Systematic control of protein interaction using a modular ER/K α-helix linker

Sivaraj Sivaramakrishnan[1,2] and James A. Spudich[1]

Department of Biochemistry, Beckman Center, B405, Stanford University, Stanford, CA 94305

Contributed by James A. Spudich, September 30, 2011 (sent for review July 8, 2011)

Cellular functions of proteins are strongly influenced by their interactions with other proteins. The frequency of protein interactions is a function of the local concentration of two proteins and their affinity for one another. When two proteins are tethered together, the link between them influences their effective concentrations and therefore the frequency of their interaction. Currently no methods exist to systematically vary the effective concentration within this intramolecular interaction. Here we outline a modular, genetically encoded linker, namely, an ER/K [genetically encoded polypeptide motif based on alternating sequence of approximately four glutamic acid (E) followed by approximately four arginine (R) or lysine (K) residues] single α-helix that can be used to regulate the frequency of interaction between two proteins, or between a protein and a peptide, one at each end. We exploit the wide range of interaction affinities between calmodulin and its binding peptides, combined with FRET to determine the effect of the ER/K α-helix in regulating protein interactions. We find that increasing the length of the ER/K α-helix reduces the on rate of the intramolecular interaction without significantly affecting the off rate, regardless of the affinity of the bimolecular interaction. We outline a genetically encoded approach to determine the dissociation constant for both moderate (micromolar K_d) and strong (nanomolar K_d) protein interactions. Our studies demonstrate the use of the ER/K α-helix to systematically engineer FRET biosensors that detect changes in concentration or affinity of interacting proteins, and modulate enzyme autoinhibition. Our findings are consistent with the ER/K α-helix as a worm-like chain with rare, stochastic breaks in the helix backbone that may account for the behavior of myosin VI stepping along actin.

systematic protein affinity strength | modulation | cell signaling | biochemistry

Interactions between proteins are essential for cellular function. The association of proteins facilitates their localization to, and therefore function in, specific subcellular compartments (1). Proteins often associate to form a macromolecular complex that can perform functions that the individual components cannot (2). Brief bimolecular interactions among a series of proteins are characteristic of signaling cascades that are essential to propagate and amplify an extracellular signal across the cell and elicit an appropriate response (3). Interactions between proteins are dictated by their binding affinity and the local cellular concentrations of the proteins involved. In addition to bimolecular interactions between proteins, many proteins have intramolecular interactions between distinct domains that regulate their function (4). Such intramolecular interactions within a protein in turn are regulated by the free solution binding affinity of the two domains and the effective concentration experienced by the two domains within the protein.

Regulation of intermolecular and intramolecular protein interactions is important to gain mechanistic insight into cellular processes and is an essential component of pharmacological intervention in disease states. However, proteins have complex structures with predictions of about 1,000 structural families (5) with 10,000 distinct types of interactions between proteins (6). Detailed insight into the intermolecular or intramolecular interaction interface gained from techniques such as X-ray crystallography, NMR, and computational modeling can be used to perturb the interaction by altering the sequence of one or both proteins/domains using site-directed mutagenesis (7, 8). This approach, however, is limited by the effects of the mutation on the structure and folding of the protein domains, which is complicated by the allosteric nature of protein structure (9). Further, site-directed mutagenesis often results in a loss of protein–protein interaction and is therefore usually an all-or-nothing approach to regulation. Given the importance of protein–protein interactions, it is necessary to have an alternate approach that can systematically modulate protein–protein interactions without perturbing the structure of the proteins or protein domains involved.

In this study, we outline the use of a modular protein linker, namely an unusual single ER/K [genetically encoded polypeptide motif based on alternating sequence of approximately four glutamic acid (E) followed by approximately four arginine (R) or lysine (K) residues] α-helix that by virtue of its structural properties can be used to dial-in the frequency of interaction of two proteins or protein domains attached to its ends. We find that varying the ER/K α-helix length systematically alters the frequency of interaction regardless of the free solution binding affinity of the interacting pair attached to the ends of the ER/K α-helix.

Results

Theoretical Estimate of the Effect of the ER/K α-Helix Linker on k_{close} and k_{open}. This study examines the use of the ER/K α-helix as a modular linker that regulates the frequency of interaction of two proteins or protein domains at its ends. In order for the ER/K α-helix to regulate the frequency of interaction, it must limit either the on rate (k_{close}) or off rate (k_{open}) of the interaction. Further, this limit should vary systematically with the length and/or structural properties of the ER/K α-helix.

Previous studies are consistent with the ER/K α-helix behaving as a worm-like chain (WLC) with a persistence length (L_p) of approximately 15 nm (10). A Monte Carlo simulation of the ER/K α-helix, modeled as an ideal WLC with $L_p = 15$ nm, provided a distribution of conformations adopted by different length ER/K α-helices (Fig. 1 *A* and *B*, *Left*). Extensive conformational sampling (>20 million) was used to estimate the probability of the two ends of the ER/K α-helix being close enough to facilitate interaction between two proteins. This probability is defined as the fraction of total number of conformations sampled with end-to-end distance less than 2 nm. This length was used as the

Author contributions: S.S. and J.A.S. designed research; S.S. performed research; S.S. analyzed data; and S.S. and J.A.S. wrote the paper.

The authors declare no conflict of interest.

See Commentary on page 20279.

[1]To whom correspondence may be addressed. Email: sivaraj@umich.edu or jspudich@stanford.edu.

[2]Present address: Department of Cell and Developmental Biology, University of Michigan, Ann Arbor, MI 48109-2200.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1116066108/-/DCSupplemental.

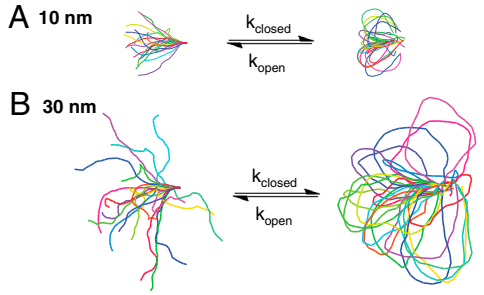


Fig. 1. ER/K α-helix as an ideal WLC. (*A*) Ten nanometer helix, (*B*) 30 nm helix. Sample conformations derived from a metropolis Monte Carlo simulation of the backbone of the ER/K α-helix. (*Left*) Helix free in solution. (*Right*) Ends of the helix are connected by a stiff spring (100 pN/nm) with resting length 2 nm. Monte Carlo simulations are used to estimate the probability of interaction between the ends, which in turn influences the on rate of the intramolecular interaction (k_{close}). Mean bending energy stored in open and closed conformations is used to predict the effect of the helix on the off rate (k_{open}).

distance between the ends of the ER/K α-helix when two small proteins [*ca.* 20 kDa with hydrodynamic radius (R_h) *ca.* 2 nm] bind to each other. However, the conclusions of different models do not vary for lengths ranging from 1 to 3 nm. This probability was computed over a range of ER/K α-helix lengths (L) and was found to be vanishingly small ($<10^{-5}$) for L comparable to the L_p. Mehraeen et al. (12) have previously reported theoretical distributions for end-to-end distance distributions for WLCs. Their theoretical distributions show that the cumulative distribution of end-to-end distances less than 2 nm are also on the order of 10^{-5} for $L \sim L_p$. These results suggest that, for a previously estimated L_p of 15 nm (11) and L ranging from 10–30 nm, the ER/K α-helix could limit the k_{close} of interaction of proteins at its ends. Given the rarity of Monte Carlo derived conformations with end-to-end distances less than 2 nm, this approach cannot accurately estimate the effect of ER/K α-helix length on the frequency of interaction. The theoretical distributions of Mehraeen et al. (12) suggest that, for an ideal WLC with $L_p = 15$ nm, the number of conformations with end-to-end distance less than 2 nm increases with ER/K α-helix length ranging from 10 to 30 nm. However, given the low frequency of occurrence of these conformations, it is possible that any deviations of the ER/K α-helix from an ideal WLC will likely influence this trend. Nonetheless, the ER/K α-helix is a good candidate linker to limit k_{close}.

The ER/K α-helix could also alter k_{open}. For the ER/K α-helix as an ideal WLC, the additional energy stored in the closed state relative to the open state decreased from 9 kcal/mol (1 kcal = 4.18 kJ) for $L = 10$ nm to 5 kcal/mol for $L = 20$ nm and 3 kcal/mol for $L = 30$ nm (see *SI Data*). For comparison, the strength of a single protein salt bridge between proteins has been estimated to range from 0.5 to 4 kcal/mol (13). The probability of dissociation of the protein–protein interaction will therefore increase as the energy needed to hold the ER/K α-helix in the closed state increases. Therefore, the ER/K α-helix modeled as an ideal WLC is a good candidate to increase k_{open} relative to that of the free bimolecular interaction.

Design of a Systematic Protein Affinity Strength Modulation (SPASM) FRET sensor. We have termed our approach SPASM. Fig. 2*A* shows a schematic of a single polypeptide chain used to detect ER/K α-helix modulation of the interaction between calmodulin (CAM) and its binding peptide (peptide). This single polypeptide chain is termed the SPASM FRET sensor to emphasize the presence of the ER/K α-helix and the FRET pair for sensing the frequency of interaction of the protein and protein/peptide at the ends of the ER/K α-helix. From the N to the C terminus, the sensor contained enhanced CFP (eCFP, the FRET donor), CAM, ER/K α-helix, mCitrine (mCit, the FRET acceptor), and a peptide with known affinity for binding Ca^{2+}-activated CAM. All components were separated by a (Gly-Ser-Gly) × 4 [(GSG) × 4] linker to ensure rotational freedom.

