The Preprotein Binding Domain of SecA Displays Intrinsic Rotational Dynamics

Peer-reviewed author version

Vandenberk, Niels; Karamanou, Spyridoula; Portaliou, Athina G.; Zorzini, Valentina; Hofkens, Johan; HENDRIX, Jelle & Economou, Anastassios (2019) The Preprotein Binding Domain of SecA Displays Intrinsic Rotational Dynamics. In: Structure, 27 (1), p. 90-101.

DOI: 10.1016/j.str.2018.10.006
Handle: http://hdl.handle.net/1942/27818

The preprotein binding domain of SecA displays intrinsic rotational dynamics

Niels Vandenberk[1], Spyridoula Karamanou[2], Athina G. Portaliou[2], Valentina Zorzini[2], Johan Hofkens[1], Jelle Hendrix[1,3,4] and Anastassios Economou[2,4]

[1]KU Leuven, Department of Chemistry, Division for Molecular Imaging and Photonics, Laboratory for Photochemistry and Spectroscopy, Celestijnenlaan 200F, B-3001 Leuven, Belgium.

[2]KU Leuven, Department of Microbiology and Immunology, Rega Institute for Medical Research, Laboratory for Molecular Bacteriology, Herestraat 49, Gasthuisberg Campus, B-3000 Leuven, Belgium.

[3]Dynamic Bioimaging Lab, Advanced Optical Microscopy Centre, Biomedical Research Institute, Agoralaan C (BIOMED), Hasselt University, B-3590, Diepenbeek, Belgium

[4]For correspondence:

Lead contact: Anastassios Economou: tassos.economou@kuleuven.be

Jelle Hendrix: jelle.hendrix@uhasselt.be

Running title: SecA PBD rotational dynamics

Characters with spaces: 55,050

Summary

SecA converts ATP energy to protein translocation work. Together with the membrane-embedded SecY channel it forms the bacterial protein translocase. How secretory proteins bind to SecA and drive conformational cascades to promote their secretion, remains unknown. To address this, we focus on the Preprotein Binding Domain (PBD) of SecA. PBD crystalizes in three distinct states, swiveling around its narrow stem. Here, we examined whether PBD displays intrinsic dynamics in solution using single molecule Förster resonance energy transfer. Unique cysteinyl pairs on PBD and apposed domains were labelled with donor/acceptor dyes. Derivatives were analyzed using pulsed interleaved excitation and multi-parameter fluorescence detection. The PBD undergoes significant rotational motions, occupying at least 3 distinct states in dimeric and 4 in monomeric soluble SecA. Nucleotides do not affect smFRET-detectable PBD dynamics. These findings lay the foundations for single molecule dissection of translocase mechanics and suggest models for possible PBD involvement during catalysis.

Keywords

Single-molecule Förster resonance energy transfer, pulsed interleaved excitation protein conformational dynamics, photon distribution analysis, Bacterial secretion pathway, SecA motor ATPase, Domain motions

Introduction

Protein export is an essential and ubiquitous process that affects >30% of the proteome (Rapoport, 2007; Tsirigotaki et al., 2017). Many bacterial pathogenicity factors are secreted proteins and some diseases result from faulty protein targeting. Despite progress, the molecular mechanism of this fundamental process is still unclear.

Bacteria such as *Escherichia coli* (*E. coli*), are ideal models to study protein secretion. All proteins necessary for translocation through the main Sec pathway are known, as are their structures, and functional assays are available (Tsirigotaki et al., 2017). The 505 secretory preproteins of *E.coli* are sorted from cytoplasmic proteins and targeted post-translationally using signals on their N-terminal peptides and mature domains and chaperones (Chatzi et al., 2017; Gelis et al., 2007; Gouridis et al., 2009; Tsirigotaki et al., 2017). Preproteins are recognized by the protein translocase comprising the dimeric peripheral SecA ATPase bound to the SecYEG membrane-embedded protein-conducting channel, and trigger their energy-dependent transport (Tsirigotaki et al., 2017).

SecA undergoes dimer to monomer equilibria on the ribosome (Huber et al., 2011), cytoplasm and SecYEG (Gouridis et al., 2013). Its intracellular concentration is ~5.7-8.2 μM (Akita et al., 1991; Woodbury et al., 2002), dimerizes with multiple arrangements of sliding protomers in equilibrium (Gouridis et al., 2013), with a dissociation constant of ~1 nM (Kusters et al., 2011; Wowor et al., 2011).

The SecA protomer comprises four domains (**Figure 1**) (Sardis and Economou, 2010): the Nucleotide binding domain (NBD), Intramolecular regulator of the ATPase2 (IRA2 or NBD2), Pre-protein binding domain (PBD) and the C-domain, containing the flexible helical sub-structure termed Wing domain (WD). NBD and IRA2 form an RNA

helicase DEAD motor, and sandwich nucleotides. The PBD and the C-domain recognize preproteins and SecYEG (Chatzi et al., 2017; Gelis et al., 2007; Zimmer and Rapoport, 2009). In various crystal structures, the PDB swivels by ~ 60°, giving rise to three states: closed, open and wide open (**Figure 1**) (Sardis and Economou, 2010). This motion forms an apparent PDB-IRA2 clamp (**Figure 1**), that is not required for initial preprotein docking to SecA (Chatzi et al., 2017) but that may accommodate translocating chains (Bauer and Rapoport, 2009). PBD motions may contribute to translocation initiation (Gold et al., 2013) but whether they occur outside crystal lattices and how they might participate in translocase catalysis, remains unknown.

To better understand protein translocase mechanics a deeper, quantitative understanding is needed of SecA conformational changes and dynamics, typically offered by state-of-the-art biophysical approaches. A notable ensemble approach is hydrogen-deuterium exchange mass spectrometry (HDX-MS) (Sardis et al., 2017; Tsirigotaki et al., 2018). At single molecule (sm) level, Förster resonance energy transfer (FRET), the near-field (1-10 nm) radiation-less transfer of energy from an excited donor fluorophore to an excitable acceptor fluorophore, is particularly powerful (Förster, 1946; Zander et al., 1996). SmFRET monitors intramolecular (conformational rearrangements/domain motions) or intermolecular (association/dissociation) dynamic processes in real time. Thus distinct conformational states, commonly lost in ensemble measurements, can be identified (Ha and Tinnefeld, 2012; Roy et al., 2008) and their physical distances derived from the measured FRET quantum efficiency (*E*) between the probes (Hellenkamp et al., 2018; Kapanidis et al., 2004; Lee et al., 2005; Vandenberk et al., 2018). In solution smFRET, fluorescent molecules diffuse freely through the confocal volume of a microscope (**Figures 2A** and **S1A**) and their fluorescent bursts, i.e. the number of photons they emit during the time they stay in

focus, are measured (Eggeling et al., 2001; Kellner et al., 2014; Zander et al., 1996). In such analyses the donor and acceptor fluorophores can be excited in an alternating fashion on the nanosecond time scale (referred to as “pulsed interleaved excitation”, PIE) and registered via multi-parameter fluorescence detection (MFD) (**Figures 2A** and **S1A**), with each photon being detected by time-correlated single-photon counting (TCSPC)(**Figure S1A**) (Hendrix and Lamb, 2013; Kudryavtsev et al., 2012; Müller et al., 2005). This yields a wealth of information on both the FRET donor and acceptor including: fluorescence intensity, lifetime, anisotropy and color (**Figure 2A**). SmFRET can detect a broad range of timescales in protein folding/conformational rearrangements (nanoseconds-minutes). These include dye rotation (ns), unfolded state dynamics (ns-μs), folding/rearrangements (μs-s) which depend on ultra-fast motions of structured protein parts (ps-μs), fast whole domain motions (μs-ms) and misfolding/oligomerization (s-min/h) (Schuler and Hofmann, 2013). SmFRET, combined with MFD-PIE and proper time-resolved analysis, discriminates between the intrinsic protein dynamics that are of biological interest from the chemical properties of the labels and the artifacts and photophysical dynamics that need to be excluded (Cotlet et al., 2005; Hofkens et al., 2003; Vandenberk et al., 2018; Vosch et al., 2003).

Here we present the first smFRET/MFD-PIE-based pipeline to study the conformational dynamics of the PBD domain of SecA on a timescale of 0.1 to 10 ms. Residues, ideally located for FRET, were mutated into unique cysteinyl pairs and tested for functionality *in vivo*. Derivative proteins were labeled and tested for functionality *in vitro*. Finally, physical distances of PBD motions were determined by photon distribution analysis (PDA) (Antonik et al., 2006; Kalinin et al., 2008).

Using smFRET we determined that the PBD displays intrinsic domain swiveling in solution. It samples at least three conformational states that approximate those in

crystal lattices but vary in their abundance. Dimer to monomer transition not only shifts the equilibrium between the observed PBD states and the physical positioning of PBD but also gives rise to a fourth state. We hypothesize that the intrinsic mobility of PBD allows: a. binding of hundreds of dissimilar, non-folded preproteins. b. mechanical motions in the SecY-bound state to promote translocation. In the functionally quiescent soluble dimeric and the monomeric SecA, these PBD motions are uncoupled from nucleotide binding to the ATPase motor. These data set the foundations for future quantitative dissection of SecA and translocase dynamics in the presence of all the reaction ligands during ongoing translocation.

Results

A 5-step pipeline for smFRET analysis of SecA

The conformational dynamics of the PBD domain of SecA were investigated using a 5-step pipeline (**Figure S2A**), detailed below: **1.** selection of optimal residues for placing donor/acceptor dye pairs, **2.** generation of mutants, testing of their *in vivo* function, gene over-expression and protein purification, **3.** optimization of labeling with fluorescent probes, **4.** smFRET measurements, data analysis and statistical treatment and **5.** derivations of physical distances of PBD motions.

Selection of SecA residues and cysteinyl mutagenesis

For residue-specific labeling via maleimide-modified dyes reacting with thiol groups of cysteinyl (cys) residues (**Figure S1C**), we used the fully functional SecA(Cys$^-$) (Chatzi et al., 2017; Sardis et al., 2017). The *secA*(cys$^-$) gene derivatives with specific cysteine pairs were generated. Based on six selection criteria and the anticipated changes in FRET according to FRET-restrained positioning and screening (FPS, see STAR methods) (**Figure S1D** and **Table S1**), 31 possible FRET pairs were identified (**Table S2).** Target residues were mutagenized (STAR methods; **Tables S3** and **S4**). All but one of the mutated genes were shown by genetic complementation to restore growth to strain BL21.19 that carries a chromosomal thermosensitive *secA* (Karamanou et al., 1999). Therefore, cys mutagenesis did not affect SecA functionality. Next, mutated SecA derivatives were purified to homogeneity (Gouridis et al., 2013; Papanikolau et al., 2007), in the presence of high concentrations of a dithiol reducing agent (e.g. dithiothreitol, DTT), to maintain optimal labeling (**Figure S2E**) and stored for up to 2 months (50% V/V glycerol, -20 °C) (**Figure S2F**). Some derivatives

displayed altered dimerization equilibria and were not studied further (**Figure S2A** and **S2B)**.

Labeling of SecA double cysteinyl derivatives

DTT was removed and SecAs were labelled using two different donor/acceptor (hereafter: D, A) maleimide-attached dye pairs under pH 6.5-7.5 (**Figure S2C**) (Brewer and Riehm, 1967): ATTO488/ATTO647N and ATTO488/Alexa647. Measurable nonspecific labeling of SecA(Cys^-) occured with ATTO647N but not with Alexa647 or ATTO488 (**Figure S2D**). Therefore, ATTO488/Alexa647 were used hereafter. Since both dyes are added simultaneously, mixed labeled populations were obtained containing both hetero-labeled (D and A or A and D on either one of the two cysteinyl residues; both usable; **Figure 2A**) and homo-labeled (D- or A-only). However, PIE allows the unwanted homo-labelled molecules to be removed during data analysis (Kudryavtsev et al., 2012). A hetero-labeling efficiency of 40-60% was typically obtained.

Labeled SecAs were re-purified by size-exclusion chromatography, collecting only the narrowest, highly homogeneous chromatographic peaks, and removing free dye and protein aggregates. Purified, labelled SecAs were found to be fully functional using an established *in vitro* assay that measures their ability to have their ATPase activity stimulated by secretory preproteins in the presence of membrane vesicles (Gouridis et al., 2010).

Accurate measurement of physical positioning of PBD states in soluble SecA

To report on all the possible anticipated PBD motions against either WD of the C-domain or the IRA2 domain, we used hereafter six of the double-Cys SecA derivatives,

which satisfied the previous preparative steps (D1-6; **Figure 2B**). To directly compare distances in the six derivatives, they shared one of the two cysteinyl residues (V280C on PBD) (**Figure 2B**).

Accurate physical distances of PBD motions were derived using photon distribution analysis (PDA) (Antonik et al., 2006; Kalinin et al., 2008). For PDA to provide reliable structural information, the donor and acceptor dye spectroscopic properties when attached to the molecule of interest need to be verified. Considerable quenching of the dye's fluorescence, independently of FRET, results in a lower lifetime and quantum yield, which complicates further analysis. Anisotropy values that are too high could reflect partial sticking of the dyes (that are overall negatively charged) on positively charged neighboring protein surfaces. This in turn causes unpredictable shifts in FRET efficiency, or even in the appearance of nonsense FRET states, and renders the derived distances and distance distributions inaccurate. Therefore, we first focused our analysis on the donor dye for proteins that were labeled only by the donor (to be able to study the dye in the absence of FRET) or on the acceptor dye for double-labeled proteins. When plotted in a r_D-τ_D or r_A-τ_A 2D histogram (with the fluorescence lifetime of the FRET donor, τ_D, that of the acceptor, τ_A, the steady-state anisotropy of the FRET donor, r_D, and of the acceptor, r_A) the data points are preferentially localized in one main fluorescence lifetime and anisotropy population, with the fluorescence lifetime preferentially close to that of the free dye, and the anisotropy preferentially low. On the other hand, if the dyes exhibit heterogeneous lifetimes or anisotropy values, then care must be taken in further analysis.

For the six SecA derivatives that were analyzed, we next investigated the donor dye of donor-only labeled molecules and the acceptor dye of double-labeled molecules. D1, D2 and D3 displayed single populations for both lifetime and anisotropy

and the exact expected lifetime of the donor dye (4.0 ns; ATTO-TEC) (**Figure S3B-S3D**). D5 displayed overall also single populations, however a minor 2nd anisotropy population is slightly appearing for the donor (**Figure S3F**).

On the other hand, derivative D4, displayed at least two anisotropy populations of the donor with the high anisotropy state being higher populated (**Figure S3E**). D4 was not used further. Derivative D6 shows a major (4.0 ns) and minor lifetime (2.5 ns) distribution of the Donor-only population, likely reflecting dye-protein interactions (data not shown). Therefore, D6 was not used further.

Taken together, our analysis of protein-attached dye behavior allowed us to retain only those double-Cys mutants that ensured deriving accurate structural information from PDA analysis (see below).

Optimization of quantitative smFRET measurements, data processing and analysis

Next, we analyzed the fluorescently hetero-labelled SecAs by MFD-PIE to detect smFRET bursts and deduce conformational behavior. Getting enough bursts from doubly hetero-labeled SecAs, in the shortest possible time yields high quality data due to the collection of several thousand single molecule samples and minimizes potential loss of SecA functionality. Condition optimization included BSA-coated coverslips and soluble agents such as free BSA or Trolox to prevent non-specific protein absorption and photo-bleaching) (**Figure S2G)** (Aitken et al., 2008; Rasnik et al., 2006; Vandenberk et al., 2018; Vogelsang et al., 2008).

Primary experimental values from thousands of bursts are graphed in 2D plots of FRET efficiency (**Figure 2C**; *y*-axis) against the fluorescence lifetime of the donor in the presence of the acceptor ($\underline{\tau}_{D(A)}$; *x*-axis). In these plots, the burst data points are

distributed in one or more “clouds” along a curved diagonal (**Figure 2C**) (STAR methods: “Burst-wise fluorescence lifetime”). The curved diagonal intersects the *x*-axis at the lifetime of the donor-only population and the *y*-axis at unity.

When FRET states do not interconvert during the single molecule observation time (0.1-10 ms), the FRET of the molecules is said to be “static” and photon bursts from the protein’s fluorescent probes will have values that fall tightly on the diagonal, commonly referred to as the ‘static FRET line’ (red). If more than one “cloud” of data co-exist on the static FRET line for the same protein, or if the clouds are poly-dispersed, the protein molecules as they travel through the confocal volume exist in and retain multiple conformational states. This “solution” analysis can define the specific conformational states sampled and how well they are represented in the whole population.

Burst values that deviate to the right of the diagonal (**Figure 2C**, green dots) would indicate that the two fluorescent probes, and hence the SecA domains on which they are carried, display conformational dynamics that occur during the timescale of the measurement, i.e. while the protein is diffusing through the focus of the laser (0.1-10 ms).

Analysis of PBD motions in soluble SecA by smFRET

We next proceeded to analyzing smFRET-derived PBD motions in dimeric SecA derivatives D1, D2, D3 and D5 by mixing ~100-200 pM of fluorescently labeled with 100 nM of unlabeled SecA, 100-fold over the dimerization K_d (**Figures 3** and **S4, left**) (Kusters et al., 2011). This allows analyzing smFRET events from a single hetero-labeled SecA protomer but within the context of the physiological SecA dimer (**Figures 3A-3B,** left, cartoon). Primary experimental values were graphed in 2D plots of FRET

efficiency against the fluorescence lifetime of the donor (as in **Figures 3** and **S4, left**), of the many parameters were obtained by MFD-PIE detection, 6 are shown for a representative SecA derivative (**Figure S3A**). In all cases, results reported in this study are consistent with “static FRET” behavior for PBD motions (**Figure 2C, red**).

The PBD of derivative D1 that probes a potential PBD to IRA2 motion, showed a high FRET state in one third of the molecules (**Figure 3A**). The structural interpretation is, therefore, that in these molecules that diffused through the confocal volume, their PBD is positioned within ~4 nm of IRA2 (State 1). However, a discernible number of bursts, display lower (State 2; ~50%) or much lower (State 3; ~10%) FRET efficiencies and, therefore, several of these SecA molecules have their PBD positioned away from IRA2 (6 or more nm; State 2 and 3). These results are corroborated and strengthened by the analysis of D2 that also probes the PBD to IRA2 motion (**Figure S4A, left**). They demonstrate that in a given population of soluble SecA, the PBD samples multiple conformational states.

In D3, that probes the PBD to WD motion, the PBD occupies low and medium FRET states for most of the bursts measured (**Figure 3B**, **left**). This implies that in most of the molecules diffusing through the confocal volume, PBD is positioned away from WD residues in this pair. D5 (**Figure S4C, left**) also probes the same inter-domain interaction. D5 showed a predominantly low and a minor-high FRET state that would be compatible with the FRET pair in some molecules having a “closed” PBD-WD interface. Because of the distributions in two distinct clouds, D5 also supports the existence of at least two stable PBD states.

Accurate measurement of physical positioning of PBD states in soluble SecA

The multiple probe pairs allow an approximate triangulation of the positions of PBD in SecA using FRET-derived structural distances by visually estimating *E* from the 2D plots (**Figure S4E, Left).** This approximation is more accurate for samples that exhibit small, tightly distributed FRET populations and not for the smFRET data of the PBD showing broadly distributed states for some derivatives. FRET-competent states may be too close to be distinguished by eye and thus accurate physical distances of PBD motions were determined by PDA (Antonik et al., 2006; Kalinin et al., 2008).

Conformational dynamics in the 0.1-10 ms timescale (**Figure 2C**, green) can be detected by cutting the burst data into time windows of specified length. This 'time window analysis' did not reveal clear differences between the FRET histograms for D1, 2, 3 and 5 (**Figure S5** and **Data S2**), indicating the PBD displays no conformational dynamics in the 0.1-10 ms timescale. In other words, PBD moves from one of its states to the other more slowly than 10 ms. Therefore, a PDA model incorporating different FRET states was fitted to the data, with each state assuming a Gaussian distance distribution (Kalinin et al., 2008; Kalinin et al., 2012; Talavera et al., 2018) (**Figures 3A-3B, right** and **S6A-S6B)**. This analysis revealed that taken collectively, D1 (**Figure 3A**), D2 (**Figure S6A, right**), D3 (**Figure 3B**) and D5 (**Figure S6B, right**) support the existence of the same three clearly defined PBD states (1, 2 and 3) (**Figure 3C**). PBD is either almost equidistant from IRA2 and WD (state 2, **Figure 3C, bottom**) or moves close to IRA2 (state 1) or to WD (state 3). The fraction of molecules that occupy these states is distinct: more than half populate state 2, followed by state 1 and 3.

In summary, D1, 2, 3 and 5 were analyzed globally using PDA. PBD occupies three distinct states in solution, with State 2 being the most populated.

Dimer to monomer transition affects PBD motions

During catalysis SecA undergoes dimer to monomer transitions (Gouridis et al., 2013; Singh et al., 2014). To investigate if monomerization affects PBD conformations and dynamics, we examined the four fluorescently labeled SecA derivatives at concentrations of 100-200 pM that, based on the determined K_d of ~1 nM (Kusters et al., 2011), should push the equilibrium mainly to monomers (**Figures 4**, **left** and **S4, right**). PDA analysis of monomeric D1, D2, D3 and D5 (**Figures 4A-4B** and **S6A-6B**) was carried out to obtain accurate measurements of the positioning of PBD states.

Monomeric D1 (**Figure 4A, left**), that probes the PBD to IRA2 inter-domain motion, exists in three states with largely similar distances to those of dimeric SecA and a fourth state, termed 2A protruding away from the main protein body, also became apparent (**Figure 4A, 4C**). The occupancy ratio between the states showed significant changes compared to that in the dimer. PBD in the majority of the monomeric molecules now occupy states 2 and 2A (> 85%) to the expense of State 1. State 3 remains poorly populated (**Figure 4C**). The PBD of monomeric D2 that also probes the PBD to IRA2 inter-domain motion (**Figure S4A, right**), yielded shifts like those of D1 (**Figure 4C**). These data demonstrated that PDB stays away from both IRA2 and WD, in most of the monomeric SecA molecules that diffused through the confocal volume, while half of the state 2 PBDs swivel away from the protein body.

In D3 and D5 (**Figures 4B** and **S4B, S4C, right**) that probe the PBD to WD motion, the PBD of the monomer samples low FRET states for most of the individual bursts measured. Therefore, the PBD is positioned far away from its respective WD residue pairs, >8 nm in most of the molecules of both derivatives. D3 and D5 displayed four distinct PBD states after PDA with similar ratios of those states, but with different distances compared to those of monomeric D1 and D2 (**Figure 3C**).

