractions were pooled and dialyzed against 25 mM Tris-HCl pH 7.5, 50 mM
435 NaCl. The concentration was determined with the BCA Protein Assay Kit (Thermo) and

436 commercial Ub (Sigma-Aldrich) as standard. The correct masses of the modified Ub variants
437 were confirmed by ESI-MS (Supplementary Figure 1).
438 Expression and purification of unmodified Ub and the Ub variants containing Q, A or R
439 instead of K at position 11 were performed similar to the modified Ub variants, except that
440 AcK/TFAcK/KEK and nicotinamide were not added to the respective expression cultures.
441
442 ***In vitro* autoubiquitination assay**
443 For *in vitro* autoubiquitination, 100 nM UBA1, 0.5 µM UbcH5b, and 0.55 µM HDM2 were
444 incubated in 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM DTT, 2 mM ATP, and 2 mM $MgCl_2$.
445 The reaction was started by addition of 7.8 µM (final conc.) of the respective Ub variant and
446 incubated at 37 °C for the times indicated. Reactions were stopped by the addition of 5x
447 reducing SDS loading buffer. Entire reaction mixtures were electrophoresed in 12.5% SDS-
448 polyacrylamide (SDS-PA) gels followed by Coomassie blue staining.
449
450 **Generation of isotopically labeled Ub for NMR analysis**
451 For preparation of ^{15}N-labeled proteins, *E. coli* harboring the expression construct for Ub with
452 a TAG stop codon at position 11 was grown in LB medium at 37 °C until OD_{600} = 1. This
453 culture was used to inoculate ^{15}N M9 medium (33.7 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.55 mM
454 NaCl, 9.35 mM $^{15}NH_4Cl$, 1 mM $MgSO_4$, 0.3 mM $CaCl_2$, 134 µM EDTA, 31 µM $FeCl_3$, 6.2 mM
455 $ZnCl_2$, 0.76 µM $CuCl_2$, 0.42 µM $CoCl_2$, 1.62 µM H_3BO_3, 0.081 µM $MnCl_2$, 1 mg/l biotin, 1 mg/l
456 thiamine, 0.4% glucose (w/v)). Upon growth overnight at 37 °C, 100 ml of the culture were
457 diluted with 900 ml of pre-warmed ^{15}N M9 medium. Cells were grown at 37 °C until OD_{600} =
458 0.4 was reached, and then either 10 mM AcK (abcr), 20 mM TFAcK (abcr) or 10 mM KeK
459 (racemate) and 20 mM nicotinamide (Sigma-Aldrich) (final conc.) were added. Gene
460 expression was induced at OD_{600} = 0.7 by the addition of 1 mM IPTG (final conc.). After
461 incubation at 25 °C for 16 h, cells were harvested by centrifugation. Purification of isotopically
462 labeled Ub was performed as described above for unlabeled Ub.
463

464 **Acquisition of NMR spectroscopic data**

465 NMR experiments were performed at *T* = 298 K on a Bruker Avance Neo 800 MHz
466 spectrometer equipped with a quadruple (QCI) cryo probe with z-gradient. Datasets were
467 processed using NMRPipe[53] and analyzed by NMRView[54]. Samples were prepared in 20 mM
468 NaH_2PO_4 (pH 6.8) supplemented with 5% (v/v) D_2O, which serves as frequency lock. All
469 measurements were performed with singly ^{15}N-labeled variants of Ub 11AcK, 11TFAcK and
470 11KeK using protein concentrations of 33 µM, 500 µM and 60 µM, respectively. For each
471 protein, a two-dimensional heteronuclear ^{1}H-^{15}N HSQC spectrum was collected with 1024
472 and 256 increments in the direct ^{1}H and indirect ^{15}N dimension, respectively, using 8-32
473 transients depending on the protein concentration. WATERGATE pulse sequence and
474 presaturation were applied for efficient solvent suppression.

475

476 **Backbone resonance assignment and calculation of chemical shift perturbations**

477 The backbone resonance assignment of Ub 11AcK was transferred from ref[11] and served as
478 basis for the assignment of backbone amide resonances of Ub 11TFAcK and 11KeK. Due to
479 the slight differences in peak position, this could be successfully accomplished by visual
480 estimate when the two-dimensional ^{1}H-^{15}N HSQC NMR spectra of Ub 11TFAcK and 11KeK
481 were directly compared to Ub 11AcK.

482 Weighted chemical shift perturbations between the different Ub variants and
483 unmodified Ub were calculated according to the following equation[55]:

$$\Delta\omega = \sqrt{\frac{(\Delta^{1}\mathrm{H})^{2}+\frac{1}{25}(\Delta^{15}\mathrm{N})^{2}}{2}}, \quad (1)$$

484 where Δ^{1}H is the peak shift in proton and Δ^{15}N is the peak shift in nitrogen dimension,
485 respectively, between two corresponding peaks. Note that the assignment of backbone
486 resonances arising in the two-dimensional heteronuclear ^{1}H-^{15}N HSQC NMR spectrum
487 related to unmodified Ub was also taken from ref[11].

488

489

490 **Cell culture and extract preparation**

491 HEK293T and H1299 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM,
492 Thermo Fisher) supplemented with 10% fetal bovine serum (FBS, Sigma) at 37 °C and 5%
493 CO_2 atmosphere.

494 For affinity enrichment experiments, HEK293T and H1299 cell pellets were
495 resuspended in cell lysis buffer (1x PBS, 2 mM $MgCl_2$, 1 mM DTT, supplemented with 1
496 mg/ml aprotinin/leupeptin, 1 mg/ml Pefabloc) and subsequently lysed by sonication. After
497 centrifugation (21,000x g, 30 min, 4 °C), the total protein concentration of the supernatant
498 was determined with the BCA Protein Assay Kit (Thermo Fisher) and freshly dissolved BSA
499 as standard.

500 For transactivation studies, H1299 cell pellets were resuspended in lysis buffer (100
501 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM DTT, 1% (v/v) Nonidet P-40, supplemented with 1
502 mg/ml aprotinin/leupeptin, 1 mg/ml Pefabloc) and incubated for 30 min at 4°C prior to
503 centrifugation (16,100x g, 30 min, 4°C). Cleared whole cell extracts were analyzed by
504 western blot and luciferase reporter assay as described below.

