Rational design of FRET-based sensor proteins

Rational Design of FRET-Based Sensor Proteins

M. Merkx

Abstract Real-time imaging of molecular events inside living cells is important for understanding the basis of physiological processes and diseases. Genetically encoded sensors that use fluorescence resonance energy transfer (FRET) between two fluorescent proteins are attractive in this respect because they do not require cell-invasive procedures, can be targeted to different locations in the cell and are easily adapted through mutagenesis and directed evolution approaches. Most FRET sensors developed so far show a relatively small difference in emission ratio upon activation, which severely limits their application in high throughput cell-based screening applications. In our work, we try to develop strategies that allow design of FRET-based sensors with intrinsically large ratiometric changes. This rational design approach requires a better understanding and quantitative description of the conformational changes in these fusion proteins. In this chapter, I first discuss some of the key factors and strategies that determine the ratiometric response of FRET sensors, followed by an overview of our recent work in this area. Important concepts that will be discussed are (1) the conformational behavior of flexible peptide linkers to quantitatively describe the dependence of energy transfer on linker length and (2) the control of intramolecular domain interactions using the concept of effective molecular concentration.

1 Introduction

Real-time imaging of molecular events inside living cells is important for understanding the molecular basis of physiological processes and diseases [1]. Fluorescence is ideally suited for this purpose because it combines high sensitivity with subcellular resolution. A wide variety of synthetic fluorescent probes have been

M. Merkx (✉)
Laboratory of Chemical Biology, Department of Biomedical Engineering, Eindhoven University of Technology
e-mail: m.merkx@tue.nl

C.D. Geddes (ed.), *Reviews in Fluorescence 2008*,
Reviews in Fluorescence 2008, DOI 10.1007/978-1-4419-1260-2_3,

developed that allow specific detection of pH, redox status, metal ions, enzymatic activity, membrane potential, etc. Although some of them have become important molecular tools to probe intracellular events such as calcium signaling, these probes need to be introduced across cellular membranes and little control is possible over their cellular localization and concentration. Protein-based sensors based on native fluorescent proteins such as GFP and its color derivatives have become attractive alternatives for these synthetic probes [2–5]. Because these sensors are genetically encoded, their concentration and subcellular localization can be easily controlled. In addition, native protein domains can be used as receptor building blocks to ensure physiologically relevant binding affinities and specificities. Finally, molecular biology approaches such as directed evolution and site-directed mutagenesis can be used to optimize various sensor properties.

Several fluorescent sensor proteins consisting of a single GFP domain have been developed in which the chromophore environment was reengineered to become sensitive to ions such as H^+, Cl^-, Ca^{2+} [2, 4, 5]. This approach is useful for small molecules that have relatively high physiological concentrations and therefore do not require specific and high-affinity binding sites. A second, more general and therefore more interesting approach takes advantage of fluorescence resonance energy transfer (FRET) between a donor and an acceptor fluorescent domain. FRET is a photophysical effect whereby the energy that is absorbed by one fluorophore (donor) is transferred non-radiatively to a second fluorophore (acceptor) [6, 7]. The distance and orientation dependence of FRET make it extremely useful to detect conformational changes on the scale of individual proteins. Any biochemical event (e.g., ligand binding) that leads to a significant conformational change can be detected, irrespective of the protein and the nature of the protein–ligand interaction. Another important characteristic of FRET for sensor design is the possibility of ratiometric detection. An increase in FRET leads to a decrease in donor emission and an increase of acceptor emission. The ratio of donor and acceptor emission therefore provides a measure of the binding state that is independent of the absolute concentration of the sensor, which is important to obtain reliable, quantitative results in biological samples. Alternatively, FRET can be monitored by determining the lifetime of either the donor or the acceptor emission, both of which are sensitive to energy transfer.

The most common approach in the development of FRET-based sensors has been one of trial and error in which many constructs (which differ in linker lengths, fluorescent and receptor domains, etc.) are screened for optimal intracellular performance. While this strategy may be suitable to obtain a FRET sensor for a single, specific target with sufficient dynamic range to allow fluorescence imaging, little insight is gained into the factors that determine sensor performance. With some notable exceptions (see below) most FRET sensors show a relatively small difference in emission ratio upon activation, which severely limits their application in high throughput cell-based screening applications. In our work, we try to develop strategies that allow design of FRET-based sensors with intrinsically large ratiometric changes. This rational design approach requires a better understanding and quantitative description of the conformational changes in these fusion proteins. In this

review, first some of the key factors and strategies will be discussed that determine the ratiometric response of FRET sensors, followed by an overview of our recent work in this area. Important concepts that will be discussed are (1) the conformational behavior of flexible peptide linkers to quantitatively describe the dependence of energy transfer on linker length and (2) the control of intramolecular domain interactions using the concept of effective molecular concentration.

2 Factors That Affect the Ratiometric Change in FRET-Based Sensor Proteins

One of the strengths of FRET-based sensor proteins is their modular design in which a conformational change in a receptor domain is translated into a change in energy transfer between donor and acceptor domains. Alternatively, two receptor domains can be employed that interact in one state but not in the other. Examples of the latter approach are more common and include the calcium sensors developed by Tsien and others [8] and a variety of sensors for protein phosphorylation [2]. Despite the strong distance dependence of the energy transfer efficiency (Eq. (1)), it has proven surprisingly difficult to translate conformational changes into ratiometric changes that are larger than 1.5-fold. To understand why this is, we need to take a look at the factors that determine energy transfer efficiency in these fusion proteins.

The relation between the energy transfer efficiency E and the distance between the donor and fluorescent domains is described by the Förster equation:

$$E = \frac{R_0^6}{R_0^6 + r^6} \tag{1}$$

whereby the Förster distance, R_0, is given by

$$R_0 = 0.21 \left[\kappa^2 Q_D n^{-4} J(\lambda) \right]^{1/6} \tag{2}$$

Thus, the energy transfer efficiency not only depends on the distance between the chromophores but is also dependent on the quantum yield of the donor (Q_D), the spectral overlap between donor emission and acceptor absorption ($J(\lambda)$), the refractive index of the medium (n), and an orientational factor κ^2, which is related to the relative orientation of the donor emission and acceptor absorption dipole moments. Of the parameters that determine R_0, κ^2 is the most likely to change upon sensor activation. Equation (3) describes how κ^2 depends on the angles (θ_D, θ_A, θ_T, and ϕ) between the dipole moments of the donor and acceptor, as depicted in Fig. 1[6]:

$$\kappa^2 = (\cos\theta_T - 3\cos\theta_D \cos\theta_A)^2 = (\sin\theta_D \sin\theta_A \cos\phi - 2\cos\theta_D \cos\theta_A)^2 \tag{3}$$

If the orientation of donor and acceptor domains undergoes complete randomization while the donor is in the excited state, κ^2 averages out to a value of 2/3. If

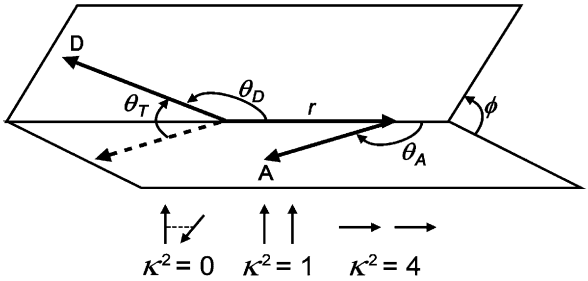

Fig. 1 Definition of angles between the donor emission dipole moment (*D*) and the acceptor absorption dipole moment (*A*), separated by distance *r*. Values for κ^2 are given for perpendicular, parallel, and collinear orientations of *D* and *A*

the orientation is fixed, κ^2 can vary between 0 (perpendicular), 1 (parallel), and 4 (collinear).

