bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Energy landscape steering in SecYEG mediates dynamic coupling in ATP driven protein translocation

Joel A. Crossley,[†,‡,¶] Tomas Fessl,[‡] Matthew A. Watson,[†] Daniel W. Watkins,[§] Robin A. Corey,[‖] William J. Allen,[§] Tara Sabir,[¶] Sheena E. Radford,[†] Ian Collinson,[*,§] and Roman Tuma[*,†,‡]

†*Astbury Centre for Structural Molecular Biology, School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, United Kingdom*
‡*Faculty of Science, University of South Bohemia in České Budějovice, České Budějovice, Czech Republic*
¶*School of Clinical and Applied Sciences, Leeds Beckett University, Leeds, United Kingdom*
§*School of Biochemistry, University of Bristol, Bristol, United Kingdom*
‖*Department of Biochemistry, University of Oxford, Oxford, United Kingdom*

E-mail: ian.collinson@bristol.ac.uk; r.tuma@leeds.ac.uk

Abstract

The Sec translocon is a transmembrane assembly highly conserved among all forms of life as the principal route for transport of polypeptides across or into lipid bilayers. In bacteria translocation involves allosteric communication between the membrane pore SecYEG and the associated SecA ATPase. Using single-molecule fluorescence we reveal that slow conformational changes associated with the ATPase SecA modulate fast opening and closure of the SecY lateral gate. Such a mismatch of timescales is not compatible with direct coupling between SecA and SecYEG. A dynamic allosteric model is proposed in which the SecA ATPase cycle 'steers' the energy landscape for SecY pore opening. We map the experimental traces onto reduced reaction coordinates derived from molecular dynamics trajectories, providing a model for the energy landscape and a structural interpretation of the associated dynamics. Dynamic allostery may be common among motor ATPases that drive conformational changes in molecular machines.

Introduction

A fascinating class of biological molecular machines are those operating upon biopolymer substrates and converting chemical energy derived from ATP binding and hydrolysis into cycles of conformational changes and mechanical work (e.g. helicases, translocases, unfoldases, chromatin remodelling complexes, primary membrane transporters and other molecular motors). The bacterial protein secretion system, composed of the integral membrane assembly SecYEG and peripherally associated cytosolic motor ATPase SecA (designated SecYEG:A here), belongs to this class, purposed to translocate unfolded polypeptides across the inner membrane.[1,2] Secretory proteins (pre-proteins) with N-terminal cleavable signal sequences (SS) are targeted to SecYEG:A for subsequent post-translational transport across the membrane.

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

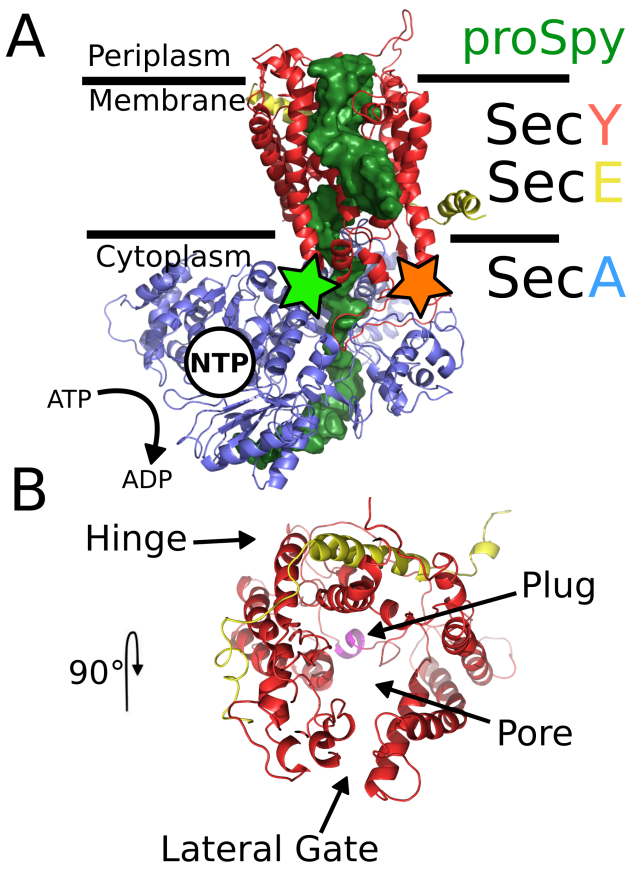


Figure 1: The Structure of Sec translocon. SecY is shown in red, SecE - yellow, SecA - blue, proSpy - Green and Plug in pink. (PDB: 5EUL[3]) (A) The structure of SecYEG:A with translocating substrate modelled as shown in Corey et al.[4]. Fluorophore placement is shown on the lateral gate in green and orange stars (see Figure S1 for structure based modelling of the dyes). Nucleotide binding pocket (NTP) of SecA is also indicated (B) Top down view of SecYEG as seen from the cytoplasm.

In the SecYEG complex, SecY forms the pore for protein translocation (Figure 1A).[1] At the side of the channel is a lateral gate (LG) (Figure 1B).[5] Opening of the LG is required for pre-protein translocation and insertion into the membrane.[6,7] SecA is a RecA-like ATPase which associates with SecYEG during translocation. It has two nucleotide binding domains forming the ATP binding site.[8] The two-helix finger (2HF) domain of SecA has been proposed to act as a sensor regulating nucleotide exchange,[9,10] or alternatively, to directly push the polypeptide across the membrane.[11,12]

The precise nature of the translocation mechanism has divided opinion owing to the inherent complexity of the system.[9,11–13] As for other molecular machines which convert the chemical potential of NTP hydrolysis into directional motion, there are two limiting cases of energy transduction: power stroke and Brownian ratchet.[14] The power-stroke mechanism involves deterministic or direct coupling in which each conformation imposed by the stage of the nucleotide hydrolysis cycle is linked with a well-defined conformation of the effector part (e.g. a mechanical lever or substrate binding site) as within a purely mechanical motor. On the other hand, a Brownian ratchet mechanism exhibits loose coupling between the conformations of the nucleotide-binding site of the motor ATPase and the effector part.[15] Both cases can be illustrated in terms of a simplified energy profile along the mechanical reaction coordinate. The power stroke profile features deep minima at either the pre-stroke or post-stroke position which directly correspond to the nucleotide state of the ATPase. In the Brownian ratchet case, the nucleotide-hydrolysis cycle biases a shallow energy profile (i.e. with low energy barriers between states) towards certain conformations, which in turn undergo rapid interconversion between the available states due to thermal fluctuations.

