

# Seesaw protein: Design of a protein that adopts interconvertible alternative functional conformations.

Toma Ikeda<sup>1,†</sup>, Tatsuya Nojima<sup>2,†</sup>, Souma Yamamoto<sup>3</sup>, Ryusei Yamada<sup>4</sup>,  
Hiroki Konno<sup>5</sup> and Hideki Taguchi<sup>1,2,\*</sup>

<sup>1</sup> School of Life Science and Technology, Tokyo Institute of Technology, S2-19, Nagatsuta 4259, Midori-ku, Yokohama, 226-8501, Japan.

<sup>2</sup> Cell Biology Center, Institute of Innovative Research, Tokyo Institute of Technology, S2-19, Nagatsuta 4259, Midori-ku, Yokohama, 226-8503, Japan.

<sup>3</sup> College of Science and Engineering, School of Biological Science and Technology, Kanazawa University, Kakuma-machi, Kanazawa 920-1192, Japan.

<sup>4</sup> Graduate School of Natural Science and Technology, Kanazawa University, Kanazawa 920-1192, Japan.

<sup>5</sup> WPI Nano Life Science Institute (WPI-NanoLSI), Kanazawa University, Kakuma-machi, Kanazawa 920-1192, Japan.

† These authors contributed equally to this work

\*Correspondence to Hideki Taguchi Email: taguchi@bio.titech.ac.jp

**Author Contributions:** T.I., T.N. S.Y., R.Y. and H.K. performed experiments; T.I., T.N. H.K. and H.T. conceived the study, designed experiments and analyzed the results; H.T. supervised the entire project; T.I., T.N., H.K and H.T. wrote the manuscript.

**Competing Interest Statement:** The authors declare no competing interest.

**Classification:** Biological sciences> Biochemistry

**Keywords:** Protein folding, metamorphic protein, GFP, DHFR, high-speed AFM

**This PDF file includes:**

Main Text

Figures 1 to 5

## **Abstract**

Classical Anfinsen's dogma states that a protein folds into a single unique conformation with minimal Gibbs energy under physiological conditions. However, recent advances have revealed that single amino acid sequences can fold into two or more conformations. Here, we propose a novel approach to design a protein that adopts interconvertible alternative functional conformations, termed "seesaw" protein (SSP). An SSP was engineered by fusing GFP lacking the C-terminal  $\beta$ -strand and DHFR lacking the N-terminal  $\beta$ -strand with an overlapping linker, which can be competitively incorporated into either the GFP or the DHFR moiety. In vivo and biochemical analysis, including AFM imaging, demonstrated that the SSP adopts two alternative conformations, which can be biased by point mutations and ligand binding. In addition, the balance of the seesaw can be reversibly changed depending on buffer conditions. In summary, our design strategy for SSP provides a new direction for creating artificial proteins with on-off behaviors.

## Introduction

The functional three-dimensional structure of a typical globular protein is established by folding the polypeptide chain, with a naturally selected specific amino acid sequence, into its unique minimum free energy state. This principle of protein structure formation—the one-to-one correspondence between a specific amino acid sequence and a specific structure—is known as Anfinsen’s dogma (1). However, the dogma that the conformation is uniquely determined by the amino acid sequence has been challenged. Protein conformation is not always uniquely defined. Pioneering work, such as the discovery of “switch peptides”, whose secondary structure varies depending on environmental factors like pH or temperature (2, 3), and the “Chameleon” sequence, where the same amino acid sequence forms distinct secondary structures when incorporated into different proteins (4, 5), has demonstrated that a particular amino acid sequence can adopt multiple secondary structures in a context-dependent manner.

In addition to those secondary structural switches, it is now clear that proteins with multiple conformations from single sequences exist in nature. For example, amyloids, which irreversibly change their conformation from a certain conformation to intermolecular  $\beta$ -sheets, are a well-known example (6). Recently, it has also been reported that there are proteins such as Morpheesins, which drastically change their tertiary structure depending on their quaternary structures (7–9), metamorphic proteins (10–12), and fold-switch proteins (13, 14), which can reversibly change their topology depending on environmental factors. In fact, more than 100 conformation-changing proteins have been reported, and about 4% of the proteins in the PDB are assumed to have multiple conformations (15). These findings imply that Anfinsen’s dogma cannot fully explain all aspects of the protein world. These proteins have been reported to undergo allosteric regulation and acquire new functions through conformational changes, suggesting that they may have physiological significance (16–18).

Beyond conformation-changing proteins in nature, the artificial design of proteins with multiple structures, including de novo design proteins (19–22), has advanced to elucidate and harness functional controls through conformational changes (23–29). Using natural proteins as templates, Solomon *et al.* designed a temperature-dependent conformational switch protein by precisely matching the sequence homology of a 56-amino-acid protein with a 3 $\alpha$  fold and a 95-amino-acid protein with an  $\alpha/\beta$ -plait topology (30). Sekhon *et al.* designed a protein in which the C-terminus of ribonuclease barnase was duplicated and fused to the

N-terminus (31). One of the C-terminal regions is inactivated by mutation, enabling the RNase activity to be controlled in an on/off manner. Sander *et al.* fused the N-terminal regions of two related fluorescent proteins, Venus and Cerulean, with a single C-terminal region that shares a common sequence (32). They succeeded in modulating the yellow-to-blue fluorescence ratio by regulating the translation elongation rate through synonymous substitutions.

Expanding on these strategies for switching conformations, a new approach to control functional switching by merging two proteins with distinct functions would be invaluable in the era of creating de novo design proteins. Consequently, we devised a method to fuse two proteins, GFP and DHFR, by overlapping their C- and N-terminal regions, thereby overlapping a segment of their C- and N-termini, respectively. We named the resulting fusion protein "Seesaw protein (SSP)" because its structure and function can be reciprocally altered by integrating either the C-terminus or the N-terminus of the design protein. The two states of the SSP were found to be adjustable by various factors, including mutagenesis and buffer conditions.

## Results

### The design of a protein capable of adopting two alternative conformations.

We designed a protein that adopts alternative functional conformations by developing a fusion method for two proteins, overlapping them at the ends. Specifically, a segment of the C-terminal region of the first protein is overlappingly fused with that of the N-terminal region of the second protein (Fig. 1A). If the overlapping region (OR) can complement both proteins, the resulting design protein is expected to possess two alternative states depending on whether the OR is incorporated into the N-terminal or C-terminal moieties. The length of the OR was adjusted to be sufficient to include essential regions for the native states and functions of both parent proteins. Consequently, the designed protein may exhibit two different functional states depending on the region of the designed protein interacting with the OR. Additionally, we anticipated that mutations altering the stability of each moiety, ligands, and buffer conditions would induce changes in the two conformations (Fig. 1B). As the two different functional states controlled by slight differences reminded us of a seesaw, we named such a fusion protein "seesaw protein (SSP)".

## Creation of a “seesaw” protein (SSP<sub>GD</sub>) using GFP and DHFR.

To create an SSP, we employed superfolder GFP and *Escherichia coli* dihydrofolate reductase (DHFR) as parent proteins (Fig. 2A). These proteins were chosen based on three criteria. First, they were selected for being monomeric proteins to simplify the equilibrium state of each moiety. Second, to facilitate the detection of the two different functional states, proteins with measurable function/activity in vivo were preferred: the GFP state can be evaluated via fluorescence imaging. For DHFR, we devised a colony assay system utilizing a DHFR inhibitor, trimethoprim (TMP). Specifically, under conditions where endogenous DHFR, essential for growth, is inhibited by TMP and therefore unable to promote growth, the growth defect is complemented only when exogenously expressed active DHFR is present (SI Appendix, Fig. S1A). Third, proteins with relatively high stability in the wild type, for which mutants with varying stabilities have been extensively studied, were chosen.

