ARTICLE

https://doi.org/10.1038/s41586-019-1432-8

De novo design of bioactive protein switches

Robert A. Langan[1,2,3,11], Scott E. Boyken[1,2,11], Andrew H. Ng[4,5,6,7,11], Jennifer A. Samson[5], Galen Dods[4], Alexandra M. Westbrook[4], Taylor H. Nguyen[4], Marc J. Lajoie[1,2], Zibo Chen[1,2,3], Stephanie Berger[1,2], Vikram Khipple Mulligan[1,2], John E. Dueber[5], Walter R. P. Novak[8], Hana El-Samad[4,9] & David Baker[1,2,10]*

Allosteric regulation of protein function is widespread in biology, but is challenging for de novo protein design as it requires the explicit design of multiple states with comparable free energies. Here we explore the possibility of designing switchable protein systems de novo, through the modulation of competing inter- and intramolecular interactions. We design a static, five-helix 'cage' with a single interface that can interact either intramolecularly with a terminal 'latch' helix or intermolecularly with a peptide 'key'. Encoded on the latch are functional motifs for binding, degradation or nuclear export that function only when the key displaces the latch from the cage. We describe orthogonal cage–key systems that function in vitro, in yeast and in mammalian cells with up to 40-fold activation of function by key. The ability to design switchable protein functions that are controlled by induced conformational change is a milestone for de novo protein design, and opens up new avenues for synthetic biology and cell engineering.

There has been considerable progress in the de novo design of stable protein structures on the basis of the principle that proteins fold to their lowest free-energy state. These efforts have focused on maximizing the free-energy gap between the desired structure and all other structures, and have resulted in a wide range of stable proteins that exclusively populate the designed state[1–4]. Designing proteins that can switch conformations is more challenging because multiple states must have sufficiently low free energies relative to the unfolded state and the free-energy differences between the states must be small enough that switching can be toggled by an external input[5,6]. Recent advances in designing systems with multiple states include a transmembrane ion transporter[7] and variants of Gβ1 that dynamically exchange between two related conformations[8]; however, a method for the de novo design of modular, tunable protein systems that switch conformational states in the presence of an external input has not yet been achieved.

We set out to design de novo switchable protein systems guided by the following considerations. First, programming free-energy differences between two states is more straightforward in a system that is governed by inter- and intramolecular competition at the same site than by allosteric activation at distant sites[9–11] because many of the residue-level interactions can be similar (if not identical). Second, a stable protein framework with an extended binding surface that is available for the competing interactions is more programmable and less likely to engage in off-target interactions than a framework that becomes ordered only upon binding[12,13]. These features are described by the abstract system in Fig. 1a, which undergoes thermodynamically driven switching between binding-competent and binding-incompetent states. A latch (blue) contains a peptide sequence (orange) that can bind a target (yellow) unless the binding surface is blocked by intramolecular interactions to a cage (cyan); a key (green) that binds more tightly to the cage outcompetes the latch, which enables the peptide to bind its target (colours refer to elements in the schematic in Fig. 1a). The behaviour of such a system is governed by binding equilibrium constants for the individual subreactions (Fig. 1a): K_{open}, the dissociation of the latch from the cage; K_{LT}, the binding of the latch to its target; and K_{CK}, the binding of the cage to key. When the latch–cage interaction is too weak (red and orange curves in Fig. 1b, Supplementary Information), the system binds target in the presence of little-to-no key and the fold induction by key is low; when the latch–cage interaction is too strong (purple curve in Fig. 1b), the system only partially binds target even at high concentrations of key. The latch–cage interaction affinity that gives optimal switching (blue curve in Fig. 1b left, green curve in Fig. 1b right) is a function of the latch–target binding affinity.

Designing 'latching orthogonal cage–key proteins'

To implement the switchable system of Fig. 1a, we chose structural features that enable the tuning of the affinities of the cage–latch and cage–key interactions over a wide dynamic range. Helical interfaces are dominated by sidechain–sidechain interactions, which can more readily be tuned than the backbone hydrogen-bonding interactions between β-strands[14]. To enable fine control over the specificity and relative affinities of the cage–latch and cage–key interactions, we chose to design interfaces containing buried hydrogen-bond networks; as in DNA base-pairing, specificity can be altered by minor changes to hydrogen-bond donors and acceptors[15]. As a starting point, we connected a designed homotrimer of α-helical hairpins mediated by hydrogen-bond networks (5L6HC3_1 (ref. [15])) into two monomeric frameworks, by designing short unstructured loops between the subunits (Fig. 1c). In the five-helix framework, there is an open binding site for a sixth helix added *in trans*; this site is filled by a sixth helix *in cis* in the six-helix framework.

The five-helix (cage) and six-helix (cage plus latch) designs expressed in *Escherichia coli* were largely monomeric by size-exclusion chromatography (Extended Data Fig. 1), and remain folded up to 5 M guanidine hydrochloride (Fig. 1d). Small-angle X-ray scattering (SAXS) spectra of the connected designs are similar to that of the starting

[1]Department of Biochemistry, University of Washington, Seattle, WA, USA. [2]Institute for Protein Design, University of Washington, Seattle, WA, USA. [3]Graduate Program in Biological Physics, Structure, and Design, University of Washington, Seattle, WA, USA. [4]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA, USA. [5]Department of Bioengineering, University of California, Berkeley, Berkeley, CA, USA. [6]The UC Berkeley–UCSF Graduate Program in Bioengineering, UCSF, San Francisco, CA, USA. [7]The UC Berkeley–UCSF Graduate Program in Bioengineering, UC Berkeley, Berkeley, CA, USA. [8]Department of Chemistry, Wabash College, Crawfordsville, IN, USA. [9]Chan-Zuckerberg Biohub, San Francisco, CA, USA. [10]Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. [11]These authors contributed equally: Robert A. Langan, Scott E. Boyken, Andrew H. Ng. *e-mail: dabaker@uw.edu

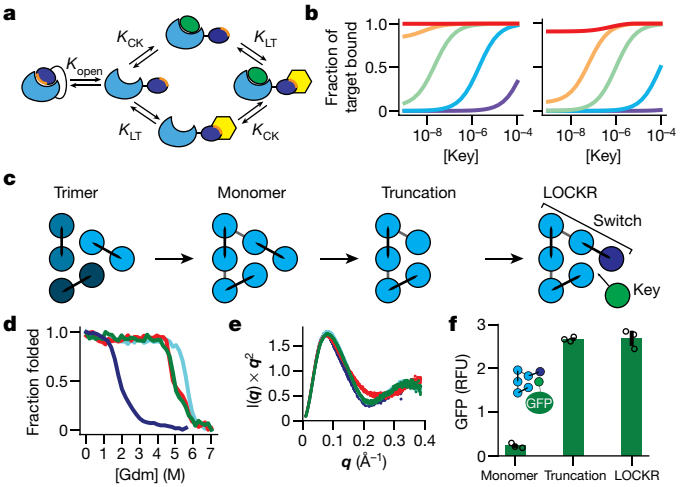


Fig. 1 | Design of the LOCKR system. a, The switch—composed of a cage (cyan) and latch (blue) with a functional motif (orange)—is in thermodynamic equilibrium between a closed state (left) and an open state that is able to bind to the key (green) and target (yellow). **b**, Numerical solutions of the model in **a** for different values of K_{LT} (left, 1 nM; right, 50 nM) and K_{open} (red, 0.1; orange, 0.001; green, 1×10^{-5}; blue, 1×10^{-7}; and purple, 1×10^{-9}), with K_{CK} fixed at 1 nM. **c**, Conversion of homotrimer 5L6HC3 to monomeric five- and six-helix frameworks by loop closure. In LOCKR (right), the double mutant V223S/I238S reduces binding affinity of latch for cage, allowing it to be displaced by the key. **d**, Guanidinium (Gdm) chloride denaturation of trimer (dark blue), monomer (cyan), truncated five-helix framework (red) and LOCKR (green), monitoring mean residue ellipticity at 222 nm; similar results were obtained in three independent experiments. **e**, SAXS Kratky plots for the monomeric frameworks are similar to that of the input trimer; the five-helix framework shows the greatest deviation from the trimer. Colours are as in **d**. **f**, Key–GFP was added to monomeric frameworks immobilized onto a plate via a hexahistidine tag; after washing, binding was measured by GFP fluorescence (mean of $n = 3$ technical replicates, error bars indicate s.d). RFU, relative fluorescence unit.

trimer and indicative of well-folded proteins[16] (Fig. 1e, Supplementary Table 1; the greater deviations for the five-helix design probably reflect the loss of a helix). The five-helix, but not the six-helix, framework bound GFP-fused sixth helix in a pull-down assay (Fig. 1f); the latter result is expected, because—if the interfaces are otherwise identical and the connecting linker unstrained—the intramolecular interaction should outcompete its intermolecular counterpart owing to the reduced entropic cost of formation of intramolecular interactions. To enable the key to outcompete the latch, we tuned K_{open} by incorporating mutations in the latch that weaken its interaction with the cage[17–19]. A cage–latch framework with two serine substitutions in the latch (V223S/I238S) bound key nearly as strongly as the five-helix cage without the latch (Fig. 1f, Extended Data Fig. 2); the two serines probably weaken the cage–latch interaction by decreasing the helical propensity of the latch and increasing the cost of desolvating the latch when it binds the cage. In the absence of key, the latch is bound to the cage as in the original monomer; the SAXS spectra of the original monomer and V223S/I238S mutant are nearly identical and closely match those of the design models (Fig. 1e, Extended Data Fig. 1), but the guanidine hydrochloride denaturation midpoint and Gibbs free energy of folding ($\Delta G_{folding}$) of the V223S/I238S mutant are more similar to the truncated five-helix design—indicating the mutations are destabilizing (Fig. 1d, e, Extended Data Fig. 1). We name such cage–latch frameworks 'switches', and name the switch–key pair 'latching orthogonal cage–key proteins' (LOCKR).

LOCKR–inducible Bim–Bcl2 binding

We reasoned that a functional peptide sequence embedded in the latch could be rendered inactive until binding with key frees the latch from the cage, and that activation could be tuned by modulating the thermodynamic parameters outlined in Fig. 1a. To install function into LOCKR, we selected as a model system the Bim–Bcl2 interaction that is central to apoptosis, and sought to cage Bim such that binding to Bcl2 occurred only in the presence of the key. We chose three Bim-related Bcl2-binding peptides[20,21] (Extended Data Fig. 3a) with a range of affinities (K_{LT}) and manipulated the base LOCKR structure to alter K_{CK} and K_{open}. The three sequences were embedded in the latch by sampling different helical registers such that residues involved in binding to Bcl2 are sequestered in the cage–latch interface (Extended Data Fig. 3b, c), optimizing for the burial of hydrophobic residues and surface exposure of polar residues (Supplementary Information). To enable sampling of a broader range of K_{open} and K_{CK} values, we took advantage of the modular nature of parametric helical bundles[2,22,23] and expanded the available interaction surface area for the cage–latch and cage–key interactions by lengthening the helices in the switch by 5, 9 or 18 residues (Extended Data Fig. 3d). Increasing cage–latch affinity (decreasing K_{open}) should make the system more 'off' in the absence of key, and extending the key to increase affinity for the cage should allow the key to better outcompete the latch once K_{open} is appropriately tuned (that is, increasing K_{open} relative to K_{CK}). A design that cages full-length Bim with the cage, latch and key each extended by 18 residues is off in the absence of the key (Fig. 2a left). The strongest inducible binding was observed with a latch that was truncated by nine residues and a full-length key (Fig. 2a right); the key buries more surface area and therefore outcompetes the latch for binding to the cage. The addition of the key activated BimSwitch by over 40-fold (Fig. 2a (right), 2b), which is comparable to—or greater than—most naturally occuring protein interactions[24–26].

According to the model in Fig. 1a, the range of concentrations of the key over which BimSwitch activates should be controllable by tuning K_{CK} and K_{LT} values. We investigated this control by using biolayer interferometry (BLI) to monitor binding to Bcl2 in response to keys of different length (and therefore different K_{CK} values) (Fig. 2b, c). With Bcl2 present on the sensor tip, and BimSwitch at 250 nM, no binding to the sensor was observed in the absence of key. As keys of different length are titrated into the solution, BimSwitch becomes activated and binds to Bcl2 on the sensor (Fig. 2c). The concentration at which activation occurs differs considerably for the different length keys: a 40-residue key provided no activation (pale green), a 45-residue key activated with a half-maximum effective concentration (EC_{50}) of 230 ± 58 nM (green), and the full length 58-residue key activated with an EC_{50} of 27.0 ± 2.8 nM (dark green, Fig. 2c).