The power-stroke mechanism has been demonstrated for many cytoskeletal motors[16] while a Brownian ratchet scheme has been implicated for other ATPases such as the ClpX polypeptide unfoldase.[17] Both mechanisms have been considered for the Sec translocon: the power stroke invokes large, piston-like motion of 2HF imparting force directly to the substrate while SecYEG is considered a passive pore.[11,12] Alternatively, SecY has been proposed to act as Brownian ratchet while allosterically communicating with SecA.[9] A recent single-molecule study detected conformational changes of the 2HF taking place on time scales (measured as dwell times in different states) of 100 - 400 ms,[12] i.e. similar to the timescale of the ATP hydrolysis cycle (∼100 ms),[18] which was interpreted to mean that SecA acts via a directly coupled power stroke. However, in all published structures of the complex[3,19–21] there is less than 1 nm available for 2HF movement, and covalent crosslinking of the 2HF to SecY

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

does not prevent translocation activity,[10] suggesting the 2HF does not move much during protein transport.

Evidence also points to the SecY pore being actively gated by SecA: (1) opening and closing of the SecY LG is linked to the nucleotide state of SecA;[9] and (2) rapid relaxation of LG to either open or closed conformation has been observed in molecular dynamics (MD) trajectories in response to changing the identity of the nucleotide bound to SecA from a transition state mimic (ADP:BeF_x) to ATP or ADP, respectively.[9] Likewise, the nucleotide turnover of SecA is affected by SecYEG: (1) the SecA ATP hydrolysis rate increases ∼10-fold when complexed to SecYEG and ∼100-fold during translocation;[18] (2) ADP release is affected by the LG conformation.[9] Further evidence for a two-way dynamic coupling and allostery comes from hydrogen-deuterium exchange experiments.[22]

Understanding dynamic coupling between SecA and SecYEG is key to reconciliation of the apparently distinct mechanisms proposed. The two principle scenarios are considered here: (1) the different conformations of SecYEG (e.g. open and closed states of the channel) deterministically follow the nucleotide state of SecA (2) SecYEG coupling to SecA is stochastic; manifested by rapid interchange between states and controlled by modulations of the underlying energy landscape by the nucleotide status of SecA.

Results

Identification of conformational states of SecYEG Previous structural studies have shown that the LG of SecYEG exhibits three discernible conformational states.[19,20] This was further confirmed by single-molecule fluorescence[9] and molecular dynamics simulations in which open, part open (here called intermediate) and closed LG conformations were identified, but no transitions between these states were previously reported, presumably due to the limited temporal resolution (200 ms per frame) of the total internal reflection fluorescence microscopy (TIRFM). Hence we employed a faster confocal approach (temporal resolution ∼0.1 ms) with the aim to detect and resolve transitions. During the SecA ATPase-coupled translocation of substrate through the translocon, the FRET signal between the dyes acts as a proxy for the the LG conformational state (Figure 1, Figure 2A and Figure S1).

The conditions most relevant to translocation - SecYEG:A alone and SecYEG:A+ADP (both at equilibrium), and SecYEG:A+ATP and SecYEG:A+ATP+substrate (proSpy) (both away from equilibrium), yielded FRET efficiency distributions well-explained by fitting to three steady-state populations corresponding to the open, intermediate and closed LG conformational states. This assignment is further supported by modelling fluorophore accessible volumes to distinct species observed by MD[4,9] that yielded three corresponding simulated FRET distributions (Figure S1B).

The relative steady-state populations of the different states could be evaluated for samples in the presence of the various nucleotides and soluble periplasmic substrate, proSpy, with FRET efficiency (E_{FRET}) values of 0.25, 0.50 and 0.72, for open, intermediate and closed state, respectively (Figure 2 and Figure S2); all of which were in good qualitative agreement with the previously published TIRFM conducted with an alternative pre-protein (proOmpA) (E_{FRET} ∼0.45, 0.59, and 0.76 for the open, intermediate and closed states respectively).[9] In fact, better time-resolution led to reduced averaging due to rapid state interconversion and helped to discern a shift in the FRET efficiency distribution for the closed state during translocation (from E_{FRET} 0.72 to 0.65, compare panels E and F in Figure S2). Presumably, this reflects signal sequence induced unlocking of SecYEG[23–25] and the presence of substrate precluding complete closure of the gate.

Dynamic behaviour of the SecYEG lateral gate In order to examine the dynamic equilibria of the LG conformational states, and its response to SecA, different nucleotides, and the pre-protein proSpy, we compiled two-