CAM is a signaling molecule that binds to Ca^{2+} and switches from an inactive state [apocalmodulin (Apo CAM)] to an active state (CAM) (14). Ca^{2+} binding causes a dramatic increase in its binding affinity for a variety of effectors including protein kinases, phosphatases (calcineurin), ion channels, and active transporters (Ca^{2+} pump). CAM binding to these proteins triggers their activation in response to Ca^{2+} influx. A wide variety of small peptides that bind specifically to Ca^{2+} bound CAM and Apo CAM have been identified, and their affinities for binding CAM have been systematically quantified (15). CAM-peptide interactions are a model system to study ER/K α-helix modulation of protein–protein interactions. CAM-peptide binding affinity can be easily controlled in vitro by the presence or absence (chelation with EGTA) of Ca^{2+}. The wide range of interaction affinities derived

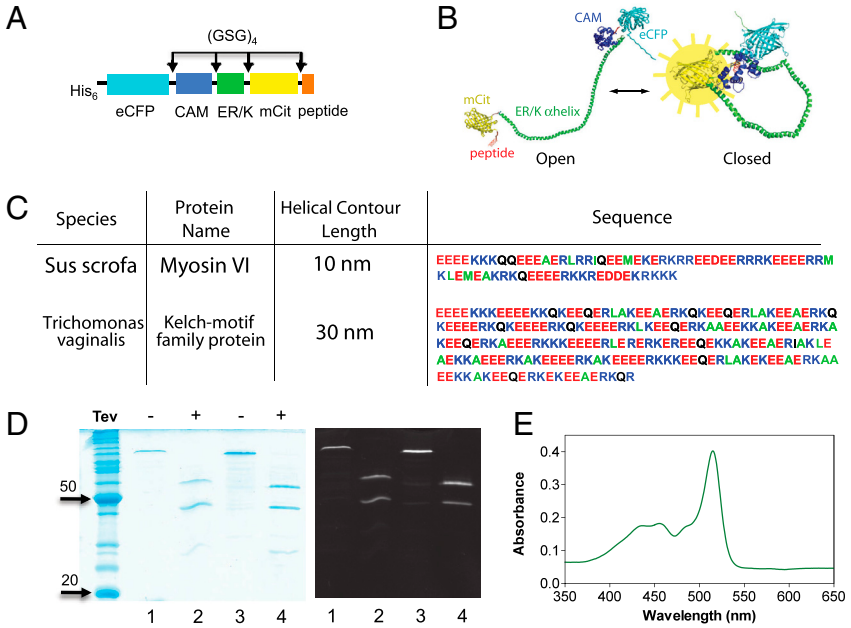


Fig. 2. SPASM sensor design. (*A*) Schematic of protein domains (His_6 is at the N terminus) in the SPASM sensor. A 12 amino acid Gly-Ser-Gly $(GSG)_4$ linker is placed between different domains to ensure rotational freedom. (*B*) Structural model of the SPASM sensor in the open (*Left*) and closed (*Right*) conformation. (*C*) Sequences of two ER/K α-helices. Note that the 20-nm helix used in this study is the first 130 amino acids of the 30-nm helix sequence. (*D*, *Left*) Coomassie gel staining of purified SPASM sensors with 30 nm (lanes 1 and 2) and 20 nm (lanes 3 and 4) ER/K α-helices. Lanes 2 and 4 were treated with TEV-protease, which cleaves at its recognition site engineered between the CAM and ER/K α-helix. (*Right*) Laser fluorescence gel scan (457-nm eCFP excitation with 520BP40 emission). This condition captures both eCFP and mCitrine fluorescent bands. (*E*) Absorption spectrum of purified SPASM sensor showing peaks at 433 nm (eCFP) and 514 nm (mCit). Note that mCit has approximately two-fold higher extinction coefficient compared to eCFP.

from CAM effector binding sites, both in the Ca^{2+} bound and Apo states, can be used to quantify the contribution of the ER/K α-helix to the on and off rates of protein–protein interactions.

FRET was used as a measure of the interaction between CAM and Ca^{2+}. FRET occurs between eCFP and mCitrine, variants of CFP and YFP, which are used because of their increased brightness and fluorophore stability (16). When CFP and YFP are in close proximity (≲8 nm) excitation of CFP results in significant FRET from CFP to YFP. FRET between CFP and YFP has been used extensively to assess interaction between two proteins that are either free in solution or tethered together with a linker (17, 18). In the absence of Ca^{2+}, the ER/K α-helix was designed to separate eCFP and mCit with the sensor adopting an "open" state (Fig. 2*B*, *Left*). In the presence of Ca^{2+}, the high affinity of the peptide for CAM was designed to hold the ER/K α-helix in a "closed" state (Fig. 2*B*, *Right*) and bring eCFP and mCit within FRET range. The emission spectrum of the sensor sample following excitation of the donor, referred to as the FRET spectrum, is a function of the relative number of sensor molecules in the open and closed states. There are several ER/K α-helical protein motifs with $L = 10$, 20, and 30 nm in naturally occurring proteins (10). In this study, a 10-nm helix from myosin VI, and a 20- and 30-nm helix derived from the Kelch-motif family protein were used (Fig. 2*C*) for our SPASM sensor constructs. The sensor was purified using a combination of affinity (Ni-nitrilotriacetate) and gel filtration, and the purity was assessed using laser fluorescence gel scanning and Coomassie staining of the PAGE run samples (Fig. 2*D*), in addition to the sample absorption spectrum showing peaks at 433 nm (eCFP) and 514 nm (mCit) (Fig. 2*E*) (see *SI Materials and Methods* for details).

FRET Ratio (R) Varies Linearly with Fraction of Sensor in the Closed State. For an ideal WLC with $L_p = 15$ nm, the mean end-to-end distances for ER/K α-helices with $L = 10$, 20, and 30 nm are 9, 16, and 23 nm, respectively (19). Therefore, in the absence of CAM-peptide interactions, no significant FRET between eCFP and mCit is expected ($R_0 \sim 5$ nm; ref. 20). Accordingly, sensor constructs with 10-, 20-, and 30-nm helices that have the CAM-binding peptide substituted by an extended linker $(GSG)\times 4$ with no known affinity for CAM showed no measurable FRET (Fig. 3*A*). For these constructs, selective excitation of the donor (eCFP at 433 nm) resulted in donor emission (eCFP emission maximum at 475 nm) without significant acceptor emission (mCit emission maximum at 525 nm) (Fig. 3*A*). A CAM-binding peptide that selectively binds to the activated Ca^{2+} bound CAM underwent a dramatic increase in FRET when CAM was transitioned from the Apo (Fig. 3*B*, EGTA) to Ca^{2+} bound state (Fig. 3*B*, calcium). The change in signal was found to be similar for 10-, 20-, and 30-nm ER/K α-helices. It must be noted that the concentrations of sensors used in these experiments ranged from 5 to 50 nM, well below the previously reported free solution K_d (equilibrium dissociation constant) of this CAM-peptide interaction (*ca.* 1 μM) (21). Therefore, the change in FRET is derived primarily from intramolecular interactions with minimal contribution from intermolecular interactions. This observation was also confirmed by showing that serial dilution of the sample does not significantly alter the ratio of mCit (acceptor) to eCFP (donor) emission (see below).

The relative number of sensors in the closed state, also termed f_c for fraction closed [fraction open (f_o) is 1-f_c], is related to the ratio of the intramolecular on and off rates by the following equation

$$\frac{k_{\text{open}}}{k_{\text{close}}} = \frac{1}{f_c} - 1. \qquad [1]$$

Here k_{open} and k_{close} are the intramolecular off and on rates, respectively. In order to derive f_c from the FRET spectrum, the equivalent bimolecular system was used where f_c can be varied systematically by changing the sensor concentration (see *SI Discussion*). A tobacco etch virus (TEV) protease site was built into the sensor between the CAM and the ER/K α-helix. Following purification of the sensor as a single polypeptide chain (see *Materials and Methods*), TEV protease was used to obtain stoichiometrically equivalent amounts of eCFP-tagged CAM and mCit-tagged binding peptide that participate in a bimolecular interaction. Complete digestion with TEV was verified by laser fluorescence gel scanning (Fig. 2*D*). TEV-digested sample was concentrated to over 10-fold (16 μM) above the K_d (*ca.* 1 μM) to facilitate the bimolecular interaction. Bimolecular interactions between Ca^{2+}-activated CAM and peptide resulted in high FRET (Fig. 4*A*). Serial dilutions of the sample were used to reduce the equilibrium fraction bound in accordance with the K_d, and therefore the fraction of molecules in the high FRET state, resulting in a reduction in mCit emission with decreasing concentration of the sample (Fig. 4*B*). Given the K_d of the interaction, the f_c for a given concentration of the sample is directly calculated from Eq. **2** (note, Fig. 4*B* top vs. bottom abscissa).

