ARTICLE

https://doi.org/10.1038/s41586-019-1188-1

Time-resolved protein activation by proximal decaging in living systems

Jie Wang[1,2,5], Yuan Liu[1,5], Yanjun Liu[1,5], Siqi Zheng[1], Xin Wang[1,2], Jingyi Zhao[1,2], Fan Yang[1], Gong Zhang[1], Chu Wang[1,3,4]* & Peng R. Chen[1,2,3,4]*

A universal gain-of-function approach for selective and temporal control of protein activity in living systems is crucial to understanding dynamic cellular processes. Here we report development of a computationally aided and genetically encoded proximal decaging (hereafter, CAGE-prox) strategy that enables time-resolved activation of a broad range of proteins in living cells and mice. Temporal blockage of protein activity was computationally designed and realized by genetic incorporation of a photo-caged amino acid in proximity to the functional site of the protein, which can be rapidly removed upon decaging, resulting in protein re-activation. We demonstrate the wide applicability of our method on diverse protein families, which enabled orthogonal tuning of cell signalling and immune responses, temporal profiling of proteolytic substrates upon caspase activation as well as the development of protein-based pro-drug therapy. We envision that CAGE-prox will open opportunities for the gain-of-function study of proteins and dynamic biological processes with high precision and temporal resolution.

Precise perturbation of protein activity with high selectivity and temporal resolution is crucial to dissect various dynamic biological processes ranging from signal transduction to programmed cell death. Using traditional genetic or pharmacological approaches, it is often challenging to fulfil the requirement of high selectivity or temporal resolution. An array of pharmacological and protein engineering-enabled approaches, including the 'bump-and-hole'[1,2], PROTAC[3], LARIAT[4], trim-away strategies[5] and engineering extrinsic disorder[6] have been successfully implemented into various proteins of interest (POIs) for studying dynamic cellular events, but they are typically conducted in a loss-of-function manner. Gain-of-function tools offer a particularly attractive 'positive' approach to probe and control dynamic protein activity. Extensive protein engineering efforts have been made to directly manipulate a POI, such as genetic fusion of switchable proteins[7–9], photo-controllable peptides[10,11], split proteins[12], light-controlled assembling domains[13–15] or homodimeric/heterodimeric proteins[16–18]. These optogenetic and chemogenetic methods allow selective investigation of dynamic biological processes and avoid potential secondary effects and phenotype compensations. Nevertheless, it is highly desirable to develop a unified and generalizable method for selective and temporal protein activation under living conditions, without genetic fusion.

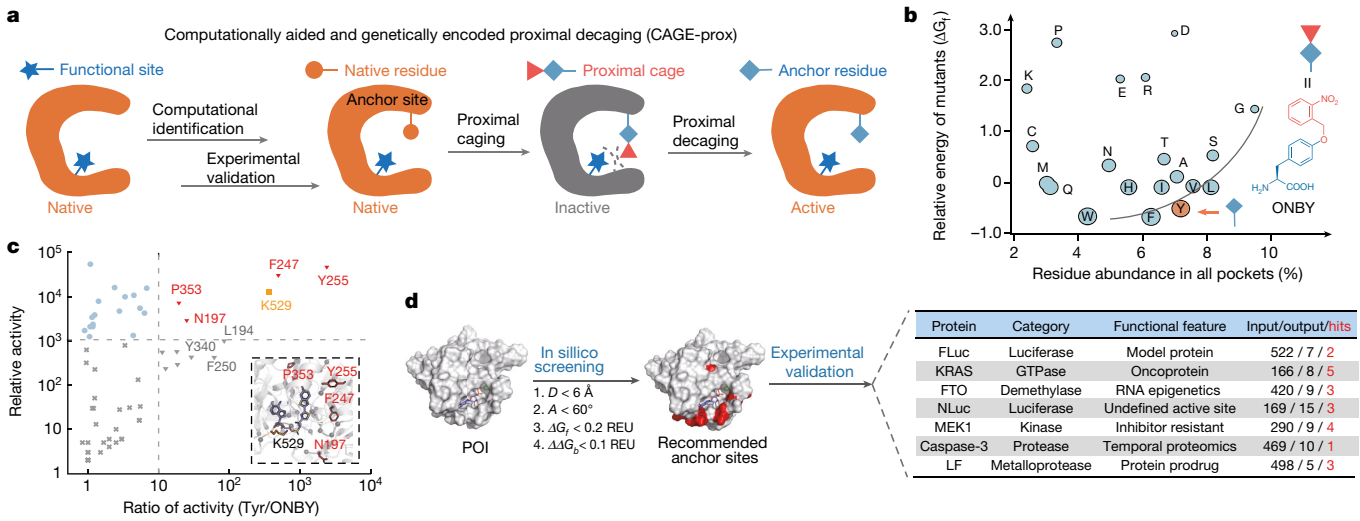
Site-specific blockage of catalytic residues through a genetically incorporated 'caged' amino acid that can be precisely decaged by chemical or optical means represents an alternative for selective and rapid protein activation[19–21]. Currently, only a limited number of caged amino acids can be genetically encoded, and different analogues need to be incorporated in a case-by-case manner depending on the chemical identity of the catalytic residues on a POI. Moreover, many enzymes do not have well-defined catalytic sites, which renders them incompatible with direct decaging. A universal decaging-based protein activation strategy in living systems that does not directly rely on catalytic residues is lacking.

Herein we address these challenges with a general 'proximal decaging' strategy, in which the activity of a protein is temporally blocked by a universal caged amino acid introduced in close proximity to its functional site and is then re-activated promptly upon photo-decaging (Fig. 1a). We name this strategy 'computationally aided and genetically encoded proximal decaging' (CAGE-prox). In contrast to direct decaging that focuses on a specific catalytic residue, CAGE-prox relies on a unified caged amino acid in vicinity of the functional site as a temporal 'bump', and therefore can be applied to block the activity of a wide range of proteins regardless of the existence and/or chemical identity of the catalytic residues. We first established a photo-caged tyrosine, ONBY[21], as the universal proximal cage and developed a computational algorithm to identify the optimal anchor site for turning off and on protein activity through caging and decaging operations. We next validated the generality of CAGE-prox in the functional manipulation of diverse POIs including GTPases, kinases, RNA demethylases, caspases and bacterial effector proteins. We further demonstrated the capability of CAGE-prox for selective, temporal probing and/or perturbing of diverse biological processes by constructing orthogonal kinase signalling cascades, temporal profiling of proteolytic substrates and the on-demand liberation of bacterial effectors as protein-based pro-drug therapy.

Establishing and optimizing the operation of CAGE-prox

According to the design of CAGE-prox, a proximal cage will be first incorporated at an anchor site on a POI to block its activity, which can later be re-activated through photo-decaging (Fig. 1a). Although blocking the protein activity with a unified proximal cage is conceptually straightforward, the following re-activation step requires that the decaged residue at the anchor site (termed anchor residue) should be either the same as the original residue or a mutation with negligible perturbation to protein activity. To survey which of the 20 natural amino acids would serve as the potential anchor residue, we analysed 3,500 residue sites within the ligand-binding pockets of 127 representative protein–ligand complexes from the Astex dataset[22]. Through systematic study, we determined that Tyr would be an ideal candidate for the

[1]Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing, China. [2]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China. [3]Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, Peking University, Beijing, China. [4]Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China. [5]These authors contributed equally: Jie Wang, Yuan Liu, Yanjun Liu. *e-mail: chuwang@pku.edu.cn; pengchen@pku.edu.cn



Protein	Category	Functional feature	Input/output/hits
FLuc	Luciferase	Model protein	522 / 7 / 2
KRAS	GTPase	Oncoprotein	166 / 8 / 5
FTO	Demethylase	RNA epigenetics	420 / 9 / 3
NLuc	Luciferase	Undefined active site	169 / 15 / 3
MEK1	Kinase	Inhibitor resistant	290 / 9 / 4
Caspase-3	Protease	Temporal proteomics	469 / 10 / 1
LF	Metalloprotease	Protein prodrug	498 / 5 / 3

Fig. 1 | Design and development of the CAGE-prox strategy. **a**, Schematic of CAGE-prox: a proximal cage can be genetically incorporated at an anchor site in close proximity to the protein functional site for temporal blockage of its activity. Proximal decaging can lead to rapid rescue of protein functions, as long as the decaged anchor residue has negligible influence on protein activity. **b**, Establishing Tyr as a universal anchor residue and photo-caged Tyr (ONBY) as the proximal cage. All the 20 natural amino acids are evaluated in terms of natural abundance and mutational stability (ΔG_f) (see details in Methods). **c**, A total of 60 sites on FLuc was randomly selected for experimental evaluation; *x* axis represents the blockage efficiency by measuring the ratio of FLuc activity before and after ONBY decaging; *y* axis represents the rescue efficiency by measuring the relative activity for each activated FLuc variant. Four positive anchor sites (see the detail criteria in Methods) are labelled in red and are also shown as stick (insert). Protein Data Bank (PDB) code 4D36. The K529-ONBK variant with direct caging of the FLuc catalytic residue was included as a control. Representative data from two biological replicates. **d**, The overall operation procedure of CAGE-prox. All the residues (sites) on a POI were computationally screened according to (i) geometric parameters (distance and direction to the ligand) and (ii) energetic parameters (folding stability and substrate binding affinity), and the recommended anchor sites were then subjected to experimental validation. CAGE-prox was applied to various enzyme families to demonstrate its general applicability. The number of anchor sites that are computationally screened ('input'), recommended after calculation ('output') and experimentally validated as positive 'hits' are listed.

anchor residue owing to its higher natural abundance in the pocket, better mutation stability and readiness for photo-decaging among the 20 natural amino acids (Fig. 1b, Extended Data Fig. 1a). This was also experimentally validated by comparing Tyr with Lys, which is another commonly used amino acid for decaging but is poorly ranked according to our in silico investigation (Extended Data Fig. 1b–d). We therefore decided to use a photo-caged Tyr, ONBY, as the unified proximal cage for CAGE-prox (Fig. 1b).

To establish a computational pipeline in CAGE-prox to rationally identify potential anchor sites for installing the proximal cage, we used firefly luciferase (FLuc) as a model to collect experimental data for training. A total of 60 sites on FLuc was randomly selected for ONBY insertion and photo-decaging with optimized conditions (Fig. 1c, Extended Data Fig. 2a–e). Out of the 60 mutants, 4 yielded desired results: the enzyme activity was efficiently blocked by the caged ONBY and promptly rescued upon photo-decaging, resulting in greater than tenfold activation (Extended Data Fig. 2h, Supplementary Data 1). We next calculated geometry and energy-based parameters using Rosetta[23] to computationally encode general properties for selecting optimal anchor sites (Supplementary Table 1). A combination of cut-off values of distance $D < 6$ Å and angle $A < 60°$ was set for the geometry parameters, and the mutation stability $\Delta G_f < 0.2$ REU (Rosetta energy unit) and substrate binding energy $\Delta\Delta G_b < 0.1$ REU were set for the energy-based parameters (Extended Data Fig. 2f, g, i, j).

Collectively, in the final streamlined operation procedure of CAGE-prox (Fig. 1d, Extended Data Fig. 3), a Rosetta-based in silico structural calculation[23] is first performed on a POI to screen and identify optimal sites for proximal caging and decaging according to the aforementioned geometry and energy-based parameters. Next, the list of recommended anchor sites is subject to experimental validation by site-specific incorporation of ONBY followed by photo-decaging. Variants that exhibit minimal blockage and maximal rescue activities are chosen for subsequent gain-of-function studies and applications.

Applying CAGE–prox for general protein activation

After optimizing the operation procedure by FLuc, we demonstrated the general applicability of CAGE-prox to control activation on a diverse set of enzymes, including those with catalytic sites that are complicated or undefined (Fig. 1d). The small GTPase KRAS is a master regulator of many essential signal transduction pathways, and its dysregulation and/or oncogenic mutations is a key driving factor for cancer[24]. Precise and temporal perturbation of KRAS is highly desired to facilitate the mechanistic investigation of KRAS-driven cancers[25]. Application of CAGE-prox on KRAS yielded eight computationally recommended variants (Supplementary Table 2), five of which were tested as effective in reactivation upon decaging (Extended Data Fig. 4a–c). We then showed that photo-activation of the KRAS-Y32-ONBY variant triggered sequential phosphorylation of the downstream kinases such as MEK1/2 and ERK in a time-dependent manner but caused no perturbation to the parallel and closely related JNK pathway (Extended Data Fig. 4d–f).

We next expanded the CAGE-prox strategy for activation of FTO, an α-ketoglutarate-dependent RNA demethylase that catalyses m^6A demethylation[26] with an active-site transition-metal ion, Fe (II). Our results showed that four out of the nine computationally identified ONBY mutants (Supplementary Table 3) exhibited desired activity as demonstrated by monitoring demethylation of an RNA substrate (Extended Data Fig. 4g–j). Furthermore, we applied CAGE-prox to nano-luciferase (NLuc), a recently developed luciferase with enhanced activity but undefined active sites[27]. Molecular docking was conducted to create initial complex structures, and efficient activation through CAGE-prox was observed on 3 out of the 15 recommended variants (Extended Data Fig. 5, Supplementary Table 4). We demonstrated that CAGE-prox is widely applicable to diverse protein families, including those with complicated or undefined functional sites.

Orthogonal kinase activation by CAGE–prox

Temporal and orthogonal control of kinase activation may allow better dissection of the dynamic features that underlie cell signal

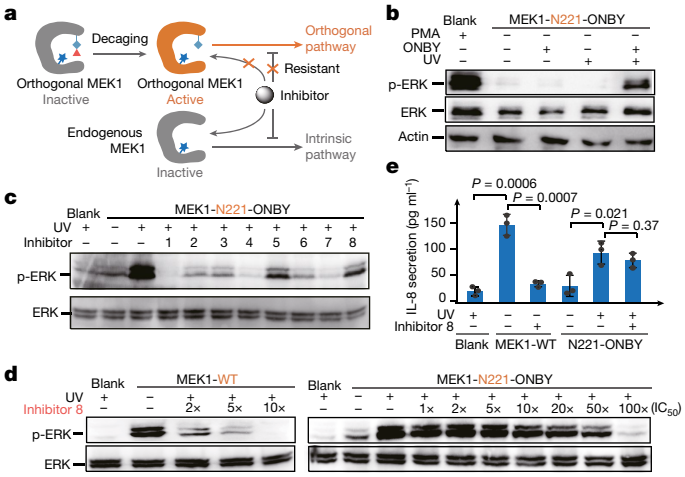


Fig. 2 | Orthogonal and mutually exclusive kinase activation by CAGE-prox. a, Schematic of CAGE-prox-enabled orthogonal activation of MEK1 with the endogenous MEK1 activity inhibited simultaneously. **b**, Photo-activation of MEK1 by CAGE-prox led to ERK phosphorylation in Jurkat cells. Electroporation was used for plasmid transfection. PMA, phorbol-12-myristate-13-acetate. $n = 2$. **c**, Identification of MEK1(N221)-ONBY as one of the photo-activatable and inhibitor-resistant MEK1 variants by CAGE-prox. See Extended Data Fig. 6 for the inhibitor-resistance profile for the other two MEK1 variants (I216 and D217) as well as the structures of inhibitors 1–8. $n = 2$. **d**, Dose-dependent treatment by inhibitor 8 in live HEK293T cells after activation of MEK1(N221Y) (right). Wild-type MEK (MEK-WT; left) was used as a control. $n = 3$. **e**, Orthogonal activation of MEK1(N221)-ONBY in HEK293T cells resulted in enhanced expression and secretion of IL-8. Mean ± s.d.; $n = 3$; two-tailed *t*-test. All above-mentioned samples are biological replicates. *P* values are shown in the figure.

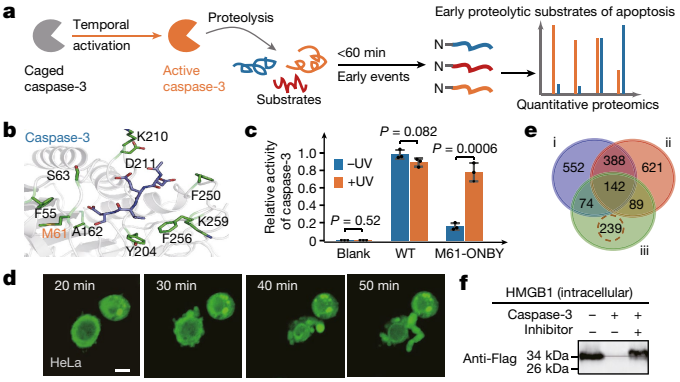


Fig. 3 | CAGE-prox-enabled temporal profiling of the proteolytic substrates upon caspase-3 activation. a, Schematic of CAGE-prox-enabled temporal profiling of the proteolytic substrates immediately after caspase-3 activation. **b**, Structural view of the relative positions of the ten computationally identified anchor sites (green sticks) on caspase-3 (grey cartoon, PDB code 4DCO) towards the substrate peptide (purple sticks). **c**, Incorporation of ONBY at M61 blocked caspase-3 activity, which can be efficiently rescued after photo-decaging. Mean ± s.d.; $n = 3$; two-tailed *t*-test. **d**, Representative images showing that HeLa cells underwent apoptosis-like morphological changes within 30 min of CAGE-prox-triggered caspase-3 activation. $n = 2$. Scale bars, 5 μm. **e**, Comparison of the identified proteolytic substrates upon caspase-3 activation with the previously reported caspase-3 substrates. i and ii denote reported substrates from chemical-induced apoptosis and recombinant caspase-3, respectively; iii denotes the proteolytic substrates after CAGE-prox-triggered caspase-3 activation. **f**, Verification of HMGB1 as a substrate of caspase-3. HMGB1–Flag and wild-type caspase-3 were co-transfected into HEK293T cells, and the cell pallets were collected for immunoblotting analysis. $n = 2$. All above-mentioned samples are biological replicates. *P* values are shown in the figure.

transductions. Although previous efforts have enabled temporal control of kinase activation through engineered constructs or direct decaging[9,17,28], the undealt activity from endogenous kinases controlling the intrinsic pathways prevents the construction of truly orthogonal signalling cascades. Creating an activatable and 'inhibitor-resistant' kinase variant could provide a solution to this problem. Inspired by the drug-resistant kinase mutations of cancer cells[29] as well as the orthogonal 'gatekeeper' mutants[30], we envisioned that the proximal Tyr mutation would gain resistance to inhibitors developed for wild-type kinases, which could be leveraged to build tailor-made signalling cascades that are orthogonal and mutually exclusive to the intrinsic pathways (Fig. 2a). Notably, because this orthogonal activation strategy relies on an anchor site instead of a gatekeeper mutation, in principle it can be extended for gain-of-function study of enzymes beyond kinases.