To further corroborate the data obtained above at the concentrations of presumed SecA monomerization and to exclude contributions from surviving dimeric molecules, we used a genetically constructed monomer derivative, mSecA, which displays a 10^5-fold loss in its dimerization K_d (~133 μM) but becomes dimeric and fully functional at high concentrations (Gouridis et al., 2013). We constructed mSecA variants with the 4 Cys-pair derivatives and determined their smFRET profiles. These profiles were highly comparable to those of kinetic monomers generated after dilution (**Figure S3G;** D1 shown as a representative example).

The detected differences between dimeric and monomeric states allowed determination of the K_d of this transition. Unlabeled SecA D1 was titrated into reactions containing 200 pM fluorescently labeled D1 (kinetic monomer conditions) (**Data S1A-S1J**), changes to the 3 observed PBD states (**Figure 3**), were monitored and a sigmoid was fitted to the data. A K_d of 2.2 nM and 3.0 nM for State 3 and 1, respectively was derived (**Data S1K**), close to the one previously obtained (0.74 nM; (Kusters et al., 2011)). Additionally, burst-wise fluorescence correlation of every single molecule event (see STAR methods) revealed a significant difference in diffusion coefficient between the monomeric and dimeric state of D1, indicative of a hydrodynamic property change of the protein (**Data S1L**).

Taken together, SecA monomerization causes significant conformational differences to the PBD relative to those in the dimer state, both in terms of number and the fraction of the population that occupies the conformational states.

Nucleotides have only a minor effect on PBD dynamics

SecA in solution binds ATP, rapidly converts it to ADP and acquires the quiescent, thermally stabilized state (Keramisanou et al., 2006; Sianidis et al., 2001).

357 However, addition of ADP to dimeric SecA leads to no changes in the FRET states of
358 the PBD (**Figures 5A-5B**). The same is seen with ATP (**Figures 5A-5B**). Similar
359 effects were seen with the kinetic SecA monomer (**Figure 5B**).
360 These data suggest that in the quiescent state of soluble SecA, nucleotide
361 interactions in the ATPase motor of SecA are not transmitted to the PBD or cannot be
362 detected in our assay.
363
364 **Visualizing the different PBD states in soluble SecA by rigid body modelling**
365 As it undergoes the motions detected here, the PBD does not lose internal
366 secondary structure as determined by HDX-MS (Krisnamurthy et al, in preparation).
367 Therefore, it largely undergoes rigid body rotations. To visualize these complex
368 motions in 3-dimensional space in SecA in solution, we rotated the PBD as a rigid body
369 around its Stem, using the beginning of the Stem as a fixed point. In the absence of
370 any currently available information on SecA structural dynamics, we made the simple
371 assumption that, as seen in crystal structures, no additional large motions occur in the
372 three other domains of SecA. Thus, while maintaining the rest of SecA as a rigid body
373 we used the FRET-derived distances as restraints, (**Figures 3C** and **4C**). The three
374 smFRET-derived PBD states of dimeric SecA differ slightly (e.g. State 1 vs the 'closed'
375 state; **Figure 6,** magenta) or more substantially (e.g. State 2 vs the 'open' state), from
376 the three PBD states seen in crystal structures (grey).
377 The four states calculated by smFRET for monomeric SecA (**Figure 4C**), were
378 also visualized (**Figure 6,** pink). States 1-3 are closer to the ones seen in the dimer,
379 while State 2A, is clearly distinct and has not been seen before in X-ray-derived
380 structures. The PBD of state 2A relates to the crystallographic open state and the
381 smFRET-derived State 2 of the dimer but projects further away from the rest of the

382 protein and is leaning closer to the IRA2. Additional restraints combined with dynamic
383 modelling of the structure will be required to trace these complex motions during
384 catalysis.
385

Discussion

We present the first, to our knowledge, complete pipeline of generating solution smFRET-compatible, hetero-labelled, while fully functional, SecA and measuring its conformational domain dynamics. This effort aims to take translocase studies beyond X-ray crystallography or NMR, focusing on structural and conformational dynamics analyzed at native conditions.

Here, we focused on the conformational domain dynamics of the PBD of SecA. This domain was an attractive target for multiple reasons: firstly, PBD binds signal peptides and, together with the C-domain, mature domains (Chatzi et al., 2017; Gelis et al., 2007). Preprotein clients are likely to affect PBD dynamics and perhaps even exploit them to convert ATP cycling to translocation-related work. Secondly, the PBD occupies different states in crystal structures. Such PBD motions may be coupled to mechanical work, although, currently, there is no direct structural evidence for this. Prior to this study, it was unknown if and to what extent PBD motions might occur in SecA in solution. Finally, seen from a technical, smFRET perspective, if PBD motions did indeed occur, they would provide distance changes that could be appropriately probed by smFRET.

Using smFRET we determined that the PBD of soluble dimeric SecA samples at least two major States (State 2 of ~40-50% and State 1 of 40-50% of the molecules) and a minor State (State 3; of ~5-10%) (**Figures 3C** and **6**). The distances of these states are similar to the ones in crystal structures, yet deviate from them by 2-15%, suggesting that the crystal lattices selected/stabilized slightly different PBD states (**Figure 6**).

Site-directed spin labeling and NMR-detected paramagnetic relaxation enhancement analysis that previously probed PBD to WD motions, suggested the

State 2 (“open”) and State 3 (“wide open”) may be occupied by 90% and 10% of the molecules, respectively (Gelis et al., 2007). Single-molecule dissection now reveals that the intrinsic dynamics of PBD are more complex, with the 90% population being split between States 2 and 1.

PBD swiveling is significantly affected when SecA monomerizes (**Figure 4**). Four distinct PBD populations are discernible but both the measured distances and distribution of molecules between them changed compared to those of the dimer. Many molecules also display a new State 2A, comparable with State 2 but with the PBD moved further away from the body of the protein. Therefore, PBD not only displays remarkable rotational dynamics but also the states that it occupies are structurally distinct. Occupancy of these apparent stable energetic minima are influenced by the quaternary state of SecA. These states interconvert slowly in tens of milliseconds, characteristic whole domain motions (Schuler and Hofmann, 2013). How they interconvert, with which rate constants, in which order and what their lifetimes are, will require future prolonged smFRET kinetic measurements of immobilized molecules (Roy et al., 2008). Diffusion measurements using our MFD-PIE set-up cannot define these slow rate constants (<0.1 ms^{-1}). The 2D plots also hint to the presence of rare kinetic exchanges that are faster than the burst duration. The latter could be investigated in more detail by using approaches like filtered FCS (Dolino et al., 2016; Felekyan et al., 2012).

Why has the PBD evolved to display such dynamics even in quiescent SecA? A major contribution of PBD at this early stage of secretion is to optimize promiscuous preprotein docking on SecA. Although enzymatically quiescent as an ATPase, solutble SecA is still a low micromolar K_d preprotein receptor and becomes a high nanomolar to low micromolar K_d receptor when bound to SecY (Gouridis et al., 2009; Gouridis et

al., 2013). SecA binds hundreds of secretory proteins that differ in size, structural folds, non-folded states and senses them as bivalent ligands, recognizing both their signal peptides and mature domain patches at different clefts (Chatzi et al., 2017; Sardis et al., 2017; Tsirigotaki et al., 2018). Presumably, PBD positioning "guides" preproteins to productively dock on SecA to proceed to secretion (Sardis et al., 2017).

In contrast to the effect of dimer to monomer transitions, nucleotides do not appear to alter intrinsic PBD swiveling detectably. This raises the possibility that while PBD is inherently dynamic in the catalytically quiescent SecA studied here, nucleotide-driven conformational cues in the helicase ATPase motor (**Figure 1**, **blue and cyan**) are not coupled to PBD (**magenta**) motions and *vice versa*. This is intriguing since, during catalysis, PBD does exert long-range effects that "break" Gate1, a salt bridge in the motor, that prevents it from acquiring elevated ATP turnovers and, moreover, mutations in the motor or the PBD do affect each other's conformation (Karamanou et al., 2007; Keramisanou et al., 2006). Our findings lead us to hypothesize the existence of a sophisticated auto-inhibitory mechanism. SecY docking, that primes SecA for high affinity preprotein binding and ATP turnover (Gouridis et al., 2013), and preprotein binding (Chatzi et al., 2017; Gelis et al., 2007) are expected to relieve this auto-inhibition and allow ATPase motor/PBD conformational cross-talk.

If and how PBD motions contribute mechanistically to actual preprotein translocation through SecYEG remains unknown. We entertain two hypotheses, both assume that PBD oscillates between the 3-4 states identified here. In dimeric SecA that initially docks on SecYEG (Gouridis et al., 2013), states 2 and 1 predominate. In these states the PBD forms a tighter PBD-IRA2 clamp and may directly contact translocating chains trapped inside it (Bauer and Rapoport, 2009). PBD may thus (**Figure 7**): **a.** act as a brake, to prevent translocating chain "back-slippage" and control

461 a Brownian “ratchet” allowing forward motion (**left**)(Allen et al., 2016; Tsirigotaki et al.,
462 2017). or **b.** bind to the translocating chain and exert a “pushing” stroke (**middle**). A
463 third, possible role relates to the channel rather than to the exiting chain. PBD contacts
464 directly the large, functionally important cytoplasmic protrusion of SecY (**right**)
465 (Zimmer et al., 2008) as can be seen in models of SecA bound to two SecY structures
466 (**Figure S7**). Thus, PBD motions could directly control both the structural dynamics of
467 SecYEG and those of the translocating chain.

468 SmFRET-derived domain dynamics analyzed here combined with HDX-MS-derived
469 structural dynamics analyses (Sardis et al., 2017; Tsirigotaki et al., 2018; Tsirigotaki et
470 al., 2017) lay the foundations for quantitative dissection of the functional translocase.
471 SecA offers an interesting example of smFRET-derived structural information in
472 proteins. This is a non-trivial, multi-disciplinary effort, requiring the design and testing
473 of multiple fluorescent pair derivatives. Compared to other methods, smFRET is
474 uniquely suited to the analysis of membrane-associated systems with short-lived,
475 interconverting states and opens numerous exciting future possibilities.

Acknowledgements

We are grateful to G. Gouridis (Rega Institute, KU Leuven) for discussions. Our research was funded through: the Research Foundation Flanders (FWO) [grant #G.0B49.15 (to JHo, JHe and SK), grants #G0683.15 and #G0A5817.N (to JHo), large infrastructure grant ZW15_09 GOH6316N (to JHo. and JHe); grant #G0C6814N RiMembR (to A.E.); the Flemish government through long term structural funding Methusalem CASAS2, Grant #Meth/15/04 (to JHo); the FWO/F.R.S.-FNRS "Excellence of Science-EOS" programme grant #30550343 (to AE)], the KU Leuven (grant C14/16/053; to JHe), the EU (FP7 KBBE.2013.3.6-02: Synthetic Biology towards applications; #613877 StrepSynth; to AE) and RUN (#RUN/16/001 KU Leuven; to AE). NV acknowledges the Agency for Innovation by Science and Technology in Flanders (IWT) for a doctoral scholarship. VZ is an FWO postdoctoral fellow.

Author contributions

NV performed protein labeling and all smFRET analysis. SK designed genetic constructs, purified proteins, analyzed them by size-exclusion/MALS and performed functional *in vivo* and *in vitro* assays. AGP designed and generated genetic constructs. VZ generated and analyzed smFRET-derived SecA models. JHe set up the MFD-PIE microscope, data analysis pipeline and supervised smFRET experiments. NV and AE wrote the paper with contributions from JHe, SK, VZ and AP. JHe and SK managed the project. JHe, JHo, SK and AE conceived and designed the project. All authors reviewed the manuscript.

Declaration of interests

The authors declare they have no competing financial interests.

References

Aitken, C.E., Marshall, R.A., and Puglisi, J.D. (2008). An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. Biophys J *94*, 1826-1835.

Akita, M., Shinkai, A., Matsuyama, S., and Mizushima, S. (1991). SecA, an essential component of the secretory machinery of Escherichia coli, exists as homodimer. Biochem Biophys Res Commun *174*, 211-216.

Allen, W.J., Corey, R.A., Oatley, P., Sessions, R.B., Baldwin, S.A., Radford, S.E., Tuma, R., and Collinson, I. (2016). Two-way communication between SecY and SecA suggests a Brownian ratchet mechanism for protein translocation. Elife *5*.

Antonik, M., Felekyan, S., Gaiduk, A., and Seidel, C.A. (2006). Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. J Phys Chem B *110*, 6970-6978.

Bauer, B.W., and Rapoport, T.A. (2009). Mapping polypeptide interactions of the SecA ATPase during translocation. Proc Natl Acad Sci U S A *106*, 20800-20805.

Brewer, C.F., and Riehm, J.P. (1967). Evidence for Possible Nonspecific Reactions between N-Ethylmaleimide and Proteins. Analytical Biochemistry *18*, 248-&.

Chatzi, K.E., Sardis, M.F., Tsirigotaki, A., Koukaki, M., Sostaric, N., Konijnenberg, A., Sobott, F., Kalodimos, C.G., Karamanou, S., and Economou, A. (2017). Preprotein mature domains contain translocase targeting signals that are essential for secretion. J Cell Biol *216*, 1357-1369.

Cordes, T., Vogelsang, J., and Tinnefeld, P. (2009). On the mechanism of Trolox as antiblinking and antibleaching reagent. J Am Chem Soc *131*, 5018-5019.

Cotlet, M., Vosch, T., Habuchi, S., Weil, T., Mullen, K., Hofkens, J., and De Schryver, F. (2005). Probing intramolecular Forster resonance energy transfer in a naphthaleneimide-peryleneimide-terrylenediimide-based dendrimer by ensemble and single-molecule fluorescence spectroscopy. J Am Chem Soc *127*, 9760-9768.

Dolino, D.M., Rezaei Adariani, S., Shaikh, S.A., Jayaraman, V., and Sanabria, H. (2016). Conformational Selection and Submillisecond Dynamics of the Ligand-binding Domain of the N-Methyl-d-aspartate Receptor. J Biol Chem *291*, 16175-16185.

Eggeling, C., Berger, S., Brand, L., Fries, J.R., Schaffer, J., Volkmer, A., and Seidel, C.A. (2001). Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. J Biotechnol *86*, 163-180.

Felekyan, S., Kalinin, S., Sanabria, H., Valeri, A., and Seidel, C.A. (2012). Filtered FCS: species auto- and cross-correlation functions highlight binding and dynamics in biomolecules. Chemphyschem *13*, 1036-1053.

Fiser, A., and Sali, A. (2003). ModLoop: automated modeling of loops in protein structures. Bioinformatics *19*, 2500-2501.

Fiser, A., and Simon, I. (2000). Predicting the oxidation state of cysteines by multiple sequence alignment. Bioinformatics *16*, 251-256.

Förster, T. (1946). Energiewanderung und Fluoreszenz. In Naturwissenschaften, pp. 166-175.

Gansen, A., Valeri, A., Hauger, F., Felekyan, S., Kalinin, S., Toth, K., Langowski, J., and Seidel, C.A. (2009). Nucleosome disassembly intermediates characterized by single-molecule FRET. Proc Natl Acad Sci U S A *106*, 15308-15313.

Gelis, I., Bonvin, A.M., Keramisanou, D., Koukaki, M., Gouridis, G., Karamanou, S., Economou, A., and Kalodimos, C.G. (2007). Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. Cell *131*, 756-769.

Gold, V.A., Whitehouse, S., Robson, A., and Collinson, I. (2013). The dynamic action of SecA during the initiation of protein translocation. Biochem J *449*, 695-705.

Gouridis, G., Karamanou, S., Gelis, I., Kalodimos, C.G., and Economou, A. (2009). Signal peptides are allosteric activators of the protein translocase. Nature *462*, 363-367.

Gouridis, G., Karamanou, S., Koukaki, M., and Economou, A. (2010). In vitro assays to analyze translocation of the model secretory preprotein alkaline phosphatase. Methods Mol Biol *619*, 157-172.

Gouridis, G., Karamanou, S., Sardis, M.F., Scharer, M.A., Capitani, G., and Economou, A. (2013). Quaternary dynamics of the SecA motor drive translocase catalysis. Mol Cell *52*, 655-666.

554 Ha, T., and Tinnefeld, P. (2012). Photophysics of Fluorescent Probes for Single-Molecule Biophysics
555 and Super-Resolution Imaging. Annu Rev Phys Chem *63*, 595-617.
556 Hellenkamp, B., Schmid, S., Doroshenko, O., Opanasyuk, O., Kühnemuth, R., Rezaei Adariani, S.,
557 Ambrose, B., Aznauryan, M., Barth, A., Birkedal, V.*, et al.* (2018). Precision and accuracy of single-
558 molecule FRET measurements—a multi-laboratory benchmark study. Nature Methods *15*, 669-676.
559 Hendrix, J., and Lamb, D.C. (2013). Pulsed interleaved excitation: principles and applications. Methods
560 Enzymol *518*, 205-243.
561 Hofkens, J., Cotlet, M., Vosch, T., Tinnefeld, P., Weston, K.D., Ego, C., Grimsdale, A., Müllen, K.,
562 Beljonne, D., Bredas, J.L.*, et al.* (2003). Revealing competitive Forster-type resonance energy-transfer
563 pathways in single bichromophoric molecules. Proc Natl Acad Sci U S A *100*, 13146-13151.
564 Huang, B., Wu, H., Kim, S., and Zare, R.N. (2005). Coating of poly(dimethylsiloxane) with n-dodecyl-
565 beta-D-maltoside to minimize nonspecific protein adsorption. Lab Chip *5*, 1005-1007.
566 Huber, D., Rajagopalan, N., Preissler, S., Rocco, M.A., Merz, F., Kramer, G., and Bukau, B. (2011).
567 SecA interacts with ribosomes in order to facilitate posttranslational translocation in bacteria. Mol Cell
568 *41*, 343-353.
569 Jomaa, A., Boehringer, D., Leibundgut, M., and Ban, N. (2016). Structures of the E. coli translating
570 ribosome with SRP and its receptor and with the translocon. Nat Commun *7*, 10471.
571 Kalinin, S., Felekyan, S., Valeri, A., and Seidel, C.A. (2008). Characterizing multiple molecular States
572 in single-molecule multiparameter fluorescence detection by probability distribution analysis. J Phys
573 Chem B *112*, 8361-8374.
574 Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P.J., Berger, S., Restle, T., Goody, R.S., Gohlke, H., and
575 Seidel, C.A. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural
576 modeling. Nat Methods *9*, 1218-1225.
577 Kapanidis, A.N., Lee, N.K., Laurence, T.A., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-
578 aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single
579 molecules. Proc Natl Acad Sci U S A *101*, 8936-8941.
580 Karamanou, S., Gouridis, G., Papanikou, E., Sianidis, G., Gelis, I., Keramisanou, D., Vrontou, E.,
581 Kalodimos, C.G., and Economou, A. (2007). Preprotein-controlled catalysis in the helicase motor of
582 SecA. EMBO J *26*, 2904-2914.
583 Karamanou, S., Vrontou, E., Sianidis, G., Baud, C., Roos, T., Kuhn, A., Politou, A.S., and Economou,
584 A. (1999). A molecular switch in SecA protein couples ATP hydrolysis to protein translocation. Mol
585 Microbiol *34*, 1133-1145.
586 Kellner, R., Hofmann, H., Barducci, A., Wunderlich, B., Nettels, D., and Schuler, B. (2014). Single-
587 molecule spectroscopy reveals chaperone-mediated expansion of substrate protein. Proc Natl Acad Sci
588 U S A *111*, 13355-13360.
589 Keramisanou, D., Biris, N., Gelis, I., Sianidis, G., Karamanou, S., Economou, A., and Kalodimos, C.G.
590 (2006). Disorder-order folding transitions underlie catalysis in the helicase motor of SecA. Nat Struct
591 Mol Biol *13*, 594-602.
592 Kudryavtsev, V., Sikor, M., Kalinin, S., Mokranjac, D., Seidel, C.A., and Lamb, D.C. (2012).
593 Combining MFD and PIE for accurate single-pair Forster resonance energy transfer measurements.
594 Chemphyschem *13*, 1060-1078.
595 Kusters, I., van den Bogaart, G., Kedrov, A., Krasnikov, V., Fulyani, F., Poolman, B., and Driessen,
596 A.J. (2011). Quaternary structure of SecA in solution and bound to SecYEG probed at the single
597 molecule level. Structure *19*, 430-439.
598 Lamb, D.C., Schenk, A., Rocker, C., Scalfi-Happ, C., and Nienhaus, G.U. (2000a). Sensitivity
599 enhancement in fluorescence correlation spectroscopy of multiple species using time-gated detection.
600 Biophys J *79*, 1129-1138.
601 Lamb, D.C., Schenk, A., Röcker, C., Scalfi-Happ, C., and Nienhaus, G.U. (2000b). Sensitivity
602 enhancement in fluorescence correlation spectroscopy of multiple species using time-gated detection.
603 Biophysical journal *79*, 1129-1138.
604 Laurence, T.A., Kwon, Y., Yin, E., Hollars, C.W., Camarero, J.A., and Barsky, D. (2007). Correlation
605 spectroscopy of minor fluorescent species: signal purification and distribution analysis. Biophys J *92*,
606 2184-2198.

Lee, N.K., Kapanidis, A.N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R.H., and Weiss, S. (2005). Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. Biophys J *88*, 2939-2953.

Moeyaert, B., Nguyen Bich, N., De Zitter, E., Rocha, S., Clays, K., Mizuno, H., van Meervelt, L., Hofkens, J., and Dedecker, P. (2014). Green-to-red photoconvertible Dronpa mutant for multimodal super-resolution fluorescence microscopy. ACS Nano *8*, 1664-1673.

Müller, B.K., Zaychikov, E., Bräuchle, C., and Lamb, D.C. (2005). Pulsed interleaved excitation. Biophys J *89*, 3508-3522.

Nir, E., Michalet, X., Hamadani, K.M., Laurence, T.A., Neuhauser, D., Kovchegov, Y., and Weiss, S. (2006). Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. J Phys Chem B *110*, 22103-22124.

Papanikolau, Y., Papadovasilaki, M., Ravelli, R.B., McCarthy, A.A., Cusack, S., Economou, A., and Petratos, K. (2007). Structure of dimeric SecA, the Escherichia coli preprotein translocase motor. J Mol Biol *366*, 1545-1557.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.

Rapoport, T.A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature *450*, 663-669.

Rasnik, I., McKinney, S.A., and Ha, T. (2006). Nonblinking and long-lasting single-molecule fluorescence imaging. Nat Methods *3*, 891-893.

Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule FRET. Nat Methods *5*, 507-516.

Sardis, M.F., and Economou, A. (2010). SecA: a tale of two protomers. Mol Microbiol *76*, 1070-1081.

Sardis, M.F., Tsirigotaki, A., Chatzi, K.E., Portaliou, A.G., Gouridis, G., Karamanou, S., and Economou, A. (2017). Preprotein Conformational Dynamics Drive Bivalent Translocase Docking and Secretion. Structure *25*, 1056-1067 e1056.