$$f_c = 1 + \frac{K_d}{2C_t} - \frac{\sqrt{K_d^2 + 4K_dC_t}}{2C_t}. \qquad [2]$$

The FRET ratio ($mCit_{max}/eCFP_{max}$) varied approximately linearly with f_c (Fig. 4*B*, $R^2 = 0.993$ for linear regression fit). The approximate linear relationship between FRET ratio (R) and f_c is consistent with a two-state model of the FRET sensor (see *Materials and Methods*). R for a given f_c was therefore expressed as a linear combination of the R_{obs} of the unbound (open) state (R_{open}) and the bound (closed) state (R_{closed}).

$$R_{obs} = R_{closed} \times f_c + R_{open} \times (1 - f_c). \qquad [3]$$

R_{obs}, unlike the FRET spectrum, is normalized for the number

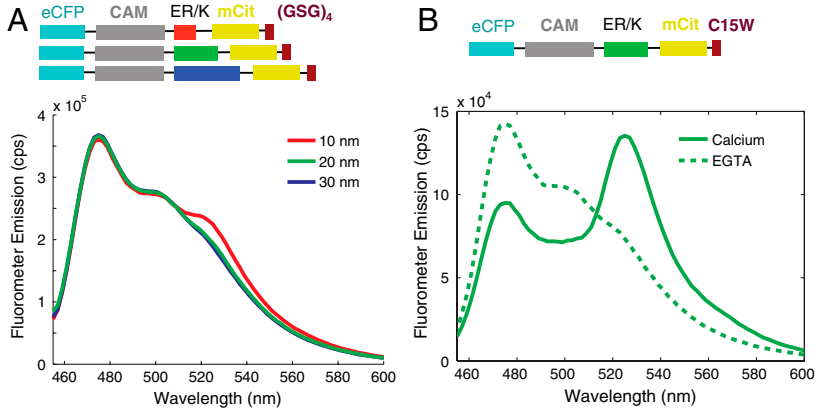


Fig. 3. FRET as a readout of protein interaction. (*A*) Emission spectra of control SPASM sensors with CAM linked to a 12 amino acid Gly-Ser-Gly peptide $(GSG)_4$ by 10-, 20-, and 30-nm ER/K α-helices. Spectra are taken at 4 mM Ca^{2+}. Excitation of eCFP (430 nm) results in eCFP emission (475 nm) without significant mCit emission (525 nm). (*B*) Sample emission spectrum of SPASM sensor with CAM linked by a 20-nm ER/K α-helix to a C15W peptide (21) that interacts with 1 μM affinity to Ca^{2+}-activated CAM. Excitation of eCFP (430 nm) results in eCFP emission (475 nm) with no detectable FRET for mCit in the absence of Ca^{2+} (chelation by 1 mM EGTA) or strong FRET (large mCit peak at 525 nm) in the presence of 4 mM Ca^{2+}.

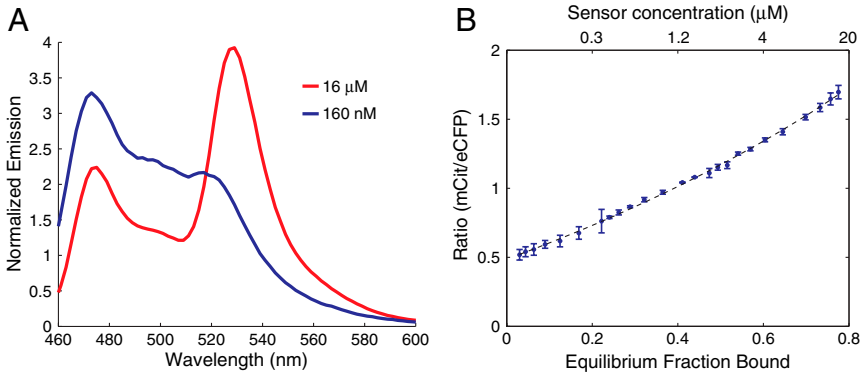


Fig. 4. FRET ratio varies linearly with fraction of interacting protein in the bound state. (*A*) Emission spectra of SPASM sensors digested with TEV-protease to cleave a site between CAM and the ER/K α-helix, resulting in a bimolecular interaction between eCFP-tagged CAM and mCit-tagged C15W peptide (*ca.* 1-μm affinity). Sample spectra corresponding to 160- and 16-μM interaction affinities are shown. (*B*) Measured FRET ratio (mCit/eCFP) as a function of sensor concentration (top abscissa) and equilibrium fraction bound (bottom abscissa). Dotted line represents fit to Eq. S2. Linear regression fit of FRET ratio to equilibrium fraction bound has R^2 of 0.996.

of sensor molecules present in the sample. Although R_{closed} and R_{open} cannot be directly measured, they can be estimated from the equivalent bimolecular interaction as discussed in the next section.

FRET Ratio (*R*) Can Be Used to Measure K_d of Medium Affinity Interactions (*ca.* Micromolar). The use of R to measure f_c was independently validated using another approach. Given the linear relationship between f_c and FRET ratio (R_{obs}), the data for the bimolecular interaction (Fig. 5*A*) can be fit to Eqs. **2** and **3** with K_d, R_{open}, and R_{closed} as fit parameters. The fit yielded a K_d of 1.3 μM, which agrees well with a previous measurement of 1.7 μM for this CAM-peptide interaction pair (21). Although R_{open} and R_{closed} were allowed to float for the fit, they are constrained by R_{obs} at high and low C_t. However, the K_d measurement and the quality of the fit (R square) were not affected significantly (<10% difference) when R_{open} and R_{closed} were constrained to R_{obs} at low and high concentrations, respectively. Therefore, a single parameter fit of the experimental data to Eqs. **2** and **3** was sufficient to measure K_d. This approach to measure K_d using R vs. C can be universally applied to any protein–peptide or protein–protein interaction. A single concentrated protein sample (*ca.* 20-fold higher than the expected K_d with 50 μL volume) is sufficient to obtain the entire curve using serial dilution (Fig. 5 represents the mean ± SD for four samples; see *Materials and Methods*). Coexpression of the protein and protein/peptide with both a TEV-protease site and an ER/K α-helix between them provides an easy method to obtain stoichiometrically equivalent amounts of the two binding partners from a single purified sample. The ER/K α-helix acts as a spacer that allows proper folding of the two proteins. In order to accurately estimate the K_d of two interacting proteins using this approach, it is essential to measure R_{obs} over a range of bimolecular concentrations that bracket the K_d. For very high-affinity interactions (<10 nM) the FRET spectrum from the fluorometer has lower signal to noise and the accuracy of R_{obs} is reduced. For such high-affinity interactions, competitive binding of either of the fluorescently tagged proteins to a nonfluorescent partner with known affinity can be used, as described in the next section.

FRET Ratio (*R*) Can Be Used to Measure the K_d of High-Affinity interactions (*ca.* Nanomolar). Competitive binding assays have been used to determine the K_d of high-affinity interactions. However, the techniques for detecting competitive binding require multiple steps, for instance, centrifugation followed by either quantitative analysis of PAGE gels by Coomassie stain, immunoblotting, or a radioimmunoassay (22). In this study, FRET was used as a quantitative readout of competitive binding. Inverse FRET ratio (*IR*), defined as the ratio of eCFP (475 nm) to mCit (525 nm) emission maxima, decreases as the vasoactive intestinal peptide (VIP)

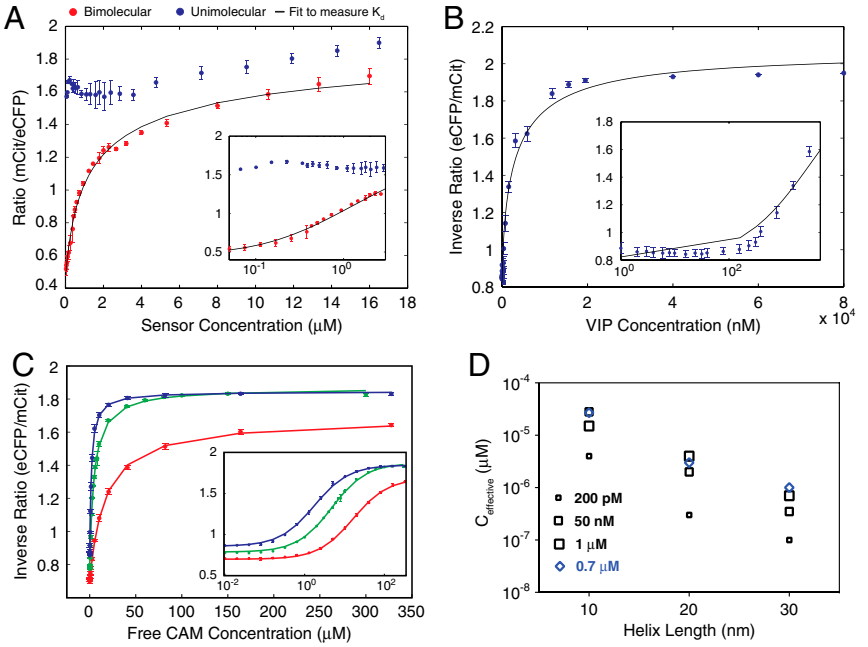


Fig. 5. Quantifying the equilibrium dissociation constant using FRET; Determination of the effective concentration of interacting partners in a SPASM sensor. (*A*) FRET ratio (mCit/eCFP) as a function of protein concentration. Blue dots, intact SPASM sensor with C15W peptide in the presence of 4 mM Ca^{2+}. Red dots, SPASM sensor with C15W peptide digested with TEV-protease to cleave SPASM sensor into eCFP-tagged CAM and mCit-tagged C15W peptide. Black line, bimolecular interaction fit to Eqs. **2** and **3** to calculate K_d of bimolecular interaction. (*B*) Competitive binding of VIP (50 nM affinity) to TEV-cleaved SPASM sensor with CAM and Trp3 peptide (0.3 nM K_d). Inverse FRET ratio (eCFP/mCit) increases as eCFP-CAM/mCit-peptide interaction is disrupted by VIP. (*C*) Competitive binding of free unlabeled CAM to intact SPASM sensor linked to C15W peptide by 10 nm (red), 20 nm (green), and 30 nm (blue) ER/K α-helices. Inverse FRET ratio (eCFP/mCit) increases as the open conformation of the sensor is favored by increased unlabeled CAM concentration. (*A–C, Insets*) Low concentration data shown on log axis for clarity. (*D*) Effective concentration of CAM-peptide interaction within the SPASM sensor measured by competitive binding to free unlabeled CAM. Black squares, peptides that bind to CAM with indicated affinities in the presence of Ca^{2+}. Blue diamonds, Apo CAM binding to peptide with 0.7-μM affinity. All data shown are mean ± SD of at least four different experiments.