We used CAGE-prox to engineer an orthogonal MAPK pathway, which is a central kinase signalling cascade in regulating diverse cellular responses. We specifically aimed at a key MAPK, MEK1, which has previously been subject to controlled activation by several approaches, but without desired mutual orthogonality to the endogenous counterparts[31]. The constitutively active form of MEK1 with S218E and S222D mutations (MEK1-ED), was used to avoid the complicated regulation by its upstream enzymes. A total of nine CAGE-prox variants was first recommended by the computation (Supplementary Table 5) and three of them (I216, D217 and N221) were validated with the expected photo-activation property in HEK293T cells (Extended Data Fig. 6a, b). The N221-ONBY variant was further shown to activate the MAPK signal pathway in Jurkat cells, a commonly used model for studying immune-cell signalling (Fig. 2b).

We then surveyed the susceptibility of our CAGE-prox activatable MEK1 variants against eight commercial MEK inhibitors (Supplementary Table 8). Four resistant enzyme–inhibitor pairs were obtained and verified by their significantly shifted half-maximal inhibitory concentration (IC_{50}) values (Fig. 2c, d, Extended Data Fig. 6c–j); part of these could also be retrospectively predicted by computation (Extended Data Fig. 6k, Supplementary Results). Given the central role of MEK–ERK signalling in fine-tuning T cell responses[32], we further demonstrated that our orthogonally activatable MEK1 construct can ultimately enhance the expression and secretion of inflammatory chemokines such as interleukin 8 (Fig. 2e). CAGE-prox allowed the photo-controlled, orthogonal activation of MEK1-mediated signalling cascades with the endogenous pathways simultaneously inhibited.

CAGE–prox–enabled temporal proteomics

We next coupled CAGE-prox with quantitative proteomics to dissect temporal proteome organization within cells. Proteolysis is a fundamental life process regulated by over 600 proteases with high spatial–temporal specificity. Among them, executioner caspases[33] —such as caspase-3, -6 and -7—are the core enzymes during apoptosis and are known to form a dynamic and cooperative network, and each individual enzyme also has distinct roles[34]. Assigning proteolytic substrates for these proteases is therefore highly warranted for comprehensive characterization of this highly dynamic cellular process[35]. Great efforts have been made to enrich the proteolytic substrates of caspases for proteomic profiling by mixing individual proteases with cell extracts[36] or triggering apoptosis with chemical inducers[35]. As a complementary approach, CAGE-prox enabled substrate profiling with high temporal resolution and selectivity inside living cells. Furthermore, considering that many proteolytic events are intermingled during apoptosis and certain caspases are also auto-activatable, we reasoned that substrate profiling immediately after the temporal activation of a specific caspase may help to reduce noise from secondary processing or parallel cascades. To prove this concept, we used CAGE-prox for temporal activation and substrates profiling of caspase-3 within a native cellular context (Fig. 3a).

Computational screening and subsequent experimental validation identified M61 as the anchor site for controlling caspase-3 activation (Fig. 3b, c, Extended Data Fig. 7c, Supplementary Table 6). Unlike

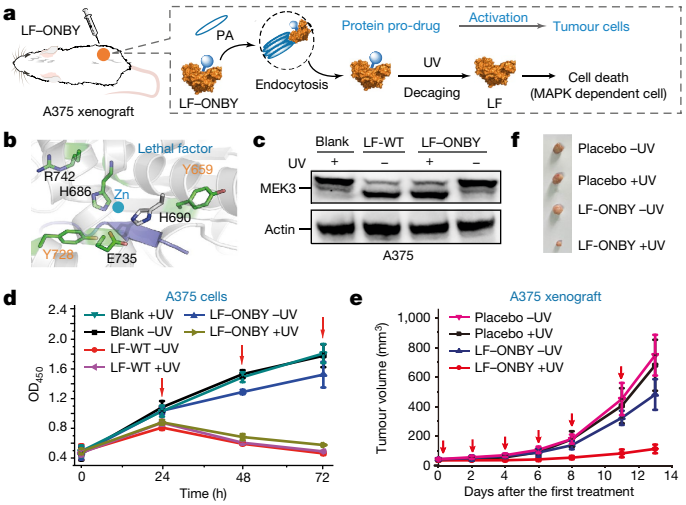


Fig. 4 | CAGE-prox-enabled protein pro-drug activation in mice. **a**, Schematic of the CAGE-prox-activated lethal factor (LF) as a protein-based pro-drug therapy. Activated LF will cause the death of MAPK-dependent tumour cells by rapid cleavage of MEK kinases. **b**, Structural view of the five recommended anchor sites (green sticks) on LF (grey cartoon) by computational screening. The catalytic Zn (II) ion and the substrate peptide were shown in the structure. PDB code 1PWV. **c**, Delivery and photo-activation of LF(Y659)-ONBY in A375 cells results in cleavage of MEK3. Purified LF(Y659)-ONBY was delivered into A375 cells by PA. $n = 2$. **d**, Photo-activated LF(Y659)-ONBY significantly inhibited the proliferation of MEK-dependent A375 cells but not MEK-independent HEK293T or HeLa cells (Extended Data Fig. 9f). Red arrows indicate the time points when the medium exchange and LF–PA treatment were performed. Mean ± s.d.; $n = 3$. **e**, Photo-activated LF(Y659)-ONBY significantly inhibited the tumour growth in mice. The red arrows indicate the time points at which LF–PA injection and ultraviolet light irradiation were performed. Each group was conducted on eight mice in parallel. Mean ± s.d.; $n = 8$. **f**, Representative images of the tumours showing that photo-activation of the LF–PA pro-drug significantly inhibited tumour growth in mice. $n = 8$. All above-mentioned samples are biological replicates.

the small-molecule compounds that are traditionally used to trigger apoptosis in hours[37], photo-activation on the M61–ONBY variant in HEK293T or HeLa cells induced the desired apoptotic phenotype within only 30 min (Fig. 3d, Extended Data Fig. 7a, b, Supplementary Videos 1, 2), which indicates that direct activation of caspase-3 could rapidly trigger apoptosis.

Using this highly specific, photo-activatable caspase-3 variant, we then profiled the proteolytic substrates after caspase-3 activation (Fig. 3a). A mass-spectrometry-based quantitative proteomics strategy was used to identify peptides that either have significant degradation rates or bear newly generated N termini (Extended Data Fig. 8a, Methods, Supplementary Results). We observed that most other caspases remained inactive for 60 min after photo-activation of caspase-3, and therefore collected cell samples within this early time window to minimize further cleavage or modifications[35] (Extended Data Fig. 7d–f). A total of 544 protein substrates was identified commonly from 3 independent experiments (Extended Data Fig. 8b–f, Supplementary Data 2), including 239 new proteolytic substrates as compared to previous profiling studies (Fig. 3e). We selected 12 proteins—including HMGB1, MSH6 and ATP6V1A—for validation as caspase-3 substrates by immunoblotting analysis (Fig. 3f, Extended Data Fig. 8g–w, Supplementary Results).

CAGE-prox-enabled protein pro-drug therapy

Finally, we explored the therapeutic potential of CAGE-prox for controlled activation of cytotoxic proteins as pro-drugs in living systems. On-demand release of small-molecule pro-drugs can effectively improve potency by increasing concentrations at the lesion site (such as solid tumours) to levels that would otherwise be too toxic to patients. As bacteria have evolved many effector proteins that possess a higher specificity towards their host targets than small molecule reagents, we envisioned that controlled activation of certain effectors by CAGE-prox may offer an in situ protein-based pro-drug therapy[38] (Fig. 4a, Extended Data Fig. 9a, b). To prove this concept, we applied CAGE-prox for targeted activation of the anthrax toxin component lethal factor (LF) in cancer cells. Upon entering host cells via the protein-protective antigen (PA) component, LF can initiate rapid cleavage of MAPK kinases and becomes highly toxic to MAPK-dependent cancer cells. It is therefore possible to combine the spatial specificity of light-controlled activation and the target specificity of LF to develop an efficient therapeutic strategy for certain cancers.

Five anchor sites were identified computationally (Fig. 4b, Supplementary Table 7) and Y659 was experimentally validated with desired efficiency for controlled activation using MEK3 as the model substrate in HEK293T, A375 and HeLa cells (Fig. 4c, Extended Data Fig. 9c–e). As expected, controlled activation of LF by CAGE-prox inhibited only the proliferation of MAPK-dependent A375 cells and not of HEK293T or HeLa cells (Fig. 4d, Extended Data Fig. 9f). We next optimized the protocols (for example, UV irradiation conditions and dosing frequency) for applying CAGE-prox to control protein activation in living mice (Extended Data Fig. 9g, k–p, Supplementary Results).

With these optimized protocols, we conducted a xenograft model of A375 cells to demonstrate the efficacy of the LF-based pro-drug therapy (Extended Data Fig. 9b). The tolerance and safety of LF–PA was first verified on healthy mice (Extended Data Fig. 9h); we then evaluated the therapeutic effects of photo-activated LF every 48 h for two weeks. As expected, in situ photo-activation of LF strongly inhibited the tumour growth as measured by tumour size and volume (Fig. 4e, f, Extended Data Fig. 9j). Of note, body weights of the mice after the LF-ONBY treatment and ultraviolet light irradiation remained unaffected in comparison with control mice (Extended Data Fig. 9i). Therefore, our controlled local ultraviolet light activation significantly expanded the dosage window of LF–PA without causing systemic damage.

Discussion

Gain-of-function tools are uniquely positioned for mechanistic study of enzymes as they offer a positive approach to investigating the activity of a given enzyme in triggering or regulating dynamic cellular processes. With high selectivity, temporal resolution and minimal changes introduced to the native enzyme, CAGE-prox offers a complementary and—in some aspects—unique approach to the aforementioned optogenetic and chemogenetic methods[6–17]. For example, CAGE-prox-enabled proteolytic substrate profilings of caspase-3 during the early stage of apoptosis were in general agreement with previous reports using the subtiligase-based N terminus enrichment[35,39] (Extended Data Fig. 8c–f, Supplementary Data 2). Therefore, coupling these two approaches may enable time-resolved and in-depth proteomics study during various proteolytic processes. Moreover, considering that new substrates and biological roles of caspases have continuously been uncovered (for example, gasdermin-E in pyroptosis)[40], it may also be possible to apply CAGE-prox in the future to reveal the cell-type-specific substrates of activated caspases.

Computational modelling has had a key role in guiding the selection of anchor residues for CAGE-prox. By systematic analysis of protein stability and substrate binding, Rosetta-based calculation significantly improved the success rate for identifying desired anchor sites than simple visual inspection (Extended Data Fig. 10a). As protein activity can also be regulated through allosteric binding, post-translational modifications and protein–protein interactions, CAGE-prox—with a tailored computational screening strategy—can in principle be applied to target these alternative regulation mechanisms. Indeed, our preliminary results showed that CAGE-prox was able to control the auto-phosphorylation of Src kinase and regulate its downstream signalling (Extended Data Fig. 10b).

Owing to the high engineering capacity of the LF–PA delivery system, fusing the N-terminal domain of LF to different bacterial toxins may permit controlled activation of diverse therapeutic proteins in 'hard-to-transfect' cells (Extended Data Fig. 10c). Furthermore, by tethering PA with various targeting elements (for example, EGF and VEGF), the CAGE-prox-activatable therapeutic proteins can be programmed to selectively recognize and kill cancer cells under clinically relevant settings (Extended Data Fig. 10c). Finally, the current CAGE-prox strategy with photo-caged unnatural amino acids can in principle be extended to use two-photon-caged or chemical-caged unnatural amino acids[19], which would allow protein activations that are more compatible with in vivo settings and/or clinical applications.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1188-1.

Received: 16 September 2018; Accepted: 4 April 2019;
Published online 8 May 2019.

1. Bishop, A. C. et al. A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* **407**, 395–401 (2000).
2. Baud, M. G. J. et al. A bump-and-hole approach to engineer controlled selectivity of BET bromodomain chemical probes. *Science* **346**, 638–641 (2014).
3. Winter, G. E. et al. Drug development. Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **348**, 1376–1381 (2015).
4. Ulman, V. et al. An objective comparison of cell-tracking algorithms. *Nature Methods* **14**, 1141–1152 (2017).
5. Clift, D. et al. A method for the acute and rapid degradation of endogenous proteins. *Cell* **171**, 1692–1706.e18 (2017).
6. Dagliyan, O. et al. Engineering extrinsic disorder to control protein activity in living cells. *Science* **354**, 1441–1444 (2016).
7. Davis, K. M., Pattanayak, V., Thompson, D. B., Zuris, J. A. & Liu, D. R. Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nature Chem. Biol*. **11**, 316–318 (2015).
8. Wu, Y. I. et al. A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* **461**, 104–108 (2009).
9. Karginov, A. V., Ding, F., Kota, P., Dokholyan, N. V. & Hahn, K. M. Engineered allosteric activation of kinases in living cells. *Nature Biotechnol*. **28**, 743–747 (2010).
10. Lungu, O. I. et al. Designing photoswitchable peptides using the AsLOV2 domain. *Chem. Biol*. **19**, 507–517 (2012).
11. Niopek, D., Wehler, P., Roensch, J., Eils, R. & Di Ventura, B. Optogenetic control of nuclear protein export. *Nature Commun*. **7**, 10624 (2016).
12. Pratt, M. R., Schwartz, E. C. & Muir, T. W. Small-molecule-mediated rescue of protein function by an inducible proteolytic shunt. *Proc. Natl Acad. Sci. USA* **104**, 11209–11214 (2007).
13. Taslimi, A. et al. An optimized optogenetic clustering tool for probing protein interaction and function. *Nature Commun*. **5**, 4925 (2014).
14. Wang, H. et al. LOVTRAP: an optogenetic system for photoinduced protein dissociation. *Nature Methods* **13**, 755–758 (2016).
15. Lee, S. et al. Reversible protein inactivation by optogenetic trapping in cells. *Nature Methods* **11**, 633–636 (2014).
16. Zhou, X. X., Fan, L. Z., Li, P., Shen, K. & Lin, M. Z. Optical control of cell signaling by single-chain photoswitchable kinases. *Science* **355**, 836–842 (2017).
17. Toettcher, J. E., Weiner, O. D. & Lim, W. A. Using optogenetics to interrogate the dynamic control of signal transmission by the Ras/Erk module. *Cell* **155**, 1422–1434 (2013).
18. Guntas, G. et al. Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins. *Proc. Natl Acad. Sci. USA* **112**, 112–117 (2015).
19. Li, J. & Chen, P. R. Development and application of bond cleavage reactions in bioorthogonal chemistry. *Nature Chem. Biol*. **12**, 129–137 (2016).
20. Deiters, A., Groff, D., Ryu, Y., Xie, J. & Schultz, P. G. A genetically encoded photocaged tyrosine. *Angew. Chem. Int. Edn Engl*. **45**, 2728–2731 (2006).
21. Arbely, E., Torres-Kolbus, J., Deiters, A. & Chin, J. W. Photocontrol of tyrosine phosphorylation in mammalian cells via genetic encoding of photocaged tyrosine. *J. Am. Chem. Soc*. **134**, 11912–11915 (2012).
22. Hartshorn, M. J. et al. Diverse, high-quality test set for the validation of protein-ligand docking performance. *J. Med. Chem*. **50**, 726–741 (2007).
23. Park, H. et al. Simultaneous optimization of biomolecular energy function on features from small molecules and macromolecules. *J. Chem. Theory Comput*. **12**, 6201–6212 (2016).
24. Simanshu, D. K., Nissley, D. V. & McCormick, F. RAS proteins and their regulators in human disease. *Cell* **170**, 17–33 (2017).
25. Cox, A. D., Fesik, S. W., Kimmelman, A. C., Luo, J. & Der, C. J. Drugging the undruggable RAS: mission possible? *Nature Rev. Drug Discov*. **13**, 828–851 (2014).
26. Zhao, B. S., Roundtree, I. A. & He, C. Post-transcriptional gene regulation by mRNA modifications. *Nature Rev. Mol. Cell Biol*. **18**, 31–42 (2017).
27. Tomabechi, Y. et al. Crystal structure of nanoKAZ: The mutated 19 kDa component of *Oplophorus* luciferase catalyzing the bioluminescent reaction with coelenterazine. *Biochem. Biophys. Res. Commun*. **470**, 88–93 (2016).
28. Zhang, G. et al. Bioorthogonal chemical activation of kinases in living systems. *ACS Cent. Sci*. **2**, 325–331 (2016).
29. Holohan, C., Van Schaeybroeck, S., Longley, D. B. & Johnston, P. G. Cancer drug resistance: an evolving paradigm. *Nature Rev. Cancer* **13**, 714–726 (2013).
30. Diaz, J. E. et al. A Split–Abl kinase for direct activation in cells. *Cell Chem. Biol*. **24**, 1250–1258.e4 (2017).
31. Zheng, S., Fan, X., Wang, J., Zhao, J. & Chen, P. R. Dissection of kinase isoforms through orthogonal and chemical inducible signaling cascades. *ChemBioChem* **18**, 1593–1598 (2017).
32. Chang, L. & Karin, M. Mammalian MAP kinase signalling cascades. *Nature* **410**, 37–40 (2001).
33. Albeck, J. G. et al. Quantitative analysis of pathways controlling extrinsic apoptosis in single cells. *Mol. Cell* **30**, 11–25 (2008).
34. Riedl, S. J. & Shi, Y. Molecular mechanisms of caspase regulation during apoptosis. *Nature Rev. Mol. Cell Biol*. **5**, 897–907 (2004).
35. Mahrus, S. et al. Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. *Cell* **134**, 866–876 (2008).
36. Agard, N. J. et al. Global kinetic analysis of proteolysis via quantitative targeted proteomics. *Proc. Natl Acad. Sci. USA* **109**, 1913–1918 (2012).
37. Putt, K. S. et al. Small-molecule activation of procaspase-3 to caspase-3 as a personalized anticancer strategy. *Nature Chem. Biol*. **2**, 543–550 (2006).
38. Liu, S. et al. Solid tumor therapy by selectively targeting stromal endothelial cells. *Proc. Natl Acad. Sci. USA* **113**, E4079–E4087 (2016).
39. Shimbo, K. et al. Quantitative profiling of caspase-cleaved substrates reveals different drug-induced and cell-type patterns in apoptosis. *Proc. Natl Acad. Sci. USA* **109**, 12432–12437 (2012).
40. Wang, Y. et al. Chemotherapy drugs induce pyroptosis through caspase-3 cleavage of a gasdermin. *Nature* **547**, 99–103 (2017).