A GFP-DHFR fusion protein was engineered to include a 13-amino-acid long OR that exclusively shares either the C-terminal strand (218-230) of GFP or the N-terminal strand (2-14) of DHFR (Fig. 2B and SI Appendix, Fig. S1B). Initially, we confirmed that monomeric GFP and DHFR lacking the OR (GFP $\Delta$ C and DHFR $\Delta$ N) exhibited loss of fluorescence and were not expressed in cells in the presence of TMP, respectively (SI Appendix, Fig. S1C). Following a series of chimera analyses (SI Appendix, Fig. S1B, D), we found that utilizing the sequence MSLIAALAVDRVI as OR, where the first M is derived from GFP, the third L is common, but the remainder is from DHFR, resulting in both GFP fluorescence and DHFR activity in the colony assay (Fig. 2B, C). We verified that the full-length fusion protein was expressed in a soluble fraction and that degradation, which could produce monomeric DHFR or GFP, was negligible (SI Appendix, Fig. S1E), suggesting that the degradation products of the fusion are unlikely to contribute to DHFR resistance. Collectively, these results suggest that the GFP-DHFR fusion protein (SSP<sub>GD</sub>) adopts two alternative functions: the GFP state, in which the GFP region incorporates the OR, and the DHFR state, in which the DHFR region incorporates the OR.

To alter the balance between the two functions in SSP<sub>GD</sub>, mutations were introduced to destabilize one state and stabilize the other (Fig. 2D). To stabilize the DHFR moiety, we substituted Met 218 with Ile (SSP<sub>GD218</sub>), which is the original residue in the corresponding DHFR region (Fig. 2D). For GFP stabilization, the 222<sup>nd</sup> residue, critical for GFP stability, was replaced with the GFP-derived

residue (SSP<sub>GD222</sub>) (Fig. 2D). The results revealed that when the SSP mutants were expressed in *E. coli*, *E. coli* expressing SSP<sub>GD218</sub> grew in the presence of higher concentrations of TMP compared to SSP<sub>GD</sub> and nearly lost GFP fluorescence (Fig. 2E and SI Appendix, Fig. S1E). Conversely, *E. coli* expressing SSP<sub>GD222</sub> failed to grow in the presence of TMP, but exhibited enhanced GFP fluorescence compared to SSP<sub>GD</sub> (Fig. 2E and SI Appendix, Fig. S1E). Thus, single amino acid mutations in the OR appear sufficient to switch between the two different functions.

Subsequently, we examined whether mutations outside the OR could influence the balance of the seesaw. To investigate this, we reverted the mutations introduced for the stabilization in sfGFP, all outside the OR, to wild-type GFP residues in SSP<sub>GD222</sub>, a GFP-biased SSP (33). Out of the six mutations tested, four point mutations exhibited both GFP fluorescence and DHFR activities (SI Appendix, Fig. S1F, G), suggesting the existence of multiple pathways for altering the seesaw balance.

### **The purified SSP<sub>GD</sub> exhibits two alternative native conformations.**

We conducted characterization of the SSPs purified from *E. coli*. Prior to analyzing the SSP fusion proteins, we evaluated the monomeric GFP and DHFR proteins, which replaced the C- or N-terminal regions, respectively, with the OR in SSP<sub>GD</sub>. The DHFR activity of DHFR $\Delta$ N with OR (OR-DHFR $\Delta$ N) closely resembled that of the wild-type DHFR monomer (SI Appendix, Fig. S2A). In contrast, the fluorescence of GFP $\Delta$ C with OR (GFP $\Delta$ C-OR) was only a few percent of that of GFP (SI Appendix, Fig. S2A), suggesting that substituting the C-terminal strand with a strand mostly derived from the DHFR sequence significantly compromised the fluorescence property of GFP $\Delta$ C-OR. Subsequently, we quantified the activity/fluorescence of SSP<sub>GD</sub>, namely GFP $\Delta$ C-OR-DHFR $\Delta$ N, in comparison to the monomeric counterparts (Fig. 3A). The DHFR activity of SSP<sub>GD</sub> was approximately 5% of that of OR-DHFR $\Delta$ N monomer, whereas the fluorescence of SSP<sub>GD</sub> was about 85% of that of the GFP $\Delta$ C-OR monomer. The diminished DHFR activity was not solely attributed to the fusion of DHFR and GFP, as the conventional fusion of GFP and DHFR with a Gly-Ser linker did not significantly reduce the DHFR activity (SI Appendix, Fig. S2A), suggesting that the proportion of the DHFR moiety in SSP<sub>GD</sub> would be low.

Next, we investigated whether SSP<sub>GD</sub> adopts two alternative conformations using bio-layer interferometry (BLI). We prepared a nanobody (LaG16) that binds



to the native GFP structure and immobilized it on a BLI biosensor on an Octet system. The GFP nanobody bound to SSP<sub>GD</sub> and SSP<sub>GD222</sub>, but not SSP<sub>GD218</sub>, indicating that SSP<sub>GD</sub> and SSP<sub>GD222</sub> possessed a native GFP conformation, which was absent in SSP<sub>GD218</sub> (Fig. 3B left). Concerning the DHFR moiety, we used a DHFR inhibitor, methotrexate (MTX), as it is known to specifically bind to the native state of DHFR (34, 35). BLI analysis using a biosensor immobilized with MTX revealed that SSP<sub>GD</sub> and SSP<sub>GD218</sub> were bound to the MTX-immobilized sensor (Fig. 3B right), indicating that SSP<sub>GD</sub> and SSP<sub>GD218</sub> possess native DHFR moiety. In summary, BLI analysis demonstrated that SSP<sub>GD</sub> and its variants have moieties with native conformations depending on the seesaw balance.

To further confirm the existence of two alternative conformations, we employed a limited proteolysis approach to distinguish between them. Typically, unstructured regions are susceptible to proteases, while stable folded structures remain intact. Under conditions where proteinase K did not degrade monomeric GFP or DHFR, the protease treatment completely digested the full-length SSP<sub>GD</sub> and its variants (42 kD) (Fig. 3C). The protease-resistant fragments were approximately 26 kD (SSP<sub>GD</sub> and SSP<sub>GD222</sub>), 21 kD (SSP<sub>GD</sub> and SSP<sub>GD218</sub>), 17 and 14 kD (SSP<sub>GD218</sub>) (Fig. 3C). Edman degradation analysis and predicted molecular weights identified that the 26 and 21 kD fragments originated from the intact GFP and DHFR moieties, respectively (SI Appendix, Fig. S2B). The 17 and 14 kD fragments in SSP<sub>GD218</sub> included the N-terminal half of GFP (SI Appendix, Fig. S2B), indicating that the GFP moiety in SSP<sub>GD218</sub> possesses a folded structure resistant to proteolysis. Limited proteolysis did not alter the intensity of GFP fluorescence or DHFR activity (SI Appendix, Fig. S2C), supporting the degradation of only the unstructured regions in SSP<sub>GD</sub> and its variants, without affecting GFP fluorescence or DHFR activity. The predominance of GFP- and DHFR-containing fragments in SSP<sub>GD222</sub> and SSP<sub>GD218</sub>, respectively, suggests an extremely biased seesaw balance in the mutants, consistent with in vivo assay and BLI analysis. Conversely, SSP<sub>GD</sub> predominantly produced a 26 kD fragment (GFP moiety) and a minor 21 kD fragment (DHFR moiety), indicating that SSP<sub>GD</sub> adopts two alternative conformations with a strongly biased GFP moiety.

Based on the ratio of the two bands in SSP<sub>GD</sub> (~9.5:~0.5) we calculated the Gibbs energy change ( $\Delta G$ ) for the seesaw slope (DHFR/GFP) to be +1.8 kcal/mol (Fig. 3D). Concerning the mutants, differences in thermodynamic energy between “wild-type” SSP<sub>GD</sub> and the mutants ( $\Delta\Delta G$ ) were predicted by FoldX (36, 37). The estimated net  $\Delta G$  values for SSP<sub>GD218</sub> and SSP<sub>GD222</sub> were -1.9 and +5.5 kcal/mol,

respectively (Fig. 3D).

