

Journal of Visualized Experiments

Förster Resonance Energy Transfer Mapping: A New Methodology to Elucidate Global Structural Features --Manuscript Draft--

Article Type:	Invited Methods Collection - JoVE Produced Video
Manuscript Number:	JoVE63433R2
Full Title:	Förster Resonance Energy Transfer Mapping: A New Methodology to Elucidate Global Structural Features
Corresponding Author:	Ishita Mukerji, Ph.D. Wesleyan University Middletown, CT UNITED STATES
Corresponding Author's Institution:	Wesleyan University
Corresponding Author E-Mail:	imukerji@wesleyan.edu
Order of Authors:	Jack Northrop Donald Oliver Ishita Mukerji, Ph.D.
Additional Information:	
Question	Response
Please specify the section of the submitted manuscript.	Biochemistry
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (\$1400)
Please indicate the city, state/province, and country where this article will be filmed . Please do not use abbreviations.	Middletown, Connecticut, United States
Please confirm that you have read and agree to the terms and conditions of the author license agreement that applies below:	I agree to the Author License Agreement
Please confirm that you have read and agree to the terms and conditions of the video release that applies below:	I agree to the Video Release
Please provide any comments to the journal here.	

TITLE:

Förster Resonance Energy Transfer Mapping: A New Methodology to Elucidate Global Structural Features

AUTHORS AND AFFILIATIONS:

Jack Northrop¹, Donald B. Oliver^{1,2}, Ishita Mukerji^{1,2}

¹Molecular Biology and Biochemistry Department

²Molecular Biophysics Program, Wesleyan University, Middletown, CT, USA

Email addresses of co-authors:

Jack Northrop (jnorthrop@wesleyan.edu)

Donald B. Oliver (doliver@wesleyan.edu)

Ishita Mukerji (imukerji@wesleyan.edu)

Corresponding author:

Ishita Mukerji (imukerji@wesleyan.edu)

KEYWORDS:

fluorescence, energy transfer, SecA, protein translocation, FRET mapping

SUMMARY:

The study details the methodology of FRET mapping including the selection of labeling sites, choice of dyes, acquisition, and data analysis. This methodology is effective at determining binding sites, conformational changes, and dynamic motions in protein systems and is most useful if performed in conjunction with existing 3-D structural information.

ABSTRACT:

Förster resonance energy transfer (FRET) is an established fluorescence-based method used to successfully measure distances in and between biomolecules *in vitro* as well as within cells. In FRET, the efficiency of energy transfer, measured by changes in fluorescence intensity or lifetime, relates to the distance between two fluorescent molecules or labels. Determination of dynamics and conformational changes from the distances are just some examples of applications of this method to biological systems. Under certain conditions, this methodology can add to and enhance existing X-ray crystal structures by providing information regarding dynamics, flexibility, and adaptation to binding surfaces. We describe the use of FRET and associated distance determinations to elucidate structural properties, through the identification of a binding site or the orientations of dimer subunits. Through judicious choice of labeling sites, and often employment of multiple labeling strategies, we have successfully applied these mapping methods to determine global structural properties in a protein-DNA complex and the SecA-SecYEG protein translocation system. In the SecA-SecYEG system, we have used FRET mapping methods to identify the peptide-binding site and determine the local conformation of the bound signal sequence peptide. This study outlines the steps for performing FRET mapping studies, including identification of appropriate labeling sites, discussion of possible labels including non-

native amino acid residues, labeling procedures, how to perform measurements, and interpreting the data.

INTRODUCTION:

For proteins, elucidation of dynamics along with 3-dimensional (3-D) structural knowledge leads to an enhanced understanding of structure-function relationships of biomolecular systems. Structural methods, such as X-ray crystallography and cryogenic electron microscopy, capture a static structure and often require the determination of multiple structures to elucidate aspects of biomolecule binding and dynamics¹. This article discusses a solution-based method for mapping global structural elements, such as binding sites or binding interactions, that are potentially more transient and less easily captured by static methods. Strong candidate systems for this methodology are ones in which a 3-D structure has been previously determined by X-ray crystallography, NMR spectroscopy, or other structural methods. In this case, we take advantage of the X-ray crystal structure of the SecA-SecYEG complex, a central player in the protein general secretory pathway, to map the location of a signal peptide binding site using Förster resonance energy transfer (FRET) prior to the transport of the preprotein across the membrane². Manipulation of the biological system through genetic modifications coupled with our knowledge of the 3-D structure enabled the determination of the conformation of the signal sequence and early mature region immediately prior to insertion into the channel³.

FRET involves the radiation less transfer of energy from one molecule (donor) to another (acceptor) in a distance-dependent fashion that is through space^{4,5}. The efficiency of this transfer is monitored through either a decrease in donor or an increase in acceptor fluorescence intensity. The efficiency of energy transfer can be described as

$$E = R_0^6 / (R_0^6 + R^6)$$

in which the R_0 value is the distance at which the transfer is 50% efficient⁶. The technique has previously been described as a molecular ruler and is effective at determining distances in the 2.5–12 nm range, depending on the identity of the donor-acceptor dyes^{4,7–9}. The donor fluorescence intensities and lifetimes with or without acceptor allow determination of transfer efficiencies and consequently, distances^{5,8}. Due to the availability of the technology, sensitivity of the method, and ease of use, FRET has also found broad application in such areas as single-molecule fluorescence spectroscopy and confocal microscopy⁶. The advent of fluorescent proteins such as green fluorescent protein has made the observation of intracellular dynamics and live-cell imaging relatively facile^{10,11}. Many FRET applications such as these are discussed in detail in this study.

In this study, we particularly focus on the use of FRET measurements to yield distance values to determine structural details. Previously, FRET measurements have been effectively used to determine the conformation of DNA molecules when bound to protein^{12–14}, the internal dynamics of proteins, and protein binding interactions^{15–17}. The advantages of this method lie in the ability to determine flexible and dynamic structural elements in a solution with relatively low amounts of material. Significantly, this method is particularly effective when used in conjunction with existing structural information and cannot be used as a means of 3-D structure determination. The method provides the best insight and refinement of structure if the work

builds on existing structural information often coupled with computational simulation^{18,19}. Here, the use of distances obtained from steady-state and time-resolved FRET measurements is described to map a binding site, the location of which was not known, on an existing crystallographic structure of the SecA-SecYEG complex, major proteins in the general secretory pathway³.

The general secretory pathway, a highly conserved system from prokaryotes to eukaryotes to archaea, mediates the transport of proteins either across or into the membrane to their functional location in the cell. For Gram-negative bacteria, such as *E. coli*, the organism used in our study, proteins are inserted into or translocated across the inner membrane to the periplasm. The bacterial SecY channel complex (termed the translocon) coordinates with other proteins to translocate the newly synthesized protein, which is directed to its correct location in the cell through a signal sequence typically located at the N-terminus^{20,21}. For proteins bound for the periplasm, the ATPase SecA protein associates with the exit tunnel of the ribosome, and with the preprotein after approximately 100 residues have been translated²². Along with the SecB chaperone protein, it maintains the preprotein in an unfolded state. SecA binds to the SecYEG translocon, and through many cycles of ATP hydrolysis, facilitates protein transport across the membrane^{23,24}.

SecA is a multi-domain protein that exists in cytosolic and membrane-bound forms. A homodimeric protein in the cytosol, SecA consists of a preprotein binding or cross-linking domain²⁵, two nucleotide-binding domains, a helical wing domain, a helical scaffold domain, and the two helix finger (THF)²⁶⁻²⁹ (**Figure 1**). In previous crystallographic studies of the SecA-SecYEG complex, the location of the THF suggested that it was actively involved in protein translocation and subsequent cross-linking experiments with the signal peptide further established the significance of this region in protein translocation^{30,31}. Previous studies, using the FRET mapping methodology, demonstrated that exogenous signal peptides bind to this region of SecA^{2,32}. To fully understand the conformation and location of the signal sequence and early mature region of the preprotein prior to insertion into the SecYEG channel, a protein chimera in which the signal sequence and residues of the early mature region were attached to SecA through a Ser-Gly linker was created (**Figure 1**). Using this biologically viable construct, it was further demonstrated that the signal sequence and early mature region of the preprotein bind to the THF in a parallel fashion². Subsequently, the FRET mapping methodology was used to elucidate the conformation and location of the signal sequence and early mature region in the presence of SecYEG as described below³.

Knowledge of the 3-D structure of the SecA-SecYEG complex³³⁻³⁵ and the possible location of the binding site allowed to judiciously place donor-acceptor labels in locations where the intersection of individual FRET distances identifies the binding site location. These FRET mapping measurements revealed that the signal sequence and the early mature region of the preprotein form a hairpin with the tip located at the mouth of the SecYEG channel, demonstrating that the hairpin structure is templated prior to channel insertion.

PROTOCOL:

1. Selection of labeling sites

1.1 Identify at least three potential labeling sites to triangulate the putative binding site on the existing protein structures. In this case, SecA, SecYEG, and preprotein attached to SecA through genetic fusion were identified².

1.1.1 Choose labeling sites within 25–75 Å of the putative binding site and in relatively static regions of the protein, the distance will determine the specific FRET dye pair to be used³⁶. Locate the labeling sites in protein regions that are relatively distinct from each other, so the sites describe the vertices of a triangle with the putative binding site located in the center (**Figure 1A–D**).

1.1.2 Introduce or identify cysteine (Cys) residues at labeling sites of interest in a protein that has no other Cys residues, to improve the labeling efficiency of the specific site^{37,38}.

1.1.3 Introduce unnatural amino acids, e.g., p-azidophenylalanine for labeling with click chemistry, in order to effectively label one protein at two distinct positions with different dyes^{39,40}.

1.1.4 Test the Cys mutants for loss of function. Verify the activity of the Cys-less mutant and the unnatural amino acid mutant using an appropriate activity assay. In this case, activity was verified with a growth assay followed by an *in vitro* malachite green ATPase assay^{32,41,42}.

2. Labeling the protein

2.1 Purify the protein or proteins of interest to at least 95% purity for accurate labeling. Purify SecA and SecYEG proteins following protocols detailed in reference³. Ensure that you have at least 5 µg of purified protein for this step, as some protein will be lost during the labeling process.

2.2 Choose two dyes for FRET measurements depending on their R_0 value and the predicted distances between labeled sites. Estimate R_0 values and observe donor emission and acceptor absorbance overlap using the information from the fluorescent protein database, which also gives spectra for commonly used dyes (<https://www.fpbases.org/spectra/>)³⁶.

NOTE: R_0 is defined as the distance at which the transfer efficiency is 50% for a given dye pair. For mapping experiments, predicted distances should be close to the R_0 value of the dye pair to ensure that distances can be measured accurately.

2.3 Label the positions identified in step 1.1.1. with the donor-acceptor dye pair.

2.3.1 Label the protein according to the manufacturer's instructions with particular attention paid to parameters such as optimal protein concentration, temperature, pH, length of time, and buffer for the specific dyes used^{43,44}.

2.3.2 Prepare the protein at a concentration of 1–2 mg/mL or approximately 10 μ M in a 25 mM Tris-HCl (pH 7.5), 25 mM KCl, 1 mM EDTA (TKE) buffer. Dissolve dye in dimethylformamide (DMF) or dimethylsulfoxide (DMSO) to a final concentration of 1 mM. Add the dye dropwise to the solution while stirring to reach a dye: protein molar ratio of 5:1 (50 μ M dye: 10 μ M protein).

2.3.3 Allow the reaction to proceed for 4 h at room temperature (RT) in a glass vial with gentle rocking or overnight at 4 °C. Stop the reaction by adding β -mercaptoethanol.

NOTE: If protein is not compatible with Tris buffer, phosphate or HEPES buffers can be used. Maintain pH in the 7.0–7.5 range. If the protein has disulfide bonds, add a reducing agent such as DTT or TCEP prior to labeling. Remove DTT by dialysis or gel filtration before adding dye.

2.4 For accurate FRET measurements remove the free dye with a centrifugal concentrator with an appropriate molecular weight cut-off (MWCO) to let the free dye flow through while retaining the labeled protein.

2.4.1 Prepare the concentrator membrane by placing ~ 1 mL of water in the upper part of a 3 mL concentrator and then centrifuge the water through the membrane (at least 10 min at 4,300 $\times g$).

2.4.2 Remove free dye by centrifuging labeled sample in concentrator (20 min at 4,300 $\times g$). Repeat 3–4 times and dispose of flow through.

2.4.3 Check labeling efficiency of the labeled protein using UV-Vis absorption spectroscopy.

NOTE: FRET measurements require labeling efficiencies of 50% or greater. Lower labeling efficiencies reduce the FRET signal and can lead to inaccuracies in measurement.

2.4.4 Obtain a UV-Vis spectrum of the labeled protein with a range from 250–700 nm to observe both the protein absorption band and the dye maximum absorption band. Measure the absorbance at the absorption peak of dye and at 280 nm for protein.

