

Chapter 1

Macromolecular Models by Single Molecule FRET

Axel T. Brunger, Pavel Strop, Marija Vrljic, Mark Bowen, Steven Chu,
and Keith R. Weninger

Abstract Single molecule fluorescence energy transfer (FRET) experiments enable investigations of macromolecular conformation and folding by the introduction of fluorescent dyes at specific sites in the macromolecule. Multiple such experiments can be performed with different labeling site combinations in order to map complex conformational changes or interactions between multiple molecules. Distances that are derived from such experiments can be used for determination of the fluorophore positions by triangulation. When combined with a known structure of the macromolecule(s) to which the fluorophores are attached, a three-dimensional model of the system can be determined by docking calculations. Here we discuss recent applications of single molecule FRET to obtain a model of the synaptotagmin-1:SNARE complex and to study the conformation of PSD-95.

A.T. Brunger (✉) • P. Strop • M. Vrljic
The Howard Hughes Medical Institute and Departments of Molecular and Cellular Physiology,
Neurology and Neurological Sciences, Structural Biology, and Photon Science, Stanford
University, James H. Clark Center, Room E300-C, Stanford, CA 94305, USA
e-mail: brunger@stanford.edu

M. Bowen
Department of Physiology and Biophysics, Stony Brook University Medical Center,
Stony Brook, NY 11794-8661, USA

S. Chu
Formerly Lawrence Berkeley National Laboratory and Physics and Cell Biology Departments,
University of California at Berkeley, Berkeley, CA 94720, USA

K.R. Weninger (✉)
Department of Physics, North Carolina State University, Raleigh, NC 27695, USA
e-mail: keith_weninger@ncsu.edu

1.1 Introduction

Detailed structural studies extending to the atomic level are effective approaches to reveal the molecular mechanisms underlying function of biological molecules. High-resolution methods such as X-ray diffraction crystallography and nuclear magnetic resonance have led the way in providing the highest spatial information, but a host of other methods, such as small angle X-ray scattering, cryo-electron microscopy, hydrodynamic assays (gel filtration chromatography, light scattering, and analytical ultra centrifugation), and spectroscopic measures of circular dichroism and fluorescence also can provide a wealth of knowledge about molecular structure. Among these approaches, fluorescence resonance energy transfer (FRET) studies have flourished in recent years as a result of the capability to detect the signal reporting intra and intermolecular distances from samples as small as single molecules [17, 57].

The benefits of studying single molecules have opened new avenues of investigation in biomolecular science. By recording properties and dynamics of single molecules one at a time, the effects of averaging that are inherent in ensemble studies are absent, allowing discovery of phenomena not otherwise observable. The single molecule approach is particularly effective at revealing heterogeneous behaviors across different individual molecules within a sample and also reporting dynamic trajectories of molecules without the need for synchronous behavior across a population. Single molecule FRET (smFRET) is well suited for structural studies because it provides a unique tool with the applicability to transient and dynamic molecular conformations and can reveal weak interactions that often are not resolvable when averaging over a larger sample.

Motivated by the dramatic successes of smFRET in the past decade, there has been a rapid development of instrumentation, sample preparation, labeling, and data analysis [14, 41, 50, 59]. By combining multiple smFRET experiments involving different labeling site combinations one obtains a network of FRET-derived distances between the labeling sites. If the distance network is augmented by structural information about the molecules to which the fluorophores are attached, powerful computational approaches can be used to obtain three-dimensional models of the entire system. As a particular example for multi-molecule docking we discuss the determination of the model of the synaptotagmin-1:SNARE complex derived from smFRET-derived distances [11]. We also discuss a recent smFRET study of the conformation of PSD-95 [42].

1.2 Förster Theory

FRET occurs between two fluorescent dyes when the emission spectrum of an excited donor fluorophore overlaps the absorption spectrum of a nearby acceptor fluorophore [23]. For applications with biomolecules, site specific chemistry is

generally used to link the fluorescent dyes at known locations on the molecule so that FRET between the fluorophores is interpretable at the distance between those points [62]. The FRET efficiency E_{FRET} depends strongly on the distance R between fluorophores:

$$E_{\text{FRET}} = \frac{1}{1 + (R/R_0)^6} \quad (1.1)$$

where the Förster radius R_0 is a donor/acceptor-pair specific constant (the distance at which the FRET efficiency is 50%). When the distance between the donor and acceptor fluorophores is $\frac{1}{2} R_0$, the FRET efficiency is nearly maximal and, so, any further decrease in distance is difficult to measure. Conversely, if the distance increases beyond $2 R_0$ then, the distance dependence is also very shallow. Thus, the most sensitive range for a typical FRET experiment is in the distance range $\frac{1}{2}$ to $2 R_0$.

The Förster radius R_0 depends on the spectroscopic properties of the FRET fluorophore pair and the surrounding medium of the fluorophores [32]. The FRET efficiency can be obtained by measuring either the donor and acceptor fluorescence intensities, or the donor lifetimes in the presence and absence of an acceptor [55]

$$E_{\text{FRET}} = \left[1 + \gamma \left(\frac{I_D}{I_A} \right) \right]^{-1} \quad (1.2)$$

where I_D and I_A are donor (D) and acceptor (A) fluorescence intensities. The factor γ combines the probabilities of the donor and acceptor fluorophores to relax to the ground state from the fluorescent excited state by emitting a photon and the likelihood of experimentally detecting emitted photons.

In summary, the Förster theory that relates measured donor and acceptor intensities (I_D and I_A) to the interfluorophore distance R is dependent on two factors involving fluorophore and instrument properties: R_0 (the Förster Radius) (Eq. 1.1) and γ (Eq. 1.2). R_0 is generally different for each specific pair of fluorescent dyes. The γ factor depends on a combination of fluorophore and instrument properties.

1.3 Corrections to Raw Single Molecule Fluorescence Intensity Data

The raw fluorescence intensity data (I_A and I_D in Eq. 1.2) must be corrected by background scattering, leakage of fluorescence intensity of donor and acceptor fluorophores into each other's detection channels, and by the γ factor. Leakage of one fluorophore's emission into the detection channel of the other fluorophore is typically characterized using measurements of samples prepared with only one of the fluorophores. The fraction of the emission for each fluorophore that appears

in the unintended channel is a fixed function of the specific configuration of the detection channel as long as the emission spectral density is constant during the experiment (see ref. [12] for a notable exception where emission spectral density changes). This fraction is subtracted from measured intensities when analyzing the FRET efficiency data. There are several approaches to background subtraction for experiments using immobilized molecules. They all are generally based around the fact that if the surface density of immobilized molecules is sufficiently dilute then background contributions to the emission intensity can be estimated from locations near observed fluorescent spots that are free of other fluorescent molecules.

