



De novo design of tyrosine and serine kinase-driven protein switches

Nicholas B. Woodall^{1,2,3}, Zara Weinberg⁴, Jesslyn Park⁴, Florian Busch^{5,6}, Richard S. Johnson⁷, Mikayla J. Feldbauer³, Michael Murphy³, Maggie Ahlrichs³, Issa Yousif^{1,8}, Michael J. MacCoss⁷, Vicki H. Wysocki^{5,6}, Hana El-Samad⁴ and David Baker^{1,2,3}✉

Kinases play central roles in signaling cascades, relaying information from the outside to the inside of mammalian cells. De novo designed protein switches capable of interfacing with tyrosine kinase signaling pathways would open new avenues for controlling cellular behavior, but, so far, no such systems have been described. Here we describe the de novo design of two classes of protein switch that link phosphorylation by tyrosine and serine kinases to protein-protein association. In the first class, protein-protein association is required for phosphorylation by the kinase, while in the second class, kinase activity drives protein-protein association. We design systems that couple protein binding to kinase activity on the immunoreceptor tyrosine-based activation motif central to T-cell signaling, and kinase activity to reconstitution of green fluorescent protein fluorescence from fragments and the inhibition of the protease calpain. The designed switches are reversible and function in vitro and in cells with up to 40-fold activation of switching by phosphorylation.

Although there has been progress in the de novo design of stable proteins with a range of folds and functions, the de novo design of proteins that modulate or are modulated by tyrosine phosphorylation has not, to our knowledge, been described. Native protein motifs such as EF-hands and coiled coils, in which aspartate or glutamate residues play critical roles, have been made responsive to serine phosphorylation by taking advantage of the approximate mimicry of these residues by phospho-serine. For example, a serine substitution in an EF-hand eliminates calcium binding, and binding activity can be restored by phosphorylation of the same serine¹. Similarly, phosphorylation of serine near the positive N-terminal end of an α -helix is electrostatically stabilizing and has been used to engineer switching². Phosphorylation control has also been achieved by introducing phosphorylation sites at naturally occurring interfaces: phosphorylation at such sites can reduce binding by introducing unfavorable steric and electrostatic interactions³. Conversely, serine protein kinase A (PKA) phosphorylation sites designed into a coiled-coil along the dimer interface in positions normally occupied by glutamate lead to an increase in stability upon phosphorylation due to electrostatic interactions with arginine residues placed across the interface^{2,4}. However, the glutamate approximate mimicry approach does not work for tyrosine-based phosphorylation, as no natural amino acid has a shape and charge distribution similar to phospho-tyrosine (pTyr). Tyrosine kinases are of particular interest for mammalian cell control, as they are involved in a large number of signaling pathways relaying information from the outside to the inside of cells⁵.

We set out to design proteins that modulate or respond to phosphorylation of tyrosine by tyrosine kinases, as outlined in Fig. 1a. In the first ‘binding-dependent phosphorylation’ scheme, a de novo designed protein has a kinase target site that is cryptic (cannot be

accessed by the kinase) until addition of a designed co-regulator. From the perspective of engineered cell control, this would enable the incorporation of additional protein control elements regulating the extent of activation of a tyrosine kinase signaling pathway. For example, T-cell activation begins with the phosphorylation of the immunoreceptor tyrosine-based activation motif (ITAM), which then binds to and activates the ZAP-70 kinase⁶; controlling the accessibility of the ITAM motif via the binding of co-regulator would hence control the extent of T-cell activation. In the second ‘phosphorylation-dependent binding’ scheme, a binding interface for a target of interest is hidden in the absence of kinase, and becomes exposed upon kinase-catalyzed phosphorylation. Such switches could allow for the incorporation of new signaling pathway branches downstream of activated kinases.

Design of binding-induced phosphorylation

We began by exploring the de novo design of protein systems implementing binding-dependent phosphorylation. We chose to focus on the ITAM central to T-cell signaling. The ITAM contains a dual-tyrosine motif that, when phosphorylated, initiates T-cell activation⁶. To make phosphorylation of the ITAM tyrosines dependent on a regulatory protein in trans, we aimed to design a system in which competition between inter- and intramolecular interactions drives a conformational switch, as in the recently described latching orthogonal cage–key proteins (LOCKR) switches^{7,8}. We first sought to design a four-helix bundle in which the C-terminal tyrosine in the ITAM sequence motif forms part of the central core and hence is not accessible, in the folded state, to the kinase. There is a considerable solvation-free energy cost to burying the polar hydroxyl groups on the tyrosines without making hydrogen bonds to other residues in the core, and hence we focused on designed proteins

¹Department of Biochemistry, University of Washington, Seattle, WA, USA. ²Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. ³Institute for Protein Design, University of Washington, Seattle, WA, USA. ⁴Department of Biochemistry & Biophysics, University of California San Francisco, San Francisco, CA, USA. ⁵Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH, USA. ⁶Resource for Native Mass Spectrometry Guided Structural Biology, The Ohio State University, Columbus, OH, USA. ⁷Department of Genome Sciences, University of Washington, Seattle, WA, USA. ⁸Department of Bioengineering, University of Washington, Seattle, WA, USA. ✉e-mail: dabaker@uw.edu

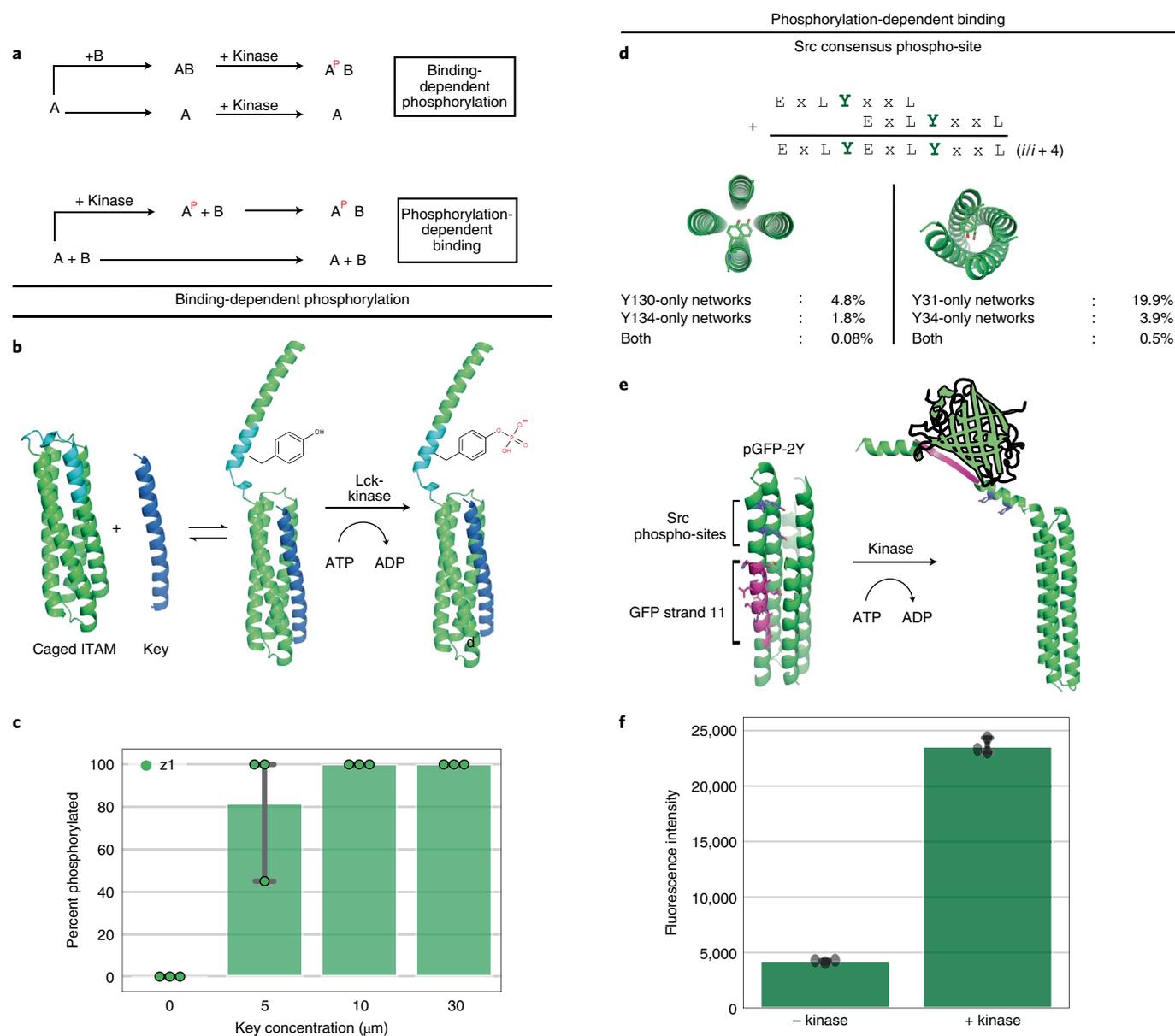


Fig. 1 | Phosphorylation switch design. a, Schematic of the design mechanisms that use binding to control a phosphorylation event (top) or phosphorylation to control a binding event (bottom). ‘A’ represents the protein switch that can be phosphorylated. ‘B’ represents a protein that can bind ‘A’. The superscript ‘p’ represents a phosphorylation event on ‘A’. **b, c**, Binding-dependent phosphorylation. Design model of the caged ζ ITAM in a DHD-1:234A_B (PDB 6DLC) based LOCKR and model mechanism for phosphorylation controlled by binding (**b**). The cage (green), ITAM threading (cyan) and key peptide (blue) are colored in the structures. **c**, Key-induced phosphorylation of the ζ ITAM motif measured by intact-protein mass spectrometry (**c**). The N-terminal tyrosine in the ζ ITAM motif is mutated to glutamate to solely follow the phosphorylation of the caged C-terminal tyrosine. Data are presented as mean values ± standard deviation from three independent experiments. **d–f**, Phosphorylation-dependent binding. Design pipeline for incorporating two-stacked hydrogen-bond networks to contain the Src-phosphorylation site (**d**). The combined phosphorylation site is shown for the (i/i + 4) arrangement used in the final switch. The percentages of backbones generated containing hydrogen-bond networks including the indicated residues, for both straight and left-handed helices, are shown. Schematic of split-GFP activation by pGFP-2Y by phosphorylation by the Src kinase (**e**). On the left is the design model for pGFP-2Y, showing the GFP11 peptide in magenta. Phosphorylation-site tyrosine residues are shown in blue. GFP fluorescence of pGFP-2Y, GFP1-10 and ATP increases with the addition of Src kinase (**f**). Data are presented as mean values ± standard deviation from three independent experiments.

in which one of the ITAM tyrosines could make hydrogen bonds. We took advantage of a series of de novo designed heterodimers containing extensive buried hydrogen-bond networks, and found that the C-terminal ζ ITAM tyrosine could be aligned onto a tyrosine, Y123, in a hydrogen-bond network in the designed heterodimer DHD-1:234A_B (PDB 6DLC; Fig. 1b)⁹. To construct a

single-chain-based switch, we connected the DHD-1:234A_B heterodimer into a four-helix bundle monomer with the embedded ITAM motif.

