Refined measurement of SecA-driven protein secretion reveals that translocation is indirectly coupled to ATP turnover

William J. Allen, Daniel W. Watkins, Mark S. Dillingham, Ian Collinson*

*Correspondence to: Ian.Collinson@bristol.ac.uk

School of Biochemistry, University of Bristol, University Walk, Bristol BS8 1TD, UK

Abstract

The universally conserved Sec system is the primary method cells utilise to transport proteins across membranes. Until recently, measuring the activity – a prerequisite for understanding how biological systems works – has been limited to discontinuous protein transport assays with poor time resolution, or reported by large, non-natural tags that perturb the process. The development of an assay based on a split super-bright luciferase (NanoLuc) changed this. Here, we exploit this technology to unpick the steps that constitute post-translational protein transport in bacteria. Under the conditions deployed, transport of a model pre-protein substrate (proSpy) occurs at 200 amino acids per minute, with SecA able to dissociate and rebind during transport. Prior to that, there is no evidence for a distinct, rate-limiting initiation event. Kinetic modelling suggests that SecA-driven transport activity is best described by a series of large, ~30 amino acid, steps each coupled to hundreds of ATP hydrolysis events. The features we describe are consistent with a non-deterministic motor mechanism, such as a Brownian ratchet.

Significance statement

The transport of proteins across membranes is fundamental to all life. The biological machinery responsible for this process has been known for some time, but exactly how it works is still contested. In this paper, we describe how a new, high precision method for measuring protein transport combined with careful data analysis can be used to solve this problem. We find that our transport data are best described by a ratchet-type mechanism, in which the protein can diffuse one way but not the other, rather that by physical pushing of the protein.

Introduction

To transport proteins from one side of a lipid bilayer to the other, cells employ specialised, membrane-embedded molecular motors. These recognise proteins for transport, then use energy from ATP binding and hydrolysis and/or the proton-motive force (PMF) to transfer them through a polypeptide conducting channel in the membrane – usually threading them through in an unfolded state. Probably the best studied protein transporter is the *Escherichia coli* version of the ubiquitous Sec system, which handles almost all proteins destined for the cell envelope and beyond. In its post-translational mode – used for exporting periplasmic, outer-membrane and extra-cellular proteins – the cytosolic ATPase SecA binds pre-proteins with a cleavable N-terminal signal sequence (SS), and translocates them through the membrane-embedded heterotrimeric core-complex SecYEG.

In many ways, the bacterial Sec system is quite well characterised: several structures of the channel complex and motor ATPase SecA are available, alone and associated (1–3); the pathway the pre-protein takes and how various domains move have been mapped extensively using biochemical, biophysical and computational approaches (4–8); and the ATPase activity of SecA and its regulation have been subject to detailed dissection (9–13). Yet despite this, there is no definitive answer to the question: how is ATP hydrolysis actually coupled to protein transport?

Perhaps the biggest barrier to elucidating the mechanism of the Sec machinery is the huge variability of the substrate: a polypeptide composed of sequences of amino acids of different size, shape and chemistry. Unlike motors that run along DNA or RNA – which have a repeating sugar-phosphate backbone to grip onto – or those that move along predictably organised cytoskeletal helical filaments, protein transporters must by turns recognise hydrophobic and hydrophilic regions, small and bulky residues, with both positive and negative charge, and varying amounts of secondary structure. Thus, no single set of domain movements is likely to work for every part of every pre-protein. Instead, current models for Sec are not purely deterministic: they allow at least some measure of pre-protein diffusion through the channel. We have previously proposed a pure ratcheted diffusion model (6, 9, 14), while others have proposed a hybrid 'push and slide' model, in which ATP-driven power strokes are complemented by an element of diffusion (15).

Very precise measurements of protein transport are required to distinguish between these types of mechanism, such as those produced by the recently published NanoLuc transport assay (16). Here we extend the use of the NanoLuc assay to reveal the elementary steps of the ATP-driven protein transport mechanism, using the model pre-protein pro-spheroplast protein Y (pSpy). The results reveal a non-deterministic transport model, with a small apparent number of steps each of which requires hundreds of ATP turnovers *in vitro*. Transport occurs at an overall rate of about 200 amino acids per minute, is apparently dependent on the pre-protein concentration gradient across the membrane, and is not limited by a distinct initiation step.

Results

Using NanoLuc to dissect translocation kinetics

To interrogate the kinetics of protein transport in sufficient detail to reveal mechanistic information we used the recently developed NanoLuc system (16). In essence, NanoLuc luciferase missing a single β-strand (11S) is encapsulated within proteo-liposomes (PLs) incorporating the Sec machinery, while a high affinity version of the missing β-strand (Pep86) is fused to a translocation substrate. These are then mixed together in the presence of the luciferase substrate furimazine and an ATP regeneration system, allowed to equilibrate, and the reaction started by the addition of ATP. As pre-protein is transported into the PLs, Pep86 complements 11S producing a luminescent signal. This signal is generally proportional to the amount of NanoLuc for the duration of the experiment, although it eventually begins to decay due to furimazine depletion and/or furimamide accumulation. An example import curve, with background subtracted (SI Appendix, Fig. S1a; 16) is shown in Fig. 1a. It can be fitted, as a fairly good approximation, to a simple delay phase (**lag**) followed by a single exponential with an apparent rate constant (**λ**) and amplitude (**A**; Fig. 1a).