Schaffer, J., Volkmer, A., Eggeling, C., Subramaniam, V., Striker, G., and Seidel, C.A.M. (1999). Identification of single molecules in aqueous solution by time-resolved fluorescence anisotropy. J Phys Chem A *103*, 331-336.

Schrimpf, W., Barth, A., Hendrix, J., and Lamb, D.C. (2018). PAM: A framework for integrated analysis of imaging, single-molecule and ensemble fluorescence data. Biophys J *in press*.

Schuler, B., and Hofmann, H. (2013). Single-molecule spectroscopy of protein folding dynamics--expanding scope and timescales. Curr Opin Struct Biol *23*, 36-47.

Sianidis, G., Karamanou, S., Vrontou, E., Boulias, K., Repanas, K., Kyrpides, N., Politou, A.S., and Economou, A. (2001). Cross-talk between catalytic and regulatory elements in a DEAD motor domain is essential for SecA function. Embo Journal *20*, 961-970.

Singh, R., Kraft, C., Jaiswal, R., Sejwal, K., Kasaragod, V.B., Kuper, J., Burger, J., Mielke, T., Luirink, J., and Bhushan, S. (2014). Cryo-electron microscopic structure of SecA protein bound to the 70S ribosome. J Biol Chem *289*, 7190-7199.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol *185*, 60-89.

Talavera, A., Hendrix, J., Versees, W., Jurenas, D., Van Nerom, K., Vandenberk, N., Singh, R.K., Konijnenberg, A., De Gieter, S., Castro-Roa, D.*, et al.* (2018). Phosphorylation decelerates conformational dynamics in bacterial translation elongation factors. Sci Adv *4*, eaap9714.

Tessler, L.A., Reifenberger, J.G., and Mitra, R.D. (2009). Protein quantification in complex mixtures by solid phase single-molecule counting. Anal Chem *81*, 7141-7148.

Tomov, T.E., Tsukanov, R., Masoud, R., Liber, M., Plavner, N., and Nir, E. (2012). Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. Biophys J *102*, 1163-1173.

Tsirigotaki, A., Chatzi, K.E., Koukaki, M., De Geyter, J., Portaliou, A.G., Orfanoudaki, G., Sardis, M.F., Trelle, M.B., Jorgensen, T.J.D., Karamanou, S.*, et al.* (2018). Long-lived folding intermediates predominate the targeting-competent secretome. Structure (in press).

Tsirigotaki, A., De Geyter, J., Sostaric, N., Economou, A., and Karamanou, S. (2017). Protein export through the bacterial Sec pathway. Nat Rev Microbiol *15*, 21-36.

662 Vandenberk, N., Barth, A., Borrenberghs, D., Hofkens, J., and Hendrix, J. (2018). Evaluation of Blue
663 and Far-Red Dye Pairs in Single-Molecule FRET Experiments. J Phys Chem B.
664 Vogelsang, J., Kasper, R., Steinhauer, C., Person, B., Heilemann, M., Sauer, M., and Tinnefeld, P.
665 (2008). A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes.
666 Angew Chem Int Ed Engl *47*, 5465-5469.
667 Vosch, T., Cotlet, M., Hofkens, J., Van Der Biest, K., Lor, M., Weston, K.D., Tinnefeld, P., Sauer, M.,
668 Latterini, L., Müllen, K*., et al.* (2003). Probing Förster Type Energy Pathways in a First Generation
669 RIgid Dendrimer Bearing Two Perylene Imide Chromophores. J Phys Chem A *107*, 6920-6931.
670 Woodbury, R.L., Hardy, S.J., and Randall, L.L. (2002). Complex behavior in solution of homodimeric
671 SecA. Protein Sci *11*, 875-882.
672 Wowor, A.J., Yu, D., Kendall, D.A., and Cole, J.L. (2011). Energetics of SecA dimerization. J Mol Biol
673 *408*, 87-98.
674 Zander, C., Sauer, M., Drexhage, K.H., Ko, D.S., Schulz, A., Wolfrum, J., Brand, L., Eggeling, C., and
675 Seidel, C.A.M. (1996). Detection and characterization of single molecules in aqueous solution. Appl
676 Phys B-Lasers O *63*, 517-523.
677 Zimmer, J., Nam, Y., and Rapoport, T.A. (2008). Structure of a complex of the ATPase SecA and the
678 protein-translocation channel. Nature *455*, 936-943.
679 Zimmer, J., and Rapoport, T.A. (2009). Conformational flexibility and peptide interaction of the
680 translocation ATPase SecA. Journal of Molecular Biology *394*, 606-612.

681

682

683

Figure legends

Figure 1. Domain organization of SecA and PBD states in solved X-ray structures

E. coli SecA structures: the indicated conformational PBD states motions (Closed, Open, Wide-open), modeled (Chatzi et al., 2017) after the *E. coli*, Open model: 2FSF (open), seen by NMR and based on crystal structures, available on PDB (Closed: 3DIN; *Thermotoga maritima*, Wide-open: 1M6N; *Bacillus subtilis*). SecYEG would bind at the back of each structure in the translocase complex (**Figure S7**).

Figure 2. Principles of smFRET using MFD-PIE

A. Illustration of burst-wise single-molecule FRET using MFD-PIE (left; **Figures S1A and S1B**) with some of the information obtained from a single burst measurement and potential derived information relating to protein structure (right). **B.** SecA derivatives (D1-6) selected (**Figure S2A**) to optimally report on potential PBD motions towards the IRA2 or the WD domains. Grey cartoon: the 2nd unlabeled protomer of the dimer. **C.** Schematic representation of anticipated MFD-PIE results plotted on a 2D graph with showing measurements distributed in the cases of hypothetical static and dynamic FRET models. A protein with three conformational states: 'closed' (high FRET), intermediate and one 'open' (low FRET), results in three donor burst-averaged fluorescence lifetimes τ_1, τ_2 and τ_3 (red dot clouds), respectively. In the case of a slow interconversion between states (i.e. >10 ms) data clouds are positioned on the 'static FRET line' (solid red) and τ_1, τ_2 and τ_3 can be determined. Faster interconversions result in averaged data clouds along the dynamic FRET line (dashed green line), and only a fluorescence-weighted average lifetime can be observed (Gansen et al., 2009; Kudryavtsev et al., 2012).

Figure 3. smFRET-derived PBD structural states in dimeric SecA

A and **B.** Analysis of PBD conformational dynamics studied within the context of dimeric SecA. **Left**, Cartoons of the derivatives. **Middle**, 3,000-4,000 individual bursts/each from the 6 indicated derivatives were plotted on 2D plots of E vs $\tau_{D(A)}$. This lifetime ($\tau_{D(A)}$) is the burst-averaged fluorescence weighted lifetime of the donor in presence of the acceptor, integrated over the collected photons per burst. Contour plots display 2D histograms of molecule counts (red = high, blue = low counts). The 1D bar charts are projections of the 2D histograms on the respective axes. Static FRET lines (red) were calculated with Eq. 13. **Right**, PDA of D1 (**A**) and D3 (**B**) for the dimeric condition in a global fit (STAR Methods and **Figure S6**). Uncorrected proximity ratio histograms (gray bars). Black stairs: total PDA model, colored dashed stairs: subpopulations. w_{res} = weighted residuals (top graph). Corresponding distance distribution plots illustrating the intricate relation of distance and FRET distribution width. Black solid line: total model, coloured dashed lines: sub-states.

C. smFRET-determined PBD States in soluble dimeric SecA summarized in a table with the fraction A (%) of each population distribution and the derived distances (Å) after PDA analysis in 3 possible states. For additional data on D2 and D5 see **Figures S4 and S6**.

Figure 4. smFRET-derived PBD states in monomeric SecA

A and **B.** Analysis of PBD dynamic conformational behaviour when studied within the context of monomeric SecA (as in **Figures 3A and 3B).** Summarized PBD physical motions derived from PDA in monomeric SecA (as in **Figure 3C**).

Figure 5. ADP-independent PBD motions in soluble SecA

A and **B.** Nucleotides have limited effect on PBD motions in soluble SecA. Effect of ADP and ATP on dimeric (above) and monomeric (below) (K_d see **Data S1**) SecA showed for D1 (**A**) and D3 (**B**). Left, Dimeric SecA: 200pM labeled SecA with addition of 100nM unlabeled SecA. Middle, Dimeric SecA with ADP: 200pM labeled SecA with addition of 100 nM unlabeled SecA, 10 μM ADP and 50 μM $MgCl_2$**.** Right, Dimeric SecA with ATP: 200 pM labeled SecA with addition of 100 nM unlabeled SecA, 10 μM ATP and 50 μM $MgCl_2$**.**

Figure 6. Visualization models of the smFRET-derived PBD states of SecA

Zoomed-in views (defined by the square in the 4-coloured SecA surface structure; top) of the 4 different states of the PBD (ribbon; coloured as indicated) modelled according to the smFRET-derived distance restraints (**Table S5**). The modelled PBDs are overlaid to those of the already solved SecA structures (dark grey; 3DIN in state 1 - closed; 2VDA in state 2 - open and state 4; 1M6N in state 3 - wide open). The body of SecA, including parts of IRA2 and WD, is shown as a white surface and the Stem antiparallel β-strands are indicated.

Figure 7. Hypothetical models of PBD function during translocation

Three hypothetical models of how PBD swiveling (magenta) might mechanistically contribute to preprotein (orange) translocation through SecYEG (see also **Figure S7**).

STAR methods:

CONTACT FOR REAGENT AND RESOURCE SHARING

Additional information or requests for resources and reagents should be directed to the Lead Contact, Anastassios Economou (tassos.economou@kuleuven.be).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

For protein purification, *E. coli* BL21 or T7 express *lys*Y/I^{q} cells transformed with pET3a plasmids carrying *secA* (P10408) or derivatives were grown in 5 L flasks (LB 2.5 L; 30 °C; OD_{600} 0.5-0.6). In each case, gene expression was induced (0.2 mM IPTG; 3 h; 30 °C. Cells were collected (5,000 g; 4 °C; 15 min; Avanti J-26S XPI, JLA 8.1000 rotor; Beckman), resuspended in 50 mM Tris/HCl pH 8; 1 M NaCl; 5% (V/V) glycerol; lysed by using a French press (8,000 psi; 55,16 MPa); 3-5 passes; pre-cooled cylinder at 4 °C; Thermo).

METHOD DETAILS

Buffers and reagents

Tris buffer consists of 50 mM Tris (Sigma-Aldrich, Diegem, Belgium) and 50 mM NaCl (Sigma-Aldrich, Diegem, Belgium) at pH 7. Aged PBS/Trolox buffer was made by dissolving 1 mM Trolox (6-Hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid, 543353, Sigma Aldrich, Denmark) overnight (~16 h) at 4 °C in the buffer. Prior to experiments, coverslips (Nunc Lab-Tek Chambered cover-glass, 155411, ThermoFisher Scientific, Ghent, Belgium) were coated with 1 mg/mL bovine serum albumin (BSA, Sigma-Aldrich, Diegem, Belgium). The BSA stock solution was made by dissolving 10 mg/mL BSA in PBS, passing the solution through a 0.45 µm filter (Reg. cellulose 0.45 µm, 5123260(K45), Grace discovery sciences, Deerfield, IL,

779 USA.) the solution and subsequent flash freezing aliquots that were stored at -20 °C.
780 The free dyes that were used for microscope calibration are Atto488 (ATTO 488-CA,
781 ATTO-TEC GmbH, Siegen, Germany) and Atto655 (ATTO 655-CA). The dyes that
782 were used for labeling the protein are ATTO 488 maleimide (ATTO-TEC GmbH,
783 Siegen, Germany) and Alexa Fluor 647 C_2-maleimide (Life Technologies Europe BV,
784 Gent, Belgium). Dye properties are summarized in **Table S1**.

785 Residue Selection

786 Six selection criteria were applied for specific residue selection. These residues
787 should: (a) localize on secondary structure elements (α-helix or β-sheet). This way,
788 dye to dye distances would report on protein conformational changes and not local
789 wobble effects. (b) have good solvent accessibility. As dyes are not infinitely small,
790 labeling of surface residues increases labelling efficiency and reduces steric
791 interference. Accessible volumes were calculated using the FRET-restrained
792 positioning and screening (FPS) software (Kalinin et al., 2012) to build models of the
793 expected FRET distances in soluble SecA and to model dye motion in space at specific
794 residues (**Figure S1D**). (c) not participate in hydrogen bonding or interactions with
795 neighboring residues. This avoids placing the dyes on residues with direct structural
796 roles. (d) not participate in the SecA dimerization or the SecY-binding or the preprotein
797 binding interfaces (e) be located such that the change in FRET between 2 residues
798 from one PBD conformation to another (Figure 1A) is high. FRET changes can be
799 calculated from the actual space-averaged distance between the dyes, attached to a
800 cysteine residue via a ~20 Å linker. (f) be located such that the attached dyes do not
801 collide or sterically hinder one another. Using the above criteria, we selected 31 FRET
802 pairs (**Table S2**).

803 Strains, genetic manipulations and mutagenesis

We used the fully functional SecA(cys$^-$) that has its 4 cysteinyl residues substituted (Chatzi et al., 2017; Sardis et al., 2017) (positions 98, 885, 887 and 896) substituted by serine (C98S) or alanines (885, 887 and 896). Gene-mutations were introduced by following the QuickChange Site-Directed Mutagenesis protocol (Stratagene-Agilent); templates and primers are listed in Supplementary **Table S3** and **S4**. For PCR mutagenesis PFU Ultra Polymerase (Agilent) was used; DpnI was used to cleave the maternal methylated DNA (Promega). All PCR-generated plasmids were sequenced (Macrogen, Europe). Plasmids were stored in DH5α cells.

Protein expression, purification and functional assays

Gene-overexpression was induced using BL21 (DE3; NEB C2527) (Studier et al., 1990) or T7 express *lys*Y/I^q (DE3; NEB C3013). Bacteria growth conditions, purification of SecA or mutant derivatives, size exclusion chromatography in line with MALS, *in vivo* complementation of with *secA* or derivatives and SecA ATP hydrolysis activity measurement *in vitro* was performed as previously described (Chatzi et al., 2017; Gouridis et al., 2013; Karamanou et al., 2007).

Fluorescent Labeling

Oxygen was removed from the buffer (50 mM Tris pH 7.0, 50 mM NaCl, termed 'Tris buffer') under vacuum and continuous stirring. TCEP was removed from the protein solution by gel filtration (PD-10 desalting columns, GE Healthcare Europe GmbH, Diegem, Belgium) and the protein was concentrated to at least 30 μM by ultrafiltration (Nominal molecular weight limit = 50 kDa Amicon Ultra-0.5, Merck Chemicals N.V., Overijse, Belgium) at 14,000 g and 4 °C. An equal molar amount (60 μM) of ATTO 488 maleimide (ATTO-TEC GmbH, Siegen, Germany) and Alexa Fluor 647 C_2-maleimide (Life Technologies Europe BV, Gent, Belgium) was mixed, the protein was added to a final concentration of 50 μM and samples were kept overnight at 4 °C. Free dye was

removed by gel filtration and ultrafiltration. A representative labeling result is presented in **Figure S3C** (DTT condition). The ATTO488-maleimide typically exhibited 30-50 % higher labeling efficiency than the Alexa Fluor 647-maleimide. This might be attributed to the dye's differences in charge (ATTO 488: 1+; Alexa Fluor 647: 4-/1+), size (ATTO 488: 710Da; Alexa Fluor 647: ~1250Da) or structure. Typically, 20-50 % of the proteins were labelled by both dyes. According to our experience, protein concentration prior to labeling, the batch of the dye, the degree of reduction of the Cys-SH moiety and the time between protein purification and labelling, all influenced this percentage. However, since FRET experiments were carried out using alternating FRET donor and acceptor excitation, the presence of both dyes can be verified per passing molecule; thus, the resulting FRET histograms represent the 100% of doubly labeled molecules. Tris buffer containing 50% (V/V) glycerol was added 1:1 and the labeled protein sample was divided in aliquots and the samples were stored at -20 °C.

Confocal multi-parameter setup

For all ensemble and single-molecule experiments, a home-built multi-parameter fluorescence detection microscope with pulsed interleaved excitation (MFD-PIE, (Kudryavtsev et al., 2012)) was used (see scheme of the setup in **Figure S1A)**. Two lasers were used: a pulsed 483-nm laser diode and a 635-nm laser diode, alternated at 26.67 MHz and delayed ~18 ns with respect to each other. The power density inside the focus was calculated via:

$$\text{Power density}\,[\text{kW/cm}^2] \;=\; 0.04\,\frac{P}{\pi\omega_r^2}, \qquad \text{Eq. 1}$$

where *P* is the laser power (in µW) measured in between the excitation polychroic mirror and the objective lens and ω_r is the lateral focus waist (in µm). Furthermore, the equation assumes that 40% of the measured light reaches the sample. Sample

emission was transmitted through a pinhole and spectrally split. Both, the blue range and red range were split by polarization on two detection channels. Photons were detected on four avalanche photodiodes: $B_{\parallel}$ (blue-parallel), $B_{\perp}$ (blue-perpendicular), $R_{\parallel}$ (red-parallel) and $R_{\perp}$(red-perpendicular) (Figure S1B), which were connected to a time-correlated single photon counting (TCSPC) device. Signals from each TCSPC channel were divided in time gates (Lamb et al., 2000a) to discern 483-nm excited FRET photons from 635-nm excited acceptor photons: $BB_{\parallel}$, $BB_{\perp}$, $BR_{\parallel}$, $BR_{\perp}$, $RR_{\parallel}$, $RR_{\perp}$ (Figure S1B). Microscope alignment (excitation light guiding, objective lens correction collar, pinhole, detectors) and determination of the lateral (ω_r) and axial (ω_z) focus waists were done using real-time fluorescence correlation spectroscopy (FCS, see further) on freely diffusing ATTO488-COOH and ATTO655-COOH in water. For more details about the used equipment the reader is referred to (Vandenberk et al., 2018).

FRET measurements

The labeled protein was diluted in Tris buffer containing 1 mM aged Trolox (Cordes et al., 2009) up to a concentration of 100-200 pM. Trolox efficiently prevented the protein from adsorbing to the sample holder, thereby decreasing the overall measurement time, yet had no influence on the functionality of the molecule, corroborated by identical FRET histograms in absence or presence of Trolox (**Figure S2G**, right). Coverslips (Nunc Lab-Tek Chambered Coverglass, ThermoFisher Scientific BVBA, Erembodegem, Belgium) were pre-coated with 1 mg/mL bovine serum albumin (BSA) and washed twice with the sample solution, after which 30 µL of the sample solution was added. Adding other agents to the sample solution such as BSA (Tessler et al., 2009) or the non-ionic detergent, n-Dodecyl-β-BD-maltopyranoside (DDM) (Huang et al., 2005) were also tested, but a significant improvement was only seen for Trolox and

DDM. Various coatings of coverslips (with Fibronectin, Laminin, Collagen or BSA) were also tested without improvement compared to BSA (data not shown). Unless explicitly stated otherwise, smFRET experiments were performed during at least 1 hour at 100 µW 483 nm and 50 µW of 635 nm excitation at room temperature (22 °C). Background and scatter information was obtained via a buffer measurement under identical condition. The background/scatter information is needed for obtaining absolute *E* (Eq. 2) and *S* (Eq. 3) parameters, but also for correct lifetime and PDA analysis (described later). Unless explicitly stated otherwise, all burst measurements were performed during at least 1 hour at 100 µW of 483 nm excitation and 50 µW of 635 nm excitation.

Accessible volume simulation

The geometry of the donor and acceptor dyes, including the linker length (measured from the C5 of Thymidine to the geometrical center of the dye), linker width and 3D radius was obtained from the FPS manual (Kalinin et al., 2012). The different parameters are summarized in **Table S1**. The FPS tool (Kalinin et al., 2012) was used to simulate the accessible volume per dye in the context of the actual dsDNA, using standard settings (*i.e.* search nodes = 3, clash tolerance = 1.0 Å). This information, together with R_0 (54.7 Å), was used to estimate the simulated FRET averaged D/A distance, $\langle R_{DA} \rangle_E$.

Generation of PBD states in SecA and SecA-SecY models

Starting structures used to generate the SecA models were 3DIN, 2VDA and 1M6N, which correspond to close, open and wide-open states of the PBD, respectively. The PBD (230-370 a.a.) from 2VDA was used and treated as rigid body to model the

different smFRET states (state 1, 2, 2A and 3). To generate the state 2A and satisfy the smFRET distances, the PBD and part of the Stem had to be used and re-positioned (216-377 a.a.). All the PBDs movements and re-localization were restrained based on the smFRET probes distances in PyMol (The PyMOL Molecular Graphics System, Version 2.0.7 Schrödinger, LLC). To obtain a single set of coordinates, the PBDs were merged to the stem using ModLoop (Fiser and Sali, 2003; Fiser and Simon, 2000). The 3DIN structure from *Thermotoga maritima* was used to model the closed helicase ATPase motor on *Escherichia coli* SecA based on superposition. Similarly, the 3DIN structure was also used to generate the SecA-SecY complexes, where the SecA is the monomeric State 2A aligned to the 3DIN SecA in complex with SecY. In this model, the 3DIN SecY remains unaltered. Alternatively, and based on superposition with the latter model, a SecA-SecY model comprising *E. coli*-only components was generated. The 5GAE SecY structure was used. All the models were energy minimized using Chimera (Pettersen et al., 2004).

QUANTIFICATION AND STATISTICAL ANALYSIS

Software

All simulations and analyses of experimental data were performed in the software package PAM (Schrimpf et al., 2018). The software is available as source code, requiring MATLAB to run, or as pre-compiled standalone distributions for Windows or MacOS at http://www.cup.uni-muenchen.de/pc/lamb/software/pam.html or hosted in Git repositories under http://www.gitlab.com/PAM-PIE/PAM and http://www.gitlab.com/PAM-PIE/PAMcompiled. Sample data is provided under http://www.gitlab.com/PAM-PIE/PAM-sampledata. A detailed manual is found under http://pam.readthedocs.io.