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

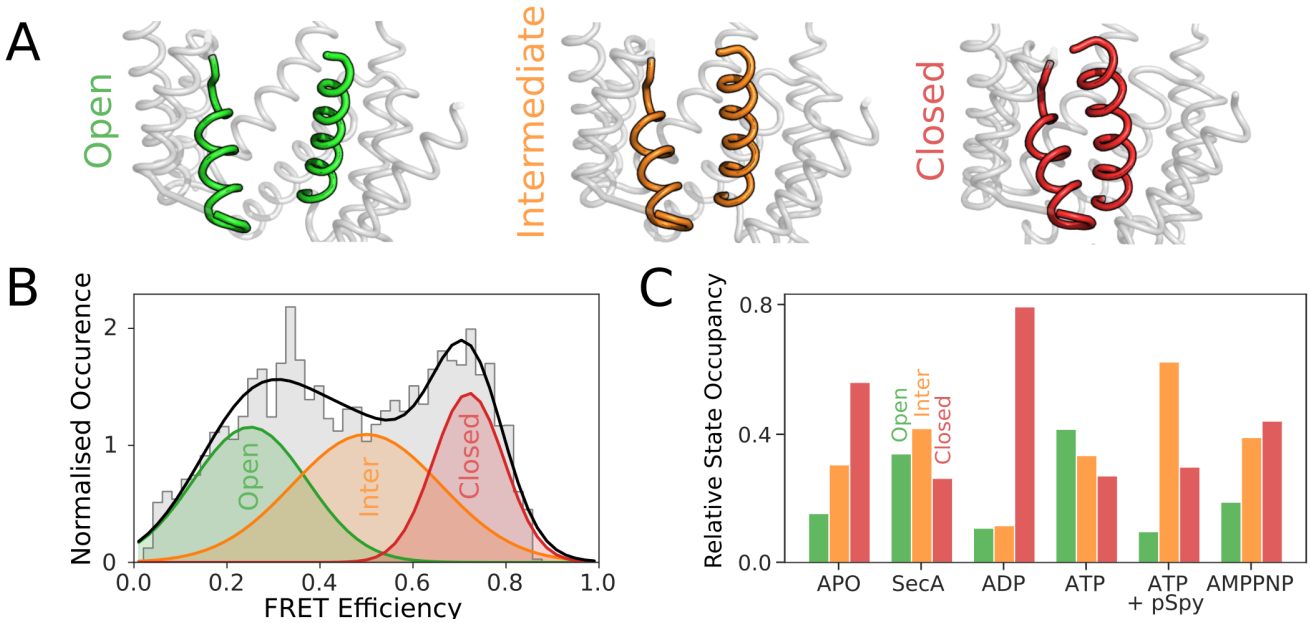


Figure 2: Conformational states of the lateral gate of SecYEG. Green = open, orange = intermediate, red = closed. (A) Side view of helices lining the lateral gate. The open, closed and intermediate states of LG are from representative frames from the equilibrium part of each MD trajectory (basins of attraction in Figure 4A). (B) Representative FRET efficiency histogram for SecYEG:A (7992 events). Grey bins represent the experimentally observed distribution of E values (corrected for the population of apo SecYEG (unable to bind SecA due to their orientation in the membrane[9]) - Figure S2A); while coloured overlaid Gaussians correspond to the components assigned to the three conformational states. The sum of the Gaussians is shown as a solid black line. Note that the Gaussian corresponding to the intermediate state represents both well defined static intermediate as well as a fraction of dynamically averaged intermediates. (C) The relative population of conformational states of the lateral gate in each experimental condition in Figures 2B and S2A-E.

dimensional (2D) histograms of FRET efficiency (E) and the corresponding donor lifetime obtained for each burst (τ_D) (Figure 3A-C and Figure S3). In the absence of direct donor quenching (Figure S4), such two-dimensional analysis allows the determination of FRET linked dynamic behaviour on the millisecond timescale, defined by the average burst duration (i.e. the time it takes for the lipid vesicle to diffuse through the confocal volume; Figure S5 and Figure S6AB). Bursts for which E and τ_D are approximately anticorrelated and fall into the vicinity of the diagonal, 'static line' exhibit no transitions, while those deviating from this line to the right indicate dynamic transitions on the millisecond timescale (for further explanations see Figure S6).

At equilibrium, the LG in both SecYEG alone and SecYEG:A exhibits intrinsic dynamic interconversion between all three states, although the closed state (E_{FRET} = 0.72, τ_D ~1.5 ns) dominates (Figure 3A, Figure S4A and Figure 2C). This interconversion was unexpectedly fast, on the diffusion timescale (represented by the mean burst width 5.75 ms, Figure S5), corresponding to a rate constant of ~175 s^{-1}. Similar dynamic behaviour was observed in SecYEG:A when saturated by AMP-PNP (a non-hydrolysable analogue of ATP) (Figure S3C). However, addition of excess ADP prevents opening of the gate almost entirely, effectively shutting down all dynamics on this timescale (Figure 3B). Interestingly, during conditions of steady state ATP turnover the dynamic behaviour is further accentuated relative to the equilibrium condition, including an increased population of the open state (Figure 3C). As noted above, in the presence of translocating substrate the closed state shifts to lower E_{FRET} (Figure 3C red dashed line).

The presence of FRET events along the 'dynamic line' in the E_{FRET}/τ_D (further referred to as E-τ) plots demonstrates that a substantial number of transitions occur on a timescale

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

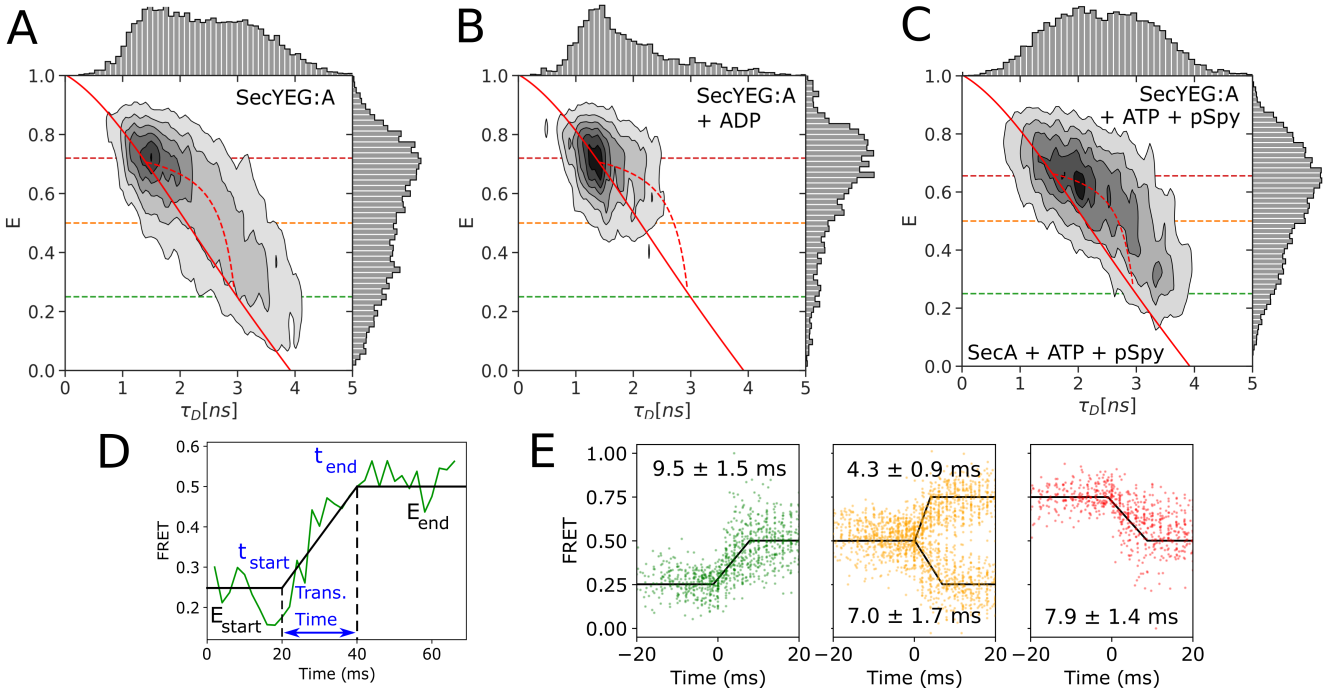


Figure 3: Dynamic behaviour of SecYEG:A LG. Two-dimensional FRET E and donor lifetime (τ_D) analysis for (A) SecYEG:A, (B) SecYEG:A+ADP and (C) SecYEG:A+ATP+proSpy. The solid diagonal (static line) and curved dashed line (dynamic line) represent the static and dynamic fits, respectively (see Figure S6 for details). Horizontal dashed coloured lines indicate the centres of the E_{FRET} values of the closed (red), intermediate (orange) and open (green) lateral gate states. (D) An example of a transition (delineated by dashed vertical lines) from experimental E_{FRET} traces (solid green line) and fitted to a three-segment line (solid black) in order to define start and end times and FRET efficiency values. Further representative transitions are shown in a gallery in Figure S7. (E) Transitions in SecYEG:A+ATP. Transitions were aligned using the mid-point ($[t_{start} + t_{end}]/2$) and classified according to the initial (E_{start}) state. Transitions originating from open are shown in green, intermediate in orange and closed in red. Average transition times (indicated) were calculated from distributions in Figure S7 and segmented models representing the average are shown as solid lines.