To drive exposure of the buried ITAM motif so that it could be phosphorylated by a tyrosine kinase, we explored switch designs analogous to LOCKR in which intermolecular interactions with a

helical ‘key’ are stronger than intramolecular interactions with a helical ‘latch’, which can hence be displaced by the key. Caging phosphorylated motifs like the ITAM in the four-helix bundle creates an additional challenge when designing the key peptide. The ideal residue choice for complementing the tyrosine hydrogen-bond network in DHD-1:234A_B by the key peptide would be a tyrosine; however, if the key peptide tyrosine becomes phosphorylated, the steric and electronic clash of the pTyr would preclude binding. To avoid inactivation of the key by phosphorylation, we replaced the tyrosine with a histidine residue that can also form a hydrogen bond across the interface.

To test the designed binding-induced phosphorylation system, we obtained synthetic genes encoding the caged ITAM motif and the key, expressed the two proteins in *Escherichia coli*, and purified them by Ni-affinity chromatography. We then tested the ability of Lck, the native T-cell kinase, to phosphorylate the C-terminal tyrosine in the ITAM motif in the presence or absence of the key peptide. Without the key peptide, no phosphorylation was detected by intact-protein mass spectrometry: the ITAM motif as designed is caged by the helical bundle. On addition of key peptide, the ITAM motif was readily phosphorylated by Lck. Use of 5 μ M key activated phosphorylation of 50 μ M of the design (which we call CD3 ζ 1 ITAM LOCKR), indicating multiple turnover activation, which could be useful for cellular control applications (Fig. 1c). Because phosphorylation is irreversible under our experimental conditions, the key can disassociate from one switch molecule and then bind to and activate another, and hence may be viewed as a multiple turnover catalyst. Once phosphorylated, the ITAM motif binds with high affinity to ZAP-70 to signal T-cell activation.

Design of phosphorylation-induced binding

We next set out to use de novo design to create a protein system in which protein-protein association is driven by a tyrosine kinase-catalyzed tyrosine phosphorylation. We sought to design a helical bundle system containing a ‘latch’ helix with multiple tyrosine phosphorylation sites that regulate the accessibility of a binding peptide motif that is inactive when sequestered within the bundle. In this design scheme, phosphorylation of one or more tyrosine residues causes the latch helix to disassociate from the bundle, allowing the binding peptide to associate with its target. We chose to design systems containing multiple tyrosines to be phosphorylated rather than a single tyrosine to obtain a greater thermodynamic driving force for conformational change following kinase activation. We found that two Src-family consensus sites, EXLYXXL, interwoven onto helix positions $i/i+4$ for a straight helix or $i/i+3$ for a left-handed helix, provided a particularly compact arrangement of tyrosine sites¹⁰. Compactness has the advantage of enabling the fitting of more phosphorylation sites in a single designed system, and leaves more flexibility in the bundle to thread any bioactive motif of choice without disrupting the phosphorylation site networks.

Aiming for switches with smaller total size, to facilitate incorporation into cell signaling studies in future work, we focused on designing four-helix-bundle (rather than the six-helix-bundle original LOCKR⁷) systems with two tyrosine-containing hydrogen-bond networks involving a terminal latch helix containing a bioactive binding peptide. We found that straight and left-hand-twisted helical topologies fit the limited rotameric preferences of tyrosine residues, with both residues pointing towards the core, with room for forming hydrogen-bond networks (Fig. 1d)—straight helices are compatible with an $i/i+4$ arrangement of tyrosine residues, while left-handed helices fit an $i/i+3$ arrangement. We parametrically generated hundreds of thousands of bundles, varying the super-helical radius, super-helical twist, chain axial offset and helical phase, and used Monte Carlo sampling with the HBNet mover to identify tyrosine-containing hydrogen-bonding networks¹¹.

We found that only large bundles with a radius greater than 6.75 Å could incorporate both tyrosine residues into hydrogen-bond networks. For the left-hand super-twisted topologies, we parametrically generated 100,000 bundles and identified 23,917 hydrogen-bond networks; 519 of the models contained hydrogen-bond networks with both tyrosines. For straight helical topologies, we parametrically generated 1,000,000 bundles, resulting in 68,095 hydrogen-bond networks with 819 networks involving both tyrosine residues. Structures without both tyrosine residues participating in a hydrogen-bond network were discarded, and combinatorial sequence optimization was used to minimize the energy of the remainder of the structure using the Rosetta energy function, which includes terms representing van der Waals packing, hydrogen bonding, electrostatics and solvation. The resulting designs were filtered based on Rosetta energy, secondary-structure side chain complementary, cavities within the core of the protein, and buried unsatisfied polar residues; polar residues without hydrogen bonds in the core of the protein are very energetically unfavorable due to the high de-solvation energies and are therefore important to avoid (Supplementary Fig. 1). Genes encoding 16 designs passing these filters were synthesized; 14 were expressed in the soluble fraction and 10 were largely monomeric by size-exclusion chromatography (SEC; Supplementary Table 1).

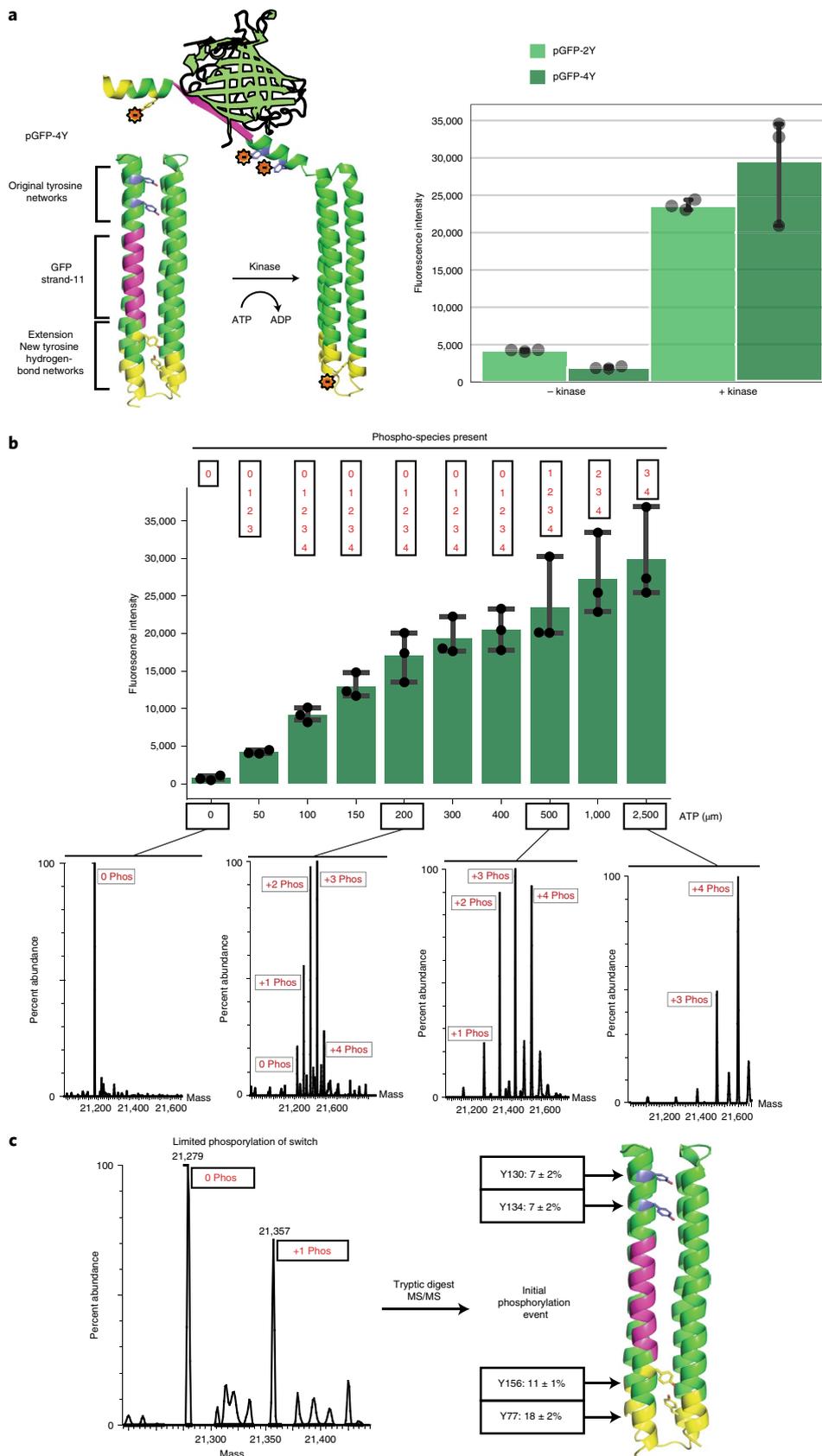
Green fluorescent protein switch

With the core phosphorylation switch mechanism in place, we sought to incorporate peptides that interact with external protein partners into the latch such that the interaction was only possible in the open state of the switch. We first sought to design a kinase-induced fluorescence switch by threading the green fluorescent protein (GFP) β -strand 11 sequence, DHMVLHERVNAAGIT, and some variants onto the latch helix of the monomeric designs. In isolation, the GFP11 peptide can assemble with the remainder of the GFP(GFP1–10) and cause the internal chromophore to mature, resulting in green fluorescence. We reasoned that this would be possible in the open (phosphorylated) form of the switch, but not the

Fig. 2 | Phosphorylation activates GFP fluorescence. **a**, The helices of pGFP-2Y were extended (yellow) and a new double-tyrosine hydrogen-bond network introduced. The caged GFP11 is shown in magenta and the original tyrosine phospho-sites in blue. The graph shows the previous pGFP-2Y in vitro switching plotted next to the new pGFP-4Y data. Data are presented as mean values \pm standard deviation from three independent experiments. **b**, The GFP fluorescence of pGFP-4Y, GFP1-10 and ATP increases with the addition of Src kinase. The correlation between phosphorylation and pGFP-4Y activation is shown. Pre-phosphorylated pGFP-4Y was added to GFP1-10. ATP concentration was used as a limiting reagent to control the amount of phosphorylation that occurred. Phosphorylation was measured by denaturing whole-protein MS. Because of the large number of possible heterogeneously phosphorylated species, we did not attempt to quantify the phosphorylation. The detected phospho-species for pGFP-4Y are shown above each bar. Examples of the deconvoluted mass spectra for the 0, 200 μ M, 500 μ M and 2.5 mM ATP samples are shown below the graph. Data are presented as mean values \pm standard deviation from three independent experiments. **c**, The pGFP-4Y initial phosphorylation event is spread across all tyrosine sites. We performed a limited phosphorylation such that only a single phosphorylation occurred per protein, representing the initial phosphorylation event. The samples were digested with trypsin and quantitatively analyzed by MS/MS to determine which sites were phosphorylated (Supplementary Table 3). The percentages reported are for the total phosphorylation for each phospho-site, individually, in the sample. The reported error is the standard deviation from three experiments.

closed state, due to steric exclusion in the latter¹². We found threadings of the GFP peptide predicted to be compatible with the structure for six of the solubly expressed buried tyrosine network proteins described above and obtained synthetic genes encoding these

designed switches. One of the six, which we refer to as pGFP-2Y in the remainder of the Article, was soluble and monomeric by SEC. It is a straight bundle with (*i/i* + 4) stacked tyrosine networks on the C-terminal helix (Fig. 1e and Supplementary Table 1). To test



