Title

De novo design of light-regulated dynamic proteins using deep learning

Authors

Scutteri, L.[1,2], Abriata, L.A. [1,3], Zhang S.[1], Chatzi Souleiman, A. [1,5], Lau, K. [3], Pojer, F. [3], Rahi, S.J. [2], Barth, P. [1,$]

Affiliations

[1] Institute of Bioengineering, EPFL, Switzerland

[2] Institute of Physics, EPFL, Switzerland

[3] Protein Production and Structure Core Facility, School of Life Sciences, EPFL, Switzerland

[5] Department of Computational Biology, UNIL, Switzerland

[$] to whom correspondence should be addressed: patrick.barth@epfl.ch

Abstract

Recent advances in deep learning have enabled accurate design of static protein structures, but the *de novo* design of protein functions controlled by programmable, intramolecular conformational changes remains an unsolved challenge. Here, we present a general deep learning–guided framework for designing dynamic, multi-domain proteins allosterically regulated by light. By integrating photoresponsive domains into de novo scaffolds, we engineered conformational switches that exhibit precise, reversible structural transitions upon illumination. Structural, spectroscopic, and functional analyses validated our designs and demonstrated precise spatiotemporal optogenetic control of diverse cellular processes in yeast, including subcellular localization, intercellular signaling, and population-level behaviors. This work establishes a broadly applicable strategy for encoding long-range allosteric control through designed intramolecular motions, and opens new avenues for programming dynamic protein functions from first principles, with implications for basic research, synthetic biology, and therapeutic development.

Highlights

- Deep-Learning guided *de novo* design of switchable protein scaffolds
- Computationally-guided control of protein conformational dynamics
- Computational design of light-induced allosteric control of multi-domain proteins
- *De novo* programmable spatio-temporal controls of cellular functions

Introduction

Protein switches play essential roles in biology by coupling input signals to output responses through structural rearrangements. Conformational transitions between distinct protein states regulate a broad array of functions, including enzyme catalysis[1], signal transduction[2,3], energy conversion[4], or protein recognition[5]. Owing to their versatility, natural protein switches have become attractive targets for *de novo* protein design. However, engineering synthetic proteins with analogous switch-like dynamics remains a major challenge as effective switches must access multiple conformational states, display fine-tuned energetics between states, and respond to specific external cues[6]. Previous advances include Ca^{2+}-responsive proteins[7], stimulus-responsive two-state hinge proteins[8], fold switches[9], tunable protein biosensors[10], and allosterically switchable protein assemblies[11]. Yet, these examples primarily rely on rigid-body movements driven by intermolecular interactions or localized intramolecular rearrangements. In contrast, natural protein switches frequently integrate stimulus sensing and response at distant sites, enabling long-range signal transmission[12–15] and the evolution of modular, multi-domain architectures[16–18]. To date, such fundamental allosteric mechanisms have not been recapitulated by *de novo* design.

A hallmark of many natural protein switches is allostery—the transmission of conformational changes from one site to a distant functional site, enabling protein activity to be modulated in response to environmental or cellular cues[19–22]. Biomolecular switches can be triggered by a variety of inputs, including ligand binding, environmental factors (e.g., temperature, heat, pH), allosteric mutations[23], and post-translational modifications. However, these conventional stimuli often pose practical limitations: they can be slow, irreversible, labor-intensive, or invasive, which limits their applicability. In contrast, optogenetics leverage light as a precise, fast, reversible, tunable, non-invasive, and cost-effective stimulus for controlling protein behavior[24–28]. Genetically

encoded light-sensitive domains—such as the LOV2 domain from *Avena sativa* phototropin 1 (AsLOV2)—can be inserted into target effectors to confer light-dependent regulation, enabling tight spatiotemporal control over protein function and cellular processes[29–36]. Despite advances in predicting insertion sites computationally[37], reliably engineering efficient allosteric coupling with LOV2 remains difficult, often requiring extensive experimental screening or rounds of in vivo evolution and optimization[38–40]. To date, the rational, computational design of optogenetic switches with tunable, allosterically controlled conformational dynamics and function has not been achieved.

Results

De novo design approach of multi-domain dynamic protein optoswitches

To address these limitations, we set out to develop a general computational platform for designing multi-domain switchable protein scaffolds, in which a stimulus-sensing domain allosterically regulates the function of a responder region through long-range intramolecular conformational changes (**Fig. 1a**). Our design strategy was guided by three key criteria: 1) To enable switching between well-folded conformations, we focused on *de novo* design of globular responder scaffolds that remain stably folded even upon disruption of intramolecular interactions between switchable elements. 2) To ensure functional relevance, we aimed to engineer conformational transitions that selectively expose or occlude a functional binding epitope. 3) To achieve allosteric regulation of the responder conformation by the sensor domain, we designed scaffolds in which the region containing the binding epitope is allosterically coupled to the switchable element of the sensor domain.

Our approach involves three main steps (**Fig. 1b, Fig. S1**).

Step 1 focuses on the *de novo* design of the switchable responder scaffold. We begin by exploring sequence–structure space to generate scaffolds that stably fold into a single-state conformation in which the functional epitope is occluded. To convert these static scaffolds into dynamic, switchable systems, we computationally design an ensemble of alternative conformational states that expose the functional epitope. This is achieved by combining steered molecular dynamics (MD) simulations with deep learning–based backbone generation. Switchable sequences capable of adopting both conformations are then identified through a combination of deep learning and physics-based sequence optimization methods.

Step 2 centers on the assembly and design of the allosterically coupled sensor–responder pair. To ensure efficient control of responder conformations, the sensor domain is inserted between the switchable elements of the responder scaffold. The strength of allosteric regulation is governed by the structural coupling between the two domains, which can be tuned by modifying the properties of the interdomain linkers. Increased linker rigidity promotes stronger coupling and vice versa. These linkers are *de novo* designed using iterative cycles of deep learning–based backbone and sequence generation, followed by evaluation of interdomain coupling using Mutual Information (MI) analyses from physics-based simulations[2,41].

Step 3 involves comprehensive experimental characterization of the designed switches (**Fig. 1c**). Responder conformations and their intrinsic dynamics are analyzed using NMR spectroscopy and physics-based simulations. The fine-tuned energetics and population distribution of the responder and responder-sensor conformational states are validated through real-time fluorescence measurements in living cells.

As a proof-of-concept application, we designed light-sensitive switches using AsLOV2 as the stimulus-sensing domain. Upon blue light irradiation, AsLOV2 undergoes a conformational

change that releases its C-terminal Jα helix from the core domain. Accordingly, we engineered responder scaffolds in which the switchable region harboring the binding epitope is allosterically coupled to the C-terminus of AsLOV2, enabling light-dependent functional control. In principle, this approach enables the custom design of diverse optogenetic control systems, enabling precise spatiotemporal regulation of programmable cellular functions (**Fig. 1d**).

Conformations and dynamics of the *de novo* responder switch

To validate the structural predictions of the *de novo* designed responder domain, we selected and expressed sequences predicted to mainly occupy either the closed (i.e. state 1, **Fig. 2a**) or open (i.e. state 2, **Fig. 2b**) states.

We first characterized their folding stability using circular dichroism (CD) spectroscopy and thermal denaturation experiments. CD spectra confirmed the presence of well-folded structures in both states, in agreement with the designed αβ topology (**Fig. S2b**). Thermal unfolding profiles indicated exceptional stability, with melting temperatures exceeding 100 °C for both closed and open states (**Fig. S2c**). Remarkably, the open state design retained proper folding at elevated temperatures, despite the exposed core of the protein and the dynamic nature of the C-terminal helix.

We then investigated the closed and open conformations of the responder scaffold using nuclear magnetic resonance (NMR) spectroscopy. We acquired 2D ^{1}H,^{15}N-heteronuclear single-quantum coherence (HSQC) spectra for each design and 3D ^{1}H,^{15}N,^{13}C NMR experiments (**Fig. S3a, b**). The NMR ensembles for both states matched closely the design models (Ca RMSD < 2A). While the switchable element is well folded onto the core of the scaffold in the closed state (**Fig. 2c**), it is largely disengaged in the open state (**Fig. 2d**). This difference stems from a large set of long-range NOEs connecting the switchable helix (residues 112-126) to the protein core in the closed

state (**Fig. 2e, S4a**), and a lack of corresponding restraints in the open state design. The fully disengaged helix conformation observed for the open state is in excellent agreement with the design models.

To directly probe the conformational dynamics of each state, we collected ^{15}N T1, ^{15}N T2 and {^{1}H,^{15}N} heteronuclear NOE relaxation data. The resulting R_2/R_1 profiles (**Fig. S4b**) displayed central values consistent with monomeric states for both proteins, along with reductions indicative of fast (ps–ns) motions at and around the C-terminal helix in the open state only. The heteronuclear NOE measurements (**Fig. 3f**) corroborated these findings, confirming that such ps–ns dynamics occur exclusively in the open state. This dynamic signature is consistent with the C-terminal helix being detached from the core, as both designed and observed in the NMR structure. In contrast, the closed-state design exhibited high rigidity at the switchable helix–core interface, in agreement with its tightly packed conformation. Notably, the R_2/R_1 data did not reveal strongly positive deviations, which would have been indicative of slower timescale dynamics in this design. Consistent with the design goals, all metrics supported a well folded and conformationally stable core of the scaffold in the open state despite the absence of stabilizing interactions with the switchable helical subdomain. Secondary structure analysis based on $\Delta C\alpha$–$\Delta C\beta$ chemical shifts revealed excellent agreement with the expected topology in both conformations (**Fig. S4c**). Importantly, the C-terminal helix retained significant helical propensity in the open state, consistent with its folded yet flexible behavior observed in the NMR ensemble. Lastly, dynamic changes between the states localized to the switchable helix were corroborated by root-mean-square fluctuation (RMSF) values obtained from molecular dynamics (MD) simulations carried out on each design model (**Fig. 3g**).

These results provide direct structural evidence that our approach can *de novo* design dynamic proteins with novel globular topologies capable of adopting distinct, programmable conformational states with near atomic-level precision.

Fine-tuning conformational equilibrium of the responder by multi state design

While the NMR experiments validated sequences that adopt one of the two conformational states of the responder switch, we next sought to identify sequences capable of toggling between both states. We reasoned that a library of switchable variants sampling a broad range of conformational energy differences (ΔE, **Fig. 3a**) would facilitate the optimization of light-inducible LOV2-grafted switches. To this end, we focused our search on residues at the core–switchable helix interface and carried out multistate design calculations, combining classical Rosetta-based energy optimization with deep learning–based sequence design approaches, to identify mutations compatible with both conformational states. This strategy yielded 15 switchable sequence variants that span a gradient of state occupancies between fully closed and fully open conformations. These designs differ from the NMR-validated closed-state sequence by 3 to 24 mutations and incorporate polar substitutions that enhance solubility and stability in the open state, as well as van der Waals packing defects that destabilize the helix–core interface, thereby promoting conformational flexibility (**Methods**).

To experimentally assess the conformational switching behavior of these *de novo* designs, we developed a real-time, non-invasive cellular assay capable of quantitatively reporting state distributions. This assay is based on split complementation of the sfCherry2 fluorescent protein, whose two fragments (1–10 and 11) spontaneously associate into a functional fluorophore when brought into close proximity[42,43]. Fluorescence is observed in the closed conformation of the scaffold, in which the N- and C-termini are juxtaposed, allowing efficient complementation. In the open conformation, displacement of the 11th β-strand disrupts complementation, leading to fluorescence loss (**Fig. 3b**). This split sfCherry2 assay thus serves as a sensitive and direct reporter of the protein's conformational state distribution.

We applied the sfCherry2 complementation assay to our *de novo* responder library. The NMR-validated sequence, which populates the closed state, exhibited the highest fluorescence, consistent with close N- and C-terminal proximity. In contrast, multistate designs showed progressively reduced fluorescence, correlating with decreasing ΔE values (**Fig. 3c, Fig. S5**). These results confirm that multistate designs function as conformational switches and demonstrate that our physics-based energy calculations accurately predicted the state distribution, as validated by steady-state fluorescence measurements.

Allosteric coupling of LOV2 sensor to de novo responder switches

We next sought to construct light-responsive switches by designing LOV2 sensor–responder fusion proteins (**Fig. 3d**). A key challenge was to preserve the intrinsic conformational equilibrium of the responder in the dark (resting state of LOV2), while enabling selective conformational switching upon blue light activation of LOV2. Direct fusion of LOV2 to the responder, without intervening linkers, strongly disrupted the closed conformation, leading to a pronounced decrease in fluorescence (**Fig. S6**).

To overcome this and achieve optimal light-inducible control, we computationally designed a set of *de novo* linker regions connecting the N- and C-termini of LOV2 to the core and switchable subdomains of the responder scaffold. This involved cycles of deep learning–based structure prediction and sequence design to evaluate the structural impact of LOV2 grafting and eliminate linker designs that significantly perturbed the responder conformation in the dark state. We generated a diverse library of optoswitches incorporating linkers of varying length and conformational flexibility to explore a wide range of allosteric coupling efficiencies.

We then performed physics-based simulations on six representative fusions to compute Mutual Information (MI) between residues in the LOV2 sensor and the switchable region of the responder.

These MI values served as a quantitative proxy for allosteric coupling strength between the two domains.

To assess functional coupling, we fused split sfCherry to the responder domains and measured fluorescence changes in yeast upon blue light stimulation. All designs exhibited significant light-dependent fluorescence shifts, indicating effective optogenetic control (**Fig. 3e**). Notably, MI values strongly correlated with fluorescence changes, validating our computational framework for achieving precise and programmable allosteric control through *de novo* designed linkers. Among the tested constructs, Linker #5, composed of a rigid, well-folded helical segment, showed the highest MI and was selected for further experimental validation.

To assess the structural integrity of the responder scaffold within the optoswitch, we characterized the linker #5 fusion construct using NMR and CD spectroscopy. The $^{1}H,^{15}N$-HSQC spectrum exhibited well-dispersed peaks, consistent with a well-folded protein (**Fig. S7**), and CD analysis confirmed proper secondary structure and high thermal stability (Tm > 100 °C) (**Fig. S8**). These results demonstrate that LOV2 insertion did not compromise the *de novo* scaffold's structure or folding stability.