### **Direct visualization of the two distinct conformations in SSPs by high-speed atomic force microscopy (HS-AFM).**

To ascertain that SSPs indeed adopt two distinct conformations, we utilized high-speed atomic force microscopy (HS-AFM) to directly visualize the alternative structures of SSPs. HS-AFM imaging of SSPs immobilized on a mica surface distinctly revealed that SSP<sub>GD222</sub> displayed a single large domain structure, whereas SSP<sub>GD218</sub> exhibited a structure with two connected domains - one large and one small (Fig. 4 left and right, Movies S1, S2). These structures correspond to an intact DHFR and a protease-resistant segment of GFP, respectively. HS-AFM observations of SSP<sub>GD</sub> predominantly displayed SSP<sub>GD222</sub>-like structures, occasionally featuring SSP<sub>GD218</sub>-like structures consisting of two domains (Fig. 4 middle, Movies S3, S4), consistent with the ratio estimated from the limited proteolysis experiment. Furthermore, the visualization of SSP<sub>GD</sub> and its variants clearly demonstrated that SSPs can exhibit two distinct states in their monomeric states, thus excluding the possibility of conformational conversion due to oligomer formation. Taken together, we conclude that SSP<sub>GD</sub> indeed adopts two alternative conformations, which can be biased by single mutations.

### **The seesaw balance can be altered by buffer conditions and ligand binding.**

Finally, we aimed to alter the seesaw balance of SSP<sub>GD</sub> using factors other than mutations. We found buffer conditions that could reversibly shift the seesaw balance of SSP<sub>GD</sub> (Fig. 5A). Through limited proteolysis and DHFR activity measurements, we found that lower salt concentrations and high pH level favored an exclusive GFP state in SSP<sub>GD</sub>, resulting in the loss of DHFR activity (Fig. 5B). Conversely, exchanging the buffer to lower pH and increase the salt concentrations significantly enhanced DHFR activity while reducing GFP fluorescence (Fig. 5B). This shift in SSP<sub>GD</sub> balance due to buffer exchange was observed repeatedly (Fig. 5B). Additionally, HS-AFM images confirmed that SSP<sub>GD</sub> exhibited distinct two-domain shapes under DHFR-biased conditions (pH 6 and 500 mM NaCl) (Fig. 5C, Movies S5, S6).

Next, we added MTX to probe the native state of the DHFR moiety of SSP<sub>GD</sub>. Evaluation of the GFP-to-DHFR state ratio in the presence of MTX through limited proteolysis showed a gradual shift from 9:1 to approximately 1:9 after 4 h (SI Appendix, Fig. S3A, B). As MTX specifically binds to the native DHFR state (34,



35), which is nearly irreversible within the timeframe employed here, the kinetics likely reflect the transition between the two states in a dynamic equilibrium.

## Discussion

Here, we proposed the concept of a fusion protein that adopts two interconvertible functional conformations, termed seesaw protein (SSP). This was specifically engineered by fusing GFP, which lacks the C-terminal  $\beta$ -strand, and DHFR, which lacks the N-terminal  $\beta$ -strand, with an overlapping linker (SSP<sub>GD</sub>). Various in vivo to in vitro analyses, including HS-AFM visualization, support the presence of two distinct states in SSP<sub>GD</sub>.

Previous studies have shown that proteins with drastic morphological switches exist in nature or can be artificially created by de novo design (7, 10–14, 23–25). SSP<sub>GD</sub> can be considered an addition to the lineage of these metamorphic proteins, with the following distinguishing features: First, two distinct proteins with completely different functions and topologies were fused; second, their different functions were designed to be evaluated in vivo; and finally, a peptide of only 13 amino acids as an overlapping region (OR) in SSP<sub>GD</sub> was sufficient to tune the seesaw balance.

Proteins that switch conformations have been designed by dividing the structures of GFP (32, 38, 39) and ribonuclease barnase (31) into two regions, duplicating one region and fusing it to the end, and then replacing the unduplicated region. However, there has been no attempt to design a protein with completely different conformations that compete for the shared region.

The key factors that enabled SSP<sub>GD</sub> to exhibit seesaw-like behavior include the choice of proteins and the optimization of the OR. GFP and DHFR, both extensively studied, are monomeric, functional proteins with established systems for measuring function in vivo (40). Both proteins are well expressed in *E. coli* and fold efficiently. They are robust to mutagenesis and truncation; for example, Ala insertion at the N-terminus of DHFR has little effect on folding efficiency (41), indicating mutation tolerance when DHFR is placed on the C-terminal side in SSP. GFP is suitable for SSP because the split GFP system allows the C-terminal  $\beta$ -strand ( $\beta$ 11) and the other region ( $\beta$ 1-10) to be expressed separately and then assembled. Additionally, GFP loses fluorescence if truncated at 226 residues and does not emit fluorescence if the OR is not incorporated (42).

Another key point is the optimization of the OR. In this design, the OR was primarily DHFR-like in sequence to allow SSP to have two different functions: the

DHFR activity within SSP<sub>GD</sub> was almost identical to that of the corresponding monomer, whereas the fluorescent properties of GFP were considerably compromised. For example, it has been reported that the E222Q mutant of eGFP slows the maturation of fluorescence by 1/25 and stabilizes the three-dimensional structure (43, 44). This may explain why GFP fluorescence in SSP<sub>GD</sub> is only ~2% of that of wild-type sfGFP.

This impaired GFP moiety is also linked to another issue. When examining the reversible seesaw balance change under altered buffer conditions (Fig. 5B), the entire protein structures were switched as evaluated by limited proteolysis, but the decrease in GFP fluorescence was only about 20%. This suggests that the GFP moiety of SSP<sub>GD</sub> comprises two populations: a minor fluorescent GFP moiety, which is relatively resistant to seesaw changes, and a major non-fluorescent GFP moiety, which changes the balance easily.

In the future, the SSP principle could be employed to create protein switches within single polypeptide chains that exclusively toggle between completely different functions.

## Materials and Methods

### ***Plasmid construction and cloning.***

Plasmids were constructed by inserting genes into the corresponding sites of the pET21c(+) vector using standard cloning procedures and Gibson assembly. Each plasmid was amplified in *E. coli* in LB medium containing appropriate antibiotics. Wild-type GFP lacking 229-238 in the C-terminus, which does not decrease fluorescence intensity or change excitation and emission peak (42), was used for SSP<sub>GD</sub> and its variants. Superfolder GFP, evolved from wild-type GFP and lacking residues 230-238 in the C-terminus, was used as full-length GFP in this study. An *E. coli* DHFR mutant, ASDHFR (DHFR-C85A/C152S), functionally and structurally equivalent to wild-type DHFR (45), was used to engineer SSP<sub>GD</sub> and its variants. The SSP mutants, SSP<sub>GD218</sub> and SSP<sub>GD222</sub>, were engineered by site-directed mutagenesis. Fusion proteins, where full-length GFP and DHFR were fused with a Gly-Ser-Gly-Gly-Gly-Gly-Ser linker, were used for control experiments. For SSP purification, proteins were tagged with 6×His at the C-termini.

### ***in vivo analysis.***

*E. coli* BL21(DE3) cells transformed with plasmids were cultured on LB agar plates. Five colonies were selected and transferred to LB medium, then incubated at 37 °C for 6 h. The cultures were streaked on LB agar plates containing 0.1 mM IPTG and 0, 1, 4 µg/ml TMP, then incubated at 18 °C for 72 h. Fluorescence at 510 nm of the colonies on plates was detected using a LAS4000 image analyzer (Fujifilm).

To assess protein expression and solubility in vivo, cells grown to an OD<sub>600</sub> of approximately 0.5 at 37 °C were further cultured overnight at 18 °C with 0.1 mM IPTG. Harvested cells were suspended in lysis buffer (20 mM Tris-HCl pH 8.0 and 1 mM EDTA). Total lysates, obtained by sonication (Branson), were centrifuged at 15,000 g for 5 min. The supernatants and the uncentrifuged lysates were subjected to SDS-PAGE. The gel was stained with CBB and imaged using a LAS4000 image analyzer (Fujifilm).