927 Single-molecule burst analysis

928 *Burst identification* - For single-molecule data, a two-color MFD all-photon burst search
929 algorithm (Nir et al., 2006) using a 500 µs sliding time window (min. 50 photons per
930 burst, min. 5 photons per time window) and a kernel-density estimator
931 (ALEX-2CDE < 15, (Tomov et al., 2012)) were used to identify single donor-acceptor
932 labeled molecules in the fluorescence trace. Data was further thresholded using a $|T_{fret}-$
933 $T_{red}|<0.07$, to remove bursts (typically 10%) exhibiting photobleaching during molecule
934 passage (Kudryavtsev et al., 2012). Additionally, a 0-20-ms burst duration cut was
935 applied to remove sparse (< 1%) slow moving long aggregates, since these can
936 significantly bias time window based analyses such as PDA.
937 *FRET efficiency* – The absolute burst-averaged FRET efficiency *E* was calculated with:
938

$$E = \frac{F_{BR}-\alpha F_{RR}-\beta F_{BB}}{F_{BR}-\alpha F_{RR}-\beta F_{BB}+\gamma F_{BB}}, \qquad \text{Eq. 2}$$

939 where $F_{BR} = S_{BR} - B_{BR}$ is the background corrected number of photons in both red
940 detection channels after blue excitation (with S_{BR} and B_{BR} the summed intensity and
941 background, respectively, in time gates $BR_{\parallel}$ and $BR_{\perp}$); $F_{BB} = S_{BB} - B_{BB}$ the background
942 corrected number of photons in the blue detection channel after blue excitation (with
943 S_{BB} and B_{BB} the summed intensity and background, respectively, in time gates $BB_{\parallel}$ and
944 $BB_{\perp}$), $F_{RR} = S_{RR} - B_{RR}$ the background corrected number of photons in the red detection
945 channel after red excitation (with S_{RR} and B_{RR} the summed intensity and background,
946 respectively, in time gates $RR_{\parallel}$ and $RR_{\perp}$), α a correction factor for direct excitation of
947 the acceptor with the 483 nm laser, β a correction factor for emission crosstalk of the
948 donor in the acceptor channel, and γ the relative detection efficiency of the donor and
949 acceptor (Kudryavtsev et al., 2012).

Stoichiometry - The corrected stoichiometry ratio *S* was calculated with:

$$S = \frac{F_{BR} - \alpha F_{RR} - \beta F_{BB} + \gamma F_{BB}}{F_{BR} - \alpha F_{RR} - \beta F_{BB} + \gamma F_{BB} + F_{RR}}, \quad \text{Eq. 3}$$

resulting in the ratio of the blue laser excited photons over all excited photons (blue and red laser). According to this calculation, D-only labeled molecules will have *S* values near unity, while A-only labeled molecules will have values near zero. Double labeled-molecules will exhibit *S* values between 0.2-0.6 depending on the used dya pair, the microscope and the laser power ratio.

Data correction – First, background was subtracted from the experimental signals. Then, the *β*- and *α*-factors were determined directly from the measurement (Kudryavtsev et al., 2012) and data was corrected. Finally, the center values of the *E*-*S* data cloud for each protein were estimated manually, plotted in an *E* vs. 1/*S* graph, and a straight line was fitted to the resulting data to obtain the *γ*-factor:

$$\gamma = \frac{\Omega - 1}{\Omega + \Sigma - 1}, \quad \text{Eq. 4}$$

where *Ω* is the intercept and *Σ* the slope of the linear fit.

Distances – FRET-averaged D/A distances (Kalinin et al., 2012) were obtained from the center *E* values with:

$$\langle R_{DA} \rangle_E = R_0 \left(\frac{1-E}{E} \right)^{1/6}, \quad \text{Eq. 5}$$

where R_0 is the Förster distance (54.7 Å), that was calculated using the measured dye spectra (Vandenberk et al., 2018), a refractive index n = 1.33, an orientation factor κ^2 = 2/3, a measured donor quantum yield Φ = 0.6 for Atto488 and acceptor extinction

coefficient (ε = 265,000 $cm^{-1}M^{-1}$) (**Table S1**). The quantum yield was determined using a home-built absorbance/fluorescence spectroscope (Moeyaert et al., 2014). For simplicity, $\langle R_{DA} \rangle_E$ will be noted *R* throughout the text.

Burstwise fluorescence lifetime - A maximum likelihood estimator approach (MLE, (Schaffer et al., 1999)) was used to estimate single-molecule burst-averaged single-exponential fluorescence lifetimes of the FRET donor, $\tau_{D(A)}$, and FRET acceptor, τ_A. For molecules that are conformationally static during transit through the laser focus, the FRET efficiency is related to the fluorescence lifetime of the donor as follows:

$$E_{static} = 1 - \frac{\tau_{D(A)}}{\tau_D}. \qquad \text{Eq. 6}$$

However, dyes are attached to the molecule of interest via flexible dye linkers, resulting in a Gaussian D/A distance distribution, even for conformationally static molecules. Especially at short distances (high FRET), this effect causes a non-linear relation between the intensity-based *E* and $\tau_{D(A)}$. We simulated this 'static FRET line including linker dynamics' as follows: we calculated *m* values for *R* between 0 and 3×R_0. For every *R*, we calculated a Gaussian distribution of *p* distances around the central *R*, with the apparent linker length as the standard deviation, resulting in a list of *m*p* values for *R*. For every *R*, we calculated which donor fluorescence lifetime would be associated with it (Eq. 6, with τ_D the mean burstwise lifetime of raw burst data with *S* > 0.8). The apparent linker length (6 Å) was obtained from a sub-ensemble donor fluorescence lifetime fitting of double-labeled molecules using a gaussian distance distribution model. Finally, we calculated the species-weighted average lifetime, and from it the intensity-based *E* (the *y*-axis of the static FRET line) and the intensity-weighted average lifetime (*x*-axis of the static FRET line).

Similarly, for molecules exhibiting multiple lifetimes during transit due to conformational FRET dynamics, the burst-averaged lifetime will be fluorescence-weighted towards the long-lifetime species that emits more photons, resulting in an even further rightward shift of the experimental data from the theoretical line (Eq. 6).

Burstwise steady-state fluorescence anisotropies of the FRET donor (r_D) and FRET acceptor (r_A) were calculated from the respective fluorescence intensities:

$$r = \frac{GF_{\parallel} - F_{\perp}}{GF_{\parallel} + F_{\perp}}, \qquad \text{Eq. 7}$$

where *G* is the correction factor for the different detection efficiencies in the two polarization channels, $F_{\parallel}$ the intensity in time gate $BB_{\parallel}$ or $RR_{\parallel}$ and $F_{\perp}$ the intensity in time gate $BB_{\perp}$ or $RR_{\perp}$. Perrin equations were calculated with:

$$r = \frac{r_0}{(1+\tau/\theta)}, \qquad \text{Eq. 8}$$

where *r* is the single molecule steady state anisotropy, r_0 = 0.4 the fundamental anisotropy, τ the fluorescence lifetime and θ the rotational correlation time.

<u>Fluorescence correlation spectroscopy (FCS)</u>

Raw data FCS was performed by cross-correlating parallelly and perpendicularly polarized photon streams from any combination of time gates. Subensemble FCS was performed by selecting particular subpopulations in burst space, correlating each burst after adding 10 ms of data before and after, and averaging the resulting data (Laurence et al., 2007). Burstwise diffusion times τ_{diff} were obtained by fitting individual burst correlations with:

$$G(\tau) = \left(1 + \frac{\tau}{\tau_{diff}}\right)^{-1}, \qquad \text{Eq. 9}$$

from which the diffusion coefficient *D* (μm²/s) was calculated:

$$D = \frac{\omega_r^2}{4\tau_D}. \quad \text{Eq. 10}$$

Photon distribution analysis

Photon distribution analysis (PDA) provides a complete statistic description of single-molecule burst experiments, allowing to discern between molecular conformational heterogeneity and any other effects (shot noise, acceptor photophysics, background…) that broaden experimental FRET histograms (Antonik et al., 2006). Here, we used an implementation that models a sum of gaussian distance distributions to the experimental data (Antonik et al., 2006). Practically, burst data was binned into constant time bins (0.2-1 ms) and first thresholded in E_{PR} (Eq. 11) vs. S_{PR} (uncorrected stoichiometry) space to remove bins with complex acceptor photophysics or photobleaching. Then, only bins with at least 20 (for displaying purposes) and maximally 250 photons (to reduce calculation time) were analyzed. For displaying purposes, uncorrected proximity ratio histograms were used:

$$E_{PR} = \frac{S_F}{S_D + S_F}, \quad \text{Eq. 11}$$

where S_D and S_F are the raw photon counts in the donor and FRET channels, respectively. For PDA analysis, data was γ- (~0.8), β- (~0.01) and direct acceptor excitation (~0.01) corrected, and background (0-1.5 kHz) was explicitly taken into account. Correction parameters were determined as described previously (Kapanidis et al., 2004; Kudryavtsev et al., 2012). Unless explicitly stated otherwise, only the 1-ms binned data was used for PDA. Model parameters were optimized using a reduced-χ^2-guided simplex search algorithm. The resulting parameters were the mean FRET-averaged distance R and standard deviation (σ_R) of each Gaussian distributed substate, and, in the case of multiple states, their area fraction A (%). Where possible, different datasets were analyzed simultaneously by optimizing relevant parameters

(e.g. area fraction) globally over all datasets to increase fitting robustness. The reader is referred to **Figure S5 and Data S2** for an exemplary analysis. Finally, the standard deviation of the distance distributions was globally optimized as a fraction $\mathcal{F}$ of the corresponding distance to further improve fitting robustness ($\sigma_R \approx 0.12 \times R$), which has been shown before to be reasonable for FRET experiments with a blinking FRET acceptor (Kalinin et al., 2008; Kalinin et al., 2012). We have validated this global fitting approach experimentally before with a dataset of nine conformational static dsDNA molecules with different D-A distances (Vandenberk et al., 2018). Interestingly, relative to these control experiments, we did notice a slightly larger $\mathcal{F}$ value for SecA (data not shown), which could be indicative of fast exchange dynamics. Criteria for a good fit were a low (< 3) reduced χ^2 value, as well as a weighted residuals plot free of trends. The uncertainty on *A* was calculated as the standard deviation from at least three independent experiments. The uncertainty on *R* (**Figure 3C** and **4C**) was calculated in two ways: (i) via error propagation using partial derivatives of Eq. 5, the uncertainty on E (as determined using a γ-factor 0.7-0.9) and the uncertainty on R_0 (as determined before for the same dye pairs (Vandenberk et al., 2018)) as input, and (ii) via the standard deviation on *R* between at least three independent experiments. The reported errors in Figure 3 and 4 are the root mean squares of both values. To display the result, the gaussian substates and their sum was plotted onto the experimental E_{PR} histogram. Probability density functions (PDF) were additionally calculated per state using the *A*, *R* and σ_R parameters obtained from PDA. The summed PDF was scaled to a total area of unity, with each state's PDF area corresponding to the fraction of molecules in that state.

KEY RESOURCE TABLE

See external document.

Supplemental information

Supplemental information includes seven figures, five tables, one data file and can be found with this article.

Supporting Citations

The following references appear in the Supplemental information: (Brewer and Riehm, 1967; Chatzi et al., 2017; Gelis et al., 2007; Jomaa et al., 2016; Kalinin et al., 2012; Kudryavtsev et al., 2012; Lamb et al., 2000b; Papanikolau et al., 2007; Sardis and Economou, 2010; Talavera et al., 2018; Zimmer et al., 2008)

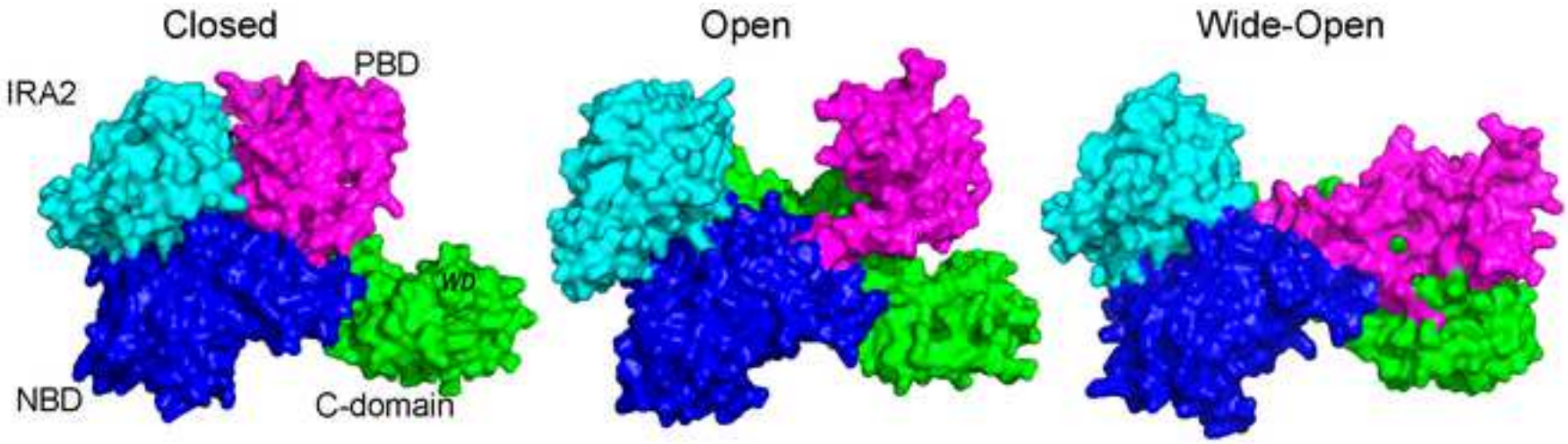
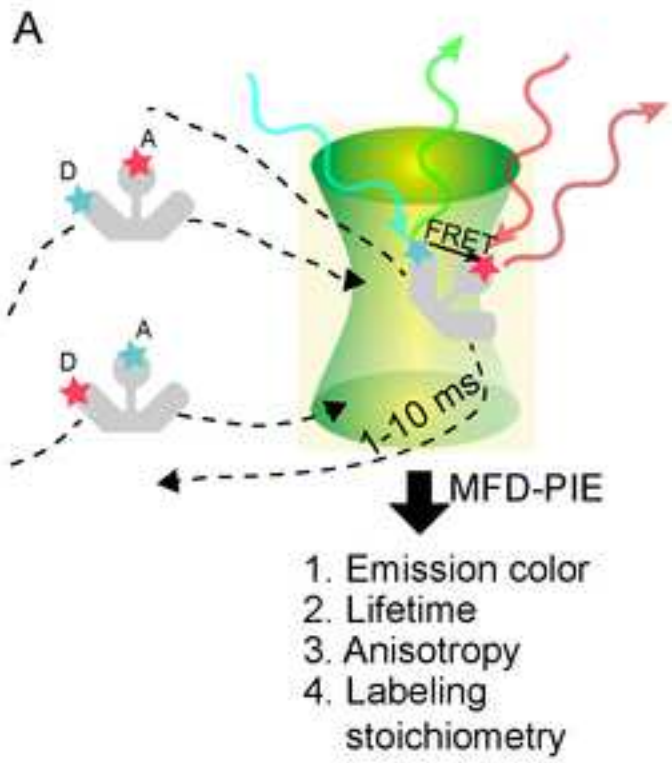
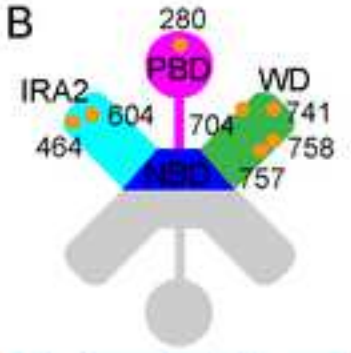


Fig. 1; Vandenberk et al.





Nr	Residues
D1	V280C/S604C
D2	V280C/L464C
D3	V280C/E757C
D4	V280C/A741C
D5	V280C/M758C
D6	V280C/P704C

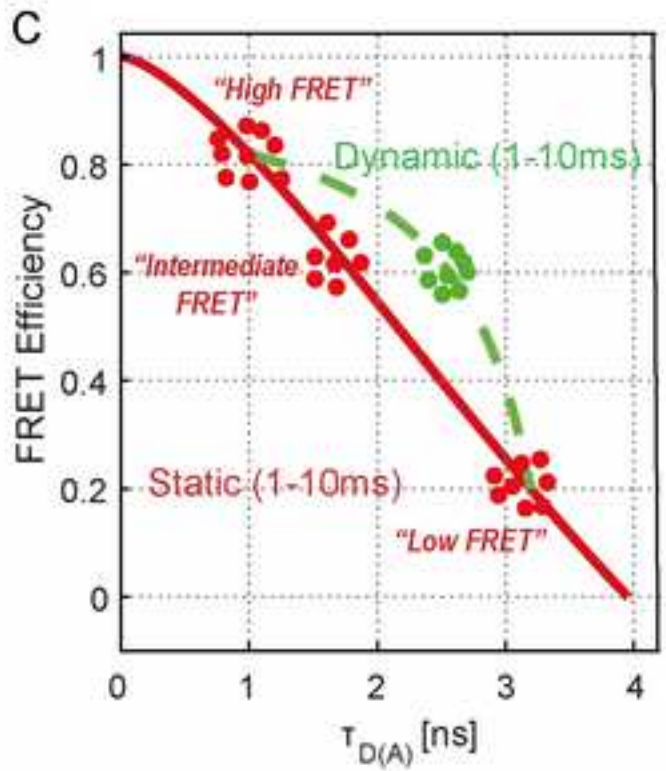
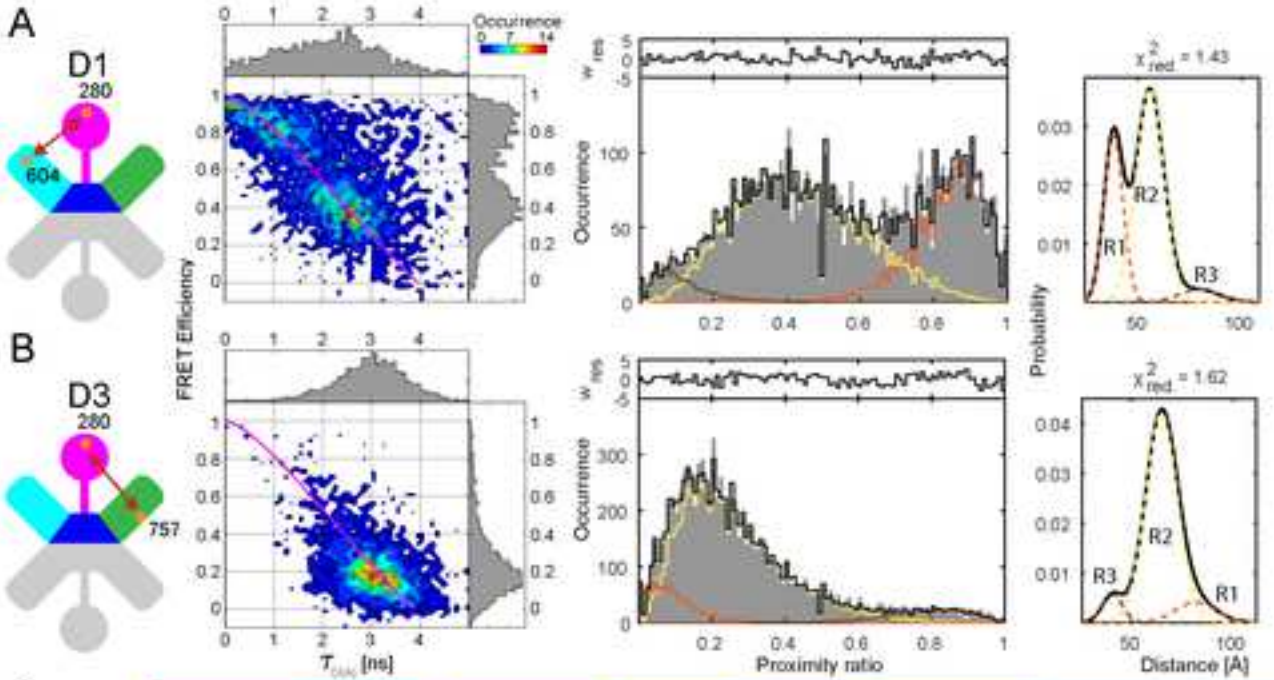


Fig. 2; Vandenberk et al.



C

	state 1				state 2				state 3			
SecA derivative	A_1 (%)	σ_{A1}	R_1 (Å)	σ_{R1}	A_2 (%)	σ_{A2}	R_2 (Å)	σ_{R2}	A_3 (%)	σ_{A3}	R_3 (Å)	σ_{R3}
D1	37.0	7.4	38.1	2.7	56.0	8.6	55.6	2.6	6.0	1.2	79.4	3.2
D2	49.0	1.5	38.7	2.5	40.0	0.8	50.2	2.5	11.0	2.3	70.4	2.8
D3	9.0	2.2	81.6	8.2	82.0	0.7	64.9	2.4	8.0	2.2	40.9	2.5
D5	5.0	0.6	86.4	3.6	88.0	1.9	61.7	2.5	8.0	2.2	39.2	2.2

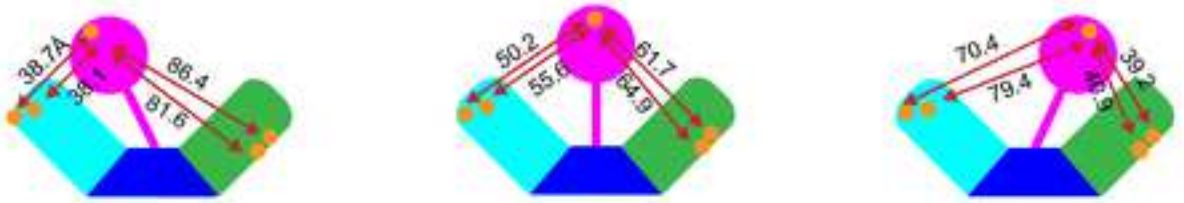
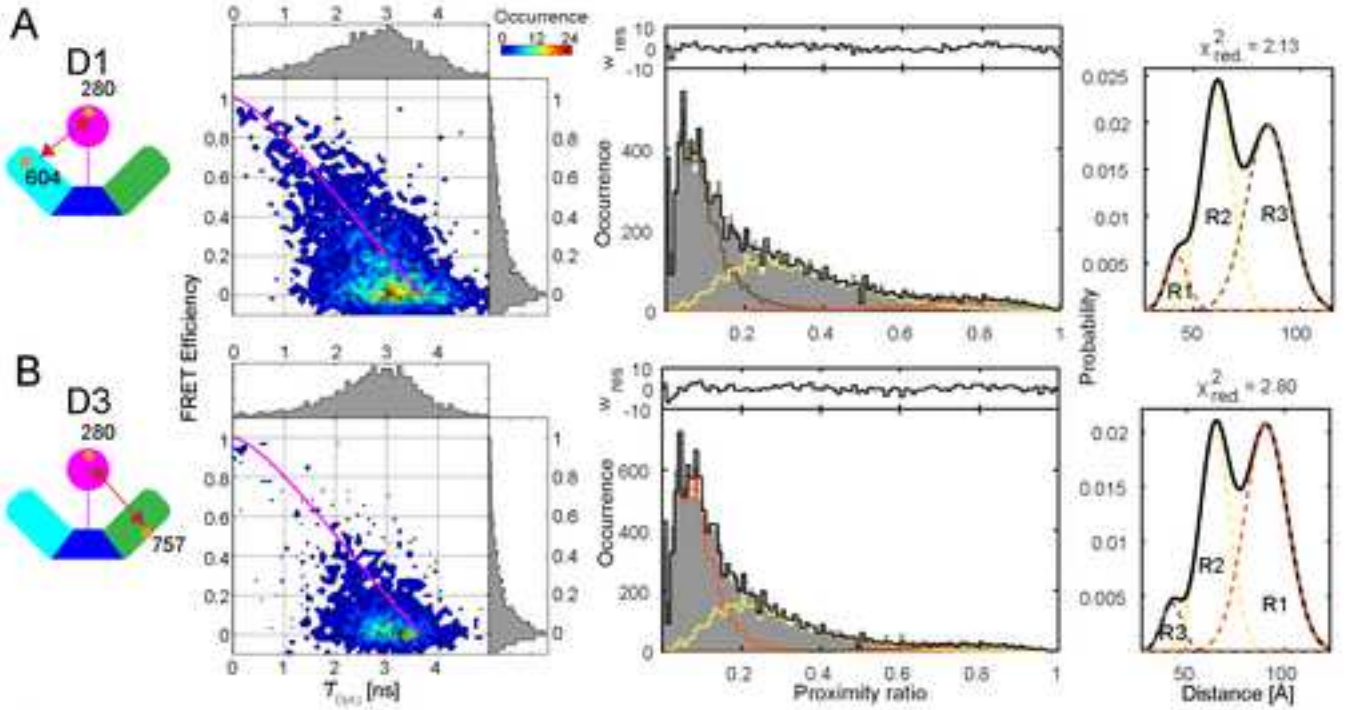


Fig. 3; Vandenberk et al.