505

506 **Affinity enrichment**

507 For affinity enrichment of Ub interacting proteins, 25 µg of each Strep-Ub variant was
508 incubated with 2.5 mg whole cell lysate for 10 min on ice. After washing 50 µl Strep-Tactin®
509 beads (iba) with three times 200 µl cell lysis buffer, the mixture was added to the beads.
510 Incubation at room temperature for 3 h in an overhead shaker was followed by 5 washing
511 steps with 200 µl cell lysis buffer each. Bound proteins were eluted by incubating the beads
512 for 5 min on ice with two times 150 µl and one time 100 µl elution buffer (100 mM Tris-HCl
513 pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin). The eluates were combined and
514 freeze-dried. Each experiment was performed in quadruplicates.

515 The freeze-dried eluate was resuspended in 100 µl 8 M urea, reduced with 5 mM
516 TCEP for 30 min at 37 °C and alkylated with 10 mM iodoacetamide for 30 min at room
517 temperature. After dilution to 1 M urea with 50 mM NH_4HCO_3, 5 µg trypsin were added, and

518 proteins were digested for 20 h at 37 °C. The digested samples were freeze-dried and kept
519 at -20 °C until they were prepared for MS measurement.
520 Prior to MS measurement, the tryptic peptides were dissolved in 50 µl 0.1% TFA in
521 H_2O, acidified with 10% TFA and desalted using Pierce™ C18 spin tips (Thermo Fisher).
522 For affinity enrichment of ectopically expressed Strep-tagged p53, transiently
523 transfected cells were harvested and lysed as described above. Whole cell extracts were
524 added to 20 µL Strep-Tactin® XT beads (iba) pre-equilibrated in lysis buffer. After incubation
525 at 4°C for 3 h in an overhead shaker, beads were washed 5 times with 200 µl lysis buffer
526 each. Bound proteins were eluted as stated above for Ub with elution buffer (100 mM Tris-
527 HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 50 mM biotin) and lyophilized. Freeze dried eluates
528 were resuspended in SDS loading buffer and subjected to SDS-PA gel electrophoresis for in-
529 gel digestion. SDS-PA gels were stained by colloidal Coomassie blue, and gel pieces
530 containing p53 were excised and digested according to ref[56]. In brief, gel pieces were
531 washed in 100 mM NH_4HCO_3 for 10 min, destained in 50 mM NH_4HCO_3, 50% (v/v)
532 acetonitrile and subsequently dehydrated in pure acetonitrile. Disulfide bonds were reduced
533 by addition of 10 mM TCEP in 100 mM NH_4HCO_3 for 60 min at 56 °C followed by
534 dehydration in pure acetonitrile and subsequent alkylation in 50 mM iodoacetamide in 100
535 mM NH_4HCO_3 for 60 min at room temperature in the dark. Following dehydration in
536 acetonitrile, 5 µg trypsin were added, and proteins digested for 20 h at 37 °C. Peptides were
537 extracted in 30% acetonitrile, 5% formic acid and 60% acetonitrile, 5% formic acid.
538 Combined extracts were freeze dried and kept at -20 °C until preparation for MS
539 measurements. Prior to MS measurement, the tryptic peptides were dissolved and desalted
540 as stated above.
541
542 **Mass spectrometric analysis**
543 Peptide samples were analyzed on a Q-Exactive HF mass spectrometer (Thermo Fisher
544 Scientific, Bremen, Germany) interfaced with an Easy-nLC 1200 nanoflow liquid
545 chromatography system (Thermo Scientific, Odense, Denmark). Samples were reconstituted

546 in 0.1% formic acid and loaded onto a C18 analytical column (75 µm × 15 cm). Peptides
547 were resolved at a flow rate of 300 nl/min using a linear gradient of 6−40% solvent B (0.1%
548 formic acid in 80% acetonitrile) over 165 min. Data-dependent acquisition with full scans over
549 a 350−1500 m/z range was carried out using the Orbitrap mass analyzer at a mass
550 resolution of 60,000 at 200 m/z, automatic gain control target value of $3e^6$, and max injection
551 time of 60 ms. The 10 most intense precursor ions were selected for further fragmentation.
552 Only peptides with charge states 2−6 were used, and dynamic exclusion was set to 30 sec.
553 Precursor ions were fragmented using higher-energy collision dissociation (HCD) with a
554 normalized collision energy (NCE) set to 28%. Fragment ion spectra were recorded at a
555 resolution of 15,000, automatic gain control target value of $1e^5$, and max. injection time 60
556 ms. Each of the biological replicates was measured in technical duplicates.
557 For label-free quantification, the raw files from LC-MS/MS measurements were
558 analyzed using MaxQuant[57] (version 1.6.8) with default settings and match between runs and
559 label-free quantification (LFQ) enabled. For protein identification, the human reference
560 proteome downloaded from the UniProt database (download date: 2018-02-22) was used.
561 For further data processing, Perseus software[58] (version 1.6.10.50) was used. Identified
562 proteins were filtered for reverse hits and common contaminants. LFQ intensities were log2
563 transformed. Only proteins that were detected in at least 4 out of 6 replicates (three biological
564 replicates, each measured as technical duplicate) were retained for further analysis. Missing
565 values were imputed from normal distribution (width = 0.3 and shift = 1.2) based on the
566 assumption that these proteins were just below the detection limit. Significantly enriched
567 proteins were identified by an ANOVA test (S0 = 2, FDR = 0.005) and Z-score normalized.
568 After averaging of groups by using the median of all replicates, enriched proteins were
569 clustered by correlation and plotted as heatmaps. Additional annotations (GO annotations,
570 KEGG, Pfam) and identification of enriched terms was done with Perseus, as described
571 previously[59]. The top 100 significantly enriched proteins were additionally used to compare
572 interactors between Ub 11AcK and Ub 11KeK and were plotted as Venn diagram.
573 For intact protein mass measurements, a micrOTOF II (Bruker) interfaced with an