Since dynamic reversibility is a key requirement for optogenetic control, we measured the kinetics of the optoswitches under brief blue light exposure (**Fig. 3f, g**). Upon illumination, the LOV2-controlled optoswitches exhibited a complete cycle of fluorescence decrease and recovery within 10 minutes, confirming dynamic, reversible, and light-dependent modulation of scaffold conformation. These findings demonstrate that our approach not only enables the design of multi-domain protein switches that transition between predefined conformational states on functionally relevant timescales, but also allows precise and tunable optogenetic control over their conformational equilibria.

Programmable spatio-temporal control of cellular behaviors via optoswitches functionalization

The precise, dynamic, and reversible light-induced activation of our *de novo* switches offers a powerful strategy for programmable spatiotemporal control of cellular behaviors, enabled by the ability to design switchable scaffolds with custom shapes and epitope properties from first principles (**Fig. 1d**).

As a proof of concept, we functionalized our optoswitches with photo-caged peptide binding motifs under the following design constraints: in the dark, the epitope must remain buried and tightly packed against the core of the *de novo* scaffold, rendering it inactive. Upon blue light exposure, the LOV2-driven conformational change should induce epitope exposure, thereby triggering a specific cellular response. This modular approach is broadly adaptable to diverse synthetic biology applications, establishing a versatile platform for engineering light-regulated cellular functions.

To demonstrate the utility of our approach, we first applied it to the optogenetic control of subcellular localization (**Fig. 4a**). We grafted a nuclear localization signal (NLS)[44] onto the C-terminal helix of the optoswitch scaffold, enabling light-inducible, reversible nuclear import with rapid kinetics on the minute timescale (**Fig. 4b, c**). In the dark, the NLS is sterically occluded, preventing its recognition by nuclear importins α and β. Upon blue light exposure, unfolding of the Jα helix unmasks the NLS, allowing engagement with the nuclear import machinery (**Fig. S9**). This design yielded robust functionality, with a 5.6-fold increase in the nuclear-to-cytoplasmic (N:C) fluorescence ratio upon light stimulation, followed by a return to baseline in the dark (**Fig. 4c, Fig. S10**).

To further demonstrate the versatility of our optoswitch platform, we next targeted a fundamentally different cellular application: the optogenetic control of intercellular communication and population-level behavior. Specifically, we focused on the well-characterized mating pathway in *Saccharomyces cerevisiae*, in which exposure to α-pheromone secreted by MATα cells triggers G_1 cell-cycle arrest in MATa cells as part of the mating response[45,46] (**Fig. 4d**).

We engineered the system so that, in the dark, the α-pheromone epitope remains sterically occluded within the optoswitch scaffold, preventing its interaction with the Ste2 G protein-coupled receptor (GPCR)[47]. Upon blue light absorption and LOV2-mediated undocking of the Jα helix, the α-pheromone becomes solvent-exposed and capable of engaging Ste2, thereby initiating the downstream signaling cascade leading to cell-cycle arrest (**Fig. S11**). To engineer a self-regulating cell population control system based on intercellular communication, we genetically programmed cells to express and secrete the optoswitch. To ensure proper secretion and surface-level activity, we appended a secretion signal peptide to the N-terminus of the optoswitch (**Methods**). Functionalization of the designed optoswitch with the α-pheromone thus enabled light-dependent control of cell-cycle progression at the G_1 phase in MATa cells. In the dark, cells proliferated normally; upon blue light illumination, however, we observed robust and reproducible cell-cycle arrest (**Fig. 4e**).

Together, these results establish our *de novo* designed multi-domain optoswitches as a modular platform for rewiring diverse cellular responses with high spatiotemporal resolution. This approach enables precise, light-dependent control over intercellular signaling pathways and communication, offering a powerful framework for dynamic cellular engineering in synthetic biology.

Discussion

The field of de novo protein design has transformed our capacity to build synthetic proteins with defined structures and tailored functions. However, the rational design of dynamic, multi-state proteins that can reversibly transition between distinct conformations has remained a longstanding challenge. Here, we introduce a general deep learning–guided framework for designing dynamic, multi-domain, light-responsive protein switches from first principles. Our strategy integrates deep learning–based multistate sequence design and structure prediction with physics-based energy calculations and molecular dynamics simulations. Using this approach, we engineered multiple protein scaffolds that undergo precise and reversible conformational transitions on functional timescales upon light stimulation. By grafting specific binding epitopes into these scaffolds, we generated optoswitches that enable programmable, spatiotemporal control of diverse cellular functions, including intracellular localization, intercellular signaling, and population-level dynamics. To our knowledge, this represents the first demonstration of optogenetically controlled *de novo* protein switches and their application in synthetic cell biology. Importantly, because scaffold shapes, binding surfaces, and conformational transitions can now be designed with atomic-level control, our framework should significantly broaden the range of light-controllable protein-based technologies (**Fig. 1d**).

Our results revealed strong agreement across deep learning–based design predictions, physics-based simulations, experimentally determined structures, and observed conformational dynamics. Notably, we could rationally tune allosteric communication between sensor and responder domains, providing direct evidence that allosteric control can be designed *de novo* in multi-domain proteins. In contrast to prior efforts that confined conformational changes to small protein regions, our designs exhibit long-range allosteric transitions spanning over 80 Å,

comparable to the global intramolecular rearrangements seen in naturally evolved systems such as signaling receptors, motor proteins, and multi-domain enzymes. Because our approach is modular and agnostic to domain structure and function, it should be broadly applicable for designing allosteric communication pathways across diverse multi-domain proteins. Collectively, this work lays the foundation for programmable synthetic proteins with precisely tunable conformational dynamics, expanding the design space of synthetic biology, optogenetics, and protein engineering.

Figures

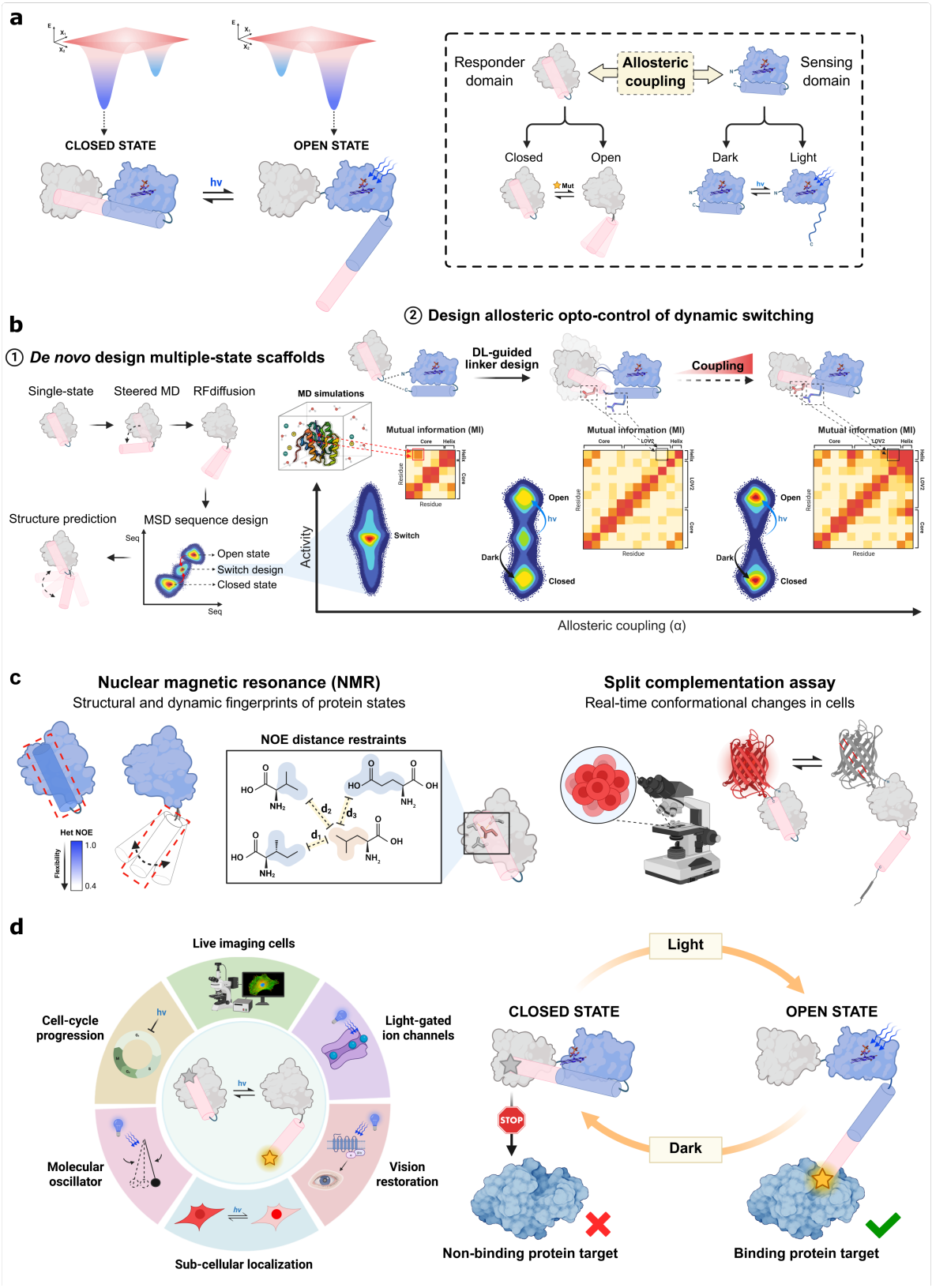

a
CLOSED STATE
OPEN STATE
hv
Responder domain
Allosteric coupling
Sensing domain
Closed
Open
Dark
Light
Mut
b
① De novo design multiple-state scaffolds
Single-state
Steered MD
RFdiffusion
MD simulations
Structure prediction
MSD sequence design
Seq
Open state
Switch design
Closed state
② Design allosteric opto-control of dynamic switching
DL-guided linker design
Coupling
Mutual information (MI)
Core
Helix
LOV2
Residue
Switch
Open
Closed
Dark
Activity
Allosteric coupling (α)
c
Nuclear magnetic resonance (NMR)
Structural and dynamic fingerprints of protein states
Het NOE
Flexibility
1.0
0.4
NOE distance restraints
Split complementation assay
Real-time conformational changes in cells
d
Live imaging cells
Cell-cycle progression
Light-gated ion channels
Molecular oscillator
Vision restoration
Sub-cellular localization
Light
Dark
CLOSED STATE
OPEN STATE
STOP
Non-binding protein target
Binding protein target

Fig. 1. Deep learning guided *de novo* design of dynamic protein optoswitches.

(**a**) Schematic of design goal to engineer multi-domain dynamic protein optoswitches in a two-state equilibrium, where the responder domain can be allosterically controlled by the stimulus-sensing domain. (**b**) Generalizable framework for the deep learning guided *de novo* design of multi-domain optoswitchable proteins: (1) Deep learning-guided *de novo* design of scaffolds switching between multiple states; (2) Design of allosteric optogenetic control of dynamic protein switching and activity through deep learning–guided linker engineering and Mutual Information (MI) analysis, enabling precise selection of linkers that control mechanical coupling between the optosensor and the switchable scaffold. (**c**) Analysis of protein structure and dynamics through nuclear magnetic resonance (NMR) spectroscopy and sfCherry2-based split complementation assay. (**d**) Widespread applications for programmable spatio-temporal optogenetic control of cellular behaviors through functionalization of protein optoswitches. In the dark, the epitope is tightly packed against the core of the *de novo* designed scaffold and is thus inactive. Upon blue light absorption, the LOV2-mediated opening of the protein switch exposes the peptide, thereby mediating the desired cellular response.

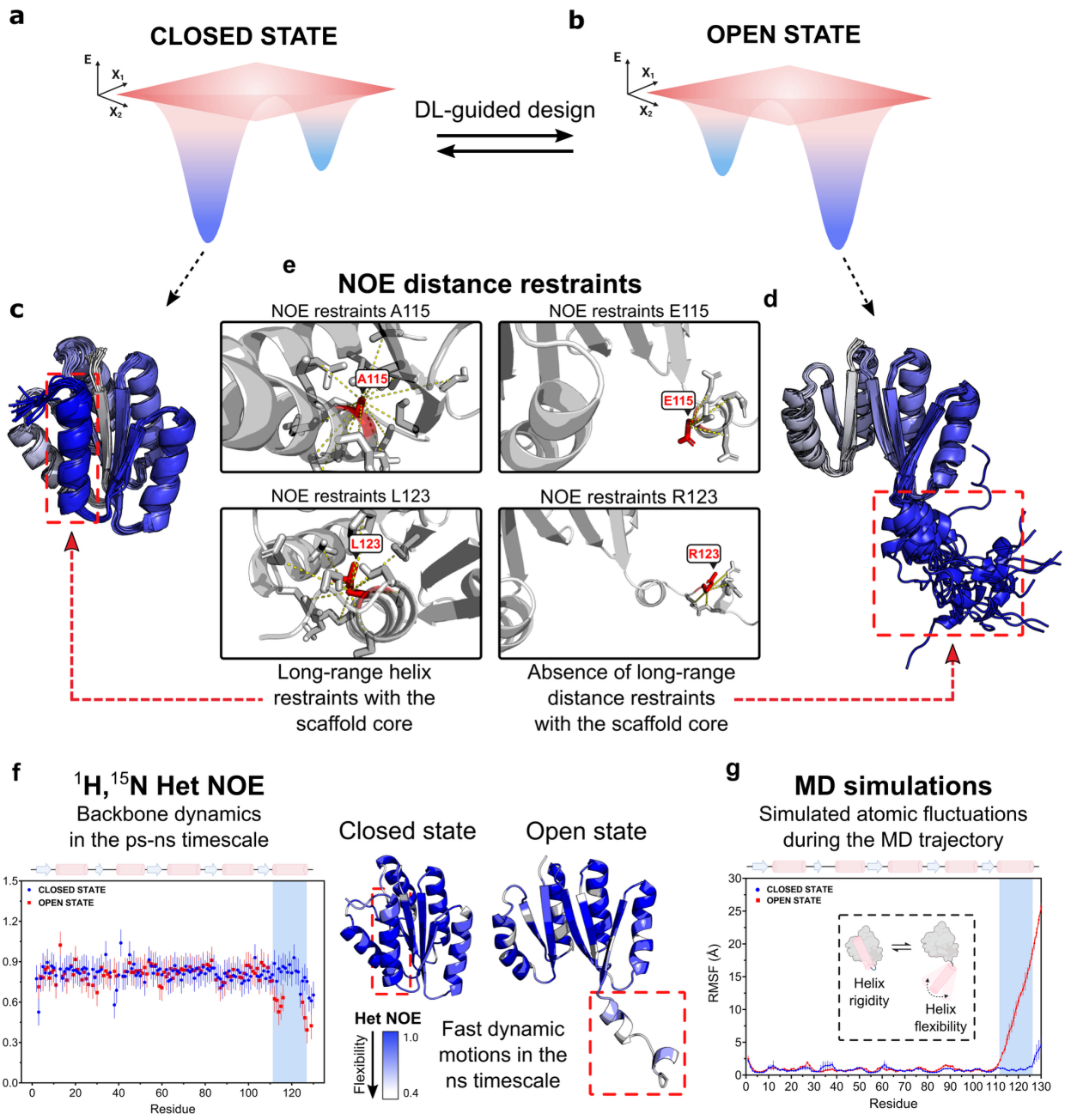

a
CLOSED STATE
b
OPEN STATE
DL-guided design
c
d
e
NOE distance restraints
NOE restraints A115
NOE restraints E115
NOE restraints L123
NOE restraints R123
A115
E115
L123
R123
Long-range helix restraints with the scaffold core
Absence of long-range distance restraints with the scaffold core
f
1H,15N Het NOE
Backbone dynamics in the ps-ns timescale
CLOSED STATE
OPEN STATE
Residue
Closed state
Open state
Het NOE
Flexibility
Fast dynamic motions in the ns timescale
g
MD simulations
Simulated atomic fluctuations during the MD trajectory
RMSF (Å)
Helix rigidity
Helix flexibility
Residue

Fig. 2. NMR analysis of conformational equilibrium and protein dynamics.