The observed kinetics are characteristic of ‘n-step sequential’ translocation mechanisms, such as those that have been applied to the analysis of DNA helicase motors (Fig. 1b; 17–19) so, the lag-exponential fit can be used to extract semi-quantitative information from the data. The amplitude (**A**) corresponds to the amount of NanoLuc that is formed when the reaction reaches completion. The **lag** before any signal occurs is the result of the accumulation of transport intermediates, because translocation of the protein involves multiple consecutive steps with comparable rates

(17). It corresponds to the sum of the time constants (i.e. $\sum 1/k_{all}$) for all steps prior to the one that yields the signal (in our case the final NanoLuc formation step; 19). The **λ** factor is complex and has a less obvious physical meaning with respect to the stepping mechanism. It contains information related to rate limiting steps in the overall pathway, as well as both the translocation step size and the static disorder in the stepping rate; we shall return to this point later.

Taking the observed kinetics into account alongside prior knowledge of the sequence of events leading to protein secretion, a minimal reaction scheme for NanoLuc-monitored protein import can be devised for use as an initial framework for analysis (Fig. 1b). Firstly, the pre-protein substrate must be recognised by the SecYEG-SecA complex (step *i*; on- and off-rates – k_{on} and k_{off}, respectively). Note that in our setup this step starts at equilibrium, as it does not require any additional input of energy. Recognition is followed by an ATP-dependent initiation step (*ii*, rate k_{init}), wherein the signal sequence unlocks the channel and primes it for transport (8, 20–23). Transport itself (step *iii*) is driven by a number (n) of ATP-dependent steps, each of which has the rate k_{step}. In a physiological context, these would presumably be assisted by the PMF (24, 25). Finally, once the Pep86 at the C-terminus of the protein has been transported into the PL, it must associate with 11S to form mature NanoLuc (step *iv*, with on- and off-rates – k_{onL} and k_{offL}, respectively).

All of these four steps must occur in order for us to measure a transport signal, but this does not necessarily mean they all contribute appreciably to the kinetics. For example, step (*ii*) is included in the model because there is ample experimental evidence that it is important for recognising secretory substrates (21, 22), but it might be too fast to affect the shape of the transport curves. Furthermore, even this relatively simple model makes some basic assumptions, e.g. that the initiated complex never dissociates (infinite processivity), which we discuss below. Note also that because only a tiny quantity of PLs are present in the reaction, the effective concentrations of all components other than SecYEG and 11S remain constant throughout the reaction.

Establishing a minimal model for transport

To explore how the parameters in the above model are related to the observed data, we carried out control experiments related to step *i* and *iv*. First, we compared reactions initiated by the addition of ATP (Fig. 1c, orange line) with those initiated by pSpy-Pep86 or SecA (Fig. 1c, purple and green lines, respectively). The only

difference between these runs is whether the step (*i*) is at equilibrium when the reaction starts (ATP), or if the SecYEG-SecA-pre-protein complex must form first (SecA or pSpy-Pep86). As the pre-equilibrated transport reaction (initiated by ATP) has a shorter lag (lag_{ATP} = 0.85 min) than the two reactions started by the addition of the pre-protein, where step *i* must also take place (lag_{pSpy} = 1.25 min; lag_{SecA} = 1.66 min), at least some transport is occurring from pre-formed complex. The difference in lag is equal to k^{-1} of assembly of the pre-initiation complex: for 2 µM pSpy this is 2.5 min^{-1} (1÷(1.25-0.85)), or 1.25 $\mu M^{-1}.min^{-1}$; while with 1 µM SecA it is 1.23 min^{-1} (1÷(1.66-0.85)) – also ~1.25 $\mu M^{-1}.min^{-1}$. Because we are most interested in extracting only k_{init}, k_{step} and n, all subsequent experiments were initiated using ATP.

We next investigated the effect of titrating pSpy-Pep86 concentration on transport. Import signal fits well to the simple exponential + lag fit (example raw data are shown in SI Appendix, Fig. S1b, and normalised in SI Appendix, Fig. S1c). The best fit parameters plotted as a function of pSpy concentration (Fig. 1d; error bars are the SEM from four repeats) show that amplitude and λ (Fig. 1d, green and pink respectively) are both strongly affected by pre-protein concentration, while lag (Fig. 1d, cyan) is affected only slightly, if at all. Assuming transport takes place according to the model in Fig. 1b (and at least some pSpy-Pep86 is pre-bound), the lag should correspond to:

Eq. 1. $$\mathrm{lag} = \frac{1}{k_{\mathrm{init}}} + \mathrm{n}\frac{1}{k_{\mathrm{step}}}$$

The observation that the lag is not pre-protein concentration dependent is therefore expected, as none of the parameters that define it are either. The plot of λ as a function of pSpy-Pep86 concentration, meanwhile, fits fairly well to a weak binding equation (Fig. 1d, pink squares) giving an apparent K_D (0.57 µM) in reasonable agreement with a previously determined affinity of pSpy for SecA (0.2 µM; 26).

The fact that signal amplitude depends on pre-protein concentration was surprising to us. PLs are diluted approximately 8000-fold (w/v) into reaction buffer for the transport experiment. So, for PLs with an internal 11S concentration of 40 µM (the maximum used here), at most 5 nM pre-protein needs to be imported for the transport reaction to reach completion. The external concentration of pSpy-Pep86 is therefore essentially unaffected by the transport reaction, and aside from one or other reaction component running out there is no obvious reason for transport to stop. Thus, the

reaction is reaching completion in a manner that is dependent on pSpy-Pep86 concentration, but is not due to it running out – we revisit this point again below.

It should be noted that signal amplitude (normalised to its maximum value within each repeat) against [pSpy-Pep86] produces a graph with large error bars (Fig. 1d, green diamonds). Presumably, these reflect the fact that signal amplitude is very sensitive to multiple different experimental parameters – especially to active SecYEG and 11S concentrations in the PLs, which are highly variable between batches. To mitigate against this, where possible all subsequent comparative experiments were performed in parallel using the same batch of PLs. It should also be mentioned that when the value for λ is very low it becomes hard to determine the x-intercept precisely; thus lags determined from transport reactions that reach completion quickly are more reliable.