To probe how LOCKR activation depends on K_{LT}, we used different affinity targets: Bcl2, BclB and Bak bind non-caged Bim with K_d values of 0.17 nM (ref. [21]), 20 nM (ref. [21]) and 4.2 μM (Extended Data Fig. 4a), respectively. Consistent with the model in Fig. 1a, the activation of BclB binding requires higher concentrations of the key than does the activation of Bcl2 binding, whereas Bak does not bind in this range of concentrations of the key (Extended Data Fig. 4b). Together with the results in Fig 2c, this demonstrates that the equilibria involved in activation are indeed sensitive to small changes in the binding free energies of the key to the cage and the target to the latch.

To enable independent caging and specific unlocking of different protein functions in the same cell, we sought to create orthogonal switch–key pairs by incorporating different hydrogen-bond networks at the interface between the cage–latch and key using Rosetta design[27,28] (Supplementary Information). The designs $BimLOCKR_b$ and $BimLOCKR_c$ show 22-fold and 8-fold activation with their cognate keys (Fig. 2d). The original BimLOCKR ($BimLOCKR_a$) and the $BimLOCKR_b$ and $BimLOCKR_c$ designs are mutually orthogonal; each switch is activated only by its cognate key at concentrations of up to 5 μM (Fig. 2e), which illustrates the power of the buried hydrogen-bond network approach to achieve specificity.

LOCKR–inducible protein degradation

We sought to couple switch activation to protein degradation in living cells by caging the cODC degron[29]. The caging strategy used for Bim was used to embed three variants of cODC into $switch_a$: the wild-type

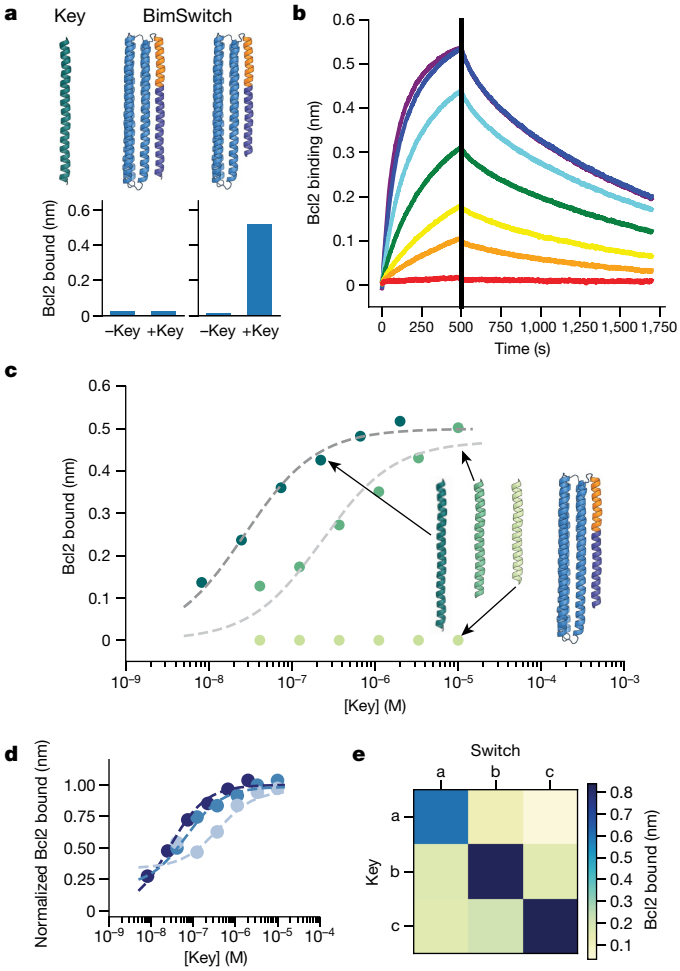


Fig. 2 | BimLOCKR design and activation. a, BLI measurement of BimSwitch (250 nM) binding to immobilized Bcl2 in the presence and absence of 5 μM key. In the lengthened BimLOCKR designs, Bim is tightly caged in the absence of the key; the introduction of the toehold (right) allows the key to outcompete the latch, which leads to Bcl2 binding. **b**, BLI measurement of key-dependent binding of 250 nM BimSwitch to Bcl2. Purple, 3 μM key; blue through cyan, green, yellow, and orange denotes successive threefold dilutions of the initial 3 μM key; red, control without key. **c**, Bcl2 binding by BimSwitch as a function of key concentration. BLI traces similar to those in **b** for the different length keys were fit to obtain predicted equilibrium sensor responses (*y* axis in **c**). BLI experiments (**b**, **c**) were repeated three times with similar results. **d**, Bcl2 binding of $BimSwitch_a$ (dark blue), $BimSwitch_b$ (blue) and $BimSwitch_c$ (light blue) designs in response to their cognate keys, measured by BLI and normalized to the maximum binding response (R_{max}). This experiment was repeated twice with similar results. **e**, Bcl2 binding in BLI experiments for each switch at 250 nM, and keys at 5 μM. Data points are the average R_{max} of two replicates.

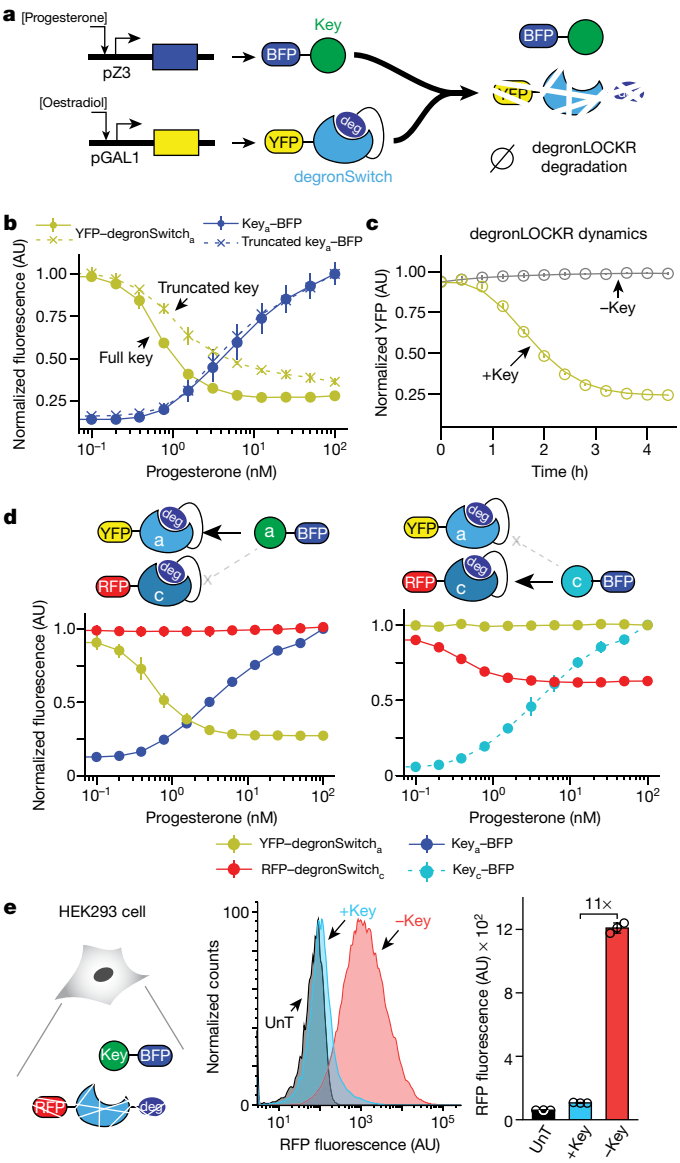


Fig. 3 | Testing functionality of degronLOCKR in live cells. a, Dual-induction system used in *S. cerevisiae* to test the function of degronLOCKR. **b**, Dose–response of YFP–$degronSwitch_a$ and key_a–BFP at 50 nM oestradiol as a function of progesterone induction; the full key is 55 residues and the truncated key is 43 residues. YFP, normalized to no progesterone; BFP, normalized to maximum progesterone. Lines connecting data are a guide for the eye. AU, arbitrary units. **c**, Dynamics of degronLOCKR degradation using an automated flow cytometry platform. Cells were grown to steady state at 50 nM oestradiol, and then induced with progesterone to express key_a–BFP at $t = 0$ h. Lines represent a moving average taken over three data points. **d**, Dose–response of orthogonal degronLOCKRs as a function of progesterone. YFP–$degronSwitch_a$ and RFP–$degronSwitch_c$ were expressed constitutively in the same cell, with expression of either key_a–BFP (left) or key_c–BFP (right) induced with progesterone. YFP–$degronSwitch_a$, RFP–$degronSwitch_c$ and either key_a–BFP or key_c–BFP were normalized to no progesterone (RFP and YFP) or maximum progesterone (BFP). Lines connecting data are a guide for the eye. **e**, Asymmetric RFP–$degronSwitch_a$ was expressed in HEK293T cells with and without key. Flow cytometry distribution of RFP fluorescence for a representative sample indicates decreased RFP expression in the presence of key. UnT, untransduced. Geometric mean of RFP expression is quantified in the bar plot. Data in all panels represent mean ± s.d. of three biological replicates.

sequence, the wild-type sequence with a proline removed (as proline destabilizes α-helices) and the central dipeptide Cys-Ala (Extended Data Fig. 5). Designs were characterized in *Saccharomyces cerevisiae* using a dual-inducible expression system[30] to independently titrate the expression of switch fused to yellow fluorescent protein (YFP) and the expression of key fused to blue fluorescent protein (BFP) (Fig. 3a). Key-induced degradation was observed with the degron caged in the switch, but not when YFP was fused to either $BimSwitch_a$ or $switch_a$ (Extended Data Fig. 5), which both lack the degron. We optimized the amount of inducible degradation by varying the latch length length to tune the K_{open} value (Extended Data Fig. 5), and found the switch with proline removed and a 12-residue shorter latch had the largest dynamic range (hereafter referred to as $degronSwitch_a$). Keys of different lengths (43 residues versus 55 residues) produced a similar dynamic range of switch activation, but a higher concentration of key was required for maximal activation when using the shorter key (Fig. 3b, Extended Data Fig. 6), as in the case of BimLOCKR in vitro

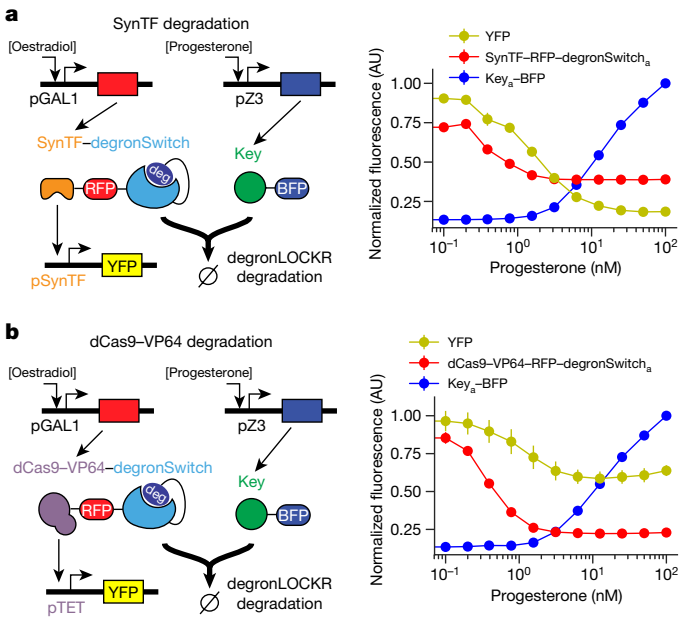


Fig. 4 | Controlling gene expression using degronLOCKR in yeast. a, Left, dual-induction system used to determine the effect of degronLOCKR$_a$ on a synthetic transcription factor (synTF). Right, dose–response of YFP, synTF–RFP–degronSwitch$_a$ and key$_a$–BFP–NLS fluorescence at 31.25 nM oestradiol as a function of progesterone induction, normalized to no progesterone (YFP and RFP) or maximum progesterone (BFP). **b**, Left, dual-induction system used to determine the effect of degronLOCKR$_a$ on a dCas9–VP64 fusion targeted to pTET promoter. Right, dose–response of YFP, dCas9–VP64–RFP–degronSwitch$_a$ and key$_a$–BFP–NLS at 31.25 nM oestradiol as a function of progesterone induction, normalized to no progesterone (YFP and RFP) or maximum progesterone (BFP). Data in all panels represent mean ± s.d. of three biological replicates. Lines connecting data are a guide for the eye.