(**a**) Schematic energy landscape of the designed scaffold in the closed state. (**b**) Schematic energy landscape of the designed scaffold in the open state. (**c**) NMR ensemble of the *de novo* designed scaffold in the closed state. (**d**) NMR ensemble of the *de novo* designed scaffold in the open state. (**e**) Representative NOE distance restraints for residues 115 and 123 in both conformational states. In the closed state, residues in the C-terminal helix engage in long-range NOEs with the scaffold core, whereas in the open state, only short-range contacts are observed. (**f**) $^{1}H,^{15}N$ Heteronuclear NOE for the closed (blue) and open (red) states of the scaffold. Structural models are colored by per-residue Het NOE values, highlighting fast-timescale dynamics of the C-terminal helix in the open state. (**g**) Per-residue root mean square fluctuation (RMSF) profiles for the closed (blue) and open (red) states of the scaffold, averaged over four independent 500 ns molecular dynamics (MD) simulations.

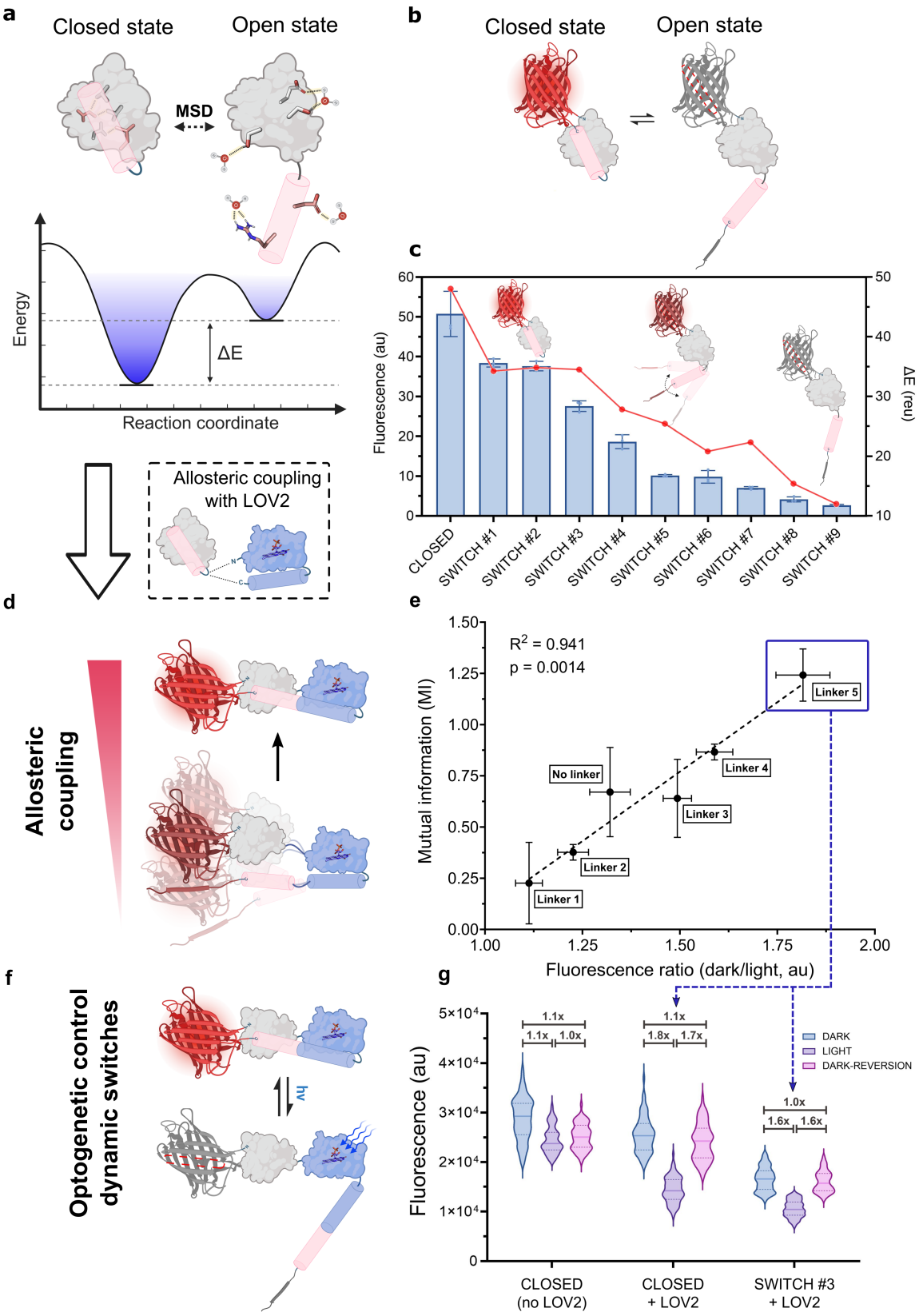

a
Closed state
Open state
MSD
Energy
ΔE
Reaction coordinate
Allosteric coupling with LOV2
b
Closed state
Open state
c
Fluorescence (au)
ΔE (reu)
CLOSED
SWITCH #1
SWITCH #2
SWITCH #3
SWITCH #4
SWITCH #5
SWITCH #6
SWITCH #7
SWITCH #8
SWITCH #9
d
Allosteric coupling
e
R2 = 0.941
p = 0.0014
Mutual information (MI)
Fluorescence ratio (dark/light, au)
Linker 1
Linker 2
Linker 3
Linker 4
Linker 5
No linker
f
Optogenetic control dynamic switches
hν
g
Fluorescence (au)
DARK
LIGHT
DARK-REVERSION
CLOSED (no LOV2)
CLOSED + LOV2
SWITCH #3 + LOV2

Fig. 3. Experimental validation of switchable dynamics through sfCherry2-based split complementation assay in cells.

(**a**) Conformational energy landscape showing the closed and open states of the dynamic switchable scaffolds. Allosteric mutations at the helix-core interface through multistate design enable precise fine-tuning of the scaffold energetics and conformational equilibrium. (**b**) Schematic of the sfCherry2-based split complementation assay. In the closed state, the spatial proximity of the N- and C-termini of the scaffold enables reconstitution of the split sfCherry2, producing a fluorescent signal. In the open state, the opening of the scaffold displaces the C-terminal β-strand, disrupting fluorescence. (**c**) On the left Y-axis, sfCherry2-based signal of designed scaffolds, representative of the scaffold's conformations across multiple states, in HEK293T cells. On the right Y-axis, computational calculation of ΔE between closed and open states upon sequence threading and Rosetta relax. (**d**) Design of LOV2 sensor-responder fusion proteins. (**e**) Correlation between Mutual Information (MI) from the LOV2 sensing domain to the C-terminal helix of the scaffold, and fluorescence ratio from dark to light, with different linker designs (R^2= 0.941, p = 0.0014). (**f**) Optogenetic control of dynamic switching behavior from the closed state (dark) to the open state (light). (**g**) validation of the reversible light-induced switching behavior; initial dark phase marked in blue, light-induction (10') marked in magenta, dark-reversion (10') marked in pink.

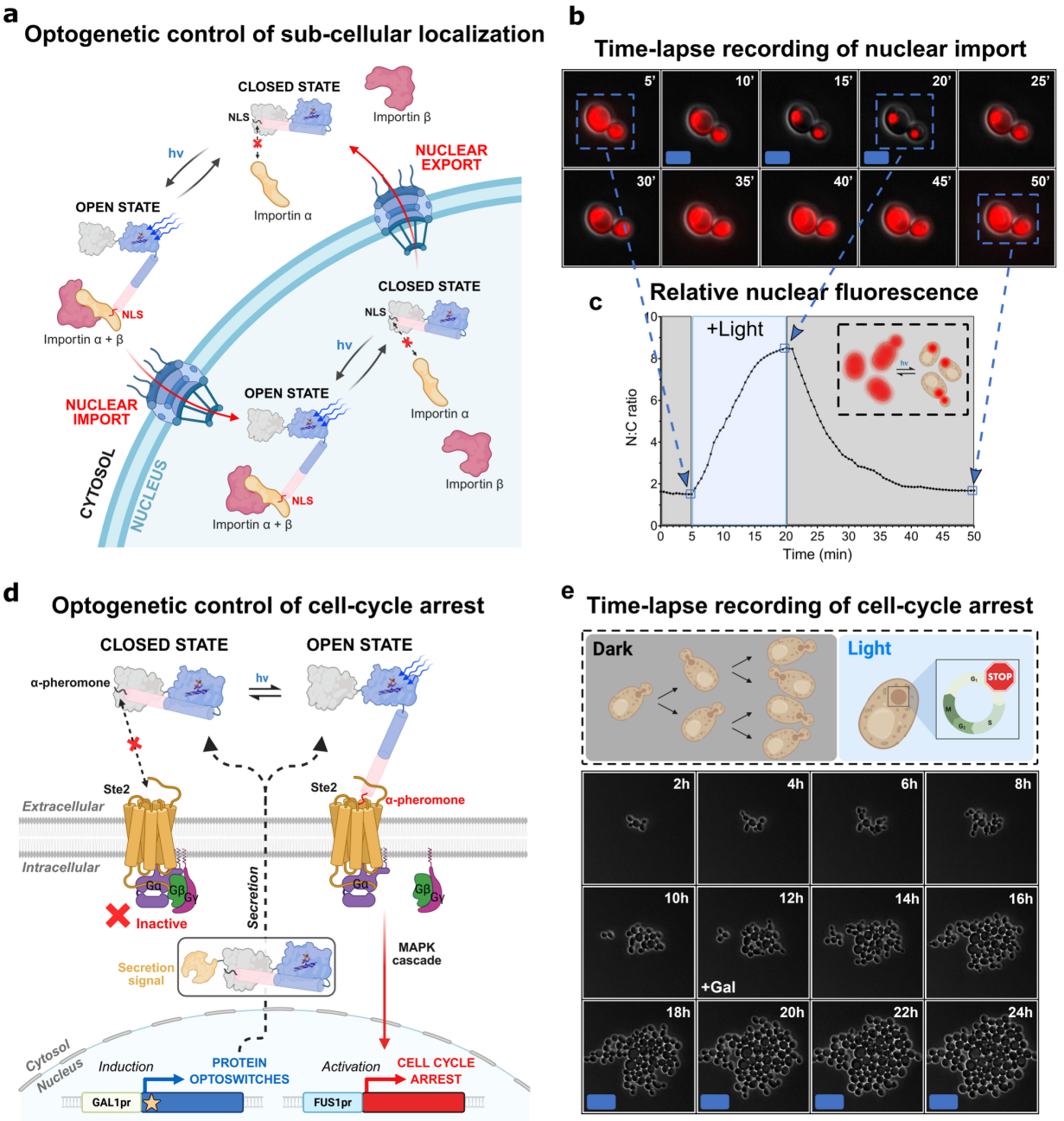

a
Optogenetic control of sub-cellular localization
CLOSED STATE
NLS
Importin β
hv
NUCLEAR EXPORT
OPEN STATE
Importin α
NLS
Importin α + β
CLOSED STATE
NLS
hv
OPEN STATE
Importin α
NUCLEAR IMPORT
Importin β
CYTOSOL
NUCLEUS
NLS
Importin α + β
b
Time-lapse recording of nuclear import
5'
10'
15'
20'
25'
30'
35'
40'
45'
50'
c
Relative nuclear fluorescence
+Light
N:C ratio
Time (min)
d
Optogenetic control of cell-cycle arrest
CLOSED STATE
OPEN STATE
α-pheromone
hv
Ste2
Extracellular
Intracellular
Gα
Gβ
Gγ
Inactive
Secretion
Secretion signal
Ste2
α-pheromone
MAPK cascade
Cytosol
Nucleus
Induction
PROTEIN OPTOSWITCHES
Activation
CELL CYCLE ARREST
GAL1pr
FUS1pr
e
Time-lapse recording of cell-cycle arrest
Dark
Light
STOP
2h
4h
6h
8h
10h
12h
14h
16h
+Gal
18h
20h
22h
24h

Fig. 4. Synthetic biology applications of functionalized protein optoswitches for programmable spatio-temporal control of cellular behaviors.

(**a**) Schematic of optogenetic control of sub-cellular localization. Photo-caging of the nuclear localization signal (NLS) enables reversible light-dependent nuclear import in the open state of the dynamic protein optoswitches. (**b**) Time-lapse recording of light-induced nuclear import. Cells were incubated in the dark for 5 min before blue light irradiation for 15 min, followed by a 30 min dark-recovery phase. The fluorescent protein ymScarlet is fused at the N-terminus of the protein optoswitch. (**c**) Relative nuclear-to-cytoplasmic (N:C) fluorescence ratio corresponding to cells in (b). (**d**) Schematic of optogenetic control of cell-cycle arrest. Photo-caging of the α-pheromone peptides enables light-dependent arrest of the cell-cycle in the open state of the dynamic protein optoswitches. (**e**) Time-lapse recording of optogenetic control of cell-cycle arrest. Cells were incubated in the dark for 10 h prior to galactose induction, followed by an additional 6 h in the dark, and finally exposed to 8 h of blue light. A secretion signal was fused to the N-terminus of the optoswitch to enable its extracellular export, allowing the light-dependent exposure of the C-terminal α-pheromone peptide to activate GPCR signaling at the cell surface.