2.4.5 Determine the concentration of protein and correct for any contributions from the dye using the correction factor, CF, and the following equations^{45,46}:

$$C = \frac{A_{280} - (A_{max} * CF)}{\epsilon_{protein}}$$

where C is the concentration of the protein (M), A_{280} is the sample absorbance at 280 nm, A_{max} is the absorbance at the dye absorption maximum, $\epsilon_{protein}$ is the extinction coefficient for the protein at 280 nm and CF is the correction factor, A'_{280}/A'_{max} , where A'_{280} is the absorbance at 280 nm and A'_{max} is the absorbance at the peak maximum for the dye only.

2.4.6 Determine labeling efficiency using the following equation:

$$E = \frac{A_{max}}{\epsilon_{dye} * C}$$

where ϵ_{dye} is the molar extinction coefficient of the dye, C is the concentration of the protein as determined in step 2.4.5, and E is the labeling efficiency. Repeat step 2.4.2 until the labeling efficiency value has plateaued and is less than 100%.

3. Determine the R_0 values

3.1. Measure the R_0 values *in situ*. Prepare two protein samples at the same concentration of total protein, 4 μ M, one with the protein labeled with the donor dye only and one with the protein labeled with the acceptor dye only. For SecA, a protein concentration of 4 μ M SecA monomer works well for these measurements.

3.1.1. Prepare sample volumes of 2.5 mL for a 1 cm x 1 cm cuvette, 600 μ L for a 5 mm x 5 mm cuvette or 200 μ L for a 3 mm x 3 mm cuvette.

3.2. Turn on the fluorometer and open the spectral acquisition and analysis program in the fluorescence software if using a spectrofluorometer. Click on the red **M** to connect the computer to the instrument (**Figure 2A**) and choose **Emission Spectra**.

3.2.1. Enter scan parameters using the **Collect Experiment** menu item, such as excitation wavelength, the range for emission scan, temperature, and sample changer position (**Figure 2B**).

3.2.2. Click on **RTC** and optimize instrument settings (e.g., spectral slits) by monitoring the fluorescence emission at the peak using an excitation wavelength set at the dye's absorption maximum. For SecA, set the following settings: bandpass as 1 nm; excitation and emission slits as 1 and 1.5 mm respectively, with the temperature at 25 °C and stirring speed at 250 rpm.

NOTE: Do not exceed the counts per second (cps) capacity of the instrument (typically 2×10^6 cps).

3.3 Place the donor labeled protein sample in the sample holder and click **Run** to generate an emission scan of the protein labeled with the donor dye only (donor only protein) by exciting the sample at the dye absorption maximum (e.g., 488 nm for AF488) and scanning over the emission peak (505–750 nm for donor only SecA protein labeled with AF488).

3.4 Establish a baseline for the scan by extending the scan 25–50 nm past the end of the peak. Measure the quantum yield of the donor-only protein, by performing absorption and fluorescence measurements on samples of different concentrations as described⁴⁷. Maintain the same slit settings for these measurements.

3.4.1 Use free donor dye as the reference for the quantum yield. Obtain at least four measurements of the donor-only protein and the free dye at different concentrations for an accurate determination.

3.4.2 Plot the fluorescence intensity or integrated area versus the absorbance for the donor-only protein and the free dye or reference. Determine the slopes for the donor-only protein ($Slope_D$) and the reference ($Slope_R$).

3.4.3 Quantum yield (ϕ) is calculated using the following equation:

$$\phi_D = \phi_R \left(\frac{Slope_D}{Slope_R} \right) \left(\frac{\eta_D^2}{\eta_R^2} \right)$$

here ϕ_D is the quantum yield of the donor only protein, ϕ_R is the quantum yield of the free dye (this can usually be obtained from the manufacturer), $Slope_D$ and $Slope_R$ are the slopes determined in step 3.4.2 for the donor only protein and reference, respectively and η_D and η_R represent the index of refraction of the donor-only protein and the reference-free dye solutions, respectively⁴⁷.

3.5 Obtain an absorption spectrum of your acceptor-only protein using a 1 cm pathlength cell. Generate an extinction coefficient spectrum of your acceptor-only protein by dividing the absorption spectrum by the dye concentration.

3.6 Generate the spectral overlap integral, $J(\lambda)$ using a graphical analysis program. A standard worksheet program (e.g., spreadsheet) can also be used for this process.

3.6.1 Multiply the fluorescence emission spectrum of the donor-only protein (step 3.4) by the extinction coefficient spectrum of the acceptor-only protein to generate the overlap spectrum.

3.6.2 Multiply the resultant overlap spectrum by λ^4 .

3.6.3 Determine the area under the curve by integration of the overlap region. The overlap region is defined as the area where the donor emission spectrum multiplied by the acceptor extinction coefficient spectrum yields positive values. The spectral overlap integral is defined as:

$$J(\lambda) = \int_0^{\infty} F_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda$$

where $F_D(\lambda)$ is the emission spectrum of the donor-only protein (obtained in step 3.4) and $\varepsilon_A(\lambda)$ is the extinction coefficient spectrum of the acceptor-only protein and has units of $M^{-1}cm^{-1}$ (obtained in step 3.5). The resulting spectral overlap integral should have units of $M^{-1}cm^{-1}nm^4$.

3.6.4 Normalize the spectral overlap integral. Divide the overlap integral by the integrated area of the donor only protein spectrum over the same spectral range:

$$J(\lambda) = \frac{\int_0^{\infty} F_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda}{\int_0^{\infty} F_D(\lambda) d\lambda}$$

3.6.5 Calculate the R_0 value in Å using the following equation:

$$R_0 = 0.211[\kappa^2\eta^{-4}Q_DJ(\lambda)]^{\frac{1}{6}}$$

where κ^2 is the orientation factor, typically taken as 2/3 for freely rotating dyes, η is the index of refraction and can be approximated as 1.33 for dilute aqueous solutions, Q_D is the quantum yield of the donor (step 3.4) and $J(\lambda)$ is the spectral overlap integral as determined in step 3.6.3⁵.

NOTE: If the dyes are not freely rotating, corrections can be introduced as described by Ivanov⁴⁸ and implemented by Auclair⁴⁹ and Zhang^{2,3}.

4. Perform FRET spectral measurements

4.1. Prepare donor-only protein, acceptor-only protein, and donor-acceptor protein samples at the same concentration; a concentration of 4 μ M is recommended. Use 200 μ L of solution, if using a 3 mm x 3 mm cuvette, 600 μ L if using a 5 mm x 5 mm cuvette, or 2.5 mL if using a 1 cm x 1 cm cuvette.

4.1.1. Prepare the donor-acceptor protein sample by using equal molar amounts of the donor only and acceptor only protein.

4.1.2. Maintain the same amount of labeled sample in control donor only and acceptor only protein samples through the introduction of unlabeled protein in equal molar amount to either the donor only or acceptor only samples. For example, for solutions of the same concentration, each donor-only FRET sample would contain 100 μ L of donor-only protein and 100 μ L of unlabeled protein for a 200 μ L volume.

4.2. Generate fluorescence emission spectra of the donor only, acceptor only, and donor-acceptor samples. Optimize the signal as described in step 3.2. Once optimized maintain the same settings for all of the samples.

4.2.1. Obtain the donor-only scan as described in step 3.3. Excite the solution at the donor dye absorption maximum and scan over the donor and (expected) acceptor emission peaks.

4.2.2. Either exchange the sample to the acceptor-only protein or change the sample changer position to the cuvette containing the acceptor-only protein.

4.2.3. Obtain an emission scan of the protein labeled with acceptor dye only (acceptor only protein) using the same settings as in step 4.2.1. Excite the sample at the donor excitation wavelength.

NOTE: This spectrum provides a correction for the amount of acceptor excited at the donor wavelength (F_A in step 5.1.2)

4.2.4. Exchange the sample to the donor-acceptor protein sample or change the sample changer position to the cuvette containing the donor-acceptor labeled protein.

4.2.5. Obtain an emission scan of the donor-acceptor protein sample using the same settings as in steps 4.2.1 and 4.2.3.

4.2.6. For all spectra, correct for background fluorescence by subtracting the background counts measured at the end of the scan.

4.3. Measure the donor lifetimes of the donor only and donor-acceptor samples prepared as described in step 4.1.2. Use a time-correlated single-photon counting fluorescence instrument capable of measuring and resolving fluorescent decays in the nanosecond (10^{-9} s) time range.

NOTE: For FRET dye pairs, match the excitation light source to the absorption maximum of the donor dye.

4.3.1. Turn on the instrument. Open the acquisition software, use the instrument control software for data acquisition with the fluorescence spectrometer.

4.3.2. For acquisition, select **TCSPC Decay**, with a time range of 55 ns, a gain of 1 and 4096 channels.

4.3.3. Obtain an instrument response function (IRF) using a solution of non-dairy creamer or commercial scattering solution and monitor the scattering at 490 nm. Adjust the slit setting and use neutral density filters as needed to maintain a low enough count rate to avoid pulse pile-up⁵. Click **Accept** and then **Start**. This will start the acquisition.

NOTE: A maximum count rate of 4000 cps is used for a 180 kHz repetition rate.

4.3.4. Collect the IRF at 490 nm until the peak channel has a maximum of 20,000 counts. Collect an IRF before and after measuring each fluorescence decay.

4.3.5. Obtain the fluorescence decays of the donor only and donor-acceptor samples by monitoring the fluorescence emission at the donor emission wavelength, 520 nm.

4.3.6. Adjust slit settings for a maximum count rate of 4000 cps or less. Slit settings are typically 15–20 nm bandpass for protein samples. Collect the decay until 20,000 counts are obtained in the peak channel.

4.4. Analyze the decay or the fluorescence intensity (I) as a function of time (t) for the fluorescence lifetime (τ). Fit the decay to a sum of exponentials with the following equation:

$$I(t) = \sum_i \alpha_i e^{-t/\tau_i}$$

where α_i is the preexponential factor of the i^{th} component and τ_i is the lifetime. The fit is reconvolved with the IRF to match the fluorescence decay. Judge the quality of the fit from the reduced χ^2 parameters.

5. Analysis of FRET data

5.1. Calculate FRET efficiency from the decrease in donor intensity of the donor-acceptor sample relative to the donor only with the following equation.

$$E = 1 - \frac{F_{DA}}{F_D}$$

where F_{DA} is the fluorescence intensity of the donor-acceptor sample and F_D is the fluorescence intensity of the donor only sample at the peak of the donor fluorescence. Use the integrated areas of the peaks if the data is noisy.

5.1.1. Correct for any differences in labeling between the donor only and donor-acceptor samples. Calculate corrections based on the donor degree of labeling as follows.

$$F'_D = F_D \frac{f_{DA}}{f_D}$$

where f_{DA} is the donor labeling efficiency in the donor-acceptor sample and f_D is the labeling efficiency in the donor-only sample.

5.1.2. Correct for any contributions of the acceptor fluorescence to the donor-excited spectrum through subtraction of the acceptor-only protein spectrum (step 4.2) from the donor-acceptor protein spectrum.

$$F'_{DA} = F_{DA} - F_A$$

5.1.3. Correct for the differences in labeling efficiency of the acceptor only protein relative to the donor-acceptor protein sample yielding the following equation for calculating efficiency:

$$E = \left(1 - \frac{F'_{DA}}{F'_D}\right) \frac{1}{f_A}$$

where f_A denotes the fractional amount of acceptor labeling. This equation includes all corrections due to dye labeling and acceptor fluorescence.

5.1.4. Calculate FRET distances from the efficiencies using the following equation:

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

using the R_0 value obtained in step 3.6.5.

5.1.5. Calculate FRET efficiency using fluorescence lifetimes of the donor only and donor-acceptor samples measured in step 4.3.3-4.3.5:

$$E = 1 - \frac{\tau_{DA}}{\tau_D}$$

5.1.6. Use the amplitude-weighted lifetime to calculate FRET efficiencies and compare with steady-state results⁵.

$$\langle \tau \rangle = \sum_i \alpha_i \tau_i$$

5.1.7. Calculate the distance as in step 5.1.4 from the efficiencies determined by fluorescence lifetime. Compare steady-state and time-resolved values for FRET efficiencies and distances and ensure that they are within error of each other.

6. Mapping the distances

6.1. Use the calculated distances to map the binding site on the three-dimensional structure. Calculate the distances and errors for all of the dye pairs and locations examined using the equation given in step 5.1.4 and R_0 values obtained in step 3.6.5 for each FRET pair.

6.1.1. Use a 3-D graphical viewing program such as PyMOL⁵⁰ to map the distances onto the structure (script given in **Supplementary File**). Commands from the script can be directly entered into the command window with the appropriate distance information.

6.1.2. Generate a shell for each distance measured and the associated error (**Figure 3, Figure 4, Supplementary Figures 1–3**).

6.1.3. Map the position through the intersection of the different shells (**Supplementary Figures 1–3**). The signal peptide binding site was mapped through the three different locations on SecA and SecYEG and four different locations on the signal peptide (**Figure 1**).