Finally, we measured NanoLuc formation (step *iv*) in solution (no membranes present) by titrating pSpy-Pep86 against a fixed concentration of 11S. The rate of formation is approximately linear up to 4 μM (above which it becomes too fast to resolve on the plate reader) with a slope (k_{onL}) of 3.7 $\mu M^{-1}.min^{-1}$ (Fig. 1e, orange squares). The fitted K_D for the interaction ($K_{D,L}$) is 58 ± 24 nM (Fig. 1e, green circles; error derived from the fit), which means that for 11S concentrations used here (generally 20 μM inside the PL, and at least 5 μM), NanoLuc formation should always be completely saturated and much faster than transport (see e.g. Fig. 1c-d, and below). Consistent with this, we have previously shown that the concentration of 11S inside the vesicles has no effect on transport kinetics of a different model substrate (proOmpA) at concentrations above 1 μM (16). Therefore, the assay reports on transport kinetics, and not the formation of the active luciferase.

Determination of the initiation and transport steps using tandem pSpys

To investigate the protein transport parameters k_{init}, k_{step} and n, we next designed a series of four nearly identical 4x tandem pSpy-Pep86 variants ($pSpy_{4x}$; Fig. 2a, SI Appendix, Fig. S2). In each substrate, three of the Pep86 sequences are scrambled so they retain the same amino acid composition but give a vastly reduced signal upon transport (SI Appendix, Fig. S3a; 'D' (for dark) in Fig. 2a). The fourth is left as active Pep86 ('L' (for light) in Fig. 2a). Thus, the resulting proteins are identical save for the length of substrate that must be translocated before the functional Pep86 becomes accessible. This eliminates any potential differences in targeting and

initiation, which we find are noticeable for short substrates (SI Appendix, Fig. S3b-c). After confirming that all four bind rapidly and with high affinity to 11S (SI Appendix, Fig. S3d-g), we carried out transport experiments initiated with saturating ATP (Fig. 2b). We observed that, as the position of the active Pep86 moves later (from LDDD to DDDL, Fig. 2b) all three parameters in the lag-exponential fit are affected in a systematic manner: the lag increases with the length of substrate before Pep86, while λ and amplitude both decrease.

From Eq. 1, a plot of lag as a function of n should fit to a straight line with a slope equal to the rate of transport, and y-axis intercept equal to k_{init}^{-1}. The experimental data do indeed fit well to a straight line (Fig. 2c, cyan line), with a slope of 0.72 min.Spy^{-1}. Because this value corresponds to n/k_{step} for a single pSpy we cannot at this stage distinguish between many fast steps or few slow ones; but as the mature domain of Spy (mSpy; with SS cleaved) is 146 amino acids long it does allow us to determine an average transport rate: ~200 amino acids per minute.

The line of best fit in Fig. 2c goes almost straight through the origin, which could be taken to mean that the value of 1/k_{init} is very small – i.e. initiation is very fast compared to k_{step}. However, this does not seem consistent with previous data, which did show a slow initiation step (21). Our alternative interpretation is that initiation is accompanied by transport of a short stretch of polypeptide, equivalent to the amount transported by k_{step}. Indeed, structural evidence suggests that insertion of SS into the lateral gate with its N-terminus facing the cytosol – a key part of initiation (1) – brings about 35 amino acids of the mature domain into the SecY-SecA channel (Fig 2d). Therefore, the simplest explanation for these results is that k_{init} (from pre-bound pre-protein) and k_{step} (in the absence of PMF) are effectively the same process. This is consistent with the notion that the catalytic cycle of SecA is primarily regulating the opening and closing of the channel through SecY (9): the same widening event permits insertion of the SS during initiation, and diffusion of the pre-protein during transport.

An unexpected observation from the pSpy$_{4x}$ series is that signal amplitude also reduces as the position of the active Pep86 moves towards the C-terminus (Fig. 2b and 2c, green circles). This suggests that a significant proportion of *in vitro* transport events initiate and begin transport, but do not reach completion. One simple explanation of this is that there is a chance for the translocating pre-protein to become irretrievably trapped within the SecYEG complex, preventing any subsequent transport at that site (hereafter 'blockage'). The rate of blockage can be estimated by

fitting amplitude (A) as a function of Pep86 position (P) to a simple exponential decay ($A = A_0 e^{-P/\Lambda}$), where A_0 is the signal in the absence of transport failure, and Λ is the average number of amino acids (aa) transported before blockage. The resulting fit (green line in Fig. 2c), is to Λ = 661 aa: in other words during *in vitro* transport of $pSpy_{4x}$ the SecYEG channel becomes irreversibly blocked after on average 3.3 minutes (from ~200 aa.min^{-1}, determined above).

This also suggests that at least for $pSpy_{4x}$, the number of active SecYEG sites limits the reaction, not 11S – a conclusion supported by the fact that transport signal is independent of internal 11S concentration down to ~5 µM (SI Appendix, Fig. S3h-j). A likely explanation of this is that although the concentrations of SecYEG and 11S supplied are fairly similar (see Methods), a large proportion of the SecYEG sites are inactive when reconstituted *in vitro* (27), so 11S is effectively always in excess. It has recently been show that resetting of the SecYEG translocon after a transport event is very slow compared to transport itself *in vitro* (28); for this reason it is likely that we only observe a single transport event per SecYEG translocon.

The processivity of transport

The data shown above allow us to determine the average rate at which successful transport events occur, and to estimate a rate of blockage where the channel is completely inactivated. However, it gives no information on processivity – how often does pre-protein completely dissociate from the channel during transport and have to reinitiate from scratch? And does SecA dissociate and rebind during the course of a single transport reaction, as has been proposed (29–31)?