574 Agilent 1260 Infinity II liquid chromatography system was used. The proteins were loaded
575 onto an analytical column (Nucleodur 300-5 C4 ec, EC 150/4, Macherey-Nagel) and resolved
576 at a flow rate of 300 µl/min using a linear gradient from 5% to 100% solvent B (0.1% formic
577 acid in acetonitrile) over 15 min. The electrospray ion source was operated in positive
578 ionization mode, scan range 650−2100 m/z at 1 Hz acquisition rate and 2x rolling average
579 was used. Compass Data Analysis software (Bruker) was used for re-calibration and spectra
580 deconvolution.
581 To quantify the p53 modification status, parallel reaction monitoring (PRM) using an
582 inclusion list and targeted MS2 mode was employed. Peptides were separated on a 15 cm
583 Acclaim PepMap C18 column (P/N 164943, Thermo Scientific) across a 50 min gradient from
584 2% to 45% solvent B (0.1% formic acid in 80% acetonitrile). A full mass spectrum was
585 acquired on an Orbitrap Tribrid Fusion (Thermo Fisher Scientific, Bremen, Germany) at 120
586 K resolution over a mass range of 330-1250 m/z (adapted to precursors in inclusion list). The
587 maximum injection time was set to 50 ms and an AGC target of 4e5. Precursors were
588 isolated sequentially according to the inclusion list with an isolation window width of 1.6 m/z
589 and fragmented by HCD at 30% NCE. Fragment mass spectra were acquired in the Orbitrap
590 at 30 K resolution and an AGC target of 5e4. PRM data were analyzed via Skyline software
591 23.1.0.268. MS1-filtering isotope precursor ion peak was set to 3 and peptides were
592 quantified on the MS2-level. Obtained values were plotted with GraphPad Prism 6.01.
593
594 ***In vitro* binding assay with recombinant proteins**
595 Verifications of the protein interactions with the respective Strep-Ub variants were performed
596 as described previously[11]. In brief, 5 µM of the respective Strep-Ub variant were incubated 10
597 min on ice in binding buffer (1x PBS, 10 mM $MgCl_2$, 5 mM DTT) with 5 µM of recombinant
598 UCHL3 or NDP52. Then, 5 µl Strep-Tactin® beads (iba) pre-equilibrated in binding buffer
599 were added and the mixture incubated for 3 h at room temperature. After washing, bound
600 proteins were eluted twice by incubating the beads with 30 µl elution buffer (200 mM Tris-HCl
601 pH 8.0, 300 mM NaCl, 4 mM EDTA, 10 mM desthiobiotin). Eluates were combined,

602 subjected to SDS-PA gel electrophoresis, and analyzed by Coomassie blue staining and/or
603 western blot analysis.
604
605 **Transient transfection and incorporation of AcK, TFAcK, and KeK into p53**
606 2-3 x 10^5 cells were seeded in 6-well plates (Sarstedt) 24 h before transfection. Cells were
607 transfected at an approximate confluency of 80% with Lipofectamine 2000 (Invitrogen-
608 Thermo Fisher Scientific) according to manufacturer's instructions with plasmids encoding
609 the AcK-RS/tRNA pair[28,29] required for incorporation of AcK and its analogs TFAcK and KeK
610 and an expression construct for p53 harboring a TAG stop codon at the position indicated. In
611 addition, an expression construct encoding β-galactosidase was cotransfected to determine
612 relative transfection efficiencies. Unless stated otherwise, 1 mM AcK, 8 mM TFAcK or 1mM
613 KeK (final conc.) were added at the time of transfection, and cells were harvested after 24
614 hours. After cell lysis, cleared whole cell extracts were subjected to β-galactosidase assay
615 and β-galactosidase normalized aliquots were electrophoresed on 12.5% SDS-PA gels
616 followed by western blot analysis with an anti-p53 antibody (DO-I) and an HRP-coupled anti-
617 mouse secondary antibody.
618
619 **Luciferase reporter assays**
620 Cells were transfected as stated above with expression constructs for the p53 variants
621 indicated, the AcK-RS/tRNA pair (if required), and β-galactosidase, and a p53-responsive
622 luciferase reporter construct (p21, PUMA). 24 h upon transfection, whole cell extracts were
623 prepared and 10 µl aliquots of each extract were transferred in duplicates to a black 96-well
624 plate and luciferase activity automatically determined by a micro-plate reader (Wallac 1420
625 VICTOR³ Multilabel plate reader, PerkinElmer). In brief, 100 µL of luciferase reaction mix
626 (470 µM D-luciferin, 530 µM ATP, 270 µM CoA, 200 µM EDTA, 20 mM tricine, 2.67 mM
627 $MgSO_4$, 1.07 mM $MgCO_3$, 33.3 mM DTT, pH 7.5) were dispensed into each well and the
628 reaction mixture incubated for 5 min. Luciferase activity was assessed by measuring light
629 emission for 10 seconds.

630 **DATA AVAILABILITY**

631 The data that support this study are available from the corresponding authors upon
632 reasonable request. The NMR solution structure and the crystal structure of ubiquitin used in
633 this study are available in the ProteinDataBank under the accession codes1D3Z and 1UBQ,
634 respectively. The mass spectrometry proteomics data have been deposited to the
635 ProteomeXchangeConsortium via the PRIDE repository[60] identifiers PXD056021 (Username:
636 reviewer_pxd056021@ebi.ac.uk; Password: AWE0DQbBpn9D) for the PRM data and
637 PXD056063 (Username: reviewer_pxd056063@ebi.ac.uk; Password: tCMcGvYVb79f) for
638 the AE-MS data.