References

1. Huse, M. & Kuriyan, J. The Conformational Plasticity of Protein Kinases. *Cell* **109**, 275–282 (2002).
2. Jefferson, R. E. *et al.* Computational design of dynamic receptor—peptide signaling complexes applied to chemotaxis. *Nature Communications 2023 14:1* **14**, 1–17 (2023).
3. Chen, K.-Y. M. *et al.* Computational design of highly signalling-active membrane receptors through solvent-mediated allosteric networks. *Nature Chemistry 2025 17:3* **17**, 429–438 (2025).
4. Schliwa, M. & Woehlke, G. Molecular motors. *Nature 2003 422:6933* **422**, 759–765 (2003).
5. Gonzalez, L., Plecs, J. J. & Alber, T. An engineered allosteric switch in leucine-zipper oligomerization. *Nature Structural Biology 1996 3:6* **3**, 510–515 (1996).
6. Alberstein, R. G., Guo, A. B. & Kortemme, T. Design principles of protein switches. *Curr Opin Struct Biol* **72**, 71–78 (2022).
7. Guo, A. B. *et al.* Deep learning-guided design of dynamic proteins. *Science* **388**, eadr7094 (2025).
8. Praetorius, F. *et al.* Design of stimulus-responsive two-state hinge proteins. *Science (1979)* **381**, 754–760 (2023).
9. Ruan, B. *et al.* Design and characterization of a protein fold switching network. *Nature Communications 2023 14:1* **14**, 1–14 (2023).
10. Quijano-Rubio, A. *et al.* De novo design of modular and tunable protein biosensors. *Nature 2021 591:7850* **591**, 482–487 (2021).
11. Pillai, A. *et al.* De novo design of allosterically switchable protein assemblies. *Nature 2024 632:8026* **632**, 911–920 (2024).
12. Rosenbaum, D. M., Rasmussen, S. G. F. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature 2009 459:7245* **459**, 356–363 (2009).
13. Zhang, M. *et al.* G protein-coupled receptors (GPCRs): advances in structures, mechanisms and drug discovery. *Signal Transduction and Targeted Therapy 2024 9:1* **9**, 1–43 (2024).

14. Wootten, D., Christopoulos, A. & Sexton, P. M. Emerging paradigms in GPCR allostery: implications for drug discovery. *Nature Reviews Drug Discovery 2013 12:8* **12**, 630–644 (2013).

15. Chen, K. Y. M. *et al.* Computational design of highly signalling-active membrane receptors through solvent-mediated allosteric networks. *Nature Chemistry 2025 17:3* **17**, 429–438 (2025).

16. Parsons, S. J. & Parsons, J. T. Src family kinases, key regulators of signal transduction. *Oncogene 2004 23:48* **23**, 7906–7909 (2004).

17. Tannous, E. A., Yates, L. A., Zhang, X. & Burgers, P. M. Mechanism of auto-inhibition and activation of Mec1ATR checkpoint kinase. *Nature Structural & Molecular Biology 2020 28:1* **28**, 50–61 (2020).

18. Lott, J. S. & Lee, T. V. Revealing the Inter-Module Interactions of Multi-Modular Nonribosomal Peptide Synthetases. *Structure* **25**, 693–695 (2017).

19. Monod, J., Changeux, J. P. & Jacob, F. Allosteric proteins and cellular control systems. *J Mol Biol* **6**, 306–329 (1963).

20. Süel, G. M., Lockless, S. W., Wall, M. A. & Ranganathan, R. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nature Structural Biology 2002 10:1* **10**, 59–69 (2002).

21. Popovych, N., Sun, S., Ebright, R. H. & Kalodimos, C. G. Dynamically driven protein allostery. *Nature Structural & Molecular Biology 2006 13:9* **13**, 831–838 (2006).

22. Swain, J. F. & Gierasch, L. M. The changing landscape of protein allostery. *Curr Opin Struct Biol* **16**, 102–108 (2006).

23. Keri, D. *et al.* Computational rewiring of allosteric pathways reprograms GPCR selective responses to ligands. *bioRxiv* 2022.03.30.486353 (2022) doi:10.1101/2022.03.30.486353.

24. Mathony, J. & Niopek, D. Enlightening Allostery: Designing Switchable Proteins by Photoreceptor Fusion. *Adv Biol* **5**, 2000181 (2021).

25. McCue, A. C. & Kuhlman, B. Design and engineering of light-sensitive protein switches. *Curr Opin Struct Biol* **74**, 102377 (2022).

26. Emiliani, V. *et al.* Optogenetics for light control of biological systems. *Nature Reviews Methods Primers 2022 2:1* **2**, 1–25 (2022).

27. Emiliani, V. *et al.* Optogenetics for light control of biological systems. *Nature Reviews Methods Primers 2022 2:1* **2**, 1–25 (2022).

28. Gligorovski, V. *et al.* Light-directed evolution of dynamic, multi-state, and computational protein functionalities. *bioRxiv* 2024.02.28.582517 (2025) doi:10.1101/2024.02.28.582517.

29. Niopek, D. *et al.* Engineering light-inducible nuclear localization signals for precise spatiotemporal control of protein dynamics in living cells. *Nature Communications 2014 5:1* **5**, 1–11 (2014).

30. Niopek, D., Wehler, P., Roensch, J., Eils, R. & Di Ventura, B. Optogenetic control of nuclear protein export. *Nature Communications 2016 7:1* **7**, 1–9 (2016).

31. Kroll, K. L., French, A. R., Sosnick, T. R. & Rock, R. S. LILAC: enhanced actin imaging with an optogenetic Lifeact. *Nature Methods 2023 20:2* **20**, 214–217 (2023).

32. Garcia-Marcos, M. *et al.* Optogenetic activation of heterotrimeric g-proteins by LOV2GIVe, a rationally engineered modular protein. *Elife* **9**, 1–21 (2020).

33. Strickland, D. *et al.* TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nature Methods 2012 9:4* **9**, 379–384 (2012).

34. Lee, S. Y. *et al.* Engineered allostery in light-regulated LOV-Turbo enables precise spatiotemporal control of proximity labeling in living cells. *Nature Methods 2023 20:6* **20**, 908–917 (2023).

35. Renicke, C., Schuster, D., Usherenko, S., Essen, L. O. & Taxis, C. A LOV2 Domain-Based Optogenetic Tool to Control Protein Degradation and Cellular Function. *Chem Biol* **20**, 619–626 (2013).

36. Bubeck, F. *et al.* Engineered anti-CRISPR proteins for optogenetic control of CRISPR–Cas9. *Nature Methods 2018 15:11* **15**, 924–927 (2018).

37. Wolf, B. *et al.* Rational engineering of allosteric protein switches by in silico prediction of domain insertion sites. *bioRxiv* 2024.12.04.626757 (2024) doi:10.1101/2024.12.04.626757.

38. Southern, N. T. *et al.* Phage-Assisted Evolution of Allosteric Protein Switches. *bioRxiv* 2025.06.12.659331 (2025) doi:10.1101/2025.06.12.659331.

39. Peter, E., Dick, B. & Baeurle, S. A. Mechanism of signal transduction of the LOV2-Jα photosensor from Avena sativa. *Nature Communications 2010 1:1* **1**, 1–7 (2010).

40. Kim, C. *et al.* Structural dynamics of protein-protein association involved in the light-induced transition of Avena sativa LOV2 protein. *Nature Communications 2024 15:1* **15**, 1–15 (2024).

41. Chen, K. Y. M., Keri, D. & Barth, P. Computational design of G Protein-Coupled Receptor allosteric signal transductions. *Nature Chemical Biology 2019 16:1* **16**, 77–86 (2019).

42. Feng, S. *et al.* Improved split fluorescent proteins for endogenous protein labeling. *Nature Communications 2017 8:1* **8**, 1–11 (2017).

43. Feng, S. *et al.* Bright split red fluorescent proteins for the visualization of endogenous proteins and synapses. *Communications Biology 2019 2:1* **2**, 1–12 (2019).

44. Kalderon, D., Roberts, B. L., Richardson, W. D. & Smith, A. E. A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499–509 (1984).

45. Jenness, D. D., Burkholder, A. C. & Hartwell, L. H. Binding of α-factor pheromone to yeast a cells: Chemical and genetic evidence for an α-factor receptor. *Cell* **35**, 521–529 (1983).

46. Bender, A. & Sprague, G. F. Yeast peptide pheromones, a-factor and α-factor, activate a common response mechanism in their target cells. *Cell* **47**, 929–937 (1986).

47. Jenness, D. D., Burkholder, A. C. & Hartwell, L. H. Binding of α-factor pheromone to yeast a cells: Chemical and genetic evidence for an α-factor receptor. *Cell* **35**, 521–529 (1983).

48. Hicks, J. B. & Herskowitz, I. Evidence for a new diffusible element of mating pheromones in yeast. *Nature 1976 260:5548* **260**, 246–248 (1976).

49. Sprague, G. F. & Herskowitz, I. Control of yeast cell type by the mating type locus: I. Identification and control of expression of the a-specific gene BAR1. *J Mol Biol* **153**, 305–321 (1981).

50. Xu, X. *et al.* Solution NMR Structure of De novo designed protein, Northeast Structural Genomics Consortium (NESG) Target OR462. *Worldwide Protein Data Bank* (2014) doi:10.2210/PDB2MR6/PDB.

51. Huang, P. S. *et al.* RosettaRemodel: A Generalized Framework for Flexible Backbone Protein Design. *PLoS One* **6**, e24109 (2011).

52. Watson, J. L. *et al.* De novo design of protein structure and function with RFdiffusion. *Nature 2023 620:7976* **620**, 1089–1100 (2023).

53. Dauparas, J. *et al.* Robust deep learning–based protein sequence design using ProteinMPNN. *Science (1979)* **378**, 49–56 (2022).

54. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nature Methods 2022 19:6* **19**, 679–682 (2022).

55. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature 2021 596:7873* **596**, 583–589 (2021).

56. Leaver-Fay, A., Jacak, R., Stranges, P. B. & Kuhlman, B. A Generic Program for Multistate Protein Design. *PLoS One* **6**, e20937 (2011).

57. Szilárd, P., Abraham, M. J., Kutzner, C., Hess, B. & Lindahl, E. Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **8759**, 3–27 (2015).

58. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).

59. Klauda, J. B. *et al.* Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *Journal of Physical Chemistry B* **114**, 7830–7843 (2010).

60. Pedregosa FABIANPEDREGOSA, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).

61. Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols 2007 2:1* **2**, 31–34 (2007).

62. Diss, G. & Landry, C. R. Combining the Dihydrofolate Reductase Protein-Fragment Complementation Assay with Gene Deletions to Establish Genotype-to-Phenotype Maps of Protein Complexes and Interaction Networks. *Cold Spring Harb Protoc* **2016**, pdb.prot090035 (2016).

63. Keller, R. L. J. The Computer Aided Resonance Assignment Tutorial. (2004).

64. Klukowski, P., Riek, R. & Guntert, P. NMRtist: an online platform for automated biomolecular NMR spectra analysis. *Bioinformatics* **39**, (2023).

65. Klukowski, P., Riek, R. & Güntert, P. Rapid protein assignments and structures from raw NMR spectra with the deep learning technique ARTINA. *Nature Communications 2022 13:1* **13**, 1–12 (2022).

66. Schmidt, E. & Güntert, P. A new algorithm for reliable and general NMR resonance assignment. *J Am Chem Soc* **134**, 12817–12829 (2012).

67. Gottstein, D., Kirchner, D. K. & Güntert, P. Simultaneous single-structure and bundle representation of protein NMR structures in torsion angle space. *J Biomol NMR* **52**, 351–364 (2012).

68. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol* **179**, 125–142 (1984).

69. Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. Crystallographic Analysis of the Recognition of a Nuclear Localization Signal by the Nuclear Import Factor Karyopherin α. *Cell* **94**, 193–204 (1998).

70. Velazhahan, V. *et al.* Structure of the class D GPCR Ste2 dimer coupled to two G proteins. *Nature 2020 589:7840* **589**, 148–153 (2020).

Acknowledgments

We thank members of the Barth lab and the Rahi lab for helpful discussions. **Funding:** L.S. is supported by the EPFLglobaLeaders doctoral fellowship; an EPFL Science Seed Fund awarded to P.B. and S.J.R.; P.B. is supported by Swiss National Science Foundation grants (31003A_182263 and 310030_208179), a Novartis Foundation for medical-biological Research grant 21C195, a Swiss Cancer Research grant KFS-4687-02-2019, funds from EPFL, and the Ludwig Institute for Cancer Research; SNSF grants CRSK-3_190526, 310030_204938, and CRSK-3_221036 awarded to S.J.R. **Authors contributions:** L.S., A.C. and P.B. designed the project; L.S. and S.Z. developed the computational framework; L.S. designed and screened dynamic protein optoswitches, ran MD simulations and Mutual Information (MI) analysis, performed experimental validation of the designs. L.S. developed applications on programmable spatio-temporal controls of cellular behaviors under the supervision of S.J.R.; L.A. performed NMR spectroscopy analysis; K.L. and F.P. performed protein purification; L.S. performed biophysical characterization, aided by K.L. and F.P.; all authors participated in the analysis and interpretation of the results; L.S. and P.B. wrote the manuscript with contributions from the other authors; P.B. supervised the entire project. **Competing interests:** P.B. holds patents and provisional patent applications in the field of engineered T cell therapies and protein design. **Data and materials availability:** all data are available in the manuscript or the supplementary materials.