C

	state 1				state 2				state 2A		state 3			
SecA derivative	A_1 (%)	σ_{A1}	R_1 (A)	σ_{R1}	A_2 (%)	σ_{A2}	R_2 (A)	σ_{R2}	$A_1^{D1,D2}$ $A_3^{D1,D3}$ (%)	$R_1^{D1,D5}$ $R_3^{D1,D3}$ (A)	A_3 (%)	σ_{A3}	R_3 (A)	σ_{R3}
D1	6.0	2.0	42.1	1.7	37.0	5.7	60.7	2.6	~51.0	85.1	~6.0	7.8	85.1	3.1
D2	7.0	3.7	39.8	3.0	49.0	2.4	58.2	3.5	~37.0	80.6	~6.0	6.0	80.6	3.3
D3	~7.0	5.2	90.6	4.9	44.0	4.9	64.5	3.4	~43.0	91.0	6.0	0.5	42.9	4.6
D5	~7.0	3.8	90.3	3.2	37.0	2.7	65.6	2.7	~52.0	90.0	4.0	1.4	43.8	5.1

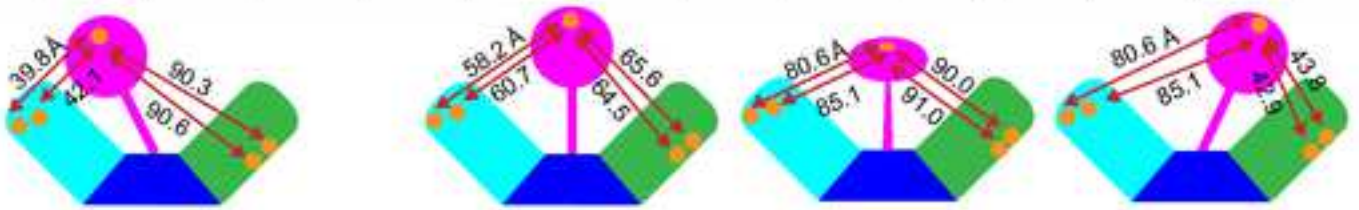


Fig. 4; Vandenberk et al.

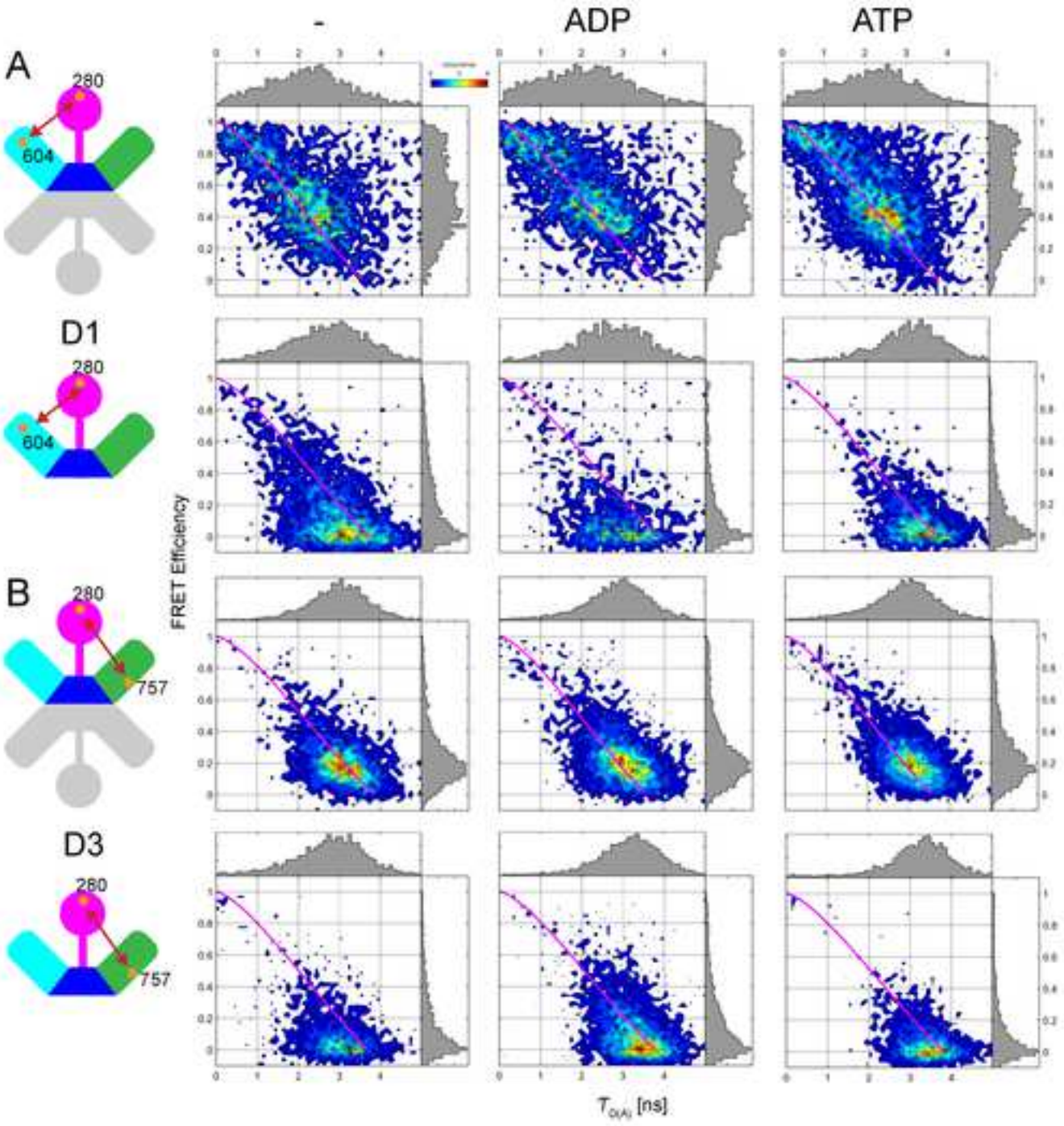


Fig. 5; Vandenberk et al.

Figure6

Click here to access/download;Figure;Fig6new_new_2ndsubmission_final2.tif

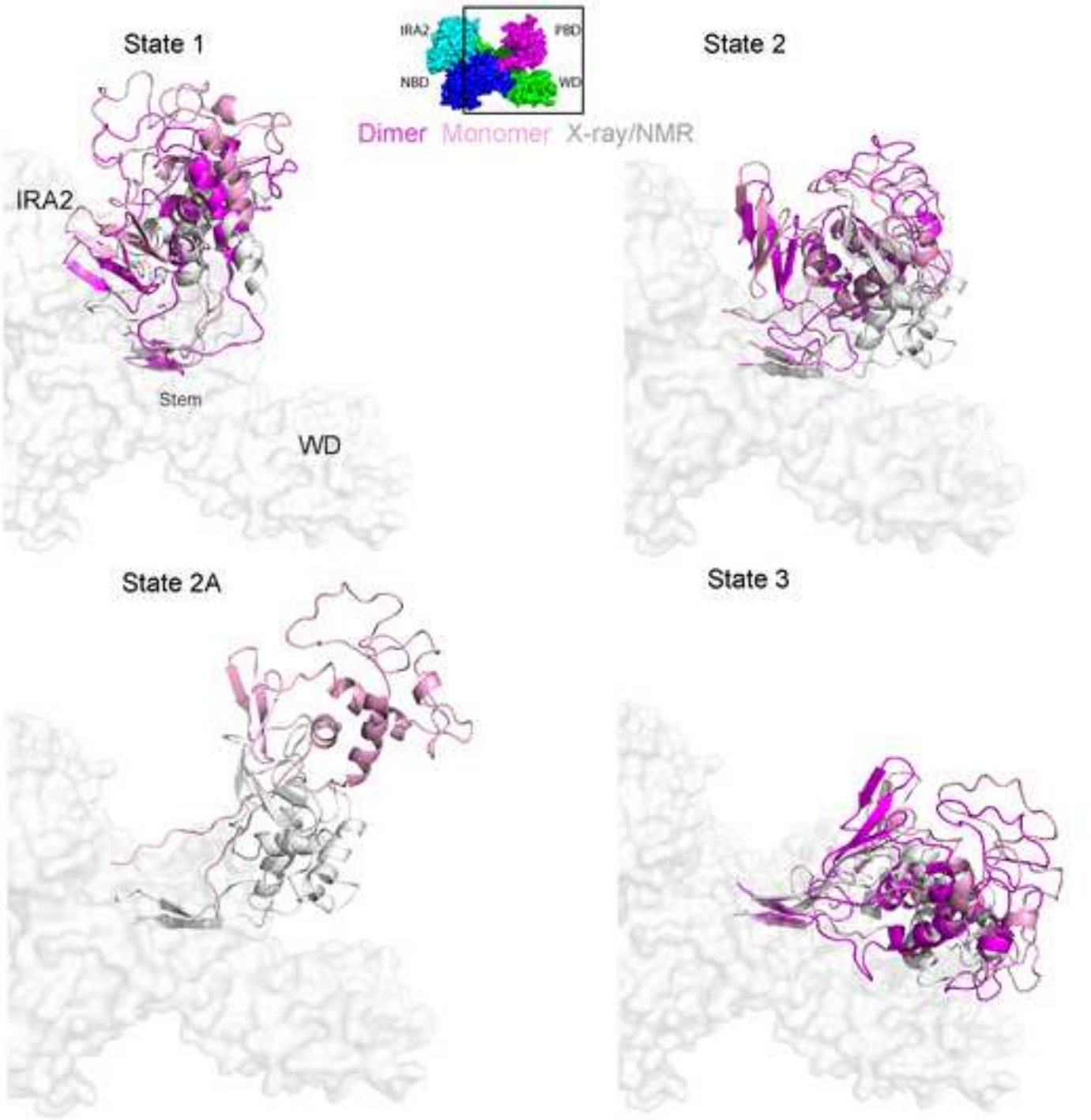


Fig. 6; Vandenberk et al.

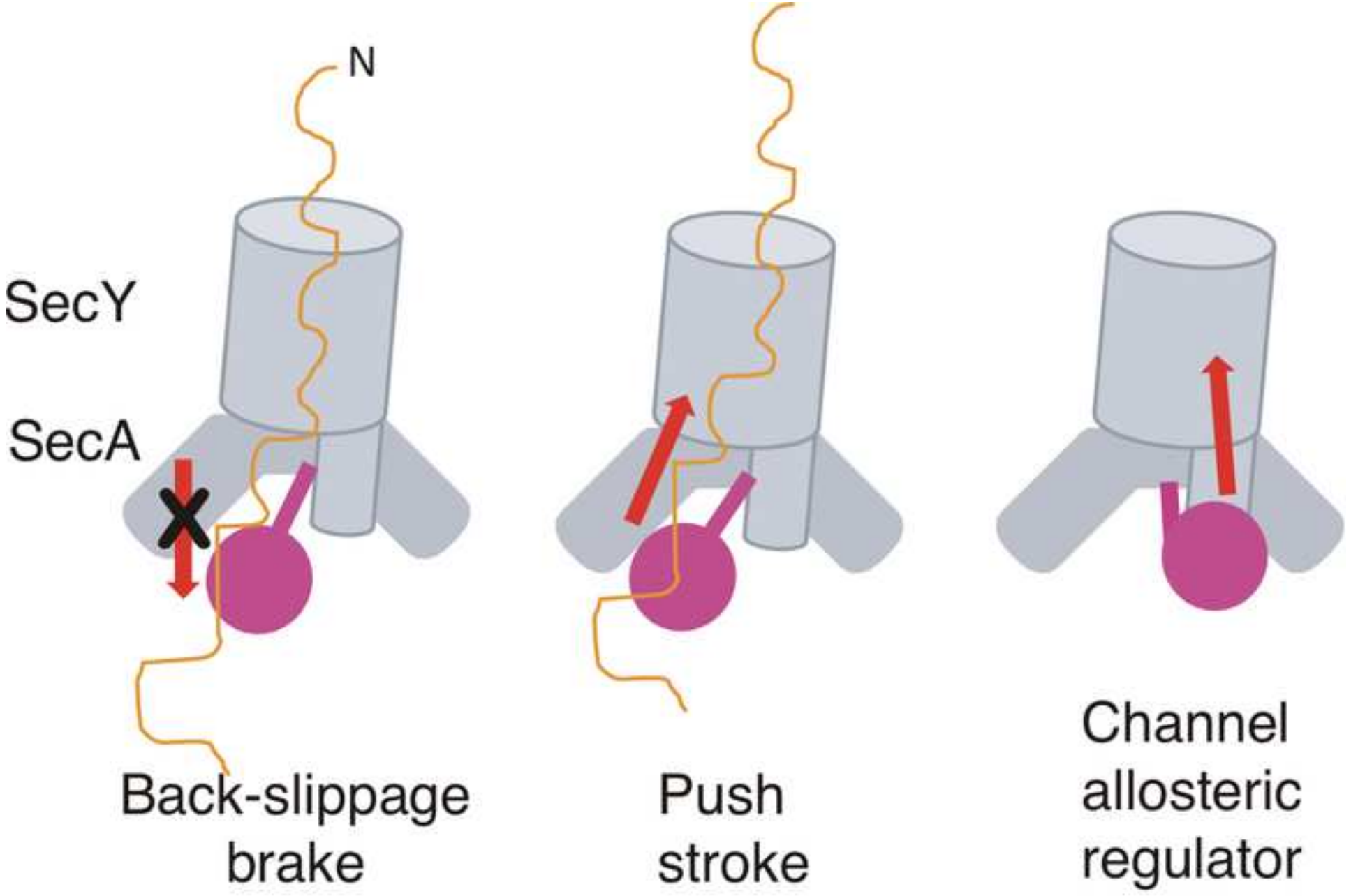


Fig. 7; Vandenberk et al.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal α-SecA	(Karamanou et al., 2008)	
Peroxidase AffiniPure Goat Anti-Rabbit IgG (H+L)	Jackson ImmunoResearch Europe Lt	Code: 111-035-003
Bacterial and Virus Strains		
DH5α: F⁻ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^- m_K^+$), λ⁻	Invitrogen	Cat#18258012
BL21.19 (DE3) (secA13 (Am) supF (Ts) trp (Am) zch::Tn10 recA::cat clpA::kan)	(Mitchell and Oliver, 1993)	N/A
BL21 (DE3) : *E. coli* str. B F⁻ *ompT gal dcm lon hsdS*$_B$($r_B^- m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S)	{Studier, 1990 #556}; NEB	NEB C2527
T7 express *lys*Y/I^q (DE3) : : MiniF lysY lacIq (CamR) / fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr73::miniTn10--TetS)2 [dcm] R(zgb-210::Tn10--TetS) endA1 Δ(mcrC-mrr)114::IS10	NEB	NEB C3013
Chemicals, Peptides, and Recombinant Proteins		
Tris base	Sigma-Aldrich	Cat#T1378
NaCl	Sigma-Aldrich	Cat#7647-14-5
Trolox	Sigma-Aldrich	Cat#53188-07-1
BSA	Sigma-Aldrich	N/A
ATTO488-CA	ATTO-TEC GmBH	Cat# AD 488-21
ATTO655-CA	ATTO-TEC GmBH	Cat# AD 655-21
ATTO488-maleimide	ATTO-TEC GmBH	Cat# D 488-45
Alexa Fluor 647 C2-maleimide	Life Technologies Europe BV	Cat#A20347
PFU Ultra Polymerase	Agilent	M7741
DpnI	Promega	R6231

Dithiotheitrol (DTT)	ApplichemPanreac	Cat#A1101
TCEP	Sigma-Aldrich	Cat#51805-45-9
Ethylenediaminetetraaceticacid, diNa salt, 2aq (EDTA)	Chemlab	Cat#CL00.0503
HEPES	Fisher	Cat#BP310
Imidazole	Carl Roth	Cat# 3899
glycerol	Sigma-Aldrich	Cat#56-81-5
Phenylmethylsulfonylfluoride (PMSF)	Roth	Cat#6367 Cat#6367
Magnesium Chloride ($MgCl_2$)	Roth	Cat#2189
Isopropylβ-D-1-thiogalactopyranoside (IPTG)	FischerScientific FischerScientific	Cat#BP1755 Cat#BP1755
Critical Commercial Assays		
Site-directed mutagenesis protocol	Stratagene-Agilent	N/A
Plasmid purification (NucleoSpin® Plasmid EasyPure)	Macherey- Nagel	Cat# 740727
Oligonucleotides		
For primers used in this study see Table S4		
Recombinant DNA		
For vectors used in this study see Table S3	This study	N/A
For genetics constructs used in this study see Table S3	This study	N/A
Software and Algorithms		
FPS software	(Kalinin et al., 2012)	http://www.mpc.hhu.de/software/fps.html
PyMol Molecular Graphics system, Version 2.07	Schrödinger	http://www.pymol.com
ModLoop	(Fiser and Sali, 2003)	https://modbase.compbio.ucsf.edu/modloop/
PAM	(Schrimpf et al., 2018)	http://www.cup.uni-muenchen.de/pc/lamb/software/pam.html
RSCB, Protein Data Bank	(Berman et al., 2000)	http://www.rcsb.org/
Other		
Nunc Lab-Tek Chambered cover-glass	Thermofisher Scientific	Cat#155411

0.45 μm filter, reg. cellulose 0.45μm	Grace discovery sciences	Cat#5123260(k45)
PD-10 desalting columns	GE Healthcare Europe GmbH	Cat#17085101
Amicon ultrafiltration columns (50 kDa, Ultra-0.5)	Merck Chemicals	Cat#UFC505024
Amicon ultrafiltration columns (3K, Ultra-15))	Merck Millipore	Cat#UFC900396
Ni^{+2}-NTA Agarose resin	Qiagen	Cat#30250
Hi-Load Superdex 200 26/60 gel filtration column	GE, Healthcare	Cat#28989336
SecA ATP hydrolysis experiment	(Chatzi et al., 2017)	N/A

Supplemental Information

The preprotein binding domain of SecA displays intrinsic rotational dynamics

Niels Vandenberk[1], Spyridoula Karamanou[2], Athina G. Portaliou[2], Valentina Zorzini[2], Johan Hofkens[1], Jelle Hendrix[1,3,4] and Anastassios Economou[2,4]

[1]KU Leuven, Department of Chemistry, Division for Molecular Imaging and Photonics, Laboratory for Photochemistry and Spectroscopy, Celestijnenlaan 200F, B-3001 Leuven, Belgium.
[2]KU Leuven, Department of Microbiology and Immunology, Rega Institute for Medical Research, Laboratory for Molecular Bacteriology, Herestraat 49, Gasthuisberg Campus, B-3000 Leuven, Belgium.
[3]Dynamic Bioimaging Lab, Advanced Optical Microscopy Centre, Biomedical Research Institute, Agoralaan C (BIOMED), Hasselt University, B-3590, Diepenbeek, Belgium

[4]For correspondence:
Lead contact: Anastassios Economou: tassos.economou@kuleuven.be
Jelle Hendrix: jelle.hendrix@uhasselt.be

Running title: SecA PBD rotational dynamics

Supplemental figures:

Supplemental tables:

Supplemental Data Files (external zip files):

Supplemental figures

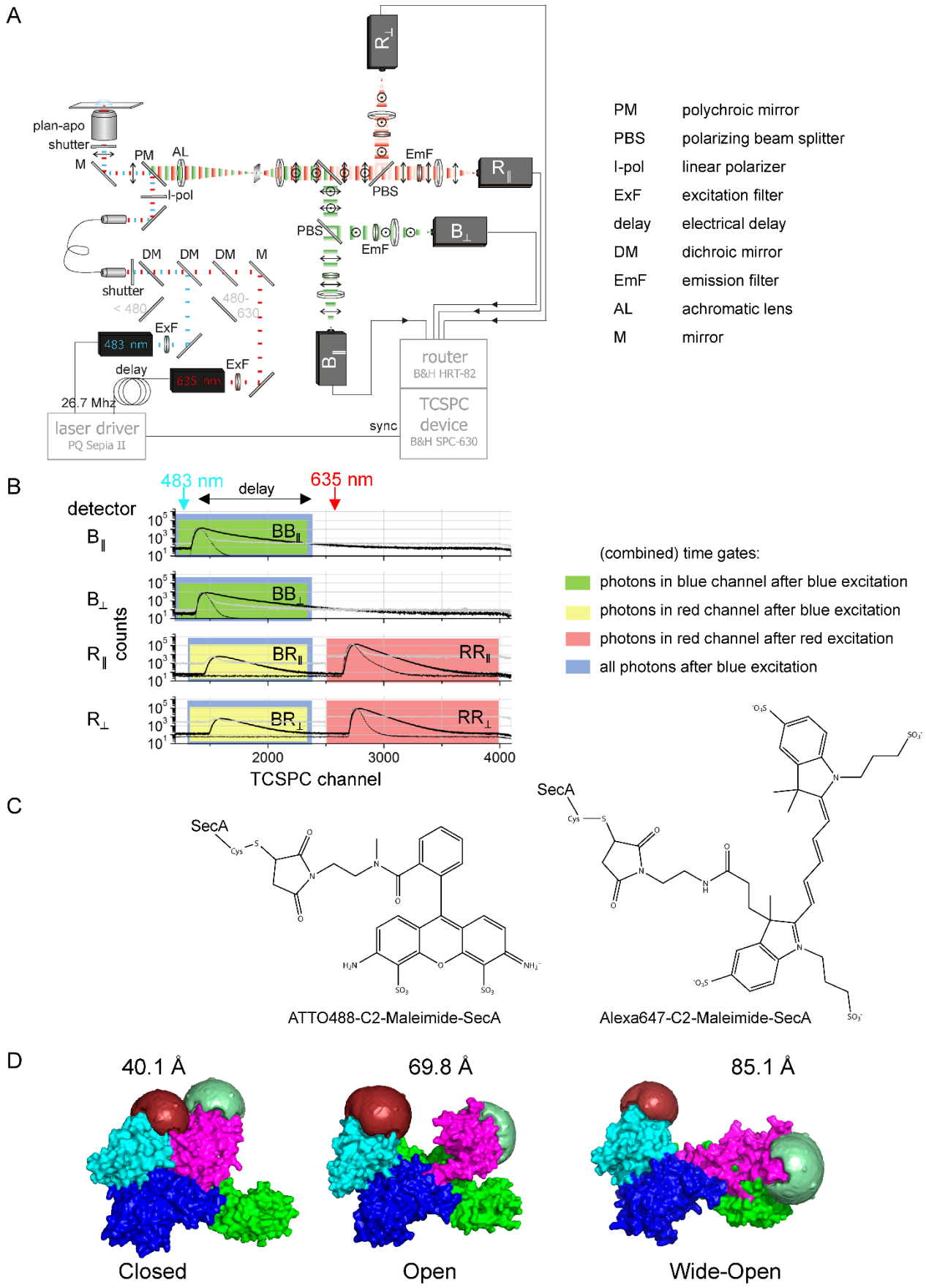


Figure S1 Employed MFD-PIE setup, photon sorting principle and accessible dye volumes. (Related to Figures 1 and 2)

A. Schematic presentation of the home-built MFD-PIE confocal setup with 483 nm and 635 nm pulsed interleaved excitation lines (Kudryavtsev et al., 2012). The two lasers, having a ~18 ns lag time, were combined in a single-mode optical fiber, the linear polarization was cleaned up and reflected into the microscope. The transmitted sample emission was focused through a pinhole and spectrally split. Each color was separately split in two detection channels according to their polarization. More specifically, two avalanche photodiodes detected blue photons; $B_{\parallel}$ (blue-parallel), $B_{\perp}$ (blue-perpendicular) while two others red photons; $R_{\parallel}$ (Red-parallel), $R_{\perp}$ (Red-perpendicular). Photodiodes were connected to a time-correlated single photon counting (TCSPC) device. Signals from each TCSPC channel were divided in time gates (Lamb et al., 2000) to discern the 483 nm – donor excited from the 635 nm acceptor excited photons. The setup has been detailed into the supplemental experimental procedures.