639

640 **REFERENCES**

641 1. Ramazi, S. & Zahiri, J. Post-translational modifications in proteins: Resources, tools and
642 prediction methods. *Database-the Journal of Biological Databases and Curation* (2021).
643 2. Aebersold, R. et al. How many human proteoforms are there? *Nat Chem Biol* **14**, 206-214
644 (2018).
645 3. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and
646 function. *Nature* **537**, 347-355 (2016).
647 4. Conibear, A.C. Deciphering protein post-translational modifications using chemical biology
648 tools. *Nat Rev Chem* **4**, 674-695 (2020).
649 5. Peng, T., Das, T., Ding, K. & Hang, H.C. Functional analysis of protein post-translational
650 modifications using genetic codon expansion. *Protein Sci* **32**, e4618 (2023).
651 6. Barber, K.W. & Rinehart, J. The abcs of ptms. *Nat Chem Biol* **14**, 188-192 (2018).
652 7. Narita, T., Weinert, B.T. & Choudhary, C. Functions and mechanisms of non-histone protein
653 acetylation. *Nat Rev Mol Cell Biol* **20**, 156-174 (2018).
654 8. Wang, Z.A. & Cole, P.A. The chemical biology of reversible lysine post-translational
655 modifications. *Cell Chem Biol* **27**, 953-969 (2020).
656 9. Knyphausen, P., Lang, F., Baldus, L., Extra, A. & Lammers, M. Insights into k-ras 4b regulation
657 by post-translational lysine acetylation. *Biol Chem* **397**, 1071-1085 (2016).
658 10. Okada, A.K. et al. Lysine acetylation regulates the interaction between proteins and
659 membranes. *Nat Commun* **12**, 6466 (2021).
660 11. Kienle, S.M. et al. Electrostatic and steric effects underlie acetylation-induced changes in
661 ubiquitin structure and function. *Nat Commun* **13**, 5435 (2022).
662 12. Liu, C.C. & Schultz, P.G. Adding new chemistries to the genetic code. *Annu Rev Biochem* **79**,
663 413-444 (2010).
664 13. Chin, J.W. Expanding and reprogramming the genetic code of cells and animals. *Annu Rev*
665 *Biochem* **83**, 379-408 (2014).
666 14. Neumann, H., Peak-Chew, S.Y. & Chin, J.W. Genetically encoding n(epsilon)-acetyllysine in
667 recombinant proteins. *Nat Chem Biol* **4**, 232-234 (2008).
668 15. Miller, G.M., Flynn, E.M., Tom, J., Song, A. & Cochran, A.G. Trifluoroacetyl lysine as a
669 bromodomain binding mimic of lysine acetylation. *ACS Chem Biol* **17**, 1022-1029 (2022).
670 16. Huang, Y. et al. Genetic incorporation of an aliphatic keto-containing amino acid into proteins
671 for their site-specific modifications. *Bioorg Med Chem Lett* **20**, 878-880 (2010).

672 17. Reed, S.M. & Quelle, D.E. P53 acetylation: Regulation and consequences. *Cancers (Basel)* **7**, 30-
673 69 (2014).
674 18. Wen, J. & Wang, D. Deciphering the ptm codes of the tumor suppressor p53. *J Mol Cell Biol* **13**,
675 774-785 (2022).
676 19. Xia, Z., Kon, N., Gu, A.P., Tavana, O. & Gu, W. Deciphering the acetylation code of p53 in
677 transcription regulation and tumor suppression. *Oncogene* **41**, 3039-3050 (2022).
678 20. Swatek, K.N. & Komander, D. Ubiquitin modifications. *Cell Res* **26**, 399-422 (2016).
679 21. Yau, R. & Rape, M. The increasing complexity of the ubiquitin code. *Nat Cell Biol* **18**, 579-586
680 (2016).
681 22. Dikic, I. & Schulman, B.A. An expanded lexicon for the ubiquitin code. *Nat Rev Mol Cell Biol* **24**,
682 273-287 (2023).
683 23. Zhang, F. et al. A genetically encoded (19)f nmr probe for lysine acetylation. *Chem Commun*
684 *(Camb)* **54**, 3879-3882 (2018).
685 24. Kruiswijk, F., Labuschagne, C.F. & Vousden, K.H. P53 in survival, death and metabolic health: A
686 lifeguard with a licence to kill. *Nat Rev Mol Cell Biol* **16**, 393-405 (2015).
687 25. Kastenhuber, E.R. & Lowe, S.W. Putting p53 in context. *Cell* **170**, 1062-1078 (2017).
688 26. Hafner, A., Bulyk, M.L., Jambhekar, A. & Lahav, G. The multiple mechanisms that regulate p53
689 activity and cell fate. *Nat Rev Mol Cell Biol* **20**, 199-210 (2019).
690 27. Liu, Y. & Gu, W. The complexity of p53-mediated metabolic regulation in tumor suppression.
691 *Semin Cancer Biol* **85**, 4-32 (2022).
692 28. Elsässer, S.J. Generation of stable amber suppression cell lines. *Methods Mol Biol* **1728**, 237-
693 245 (2018).
694 29. Schmied, W.H., Elsässer, S.J., Uttamapinant, C. & Chin, J.W. Efficient multisite unnatural amino
695 acid incorporation in mammalian cells via optimized pyrrolysyl trna synthetase/trna
696 expression and engineered erf1. *J Am Chem Soc* **136**, 15577-15583 (2014).
697 30. Mitsudomi, T. et al. P53 gene mutations in non-small-cell lung cancer cell lines and their
698 correlation with the presence of ras mutations and clinical features. *Oncogene* **7**, 171-180
699 (1992).
700 31. el-Deiry, W.S. et al. Waf1, a potential mediator of p53 tumor suppression. *Cell* **75**, 817-825
701 (1993).
702 32. Arbely, E. et al. Acetylation of lysine 120 of p53 endows DNA-binding specificity at effective
703 physiological salt concentration. *Proc Natl Acad Sci U S A* **108**, 8251-8256 (2011).
704 33. Bullock, A.N., Henckel, J. & Fersht, A.R. Quantitative analysis of residual folding and DNA
705 binding in mutant p53 core domain: Definition of mutant states for rescue in cancer therapy.
706 *Oncogene* **19**, 1245-1256 (2000).
707 34. Peterson, A.C., Russell, J.D., Bailey, D.J., Westphall, M.S. & Coon, J.J. Parallel reaction
708 monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol*
709 *Cell Proteomics* **11**, 1475-1488 (2012).
710 35. Julier, A., Radtke, V., Marx, A. & Scheffner, M. Generation and characterization of site-
711 specifically mono-ubiquitylated p53. *Chembiochem* **23**, e202100659 (2022).
712 36. Glozak, M.A., Sengupta, N., Zhang, X. & Seto, E. Acetylation and deacetylation of non-histone
713 proteins. *Gene* **363**, 15-23 (2005).
714 37. Choudhary, C., Weinert, B.T., Nishida, Y., Verdin, E. & Mann, M. The growing landscape of
715 lysine acetylation links metabolism and cell signalling. *Nat Rev Mol Cell Biol* **15**, 536-550 (2014).
716 38. Dancy, B.M. & Cole, P.A. Protein lysine acetylation by p300/cbp. *Chem Rev* **115**, 2419-2452
717 (2015).
718 39. Drazic, A., Myklebust, L.M., Ree, R. & Arnesen, T. The world of protein acetylation. *Biochim*
719 *Biophys Acta* **1864**, 1372-1401 (2016).
720 40. Seto, E. & Yoshida, M. Erasers of histone acetylation: The histone deacetylase enzymes. *Cold*
721 *Spring Harb Perspect Biol* **6**, a018713 (2014).
722 41. Smith, B.C. & Denu, J.M. Mechanism-based inhibition of sir2 deacetylases by thioacetyl-lysine
723 peptide. *Biochemistry* **46**, 14478-14486 (2007).