To investigate these questions, we carried out transport experiments under conditions where multiple turnovers were prevented. When a large excess of unlabelled ('cold') $pSpy_{2x}$ is present, it competes with pSpy-Pep86 for import sites and blocks transport of Pep86 (SI Appendix, Fig. S4a). If the excess $pSpy_{2x}$ is instead added together with the ATP, any pre-formed SecYEG-SecA-pre-protein complex will continue to translocate, but no new transport events can start – i.e single turnover conditions with respect to transport. This will reduce the total amplitude in two main ways: (i) by a fixed amount, as new transport events cannot start after the addition of ATP; and (ii) in a length-dependent manner, from any pre-protein that dissociates during transport.

When we performed this experiment using the $pSpy_{4x}$ series, we did indeed observe both length-dependent and length-independent decreases in signal (Fig. 3a, light *versus* dark pink, and SI Appendix, Fig. S4b). Fitting the quenched transport amplitudes to exponential decay (see above) gives Λ = 1040 aa when excess $pSpy_{2x}$ is added (light pink line in Fig. 3a) and Λ = 2060 aa in a parallel experiment without the cold substrate (dark pink line in Fig. 3a). From these numbers it appears that for transport of $pSpy_{4x}$ into PLs, transport failure has roughly a 50% chance of permanently blocking the channel, and a 50% chance of freeing the channel for another round of transport. However, as the fitting error is significant, we cannot accurately state from this how often complex dissociation occurs.

To investigate the processivity of SecA, we performed a similar experiment but using a large excess of the catalytically inactive SecA mutant $SecA_{D209N}$ (SI Appendix, Fig. S4c; 32). Because SecA is required to initiate transport, this prevents reinitiation just as with the competing substrate; but it also prevents transport restart after SecA dissociation and rebinding. Any additional length-dependent signal decrease over and above the competing substrate is therefore indicative of multiple SecAs interacting with a single pre-protein.

The results (Fig. 3a, light blue line; SI Appendix, Fig. S4d) give Λ = 460 aa, corresponding to a channel dwell time for SecA of 2.3 min. This is substantially lower than with a competing substrate (pink line), suggesting that one SecA can indeed be fully released from the translocating pre-protein, followed by binding of another, without the pre-protein being released from the channel in the interim. From this, we conclude that the number of SecAs used to transport a single substrate is determined kinetically: only one is needed, but multiples can be used if transport takes longer than the dwell time of SecA on the machinery. An extended schematic model incorporating the additional possible fates of a translocating pre-protein is shown in SI Appendix, Fig. S4e.

The ATP dependence of pre-protein transport

Both initiation and transport of pre-proteins are driven by cycles of ATP binding and hydrolysis in SecA (10). The ATP turnover reaction itself has been well characterised in the past (11, 13), but how it is coupled to transport is less well understood. We therefore measured import of each of the Spy_{4x} series at a range of ATP concentrations: all three parameters (lag, λ and amplitude) are affected, in a

similar manner for all four substrates (Fig. 3b-d and S4f). As expected from Eq. 1, the lag remains proportional to the number of Spys before the active Pep86 (and thus the number of steps n) as [ATP] (and thus k_{step}) is lowered, with the slope of the line becoming steeper (Fig. 3b). The corresponding transport rates and x-intercepts for these data are plotted in Fig. 3c. Rate as a function of [ATP] fits well to the Michaelis-Menten equation, giving an apparent K_M for ATP of 20 µM. The x-intercept is close to the origin and does not change significantly with [ATP], despite k_{init} also requiring ATP turnover – consistent with k_{init} simply being k_{step}. Note that as mentioned above, determining lag accurately becomes more difficult for low values of λ, hence the scatter at low [ATP].

Both amplitude and λ vs [ATP] also fit well to the Michaelis-Menten equation for all four substrates (Fig. 3d and SI Appendix, Fig. S4f). The K_M for ATP determined from λ (15.4 µM, obtained by globally fitting all four data sets; SI Appendix, Fig. S4f) is very similar to that determined from lag (20 µM; Fig 3c), and fairly close to the value of 46 µM determined for the ATPase activity of translocating SecA (13). The small discrepancy perhaps reflects the fact that K_M determined here only reports on successfully translocated pre-proteins, whereas bulk ATPase activity includes all SecA. For amplitude (Fig. 3d), the apparent K_M for ATP is much lower, at 2.6 µM; the discrepancy suggests that the overall amount of pre-protein transported is not directly correlated with transport rate.

Evaluating possible transport models numerically using Berkeley Madonna

The 'single exponential plus lag' equation used thus far (Fig. 1a) is straightforward to fit and describes each data set with reasonable accuracy, however it is not immediately evident what λ actually corresponds to in physical terms. Furthermore, the fits deviate significantly from the data at the point where the lag and exponential meet (which we will refer to as the 'start phase') – the part of the curve that should contain information about how the motor is distributed along its substrate (19). We therefore sought to fit the data more directly to physical models of transport using numerical integration techniques.

In a biochemical reaction scheme such as the one in Fig. 1b, the concentration of each component changes as a function of time, dependent on the processes that populate it and depopulate it. Because these processes are themselves concentration dependent, the overall reaction can be described by a set of differential equations.

Analytical solutions to such problems become highly complex even for fairly simple reaction schemes, so to evaluate different transport models we used numerical integration, as implemented by the software package Berkeley Madonna. In this method, the complete set of differential equations for a given model is defined (see SI Appendix, Section S2), along with all rates, and the initial concentrations of each species. Next, each concentration is recalculated in very small time increments, and formation of the measured component – in this case NanoLuc – determined as a function of time. These simulated data can then be compared to experimental data, varying the unknown values to try to obtain a reasonable fit.