Having shown how the energy transfer efficiency depends on the distance and orientation of the donor and acceptor domains, we can identify several reasons for the low dynamic range of many genetically encoded FRET-based sensors:

(1) Changes in the distance between donor and acceptor are detected with the greatest sensitivity around the Förster distance, which is 48 Å for the most common pair of cyan and yellow fluorescent domains [9, 10]. However, because the fluorophores are buried inside the fluorescent protein's β-barrel and the intervening receptor domain is often also large, the energy transfer is relatively inefficient in most sensors, resulting in small changes in FRET.
(2) Flexible linkers of at least a few amino acids are required between two domains to allow proper folding of the individual domains. As a result, effective translation of the conformational change in the receptor domain into a significant conformational change for the entire sensor protein is hard to achieve.
(3) Because κ^2 can vary between 0 and 4, relatively small differences in orientation can in principle have a large effect on energy transfer efficiency. However, because of the limited possibilities to predict conformational preferences in fusion proteins, taking advantage of the orientational factor using rational design approaches is very difficult. In addition, transition dipole moments have been determined experimentally only for green fluorescent proteins [11, 12]. While the transition dipole moment of the yellow fluorescent domains are likely to be similar to that of EGFP, the transition dipole moments of frequently used donor domains such as ECFP and Cerulean are not known.

Despite the lack of true rational design approaches, two more empirical approaches are worth mentioning. A strategy that has been successfully applied in the optimization of several FRET sensors is the use of circularly permuted fluorescent protein domains [13–15]. In a circular permuted protein, the original C- and N-termini are connected by a short linker and new termini are generated at a solvent accessible loop [16]. Employment of a circular permuted form of Venus (a YFP variant) was shown to increase the ratiometric response of the cameleon calcium

sensors by 560%, probably due to a change in the orientational factor. However, no clear correlation between the type of circular permutation and the effect on the ratiometric change was observed, making this approach still (to some extent) a process of trial and error. A second strategy that was reported by Daugherty and coworkers is the use of directed evolution combined with FACS screening to select FRET sensors with an increased ratiometric change [17]. Variants of ECFP and EYFP, CyPet and YPet, were identified that, when linked via a flexible caspase recognition sequence, yielded a FRET sensor with a 20-fold ratiometric change upon cleavage by caspase-3. Although YPet and CyPet have been advertised as the donor and fluorescent domains of choice for FRET sensors [18, 19], we and others have recently shown that their improved FRET properties are caused by an increased tendency to form an intramolecular complex when linked via a flexible peptide linker, limiting their potential as general applicable fluorescent domains for FRET sensors [20, 21].

3 Quantitative Understanding of Energy Transfer by Modeling the Conformational Behavior of Flexible Linkers

Despite their importance in determining the functional properties of FRET sensors, little attention has been given to understand the effects of linker length and constitution on the energy transfer properties. Many studies reported a decrease in energy transfer upon increasing the linker length between donor and acceptor domains, but no attempt was made to quantitatively understand the efficiency of energy transfer as a function of linker length [22, 23]. Repeats of flexible (Gly) and hydrophilic (Ser) residues are popular linker peptides, because they are assumed to form a random coil structure in solution and to not interact with the protein domains [24]. We constructed a series of ECFP-linker-EYFP proteins in which the linker length was systematically varied between 1 and 9 GGSGGS repeats (CLY1–9) and carefully determined the amount of energy transfer for each linker length (Fig. 2) [10]. The energy transfer efficiency decreased from 71 to 43% when going from CLY1 (23 residue linker) to CLY9 (71 residue linker). Direct application of the Förster equation (assuming $R_0 = 48$ Å), would yield interchromophore distances ranging from 41 Å for CLY1 to 50 Å for CLY9. In other words, an increase in linker length of 48 amino acids results in an apparent increase in distance of only 9 Å! This analysis assumes that the energy transfer is due to a single fixed distance, however, which is clearly not the case when the fluorescent domains are linked via a flexible linker and a distribution of interchromophore distances is present. We therefore analyzed the energy transfer by considering the distribution of conformations available to the two fluorescent domains, which in turn is based on the conformational properties of the linker peptide.

Two models frequently used to describe the conformational behavior of random coil structures are the Gaussian chain model developed by Flory [25] and the Worm-Like Chain (WLC) model [26, 27]. For long flexible chains (as is the case here) both models are actually equivalent and only differ in the definition of the parameters

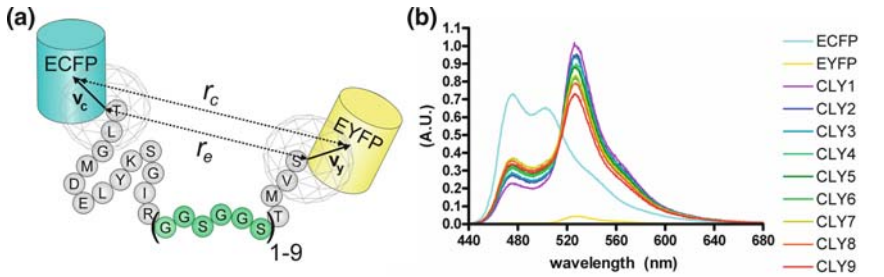

Fig. 2 (**A**) ECFP-linker-EYFP proteins containing a linker of 1–9 copies of the GGSGGS repeat (CLY1–9) that were used to understand the influence of linker length on energy transfer efficiency. Shown are the residues that are assumed to be part of the linker, as well as the spherical distribution of orientations that are accessible to the vectors that connect the ends of the peptide linkers and the chromophores ($\boldsymbol{v}_c$ and $\boldsymbol{v}_y$). The end-to-end distance of the linker (r_e) is translated into a chromophore-to-chromophore distance (r_c) by averaging over all possible orientations of $\boldsymbol{v}_c$ and $\boldsymbol{v}_y$. (**B**) Emission spectra of CLY1-9, ECFP, and EYFP were obtained using 420 nm excitation. Measurements were done using 0.2 μM protein in 50 mM Tris-HCl, 100 mM NaCl, 20 μM EDTA, 10% (v/v) glycerol, pH 8.0. Figures reprinted with permission from reference [10]. Copyright 2006 American Chemical Society

used to describe chain stiffness. The Gaussian chain model uses the characteristic ratio (C_∞), whereas the WLC model uses the persistence length (l_p). Here, we will analyze our data using the WLC model, because it has the advantage that it can be applied more easily to short and stiff polymers. In the WLC model the probability (P) of finding a certain end-to-end distance (r_e) for a chain with a total contour length l_c is described by Eq. (4).

$$\begin{aligned} P(r_e) = {} & 4\pi r_e^2 (3/4\pi l_p l_c)^{3/2} \exp(-3r_e^2/4l_p l_c)(1 - 5l_p/4l_c - 2r_e^2/l_c^2 \\ & + 33r_e^4/80l_p l_c^3 + 79l_p^2/160l_c^2 + 329r_e^2 l_p/120l_c^3 \\ & - 6799r_e^4/1600l_c^4 + 3441r_e^6/2800l_p l_c^5 - 1089r_e^8/12800l_p^2 l_c^6) \end{aligned} \tag{4}$$