REPRESENTATIVE RESULTS:

This study focused on determining the location of the preprotein binding site on SecA prior to insertion of the preprotein into the SecYEG channel. To map the binding site, FRET experiments were performed between different regions of the preprotein and three distinct locations on the SecA and SecYEG proteins (**Figure 1A–D**). From the distances obtained and three-dimensional structures of SecA, SecYEG, and the preprotein, the location of the preprotein binding site was predicted. Rather than employing three separate entities (SecA, SecYEG, and preprotein) to perform these measurements, the PhoA signal sequence was attached to SecA through genetic modification after incorporation of a Gly-Ser linker^{2,3}. For facile labeling with dyes, Cys residues or amber mutations were introduced at residues 2, 22, 35, and 45 in the PhoA preprotein (**Figure 1E**).

Identification of sites and labeling

A putative binding site for the signal sequence had been previously identified using similar FRET mapping methods^{2,32}. These and other studies had identified the two-helix finger (THF) and the preprotein cross-linking domain as possible binding sites of the signal peptide with a suggested orientation in a parallel position to the THF^{33,51–54} (**Figure 1F**). Thus, identification of potential labeling sites on the SecA and SecYEG proteins was done based on the location of the putative

binding site in the SecA and SecYEG crystal structure (shown in green in **Figure 1A–D**). Three sites were chosen to triangulate the position of the putative binding site, where essentially the three sites form a triangle around the putative binding site. As shown in **Figure 1**, the three sites were within the FRET range of the binding site (50–70 Å). The dye pair of Alexa Fluor 488 (AF488) and Alexa Fluor 647 (AF647) was chosen, as the R_0 value of 55.7 Å³⁶ corresponds well with the expected distances between the labeled sites and the putative binding site ensuring measurement accuracy.

The three sites chosen for labeling, SecA37, SecA321, and SecY292 (shown as magenta, violet, and cyan spheres in **Figure 1A–D**) are located throughout the protein complex forming a triangle around the putative binding site. The three sites were separately mutated to Cys residues in a Cys-less mutant to ensure that only the correct position was labeled^{2,37}. For the SecY292 experiments, the PhoA preprotein sites were labeled with AF647 and SecY residue 292 was labeled with AF488 using maleimide chemistry. In the chimeric protein, sites SecA37 and SecA321 were labeled with AF647 and the preprotein was labeled with AF488. In the SecA-PhoA chimeric protein, residues 2, 22, 37, and 45 of the PhoA preprotein segment were each mutated to an amber codon in individual proteins. The amber codon mutations allowed the introduction of unnatural amino acid, p-azidophenylalanine, at those positions, which were subsequently labeled with the AF488 using click chemistry^{39,40}. Each mutation was generated and labeled independently to ensure correct, differential labeling of the protein components. The degree of labeling was determined for all proteins and generally needed to be 50% or better in order to proceed with the sample.

Determination of transfer efficiencies and distances

Prior to performing the energy transfer measurements, the donor quantum yield, overlap integral and R_0 values were determined (steps 3.4–3.6). The donor quantum yield was measured relative to the dye, fluorescein, which has a quantum yield of 0.79 in 0.1 M NaOH⁴⁷. Absorption and fluorescence emission spectra were obtained at a series of concentrations to generate a linear plot of absorption relative to fluorescence emission intensity to determine the quantum yield. In these measurements, it is critical to measure the absorption in the linear range (0.1–1.0) and all emission measurements need to be generated with the same slit settings. As these values are used to determine overlap integrals and R_0 values, they should be measured under FRET conditions. Protein local environment profoundly affects dye emission and consequently, donor quantum yields should be measured for each of the sites investigated. We note that sites on SecA and SecYEG influence the R_0 values more strongly than those on the PhoA portion of the chimera. For dye pairs with the same SecA or SecYEG site, the R_0 values are typically within 5 Å of each other; whereas the R_0 values can differ by as much as 20 Å for the two different SecA locations (residues 37 vs. 321), underscoring the importance of determining R_0 values for each dye pair (**Table 1**).

The calculation of R_0 assumes that the donor and acceptor dyes are freely rotating and the degree to which the dyes do not rotate contributes to the overall uncertainty in the measurement. To appropriately take into account the relative motion of the dyes and their orientations, steady-state fluorescence anisotropy measurements were performed on all the donor and acceptor dyes

in the different labeling positions. These values, which were in the 0.10–0.21 range, were used to calculate the error associated with the distance measurements^{2,3,48}. The relatively high anisotropy values observed for both the donor and acceptor dyes correspond to a reduction in dye rotation, which is inconsistent with the assumption of free rotation. The lack of free rotation generates an error of 19%–25% in the distance calculations. As shown in **Table 1**, this led to an average uncertainty in the measured distances of at most ± 15 Å. When mapping the FRET distances, these uncertainties in the distance measurements are an important consideration, as discussed below.

The calculated distance between donor-acceptor pairs is based upon the relationship between efficiency and distance, in which a higher efficiency is indicative of donor-acceptor pairs separated by a shorter distance. To determine FRET efficiencies, fluorescence emission spectra excited at the donor excitation wavelength (488 nm) are obtained on donor-only and donor-acceptor samples. Typically, a reduction in donor emission intensity signifies the presence of energy transfer (step 5.1). **Figure 1G** depicts the donor-acceptor spectra for the SecA 37 residue with either residue 2 or 22 of the PhoA chimera. The SecA37 residue is labeled with AF647 or the acceptor dye, and the PhoA residues are labeled with AF488 or the donor dye. At either position, the donor fluorescence is reduced and a small increase in acceptor fluorescence intensity can be seen in the donor-acceptor samples. Since excitation is done at the donor excitation wavelength of 488 nm, which does not directly excite the acceptor, any acceptor fluorescence observed results from energy transfer. Thus, the decrease in donor intensity and concomitant increase in acceptor intensity results from energy transfer between the two dyes. Significantly, the donor fluorescence intensity is higher for the PhoA2 position (blue) relative to the PhoA22 position (yellow) in the presence of the acceptor. This relative difference in the donor intensity decrease indicates that energy transfer between the PhoA2 residue and the SecA37 residue is weaker than the transfer between the PhoA22 and SecA37 residues, which implies that the PhoA2 residue is located further away from SecA37 than PhoA22. Distances are determined from the relationship between efficiency and R_0 values (step 5.1.4).

Since the steady-state fluorescence measurements could represent an average of two or more distances, we also performed time-resolved fluorescence measurements. For these experiments, the lifetime of the donor dye is measured in the presence and absence of the acceptor (**Figure 1G**). If there were two distinct energy transfer processes contributing to the measured steady-state fluorescence efficiency, they would be observed as discrete lifetimes, provided they were resolvable within the time resolution of the instrument. To enhance the ability to resolve the lifetimes, 10,000 counts or more should be collected at the peak; however, the peak height or peak channel counts needs to be balanced with the time of acquisition and potential damage to the sample. Time-resolved measurements yielded single lifetimes for each donor-acceptor pair consistent with only one orientation or distance between the dyes. We note that small differences in the distance as observed in the areas revealed by our FRET mapping technique would not lead to resolvable lifetimes in our system. Moreover, the efficiencies as determined by the time-resolved fluorescence measurements were in good agreement with those determined from the steady-state measurements, providing further support that the measured efficiencies arise from only one distance between the dye pairs³.

Mapping the FRET distances onto the 3-dimensional structure

The resonance energy transfer measurements yield sufficient distance information to identify the binding site and orientation of the signal sequence on SecA. The three locations on SecA and SecYEG along with the four positions on the PhoA region of the SecA-PhoA chimera provide the 12 different distances used to map the binding site (**Table 1**). The twelve distances were mapped onto the three-dimensional X-ray cocrystal structure of the *Thermotoga maritima* SecA-SecYEG complex (PDBID: 3DIN) to identify the binding site of the signal sequence³³. The structure of the SecA-SecYEG complex is similar to that observed in *E. coli* as evidenced by an *in vivo* photocrosslinking study⁵⁵.

We use the beginning (PhoA2) and end (PhoA22) residues of the PhoA signal sequence in the SecA-PhoA chimera to illustrate how residue locations were identified on the SecA-SecYEG complex. As energy transfer can occur in all directions, the FRET distances and associated errors describe a spherical shell, with one of the dye locations from the donor-acceptor pair designated as the center. In this study, the residues SecA37, SecA321, and SecY292 form the centers of three spherical shells that describe the location of the PhoA2 residue of the signal sequence. Visualization of the overlapping regions arising from the three separate locations, SecA37 (magenta), SecA321 (purple), and SecY292 (cyan) are shown in **Figure 3**. Only a portion of each FRET shell intersects with the protein structure, and the residues and backbone that fall within that shell are highlighted. Thus, the protein regions within the shell defined by the SecA37-PhoA2 distance are shown in magenta (**Figure 3A,E**), while the regions defined by the SecA321-PhoA2 and SecY292-PhoA2 shells are shown in purple (**Figure 3B,F**) and cyan (**Figure 3C,G**), respectively. The putative binding site, consisting mainly of the two-helix finger, is shown in green.

As shown in **Figure 3**, each of these FRET shells defines a relatively large section of the protein complex. For all three locations, the FRET shell does intersect with the putative binding site; however, for the SecA321 residue, for example, the intersected area is smaller and lies towards the ends of the fingers with significant overlap with the helical scaffold domain. The intersection or the common area of all three FRET shells (**Figure 3D,H**), defines the location of the PhoA2 residue. This area is considerably smaller than each FRET shell and includes only a small portion of the THF with a large contribution from the helical scaffold. The scripts used for generating the FRET shells and the intersected areas for the molecular visualization program, PyMOL, are given in the Supplemental Information. Portions of the shells are visualized as pink dots on the SecA-SecYEG complex in **Supplementary Figures 1–3**.

A similar strategy was used to identify the location of the PhoA22 residue. The FRET shells defined by the PhoA22 FRET distances (**Figure 4A–C,E–G**) describe a smaller area relative to the PhoA2 residue (**Figure 4D,H** vs. **Figure 3D,H**). We interpret this difference to suggest that the PhoA2 residue and associated region are more flexible and labile than PhoA22. Significantly, the area ascribed to the PhoA22 residue is located closer to the tip of the THF and the mouth of the SecYEG channel, with regions of SecY identified in the common area (**Figure 3D,H**). All three dye pairings identify regions along with the putative binding site; however, the intersected common areas center the PhoA22 location (**Figure 4D,H**) at the opposite end of the THF relative to the PhoA2 location (**Figure 3D,H**). These findings would suggest that the signal sequence of the preprotein

which extends from residues 2-22, lies along the THF in a relatively unstructured state. This result is consistent with earlier studies suggesting the signal sequence binds to the protein along the THF in an extended state and that the C-terminal end of SecA in the *B. subtilis* crystal structure essentially models the structure of the signal peptide and occupies the same location (shown in red **Figure 1F**)^{2,26}. We employed a similar approach to identify two locations in the early mature region (residues 37 and 45) of the SecA-PhoA preprotein chimera to further define the binding and orientation of the signal sequence and early mature region as discussed below. In other studies FRET distances have been used effectively to refine an existing structure or molecular dynamics simulation-derived model^{18,19,56}; we were not able to do this, as no structure for the signal sequence bound to SecA exists.

FIGURE AND TABLE LEGENDS:

Figure 1: Labeling sites in SecA-SecYEG complex with representative FRET spectra. (A–D) Four different views of the SecA-SecYEG co-crystal structure (PDBID: 3DIN)³³ in which the labeling sites of SecA37, SecA321, and SecY292 are shown as magenta, violet, and cyan spheres, respectively. A-C are side views of the complex and D is a top view. SecA is shown in light grey, SecYEG is shown in dark grey and the putative binding site, the THF, is shown in green. **(E)** Schematic of the SecA-PhoA chimera construct, which connects the SecA protein to the PhoA preprotein through a Ser-Gly linker (not drawn to scale). The labeling sites on the PhoA portion of the chimera are given in blue, green, yellow, and red, corresponding to residues 2, 22, 37, and 45. **(F)** Ribbon diagram of the crystal structure of *B. subtilis* SecA protein (PDBID: 1M6N) colored by domain where nucleotide-binding domains 1 and 2 are shown in blue and light blue, respectively, the preprotein cross-linking domain is shown in gold, the central helix in green, the two-helix finger in cyan, the helical wing domain in dark green and the C-terminal linker in red²⁶. The unstructured C-terminus serves as a model of the bound PhoA signal peptide based on a previous FRET mapping study². **(G)** Steady-state fluorescence spectra of the donor only and donor-acceptor samples of the SecA37-AF647 and PhoA2-AF488 FRET pair and the SecA37-AF647 and PhoA22-AF488 FRET pair. The reduction in donor intensity for the donor-acceptor sample is indicative of energy transfer. Greater energy transfer occurs from PhoA22 relative to the PhoA2 site based on the decrease in donor intensity. **(H)** Time-resolved fluorescence donor only (magenta) and donor-acceptor (light magenta) decay spectra of the SecA37-AF647 and PhoA22-AF488 FRET pair. The instrument response function is shown in grey. The donor-acceptor complex gives a shorter decay and consequently a faster lifetime consistent with energy transfer. All molecular structures were generated with the indicated PDB file and PyMOL⁵⁰. Figure 1E-H has been modified from Zhang et al.³.