B. Different time gates ('PIE channels') in the previously described detection channels depending on the blue or red excitation line. Data (thick black line), IRF (thin black line), background recording (gray) are illustrated, along with the nomenclature of the different detectors and time gates. Since the 635 nm laser is delayed, first photons after 483-nm excitation will arrive on the specific detector. Photons are detected in $BB_{\parallel}/BB_{\perp}$ and $BR_{\parallel}/BR_{\perp}$ time gates. After the delay, 635 nm excitation triggers photons in the $RR_{\parallel}/RR_{\perp}$ time gates.

C. Lewis structure composition of dyes (ATTO488 and Alexa647) attached to a Cysteine of SecA via maleimide coupling

D. Pymol illustration of the accessible volume the dyes can sample via their flexible linker (Kalinin et al., 2012) for one SecA derivative (D1) in three conformations. Dye properties and simulated geometry are summarized in **Table S1**. The indicated expected distances were calculated (**Fout! Verwijzingsbron niet gevonden.**) from the simulated FRET efficiency (R_0 = 54.7 Å). The 3 structures correspond to the available X-ray structures of homologues: Closed and Wide-open, template-based models (Chatzi et al., 2017), and Open: 2FSF, seen by X-ray crystallography (Papanikolau et al., 2007) and NMR (Gelis et al., 2007). SecYEG would bind at the back of each structure in the translocase complex.

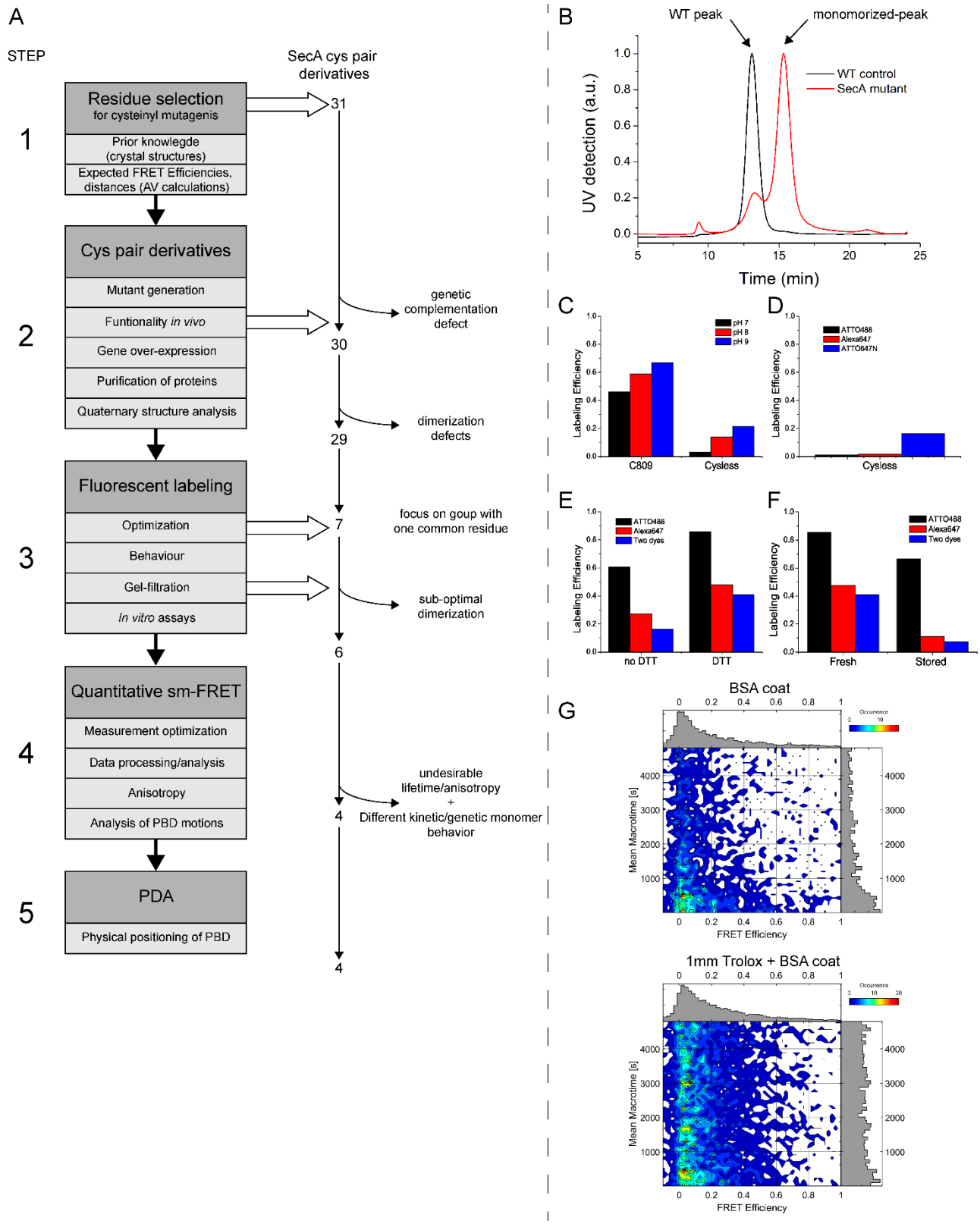


Figure S2 SmFRET pipeline to investigate the PBD conformational dynamics and labeling and measurement optimization. (Related to Figures 1-7)

A. Left, 5-step process as indicated. Middle, Detailed description of the steps. Right, Cys pair SecA derivatives that were constructed and taken to subsequent steps or eliminated.

B. Size-exclusion chromatography example of a fluorescently labelled protein. If derivatives show a single main peak comparable with the elution time of the WT peak the derivative is said to be well-behaved. On the other hand, if a derivative shows 2 peaks in close vicinity and one comparable

with the WT-peak, we can say that this is due to a monomerization process performed by fluorescent labeling and this derivative was excluded.

C-F) Labeling optimization.

C) pH-dependence of specific maleimide-Cys labeling was tested for pH7, 8 and 9 on a C809 and a Cys^- derivative. Labeling conditions can influence the efficiency of this reaction. Higher pH (>8.0) favors the reaction of maleimides with primary amines, such as lysinyl and argininyl residues (Brewer and Riehm, 1967), resulting in unspecific labeling, more dye molecules per protein, and at unwanted positions. Ideally, this reaction should be performed under pH conditions between 6.5 and 7.5, since amines remain protonated and are not nucleophilic to react with maleimides (Brewer and Riehm, 1967).

D. Dye choice. Commonly used FRET dyes: ATTO488 (donor), Alexa647 (acceptor) and ATTO647N (acceptor) were tested in terms of labeling specificity on a Cys^- derivative. The main reason for unspecific labeling of ATTO647N to the Cys^- may be due to the differences in the net charge of the two dyes (ATTO647N: 1^+, Alexa647: 3^-, ATTO488: 2^-).

E. Purification conditions. Reducing steps during purification can influence the labeling efficiency positively, therefore a labeling reaction was performed on a strongly reduced (DTT included during purification) and a weakly reduced (no DTT) protein. The data clearly show that maintaining Cysteine reduction until right before maleimide labeling considerably enhances the labeling efficiency.

F. Effect of storage condition.

G. Example of measurement optimization for mutant D2. Shown are FRET efficiency versus (laboratory) time plots, which illustrate how the number of D/A bursts, and the population FRET efficiency evolve over time. Top: Addition of only a coverslip BSA coat still results in a decreased number of bursts over time due to absorption (although less as compared to no-BSA samples; data not shown). Bottom: Additionally adding 1 mM aged Trolox in the sample solution caused the number of double labeled bursts to remain constant over time by improving the solubility of the protein.

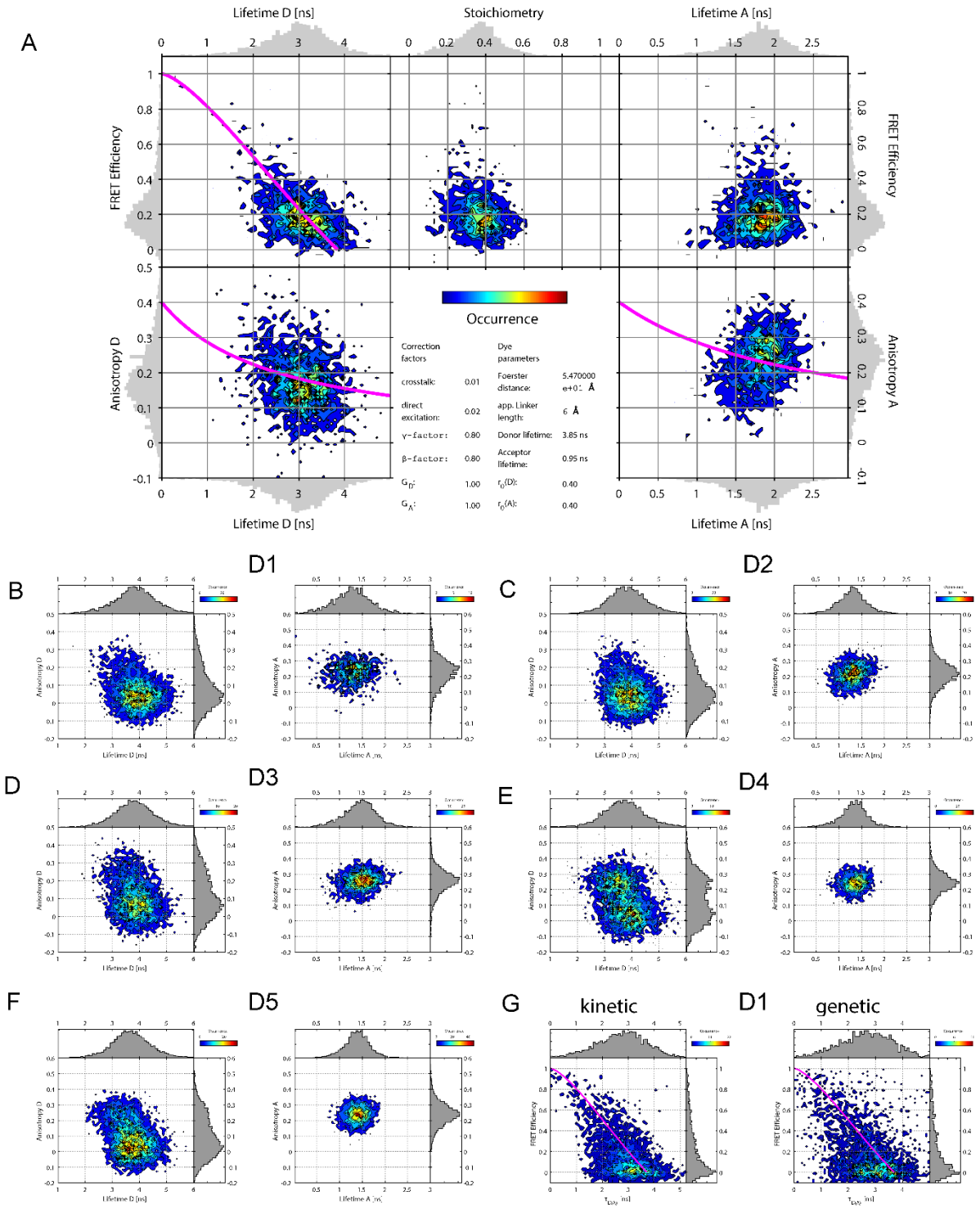


Figure S3 Multi-parameter analysis of double and single-labeled proteins (Related to Figures 3-6).

A. Exemplary multi-dimensional smFRET histogram for derivative D3 displaying the intensity, lifetime and anisotropy information for both donor and acceptor dyes of all double-labeled single molecules. Displayed are *E*-*S* (center), *E*-τ_D (top left), *E*-τ_A (top right), r_D-τ_D (bottom left) and r_A-τ_A (bottom right) 2D histograms (red = high histogram counts, blue = low histogram counts). The 1D bar charts are projections of the 2D histograms on the respective axes. The pink line in the *E*-τ_D plot is the static FRET line calculated as described in the STAR Methods section. The pink lines in the *r*-τ plots are Perrin equations calculated with Eq. 8.

B-F. Donor-only and acceptor fluorescence lifetimes and steady-state anisotropies of 5 SecA derivatives. Left: 2D histograms of r_D-τ_D. Right, 2D histograms of r_A-τ_A.

G. Comparison of kinetic and genetic SecA monomers.

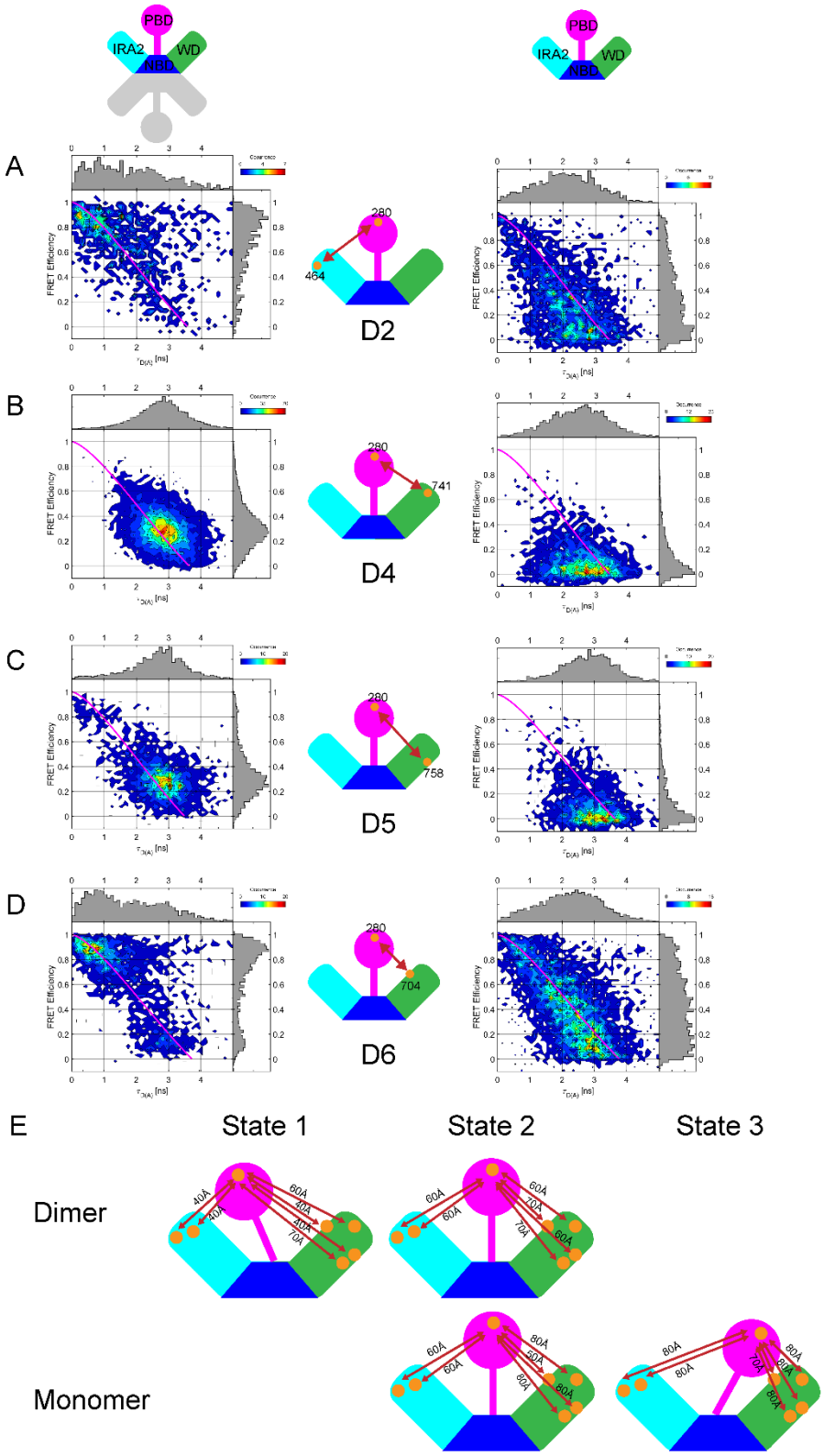


Figure S4 Analysis of PBD motions by smFRET. (Related to Figures 3 and 4)
A-D. Analysis of PBD dynamic conformational behavior when studied within the context of dimeric (**left**) and monomeric (**right**) SecA, with cartoons of the derivatives in the middle. 3,000-4,000 individual bursts/each from the 4 indicated derivatives were plotted on 2D plots (as in Figure 2C). Contour plots display 2D histograms of molecule counts (red = high, blue = low counts). The 1D bar charts are projections of the 2D histograms on the respective axes. Static FRET lines (red) were calculated with Eq. 13. Based on the structures, D6 would be expected to display a FRET efficiency of ~0.2 for the "closed" state and 0.8 for both the "open" and "wide-open" states. Thus, it agrees with

the other probes of the same interdomain interface but due to its different position on the WD, results in different FRET behavior.

E. Models of the different states of Dimer and Monomer based on distances directly calculated from the Effiency (2D plots).

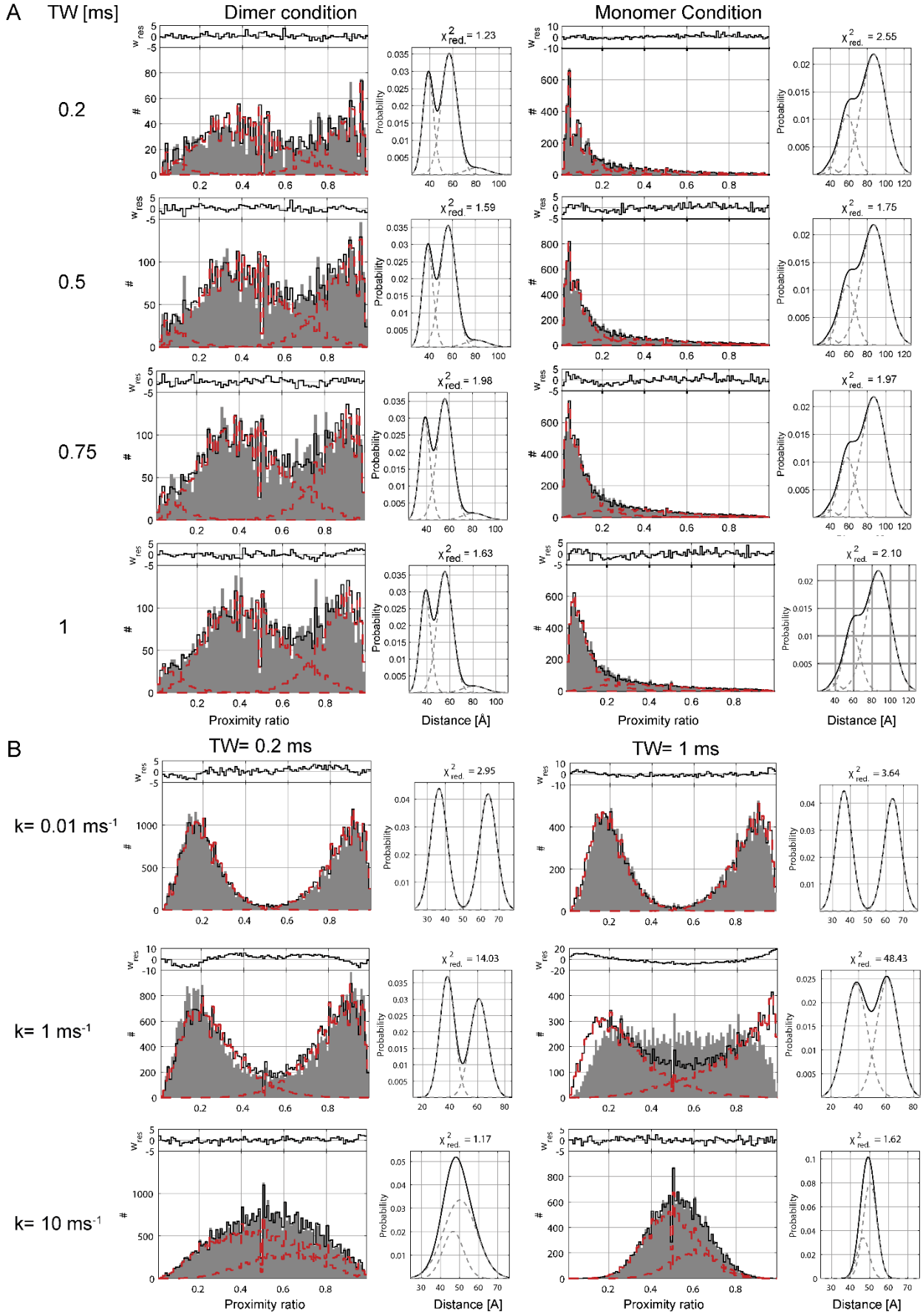


Figure S5 PDA analysis of SecA and control experiments on simulated molecules and dsDNA. (Related to Figures 3 and 4)

A. Detailed illustration of PDA of derivative D1 in dimeric and monomeric conditions. Burst data were binned into either of four time windows (TW, left of the histogram). PDA analysis was performed as described in the STAR Methods. Correction parameters were γ-factor = 0.8, β-factor = 0.010, B_{BB} = 1.23 kHz, B_{RR} = 0.43 kHz and R_0 = 54.7 Å. Individual χ^2 values are indicated in the figure, the overall χ^2 was 2.10. Both for the dimer and monomer, no extra states visually appeared when decreasing the bin size, and the 4 histograms could be modelled perfectly with the same state fractions, distances and distance distribution widths. Both of these observations are strongly indicative of the absence of FRET dynamics on the 1-10-ms time scale, yet do not exclude FRET dynamics below or above that time scale.