The model in Fig. 1b is defined for Berkeley Madonna in SI Appendix, Section S2. For simplicity, two additional assumptions are made: that step (*i*) is at equilibrium when the reaction starts, and that NanoLuc formation is instant. The first is reasonable, given that the assay setup includes an 8 min incubation step prior to the addition of ATP and k_{on} is of the order of 1.25 $\mu M^{-1}.min^{-1}$ (see above), while the latter is effectively true under the conditions used here (see Fig. 1e). The value for k_{off} was set to 0.7125 min^{-1}, to give K_d = 0.57 (as estimated from Fig. 1d; where $K_d = k_{off} \div k_{on}$). We also include two additional rate constants: dissociation of the translocating complex, allowing reinitiation (k_{fail}); and blockage of the channel, preventing any more transport (k_{block}). This complete model is illustrated in SI Appendix, Fig. S5a. Note that the value k_{init} is also set to equal k_{step}, as per the results above, however the same results are produced if k_{init} is set very fast and n is increased by 1. For modelling purposes, we set the concentrations of SecYEG and 11S to 4 nM and 5 nM respectively. Although we do not know the exact concentration of active SecYEG, it makes no difference to resulting traces unless it is higher than 11S (which we know is not the case, see SI Appendix, Fig. S3h-j).

A simulated transport curve – using reasonable values for each parameter based on the simpler fitting results above, then optimised to fit a real transport data set – is shown in Fig. 4a. Just as with the real data, simulated transport can be divided into four phases: a lag, characteristic of n-step sequential mechanisms; a start phase, which contains information about the number of steps and static disorder; a burst phase, produced from pre-formed SecYEG-SecA-pre-protein complex; and an end phase, as transport slows to a halt. Note that in this model each SecYEG only turns over a single pre-protein; however, in practice we find that allowing multiple turnovers

only affects the very end of the transport curves, where it is swamped by signal decay – particularly as subsequent turnovers are likely to be significantly slower [28].

While the above model has a lot of parameters, each one only affects a limited part of the transport curve (SI Appendix, Fig. S5b). Most importantly, k_{step} and n are the only parameters that make any appreciable difference to the lag and start phases, with many fast steps giving rise to a sharp start phase and a few slow steps producing a more diffuse start (SI Appendix, Fig. S5b). The other parameters each affect the size of the burst phase and the shape of the end phase in a subtly different way, and we were unable to reproduce experimental data as well if any one of them is eliminated. We are therefore confident that this model represents the simplest solution that adequately describe the data. As an illustration of this, the experimental ATP concentration dependence can be reproduced down to 10 µM ATP (Fig. 4b) only by varying k_{step} (the only ATP-dependent rate constant).

Using the Berkeley Madonna model to estimate elementary step size

Because the shape of the beginning of the transport curve is affected differently by n and k_{step} (and not by any other parameter), our model allows us to estimate the number of individual steps that make up transport. To do this we fixed k_{block}, k_{on} and k_{off} at the values approximated above, then used the Berkeley Madonna curve fit algorithm to find best fit values for k_{step}, k_{fail} and brightness at a range of different values for n (Fig. 5a, SI Appendix, Fig. S6). The best fit is to n = 5, which equates to about one step every ~30 amino acids, with a k_{step} of around 3.0 min^{-1}. The corresponding k_{fail}, 0.13 min^{-1}, is well within the expected range from the single turnover transport reaction (Fig. 3a).

In practice, this analysis is complicated by the phenomenom of static disorder, whereby different, ostensibly identical motors can transport at different rates (33, 34). This can lead to an underestimation of the true number of steps. For example, ensemble measurements on the helicase PcrA overestimate step size about 4-fold compared to the true value determined by single molecule analysis (33). Notwithstanding this effect, we conclude that transport is best described by a relatively small number of steps – while bearing in mind that the exact number may be an underestimate.

So what are these steps? The measured k_{cat} of the SecA ATPase activity when transporting pSpy is ~850 min^{-1} for *in vitro* transport with purified components (SI

Appendix, Fig. S7); not too far from the previously determined value of 450 min^{-1} for proOmpA (13). This corresponds to around 610 ATP turnovers per Spy, or 120 ATP turnovers per step. Clearly, therefore, each step represents many turnovers of ATP: this is also illustrated in Fig. 5b by the poor fitting of the data to simulations with k_{step} = 850 min^{-1} and n = 610 (i.e. where each step corresponds to a single turnover of ATP, as would be expected of a purely deterministic 'power-stroke' mechanism).

In summary, each elementary transport step transports around 30 amino acids – or somewhat fewer, depending on static disorder – and consumes around 120 ATP molecules. The most plausible explanation for this observation is that not every ATP turnover gives rise to a transport event – i.e. the ATPase activity of SecA is coupled to polypeptide movement indirectly, not directly. Indirect coupling is a hallmark of our previously proposed Brownian ratchet transport model (9, 14), wherein blockages at the entrance to the SecYEG channel trigger nucleotide exchange, giving the blockage the opportunity to diffuse through (Fig. 5c). In this interpretation, each step is a blockage, and the probability of any given ATP turnover resolving this blockage (p_{res}) is 0.8% (1/120) for *in vitro* transport of pSpy into proteo-liposomes. By contrast, for a tightly coupled power stroke motor to produce the same kinetic profile, one cycle would have to consume 120 ATP molecules (SI Appendix, Fig. S8) – even ignoring the fact that the proposed piston (the two-helix finger; 15) can be crosslinked in place without preventing transport (35).