1.4 Empirical Determination of γ

The γ factor accounts for differences between the donor and acceptor fluorophores in the probability that emitted photons will be detected (detector efficiency) and the probability of photon emission upon excitation (quantum yield). These properties can be determined experimentally by measuring the detected intensities of acceptor and donor fluorophores separately as a function of illumination power, and by measuring the relative quantum yields of the fluorophores attached to the particular biomolecule under study. Quantum yields are commonly measured by comparing ensemble fluorescence measurements of samples with known concentrations to the emissions of quantum yield standards (for example, rhodamine 101 in ethanol). For single molecule studies of freely diffusing molecules, alternative approaches to determine the γ factor that do not require independent determination of detection efficiencies and quantum yields are available. One such method exploits the linear relationship between the apparent FRET ratio and the stoichiometry by using an alternative laser excitation scheme and anti-correlated photobleaching events [35]. Another approach relies on measurement of the lifetime of the fluorescent dye excited states [55].

A systematic study of γ normalization methods using DNA found that determination of γ from the magnitude of anti-correlated photobleaching events was the most effective at achieving convergence of measured FRET values and high resolution structures of duplex DNA [41]. Variability in γ raises questions as to how normalization should be applied to recover FRET efficiency effectively. It is possible to measure γ once for an instrument and apply this as a “universal” normalization to all measurements using the same dyes and optical path. This method does not account for the observed sample-to-sample variation in γ . One could normalize an individual data set using the data set specific mean or “global” γ factor. This does not account for variation within a data set or the outliers with γ values significantly different from the mean. To account for these outliers, one would have to normalize each molecule with an “individual” γ factor. These approaches have all been previously reported [13, 27, 28, 56, 67].

A systematic comparison of these approaches [41] found that the global γ factor was sufficient to correct the mean FRET efficiency. A universal γ factor was less effective because it fails to account for actual variability in γ between samples. However, only normalization with individual γ factors for each molecule resulted a narrowing of width of smFRET distributions. Both systematic factors, including instrumental or photophysical effects as well as dynamic molecular motion, can contribute to broadening the widths of FRET efficiency distribution peaks in smFRET experiments [8, 43, 44]. This study showed that γ variability also contributes to broadening FRET histogram widths. Molecules with FRET values near the edges of the FRET distribution also commonly had outlying γ values and per molecule γ normalization brings these values closer to the mean. The γ outliers may be due to differences in detection efficiency introduced during image recording or processing. In agreement with this notion, it has been noted that one can change the value of γ by misaligning the detectors when measuring diffusing single molecules [35, 44]. Empirical measurements of the terms composing γ cannot account for aberration this of kind. As a consequence of the fact that γ outliers are not representative of the population, applying a γ cutoff as a means of selecting accepted molecules could further affect peak width and shape.

1.5 Empirical Determination of R_0

In principle, R_0 can be calculated from the spectral overlap of the fluorophore pairs, the donor quantum yield, and the orientation factor [32]. Some of these parameters can be obtained experimentally, but the orientation factor requires some knowledge of the fluorophores's dynamics with respect to the attached molecule. For the synaptotagmin-1:SNARE complex [11], an entirely empirical approach was used to calibrate R_0 inspired by previous work [1, 2]. For a fluorophore pair (Alexa555:Alexa647) attached to one of the two rigid C2 domains of synaptotagmin-1 FRET efficiency was measured and compared with the calculated value using the fluorophore separation from a crystal structure of Syt1 [24]. The resulting FRET efficiency distributions produced a single Gaussian peak with a width consistent with shot noise [11]. The FRET efficiency and fluorophore separation yielded an empirical value of $R_0 = 5.55$ nm (using $\gamma = 1$, Eqs. 1.1 and 1.2), compared to the theoretical R_0 for the Alexa555:Alexa647 fluorophore pair of 5.1 nm [30]. R_0 is expected to deviate from the theoretical value due to changes in the fluorophore microenvironment when conjugated to the molecule of interest. In the work with synaptotagmin-1, the spread of the R_0 values derived from three label-site combinations (using the same fluorescent dyes but different amino acid attachment sites) was 0.23 nm, which is smaller than the error bounds used in the docking calculation [11].

1.6 Simulation of Fluorophore Center Positions

As a pre-requisite for docking and fitting calculations, the average fluorophore center positions have to be computed relative to the position of the molecule or domain that the fluorophore is attached to. Depending on the particular fluorophore and linker, the fluorophore center position is separated from the coordinates of the covalently attached residue (often generated by site-specific mutagenesis) by ~ 1 nm. In earlier work, the fluorophore center position was simply taken at the protein residue site or displaced by a certain amount away from the molecule's center [49]. To obtain a more precise estimate, a molecular dynamics simulation can be used to obtain the average position of the fluorophore center using an atomic model of the fluorophore linked to a protein at the residue position used for labeling [11, 66] (Fig. 1.1).

For these particular simulations, the protein atoms of the molecule to which the fluorophore is attached were kept fixed while the linker and the fluorophore

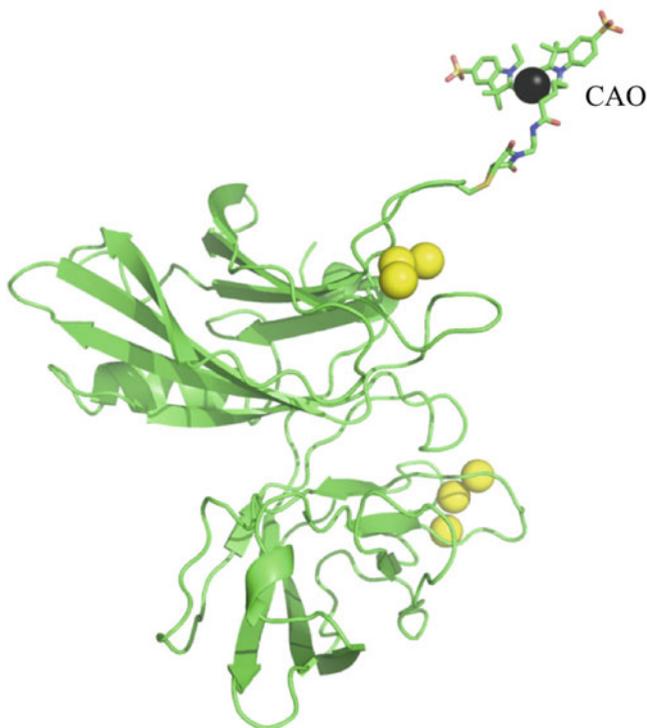


Fig. 1.1 Cy3 fluorescent dye (*sticks*) attached to the C2B domain of synaptotagmin-3 (*green cartoon*) [66]. *Yellow spheres* indicate the Ca^{2+} positions of the crystal structure of synaptotagmin-3. *Black sphere* indicates the position of atom CAO of the dye that is used to define the center of the fluorescent dye (Color figure online)

atoms were allowed to move. Chemical structures are available for some of the commonly used fluorescent fluorophores, such as Cy3, Cy5, and Alexa 647 [66]. A large number of simulations were performed starting from different initial velocities in order to obtain good conformational sampling.