42. Farkas, M. et al. Distinct mechanisms control genome recognition by p53 at its target genes linked to different cell fates. *Nat Commun* **12**, 484 (2021).
43. Sykes, S.M. et al. Acetylation of the p53 DNA-binding domain regulates apoptosis induction. *Mol Cell* **24**, 841-851 (2006).
44. Vainer, R., Cohen, S., Shahar, A., Zarivach, R. & Arbely, E. Structural basis for p53 lys120-acetylation-dependent DNA-binding mode. *J Mol Biol* **428**, 3013-3025 (2016).
45. Tang, Y., Luo, J., Zhang, W. & Gu, W. Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. *Mol Cell* **24**, 827-839 (2006).
46. Mujtaba, S., Zeng, L. & Zhou, M.M. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **26**, 5521-5527 (2007).
47. Marchenko, N.D., Wolff, S., Erster, S., Becker, K. & Moll, U.M. Monoubiquitylation promotes mitochondrial p53 translocation. *EMBO J* **26**, 923-934 (2007).
48. Nie, L., Sasaki, M. & Maki, C.G. Regulation of p53 nuclear export through sequential changes in conformation and ubiquitination. *J Biol Chem* **282**, 14616-14625 (2007).
49. Zong, Z. et al. Alanyl-trna synthetase, aars1, is a lactate sensor and lactyltransferase that lactylates p53 and contributes to tumorigenesis. *Cell* **187**, 2375-2392 e2333 (2024).
50. Stuber, K. et al. Structural and functional consequences of nedd8 phosphorylation. *Nat Commun* **12**, 5939 (2021).
51. Hammler, D. et al. Fluorescently labelled atp analogues for direct monitoring of ubiquitin activation. *Chemistry* **26**, 6279-6284 (2020).
52. Götz, K.H. et al. Formation of the alarmones diadenosine triphosphate and tetraphosphate by ubiquitin- and ubiquitin-like-activating enzymes. *Cell Chem Biol* **26**, 1535-1543.e1535 (2019).
53. Delaglio, F. et al. Nmrpipe: A multidimensional spectral processing system based on unix pipes. *J Biomol NMR* **6**, 277-293 (1995).
54. Johnson, B.A. & Blevins, R.A. Nmr view: A computer program for the visualization and analysis of nmr data. *J Biomol NMR* **4**, 603-614 (1994).
55. Grzesiek, S., Stahl, S.J., Wingfield, P.T. & Bax, A. The cd4 determinant for downregulation by hiv-1 nef directly binds to nef. Mapping of the nef binding surface by nmr. *Biochemistry* **35**, 10256-10261 (1996).
56. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* **1**, 2856-2860 (2006).
57. Tyanova, S., Temu, T. & Cox, J. The maxquant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319 (2016).
58. Tyanova, S. et al. The perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740 (2016).
59. Krüger, L. et al. Chemical proteomic profiling reveals protein interactors of the alarmones diadenosine triphosphate and tetraphosphate. *Nat Commun* **12**, 5808 (2021).
60. Perez-Riverol, Y. et al. The pride database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* **50**, D543-D552 (2022).
61. Cornilescu, G., Marquardt, J.L., Ottiger, M. & Bax, A. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J Am Chem Soc* **120**, 6836-6837 (1998).

ACKNOWLEDGEMENTS

We thank Andreas Marquardt and Anna Sladewska-Marquardt of the Proteomics Center of the University of Konstanz for assistance with mass spectrometric experiments, and Silke Büstorf and Nicole Richter-Müller for technical support. We are grateful to Dr. Kathrin Lang

and Dr. Simon Elsässer for providing cDNAs and plasmids encoding the acetyllysyl-tRNA synthetase and the corresponding tRNA. This work was supported by the DFG (SFB969, Projects B3 and B9). S.M.K. thanks the doctoral fund of the University of Konstanz and the Konstanz Research School Chemical Biology for financial support. K.S. acknowledges the Zukunftskolleg of the University of Konstanz for a doctoral fellowship. F.S. is grateful for funding by the DFG (project grants 516836828 and 496470458).

AUTHOR CONTRIBUTIONS

S.M.K., M.S., M.Sch. conceived the study and experimental approach, S.M.K. generated all Ub variants, K.S. provided E1, E2, UCHL3, and established the AE-MS workflow for Ub variants, M.S. generated all p53 expression constructs, J.L. synthesized KeK. S.M.K. performed HDM2 autoubiquitination assays, AE-MS with the Ub variants, and interaction studies for AE-MS validations. T.S. conducted NMR spectroscopic analyses. M.S. performed all transient transfection experiments with p53. J.J. performed the PRM analysis. S.M.K., M.S., T.S., J.L., J.J., F.S., A.M., M.K. and M.Sch. analyzed the data. S.M.K., M.S. and M.Sch. wrote the paper. All authors provided critical feedback on the paper.

COMPETING INTERESTS

The authors declare no competing interests.