Methods

Open state scaffold design

Starting from a scaffold with a closed state conformation (state 2), the goal was to computationally design an open state variant (state 1), in which the C-terminal helix is unlocked from the core. To this end, we used the NMR ensemble of the *de novo* designed closed state scaffold[50]. All 20 conformers were idealized and relaxed using Rosetta, and conformer 4 - exhibiting the lowest Rosetta energy - was selected as state 1. To generate an open conformation (state 2), we employed steered molecular dynamics (MD) simulations (see Methods, molecular dynamics (MD) simulations), applying a pulling force along a reaction coordinate perpendicular to the C-terminal helix, while fixing the protein core as an anchor. Representative frames in which the C-terminal helix was displaced, yet the scaffold remained folded, were extracted from the simulation trajectory. To preserve the integrity of the core structure from any disruptions potentially introduced by the pulling force, we aligned the open state structure to state 1, and grafted the C-terminal helix (residues 112-130) from the open conformation back onto the scaffold core structure from state 1 (residues 1-109). The region linking the helix to the core (residues 110-111) was redesigned using RosettaRemodel[51]. To explore structural diversity around the open state conformation, we used the partial diffusion mode of RFdiffusion[52] to generate 10 structural models. Subsequent sequence design was performed with ProteinMPNN[53], yielding 100 unique sequences per model. These sequences were submitted without MSA generation to ColabFold for structure prediction[54]. Predicted models were ranked based on the AlphaFold2[55] confidence metric (pLDDT), and the top-scoring design representing state 2 was selected for experimental characterization.

Multistate design of switchable scaffolds

We initially applied a Rosetta-based multistate design protocol, combining atomistic modeling with a genetic algorithm to optimize amino acid sequences across multiple conformations[56]. State 1 corresponded to the open state and was generated via steered molecular dynamics simulations in combination with deep learning-guided single-state design (see Methods, single-state design). State 2 was represented by lowest energy conformer of the closed state NMR ensemble[50], after idealization and relaxation using Rosetta (see Methods, single-state design). The objective of Rosetta multistate design was to introduce mutations that minimized the energy difference (ΔE) between the two states, thereby enabling reversible switching. The following fitness function was employed, where M1 represents the initial energy of the model in the open state and M2 represents the initial energy of the model in the closed state:

$$((M_1 - E_{M1}) - (M_2 - E_{M2})) + \left(\left(\frac{(M_2 - (E_{M2} + E_{M1}))}{\sqrt{(1 + (M_2 - (E_{M2} + E_{M1}))^2)}}\right) + 1\right) * \left(\frac{(M_2 - (E_{M2} + E_{M1}))^2}{-(E_{M2} + E_{M1})}\right))$$

We also carried out multistate design using ProteinMPNN[53], a deep learning sequence design method capable of directly generating designs compatible with distinct input states.

To generate switchable sequences that modulate the conformational equilibrium specifically at the helix–core interface, we restricted the designable space to 24 residues in this region that undergo substantial environmental changes between states. To comprehensively explore and program the conformational equilibrium between states, we tested a broad range of state-weighting schemes: 100–0%, 75–25%, 50–50%, 25–75%, 10–90%, and 0–100%. For each weighting scheme, ProteinMPNN was used to generate 100 sequences, which were then subjected to further computational evaluation. First, we threaded the designed sequences back

onto both state 1 and state 2 structures using Rosetta CoupledMoves, followed by model relaxation and energy evaluation to compute the energy difference (ΔE) between the two states. Representative energy values for each weight class are shown in Figure S6a. Second, we predicted the structures of the multistate designs using AlphaFold3, assessing switchability by biasing the input with either the state 1 or state 2 model as a template. Predicted structures, colored by pLDDT confidence scores, are shown in Figure S6b. Interestingly, only switches from the 10–90% weighting scheme exhibited switchable behavior according to AlphaFold3 predictions, depending on the template provided.

Molecular dynamics (MD) simulations

Simulations were performed using GROMACS 2020.5[57,58] with the CHARMM36 all-atoms forcefield[59]. Systems were simulated in an NPT ensemble at 300 K and 1 bar, using a Nosé–Hoover thermostat for temperature control and a Parrinello–Rahman barostat with semi-isotropic pressure coupling (relaxation time: 5 ps). As for equilibrium Molecular dynamics (MD) simulations, following constrained equilibration, five independent production runs were performed for each system, each consisting of 500 ns trajectories. As for steered Molecular dynamics (MD) simulations, performed to generate open state conformations during the single-state design process, a pulling force was applied along a reaction coordinate perpendicular to the C-terminal helix, using the protein core as a fixed anchor. Simulations were carried out with a pulling rate of $0.8\,\frac{\mathrm{nm}}{\mathrm{ns}}$, and a spring constant of $1000\,\frac{\mathrm{kJ}}{\mathrm{mol\,nm^2}}$, with each SMD trajectory lasting 10 ns. To quantify dynamic correlations between residues, we computed Mutual Information (MI) between time series of Cα atom distances. The trajectories were discretized into 15 bins, and MI was calculated for each residue pair using the scikit-learn Python package[60].

Molecular cloning

The DNA sequence of the designs was synthesized by GenScript. DNA digestions were performed using restriction endonucleases (New England Biolabs, USA). PCRs were performed using high-fidelity Phusion Polymerase (New England Biolabs, USA). Cloning was performed using restriction enzyme digestion and ligation with T4 ligase (New England Biolabs) or Gibson assembly (New England Biolabs). The DNA sequences of the constructs were verified by Sanger sequencing (Microsynth AG, Switzerland).

Protein expression and purification

Plasmids were propagated in BL21(DE3) or NEB T7 Express lysY/Iq *Escherichia coli* cells cultured in LB medium, which was supplemented with kanamycin to select for plasmid-containing bacteria. Cells were cultured in Luria-Bertani (LB) medium with 50 µg/mL kanamycin at 37 °C overnight. The overnight culture was inoculated to 2-4 L of Terrific Broth (TB) medium (20 mL per L) with 50 µg/mL kanamycin and cultivated at 37 °C and 200 rpm until an optical density at 600 nm (OD600) of ~0.8-1.0 was reached. Recombinant protein expression was induced upon addition of 0.5 mM isopropyl-β-D-thio-galactopyranoside (IPTG) and the culture was further incubated at 18 °C and 200 rpm for 18 h. The cells were harvested by centrifugation at 4,000 g for 20 min at 4 °C. The cell pellets were stored at -80 °C until further purification. For purification, the harvested cell pellet was resuspended in a lysis buffer (700 mM NaCl, 20 mM HEPES, pH 7.5, protease inhibitor tablet), and lysed by sonification (2.5 min, cycle 10s, 70% amplitude). The lysate was centrifuged at 20,000 g for 20 min at 4 °C. 25 mM Imidazole was added to the supernatant, and mixed with 5 mL HisPur Ni-NTA Resin beads (Thermo Scientific) or 5 mL of Ni-Indigo resin (Cube Biotech; 25 mM imidazole was omitted in this instance) for 2 hours at 4 °C. The supernatant was loaded onto a column, washed with increasing concentrations of imidazole and eluted in elution buffer (700 mM NaCl, 250 - 500 mM imidazole, 20 mM HEPES, pH 7.5). The eluted protein solution was buffer-exchanged and further purified by Size Exclusion

Chromatography using a Superdex 200 Increase 10-300 GL column in 111 mM NaCl, 14.8 mM HEPES 7.5. Protein was concentrated to the appropriate concentrations and flash frozen in liquid N2 and finally stored at -80 °C.

Yeast strains and media

All experiments were conducted using the wild-type haploid W303 yeast strain background (*MATa ade2-1 leu2-3 ura3-1 trp1-1 his3-11,15 can1-100*). Yeast transformations were performed following the LiAc/DNA carrier/PEG (polyethylene glycol) protocol[61]. For genomic integration, 1 µg of linearized plasmid DNA was used to achieve single-copy insertion at the non-essential YCR051W locus. Cells were grown in synthetic complete (SC) medium lacking methionine (-Met), supplemented with 2% (w/v) glucose for transformation and selection purposes[62]. For preparing solid SC plates, 2% agar was added to the medium. Transformed strains were selected on drop-out plates lacking uracil. To induce expression of the protein of interest, which was placed under the control of the inducible GAL1 promoter, strains were cultured in SC -Met medium supplemented with 2% (w/v) raffinose. Protein expression was induced by the addition of 2% (w/v) galactose.

Concerning the pheromone-induced cell cycle control application, genomic integration of the construct under the control of the GAL1 promoter failed to produce sufficient arrest, likely due to low expression levels and extracellular dilution of the secreted optoswitch. To overcome this limitation, we used a high-copy 2µ plasmid to drive strong overexpression, which restored light-dependent arrest. In addition, deletion of *BAR1*[48,49], an aspartyl protease responsible for α-pheromone degradation, was necessary to prevent loss of the signal and enable sustained pathway activation.

HEK293T cell culture

HEK293T cells were cultured in DMEM (Thermo Fisher Scientific) supplemented with 10% FBS (Thermo Fisher Scientific), 50 U/ml Penicillin-Streptomycin (Thermo Fisher Scientific) at 37 °C with 5% CO_2. Cells were passaged every 2 or 3 days with 0.05% trypsin/EDTA. Lipofectamine 2000 (Thermo Fisher Scientific) was used to transiently transfect the HEK293T cells with 100 ng plasmid DNA.

sfCherry2-based split complementation assay

To monitor conformational changes of the designed protein optoswitches, we employed a split complementation assay based on the engineered sfCherry2$_{1–10/11}$ fluorescent protein, whose two fragments spontaneously associate to re-form a functional fluorescent protein when brought into close spatial proximity[42,43]. In this system, the N-terminal β-barrel of sfCherry2 (sfCherry2$_{1-10}$, 208 amino acids) is genetically fused to the N-terminus of the scaffold, while the eleventh β-strand of sfCherry2 (sfCherry2$_{11}$, 16 amino acids) is genetically fused to the C-terminus of the scaffold. Specific fluorescence signal is detected in the closed conformation of the scaffold, where the N- and C-termini are juxtaposed and allow for complementation of the sfCherry2 subunits. Constructs were expressed in HEK293T cells and in *Saccharomyces cerevisiae*. Fluorescence in HEK293 cells was quantified using a FlexStation 3 multi-mode microplate reader (Molecular Devices) with excitation at 560 nm and emission at 610 nm. Yeast fluorescence was analyzed by microscopy using a Nikon Ti2-E inverted fluorescence microscope equipped with appropriate filters for sfCherry2 detection. Fluorescence values were normalized by subtracting the background signal from the negative control, consisting of mock-transfected (HEK293T) or mock-transformed (yeast) cells carrying the empty vector.

Circular Dichroism (CD) spectroscopy

Purified protein variants were diluted to a final concentration of 10 μM in buffer containing 10 mM HEPES and 10 mM NaCl at pH 7.5. Circular dichroism (CD) spectra were acquired using a

Chirascan spectropolarimeter (Applied Photophysics) to monitor thermal denaturation. Measurements were performed from 20 °C to 98 °C in 2 °C increments, recording spectra between 200 nm and 260 nm.

Samples for solution NMR

All NMR experiments were carried out in a Bruker spectrometer operating at 800.13 MHz ^{1}H frequency and equipped with a CPTC ^{1}H,^{13}C,^{15}N 5 mm cryoprobe controlled via an Avance Neo console. All experiments were carried out at 298 K on min. 800 µM protein solutions prepared in 111 mM NaCl, 14.8 mM HEPES at pH 7.5. All spectra were acquired and processed with Bruker TopSpin 4 using standard pulse sequences, and analyzed with CARA, NMRtist/Cyana and Sparky-NMRFAM.

Protein resonance assignment and structure determination on the open state scaffold

Experiments for backbone resonance assignment consisted in standard triple resonance spectra HNCA, HN(CO)CA, HNCO, HN(CO)CA, CBCA(CO)NH and HNCACB acquired on a ca. 1 mM sample prepared from cultures grown in M9 medium supplemented with ^{13}C-enriched glucose and ^{15}N-enriched ammonium chloride. HCCH-TOCSY and ^{13}C-resolved NOESY experiments were recorded on the same sample used for backbone assignments; they served for side chain assignment and structure determination together with HNHA, ^{15}N-resolved NOESY and ^{15}N-resolved TOCSY spectra acquired on a ^{15}N-labeled sample plus 2D NOESY and 2D TOCSY spectra collected on an unlabelled sample. Spectra for backbone assignments were acquired with 40 increments in the ^{15}N dimension and 128 increments in the ^{13}C dimension, and processed with 128 and 256 points using forward linear prediction. The HCCH-TOCSY spectrum was recorded with 128 increments in the ^{13}C dimensions and processed with 256 and 512 points. ^{15}N-resolved NOESY and TOCSY spectra were acquired with 40 increments in ^{15}N and 128 increments in the indirect ^{1}H dimension, and processed with 128 and 256 points respectively. ^{1}H-^{1}H 2D-NOESY

and 2D TOCSY spectra were acquired with 512 increments in the indirect dimension, processed with 2048 points. Mixing times for NOESY spectra were 120 or 200 ms, and TOCSY spin lock times were 60 ms.

All spectra were acquired and processed with Bruker's TopSpin 4.0 using the standard pulse programs with sensitivity improvement. Backbone assignments were obtained in parallel through manual analysis with the program CARA[63] and automatically with NMRtist[64], a webserver implementation of the ARTINA[65] pipeline that uses a deep neural network to identify signals in NMR spectra, FLYA[66] for resonance assignment, and automatic NOE identification and analysis for structure determination using CYANA[67]. Both procedures converged to 90% of the $^{1}H,^{15}N$ assignments for backbone units; the 14 residues without full backbone $^{1}H,^{15}N$ assignments being residues 1 and 2, most of the C-terminal his-tag, the only proline, and the stretch Arg_{117}-Arg_{124} which covers part of the helix engineered to detach from the protein core. This stretch consists of the sequence REREEERR and is itself surrounded by other E and R residues, a complexity that prevented full assignments (not only because of the repetitive nature itself but also because ^{13}CA and ^{13}CB shifts for Arg and Glu are quite similar). However, the central region of the HSQC spectrum displays a set of crosspeaks whose ^{13}C projections in the CA and CB axes of HNCA, HNCOCA, CBCACONH and HNCACB spectra are consistent with multiple consecutive pairs of residues that have similar CA/CA-1 (55.04 ppm) and similar CB/CB-1 (27.245 ppm), all around the values expected for glutamates and arginines and hence very likely belonging to residues of the Arg_{117}-Arg_{124} stretch. While we cannot assign these spin systems to specific residues of the stretch because of the excessive overlap, we can safely assume that all the CA and CB shifts recorded apply to all 8 residues of the region. These shifts were utilized to derive the ΔCA-ΔCB values for this region shown in Fig. 3h.