Figure 2: User interface of the FluorEssence program. (A) The opening window is shown with a circle around the red M. This must be clicked to connect the program with the fluorometer. **(B)** The experiment set-up window illustrates the different areas (monos, detectors, accessories) where scan relevant parameters are entered.

Figure 3: FRET distance shells determined for the SecA-PhoA2 position. FRET distance shells constructed from the FRET distances and the associated uncertainties (**Table 1**) are depicted on the SecA-SecYEG complex (PDBID: 3DIN). **(A–C)** FRET distance shells for the PhoA2 location

constructed with SecA37 (magenta), SecA321 (violet), and SecY292 (cyan), respectively at the center position. The shells are colored according to the center residue. (D) The intersection of the three FRET shells defines the location of PhoA2, shown in blue. (E–H) Views are rotated approximately 180° from A–D. All molecular structures were generated with the indicated PDB file and PyMOL⁵⁰.

Figure 4: FRET distance shells determined for the SecA-PhoA22 position. FRET distance shells constructed from the FRET distances and associated uncertainties (Table 1) are depicted on the SecA-SecYEG complex (PDBID: 3DIN). (A–C) FRET distance shells for the PhoA22 location constructed with SecA37 (magenta), SecA321 (violet) and SecY292 (cyan), respectively at the center position. The shells are colored according to the center residue. (D) The intersection of the three FRET shells defines the location of PhoA22, shown in yellow. (E–H) Views are rotated approximately 180° from A–D. All molecular structures were generated with the indicated PDB file and PyMOL⁵⁰.

Figure 5: FRET-mapped locations of PhoA2, PhoA22, PhoA37, and PhoA45 projected onto the *B. subtilis* SecA – *Geobacillus thermodenitrificans* SecYE cocrystal structure (PDBID: 5EUL). (A) Coloring of SecA-SecYE as in Figure 1. The OmpA peptide substrate inserted at the end of the THF is shown in pink. FRET-mapped regions were generated in the presence of ATP-γS, with PhoA2 shown in blue, PhoA22 in green, PhoA37 in yellow, and PhoA45 in red. Overlap regions are shown in olive (PhoA22 and PhoA37) and orange (PhoA37 and PhoA45). The peptide substrate (residues 749-791, cyan) was excised from the original structure and modeled into the putative binding region without any alterations to the structure (circled in red). (B) Enlarged view of the modeled peptide substrate. Residues 2 (Lys), 22 (Tyr), and 37 (Gly) of the OmpA peptide substrate are depicted in stick form in blue, green, and yellow, respectively. These residues in the modeled peptides exhibit excellent agreement with the predicted FRET mapped locations. For clarity, the nanobody in the original structure has been omitted. This figure has been modified from Zhang et al.³.

Table 1: Transfer Efficiencies and Distances Determined for the SecA-PhoA-SecYEG complex. FRET efficiencies, distances, and R_0 values are given for the 12 distances used for mapping the preprotein binding site.

Supplementary Figure 1: FRET distance shell, shown in pink dots, determined for the SecA37 residue and the PhoA 37 residue on the SecA-SecYEG complex (PDBID: 3DIN). Sec A is shown in light grey, SecYEG is shown in dark grey and the SecA37 residue is shown in magenta.

Supplementary Figure 2: FRET distance shell, shown in pink dots, determined for the SecA321 residue and the PhoA 37 residue on the SecA-SecYEG complex (PDBID: 3DIN). Sec A is shown in light grey, SecYEG is shown in dark grey and the SecA321 residue is shown in violet.

Supplementary Figure 3: FRET distance shell, shown in pink dots, determined for the SecY292 residue and the PhoA 37 residue on the SecA-SecYEG complex (PDBID: 3DIN). Sec A is shown in light grey, SecYEG is shown in dark grey and the SecY292 residue is shown in cyan.

DISCUSSION:

Through the use of the FRET mapping methodology, we identified the signal sequence binding site on the SecA protein. Importantly, the presence of a 3-D crystal structure of the complex greatly facilitated our study. The strength of this mapping methodology lies in the ability to use an existing structure to identify locations for labeling. This methodology cannot be used to determine a 3-D structure; however, determination of structural elements⁵⁶, refinement of an existing structure⁴⁹, determination of a binding site location^{2,32}, or elucidation of dynamic motion⁵⁷, are all possible applications of this method.

In the SecYEG-SecA-PhoA complex, the three labeling sites form a triangle around the putative binding site (**Figure 1**). The employment of multiple distance measurements from the vertices of this triangle to the same residue refines the location information similar to GPS navigation methods. Three sites, SecA37, SecA341, and SecY292, were identified on SecA and SecY along with four sites, PhoA2, PhoA22, PhoA37, PhoA45 in the signal sequence and early mature region of the preprotein to give a total of 12 distances for mapping the location of the signal peptide (**Table 1**). Importantly, to improve the accuracy of the distance measurements, labeling sites should be located in relatively static regions of the protein, such as in secondary structure elements rather than loops. Furthermore, sites should also be in locations that are relatively accessible to solvent for ease and increased efficiency of labeling. Within the SecA-PhoA-SecYEG complex, the performance of distance measurements from the triangle sites or vertices to residues in the binding substrate were sufficient for locating the residues in the binding substrate to a relatively small area (**Figure 3D,H** and **Figure 4D,H**). Identification of the intersected area from the multiple distance measurements significantly refines the location from that of single distance measurement, as shown in **Figure 3** and **Figure 4**. Thus, when using this method, the measurement of multiple distances is strongly recommended. Although this method can identify regions of molecules involved in binding, for example, it does not provide precise structural information; such information is best obtained from other structural methods such as X-ray, NMR, and cryo-EM. FRET distances can be used to effectively refine an existing structure^{18,19} or model, in this case, that was not possible as no model of the signal sequence bound to SecA exists.

To ensure that dye labeling occurs at only desired locations and FRET distance measurements are accurate, the use of mutagenesis for labeling is preferred. Generation of a Cys-less mutant requires relatively conservative mutation of Cys residues to Ser or similar residues in the otherwise wild-type protein using site-directed mutagenesis methods. In the current study, Cys mutations were introduced into Cys-free versions of SecA and SecYEG for labeling³⁷. The activity of the mutant protein was verified with a growth assay followed by an *in vitro* malachite green ATPase assay^{41,42}. Relevant activity assays depend on the function of the protein, for example for DNA binding proteins, a DNA binding assay would be appropriate⁵⁸. Failure to ensure labeling at only one location can lead to labeling of more than one site with the same dye, which significantly complicates the distance determinations. Thus, the introduction of a second label on the same protein can be done through the incorporation of unnatural amino acid by site-directed mutagenesis. We employed this methodology to label the SecA-PhoA chimera at locations distinct from Cys residues by introducing the unnatural amino acid, p-azidophenylalanine and

labeling with click chemistry^{39,40}.

An additional important consideration of this method is the choice of dyes used and the associated R_0 value. After identification of the potential labeling sites, the distances to be measured can be estimated from the 3-D structure. With this information, investigators can choose dye pairs with R_0 values that span the desired range of expected distances. For example, the AF488-AF647 dye pair has a calculated R_0 of 55.7 Å which provides a good range for labeling sites located an estimated 40–75 Å away from the putative binding site. Although the R_0 values calculated in step 2.2 are useful for choosing which dye pair to use for your system, attachment of the dyes to the protein can alter their properties significantly. For greater accuracy, R_0 values should be calculated from *in situ* experiments performed with labeled protein (steps 3.1 – 3.6.5).

Measurement of transfer efficiency can be done either by monitoring steady-state fluorescence emission and observing either a decrease in donor emission or an increase in acceptor emission. Although observation of both effects is desirable, the efficiency can be calculated from either as described^{5,8}. Efficiency can also be calculated from the decrease in donor lifetime in the donor-acceptor sample relative to the donor-only sample. Determination of efficiency by more than one method is recommended, particularly the use of time-resolved methods to establish the relative homogeneity of the efficiencies measured.

The mapping method also allowed us to determine the relative *orientation* of the signal sequence and the early mature regions of the PhoA preprotein relative to SecA and the putative binding site. A SecA-SecYEG X-ray crystal structure and subsequent cryo-EM study provided clarity regarding the structure of the signal sequence and the early mature region of the preprotein with respect to the channel and SecA^{34,35}. In the X-ray structure, residues 1–41 of the OmpA preprotein were attached to the tip of the two-helix finger and visualized in a hairpin structure in the channel (shown in pink, **Figure 5**). The locations of the four PhoA residues in the SecA-PhoA chimera were mapped onto this structure using the same protocol as described above. As shown in **Figure 5**, the location of the PhoA37 and PhoA45 residues (yellow, orange, red) is in between PhoA2 and PhoA22, with PhoA45 closer to PhoA2. These findings, particularly the location of PhoA45, suggested that the PhoA preprotein was forming a hairpin structure.

To further validate our FRET-identified binding site, we performed a comparison of our mapped locations with that of the OmpA preprotein, by excising the 41 residue preprotein structure from the channel and modeling it into the regions defined by FRET mapping (**Figure 5**, cyan). Without any alteration of the preprotein X-ray structure, we find that the locations of residues 2, 22, and 37 (shown in blue, green, and yellow) on the OmpA preprotein fragment excised structure agree remarkably well with the FRET mapped locations (**Figure 5B**) and suggest that the hairpin forms prior to channel entry. The OmpA preprotein ends at residue 41 in the X-ray crystal structure; however, the C-terminal of SecY, which is unstructured, provides an indication of the possible location of PhoA45. In our modeled structure, the hairpin loop sits at the mouth of the channel, poised to facilitate the translocation of the preprotein across the membrane. Thus, in this example, the FRET mapping methodology enhances existing information to illuminate the current understanding of structure and function, by providing dynamic information to existing static

structures. Although not suitable for *de novo* structure determinations, if 3-dimensional structural information is available, the FRET mapping methodology can further the current understanding of structure-function relationships, through elucidation of binding sites and dynamic motions.

ACKNOWLEDGMENTS:

This work was supported by National Institutes of Health grant R15GM135904 (awarded to IM) and National Institutes of Health Grant GM110552 (awarded to DBO).

DISCLOSURES:

The authors have nothing to disclose.

REFERENCES:

1. Thompson, M. C., Yeates, T. O., Rodriguez, J. A. Advances in methods for atomic resolution macromolecular structure determination. *F1000Research*. **9**, F1000 Faculty Rev-667 (2020).
2. Zhang, Q., Li, Y., Olson, R., Mukerji, I., Oliver, D. Conserved SecA signal peptide-binding site revealed by engineered protein chimeras and Forster resonance energy transfer. *Biochemistry*. **55** (9), 1291–1300 (2016).
3. Zhang, Q. et al. Alignment of the protein substrate hairpin along the SecA two-helix finger primes protein transport in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. **114** (35), 9343–9348 (2017).
4. Stryer, L. Fluorescence energy transfer as a spectroscopic ruler. *Annual Review of Biochemistry*. **47**, 819–846 (1978).
5. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. Springer, Boston, MA (2006).
6. Algar, W. R., Hildebrandt, N., Vogel, S. S., Medintz, I. L. FRET as a biomolecular research tool - understanding its potential while avoiding pitfalls. *Nature Methods*. **16** (9), 815–829 (2019).
7. Magde, D., Wong, R., Seybold, P. G. Fluorescence quantum yields and their relation to lifetimes of rhodamine 6G and fluorescein in nine solvents: improved absolute standards for quantum yields. *Photochemistry and Photobiology*. **75** (4), 327–334 (2002).
8. Clegg, R. M. Fluorescence resonance energy transfer and nucleic acids. *Methods in Enzymology*. **211**, 353–388 (1992).
9. Scientific, T. F. *R0 Values from Some Alexa Fluor Dyes - Table 1.6*. <https://www.thermofisher.com/us/en/home/references/molecular-probes-the-handbook/tables/r0-values-for-some-alexa-fluor-dyes.html> (2021).
10. Bajar, B. T., Wang, E. S., Zhang, S., Lin, M. Z., Chu, J. A Guide to Fluorescent Protein FRET Pairs. *Sensors (Basel, Switzerland)*. **16** (9), 1488 (2016).
11. Day, R. N., Davidson, M. W. Fluorescent proteins for FRET microscopy: monitoring protein interactions in living cells. *BioEssays : News and Reviews in Molecular, Cellular, and Developmental Biology*. **34** (5), 341–350 (2012).
12. Lee, S. J., Syed, S., Ha, T. Single-Molecule FRET Analysis of Replicative Helicases. *Methods in Molecular Biology (Clifton, N.J.)*. **1805**, 233–250 (2018).
13. Uhm, H., Hohng, S. Single-Molecule FRET Assay for Studying Cotranscriptional RNA Folding. *Methods in Molecular Biology*. **2106**, 271–282 (2020).
14. Globy, V., Joo, C. Single-molecule FRET studies of Cas9 endonuclease. *Methods in*

816 *Enzymology*. **616**, 313–335 (2019).