Acknowledgements We thank G. Jia, X. Zhang, W. Wang, Y. Wang and H. Song for providing the RNA oligonucleotide, expert technical assistance and helpful discussions. We thank the Computing Platform of the Center for Life Science for supporting the proteomic data analysis. This work was supported by research grants from the National Key Research and Development Program of China (2016YFA0501500 to P.R.C. and C.W.), the National Natural Science Foundation of China (21521003 to P.R.C. and C.W., 21740001 and 91753000 to P.R.C., 81490741 and 21778004 to C.W.), and a 'Young 1000-Talent Plan' Award to C.W.

Reviewer information *Nature* thanks Martin Schnermann, Klaus Michael Hahn, Nanda Vikas and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions P.R.C. and C.W. conceived the study. J.W. conducted most of the experiments unless otherwise specified. Y.L. developed the computational pipeline and Y.-J.L. conducted the LF-based pro-drug experiments in mice. S.-Q.Z., J.-Y.Z. and G.Z. contributed to the biochemical experiments. X.W. contributed to the chemical synthesis. F.Y. contributed to the proteomic experiments. J.W., Y.L., Y-J.L., C.W. and P.R.C. wrote the paper with inputs from all authors.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1188-1.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1188-1.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to C.W. or P.R.C.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Data reporting. No statistical methods were used to predetermine sample size. Xenograft mice were randomly selected before treatment. The investigators were blinded when grouping these mice and were not blinded for other experiments.

Chemical synthesis of ONBY. ONBY was synthesized according to the reported procedure[41]. ONBY used in cell-based experiments was further purified by HPLC with the 5–80% gradient of MeCN in H_2O. ^{1}H NMR (400 MHz, DMSO) δ 8.12 (d, J = 8.0 Hz, 1H), 7.83 – 7.76 (m, 2H), 7.66 – 7.59 (m, 1H), 7.22 (d, J = 8.5 Hz, 2H), 6.97 (d, J = 8.6 Hz, 2H), 5.43 (s, 2H), 4.00 (t, J = 6.0 Hz, 1H), 3.15 – 3.01 (m, 2H). ^{13}C NMR (101 MHz, DMSO) δ 170.88, 157.51, 148.05, 134.41, 132.87, 131.24, 129.85, 129.68, 128.39, 125.29, 115.26, 66.84, 54.10, 35.46. (d, double; J, coupling constant; m, multiple; t, triple).

Cell transfection and ONBY incorporation. The plasmid encoding a POI bearing the amber codon was first mixed with the plasmid encoding ONBY–RS–$tRNA^{Pyl}_{CUA}$ (CUA, the anticodon of the amber codon; Pyl, pyrrolysine) pair, followed by adding OPTI-MEM and XtremeGENE HP (Roche). These components are mixed at a ratio of 1.5 μg total plasmid to 100 μl OPTI-MEM, 2 μl XtremeGENE HP for 30 min and then added into 1 ml DMEM with 500 μM ONBY. Subsequently, the cell culture medium was exchanged to this mixture and cultured for another 24 h before the following experiments.

Photo-activation of POIs in living cells. Cells expressing POIs were subjected to UV irradiation without medium exchange. The plate or dish was put on the ChemiDoc XRS+ and irradiated from the bottom by UV light at intensity of 0.5 milliwatt per cm^2 for 3 min.

Evaluation and selection of the universal anchor residue for CAGE-prox. Although blocking the protein activity with a unified proximal cage is conceptually straightforward, the following re-activation step requires that the decaged residue at the anchor site (termed 'anchor residue') to be either the same as the original native residue or a mutation with negligible perturbation to protein activity. To survey which of the 20 natural amino acids would serve as the potential anchor residue, a total of 127 (details in Supplementary Notes) representative protein-ligand complexes with distinct binding pockets from the Astex dataset[22] was collected and a total number of 3,500 proximity residues was defined based on the minimum atom-atom distance to the ligand (<6 Å) with the percentage of natural abundance of each amino acid calculated. We also performed in silico site-saturation mutagenesis on these 3,500 positions and computed the averaged mutation stability change of each type of amino acid. When these values were plotted with the circle diameter showing the number of stable mutations ($\Delta G_f < 0$, Fig. 1b), Tyr, Phe and Leu were among the top three candidates for anchor residues that satisfy both of the abovementioned criteria (Fig. 1b, Extended Data Fig. 1). Because Phe and Leu are chemically inert for caging and decaging operations, we focused on Tyr. We next experimentally validated our computational analysis by comparing the performance of Tyr in parallel with Lys, which is another commonly used amino acid for decaging but is poorly ranked according to our in silico calculation. We mutated each of the 30 randomly selected sites on two model enzymes—FLuc and NLuc—into Tyr or Lys, and experimentally confirmed our prediction that the introduced Tyr mutations can largely maintain the enzyme activity, by comparison to the corresponding Lys mutations (Extended Data Fig. 1a–d). These results could be explained by the chemical basis of these two amino acids—owing to the strong hydrophilicity, terminal positive charge and flexibility of its side-chain, Lys can cause more marked changes in the chemical environment of the active-site pocket and thus significantly perturb the enzyme activity, as compared with Tyr.

General design and workflow of CAGE-prox. As shown in Extended Data Fig. 3, CAGE-prox started from a complex structure of the POI from the PDB. If there is no solved complex structure, homology model can be built by using Rosetta server[42] and a ligand-docking protocol[43] was used for predicting the possible binding mode (see 'Docking procedure'). The complex structure was relaxed in torsion space with constraints to eliminate all sterical clashes within experimental structure and optimized again in Cartesian space with ramped constraints to locate the energy minimum near the native state. After the initial structure was prepared, we applied two filters to identify potential anchor sites. First, all the residues within 6 Å around the ligand and that have reasonable direction (see the distance and angle definition in 'Definition of computational parameters in CAGE-prox and determination of their cut-offs') are defined as proximal residues. Second, for each mutant of proximal residue, Rosetta scores can be evaluated for both the apo and holo states and the changes of protein stability (ΔG_f) and binding affinity ($\Delta\Delta G_b$) are calculated by equations shown in Extended Data Fig. 2f, g (scripts for calculation: https://github.com/wendao/CAGE-prox). Two criteria are then applied to filter out mutants that are not stable enough ($\Delta G_f > 0.2$) or deleterious to the binding ligand ($\Delta\Delta G_b > 0.1$). About 6–10 mutants are typically labelled as the 'recommended' sites that are subject to site-specific incorporation of caged UAA[44–46] such as ONBY, and activity measurement before and after photo-decaging.

Definition of computational parameters in CAGE-prox and determination of their cut-offs. We used a recently developed protocol that can predict protein stability change introduced by a single point mutation accurately[23]. Optimized Rosetta scores can be calculated for both the apo and holo states after the 'relax' procedure. Protein stability change (ΔG_f) is defined as the Rosetta score difference between the wild-type and mutant apo state. Binding affinity (ΔG_b) is defined as the Rosetta score difference between the apo state and the holo state, and $\Delta\Delta G_b$ is then calculated as the difference in binding affinity between wild-type and mutant. The distance (D) between an anchor residue and the ligand is defined as the minimal atom–atom distance from all atoms of the residue to those of the ligand. The angle (A) that represents the residue direction is defined as the minimal axis crossing the angle between C_α–C_β and C_α to any atom of the ligand. 6 Å and 60 degrees are set as cut-off values for the distance and angle parameters for selecting anchor residues.

Docking procedure. AutoDock vina[47] was used to generate an initial binding conformation and the Rosetta ligand-docking protocol[43] was then applied. Ten thousand configurations were sampled and the top five structures were selected in a Pareto Optimality manner in terms of the total Rosetta score and the buried solvent accessible surface area (SASA) as shown in Extended Data Fig. 5b. Amber tools[48] and Acype software[49] were used to construct the force field parameters for the ligand-docking protocol.

Activity measurement of FLuc. The list of Fluc variants can be found in Supplementary Table 1. HEK293T cells were grown to 90% confluency for transfection in 12-well plates. Plasmids encoding the FLuc-TAG mutants were co-transfected with the plasmid bearing ONBY–RS–$tRNA^{Pyl}_{CUA}$ pair into cells via XtremeGENE HP (Roche) and cultured in the presence of 500 μM ONBY for another 24 h. Next, cells of each well were collected, washed by PBS buffer, lysed in 100 μl of the lysis buffer on ice for 15 min and centrifuged at 10,000g for 5 min. After the supernatant of each sample was divided into two 96-well plates, 1 plate was irradiated by UV light for 3 min and the other plate was kept in dark. Subsequently, each sample were mixed with 50 μl of the assay reagent and activity of each FLuc variant was measured by monitoring the bioluminescence of each well. The assay reagent was prepared by combining 1 volume of the luciferin solution with 25 volumes of the assay buffer from the firefly luciferase assay kit (Vigorous biotechnology, T003). The expression level of each variant was verified by immunoblotting analysis. Two cut-offs were introduced: (i) the activity readout by the plate reader equal to or higher than 1,000 was considered as an 'active' variant, and (ii) the activity ratio between caging and decaging equaled to or higher than 10 was considered as an 'activatable' variant.

Activity measurement of KRAS. The list of KRAS variants can be found in Supplementary Table 2. Plasmids encoding the 8 KRAS-TAG mutants were co-transfected with the plasmid bearing ONBY–RS–$tRNA^{Pyl}_{CUA}$ pair into cells as described above, and cultured for another 24 h in the presence of 500 μM ONBY. After activation by UV light as described above, cells were cultured at 37 °C for another 2 h. Culture medium of each well was removed and 100 μl of the loading buffer was added directly to prepare the lysate samples without other operations (to avoid further phosphorylation changes). The samples were heated at 95 °C for 20 min before being subject to SDS–PAGE and immunoblotting analysis.

Temporal monitoring of downstream signal transductions of KRAS. The samples were prepared as described above at different time points after UV irradiation (2, 5, 10, 20, 40, 60, 120 and 180 min), and the phosphorylation level of downstream substrates of KRAS was monitored by immunoblotting analysis. The EGF stimulation of HEK293T cells were conducted as a control. Cells without transfection were added with 100 ng of EGF for each well and the samples were prepared and analysed accordingly as described above.

Expression and purification of FTO. HEK293T cells were grown to 90% confluency for transfection. After expression, cells of each well were then collected, lysed and subjected to immunoblotting analysis to demonstrate the expression level of each mutant. Two sites (P93 and N205) with low incorporation efficiency of ONBY were excluded in the following experiments. Next, the rest of the seven FTO-TAG mutants bearing ONBY were expressed and purified from HEK293T cells as described: each FTO mutant was expressed in HEK293T cells (10^8 cells from two 15-cm dishes) for 36 h and then the cells were resuspended in 3 ml PBS contain 200 μM DTT for ultrasound lysis. The lysate was then centrifuged and to the supernatant was added 200 μl Flag beads for 2 h incubation at 4 °C. Next, the beads were washed rapidly in tris buffer (50 mM Tris-HCl, 150 mM NaCl, pH = 8, 200 μM DTT) for three times and eluted by 200 μl 3× Flag peptide solution (2 mg ml^{-1}, dissolved in PBS containing 100 μM DTT). The eluted protein was concentrated and the purity was verified by SDS–PAGE analysis.

Activity measurement of FTO. The list of FTO variants can be found in Supplementary Table 3 and the activity assay was performed according to a previously published protocol[50,51] with a 15-mer synthetic RNA oligonucleotide substrate (5′-AUUGAG(m^{6}A)CAGCAGC-3′). In brief, each purified mutant was subject to UV irradiation in a 96-well plate for 3 min or not, and then tested by the demethylation assay as described below. The reaction was performed at 37 °C for 30 min in a mixture which contains 0.2 μM RNA oligonucleotide bearing N6-methyladenine, 0.4 μM FTO, 200 μM of $(NH_4)_2Fe(SO_4)_2{\cdot}6H_2O$, 200 μM of

2-KG, 1 mM of L-ascorbic acid in 100 μl 50 mM HEPES buffer (pH = 7). After 30 min, the reaction was quenched by adding 5 mM EDTA and heating for 5 min at 95 °C. The mixture was then treated with 11 μl of a 100 mM NH_4Ac and 1 unit of nuclease P1 (Sigma, N8630), and incubated for 3 h at 42 °C. The reaction mixture was then added with 13 μl of 1 M NH_4HCO_3 and 1 unit of alkaline phosphatase (Sigma, P4252), and incubated for another 3 h at 37 °C. After centrifugation, the contents of A and m^6A in the supernatant were measured by the liquid chromatography with tandem mass spectrometry (LC/MS–MS) analysis with a ZORBAX SB-Aq column (Agilent) and a Triple Quad 5500 (AB SCIEX) mass spectrometer. A relative quantification of each nucleoside was performed by calculating the areas of corresponding chromatographic peaks.

Activity measurement of NLuc. The list of Nluc variants can be found in Supplementary Table 4. HEK293T cells were grown to 90% confluency for transfection in 12-well plates. Plasmids encoding the NLuc-TAG mutants were co-transfected with the plasmid bearing ONBY–RS–$tRNA^{pyl}_{CUA}$ pair into cells via XtremeGENE HP (Roche) and cultured in the presence of 500 μM of ONBY for another 24 h. After activation by UV light as described above, cells were collected from each well, washed by PBS and lysed in 100 μl of the lysis buffer (Thermo Fisher, 78503) on ice for 15 min. The samples were added with 100 μl of the assay reagent and the bioluminescence was measured. The assay reagent was prepared by combining 1 volume of the NLuc substrate with 100 volumes of the assay buffer from the nano luciferase assay kit (Nano-Glo Luciferase Assay System, Promega, N1110).

Activity measurement of MEK1. The list of MEK1 variants can be found in Supplementary Table 5. HEK293T cells were grown to 70–80% confluency for transfection in 12-well plates. Plasmids encoding the 9 MEK1-TAG mutants were co-transfected respectively with the plasmid bearing ONBY–RS–$tRNA^{pyl}_{CUA}$ pair into cells and cultured for another 24 h in the presence of 500 μM ONBY. After activation by UV light as described above, cells were cultured at 37 °C for another 2 h. Culture medium of each well was removed and 100 μl of the loading buffer was added directly to prepare the lysate samples, which were then heated at 95 °C for 20 min and subjected to immunoblotting analysis.

Profiling the inhibitor-resistant activities of MEK1 variants. The susceptibility of three CAGE-prox-activatable MEK1 variants against eight commercial MEK inhibitors are surveyed[52–59] (Extended Data Fig. 6c). The MEK1 variants were expressed and activated as described above, and the inhibitor was added with a final concentration fivefold higher than its corresponding IC_{50} value. Wild-type MEK1 was used as a control to demonstrate the efficacy of these inhibitors. The activity of each MEK1 variant was measured as described above. The detailed information of each inhibitor can be found in Supplementary Table 8.

Evaluation of the binding affinity of each inhibitor–MEK1 pair by computational modelling. The computational modelling method was used as previously reported[26]. In brief, to investigate interactions between the inhibitors (I1–I8) and selected mutants (I216, D217, N221), we first collected MEK1 complex structures co-crystallized with ligands from the PDB that are either the same as or similar to each of the eight inhibitors. Then, we aligned our predicted structure (with ED mutations) to the complex and calculated ΔG_f and $\Delta\Delta G_b$ for each mutant–inhibitor pair, which allows calculation of the final binding energy using the equation from ref. [60], with $\Delta G_0 = -1.0$, kT = 1.0 as the default parameters and 0.5 as the cut-off.