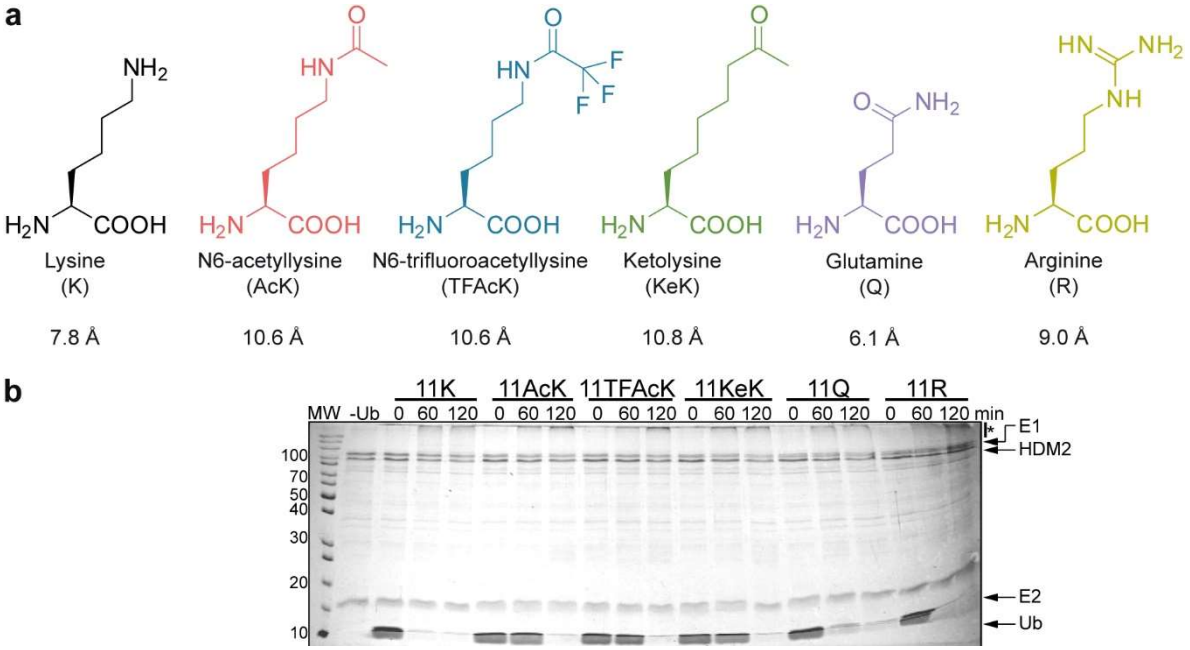
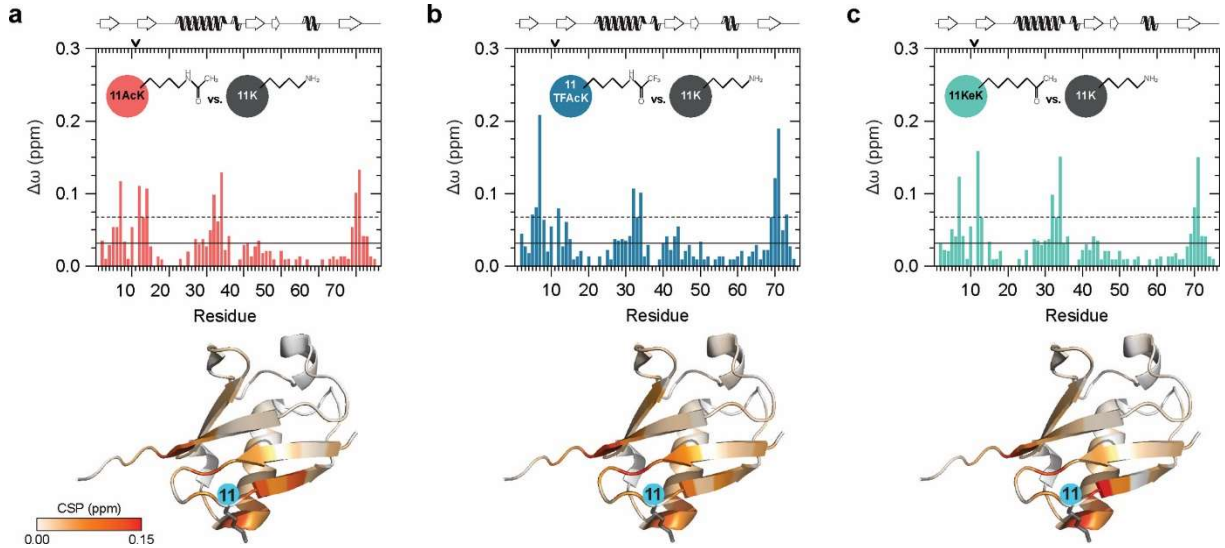


Fig. 1. The length of the side chain at position 11 of Ub affects HDM2 autou-biquitination. a, Comparison of the length (Å) of the side chain of amino acids incorporated at position 11 of Ub. **b,** HDM2 autoubiquitination assay with Ub variants, in which K11 was replaced by AcK, TFAcK, KeK, Q or R, indicate that the length of the side chain at position 11 is the main determinant for the performance of Ub variants in HDM2 autoubiquitination. Autoubiquitination reactions were started by addition of the respective Ub variant. Reactions were stopped at the times indicated. Reactions in the absence of Ub (-Ub) were stopped at 60 min. All reactions were analyzed by SDS-PAGE followed by Coomassie blue staining. Running positions of molecular mass markers (kDa), HDM2, UBA1 (E1), UbcH5b (E2), free Ub (Ub), and autoubiquitinated forms of HDM2 (*) are indicated. Quantitative incorporation of the respective amino acids and the homogeneity of the purified Ub variants were confirmed by ESI-MS analysis (Supplementary Fig. 1).



803

804 **Fig. 2. Incorporation of AcK, TFAcK or KeK at position 11 of Ub results in similar structural**
805 **changes. a-c,** Weighted chemical shift perturbations (CSPs, Δω) of backbone amide resonances
806 were calculated comparing the differently modified Ub variants (11AcK, 11TFAcK, 11KeK) with
807 unmodified Ub (11K). The corresponding two-dimensional heteronuclear ^{1}H-^{15}N HSQC NMR spectra
808 are shown in Supplementary Fig. 2. In the upper panel, the corresponding CSP plots are presented for
809 Ub 11AcK (a), Ub 11TFAcK (b), Ub 11KeK (c), while in the lower panel the respective CSP values are
810 highlighted on the NMR solution structure of unmodified Ub (PDB ID 1D3Z)[61]. The Ub structure is
811 illustrated in cartoon mode and the respective CSP values are indicated by using a continuous color
812 palette ranging from white (Δω = 0.00 ppm) to red (Δω ≥ 0.15 ppm). Additionally, the position of K11
813 and its side chain are depicted by a blue circle and by anthracite sticks, respectively. In the CSP plots,
814 the horizontal solid and dashed lines are indicating the mean and the mean plus one standard
815 deviation, respectively, calculated over all CSP values from (a) to (c) excluding values related to
816 position 11. On top of the plots, position 11 is indicated by an arrow and the secondary structural
817 elements are schematically depicted according to the NMR solution structure of unmodified Ub.
818