1.7 Docking Calculations

Determination of a three-dimensional model from smFRET derived distances is reminiscent of rigid body docking approaches using nuclear magnetic resonance (NMR)-derived distances between protons [19, 61]. We discuss here the method that was used to determine a model of the synaptotagmin-1:SNARE complex [11]. In the case of smFRET data, the distances refer to distances between fluorophore centers. The fluorophore center position is treated as a “pseudoatom” that is rigidly associated with the molecule to which the fluorophore is covalently attached; the position of the pseudoatom is taken as the average position of the fluorophore center relative to the molecule, as obtained from a molecular dynamics simulation (see above) [11]. The term pseudoatom is used since these points have no chemical energy terms associated with them during the docking calculation, but are rather restricting the possible conformations of the rigid molecule or domain that the pseudoatom is attached to. Extensive torsion angle/rigid body molecular dynamics simulations [52] were performed using a simulated annealing slow-cooling protocol. The total energy function consists of a repulsive term for the nonbonded interactions (i.e., excluding electrostatic and attractive van der Waals terms) [20] and the distance restraints term. This type of energy function is widely used for three-dimensional structure determination based NMR data [6]. A harmonic square-well potential was used to restrain the distances between fluorophore center pseudoatom positions [4]. The smFRET efficiencies were converted to distances (Eqs. 1.1 and 1.2). FRET efficiency becomes less sensitive to fluorophore separation at distances much less and much greater than R_o (see Eq. 1.1). Therefore, variable bounds were used for the square well potential depending how close the expected distance is to R_o [11].

Many trials (typically $\sim 1,000$) with different randomly assigned orientations of domains and molecules, different relative conformations of flexible domains, and initial velocities were performed for each set of calculations (Schema 1.1). Each resulting model was then characterized by the root mean square (rms) deviation between the distances predicted by the model and the distance ranges used as the square well potentials derived from the FRET efficiency measurements. The solutions of the docking simulations were sorted by rms deviation satisfaction in increasing order. The solutions were clustered according to the rms deviation using an algorithm implemented in the program HADDOCK [15]. For each cluster, the structure with the best distance satisfaction was used for subsequent analysis. The derived models could also be further refined using local perturbation and refinement in docking programs to optimize local interactions. This step would allow for side chain refinement and inclusion of electrostatics and van der Waals energy terms.

- **SNARE complex and C2 domains kept rigid**
- **C2A-C2B linker flexible**
- **distance bounds +/- 2.5 Å, +/- 5 Å**
- **1000 simulated annealing docking calculations starting from random placements of the domains**
- **Cluster analysis of the top solutions**

Schema 1.1 Key points of the docking calculations for the synaptotagmin-1:SNARE complex. Further details can be found in the single molecule FRET tutorial of CNS, version 1.3, section on “Docking calculations with single-molecule FRET derived distances”, <http://cns-online.org> [60]

1.8 smFRET Derived Model of the Synaptotagmin-1:SNARE Complex

Ca²⁺-induced membrane fusion of synaptic vesicles at synapses is the central phenomenon that results in triggered inter-neuron signaling. The membrane protein synaptotagmin-1 is the Ca²⁺ sensor for synchronous neurotransmitter release [22, 51]. Highly coordinated interactions among synaptotagmin, SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptor) proteins and other neuronal factors are required to create robust and adaptive neural circuits [5, 54]. Synaptotagmin-1 is primarily located on synaptic vesicles and contains two independent C2-type Ca²⁺-sensing domains [9] (termed C2A and C2B, respectively) that are connected by a linker (the fragment containing both domains is designated C2AB). Synaptotagmin 1 interacts with both anionic membranes and SNARE complexes, and both interactions are physiologically relevant [22, 45].

A general model has emerged where inhibitory and activating interactions among synaptotagmin, complexin, and the SNARE complex (that juxtaposes synaptic vesicles and the plasma membrane) yield a stall in membrane fusion that is released by Ca²⁺ influx following an action potential [25, 40]. Understanding the molecular mechanism underlying the release of the stall requires knowledge of the structures and dynamics of the complexes formed by these proteins. However, the structure of the complex between synaptotagmin and SNAREs has remained elusive. Extensive smFRET measurements were performed between a set of 34 fluorophore pairs located at various amino acid positions in the C2AB fragment of synaptotagmin-1 and the cis (post-fusion) state of the neuronal SNARE complex [11] (Fig. 1.2).

The SNARE complex, and the two C2 domains of synaptotagmin-1 were treated as independent rigid bodies (residues 140:262 and 273:418, respectively), while the torsion angles of the linker connecting the two domains (residues 263–272) were simulated in torsion angle space, i.e., with bond lengths and bond angles fixed. The coordinates for the SNARE complex were obtained from the crystal structure of the neuronal SNARE complex (PDB ID 1SFC) [63] and those of the C2 domains of synaptotagmin-1 from the Ca²⁺-free crystal structure of synaptotagmin-1 (PDB ID 2R83) [24].

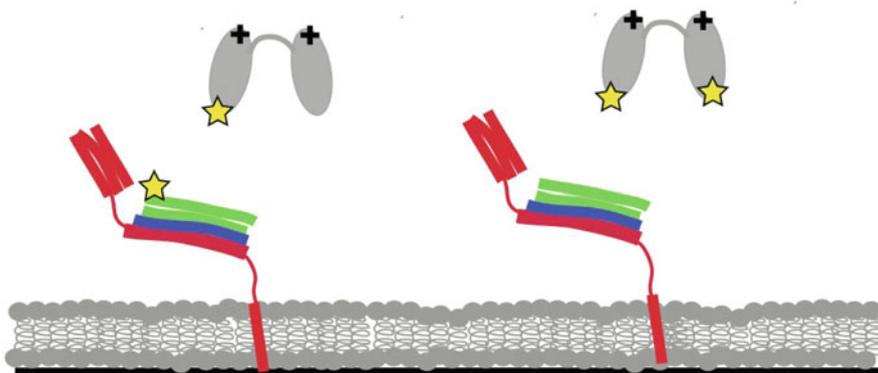


Fig. 1.2 SNARE complex is anchored to a supported bilayer through the transmembrane domain of syntaxin (*red*). The cytoplasmic domain of synaptobrevin (*blue*) and soluble SNAP-25 (*green*) were added and the complex extensively purified prior to membrane reconstitution to obtain mostly parallel complex [68]. After reconstitution into liposomes, a deposited bilayer was formed by liposome condensation. Soluble C2A-C2B fragment of synaptotagmin-1 was added. The “+” signs indicate the calcium binding regions. FRET label pairs were placed in several positions in the C2A, C2B domains and the SNARE complex [11] (*yellow stars*)