Quantification of the downstream IL-8 secretion. After activation of MEK1-N221-ONBY for 12 h in the presence or absence of inhibitor 8, cells were centrifuged and the supernatant was collected. Culture medium was concentrated and subjected for IL-8 concentration measuring using Human CXCL8/IL-8 Quantikine ELISA Kit (R&D, D8000C), according to the manufacturer's instructions.

Activation of MEK1 in Jurkat cells. For transfection by electroporation, Jurkat cells were cultured in RPMI 1640 supplemented with 10% fetal bovine serum (FBS) for at least 1 day. Cells (4×10^7) were spun down, washed once with PBS buffer and re-suspended in 300 μl of Gene Pulser Electroporation Buffer (Bio-rad 1652676). Twelve micrograms of the plasmid encoding MEK-N221TAG mutant and 9 μg of the plasmid bearing ONBY–RS–$tRNA^{pyl}_{CUA}$ pair were added to each transfection, mixed briefly and incubated at room temperature for 2 min. The cell–DNA mixture was then subjected to electroporation (BioRad, square pulse, 300 V, 10-ms pulse, 0.4-cm cuvette). Cells were rested at room temperature for 10 min before being re-suspended in fresh RPMI 1640 medium and cultured for another 36–48 h for further experiments. After transfection, cells containing MEK1-221-ONBY were irradiated with UV for 5 min, cultured at 37 °C for another 1 h and lysed in TNE buffer containing protease inhibitor cocktail (Roche, 4693132001). The lysate was subjected to immunoblotting analysis using corresponding antibodies.

Activity measurement of caspase-3. The list of caspase-3 variants can be found in Supplementary Table 6. HEK293T cells were grown to 70% confluency for transfection in 12-well plates. Cell culture medium containing 300 μM ONBY was heated to 37 °C and filtered through a 0.22-μm nylon membrane before use. Plasmids encoding the 10 caspase-3–TAG mutants were co-transfected with the plasmid bearing ONBY–RS–$tRNA^{pyl}_{CUA}$ pair into cells and cultured for another 24 h. After activation by UV light as described above, cells of each well were collected and suspended in 100 μl of cell lysis buffer on ice for 15 min. Subsequently, the lysate was centrifuged at 10,000*g* for 5 min and 50 μl of supernatant was mixed with 50 μl of 2× reaction buffer (containing 5 mM DTT and 50 μM DEVD-AFC substrates) followed by incubation at 37 °C for 15 min. Activity of caspase-3 was measured by monitoring the fluorescence of each well (ex: 400 nm, em: 505 nm). Lysis buffer, 2× reaction buffer and DEVD-AFC substrates are all available from the Caspase-3 Assay Kit (abcam, ab39388). Similarly, the activities of caspase-1, -2, -4, -5, -6, -8, -9 and -10 were monitored via the same procedure by the Caspase Family Fluorometric Substrate Kit II Plus (abcam, ab102487) before and after caspase-3 activation.

Monitoring cell morphological changes after caspase-3 activation. HeLa cells or HEK293T cells were grown to 40% confluency in an eight-well chamber for transfection and expression of caspase-3-EGFP-M61-ONBY. After 24 h expression and 3 min UV irradiation as described above, cell imaging was then taken on a LSM700 laser scanning confocal microscope (Zeiss) in the time-lapse manner under normal cell culturing conditions (37 °C with 5% CO_2). Frames were captured with 2-min or 5-min intervals and the initiation time of UV irradiation was set as time zero.

Temporal substrate profiling of CAGE-prox-activated caspase-3[61–64]. Control cells (2×10^7) or caspase-3 activated cells (2×10^7) were resuspended in 1 ml of the cell lysis buffer (Thermo Fisher, 78503) and incubated with the protease inhibitor cocktail (Roche, 4693132001) on ice for 30 min. Cell lysates were then centrifuged (10,000*g*, 10 min) to collect soluble supernatants. Protein concentrations were determined by Protein quantitation assay kit (Pierce BCA Protein Assay Reagent A, 23223) and diluted to 2 mg ml^{-1}. Proteomes were denatured in 6 M urea–TEAB with 10 mM of DTT for 1 h and then incubated with 20 mM iodoacetamide for 30 min. Subsequently, 100 μl samples were added with 10 μg of Trypsin Gold and rLys-C mixture (Promega, V5072) for 16 h incubation at 37 °C. After trypsin digestion, 8 μl of 4% heavy formaldehyde ($D^{13}CDO$, Sigma, 596388) was added into the control sample and 8 μl of 4% light formaldehyde (HCHO, Sigma, F8775) was added into the caspase-3 activated sample, followed by addition of 5 μl of 1 M sodium cyanoborohydride into each sample and incubation at room temperature for 1 h. The dimethyl labelling reactions were quenched by adding 30 μl of 1% ammonia into each sample. Finally, 2 μl of formic acid was added to each sample before they were combined.

The samples were centrifuged (100,000*g*, 10 min, 4 °C) and fractionated by a two-dimensional HPLC (2D-HPLC). The first dimension of high pH was performed on an Agilent 1260 infinity quaternary LC using a Durashell RP column (5 μM, 150 Å, 250 mm × 4.6 mm, Agela) and an alkaline mobile phase (0.5mM $NH_3{\cdot}H_2O$). Eluents were collected every minute in the 30 min gradient of 1.5 ml min^{-1} flow rate: (0.1 min, 2% ACN; 2.0 min, 8% ACN; 15 min, 20% ACN; 25 min, 30% ACN; 26 min, 95% ACN; 28 min, 95% ACN; 29 min, 15% ACN; 30 min, 15% ACN). A total of 30 samples was collected and then combined into 10 fractions. Each fraction was dried under vacuum and reconstituted in 20 μl of 0.1% (v/v) formic acid in water. Subsequently, the second dimensional low pH reverse-phase chromatography of each faction was performed on an Ultimate 3000 LC system with an adjusted linear gradient of 350 nl min^{-1} and subject to MS/MS analysis on a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific).

The MS/MS analysis was performed under the positive-ion mode with a full-scan *m/z* range from 350 to 1,800 and a mass resolution of 70,000. MS/MS fragmentation was performed in a data-dependent mode, of which TOP 20 most intense ions are selected for MS2 analysis with a resolution of 17,500 using the collision mode of HCD. Other important parameters: isolation window, 2.0 *m/z* units; default charge, +2, charge exclusion, unassigned, 8, >8 ; normalized collision energy, 28%; maximum IT, 50 ms; dynamic exclusion, 20.0 s.

Dimethylation labelled peptides were identified by ProluCID[48] and DTASelect[65] with a precursor mass tolerance of 20 p.p.m. and fragment ion mass tolerance of 20 p.p.m. and static modification of cysteine (+57.0215 Da) and variable oxidation of methionine (+15.9949 Da). The isotopic modifications (28.0313 and 34.0631 Da for light and heavy labelling, respectively) were set as fixed modifications on the peptide N termini and variable modifications on lysine residues. Half-tryptic termini and up to two missing trypsin cleavages are allowed. The quantification of light/heavy ratios was performed by an in-house software CIMAGE[66] with precursor mass tolerance of 25 p.p.m.

Analysis of proteolytic substrates upon caspase-3 activation. To define proteolytic substrates at the protein level, all the non-Asp-containing tryptic peptides are considered as 'background' and a Gaussian distribution fitted by histogram of the log value of ratios are rescaled to have the mean value at 0 and standard deviation equal to 1. The cut-off of ±0.693 (corresponding to ratios at 0.5 and 2.0, respectively) are used for defining a 'significant change'. A protein is considered as a candidate substrate if either of the two conditions is met: (i) it has a tryptic peptide with a heavy-to-light ratio larger than 2.0; or (ii) it has a newly generated half-tryptic

peptide and a light-to-heavy ratio larger than 5.0. Only if the protein is identified as a candidate substrate commonly from all the three replicates, is it included in the final list of proteolytic substrates as reported in Supplementary Data 2. In this list, we also cross-checked with DegraBase[67] to provide information on potential cleavage sites. Considering the specificity of caspase-3, we also extracted those proteins with the Asp-containing tryptic peptides or Asp-terminated half tryptic peptides and reported them in the Supplementary Data 2 as the 'high-confidence' substrates.

To define the cleavage sites explicitly, we extracted from the aforementioned protein substrate list all the half-tryptic peptides that have light-to-heavy ratios larger than 5.0 in 2 out of the 3 replicates and generated sequence logos using IceLogo with the human genome as the background.

The previously reported substrates of caspase-3 used for the comparison in Fig. 3e were defined by combining two datasets of caspase-3 substrates by in vitro cleavage experiments[36,68]. The total number of these defined substrates is 1,240. The previously reported substrates of apoptosis were defined by combining four datasets of proteolytic substrates identified during apoptosis[35,63,69,70]. The total number of defined substrates of apoptosis is 1,156.

Validation of the newly identified substrates of caspase-3 by CAGE-prox. Bio-Spin and desalting columns (BIO-RAD, 7326227) were washed three times with water and equilibrated 3 times with 2 × assay buffer (from the Caspase-3 Assay Kit, abcam) before use. HEK293T cells (1×10^7) were collected and suspended in 200 μl of lysis buffer (ThermoFisher, 78503) on ice for 30 min. After centrifugation at 10,000*g* for 5 min, 100 μl of the collected supernatant was added to the column and the eluted solution was mixed with the recombinant wild-type caspase-3 protein (Sino Biological, 10050-H08E) with a 20 μg ml^{-1} final concentration in presence or absence of a caspase-3 inhibitor (2 μM). After incubation at 37 °C for 2 h, the samples were added with the gel-loading buffer and subjected to immunoblotting analysis with the corresponding antibody. The epitope of each antibody was carefully chosen to be away from the cleavage site of the substrates.

Validation of the kinetics of V-ATPase cleavage by caspase-3. HEK293T cells were collected and lysed as described above. Two hundred microlitres of the cell lysates was added with the recombinant wild-type caspase-3 protein (Sino Biological, 10050-H08E) at a final concentration of 40 μg ml^{-1}. Subsequently, the sample was divided into 7 aliquots and incubated at 37 °C for 0, 0.5, 1, 2, 3, 4 and 6 h respectively. The reactions were stopped by adding the gel-loading buffer directly and the samples were subject to immunoblotting analysis with antibodies for V-ATPase subunit A and B, respectively[71,72].

Validation of the cleavage site by caspase-3 on HMGB1. Flag–HMGB1 and wild-type caspase-3 was co-transfected into HEK293T cells and cultured for 24 h, and the cells and culture medium were collected, respectively, for measurement of cleavage of intracellular and extracellular HMGB1 by caspase-3. One millilitre of the collected culture medium was centrifuged at 10,000*g* for 5 min and concentrated to 100 μl before immunoblotting analysis. The cells were lysed directly by loading buffer and subjected to immunoblotting analysis. Six 'D-to-A' mutants of HMGB1 were constructed and expressed in HEK293T cells for determining the cleavage site by caspase-3[73].

Expression and purification of LF and PA. The list of LF variants can be found in Supplementary Table 7. Plasmids encoding the LF-TAG mutants (pET-21a vector) were co-transformed with the plasmid bearing the ONBY–RS–$tRNA^{pyl}_{CUA}$ pair into BL21 (DE3) cells. Colonies were picked and grown in LB medium overnight at 37 °C. After 1:100 dilutions in 1 l LB medium, cells were left to grow at 37 °C to an OD_{600} of 0.6, and then 500 μM ONBY was added to the culture (final concentration). The bacteria were grown at 22 °C for another 30 min before the addition of 0.2% arabinose (final concentration) and grown at 22 °C for 16 h. Cells were then collected and suspended in lysis buffer (20 mM Tris-HCl, pH = 7.5, 500 mM NaCl). After sonication and centrifugation (15,000*g*, 30 min), lysate was loaded onto a Ni-NTA column (Histrap 5 mL, GE Healthcare), which was washed with 50 ml washing buffer (20 mM Tris-HCl, pH = 7.5, 500 mM NaCl, 30 mM imidazole) and eluted with elution buffer (20 mM Tris-HCl, pH 7.8, 500 mM NaCl and 200 mM imidazole) to yield LF variants bearing ONBY. Target proteins were finally desalted into PBS buffer and concentrated to 5 mg ml^{-1} before usage. PA was purified according to the previously reported protocols[74].

Activity measurement of LF. HEK293T cells were grown to 90% confluency and LF–PA or LF-ONBY/PA were added with 1 μg ml^{-1} final concentration of each component. After 3 min UV irradiation and 1 h incubation, cells were collected and lysed, followed by SDS–PAGE and immunoblotting analysis with a polyclonal antibody against MEK3 (C-15, Santa Cruz Biotechnology).

Cell proliferation assay under LF treatment. A375 and HEK293T cells were plated in 96-well plates with a density of 2,000 cells per well and cultured for 24 h before the following experiments. The medium was changed to fresh DMEM with 10% FBS, which contains LF–PA or LF-ONBY/PA at a 1 μg ml^{-1} final concentration. Medium change and UV irradiation (3 min) were performed 24 h and 48 h afterwards and cell proliferations were measured by the CCK8 reagent. Ten microlitres of CCK8 was added to each well and cells were incubated at 37 °C for 1 h, followed by absorbance measurement at 450 nm by Microplate Reader.

CAGE-prox-enabled LF activation in living mice. Male BALB/c nude mice (6–8 weeks old from Vital River Laboratories) were used in this study. All animal experiments were carried out in accordance with the national guidelines for housing and care of laboratory animals (Ministry of Health, China) and the protocol is in accordance with institutional regulations after review and approval by the Animal Care and Use Committee of the Peking University, Beijing.

The light source used was a hand-held UV lamp with a 15-W device power. The light intensity was 7.5 mW cm^{-2} when the irradiated spot was about 1 cm away from the light without focusing. HEK293T cells (1×10^7), which expressed FLuc-Y255-ONBY, were subcutaneously injected into each leg. After 1 h, the left leg was irradiated by 365-nm UV light for 5 min. Next, 150 μl of luciferin (15 mg ml^{-1}) was administered at a dose of 90 mg kg^{-1} by intraperitoneal injection. After another 20 min, the in vivo bioluminescence images were captured on an IVIS Lumina II imaging system (PerkinElmer). For time-dependent activation of FLuc-Y255-ONBY in mice, 1×10^7 HEK293T cells expressing FLuc-Y255-ONBY were subcutaneously injected and each leg was irradiated by 365-nm UV light for 0, 0.5, 2.5, 5 and 10 min.

In vitro and in vivo evaluation of the penetration efficiency of UV. For the in vitro assay, the stripped mice skin was used to cover a 96-well plate containing lysate of cells expressing FLuc-Y255-ONBY. After UV irradiation for 5 min, each sample was mixed with 50 μl of the assay reagent and the activity of each FLuc variant was measured by monitoring the bioluminescence of each well as described above. For the in vivo assay, 5×10^5 of A375 cells were injected in the armpit of each Nu/Nu nude mice to prepare a xenograft model of human melanoma. When the volume of tumours were grown to around 100 mm^3, the mice were intratumorally injected with 10 μl of ONB-caged resorufin (at 10 mM) and followed by a 5-min UV irradiation. After 2 h, tumours were then excised and cut as 10-μm slices at 4 °C on Leica CM 1950. The slices were transferred into PBS solution, washed three times at 4 °C, stained with 1 μg ml^{-1} Hoechst for 20 min and washed 3 times again at room temperature. For imaging, the slices were mounted between coverslips and glass slides, and the fluorescence imaging was conducted on a LSM700 laser scanning confocal microscope (Zeiss) in the Tile-scan manner.

Bio-distribution of LF-ONBY in mice. Mice xenograft models were prepared as described above. When tumours were grown to around 200–500 mm^3, Cy5-labelled LF-ONBY and PA were intratumorally injected at 0.5 mg kg^{-1} and 1 mg kg^{-1}, respectively. The in vivo fluorescence imaging on mice was then conducted by using the IVIS Lumina II imaging system (PerkinElmer) after 2, 8, 24, 30 and 48 h, respectively. The Cy5-labelled LF-ONBY was prepared by adding Cy5-NHS (in DMF) to the LF-ONBY (4 mg ml^{-1} in PBS, pH = 8.0) to reach a final concentration of 50 μM. After incubation overnight at 4 °C, the Cy5-labelled LF-ONBY was desalted twice by Bio-Spin.

Immunogenicity measurement of LF–PA. The immunogenicity measurement of LF–PA fused with EGF or ZHer tag was conducted according to the previously reported protocol[75]. In brief, male C57BL/6 mice (3–4 weeks old) were injected intraperitoneally with 6 doses of wild-type PA, PA-EGF or PA-ZHer (0.5 mg kg^{-1}) within a period of 2 weeks. Ten days later, blood samples were collected from the eye socket vEin in mice and centrifuged at 1,000*g* to obtain the serum fraction. Biotinylated PA or LF (0.2 ml; 0.1 mg ml^{-1}) was added to each well of a streptavidin coated 96-well plate (Sigma, S6940). After incubation overnight at 4 °C, the plate was washed 3 times with PBST (0.1% Tween 20) to remove unbound PA or LF. 5% BSA was added to the plate and incubated for 1 h at room temperature and the plate was washed again with PBST 3 times. Mouse serum was diluted by 5,000 folds and incubated in the plate for 1 h. The plate was washed three times and incubated with a 5,000× diluted mouse antibody (CST, 7076S) for 1 h. The substrate solution (R&D, D8000C) was added to each well and, after 0.5 h, the reaction was stopped by adding the stop solution and measured for OD_{450} by Microplate Reader.