similar to, or faster than that of diffusion of the proteoliposomes through the confocal volume (burst duration), i.e. a few milliseconds (Figure S6). Given the broad distribution of diffusion times a subset of events (bursts) with discernible transitions between well defined states were identified (Figure S6D). In addition, many events also contain multiple fast transitions beyond the present time resolution (Figure S6E). We employed a Hidden Markov Model (HMM) analysis solely as a tool to select a subset of bursts containing the well-defined transitions. While HMM also yields dwell times this is not applicable to diffusion-based experiments due to the limited observation (burst duration) time. HMM assumes that the transitions are instantaneous (Figure 3D) but many of the detected transitions had a finite duration, i.e. measurable transition path time (Figure 3D, Figure S7). Hence, we further analysed the identified transitions by fitting a three-segment line to each transition and the flanking regions (Figure 3D). Alignment of the FRET traces at the mid-point ($[t_{start} + t_{end}]/2$) facilitated comparison of characteristic trasition path times for transition under different conditions (Figure 3E and histograms in Figure S7 for SecYEG:A + ATP).

From this representation it is clear that in the presence of ATP these transitions exhibit similar transition path times (Figure 3E and S7). However, it is important to note that many events which map along the 'dynamic line' in the E-τ plot represent faster dynamics on millisecond and sub-millisecond time scale with multiple transitions that cannot be directly resolved in our experiment. In contrast, in the presence of ADP such transitions are rare on

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the accessible time scale (up to 50 ms) (Figure S8).

The pattern of transitions in Figure 3E (ATP) suggests that opening and closing occurs via the intermediate state. This is further supported by histograms of transition frequencies obtained from HMM analysis in which direct open to closed and closed to open transitions are relatively rare (Figure S9).

Thus, while the intermediate state is only transiently populated, it seems to be an obligatory step along the LG reaction coordinate. Thus, the dynamic fraction of the intermediate state, i.e. events that map to the vicinity of the dynamic line in E-τ plots, still represents the intermediate population and not the dynamically averaged open and closed states.

In summary, the results presented here show that the LG undergoes dynamic transitions on 5-10 ms time scale (defined by the diffusion time), some of which exhibit discernible, millisecond transition path time. These timescales are an order of magnitude faster than the ATPase cycle, as well as the conformational changes of SecA recently observed by single molecule fluorescence.[12,26] However, the average population of the states is clearly linked to the SecA ATPase cycle (Figure 2C) and thus the 'fast' LG opening reaction coordinate must be stochastically coupled to the 'slow' SecA chemical reaction coordinate.

Mapping the lateral gate dynamics into reduced configuration space There are now multiple high resolution structures available of the SecYEG:A complex[3,19] as well as MD trajectories exhibiting transitions from the 'high-energy' intermediate SecYEG:A:ADP:BeF$_x$ to structures with open, part-open or closed LG.[4,9] Therefore, we reasoned that they might together contain enough information to give structural insights into the dynamic coupling between SecYEG and SecA through mapping of the dynamics along the experimentally observed LG reaction coordinate onto a reduced set of configuration space coordinates that represent the structural changes.

Principal component analysis (PCA) is a useful tool for projecting variation in complex,

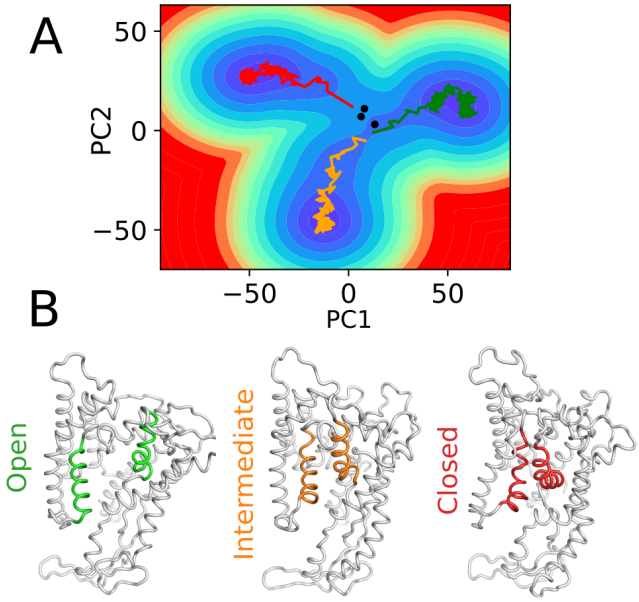


Figure 4: SecYEG:A pseudo energy landscape and associated conformers. (A) Molecular dynamics trajectories (solid lines) which were initiated from a 'high-energy' intermediate structures (black dots), are mapped onto the two most significant principal components PC1 and PC2, and relax into the basins associated with the closed (red), intermediate (yellow) and open (green) state, respecively. Black dots correspond to the several high-resolution structures (see Methods) which were also included in principal component analysis data set. A pseudo free energy landscape in the background (contours) was computed from the density of molecular dynamics states when projected onto the PC1-PC2 space. (B) SecYEG structures representing the basins of attraction. The lateral gate helices are highlighted in corresponding colors. While PCA analysis recapitulates well the experimentally observed FRET transitions it was performed for a limited number of MD frames and thus is unlikely to describe all changes associated with the ATPase cycle.

multidimesional data onto a reduced set of components (vectors) and ranking their contributions. In order to apply this method to the structural data (atomic models obtained by crystallography or MD simulations) we first have to use invariant core analysis (ICA) to identify SecYEG regions that are structurally invariant across the data set (e.g. invariant to presence and nucleotide state of SecA, see Supporting Material and Methods section for details).

This "invariant core" is then used for structural alignment of the datasets prior to PCA (Supporting Information Materials and Meth-

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ods). ICA identified the two transmembrane helices (TMHs) lining LG (TMHs 2+7, shown in pink and red, respectively in Figure S8) as mobile while these changes were also accompanied by bending of TMHs 4,5+8. PCA ranking yielded two dominant components which account for 77% structural variability among the SecY structures (Figure S9).