competes for binding in the bimolecular FRET interaction between eCFP-tagged CAM and mCit-tagged Trp3 (Fig. 5*B*). We measured a K_d of 53 nM for the VIP-based on a 0.3-nM affinity of the Trp3 peptide (see *Materials and Methods*), which is nearly identical to the previously reported 54-nM affinity for this interaction (23).

Effective Concentration of the Intramolecular Interaction Decreases with Increasing ER/K α-Helix Length. Studies on the equivalent bimolecular interaction can be used to measure the FRET ratio corresponding to the open (R_{open}) and closed (R_{closed}) states, which in turn can be used to calculate f_c (Eq. **3**) and hence k_{open}/k_{close} for different SPASM sensors. One drawback of this approach is that, as f_c approaches one, the k_{open}/k_{close} is very sensitive to the accuracy of f_c (Eq. **1**); f_c approaches one for protein/peptide pairs that interact with high affinity (*ca.* nanomolar). In addition, for such high-affinity interactions we cannot use the serial dilution approach, outlined earlier, to measure R_{open} and R_{closed}. Hence, an alternative approach to measure f_c-based competitive binding of the sensor to free unlabeled CAM was used (Fig. 5*C*, *Materials and Methods*). This technique of measuring f_c can be applied to both weak- and strong-affinity interactions between the protein and protein/peptide within the sensor.

An alternate measure of the strength of interaction between the two proteins is the effective concentration (C_{eff}). Effective concentration is the concentration of the protein/peptide pair in the equivalent bimolecular interaction that has an f_c equal to that of the intramolecular interaction within the SPASM sensor. The effective concentration can be estimated from the competitive binding assay (see *Materials and Methods*). The higher the C_{eff} or f_c, the stronger the interaction between the protein–peptide pair within the sensor and vice versa. The advantage of using C_{eff} over f_c is that, as f_c approaches one, small changes in f_c can cause large changes in C_{eff}. Therefore, comparing interaction strengths based on C_{eff} is easier than with f_c. Fig. 5*D* is a summary of C_{eff} for the intramolecular interaction between CAM and corresponding peptide, measured using five different interaction affinities (four different peptides with one in the Apo CAM state), for three ER/K α-helices with length 10, 20, and 30 nm. Regardless of the affinity of interaction between CAM and peptide, or the state of activation of CAM, the C_{eff} for a given interaction was found to decrease with increasing ER/K α-helix length (Fig. 5*D*). Specifically, C_{eff} decreased about 10-fold when helix length increased from 10 to 20 nm and about threefold when the helix length increased from 20 to 30 nm. Hence the ER/K α-helix can be used to systematically effect large changes in C_{eff} between interacting proteins. Further, we estimated the off rate (k_{open}) of the CAM-Trp3 peptide sensor by monitoring the quench of FRET signal following incubation of vast excess of unlabeled CAM (Fig. S1). These measurements (see *SI Data*) show the ER/K α-helix does not affect the off rate (k_{open}). Therefore the effect of the ER/K α-helix on the effective concentration is mediated through the on rate (k_{close}).

Discussion

Structural Role of the ER/K Motif in Naturally Occurring Proteins. $Glu_4(Arg/Lys)_4$ (ER/K) is a naturally occurring motif that we had identified to be present in numerous proteins from divergent species with diverse functions (10). We have previously used a variety of biophysical approaches to summarily characterize its behavior as a semirigid spacer between protein domains (11). In this study, we find that, although the ER/K α-helix predominantly adopts an extended conformation (Fig. 3*A*), rare stochastic breaks in it enable two protein domains placed one at each end to interact (see *SI Data*). Although the frequency of breaks increases linearly with increasing ER/K α-helix length, the probability of end-to-end interaction decreases due to the lack of spatial coordination between the breaks (see Fig. S2 for simplified model). Taken together, the rare stochastic breaks result in an approximately 30-fold decrease in effective concentration of the ends when the ER/K α-helix length is increased threefold (Fig. 5*D*). In the context of the ER/K motif in naturally occurring proteins, the findings of this study have profound implications for the contribution of this domain to protein function. Specifically, if the protein domains separated by the ER/K α-helix have moderate (*ca.* micromolar) to high (*ca.* nanomolar) affinity for each other, the length of the ER/K motif will dictate the frequency of the interaction. For proteins such as the molecular motor myosin, kinases, or translation initiation factors, the frequency of interaction is likely to influence enzymatic activity (10). In this study, we quantify the effective concentrations for a range of bimolecular interaction affinities (0.2 nM–1 μM) that will serve as a measure of effective concentration of different interactions separated by an ER/K α-helix of defined length (Fig. 5*D*). Further, we isolate ER/K α-helix dependent on-rates that when combined with the kinetics of individual interactions will enable predictions of the preferred conformation (open vs. closed) of the protein.

We have previously reported that the ER/K α-helix within myosin VI forms an extended structure that extends the lever arm stroke of this molecular motor (10). The lever arm stroke of several myosins, including myosin VI has been recorded using optical tweezers (24). The upper bound on the stroke time of the lever arm is 100 μs, which is determined by the time resolution of the optical trapping measurements (10 kHz). The fact that myosin VI has a large stroke is therefore consistent with the time between helix breaks being larger than 100 μs, such that helix breaks do not influence the stroke size. In a processive myosin VI dimer, the lever arm stroke of the leading head is followed by a dwell, when the trailing head searches for the next binding site. Stochastic breaks in the ER/K α-helix during this dwell would allow the trailing head to bind actin at different sites, corresponding to the location of the break in the helix. Interestingly, the model with stochastic breaks is consistent with the wide distribution of step sizes in myosin VI, relative to myosin V, which does not have an ER/K α-helix (25). Therefore, the influence of the ER/K α-helix linker on protein function appears to be different at different timescales.

Engineering FRET Biosensors. Since the first genetically encoded protein FRET biosensor to detect calcium signaling over a decade ago (26), a large number of similar designs have been developed to directly detect the activation of proteins involved in cell signaling pathways (18, 27, 28). These FRET biosensor designs are based on a single polypeptide chain with CFP and YFP variants separated by two protein domains that interact in response to cell signaling events such as calcium influx or phosphorylation of one of the protein domains. The two protein domains in these biosensors are separated by a short (approximately three) amino acid linker composed of Gly and Ser. As a consequence, the CFP/YFP pair is in close proximity with substantial FRET under basal conditions. Activation of the sensor causes the protein domains to interact, which results in a decrease in distance and rotational freedom between the FRET pair and therefore increased FRET. The change in FRET signal, assessed from the emission spectrum following donor excitation, is limited by the close proximity of the CFP/YFP pair enforced by the short linker between the protein domains. Enhancing the change in FRET signal involves either changing the specific linker sequence by trial and error (29) or altering the arrangement of protein elements in the biosensor using circular permutation (30). In this study, we find that the ER/K α-helix minimizes baseline FRET by virtue of its extended structure (Fig. 3*A*), resulting in a substantial increase in the dynamic range of the biosensor following activation (Fig. 3*B*). The availability of different length ER/K α-helices allows systematic alteration of the effective concentration of the interacting proteins (Fig. 5*B*), such that the ER/K α-helix provides an affi-

nity barrier that separates the protein domains in the inactive conformation. The biosensor constructs with the ER/K α-helix are designed with $(Gly\text{-}Ser\text{-}Gly)_4$ linkers between the proteins to ensure rotational freedom between the protein domains. This feature, combined with the wide range of conformations explored by the ER/K α-helix (Fig. 1), should allow protein domains with diverse binding interfaces to interact without steric hindrance. Further, the plasmid constructs for the biosensors containing the ER/K α-helix have been designed with unique restriction enzyme sites flanking different protein domains to allow easy cut-and-paste cloning using PCR.

Determining the Dissociation Constant of Protein–Protein Interactions. Several approaches exist to determine the dissociation constant (K_d) of a pair of interacting proteins. In this study, we use the ER/K α-helix to make stoichiometric amounts of the interacting partners. TEV-protease is used to separate the interacting pair. We show that the FRET ratio correlates linearly with equilibrium fraction bound (Fig. 4*B*). A straightforward dilution series with FRET ratio measured on a fluorometer is therefore sufficient to measure the K_d of moderate affinity (*ca.* micromolar) interactions (Fig. 5*A*). This technique can also be extended to measure the K_d of high-affinity interactions (*ca.* nanomolar) using a competitive binding assay (Fig. 5*B*). Key advantages of this approach over other spectroscopic techniques (measurement of tryptophan fluorescence or dansyl labeling) is that it is both genetically encoded and generally applicable to any protein interacting pair. The use of the ER/K α-helix involves expressing a single polypeptide chain with both proteins in stoichiometric amounts.