### ***Protein purification.***

To purify the sfGFP, DHFR, SSP<sub>GD</sub>, SSP<sub>GD218</sub>, and SSP<sub>GD222</sub>, these proteins were overexpressed in *E. coli* BL21(DE3) cells under the same conditions as described above. The harvested cells were lysed by sonication in 20 mM

imidazole buffer (50 mM Tris-HCl pH 7.5, 20 mM imidazole, 200 mM NaCl, and 1 mM DTT). Since all proteins were soluble in *E. coli* cells, supernatants of the lysates were collected by low-speed centrifugation (10,000 g, 5 min, 4 °C) to remove the pellet fraction. Subsequently, another centrifugation at high speed (30,000 rpm, 30 min, 4 °C) was performed to eliminate the remaining pellet. The supernatants were then applied to Ni-NTA agarose (Qiagen) columns that had been pre-equilibrated with 20 mM imidazole buffer. The column was washed with 10 column volumes of 20 mM imidazole buffer. Proteins were eluted using a step gradient of 40-200 mM imidazole in 50 mM Tris-HCl pH 7.5, 200 mM NaCl, and 1 mM DTT. The eluted fractions were analyzed by SDS-PAGE. The buffer content of the fractions containing the target proteins was exchanged to HKM buffer (25 mM HEPES-KOH pH 7.5, 100 mM KCl, 5 mM MgCl<sub>2</sub>) through repeated concentration and dilution using an ultrafiltration apparatus. The purified proteins were then stored at -80 °C. Protein concentrations were determined by absorbance at 280 nm using Jasco V-750 spectrophotometer, with estimated extinction coefficients obtained from Expasy ProtParam.

### ***Measurement of GFP fluorescence and DHFR enzyme activity.***

GFP fluorescence was measured as previously described (33). Purified proteins were diluted in HKMD buffer (25 mM HEPES-KOH pH7.5, 100 mM KCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT) to a final concentration of 0.25 µM in a 96-well plate. Fluorescence intensity at 513 nm was measured with an excitation wavelength of 480 nm. DHFR activity was measured as previously described (46). Purified proteins were added to a 96-well plate at a final concentration of 0.2 µM in HKMD buffer containing 400 µM dihydrofolate (DHF) and 400 µM NADPH. The concentrations of DHF and NADPH were determined by absorbance at 282 nm and 340 nm, respectively. The decrease in NADPH was measured spectroscopically at 340 nm ( $\epsilon_{340} = 11,800 \text{ M}^{-1} \text{ cm}^{-1}$ ). Both fluorescence and enzyme activity measurements were performed using a Varioskan LUX Multimode Microplate Reader (Thermo Fisher Scientific). Following these activity measurements, absorbance at 280 nm (protein concentration) and 490 nm (GFP chromophore) was measured on a Jasco V-750 spectrophotometer. Dilution errors were corrected using the estimated molar absorption coefficient.

### ***Limited proteolysis.***

Purified proteins were diluted to 5 µM in HKMD buffer. Proteinase K (Merck) was

added to a final concentration of 1  $\mu\text{g/ml}$  and incubated at 18  $^{\circ}\text{C}$  for 40 min. The reaction was then quenched with phenylmethanesulfonyl fluoride (PMSF) at a final concentration of 5 mM and separated by SDS-PAGE. The gels were stained with CBB. For N-terminal amino acid sequencing of proteolyzed protein fragments, proteins were transferred onto a polyvinylidene difluoride membrane and analyzed with a gas phase peptide sequencer (PPSQ-21, Shimadzu).

To monitor the transition of  $\text{SSP}_{\text{GD}}$  to a DHFR-biased state,  $\text{SSP}_{\text{GD}}$  (10  $\mu\text{M}$ ) was incubated in HKMD buffer in the presence of 50  $\mu\text{M}$  MTX at 30  $^{\circ}\text{C}$ . At the indicated times, aliquots were transferred to 18  $^{\circ}\text{C}$  for 1 min, and proteinase K was then added to a final concentration of 10  $\mu\text{g/ml}$  for 5 min. PMSF was added to a final concentration of 5 mM to quench proteolysis, and the proteins were separated by SDS-PAGE. Gels were stained with CBB and photographed using a LAS4000 image analyzer (Fujifilm).

### ***Bilayer interferometry (BLI).***

The binding of proteins to LaG16 or MTX was measured using the Octet K2 (Pall FortéBio, Menlo Park, CA). All measurements were conducted at 30  $^{\circ}\text{C}$  with an agitation rate of 1,000 rpm. An Amine Reactive Second-Generation (AR2G) sensor was used to measure the interaction between LaG16 and SSPs. LaG16 were immobilized onto the AR2G biosensor as follows: each AR2G sensor was equilibrated in NaOAC pH 6 buffer for 2 min, then activated with a buffer containing 20 mM 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) and 10 mM N-hydroxysuccinimide. The activated sensors were modified with the amine group of 1  $\mu\text{M}$  LaG16, then quenched with 500 mM Tris-HCl pH 7.5. After establishing a baseline in Kinetics buffer (PBS supplemented with 0.1% BSA and 0.02% Tween20), the sensor was dipped into Kinetics buffer containing 50 nM SSPs for 600 s, then into Kinetics buffer without proteins for 300 s. A streptavidin (SA) sensor was used for the protein-MTX binding experiment. The sensor was constructed by loading biotin- $\text{NH}_2$  onto the SA sensor, followed by a reaction in 10 mM NaOAC (pH 6) containing 1 mM MTX and 20 mM EDC, leading to the conjugation of the carboxy group of MTX to the amino group derived from biotin- $\text{NH}_2$  or SA. The sensor was dipped into Kinetics buffer containing 2  $\mu\text{M}$  GFP $\Delta\text{C\_OR}$ ,  $\text{SSP}_{\text{GD}}$ ,  $\text{SSP}_{\text{GD}222}$ , or 20 nM  $\text{SSP}_{\text{GD}218}$  for 300 s, then into Kinetics buffer without proteins for 300 s.

### ***HS-AFM observation of SSPs.***

AFM imaging of the SSPs in solution was performed using a laboratory-built HS-AFM setup (47, 48). Immobilization of SSPs for AFM imaging was achieved via a His-tag introduced at the C-terminus of the SSPs and a Ni-coated mica. Two microliters of 100 nM SSP<sub>GD</sub> or the mutants were placed on a Ni-coated mica substrate and incubated for 10 min at room temperature (22–28 °C). Unattached molecules were removed by washing with observation buffer (HKM buffer or 100 mM sodium phosphate (NaPi) pH 6.5, 500 mM NaCl). Imaging was carried out in tapping mode using small cantilevers (BL-AC10DS-A2 or custom-made BL-AC7DS-KU4; Olympus, Tokyo, Japan). The cantilever-free oscillation amplitude was approximately 1.5 nm, and the set-point amplitude was 80–90% of the free oscillation amplitude. The imaging rate, scan size, and feedback parameters were optimized to enable visualization using a minimum tip force.

### ***Buffer exchange experiment.***

SSP<sub>GD</sub> was diluted to 100  $\mu$ M with a GFP-stabilized buffer (25 mM NaPi pH 8.0, 50 mM NaCl) or a DHFR-stabilized buffer (25 mM NaPi pH 6.0, 500 mM NaCl). The buffer containing SSP<sub>GD</sub> was exchanged with a different-state-stabilized buffer on a NAP-5 column (Cytiva). SSP<sub>GD</sub> was then concentrated by ultrafiltration. Buffer exchange and protein concentration were repeated alternately. SSP<sub>GD</sub> concentrations were measured by absorbance at 280 nm. For the limited proteolysis experiment, SSP<sub>GD</sub> was diluted to 5  $\mu$ M with the buffer. Proteinase K was added a final concentration of 10  $\mu$ g/ml and incubated for 3 min at room temperature. The reaction was then quenched with PMSF (5 mM final concentration), separated by SDS-PAGE, and stained with CBB. GFP fluorescence and DHFR enzyme activity of the proteolyzed products were measured as described above, but with a high buffering capacity buffer (400 mM HEPES-KOH pH 7.5, 100 mM KCl, 5 mM MgCl<sub>2</sub>) to maintain constant pH.