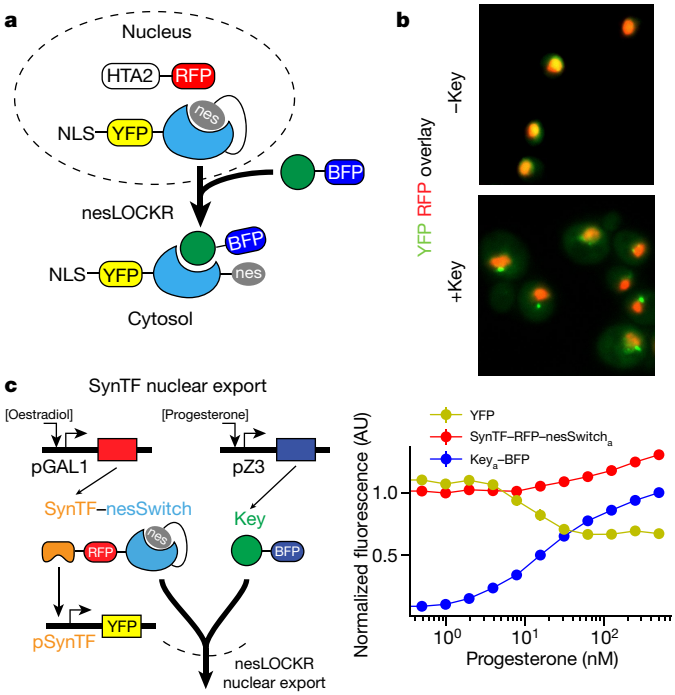


Fig. 5 | Controlling protein localization using nesLOCKR in yeast. a, Key-induced nuclear export of NLS–YFP–nesSwitch$_a$. The nucleus is marked by the histone HTA2–RFP. **b**, Fluorescence microscopy showing co-localization of NLS–YFP–nesSwitch$_a$ (green) with nuclear HTA2–RFP (red) fluorescence when no key$_a$–BFP is expressed (top), compared to the more-diffuse fluorescent signal for NLS–YFP–nesSwitch$_a$ outside of the nucleus when key$_a$–BFP is expressed (bottom). Images shown are representative of $n = 3$ biological replicates. **c**, Left, dual-induction system used to determine the effect of nesLOCKR$_a$ on a synthetic transcription factor. Right, dose–response of YFP, synTF–RFP–nesSwitch$_a$ and key$_a$–BFP at 31.25 nM oestradiol as a function of progesterone induction, normalized to no progesterone (RFP and YFP) or maximum progesterone (BFP). Data represent mean ± s.d. of three biological replicates. Lines connecting data are a guide for the eye.

(Fig. 2d); this suggests that the model in Fig. 1a holds in living cells. The fluorescence of key–BFP was independent of degronSwitch$_a$ concentration (Extended Data Fig. 6), which suggests that the key is not co-degraded with degronSwitch$_a$. We next examined the dynamics of activation, and found that the amount of YFP–degronSwitch$_a$ starts decreasing shortly after induction of key, reaching a new steady state in about three hours (Fig. 3c). Taking into account the rates of synthesis (Supplementary Information), we estimate a 24-min half-life for activated degronSwitch$_a$, which is similar to the previously measured 11–30 min[29] for the cODC degron.

We next installed the cODC degron with the proline removed into switch$_b$ and switch$_c$ to produce orthogonal degronLOCKRs to enable key-induced degradation of different proteins in the same cell. YFP fusions of degronSwitch$_a$, degronSwitch$_b$ and degronSwitch$_c$ were expressed together with key$_a$, key$_b$ and key$_c$ fused to cyan fluorescent protein (CFP). degronSwitch$_a$ and degronSwitch$_c$ were strongly activated by their cognate keys, but not by each other's key (degronLOCKR$_b$ did not activate) (Extended Data Fig. 7). To test the orthogonality of degronLOCKRs in the same cell, we constitutively co-expressed degronSwitch$_a$ and degronSwitch$_c$ fused to YFP and red fluorescent protein (RFP), respectively, and titrated expression of each key. Expression of key$_a$ led to selective degradation of YFP but not RFP, and expression of key$_c$ led to selective degradation of RFP but not YFP (Fig. 3d).

To evaluate degronLOCKR function in mammalian cells, we expressed degronSwitch$_a$ fused to mCherry RFP in human HEK293T cells, and measured RFP fluorescence in the presence and absence of key. A redesigned asymmetric degronSwitch$_a$ with an 8-residue toehold (Extended Data Fig. 8, Supplementary Information) triggered an 11-fold reduction in mean RFP fluorescence in the presence of key (Fig. 3e).

degronLOCKR control of gene expression in live cells

We sought to use degronLOCKR to modulate the intracellular concentration of a synthetic transcription factor and dCas9 in yeast. We independently expressed a zinc-finger-based synthetic transcription factor (synTF)[31] fused to RFP–degronSwitch$_a$, and a key$_a$–BFP–nuclear localization signal (NLS) fusion. Activity of the synTF was measured using YFP produced by the promoter of the synTF (pSynTF; p- prefix denotes promoter throughout) (Fig. 4a left). Induction of key triggered a 61% decrease in RFP (the transcription factor) and an 82% decrease in YFP (the transcription-factor target) (Fig. 4a right, Extended Data Fig. 9a). To investigate the generality of transcriptional control by degronLOCKR, we fused an activating dCas9–VP64 fusion[32] to RFP-degronSwitch$_a$ and targeted it to the Tet operator site on the pTET promoter with a constitutively expressed single-guide RNA to induce expression of YFP (Fig. 4b left). A 78% reduction of RFP and 41% reduction of YFP was observed upon induction of key (Fig. 4b right, Extended Data Fig. 9b). Together, these results demonstrate the modularity and functionality of degronLOCKR for controlling the stability of proteins in live cells.

LOCKR-inducible nuclear export

To investigate inducible control over nuclear localization, we caged a nuclear export sequence[33] within switch$_a$ (using the same strategy as for Bim and cODC, Extended Data Fig. 10a, b) to produce nesSwitch$_a$, and fused the resulting nesSwitch$_a$ to YFP with a nuclear localization sequence[34]. An RFP–histone fusion (HTA2) was constitutively expressed in the same yeast cells to act as a nuclear marker (Fig. 5a). YFP co-localized with RFP in the nucleus in the absence of key$_a$–BFP, but upon expression of key$_a$–BFP, the YFP fluorescence becomes more

cytosolic, probably owing to uncaging of the nuclear export signal (Fig. 5b, Extended Data Fig. 10c, d).

Next, we used nesLOCKR to control the nuclear localization of synTF to modulate its activation of the pSynTF. Using the dual-induction system, we expressed synTF–RFP–nesSwitch$_a$ and key$_a$–BFP in the same cell as a pSynTF–YFP reporter, and observed that induction of key caused a 33% decrease in YFP signal, indicating successful activation of nesLOCKR and exclusion of synTF from the nucleus (Fig. 5c, Extended Data Fig. 10e). Together, these results demonstrate our ability to cage different functional peptide motifs in live cells, which highlights the modularity and utility of LOCKR.

Conclusions

The design of tunable and generalizable protein switches is a considerable advance for de novo protein design. In the switchable LOCKR system, a designed key that is added *in trans* induces a large conformational change in a designed cage that unlocks protein function. We demonstrate the power and generality of LOCKR by caging three distinct functions: the pro-apoptotic peptide Bim binding to Bcl2, protein degradation mediated by the cODC degron, and protein localization via a nuclear export sequence. The modularity and hyperstability of de novo-designed proteins enables the tuning of switch activation over a broad dynamic range by modulating the strength of the competing cage–key and cage–latch interfaces. Moving forward, LOCKR provides a general approach for controlling function that should be applicable to a wide range of proteins and challenges in synthetic biology.

It is instructive to compare LOCKR to regulatory systems in nature that use autoinhibition, and to efforts to co-opt these systems for engineered protein switches. The activation of apoptosis by the pro-apoptotic proteins Bak and Bax can be triggered by displacement of auto-inhibitory interactions[35], analogous to activation by keys in LOCKR. Actin nucleation is modulated by N-WASP, which has an auto-inhibited actin nucleating Arp2/3 binding domain that is released upon binding to the activators Cdc42 and PIP2[36]. A number of proteins—including N-WASP—have previously been repurposed to control non-cognate functions in a switchable, inducible manner[37,38], but the LOCKR system has several advantages. First, LOCKR is a quite general platform for caging and then activating functionalities at will, ranging from the inducible activation of high-affinity protein–protein interactions to the controlled degradation or localization of an attached cargo. Second, for any functional modality, many cargoes can be regulated: here we couple key-induced LOCKR-gated degradation to fluorescent protein levels, both directly through fusion and indirectly through fusion to an activating transcription factor and kinases can be controlled in the same way[39]. Modularity and tunability in strategies that rely on repurposing natural proteins are limited by the evolved functions and ligands of the existing proteins, whereas altering the affinities of LOCKR components is tunable on the basis of simple design principles that are, in general, irrespective of the functional modality or application. Our use of a toehold for tuning helical displacement is reminiscent of DNA strand displacement technology[40,41] but—unlike approaches based on nucleic acids such as genetic toggles[42] or riboswitches[43,44] (which have largely focused on controlling transcription)—LOCKR systems can be readily integrated with the many diverse processes controlled by proteins. Viewed in this light, LOCKR brings the programmability of DNA switching technology to proteins, with the added advantages of tunability and flexibility over rewired natural protein systems, and ready interfacing with biological machinery over DNA nanotechnology[39]. More generally, the domain of sophisticated environmentally sensitive and switchable function no longer belongs exclusively to naturally occurring proteins.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1432-8.

Received: 12 October 2018; Accepted: 19 June 2019;
Published online 24 July 2019.