Protein Engineering with the ER/K α-Helical Linker. Three key features of the ER/K α-helix for protein engineering are (*i*) it is genetically encoded, (*ii*) varying length ER/K α-helices are readily available from naturally occurring ER/K motifs of different lengths, and (*iii*) the ER/K α-helix by virtue of its alternating charge is unlikely to interact with the protein domains at its ends and is stable under extremes of pH, salt concentration, and temperature. Regardless of the type of linker used to regulate interactions between proteins, detailed characterization of the effect of the linker on the on and off rates of the interaction is necessary before it can be generally applied. In this study, we use the wide range of interaction affinities between calmodulin and its binding peptides to characterize the effect of different length ER/K α-helices. Specifically, we find that the ER/K α-helix influences the on rate without significantly impacting the off rate relative to the bimolecular interaction. Our detailed characterization, combined with the unique properties of the ER/K α-helix, should enable its use as a modular linker in regulating protein–protein interactions.

Materials and Methods

Reagents. Vasoactive intestinal peptide (synthetic, human sequence) was purchased from Sigma. All other reagents, unless otherwise stated, were also purchased from Sigma.

Molecular Biology. All elements were cloned into the multiple cloning sites of the pET47b vector (Novagen). From N to C terminus, the constructs consist of eCFP, CAM, ER/K α-helix, mCitrine, and CAM-binding peptides (see *SI Materials and Methods* for details). All elements were separated by a $(Gly\text{-}Ser\text{-}Gly)\times 4$ linker.

Protein Expression and Purification. Protein expression and purification are based on previously published protocols (31). Details can be found in *SI Materials and Methods*.

Data Acquisition and Analysis. FRET spectra of samples diluted in buffer A with indicated Ca^{2+} ($CaCl_2$) or EGTA (stock pH 8) concentration were acquired on a Fluorolog fluorometer (Horiba Scientific). Quartz microcuvettes (Starna Cell, 16.40F-Q-10/Z8.5) were used with sample volume of 50 μL. For FRET spectra, samples were excited at 430 nm with 2- to 8-nm bandpass on excitation and emission (4- to 8-nm bandpass) was acquired from 450–650 nm. Further details can be found in *SI Materials and Methods*. Data from fluorometer was analyzed using custom Matlab software (Mathworks).

ACKNOWLEDGMENTS. The authors thank Mark Tsuchida and Pehr Harbury for useful discussions. S.S. is supported by an American Cancer Society Postdoctoral Fellowship and J.A.S. by Grant GM33289 from the National Institutes of Health.

1. Geiger B, Spatz JP, Bershadsky AD (2009) Environmental sensing through focal adhesions. *Nat Rev Mol Cell Biol* 10:21–33.
2. Steitz TA (2008) A structural understanding of the dynamic ribosome machine. *Nat Rev Mol Cell Biol* 9:242–253.
3. Brown MD, Sacks DB (2009) Protein scaffolds in MAP kinase signalling. *Cell Signal* 21:462–469.
4. Pufall MA, Graves BJ (2002) Autoinhibitory domains: Modular effectors of cellular regulation. *Annu Rev Cell Dev Biol* 18:421–462.
5. Chothia C (1992) Proteins. One thousand families for the molecular biologist. *Nature* 357:543–544.
6. Aloy P, Russell RB (2004) Ten thousand interactions for the molecular biologist. *Nat Biotechnol* 22:1317–1321.
7. Shoemaker BA, Panchenko AR (2007) Deciphering protein–protein interactions. Part II. Computational methods to predict protein and domain interaction partners. *PLoS Comput Biol* 3:e43.
8. Shoemaker BA, Panchenko AR (2007) Deciphering protein–protein interactions. Part I. Experimental techniques and databases. *PLoS Comput Biol* 3:e42.
9. Gunasekaran K, Ma B, Nussinov R (2004) Is allostery an intrinsic property of all dynamic proteins? *Proteins* 57:433–443.
10. Sivaramakrishnan S, Spink BJ, Sim AY, Doniach S, Spudich JA (2008) Dynamic charge interactions create surprising rigidity in the ER/K alpha-helical protein motif. *Proc Natl Acad Sci USA* 105:13356–13361.
11. Sivaramakrishnan S, et al. (2009) Combining single-molecule optical trapping and small-angle X-ray scattering measurements to compute the persistence length of a protein ER/K alpha-helix. *Biophys J* 97:2993–2999.
12. Mehraeen S, Sudhanshu B, Koslover E, Spakowitz A (2008) End-to-end distribution for a wormlike chain in arbitrary dimensions. *Phys Rev E Nonlin Soft Matter Phys* 77:061803.
13. Baldwin RL (2007) Energetics of protein folding. *J Mol Biol* 371:283–301.
14. Chin D, Means AR (2000) Calmodulin: A prototypical calcium sensor. *Trends Cell Biol* 10:322–328.
15. O'Neil KT, DeGrado WF (1990) How calmodulin binds its targets: Sequence independent recognition of amphiphilic alpha-helices. *Trends Biochem Sci* 15:59–64.
16. Sample V, Newman RH, Zhang J (2009) The structure and function of fluorescent proteins. *Chem Soc Rev* 38:2852–2864.
17. Kenworthy AK (2001) Imaging protein–protein interactions using fluorescence resonance energy transfer microscopy. *Methods* 24:289–296.
18. Zhang J, Allen MD (2007) FRET-based biosensors for protein kinases: Illuminating the kinome. *Mol Biosyst* 3:759–765.
19. Howard J (2001) *Mechanics of Motor Proteins and the Cytoskeleton* (Sinauer, Sunderland, MA), Chap 6.
20. Siegel RM, et al. (2000) Measurement of molecular interactions in living cells by fluorescence resonance energy transfer between variants of the green fluorescent protein. *Sci STKE* 2000:lp1.
21. Vorherr T, et al. (1990) Interaction of calmodulin with the calmodulin binding domain of the plasma membrane Ca^{2+} pump. *Biochemistry* 29:355–365.
22. Zettner A (1973) Principles of competitive binding assays (saturation analysis). 1. Equilibrium techniques. *Clin Chem* 19:699–705.
23. Malencik DA, Anderson SR (1983) Binding of hormones and neuropeptides by calmodulin. *Biochemistry* 22:1995–2001.
24. Sung J, Sivaramakrishnan S, Dunn AR, Spudich JA (2010) Single-molecule dual-beam optical trap analysis of protein structure and function. *Methods Enzymol* 475:321–375.
25. Rock RS, et al. (2001) Myosin VI is a processive motor with a large step size. *Proc Natl Acad Sci USA* 98:13655–13659.
26. Miyawaki A, et al. (1997) Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. *Nature* 388:882–887.
27. Kiyokawa E, Aoki K, Nakamura T, Matsuda M (2011) Spatiotemporal regulation of small GTPases as revealed by probes based on the principle of Forster Resonance Energy Transfer (FRET): Implications for signaling and pharmacology. *Annu Rev Pharmacol Toxicol* 51:337–358.
28. Violin JD, Zhang J, Tsien RY, Newton AC (2003) A genetically encoded fluorescent reporter reveals oscillatory phosphorylation by protein kinase C. *J Cell Biol* 161:899–909.
29. Miyawaki A, Tsien RY (2000) Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein. *Methods Enzymol* 327:472–500.
30. Baird GS, Zacharias DA, Tsien RY (1999) Circular permutation and receptor insertion within green fluorescent proteins. *Proc Natl Acad Sci USA* 96:11241–11246.
31. Spink BJ, Sivaramakrishnan S, Lipfert J, Doniach S, Spudich JA (2008) Long single alpha-helical tail domains bridge the gap between structure and function of myosin VI. *Nat Struct Mol Biol* 15:591–597.

Supporting Information

Sivaramakrishnan and Spudich 10.1073/pnas.1116066108

SI Text

SI Data. ***The ER/K α-helix limits the on rate (k_{close}) of the intramolecular interaction without significantly affecting the off rate (k_{open}).*** The large changes in C_{eff} (effective concentration) with helix length, for a given protein–peptide interaction, strongly suggest that the ER/K [genetically encoded polypeptide motif based on alternating sequence of approximately four glutamic acid (E) followed by approximately four arginine (R) or lysine (K) residues] α-helix successfully limits the on and/or off rate (k_{close} and k_{open}) of the intramolecular interaction. C_{eff} and f_c (fraction closed) can be converted to a ratio of intramolecular on and off rates (k_{close}/k_{open}, Eq. **S1**). To individually determine the effect of the ER/K α-helix on k_{close} and k_{open}, one of the rates needs to be independently measured. We used a FRET-based approach to determine k_{open} for the high-affinity interaction (0.3 nM) between calmodulin (CAM) and the Trp3 peptide. CAM-Trp3 FRET, in the presence of intact or tobacco etch virus (TEV) cleaved sensor, was quenched by a vast excess of unlabeled CAM and the change in R_{obs} over time was monitored (Fig. S1, *SI Materials and Methods*). The change in R_{obs} with time fit well to a single exponential decay, characteristic of a first-order, rate-limiting step corresponding to the opening rate (k_{open}) of the intact sensor (Fig. S1). The sensor opening rate k_{open} was not significantly affected by the presence of the ER/K α-helix (<20% change relative to the bimolecular interaction). Knowledge of k_{open} of the sensor was used to compute the k_{close} from Eq. **S1**. For the 10-, 20-, and 30-nm helices the k_{close} was calculated to be 27, 1.4, and 0.5 s^{-1}, respectively. Therefore, increasing ER/K α-helix length from 10 to 20 and 20 to 30 nm for the CAM-Trp3 peptide sensor decreased k_{close} ~ 20- and threefold, respectively. The decrease in k_{close} correlates well with the ~10- and threefold decrease in C_{eff} for all of the CAM-peptide interactions when the ER/K α-helix length is increased from 10 to 20 and 20 to 30 nm. Taken together, the data strongly suggest that the ER/K α-helix regulates protein–protein/peptide interactions by systematically reducing the intramolecular on rate, without affecting the off rate.