### ***Statistical Analysis.***

Student's *t*-test was used to calculate statistical significance, with a two-tailed distribution and unequal variance. All experiments were conducted at least three times independently, and mean values  $\pm$  standard deviation (SD) are represented in the figures.



## Acknowledgments

We thank Keisuke Yoshida and Toru Hisabori for peptide sequencing; Zhu Bo for BLI; Toshio Ando (Kanazawa University) for providing us with HS-AFM instruments; Integrative Biosciences and the Biomaterials Analysis Division, Open Facility Center at the Tokyo Institute of Technology for DNA sequencing. This work was supported by MEXT Grants-in-Aid for Scientific Research (Grant Numbers JP18H03984, JP20H05925, and JP21H04763 to HT), the World Premier International Research Center Initiative (WPI), MEXT, Japan, and the Kanazawa University CHOZEN project to HK.

## Competing interests

The authors declare no competing interest.

## References

1. C. B. Anfinsen, Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
2. A. Miiller, *et al.*, Switch Peptides: pH-Induced  $\alpha$ -Helix to  $\beta$ -Sheet Transitions of Bis-amphiphilic Oligopeptides. *Angewandte Chemie International Edition in English* **30**, 1514–1516 (1991).
3. S. Zhang, A. Rich, Direct conversion of an oligopeptide from a beta-sheet to an alpha-helix: a model for amyloid formation. *Proc Natl Acad Sci U S A* **94**, 23–28 (1997).
4. D. L. Minor, P. S. Kim, Context-dependent secondary structure formation of a designed protein sequence. *Nature* **380**, 730–734 (1996).
5. W. Kabsch, C. Sander, On the use of sequence homologies to predict protein structure: identical pentapeptides can have completely different conformations. *Proc Natl Acad Sci U S A* **81**, 1075–1078 (1984).
6. J. W. Kelly, The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr Opin Struct Biol* **8**, 101–106 (1998).
7. E. K. Jaffe, Morpheesins--a new structural paradigm for allosteric regulation. *Trends Biochem Sci* **30**, 490–497 (2005).
8. Z. A. Bornholdt, *et al.*, Structural Rearrangement of Ebola Virus VP40 Begets Multiple Functions in the Virus Life Cycle. *Cell* **154**, 763–774 (2013).
9. S. H. Lawrence, *et al.*, Shape shifting leads to small-molecule allosteric drug discovery. *Chem Biol* **15**, 586–596 (2008).

10. A. G. Murzin, Biochemistry. Metamorphic proteins. *Science* **320**, 1725–1726 (2008).
11. M. Lella, R. Mahalakshmi, Metamorphic Proteins: Emergence of Dual Protein Folds from One Primary Sequence. *Biochemistry* **56**, 2971–2984 (2017).
12. A. F. Dishman, B. F. Volkman, Unfolding the Mysteries of Protein Metamorphosis. *ACS Chem Biol* **13**, 1438–1446 (2018).
13. P. N. Bryan, J. Orban, Proteins that switch folds. *Curr Opin Struct Biol* **20**, 482–488 (2010).
14. A. LiWang, L. L. Porter, L. P. Wang, Fold-switching proteins. *Biopolymers* **112** (2021).
15. L. L. Porter, L. L. Looger, Extant fold-switching proteins are widespread. *Proc Natl Acad Sci U S A* **115**, 5968–5973 (2018).
16. X. Luo, H. Yu, Protein metamorphosis: the two-state behavior of Mad2. *Structure* **16**, 1616–1625 (2008).
17. A. F. Dishman, *et al.*, Evolution of fold switching in a metamorphic protein. *Science* **371**, 86–90 (2021).
18. B. M. Burmann, *et al.*, An  $\alpha$  helix to  $\beta$  barrel domain switch transforms the transcription factor RfaH into a translation factor. *Cell* **150**, 291–303 (2012).
19. S. E. Boyken, *et al.*, De novo design of tunable, pH-driven conformational changes. *Science* **364**, 658–664 (2019).
20. R. A. Langan, *et al.*, De novo design of bioactive protein switches. *Nature* **572**, 205–210 (2019).
21. F. Praetorius, *et al.*, Design of stimulus-responsive two-state hinge proteins. *Science* **381**, 754–760 (2023).
22. X. I. Ambroggio, B. Kuhlman, Computational design of a single amino acid sequence that can switch between two distinct protein folds. *J Am Chem Soc* **128**, 1154–1161 (2006).
23. X. I. Ambroggio, B. Kuhlman, Design of protein conformational switches. *Curr Opin Struct Biol* **16**, 525–530 (2006).
24. R. G. Alberstein, A. B. Guo, T. Kortemme, Design principles of protein switches. *Curr Opin Struct Biol* **72**, 71–78 (2022).
25. A. F. Dishman, B. F. Volkman, Design and discovery of metamorphic proteins. *Curr Opin Struct Biol* **74** (2022).
26. Y. Hori, Y. Sugiura, Conversion of antennapedia homeodomain to zinc finger-like domain: Zn(II)-induced change in protein conformation and DNA binding. *J Am Chem Soc* **124**, 9362–9363 (2002).
27. J. A. Davey, A. M. Damry, N. K. Goto, R. A. Chica, Rational design of proteins that exchange on functional timescales. *Nat Chem Biol* **13**, 1280–1285 (2017).

28. M. J. Pandya, *et al.*, Sequence and structural duality: designing peptides to adopt two stable conformations. *J Am Chem Soc* **126**, 17016–17024 (2004).
29. E. Cerasoli, B. K. Sharpe, D. N. Woolfson, ZiCo: a peptide designed to switch folded state upon binding zinc. *J Am Chem Soc* **127**, 15008–15009 (2005).
30. T. L. Solomon, *et al.*, Reversible switching between two common protein folds in a designed system using only temperature. *Proc Natl Acad Sci U S A* **120** (2023).
31. H. Sekhon, J. H. Ha, S. N. Loh, Enhancing response of a protein conformational switch by using two disordered ligand binding domains. *Front Mol Biosci* **10** (2023).
32. I. M. Sander, J. L. Chaney, P. L. Clark, Expanding Anfinsen's principle: contributions of synonymous codon selection to rational protein design. *J Am Chem Soc* **136**, 858–861 (2014).
33. J. D. Pédelacq, S. Cabantous, T. Tran, T. C. Terwilliger, G. S. Waldo, Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* **24**, 79–88 (2006).
34. N. A. Touchette, K. M. Perry, C. R. Matthews, Folding of dihydrofolate reductase from *Escherichia coli*. *Biochemistry* **25**, 5445–5452 (1986).
35. C. Frieden, Refolding of *Escherichia coli* dihydrofolate reductase: sequential formation of substrate binding sites. *Proc Natl Acad Sci U S A* **87**, 4413–4416 (1990).
36. J. Schymkowitz, *et al.*, The FoldX web server: an online force field. *Nucleic Acids Res* **33**, W382–W388 (2005).
37. J. van Durme, *et al.*, A graphical interface for the FoldX forcefield. *Bioinformatics* **27**, 1711–1712 (2011).
38. K. Do, S. G. Boxer, GFP variants with alternative  $\beta$ -strands and their application as light-driven protease sensors: a tale of two tails. *J Am Chem Soc* **135**, 10226–10229 (2013).
39. A. M. John, H. Sekhon, J. H. Ha, S. N. Loh, Engineering a Fluorescent Protein Color Switch Using Entropy-Driven  $\beta$ -Strand Exchange. *ACS Sens* **7**, 263–271 (2022).
40. M. Iwakura, T. Nakamura, C. Yamane, K. Maki, Systematic circular permutation of an entire protein reveals essential folding elements. *Nat Struct Biol* **7**, 580–585 (2000).
41. R. Shiba, *et al.*, Systematic alanine insertion reveals the essential regions that encode structure formation and activity of dihydrofolate reductase. *Biophysics (Nagoya-shi)* **7**, 1–10 (2011).
42. H. K. Kim, B. K. Kaang, Truncated green fluorescent protein mutants and their expression in *Aplysia* neurons. *Brain Res Bull* **47**, 35–41 (1998).
43. A. Royant, M. Noirclerc-Savoye, Stabilizing role of glutamic acid 222 in the structure of Enhanced Green Fluorescent Protein. *J Struct Biol* **174**, 385–390 (2011).