The total contour length is equal to $b_0 n$, with n being the number of peptide bonds and $b_0 = 3.8$ Å (the average distance between adjacent C_α atoms). Using a persistence length l_p of 4.5 Å, the probability of finding a certain end-to-end distance can be calculated for each of the peptide linkers present in CLY1-9 (Fig. 3). To calculate the amount of energy transfer at a certain r_e, it should be realized that ECFP and EYFP can assume many different orientations, which differ in the interchromophore distance r_c and can therefore all contribute differently to the energy transfer efficiency. Consequently, the average energy transfer efficiency, <*E*>, was determined by averaging the value of E calculated for each combination of ECFP and EYFP orientations, assuming distances of 20 and 24 Å between the ends of the peptide linker and the ECFP and EYFP chromophores, respectively (Fig. 3b). In doing so, we can account for the fact that orientations with a short interchromophore distance contribute more to the energy transfer efficiency than orientations in which ECFP and EYFP are relatively far apart. Finally, the overall energy transfer

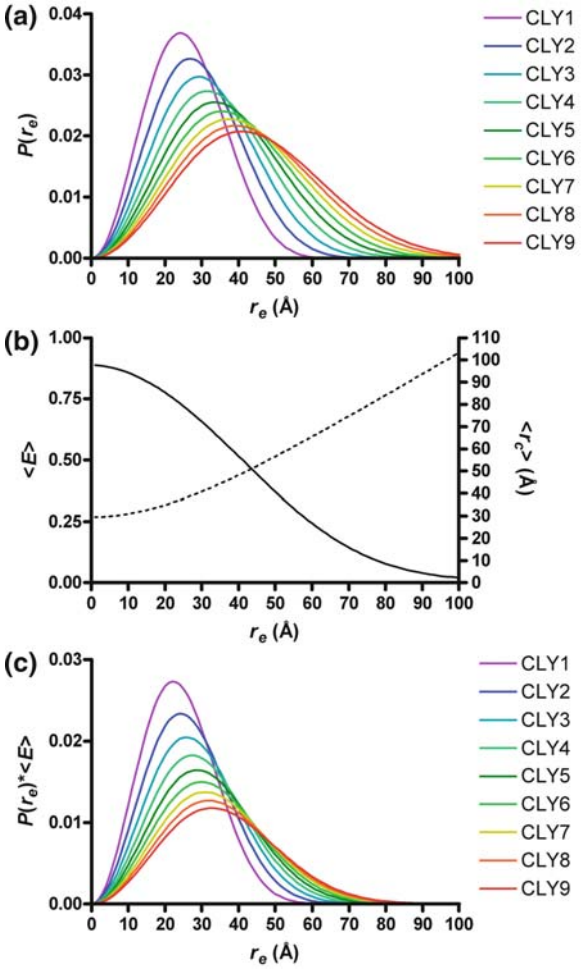

Fig. 3 Modeling the energy transfer efficiency for CLYx using the WLC model to describe the peptide linker. (**A**) Distribution functions showing the probability $P(r_e)$ for each end-to-end distance of the peptide linker for CLY1–9 calculated using the WLC model assuming a persistence length of 4.5 Å. (**B**) Plot showing the average energy transfer ($\langle E \rangle$, *solid line*) and the average interchromophore distance ($\langle r_c \rangle$, *dashed line*) as a function of r_e. (**C**) The contribution of each r_e to the overall energy transfer $\langle E \rangle_{ensemble}$, calculated by multiplying the probability of each r_e with the corresponding $\langle E \rangle$. Figure reprinted with permission from reference [10]. Copyright 2006 American Chemical Society

efficiency was calculated by multiplying the function describing $<E>$ as a function of r_e by the probability of finding a certain r_e. Integration of this curve yields the predicted overall energy transfer efficiency ($<E>_{ensemble}$) for each of the CLY proteins.

Figure 4 shows that there is excellent agreement between the predicted value for $<E>_{ensemble}$ using the random coil model and the experimentally determined energy transfer efficiency over the entire range of linker lengths. To show the strong dependence of the predicted energy transfer on the persistence length, Fig. 4 also shows the predicted curves for l_p values of 4.0 and 3.0 Å. The value of 4.5 Å corresponds to a C_∞ value of 2.3 in the Gaussian chain model, which is close to the C_∞ value of 2.16 previously reported for poly(Gly) [28].

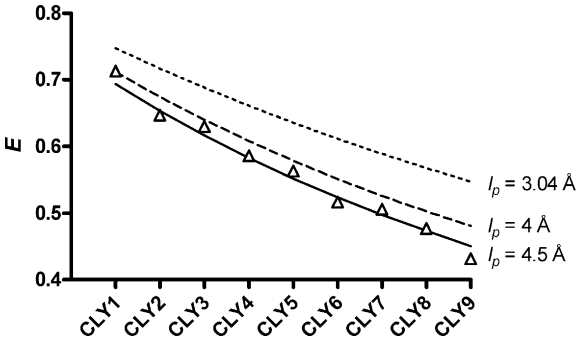

Fig. 4 Comparison of the experimentally determined energy transfer efficiency (*triangles*) with the energy transfer efficiency predicted by the WLC model using persistence lengths of 3.0, 4.0, and 4.5 Å. Figure reprinted with permission from reference [10]. Copyright 2006 American Chemical Society

The fact that the energy transfer efficiency between ECFP and EYFP can be quantitatively understood by describing the flexible linker as a random coil has several important implications for FRET sensor design. First, this study shows that random coil peptide linkers form relatively compact assemblies, giving rise to relatively high energy transfer efficiencies. For example, while the long peptide linker used in CLY9 (71 amino acids) can span a distance of 252 Å when fully extended, the most abundant end-to-end distance is only 41 Å. A second issue that is sometimes misunderstood is that the energy transfer efficiency between fluorescent domains separated by flexible linkers reflects a broad distribution of distances and not a single distance. Because of the nonlinear dependency of E on distance, it is not correct to calculate an average distance based on the energy transfer efficiency. The apparent average distance will be underestimated for long linkers and overestimated for short linkers.

4 Quantitative Understanding of the Effect of Flexible Peptide Linkers on Effective Concentration

In the previous section, we showed that the conformational distribution of two protein domains connected via Ser/Gly linkers can be accurately described by modeling the linkers as flexible chains with a persistence length of 4.5 Å [29]. The same model can also be used to calculate the effective local concentration of one protein domain with respect to another in a system in which the two protein domains are linked [27, 30, 31]. A clear demonstration of the predictive power of this model, and the importance of considering conformational distributions in the design of FRET-based sensor proteins, is our work on the Zn(II) sensor CALWY [29, 32]. This sensor is based on a serendipitous discovery that Zn(II) mediates an interaction between ATOX1 and WD4, two metal binding domains involved in Cu(I) homeostasis. A sensor system consisting of ECFP-Atox1 and WD4-EYFP was shown to bind Zn(II) at subnanomolar levels using equimolar amounts of both proteins (2 μM) [32]. Next, a series of ECFP-Atox1-linker-WD4-EYFP fusion proteins were

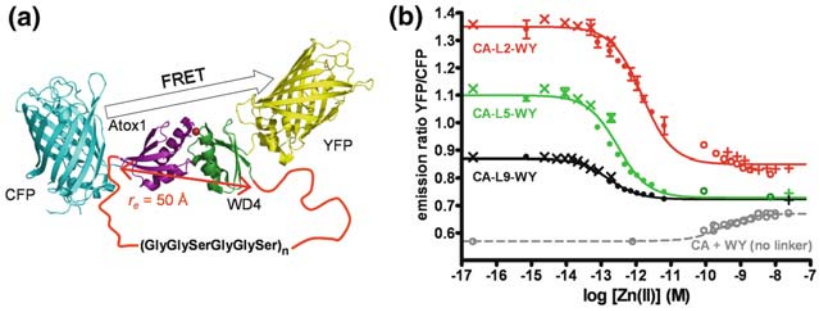

Fig. 5 Variation of the linker length in CA-Ln-WY allows rational tuning of Zn(II) affinity in the picomolar to femtomolar range. (**A**) Structural model of CA-Ln-WY in the Zn(II) bound state. (**B**) Zn(II) titration for CFP-Atox1-linker-WD4-YFP proteins with 2, 5, and 9 GGSGGS repeats. Titrations were performed in 50 mM Tris, 100 mM NaCl, 1 mM DTT, 10% glycerol at pH 7.5 using various Zn(II) buffers. Also shown is a titration for a 1:1 mixture of CFP-Atox1 and WD4-YFP (2 μM each). Figures reprinted with permission from reference [29]. Copyright 2007 American Chemical Society

constructed in which the linker length was systematically varied between 2 and 9 GGSGGS repeats (CA-Ln-WY, with n = the number of GGSGGS repeats; Fig. 5a) [29]. This series of sensor proteins allowed us to systematically study the effect of the linker length on both the relative ratiometric change and the affinity for Zn(II), which was expected to increase due to the chelate effect.