The immunogenicity measurement of LF–PA after small molecule immunosuppressor treatment was conducted according to the previously reported protocol[75]. Male C57BL/6 mice (3–4 weeks old) were injected intraperitoneally with 3 doses of 0.5 mg kg^{-1} of pentostatin (Targetmol, T4006) and 10 mg kg^{-1} of cyclophosphamide (Selleck, S2057)[60]. For the control group, 100 μl of PBS was injected. The rest of the procedures were conducted as described above.

Dose tolerance of LF–PA on mice. The maximum tolerated doses of LF–PA (ratio 2:1) and LF-ONBY:PA (ratio 2:1) were determined using a dose escalation protocol aimed at minimizing the number of mice used. Male C57BL/6 mice (6–8 weeks old, $n = 4$) in each group were intraperitoneally injected of LF–PA 6 times at a frequency of every 2 days within 2 weeks. The survival rates of each group were monitored every day during this two-week period[38,75].

Controlled LF activation as a pro-drug therapy for cancer. Male Nu/Nu nude mice (3–4 weeks old) were used in this study. To prepare the xenograft model of human melanoma on mice, 5×10^5 A375 cells were injected in the armpit of each Nu/Nu nude mouse. Tumour-bearing mice were randomized into groups when the

volume of tumours was grown to about 50 mm^3. Each mouse bearing a tumour was treated with LF–PA or control reagents every two days during a two-week period at the dose of 0.8 mg kg^{-1} PA and 0.4 mg kg^{-1} LF in 100 μl PBS buffer. A total of 6 injections of LF–PA or control reagents was performed and a 5-min UV irradiation was performed 2 h after each injection. After two weeks, the mice were euthanized by CO_2 inhalation and tumours were collected, with the tumour volume measured as described (volume = length × ($width^2$)/2).

Manipulation of Src phosphorylation by CAGE-prox. The anchor sites of Src-Y416 are D413, N414, E415, T417, A418 and R419, and they were selected directly near the phosphorylation site of Y416. HEK293T cells were grown to 90% confluency and co-transfected with plasmids encoding the Src-TAG mutants and ONBY–RS–$tRNA^{pyl}_{CUA}$ pair in a 12-well plate. After 24 h expression followed by UV triggered decaging, cells were cultured at 37 °C for another 3 h. Next, medium of each well was then removed and 100 μl of loading buffer was added directly to prepare the lysate samples. Subsequently, the samples were heated at 95 °C for 20 min and then subject to immunoblotting analysis.

Statistical analysis. Data were summarized using descriptive statistics. Comparisons between groups were carried out using the two tailed *t*-test. In the statistical analysis, the data were unpaired, and assumed that both population of the two compared groups are Gaussian distributed and have the same standard deviations. Prism 7 (GraphPad) was used for statistical analysis. *P* value of less than 0.05 was considered statistically significant.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

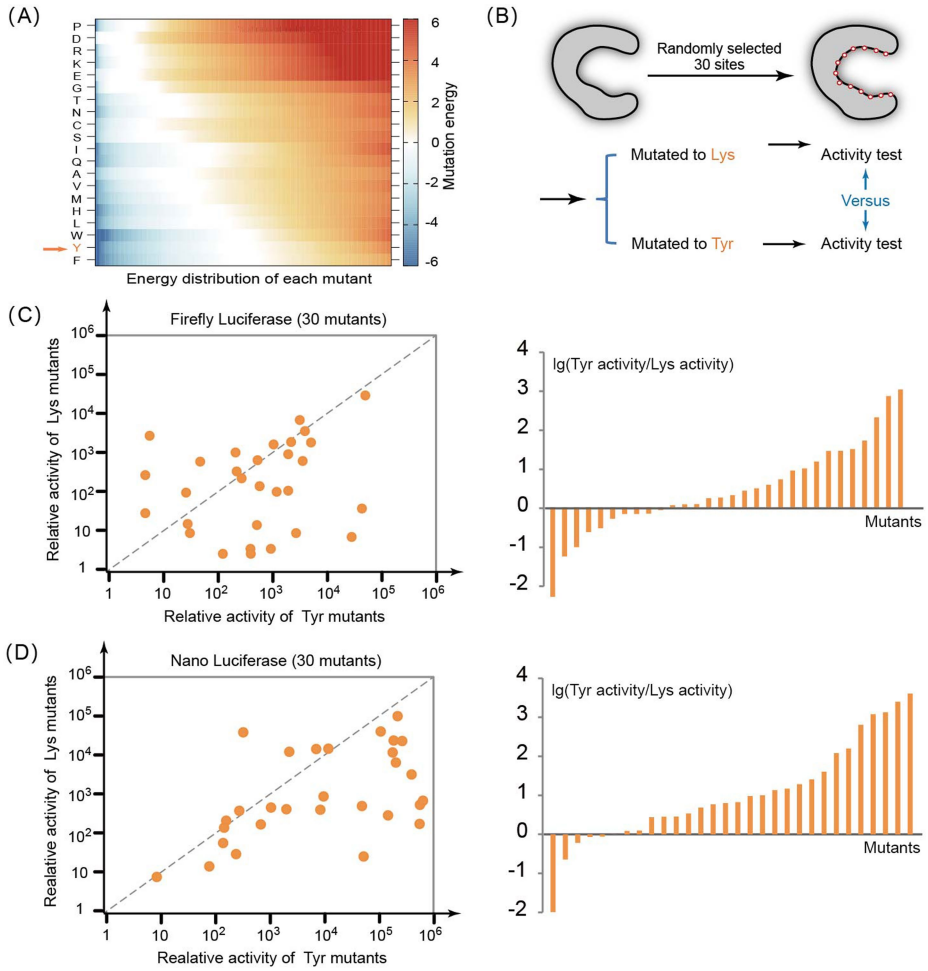
Data availability

All relevant data are included in the Article or its Supplementary Information. More details are available from the corresponding authors upon request.

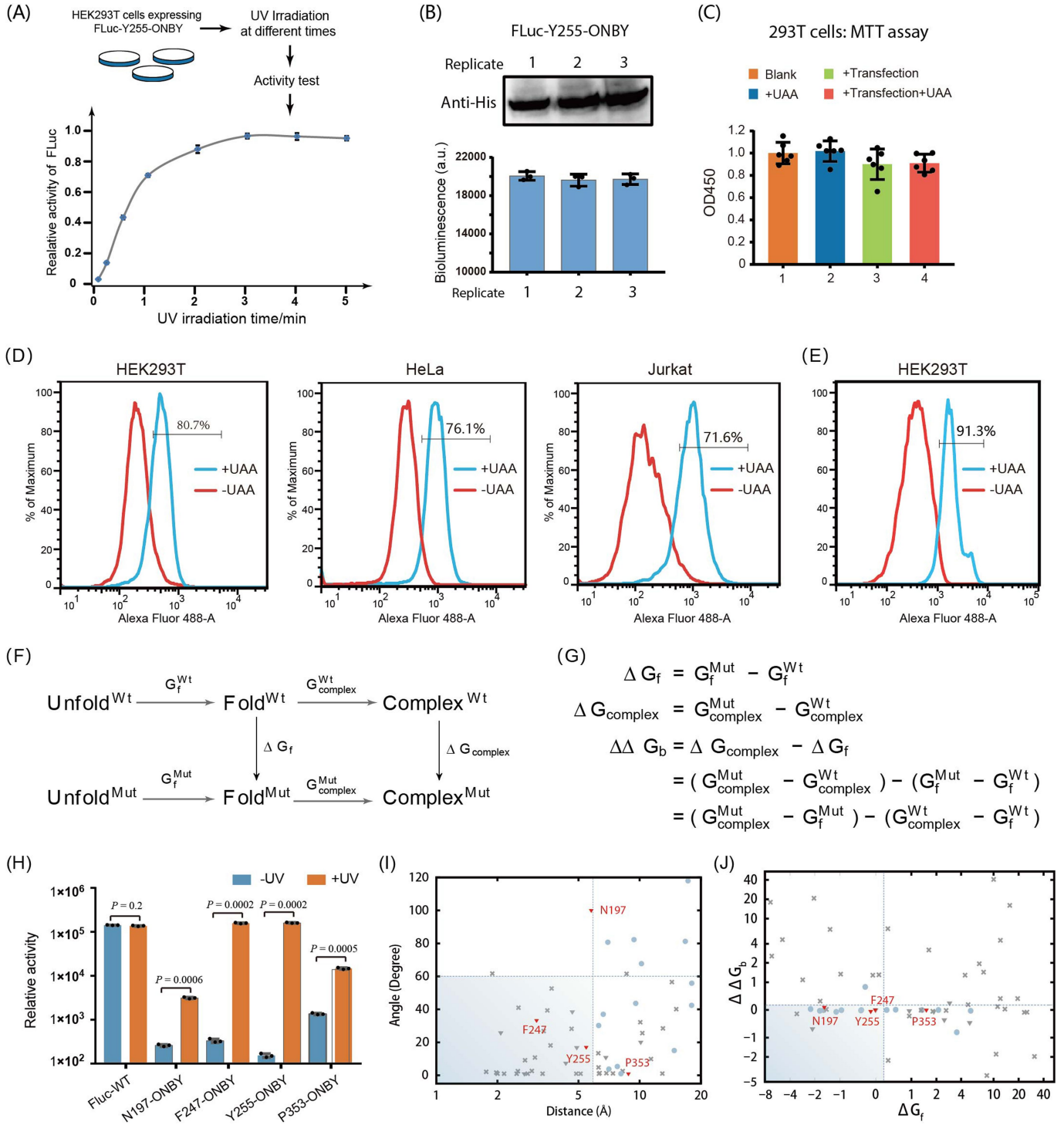
Code availability

All relevant code is available on GitHub at https://github.com/wendao/CAGE-prox. All other code is available from the corresponding authors upon reasonable request.

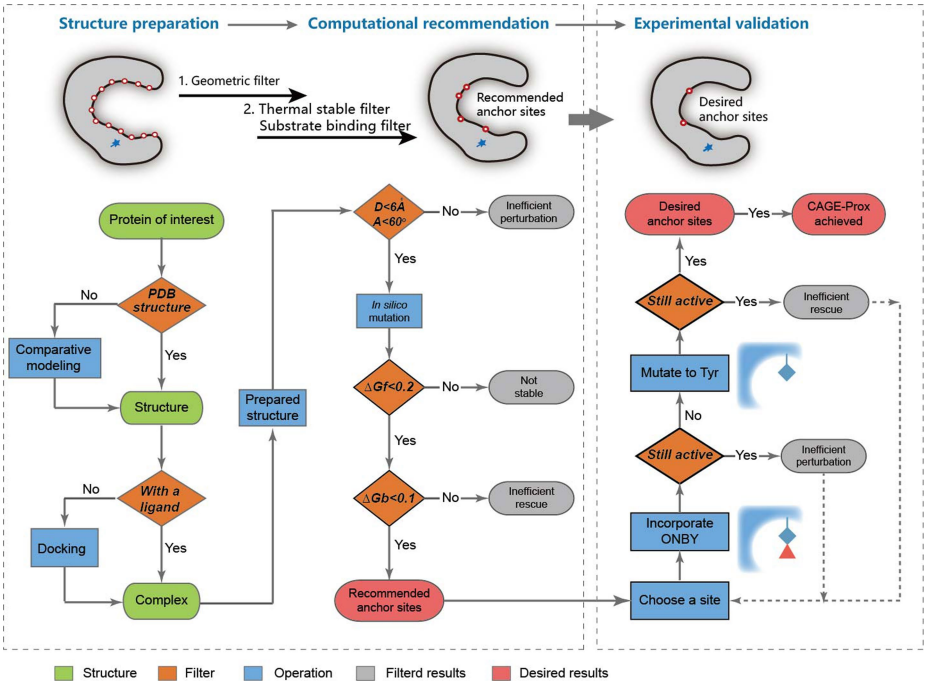
41. Miller, J. C., Silverman, S. K., England, P. M., Dougherty, D. A. & Lester, H. A. Flash decaging of tyrosine sidechains in an ion channel. *Neuron* **20**, 619–624 (1998).
42. Kim, D. E., Chivian, D. & Baker, D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* **32**, W526-31 (2004).
43. Meiler, J. & Baker, D. ROSETTALIGAND: protein-small molecule docking with full side-chain flexibility. *Proteins* **65**, 538–548 (2006).
44. Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **550**, 53–60 (2017).
45. Li, J., Jia, S. & Chen, P. R. Diels-Alder reaction-triggered bioorthogonal protein decaging in living cells. *Nature Chem. Biol.* **10**, 1003–1005 (2014).
46. Luo, J. et al. Genetically encoded optochemical probes for simultaneous fluorescence reporting and light activation of protein function with two-photon excitation. *J. Am. Chem. Soc.* **136**, 15551–15558 (2014).
47. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
48. Salomon-Ferrer, R., Case, D. A. & Walker, R. C. An overview of the amber biomolecular simulation package. *Wires. Comput. Mol. Sci.* **3**, 198–210 (2013).
49. Sousa da Silva, A. W. & Vranken, W. F. ACPYPE - AnteChamber Python parser interface. *BMC Res. Notes* **5**, 367 (2012).
50. Jia, G. et al.; J. G et al. N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature Chem. Biol.* **7**, 885–887 (2011).
51. Fu, Y. et al. FTO-mediated formation of N6-hydroxymethyladenosine and N6-formyladenosine in mammalian RNA. *Nature Commun.* **4**, 1798 (2013).
52. Wallace, E. M., Lyssikatos, J. P., Yeh, T., Winkler, J. D. & Koch, K. Progress towards therapeutic small molecule MEK inhibitors for use in cancer therapy. *Curr. Top. Med. Chem.* **5**, 215–229 (2005).
53. Duncia, J. V. et al. MEK inhibitors: the chemistry and biological activity of U0126, its analogs, and cyclization products. *Bioorg. Med. Chem. Lett.* **8**, 2839–2844 (1998).
54. Yamaguchi, T., Kakefuda, R., Tajima, N., Sowa, Y. & Sakai, T. Antitumor activities of JTP-74057 (GSK1120212), a novel MEK1/2 inhibitor, on colorectal cancer cell lines in vitro and in vivo. *Int. J. Oncol.* **39**, 23–31 (2011).
55. Dong, Q. et al. Discovery of TAK-733, a potent and selective MEK allosteric site inhibitor for the treatment of cancer. *Bioorg. Med. Chem. Lett.* **21**, 1315–1319 (2011).
56. Hatzivassiliou, G. et al.; H. G et al. Mechanism of MEK inhibition determines efficacy in mutant KRAS- versus BRAF-driven cancers. *Nature* **501**, 232–236 (2013).
57. Phadke, M. S., Sini, P. & Smalley, K. S. The novel ATP-competitive MEK/Aurora kinase inhibitor BI-847325 overcomes acquired BRAF inhibitor resistance through suppression of Mcl-1 and MEK expression. *Mol. Cancer Ther.* **14**, 1354–1364 (2015).
58. Hoeflich, K. P. et al. Intermittent administration of MEK inhibitor GDC-0973 plus PI3K inhibitor GDC-0941 triggers robust apoptosis and tumor growth inhibition. *Cancer Res.* **72**, 210–219 (2012).
59. Ohren, J. F. et al. Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. *Nature Struct. Mol. Biol.* **11**, 1192–1197 (2004).
60. Whitehead, T. A. et al. Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing. *Nature Biotechnol.* **30**, 543–548 (2012).
61. Lakhani, S. A. et al. Caspases 3 and 7: key mediators of mitochondrial events of apoptosis. *Science* **311**, 847–851 (2006).
62. Perinpanayagam, M. A. et al. Regulation of co- and post-translational myristoylation of proteins during apoptosis: interplay of N-myristoyltransferases and caspases. *FASEB J.* **27**, 811–821 (2013).
63. Dix, M. M., Simon, G. M. & Cravatt, B. F. Global mapping of the topography and magnitude of proteolytic events in apoptosis. *Cell* **134**, 679–691 (2008).
64. Dix, M. M. et al. Functional interplay between caspase cleavage and phosphorylation sculpts the apoptotic proteome. *Cell* **150**, 426–440 (2012).
65. Tabb, D. L., McDonald, W. H. & Yates, J. R., III. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **1**, 21–26 (2002).
66. Weerapana, E. et al. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790–795 (2010).
67. Crawford, E. D. et al. The DegraBase: a database of proteolysis in healthy and apoptotic human cells. *Mol. Cell. Proteomics* **12**, 813–824 (2013).
68. Wang, C. et al. A bead-based cleavage method for large-scale identification of protease substrates. *Sci. Rep.* **6**, 22645 (2016).
69. Stoehr, G., Schaab, C., Graumann, J. & Mann, M. A SILAC-based approach identifies substrates of caspase-dependent cleavage upon TRAIL-induced apoptosis. *Mol. Cell. Proteomics* **12**, 1436–1450 (2013).
70. Lüthi, A. U. & Martin, S. J. The CASBAH: a searchable database of caspase substrates. *Cell Death Differ.* **14**, 641–650 (2007).
71. Young, L. C., Thulien, K. J., Campbell, M. R., Tron, V. A. & Andrew, S. E. DNA mismatch repair proteins promote apoptosis and suppress tumorigenesis in response to UVB irradiation: an in vivo study. *Carcinogenesis* **25**, 1821–1827 (2004).
72. Nakashima, S. et al. Vacuolar H^+-ATPase inhibitor induces apoptosis via lysosomal dysfunction in the human gastric cancer cell line MKN-1. *J. Biochem.* **134**, 359–364 (2003).
73. Victor, K. G. et al. Proteomic identification of synaptic caspase substrates. *Synapse* **72**, e22014 (2018).
74. Miller, C. J., Elliott, J. L. & Collier, R. J. Anthrax protective antigen: prepore-to-pore conversion. *Biochemistry* **38**, 10432–10441 (1999).
75. Liu, S. et al. Matrix metalloproteinase-activated anthrax lethal toxin demonstrates high potency in targeting tumor vasculature. *J. Biol. Chem.* **283**, 529–540 (2008).