B. PDA analysis in two time windows (TW = 0.2 ms and 1 ms) of simulated data (R_0 = 52 Å) of molecules interconverting between two states (65 Å and 35 Å D-A distance) with different rate constants (k_{12} and k_{21} equal to 0.01 ms^{-1}, 1 ms^{-1} and 10 ms^{-1}). For a slow interconversion (k_{12} and k_{21} equal to 0.01 ms^{-1}), two populations appeared for both time windows. Thus, in the case of the 0.01 ms^{-1} rate constant, we can assume the conformation of the molecule is "static" during its passage through the focus. For larger rate constants, a single broad population appeared, with FRET values in between the two extreme states. In this case, the data can clearly not be modeled with two state of 65 and 35 Å, respectively, and other PDA methods should be used to describe the data.

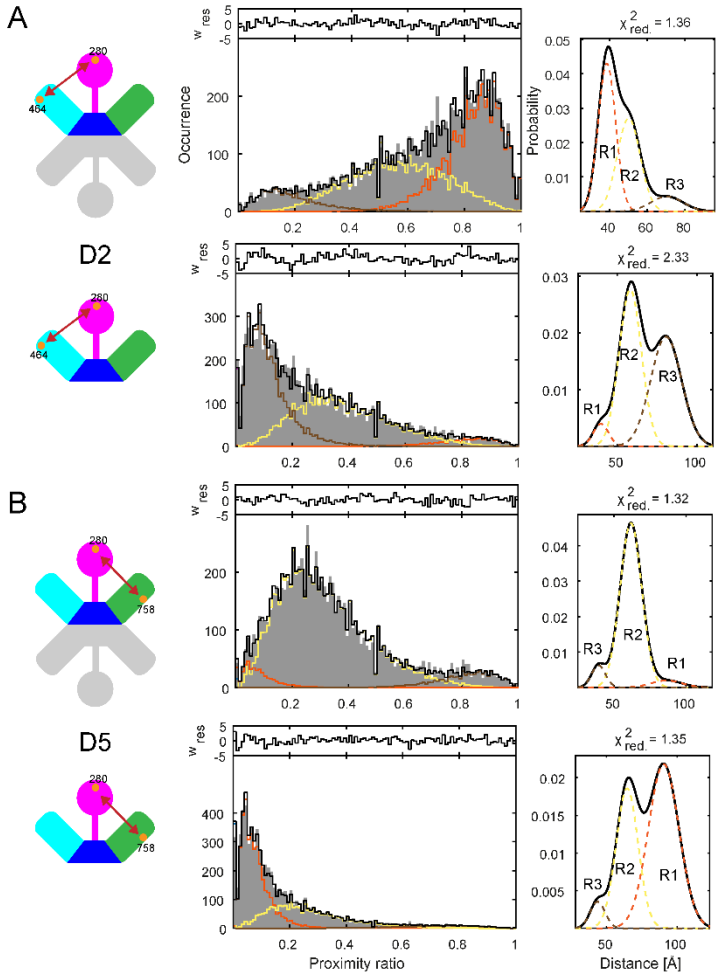


Figure S6 PDA analysis of dimeric and monomeric SecA. (Related to Figures 3 and 4)
PDA of (**A**) D2 and (**B**) D5. Data of four time windows was globally analyzed (0.2, 0.5, 0.75 and 1 ms; STAR methods and Figure S5), but only the 1-ms time bin data is visualized. **Left**, PDA analysis with Gaussian distribution widths globally determined as a distance fraction F over the total range of R values. **Right**, Corresponding distance distribution plots illustrating the intricate relation of distance and FRET distribution width.

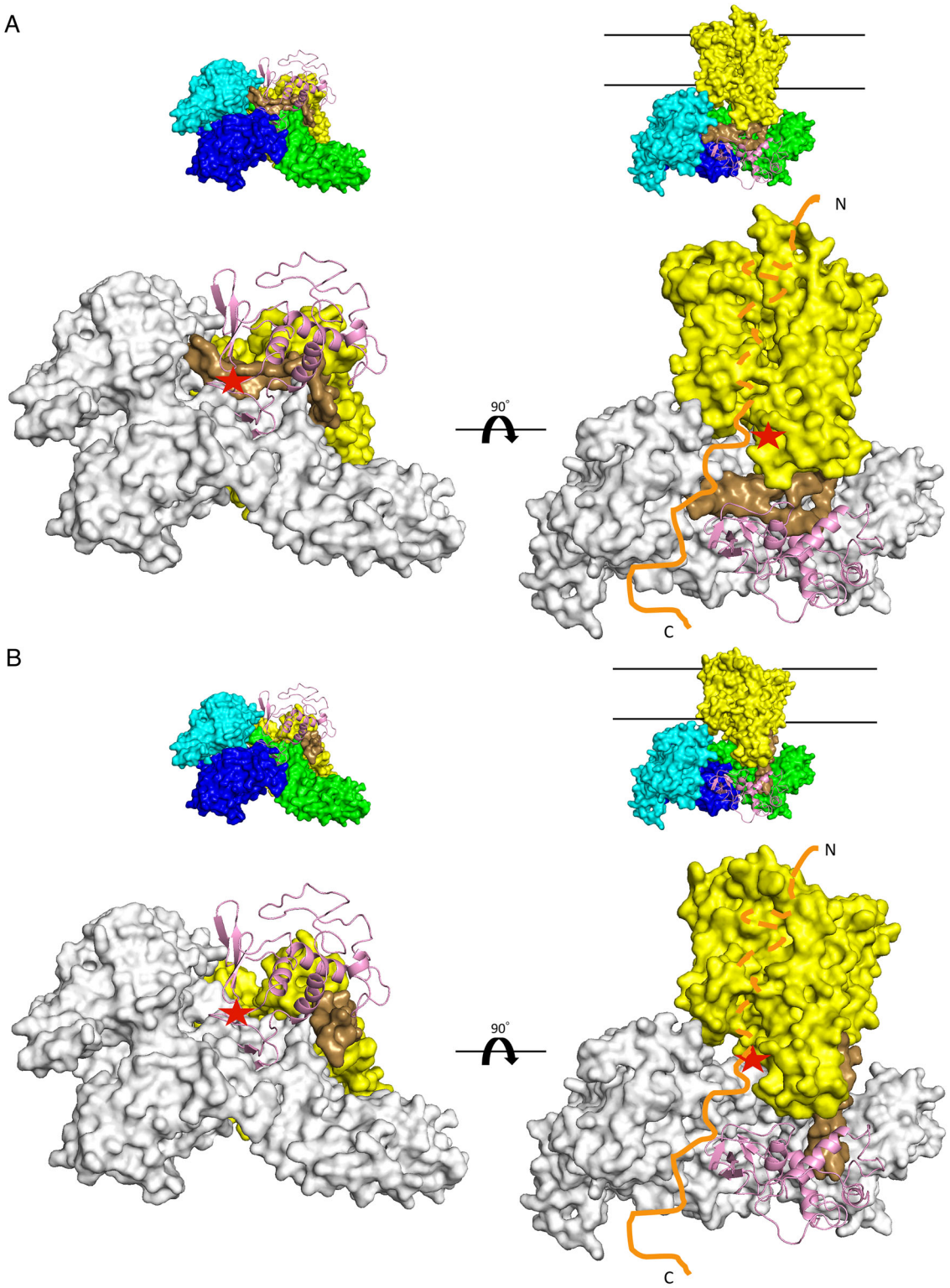


Figure S7 Structural visualization of SecA bound to SecY. (Related to Figures 1, 6 and 7)
A. Two enlarged views, rotated by 90º, of the SecA structure (top; 4-coloured, surface) of the sm-FRET-derived, monomeric State 2A SecA modelled on the SecY structure from *Thermotoga maritima* (pdb: 3DIN; (Zimmer et al., 2008)).
B. Two enlarged views, rotated by 90º, of the sm-FRET-derived, monomeric State 2A SecA modelled on the SecY structure from *Escherichia coli* (pdb: 5GAE;(Jomaa et al., 2016)). 3DIN was derived as a co-complex with monomeric SecA in the presence of ADP and BeF_x; in this structure the cytoplasmic protrusion of SecY (brown) has little secondary structure and is tilted towards the translocation channel. In 5GAE, the cytoplasmic protrusion has a prominent anti-parallel beta stranded sheet and is facing the cytoplasm. In either case the PBD would make intimate contacts with the cytoplasmic protrusion. The SecA surface is represented in grey, the PBD is in pink ribbon, the

SecY surface in yellow. The SecY cytoplasmic domain is coloured in brown. A red star indicates the translocation channel inside SecY thought to be occupied by preproteins (orange) as they translocate. Parallel lines demarcate the position of the inner membrane lipid bilayer.

Supplemental tables:

Table S1 Dye properties and simulated geometry. (Related to Figure 2)
See Kalinin et al. for an illustration of the different parameters concerning linker and dye radius (Kalinin et al., 2012). $\lambda_{exc,max}$ is the absorption maximum, $\lambda_{em,max}$ is the emission maximum, $\varepsilon_{abs.max}$ is the extinction coefficient at maximum absorption, $\varepsilon_{483/635nm}$ is the extinction coefficient at the used excitation wavelengths (483 nm for the donor dyes and 635 nm for the acceptor dyes), ϕ is the measured fluorescence quantum yield of ATTO488 and the literature fluorescence quantum yield of Alexa647. τ_{fluo} is the literature fluorescence lifetime. All dye photophysical parameters were obtained from the respective manufacturers' websites.

	Atto488	Alexa647
Charge	3-/1+	4-/1+
$\lambda_{exc,max}$ (nm)	500	651
$\varepsilon_{abs.max}$ ($M^{-1}cm^{-1}\times10^3$)	90	265
$\varepsilon_{483nm\ or\ 635nm}$ ($M^{-1}cm^{-1}\times10^3$)	46	183
$\lambda_{em,max}$ (nm)	520	672
ϕ	0.6	0.33
τ_{fluo} (ns)	4.0	1.0
length [Å]	18.7	20.3
width [Å]	4.5	4.5
radius 1 [Å]	5.0	11.0
radius 2 [Å]	4.5	4.7
radius 3 [Å]	1.5	1.5

Table S2 List of designed double Cys SecA derivatives with their expected FRET efficiencies calculated from FPS analysis. (Related to Figures 2-5)

Colors refer to the specific domain where the residue is located. Purple = PBD, Blue = NBD, Cyan= IRA2 and Green= C-domain.

Nomenclature of characterized derivatives	Nomenclature of designed derivatives	Residue 1	Residue 2	Closed	Open	Wide-Open
	21	E284	P704	0.33	0.89	0.75
D2	22	V280	L464	0.87	0.19	0.07
D6	23	V280	P704	0.31	0.83	0.82
	24	E283	A741	0.09	0.47	0.84
	25	E283	M758	0.21	0.34	0.8
	26	E148	E752	0.43	0.43	0.43
	27	E148	D330	0.67	0.35	0.17
	28	E284	L464	0.71	0.18	0.2
	29	E284	A432	0.57	0.15	0.11
	30	E284	A741	0.07	0.38	0.59
	31	E284	E757	0.12	0.22	0.6
	32	E284	M758	0.22	0.32	0.74
	33	K241	P704	0.43	0.9	0.64
	34	V280	E752	0.04	0.16	0.73
D4	35	V280	A741	0.07	0.32	0.71
D3	36	V280	E757	0.09	0.20	0.71
D5	37	V280	M758	0.13	0.24	0.85
D1	38	V280	S604	0.91	0.28	0.075
	39	K268	E591	0.98	0.69	0.33
	40	P145	E752	0.27	0.27	0.27
	41	E252	L464	0.96	0.37	0.27
	42	E252	P704	0.12	0.64	0.69
	43	E252	S604	0.95	0.55	0.21
	44	K251	P704	0.18	0.85	0.75
	45	K251	L464	0.91	0.2	0.35
	46	K251	S604	0.91	0.36	0.32
	47	E283	L464	0.57	0.12	0.08
	48	E283	P704	0.37	0.89	0.9
	49	E283	E757	0.09	0.2	0.66
	50	E141	P704	0.1	0.09	0.05
	51	E141	D330	0.86	0.53	0.27

Table S3 Genetic constructs (Related to Figure 2).

Gene/mutation	Plasmid ID	Vector	Cloning / PCR strategy or source
His-*secA*(6-901)			
His-*secA(6-901cys$^-$)*	pLMB0080	pET5a	The Cys98 was mutated to a serine in pIMBB258 (His SecA(N6-901)3cys-; (Chatzi et al., 2017)) using the primers X1913-X1914
***secA*(1-901)**			
cys$^-$	pLMB0092	pET3a	The 2.5kb NcoI fragment from pLMB0080 replaced the corresponding one in pIMBB1280 (Gouridis et al., 2013)
E284C	pLMB1654	pET3a	The mutation was introduced in pLMB0092 using primers X1820-X1821
V280C	pLMB0094	pET3a	The mutation was introduced in pLMB0092 using primers X1824-X1825
E283C	pLMB0095	pET3a	The mutation was introduced in pLMB0092 using primers X1832-X1833
E148C	pLMB0096	pET3a	The mutation was introduced in pLMB0092 using primers X1834-X1835
P704C	pLMB1658	pET3a	The mutation was introduced in pLMB0092 using primers X1772-X1773
L464C	pLMB1659	pET3a	The mutation was introduced in pLMB0092 using primers X1768-X1769
A741C	pLMB1660	pET3a	The mutation was introduced in pLMB0092 using primers X1774-X1775
M758C	pLMB1661	pET3a	The mutation was introduced in pLMB0092 using primers X1780-X1781
E752C	pLMB1662	pET3a	The mutation was introduced in pLMB0092 using primers X1776-X1777
D330C	pLMB1663	pET3a	The mutation was introduced in pLMB0092 using primers X1838-X1839
K268C	pLMB1655	pET3a	The mutation was introduced in pLMB0092 using primers X1032-X1033
P145C	pLMB1656	pET3a	The mutation was introduced in pLMB0092 using primers X1826-X1827
E252C	pLMB1657	pET3a	The mutation was introduced in pLMB0092 using primers X1828-X1829
K251C	pLMB1644	pET3a	The mutation was introduced in pLMB0092 using primers X1830-X1831
E141C	pLMB1648	pET3a	The mutation was introduced in pLMB0092 using primers X1836-X1837
A432C	pLMB1683	pET3a	The mutation was introduced in pLMB0092 using primers X1770-X1771
E757C	pLMB1684	pET3a	The mutation was introduced in pLMB0092 using primers X1778-X1779
S604C	pLMB1685	pET3a	The mutation was introduced in pLMB0092 using primers X1782-X1783
E284C P704C	pLMB1645	pET3a	The mutation was introduced in pLMB1654 using primers X1772-X1773
V280C L464C	pLMB1646	pET3a	The mutation was introduced in pLMB0094 using primers X1768-X1769
V280C P704C	pLMB1647	pET3a	The mutation was introduced in pLMB0094 using primers X1772-X1773
E283C A741C	pLMB0097	pET3a	The mutation was introduced in pLMB1660 using primers X1832-X1833
E283C M758C	pLMB1649	pET3a	The mutation was introduced in pLMB1661 using primers X1832-X1833
E148C E752C	pLMB1650	pET3a	The mutation was introduced in pLMB0096 using primers X1776-X1777
E148C D330C	pLMB0098	pET3a	The mutation was introduced in pLMB0096 using primers X1838-X1839
E284C L464C	pLMB1651	pET3a	The mutation was introduced in pLMB1654 using primers X1768-X1769
E284C A432C	pLMB1652	pET3a	The mutation was introduced in pLMB1654 using primers X1770-X1771
E284C A741C	pLMB1653	pET3a	The mutation was introduced in pLMB1654 using primers X1774-X1775

E284C_E757C	pLMB1686	pET3a	The mutation was introduced in pLMB1654 using primers X1778-X1779
E284C_M758C	pLMB1687	pET3a	The mutation was introduced in pLMB1654 using primers X1780-X1781
K241C_P704C	pLMB1688	pET3a	The mutation was introduced in pLMB1658 using primers X1772-X1773
V280C_E752C	pLMB1689	pET3a	The mutation was introduced in pLMB0094 using primers X1776-X1777
V280C_A741C	pLMB0099	pET3a	The mutation was introduced in pLMB1660 using primers X1824-X1825
V280C_E757C	pLMB1690	pET3a	The mutation was introduced in pLMB1684 using primers X1824-X1825
V280C_M758C	pLMB0100	pET3a	The mutation was introduced in pLMB1661 using primers X1824-X1825
V280C_S604C	pLMB1691	pET3a	The mutation was introduced in pLMB1685 using primers X1824-X1825
P145C_E752C	pLMB1692	pET3a	The mutation was introduced in pLMB1656 using primers X1776-X1777
E252C_L464C	pLMB1693	pET3a	The mutation was introduced in pLMB1659 using primers X1828-X1829
E252C_P704C	pLMB1694	pET3a	The mutation was introduced in pLMB1657 using primers X1772-X1773
E252C_S604C	pLMB1695	pET3a	The mutation was introduced in pLMB1657 using primers X1782-X1783
K251C_ P704C	pLMB1696	pET3a	The mutation was introduced in pLMB1644 using primers X1772-X1773
E283C_L464C	pLMB1697	pET3a	The mutation was introduced in pLMB1659 using primers X1832-X1833
E283C_P704C	pLMB1698	pET3a	The mutation was introduced in pLMB1658 using primers X1832-X1833
E283C_E757C	pLMB1699	pET3a	The mutation was introduced in pLMB1684 using primers X1832-X1833
E141C_P704C	pLMB1700	pET3a	The mutation was introduced in pLMB1648 using primers X1772-X1773
E141C_D330C	pLMB1701	pET3a	The mutation was introduced in pLMB1648 using primers X1838-X1839
***secA*(15-901/6A) (encoding monomeric SecA)**			
V280C_E752C	pLMB1729	pET3a	The 2.5 kb NcoI fragment from pLMB1689 replaced the corresponding one in pIMBB1286.
V280C_A741C	pLMB1730	pET3a	The 2.5 kb NcoI fragment from pLMB0099 replaced the corresponding one in pIMBB1286.
V280C_E757C	pLMB1731	pET3a	The 2.5 kb NcoI fragment from pLMB1690 replaced the corresponding one in pIMBB1286
V280C_M758C	pLMB1732	pET3a	The 2.5 kb NcoI fragment from pLMB0100 replaced the corresponding one in pIMBB1286
V280C_S604C	pLMB1733	pET3a	The 2.5 kb NcoI fragment from pLMB1691 replaced the corresponding one in pIMBB1286
V280C_L464C	pLMB1727	pET3a	The 2.5 kb NcoI fragment from pLMB1646 replaced the corresponding one in pIMBB1286
V280C_P704C	pLMB1728	pET3a	The 2.5 kb NcoI fragment from pLMB1647 replaced the corresponding one in pIMBB1286

Table S4 List of primers (Related to Figure 2).