The Zn(II) affinity showed a consistent trend with K_d values ranging from 170 fM for the protein with the longest linker (CA-L9-WY) to 1.4 pM for the shortest linker (CA-L2-WY). Remarkably, the Zn(II)-induced intramolecular complex formation between Atox1 and WD4 resulted in a decrease in energy transfer for all sensor constructs (Fig. 5b). The emission ratio in the unbound state was strongly dependent on the linker length, whereas similar emission ratios were observed for the Zn(II)-bound state. This behavior can be readily understood by considering the conformational behavior of the linker. The distance between the C-terminus of Atox1 and the N-terminus of WD4 in the Zn(II) complex has been estimated to be 50 Å, based on the X-ray structure of an analogous Cd(II)-bridged Atox1–Atox1 homodimer complex [33]. Figure 6 shows the distribution of end-to-end distances of the

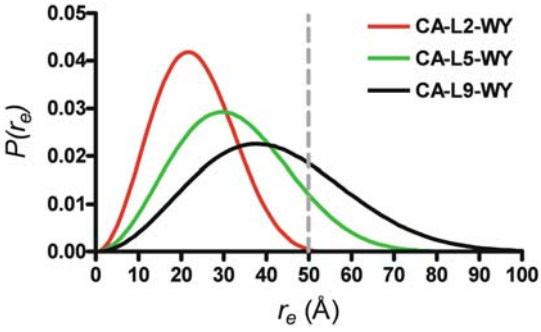

Fig. 6 Distribution functions showing the probability $P(r_e)$ of a certain end-to-end distance r_e for the peptide linkers used in CA-L2-WY, CA-L5-WY, and CA-L9-WY. Figure reprinted with permission from reference [29]. Copyright 2007 American Chemical Society

peptide linker as predicted by the WLC model for linkers containing 2, 5, and 9 GGSGGS repeats. Even for the longest linker, the average end-to-end distance is smaller than 50 Å, resulting in a predicted net decrease in energy transfer upon formation of the Zn(II)-bound complex.

The WLC model can also be used to predict the influence of linker length on the Zn(II) affinity. When both metal binding domains are not connected, the formation of the ternary CFP-Atox1-Zn(II)-WD4-EYFP can be described using the following equilibria:

$$\mathrm{CA} + \mathrm{Zn(II)} \leftrightarrows \mathrm{CA\text{-}Zn(II)} \qquad K^{\mathrm{s}}_{\mathrm{CA}} = \frac{[\mathrm{CAZn(II)}]}{[\mathrm{CA}][\mathrm{Zn(II)}]} \tag{5a}$$

$$\mathrm{CA\text{-}Zn(II)} + \mathrm{WY} \leftrightarrows \mathrm{CA\text{-}Zn(II)\text{-}WY} \qquad K^{\mathrm{s}}_{\mathrm{WY'}} = \frac{[\mathrm{CAZn(II)WY}]}{[\mathrm{CAZn(II)}][\mathrm{WY}]} \tag{5b}$$

$$\mathrm{CA} + \mathrm{WY} + \mathrm{Zn(II)} \leftrightarrows \mathrm{CA\text{-}Zn(II)\text{-}WY} \quad K^{\mathrm{s}} = K^{\mathrm{s}}_{\mathrm{CA}} K^{\mathrm{s}}_{\mathrm{WY'}} = \frac{[\mathrm{CAZn(II)WY}]}{[\mathrm{CA}][\mathrm{WY}][\mathrm{Zn(II)}]} \tag{5c}$$

Binding of Zn(II) to the single, peptide-linked protein sensors may be described analogously, using the following equilibria:

$$\overbrace{\mathrm{CA\quad WY}} + \mathrm{Zn(II)} \leftrightarrows \overbrace{\mathrm{CA\text{-}Zn(II)\quad WY}} \qquad K^{l}_{CA} = \frac{[\overbrace{\mathrm{CAZn(II)\quad WY}}]}{[\overbrace{\mathrm{CA\quad WY}}][\mathrm{Zn(II)}]} \tag{6a}$$

$$\overbrace{\mathrm{CA\text{-}Zn(II)\quad WY}} \leftrightarrows \overbrace{\mathrm{CA\text{-}Zn(II)\text{-}WY}} \qquad K^{l}_{WY'} = \frac{[\overbrace{\mathrm{CAZn(II)WY}}]}{[\overbrace{\mathrm{CAZn(II)\quad WY}}]} \tag{6b}$$

$$\overbrace{\mathrm{CA\quad WY}} + \mathrm{Zn(II)} \leftrightarrows \overbrace{\mathrm{CA\text{-}Zn(II)\text{-}WY}} \qquad K^{l} = K^{l}_{CA} K^{l}_{WY'} = \frac{[\overbrace{\mathrm{CAZn(II)WY}}]}{[\overbrace{\mathrm{CA\quad WY}}][\mathrm{Zn(II)}]} \tag{6c}$$

Binding of Zn(II) to the first protein domain (CA) can be assumed to be the same in both cases ($K^{\mathrm{s}}_{\mathrm{CA}} = K^{\mathrm{l}}_{\mathrm{CA}}$). Binding of the second metal binding domain (WY) is intramolecular for the linked protein (Eq. (6b)) and intermolecular for the ternary system (Eq. (5b)). The intramolecular binding constant $K^{\mathrm{l}}_{\mathrm{WY'}}$ is related to the intermolecular binding constant $K^{s}_{\mathrm{WY'}}$ by $K^{\mathrm{l}}_{\mathrm{WY'}} = c_{\mathrm{eff}} K^{\mathrm{s}}_{\mathrm{WY'}}$ where c_{eff} is the effective concentration of the second protein domain after formation of the initial Zn(II) complex. The effective concentration for the formation of an intramolecular complex is proportional to the probability density $p(r_{\mathrm{e}})$ for the end-to-end distance, r_{e}, that the linker needs to span in the complex, according to Eqs. (7) and (8) [27, 30, 31].