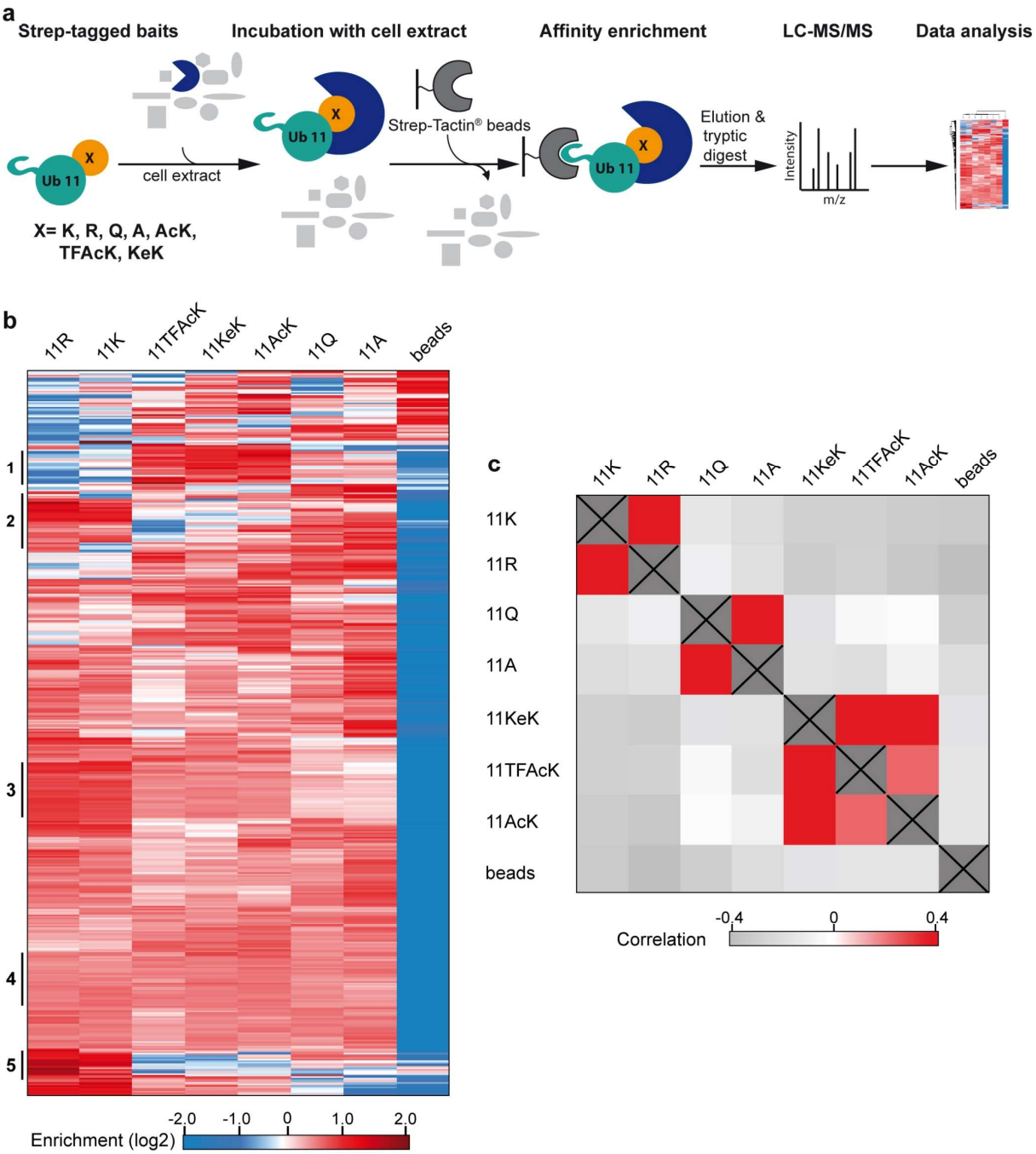


Fig. 3. Ub variants containing AcK or its analogs TFAcK and KeK resemble one another with respect to their protein interaction properties. a, Schematic overview of the AE-MS workflow for the identification of interactors of the Ub variants indicated. Unmodified Ub (11K) and each of the six Ub variants were used as bait molecules for affinity enrichment and incubated with HEK293T cell extract. Empty beads (beads) were used as control. **b,** Hierarchical clustering of statistically significant interactors (rows) of the different bait molecules (columns). Red indicates enrichment, whereas blue indicates lack of enrichment. Thresholds for ANOVA statistics were set to S0=2 and FDR=0.005. Cluster numbers, indicating binding to one or multiple Ub variants, are indicated on the left. **c,** Correlation heatmap of significantly enriched interactors of the differently modified Ub variants. Strong correlation is indicated in red, no correlation in white, and inverse correlation in gray.

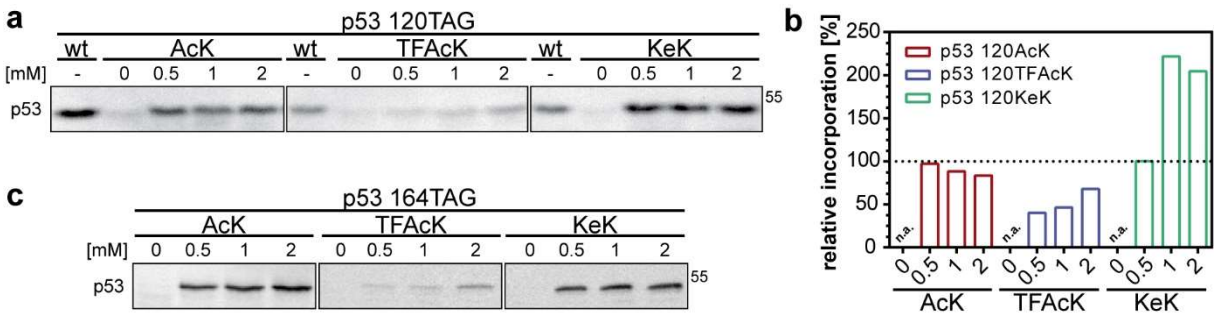


Fig. 4. Incorporation of AcK and its analogs TFAcK and KeK into p53 by genetic code expansion in H1299 cells. a, Incorporation of AcK, TFAcK, and KeK at position 120 of p53. **b,** Quantification of relative p53 levels in the presence of AcK, TFAcK, and KeK shown in **a**. Levels of wild-type p53 (wt) were set to 100 percent. **c,** Incorporation of AcK, TFAcK, and KeK into position 164 of p53. H1299 cells were transfected with expression constructs encoding wild-type p53 (wt) or p53 with a TAG stop codon at position 120 and 164, respectively, and the AcK-RS/tRNA pair as described in Methods. Growth medium was supplemented with increasing concentrations of AcK, TFAcK, or KeK as indicated. 24 h upon transfection**,** p53 levels were determined by western blot analysis with the p53-specific DO-1 antibody (**a, c**).