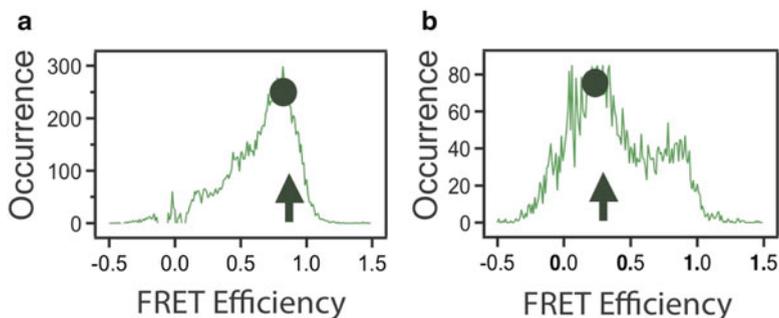


Fig. 1.3 smFRET efficiency histograms of complex between synaptotagmin-1 and SNARE complex with labels placed at position 350 of synaptotagmin-1 (Cy3) and position 61 of synaptobrevin (Cy5) [11]. *Black circles* indicate the FRET efficiency value used to derive the distance restraint for the docking calculations. *Arrows* indicate the FRET efficiency value calculated from best model of the synaptotagmin-1:SNARE complex. (a) Example of a unimodal FRET efficiency distribution. (b) Example of a bimodal FRET efficiency distribution

The FRET measurements used for the synaptotagmin-1:SNARE complex study [11] were acquired with a 0.1 s integration period (referred to as “time bin”). The histograms of FRET efficiency for the different sets of label attachment sites were characterized by either one or two well defined Gaussian peaks whose width was near the shot noise expected width or up to a factor of two wider (Fig. 1.3). The center of a Gaussian function fitted to the histograms was used to represent the

FRET efficiency value for a defined state of the complex. The fraction of area under that Gaussian compared to the area under the complete histogram indicated the fractional population within that FRET efficiency state. In this system, 26 out of 34 label site combinations had dominant states with at least 70% fractional population. These states were used to generate distance restraints for molecular modeling. The remaining label site combinations indicated a mixture of states with none reaching 70% occupancy. For these cases the state was selected using an iterative modeling approach to make use of these mixed-state FRET efficiency distributions. An intermediate model was generated using only the restraints derived from FRET states with >70% in one Gaussian and the mixed states were compared to this preliminary model. The FRET value in the mixed states that was more consistent with the intermediate model was selected and then the entire set of restraints was used for a final round of modeling.

The presence of distinct FRET states for some label pairs along with the fact that the widths of some FRET efficiency peaks were wider than expected from shot noise alone indicate some degree of heterogeneity within the synaptotagmin-1:SNARE complex. For most fluorophore attachment site combinations a dominant configuration could be identified, highlighting one of the advantages of the single molecule approach. On the other hand, this example serves as a warning that in future applications possible heterogeneity within a configurational ensemble may prevent convergence of the modeling calculations to a single solution. Advances in single-molecule imaging technology that allow improved temporal resolution in single molecule FRET studies will allow multiple interconverting states to be better resolved and will extend the applicability of the modeling approaches discussed here.

For 26 of the 34 measured FRET pairs, the major fitted peaks capture 70% or more of the total non-zero smFRET distribution; ten label pairs have distributions with a major peak comprising 90% of the total distribution (Fig. 1.3a), seven label pairs between 90% and 80%, nine between 80% and 70%. The assignments of the dominant FRET states at the 70% level were robust against run-to-run variation. Repeating measurements of single label pair combinations generated the same central FRET efficiency values within experimental error for all label pairs and for most label pairs the dominant population was consistently above the 70% value. For all pairs the dominant population was observed at levels above 70% in at least 66% of the repeated experiments. Many were confirmed greater than 70% dominant in all repeats. Some of the 34 distance restraints involved multiple distinct FRET populations (Fig. 1.3b). Therefore, the docking calculations proceeded in two steps.

For final docking calculations all of the repeated experiments for each label pair were pooled into a single histogram to address the most probable configuration observed across multiple repeated experiments (at least three repeats for every label site pair). For label pairs with FRET efficiency distributions with a dominant peak of >70% of the total area, the FRET measurements were converted to distances and used as restraints for a first round of docking calculations (26 out of 34 pairs were included at this first step). The FRET histograms for the remaining eight label combinations required sums of two Gaussians to fit where neither comprised more

than 70% of the total population. These measured distances were compared to the best model from the first simulation using the first 26 distances and the measured distance that was closer to the model distance was selected. Then the simulation was repeated using all 34 restraints. The resulting models did not change significantly upon inclusion of the eight additional restraints.

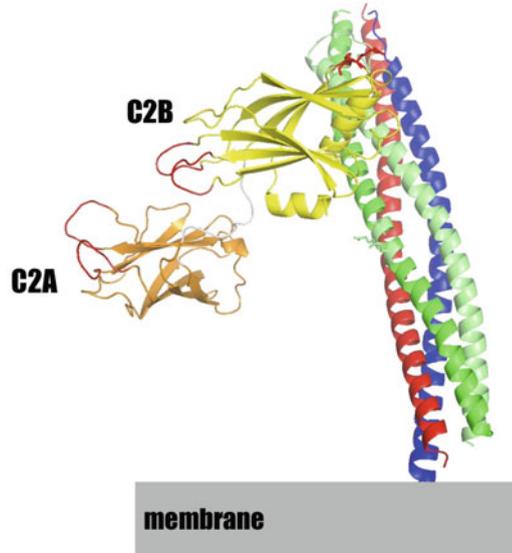
A docking simulation using the minor populations from each of the FRET efficiency distributions was also performed (if only one FRET population was present, it was used for both the major and minor population simulation) but convergence was much poorer than for the calculations with the major population. Thus, the conformations arising from docking using the major FRET population restraints are much more likely to occur than those derived using the minor FRET populations.

Uniformly increasing or decreasing all major FRET derived distance restraints by 1 nm led to non-physical results where the proteins were far away from each other or overlapped in space respectively. If the intra-C2 domain restraints were released then the models converged to the same C2B docking state, but the location of the C2A domain became variable.

In order to cross-validate the model, the docking calculations were repeated omitting all of the restraints that involved one particular fluorophore position (residue 383) on synaptotagmin-1. This site showed some of the highest FRET efficiency values when combined with acceptor fluorophores on the SNARE complex. Remarkably, the resulting top model was very similar to the docking calculation using all distances. Because the number of restraints exceeds the number of degrees of freedom for the docking calculations, it is reasonable that omitting a few restraints does not lead to drastically different solutions. The similarity of the top models with and without distances involving the 383 site illustrates the robustness of the top solution with respect to such cross-validation.