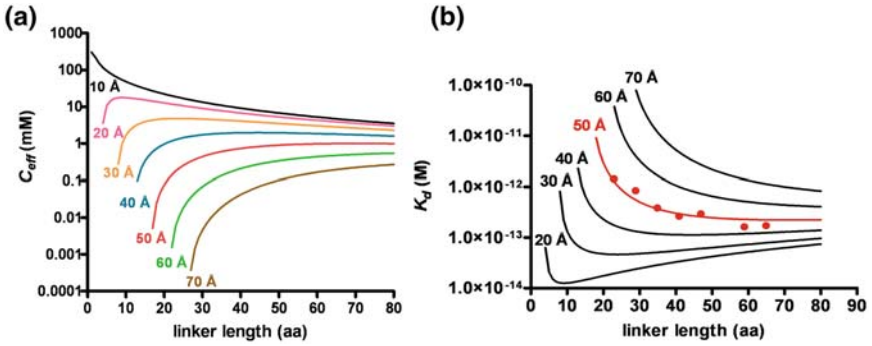

Fig. 7 (**A**) Calculated effective concentration (c_{eff}) as a function of linker length for different values of the end-to-end distance in the complex. (**B**) K_{D} values predicted by the WLC model assuming various values for the distance between the C-terminus of ATOX1 and the N-terminus of WD4 in the Zn(II) complex. For comparison, K_{D} values that were experimentally determined for various CA-Lx-WY proteins are shown as *circles*

$$p(r_e) = \frac{p(r_e)}{4\pi r_e^2} \tag{7}$$

$$c_{\text{eff}} = \frac{p(r_{\text{e}})}{N_{\text{AV}}} \tag{8}$$

where r_{e}, l_{p}, and l_{c} are in dm and N_{AV} is Avogadro's number.

Figure 7a shows the dependence of the effective concentration as a function of linker length for several values of r_{e}. The model predicts that an increase in linker length can give rise to either an increase or a decrease in effective concentration, depending on the distance that the linker needs to bridge in the complex. The dissociation constant for complex formation (K_{D}) can be calculated using c_{eff} and the formation constant (K^{s}) for the ternary complex, which was previously determined to be $4.5 \times 10^{15}\,\text{M}^{-2}$ using Eq. (9).

$$\begin{aligned} K^{\text{l}} &= c_{\text{eff}} K^{\text{s}} \\ K_{\text{D}} &= \frac{1}{K^{\text{l}}} = \frac{1}{c_{\text{eff}} 4.5 \times 10^{15}\text{M}^{-2}} \end{aligned} \tag{9}$$

The experimental K_{D} values agree very well with the calculated values predicted by the WLC model for an end-to-end distance of 50 Å (Fig. 7b). Thus, analysis of the distance distribution of the linker in terms of random coil behavior fully explains both the magnitude and the linker length dependence of the affinity enhancement.

5 Chelating Fluorescent Protein Chimeras as Efficient Zn(II) Sensor Proteins

The example discussed above shows that the sensor properties of FRET sensors with long flexible peptide linkers can be rationalized and even be predicted [34].

However, the CALWY sensors showed a relatively small change in emission ratio upon Zn(II) binding. Although the dynamic range can be improved by further decreasing the linker length, the WLC models also predict a significant decrease in Zn(II) affinity when using shorter peptide linkers. One problem encountered here (and in many other FRET sensor systems) is that the fluorophores are relatively far apart in the ligand-bound state, resulting in a low-energy transfer efficiency. In the ligand-free state, a distribution of conformations is present that has an average energy transfer efficiency that is also rather low, resulting in a small net change upon ligand binding.

One approach to obtain sensor proteins with an intrinsically large difference in FRET between the two sensor states is to create complementary target binding sites directly at the surface of the donor and the acceptor fluorescent domains. This approach is particularly attractive for metal binding sensors. The factors that determine metal complex stabilities are relatively well understood [35], thus allowing successful creation of de novo metal binding sites on protein surfaces [36]. Metal binding to each of the two fluorescent domains should result in a complex in which the two fluorophores are at a fixed position and as close together as sterically possible. To demonstrate the feasibility of this approach we took advantage of a report by Jensen and coworkers who found that the interaction between ECFP and EYFP can be enhanced by the introduction of Zn(II)-coordinating residues (Ser208C and Tyr39His) at their 'natural' dimerization interface [37]. Wild-type GFP has a well-known tendency to dimerize at high concentrations ($K_{dim} = 0.11$ mM) [38], forming a complex in which both β-barrels are in an anti-parallel orientation [39]. Jensen and coworkers found a weak Zn(II) affinity of 0.7 mM when mutant ECFP and EYFP were present at 10 μM, but we reasoned that an ECFP-linker-EYFP sensor should display a substantially higher affinity due to the high effective concentration of the two domains.

Introduction of the Ser208Cys and Tyr39His mutations on both ECFP and EYFP in CLY9 yielded a zinc sensor (called ZinCh-9) that displays an impressive 4-fold change in emission ratio upon addition of Zn(II) [34]. Unexpectedly, Zn(II) titrations showed a biphasic Zn(II) sensitivity over a large range of Zn(II) concentrations between 10 nM and 1 mM (Fig. 8). The presence of high ($K_D = 100$ nM)- and low-affinity Zn(II) binding ($K_D = 50$ μM) sites is inconsistent with the formation of an anti-parallel complex between ECFP and EYFP. Such a complex would be expected to have two equivalent Zn(II) binding sites, each containing one Cys208 and one His39 ligand. A second reason to believe that the Zn(II) complex is different from that observed for the GFP dimer is that formation of this complex is highly unlikely for sensor constructs with short linkers such as ZinCh-1, which was shown to form similar Zn(II) complexes as ZinCh-9. We therefore examined the Zn(II) binding properties of sensor variants that contained only the Ser208Cys mutation (CLY9-S208C) or the Tyr39His mutation (CLY9-Y39H). CLY9-S208C showed the same high affinity for Zn(II) as ZinCh-9, but not the additional increase in emission ratio at higher Zn(II) concentrations (Fig. 9). CLY9-Y39H on the other hand showed very weak Zn(II) binding. These results are consistent with a model in which ECFP and EYFP interact in a parallel orientation that allows the two Cys208 residues to

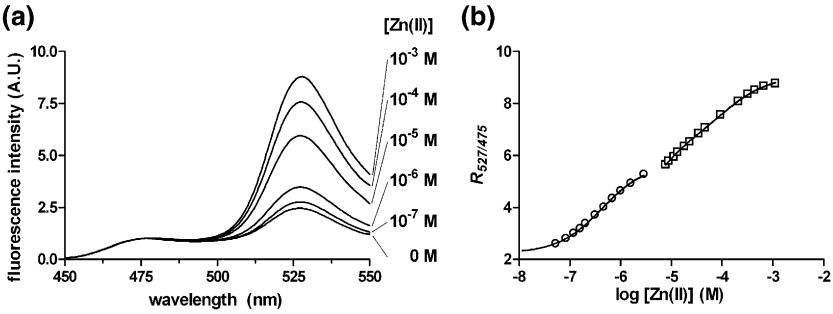

Fig. 8 Zn(II) titration experiments for ZinCh-9. (**a**) Emission spectra of ZinCh-9 in the presence of increasing amounts of Zn(II), normalized at the emission maximum of ECFP at 475 nm. (**b**) The EYFP/ECFP emission ratio of ZinCh-9 as a function of Zn(II) concentration shows a biphasic response. Data were obtained using Ba(II)-EGTA as a Zn(II) buffer for low Zn(II) concentrations (*circles*) and a non-buffered titration for high Zn(II) concentrations (*squares*). Figures reprinted with permission from reference [34]. Copyright 2007 Elsevier

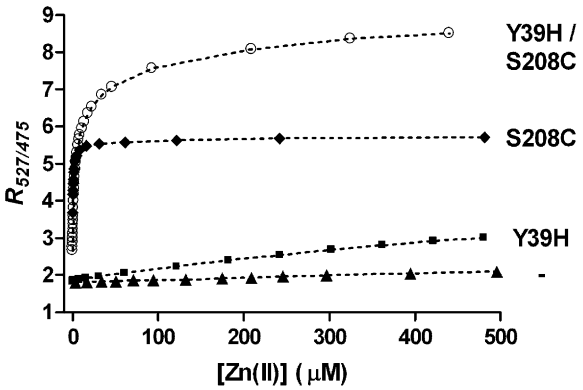

Fig. 9 Comparison of the Zn(II) binding properties of ZinCh-9 (*open circles*), CLY9-S208C (*closed circles*), CLY9-Y39H (*squares*), and CLY9 (*triangles*) showing the EYFP/ECFP emission ratio as a function of Zn(II) concentration. Figure reprinted with permission from reference [34]. Copyright 2007 Elsevier

form a high-affinity Zn(II) binding site. Following the formation of the initial Zn(II) complex, the two His39 residues are pre-organized to form a second low-affinity binding site.