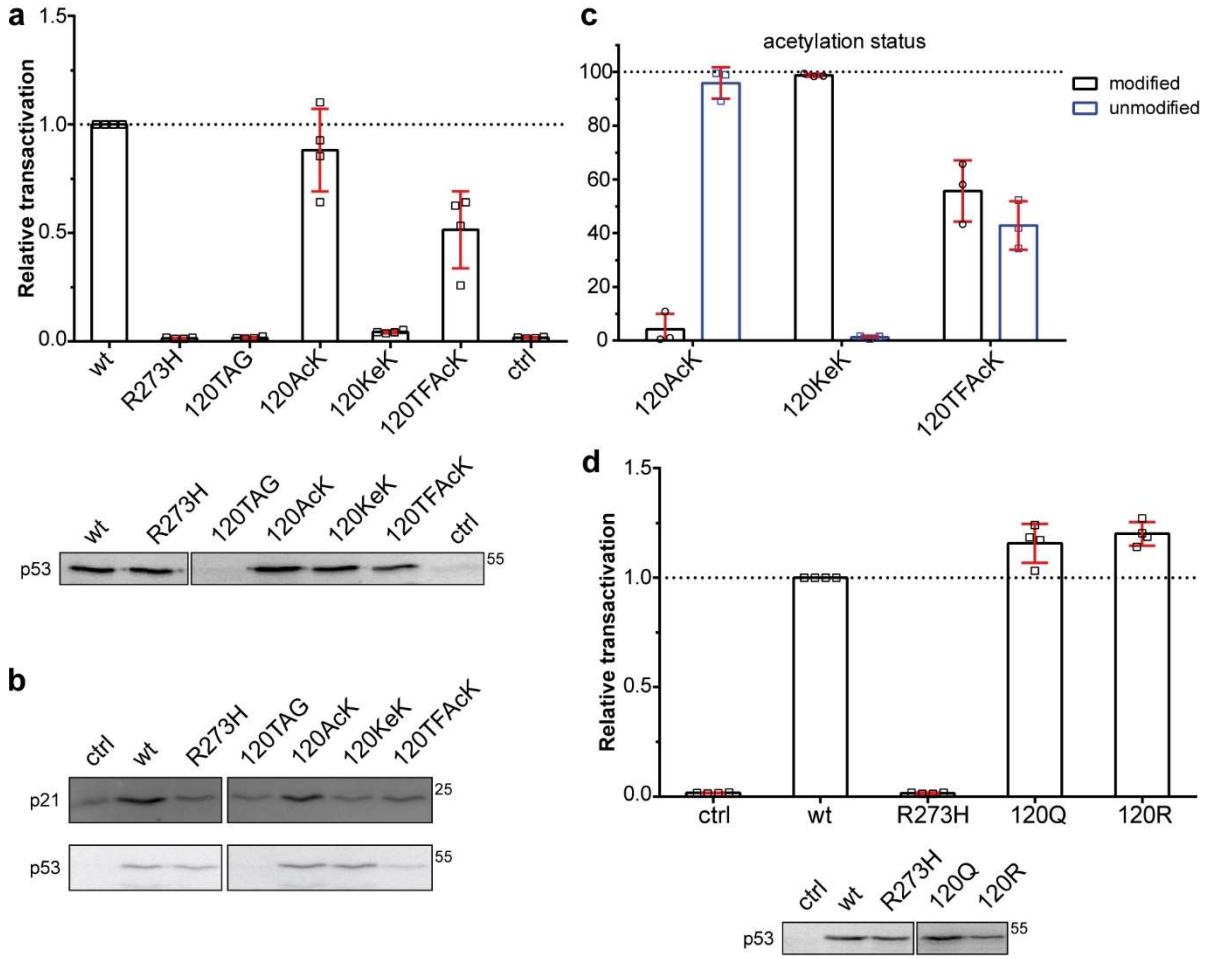


Fig. 5. Transactivation activity of p53 variants. a, H1299 cells were transfected with a p53-responsive expression construct encoding luciferase in the absence (ctrl) or presence of constructs encoding wild-type p53 (wt), the tumor-derived p53 mutant R273H, p53 with a TAG stop codon at position 120, and the AcK-RS/tRNA pair. In addition, cells transfected with the p53-120TAG construct were grown in the absence (120TAG) or presence of AcK (120AcK), TFAcK (120TFAcK) or KeK (120KeK). 24 h upon transfection, luciferase activity was determined as a measure for p53 activity and the relative values obtained were adjusted for transfection efficiency (n = 4 ± SD independent experiments). 20 percent of the extracts were subjected to western blot analysis using the p53-specific DO-1 antibody. **b,** H1299 cells were transfected as in **a** in the absence of the p53-responsive reporter construct. 24 h upon transfection, levels of endogenous p21 protein and of ectopically expressed p53 were determined by western blot analysis using a p21-specific antibody and the p53-specific DO-1 antibody, respectively. **c,** H1299 cells were transfected with expression constructs encoding p53-120TAG and the AcK-RS/tRNA pair. Cells were grown in the presence of AcK, KeK, and TFAcK as indicated and the presence of AcK, KeK, TFAcK or lysine (unmodified) at position 120 of p53 was determined by PRM analysis (n = 3 ± SD independent experiments). **d,** Transactivation activity of the p53 variants 120Q and 120R was determined as described in **a** (n = 4 ± SD independent experiments).