817 15. Qiao, Y., Luo, Y., Long, N., Xing, Y., Tu, J. Single-Molecular Förster Resonance Energy Transfer

818 Measurement on Structures and Interactions of Biomolecules. *Micromachines*. **12** (5), 492 (2021).

819 16. Catipovic, M. A., Bauer, B. W., Loparo, J. J., Rapoport, T. A. Protein translocation by the SecA

820 ATPase occurs by a power-stroke mechanism. *The EMBO Journal*. **38** (9), e101140 (2019).

821 17. Seinen, A. B., Spakman, D., van Oijen, A. M., Driessen, A. J. M. Cellular dynamics of the SecA

822 ATPase at the single molecule level. *Scientific Reports*. **11** (1), 1433 (2021).

823 18. Dimura, M. et al. Quantitative FRET studies and integrative modeling unravel the structure

824 and dynamics of biomolecular systems. *Current Opinion in Structural Biology*. **40**, 163–185 (2016).

825 19. Kalinin, S. et al. A toolkit and benchmark study for FRET-restrained high-precision structural

826 modeling. *Nature Methods*. **9** (12), 1218–1225 (2012).

827 20. Paetzel, M. Structure and mechanism of Escherichia coli type I signal peptidase. *Biochimica*

828 *et Biophysica Acta*. **1843** (8), 1497–1508 (2014).

829 21. Ng, D., Brown, J., Walter, P. Signal sequences specify the targeting route to the endoplasmic

830 reticulum membrane. *The Journal of Cell Biology*. **134** (2), 269–278 (1996).

831 22. Huber, D. et al. SecA Cotranslationally Interacts with Nascent Substrate Proteins In Vivo.

832 *Journal of Bacteriology*. **199** (2), e00622-16 (2017).

833 23. Lill, R. et al. SecA protein hydrolyzes ATP and is an essential component of the protein

834 translocation ATPase of *Escherichia coli*. *The EMBO Journal*. **8** (3), 961–966 (1989).

835 24. Lill, R., Dowhan, W., Wickner, W. The ATPase activity of SecA is regulated by acidic

836 phospholipids, SecY, and the leader and mature domains of precursor proteins. *Cell*. **60** (2), 271–

837 280 (1990).

838 25. Kimura, E., Akita, M., Matsuyama, S., Mizushima, S. Determination of a region of SecA that

839 interacts with presecretory proteins in *Escherichia coli*. *The Journal of Biological Chemistry*. **266**

840 (10), 6600–6606 (1991).

841 26. Hunt, J. F. et al. Nucleotide control of interdomain interactions in the conformational reaction

842 cycle of SecA. *Science (New York, N.Y.)*. **297** (5589), 2018–2026 (2002).

843 27. Sharma, V. et al. Crystal structure of *Mycobacterium tuberculosis* SecA, a preprotein

844 translocating ATPase. *Proceedings of the National Academy of Sciences of the United States of*

845 *America*. **100** (5), 2243–2248 (2003).

846 28. Vassilyev, D. et al. Crystal structure of the translocation ATPase SecA from *Thermus*

847 *thermophilus* reveals a parallel, head-to-head dimer. *Journal of Molecular Biology*. **364** (3), 248–

848 258 (2006).

849 29. Zimmer, J., Li, W., Rapoport, T. A. A novel dimer interface and conformational changes

850 revealed by an X-ray structure of *B. subtilis* SecA. *Journal of Molecular Biology*. **364** (3), 259–265

851 (2006).

852 30. Zimmer, J., Rapoport, T. A. Conformational flexibility and peptide interaction of the

853 translocation ATPase SecA. *Journal of Molecular Biology*. **394** (4), 606–612 (2009).

854 31. Bauer, B. W., Rapoport, T. A. Mapping polypeptide interactions of the SecA ATPase during

855 translocation. *Proceedings of the National Academy of Sciences of the United States of America*.

856 **106** (49), 20800–20805 (2009).

857 32. Auclair, S. et al. Mapping of the signal peptide-binding domain of *Escherichia coli* SecA using

858 Förster resonance energy transfer. *Biochemistry*. **49** (4), 782–792 (2010).

859 33. Zimmer, J., Nam, Y., Rapoport, T. A. Structure of a complex of the ATPase SecA and the protein-

translocation channel. *Nature*. **455** (7215), 936–943 (2008).

34. Li, L. et al. Crystal structure of a substrate-engaged SecY protein-translocation channel. *Nature*. **531** (7594), 395–399 (2016).

35. Ma, C. et al. Structure of the substrate-engaged SecA-SecY protein translocation machine. *Nature Communications*. **10** (1), 2872 (2019).

36. Lambert, T. J. FPbase: a community-editable fluorescent protein database. *Nature Methods*. **16** (4), 277–278 (2019).

37. Jilaveanu, L. B., Oliver, D. In vivo membrane topology of Escherichia coli SecA ATPase reveals extensive periplasmic exposure of multiple functionally important domains clustering on one face of SecA. *The Journal of Biological Chemistry*. **282** (7), 4661–4668 (2007).

38. Ramamurthy, V., Oliver, D. Topology of the integral-membrane form of Escherichia coli SecA protein. *The Journal of Biological Chemistry*. **272** (37), 23239–23246 (1997).

39. Chin, J. et al. Addition of p-Azido-L-phenylalanine to the genetic code of Escherichia coli. *Journal of the American Chemical Society*. **124** (31), 9026–9027 (2002).

40. Deiters, A. et al. Adding amino acids with novel reactivity to the genetic code of Saccharomyces cerevisiae. *Journal of the American Chemical Society*. **125** (39), 11782–11783 (2003).

41. Lanzetta, P. A., Alvarez, L. J., Reinach, P. S., Candia, O. A. An improved assay for nanomole amounts of inorganic phosphate. *Analytical Biochemistry*. **100** (1), 95–97 (1979).

42. Mitchell, C., Oliver, D. B. Two distinct ATP-binding domains are needed to promote protein export by Escherichia coli SecA ATPase. *Molecular Microbiology*. **10** (3), 483–497 (1993).

43. Scientific, T.F. *Thiol-reactive Probe Labeling Protocol*. <https://www.thermofisher.com/us/en/home/references/protocols/cell-and-tissue-analysis/labeling-chemistry-protocols/thiol-reactive-probe-labeling-protocol.html>. (2021).

44. Scientific, T.F. *Click Chemistry - Section 3.1*. <https://www.thermofisher.com/us/en/home/references/molecular-probes-the-handbook/reagents-for-modifying-groups-other-than-thiols-or-amines/click-chemistry.html>. (2021).

45. Bioquest, A. *Correction Factor*. <https://www.aatbio.com/resources/correction-factor/>. (2019).

46. Scientific, T.F. *Calculate dye:protein (F/P) molar ratios*. <https://tools.thermofisher.com/content/sfs/brochures/TR0031-Calc-FP-ratios.pdf>. (2011).

47. Scientific, H. *A Guide to Recording Fluorescence Quantum Yields*. https://static.horiba.com/fileadmin/Horiba/Application/Materials/Material_Research/Quantum_Dots/quantumyieldstrad.pdf.

48. Ivanov, V., Li, M., Mizuuchi, K. Impact of emission anisotropy on fluorescence spectroscopy and FRET distance measurements. *Biophysical Journal*. **97** (3), 922–929 (2009).

49. Auclair, S., Oliver, D., Mukerji, I. Defining the solution state dimer structure of Escherichia coli SecA using Forster resonance energy transfer. *Biochemistry*. **52** (14), 2388–2401 (2013).

50. *The PyMOL Molecular Graphics System, Version 2.4*. <https://pymol.org/2/>. Schrodinger, LLC, New York (2021).

51. Musial-Siwiek, M., Rusch, S.L., Kendall, D.A. Selective photoaffinity labeling identifies the signal peptide binding domain on SecA. *Journal of Molecular Biology*. **365** (3), 637–648 (2007).

52. Miller, A., Wang, L., Kendall, D.A. Synthetic signal peptides specifically recognize SecA and

904 stimulate ATPase activity in the absence of preprotein. *The Journal of Biological Chemistry*. **273**
905 (19), 11409–11412 (1998).

906 53. Gelis, I. et al. Structural basis for signal-sequence recognition by the translocase motor SecA
907 as determined by NMR. *Cell*. **131** (4), 756–769 (2007).

908 54. Erlandson, K. J. et al. A role for the two-helix finger of the SecA ATPase in protein
909 translocation. *Nature*. **455** (7215), 984–988 (2008).

910 55. Das, S., Oliver, D. Mapping of the SecA-SecY and SecA-SecE interfaces by site-directed in vivo
911 photocross-linking. *The Journal of Biological Chemistry*. **286** (14), 12371–12380 (2011).

912 56. Wheatley, E. G., Pieniazek, S. N., Vitoc, I., Mukerji, I., Beveridge, D. L. Molecular Dynamics
913 Structure Prediction of a Novel Protein-DNA Complex: Two HU Proteins with a DNA Four-way
914 Junction. in *Innovations in Biomolecular Modeling and Simulations: Volume 2*. The Royal Society
915 of Chemistry. 111–128 (2012).

916 57. Vitoc, C. I., Mukerji, I. HU binding to a DNA four-way junction probed by Förster resonance
917 energy transfer. *Biochemistry*. **50** (9), 1432–1441 (2011).

918 58. Hellman, L. M., Fried, M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein-
919 nucleic acid interactions. *Nature Protocols*. **2** (8), 1849–1861 (2007).