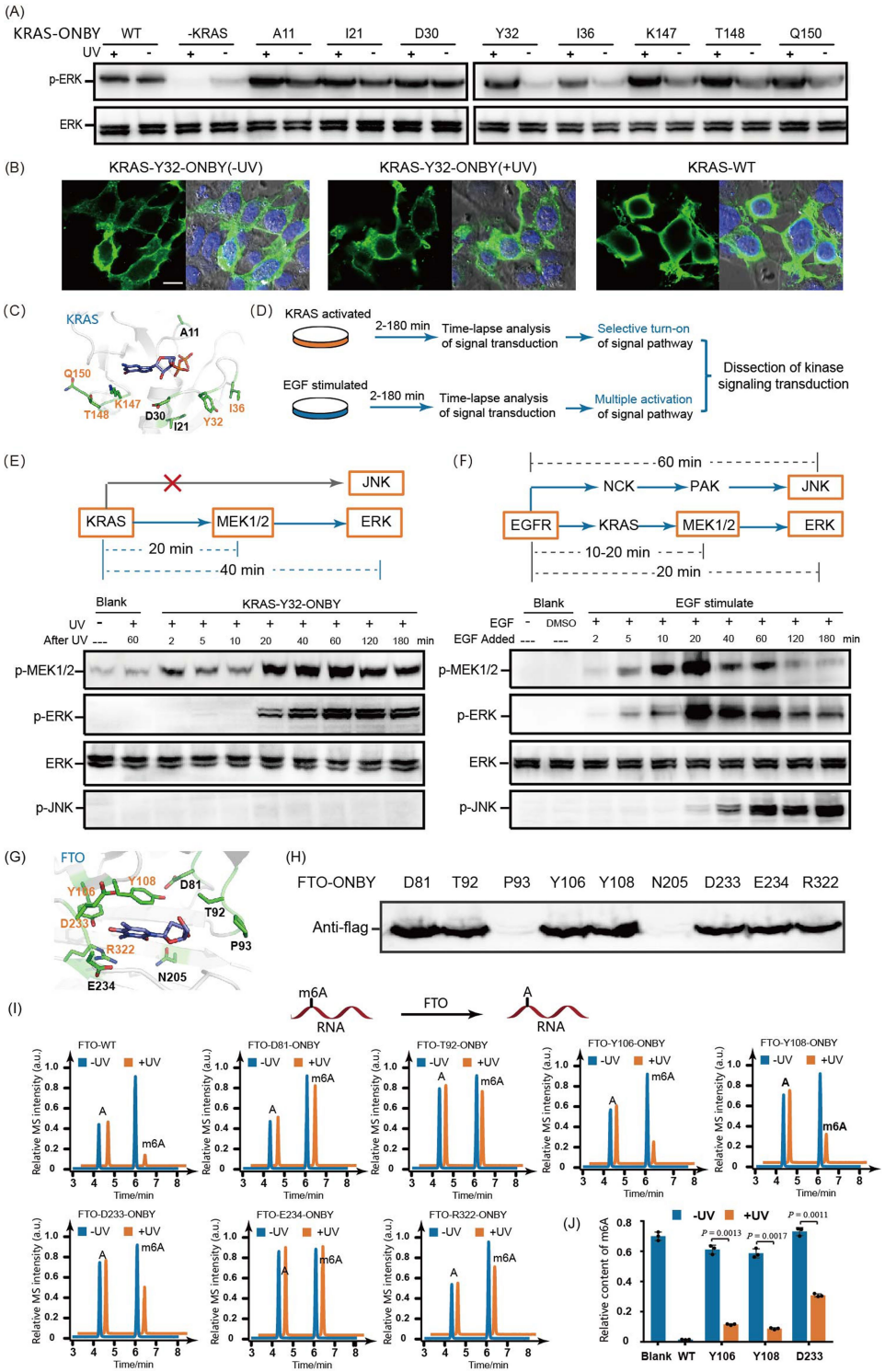
Extended Data Fig. 1 | Computational analysis and validation of Tyr as an optimal anchor residue. a, Surveying the mutational stability of each of the 20 amino acids. The mutation stability of each residue was obtained by the in silico site-saturated mutagenesis of these 3,500 residues and energy calculation of each resulting mutant. Distribution of folding stability for each candidate amino acid is shown in the heat map. The blue and red colours represent the stable and deleterious mutations, respectively. **b**, Schematic of the validation workflow. According to our computational analysis and ranking, Tyr is a suitable anchor residue, whereas the introduced Lys mutations were more likely to affect the enzyme activity. To validate the computational results, a total of 30 randomly selected sites on each of the two model enzymes—FLuc and Nluc—was mutated to either Lys or Tyr, and the activities of the resulting mutants were systematically compared by measuring the resulting luminescence intensity. **c**, **d**, Results from tested mutations on FLuc and NLuc model proteins. Tyrosine mutations showed more stability in general than the lysine mutants. The randomly selected 30 mutation sites of FLuc: N84, F89, L194, N197, T214, R218, N229, H244, H245, F247, F250, Y255, S284, L286, E311, A313, Q338, Y340, I351, E354, V362, L411, L418, I434, R437, L441, Q448, T527, G528 and L530. The randomly selected 30 mutation sites of NLuc: F8, Q14, L20, V23, L24, S31, F33, Q34, P42, Q44, I46, I56, I58, V60, I62, I78, F79, V92, L94, Y96, I109, Y111, F112, Y116, V129, L133, R143, L151, F153 and I157. Results are the representative data of two biological replicates.



Extended Data Fig. 2 | Optimizing the experimental and computational procedures of CAGE-prox. a, Optimization of the UV irradiation time for protein activation in living cells. HEK293T cells expressing FLuc-Y255-ONBY were irradiated by UV light for different times and were further tested for FLuc activity. $n = 3$. **b**, Demonstration of the experimental variation from different batches. The expression level and photo-release efficiency are consistent in different batches of experiments. FLuc was used as an example. Mean ± s.d.; $n = 3$. **c**, Verification of the cell viability under the condition of ONBY incorporation. Mean ± s.d.; $n = 3$. **d**, **e**, The heterogeneity analysis of transient or stable transfection. FLuc-Y255-TAG was used as a model protein for evaluation and cells were stained using an anti-His tag antibody followed by an Alexa-Fluor-488-conjugated anti-mouse IgG. $n = 2$. The high transfection and expression efficiency were observed in HEK293T, HeLa and Jurkat cells (**d**) as well as in ONBY–RS–$tRNA^{pyl}_{CUA}$ pair stable-expressed HEK293T cells expressing FLuc-Y255-TAG (**e**). **f**, Thermodynamic cycle of calculating protein stability (ΔG_f) and binding affinity ($\Delta\Delta G_b$) as introduced by single point mutation. **g**, Equations for calculating protein stability (ΔG_f) and binding affinity ($\Delta\Delta G_b$). The ΔG_f and $\Delta\Delta G_b$ of a given protein are approximated by Rosetta energy scores. **h**, Relative activity of the positive variants of FLuc before and after photo-decaging. Mean ± s.d.; $n = 3$; two-tailed t-test. **i**, Distribution of the distance (Å) and angular direction (degree) from the 60 anchor sites to the FLuc substrate. **j**, Distribution of the mutation stability (ΔG_f) and substrate binding energy change ($\Delta\Delta G_b$) of all the 60 variants calculated by Rosetta. All above-mentioned samples are biological replicates. P values are shown in the figure.

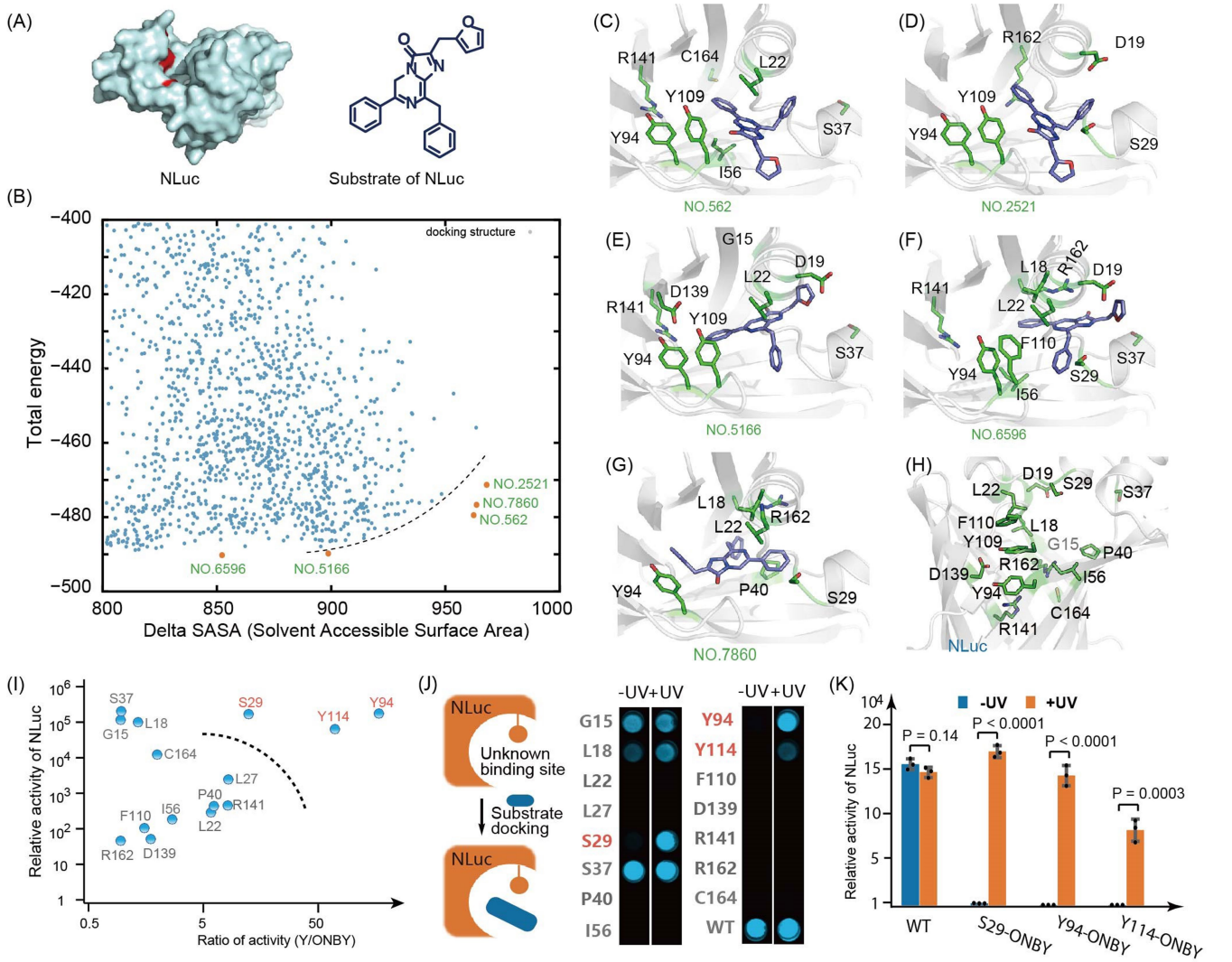


Extended Data Fig. 3 | Detailed workflow of the CAGE-prox strategy. To prepare the initial structure for CAGE-prox calculation, an optimized protein–ligand complex structure was obtained from the PDB, and comparative modelling and/or ligand-docking were used when necessary. Geometry and energy parameters (D, A, ΔG_f and $\Delta\Delta G_b$) were calculated by Rosetta and used as filters to exclude undesired sites. The sites that successfully passed the filters are considered as the 'recommended sites'. In the next experimental validation step, ONBY was incorporated into the recommended sites followed by protein activity characterization with and without photo-activation.

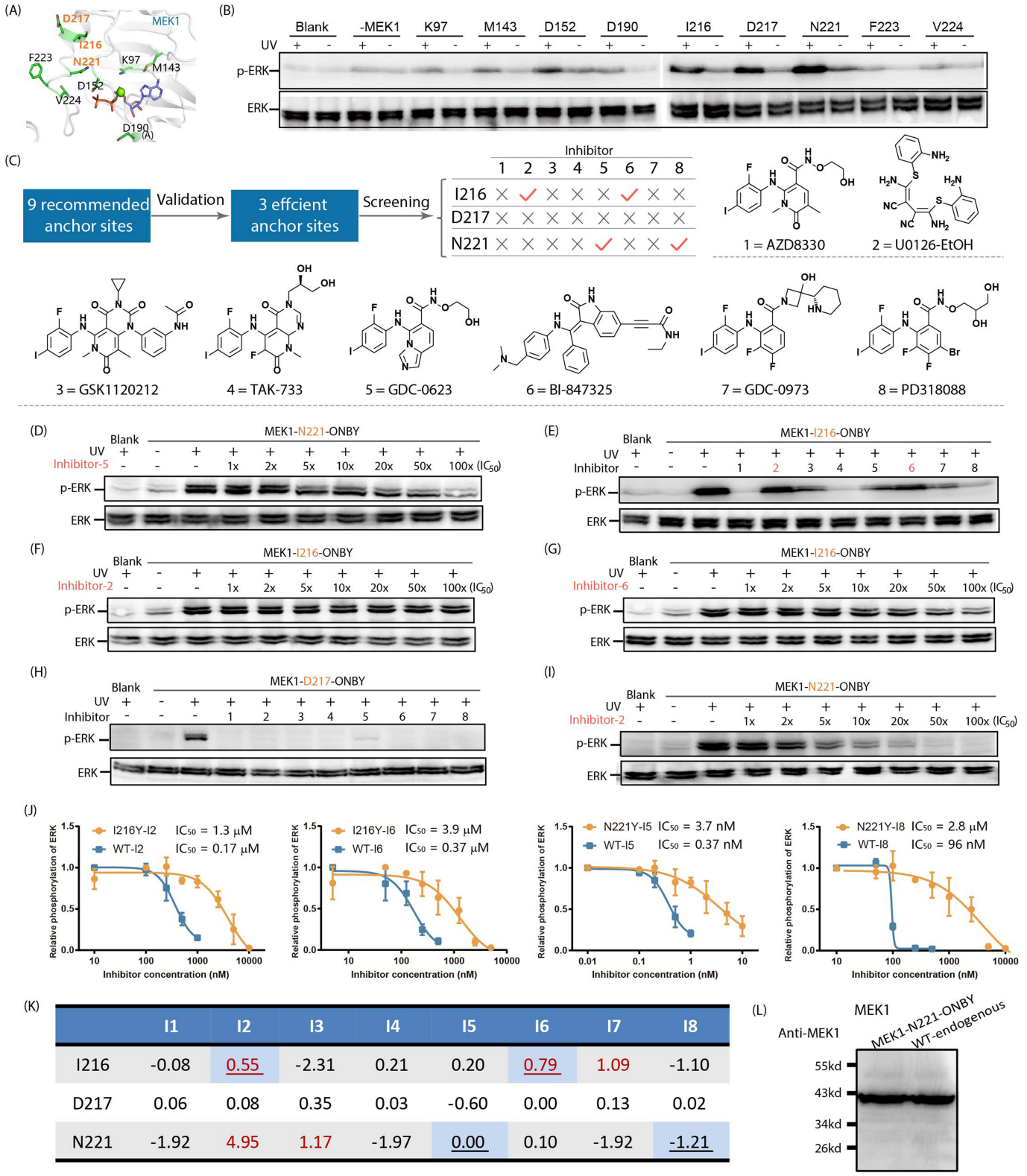


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | CAGE-prox-enabled activation of KRAS and FTO. a, Experimental validation of all the eight recommended anchor sites of KRAS by ERK phosphorylation upon photo-activation. $n = 2$. **b**, Fluorescence imaging of the subcellular localization of wild-type KRAS and KRAS-Y32-ONBY before and after photo-activation, which confirmed correct membrane location for these variants inside cells. Scale bars for all frames, 10 μm. $n = 2$. **c**, Structural view of the eight computationally recommended anchor sites of KRAS by CAGE-prox. The five positive anchor sites are labelled in orange. **d**, Schematic of the workflow for dissecting KRAS-mediated intracellular signal transduction. Both the KRAS activation-triggered signal transduction and EGF stimulation-triggered signal transduction were compared by monitoring the downstream phosphorylation responses. **e**, Upon temporal activation of KRAS, the phosphorylation levels of MEK1, ERK and JNK were analysed by immunoblotting. $n = 2$. **f**, Upon temporal stimulation by EGF, the phosphorylation levels of MEK1, ERK and JNK were analysed and compared by immunoblotting. Distinct patterns of downstream kinase phosphorylation were observed with CAGE-prox-enabled temporal KRAS activation as compared to EGF-stimulated samples. $n = 2$. **g**, Structural view of the nine computationally recommended anchor sites on FTO by CAGE-prox. **h**, Incorporation efficiency of ONBY in each of the recommended anchor sites of FTO as measured by immunoblotting. As the incorporation efficiency at the P93 and N205 sites is extremely low, they were excluded from further experimental validations. $n = 2$. **i**, The rest of seven FTO variants were validated by the LC–MS/MS-based RNA demethylation assay. Four variants (Y106, Y108, D233 and R322) exhibited desired demethylation activity on m^6A upon photo-activation. $n = 3$. **j**, The activities of three positive FTO variants (Y106, Y108, and D233) were measured by a substrate demethylation assay. Mean ± s.d.; $n = 3$; two-tailed *t*-test. All above-mentioned samples are biological replicates. *P* values are shown in the figure.