To gain more insight into the changes that accompany each Zn(II) binding step, Zn(II) titrations were also monitored using fluorescence anisotropy measurements (Fig. 10). The fluorescence emission anisotropy observed using direct excitation of EYFP showed that the high-affinity Zn(II) binding results in an increase in anisotropy from 0.304 ± 0.004 to 0.328 ± 0.004. This increase in anisotropy indicates an increase in the apparent molecular volume of $\sim$100% and is thus consistent with complex formation between ECFP and EYFP. To learn about the changes in the relative orientation of the ECFP and EYFP chromophores, the change in anisotropy for EYFP emission was determined after excitation of ECFP. In the absence of Zn(II), the anisotropy is low (0.09), which is consistent with a random orientation of the EYFP domain with respect to the ECFP domain. Binding of Zn(II) to the

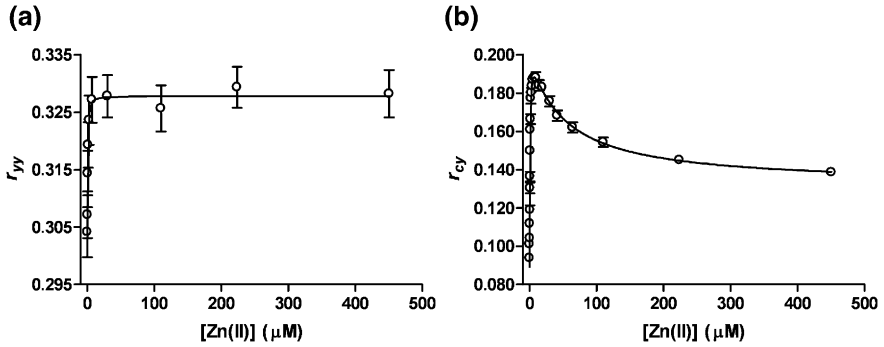

Fig. 10 Zn(II) binding to ZinCh-6 monitored using fluorescence anisotropy. (**A**) Fluorescence emission anisotropy at 527 nm using 500 nm excitation (r_{yy}). (**B**) Fluorescence emission anisotropy at 527 nm using 420 nm excitation (r_{cy}). The solid line is a fit of the data using Eq. (5.11) with $K_{d1} = 0.35\,\mu M$ and $K_{d2} = 50\,\mu M$. Figures reprinted with permission from reference [34]. Copyright 2007 Elsevier

high-affinity site results in a substantial increase in anisotropy to ~0.19, indicating that a more fixed orientation between ECFP and EYFP is formed. Population of the second Zn(II) site results in a decrease in anisotropy to 0.13. This decrease in anisotropy is most likely due to a second conformational change, in which the transition dipole moments of ECFP and EYFP are oriented in a way that results in an increase in energy transfer efficiency due to a more favorable κ^2, and, at the same time, a larger amount of depolarization (Fig. 11).

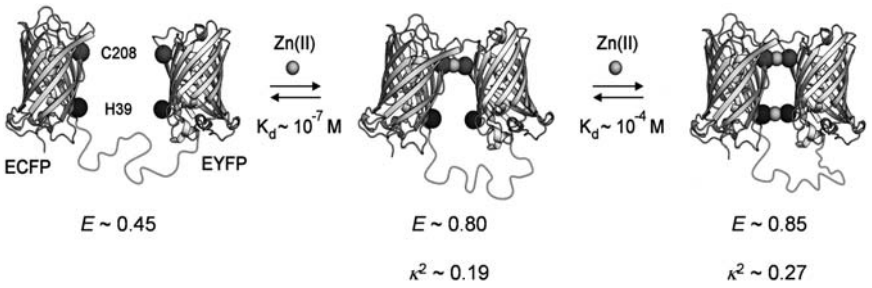

Fig. 11 Model of Zn(II) binding to ZinCh sensor proteins. Zinc binding to the two Cys208 residues stabilizes the formation of an intramolecular complex of ECFP and EYFP in a parallel orientation, resulting in a large increase in FRET. After binding of Zn(II) to the high-affinity site, the two His39 s are pre-organized to form a second, low-affinity Zn(II) binding site. Binding to this low-affinity site is accompanied by a small conformational change, resulting in a further increase in FRET. Figure adapted from reference [34] with permission. Copyright 2007 Elsevier

Assuming that the fluorescent domains adopt a conformation in which the two Cys208 residues are facing each other and the domains are in a parallel orientation, the distance between the fluorophores can be estimated to be ~31 Å. The energy transfer efficiencies of 80 and 85% that are observed for binding of the first and the

second Zn(II) ions then imply values of κ^2 of 0.19 and 0.27, respectively. Since κ^2 can vary between 0 and 4 it is clear that orientations of the transition dipole moments in these complexes are still far from optimal for energy transfer. Further optimization of the energy transfer in the complex form might be achieved using circular permuted forms of one of the fluorescent domains. However, rational prediction of these effects is difficult, particularly because the transition dipole moment of ECFP is still unknown.

6 Taking Advantage of 'Stickiness': FRET Sensor Proteins Based on Conformational Switching

GFP and many of its color variants are known to dimerize at high (local) protein concentrations. The use of A206K or A206R mutants has therefore been advocated to prevent the formation of intramolecular interactions between fluorescent domains in FRET-based sensors [18]. Before we show how deliberate promotion of intramolecular interactions can actually be used to enhance the ratiometric response of FRET-based protease sensors, it is useful to first discuss an apparent contradiction between our findings and results obtained by Tsien and coworkers who reported a dissociation constant for the GFP dimer (K_{dim}) of 0.11 mM [38]. The WLC model predicts effective concentrations in the millimolar range for the ECFP and EYFP domains when connected via a flexible GlyGlySer-linker, yet our studies clearly show the absence of significant interactions between ECFP and EYFP. The solution to this apparent contradiction may be found in the type of complex that can be formed. In the absence of a linker, GFP dimerizes in an anti-parallel orientation, at least according to the X-ray structure [39]. In our sensor proteins, the anti-parallel orientation is strongly disfavored, because this orientation would result in an entropically unfavorable extended linker conformation. This explanation implies that K_{dim} for the formation of the parallel complex is significantly higher than the predicted effective concentration for this complex, which is $\sim$3 mM for the longest linker.

Although wild-type ECFP and EYFP do not interact when connected by flexible SGGSGG linkers, we showed that subtle redesign of the dimerization interface by substitution of the hydrophilic Ser208 by a hydrophobic phenylalanine promotes intramolecular complex formation between the two fluorescent domains, resulting in a large increase in energy transfer (Fig. 12) [21]. The introduction of a second mutation at the dimerization interface, V224L, did not by itself enhance intramolecular interactions between ECFP and EYFP, but it did induce a further increase in energy transfer efficiency in the presence of the S208F mutation. Cleavage of the linker by a protease resulted in a 16-fold decrease in emission ratio (Fig. 13). The absence of any FRET after cleavage shows that the complex between these more 'sticky' variants of ECFP and EYFP is readily disrupted upon cleavage of the linker and no intermolecular complex is formed at the low micromolar concentrations used in these experiments.