1. Huang, P.-S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320–327 (2016).
2. Huang, P.-S. et al. High thermodynamic stability of parametrically designed helical bundles. *Science* **346**, 481–485 (2014).
3. Brunette, T. J. et al. Exploring the repeat protein universe through computational protein design. *Nature* **528**, 580–584 (2015).
4. Rocklin, G. J. et al. Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science* **357**, 168–175 (2017).
5. Ambroggio, X. I. & Kuhlman, B. Design of protein conformational switches. *Curr. Opin. Struct. Biol.* **16**, 525–530 (2006).
6. Choi, J. H., Laurent, A. H., Hilser, V. J. & Ostermeier, M. Design of protein switches based on an ensemble model of allostery. *Nat. Commun.* **6**, 6968 (2015).
7. Joh, N. H. et al. De novo design of a transmembrane Zn^{2+}-transporting four-helix bundle. *Science* **346**, 1520–1524 (2014).
8. Davey, J. A., Damry, A. M., Goto, N. K. & Chica, R. A. Rational design of proteins that exchange on functional timescales. *Nat. Chem. Biol.* **13**, 1280–1285 (2017).
9. Ha, J.-H. & Loh, S. N. Protein conformational switches: from nature to design. *Chemistry* **18**, 7984–7999 (2012).
10. Liu, J. & Nussinov, R. Allostery: an overview of its history, concepts, methods, and applications. *PLoS Comput. Biol.* **12**, e1004966 (2016).
11. Raman, S., Taylor, N., Genuth, N., Fields, S. & Church, G. M. Engineering allostery. *Trends Genet.* **30**, 521–528 (2014).
12. Perkins, J. R., Diboun, I., Dessailly, B. H., Lees, J. G. & Orengo, C. Transient protein–protein interactions: structural, functional, and network properties. *Structure* **18**, 1233–1243 (2010).
13. Singh, G. P., Ganapathi, M. & Dash, D. Role of intrinsic disorder in transient interactions of hub proteins. *Proteins* **66**, 761–765 (2007).
14. Dou, J. et al. De novo design of a fluorescence-activating β-barrel. *Nature* **561**, 485–491 (2018).
15. Boyken, S. E. et al. De novo design of protein homo-oligomers with modular hydrogen-bond network-mediated specificity. *Science* **352**, 680–687 (2016).
16. Dyer, K. N. et al. High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. *Methods Mol. Biol.* **1091**, 245–258 (2014).
17. Fleming, P. J. & Rose, G. D. Do all backbone polar groups in proteins form hydrogen bonds? *Protein Sci.* **14**, 1911–1917 (2005).
18. Chothia, C. & Janin, J. Principles of protein–protein recognition. *Nature* **256**, 705–708 (1975).
19. Moreira, I. S., Fernandes, P. A. & Ramos, M. J. Hot spots—a review of the protein–protein interface determinant amino-acid residues. *Proteins* **68**, 803–812 (2007).
20. Delgado-Soler, L., Pinto, M., Tanaka-Gil, K. & Rubio-Martinez, J. Molecular determinants of Bim(BH3) peptide binding to pro-survival proteins. *J. Chem. Inf. Model.* **52**, 2107–2118 (2012).
21. Berger, S. et al. Computationally designed high specificity inhibitors delineate the roles of BCL2 family proteins in cancer. *eLife* **5**, e20352 (2016).
22. Grigoryan, G. & Degrado, W. F. Probing designability via a generalized model of helical bundle geometry. *J. Mol. Biol.* **405**, 1079–1100 (2011).
23. Crick, F. H. C. The Fourier transform of a coiled-coil. *Acta Crystallogr.* **6**, 685–689 (1953).
24. Ikeda, K., Watanabe, Y., Ohto, H. & Kawakami, K. Molecular interaction and synergistic activation of a promoter by Six, Eya, and Dach proteins mediated through CREB binding protein. *Mol. Cell. Biol.* **22**, 6759–6766 (2002).
25. Giesecke, A. V., Fang, R. & Joung, J. K. Synthetic protein–protein interaction domains created by shuffling Cys_2His_2 zinc-fingers. *Mol. Syst. Biol.* **2**, 2006.0011 (2006).
26. Rehtanz, M., Schmidt, H.-M., Warthorst, U. & Steger, G. Direct interaction between nucleosome assembly protein 1 and the papillomavirus E2 proteins involved in activation of transcription. *Mol. Cell. Biol.* **24**, 2153–2168 (2004).
27. Kuhlman, B. & Baker, D. Native protein sequences are close to optimal for their structures. *Proc. Natl Acad. Sci. USA* **97**, 10383–10388 (2000).
28. Leaver-Fay, A. et al. in *Methods in Enzymology* (eds Johnson, M. L. & Brand, L.) 545–574 (Academic, 2011).
29. Takeuchi, J., Chen, H., Hoyt, M. A. & Coffino, P. Structural elements of the ubiquitin-independent proteasome degron of ornithine decarboxylase. *Biochem. J.* **410**, 401–407 (2008).
30. Aranda-Díaz, A., Mace, K., Zuleta, I., Harrigan, P. & El-Samad, H. Robust synthetic circuits for two-dimensional control of gene expression in yeast. *ACS Synth. Biol.* **6**, 545–554 (2017).
31. Khalil, A. S. et al. A synthetic biology framework for programming eukaryotic transcription functions. *Cell* **150**, 647–658 (2012).
32. Perez-Pinera, P. et al. RNA-guided gene activation by CRISPR–Cas9-based transcription factors. *Nat. Methods* **10**, 973–976 (2013).
33. Güttler, T. et al. NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat. Struct. Mol. Biol.* **17**, 1367–1376 (2010).
34. Kosugi, S. et al. Six classes of nuclear localization signals specific to different binding grooves of importin α. *J. Biol. Chem.* **284**, 478–485 (2009).
35. Westphal, D., Dewson, G., Czabotar, P. E. & Kluck, R. M. Molecular biology of Bax and Bak activation and action. *Biochim. Biophys. Acta* **1813**, 521–531 (2011).
36. Prehoda, K. E., Scott, J. A., Mullins, R. D. & Lim, W. A. Integration of multiple signals through cooperative regulation of the N-WASP–Arp2/3 complex. *Science* **290**, 801–806 (2000).

37. Dueber, J. E., Yeh, B. J., Bhattacharyya, R. P. & Lim, W. A. Rewiring cell signaling: the logic and plasticity of eukaryotic protein circuitry. *Curr. Opin. Struct. Biol.* **14**, 690–699 (2004).
38. Stein, V. & Alexandrov, K. Synthetic protein switches: design principles and applications. *Trends Biotechnol.* **33**, 101–110 (2015).
39. Ng, A. H. et al. Modular and tunable biological feedback control using a de novo protein switch. *Nature* https://doi.org/10.1038/s41586-019-1425-7 (2019).
40. Zhang, D. Y. & Winfree, E. Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc.* **131**, 17303–17314 (2009).
41. Linko, V. & Dietz, H. The enabled state of DNA nanotechnology. *Curr. Opin. Biotechnol.* **24**, 555–561 (2013).
42. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
43. Wittmann, A. & Suess, B. Engineered riboswitches: expanding researchers' toolbox with synthetic RNA regulators. *FEBS Lett.* **586**, 2076–2083 (2012).
44. Isaacs, F. J., Dwyer, D. J. & Collins, J. J. RNA synthetic biology. *Nat. Biotechnol.* **24**, 545–554 (2006).

Acknowledgements R.A.L. was supported by Bruce and Jeannie Nordstrom, thanks to the Patty and Jimmy Barrier Gift for the Institute for Protein Design Directors Fund. S.E.B. is supported by a Career Award at the Scientific Interface from Burroughs Wellcome Fund. A.H.N. was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program. M.J.L. was supported by a Washington Research Foundation Innovation Postdoctoral Fellowship and a Cancer Research Institute Irvington Fellowship from the Cancer Research Institute. SAXS data were collected at the Advanced Light Source (ALS) at LBNL, supported by the following grants from NIH (P30 GM124169-01, ALS-ENABLE P30 GM124169 and S10OD018483), NCI SBDR (CA92584) and DOE-BER IDAT (DE-AC02-05CH11231). This work was supported by the Defense Advanced Research Projects Agency, contract no. HR0011-16-2-0045 to H.E.-S and by Open Philanthropy to D.B. The content and information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred. H.E.-S. is a Chan-Zuckerberg investigator. We thank L. Carter and the Protein Production Facility in the Institute for Protein Design for protein used in this study, and L. Stewart, S. Bermeo, A. Quijano Rubio, B. Basanta, M. Chevalier, A. Bonny and J. Pedro-Fonseca for help and advice.

Reviewer information *Nature* thanks Vincent J. Hilser and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Author contributions R.A.L., A.H.N. and S.E.B. contributed equally to this publication. R.A.L., S.E.B., Z.C., W.R.P.N. and D.B. conceived the idea and initial steps for designing protein switches from de novo-designed helical bundles. R.A.L. and D.B. developed the thermodynamic model and the code upon which it works. R.A.L., S.E.B. and W.R.P.N. designed and biophysically characterized LOCKR scaffolds and BimLOCKR. R.A.L. performed mutagenesis and BLI experiments. S.B. characterized Bim interactions to Bcl2 homologues, and aided with experimental design. R.A.L. performed design calculations for orthogonal LOCKR designs using code from S.E.B. and V.K.M. A.H.N. and R.A.L. conceived caging cODC. R.A.L. performed design calculations to cage cODC and tune degronLOCKR. A.H.N. conceived and contributed to all experiments with degronLOCKR. T.H.N. performed dynamic measurement of degronLOCKR. A.M.W. tested degronLOCKR in HEK293T cells. M.J.L., S.E.B. and R.A.L. performed design calculations for asymmetric LOCKR. G.D. performed experiments with degronLOCKR and dCas9. G.D. contributed to plasmid and strain construction. R.A.L., S.E.B. and M.J.L. conceived caging sequences to control subcellular location and R.A.L. performed design calculations for nesLOCKR. J.A.S. and A.H.N. performed all experiments for nesLOCKR. D.B., H.E.-S. and J.E.D. supervised research. R.A.L., S.E.B., A.H.N., H.E.-S. and D.B. wrote the manuscript, all authors edited and approved.

Competing interests : R.A.L., S.E.B., D.B., W.R.P.N. and M.J.L. have filed a provisional patent describing the design, composition and function of LOCKR switches, keys and scaffolds. R.A.L., A.H.N., S.E.B., M.J.L., D.B. and H.E.-S. have filed a provisional patent application describing the design, composition and function of degronLOCKR constructs. D.B., R.A.L., S.E.B. and M.J.L. hold equity in Lyell Immunopharma. D.B. holds equity in Sana Biotechnology.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1432-8.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1432-8.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to D.B.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

PCR mutagenesis and isothermal assembly. All primers for mutagenesis were ordered from Integrated DNA Technologies (IDT). Mutagenic primers were designed to anneal >18 bp on either side of the site for mutagenesis, with the desired mutation encoded in the primer. PCR was used to create fragments upstream and downstream of the mutation site with >20 bp overlap with the desired pET vector. The resulting amplicons were isothermally assembled into either pET21b, pET28b or pET29b restriction sites, digested with XhoI and NdeI and transformed into chemically competent *E. coli* XL1-Blue cells. Monoclonal colonies were sequenced with Sanger sequencing. Sequence-verified plasmid was purified using Qiagen miniprep kit and transformed into chemically competent *E. coli* BL21(DE3)Star, BL21(DE3)Star-pLysS cells (Invitrogen) or Lemo21(DE3) cells (NEB) for protein expression.

Synthetic gene construction. Synthetic genes were ordered from Genscript and delivered in pET 28b+, pET21b+ or pET29b+ *E. coli* expression vectors, inserted at the NdeI and XhoI sites of each vector. For pET28b+ constructs, synthesized DNA was cloned in frame with the N-terminal hexahistidine tag and thrombin cleavage site and a stop codon was added at the C terminus. For pET21b+ constructs, a stop codon was added at the C terminus such that the protein was expressed with no hexahistidine tag. For pET29b+ constructs, the synthesized DNA was cloned in frame with the C-terminal hexahistidine tag. Plasmids were transformed into chemically competent *E. coli* BL21(DE3)Star, BL21(DE3)Star-pLysS cells (Invitrogen) or Lemo21(DE3) cells (NEB) for protein expression.

Bacterial protein expression and purification. Starter cultures were grown in lysogeny broth (LB) or terrific broth II (TBII) overnight in the presence of 50 μg/ml carbenicillin (pET21b+) or 30 μg/ml (for LB) to 60 μg/ml (for TBII) kanamycin (pET28b+ and pET29b+). Starter cultures were used to inoculate 500 ml of Studier TBM-5052 autoinduction medium containing antibiotic, and grown at 37 °C for 24 h. Cells were collected by centrifugation at 4,000 rcf for 20 min at 4 °C and resuspended in lysis buffer (20 mM Tris, 300 mM NaCl, 20 mM imidazole, pH 8.0 at room temperature), then lysed by microfluidization in the presence of 1 mM PMSF. Lysates were cleared by centrifugation at 24,000 rcf for at least 30 min at 4 °C. Supernatant was applied to Ni-NTA (Qiagen) columns pre-equilibrated in lysis buffer. The column was washed twice with 15 column volumes of wash buffer (20 mM Tris, 300 mM NaCl, 40 mM imidazole, pH 8.0 at room temperature), followed by 15 column volumes of high-salt wash buffer (20 mM Tris, 1 M NaCl, 40 mM imidazole, pH 8.0 at room temperature), and then 15 column volumes of wash buffer. Protein was eluted with 20 mM Tris, 300 mM NaCl, 250 mM imidazole, pH 8.0 at room temperature. Proteins were further purified by gel filtration using FPLC and a Superdex 75 Increase 10/300 GL (GE) size-exclusion column, pooling fractions that contained monomeric protein.

Size-exclusion chromatography and multi-angle light scattering. Size-exclusion chromatography and multi-angle light scattering (SEC-MALS) experiments used a Superdex 75 Increase 10/300 GL (GE) size-exclusion column connected to a miniDAWN TREOS multi-angle static light scattering instrument and an Optilab T-rEX (refractometer with extended range) detector (Wyatt Technology). Protein samples were injected at concentrations of 3–5 mg/ml in TBS (pH 8.0). Data were analysed using ASTRATM (Wyatt Technologies) software to estimate the weight-average molar weight of eluted species, as well as the number-average molar weight to assess monodispersity by polydispersity index (obtained by dividing the weight-average molar weight of eluted species by the number-average molar weight).

Circular dichroism measurements. Circular dichroism wavelength scans (260 to 195 nm) and temperature melts (25 to 95 °C) were measured using an AVIV model 420 circular dichroism spectrometer. Temperature melts monitored absorption signal at 222 nm and were carried out at a heating rate of 4 °C/min. Protein samples were at 0.3 mg/ml in PBS, pH 7.4, in a 0.1-cm cuvette. Guanidinium chloride (GdmCl) titrations were performed on the same spectrometer with an automated titration apparatus in PBS, pH 7.4 at 25 °C, monitored at 222 nm with protein sample at 0.03 mg/ml in a 1-cm cuvette with stir bar. Each titration consisted of at least 40 evenly distributed concentration points with a 1-min mixing time for each step. Titrant solution consisted of the same concentration of protein in PBS + GdmCl. GdmCl concentration was determined by refractive index.