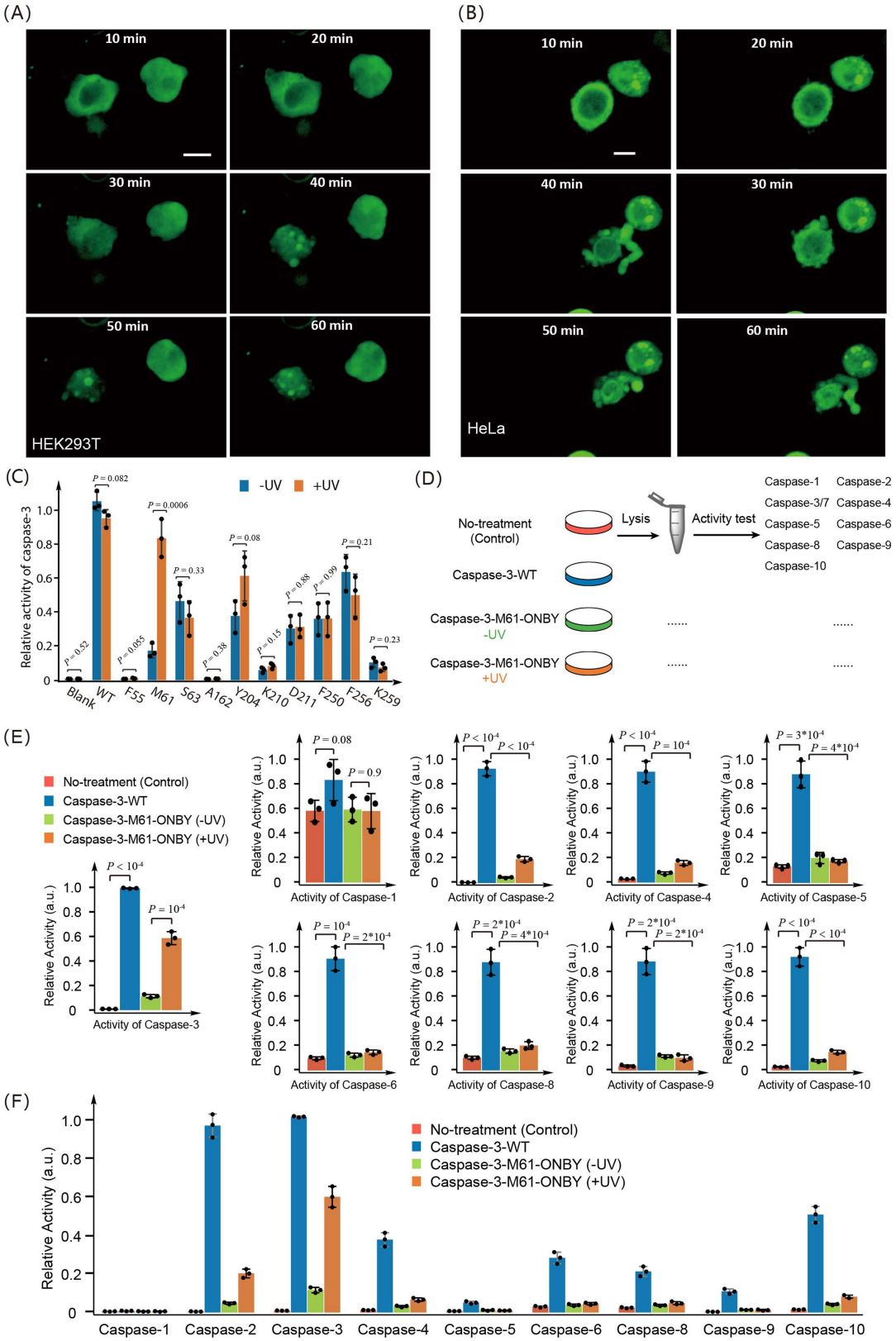
Extended Data Fig. 5 | CAGE-prox-enabled activation of NLuc, an enzyme with undefined active sites. a, Structures of NLuc and its substrate. **b**, Energy distribution of NLuc complex models with the substrate docked. As NLuc does not have a defined active site, protein–ligand docking was performed to generate a total of 10,000 molecular docking models of the enzyme–substrate complexes. The computed parameters were shown in a scatter plot. *x* axis represents the interface area of each docking model by calculating the difference in solvent accessible surface areas between the apo state and the holo state. *y* axis represents the total energy of the docking complex calculated by Rosetta. The five docking models with either the best total energy or the largest difference in solvent-accessible surface area (ΔSASA, plots labelled in orange) were selected for CAGE-prox calculation as potential anchor residues. **c**–**g**, Structural view of the five selected docking model with the substrate shown in purple sticks and the recommended CAGE-prox anchor residues in green sticks. **h**, Structural view of all the 15 recommended anchor sites (in green sticks) shown on the apo structure of NLuc. PDB code 5B0U. **i**, Experimental validation of all the 15 recommended CAGE-prox variants of NLuc. Three variants (marked in orange labels) showed strong luciferase activity (*x* axis) and high activation ratio (*y* axis) after photo-decaging. **j**, Validation of the recommended anchor sites on NLuc. Three out of fifteen recommended sites (S29, Y94 and Y114) were efficiently blocked by ONBY incorporation and rapidly rescued after photo-decaging. $n = 2$. **k**, The luciferase activity of three positive NLuc variants before and after photo-decaging. Mean ± s.d.; $n = 3$; two-tailed *t*-test. All above-mentioned samples are biological replicates. *P* values are shown in the figure.



	I1	I2	I3	I4	I5	I6	I7	I8
I216	-0.08	0.55	-2.31	0.21	0.20	0.79	1.09	-1.10
D217	0.06	0.08	0.35	0.03	-0.60	0.00	0.13	0.02
N221	-1.92	4.95	1.17	-1.97	0.00	0.10	-1.92	-1.21

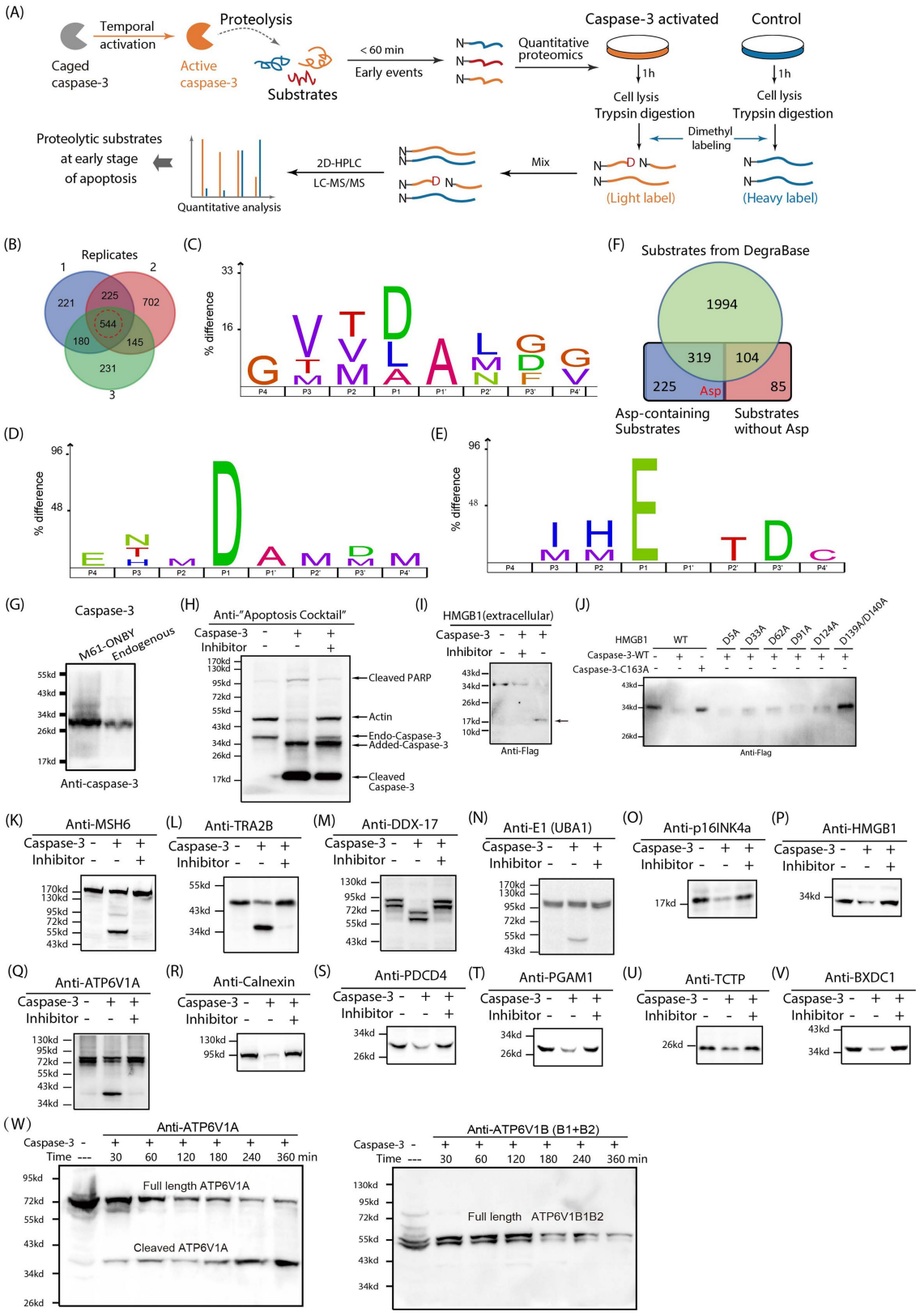
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Orthogonal activation of the inhibitor-resistant MEK1 variants. a, Structural view of the nine recommended anchor sites of MEK1 by CAGE-prox (PDB code 1S9I). The three experimentally validated positive variants are labelled in orange. **b**, Experimental validation of the nine recommended MEK1 variants by monitoring downstream ERK phosphorylation. Three positive variants (I216, D217 and N221) were identified as efficient anchor sites. $n = 2$. **c**, The screening strategy for identifying the inhibitor-resistance MEK1 mutant. The three identified CAGE-prox-activatable MEK1 variants were inspected which yielded four enzyme variant–inhibitor 'resistant pairs'. **d**, Dose-dependent inhibition of the MEK1(N221)-ONBY variant showed that this variant is not resistant to inhibitor 5, with observable inhibition starting only at the 5× IC_{50} concentration. $n = 2$. **e**, Inhibitors screening of the MEK1(I216Y) variant identified two 'resistant' pairs. $n = 2$. **f**, **g**, Dose-dependent inhibition of the MEK1(I216Y) variant by inhibitor 2 (**f**) and inhibitor 6 (**g**). $n = 3$. **h**, Inhibitor screening on the MEK1(D217Y) variant yielded no 'resistant' pairs. $n = 2$. **i**, Dose-dependent inhibition the MEK1(N221Y) variant by inhibitor 2. $n = 2$. **j**, IC_{50} evaluation of all the four MEK1 mutant–inhibitor pairs in comparison with that of the wild-type MEK1–inhibitor pair. Mean ± s.d.; $n = 3$. **k**, The predicted binding affinities among the three mutants and eight inhibitors. The scores represent the relative binding energy of each inhibitor–mutant complex: the higher the score, the weaker the binding. A total of five inhibitor–mutant pairs (score > 0.5, coloured in red) were calculated and predicted as 'resistant pairs'. The four resistant pairs that were validated experimentally were shown in light blue background with the numbers underlined. **l**, The expression level of ONBY-incorporated MEK1 (MEK1(N221)-ONBY) and the corresponding endogenous MEK1 protein. $n = 2$. All above-mentioned samples are biological replicates.



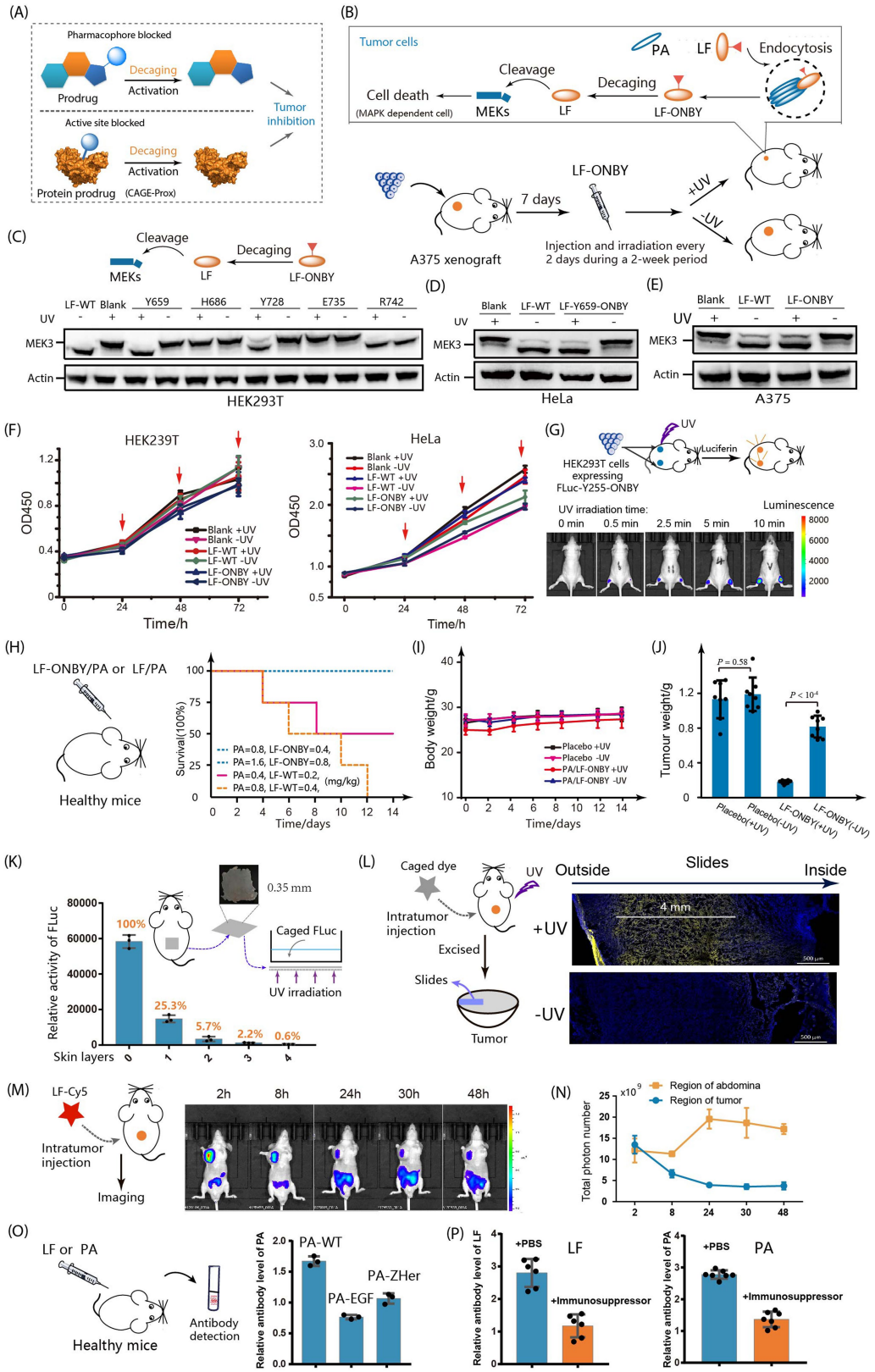
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Cell morphological changes and activity measurement of caspase family members after photo-activation of caspase-3. a, Apoptosis of HEK293T cells after photo–induced caspase-3 activation. Scale bars for all frames, 10 μm. $n = 2$. **b**, Apoptosis of HeLa cells after caspase-3 activation. Scale bars of all frames, 5 μm. In both cells, morphological changes upon apoptosis were observed as early as 30 min after photo-activation of caspase-3. $n = 2$. **c**, Experimental validation revealed that incorporation of ONBY at M61 blocked caspase-3 activity, which can be efficiently rescued after photo-decaging. Mean ± s.d.; $n = 3$; two-tailed *t*-test. **d**, Schematic of the experimental design showing that the activities of nine caspase family members were measured in HEK293T cells with either overexpression or temporal photo-activation of caspase-3. **e**, Normalized activity of each caspase from control cells (red), cells with transient overexpression of wild-type (WT) caspase-3 (blue), cells with overexpression of caspase-3-M61-ONBY before (green) and after (orange) photo-activation. Overexpression of the wild-type caspase-3 resulted in markedly increased activities for most of the other caspases, whereas direct photo-activation of the caspase-3 variant enabled by CAGE-prox only activated this specific enzyme without much influence on other caspases. It allows the temporal profiling of proteolytic substrates immediately after caspase-3 activation. Error bars represent mean ± s.d.; $n = 3$; two-tailed *t*-test. **f**, Replotting of **d** with activities of all caspase members shown in the same scale. Error bars represent mean ± s.d.; $n = 3$. All above-mentioned samples are biological replicates. *P* values are shown in the figure.