920

Figure 1

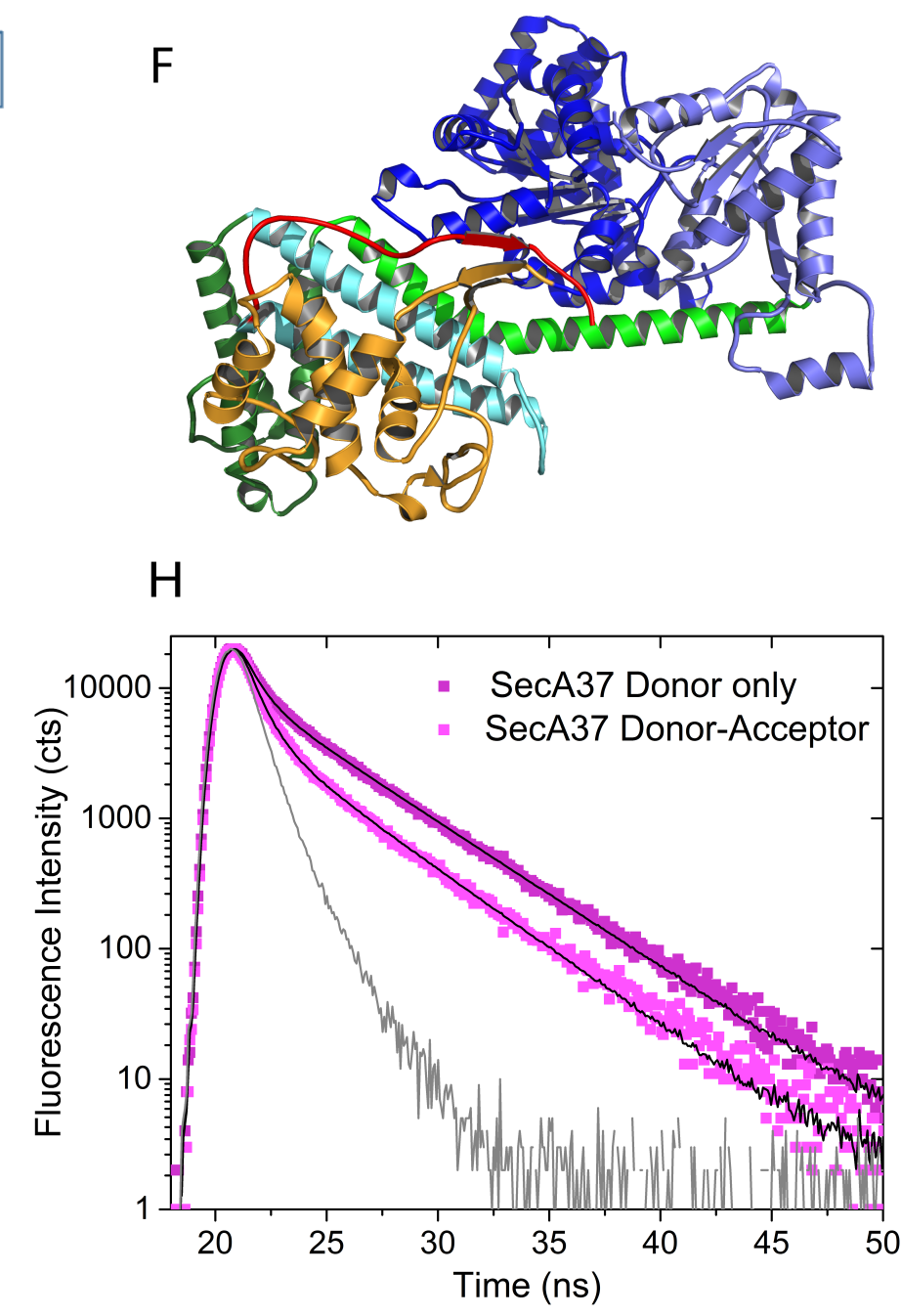
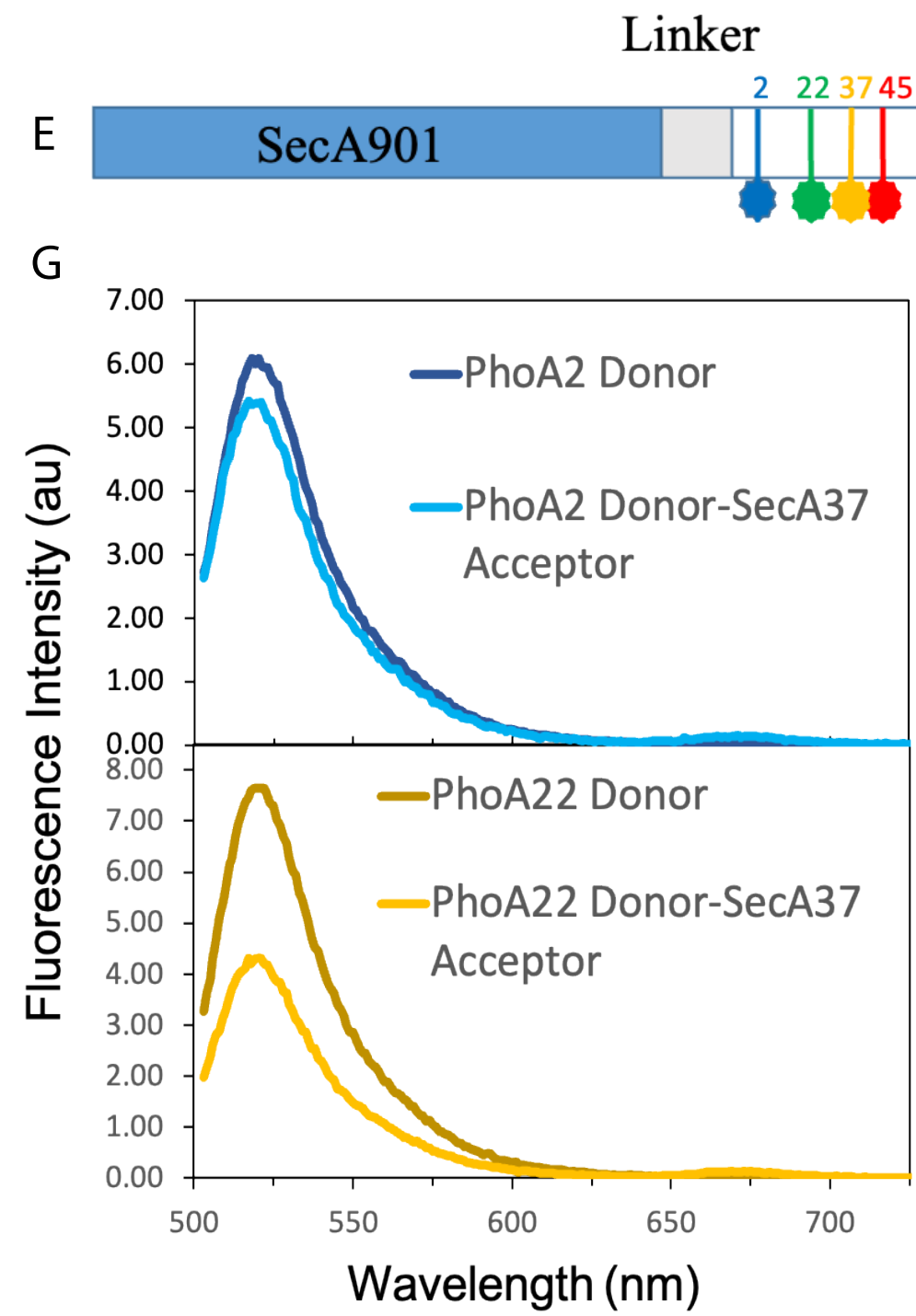
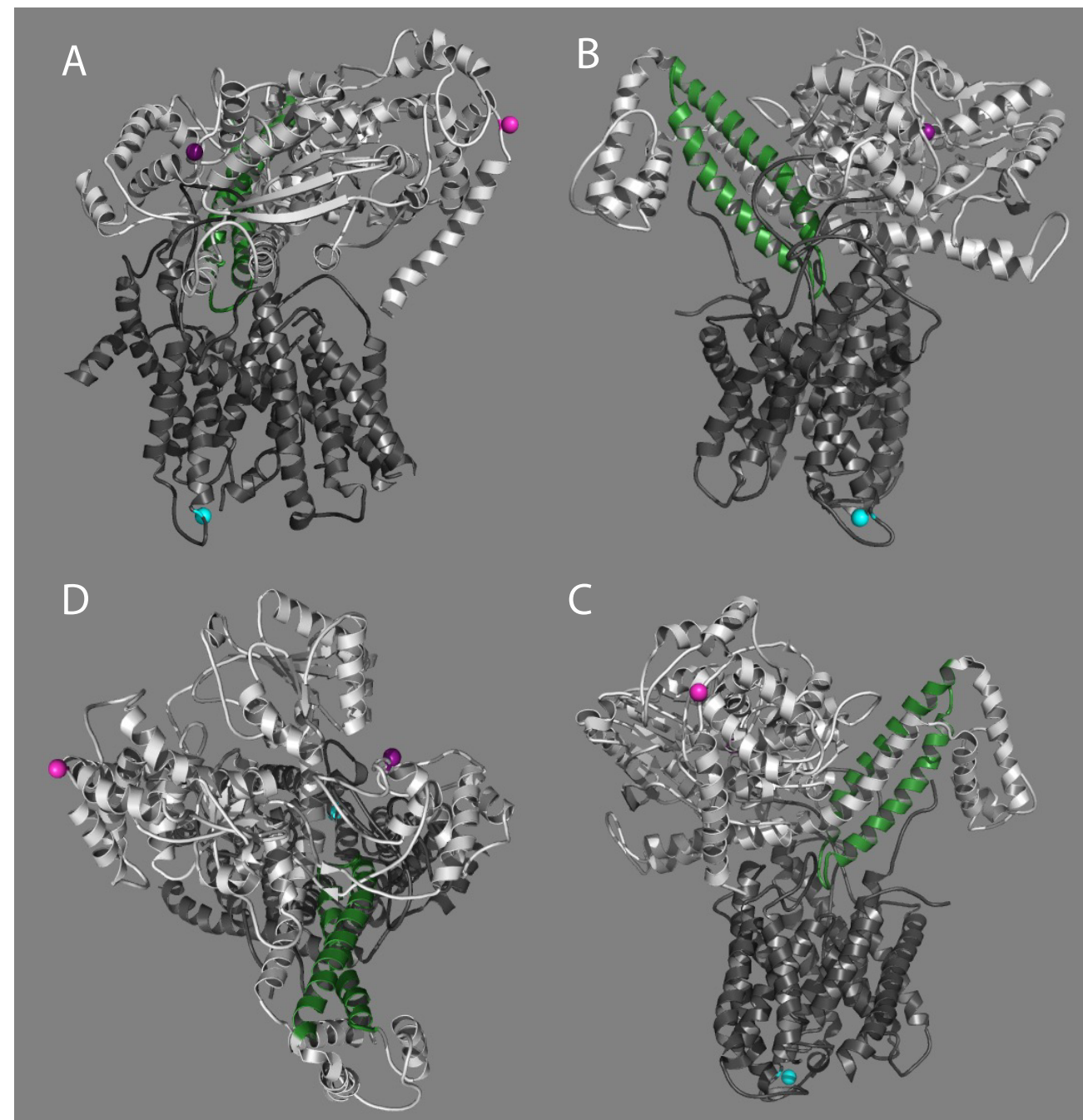
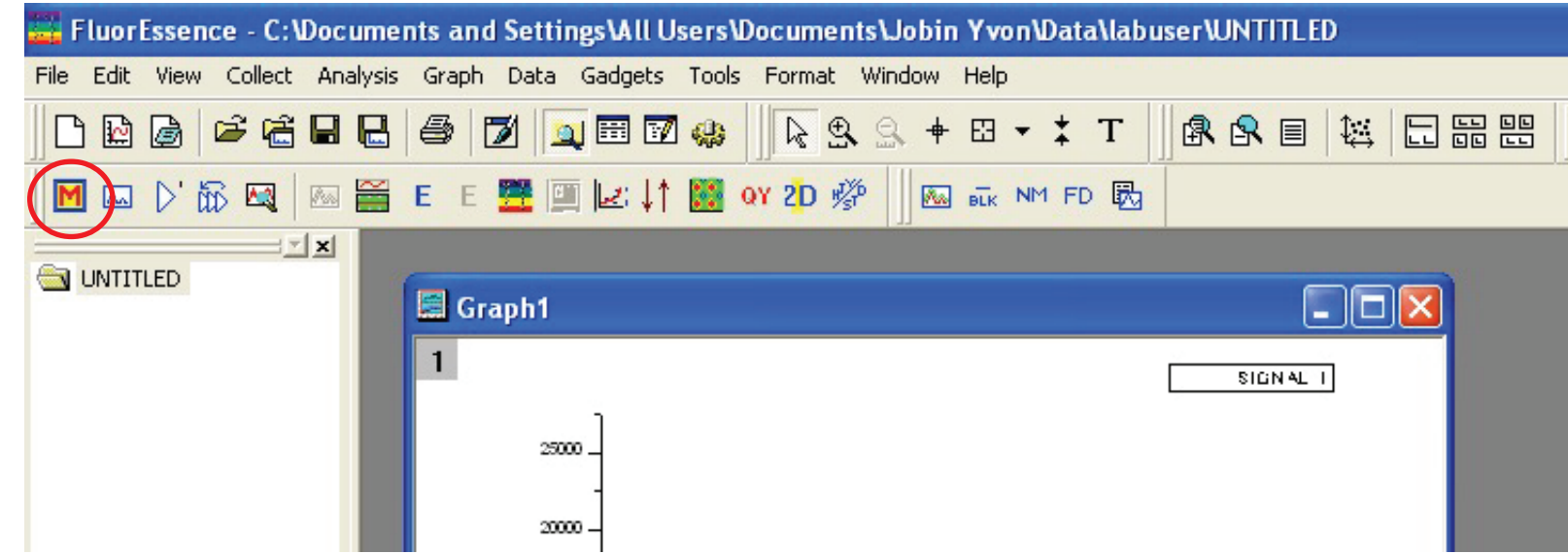


Figure 2

[Click here to access/download;Figure;new_Figure2.pdf](#)

A



B

The screenshot shows the "Fluorescence Division - Experiment Setup (Spectral Acquisition[Emission])" dialog box. The "General information" tab is selected. The "Experiment" section includes a file name "Trp_Emission.xml" and a directory "C:\Documents and Settings\All Users\Documents\Jobin Yvon\". The "Data Storage" section includes a "Data Identifier" field with the value "Trp" and a "Comment" field with the value "Trp". The "Accessories" section includes tabs for "Pol(EX)", "Pol(EM)", "Sample Changer", and "Temperature Controller". The "Display Options" section includes radio buttons for "Dynamic", "Fixed", and "Scan". The "Fixed" option is selected. The "Axis Value" and "Position" fields are also visible.