The best model from the docking calculations is shown in Fig. 1.4. It showed up consistently in all docking calculations as the top solution. One helix of synaptotagmin-1 (residues 385–395) is directly positioned at the interface with the SNARE complex (Fig. 1.4). The central region of the SNARE complex that mediates synaptotagmin-1 binding (as predicted by the smFRET-derived model) is essential for function. Mutations of glutamates near this area of SNAP-25 (Glu51, Glu52, Glu55) to lysines eliminated *in vitro* binding of synaptotagmin to the SNARE complex and greatly reduced Ca^{2+} stimulated release in PC12 cells [53]. These same SNAP-25 mutations as well as additional SNAP-25 mutations directly adjacent to this region (Leu50 and Ile171) are critical in the context of docking vesicles in adrenal chromaffin cells [16]. Additionally, in the smFRET-derived model the conserved arginine residues at the bottom of C2B [69] are close to the interface with the SNARE complex but also are sufficiently exposed to allow potential interactions with membranes. Mutation of these residues results in decreased synchronous neurotransmitter release in hippocampal glutamatergic neurons [69]. These independent functional assessments of the synaptotagmin-1 - SNARE interactions lend further credence to the smFRET-derived model.

Fig. 1.4 Model of the synaptotagmin-1:SNARE complex obtained from docking calculations using 34 smFRET derived distances [11]. Shown are cartoon representations of synaptobrevin (*blue*), syntaxin (*red*), SNAP-25 (*green*), C2A domain of synaptotagmin-1 (*orange*), C2B domain of synaptotagmin-1 (*yellow*) (Color figure online)



The synaptotagmin-1:SNARE complex is not a rigid structure since occasional transitions between different FRET efficiency states were observed, and some of the smFRET efficiency distributions have a multimodal appearance (Fig. 1.3b). This intrinsic flexibility of the synaptotagmin-1:SNARE complex may allow the complex to adapt to the particular geometry of the interacting membranes in the pre-fusion state.

1.9 Application to PSD-95

Multidomain scaffold proteins are critical organizers of signal transduction and junctional communication [46]. Often these proteins contain a series of archetypal protein interaction domains connected by flexible linkers into a larger protein. It is convenient to try and understand scaffold proteins by studying each individual domain as an isolated unit. However, dimerization and incorporation into larger proteins can alter the structure as well as the binding specificity relative to the isolated domain [33, 47, 65, 70]. Thus, studies of isolated domains can be only partially successful at explaining function in their biological context.

PDZ (PSD-95/Dlg/ZO-1) domains are the most common protein-interaction domains in the human genome [3, 29]. PDZ domains typically form part of a larger multidomain protein and often appear in tandem with instances of up to 13 PDZ domains in a single protein [31]. PSD-95 was one of the first PDZ-containing proteins identified [10]. PSD-95 contains three tandem PDZ domains followed by an SH3 domain and an enzymatically inactive guanylate kinase-like domain. The first

two tandem PDZ domains are connected by a five residue linker. Such tandem domain arrays are often conceptualized as folded beads on a disordered string. However, recent structural studies have found that tandem PDZ domains can adopt a fixed interdomain orientation [21, 26, 34, 37–39].

Most studies of PDZ tandems to date have examined portions of much larger proteins so the degree to which the structure of the tandem PDZ supramodules depends on the remainder of the protein is therefore unknown. The crystallization of flexible, multidomain scaffold proteins is challenging. The PDZ tandem from PSD-95 “yielded poorly diffracting crystals” so a self-interacting PDZ ligand sequence appended to the C-terminus that induced non-native protein interactions which lead to diffracting crystals [58]. The resulting crystals contained two different conformations with slightly different orientations of the PDZ domains. The large molecular weight of scaffold proteins often necessitates the use of truncated constructs. The limited interdomain contact provides few unambiguous NMR restraints. In contrast, FRET experiments allow for the accurate measurement of long intramolecular distances (3–8 nm) so it is particularly useful for studying the orientation of domains that are not in intimate contact. In addition, fluorescence measurements are not subject to molecular weight restrictions, which makes measurements in full-length scaffold proteins possible.

Single molecule FRET observations and other data were used to characterize the interdomain orientation of the first two PDZ domains in full-length PSD-95 [42]. Using the high-resolution structures of the first two PDZ domains [37, 48, 64] four surface-exposed sites in PDZ1 and five in PDZ2 were selected for fluorescent labeling and mutated to cysteine (Fig. 1.5a). Accepted labeling sites showed maximal solvent exposure and minimal near neighbor interactions, which should minimize any positional dependence of the photophysical properties. This expectation was confirmed by measuring the anisotropy and quantum yield of the attached dyes, which showed minimal environmental impacts on fluorescence emission [42].

The donor anisotropy was similar for all samples and was minimally higher than that measured for unconjugated Alexa 555. Using the 9 single labeling sites, 11 pairwise combinations were generated in full-length PSD-95 to create 11 double-cysteine mutants for smFRET measurements [42] (Fig. 1.5a). This ensemble of labeling combinations was chosen to sample any possible orientation between these two domains.

All 11 labeling combinations produced smFRET histograms containing a single peak of narrow width that was well described by a single Gaussian curve (Fig. 1.5b). The narrow widths are similar to those observed in duplex DNA and could either arise from a fixed interdomain orientation or time averaging of rapid motions. Paradoxically, these two extremes are indistinguishable in this assay. The different labeling site combinations showed a wide dispersion in their mean smFRET efficiencies (from 0.37 to 0.94), which would not be expected if the two domains were undergoing isotropic motion. To see if domain orientation is dependent on interdomain contacts within the full-length protein, smFRET efficiencies were measured for all 11 labeling site combinations in truncated PSD-95 constructs