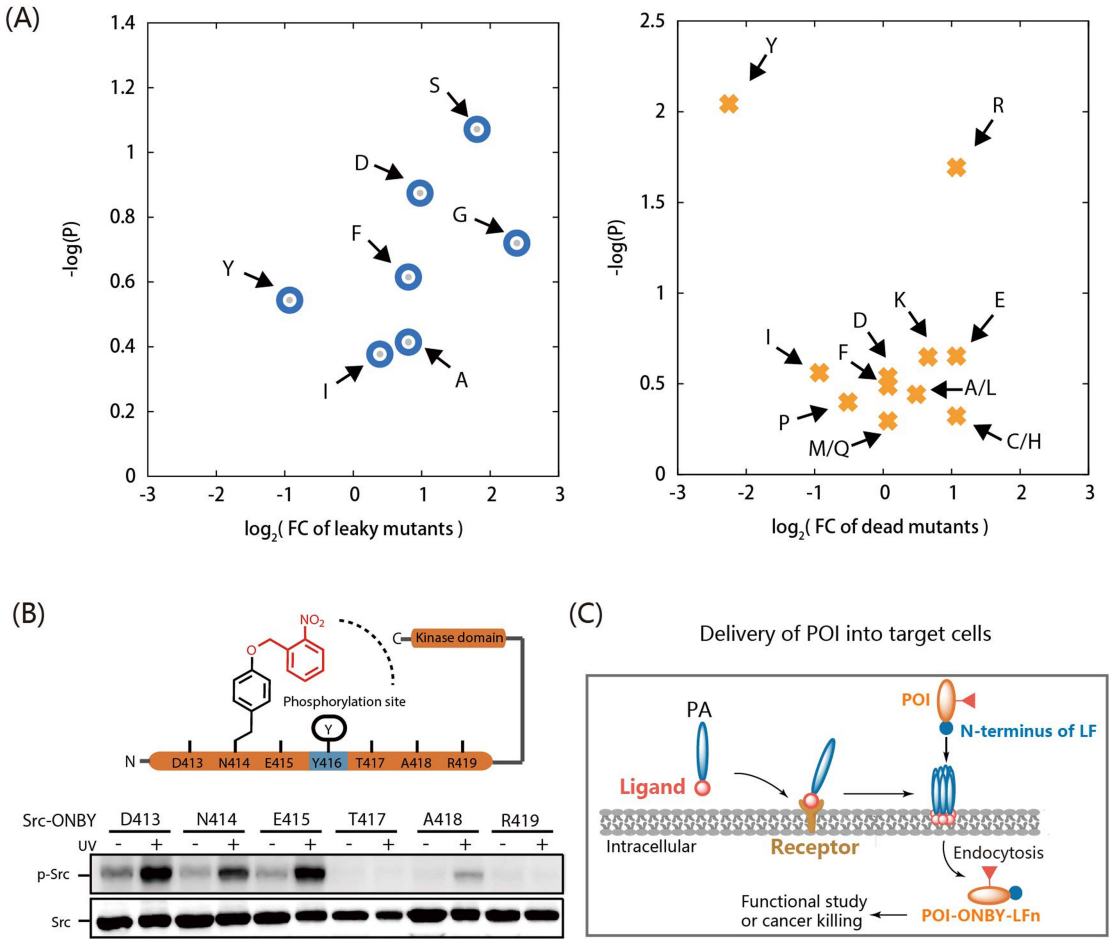
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Analysis and verification of the identified proteolytic substrates upon caspase-3 activation. a, Schematic of the workflow for CAGE-prox-enabled temporal profiling of the proteolytic substrates immediately after caspase-3 activation. **b**, Venn diagram showing a summary of numbers of the proteolytic substrates identified from three independent temporal profiling experiments. A total of 544 proteins was commonly identified from all the three experiments. **c**–**e**, Sequence analysis of the cleavage sites in the identified proteolytic substrates. The logos were generated using Icelogo (https://iomics.ugent.be/icelogoserver/), with all the identified cleavage sequences aligned (the cleavage site is between the P1 and P1′ position). The sequence logos of cleavage sites at all 20 amino acids (526), Asp only (79) and Glu only (29) are shown in **c**, **d** and **e**, respectively. Although Asp is the most abundant cleavage site, the background of other non-caspase sites is relatively high in our results, owing to the lack of N terminus enrichment. The typical DEVD/E motifs were not observed directly, but the amino acid pattern at the P1′ position is similar to that found in previous studies[41]. **f**, Venn diagram comparing the substrates identified from this work with the proteolytic sites recorded in DegraBase (https://wellslab.ucsf.edu/degrabase/). A total of 773 overlapping substrates was found, including 544 substrates that were identified with Asp-containing peptides. **g**, The expression level of ONBY-incorporated caspase-3 (M61–ONBY) and the corresponding endogenous caspase-3. $n = 2$. **h**, Verification of the in vitro cleavage assay using actin, caspase-3 and PARP as positive controls. $n = 2$. **i**, Verification of the cleaved and secreted HMGB1 in the culture medium. HMGB1–Flag and wild-type caspase-3 were co-transfected into HEK293T cells, and the cell culture medium were concentrated before immunoblotting analysis. $n = 2$. **j**, Alanine screening mapped the caspase-3's cleavage site on HMGB1 as D139/D140. $n = 2$. **k**–**v**, Verification of the newly identified proteolytic substrates of caspase-3. Recombinant caspase-3 protein was added into cell lysate of HEK293T followed by immunoblotting analysis. $n = 2$. **w**, Comparison of the cleavage kinetics of ATP6V1A versus ATP6V1B by caspase-3. After the recombinant wild-type caspase-3 was added into the cell lysate, the mixture was incubated at 37 °C for 0, 0.5, 1, 2, 3, 4 and 6 h, followed by immunoblotting analysis. $n = 2$. All above-mentioned samples are biological replicates.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | CAGE-prox-enabled protein activation as a pro-drug therapy in living mice. a, The traditional pro-drug strategy on small-molecule drugs can be extended to proteins by adopting the CAGE-prox-enabled in situ activation of therapeutic proteins. **b**, Schematic of the CAGE-prox-activated LF as a potential pro-drug therapy. MEK1-dependent A375 human melanoma cells were planted into mice as a xenograft model; activated LF will cause the death of MAPK-dependent tumour cells by rapid cleavage on MEK kinases. **c**, Experimental validation of the CAGE-prox-activatable LF variants, with the activity evaluated by cleavage efficiency on MEK3 in HEK293T cells. Incorporation of ONBY at Y659 or Y728 blocked LF's activity, which can be efficiently rescued by photo-decaging. $n = 2$. **d**, **e**, Delivery and activation of LF(Y659)-ONBY was demonstrated in HeLa cells and HEK293T cells. LF(Y659)-ONBY was expressed and purified in *Esherichia coli* and delivered into the target cell lines by PA. $n = 2$. **f**, The growth curves of HEK293T or HeLa cells treated with LF(Y659)-ONBY/PA. Activated LF had negligible influence on the proliferation of MEK-independent HEK293T and HeLa cells. The red arrows represent the time point at which the medium was exchanged and LF–PA treatments were performed. Mean ± s.d.; $n = 3$. **g**, Photo-activation of the CAGE-prox variant of FLuc in BALB/c nude mice and optimization of the UV irradiation protocol. Increased FLuc activity were observed after UV irradiation (bottom image). $n = 2$. **h**, The caged LF-ONBY was safer than the wild-type LF (LF-WT) as determined by a dose escalation protocol. Intraperitoneal injection of wild-type LF into healthy mice every 2 days for a 2-week period caused about 50% of animal death. By contrast, injection of LF(Y659)-ONBY with a fourfold-higher dosage during the same period had negligible adverse effects on mice ($n = 4$). **i**, Photo-activation of the LF-ONBY variant has negligible influence on mouse body weight in a two-week period treatment. Mean ± s.d.; $n = 8$. **j**, Photo-activation of the LF-ONBY variant markedly reduced the tumour weight in the xenograft model. Mean ± s.d.; $n = 8$; two-tailed *t*-test. **k**, Evaluation of UV light penetration in vitro. UV light was irradiated through the skin and re-activation of FLuc was used to evaluate the penetration ability. Mean ± s.d.; $n = 3$. **l**, Evaluation of UV light penetration in vivo. A caged fluorescent dye was intra-tumourally injected into mice followed by UV irradiation for 5 min. The tumour was then excised and the depth of light penetration was examined. $n = 2$. **m**, **n**, Bio-distribution of the injected caged LF in mice. The Cy5-labelled LF protein was injected into mice, followed by imaging the whole mouse body. Mean ± s.d.; $n = 3$. **o**, The immunogenicity of LF–PA can be reduced by fusing cell-surface-targeting elements. EGF, epidermal growth factor that can target its cell-surface receptor EGFR; ZHer, an affibody that can target its receptor HER2. Mean ± s.d.; $n = 3$. **p**, The immunogenicity of LF–PA can be reduced by the addition of immunosuppressors (pentostatin combined with cyclophosphamide). Mean ± s.d.; $n = 3$. All above-mentioned samples are biological replicates. *P* values are shown in the figure.



Extended Data Fig. 10 | A retrospective analysis of 'failed' CAGE-prox predictions and other potential expanded applications of CAGE-prox. **a**, The CAGE-prox predicted mutations that have been experimentally evaluated as failed can be classified into two categories: 'leaky' (fail to block the protein activity with the inserted caged ONBY) and 'dead' (fail to restore the protein activity after photo-decaging). We calculated the frequency fold change (FC) of each native amino acid in each category. The fold change was defined as the frequency of a residue in a certain category/frequency in all predicted residues. *x* axis represents the fold change (expressed in $\log_2$) and *y* axis represents the log(*P* value). As shown in the plot, ONBY insertions at native Gly and Ser positions are more likely to result in leaky mutants (left) whereas insertion at native Arg positions is more likely to result in dead mutants (right). As expected, a native Tyr position is less likely to fail. A total of 56 anchor residues from 7 different proteins was analysed using hypergeometric test. **b**, CAGE-prox-enabled control of auto-phosphorylation of Src kinase. HEK293T cells were co-transfected with Src-TAG mutants and the ONBY–RS–$tRNA^{pyl}_{CUA}$ pair, and cultured for 24 h in the presence of ONBY. After UV-triggered decaging of the Src-OBNY variants, cells were cultured at 37 °C for another 3 h before the auto-phosphorylation level of each Src mutant was detected by immunoblotting. $n = 2$. **c**, Cell-specific targeting of a POI by CAGE-prox by adding the cancer-cell-targeting ligand to PA and the N-terminal domain of LF (LFn) to a POI, respectively. All above-mentioned samples are biological replicates.

Corresponding author(s): pengchen@pku.edu.cn, chuwang@pku.edu.cn

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Robetta Server (http://robetta.bakerlab.org/) for homology modeling. Rosetta Software (v3.6) for mutant energy calculation and ligand docking. ACPYPE and Ambertools (v16) for generating ligand atom parameters and charge. AutoDock Vina (v1.1.2) for generating initial docking configuration.
Data analysis	Pymol and Avogadro for showing and editing protein structures. Gnuplot for ploting data. GraphPad Prism 5.0 and Microsoft Excel for statistics. Custom python code for data analysis (https://github.com/wendao/CAGE-Prox). ProLuCID (v1.3.3) for peptide identification. DTASelect (v2.0.47) for protein inference. CIMAGE (source code available upon request) for protein quantification.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All relevant data are included in the manuscript, source data or supplementary information. More details are available from the corresponding author upon request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine the sample sizes. As for the western blotting experiment or the experiments that do not need statistic analysis repeated twice, for the experiments need that need statistic analysis repeated three times, for the animal experiments, we chose sample size as 6-8 just in case that the mice were dead during the experiments.
Data exclusions	There were no data exclusions.
Replication	All experiments were confirmed with multiple biological replicates as indicated in the Figure legends, and the representative results are shown.
Randomization	Tumor bearing mice were randomly selected before treatment.
Blinding	Investigators are blinded when grouping tumor bearing mice, Investigators for experiments and Investigators for mass-spec analysis are blinded with each other Experiments with internal controls are not blinded.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☐	☒ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Unique biological materials

Policy information about availability of materials

Obtaining unique materials	Chemical reagents: All chemical reagents are in analytical grade, obtained from commercial suppliers, and used without further purification. Others are listed as follows: Inhibitors of MEK1 (1-8: AZD8330, U0126-EtOH, GSK1120212, TAK-733, GDC-0623, BI-847325, GDC-0973 and PD318088) are purchased from Selleck. Heavy formaldehyde (D13CHD, 596388) and light formaldehyde (HCHO, F8775) were purchased from Sigma.

Enzymes and protein reagents:
Enzymes for subcloning and site-specific mutagenesis were ordered from New England Biolabs (NEB). Primers were ordered from Sangon Biotech (Shanghai) Co. Ltd. Others are listed as follows: recombinant human EGF (Promega, G5021), recombinant caspase-3 (Sino Biological, 10050-H08E), nuclease P1 (Sigma, N8630), alkaline phosphatase (Sigma, P4252), Trypsin Gold and rLys-C mixture (Promega, V5072)

Assay kits:
Firefly luciferase assay kit (Vigorous biotechnology, T003), Firefly & Renilla Dual-Luciferase Assay (Dual-Glo® Luciferase Assay System, Promega, E2920), NanoLuc assay kit (Nano-Glo® Luciferase Assay System, promega, N1110), Caspase-3 Assay Kit (Abcam, ab39388), Caspase Family Fluorometric Substrate Kit II Plus (Abcam, ab102487), Protein quantitation assay kit (Pierce™ BCA Protein Assay Reagent A, 23223). Cell Proliferation Assay Kit (CellTiter 96 Aqueous One Solution Reagent, Promega, G3582), IL-8 ELISA Kit (R&D Systems, D8000C), cell lysis buffer (ThermoFisher, 78503)

Other materials:
Flag beads (Pierce™ Anti-DYKDDDDK Magnetic Agarose, A36797), flag peptide (Pierce™ 3x DYKDDDDK Peptide, A36806), protein desalting column （Bio-Spin®, BIO-RAD, 7326227） Peptide Desalting Columns (Pierce™ Peptide Desalting Spin Columns, 89852). Transfection reagent (XtremeGENE HP, Roche, 6366236001). DMEM, fetal bovine serum and OPTI-MEM were purchased from Thermo Fisher Scientific (gibco™), HEK293T, HeLa and A375 cell lines were obtain from ATCC. RNA oligonucleotides was obtain from Prof. Guifang Jia's group as a gift, which were synthesized by Expedite 8909 DNA/RNA synthesizer and purified by RNA purification cartrige (Glen Research). 6 to 8-week-old male Balb/c nude mice, 6 to 8-week-old male C57BL/6 mice and 3 to 4-week-old male Nu/Nu nude mice, were all purchased from Vital River Laboratories. Database search was performed on an in-house Mascot server (Matrix Science Ltd., London, UK).

Antibodies

Antibodies used

All the antibody used in this study are commercially available :
His tag (Abcam, ab18184)，MEK1(Abcam, ab32091), Caspase-3(Abcam, ab32351), pERK (CST, 9101)，pMEK (CST, 9154)，ERK (CST, 9102)，Actin (Abcam, 124964)，p-JNK (CST, 9251)，Apoptosis Western Blot Cocktail (Abcam, ab136812), DDX-17 (Abcam, ab180190), ATP6V1A (Abcam, ab199325), ATP6V1B1+ATP6V1B2 (Abcam, ab200839), TRA2B (abcam, ab171082), MSH6 (abcam, ab92471), E1-Ubiquitin Activating Enzyme, UBA1 (abcam, ab181225), CDKN2A/p16INK4a (abcam, ab108349), HMGB1 (abcam, ab79823), Calnexin (Santa Cruz Biotechnology, sc-46669)，PDCD4 (abcam, ab79405), PGAM1 (abcam, ab129191), TCTP (abcam, ab133568), BXDC1 (abcam, ab180604), , MEK3 (C-15, Santa Cruz Biotechnology, sc-961)，pSrc-Y416 (CST, 6943), Alexa Fluor 488-conjugated anti-mouse IgG (abcam, ab150113)

Validation

Validation of all antibodies can be found on the manufacturer's website, all the validation results can be founded on the websites by search the catalog numbe. Website of abcam: https://www.abcam.com/, Website of CST: https://www.cellsignal.com Website of Santa Cruz Biotechnology: https://www.scbt.com

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

A375, HEK293T, HeLa and Jurkat cells were obtained from the American Type Culture Collection (ATCC).

Authentication

The cell lines were frequently checked by their morphological features and the cell lines are not been authenticated by the short tandem repeat (STR) profiling.

Mycoplasma contamination

All cell lines were tested to be mycoplasma-negative by the standard PCR method

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines are used in this study.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

6 to 8-week-old male Balb/c nude mice and male C57BL/6 mice (purchased from Vital River Laboratories, China.) were used in this study.

Wild animals

No wild animals are used in this study.

Field-collected samples

The study did not involve samples collected from the field