The high-resolution structure, which was used as starting point for MD simulations and corresponds to a trapped high-energy intermediate state, maps in the middle of the projection space defined by the two dominant principal coordinates (black dots in Figure 4A). Depending on the SecA nucleotide state, MD simulation frames then 'drift' away from this starting point towards three basins of attraction (Figure 4A). As expected, the average potential energy within these basins is lower but comparable for all conditions and fluctuates considerably (Figure S10).

FRET efficiencies which were estimated for LG pair for each MD frame quickly reach stable averages (Figure S11) which correspond to the experimentally observed values for open, intermediate and closed states, respectively. The structures representing the three basins (Figure 4B) further illustrate that most of the changes are associated with TMHs lining the LG. However, when the changes are separated according to the two PC coordinates (Supplementary Video S1 and Video S2) motion along PC1 is mostly producing LG widening and thus maps well onto the experimental FRET reaction coordinate. On the other hand, PC2 is characterized by sliding of TMHs that line the LG. The 2D representation (Figure 3A) also shows that the intermediate state is reached mostly by progression along both coordinates. While the dynamic interchange of all three states is well represented by experimental FRET coordinate (and PC1), the underlying configuration space changes are more complex and require a minimum of two coordinates.

For relatively simple systems, like equilibrium folding dynamics of small proteins, it is now possible to reconstruct free energy landscapes associated with reduced coordinates, e.g. using extensive MD sampling schemes.[27] However, this is still difficult for larger systems and combinations of experimental mapping and empirical potentials in the reduced configuration space offer a viable alternative to gain further mechanistic insights into the dynamics of such systems.[28,29]

Here, we resorted to a similar approximation and converted the density of MD states within the reduced 2D representation into a probability kernel density surface. This surface was then converted into pseudo free energy landscape as shown by the contour plot in Figure 4A. Due to the paucity of higher energy states this does not represent a true free-energy landscape but provides a convenient backdrop for discussion of the coupling between the chemical reaction coordinate (ATP turnover), the experimentally observed LG dynamics and conformational changes in the reduced configuration space.

Discussion

The stark difference in timescales between SecYEG LG dynamics (<10 ms) shown here and those of ATP hydrolysis by SecA (>100 ms) demonstrate that SecYEG and SecA cannot be coupled directly and that thermal fluctuations play an important role in the pre-protein translocation mechanism. Indeed, lateral gating by SecY has been shown previously to be strongly dependent on temperature.[30] Rapid dynamic interchange is an intrinsic property of the SecY LG, even in the absence of SecA; although the closed state is the most highly populated state (Figure 3 and 2B). The presence of SecA alone allosterically changes this rapid equilibrium by promoting the open LG conformation (Figure 2C). This allosteric effect further depends on the nucleotide state of the ATPase (Figure 2 and Figure S2). We propose that this is achieved via modulation of the energy landscape associated with LG opening and closing and this is illustrated within the reduced configuration space obtained by PCA from MD trajectories and available structures (Figure 4).[19] Given that the MD starts in a high-energy transient intermediate (repre-