SAXS. Samples were exchanged into SAXS buffer (20 mM Tris, 150 mM NaCl, 2% glycerol, pH 8.0 at room temperature) via gel filtration. Scattering measurements were performed at the SIBYLS 12.3.1 beamline at the Advanced Light Source. The X-ray wavelength (λ) was 1.27 Å and the sample-to-detector distance of the Mar165 detector was 1.5 m, corresponding to a scattering vector $\boldsymbol{q}$ ($\boldsymbol{q} = 4\pi \times \sin(\theta/\lambda)$, in which 2θ is the scattering angle) range of 0.01 to 0.59 $Å^{-1}$. Datasets were collected using 34 0.2-s exposures over a period of 7 s at 11 keV with protein at a concentration of 6 mg/ml. Data were also collected at a concentration of 3 mg/ml to determine concentration-dependence; all data presented in this Article were collected at the higher concentration, as no concentration-dependent aggregation was observed. Data from 32 exposures were averaged separately over the Gunier, Parod and wide-$\boldsymbol{q}$ regions depending on signal quality over each region and frame. The averages were analysed using the ScÅtter software package to analyse data and report statistics (Supplementary Data Table 1). FoXS was used to compare design models to experimental scattering profiles and calculate quality-of-fit (X) values. The hexahistidine tags and thrombin cleavage sites on the N termini of LOCKR proteins were modelled using Rosetta Remodel, such that the design sequence matched that of the experimentally tested protein. To capture conformational flexibility of these residues, 100 independent models were generated and clustered, and the cluster centre of the largest cluster was selected as a representative model for FoXS fitting without bias.

GFP pulldown assay. His-tagged LOCKR was expressed per the protocol described in 'Bacterial protein expression and purification' from pET28b+ while the key was expressed fused to superfolder GFP (sfGFP) without a His-tag in pET21b+. The His-tagged switch was purified to completion and dialysed into TBS (20 mM Tris, 150 mM NaCl, pH 8.0 at room temperature); the key–GFP remained as lysate for this assay. One hundred microlitres switch at >1 μM was applied to a 96-well black Pierce Nickel Coated Plate (Thermo Fisher) and incubated at room temperature for 1 h. Sample was discarded from the plate and washed 3× with 200 μL TBST (TBS + 0.05% Tween-20). One hundred microlitres of lysate containing key–GFP was added to each well and incubated at room temperature for 1 h. Sample was discarded from the plate and washed 3× with 200 μl TBST (TBS + 0.05% Tween-20). The plate was washed 1× with TBS, and 100 μl of TBS was added to each well. sfGFP fluorescence was measured on a Molecular Devices SpectraMax M5 plate reader or BioTek Synergy Neo2 plate reader; fluorescence was measured at 485-nm excitation and 530-nm emission, with a bandwidth of 20 nm for excitation and emission.

BLI. BLI measurements were made on an Octet RED96 System (ForteBio) with streptavidin (SA) coated biosensors and all analysis was performed within ForteBio Data Analysis Software version 9.0.0.10. Assays were performed with protein diluted into HBS-EP+ buffer from GE (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% v/v surfactant P20, 0.5% non-fat dry milk, pH 7.4 at room temperature). Biotinylated Bcl2 was loaded onto the streptavidin tips to a threshold of 0.5 nm programmed into the protocol of the machine. The baseline was obtained by dipping the loaded biosensors into HBS-EP+ buffer; association kinetics were observed by dipping into wells containing defined concentrations of switch, and dissociation kinetics of key were then observed by dipping into the buffer used to obtain the baseline. Kinetic constants and response at equilibrium were computed by fitting a 1:1 binding model.

Construction of DNA circuits. Hierarchical golden gate assembly was used to assemble plasmids for yeast strain construction using a previously published method[45]. Individual parts had their BsaI, BsmBI and NotI cut sites removed to facilitate downstream assembly and linearization. Parts were either generated via PCR or purchased as gBlocks from IDT. These parts were assembled into transcriptional units (promoter–gene–terminator) on cassette plasmids. These cassettes were then assembled to form multi-gene plasmids for insertion into the yeast genome.

Yeast strains and growth medium. The base *S. cerevisiae* strain used in all experiments was BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0). All yeast cultures were grown in YPD medium (10 g/l bacto-yeast extract, 20 g/l bacto peptone, 20 g/l dextrose) or synthetic complete medium (SDC) (6.7 g/l bacto-yeast nitrogen base without amino acids, 2 g/l complete supplement amino acid mix, 20 g/l dextrose). Selection of auxotrophic markers (Ura3, Leu2 and/or His3) was performed on synthetic complete medium with the appropriate dropout amino acid mix. All yeast strains used in this work are listed in Supplementary Tables 4, 5.

Oestradiol and progesterone induction. Yeast strains were grown overnight by picking a single colony from a plate into YPD medium. Saturated culture was diluted 1:500 in fresh SDC and aliquoted into individual wells of a 2-ml 96-well storage block (Corning) for a three-hour outgrowth at 30 °C and 900 rpm in a Multitron shaker (Infors HT). Oestradiol (Sigma-Aldrich) and progesterone (Fisher Scientific) were prepared at a 10× concentration by making the appropriate dilutions into SDC from a 3.6-mM oestradiol and 3.2-mM progesterone stock solution. After the three-hour outgrowth, 50 μl of oestradiol and progesterone inducer were added to the 96-well block in the appropriate combinations and the block was returned to the shaker. Flow cytometry measurement was performed after 6 h of incubation for all experiments, except for those involving synTF or dCas9, which were allowed to incubate for 12 h owing to the additional transcriptional step in the system.

Mammalian cell culture and lentiviral transduction. HEK293T cells (from ATCC CRL-3216) were maintained in DMEM (Dulbecco's Modified Eagle Medium, Gibco) supplemented with 10% fetal calf serum (SAFC) and passaged approximately every 3 days. The cell line tested negative for mycoplasma contamination.

Pantropic VSV-G pseudotyped lentivirus was produced via transfection of Lenti-X 293T cells (Clontech #11131D) with a custom pHR'SIN:CSW transgene expression vector and the viral packaging plasmids pCMVdR8.91 and pMD2.G using Fugene HD (Promega). At 48 h, viral supernatant was collected and the HEK293T cells were exposed to the virus for 24 h. Transductions were performed in triplicate. All plasmids used for HEK293T experiments are listed in Supplementary Table 6.

Description of automated flow cytometry and continuous culture system. *Hardware.* We adapted an existing automated experimental platform[46] to perform variable-concentration small-molecule induction and long-term culturing. Yeast cultures were grown in 50-ml optically clear conical tubes (Falcon) that were held in 8 custom temperature-controlled, magnetically stirred chambers. Liquid handling was accomplished using a 14 position stream selector (VICI Cheminert) and 2 syringe pumps (Cavro XCalibur Pump, TECAN) of a BD High-Throughput Sampler (HTS). Commands to the HTS were controlled using LABVIEW 2013. This setup enabled periodic sampling and dilution of individual cultures. Each sampling period consisted of three main steps: (1) send the sample to the flow cytometer for measurement, (2) extract the culture and send to waste and (3) replenish the culture with fresh medium at the desired hormone concentration. Each sampling period can be designated to either induce cultures to a new, higher hormone concentration or to maintain the desired hormone concentration. A sampling frequency of 24 min and a dilution volume of 3 ml were used.

Yeast culture. Yeast strains were grown overnight by picking a single colony from a plate into YPD medium. Saturated culture was diluted 1:200 into fresh SDC. Cultures were grown for 2 h in glass tubes at 30 °C and 250 rpm in a Innova 44 shaker (New Brunswick). Cultures were then diluted to optical density at 600 nm of 0.01 in fresh SDC and aliquoted into individual 50-ml optically clear conical tubes (Falcon) at a total volume of 30 ml YPD. Another one-hour outgrowth was performed in bioreactors with magnetically controlled stir bars at 30 °C. All SDC medium was supplemented with 5,000 U/ml penicillin streptomycin (Thermo Fisher).

Oestradiol and progesterone induction to test degronLOCKR dynamics. A 1× concentration was determined by the highest-desired hormone concentration at which to test strains (30 nM oestradiol and 50 nM progesterone, respectively). A solution of oestradiol and SDC medium was created at a 10× concentration to bring pre-induced cultures to a desired concentration in one sampling period. A second solution of progesterone and SDC medium was created at a 10× concentration to induce expression of key after degronSwitch–YFP expression reached steady state. SDC medium was prepared at three different concentrations of hormone: (1) 10× oestradiol and no progesterone, (2) 1× oestradiol and no progesterone, (3) 1× oestradiol and 10× progesterone, and (4) 1× oestradiol and 1× progesterone. After a one-hour outgrowth in bioreactors ($t = -6$ h), the first induction was performed to achieve oestradiol concentration by extracting 3 ml from all cultures and replenishing with concentration (1). After oestradiol induction, sampling proceeded as described in 'Hardware'. All sampling periods following the first induction time point included sending a sample to the cytometer for measurement, extracting 3 ml from all cultures, and replenishing cultures with concentration (2). At the second induction time point ($t = 0$ h), cultures were induced with concentration (3) to activate expression of key. This induction was followed by the same procedure as the first induction, except that hormone concentrations were maintained at concentration (4). Controls (no activated expression of key) did not undergo a second induction, and instead continued to be replenished by concentration (2).

Flow cytometry. *Yeast experiments.* Analysis of fluorescent protein expression was performed using a BD LSRII flow cytometer (BD Biosciences) equipped with a high-throughput sampler. Yeast cultures were diluted in TE before running through the instrument to obtain an acceptable density of cells. YFP (Venus) fluorescence was measured using the FITC channel, RFP (mScarlet) was measured using the PE–Texas Red channel and BFP (mTagBFP2) was measured using the DAPI channel. For steady-state measurements, 5,000–10,000 events were collected per sample. For dynamic measurements, 2,000–10,000 events were collected per sample. Fluorescence values were calculated as the height measurement for the appropriate channel and normalized to cell size by dividing by side scatter (SSC-H). All analysis of yeast flow-cytometry data was performed in Python 2.7 using the package FlowCytometryTools v.0.5.0 and custom scripts.

HEK293T experiments. Analysis of fluorescent protein expression was performed using a BD Fortessa flow cytometer (BD Biosciences) equipped with a high-throughput sampler. Cells were collected and washed twice in PBS before running through the instrument in PBS + 5% FBS. RFP (mCherry) fluorescence was measured using the PE–CF594 channel and BFP (tagBFP) was measured using the BV 421 channel. Fifty thousand events were collected per sample. Live cells were gated according to FSC-A and SSC-A, and singlets were gated according to SSC-A and SSC-H. Analysis of HEK293T flow-cytometry data was performed using FlowJo v10.

Fluorescence microscopy. Saturated culture was diluted 1:100 in fresh SC medium followed by a 3-h outgrowth at 30 °C with shaking at 700 rpm in a Multitron shaker (Infors HT). Oestradiol (Sigma-Aldrich) and progesterone (Fisher Scientific) were prepared at a 20× concentration by making the appropriate dilutions into SC medium from a 3.6-mM oestradiol and 3.2-mM progesterone stock solution. Cells were induced with oestradiol and/or progesterone at a final concentration of 200 μM and 250 μM, respectively. After 8 h of growth, cells were resuspended in 1× PBS and imaged on a Zeiss Axio Observer Z1 microscope with X-Cite Series 120 fluorescent lamp and Hamamatsu Orca-Flash 4.0 Digital Camera.

Structural visualization and figures. All structural images for figures were generated using PyMOL[47].