44. J. A. Sniegowski, J. W. Lappe, H. N. Patel, H. A. Huffman, R. M. Wachter, Base catalysis of chromophore formation in Arg96 and Glu222 variants of green fluorescent protein. *J Biol Chem* **280**, 26248–26255 (2005).
45. M. Iwakura, B. E. Jones, J. Luo, C. R. Matthews, A strategy for testing the suitability of cysteine replacements in dihydrofolate reductase from *Escherichia coli*. *J Biochem* **117**, 480–488 (1995).
46. C. A. Fierke, K. A. Johnson, S. J. Benkovic, Construction and evaluation of the kinetic scheme associated with dihydrofolate reductase from *Escherichia coli*. *Biochemistry* **26**, 4085–4092 (1987).
47. T. Uchihashi, N. Kodera, T. Ando, Guide to video recording of structure dynamics and dynamic processes of proteins by high-speed atomic force microscopy. *Nat Protoc* **7**, 1193–1206 (2012).
48. T. Ando, T. Uchihashi, S. Scheuring, Filming biomolecular processes by high-speed atomic force microscopy. *Chem Rev* **114**, 3120–3188 (2014).

## Figure legends

### Figure 1. Design principle of a Seesaw Protein (SSP): a protein with interconvertible two alternative functional states.

(A) Fusion of Two Proteins: Two proteins with different functions, A and B, are joined together by overlapping their ends. In this illustration, the C-terminal region of protein A overlaps with the N-terminal region of protein B and is fused together. If the overlapping region (OR) is crucial for the function of each parent protein, then only the moiety to which the OR binds should exhibit a function. Therefore, two functional states, states A and B, are achieved depending on whether the OR binds to either moiety. This fusion protein is termed a seesaw protein (SSP) because it exhibits seesaw-like behavior, switching between the two functional states within a single amino acid sequence. (B) Switchable States of SSP: The SSP can switch between two states. The seesaw tilts toward the more stable state of the two states of the designed protein. The balance can be adjusted by mutations, ligands, buffer conditions, and other factors.

### Figure 2. Creation of an SSP (SSP<sub>GD</sub>) composed of GFP and DHFR, exhibiting two distinct functions in the cell.

(A) Design of the SSP (SSP<sub>GD</sub>) composed of GFP and DHFR. SSP<sub>GD</sub> consists of three components: GFP lacking a C-terminal region (GFP\_ΔC), an overlapping region (OR), and DHFR lacking an N-terminal region (DHFR\_ΔN). (B) Alignment of the OR with the C-terminal region of GFP and the N-terminal region of DHFR. Amino acids derived from GFP or DHFR are colored green and blue, respectively. The numbers indicate the amino acid residue numbers of GFP and DHFR, with the residue 220th (Leu) common to both. The residue numbers of the SSP<sub>GD</sub> correspond to those of GFP. (C) In vivo colony assay to assess SSP<sub>GD</sub> functions. The positions of overexpressed proteins in the *E. coli* cells are shown in the circle (*left*). "Empty" refers to cells not overexpressing SSP<sub>GD</sub> or the parent proteins. After incubation at 37 °C for 6 h, proteins were overexpressed for 72 h at 18 °C in plates with varying TMP concentrations (0, 1, and 4 μg/ml). Plates were photographed (*top*) and imaged for fluorescence (*bottom*). (D) Amino acid sequences of ORs in SSP<sub>GD</sub> and its mutants, SSP<sub>GD222</sub> and SSP<sub>GD218</sub>, affecting the seesaw balance. The color code is consistent with (B). The expected seesaw balances for each SSP protein are illustrated on the right. (E) In vivo colony assay to assess SSP<sub>GD222</sub> and SSP<sub>GD218</sub> functions. The explanation is the same as described in (C).

### Figure 3. SSP<sub>GD</sub> exhibited two different conformations.

(A) Measurement of GFP fluorescence (*left*) and DHFR enzymatic activity (*right*) of SSP<sub>GD</sub> and its parent protein monomers, depicted schematically as in Fig. 2A. Fluorescence excitation and emission wavelength were 480, 513 nm, respectively. Unit is defined as the amount of enzyme degrading 1  $\mu$ mol of substrate per minute. Data represent the means ( $\pm$ S.D.) of three independent experiments. (B) Bio-layer interferometry (BLI) analysis. (*left*) Binding between SSPs and a GFP nanobody (LaG16)-immobilized sensor. The sensor was dipped in a kinetics buffer (PBS supplemented with 0.1% BSA and 0.02% Tween20) containing 50 nM SSPs and then transferred to the kinetics buffer without SSPs. (*right*) Binding between SSPs or GFP $\Delta$ C-OR and an MTX-labeled sensor. MTX, a DHFR inhibitor, was immobilized onto the SA sensor. The MTX-labeled sensor was immersed in the kinetics buffer containing 2  $\mu$ M GFP $\Delta$ C\_OR, SSP<sub>GD</sub>, SSP<sub>GD222</sub>, or 20 nM SSP<sub>GD218</sub>, and then transferred to the kinetics buffer without SSPs or GFP $\Delta$ C\_OR. (C) Limited proteolysis of SSPs. Purified SSP or mutants (5  $\mu$ M) were incubated with 1  $\mu$ g/ml proteinase K at 18 °C for 40 min. After quenching with PMSF, the limited proteolysis products were evaluated by SDS-PAGE. The N-terminal sequences of the limited proteolysis products were determined by Edman degradation (Fig. S2B). Schematics on the right side of the gel show the limited proteolysis products corresponding to the bands. The asterisk indicates a contaminant that could not be removed by purification. (D) The thermodynamic balance of the seesaw. The free energy difference between the GFP and DHFR states in SSP<sub>GD</sub> was analyzed using band intensities of the limited proteolysis products. In addition, changes in thermodynamic stabilities of the two states of SSP mutants relative to SSP<sub>GD</sub> were predicted by FoldX. The predicted change in thermodynamic stability of each state corresponds to the tilt and height of the seesaw.

### Figure 4. Direct visualization of SSP<sub>GD</sub> and its mutants using high-speed atomic force microscopy (HS-AFM).

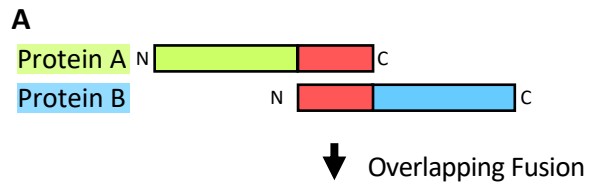
HS-AFM imaging of the structural dynamics of SSPs in solution. The top three images show low magnification (wide range) to depict multiple SSPs, while the bottom three present representative magnified images. In the wide range SSP<sub>GD</sub> image, particles that were considered by visual inspection from the movie to consist of two domains or one domain are indicated by white and green