Alternate physical model of the ER/K α-helix. The experimental data show that the strength of protein–protein interactions (measured by C_{eff} or f_c) decreased with increasing length of the ER/K α-helix (Fig. 5*D*). These experimental data are not consistent with the model of the ER/K α-helix as an ideal worm-like chain (WLC) with $L_p \sim 15$ nm (1). Specifically, for $L_p \sim 15$ nm an ideal WLC with $L = 30$ nm would have a higher probability than $L = 10$ nm for the ends to be close enough to facilitate interaction (*ca.* 2 nm). An alternate model that is consistent with previous measurements and the data reported in this study is that of the ER/K α-helix as a WLC with rare, stochastic breaks in the helix backbone that enable interaction between the two ends. This alternate model explains three inconsistencies between experimental data presented in this study with an ideal WLC. First, for all three helices (10, 20, and 30 nm) there is a very low probability ($<2 \times 10^{-5}$) of the two ends of the helix being close enough (*ca.* 2 nm) to facilitate interactions. With stochastic breaks, the interactions between the ends are more likely. Second, the frequency of the breaks should be a linear function of the helix length. Therefore, the 30-nm helix should have three times as many stochastic breaks as the 10-nm helix. However, these helix breaks are not coordinated in space. Hence, multiple breaks in a 30-nm helix do not imply an increase in likelihood of the two ends coming closer together for a longer helix. A simplified model (see Fig. S2) predicts that the ends of a helix with "n" stochastic breaks, which are uncorrelated in space and time, have a probability of interaction of $(n+1)/2^{n+1}$ for an odd number of breaks and $n/2^{n+1}$ for an even number of breaks. For $n \geq 1$ over a 10-nm length of ER/K α-helix, the probability of interaction decreases rapidly with increasing helix length. Specifically, for two stochastic breaks along a 10-nm helix, this simplified model predicts that the probability of the two ends interacting for the 20- and 30-nm helices are fourfold and 21-fold lower than the 10-nm helix, respectively. Third, for an ideal WLC with the two proteins at the ends interacting with each other, the helix occupies a higher energy state resulting from the bending energy stored in the closed conformation (Fig. 1*B*). The energy stored decreases with helix length (9 kcal/mol for a 10-nm helix vs. 3 kcal/mol for a 30-nm helix). Transition from the closed to open conformation reduces the free energy of the sensor molecule. Therefore, an ideal WLC is likely to increase the off rate of the protein–protein interaction relative to the off rate of the free bimolecular interaction. Stochastic breaks, however, dissipate the bending energy stored in the WLC. The observation that the ER/K α-helix limits the on rate and does not affect the off rate (Fig. S1) is therefore consistent with the alternate model of the ER/K α-helix as a WLC with stochastic breaks along its length. Previously reported small angle X-ray scattering and optical trapping experiments were consistent with the ER/K α-helix as an extended structure (1, 2). The FRET spectra measured with control peptide [(Gly-Ser-Gly) × 4] (Fig. 3*A*) were also consistent with an extended structure. The interaction between the protein–peptide pair at the ends of the ER/K α-helix, therefore, captures rare transitions of the ER/K α-helix from the predominant extended conformation. Thus, the stochastic breaks along the length of the helix must be relatively infrequent and the helix reanneals into an extended structure following these breaks, such that the most frequently observed conformations of the helix are consistent with an extended structure.

SI Discussion. ***ER/K α-helix as a modular linker between two protein domains.*** The ER/K motif is based on an alternating sequence of four negatively charged residues (glutamic acid, *E*) and four positively charged residues (arginine, *R*; or lysine, *K*). We have previously identified the ER/K α-helical protein motif to be recurrent in naturally occurring proteins (2). Molecular dynamics simulations reveal an oscillating pattern of interactions between the *E* and *R/K* side chains that stabilize the α-helical conformation (2). In addition, the long *E* and *R/K* side chains exclude water molecules from the α-carbon backbone and stabilize its structure. Using a combination of single molecule optical trapping, small angle X-ray scattering and Monte Carlo simulations, we have estimated the persistence length (length scale over which a structure appears rigid) of the ER/K α-helix as a worm-like chain to be approximately 15 nm (1). Knowledge of the biophysical properties of this structure and their variation with the length of the ER/K α-helix forms the basis of the technique outlined in this study. Varying length ER/K α-helices are easily obtained from the large number of naturally occurring ER/K motifs identified in a previous study (2).

Beyond the well-characterized structural properties, there are three features of the ER/K α-helix that make it an excellent candidate for a structural linker between two proteins or protein domains. The alternating charge sequence makes the ER/K α-helix highly soluble, with a high entropic penalty to be buried in the hydrophobic core of large proteins. Therefore, the ER/K

α-helix is unlikely to cause a protein attached to its ends to engulf it and adopt a different folded state. The oscillating pattern of side-chain interactions in the ER/K α-helix does not lend itself to a stable binding interface based on electrostatic or hydrophobic interactions. As a consequence, the ER/K α-helix is unlikely to interact with proteins attached to its ends. Finally, the ER/K α-helix is a very stable structure and its α-helical conformation is not destabilized over a range of pH (4–9) and high concentrations (*ca.* 1 M) of salt or denaturants such as urea (3). Thus, the ER/K α-helix will retain its structural properties in diverse subcellular environments where the attached proteins localize and function.

Short polypeptides (*ca.* 10 amino acids) that link domains within proteins have been studied extensively and classified based on their propensity to form secondary structure (4). Although these short linkers significantly influence protein function, they often interact with the protein domains involved, by virtue of their proximity to them, and therefore are highly specific to the particular domains involved. Increasing the amino acid length of unstructured linkers (usually a combination of glycine, serine, and proline residues) does not significantly alter the spacing between protein domains. The lack of additional spacing between the protein domains results from the fact that the end-to-end distance of unstructured peptides does not scale linearly with the length of their amino acid sequence (5). With the exception of the ER/K motif, long polypeptides (>12 residues) do not form stable α-helices due to the effect of polar water molecules on the hydrogen bonds in the α-helical backbone. Thus, unless protein linkers with tertiary structure (multiple α-helices and/or β-sheets) are used, effective spacing of protein domains cannot be achieved. However, the use of linker sequences with tertiary structure does not lend itself to a modular linker that can be used to systematically vary the spacing and therefore regulate the interaction between protein domains involved.

FRET ratio as a measure of the fraction of sensors in the bound state (f_c). The FRET spectrum is an ensemble of the FRET signals of individual sensor molecules that have different levels of FRET between enhanced CFP (eCFP) and mCitrine (mCit). Assuming that the individual sensor molecules occupy one of two states, namely the open and closed state (Fig. 2*B*), the FRET spectrum should be a linear combination of the FRET spectra of these two states. In the two-state system, knowledge of the FRET spectrum corresponding to the open and closed states is sufficient to derive the relative number of sensors in the open and closed states. The relative number of sensors in the closed state, also termed f_c for fraction closed (fraction open, f_o, is $1-f_c$), is related to the ratio of the intramolecular on and off rates by the following equation:

$$\frac{k_{\text{open}}}{k_{\text{close}}} = \frac{1}{f_c} - 1. \qquad \textbf{[S1]}$$

Here k_{open} and k_{close} are the intramolecular off and on rates, respectively. Because the goal of this study was to understand the effect of the ER/K α-helix on k_{open} and k_{close} of the intramolecular interaction, it is necessary to convert the measured FRET spectrum to f_c. In practice, the FRET spectrum corresponding to the state when all sensor molecules are in the open or closed conformation cannot be directly measured. An alternative is to convert the FRET spectrum to FRET efficiency, which in turn is a measure of the average distance between the eCFP/mCit fluorophores (6). In this study, the sensors are distributed between the open and closed states that have two different FRET pair distances. Hence, FRET efficiency cannot be easily correlated to f_c. Furthermore, calculating FRET efficiency requires comparison of sensor signal in each sample to a no-FRET state that has identical donor concentration with the donor fluorophores (eCFP) tethered to the same protein domains. Although several techniques exist to measure the FRET spectrum in the no-FRET state, all of them pose significant challenges (6).