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

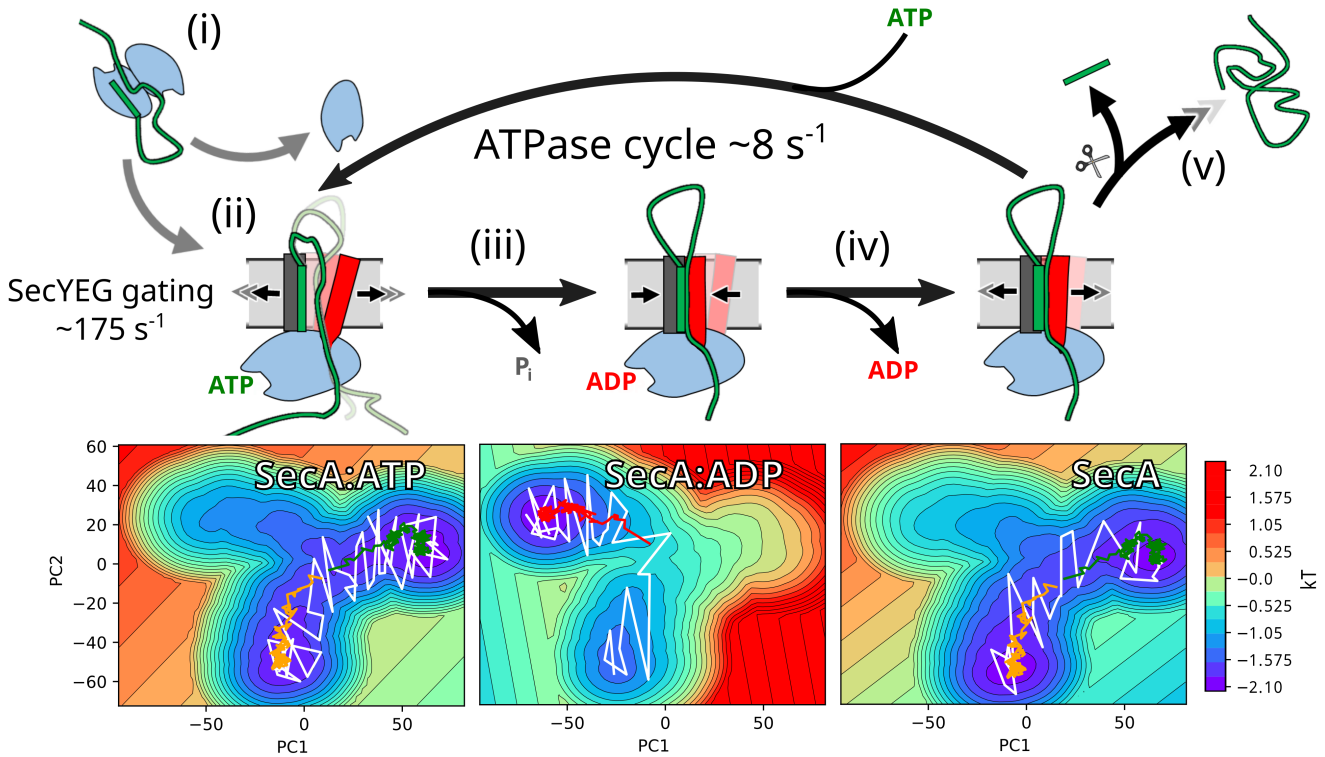


Figure 5: Coupling of the SecYEG LG dynamics to the SecA ATPase cycle via energy landscape steering. (i) The SecA dimer delivers pre-protein to the SecYEG translocon requiring nucleotide exchange (initiation, see[23]). (ii) In the SecA:ATP state, SecYEG:A steers the energy landscape associated with LG movement (as measured by the E_{FRET} reaction coordinate) to favour the SecY open and intermediate states and rapid interconversion between lateral gate states. (iii) ATP hydrolysis to ADP results in a conformational change in SecA steering the energy landscape to a dominant minimum of the closed state. (iv) ADP dissociation from SecA permits lateral gate dynamics between open and intermediate states. (v) After repeated cycles of ATP hydrolysis, protein translocation is terminated by substrate dissociation and cleavage by signal peptidase, making the overall process irreversible. We used the steady-state population of the three states (Figure 1E) to add realistic tilt (reflected in the colour scale bar) to the two-dimensional landscape. Note that the surface does not represent a true energy landscape that would allow one to make rigorous predictions e.g. about rates and energy barriers. Coloured trajectories represent the MD trajectories as in Figure 4A while the white lines illustrate hypothetical trajectories which drift along the 'orthogonal' PC2 coordinate and thus may produce longer transition path times as measured by FRET mapped mainly onto PC1.

sented by the SecYEG:A:BeFx:ADP structure) and quickly relaxes towards the three basins of attraction, this approach has limited sampling of higher energy states, e.g. those on the saddle points of the landscape and corresponding to transition states. One manifestation of this limitation is the apparent possibility of direct transition between open and closed state while such transitions are rarely observed in experiment (Figure S9). However, by linking the short MD trajectories with experimentally observed population of states (obtained from FRET efficiency histograms, Figure 2C) we can estimate the relative depth of the basins as they are being modulated by the nucleotide state of SecA, effectively tilting the energy landscape (Figure 5).

Figure 5 illustrates 'steering' of SecYEG energy landscape in the context of SecA driven polypeptide translocation. After loading the substrate the ATP state tilts the landscape in favour of dynamic equilibrium between the open and intermediate states and permits substrate diffusion. Upon hydrolysis and phosphate release (ADP state), the landscape tilts towards the closed state, disfavouring dynamics and abrogating polypeptide substrate diffusion. Since ADP release is the rate limiting step of

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

SecA ATPase cycle[18] this constitutes the long lived state and indeed very few transitions are being observed (Figure S6). Upon ADP dissociation, the landscape tilts back to a similar level as seen for the ATP state, perhaps priming the system for rapid re-binding of ATP and also permitting substrate diffusion.

While the pseudo-free energy landscape as derived here offers little quantitative insight and is merely introduced to illustrate the mechanistic concepts experimental observations can add more insights. Rapid, millisecond time scale fluctuations between states suggests that the barriers between the states are low. Such low energy barriers might be a hallmark of membrane proteins due to their high residual conformational entropy.[31] In a parabolic approximation for the barrier shape the observed millisecond transition times (Figure 3DE and S7) would implicate low curvature of the barrier top and slow diffusion across the barrier.[32] However, such one dimensional approximation is most likely too simplistic for the SecYEG system, which in reality exhibits many degrees of freedom. In the PCA reduced configuration space the interconversion between the states involves both degrees of freedom while only one (PC1) maps well onto the observable FRET changes. Thus, the system might spend considerable time wandering along 'orthogonal' degrees of freedom and producing projected trajectories with long transition path times (see schematic alternative trajectories in Figure 5, white lines).

The observed millisecond transition path times are much longer than the microsecond time scale paths previously detected in equilibrium folding trajectories of small proteins.[33,34] The folding of such model proteins is a rapid, essentially two state reaction and is being monitored by a FRET pair optimally positioned to reflect the folding reaction coordinate. In contrast, our system involves many degrees of freedom and interconversion between multiple states within a large quaternary complex. Hence, the expected transition times are likely to be longer and span multiple time scales.

Projection of the overall LG opening into reduced configuration space (see Video S1 and S2) provides further structural insights into possible coupling between SecA conformation and SecYEG dynamics. Subtle changes at the cytoplasmic SecA interface, namely repositioning of loop C2 (∼3 Å) and the C-terminal helices TMHs 5-10, lead to an amplification into substantial lateral motion of the LG TMH 7 at the periplasmic side (PC1, Video S1) as well as sliding of TMH 2 with respect to TMH 7 (PC2, Video S2). Both motions lead to an enlargement of the periplasmic cavity in the open state while the cytoplasmic cavity, which is in the vicinity of 2HF and TMHs 6-9, is significantly smaller in the ADP state.[4] This suggests that subtle movements within SecA at the SecYEG interface (e.g. the helical scaffold domain and 2HF) are amplified and leveraged through transmembrane helices leading to large changes in the LG and cytoplasmic and periplasmic cavity sizes.

So why are SecA and SecYEG not directly coupled in a power stroke-like mechanism? Our data and recent SecA studies[4,12,35] indicate that SecYEG and SecA operate on separate, but stochastically coupled timescales. Perhaps, while SecA, during its ∼500 ms cycle, can 'grab' the pre-protein substrate and translocate about 20 amino acids (based on ∼40 aa/s processive rate[23]), SecYEG operates on a much shorter time scale of <10 ms, which is less than the average passage time of a single amino acid. Thus, the two machines 'process' the translocating polypeptide on different length scales which need to be coupled during translocation. In addition, SecYEG is a versatile machinery, serving in the SecA ATP hydrolysis-driven translocation across the membrane, as well as ribosome GTP hydrolysis-driven co-translational folding into the membrane, with both processes requiring LG dynamics. Since the latter process relies on intimate probing of the lipid environment by the nascent polypeptide chain during initial recognition of the signal anchor and subsequent insertion of TMHs, it may require multiple LG configurations that are readily accessible due to the shallow energy landscape.[7,36,37] Thus, SecYEG may be 'programmed' to exhibit fast LG dynamics which then are further modulated by either SecA in nucleotide-dependent manner, or

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

by the ribosome in a GTP- and nascent chain sequence-dependent manner. We propose that these mechanisms each occur via a general energy landscape biasing, mediated by interactions of SecY cytoplasmic loops with SecA or the ribosome. Such a dynamic allosteric mechanism is not unique to SecYEG:A, since similar nucleotide-dependent steering of rapid dynamics has also been shown recently for adenylate kinase,[38] AAA+ ring ATPase,[38–40] and membrane peptide transporter,[41] all of which relied on advances in single-molecule technology.