Discussion

The advent of an assay capable of measuring protein transport accurately, with a time resolution of seconds, has opened the door to investigating the process on a detailed functional as well as a structural level. While well-conducted studies of the kinetics of protein transport have been performed previously (27, 36), their interpretation has always been limited by the poor time resolution inherent to end point measurements. Using NanoLuc, we have here built up a detailed model of transport, and a thorough understanding of how each transport parameter affects the measured luminescence signal.

To analyse the data, we have developed two different approaches. Fitting to a simple model requires no specialist software and gives two useful parameters: lag, which corresponds to the minimum time required for transport; and amplitude, which correlates with the total amount of substrate transported. A complete numerical

integration solution, meanwhile, validates the overall model and provides estimates for each individual rate constant. It should be noted that lag is relatively independent of experimental variables, and is thus a robust measure of transport rate. It will therefore be particularly useful for evaluating differences in transport by Sec mutants, or of different pre-protein substrates, with high sensitivity. Amplitude, meanwhile, is highly susceptible to experimental error, and so should ideally be interpreted from experiments run in parallel, using the same reagents.

Surprisingly, we find that initiation – although clearly a critical part of the mechanism for recognising genuine Sec substrates – does not seem to contribute appreciably to the overall kinetics of transport. Indeed, if anything the first few amino acids of mSpy are transported faster than the rest (Fig. 2c, 3c). A likely explanation for this is that insertion of the signal sequence into the LG of SecA – oriented to keep the positively charged N-terminus in the cytosol and bring the narrow, hydrophilic C-terminus through the channel (Fig. 2d) – provides an extra driving force to pull the first amino acids across the membrane. The idea that initiation and transport are effectively the same process is further supported by a structural model of the pre-initiation complex, based on FRET constraints, in which signal sequence and the beginning of the mature domain form a hairpin poised at the entrance to the channel through SecY (37).

The net transport rate of ~200 amino acids per minute determined here is similar to one previous estimate of transport rate, for fluorescently labelled proOmpA into IMVs (36), but roughly 10-fold slower than translocation of unlabelled proOmpA into PLs (21). This can partly be ascribed to the nature of the substrate: proOmpA was originally chosen as a model translocation substrate precisely because it is secreted very efficiently. However it should also be noted that the NanoLuc assay reports only on successful transport events, whereas the single molecule assay in Fessl *et al.* uses movement of the plug domain of SecY as a proxy for transport. Therefore, the higher rates reported in Fessl *et al*. may also partially reflect a fraction of initiated, but subsequently aborted events (represented by by k_{fail} in the model in SI Appendix, Fig. S5a).

Almost all the observed experimental data can be reproduced by the Berkeley Madonna transport model (SI Appendix, Fig. S5a) using reasonable values for each of the six rate constants. However, the strong dependence of signal amplitude on pre-protein concentration (Fig. 1d and S1b) is not explicable within this framework, as

pSpy is always in huge excess, and there is nothing else to stop transport except the gradual blocking of translocons (Fig. 3a; SI Appendix, Fig. S4a-d). A plausible explanation for this is that the Sec complex behaves like a classical membrane transporter described by Peter Mitchell (38): whereby transport is partially driven by a concentration gradient, in this case of pre-protein, across the membrane; this could be either total protein or some particular feature of the translocating pre-protein such as charged residues. This is, again, a feature expected of a Brownian ratchet style mechanism – where transport is influenced by the relative rates of inward and outward diffusion (see SI Appendix, Fig. S9 for more detail) – but not of a directly coupled power stroke motor.

A slowing of transport as pre-protein accumulates inside the PLs would go some way towards explaining why transport assays performed *in vitro* are so much slower than the rates expected *in vivo* (39, 40). But it cannot explain a predicted difference of nearly 2 orders of magnitude (39, 40). A clue to this may come from the extremely low probability that any given ATP turnover event gives rise to transport *in vitro* (p_{res} = 0.8%; see Fig. 5c). This value seems implausibly low, so it is very likely that other factors will increase this value substantially *in vivo*. These might include auxiliary drivers of transport, particularly the PMF (24, 25), and some of the many other proteins that associate with the Sec system such as SecDF, PpiD and YfgM (41, 42). It does, however, make sense that all these factors affect p_{res}, not ATP turnover itself: it is of course far easier to add additional driving forces to a ratchet than to a directly-coupled motor.

One additional factor we believe will prove particularly critical to understanding the slow *in vitro* transport rates is the folding state of the pre-protein, which is known to be important for enabling transport (6, 43). Chaperones generally capture pre-proteins *in vivo* as they are translated and deliver them to the membrane in an optimally translocation competent state. *In vitro*, meanwhile, pre-proteins are diluted out of urea, and so have far more opportunity to form folding intermediates that delay transport. Without extra assistance, the diffusion-based transport motor has little power to unfold pre-proteins: instead it must wait for a spontaneous unfolding event prior to trapping them in an unfolded state within the channel. Moreover, it seems that some secretory proteins are delivered directly to SecA, lurking at the ribosome exit site, during their translation (44, 45). Thus, it seems plausible that in the presence of

a stimulatory PMF (24), careful pre-protein management and bespoke ancillary factors, transport could easily be sped up by at least an order of magnitude.

The twin developments described here: an assay that generates high quality transport data and a fitting process capable of describing it, together provide the first fully quantitative framework for understanding the mechanism of ATP-driven transport through Sec. We anticipate that the experimental and data analysis approaches developed here will be very useful in the future both for furthering our understanding of the bacterial Sec machinery, and also far more broadly to study many other membrane transport processes.

Acknowledgements

This work was funded by the BBSRC (BB/S008349/1 and BB/N015126/1 to DWW and IC) and Wellcome (104632 to WJA and IC). Thanks to Robin Corey for critical reading of the manuscript.

Author Contributions

WJA, DWW, MSD and IC designed experiments; WJA and DWW conducted experiments; WJA, DWW, LT and IC wrote the manuscript; IC secured funding and led the project.