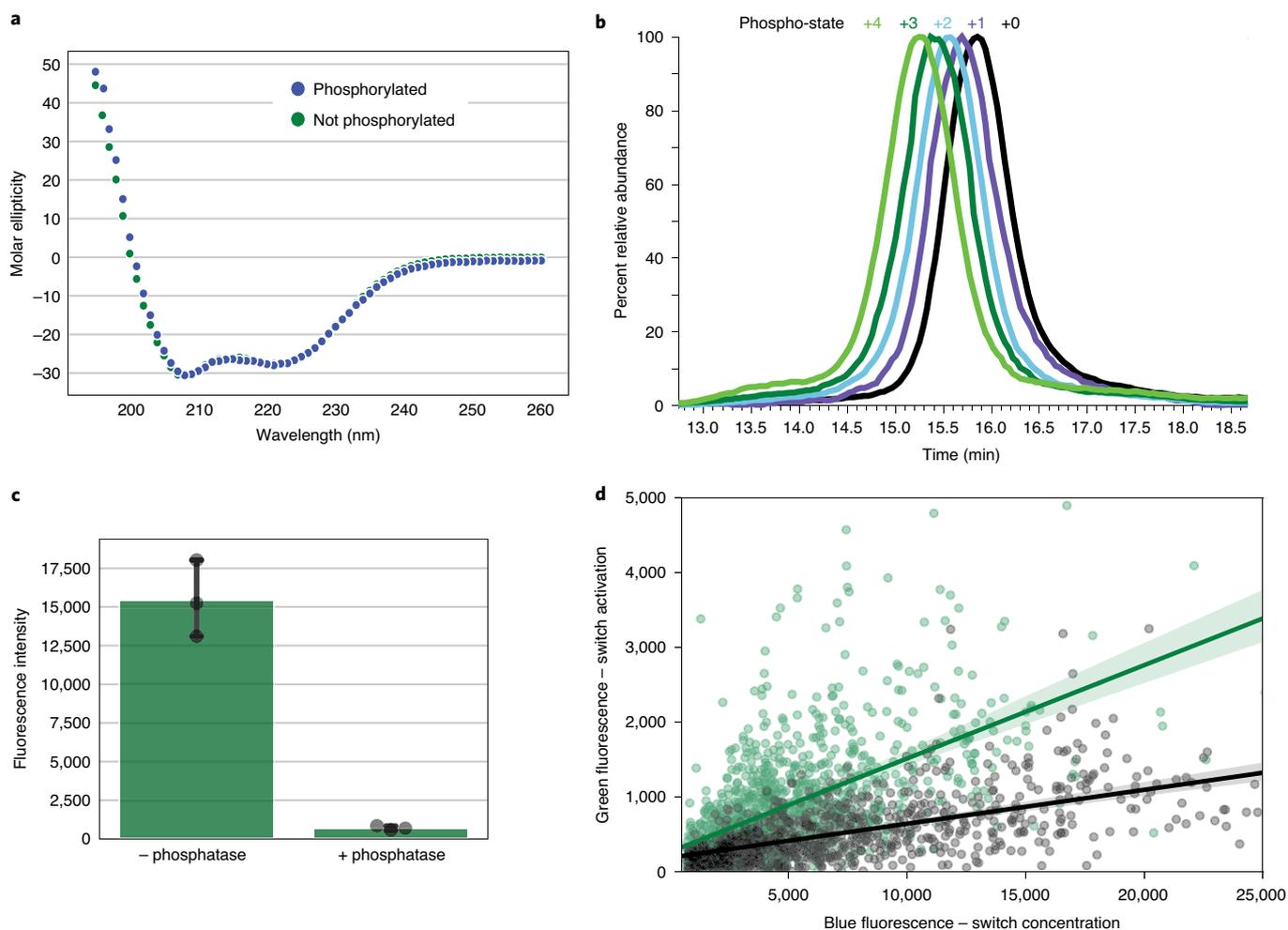


Fig. 3 | Characterization of the phospho-states of pGFP-4Y. **a**, Both phosphorylated and unphosphorylated pGFP-4Y have helical CD spectra. **b**, Increasing phosphorylation results in a larger hydrodynamic radius. A mixed sample of phosphorylated and unphosphorylated pGFP-4Y was separated over a size-exclusion column and detected by nMS. Smoothed and overlaid extracted ion chromatograms (XICs) for m/z 2,550–2,551, 2,560–2,561, 2,570–2,571, 2,580–2,581 and 2,590–2,591 correspond to 8+ pGFP-4Y with 0, 1, 2, 3 and 4 phosphorylations, respectively. **c**, GFP fluorescence of phosphorylated or dephosphorylated pGFP-4Y by the lambda phosphatase added to GFP1-10. Data are presented as mean values \pm standard deviation from three independent experiments. **d**, Activation of pGFP-4Y-Tsk4 in HEK 293T cells by constitutively active Src-530F. HEK 293T cells expressing GFP1-10 from the chromosome were transiently transfected with pGFP-4Y-Tsk4 (black) or Src-530F with pGFP-4Y-Tsk4 (green). Enhanced blue fluorescent protein (EBFP) co-expressed with pGFP-4Y-Tsk4 via P2A peptide monitors the concentration of pGFP-4Y-Tsk4 via blue fluorescence. Each point represents the fluorescence of a single cell from three separate experiments measured 48 h after transfection. A linear regression with 95% confidence interval is plotted for each dataset.

for phosphorylation-induced switching, 2 μ M pGFP-2Y was mixed with 1 μ M GFP1-10 and 500 μ M adenosine triphosphate (ATP) (Fig. 1e). Addition of the Src-kinase catalytic domain resulted in a fivefold increase in GFP fluorescence, consistent with the designed phosphorylation switch concept (Fig. 1f). However, pGFP-2Y was also observed to equilibrate with a larger oligomeric species, which probably limits the extent of activation (Supplementary Fig. 2).

We sought to stabilize the bundle by extending each helix by one turn while maintaining the loops. To counteract the decrease in switching expected from the concomitant increase in the free energy of latch disassociation, we searched the newly introduced positions for additional hydrogen-bond network placements compatible with the Src-kinase phosphorylation site to provide a larger destabilization upon phosphorylation and increase switching. We were able to incorporate another hydrogen-bond network with two additional Src-phosphorylation sites. The resulting pGFP-4Y design is a four-helix bundle, with 35 amino acids per helix, containing four Src phospho-sites (Fig. 2a).

The pGFP-4Y design was expressed in *E. coli* at higher levels than pGFP-2Y, and was found to remain monomeric for over a week when stored at 4 °C (Supplementary Fig. 2). pGFP-4Y has a highly helical content, as indicated by CD spectroscopy, consistent with the design model, and was found to be monomeric by SEC-multiple angle light scattering (SEC-MALS; Supplementary Fig. 2). The small-angle X-ray scattering (SAXS) spectrum of pGFP-4Y matches that expected from the computational design model, suggesting that the overall shape of the protein is close to that designed (Supplementary Fig. 2 and Supplementary Table 2). We investigated pGFP-4Y switching by comparing fluorescence in the presence and absence of Src kinase. Addition of the Src kinase to pGFP-4Y, ATP and GFP1-10 resulted in a 10-fold increase in GFP fluorescence (Fig. 2a).

Characterization of the pGFP-4Y switching mechanism

To characterize the mechanism by which Src kinase activates the pGFP-4Y switch, we investigated the relationship between the extent of phosphorylation and activation of GFP fluorescence. We

generated pGFP-4Y with different extents of phosphorylation in the absence of GFP1–10, utilizing ATP as a limiting reagent, and analyzed the protein using intact-protein MS. As there are four tyrosines that can each be either phosphorylated or not phosphorylated, there are 16 possible species, and we were not able to quantify the populations of each of these species; we focused instead on the overall properties of the distribution. The distribution of pGFP-4Y phosphorylation states goes smoothly from completely unphosphorylated at zero ATP to nearly completely phosphorylated at 5 mM ATP (Supplementary Fig. 3 and Fig. 2a). There was a strong correlation between the amount of ATP/phosphorylation and the extent of activation of GFP fluorescence following the addition of GFP1–10. The most heavily phosphorylated population of pGFP-4Y activates GFP 40-fold more than the unphosphorylated protein (Fig. 2b).

We next sought to determine whether there was an order in which the tyrosine residues become accessible to the kinase. We performed a limited phosphorylation of pGFP-4Y such that each protein was phosphorylated on no more than one tyrosine (Fig. 2c). Using tryptic digestion and MS/MS on the unphosphorylated and singly phosphorylated states, we found that, despite the different accessibility of tyrosines in the designed folded state, there was a roughly equal probability for each of the four to be the first to be phosphorylated (Fig. 2c and Supplementary Table 3). The simplest explanation for this lack of preference is that the entire switch must be in the unfolded state for the tyrosine residues to become phosphorylated—in this case, all tyrosines would be approximately equally accessible (in native protein systems, randomly ordered phosphorylation has also been observed to an unfolded state of the region being phosphorylated¹³). In vitro, the switch takes ~2 h at room temperature to become fully phosphorylated; this probably reflects the slow intrinsic opening rate of the pGFP-4Y latch (Supplementary Fig. 4).

How does the addition of four large and negatively charged phosphate groups influence the structure of pGFP-4Y? We found that the fully phosphorylated form of pGFP-4Y is monomeric by native MS (nMS) and has a helical CD spectrum identical to that of the unphosphorylated form (Fig. 3a and Supplementary Fig. 4). To probe the effect of phosphorylation on the hydrodynamic radius (which reflects both the size and shape of the protein), we analyzed a mixture of 0–4 phosphorylated pGFP-4Y by SEC coupled to nMS. The eluting monomeric species differ in their m/z due to the different amounts of phosphorylation, allowing us to distinguish co-eluting species and extract their individual elution profiles (Fig. 3b and Supplementary Fig. 4). We found a decrease in elution time with increasing phosphorylation of pGFP-4Y that is not attributed to a change in oligomerization, indicating an increase in hydrodynamic radii due to a shift in protein conformation.

We next investigated whether the pGFP-4Y phosphorylation switch is reversible. Because the maturation of the GFP chromophore is an irreversible process, we examined whether, in the absence of GFP1–10, the addition of phosphatase could return phosphorylated pGFP-4Y to a caged GFP11 state. First, we fully phosphorylated pGFP-4Y and measured its phospho-state by MS. We then split the sample in two, and for one half used lambda phosphatase to dephosphorylate the tyrosines (Supplementary Fig. 4). When mixed with GFP1–10, the dephosphorylated pGFP-4Y does not induce GFP fluorescence, indicating that the phosphorylation switch is reversible (Fig. 3c).

Because tyrosine phosphorylation-based cellular signaling plays a central role in mammalian cell signaling, we tested if pGFP-4Y could switch in HEK 293T cells. The regulation of tyrosine phosphorylation in mammalian cells occurs primarily through co-localization of kinase and the substrate, with phosphatase activity dominating elsewhere¹⁴. We appended an Src-binding peptide from a native Src substrate, Tsk4, that recruits the Src kinase to

create pGFP-4Y-Tsk4¹⁵. In a GFP1–10 expressing background, we transiently transfected pGFP-4Y-Tsk4 with or without Src-530F, a constitutively active variant of Src, to test phosphorylation-based switching. To control for the variable expression from transient transfection, we expressed EBFP with a P2A site behind pGFP-4Y-Tsk4 to monitor the concentration of pGFP-4Y-Tsk4 in the cell, and normalized the switch activation signal (green fluorescence) by the total amount of sensor (blue fluorescence). In the presence of added Src, there was, on average, 2.5 ± 0.1 -fold more switch activation (green fluorescence) than in the absence of added Src (Fig. 3d). Thus, the phosphorylation switch functions as designed in mammalian cells. HEK 293T cells natively contain the Src kinase, which can be activated in normal growth media, probably contributing (through increasing the background switching) to the reduced fold activation relative to the in vitro system.