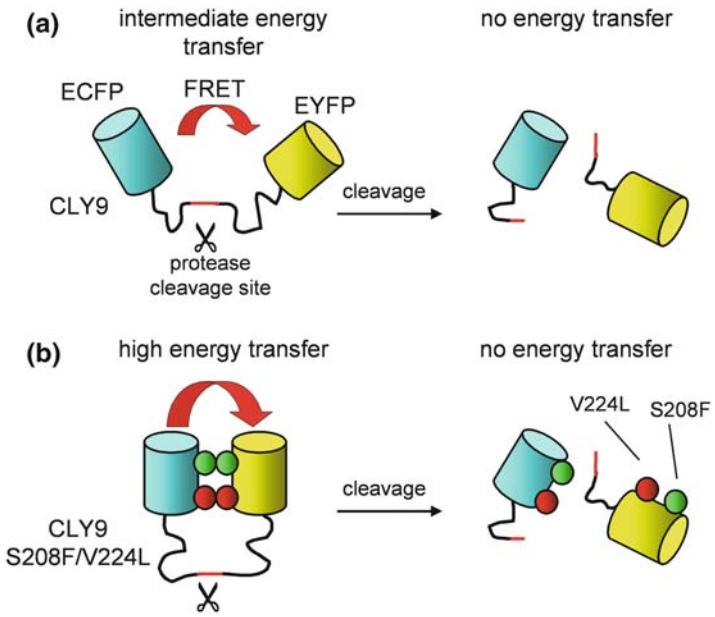

Fig. 12 Improvement of FRET-based sensor proteins by redesign of the dimerization interface between ECFP and EYFP. (**A**) In the classical sensor design the random coil behavior results in a distribution of conformations with a moderate energy transfer efficiency. (**B**) Introduction of S208F and V224L mutations promotes intramolecular complex formation, resulting in a substantial increase in energy transfer efficiency prior to protease cleavage. Figures reproduced with permission from reference [21]. Copyright Wiley-VCH Verlag GmbH & Co. KGaA

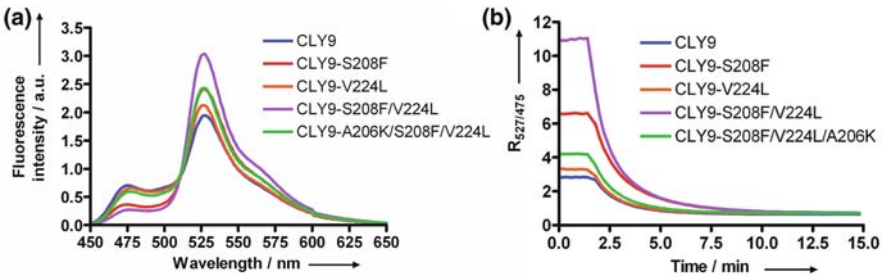

Fig. 13 (**A**) Emission spectra of CLY9 and various mutants before linker cleavage (excitation at 420 nm). (**B**) EYFP/ECFP emission ratios of CLY9 and mutant variants, monitored over time after the addition of proteinase K. Figures reproduced with permission from reference [21]. Copyright Wiley-VCH Verlag GmbH & Co. KGaA

Promoting intramolecular domain interactions in either the on or the off state of a sensor as we did in the protease sensor could be a more general approach to rationally increase the ratiometric response of FRET-based sensor proteins. Given the high sequence and structural similarities among the many color variants of GFP,

these mutations are generally applicable. In fact, S208F and V224L are both present in YPet, suggesting that the improved FRET properties of YPet (and CyPet) are largely due to an increased 'stickiness' [20, 21]. To allow more rational application of these mutations, it will be useful to actually determine the dissociation constant for the S208F/V224L mutants. For now, the K_D for this interaction can be estimated to lie somewhere in the range between 10 μM and 1 mM. A K_D lower than 10 μM would result in complex formation in the absence of the linker, while a K_D higher than 1 mM would be inconsistent with the intramolecular complex formation that is observed in the presence of the linker.

7 Conclusion and Outlook

FRET between fluorescent protein domains is a powerful and generally applicable approach to translate a protein conformational change into a ratiometric fluorescent read-out for intracellular imaging. Rational design strategies to improve the ratiometric response of FRET sensors are important in order to broaden the scope of this technology, for example, to allow high throughput applications based on fluorescence plate readers and fluorescence-assisted cell sorting (FACS) [40]. Long, flexible peptide linkers consisting of GlyGlySer repeats are attractive building blocks for FRET-based sensors, because simple models exist that can quantitatively describe their random conformational behavior. We have shown how these models can be used to understand two important properties of FRET-based sensor proteins, energy transfer efficiency and the strength of intramolecular interactions between receptor domains. The concept of effective concentration provides a useful framework to understand and predict conformational preferences in multidomain proteins, such as FRET sensors. Redesigning the GFP dimerization interface was shown to be an attractive strategy to obtain FRET sensors with an intrinsically large change in emission ratio. Introduction of Zn(II) binding residues yielded a Zn(II) sensor with a 4-fold change in emission ratio, whereas the introduction of hydrophobic residues enhanced the ratiometric change of a protease sensor from 4- to 16-fold. These examples illustrate that the distance dependence of energy transfer efficiency can be employed successfully in rational design approaches. Given the difficulty of controlling the precise orientation of protein–protein complexes, it will probably remain difficult to also include the orientational dependence of energy transfer into the rational design of FRET sensors. Instead, directed evolution approaches combined with high throughput screening methods such as FACS are probably best suited to optimize the orientation between donor and acceptor dipole moments.

Acknowledgments I would like to thank all current and previous members of my group working on FRET-based sensor proteins. In particular, I would like to acknowledge Ir. Jan Vinkenborg, Dr Toon Evers, Dr Melissa Koay, and Dr Misha Golynskiy for critically reading this chapter. This work was supported by a Human Frontier of Science Program Young Investigator Grant (RGY0068-2006C)