Figure 3

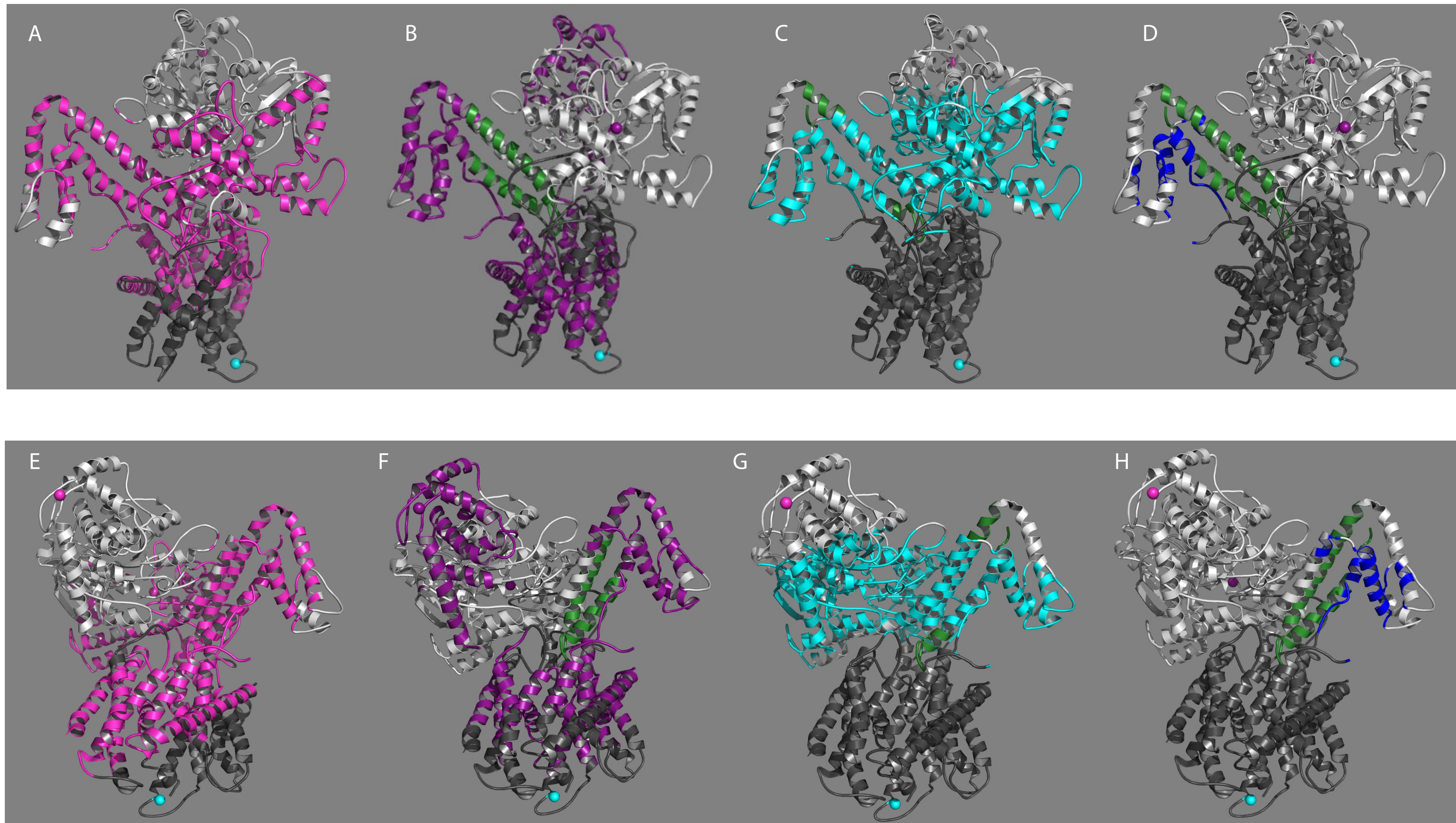
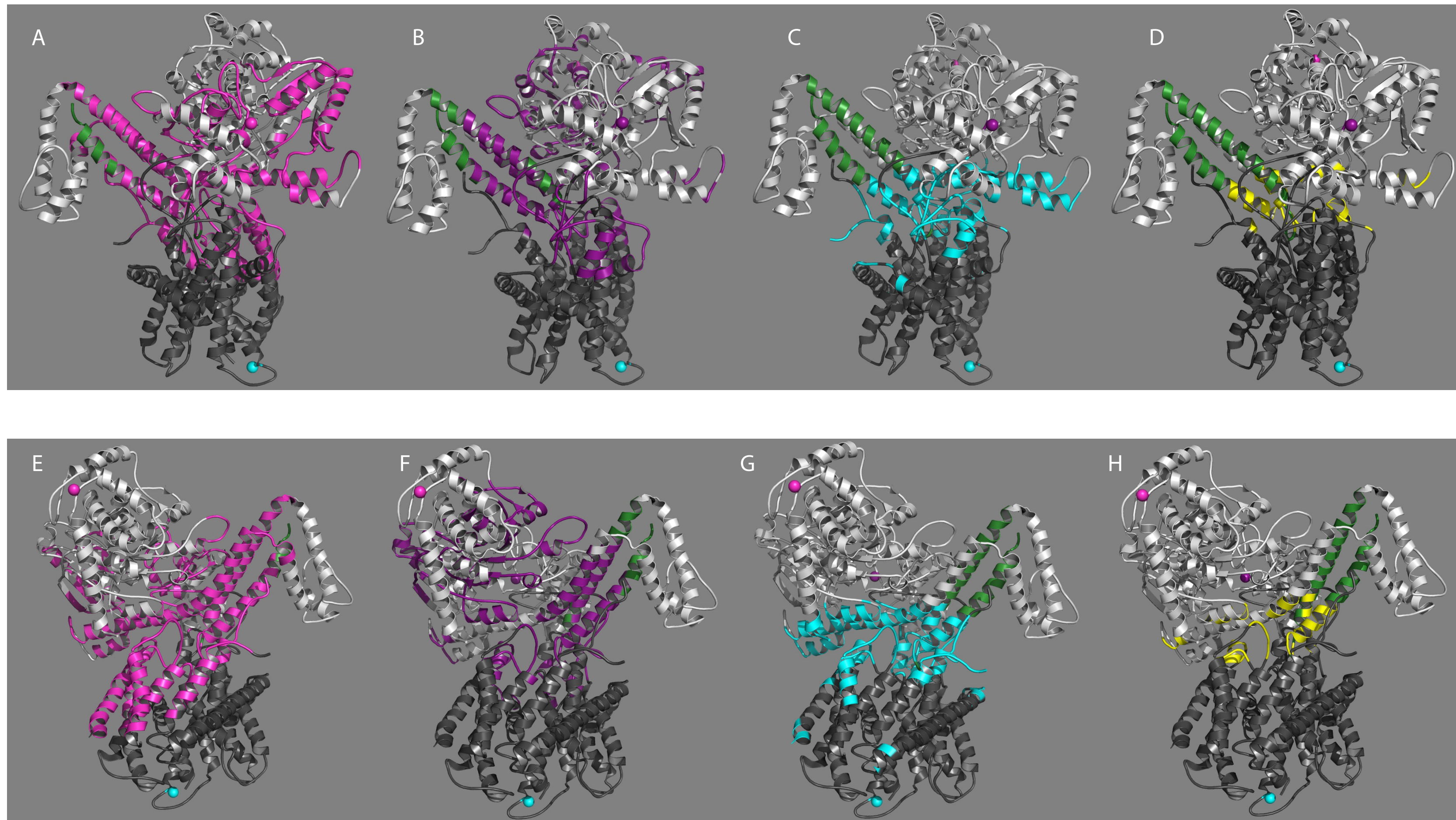


Figure 4



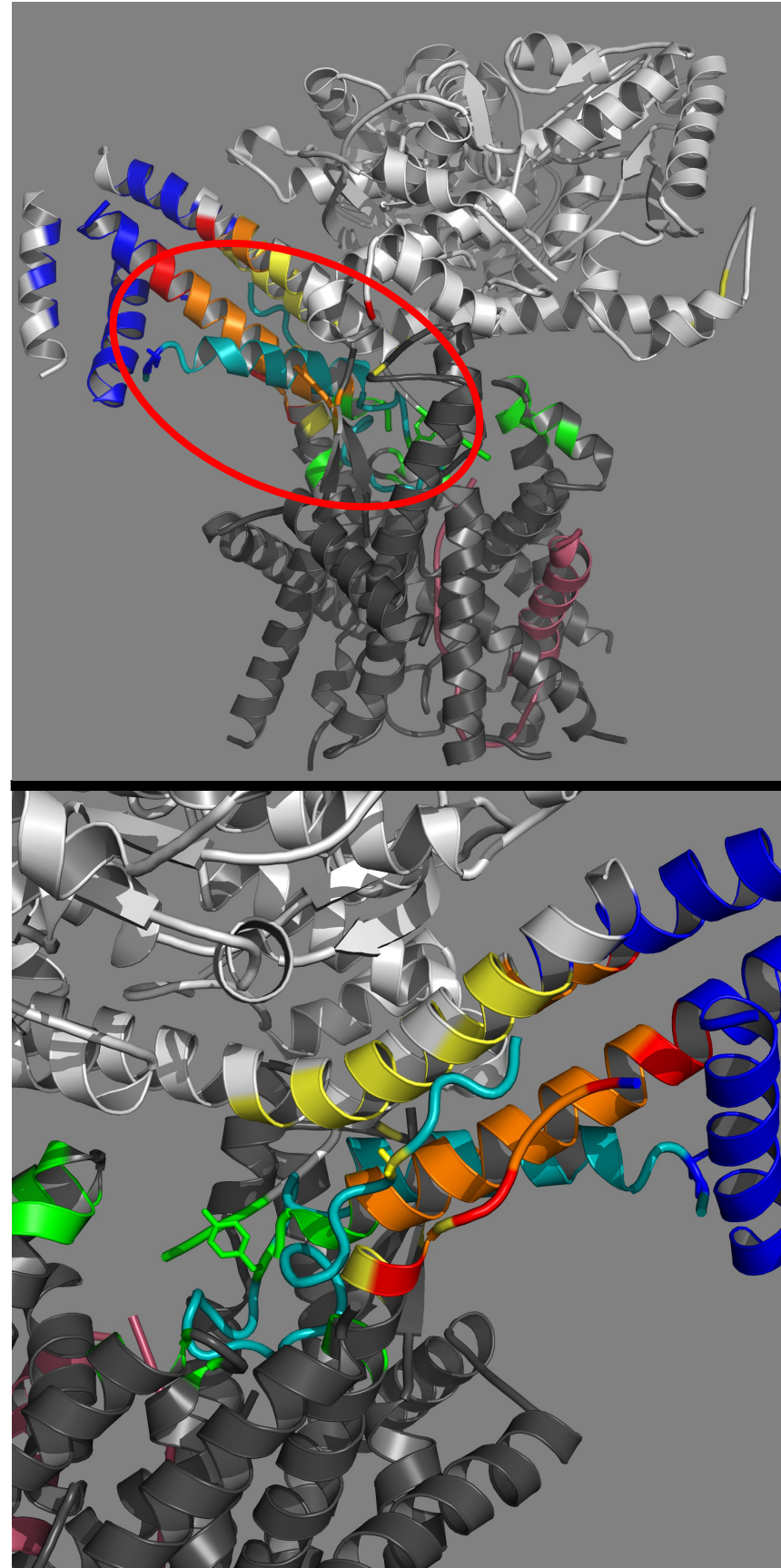


Table 1: Efficiencies and Distances determined for SecA-PhoA-SecYEG complex

labeled site on SecA-PhoA-SecYEG complex	Labeled site on PhoA peptide portion of SecA-PhoA complex		
	PhoA2-AF488	PhoA22-AF488	PhoA37-AF488
SecA 37-AF467-PhoA			
	R0 = 57	R0 = 57	R0 = 57
FRET efficiency	0.27 +/- .01	0.52 +/- .02	0.39 +/- .03
distance	67 +/- 15	56 +/- 12	62 +/- 13
SecA321-AF647-PhoA			
	R0 = 40	R0 = 38	R0 = 37
FRET efficiency	0.16 +/- .04	0.62 +/- .01	0.39 +/- 0.02
distance	53 +/- 11	34.9 +/- 7.3	39.7 +/- 7.9
	PhoA2-AF647	PhoA22-AF647	PhoA37-AF647
SecY292-AF488 EG			
	R0 = 57	R0 = 50	R0 = 53
FRET efficiency	0.25 +/- .06	0.46 +/- .06	0.24 +/- .05
distance	68 +/- 15	51.3 +/- 9.2	64 +/- 14

adapted from reference 3.

R_0 values given in angstroms were calculated as described in the text

The FRET efficiency was calculated from the decrease in donor fluorescence intensity in the presence of the acceptor as described. The error is reported as the SD from three independent measurements.

Donor-acceptor distances (R) are given in angstroms and calculated as described in the text. The error reported results from a consideration of the experimental error and that arising from the orientation of the dyes. Dye orientation is estimated from the steady state fluorescence anisotropy

Chimera

PhoA45-AF488

R0 = 60

0.36 +/- .03

66 +/- 15

R0 = 38

0.43 +/- 0.02

39.7 +/- 7.5

PhoA45-AF647

R0 = 54

0.30 +/- .01

62 +/- 13



[Click here to access/download](#)

Table of Materials

[Table of Materials_Northrop et al_JOVE_RE.xls](#)



**Molecular Biology and Biochemistry Department**

Room 157, Hall Atwater Laboratories
52 Lawn Ave
Middletown, CT. 06459-0175

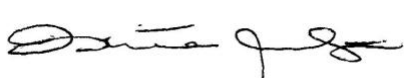
December 24, 2021

Dear Dr. Mittal,

Thank you for giving us the opportunity to revise our manuscript entitled, "Förster Resonance Energy Transfer Mapping: A New Methodology to Elucidate Global Structural Features," JoVE submission 63433. We are also pleased to address the editorial comments and have revised the manuscript accordingly. We trust that with these revisions the manuscript is now acceptable for publication in JoVE.