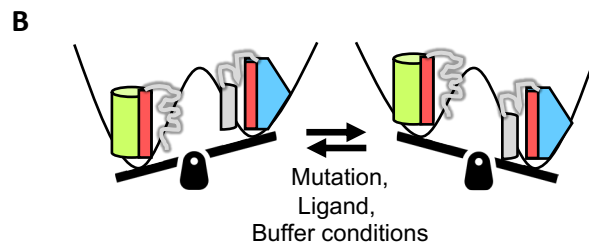
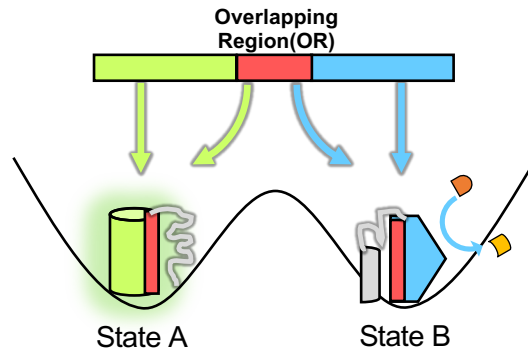
arrowheads, respectively. (*Left*) HS-AFM images of SSP<sub>GD218</sub>. For the top and bottom images, operational parameters were as follows: scanning range, 200 × 200 nm (100 × 100 pixels) and 70 × 700 nm (70 × 70 pixels); scan rate, 0.5 and 0.1 s/frame, respectively. (*Middle*) HS-AFM images of SSP<sub>GD</sub>. For the top and bottom images, operational parameters were: scanning range, 150 × 150 nm (100 × 100 pixels) and 80 × 64 nm (80 × 80 pixels); scan rate, 0.2 and 0.07 s/frame, respectively. (*Right*) HS-AFM images of SSP<sub>GD222</sub>. For the top and bottom images, operational parameters were: scanning range, 200 × 200 nm (100 × 100 pixels) and 80 × 65 nm (80 × 65 pixels); scan rate, 1 and 0.07 s/frame, respectively.

### **Figure 5. Reversible alteration of seesaw balance by buffer conditions.**

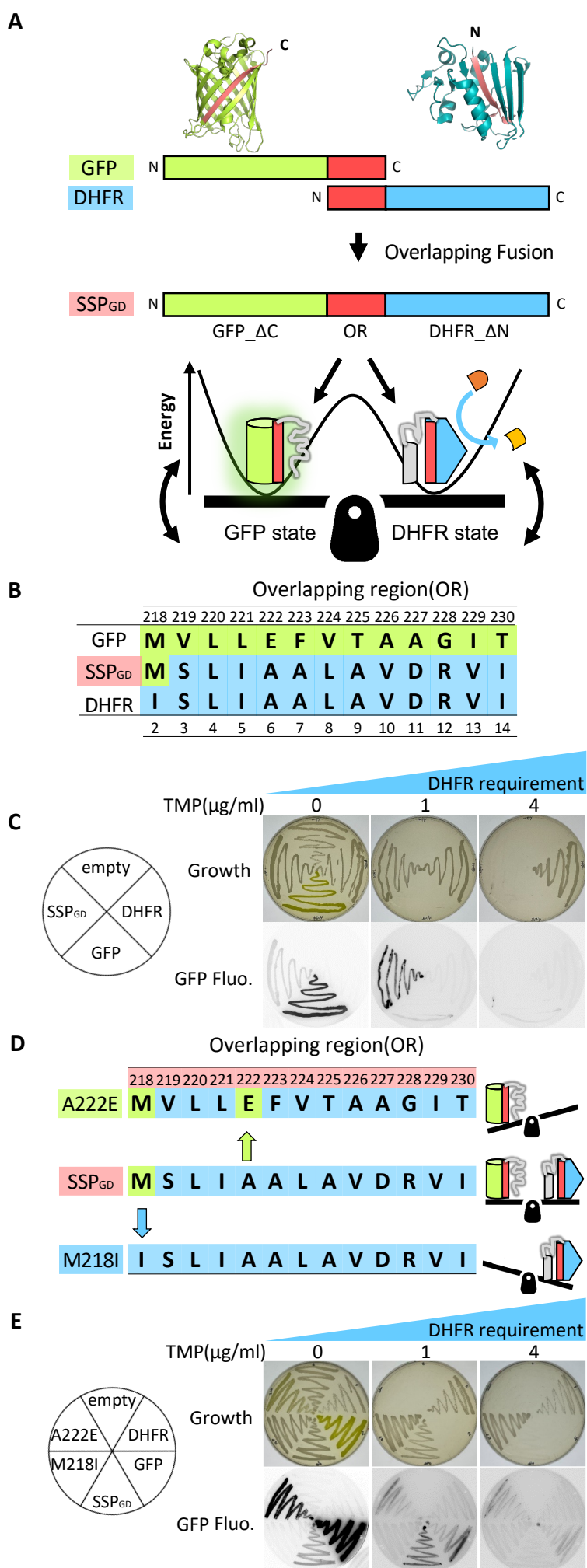
(A) Transition between two SSP<sub>GD</sub> states by adjusting pH and NaCl concentration. High pH and low NaCl concentration stabilize the GFP-biased state, while low pH and high NaCl concentration stabilize the DHFR-biased state. (B) Alteration of SSP<sub>GD</sub> conformation through buffer exchange. SSP<sub>GD</sub> was diluted to 100 μM using either a GFP-stabilized buffer (25 mM NaPi pH 8, 50 mM NaCl) or a DHFR-stabilized buffer (25 mM NaPi pH 6, 500 mM NaCl). The buffer conditions were alternated through the cycle of concentration and dilution. Light-colored backgrounds represent the respective conditions (green for GFP-biased and blue for DHFR-biased). GFP fluorescence (*green*) and the DHFR enzymatic activity (*blue*) were measured. The conformational states of SSP<sub>GD</sub> under each condition were verified by limited proteolysis, as shown in Fig. 3C. The asterisk (\*) indicates the position of proteinase K. (C) HS-AFM imaging of SSP<sub>GD</sub> structural dynamics under the two buffer conditions. *Left*, HS-AFM imaging of the SSP<sub>GD</sub> in 25 mM NaPi (pH8.0), 50 mM NaCl. Operational parameters: scanning range, 100 × 100 nm (100 × 100 pixels); scan rate, 0.34 s/frame. *Right*, HS-AFM imaging of the SSP in 25 mM NaPi (pH 6.0), 500 mM NaCl. Operational parameters: scanning range, 150 × 150 nm (100 × 100 pixels); scan rate, 0.2 s/frame.