Calpain inhibition switch

We next explored the versatility of the phosphorylation-induced binding mechanism by changing the functionality of the switch by incorporating in the latch the IA peptide from the calpain inhibitor, calpastatin (Supplementary Fig. 5). Calpain is a regulatory cysteine protease involved in cell motility and cell cycle progression. Calpastatin tightly binds calpain with its IA domain, while the IB domain inhibits the active site^{16,17}. We found a threading of the IA peptide, MDAALDDLITLGG, predicted to not destabilize the closed form of the switch. We call this design pDIA-4Y.

We expressed pDIA-4Y in *E. coli* and found the purified protein to be soluble and monomeric. The level of activation by phosphorylation of the initial design was low, suggesting the binding affinity of the latch for the cage was too strong. We tuned the latch-cage affinity by introducing mutations of Leu and Ile residues on both sides of the interface to Ala or Ser (L17S, L28A, L53A, I56A, I71A, L94A, L108A, L161S). We observed an increase in the degree of phosphorylation of the switch with increasing numbers of mutations, finally yielding pDIA-4Y_M8 (Supplementary Fig. 5). We measured the relative binding of phosphorylated pDIA-4Y_M8 and unphosphorylated pDIA-4Y_M8 to their binding partner domain IV (DIV) of calpain by biolayer interferometry (Supplementary Fig. 5). Phosphorylation of pDIA-4Y_M8 results in an (8.5 ± 1.6) -fold increased binding to DIV of calpain immobilized on the tip.

To convert pDIA-4Y_M8 into a phosphorylation-inducible inhibitor of calpain, we fused the calpain-inhibitory IB peptide downstream of the latch to create pDIA-4Y_M8_IB. In this scheme, phosphorylation of pDIA-4Y_M8 opens the latch, enabling the caged IA peptide to bind calpain, and then the adjacent IB peptide can block the enzyme active site. We assayed the ability of phosphorylated pDIA-4Y_M8_IB to inhibit calpain. Calpain activity was monitored spectroscopically by the fluorescence increase resulting from cleavage of a calpain substrate peptide (Supplementary Fig. 5)¹⁸. We compared the phosphorylated and unphosphorylated switch to the uncaged IAB calpastatin positive control and the unlocalized IB peptide negative control. Phosphorylation of pDIA-4Y_M8_IB increased calpain inhibition; the fold activation is lower than that observed for the binding interaction in part because the unphosphorylated pDIA-4Y_M8_IB retains considerable inhibitory activity (localization is not essential for inhibition under the conditions we tested).

Design of the phospho-serine PKA switch

Given our success at using the phosphorylation of tyrosine residues in core hydrogen-bonding networks to drive switching, we reasoned the same switching mechanism should be extensible to serine-based phosphorylation by designing protein hydrogen-bond networks incorporating a serine-based phosphorylation site. We chose to design a switch using the consensus phosphorylation site RRASL from the cyclic-AMP-dependent PKA due to its centrality

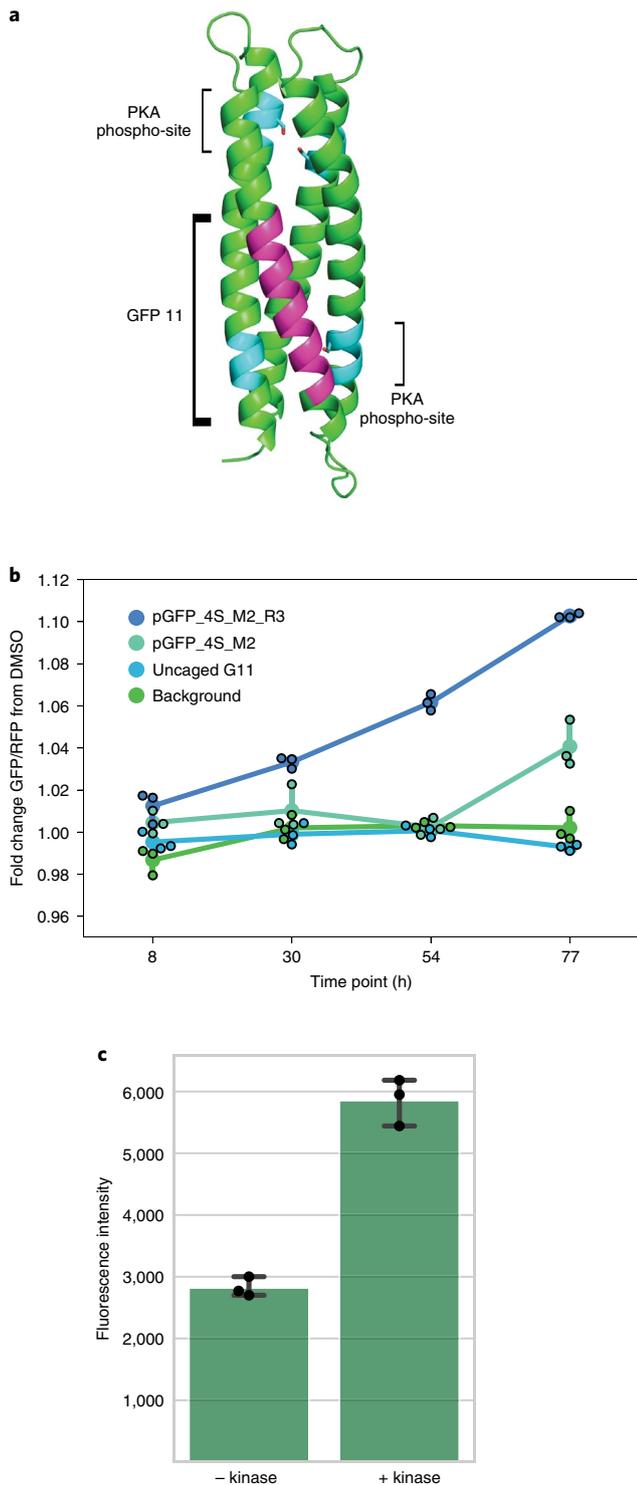


Fig. 4 | Serine-based phosphorylation switch. **a**, Design model for pGFP-4S. The GFP11 threading (magenta) and PKA phosphorylation sites (cyan) are shown on the structure. **b**, Activation of pGFP-4S in K562 cells by the PKA kinase. The serine switch monitored by the red fluorescent protein (RFP) fluorescence was used to normalize the GFP fluorescence for each cell. The background (GFP1-10 expression alone) and GFP1-10 activated with an uncaged version of GFP11 showed no changes upon forskolin and IBMX activation of the PKA kinase. Data are presented as mean values \pm standard deviation from three independent experiments. **c**, pGFP-4S_M2 phosphorylated by the PKA kinase in vitro preferentially activates GFP1-10. Data are presented as mean values \pm standard deviation from three independent experiments.

in G-protein-coupled receptor signaling^{19,20}. Using a strategy analogous to that for pGFP-4Y, we set out to design a four-helix bundle that incorporates four ‘RRASL’ motifs with phosphorylated serine involved in core hydrogen-bond networks. Upon phosphorylation of the serine residue, the core of the protein should be disrupted sterically and electronically by the phosphoryl group.

Following the same protocol as the tyrosine phosphorylation switch, we first identified four positions in a four-helix bundle for the RRASL site that both positions the serine in the core and the charged arginine residues on the surface. Based on the results with pGFP-4Y suggesting phosphorylation through global unfolding, the phosphorylation sites were not limited to the C-terminal helix (Fig. 4a). A total of 100,000 nearby helical backbones were generated from helical parameters, and the HBNet mover was used to search for hydrogen-bond networks that satisfied all four serine residues in the RRASL sites. For the 350 backbones with such networks, the remainder of the hydrophobic core and surface residue were designed and filtered by Rosetta energy, secondary-structure side chain complementarity, buried unsatisfied polar residues and cavities. Genes encoding 12 designs passing these filters were synthesized. All designs expressed in the soluble fraction and four designs were monomeric by SEC (Supplementary Table 4). We sought to introduce functionality into the three designs with the best solution properties by threading the GFP β -strand 11 sequence, DHMVLLERVNAAGIT (or a closely related variant), onto a helix of the design. From seven threadings, six designs expressed, and one of those, denoted pGFP-4S, maintained the monomeric form of the protein by SEC (Supplementary Table 4). pGFP-4S showed limited phosphorylation by PKA at room temperature (Supplementary Fig. 5).

We hypothesized that, as in the case of pDIA-4Y, destabilizing pGFP-4S mutations would increase phosphorylation, so we generated and screened variants for PKA-driven GFP reconstitution by co-expressing them with GFP1-10 in K562 cells. Reasoning that high local PKA concentrations could enhance activation of the switch in cells, we also tested switches that were C-terminally tagged with a 3X tandem repeat of the PKA consensus sequence (RRASL) as a recruitment motif (denoted by R3 in the name). To activate PKA, we co-treated cells with 50 μ M forskolin and 100 mM 3-isobutyl-1-methylxanthine (IBMX) and screened cells at 8, 30, 55 and 77 h post treatment for GFP fluorescence. The optimal switch variant, I48A and L49A (called pGFP-4S_M2_R3), demonstrated a small but reproducible fold change of 1.10 (± 0.001) (Fig. 4b), which is on par with most current PKA biosensors, albeit on a substantially longer timescale^{21–23}. The recruitment motif dramatically increased the response to PKA activation. When phosphorylated in vitro, pGFP-4S_M2 activates green fluorescence 2.1 (± 0.2) times more than unphosphorylated pGFP-4S_M2 (Fig. 4c and Supplementary Fig. 5). Thus, our phosphorylation switch concept extends to serine phosphorylation and can be activated by endogenous kinases. The reduced activation of pGFP-4S_M2 relative to the tyrosine switch may reflect the greater concentration of phosphorylation sites in the latch helix for pGFP-4Y relative to the dispersed phospho-sites in pGFP-4S.

Discussion

It is instructive to compare our designed binding-induced phosphorylation and phosphorylation-induced binding switches to naturally occurring kinase switches that are the product of natural selection. Signaling networks utilizing protein phosphorylation have phospho-sites that frequently occur in intrinsically disordered regions such that the phospho-acceptor is always accessible in the unphosphorylated state²⁴. Often, the additional negative charge of the phosphate group promotes a disorder to order transition, for example, the activation of ZAP-70 kinase via binding a doubly-phosphorylated and previously disordered ITAM motif. Our binding-induced phosphorylation switch reverses the order

of events: binding of the 'key' peptide releases the latch helix containing the ITAM motif, which probably becomes disordered in isolation. In this way, the expression of 'key' peptide modulates the functional change brought about by kinase activation (in the case of ITAM, T-cell activation) by regulating the accessibility of the substrate to the kinase. Analogously, in natural systems, phosphorylation of the Y291 in WASP is controlled by the binding of activated Cdc42. Once Cdc42 is bound, Y291 becomes unstructured and available for phosphorylation, resulting in increased actin remodeling through activation of the Arp2/3^{25,26}. Our designed phosphorylation-induced binding switch employs structured phospho-sites that are disrupted by the electronic and steric clash of the added phosphoryl groups. This is analogous to the mechanism of phosphorylation-based switching for the voltage-dependent potassium channel, K_v3.4, the cell cycle inhibitor p27^{Kip1} and the KH domain of the messenger RNA regulatory protein, KSRP^{27–29}. For K_v3.4, phosphorylation of serine residues disrupts the structure of the inactivation domains, altering the inactivation kinetics of the channel³⁰. Our current switches activate slowly due to the high cooperativity of unfolding; to increase the rate of activation it should be possible to design switches with low-energy intermediates that expose individual tyrosine phosphorylation sites similar to p27^{Kip1}, although designing multiple intermediates remains a challenge for de novo design²⁸. Although slow kinetics increases the response time after a signaling event, they can be advantageous for molecular timing and memory, as in the slow re-phosphorylation of the phosphoducin protein upon complex formation after light stimulation in rod cells³¹.