SI Materials and Methods. ***Molecular biology.*** A modular cloning approach was used to integrate elements of the Systematic Protein Affinity Strength Modulation (SPASM) sensor. This approach allows easy restriction enzyme cloning of any two interacting proteins or protein domains to form a SPASM sensor. All elements were cloned into the multiple cloning sites of the pET47b vector (Novagen). A $(\text{Gly-Ser-Gly})_4$ linker is inserted between all protein domains as part of the primer sequence. The vector has an N-terminal His_6 tag followed by a PreScission protease site. A fusion between monomeric eCFP (A206K mutation) and CAM (*Pan troglodytes*), containing a $(\text{GSG})_4$ linker in between, was amplified by crossover PCR and cloned between BamHI and EcoRI sites. The placement of protein domains on either side of the ER/K α-helix was based on the modeled structure (Fig. 2*B*). Alternate configurations were not tested. Based on the modeled structures, these alternate configurations are equally likely to work. ER/K α-helices (see Fig. 2*C* for details) with N-terminal ENFLYQG amino acid sequence as a TEV-protease recognition site were cloned between EcoRI and AscI sites; mCitrine was cloned between AscI and SalI sites. Oligos that translate into the different calmodulin binding peptide sequences were purchased from Integrated DNA Technologies, corresponding to the forward and reverse strands and containing 5′ phosphate groups. CAM-peptide sequences were cloned between SalI and PacI sites. Sequences of peptides used are (*i*) C15W, LRRGQILWFRGLNRI, (*ii*) Trp3, LKWKKLLKLLKKLLKLG, (*iii*) Melittin, GIGAVLKVLTTGLPALISWIKRKRQQ, (*iv*) control, $(\text{GSG})_4$ extension.

Protein expression and purification. pET47b vectors with different SPASM sensors were transformed into JM109(DE3). Single colonies were inoculated in Terrific Broth (TB) media (Invitrogen) and grown at 30 °C to an OD (600 nm) of 1.0 and transferred to room temperature (22 °C) where protein expression was induced with 500 μM IPTG. Cells were incubated at 22 °C for 4 h and harvested by spinning at 4,000 × *g* for 20 min. Cells were resuspended in lysis buffer (20 mM phosphate at pH 7.4, 0.5 M NaCl, 10 mM imidazole at pH 7.5, 10% glycerol) plus Complete Protease Inhibitor Tabs (Roche) and lysed by adding lysozyme (Sigma) to 1 mg/mL, incubating for 15 min, and sonicating. Lysates were clarified by centrifugation at 100,000 × *g* for 30 min and supernatants were bound to Ni-nitrilotriacetate (Ni-NTA) resin (Qiagen) for 30 min. The resin was washed with 8 column volumes of 20 mM phosphate (pH 7.4), 0.5 M NaCl, 20 mM imidazole (pH 7.5), and eluted with 20 mM phosphate (pH 7.4), 0.5 M NaCl, 0.5 M imidazole (pH 7.5). Ni-NTA eluate was desalted on PD10 column (GE Healthcare) into buffer A (20 mM Tris at pH 8, 50 mM NaCl, 0.5 mM EDTA). Sample was concentrated four- to eightfold using 30 kDa Vivaspin concentrators (GE Healthcare). For control and C15W peptide sensors, sample was purified by size-exclusion on superdex 200 column in buffer A. For high affinity constructs (Trp3 and Melittin) the sample was mixed with equal volume of 20 mM Tris (pH 8), 50 mM NaCl, 0.5 mM EDTA, 8 M Urea. Urea is used to reversibly denature the proteins and ensure separation between CAM and peptide that interact outside of the intramolecular interaction. The samples were then purified by size-exclusion on superdex 200 column in buffer B (20 mM Tris at pH 8, 50 mM NaCl, 0.5 mM EDTA, 4 M Urea). Gel filtration fractions were run on SDS-PAGE gels and the fractions with a single band on Coomassie gel, which contain both eCFP and mCit (assessed by gel scan on Typhoon gel image; GE Healthcare) were pooled. Samples in buffer B were dialyzed into buffer A (three sequential steps with 500-fold dilution each step). Samples were spin concentrated to obtain the final fraction that was used for fluorom-

eter measurements. The FRET spectrum is sensitive to the stoichiometry of eCFP and mCit (7). Degradation of the sensor during protein purification or an excess of eCFP resulting from partial translation of the sensor will affect the relative signal from eCFP and mCit. Furthermore, proteolysis of the sensor during protein purification such that the eCFP and mCit are no longer tethered can lead to misinterpretation of the contribution of the ER/K α-helix to the FRET spectrum. Hence, the sensor purity was assessed using laser fluorescence gel scanning and Coomassie staining of the PAGE run samples (Fig. 2*D*), in addition to the sample absorption spectrum showing peaks at 433 nm (eCFP) and 514 nm (mCit) (Fig. 2*E*). Sample concentrations were obtained from absorption spectra (Fig. 2*E*) based on extinction coefficients of eCFP (32,500 $M^{-1}\,cm^{-1}$ at 475 nm) and mCitrine (77,000 $M^{-1}\,cm^{-1}$ at 526 nm). Concentrations obtained using eCFP and mCitrine were within 10% of each other for all samples. In fact, the ratio of eCFP to mCitrine absorption correlated well to the purity of the samples observed on SDS-PAGE. CAM was cloned into pET28a vector (Novagen). CAM construct was transformed into Rosetta(DE3) (EMDChemicals) cells. Single colonies were inoculated in TB media (Invitrogen) and grown at 37 °C to an OD (600 nm) of 1.0 and transferred to room temperature (22 °C) where protein expression was induced with 500 μM IPTG. Cells were incubated at 22 °C for 4 h and harvested by spinning at 4,000 × *g* for 20 min. Cells were resuspended in 50 mM Tris (pH 7.5), 2 mM EDTA, 1 mM DTT, and lysed by high-pressure homogenization (EmulsiFlex, Avestin) and centrifuged 35,000 × *g* for 20 min. $CaCl_2$ was added to the supernatant to a final concentration of 5 mM and bound to a phenylsepharose column equilibrated in C buffer (50 mM Tris at pH 7.5, 0.5 mM DTT, 0.1 mM $CaCl_2$). Column was washed sequentially with C buffer followed by C buffer with 0.5 M NaCl. CAM was eluted with 50 mM Tris (pH 7.5), 0.5 mM DTT, 1 mM EDTA, 150 mM NaCl. Fractions containing purified CAM were pooled and spin concentrated (Vivaspin, 10-kDa cutoff). CAM concentration was measured by OD at 280 nm. His_6-tagged TEV-protease was expressed and purified as described elsewhere (8).

Fluorometer data acquisition. The bandpass width was selected to maximize signal without saturation of the detector. Photomultiplier voltage was varied from 600–950 V depending on signal strength. For measurement of K_d of bimolecular interactions, serial dilutions were used. Serial dilution minimizes pipetting error from samples prepared at different concentrations. Quartz submicrocuvettes allow measurements with small volumes (*ca.* 50 μL) on a standard fluorometer were used. These cuvettes limit the volume of sample excited during each scan. Serial dilution was performed by addition of sample buffer directly to the cuvette. Therefore, the fraction of sample that is subject to repeated scans decreases with each subsequent scan and minimizes photobleaching to the sample. Further, an equal number of repetitive scans on a 50-μL sample without serial dilution showed that, although there is photobleaching of the sample over the coarse of the experiment, the *R* values do not change unless the concentration is altered. For measurement of K_d of high-affinity interactions, the purified CAM-Trp3 FRET protein is digested with TEV-protease to unlink the eCFP-tagged CAM and mCit-tagged Trp3, and then diluted to 5-nM concentration. The sample is mixed with increasing concentration (1 nM to 20 μM) of unlabeled vasoactive intestinal peptide (VIP), which binds Ca^{2+} activated CAM with 50-nM affinity. Samples with increasing concentration of VIP are scanned on the fluorometer. The inverse FRET ratio (*IR*), defined as the ratio of eCFP (475 nm) to mCit (525 nm) emission maxima. Similar to the FRET ratio (R_{obs}), IR_{obs} can be expressed as a linear combination of *IR* for the closed and open states. For known concentrations of CAM, Trp3, VIP, and previously measured K_d of one of the interactions, the K_d of the other interaction can be computed by an iterative fit of equilibrium competitive binding equations to the experimental data.

SI Data Analysis. ***Estimation the effect of the ER/K α-helix on the off rate (k_{open}).*** The mean bending energy stored in the ER/K α-helix can be computed from our Monte Carlo derived conformations of an ideal WLC. When two proteins at the ends of the ER/K α-helix interact, the helix transitions from an extended "open" conformation to a "closed" conformation. The effect of protein interaction was modeled in our Monte Carlo simulations by an arbitrarily very high stiffness spring (100 pN/nm) with resting length 2 nm connecting the two ends of the ER/K α-helix (Fig. 1, *A* and *B*, *Right*). These simulations were used to estimate mean bending energy stored in the ER/K α-helix in the closed conformation. The difference in mean bending energy between the closed and open conformation is a measure of the free energy change associated with helix transition from the open to the closed state. For the ER/K α-helix as an ideal WLC, the additional energy stored in the closed state relative to the open state decreased from 9 kcal/mol for $L = 10$ nm to 5 kcal/mol for $L = 20$ nm and 3 kcal/mol for $L = 30$ nm.