Thank you for your consideration,

Sincerely yours,



Ishita Mukerji
Fisk Professor of Natural Science
Professor, Molecular Biology and Biochemistry
Director, Molecular Biophysics Program

FRET Mapping: A New Methodology to Elucidate Global Structural Information

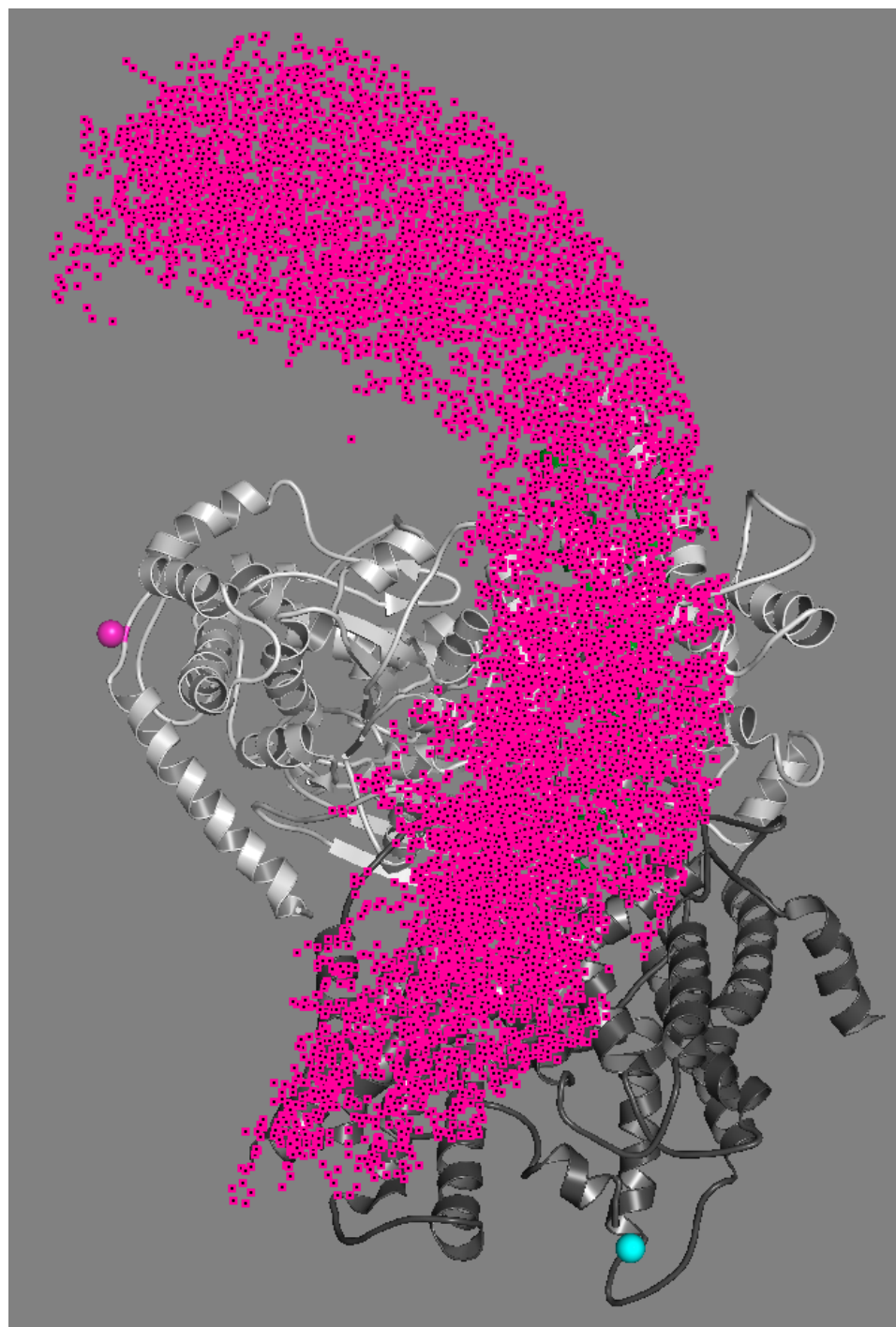
Supplemental Information: Pymol Script

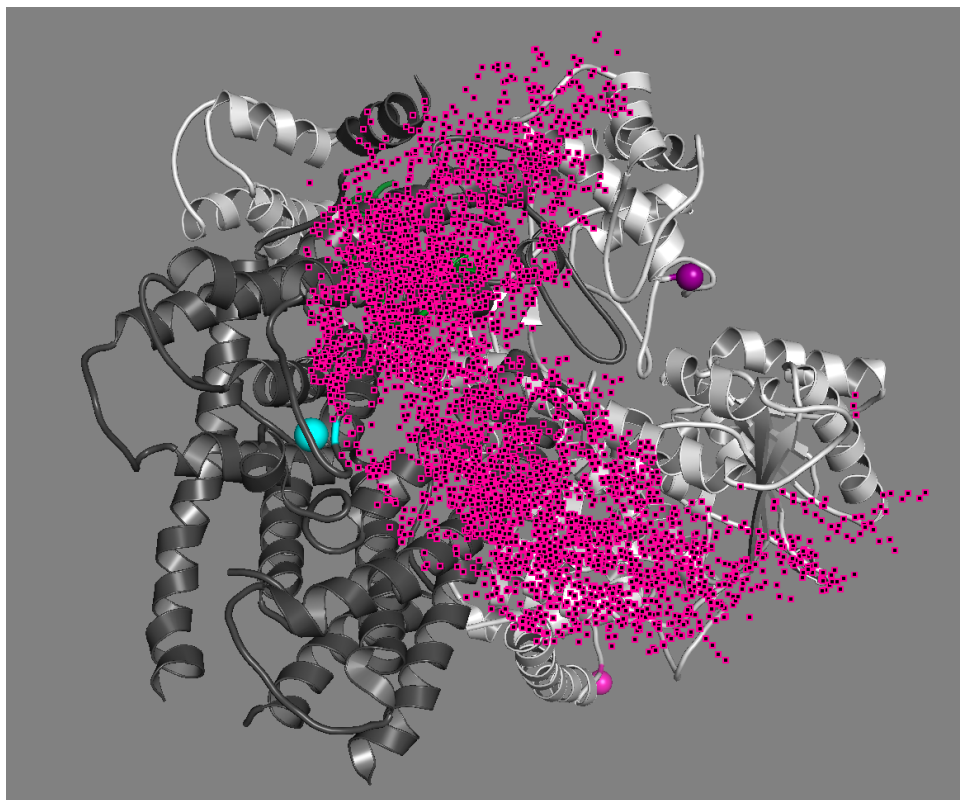
Mapping of sites within the SecA-SecYEG complex using FRET distances. The overlap regions between the SecA-SecYEG structure (PDBID: 3DIN) and the FRET dataset (see Table 1) were determined from the intersection of the three spherical shells generated from the FRET distances for a specific labeled position on the SecA-PhoA chimera protein and the selected atoms of the SecA-SecYEG structure (Figs. 2 and 3). The thickness of the shell was defined by the maximum and minimum distances determined for that FRET pair based on the measured error. The shells were generated using standard selection algebra commands in the program Pymol (Schrodinger, LLC) (https://pymolwiki.org/index.php/Selection_Algebra)¹. For each shell, we made use of the "around" command for the maximum distance and "beyond" commands for the minimum distance and the Boolean "and" logical operator to create the set of atoms. A sample script for identifying the overlapping region would be as follows (adapted from Zhang et al.²:

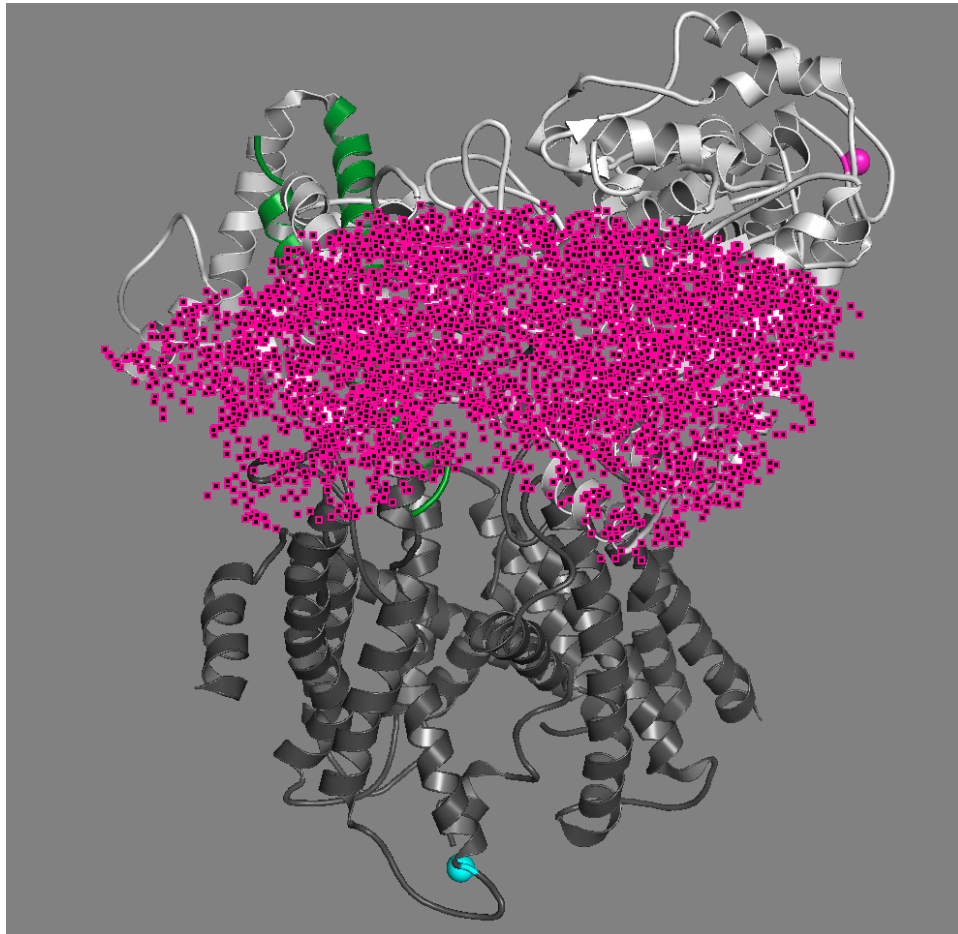
```
PyMOL>select a37PhoA2-1, chain A and resi 30 around 82
Selector: selection "a37PhoA2-1" defined with 14843 atoms.
PyMOL>select a37PhoA2-2, a37PhoA2-1 beyond 52 of chain A and resi 30
Selector: selection "a37PhoA2-2" defined with 9699 atoms.
PyMOL>select a321PhoA2-1, chain A and resi 346 around 64
Selector: selection "a321PhoA2-1" defined with 11367 atoms.
PyMOL>select a321PhoA2-2, a321PhoA2-1 beyond 42 of chain A and resi 346
Selector: selection "a321PhoA2-2" defined with 5614 atoms.
PyMOL>select y292PhoA2-1, chain C and resi 288 around 83
Selector: selection "y292PhoA2-1" defined with 8822 atoms.
PyMOL>select y292PhoA2-2, y292PhoA2-1 beyond 53 of chain C and resi 288
Selector: selection "y292PhoA2-2" defined with 4932 atoms.
PyMOL>select PhoA2res, a37PhoA2-2 and a321PhoA2-2 and y292PhoA2-2
Selector: selection "PhoA2res" defined with 442 atoms.
```

References:

- 1 The PyMOL Molecular Graphics System, Version 2.4 (Schrodinger, LLC, New York, 2021).
- 2 Zhang, Q. *et al.* Alignment of the protein substrate hairpin along the SecA two-helix finger primes protein transport in Escherichia coli. *Proc Natl Acad Sci U S A.* **114** (35), 9343-9348, (2017).







Rights and permissions of PNAS

Copied from <https://www.pnas.org/page/about/rights-permissions>

The author(s) retains copyright to individual PNAS articles, and the National Academy of Sciences of the United States of America (NAS) holds copyright to the collective work and retains an [exclusive License to Publish](#) these articles, except for open access articles submitted beginning September 2017. For such [open access](#) articles, NAS retains a [nonexclusive License to Publish](#), and these articles are distributed under either a [CC BY-NC-ND](#) or [CC BY](#) license.

For volumes 106–114 (2009–September 2017), the author(s) retains copyright to individual articles, and NAS retains an exclusive License to Publish these articles and holds copyright to the collective work. Volumes 90–105 (1993–2008) are copyright National Academy of Sciences. For volumes 1–89 (1915–1992), the author(s) retains copyright to individual articles, and NAS holds copyright to the collective work.

Authors whose work will be reused should be notified. Use of PNAS material must not imply any endorsement by PNAS or NAS. The full journal reference must be cited and, for articles published in Volumes 90–105 (1993–2008), "Copyright (copyright year) National Academy of Sciences" must be included as a copyright note.

Please visit the [Permissions FAQ](#) for detailed information about PNAS copyright and self-archiving guidelines. The PNAS listing on the Sherpa RoMEO publisher policies pages can be found [here](#).

Additional information and answers to frequently asked questions about author rights and permissions are available on our [FAQ page](#).

PNAS authors need not obtain permission for the following cases:

1. to use their original figures or tables in their future works;
2. to make copies of their articles for their own personal use, including classroom use, or for the personal use of colleagues, provided those copies are not for sale and are not distributed in a systematic way;
3. to include their articles as part of their dissertations; or
4. to use all or part of their articles in printed compilations of their own works.

The full journal reference must be cited and, for articles published in volumes 90–105 (1993–2008), "Copyright (copyright year) National Academy of Sciences" must be included as a copyright note.