Our designed switches provide a mechanism for designed proteins to ascertain or alter the state of the cell through direct interactions with signaling networks. As an illustrative example, the advent of chimeric antigen receptor T cells (CAR-T) as a cancer therapy has created new synthetic biology challenges such as mitigating T-cell exhaustion from excess tyrosine kinase signaling through the chimeric receptor; exhausted cells have a reduced ability to perform effector T-cell functions^{32,33}. CAR-T expansion and action occur over many days, making switch kinetics less relevant³⁴. For our binding-dependent phosphorylation switches, the reduction of the level of key peptide by placement under the control of repressors associated with T-cell exhaustion could inhibit caged ITAMs and mute CAR signaling during times of excess signaling^{35,36}. Alternately, a phosphorylation-dependent binding switch that responds to the tyrosine kinase that starts T-cell activation, Lck, could cage an inhibitor of T-cell activation, and dampen excess signaling from the CAR. Going beyond previous designed phospho-switches that control coiled-coil dimerization⁴, our approach enables caging of peptides that interact with other native cellular proteins, including potentially toxic peptides like the BCL-2-interacting mediator of cell death (BIM), which drive apoptosis and are difficult to sequester within a coiled-coil. There is still considerable room for improvement in the dynamic range of the switches described in this Article, both by decreasing signal in the absence of kinase and increasing the extent of activation in the presence of kinase. These improvements will facilitate applications in engineered T cells and other synthetic biology systems.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41594-021-00649-8>.

Received: 18 December 2020; Accepted: 26 July 2021;
Published online: 13 September 2021

References

- Balalashnan, S. & Zondlo, N. J. Design of a protein kinase-inducible domain. *J. Am. Chem. Soc.* **128**, 5590–5591 (2006).
- Szilak, L., Moitra, J. & Vinson, C. Design of a leucine zipper coiled coil stabilized by 1.4 kcal mol⁻¹ by phosphorylation of a serine in the e position. *Protein Sci.* **6**, 1273–1283 (1997).
- Smith, C. A. et al. Design of a phosphorylatable PDZ domain with peptide-specific affinity changes. *Structure* **21**, 54–64 (2013).
- Winter, D. L., Iranmanesh, H., Clark, D. S. & Glover, D. J. Design of tunable protein interfaces controlled by post-translational modifications. *ACS Synth. Biol.* **9**, 2132–2143 (2020).
- Schlessinger, J. & Ullrich, A. Growth factor signaling by receptor tyrosine kinases. *Neuron* **9**, 383–391 (1992).
- Hatada, M. et al. Molecular basis for interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor. *Nature* **377**, 32–38 (1995).
- Langan, R. A. et al. De novo design of bioactive protein switches. *Nature* **572**, 205–210 (2019).
- Ng, A. H. et al. Modular and tunable biological feedback control using a de novo protein switch. *Nature* **572**, 265–269 (2019).
- Chen, Z. et al. Programmable design of orthogonal protein heterodimers. *Nature* **565**, 106–111 (2019).
- Shah, N., Lobel, M., Weiss, A. & Kuriyan, J. Fine-tuning of substrate preferences of the Src-family kinase Lck revealed through a high-throughput specificity screen. *eLife* **7**, e35190 (2018).
- Maguire, J. B., Boyken, S. E., Baker, D. & Kuhlman, B. Rapid sampling of hydrogen bond networks for computational protein design. *J. Chem. Theory Comput.* **14**, 2751–2760 (2018).
- Cabantous, S., Terwilliger, T. C. & Waldo, G. S. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat. Biotechnol.* **23**, 102–107 (2005).
- Salazar, C. & Höfer, T. Multisite protein phosphorylation—from molecular mechanisms to kinetic models. *FEBS J.* **276**, 3177–3198 (2009).
- Mayer, B. J. Perspective: dynamics of receptor tyrosine kinase signaling complexes. *FEBS Lett.* **586**, 2575–2579 (2012).
- Dülk, M. et al. EGF regulates the interaction of Tks4 with Src through its SH2 and SH3 domains. *Biochemistry* **57**, 4186–4196 (2018).
- Hanna, R. A., Campbell, R. L. & Davies, P. L. Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin. *Nature* **456**, 409–412 (2008).
- Todd, B. et al. A structural model for the inhibition of calpain by calpastatin: crystal structures of the native domain VI of calpain and its complexes with calpastatin peptide and a small molecule inhibitor. *J. Mol. Biol.* **328**, 131–146 (2003).
- Cuerrier, D. et al. Development of calpain-specific inactivators by screening of positional scanning epoxide libraries. *J. Biol. Chem.* **282**, 9600–9611 (2007).
- Robinson-White, A. & Stratakis, C. Protein kinase A signaling 'cross-talk' with other pathways in endocrine cells. *Ann. NY Acad. Sci.* **968**, 256–270 (2002).
- Kemp, B., Graves, D., Benjamini, E. & Krebs, E. Role of multiple basic residues in determining the substrate specificity of CAMP-dependent protein kinase. *J. Biol. Chem.* **252**, 4888–4894 (1977).
- Depry, C., Allen, M. D. & Zhang, J. Visualization of PKA activity in plasma membrane microdomains. *Mol. Biosyst.* **7**, 52–58 (2011).
- Chen, Y., Saulnier, J. L., Yellen, G. & Sabatini, B. L. A. PKA activity sensor for quantitative analysis of endogenous GPCR signaling via 2-photon FRET-FLIM imaging. *Front. Pharmacol.* **5**, 56 (2014).
- Mehta, S. et al. Single-fluorophore biosensors for sensitive and multiplexed detection of signalling activities. *Nat. Cell Biol.* **20**, 1215–1225 (2018).
- Iakoucheva, L. M. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037–1049 (2004).
- Torres, E. & Rosen, M. Contingent phosphorylation/dephosphorylation provides a mechanism of molecular memory in WASP. *Mol. Cell* **11**, 1215–1227 (2003).
- Kim, A., Kakalis, L., Abdul-Manan, N., Liu, G. & Rosen, M. Autoinhibition and activation mechanisms of the Wiskott–Aldrich syndrome protein. *Nature* **404**, 151–158 (2000).
- Antz, C. et al. NMR structure of inactivation gates from mammalian voltage-dependent potassium channels. *Nature* **385**, 272–275 (1997).
- Tsytonok, M. et al. Dynamic anticipation by Cdk2/Cyclin A-bound p27 mediates signal integration in cell cycle regulation. *Nat. Commun.* **10**, 1676 (2019).
- Diaz-Moreno, I. et al. Phosphorylation-mediated unfolding of a KH domain regulates KSRP localization via 14-3-3 binding. *Nat. Struct. Mol. Biol.* **16**, 238–246 (2009).
- Antz, C. et al. Control of K_v channel gating by protein phosphorylation: structural switches of the inactivation gate. *Nat. Struct. Biol.* **6**, 146–150 (1999).
- Wilkins, J., Bitensky, M. & Willardson, B. Regulation of kinetics of phosphoducin phosphorylation in retinal rods. *J. Biol. Chem.* **271**, 19232–19237 (1996).

32. Majzner, R. G. & Mackall, C. L. Clinical lessons learned from the first leg of the CAR T cell journey. *Nat. Med.* **25**, 1341–1355 (2019).
33. Salter, A. et al. Phosphoproteomic analysis of chimeric antigen receptor signaling reveals kinetic and quantitative differences that affect cell function. *Sci. Signal.* **11**, eaat6753 (2018).
34. Sahoo, P. et al. Mathematical deconvolution of CAR T-cell proliferation and exhaustion from real-time killing assay data. *J. R. Soc. Interface* **17**, 20190734 (2020).
35. Shin, H. et al. A role for the transcriptional repressor Blimp-1 in CD8⁺ T cell exhaustion during chronic viral infection. *Immunity* **31**, 309–320 (2009).
36. Lu, P. et al. Blimp-1 represses CD8 T cell expression of PD-1 using a feed-forward transcriptional circuit during acute viral infection. *J. Exp. Med.* **211**, 515–527 (2014).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021

Methods

Design of ITAM LOCKR. The designed heterodimer DHD-1:234A_B was used as base for the ITAM LOCKR (PDB 6DLC)⁹. Originally crystallized as three helices + one helix heterodimer, we connected the helices with a loop placing the single helix as C-terminal. The loop was composed of the ξ 1 ITAM sequence, NQLYNELNLGRREEYDVL, which was transitioned into the C-terminal helix such that the C-terminal tyrosine of the ITAM aligned with the Y123 in the original single helix in DHD-1:234A_B. We incorporated this loop sequence and helical segment into the DHD-1:234A_B model using the remodel program from the Rosetta suite³⁷. Design was carried out using Rosetta 2018.19. Using a blue print file we assigned a 'loop' secondary structure for residues NQLYNELNLGR and a 'helical' secondary structure for residues EEYDVL, overwriting the native residues of DHD-1:234A_B. We chose the lowest-energy structure from 30 runs as the design model. The key peptide for this ITAM LOCKR is the +1 helix from DHD_37. The single tyrosine residue for the key peptide was mutated to histidine to prevent phosphorylation by Lck.

Design of the tyrosine phosphorylation switch scaffold. Parametric backbones were generated using the Rosetta mover BundleGridSampler. Design was carried out using Rosetta 2018.19. For the straight helices, we sample a normal distribution with the given values being 2σ from the mean of the two values around the following parameters: Omega0 (-0.3, 0.3), r0 (6.5, 7), helix1 delta omega1 (-15, 15), helix1 z0 offset (-1, 1), helix2 delta omega1 (-15, 15), helix2 z0 offset (1, 4), helix3 delta omega1 (-15, 15), helix3 z0 offset (-1, 1), helix4 delta omega1 (-30, -40) and helix4 z0 offset (3, 4). Degrees were used for any angle and ångströms for any distance. For the left-handed helices, the following parameters were used: Omega0 (-3.9, -3.3), r0 (6.75, 7.1), helix1 delta omega1 (-10, 10), helix1 z0 offset (-4, -1), helix2 delta omega1 (-15, 15), helix2 z0 offset (1, 4), helix3 delta omega1 (-15, 15), helix3 z0 offset (-2, 1), helix4 delta omega1 (-15, 15) and helix4 z0 offset (1, 4). For both types of bundle, helices 2 and 4 were inverted for the N to C direction of the helices. Delta omega0 for was 0° for helix 1, 90° for helix 2, 180° for helix 3 and 270° for helix 4. Each helix contained 35 amino acids. We mutated in the tyrosine residues—positions 107 and 111 for the straight helices and positions 31 and 34 for the left-handed helices. Using 100,000 left-handed helices and 1,000,000 straight helices, we seeded the creation of hydrogen-bond networks off the tyrosine residues using the HBNet mover¹¹. We then identified the models that included a hydrogen-bond network with both tyrosine residues and only moved these models to the next stage of design. We maintained the constraints on the hydrogen-bond networks throughout the design process. In this stage, we added an additional hydrogen-bond network to the opposite end of the helices from the tyrosine residues. After generating these networks, we carried out sequence design on the remaining residues. Essentially, we used layer design and used PackRotamers mover to pack hydrophobic residues around hydrogen-bond networks. We then performed five cycles of limited use of the FastDesign mover only designing alanine, methionine and phenylalanine residues in a generic Monte Carlo mover optimizing for the least buried unsaturated hydrogen bonds and secondary structure side chain compatibility^{38–40}. We also used amino acid composition constraints to keep the fraction of alanine residues less than 0.07, and limited the total phenylalanine and methionine residues to eight each. From these structures, we then filtered on Rosetta score (<-3.0), secondary structure side chain compatibility (>0.65) and cavities by visual inspection using PyMOLs cavity display feature to reduce the number of sequences to 30 of each type (left-handed or straight). We connected the individual helices with loops using the ConnectChainsMover. We used FastRelax on the loop segments to get rid of clashes and designed the sequence with the PackRotamers mover. We then picked 12 left-handed helical models structure and nine straight helical models for experimental validation based on maximizing hydrogen-bond network diversity. We threaded the caged peptides onto the phosphorylation switch scaffolds by visual inspection using PyMOL.