Declaration of Interests

The authors declare no competing interests

Methods

Reagents

All previously reported reagents for transport assays, including SecYEG, SecA, pSpy and PLs were produced exactly as described previously (16). Proteoliposomes containing only SecYEG, for ATPase assays, were prepared as described in (11). The $pSpy_{2x}$ gene was synthesised commercially (GeneArt Gene Synthesis service, Thermo Fisher Scientific), then cloned into a pBAD vector and expressed exactly as pSpy-Pep86 (16).

The Spy_{4x} series was constructed by first ordering 3 fragments of Spy containing 2 tandem repeats of the mature region, each with a different combination of 'light' (L, active Pep86, VSGWRLFKKIS) or 'dark' (D, inactive pep86, VSWGLRKFKIS) Pep86 sequences in the following combinations: DD, LD and DL, where for example, DD contained two inactive HiBiT sequences at the C terminus of each mature region of Spy (produced by the GeneArt Gene Synthesis service, Thermo Fisher Scientific). For cloning purposes, each fragment began with residue A24 of pSpy and ended with a GSG linker immediately following the second Pep86 sequence (sequence 1 in SI Appendix, Fig. S2). The fragment was cloned into pBAD-pSpy-V5-pep86-TEV-His (16; sequence 2 in SI Appendix, Fig. S2) using site directed ligase independent mutagenesis (46). More specifically, the fragments were cloned in the place of Spy-V5, using the same primers to introduce a ZraI site (GACGTC) immediately after the fragment and before the TEV cleavage site of the template, to give pBAD-(LL, LD or DL)-ZraI-TEV-His (sequence 3 in SI Appendix, Fig. S2). The synthesised fragments were then amplified using linear PCR and cloned into the ZraI site of pSpy LL, LD and DL to give DDDD, LDDD, DLDD, DDLD and DDDL (Sequence 4 in SI Appendix, Fig. S2).

NanoLuc formation assays

NanoLuc formation was measured at 25 ˚C in a BioTek Synergy Neo2 plate reader. A dilution series of pSpy-Pep86 was prepared in the wells using TKM with furimazine (to 1/500) and Prionex (to 0.125%), with a volume of 100 µl per well. Reactions were started by injecting 25 µl 11S at 1 nM in TKM (to give 200 pM final), then shaken for 2 s and the luminescence monitored with no emission filter.

NanoLuc transport experiments

Standard transport experiments were performed at 25 ˚C exactly as in our recent methods paper (16). Essentially, master mixes were assembled containing all transport components except pre-protein and ATP (unless otherwise stated). Unless otherwise specified, we used PLs with an internal 11S concentration of 20 µM, resuspended ~7-fold (w/v) to give a final concentration of 9.2 µM SecYEG (of which half ~4.6 µM is oriented correctly to participate in transport; see ref 9). These were then diluted 550-fold into reaction buffer, giving [SecYEG] ~ 8.4 nM and [11S] ~ 5.2 nM. Note, however, that a large fraction of the SecYEG translocons are expected to be inactive (27), and empirically we observe that 11S is in excess over active SecYEG translocons even at 5 µM internal concentration (equivalent to 1.3 nM total; see SI Appendix, Fig. S3h-j). SecA and pSpy were generally provided at 1 µM, unless otherwise specified.

To follow transport, we first added pre-protein and measured background for 8 minutes. Transport itself was then initiated by the addition of ATP (to 1 mM unless otherwise stated), then monitored for 25 minutes, or until all reactions had reached completion. In most cases, 8 reactions were performed in parallel and luminescence was measured using a BioTek Synergy Neo2 plate reader. These luminescence values reflect the rate of photon emission, which is generally proportional to NanoLuc concentration over the measured time ranges (16). For the earlier reactions with different initiation conditions (Fig. 1c), reactions were instead performed one at a time in a Jobin Yvon Fluorolog (Horiba) with the lamp turned off.

For subsequent reactions where additional reagents were added together with the ATP (Fig. 3a-b), we modified the plate reader protocol to allow manual injection. 100 µl reactions with all components at 1.2x final concentration were assembled as above (except with SecA at 100 nM final instead of 1 µM), and measured for 8 mins after the addition of pre-protein. The plate was then ejected, and 20 µl ATP together with the other reagent – both at 6x final concentration – added and mixed immediately using a multi-channel pipette. Measurement was then resumed as fast as possible. The manual mixing step adds an additional constant error to reaction time (t), which we estimate to be less than 5 s.

Data analysis

Initial data processing was performed using pro Fit 7 (Quansoft). Raw data before the addition of ATP were fitted to the single molecule plus lag model (SI Appendix, Section S1.1), and the fits subtracted to give transport signal (see SI Appendix, Fig. S1a). Corrected data were then fitted to the same model, to give lag, λ and amplitude for transport.

The model for Berkeley Madonna is described in detail in SI Appendix, Section S2.

ATPase assays

Steady-state ATPase assays were performed as in (11), with ATP consumption calculated from the decrease in NADH absorbance at 340 nm, measured in a Lambda 25 spectrophotometer (PerkinElmer). Reactions were conducted in TKM buffer with final concentrations: 230 nM SecYEG in proteoliposomes, 60 nM SecA, 2 mM phosphoenol pyruvate, 6-10 units of pyruvate kinase and 9-14 units of lactate dehydrogenase (Merck), 0.2 mM NADH, pSpy at specified concentration, and 1 mM ATP. Reactions were incubated at 25°C for 5 minutes prior to the addition of ATP or pSpy, then ATP was added and the basal SecA ATPase rate measured for 10 minutes. Finally, the translocation ATPase rate was measured following addition of pSpy. ATPase rates were calculated from the slope of the line as in (11).