Two-state model of the FRET sensor. Assuming that the sensor exists in only two states, open and closed, with a defined FRET spectrum associated with each state, the FRET ratio (*R*) can be expressed as

$$R = \frac{f_c \times A_c + (1 - f_c) \times A_o}{f_c \times D_c + (1 - f_c) \times D_o}. \qquad \textbf{[S2]}$$

Here, A_c and D_c correspond to the emission maxima of the acceptor (525 nm) and donor (475 nm) in the closed state and A_o and D_o correspond to the emission maxima of the acceptor and donor emission in the open state. Eq. **S2** was fit to the measured *R* vs. f_c data with A_c, D_c, A_o, and D_o as parameters (dotted line Fig. 4*B*). Although the absolute values of the parameters are not meaningful, the pairwise ratios between them should correspond to observed values for the high and low FRET states of the interaction. Accordingly, the curve fit provided $A_c/D_c = 2.2$ and $A_o/D_o = 0.5$, which were consistent with the R_{obs} for the high- (1.8 for $C_t = 16$ μM) and low-FRET (0.55 for $C_t = 160$ nM) states (Fig. 4*A*). It must be noted that A_c/D_c and A_o/D_o represent the upper and lower bounds, respectively, and therefore should be higher and lower, respectively, than the corresponding observed *R* values. Finally, the curve fit provided a third ratio of $A_c/A_o \sim 3$, which is consistent with the trend in R_{obs} following a transition from high- to low-FRET state. Unlike A_c/D_c and A_o/D_o, the last parameter A_c/A_o requires comparison across samples with different concentrations and the corresponding *R* value cannot be estimated from our dataset. Taken together, the experimental data are consistent with a two-state model of the FRET sensor that results in a linear relationship between *R* and f_c.

Measurement of effective concentration. SPASM sensor at a fixed concentration between 5 and 25 nM was mixed with increasing concentrations of unlabeled CAM (0–300 μM). Competition between unlabeled CAM and sensor CAM for binding to the sensor peptide results in a decrease in FRET (increased IR_{obs}) with increasing unlabeled CAM concentration. Competitive binding is expressed by the following equation:

$$CP_c \underset{k_{closed}}{\overset{k_{open}}{\rightleftharpoons}} CP_o + F \overset{K_d}{\rightleftharpoons} CP_o \bullet F. \quad \textbf{[S3]}$$

Here, C represents sensor CAM, P represents peptide, F represents free unlabeled CAM, subscript "o" represents the open, low-FRET state of the sensor, and "c" represents the closed, high-FRET state of the sensor. K_d is the equilibrium dissociation constant for F binding to P. The first-order on rate is k_{close}, and k_{open} is the first-order off rate of the interaction between C and P within the sensor. Eq. **4** can be used to solve for the amount of $CP_o \bullet F$

$$CP_o \bullet F = \frac{(C_t + F_t + K_{app}) - \sqrt{(C_t + F_t + K_{app})^2 - 4C_tF_t}}{2}. \quad \textbf{[S4]}$$

Here C_t and F_t are the total concentration of sensor and free unlabeled CAM. We define the apparent dissociation constant K_{app} as

$$K_{app} = \frac{K_d}{1 - f_c}. \quad \textbf{[S5]}$$

Here f_c is the fraction of the sensor in the closed state. The IR value (IR_{obs}) can be expressed as a linear combination of IR_{open} and IR_{closed}.

$$IR_{obs} = IR_{closed} \times \frac{CP_c}{C_t} + IR_{open} \times \frac{(CP_o + CP_o \bullet F)}{C_t}. \quad \textbf{[S6]}$$

Eq. **S6** can be simplified to express IR_{obs} in terms of IR_{noCAM}, IR_{open}, the known value of C_t and $CP_o \bullet F$ calculated from Eq. **S4**:

$$IR_{obs} = IR_{noCAM} \frac{(C_t - CP_o \bullet F)}{C_t} + IR_{open}\left(\frac{CP_o \bullet F}{C_t}\right). \quad \textbf{[S7]}$$

Eqs. **S4** and **S7** were fit to the experimental measurements with K_{app} and IR_{open} as the two floating parameters. For a given K_d of the bimolecular interaction, f_c can then be calculated from Eq. **S5**. The effective concentration can be estimated from the competitive binding assay described in Eqs. **S3–S7** as the concentration of free unlabeled CAM (F_t) when the concentration of $CP_o \bullet F = CP_c$. Effective concentration is related to K_d and f_c by

$$C_{eff} = \frac{f_c}{(1 - f_c)} K_d. \quad \textbf{[S8]}$$

Measurement of the opening rate (k_{open}) of the intramolecular interaction. CAM-Trp3 sensor at a concentration of 10 nM, containing either the 10-, 20-, or 30-nm helices, was activated with 4 mM Ca^{2+}. The C_{eff} for these sensors were 4, 0.3, and 0.1 μM, respectively, and their f_c values were nearly one. Therefore, at the 10-nM sensor concentration, a majority (>90%) of the sensor was in the closed state. The sensor was mixed with free unlabeled CAM to a final concentration of 300 μM ($\gg C_{eff}$) and the time course of FRET ratio (R_{obs}) was recorded. The sensor transitions to the open state at a rate of k_{open}. The vast excess of free unlabeled CAM should prevent subsequent interaction between sensor CAM and Trp3 peptide and result in loss of FRET signal.

1. Sivaramakrishnan S, et al. (2009) Combining single-molecule optical trapping and small-angle X-ray scattering measurements to compute the persistence length of a protein ER/K alpha-helix. *Biophys J* 97:2993–2999.
2. Sivaramakrishnan S, Spink BJ, Sim AY, Doniach S, Spudich JA (2008) Dynamic charge interactions create surprising rigidity in the ER/K alpha-helical protein motif. *Proc Natl Acad Sci USA* 105:13356–13361.
3. Lyu PC, Gans PJ, Kallenbach NR (1992) Energetic contribution of solventexposed ion pairs to alpha-helix structure. *J Mol Biol* 223:343–350.
4. George RA, Heringa J (2002) An analysis of protein domain linkers: Their classification and role in protein folding. *Protein Eng* 15:871–879.
5. Ohashi T, Galiacy SD, Briscoe G, Erickson HP (2007) An experimental study of GFP-based FRET, with application to intrinsically unstructured proteins. *Protein Sci* 16:1429–1438
6. Vogel SS, Thaler C, Koushik SV (2006) Fanciful FRET. *Sci STKE* 2006:re2.
7. Hoppe A, Christensen K, Swanson JA (2002) Fluorescence resonance energy transfer-based stoichiometry in living cells. *Biophys J* 83:3652–3664.
8. Spink BJ, Sivaramakrishnan S, Lipfert J, Doniach S, Spudich JA (2008) Long single alpha-helical tail domains bridge the gap between structure and function of myosin VI. *Nat Struct Mol Biol* 15:591–597.

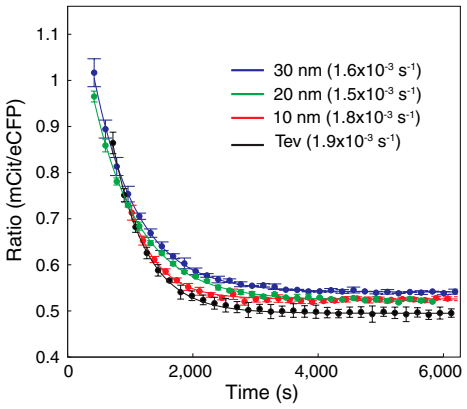
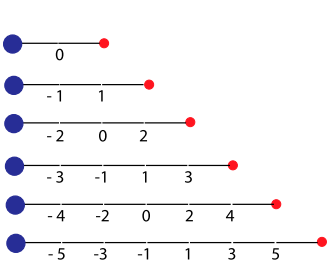


Fig. S1. Measuring off rate (k_{open}) using FRET. Change in FRET ratio (mCit/eCFP) following the addition of 100 μM unlabeled CAM at time = 0 s to 10 nM of intact SPASM sensor with 10-, 20-, and 30-nm ER/K α-helix. SPASM sensor switches from the closed to the open conformation at the rate of k_{open} and is held in the open conformation by binding of unlabeled CAM to sensor peptide. There is no significant change in k_{open} relative to the bimolecular interaction (TEV cleaved sensor). Note that zero time on the abscissa corresponds to the time of addition of excess CAM. The time lag in the dataset represents the time from mixing of unlabeled CAM with sensor to the beginning of data collection. The rate constants are mean ± SD of three separate datasets.



n	i	N	
1	1	2	
2	1	4	*Odd* $(n+1)/2^{n+1}$
3	2	8	
4	2	16	*Even* $n/2^{n+1}$
5	3	32	
6	3	64	

Fig. S2. Simplified model of stochastic breaks in the ER/K α-helix A simplified one-dimensional model is used to assess the effect of stochastic breaks on the rate of interaction between the ends of the ER/K α-helix. Blue and red circles represent CAM and the peptide separated by an ER/K α-helix. The number of stochastic breaks is *n*. *N* is the number of possible configurations of those breaks, assuming they are not correlated in space or time. The numbers below the helix represent the distance of the peptide from CAM following a break at that location. Negative numbers correspond to configurations where the peptide is to the left of CAM and vice versa. The number of configurations out of a total of *N* where CAM and peptide interact is represented by *i*. Interaction is possible when the sum of the numbers corresponding to the breaks equals zero. As witnessed from the table the fraction of configurations, and hence the probability of the two ends interacting, follows $n/2^{n+1}$ for even number of breaks and $(n+1)/2^{n+1}$ for odd number of breaks.