Design of the PKA switch scaffold. Parametric backbones were generated using the Rosetta mover BundleGridSampler using the following parameters sampling a normal distribution, with the given values being 2σ from the mean of the two values. Omega0 (-4, -3.3), r0 (6.4, 7.0), helix1 delta omega1 (-10, 10), helix1 z0 offset (-4, -1), helix2 delta omega1 (-15, 15), helix2 z0 offset (1, 4), helix3 delta omega1 (-15, 15), helix3 z0 offset (-2, 1), helix4 delta omega1 (-15, 15) and helix4 z0 offset (1, 4). For both types of bundle, helices 2 and 4 were inverted for the N-to-C direction of the helices. Delta omega0 for was 0° for helix 1, 90° for helix 2, 180° for helix 3 and 270° for helix 4. Each helix contained 35 amino acids. We mutated in RRASL sites to the poly-alanine backbones with the serine positioned at residues 10, 45, 80 and 101. We performed the hydrogen-bond network and the remainder of the design as described in the previous section. After filtering, we selected 12 designs based on achieving hydrogen-bond network diversity. We threaded the caged peptides onto the phosphorylation switch scaffolds by visual inspection.

Gene construction. Synthetic genes were constructed for bacterial expression by either Genscript into a pET21b+ vector or with Integrated DNA Technologies (IDT) into the pET29b+ vector using the NdeI and XhoI restriction sites. For

the pET21b+ vector, a stop codon was included to prevent the expression of the C-terminal His-tag. For purification, an N-terminal His-tag was synthesized with the gene, GSSHHHHHHSSGS. The original designs for phosphorylation-dependent binding (no caged peptide), psGFP_2Y and psGFP_4Y, were expressed using a pET21b+ vector. The remaining design used the pET29b+ vector. For the pET29b+ vector, a stop codon was not included, allowing expression of the C-terminal His-tag after the XhoI site. Plasmids were transformed into chemically competent Lemo21(DE3) cells (NEB). Constructs that required His-tag cleavage were synthesized with an N-terminal His-tag followed by the tobacco etch virus (TEV) site in the pet29b+ vector with a stop codon to prevent expression of the C-terminal His-tag. Maltose binding protein (MBP) fusions (key peptides and uncaged control peptides) were synthesized into the pET28b+ vector from IDT that includes an N-terminal His-tag followed by a thrombin site before MBP. The fused gene was placed downstream of MBP after a TEV cleavage site. The DIV of calpain (UniProt P07384) was expressed with an N-terminal Avi-tag, LNDIFEAKQIEWHE, for *in vivo* biotinylation by BirA. For mammalian cell expression of pGFP_4Y and derivatives, genes were synthesized by GenScript in the pCMV vector using the KpnI/XhoI restriction sites. An EBFf was added to the C-terminal end of pGFP_4Y_Tsk4 behind a P2A site to monitor concentration in the cell. The Src-kinase domain (Addgene 79700), YopH phosphatase (Addgene cat. no. 79749)⁴¹, Src-530F (Addgene 124659)⁴² and pEGFP-GFP11 clathrin light chain (Addgene 70217)⁴³ were ordered from Addgene. The Lck kinase domain in the pET28b+ was a gift from the Andreotti laboratory⁴⁴. GFP strands 1–10 were taken from sfGFP. All numbering reflects the design model without the histidine tags. For the caged peptides, we used residues 216–230 from sfGFP. We used residues 153–166 from DIA of calpastatin (UniProt P20810) threaded at residue 135 to make psDIA_dY. We fused residues 170–222 from calpastatin IB to make psDIA_M8_IB from psDIA_M8. Residues 466–516 of Tsk4 were appended to pGFP_4Y to make pGFP_4Y_Tsk4⁴⁵. We removed any tyrosine residues that could be phosphorylated but were not part of the designed phosphorylation sites. In this way, we could assign any phosphorylation events identified by LC-MS to the designed phosphorylation sites. For this, we made a Y225F (native sequence numbering) mutation in the DIB peptide from calpastatin. Additionally, the N-terminal ITAM tyrosine (in the loop) was mutated to glutamate to reduce the construct to one tyrosine and simplify analysis by MS. ZAP-70 binding and activation strictly requires a doubly-phosphorylated ITAM, so controlling the phosphorylation of a single tyrosine in the ITAM is sufficient to control ZAP-70 activation⁴⁵.

Plasmids for testing pGFP_4S and variants were constructed by using the Mammalian Cloning Toolkit (MTK)⁴⁶. Human codon-optimized gene blocks for all de novo designed phosphoswitch variants were ordered from IDT. Phosphoswitch variants were assembled into multi-cistronic transcriptional units containing an Efla promoter driving GFP1–10^{15,43} with or without a 32 amino acid linker upstream of helix 2 of the phosphoswitch, a P2A element, either the phosphoswitch variant or IREFP713⁴⁷, an optional recruitment motif, a T2A element, mCherry as a transduction marker, and the WPRE transcriptional terminator, all in a lentiviral packaging backbone. All plasmids using the lentiviral backbone were grown in the Stb3 *E. coli* strain, and all other constructs were grown in Mach1 *E. coli* (QB3-Berkeley Macrolab). Protein sequences are included in the Supplementary Information.

Bacterial protein expression and purification. Starter cultures were grown overnight in Luria–Bertani (LB) medium with 50 µg ml⁻¹ carbenicillin (pET21b+), 50 µg ml⁻¹ kanamycin (pET28b+, pET29b+) or 10 µg ml⁻¹ spectinomycin (YopH expression, Addgene 79749). Starter cultures were used to inoculate 500 ml of Studier TBM-5052 auto-induction medium without antibiotic at a 1:100 ratio. The cells were grown at 37 °C for 3.5 h and switched to 18 °C for expression for 18 h. The cells were collected by centrifugation and lysed by sonication in lysis buffer (20 mM Tris, 300 mM NaCl, 30 mM imidazole pH 8.0) on ice. De novo proteins were lysed by sonication. Cells expressing native proteins were lysed at 80% amplitude for 5 s on/25 s off (cool) while cells expressing designed proteins were lysed for 10 s on/10 s off (warm). Lysates were cleared by centrifugation at 22,000 r.c.f. for 25 min. Supernatants were applied to (1 ml) Ni-NTA columns equilibrated in lysis buffer. The columns were washed twice with 10 column volumes of wash buffer (20 mM Tris, 300 mM NaCl, 50 mM imidazole pH 8.0). The columns were eluted with 5 ml of elution buffer (20 mM Tris, 300 mM NaCl, 250 mM imidazole pH 8.0). If the His-tags were removed by TEV cleavage, the sample was buffer-exchanged into 50 mM Tris pH 8.0, 0.5 mM EDTA and 1 mM DTT over a PD-10 column. Then, 0.5 mg of TEV protease was added to the sample and allowed to sit overnight at room temperature (4 °C for the Src-kinase domain). The TEV protease was removed by batch binding with 1 ml of Ni-NTA resin equilibrated in 10 mM HEPES pH 8.0, 150 mM NaCl for 10 min. The flow through of the column was collected and the protein was further purified over a Superdex 75 Increase 10/300 GL (GE) size-exclusion column into pooling fractions with the desired monomeric protein. Split-GFP was purified using Ni-affinity as above but with 10% glycerol added to the buffer. The Src kinase domain was co-expressed with the YopH phosphatase to prevent toxicity⁴¹. DIV of calpain was co-expressed with soluble BirA ligase and with 50 mM biotin added to the cell culture after the culture grew to an optical density at 600 nm of ~0.6 (ref. 48). Lck was expressed in

ArcticExpress (DE3) cells at 16 °C overnight. The cells were lysed with lysozyme and subjected to freeze–thaw in three cycles, purified at 4 °C by Ni-affinity resin, and the His-tag was removed using the TEV protease⁴⁴.

In vitro phosphorylation assay. For Lck phosphorylation, 500 nM Lck kinase domain were mixed with 50 μ M ITAM LOCKR in 50 mM HEPES pH 7.0, 10 mM MgCl₂, 1 mM DTT, 1 mg ml⁻¹ BSA, 1 mM Pefablock and 500 μ M ATP for at least 2 h. The amount of key peptide (fused to MBP) was included as indicated.

For Src phosphorylation, 500 nM Src kinase domain was mixed with 50 μ M phosphate switch in 10 mM HEPES pH 8.0, 10 mM MgCl₂, 2 mM ATP and allowed to react for at least 5 h. The non-phosphorylated samples were generated by incubating in the same solution without the Src kinase domain. To generate the singly phosphorylated sample of pGFP-4Y, the reaction was quenched with 20 mM EDTA after 10 min.

For PKA phosphorylation, 200 nM PKA kinase (NEB) was mixed with 25 μ M switch in 50 mM Tris pH 7.5, 10 mM MgCl₂, 2 mM ATP and allowed to react for at least 5 h at room temperature. The non-phosphorylated samples were generated by incubating in the same solution without the kinase.

Liquid chromatography mass spectrometry. Phosphorylation was monitored by whole-protein denaturing MS by means of electron spray ionization using a Waters Synapt-QTOF system at the University of Washington. The protein was desalted in line to the mass spectrometer using an AdvanceBio RP-Desalting (2.1 mm, 12.5 mm) column using an acetonitrile/water gradient in 0.1% formic acid with a flow rate of 0.5 ml min⁻¹ on a Waters Acquity Ultraperformance LC instrument. The sample was bound to the column using 10% acetonitrile isocratic flow for 1 min and then eluted by shifting to 95% acetonitrile using a linear gradient during the next minute. The concentration was held at 95% acetonitrile for 2 min. The column was shifted back to 10% acetonitrile for the next minute and then held at 10% acetonitrile for another minute. The spectrum was deconvoluted using Mass Lynx 4.1 software using the *m/z* range 600–1,500. A single phosphorylation event on a protein can be accurately quantified using peak ratios⁴⁹. We thus quantified the percent phosphorylation for the ITAM LOCKR by using the ion ratios between the phosphorylated and non-phosphorylated peaks.