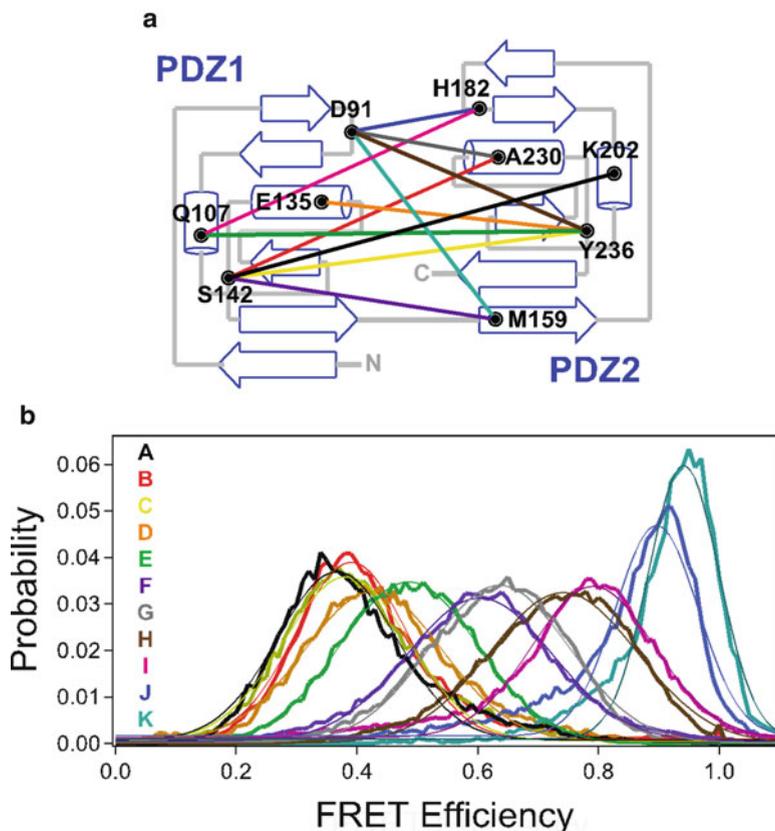


Fig. 1.5 Single molecule FRET measurements between PDZ1 and PDZ2 in full length PSD-95. (a) Topology diagram showing the position of the cysteine labeling sites in each domain and the 11 combinations of labeling sites used for fluorescence studies [42]. (b) smFRET histograms for all 11 FRET pairs made in full-length PSD-95. Letter codes (indicated within the panel) were assigned to each mutant in order of increasing FRET efficiency. The lettering of the FRET distributions corresponds to the *lettering* denoting labeling combinations in panel A. *Thin lines* indicate the fit to a single Gaussian function (Color figure online)

lacking the 3 additional domains. None of the labeling site combinations showed significant changes in mean FRET efficiency or the width of the FRET distribution relative to full length PSD-95. Thus, domain orientation in the PDZ tandem is not affected by intramolecular interactions with other domains present in full length PSD-95.

Molecular dynamics simulations were performed to calculate the mean dye position relative to the protein backbone for each of the nine labeling sites similar to the studies of the synaptotagmins [11, 66]. The mean FRET efficiencies were converted to distances using a calibrated Förster's radius. The γ factor was determined from photobleaching and applied as individual γ correction at the

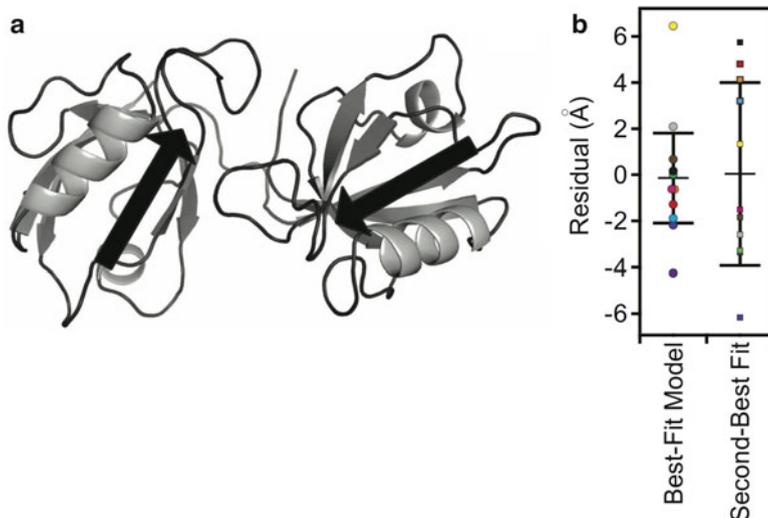


Fig. 1.6 (a) Cartoon representation of the best fit model to the 11 FRET distance restraints [42]. The model shows a relatively compact orientation without interdomain contacts. The position and direction of canonical PDZ ligand binding sites are represented as *arrows*. (b) Residuals (r.m.s. distance differences) for the first and second best-fit models based on the smFRET distance restraints. Residuals are colored for each labeling site according to Fig. 1.5

single molecule level [41]. Extensive torsion angle/rigid body molecular dynamics simulations were performed with atoms in PDZ1 and PDZ2 fixed while atoms in the interdomain linker were unrestrained. FRET-derived distance restraints were applied to guide domain docking [42].

From 500 trials with randomly assigned starting orientations, cluster analysis identified a single model was found that best satisfied the FRET distance restraints (Fig. 1.6a) [42]. The best-fit model for the PDZ tandem showed a relatively closed configuration but with the distance between domains greater than a single bond. The orientation in the model shows an antiparallel alignment of the two ligand binding pockets (Fig. 1.6a). The root mean square error (E_{RMS}) for the best-fit model was 2.54 Å, while in the second model E_{RMS} was 5.21 Å (Fig. 1.6b). Outlying data points were within loops that were held rigid during the simulation but are expected to fluctuate in the native protein. The second-best fit model was more compact, with the centers of mass for the two domains positioned 7.7 Å closer. The second model had the relative orientation flipped with PDZ2 positioned on the opposite face of PDZ1. To insure that FRET data were not misfit or overfit, docking simulations were repeated with each of the FRET restraints omitted from the refinement. The final results from all 11 simulations were highly similar structures, indicating that FRET restraints oversampled the domain orientation in the PDZ tandem. This represents the first structural model for the PDZ tandem of PSD-95 in the context of the full-length protein.

1.10 Conclusions

Many important biological structures have resisted analysis by high-resolution structural methods. The causes of these difficulties are often multiple and wide-ranging. Some proteins are not stable at the high concentrations required; some molecular systems do not reside in a single stable configuration; other interesting complexes are present only rarely within an ensemble. Single molecule FRET has proven to be an effective tool that can provide a window into these difficult systems. Although the distance resolution provided in any single FRET pair measurement does not approach the atomic dimensions, detailed structural information can be gleaned by oversampling in distance space with combinations of multiple FRET measurements across different locations. Easier access to such oversampling will be provided by the recent development of four-color single molecule FRET measurements of individual complexes [18, 36]. If four distinct fluorescent dyes can be attached at specific locations to the domains of larger multi-molecular complexes, then the FRET coupling between the dyes can simultaneously report 6 inter-dye distances [36]. Such advances will allow more confident and efficient application of the FRET-restrained docking approaches we have described.

These studies of the synaptotagmin-1:SNARE complex and the conformation of PDZ domains within full-length PSD-95 we have described are examples of particularly effective combinations of single molecule FRET with available high-resolution structures of individual domains. The domains are either connected by flexible linkers or are bound to other stable domains in the final complex. Single molecule FRET measurements restrained docking calculations of the known high-resolution structures of the individual domains and allowed three-dimensional models ranked by distance satisfaction to be obtained. This hybrid strategy will likely be useful to determine the configuration of other biological complexes.

Acknowledgments We thank the National Institutes of Health for support (to A.T.B., RO1-MH63105), and a Career Award at the Scientific Interface from the Burroughs Wellcome Fund (to K.W.). Part of this material has been published in modified form in the *Journal of Structural Biology* [7].