Conclusions

We have used time-resolved smFRET combined with analysis of existing crystal structures and MD simulations to directly observe dynamic interchange between different conformational states of the SecYEG LG. The rates of dynamic exchange between states of the LG are on the order 100-200 s^{-1}, i.e. a time scale at least an order of magnitude faster than the ATPase cycle ($\approx$ 8 s^{-1} in the presence of substrate and $\approx$ 0.3 s^{-1} without[18]). While this observation adds further evidence in support of the two-way communication model presented previously,[9] it is clear that the LG opening and closing (i.e. the mechanical reaction coordinate) cannot be directly coupled to the SecA ATPase state (i.e. the chemical reaction coordinate). Thus, even if SecA itself undergoes a power-stroke like cycle, SecYEG LG does not follow it directly state-by-state, but is instead dynamically coupled to SecA conformations.

Dynamic allostery involving motions occurring across multiple time scales coupled via energy landscape bias could be ubiquitous among complex molecular machines. Further experiments exploiting the wide range of timescales and motions now available via smFRET measurements will enable discovery of the extent to which such dynamic allostery is exploited in biology.

Acknowledgement This work was funded by the BBSRC: MAW, RT and SER (BB/N017307/1), DW, WJA and IC (BBSRC: BB/N015126/1 and BB/I008675/1), RAC (BBSRC South West Bioscience Doctoral Training Partnership and BB/M003604/I). Additional support was provided by the Wellcome Trust (104632) to WJA and IC, and ERC ((FP7/2007-2013) / ERC grant agreement 32240) to SER. JAC, TF and RT are supported from European Regional Development Fund-Project "Mechanisms and dynamics of macromolecular complexes: from single molecules to cells" (No. CZ.02.1.01/0.0/0.0/15_003/0000441). JAC and TS were funded by the School of Clinical and Applied Sciences, Leeds Beckett University. TF and JAC were also supported by the The Czech Science Foundation (20-11563Y).

Supporting Information Available

- Materials and Methods, Supplementary Figures and Description of Videos (.pdf)
- Video S1: SecY projected onto PC1 (.mp4)
- Video S2: SecY projected onto PC2 (.mp4)

References

(1) Brundage, L.; Hendrick, J. P.; Schiebel, E.; Driessen, A. J.; Wickner, W. The purified E. coli integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. *Cell* **1990**, *62*, 649–57.

(2) Arkowitz, R. A.; Joly, J. C.; Wickner, W. Translocation can drive the unfolding of a preprotein domain. *EMBO J.* **1993**, *12*, 243–253.

(3) Li, L.; Park, E.; Ling, J. J.; Ingram, J.; Ploegh, H.; Rapoport, T. A. Crystal structure of a substrate-engaged SecY protein-translocation channel. *Nature* **2016**, *531*, 395–399.

(4) Corey, R. A.; Ahdash, Z.; Shah, A.; Pyle, E.; Allen, W. J.; Fessl, T.; Lovett, J. E.; Politis, A.; Collinson, I. ATP-induced asymmetric pre-protein folding as a driver of protein translocation through the Sec machinery. *Elife* **2019**, *8*, e41803.

(5) Van den Berg, B.; Clemons, J., W. M.; Collinson, I.; Modis, Y.; Hartmann, E.; Harrison, S. C.; Rapoport, T. A. X-ray structure of a protein-conducting channel. *Nature* **2004**, *427*, 36–44.

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(6) du Plessis, D. J.; Berrelkamp, G.; Nouwen, N.; Driessen, A. J. The lateral gate of SecYEG opens during protein translocation. *J Biol Chem* **2009**, *284*, 15805–14.

(7) Kater, L.; Frieg, B.; Berninghausen, O.; Gohlke, H.; Beckmann, R.; Kedrov, A. Partially inserted nascent chain unzips the lateral gate of the Sec translocon. *EMBO Rep* **2019**, e48191.

(8) Hunt, J. F.; Weinkauf, S.; Henry, L.; Fak, J. J.; McNicholas, P.; Oliver, D. B.; Deisenhofer, J. Nucleotide control of interdomain interactions in the conformational reaction cycle of SecA. *Science* **2002**, *297*, 2018–26.

(9) Allen, W. J.; Corey, R. A.; Oatley, P.; Sessions, R. B.; Baldwin, S. A.; Radford, S. E.; Tuma, R.; Collinson, I. Two-way communication between SecY and SecA suggests a Brownian ratchet mechanism for protein translocation. *Elife* **2016**, *5*, e15598.

(10) Whitehouse, S.; Gold, V. A.; Robson, A.; Allen, W. J.; Sessions, R. B.; Collinson, I. Mobility of the SecA 2-helix-finger is not essential for polypeptide translocation via the SecYEG complex. *J Cell Biol* **2012**, *199*, 919–29.

(11) Bauer, B. W.; Shemesh, T.; Chen, Y.; Rapoport, T. A. A "push and slide" mechanism allows sequence-insensitive translocation of secretory proteins by the SecA ATPase. *Cell* **2014**, *157*, 1416–29.

(12) Catipovic, M. A.; Bauer, B. W.; Loparo, J. J.; Rapoport, T. A. Protein translocation by the SecA ATPase occurs by a power-stroke mechanism. *EMBO J* **2019**, e101140.

(13) Collinson, I. The Dynamic ATP-Driven Mechanism of Bacterial Protein Translocation and the Critical Role of Phospholipids. *Frontiers in Microbiology* **2019**, *10*, 1217.

(14) Wagoner, J. A.; Dill, K. A. Molecular Motors: Power Strokes Outperform Brownian Ratchets. *Journal of Physical Chemistry B* **2016**, *120*, 6327–6336.

(15) Astumian, R. D. Design principles for Brownian molecular machines: how to swim in molasses and walk in a hurricane. *Physical Chemistry Chemical Physics* **2007**, *9*, 5067–5083.

(16) Lin, J. F.; Okada, K.; Raytchev, M.; Smith, M. C.; Nicastro, D. Structural mechanism of the dynein power stroke. *Nature Cell Biology* **2014**, *16*, 479–U208.

(17) Bell, T. A.; Baker, T. A.; Sauer, R. T. Hinge-Linker Elements in the AAA plus Protein Unfoldase ClpX Mediate Intersubunit Communication, Assembly, and Mechanical Activity. *Biochemistry* **2018**, *57*, 6787–6796.

(18) Robson, A.; Gold, V. A.; Hodson, S.; Clarke, A. R.; Collinson, I. Energy transduction in protein transport and the ATP hydrolytic cycle of SecA. *Proc Natl Acad Sci U S A* **2009**, *106*, 5111–6.

(19) Zimmer, J.; Nam, Y.; Rapoport, T. A. Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* **2008**, *455*, 936–43.

(20) Park, E.; Menetret, J. F.; Gumbart, J. C.; Ludtke, S. J.; Li, W.; Whynot, A.; Rapoport, T. A.; Akey, C. W. Structure of the SecY channel during initiation of protein translocation. *Nature* **2014**, *506*, 102–6.