In vitro GFP switching assay. For the tyrosine switch pre-phosphorylated assay, 50 μ M pGFP-4Y was mixed with 500 nM Src kinase in 10 mM HEPES pH 8.0, 10 mM MgCl₂, 150 mM NaCl and 2 mM ATP for a total of 200 μ l. The reaction was allowed to sit for 24 h in a covered 96-well plate. Reactions with or without kinase were mixed at 2 μ M pGFP-4Y with 1 μ M GFP1–10 for a total of 200 μ l and allowed to equilibrate for 1 day. For the tyrosine switch continuous phosphorylation assay, 2 μ M pGFP-4Y was mixed with 1 μ M GFP1–10 and 500 nM Src kinase in 10 mM HEPES pH 8.0, 10 mM MgCl₂, 150 mM NaCl and 500 μ M ATP to a total volume of 200 μ l. The reaction sat in the dark at room temperature for 3 days in a covered 96-well plate. For the serine switch pre-phosphorylated assay, 25 μ M pGFP-4S_M2 was mixed with 200 nM PKA kinase in 50 mM Tris pH 8.0, 10 mM MgCl₂ and 2 mM ATP for at least 5 h. Reactions with or without kinase were mixed at 1 μ M pGFP-4S_M2 with 1 μ M GFP1–10 for a total of 200 μ l and equilibrated for 1 day. For all preparations, GFP fluorescence was measured on a BioTek Synergy Neo2 Multi-mode plate reader with excitation at 488 nm and emission measured at 509 nm using a 10-nm bandwidth.

Phosphatase reaction. We purchased lambda protein phosphatase from New England BioLabs (NEB). We mixed 50 μ l of 50 μ M phosphorylated pGFP-4Y in the supplied reaction buffer with 400 units of the lambda phosphatase and let the reaction proceed overnight at 30 °C.

Circular dichroism measurements. Proteins were buffer-exchanged into 10 mM phosphate, 100 mM NaCl pH 7.4 via SEC. The CD spectrum from 260 nm to 190 nm was recorded with 0.2 mg ml⁻¹ protein in a 1 mm path length cuvette on a Jasco J-1500 CD spectrometer.

Solution X-ray scattering. Samples were separated by SEC. psGFP-4Y at 75 μ M in 10 mM HEPES pH 8.0 100 mM NaCl was used for SAXS analysis. Data were collected at the SIBYLS 12.3.1 beamline at the Advanced Light Source. The X-ray wavelength was 1.27 Å at 11 keV. The sample-to-detector distance of the Pilatus3 2M detector was 2 m, corresponding to a scattering vector q ($q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle) range of 0.008–0.36 Å⁻¹. Datasets were collected at 10 °C with 10-s exposure divided into 33 frames every 0.3 s. Frames were merged using the FrameSlice utility from SIBYLS to exclude radiation damage at low q values but reduce noise over the rest of the curve. Data were analyzed using the ScÅtter (version 2.5) software package to calculate SAXS statistics^{50,51}. The FoXS web server was used to compare design models to experimental scattering profiles and calculate quality of fit χ -values^{52,53}.

Size exclusion chromatography–multiple angle light scattering. We ran 3 mg ml⁻¹ of the indicated sample in 50 mM Tris-HCl, 150 mM NaCl pH 8.0 at 1 ml min⁻¹ over a Superdex 6 10/300 column in line with a Heleos multi-angle static light scattering system and an Optilab T-REX detector (Wyatt Technology

Corporation). The data were then analyzed using ASTRA (Wyatt Technologies) to calculate the weighted average molar mass (M_w) of the selected species and the number average molar mass (M_n) to determine the monodispersity to polydispersity index (PDI) = M_w/M_n .

Native mass spectrometry. Sample purity and integrity were analyzed by online buffer exchange MS using a Vanquish UHPLC system (Thermo Fisher Scientific) coupled to a Q Exactive UHMR Hybrid Quadrupole-Orbitrap instrument (Thermo Fisher Scientific); 100 pmol protein was injected and online buffer-exchanged to 200 mM ammonium acetate pH 6.8 (AmAc) by a self-packed buffer exchange column (P6 polyacrylamide gel, BioRad) at a flow rate of 100 μ l min⁻¹ (ref. 54). Mass spectra were recorded for 800–16,000 *m/z* at 17,500 resolution as defined at 400 *m/z*. The injection time was set to 200 ms. Voltages applied to the transfer optics were optimized to allow for ion transmission while minimizing unintentional ion activation. Mass spectra were deconvoluted with UniDec version 4.0.0 beta⁵⁵. For SEC coupled to nMS, unphosphorylated and phosphorylated samples were buffer-exchanged into 200 mM AmAc and combined to ensure the presence of 0–4 phosphorylated pGFP-4Y at sufficient concentration for subsequent MS detection. Protein (40 pmol) was injected onto a MabPac SEC-1 column (5 μ m 300 Å, 2.1 \times 300 mm; Thermo Fisher Scientific) and eluted with 200 mM AmAc at a flow rate of 50 μ l min⁻¹. RNaseA, cytochrome *c*, myoglobin, carbonic anhydrase and alcohol dehydrogenase were used as standards to calibrate the SEC column.

Tryptic digest tandem mass spectrometry. pGFP-4Y samples (25 μ l of 1 mg ml⁻¹ or 50 μ M) in 10 mM HEPES pH 7.5 containing 5 mM MgCl₂ and 100 mM NaCl were mixed 1:1 with 0.2% PPS Silent Surfactant (Expediton) in 50 mM Tris pH 8. Tryptic digestion was initiated by the addition of 1 μ g of trypsin overnight at 37 °C. Samples were diluted 100-fold to a final concentration of 0.25 pmol μ l⁻¹ in LC loading buffer (2% acetonitrile with 0.1% trifluoroacetic acid), and 3 μ l was injected onto the LC-MS system.

Synthetic peptides to be used as standards were ordered from GenScript. Dilutions of the synthetic peptides were solubilized in LC loading buffer, with the non-phosphorylated peptide kept at a constant 400 fmol μ l⁻¹ and the phosphopeptide diluted 1-, 4-, 20- and 40-fold to assess the quantitative change in mass spectrometric signal following addition of a phosphate group, then 3 μ l of the peptide mixtures was injected onto the LC-MS system.

Using a Waters Nanoacquty HPLC system, samples were injected via the autosampler onto a 150- μ m Kasil fritted trap packed with Reprosil-Pur C18-AQ (3- μ m bead diameter, Dr. Maisch) to a bed length of 2 cm at a flow rate of 2 μ l min⁻¹. After loading and desalting using a total volume of 8 μ l of loading buffer, the trap was brought online with a pulled fused-silica capillary tip (inner diameter, 75 μ m) packed to a length of 30 cm with the same Dr. Maisch beads. The column and trap were mounted to a homemade ion source that was operated at room temperature. Peptides were eluted off the column using a gradient of 2 to 32% acetonitrile in 0.1% formic acid over 1 h, followed by 32 to 60% acetonitrile over 5 min at a flow rate of 250 nl min⁻¹.

The mass spectrometer was operated using electrospray ionization (2 kV) with the heated transfer tube at 275 °C using data-dependent acquisition, whereby one ms1 scan (*m/z* 400–1,600) was acquired with up to 10 ms2 spectra. All data were acquired at unit resolution. The automatic gain control targets for MS was 3e4, and 8e3 for MS/MS. The maximum fill times were 10 and 50 ms, respectively. The MS/MS spectra were acquired with an isolation width of 2.0 *m/z* and collision-induced dissociation collision energy of 35%.

Biolayer interferometry. The affinity for pGFP_DIA to domain IV of calpain was measured on an Octet Red 96 instrument by ForteBio. We used 10 mM HEPES pH 7.0, 150 mM NaCl, 0.05% (vol/vol) Tween, 1 mM CaCl₂, 0.25% (wt/vol) BSA as binding buffer. Domain IV was bound to streptavidin-functionalized tips (18–5020 ForteBio) using the biotinylated Avi-tag. Samples were diluted from 50 μ M phosphorylation buffer into the binding buffer to a final concentration of 500 nM. The tips were equilibrated in binding buffer before initiating the binding reaction by dipping the tips into a well containing pDIA-4Y at 500 nM. The signal response was reported after 500 s of association to the tip.

Calpain inhibition assay. Calpain protease kinetics were monitored by incubation of calpain-1 (C6108) with calpain-1 substrate II (208772 Sigma). Calpain-1 substrate is an internally quenched peptide functionalized with an EDANS group on the N terminus and DABCYL group on the C terminus. After cleavage by calpain, the peptide becomes fluorescent. Calpain (30 nM) was mixed with 20 μ M calpain-1 substrate II in 50 mM HEPES pH 7.7, 0.1% (vol/vol) 2-mercaptoethanol and 1 mM CaCl₂ for 2 min, then 500 nM of the indicated pDIA_4Y or uncaged control peptides were mixed and fluorescence excited at 335 nm, and the emission was measured at 505 nm by a BioTek Synergy Neo2 Multi-mode plate reader for 2 h.

Mammalian cell activation of pGFP_4Y. Human embryonic kidney cells (HEK 293T) were obtained from ATCC (CRL-3216) and were not authenticated after acquisition. To generate a stable cell line expressing GFP1–10, HEK 293T cells were transfected with lentiviral particles delivering the GFP1–10-P2A-Puro gene driven by the MND promoter. Cells were then selected with puromycin

over a period of two weeks. GFP1–10 expression and complementation with GFP11 was verified through transfection of plasmid encoding GFP11, followed by fluorescence detection. Cells were cultured in DMEM (Gibco) supplemented with 1 mM L-glutamine (Gibco), 4.5 g l⁻¹ D-glucose (Gibco), 10% FBS and (1×) non-essential amino acids (Gibco). Cells were cultured at 37 °C and 5% CO₂ and passaged three times per week. For the transfection, HEK 293T GFP1–10 cells were plated at 40% confluence. After 24 h, the cells were transiently transfected with SRC-530F, pGFP_4Y_Tsk4 and pGFP_4Y_Tsk4/SRC-530F using linear 40-kDa polyethyleneimine (PEI-MAX, Polysciences).

Cell images were taken using the IN Cell Analyzer 2500 (GE) on a Greiner Screenstar microplate 655866 using a Nikon Plan Apo ×60/0.95-mm objective. The EBFP was excited at 390 nm and recorded with 0.008-s exposure. Activated split-GFP was excited at 473 nm and recorded with 0.1-s exposure time. Pictures were analyzed using INCarta software (GE, Version: 1.11.3667461). Cell areas were identified using blue fluorescence with a sensitivity of 50 (out of 100). These identified cells were used to record the fluorescence in both the green and blue channels. A region without cells was used for a background subtraction. The linear regression and confidence interval were produced by the SciPy stats library using the linregress method. The mean green fluorescence was normalized to the mean blue fluorescence for each of three separate experiments, before determining the fold activation from Src-530F expression.