References

1. Amir D, Haas E (1987) Estimation of intramolecular distance distributions in bovine pancreatic trypsin inhibitor by site-specific labeling and nonradiative excitation energy-transfer measurements. *Biochemistry* 26:2162–2175
2. Amir D, Haas E (1988) Reduced bovine pancreatic trypsin inhibitor has a compact structure. *Biochemistry* 27:8889–8893
3. Bhattacharyya RP, Remenyi A, Yeh BJ, Lim WA (2006) Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu Rev Biochem* 75:655–680
4. Brunger AT (1992) X-PLOR, version 3.1. A system X-ray crystallography and NMR. Yale University Press, New Haven

5. Brunger AT (2005) Structure and function of SNARE and SNARE-interacting proteins. *Q Rev Biophys* 38:1–47
6. Brunger AT, Nilges M (1993) Computational challenges for macromolecular structure determination by X-ray crystallography and solution NMR-spectroscopy. *Q Rev Biophys* 26:49–125
7. Brunger AT, Strop P, Vrljic M, Chu S, Weninger KR (2011) Three-dimensional molecular modeling with single molecule FRET. *J Struct Biol* 173:497–505
8. Cherny DI, Eperon IC, Bagshaw CR (2009) Probing complexes with single fluorophores: factors contributing to dispersion of FRET in DNA/RNA duplexes. *Eur Biophys J* 38:395–405
9. Cho W, Stahelin RV (2006) Membrane binding and subcellular targeting of C2 domains. *Biochim Biophys Acta* 1761:838–849
10. Cho KO, Hunt CA, Kennedy MB (1992) The rat brain postsynaptic density fraction contains a homolog of the drosophila discs-large tumor suppressor protein. *Neuron* 9:929–942
11. Choi UB, Strop P, Vrljic M, Chu S, Brunger AT, Weninger KR (2010) Single-molecule FRET-derived model of the synaptotagmin 1-SNARE fusion complex. *Nat Struct Mol Biol* 17:318–324
12. Chung HS, Louis JM, Eaton WA (2009) Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories. *Proc Natl Acad Sci U S A* 106:11837–11844
13. Dahan M, Deniz AA, Ha T, Chemla DS, Schultz PG, Weiss S (1999) Ratiometric measurement and identification of single diffusing molecules. *Chem Phys* 247:85–106
14. Dave R, Terry DS, Munro JB, Blanchard SC (2009) Mitigating unwanted photophysical processes for improved single-molecule fluorescence imaging. *Biophys J* 96:2371–2381
15. de Vries SJ, van Dijk AD, Krzeminski M, van Dijk M, Thureau A, Hsu V, Wassenaar T, Bonvin AM (2007) HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins* 69:726–733
16. de Wit H, Walter AM, Milosevic I, Gulyas-Kovacs A, Riedel D, Sorensen JB, Verhage M (2009) Synaptotagmin-1 docks secretory vesicles to syntaxin-1/SNAP-25 acceptor complexes. *Cell* 138:935–946
17. Deniz AA, Mukhopadhyay S, Lemke EA (2008) Single-molecule biophysics: at the interface of biology, physics and chemistry. *J R Soc Interface* 5:15–45
18. DeRocco V, Anderson T, Piehler J, Erie DA, Weninger K (2010) Four-color single-molecule fluorescence with noncovalent dye labeling to monitor dynamic multimolecular complexes. *Biotechniques* 49:807–816
19. Dominguez C, Boelens R, Bonvin AM (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* 125:1731–1737
20. Engh RA, Huber R (1991) Accurate bond and angle parameters for X-Ray protein- structure refinement. *Acta Crystallogr A* 47:392–400
21. Feng W, Shi Y, Li M, Zhang M (2003) Tandem PDZ repeats in glutamate receptor-interacting proteins have a novel mode of PDZ domain-mediated target binding. *Nat Struct Mol Biol* 10: 972–978
22. Fernandez-Chacon R, Konigstorfer A, Gerber SH, Garcia J, Matos MF, Stevens CF, Brose N, Rizo J, Rosenmund C, Sädhof TC (2001) Synaptotagmin I functions as a calcium regulator of release probability. *Nature* 410:41–49
23. Förster T (1948) Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann Phys* 2:55–57
24. Fuson KL, Montes M, Robert JJ, Sutton RB (2007) Structure of human synaptotagmin I C2AB in the absence of Ca²⁺ reveals a novel domain association. *Biochemistry* 46:13041–13048
25. Giraudo CG, Garcia-Diaz A, Eng WS, Chen Y, Hendrickson WA, Melia TJ, Rothman JE (2009) Alternative zippering as an on-off switch for SNARE-mediated fusion. *Science* 323:512–516
26. Goult BT, Rapley JD, Dart C, Kitmitto A, Grossmann JG, Leyland ML, Lian LY (2007) Small-angle X-ray scattering and NMR studies of the conformation of the PDZ region of SAP97 and its interactions with Kir2.1. *Biochemistry* 46:14117–14128
27. Ha T, Ting AY, Liang J, Caldwell WB, Deniz AA, Chemla DS, Schultz PG, Weiss S (1999) Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc Natl Acad Sci USA* 96:893–898