(21) Ma, C.; Wu, X.; Sun, D.; Park, E.; Catipovic, M. A.; Rapoport, T. A.; Gao, N.; Li, L. Structure of the substrate-engaged SecA-SecY protein translocation machine. *Nat Commun* **2019**, *10*, 2872.

(22) Ahdash, Z.; Pyle, E.; Allen, W.; Corey, R. A.; Collinson, I.; Politis, A. HDX-MS reveals nucleotide-regulated, anti-correlated opening and closure of SecA and SecY channels of the bacterial translocon. *Elife* **2019**, *8*, e47402.

(23) Fessl, T.; Watkins, D.; Oatley, P.; Allen, W. J.; Corey, R. A.; Horne, J.; Baldwin, S. A.; Radford, S. E.; Collinson, I.; Tuma, R. Dynamic action of the Sec machinery during initiation, protein translocation and termination. *Elife* **2018**, *7*, e35112.

(24) Hizlan, D.; Robson, A.; Whitehouse, S.; Gold, V. A.; Vonck, J.; Mills, D.; Kuhlbrandt, W.; Collinson, I. Structure of the SecY complex unlocked by a preprotein mimic. *Cell Rep* **2012**, *1*, 21–8.

(25) Gouridis, G.; Karamanou, S.; Gelis, I.; Kalodimos, C. G.; Economou, A. Signal peptides are allosteric activators of the protein translocase. *Nature* **2009**, *462*, 363–7.

(26) Ernst, I.; Haase, M.; Ernst, S.; Yuan, S.; Kuhn, A.; Leptihn, S. Large conformational changes of a highly dynamic pre-protein binding domain in SecA. *Commun Biol* **2018**, *1*, 130.

(27) Noé, F.; Schütte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 19011–19016.

(28) Kerns, S. J.; Agafonov, R. V.; Cho, Y.-J.; Pontiggia, F.; Otten, R.; Pachov, D. V.; Kutter, S.; Phung, L. A.; Murphy, P. N.; Thai, V.; Alber, T.; Hagan, M. F.; Kern, D. The energy landscape of adenylate kinase during catalysis. *Nat. Struct. Mol. Biol.* **2015**, *22*, 124–131.

(29) Lu, Q.; Wang, J. Single molecule conformational dynamics of adenylate kinase: energy landscape, structural correlations, and transition state ensembles. *J. Am. Chem. Soc.* **2008**, *130*, 4772–4783.

(30) Ge, Y.; Draycheva, A.; Bornemann, T.; Rodnina, M. V.; Wintermeyer, W. Lateral opening of the bacterial translocon on ribosome binding and signal peptide insertion. *Nat Commun* **2014**, *5*, 5263.

(31) O'Brien, E. S.; Fuglestad, B.; Lessen, H. J.;

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Stetz, M. A.; Lin, D. W.; Marques, B. S.; Gupta, K.; Fleming, K. G.; Wand, A. J. Membrane Proteins Have Distinct Fast Internal Motion and Residual Conformational Entropy. *Angew. Chem. Int. Ed Engl.* **2020**, *59*, 11108–11114.

(32) Chung, H. S.; Gopich, I. V. Fast single-molecule FRET spectroscopy: theory and experiment. *Phys. Chem. Chem. Phys.* **2014**, *16*, 18644–18657.

(33) Chung, H. S.; Eaton, W. A. Protein folding transition path times from single molecule FRET. *Curr. Opin. Struct. Biol.* **2018**, *48*, 30–39.

(34) Chung, H. S.; Louis, J. M.; Eaton, W. A. Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 11837–11844.

(35) Chada, N.; Chattrakun, K.; Marsh, B. P.; Mao, C. F.; Bariya, P.; King, G. M. Single-molecule observation of nucleotide induced conformational changes in basal SecA-ATP hydrolysis. *Science Advances* **2018**, *4*.

(36) Bischoff, L.; Wickles, S.; Berninghausen, O.; van der Sluis, E. O.; Beckmann, R. Visualization of a polytopic membrane protein during SecY-mediated membrane insertion. *Nat Commun* **2014**, *5*, 4103.

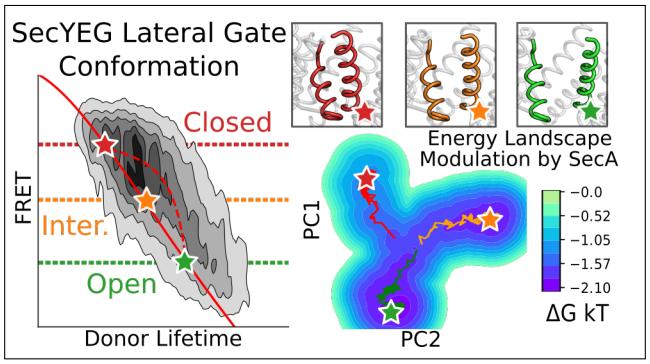
(37) Wickles, S.; Singharoy, A.; Andreani, J.; Seemayer, S.; Bischoff, L.; Berninghausen, O.; Soeding, J.; Schulten, K.; van der Sluis, E. O.; Beckmann, R. A structural model of the active ribosome-bound membrane protein insertase YidC. *Elife* **2014**, *3*, e03035.

(38) Aviram, H. Y.; Pirchi, M.; Mazal, H.; Barak, Y.; Riven, I.; Haran, G. Direct observation of ultrafast large-scale dynamics of an enzyme under turnover conditions. *Proceedings of the National Academy of Sciences* **2018**, *115*, 3243–3248.

(39) Mazal, H.; Iljina, M.; Barak, Y.; Elad, N.; Rosenzweig, R.; Goloubinoff, P.; Riven, I.; Haran, G. Tunable microsecond dynamics of an allosteric switch regulate the activity of a AAA+ disaggregation machine. *Nat Commun* **2019**, *10*, 1438.

(40) Pirchi, M.; Tsukanov, R.; Khamis, R.; Tomov, T. E.; Berger, Y.; Khara, D. C.; Volkov, H.; Haran, G.; Nir, E. Photon-by-Photon Hidden Markov Model Analysis for Microsecond Single-Molecule FRET Kinetics. *J Phys Chem B* **2016**, *120*, 13065–13075.

(41) Lasitza-Male, T.; Bartels, K.; Jungwirth, J.; Wiggers, F.; Rosenblum, G.; Hofmann, H.; Löw, C. Membrane Chemistry Tunes the Structure of a Peptide Transporter. *Angew. Chem. Int. Ed Engl.* **2020**,

bioRxiv preprint doi: https://doi.org/10.1101/793943; this version posted August 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Graphical TOC Entry