Mammalian cell culture and cell line generation for pGFP_4S constructs. K562 cells were acquired from the UCSF Cell and Genome Engineering Core (UCSF CGEC cat. no. CCLZR466) and were not authenticated after acquisition. K562 cells were maintained in RPMI 1640 (Life Technologies) medium, supplemented with 10% FBS and 1% Anti-Anti. Cells were grown at 37 °C and 5% CO₂. Lenti-X 293T cells (Takara Bio) were transfected at 50% confluency in a 24-well plate (Corning) with lentiviral plasmids mixed with packaging plasmids (pMD2.G and pCMV-dR8.91), Lipofectamine 2000 (Thermo Fisher) and Opti-MEM (Life Technologies) according to the manufacturers' instructions. After 72 h of virus production, the supernatant containing accumulated virus was filtered through a 0.45-µm filter and dispensed onto K562 cells seeded at 50% confluency. The medium was replenished after 24 h, and transduced cells were expanded to a six-well plate (Corning). Using a BD FACSAria II cell sorter, the top 5% of mCherry-expressing cells were sorted and propagated in a 24-well plate. Sufficiently expanded cell lines were seeded into round-bottom Corning 96-well plates for the K562 cells at 2,000 cells per well for induction and flow cytometry experiments.

Drug preparation and cell line induction. Forskolin (Tocris) and IBMX (Tocris) were dissolved in DMSO (Sigma-Aldrich) in stock concentrations of 50 mM and 500 mM, respectively. Drugs were aliquoted in volumes suitable for single time-point inductions and stored at -20 °C, then thawed immediately before usage. At time points of 77, 54, 30 and 8 h before analysis, drug stocks were vortexed and dissolved in appropriate media for each cell type and added on top of triplicate wells to final concentrations of 50 µM and 100 µM for forskolin and IBMX, respectively. Corresponding volumes of DMSO were added to triplicate control wells.

Flow cytometry. All cells were washed with Dulbecco's phosphate buffered saline (DPBS) and resuspended in DPBS + 10% FBS. Flow cytometry experiments were conducted on an LSRFortessa system (BD) with a four-laser configuration (488 nm, 635 nm, 355 nm, 405 nm) using the software FACSDiva 8.0.1. GFP1–10 (excitation at 488 nm, 530/30 emission filter) and mCherry (excitation at 561 nm, 610/20 emission filter) fluorescence was collected for at least 10,000 single-cell events, recorded at a flow rate of 3 µl s⁻¹.

Data processing and analysis. Exported FCS files were analyzed with a custom Python script. Cells were density-gated on SSC-A versus SSC-H distributions to capture singlets. For every gated cell, the phosphoswitch response level was calculated by dividing the signal of interest (GFP1–10) by cassette expression level (mCherry), and the mean of the response was determined for each individual triplicate at each time point. The mean response of each DMSO-treated well was determined for all cells at every time point, representing a baseline signal for each construct and time point pair. To determine the fluorescence fold change upon forskolin and IBMX induction of PKA activity, the mean normalized GFP1–10 responses were divided by the mean baseline signal for each triplicate well. Analysis depended on the Python libraries numpy (version 1.18.1)⁵⁶, pandas, matplotlib, seaborn, scipy (version 1.4.1)⁵⁷ and fcsparser.

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

Data availability

The PKA switch in vivo data are available at <https://zenodo.org/record/5095560> and <https://doi.org/10.5281/zenodo.5095560>. Tryptic digest MS/MS data are available from the ProteomeXchange with identifier PXD027295. All protein sequences used in experiments are available as a Supplementary file, Protein_Sequences.xlsx. The design models experimentally tested are available as a supplementary protein databank file. The source backbone for the caged ITAM is available from the

Protein Data Bank under ID 6DLC. The remainder of the experimental data and computational data are available at <https://files.ipd.uw.edu/pub/PhosphateSwitch/DenovoPSwitch.zip>. Unique biological materials (plasmids) are available upon request to the corresponding author. Source data are provided with this paper.

Code availability

The custom Rosetta Scripts protocol for design with Rosetta 2018.19 is provided at https://github.com/NickWoodall/PhosphoSwitch_Design. Source code for PKA in vivo analysis is available at <https://github.com/weinberz/phosphoswitch>.

References

- Huang, P.-S. et al. RosettaRemodel: a generalized framework for flexible backbone protein design. *PLoS ONE* **6**, e24109 (2011).
- Fleishman, S. J. et al. RosettaScripts: a scripting language interface to the Rosetta Macromolecular Modeling Suite. *PLoS ONE* **6**, e20161 (2011).
- Bhardwaj, G. et al. Accurate de novo design of hyperstable constrained peptides. *Nature* **538**, 329–335 (2016).
- Hosseinizadeh, P. et al. Comprehensive computational design of ordered peptide macrocycles. *Science* **358**, 1461–1466 (2017).
- Albanese, S. K. et al. An open library of human kinase domain constructs for automated bacterial expression. *Biochemistry* **57**, 4675–4689 (2018).
- Zhang, Y. et al. Tail domains of myosin-1e regulate phosphatidylinositol signaling and F-actin polymerization at the ventral layer of podosomes. *Mol. Biol. Cell* **30**, 622–635 (2019).
- Kamiyama, D. et al. Versatile protein tagging in cells with split fluorescent protein. *Nat. Commun.* **7**, 11046 (2016).
- Boyken, S. E. et al. A conserved isoleucine maintains the inactive state of Bruton's tyrosine kinase. *J. Mol. Biol.* **426**, 3656–3669 (2014).
- Deindl, S. et al. Structural basis for the inhibition of tyrosine kinase activity of ZAP-70. *Cell* **129**, 735–746 (2007).
- Fonseca, P. et al. A toolkit for rapid modular construction of biological circuits in mammalian cells. *ACS Synth. Biol.* **8**, 2593–2606 (2019).
- Filonov, G. S. Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat. Biotechnol.* **29**, 757–761 (2011).
- Graslund, S., Savitsky, P. & Muller-Knap, S. in *Heterologous Gene Expression in E. coli* Vol. 1586 (ed. Burgess-Brown, N. A.) 337–344 (Springer, 2017).
- Wu, Z. et al. Impact of phosphorylation on the mass spectrometry quantification of intact phosphoproteins. *Anal. Chem.* **90**, 4935–4939 (2018).
- Dyer, K. N. et al. in *Structural Genomics: General Applications* (ed. Chen, Y. W.) 245–258 (Humana Press, 2014).
- Rambo, R. P. & Tainer, J. A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod–Debye law. *Biopolymers* **95**, 559–571 (2011).
- Schneidman-Duhovny, D., Hammel, M. & Sali, A. FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res.* **38**, W540–W544 (2010).
- Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. & Sali, A. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.* **105**, 962–974 (2013).
- VanAernum, Z. L. et al. Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry. *Nat. Protoc.* **15**, 1132–1157 (2020).
- Marty, M. T. et al. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).
- Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).
- Virtanen, P. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

Acknowledgements

N.B.W. and D.B. were supported by the Howard Hughes Medical Research Institute. M.M. was supported by The Audacious Project at the Institute for Protein Design. M.A. was supported by a gift from Amgen. I.Y. was supported by the NSF Graduate Research Fellowships Program (GRFP). Z.W. is supported by NIH 5K12GM081266. J.P. was supported by NIH 5T32GM007810. M.J.M. and R.S.J. are supported by NIGMS grant no. P41GM103533. D.B. and M.J.M. are supported by NIGMS U19 AG065156. Native mass spectrometry measurements were provided by the NIH-funded Resource for Native Mass Spectrometry Guided Structural Biology at The Ohio State University (NIH P41 GM128577 awarded to V.H.W.). We thank the staff at the Advanced Light Source SIBYLS beamline at Lawrence Berkeley National Laboratory, including K. Burnett, G. Hura, M. Hammel, J. Tanamachi and J. Tainer, for the services provided through the mail-in SAXS program, which is supported by the DOE Office of Biological and Environmental Research Integrated Diffraction Analysis program DOE BER IDAT grant (DE-AC02-05CH11231), NIGMS-supported ALS-ENABLE (GM124169-01) and National Institute of Health project MINOS (R01GM105404). Z.W. and J.P. thank A. Ng and A. Bonny, M. Kim for cloning advice and S. Allison for essential advice. We thank the Andreotti laboratory for their gift of the Lck kinase plasmid.

Author contributions

D.B. and N.B.W. conceived the switch concept. N.B.W. and M.J.F. designed the switches. Z.W., J.P. and H.E.-S. characterized the serine switch in cells. M.M., M.A., N.B.W. and I.Y. characterized the tyrosine switch in cells. F.B. and V.H.W. characterized the tyrosine switch with native MS. R.S.J. and M.J.M. characterized the tyrosine switch with MS/MS. N.B.W. performed all other experiments. All authors contributed to the writing of the manuscript.

Competing interests

N.W. and D.B. are inventors on US patent application PCT/US2020/038048. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41594-021-00649-8>.

Correspondence and requests for materials should be addressed to David Baker.

Peer review information *Nature Structural & Molecular Biology* thanks Dominic Glover and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available. Beth Moorefield was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

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 our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the 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
- A statement on 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 statistical parameters 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

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Rosetta 2018.19 used for design of proteins. Custom Rosetta Scripts protocol included as supplementary file.

Data analysis pGFP_4Y cell data analyzed using INCarta (GE version 1.11.3667461). FACs and analysis custom program for pGFP_4S at <https://github.com/weinberz/phosphoswitch> as well as numpy (version 1.18.1) and scipy (version 1.4.1). SAX analysis done with ScÅtter (version 2.5).

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 and reviewers. 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

Source data for Figures 1,2,3,4 and Supp. Fig 5 has been uploaded to the journal. The PKA switch in vivo data is available <https://zenodo.org/record/5095560> DOI: 10.5281/zenodo.5095560. Tryptic digest MS/MS available via ProteomeXchange with identifier PXD027295. All protein sequences used in experiments are available as a supplementary file Protein_Sequences.xlsx. The design models experimentally tested are available as a supplementary protein databank file. The source backbone for the caged ITAM is PDB ID: 6DLC. The remainder of the experimental data and computational data are available at <https://files.ipd.uw.edu/pub/PhosphateSwitch/DenovoPSwitch.zip>

Field-specific reporting

Please select the one below that is 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/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-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 sample size. Sample size of 3 was used. These sample sizes are sufficient due to use of purified protein and genetically identical samples.

Data exclusions: No data was excluded.

Replication: All attempts to replicate experimental results were successful. Separate transient transfections started the in vivo cell work and In vitro protein experiments samples were independently initiated from a purified protein prep on separate days.

Randomization: Not relevant to this study. All experimental components are genetically identical.

Blinding: Blinding was not used for this study because experiments were carried out by individual researchers.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s): HEK293T (ATCC® CRL-3216™) were obtained from ATCC. K562 cells were acquired from UCSF Cell and Genome Engineering Core

Authentication: Cell lines were not authenticated after acquisition from UCSF CGEC or ATCC

Mycoplasma contamination: HEK293T tested negative for mycoplasma. K562 cells were not tested for mycoplasma after acquisition from UCSF CGEC

Commonly misidentified lines (See [ICLAC](#) register): No commonly misidentified lines were used in this study.

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Cultured cells were prepared for flow cytometry by enzymatic dissociation (for adherent cell lines), and the DPBS washes before resuspension in final flow buffer.

Instrument

BD LSRFortessa X-20

Software

Cytometry data were collected using BD FACSDiva software. Data were analyzed using custom python code, available at <https://github.com/weinberz/phosphoswitch> after publication

Cell population abundance

Abundance: Cultured cell lines were pure and represented 100% of the sample run on the flow cytometer, no sorting was necessary.

Gating strategy

Cells were density gated on SSC-A/SSC-H to retain 90% of the population near the center of a 2D-gaussian fit against a random sampling of the cell population. No subsequent gating was necessary.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.