**SeeSaw Protein(SSP)**

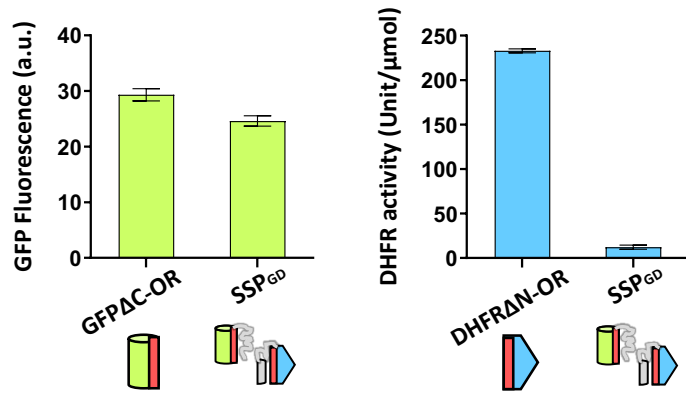


**Figure 1**

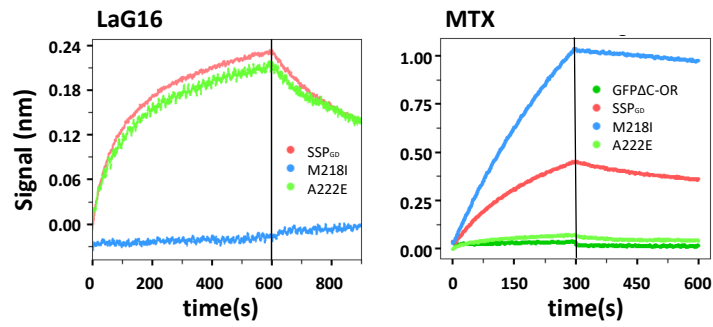


**Figure 2**

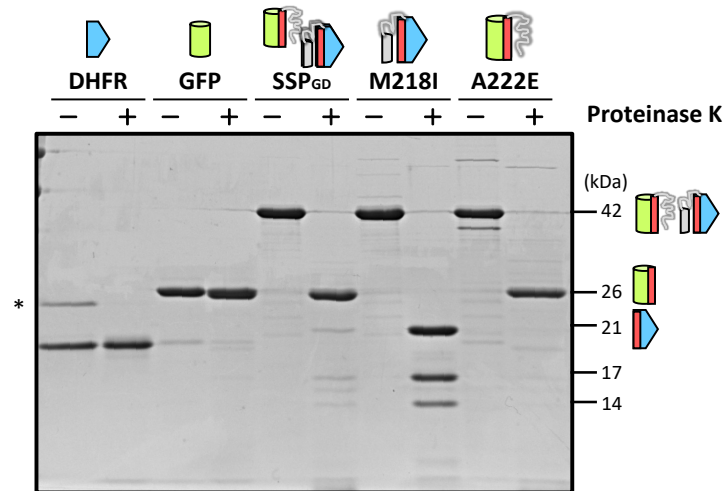
**A**



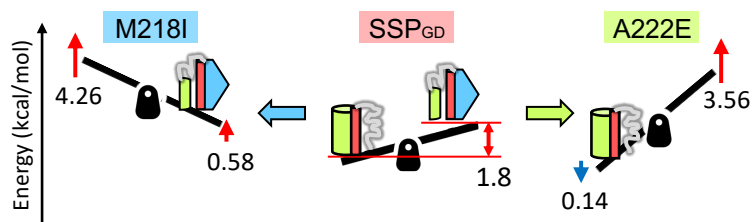
**B**



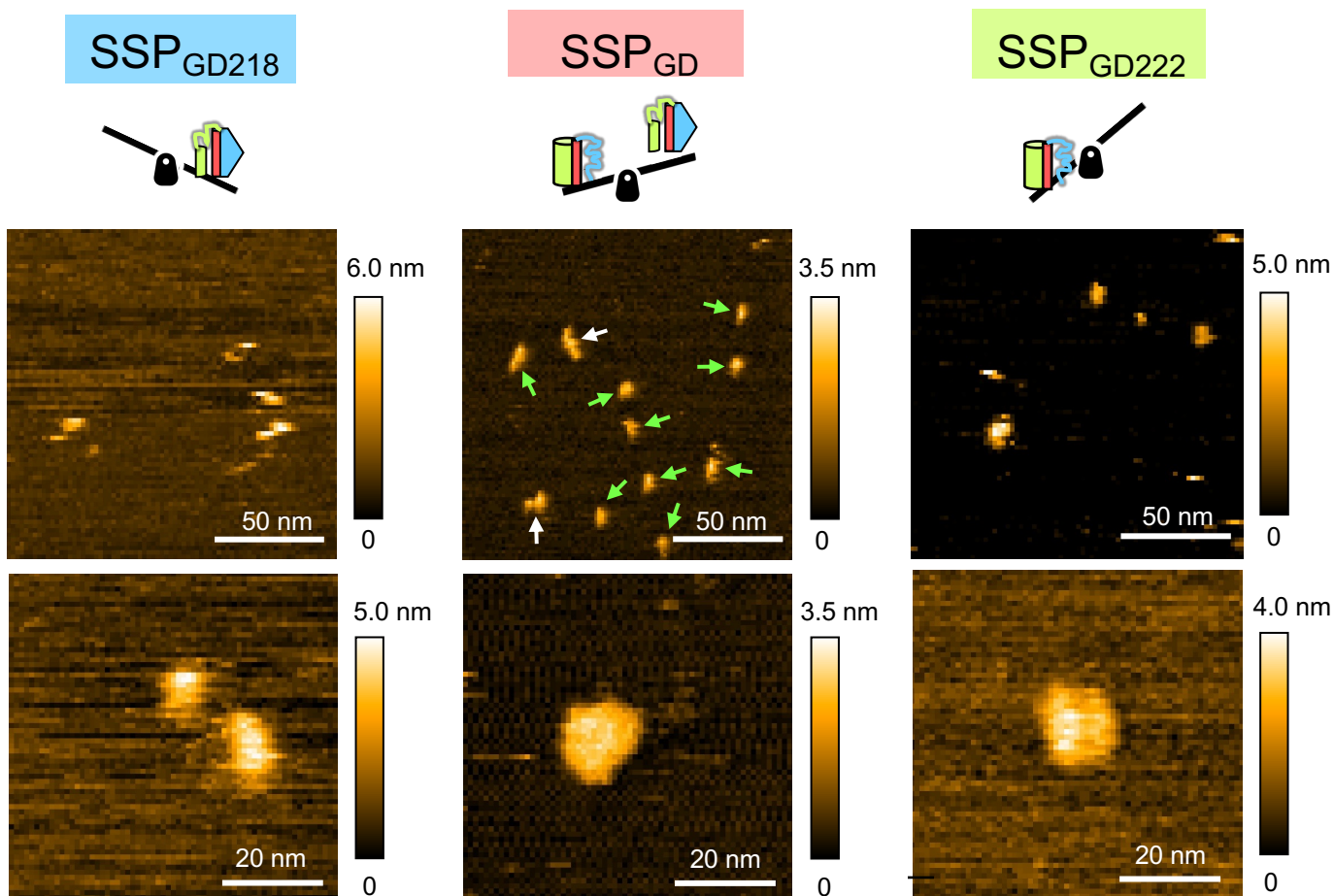
**C**



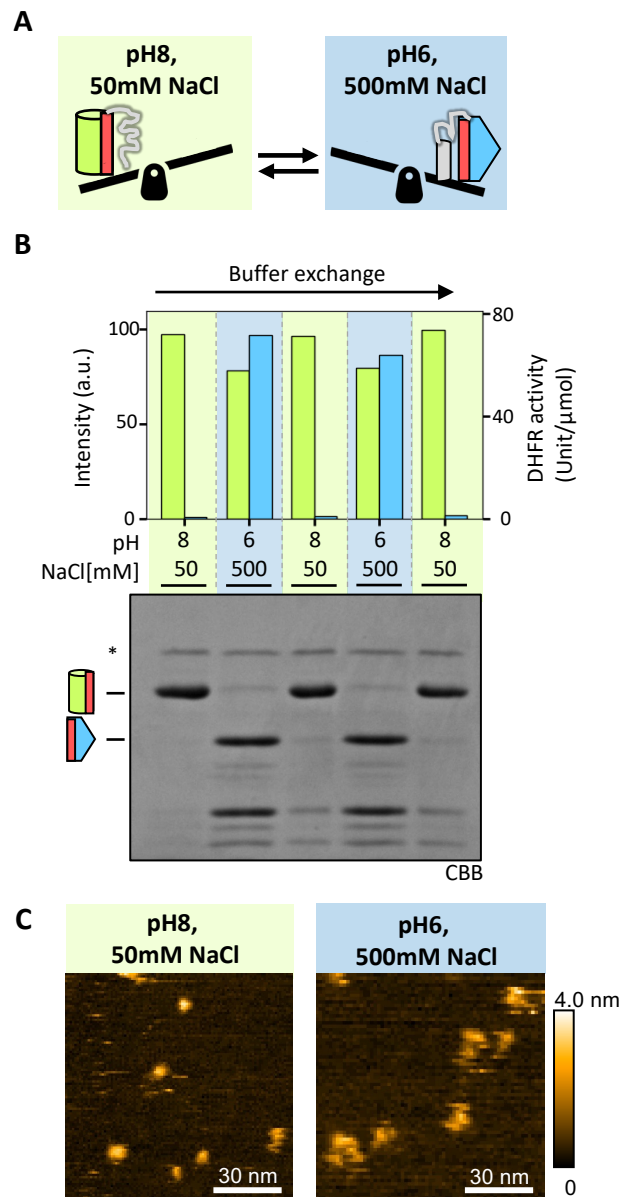
**D**



**Figure 3**



**Figure 4**



**Figure 5**