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

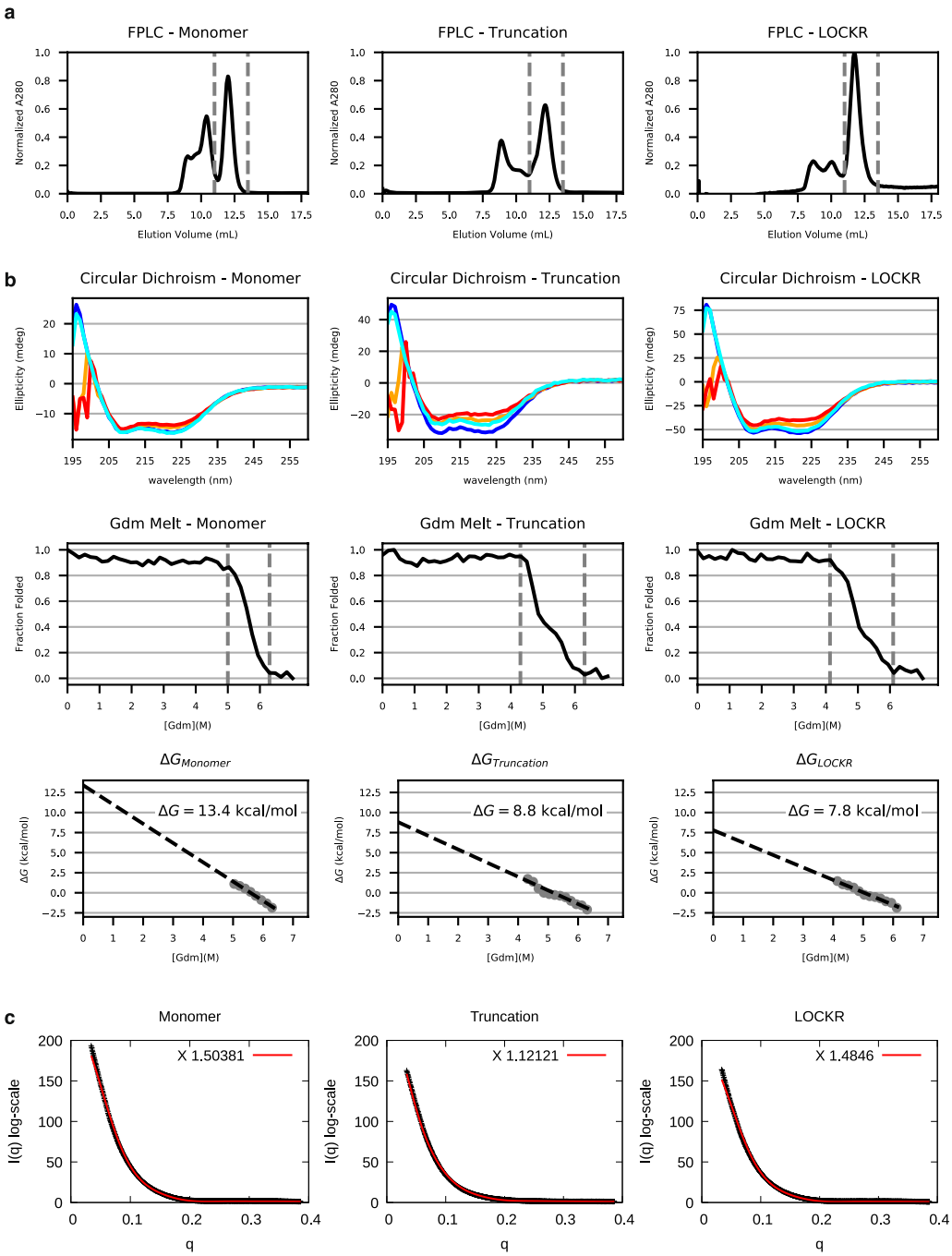
Data availability

All data that support the findings of this study are available within the Article and its Supplementary Information. The original data that support the findings are available from the corresponding author upon reasonable request. Plasmids that encode the LOCKR scaffolds (non-functional switches and keys), BimLOCKR, degronLOCKR and nesLOCKR can be found on Addgene (plasmids 127416-127424, 127200–127206 and 127246).

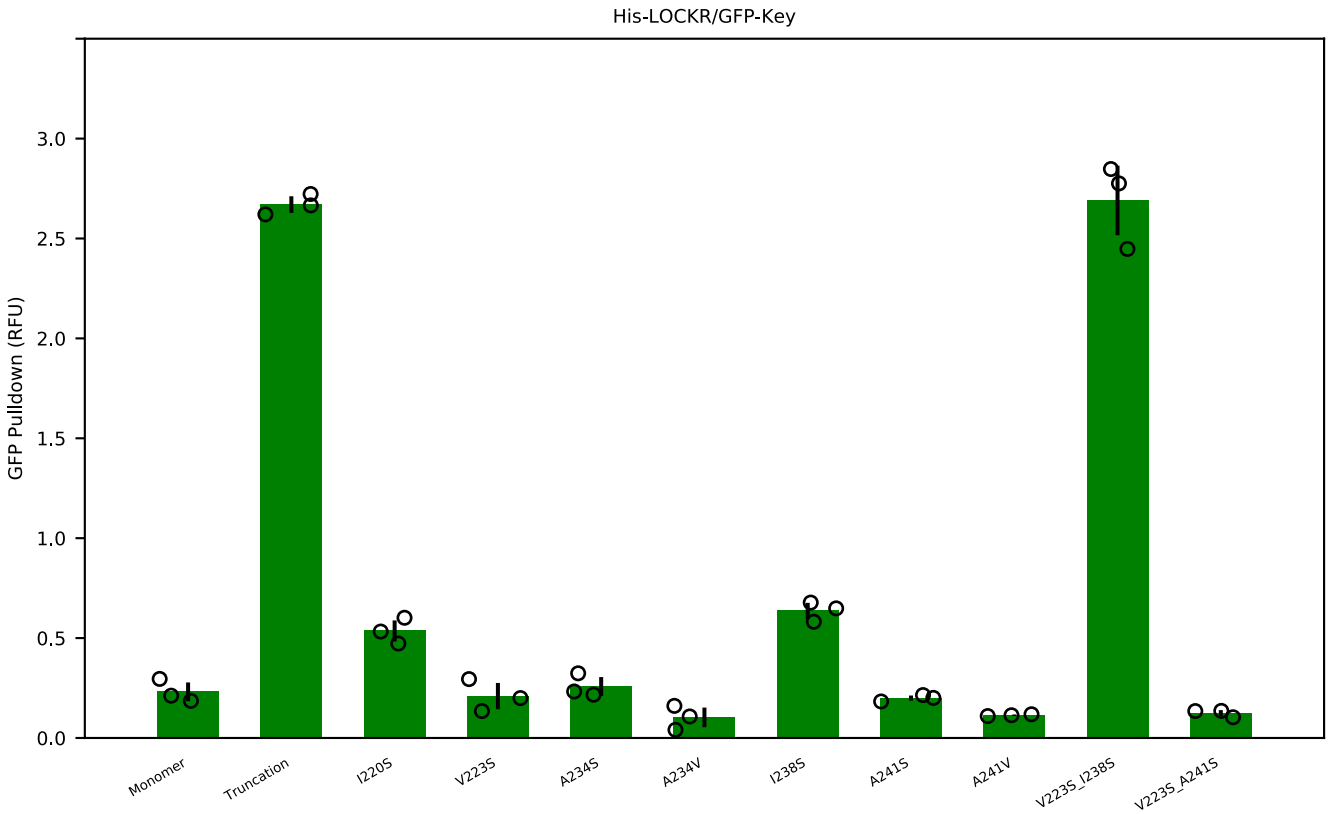
Code availability

Python scripts, bash scripts, and Rosetta Design XML files are available for download at https://github.com/BobbyLangan/DeNovoDesignofBioactiveProteinSwitches.

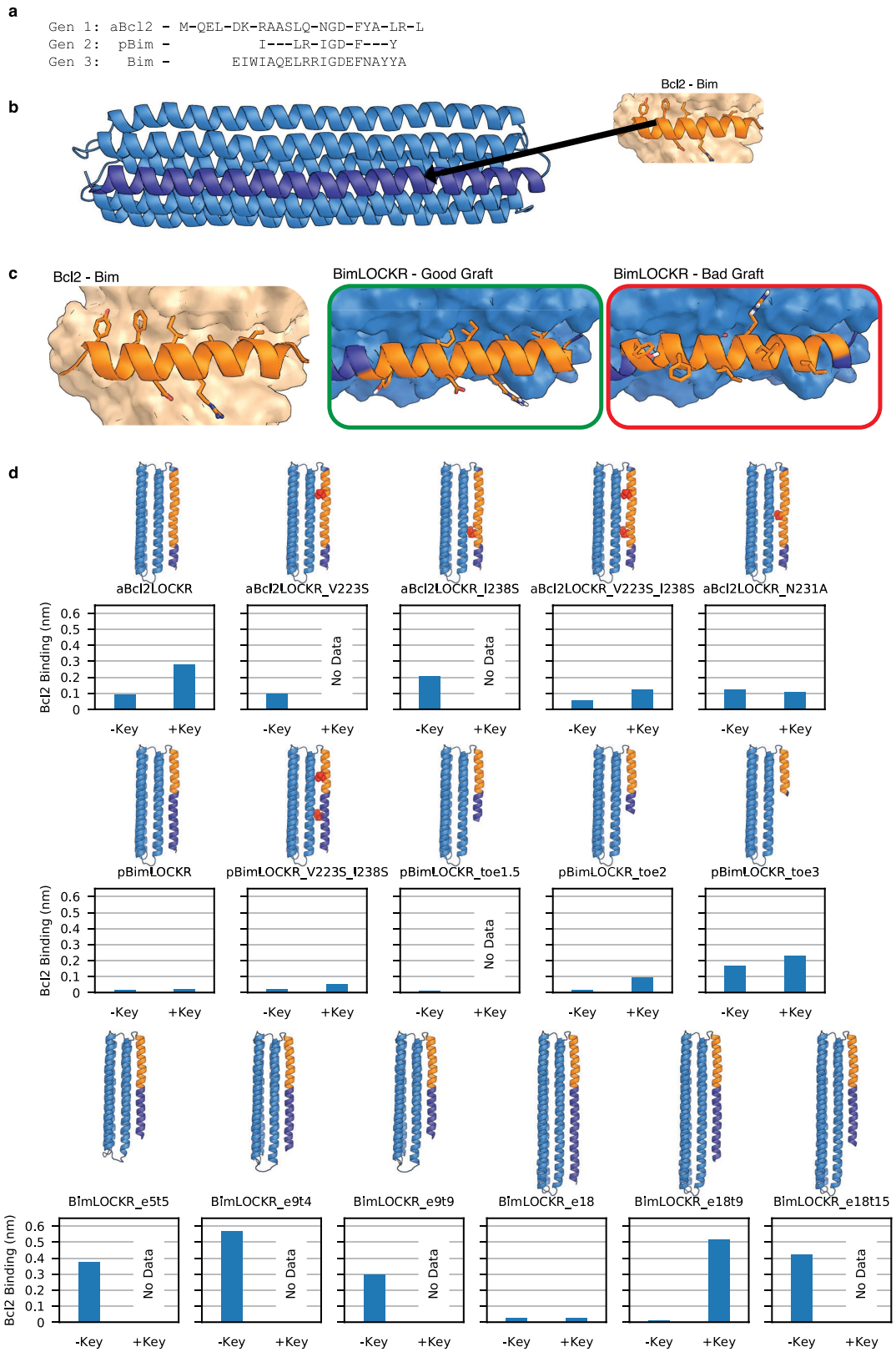
45. Lee, M. E., DeLoache, W. C., Cervantes, B. & Dueber, J. E. A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synth. Biol.* **4**, 975–986 (2015).
46. Harrigan, P., Madhani, H. & El-Samad, H. Real time genetic compensation operationally defines the dynamic demands of feedback control. Preprint at https://www.biorxiv.org/content/10.1101/244020v1 (2018).
47. DeLano, W. L. The PyMOL Molecular Graphics System. http://www.pymol.org (2002).



Extended Data Fig. 1 | Biophysical data from LOCKR design. a, Size-exclusion chromatography for the monomer, truncation and LOCKR designs on Superdex 75. Peaks indicated by vertical dashed lines represent monomeric protein used in downstream characterization and functional assays. Size-exclusion chromatography was repeated three times with similar results. **b**, Circular dichroism spectroscopy to determine protein stability upon heating and treatment with the chemical denaturant guanidinium chloride. Top, full wavescan at 25 °C (blue), 75 °C (orange) and 95 °C (red), then cooled to 25 °C (cyan). Middle, guanidinium chloride melts (also shown overlapped in Fig. 1d). Bottom, fraction of folded protein was converted to the equilibrium constant, and then to the conformational stability for protein unfolding ($\Delta G_{\text{unfolding}}$) value. The linear unfolding region, marked by vertical lines in the middle panels, was fit to determine the $\Delta G_{\text{folding}}$ for each design. The experiment was repeated four times with similar results. **c**, SAXS spectra (black; referenced in Fig. 1e) fit to the Rosetta design models (red) using FoXS with χ-values referenced in the top right.