References

1. R.Y. Tsien, Building and breeding molecules to spy on cells and tumors, *FEBS Lett.* **579**, 927–932 (2005).
2. S.B. Van Engelenburg, and A.E. Palmer, Fluorescent biosensors of protein function, *Curr. Opin. Chem. Biol.* **12**, 60–65 (2008).
3. A. Miyawaki, Visualization of the spatial and temporal dynamics of intracellular signaling, *Dev. Cell* **4**, 295–305 (2003).
4. D.M. Chudakov, S. Lukyanov, and K.A. Lukyanov, Fluorescent proteins as a toolkit for in vivo imaging, *Trends Biotechnol.* **23**, 605–613 (2005).
5. N.C. Shaner, G.H. Patterson, and M.W. Davidson, Advances in fluorescent protein technology, *J. Cell. Sci.* **120**, 4247–4260 (2007).
6. J.R. Lakowicz, Principles of fluorescence spectroscopy 2nd ed. 1999, New York: Kluwer Academic/Plenum Publishers.
7. R. Heim, Green fluorescent protein forms for energy transfer, *Methods Enzymol.* **302**, 408–423 (1999).
8. M. Mank, and O. Griesbeck, Genetically encoded calcium sensors, *Chem. Rev.* **108**, 1550–1564 (2008).
9. G.H. Patterson, D.W. Piston, and B.G. Barisas, Förster distances between green fluorescent protein pairs, *Anal. Biochem.* **284**, 438–440 (2000).
10. T.H. Evers, E.M.W.M. Van Dongen, A.C. Faesen, E.W. Meijer, and M. Merkx, Quantitative understanding of the energy transfer between fluorescent proteins connected via flexible peptide linkers, *Biochemistry* **45**, 13183–13192 (2006).
11. X.H. Shi, J. Basran, H.E. Seward, W. Childs, C.R. Bagshaw, and S.G. Boxer, Anomalous negative fluorescence anisotropy in yellow fluorescent protein (YFP 10C): Quantitative analysis of FRET in YFP dimers, *Biochemistry* **46**, 14403–14417 (2007).
12. D. Stoner-Ma, E.H. Melief, J. Nappa, K.L. Ronayne, P.J. Tonge, and S.R. Meech, Proton relay reaction in green fluorescent protein (GFP): Polarization-resolved ultrafast vibrational spectroscopy of isotopically edited GFP, *J. Phys. Chem. B* **110**, 22009–22018 (2006).
13. T. Nagai, S. Yamada, T. Tominaga, M. Ichikawa, and A. Miyawaki, Expanded dynamic range of fluorescent indicators for Ca^{2+} by circularly permuted yellow fluorescent proteins, *Proc. Natl. Acad. Sci. U S A* **101**, 10554–10559 (2004).
14. M. Mank, D.F. Reiff, N. Heim, M.W. Friedrich, A. Borst, and O. Griesbeck, A FRET-based calcium biosensor with fast signal kinetics and high fluorescence change, *Biophys. J.* **90**, 1790–1796 (2006).
15. A.E. Palmer, M. Giacomello, T. Kortemme, S.A. Hires, V. Lev-Ram, D. Baker, and R.Y. Tsien, Ca^{2+} indicators based on computationally redesigned calmodulin-peptide pairs, *Chem. Biol.* **13**, 521–530 (2006).
16. G.S. Baird, D.A. Zacharias, and R.Y. Tsien, Circular permutation and receptor insertion within green fluorescent proteins, *Proc. Natl. Acad. Sci. U S A* **96**, 11241–11246 (1999).
17. A.W. Nguyen and P.S. Daugherty, Evolutionary optimization of fluorescent proteins for intracellular FRET, *Nat. Biotechnol.* **23**, 355–360 (2005).
18. N.C. Shaner, P.A. Steinbach, and R.Y. Tsien, A guide to choosing fluorescent proteins, *Nat. Methods* **2**, 905–909 (2005).
19. X. You, A.W. Nguyen, A. Jabaiah, M.A. Sheff, K.S. Thorn, and P.S. Daugherty, Intracellular protein interaction mapping with FRET hybrids, *Proc. Natl. Acad. Sci. U S A* **103**, 18458–18463 (2006).
20. T. Ohashi, S.D. Galiacy, G. Briscoe, and H.P. Erickson, An experimental study of GFP-based FRET, with application to intrinsically unstructured proteins, *Protein Sci.* **16**, 1429–1438 (2007).
21. J.L. Vinkenborg, T.H. Evers, S.W.A. Reulen, E.W. Meijer, and M. Merkx, Enhanced sensitivity of FRET-based protease sensors by redesign of the GFP dimerization interface, *ChemBioChem.* **8**, 1119–1121 (2007).

22. L.M. Felber, S.M. Cloutier, P. Jichlinski, H.-J. Leisinger, D. Deperthes, C. Kündig, V. Brossard, and T. Kishi, Evaluation of the CFP-substrate-YFP system for protease studies: Advantages and limitations, *Biotechniques* **36**, 878–885 (2004).
23. R. Arai, H. Ueda, A. Kitayama, N. Kamiya, and T. Nagamune, Design of the linkers which effectively separate domains of a bifunctional fusion protein, *Protein. Eng.* **14**, 529–532 (2001).
24. D. Neri, M. Momo, T. Prospero, and G. Winter, High-affinity antigen binding by chelating recombinant antibodies (CRAbs), *J. Mol. Biol.* **246**, 367–373 (1995).
25. P.J. Flory, Statistical mechanics of chain molecules. 1969, New York: Interscience Publishers.
26. O. Kratky and G. Porod, Röntgenuntersuchung gelöster fadenmoleküle, *Recl. Trav. Chim. Pays Bas* **68**, 1106–1122 (1949).
27. H.-X. Zhou, Polymer models of protein stability, folding, and interactions, *Biochemistry* **43**, 2141–2154 (2004).
28. D.A. Brant, W.G. Miller, and P.J. Flory, Conformational energy estimates for statistically coiling polypeptide chains, *J. Mol. Biol.* **23**, 47–65 (1967).
29. E.M.W.M. Van Dongen, T.H. Evers, L.M. Dekkers, E.W. Meijer, L.W.J. Klomp, and M. Merkx, Variation of linker length in ratiometric fluorescent sensor proteins allows rational tuning of Zn(II) affinity in the picomolar to femtomolar range, *J. Am. Chem. Soc.* **129**, 3494–3495 (2007).
30. H.-X. Zhou, The affinity-enhancing roles of flexible linkers in two-domain DNA-binding proteins, *Biochemistry* **40**, 15069–15073 (2001).
31. H.-X. Zhou, Quantitative account of the enhanced affinity of two linked scFvs specific for different epitopes on the same antigen, *J. Mol. Biol.* **329**, 1–8 (2003).
32. E.M.W.M. Van Dongen, L.M. Dekkers, K. Spijker, E.W. Meijer, L.W.J. Klomp, and M. Merkx, Ratiometric fluorescent sensor proteins with subnanomolar affinity for Zn(II) based on copper chaperone domains, *J. Am. Chem. Soc.* **128**, 10754–10762 (2006).
33. A.K. Wernimont, D.L. Huffman, A.L. Lamb, T.V. O′Halloran, and A.C. Rosenzweig, Structural basis for copper transfer by the metallochaperone for the Menkes/Wilson disease proteins, *Nat. Struct. Biol.* **7**, 766–771 (2000).
34. T.H. Evers, M.A.M. Appelhof, P.T.H.M. De Graaf-Heuvelmans, E.W. Meijer, and M. Merkx, Ratiometric detection of Zn(II) using chelating fluorescent protein chimeras, *J. Mol. Biol.* **374**, 411–425 (2007).
35. T. Dudev and C. Lim, Principles governing Mg, Ca, and Zn binding and selectivity in proteins, *Chem. Rev.* **103**, 773–787 (2003).
36. Y. Lu, S.M. Berry, and T.D. Pfister, Engineering novel metalloproteins: Design of metal-binding sites into native protein scaffolds, *Chem. Rev.* **101**, 3047–3080 (2001).
37. K.K. Jensen, L. Martini, and T.W. Schwartz, Enhanced fluorescence resonance energy transfer between spectral variants of green fluorescent protein through zinc-site engineering, *Biochemistry* **40**, 938–945 (2001).
38. D.A. Zacharias, J.D. Violin, A.C. Newton, and R.Y. Tsien, Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells, *Science* **296**, 913–916 (2002).
39. F. Yang, L.G. Moss, and G.N. Phillips Jr., The molecular structure of green fluorescent protein, *Nat. Biotechnol.* **14**, 1246–1251 (1996).
40. M.D. Allen, L.M. DiPilato, M. Rahdar, Y.R. Ren, C. Chong, J.O. Liu, and J. Zhang, Reading dynamic kinase activity in living cells for high-throughput screening, *ACS Chem. Biol.* **1**, 371–376 (2006).