Primer ID	Forward (F)/ Reverse (R)	Mutation introduced in *secA*	Sequence (3') (Mutated codons are shown in bold)
X1032	F	K268C	TTC TCG GTG GAC GAA **TGC** TCT CGC CAG GTG AAC
X1033	R	K268C	GTT CAC CTG GCG AGA **GCA** TTC GTC CAC CGA GAA
X1768	F	L464C	CC ATC GAA AAA TCG GAG **TGC** GTG TCA AAC GAA CTG
X1769	R	L464C	CAG TTC GTT TGA CAC **GCA** CTC CGA TTT TTC GAT GG
X1770	F	A432C	C CTG GTC TAC ATG ACT GAA **TGC** GAA AAA ATT CAG GCG
X1771	R	A432C	CGC CTG AAT TTT TTC **GCA** TTC AGT CAT GTA GAC CAG G
X1772	F	P704C	GAA ATG TGG GAT ATT **TGC** GGG CTG CAG GAA CGT C
X1773	R	P704C	G ACG TTC CTG CAG CCC **GCA** AAT ATC CCA CAT TTC
X1774	F	A741C	CGT GAG CGC ATT CTG **TGC** CAG TCC ATC GAA GTG
X1775	R	A741C	CAC TTC GAT GGA CTG **GCA** CAG AAT GCG CTC ACG
X1776	F	E752C	GTG TAT CAG CGT AAA GAA **TGC** GTG GTT GGT GCT GAG
X1777	R	E752C	CTC AGC ACC AAC CAC **GCA** TTC TTT ACG CTG ATA CAC
X1778	F	E757C	GAA GTG GTT GGT GCT **TGC** ATG ATG CGT CAC TTC G
X1779	R	E757C	C GAA GTG ACG CAT CAT **GCA** AGC ACC AAC CAC TTC
X1780	F	M758C	GTG GTT GGT GCT GAG **TGC** ATG CGT CAC TTC GAG
X1781	R	M758C	CTC GAA GTG ACG CAT **GCA** CTC AGC ACC AAC CAC
X1782	F	S604C	GCT TCC GAC CGA GTA **TGC** GGC ATG ATG CGT AAA
X1783	R	S604C	TTT ACG CAT CAT GCC **GCA** TAC TCG GTC GGA AGC
X1820	F	E284C	CTG GTG CTG ATT GAA **TGC** CTG CTG GTG AAA GAG GGC
X1821	R	E284C	GCC CTC TTT CAC CAG CAG **GCA** TTC AAT CAG CAC CAG
X1824	F	V280C	G ACC GAA CGT GGT CTG **TGC** CTG ATT GAA GAA CTG C
X1825	R	V280C	G CAG TTC TTC AAT CAG **GCA** CAG ACC ACG TTC GGT C
X1826	F	P145C	C GCC GAA AAC AAC CGT **TGC** CTG TTT GAA TTC CTT GGC C
X1827	R	P145C	G GCC AAG GAA TTC AAA CAG **GCA** ACG GTT GTT TTC GGC G
X1828	F	E252C	G ATC CGT CAG GAA AAA **TGC** GAC TCC GAA ACC TTC C
X1829	R	E252C	G GAA GGT TTC GGA GTC **GCA** TTT TTC CTG ACG GAT C
X1830	F	K251C	C CTG ATC CGT CAG GAA **TGC** GAA GAC TCC GAA ACC
X1831	R	K251C	GGT TTC GGA GTC TTC **GCA** TTC CTG ACG GAT CAG G
X1832	F	E283C	GGT CTG GTG CTG ATT **TGC** GAA CTG CTG GTG AAA G
X1833	R	E283C	C TTT CAC CAG CAG TTC **GCA** AAT CAG CAC CAG ACC
X1834	F	E148C	C AAC CGT CCG CTG TTT **TGC** TTC CTT GGC CTG ACT G
X1835	R	E148C	C AGT CAG GCC AAG GAA **GCA** AAA CAG CGG ACG GTT G
X1836	F	E141C	GCG CAA CGT GAC GCC **TGC** AAC AAC CGT CCG CTG

X1837	R	E141C	CAG CGG ACG GTT GTT **GCA** GGC GTC ACG TTG CGC
X1838	F	D330C	C GTC GAC TAC ATC GTT AAA **TGT** GGT GAA GTT ATC ATC G
X1839	R	D330C	C GAT GAT AAC TTC ACC **ACA** TTT AAC GAT GTA GTC GAC G
X1913	F	C98S	GTT CTT AAC GAA CGC **AGC** ATC GCC GAA ATG CGT
X1914	R	C98S	ACG CAT TTC GGC GAT **GCT** GCG TTC GTT AAG AAC

Table S5 Extrapolated distances from X-ray data versus distances measured in solution by smFRET and PDA (Related to Figure 6).

Comparison of physical distances between the same residues of PBD as it undergoes motions. Left, values were obtained from the three defined modeled states from the *E. coli* (Open model, 2FSF) (Sardis and Economou, 2010, Chatzi et al., 2017)) after deriving $\langle R_{DA} \rangle_E$ from accessible volume calculations. **Right**, values obtained from smFRET of the three indicated SecA derivatives in solution, followed by PDA.

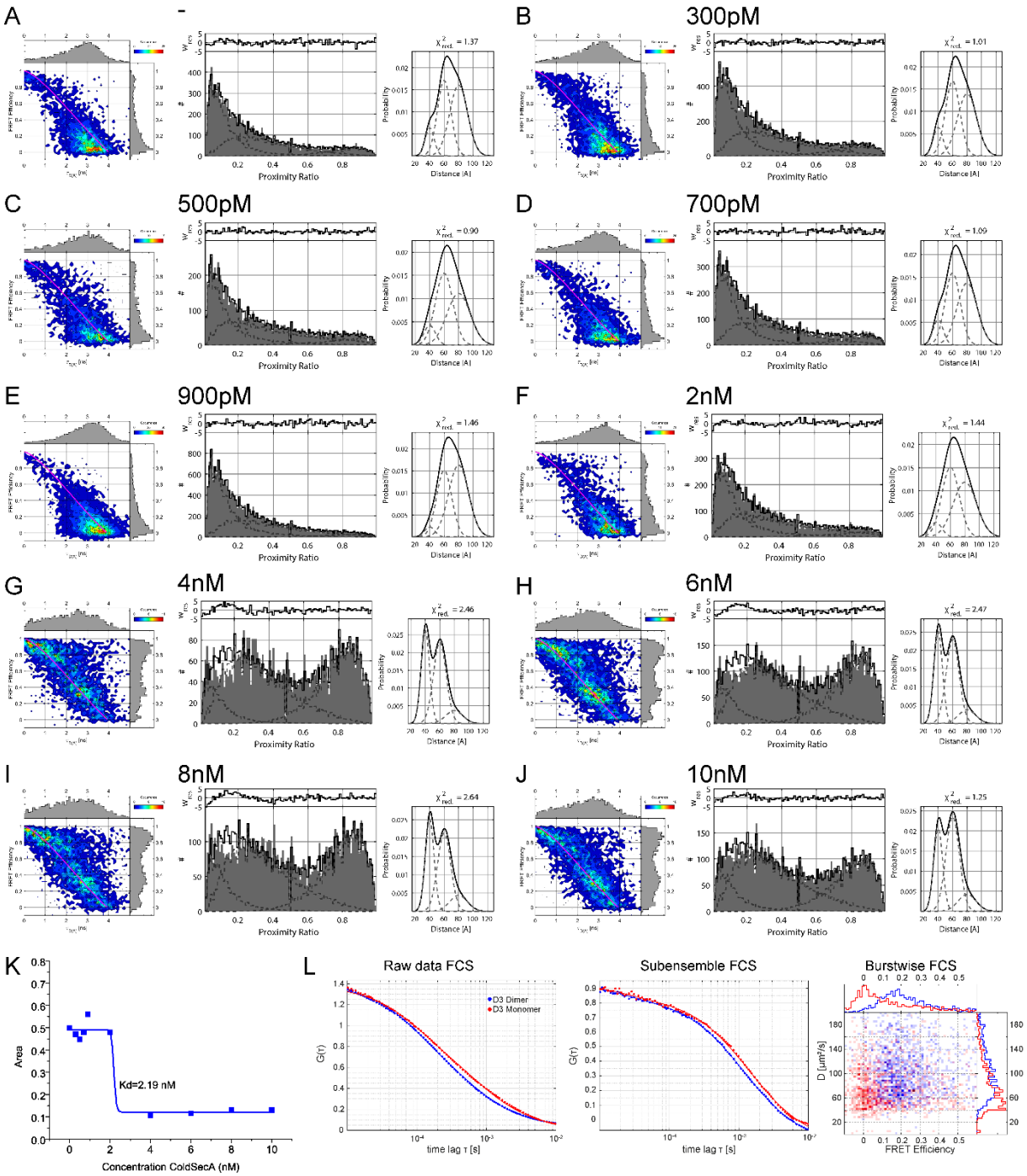
Distances (Å)										
Probe	Expected FPS			Dimer			Monomer			
	Closed	Open	Wide-open	State 1	State 2	State 3	State 1	State 2	State 2a	State 3
D1	37.3	64.0	83.2	38	56	79	42	61	85	85
D2	40.1	69.8	85.1	39	50	70	40	58	81	81
D3	80.3	68.8	46.9	82	65	41	91	65	91	43
D5	74.7	66.2	40.9	86	62	39	90	66	90	44

Supplemental Data Files

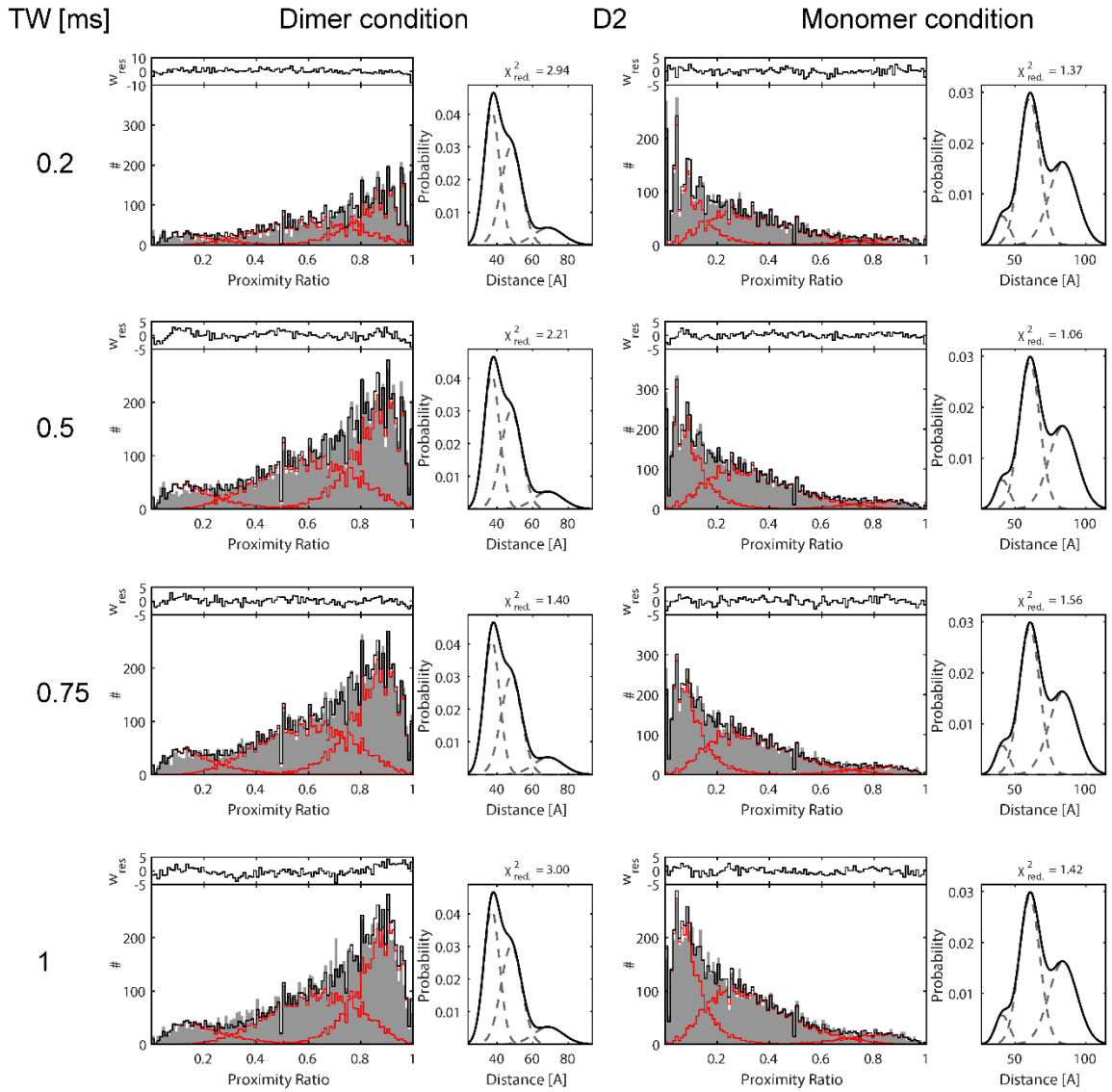
Data File 1: Estimation of the dimerization K_d of SecA. (Related to Figures 3-5)
A-J) To estimate the dimerization dissociation constant of SecA, different concentrations of unlabeled SecA were added to 200 pM of fluorescently labeled SecA (D1) and measured using smFRET. Increasing amounts of unlabeled SecA up to 2 nM did not significantly affect the distribution of the 3 states (**Figure 3A, left**). When more than 2 nM unlabeled SecA was added, a clear shift of the area percentage of state 1 to high ratio and state 3 to low ratio was seen. The high FRET state saturated at 4 nM unlabeled SecA and remained unaltered even after increasing the concentration until 10 nM. The dimerization process is completed after adding 4-10 nM of unlabeled SecA. Left, E-τ_D histogram; Right, PDA analysis. The different concentrations were globally analyzed in PDA (global R, free A). **K)** Binding curve obtained by plotting the A3 parameter after PDA analysis of the data in panels A-J. State 3 is the low FRET state, which decreases when SecA is dimerizing. **L)** FCS analysis of the monomer (100 pM) and dimer (100 pM + 10 nM unlabeled SecA) form of derivative D3. Left, Cross-correlation of raw data using all photons obtained after donor excitation. Middle, subensemble FCS using all photons obtained after donor excitation. Right, burstwise diffusion coefficients from subensemble data. Dimer data is in blue, monomer data in red. For the very simplified case of a globular molecule, the diffusion constant would be 1.2-fold smaller if M_r doubles, because of the increased hydrodynamic radius. Experimentally, the diffusion constant of the dimer condition was consistently higher (about ~1.4-fold) than that of the monomer, indicating the hydrodynamic radius of the dimer is smaller than that of the monomer. Although surprising at first, this observation is in line with the FRET data that also suggest an overall compaction of the protein in its dimeric state (overall higher FRET values were observed). Similar observations were done for the other derivatives (D1, D2 and D5, data not shown), or when all photons after acceptor excitation were used.

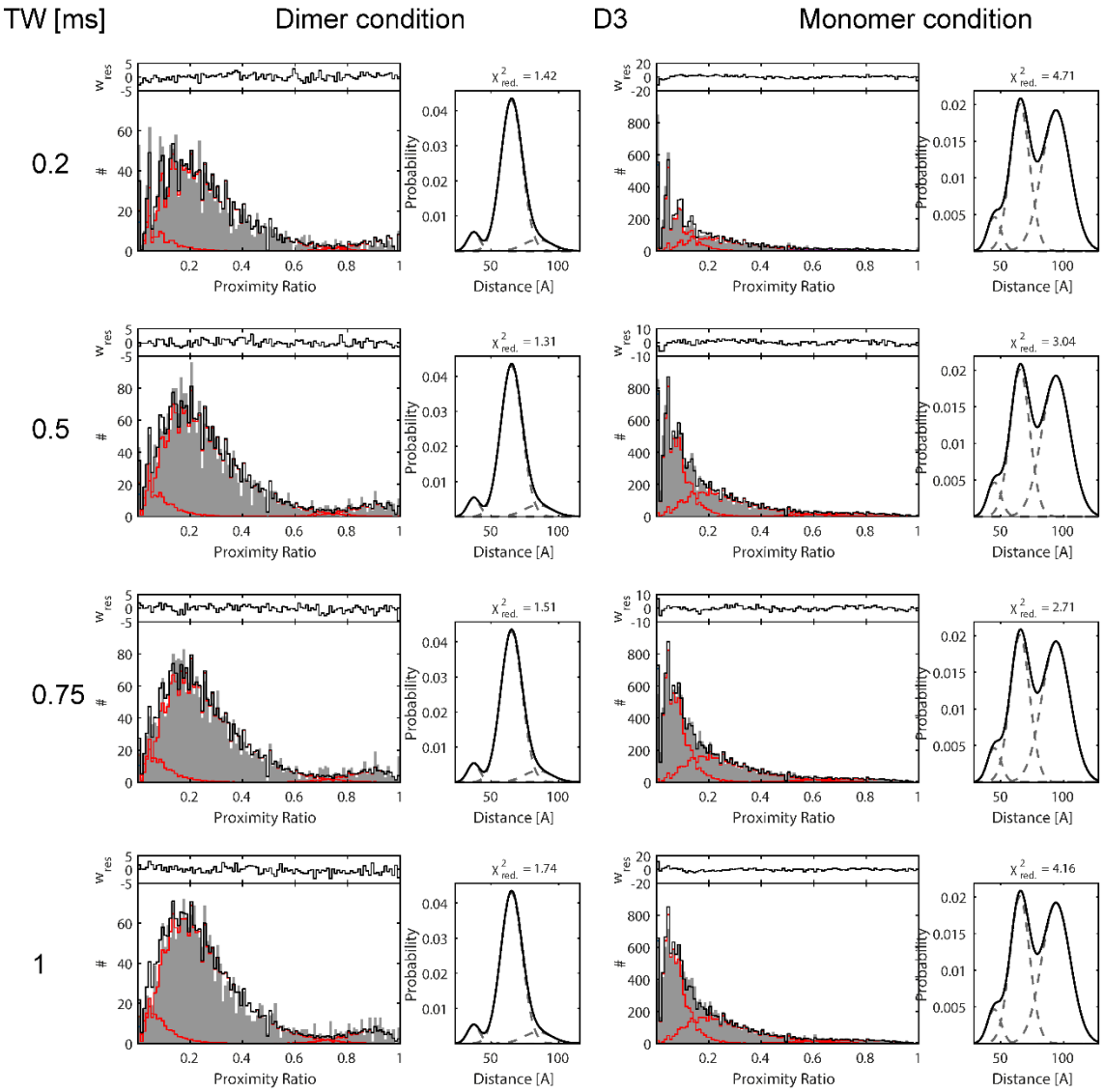
Data File 2: Global fitting of different time windows from the same derivative in both dimer and monomer condition of respectively D2, D3 and D5. (Related to figure S5)
Detailed illustration of PDA of derivatives D2, D3 and D5 in dimeric and monomeric conditions. Burst data were binned into either of four time bins (left of the histogram). PDA analysis was performed as described in the STAR Methods. Correction parameters were γ-factor = 0.8, β-factor = 0.010, B_{BB} = 1.23 kHz, B_{RR} = 0.43 kHz and R_0 = 54.7 Å. Individual χ^2 values are indicated in the figure. Both for the dimer and monomer, no extra states visually appeared when decreasing the bin size, and the 4 histograms could be modelled perfectly with the same state fractions, distances and distance distribution widths. Both of these observations are strongly indicative of the absence of FRET dynamics on the 1-10-ms time scale, yet do not exclude FRET dynamics below or above that time scale.

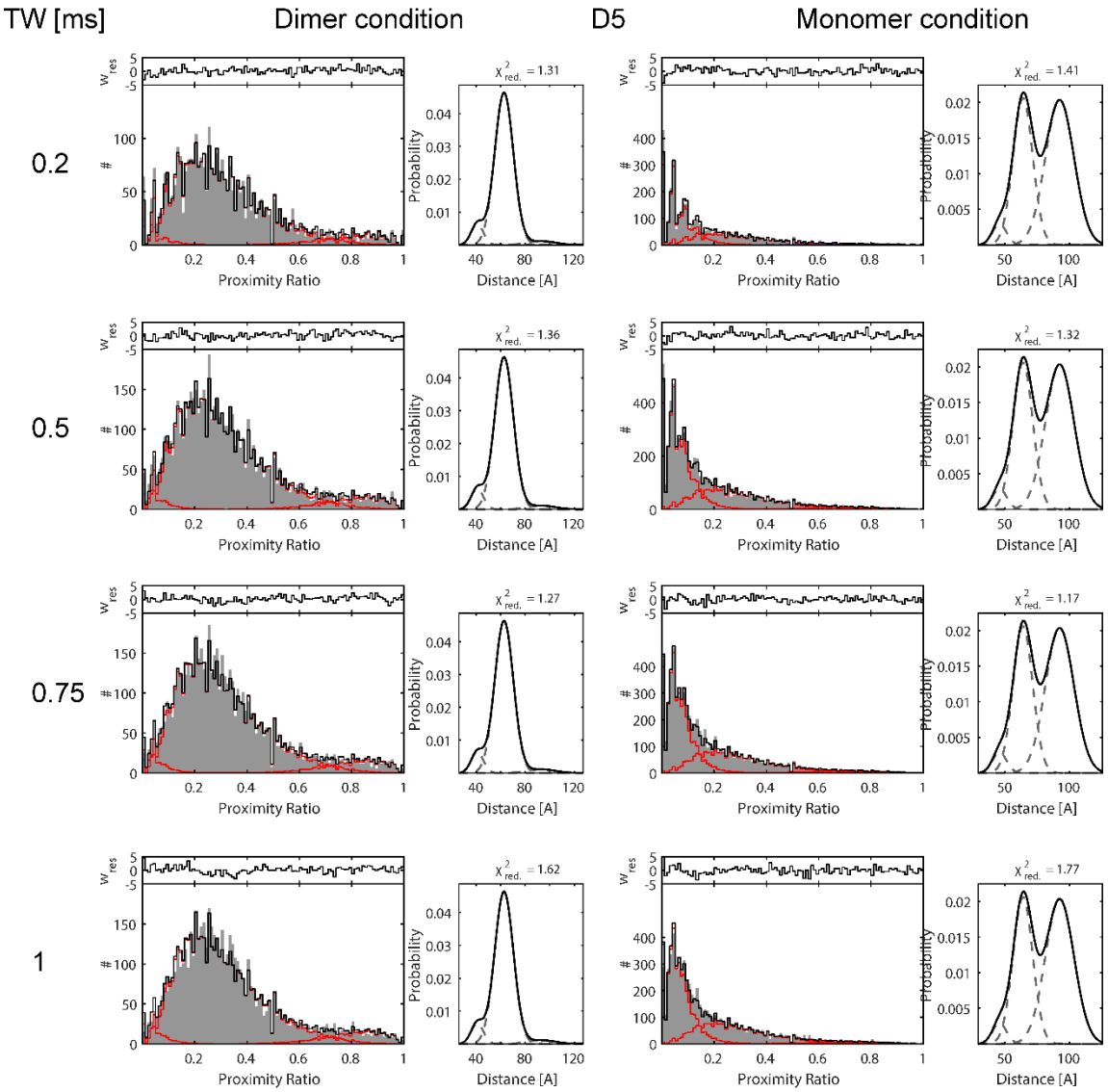
DATA FILE 1


A
-
B
300pM
C
500pM
D
700pM
E
900pM
F
2nM
G
4nM
H
6nM
I
8nM
J
10nM
Proximity Ratio
Distance [A]
Probability
K
Area
Kd=2.19 nM
Concentration ColdSecA (nM)
L
Raw data FCS
Subensemble FCS
Burstwise FCS
D3 Dimer
D3 Monomer
G(τ)
time lag τ [s]
D [µm²/s]
FRET Efficiency

DATA FILE 2


TW [ms]
Dimer condition
D2
Monomer condition
0.2
0.5
0.75
1
W res
#
Proximity Ratio
Probability
Distance [A]
χ²red. = 2.94
χ²red. = 1.37
χ²red. = 2.21
χ²red. = 1.06
χ²red. = 1.40
χ²red. = 1.56
χ²red. = 3.00
χ²red. = 1.42


TW [ms]
Dimer condition
D3
Monomer condition
0.2
0.5
0.75
1
W res
#
Proximity Ratio
Probability
Distance [A]
χ²red. = 1.42
χ²red. = 4.71
χ²red. = 1.31
χ²red. = 3.04
χ²red. = 1.51
χ²red. = 2.71
χ²red. = 1.74
χ²red. = 4.16


TW [ms]
Dimer condition
D5
Monomer condition
0.2
0.5
0.75
1
W res
#
Proximity Ratio
Probability
Distance [A]
X²red. = 1.31
X²red. = 1.41
X²red. = 1.36
X²red. = 1.32
X²red. = 1.27
X²red. = 1.17
X²red. = 1.62
X²red. = 1.77