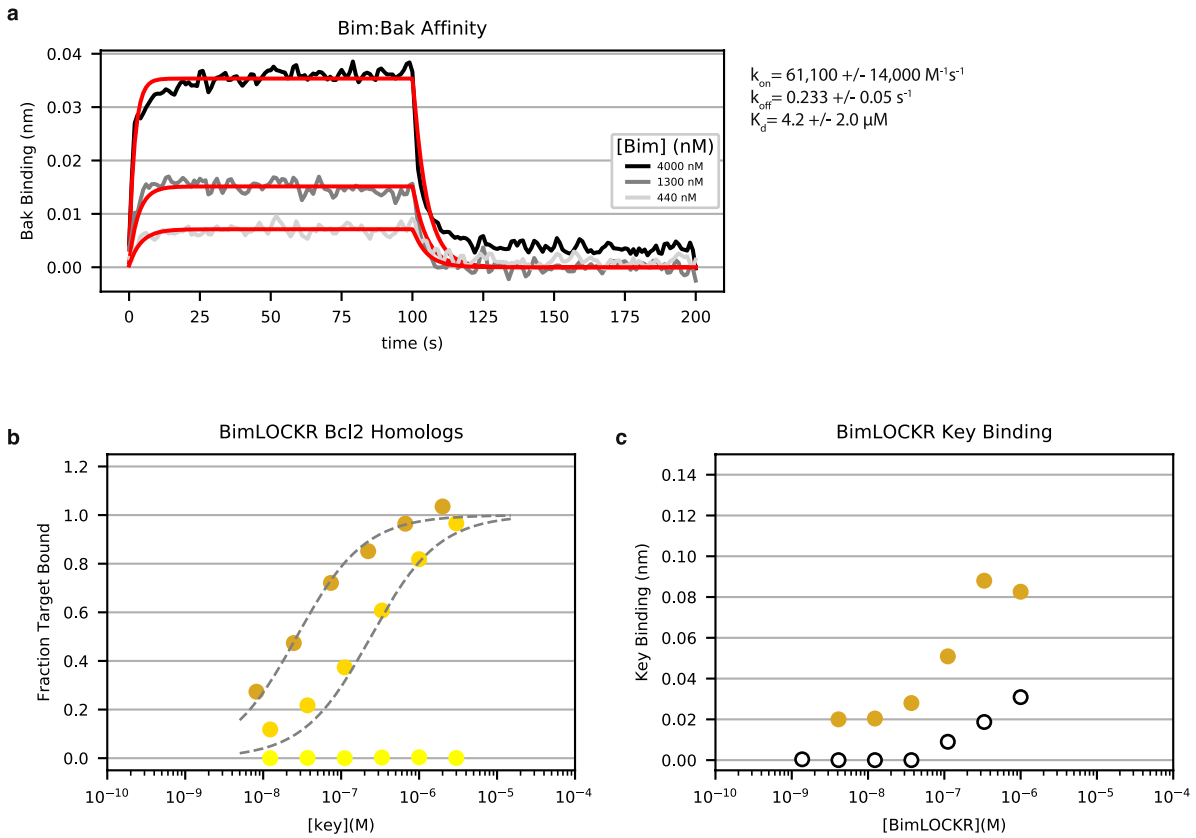


Extended Data Fig. 2 | GFP pulldown assay finds mutations for LOCKR. Different putative LOCKR constructs were adhered via the 6×His tag to a Ni-coated 96-well plate, key–GFP was applied and excess was washed away. The resulting mean fluorescence values represent key–GFP bound to LOCKR constructs. The truncation was used as a positive control, because the key binds to the open interface. The monomer was used as a negative control, because it does not bind the key. Error bars represent the s.d. of three technical replicates (key–GFP was not purified from bacterial lysate. which led to minor technical variability).



Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Caging Bim-related sequences. a, Three Bcl2-binding sequences were grafted onto the latch. aBcl2 is a single helix from a designed Bcl2 binder (RCSB Protein Data Bank code (PBD) 5JSN) in which non-Bcl2-interacting residues were reverted back to the standard LOCKR latch sequence (shown as dashes). pBim is the partial Bim sequence in which only Bcl2-interacting residues are grafted onto the latch. Bim is the full consensus sequence of the BH3 domain. **b**, LOCKR (left) with the latch in dark blue. The helical Bim sequence is taken from the Bim–Bcl2 interaction and grafted onto the latch. **c**, Left, Bcl2 (tan) binding to Bim (orange) from PDB 2MV6, with pBim residues shown as sticks. Centre, a well-caged graft in which important binding residues are caged. Right, a poor graft in which Bcl2-binding residues are exposed and polar surface residues are against the cage interface. **d**, Tuning BimLOCKR. aBcl2, pBim and Bim were caged to varying degrees of success. Early versions of the switch (with aBcl2 and pBim) did not efficiently cage Bcl2 binding in the off state. These version also bound the key only weakly, which led to a small dynamic range. The cage and key were extended by 5, 9 and 18 residues in an attempt to provide a larger interface to tightly hold the latch in the off state, and to provide a larger interface for binding of key to increase the dynamic range of activation. Mutations on the latch (identified in Extended Data Fig. 2) and providing toeholds for binding of key were the two strategies used to tune the switch. In graphs, 'off' refers to 250–310-nM switch and an absence of key, whereas 'on' refers to excess key added. The height of the bar graph shows the equilibrium response (R_{eq}) as measured by BLI.

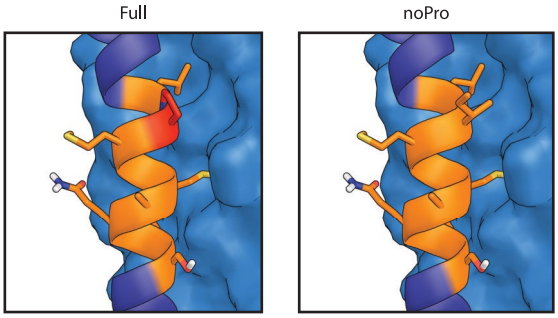


Extended Data Fig. 4 | Validation of model. This is a validation of the model shown in Fig. 1a. **a**, Measurement of Bim–Bak affinity. BLI at three concentrations gave the on and off rates for Bim–Bak binding, which yielded the constants shown on right. Mean shown with s.d. of four technical replicates (to account for variability in drift on the BLI instrument). **b**, BLI measurement of $BimLOCKR_a$ (400 nM) binding to Bcl2 (gold), BclB (yellow) and Bak (lighter yellow, BimLOCKR at 1 µM) as key is added to the solution. Data are normalized, owing to differences in R_{max} for Bcl2 and BclB on the tip. **c**, BLI measurement of $BimLOCKR_a$ binding to key_a immobilized on the tip. Open circles, no Bcl2 present; gold points, Bcl2 present at 500 nM.

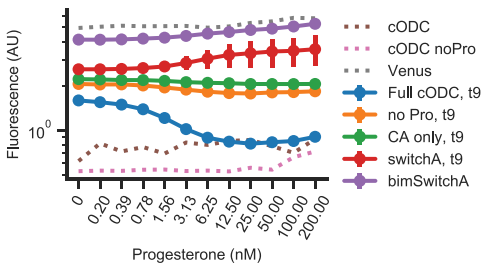
a

```
cODC Full:     LPMSCAQES
cODC noPro:    L-MSCAQES
cODC CA_only: ----CA---
```

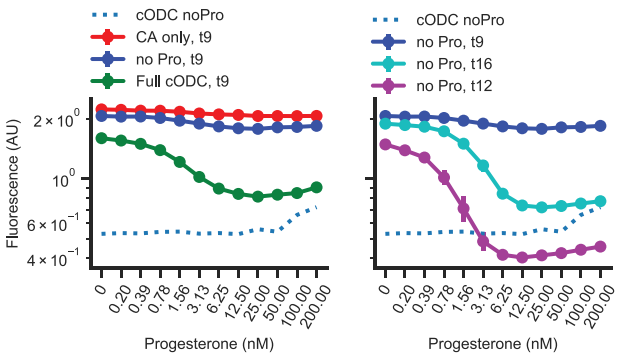
b



c

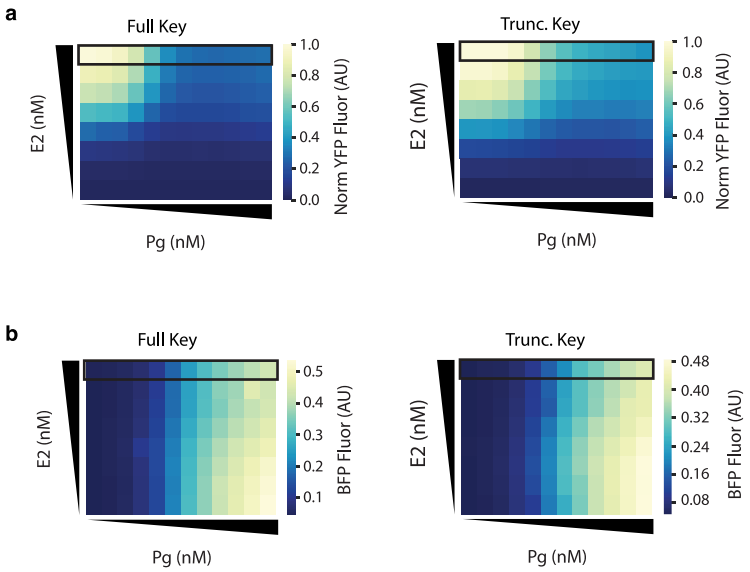


d

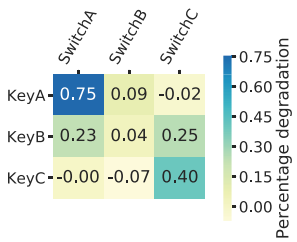


Extended Data Fig. 5 | See next page for caption.

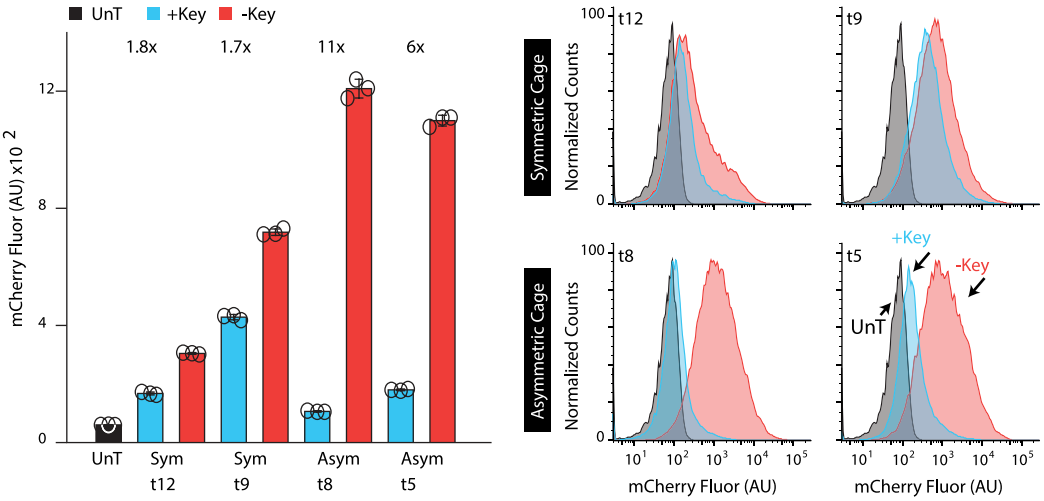
Extended Data Fig. 5 | Caging cODC sequences. a, Three variations of the cODC degron to cage. Variations meant to tune the K_{open} value, by removing the destabilizing proline (no Pro) and minimizing mutations to the latch (Cys-Ala (CA) only). **b**, Predicted models of the full and no-Pro cODC sequences (orange) threaded onto the latch (dark blue). Thread position were chosen such that the cysteine residue needed for degradation is sequestered against the cage (light blue). Proline highlighted in red in the full cODC was mutated to an isoleucine in the no-Pro variant. **c**, Comparing the stability of YFP fused to cODC variants caged in $switch_a$ to an empty $switch_a$, and to $BimSwitch_a$. The dual-inducible system from Fig. 3a was used to express the various YFP–$switch_a$ fusions (solid lines and dots) via pGAL1 and oestradiol, and key_a–BFP via pZ3 and progesterone. YFP (Venus) alone, YFP fused to the wild-type cODC (cODC) or YFP fused to the proline-removed cODC (cODC no Pro), were also expressed using pGal1 and oestradiol (dashed lines). Cells were induced with a saturating dose of oestradiol (50 nM) and progesterone was titrated in from 0–200 nM. Fluorescence was measured at steady state using a flow cytometer; data represent mean ± s.d. of three biological replicates. Lines connecting data are a guide for the eye. A moderate decrease in YFP fluorescence was observed as a function of progesterone for the full cODC variant, whereas only a small decrease was observed for the proline-removed and Cys-Ala only. No decrease in fluorescence was observed as a function of induction of key for YFP alone, empty $switch_a$ or $BimSwitch_a$. **d**, Tuning toehold lengths of $degronLOCKR_a$. The dual-inducible system from Fig. 3a was used to express the various YFP–$switch_a$ fusions via pGal1 and oestradiol, and key_a–BFP via pZ3 and progesterone. YFP fused to the proline-removed cODC (cODC no Pro) was also expressed using pGal1 and oestradiol (dashed line). Cells were induced with a saturating dose of oestradiol (50 nM) and progesterone was titrated in from 0–200 nM. Fluorescence was measured at steady state using a flow cytometer; data represent mean ± s.d. of three biological replicates. Lines connecting data are a guide for the eye. Left, cODC variants alone to show the dynamic range of full cODC. Right, extending toehold on proline-removed version from 9 to 12 and 16 amino acids. Proline-removed cODC with 12-amino-acid toehold shows the greatest dynamic range of all the switches that we tested.