28. Ha T, Ting AY, Liang J, Deniz AA, Chemla DS, Schultz PG, Weiss S (1999) Temporal fluctuations of fluorescence resonance energy transfer between two dyes conjugated to a single protein. *Chem Phys* 247:107–118
29. Harris BZ, Lim WA (2001) Mechanism and role of PDZ domains in signaling complex assembly. *J Cell Sci* 114:3219–3231
30. Haugland RP (2005) *The handbook: A guide to fluorescent probes and labeling technologies*. Molecular Probes, Carlsbad
31. Hung AY, Sheng M (2002) PDZ domains: structural modules for protein complex assembly. *J Biol Chem* 277:5699–5702
32. Iqbal A, Arslan S, Okumus B, Wilson TJ, Giraud G, Norman DG, Ha T, Lilley DM (2008) Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids. *Proc Natl Acad Sci U S A* 105:11176–11181
33. Irie M, Hata Y, Takeuchi M, Ichtchenko K, Toyoda A, Hirao K, Takai Y, Rosahl TW, Sudhof TC (1997) Binding of neuroligins to PSD-95. *Science* 277:1511–1515
34. Kang BS, Cooper DR, Jelen F, Devedjiev Y, Derewenda U, Dauter Z, Otlewski J, Derewenda ZS (2003) PDZ tandem of human syntenin: crystal structure and functional properties. *Structure* 11:459–468
35. Lee NK, Kapanidis AN, Wang Y, Michalet X, Mukhopadhyay J, Ebricht RH, Weiss S (2005) Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys J* 88:2939–2953
36. Lee J, Lee S, Ragunathan K, Joo C, Ha T, Hohng S (2010) Single-molecule four-color FRET. *Angew Chem Int Ed Engl* 49:9922–9925
37. Long JF, Tochio H, Wang P, Fan JS, Sala C, Niethammer M, Sheng M, Zhang M (2003) Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95. *J Mol Biol* 327:203–214
38. Long JF, Feng W, Wang R, Chan LN, Ip FC, Xia J, Ip NY, Zhang M (2005) Autoinhibition of X11/Mint scaffold proteins revealed by the closed conformation of the PDZ tandem. *Nat Struct Mol Biol* 12:722–728
39. Long J, Wei Z, Feng W, Yu C, Zhao YX, Zhang M (2008) Supramodular nature of GRIP1 revealed by the structure of its PDZ12 tandem in complex with the carboxyl tail of Fras1. *J Mol Biol* 375:1457–1468
40. Maximov A, Tang J, Yang X, Pang ZP, Sudhof TC (2009) Complexin controls the force transfer from SNARE complexes to membranes in fusion. *Science* 323:516–521
41. McCann J, Choi UB, Zheng L, Weninger K, Bowen ME (2010) Optimizing methods to recover absolute FRET efficiency from immobilized single molecules. *Biophys J* 99:961–970
42. McCann J, Zheng L, Chiantia S, Bowen ME (2011) Domain orientation in the tandem PDZ supramodule from PSD-95 is maintained in the full-length protein. *Structure* 19:810–820
43. Merchant KA, Best RB, Louis JM, Gopich IV, Eaton WA (2007) Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations. *Proc Natl Acad Sci* 104:1528–1533
44. Nir E, Michalet X, Hamadani KM, Laurence TA, Neuhauser D, Kovchegov Y, Weiss S (2006) Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. *J Phys Chem B* 110:22103–22124
45. Pang ZP, Shin OH, Meyer AC, Rosenmund C, Sudhof TC (2006) A gain-of-function mutation in synaptotagmin-1 reveals a critical role of Ca²⁺-dependent soluble N-ethylmaleimide-sensitive factor attachment protein receptor complex binding in synaptic exocytosis. *J Neurosci* 26:12556–12565
46. Pawson T, Scott JD (1997) Signaling through scaffold, anchoring, and adaptor proteins. *Science* 278:2075–2080
47. Peterson FC, Penkert RR, Volkman BF, Prehoda KE (2004) Cdc42 regulates the Par-6 PDZ domain through an allosteric CRIB-PDZ transition. *Mol Cell* 13:665–676
48. Piserchio A, Pellegrini M, Mehta S, Blackman SM, Garcia EP, Marshall J, Mierke DF (2002) The PDZ1 domain of SAP90. Characterization of structure and binding. *J Biol Chem* 277:6967–6973

49. Rasnik I, Myong S, Cheng W, Lohman TM, Ha T (2004) DNA-binding orientation and domain conformation of the *E. coli* rep helicase monomer bound to a partial duplex junction: single-molecule studies of fluorescently labeled enzymes. *J Mol Biol* 336:395–408
50. Rasnik I, McKinney SA, Ha T (2006) Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods* 3:891–893
51. Rhee JS, Li LY, Shin OH, Rah JC, Rizo J, Sudhof TC, Rosenmund C (2005) Augmenting neurotransmitter release by enhancing the apparent Ca^{2+} affinity of synaptotagmin 1. *Proc Natl Acad Sci U S A* 102:18664–18669
52. Rice LM, Brunger AT (1994) Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins* 19:277–290
53. Rickman C, Jimenez JL, Graham ME, Archer DA, Soloviev M, Burgoyne RD, Davletov B (2006) Conserved prefusion protein assembly in regulated exocytosis. *Mol Biol Cell* 17:283–294
54. Rizo J, Rosenmund C (2008) Synaptic vesicle fusion. *Nat Struct Mol Biol* 15:665–674
55. Rothwell PJ, Berger S, Kensh O, Felekyan S, Antonik M, Wöhrl BM, Restle T, Goody RS, Seidel CAM (2003) Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase:primer/template complexes. *Proc Natl Acad Sci U S A* 100:1655–1660
56. Roy R, Kozlov AG, Lohman TM, Ha T (2007) Dynamic structural rearrangements between DNA binding modes of *E. coli* SSB protein. *J Mol Biol* 369:1244–1257
57. Roy R, Hohng S, Ha T (2008) A practical guide to single-molecule FRET. *Nat Methods* 5:507–516
58. Sainlos M, Tigaret C, Poujol C, Olivier NB, Bard L, Breillat C, Thiolon K, Choquet D, Imperiali B (2010) Biomimetic divalent ligands for the acute disruption of synaptic AMPAR stabilization. *Nat Chem Biol* 7:81–91
59. Sakon JJ, Weninger KR (2010) Detecting the conformation of individual proteins in live cells. *Nat Methods* 7:203–205
60. Schroder GF, Levitt M, Brunger AT (2010) Super-resolution biomolecular crystallography with low-resolution data. *Nature* 464:1218–1222
61. Schwieters CD, Clore GM (2001) Internal coordinates for molecular dynamics and minimization in structure determination and refinement. *J Magn Reson* 152:288–302
62. Stryer L, Haugland RP (1967) Energy transfer: a spectroscopic ruler. *Proc Natl Acad Sci USA* 58:719–726
63. Sutton RB, Fasshauer D, Jahn R, Brunger AT (1998) Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* 395:347–353
64. Tochio H, Hung F, Li M, Bredt DS, Zhang M (2000) Solution structure and backbone dynamics of the second PDZ domain of postsynaptic density-95. *J Mol Biol* 295:225–237
65. van den Berk LC, Landi E, Walma T, Vuister GW, Dente L, Hendriks WJ (2007) An allosteric intramolecular PDZ-PDZ interaction modulates PTP-BL PDZ2 binding specificity. *Biochemistry* 46:13629–13637
66. Vrljic M, Strop P, Ernst JA, Sutton RB, Chu S, Brunger AT (2010) Molecular mechanism of the synaptotagmin-SNARE interaction in Ca^{2+} -triggered vesicle fusion. *Nat Struct Mol Biol* 17:325–331
67. Watkins LP, Chang H, Yang H (2006) Quantitative single-molecule conformational distributions: a case study with poly-(L-proline). *J Phys Chem A* 110:5191–5203
68. Weninger K, Bowen ME, Chu S, Brunger AT (2003) Single-molecule studies of SNARE complex assembly reveal parallel and antiparallel configurations. *Proc Natl Acad Sci USA* 100:14800–14805
69. Xue M, Ma C, Craig TK, Rosenmund C, Rizo J (2008) The Janus-faced nature of the C(2)B domain is fundamental for synaptotagmin-1 function. *Nat Struct Mol Biol* 15:1160–1168
70. Zhang Q, Fan JS, Zhang M (2001) Interdomain chaperoning between PSD-95, Dlg, and Zo-1 (PDZ) domains of glutamate receptor-interacting proteins. *J Biol Chem* 276:43216–43220