Given NMRtist's accuracy on backbone assignment on our spectra, fully matching our manual assignment, we then ran it on the full set of spectra to obtain automated side chain assignments and structure determination. NMRtist's structure calculation procedure converged in the

program's two independent runs to the same structure, indicative of sufficient data quality even at the regions with partial assignments or with low numbers of NOE due to dynamics. We took Proposal 1 as the final structure, an ensemble calculated from 2060 NOE restraints and 242 dihedral angle restraints. More statistics related to structure determination are provided in supplementary table S1.

^{15}N relaxation on the closed and open states of the scaffold

Heteronuclear {^{1}H}-^{15}N NOE was measured with 128 ^{15}N increments processed with 256 points, using 16 scans, a saturation time of 6 seconds and a total recycle time of 10 seconds. ^{15}N T_1 and T_2 were measured with 128 ^{15}N increments processed with 256 points and a total recycle time of 10 seconds, with the following delays: 8, 64 (twice), 136, 232, 336, 472, 664 (twice), 800, 1200 and 1600 ms (for T_1) and 16, 32 (twice), 64, 96, 128, 160, 192 (twice), 224, 256 and 288 ms (for T_2). Spectra were acquired and processed in Bruker TopSpin and then analyzed with the relaxation module in Sparky-NMRFAM.

Fluorescence microscopy

Images were acquired using a Nikon Ti2-E inverted microscope (Nikon, Switzerland) equipped with a 60x oil, N.A. 1.40 objective and a Hamamatsu Orca-Flash 4.0 camera (Hamamatsu, Japan). The system was controlled via NIS-Elements software and maintained in focus using the Nikon Perfect Focus System. For time-lapse imaging and induction with diascopic light, a green band-pass interference filter (Nikon, Switzerland) was placed between the light source and the sample to prevent unintentional activation of the LOV2 optogenetic domain. During induction, diascopic light intensity was set to 80% of maximum.

Supplementary Figures

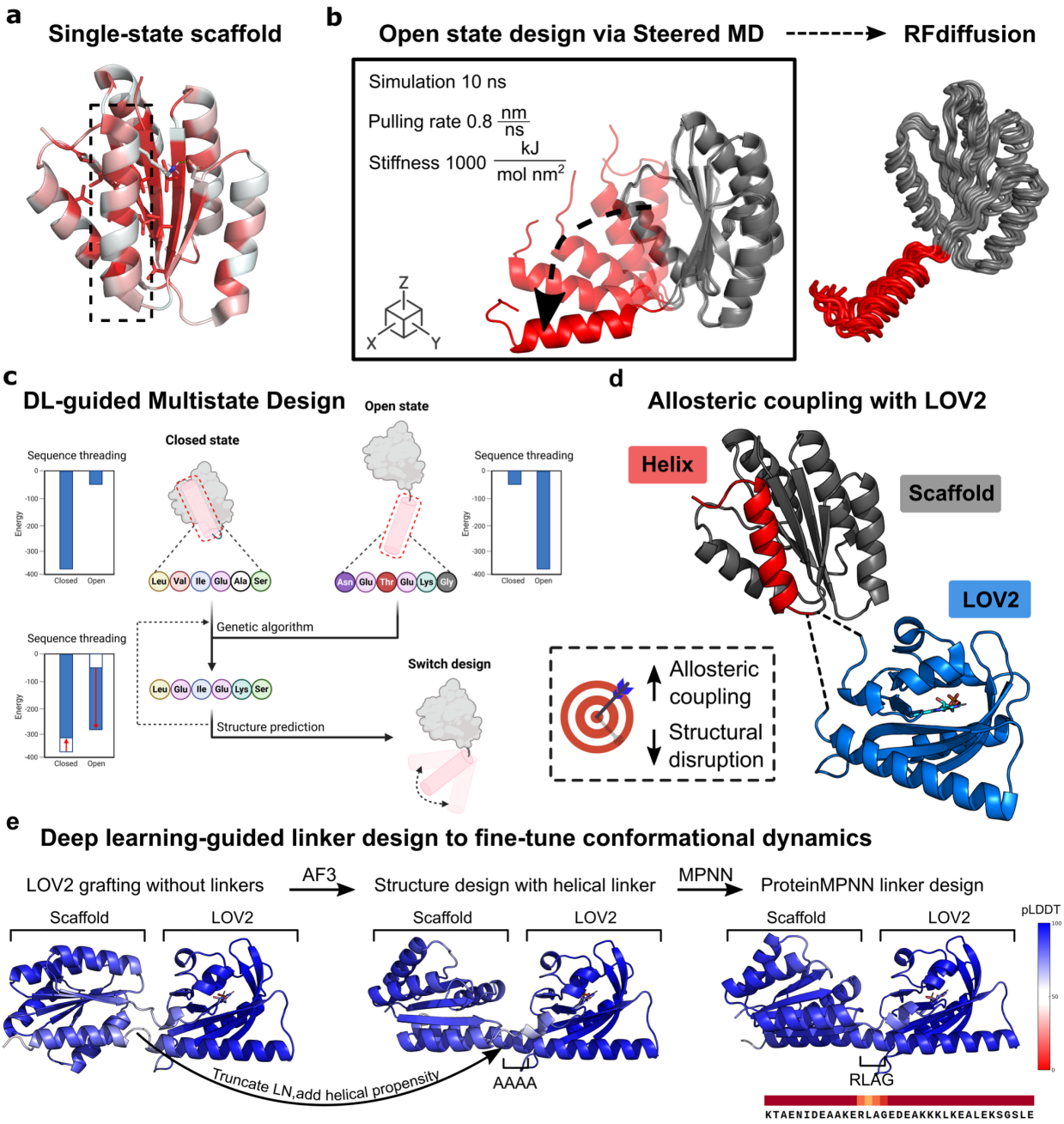


Fig. S1. Computational pipeline for the deep learning-guided *de novo* design of multi-domain optoswitchable proteins.

(**a**) Single-state *de novo* designed scaffold in the closed conformation, colored according to the Eisenberg hydrophobicity scale[68]. (**b**) Open state structure design via steered molecular dynamics (MD) simulations applying a pulling force along a reaction coordinate orthogonal to the C-terminal switchable helix. The resulting structure was subjected to design diversification using RFdiffusion through partial noising and de-noising. (**c**) Deep learning-guided multistate sequence design to engineer dynamic switchable variants capable of transitioning between closed and open conformations. The closed and open state models were input into a genetic algorithm that iteratively explored mutations aimed at minimizing the initial ΔE between states. The resulting designs were further evaluated via sequence threading. (**d**) LOV2 was allosterically coupled to the scaffold at the junction between the core scaffold and the C-terminal switchable helix to enable light control of the scaffold's conformational switching. The grafting was designed to transmit allosteric signals from the LOV2 Jα helix to the switchable helix while preserving the integrity of the scaffold's closed conformation. (**e**) Deep learning-guided linker design to fine-tune conformational dynamics. In this illustration, the original flexible junction between the Jα helix and the switchable helix was replaced with a 4-alanine stretch (high helical propensity) to introduce a helical turn and maximize allosteric coupling. AlphaFold3 was then used to predict the optoswitch structure, resulting in a helical linker connecting the Jα helix to the switchable helix. Final sequence design of the linker was performed using ProteinMPNN, keeping the rest of the protein fixed. The introduction of a continuous helical connection is expected to enhance mechanical signal transmission, thereby maximizing allosteric coupling between the LOV2 domain and the C-terminal switchable helix. A wide range of linker structures and sequences were designed using RF diffusion-Protein MPNN to explore the relationship between linker features and allosteric control of the scaffold conformational dynamics.

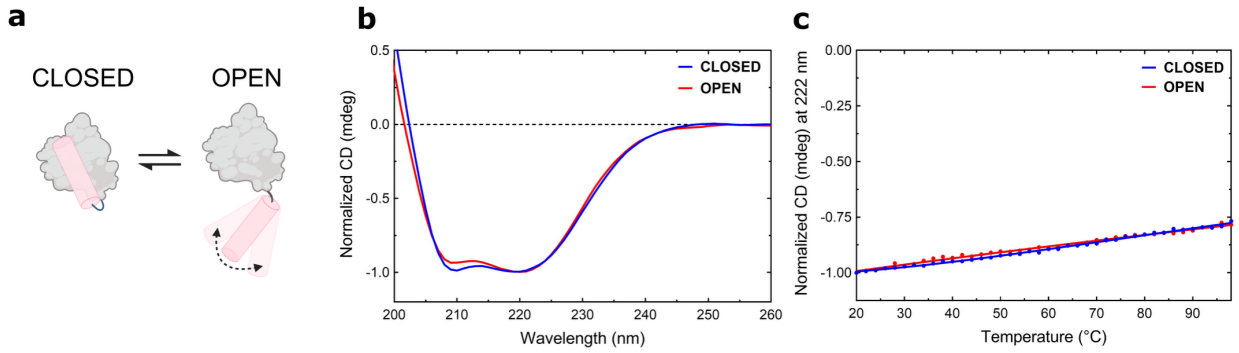


Fig. S2. Biophysical characterization of *de novo* designed scaffolds through circular dichroism and thermo-melting curve.

(**a**) Schematics of *de novo* designed scaffolds in the closed and open states. (**b**) Circular dichroism spectra of *de novo* designed scaffolds in the closed (blue) and open (red) states, measured at 20 °C. (**c**) Thermo-melting curve of *de novo* designed scaffolds in the closed (blue) and open (red) states, with the circular dichroism signals measured at 222 nm from 20 °C to 95 °C.

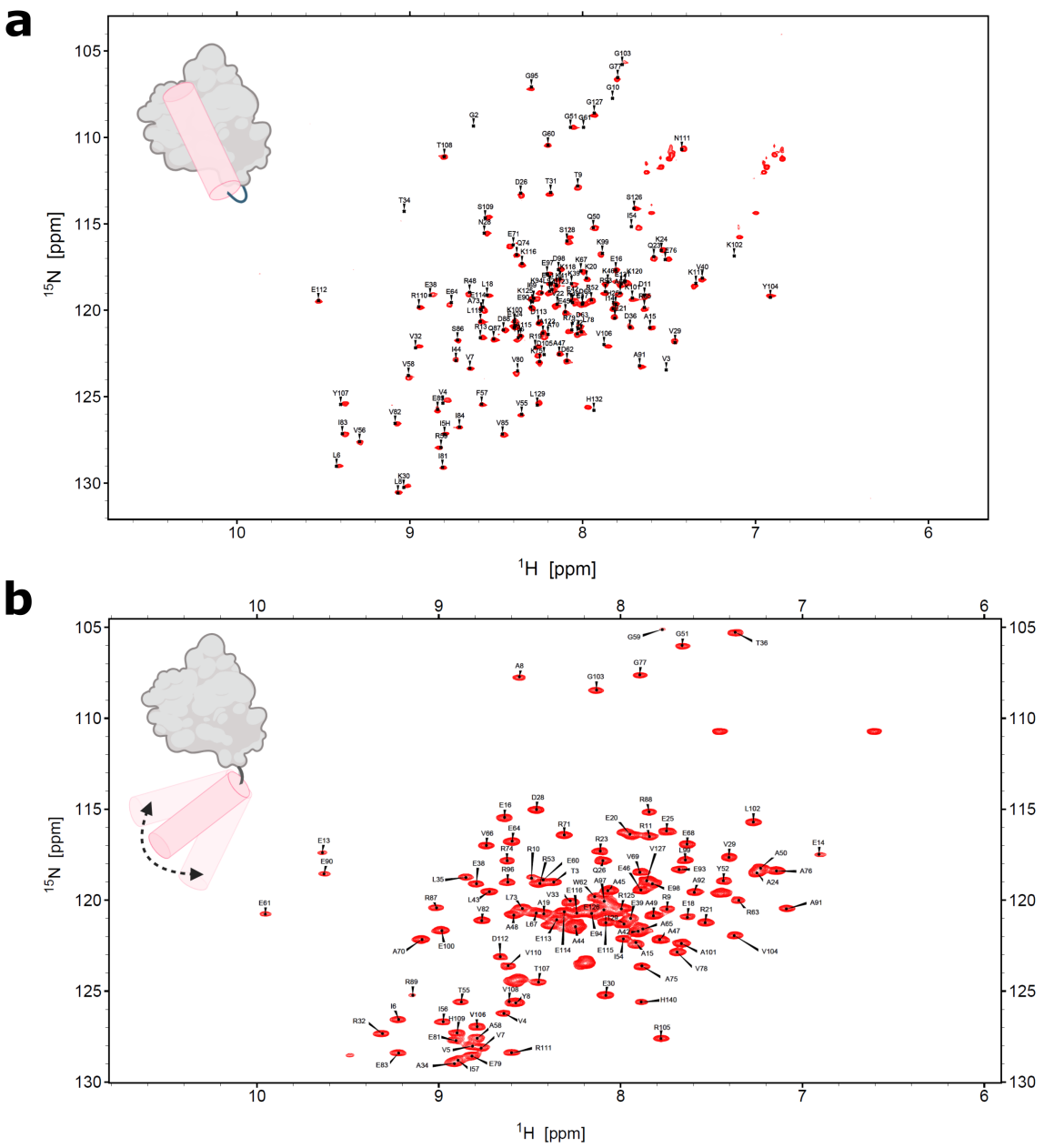


Fig. S3. 2D ^{1}H,^{15}N-HSQC spectra of *de novo* designed scaffolds.

(**a**) 2D ^{1}H,^{15}N-HSQC spectra of *de novo* designed scaffold in the closed state. The resolved peaks are superimposed with the deposited assignments from the Biological Magnetic Resonance Bank for PDB ID: 2MR6[50]. (**b**) 2D ^{1}H,^{15}N-HSQC spectra with backbone assignments of *de novo* designed scaffold in the open state. The spectra were recorded at 800.13 MHz ^{1}H frequency.