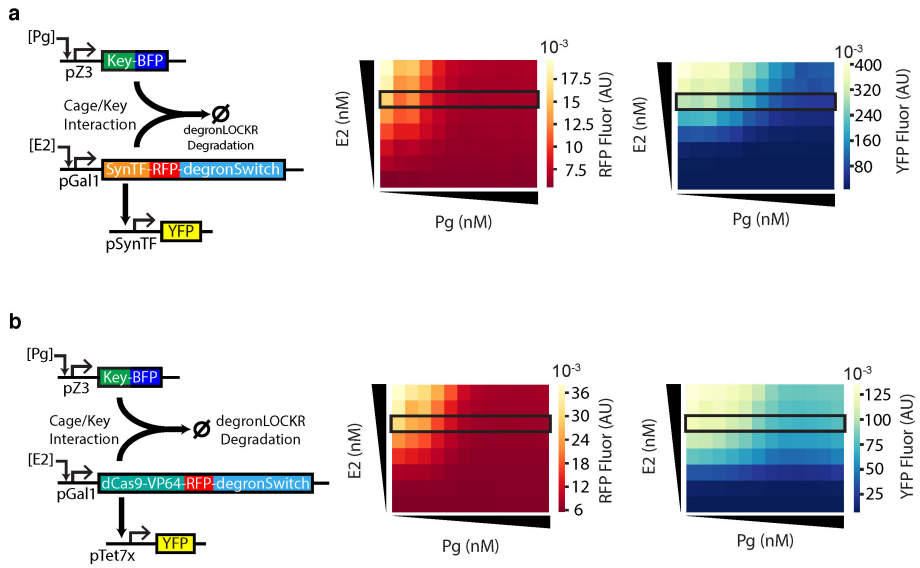
Extended Data Fig. 6 | YFP and BFP expression. a, **b**, YFP (**a**) and BFP (**b**) expression, corresponding to Fig. 3b. We used 0–50 nM oestradiol and 0–200 nM progesterone to induce expression of YFP–degronSwitch$_a$ and key$_a$ (full-length or truncated)–BFP, respectively. Fluorescence was measured at steady state using a flow cytometer. Heat maps depict mean fluorescence and are a representative sample of three biological replicates. The oestradiol dose (50 nM) depicted in Fig. 3b is indicated with the black rectangle on the heat maps. YFP fluorescence was normalized to the maximum fluorescence (50 nM oestradiol or 0 nM progesterone). BFP expression was not dependent on expression of the switch, which suggests that the key does not co-degrade with the switch.



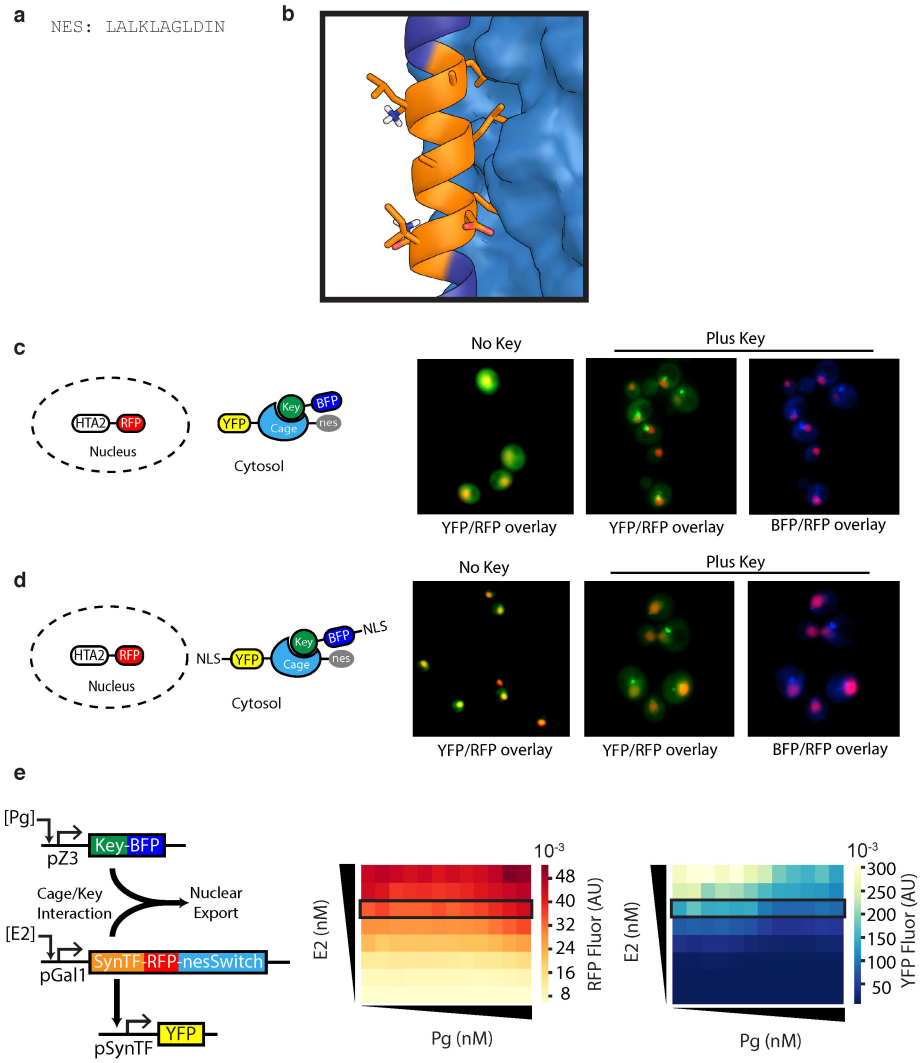
Extended Data Fig. 7 | degronLOCKR$_{a–d}$ orthogonality. All combinations of pTDH3–YFP–degronSwitch and pTDH3–key–CFP were tested. Fluorescence was measured at steady state using a flow cytometer. YFP fluorescence was averaged across three biological replicates. The percentage degradation was calculated by subtracting the mean YFP–degronSwitch fluorescence with the given key–CFP coexpressed from the YFP–degronSwitch fluorescence without any key expressed, and normalizing by the YFP–degronSwitch fluorescence without any key expressed. degronSwitch$_a$ is activated strongly by key$_a$ and activated weakly by key$_b$. degronSwitch$_c$ is activated strongly by key$_c$ and activated weakly by key$_b$. Because degronSwitch$_a$ and degronSwitch$_c$ are not activated by key$_c$ and key$_a$, respectively, we consider these two to be an orthogonal pair.



Extended Data Fig. 8 | Comparison of variants of degronSwitch in HEK293T cells. Fluorescence of variants of RFP–degronSwitch in the presence and absence of key–BFP were measured using flow cytometry. The original symmetric design was compared against an asymmetric design. Two toehold lengths were tested for each variant. Data in the bar graph represent the geometric mean ± s.d. of three biological replicates. Histograms are depicted for a representative sample. Asymmetric cage with an 8-residue toehold (signified by t8) demonstrates the largest dynamic range.



Extended Data Fig. 9 | YFP and RFP expression for synTF and dCas9–VP64. a, **b**, Assay for synTF (**a**) and dCas9–VP64 (**b**), corresponding to Fig. 4, as a function of oestradiol (E2) (0–125 nM) and progesterone (Pg) (0–100 nM). YFP fluorescence represents the transcriptional output of either synTF or dCas9–VP64 and RFP fluorescence represents fluorescence of either synTF or dCas9–VP64. Fluorescence was measured at steady state using flow cytometry. Heat maps depict mean fluorescence, and are a representative sample of three biological replicates. The oestradiol dose (31.25 nM) depicted in Fig. 5 is indicated with the black rectangle on the heat maps.



Extended Data Fig. 10 | Design and characterization of nesLOCKR. **a**, Nuclear export sequence used in this Article. **b**, The nuclear export sequence (orange) caged on the helical latch (dark blue, cartoon) with hydrophobic residues sequestered against the cage (light blue, surface). **c**, Left, schematic of cytosolic YFP–nesSwitch$_a$ and key–BFP with nuclear marker HTA2–RFP. Right, YFP fluorescence shows the expected cytosolic distribution when YFP–nesSwitch$_a$ is expressed with no NLS (left), but punctae of YFP fluorescence are observed when both YFP–nesSwitch$_a$ and key–BFP are expressed in the cytosol—which we assume is due to aggregation of the nesSwitch$_a$. Key–BFP fluorescence is co-localized with YFP–nesSwitch$_a$ fluorescence. **d**, Left, schematic of NLS–YFP–nesSwitch$_a$ with key–BFP–NLS, and with nuclear marker HTA2–RFP. Right, YFP–nesSwitch$_a$ is localized to the nucleus when expressed with the strong (SV40) NLS. When key–BFP is expressed with a moderately strong NLS, the same pattern of cytosolic YFP punctae formation is observed as when key–BFP is expressed without a NLS (Fig. 5b), which indicates that uncaging of the nuclear export sequence is independent of NLS on key–BFP localization. Key–BFP–NLS fluorescence is co-localized to NLS–YFP–nesSwitch$_a$ fluorescence **e**, YFP and RFP expression for synTF assay (corresponding to Fig. 5c) as a function of oestradiol (0–125 nM) and progesterone (0–500 nM). Fluorescence was measured at steady state using flow cytometry. Heat maps depict mean fluorescence, and are a representative sample of three biological replicates. The oestradiol dose (31.25 nM) depicted in Fig. 5c is indicated with the black rectangle on the heat maps.

nature research

Corresponding author(s): Dr. David Baker

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	CD Data collected on an AVIV CD Spectrometer. SEC data were collected using an Akta Pure with Superdex S75 Increase 10/30 column. MALS data collected with miniDAWN TREOS multi-angle static light scattering and an Optilab T-rEX (refractometer with EXtended range) detector. SAXS was collected at the SIBYLS beamline at LBNL in 2016. Bio Layer Interferometry data was collected on a ForteBio Octet Red96. Flow cytometry data was collected using a combination of BD FACS Diva v6.1.3 and v8.0.1 and custom software written in LabView 2013.
Data analysis	CD Data analyzed in Excel. BLI data was analyzed using ForteBio Data Analysis Software version 9.0.0.10. MALS data was analyzed using ASTRATM software. SAXS statistics determined using ScÅtter3. CD, SEC, SAXS, fluorescence, and BLI data were plotted using MatPlotLib 2.1.2 within Python 3.6.4 within an Anaconda environment. All analysis of flow cytometry data was performed using the FlowCytometryTools v0.5.0 package in Python 2.7 and FlowJo v10.

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

The data that support the findings of this study are available from the corresponding author upon reasonable request. Accession codes for Addgene plasmids are listed in the manuscript

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 sample size calculation was performed. Sample size of n=3 was selected. These sample sizes are sufficient due to use of purified protein and genetically identical samples.
Data exclusions	No data was excluded
Replication	All experimental findings were replicated successfully using biological and technical replicates taken on different days.
Randomization	Randomization was not performed because genetically identical samples were used
Blinding	Blinding was not performed because experiments were carried out by individual researchers

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	The unique materials (plasmids) that support the findings of this study are available on Addgene and from the corresponding authors upon reasonable request.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	ATCC® CRL-3216™

Authentication	Cell lines were not authenticated
Mycoplasma contamination	Cell lines tested negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	Commonly misidentified lines were not used in this study

Flow Cytometry

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	Yeast were grown in yeast media and diluted in TE solution to obtain an appropriate density or flowed directly into the instrument. Human HEK293T cells were grown in DMEM and washed in PBS+5%FBS before flow.
Instrument	BD LSR II and BD Fortessa
Software	BD FACS Diva v6.1.3 (LSR II) and BD FAC Diva v8.0.1 (Fortessa)was used to collect data and analysis was performed using the FlowCytometryTools v0.5.0 package in Python 2.7. Gating strategy was illustrated using FlowJo Version 10
Cell population abundance	Relevant yeast cell fractions were above 90% for all samples. Relevant HEK293T cell fractions were above 70% for all samples.
Gating strategy	For yeast, debris was removed by thresholding on SSC-H > 1000 and FITC-H > 10. For HEK293T cells, cells were gated on size based on SSC-A vs. FSC-A. Singlets were selected from this population based on SSC-A vs. SSC-H.

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