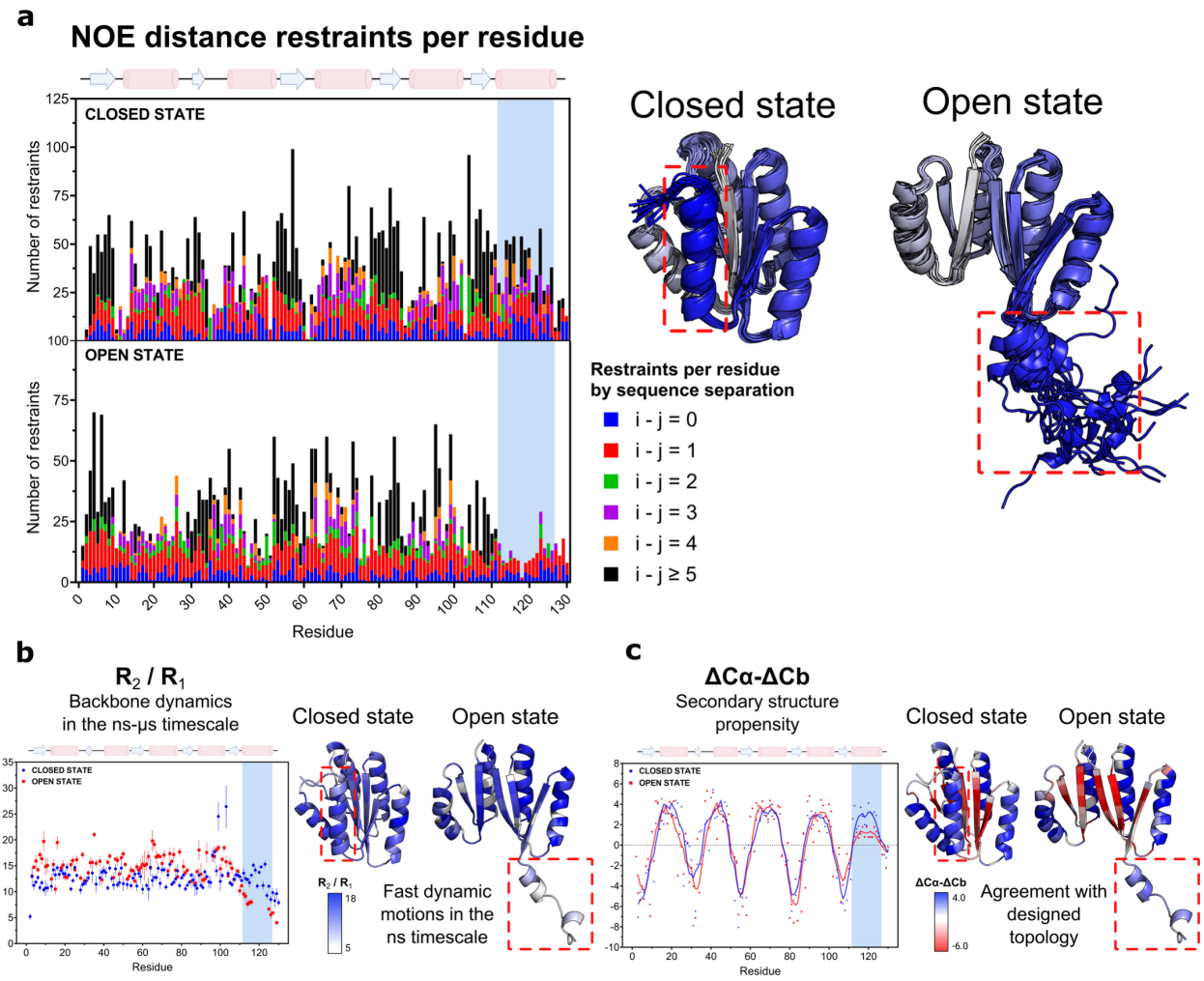
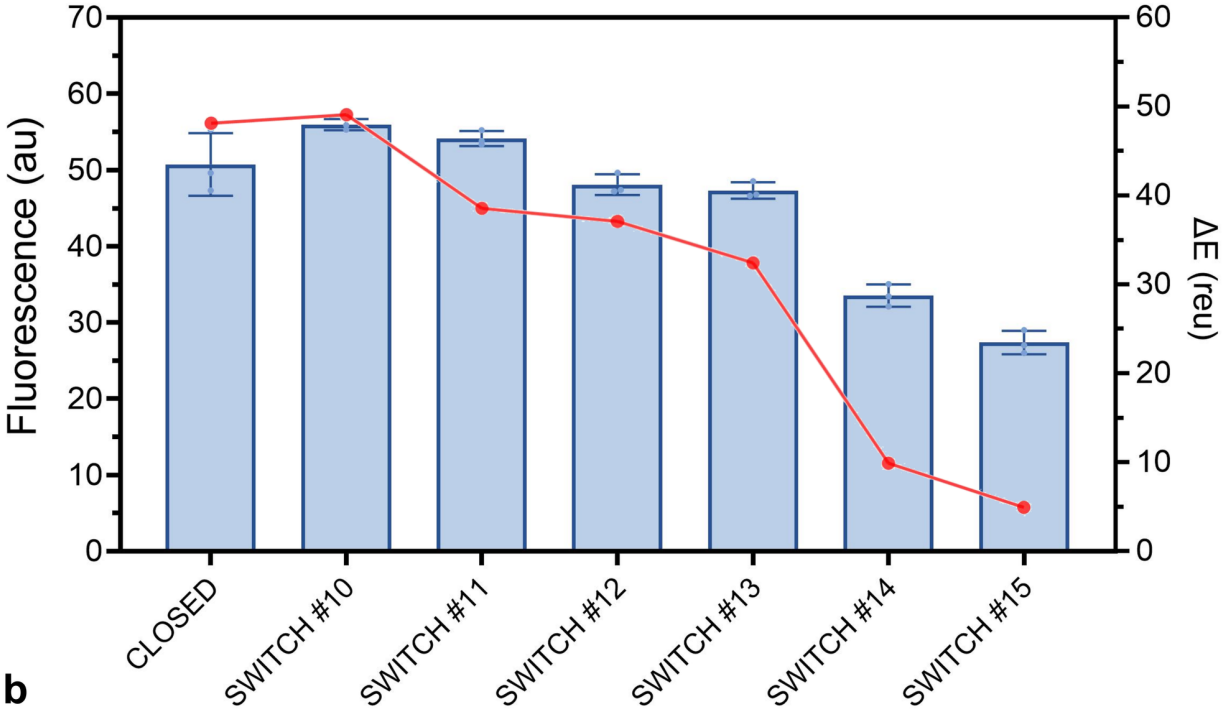

a
NOE distance restraints per residue
CLOSED STATE
OPEN STATE
Number of restraints
Residue
Closed state
Open state
Restraints per residue
by sequence separation
i - j = 0
i - j = 1
i - j = 2
i - j = 3
i - j = 4
i - j ≥ 5
b
R2 / R1
Backbone dynamics
in the ns-μs timescale
Closed state
Open state
Fast dynamic
motions in the
ns timescale
c
ΔCα-ΔCb
Secondary structure
propensity
Closed state
Open state
Agreement with
designed
topology

Fig. S4. NMR analysis of conformational equilibrium and protein dynamics.

(**a**) Number of NOE distance restraints per residue for the closed (top) and open (bottom) states of the responder scaffold, color-coded by sequence separation. The C-terminal helix in the open state (residues 112-126) lacks long-range NOEs (i.e., i - j > 3), consistent with the absence of contacts with the scaffold core. (**b**) R_2/R_1 relaxation rate ratios for the closed (blue) and open (red) states of the scaffold. Structural models are colored by per-residue R_2/R_1 values, highlighting fast-timescale dynamics of the C-terminal helix in the open state. (**c**) ΔCα-ΔCβ chemical shift indices for the closed (blue) and open (red) states of the scaffold. Structural models are colored by per-residue ΔCα-ΔCβ values. Positive values indicate helical propensity, near-zero values indicate disorder, and negative values reflect β-sheet propensity. The observed profiles show excellent agreement between the designed topology and the experimentally-observed secondary structure propensity.

a


Fluorescence (au)
ΔE (reu)
0
10
20
30
40
50
60
70
CLOSED
SWITCH #10
SWITCH #11
SWITCH #12
SWITCH #13
SWITCH #14
SWITCH #15

b

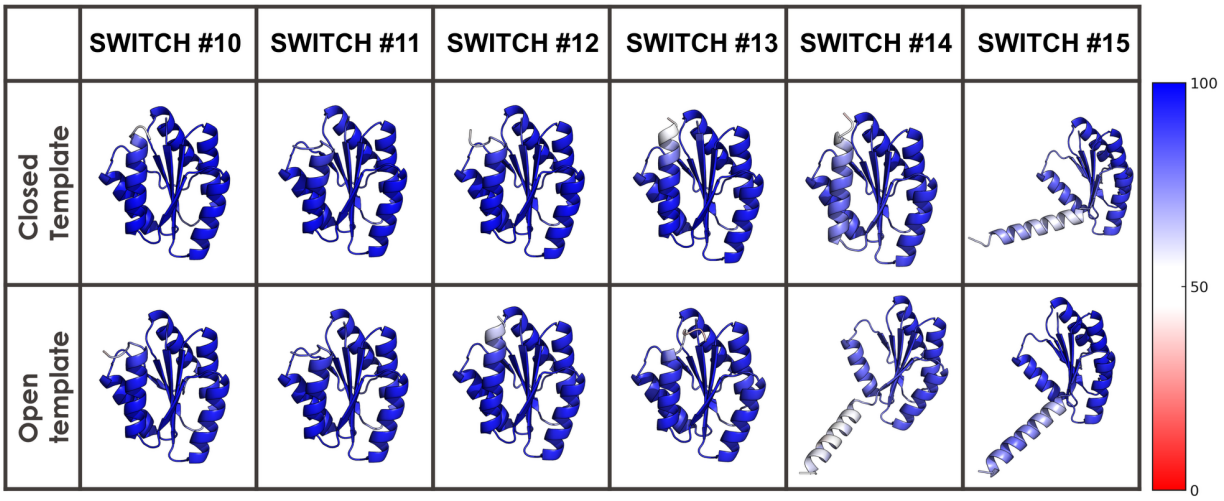

SWITCH #10
SWITCH #11
SWITCH #12
SWITCH #13
SWITCH #14
SWITCH #15
Closed Template
Open template
100
50
0

Fig. S5. ProteinMPNN-based multistate design (MSD-MPNN) mutants.

(**a**) On the left Y-axis, sfCherry2-based signal of the ProteinMPNN-based multistate design (MSD-MPNN) switches, representative of the scaffold's conformations across multiple states, in HEK293T cells. On the right Y-axis, computational calculation of ΔE between closed and open states upon sequence threading and Rosetta relax. The set of designable residues for multistate design included key residues laying at the helix-core interface of the scaffold (n = 24 residues, < 20% of the protein sequence). To explore the effect of design bias, we tested multiple state-weighting schemes (close-to-open percentages): SWITCH#10: 100–0%, SWITCH#11: 75–25%, SWITCH#12: 50–50%, SWITCH#13: 25–75%, SWITCH#14: 10–90%, SWITCH#15: 0–100%. (**b)** AlphaFold3 predictions of the ProteinMPNN-based multistate design (MSD-MPNN) switches, assessing switchability by biasing the input with either the closed state or open state model as a template. The predicted structures are colored by pLDDT confidence scores.

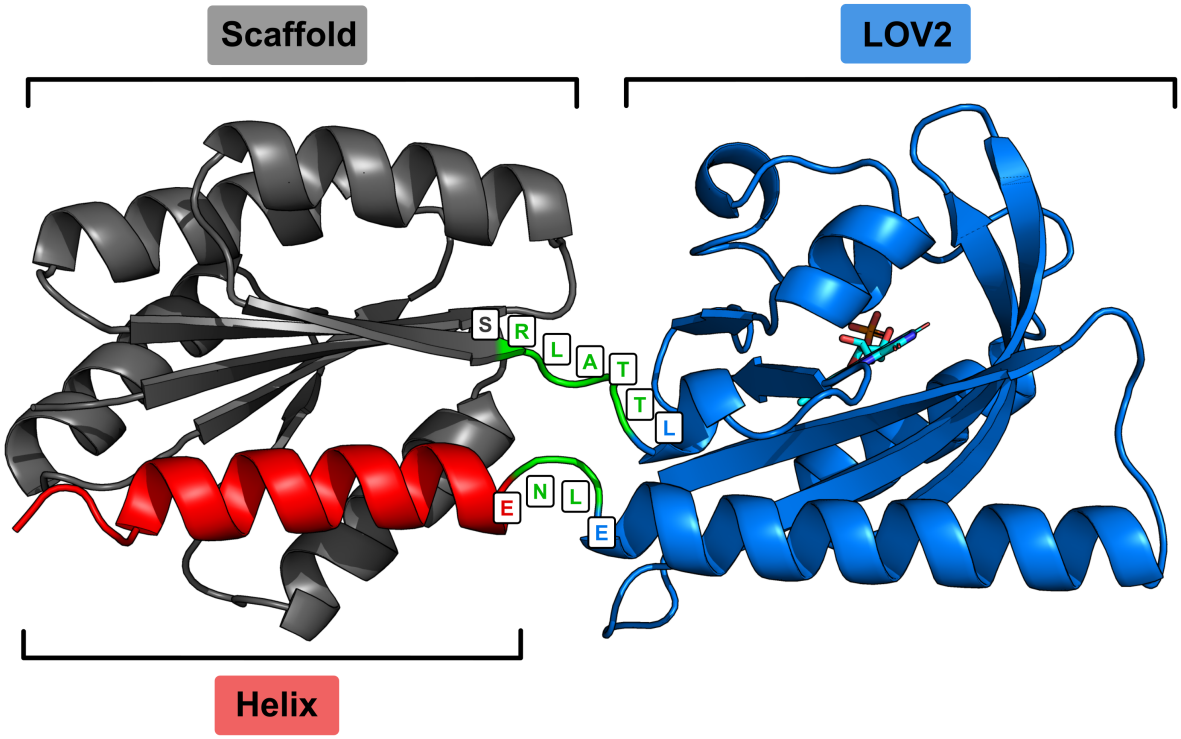


Fig. S6. LOV2 grafting into the *de novo* designed scaffold in the closed state.
Structural model of the LOV2 optogenetic domain grafted into the *de novo* designed scaffold in the closed state, inserted at the junction between the core and the C-terminal switchable helix. Residues at the insertion site prior to any linker optimization are highlighted: flexible residues in green, α-helical residues from LOV2 in blue, α-helical residues from the scaffold core in gray, and those from the switchable helix in red.

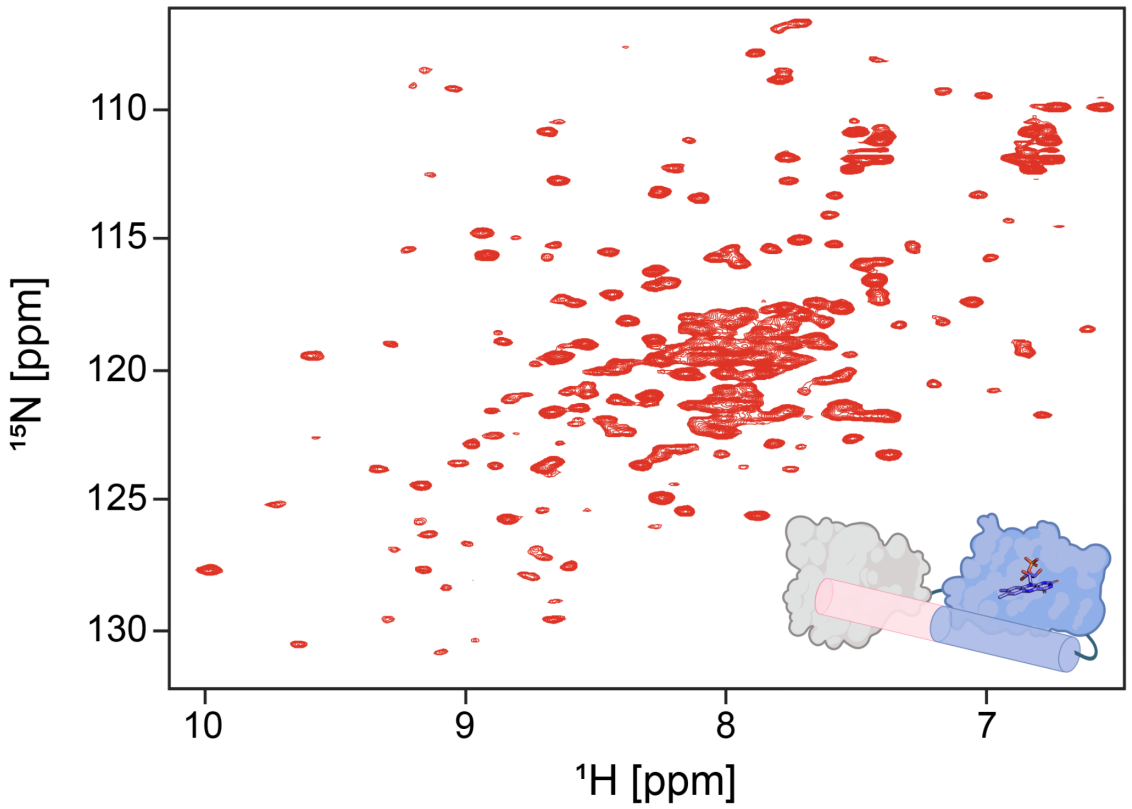


Fig. S7. 2D ^{1}H,^{15}N-HSQC spectrum of *de novo* designed multi-domain protein optoswitch. 2D ^{1}H,^{15}N-HSQC spectra of *de novo* designed multi-domain protein optoswitch. The spectrum was recorded at 25 °C in the dark and 800.13 MHz ^{1}H frequency.

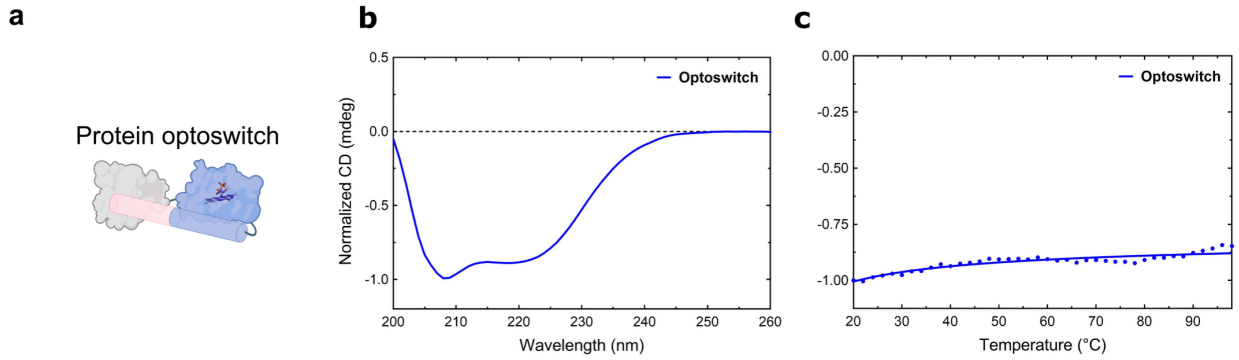


Fig. S8. Biophysical characterization of *de novo* designed multi-domain protein optoswitch through circular dichroism and thermo-melting curve.

(**a**) Schematics of *de novo* designed multi-domain protein optoswitch. (**b**) Circular dichroism spectra of *de novo* designed multi-domain protein optoswitch, measured at 20 °C. (**c**) Thermo-melting curve of *de novo* designed multi-domain protein optoswitch, with the circular dichroism signals measured at 222 nm from 20 °C to 98 °C.

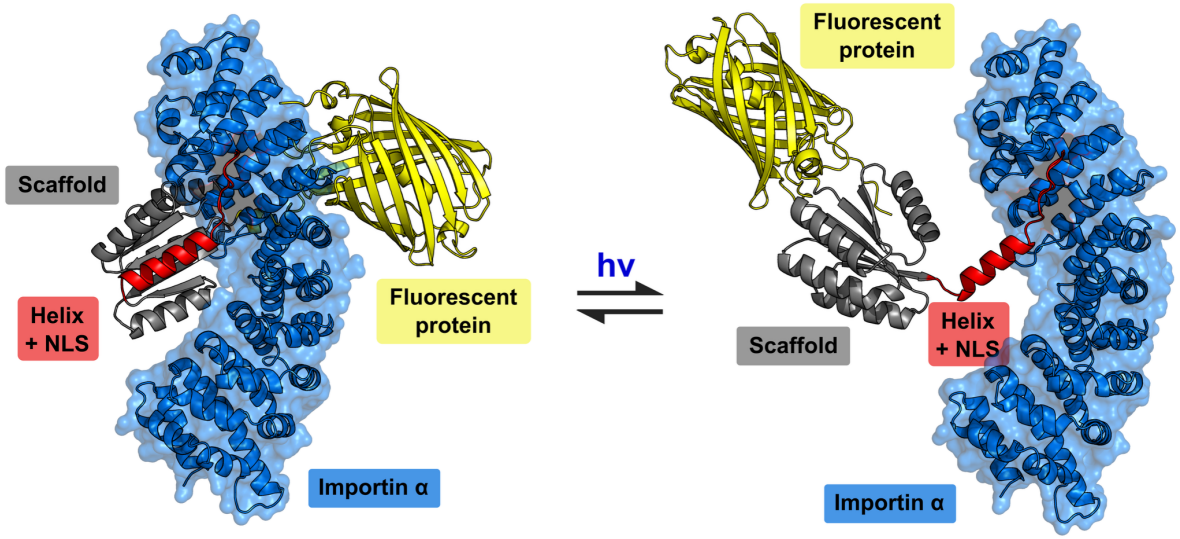


Fig. S9. Modeling NLS photo-caging.

Structural representation of NLS photo-caging in the closed (dark) and open (light) states of the multi-domain dynamic optoswitchable proteins. In the dark, the NLS is sterically occluded, preventing interaction with the nuclear import factor karyopherin α (PDB ID: 1BK6)[69]. Upon blue light absorption, the opening of the system enables NLS to become exposed and engage with the nuclear import machinery.

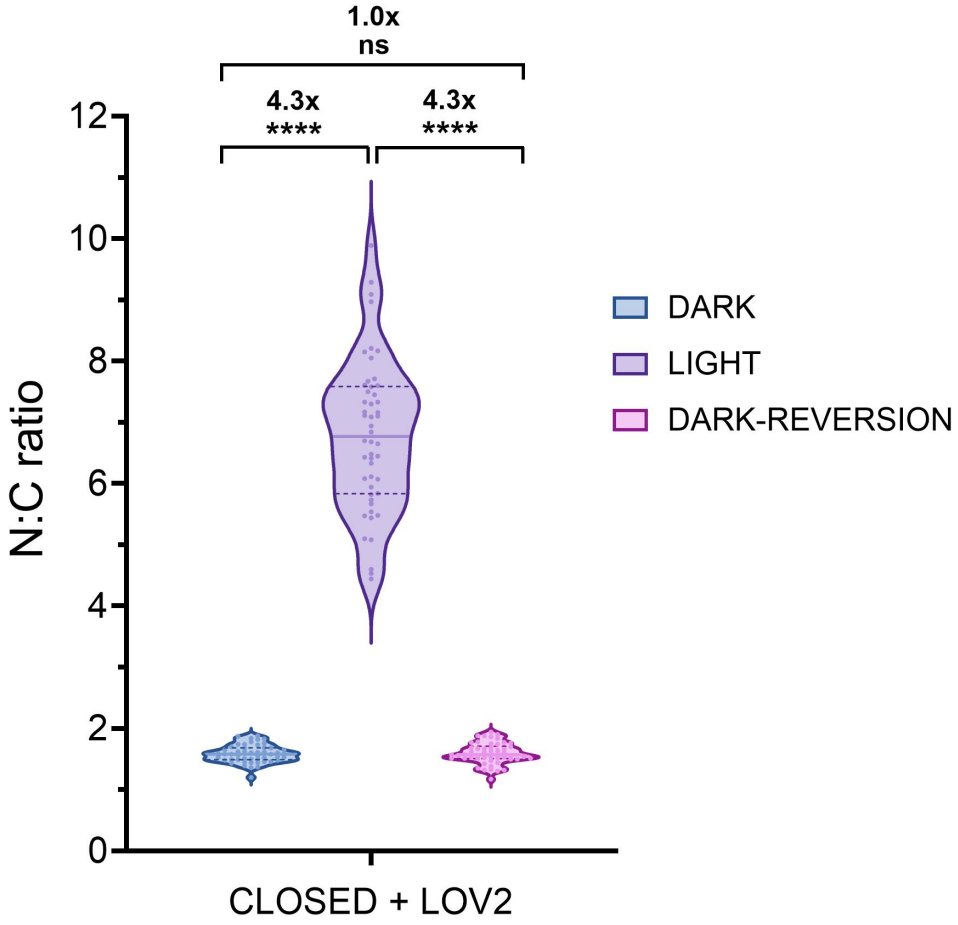


Fig. S10. Quantification of light-induced nuclear import at the population level.

Violin plot showing the nuclear-to-cytoplasmic (N:C) fluorescence intensity ratio for the dynamic multi-domain protein optoswitch derived from the closed state scaffold, functionalized with a nuclear localization signal (NLS). Measurements were performed under three conditions: dark (blue), light-induction (purple, 15'), and dark-reversion (magenta, 30'). 'ns' = not significant, ****p-value < 0.0001. n = 50 independent cells.

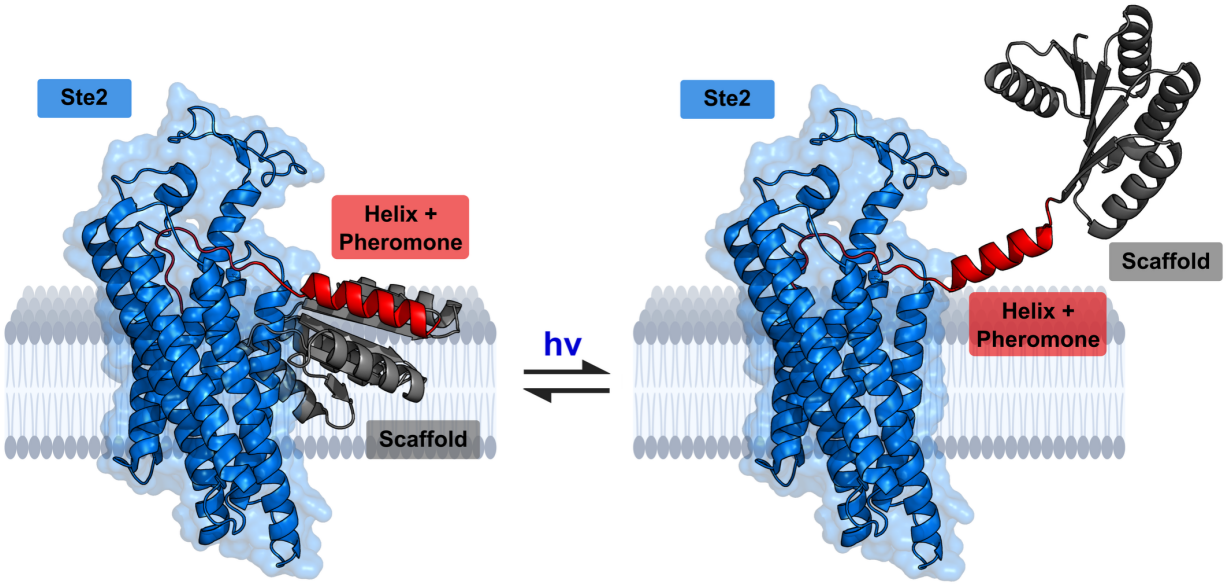


Fig. S11. Modeling α-pheromone photo-caging.

Structural representation of α-pheromone photo-caging in the closed (dark) and open (light) states of the multi-domain dynamic optoswitchable proteins. In the dark, the α-pheromone is sterically occluded and clashing with the cell membrane, preventing interaction with the class D GPCR Ste2 (PDB ID: 7AD3)[70]. Upon blue light absorption, the opening of the system enables α-pheromone to become exposed and engage with the membrane receptor.

NMR restraints	
Total NOEs	***2060***
Intraresidual	**508**
Interresidual	**1544**
Sequential (i – j = 1)	617
Medium-range (1 < i – j < 5)	403
Long-range (i – j ≥ 5)	532
Torsion angle restraints	***242***
Structure calculation statistics	
CYANA target function value	1.47 +/- 0.16 $Å^2$
Restraint Violations	
Distance restraints violations > 0.2 Å	5 ± 2
Maximal distance restraints violation	0.28 +/- 0.07 Å
Angle restraint violations	2 ± 0
Maximal dihedral angle restraints violation	5.93 +/- 0.18 °
Ramachandran plot	
Residues in most favoured region	91.7 %
Additionally allowed regions	8.2 %
Generously allowed regions	0 %
Disallowed regions	0 %

Supplementary table S1. Refinement statistics of the PIPS2 NMR structure.