Figure Legends

Figure 1: The NanoLuc transport assay

a) An example NanoLuc transport curve, with simple fitting.

b) The minimal model used to describe pre-protein import in the NanoLuc assay.

c) Transport initiated by the addition ATP (orange), pSpy-Pep86 (purple) or SecA (green). Lines represent best fit to the single exponential + lag model.

d) Fitted λ (pink squares), amplitude (green diamonds) and lag (cyan circles) as a function of pSpy-Pep86 concentration. Error bars represent average and SEM from 4 repeats. λ is fitted to a weak binding equation.

e) Secondary data from a titration of pSpy-Pep86 against 200 pM 11S in solution. Fits are to a weak binding equation for amplitude (green circles) and a straight line for rate (orange squares).

Figure 2: The tandem pSpy-Pep86 series.

a) Schematic of the tandem pSpy-Pep86 series.

b) Average transport of 1-2 µM of the $pSpy_{4x}$ series into vesicles containing 10-20 µM 11S, normalised to LDDD for each run. LDDD is pink, DLDD green, DDLD beige and DDDL blue. Error bars are the SEM of 8 repeats.

c) Normalised amplitude (green circles) and lag (cyan circles) as a function of active Pep86 position (and equivalent in amino acids) for the Spy_{4x} series, extracted from the data in panel **b**. Fits are to straight lines (lag) and exponential decay (amplitude).

b) Model of SecYEG (red) and SecA (blue) with a pre-protein in the channel (SS in black, mature in green; 39). Initiation brings ~35 amino acids of mature domain into the SecY-SecA complex.

Figure 3: Processivity and ATP-dependence of transport

a) Amplitude of transport signal for tandem $pSpy_{4x}$ series with 50 µM $pSpy_{2x}$ (pink) or 10 µM $SecA_{D209N}$ (cyan) added at the same time as the ATP. Equivalent data with only ATP, performed in parallel, are shown in dark pink and blue, respectively (n.b. these overlay very closely, largely obscuring the blue data). Fits are to exponential decay, and normalised to the fitted value of A_0 with only ATP.

b) Lag as a function of active Pep86 position in the Spy_{4x} series, at a range of ATP concentrations: red = 320 µM, orange = 160 µM, yellow = 80 µM, green = 40 µM, teal = 20 µM, blue = 10 µM and purple = 5 µM. Fits are to straight lines.
c) Transport rate (black circles) and x-intercept (orange squares) extracted from the fits in panel **b** as a function of ATP concentration. Transport is fitted to the Michaelis-Menten equation (black line), while the orange line is the mean x-intercept.
d) Amplitude as a function of [ATP] for the Spy_{4x} series. Lines are global fits to a weak binding equation.

Figure 4: Numerical modelling of transport with Berkeley-Madonna
a) Example simulated transport curve for the model in SI Appendix, Section S2.1 and Fig. S5a. Parameters are: k_{step} = 6.3 min^{-1}; n = 6; k_{on} = 2 µM.min^{-1}; k_{off} = 1.6 min^{-1}; k_{fail} = 0.19 min^{-1}; k_{block} = 0.13 min^{-1}; brightness = 283.
b) Transport signal for $pSpy_{4x}$ LDDD with 320, 40, 20, 10 and 5 µM ATP (black lines). Dotted lines are simulated data (as in panel a) with k_{step} = 6.3 min^{-1} (red), 5.05 min^{-1} (orange), 3.85 min^{-1} (yellow), 2.65 min^{-1} (green) and 1.9 min^{-1} (blue).

Figure 5: The elementary transport step
a) Root mean square deviation (RMSD) for the best fits to the transport data for pSpy-Pep86. In each case, k_{block} was fixed at 0.31 min^{-1} (from the green fit in Fig. 2c), k_{on} at 1.25 $\mu M^{-1}.min^{-1}$ (from the purple vs orange data in Fig. 1c), and k_{off} at 0.7125 min^{-1} (to give K_d = 0.57 µM, as in Fig. 1d, pink fit). The fits themselves are shown in SI Appendix, Fig. S6.
b) Attempting to fit a transport model where k_{step} is the rate of ATP turnover, presented as in SI Appendix, Fig. S6. With many steps (n = 610 shown, chosen to give the correct lag; pink line), the start and burst phases become extremely sharp, making it impossible to fit the experimental data. The best fit trace (n = 5; green line) is shown for comparison.
c) Schematic of the action of ATP in the Brownian ratchet model of pre-protein transport, with components coloured as in Fig. 1b. In the ATP bound state (middle), blockages at the entrance to SecY have a probability of resolving (p_{res}) determined by their diffusion through the channel.

Bibliography

1. B. V. den Berg, *et al.*, X-ray structure of a protein-conducting channel. *Nature* **427**, 36–44 (2004).

2. J. F. Hunt, *et al.*, Nucleotide control of interdomain interactions in the conformational reaction cycle of SecA. *Science (New York, N.Y.)* **297**, 2018–2026 (2002).

3. J. Zimmer, Y. Nam, T. A. Rapoport, Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* **455**, 936–943 (2008).

4. B. W. Bauer, T. A. Rapoport, Mapping polypeptide interactions of the SecA ATPase during translocation. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 20800–20805 (2009).

5. K. S. Cannon, E. Or, W. M. Clemons, Y. Shibata, T. A. Rapoport, Disulfide bridge formation between SecY and a translocating polypeptide localizes the translocation pore to the center of SecY. *The Journal of cell biology* **169**, 219–225 (2005).

6. R. A. Corey, *et al.*, ATP-induced asymmetric pre-protein folding as a driver of protein translocation through the Sec machinery. *eLife* **8** (2019).

7. V. A. M. Gold, S. Whitehouse, A. Robson, I. Collinson, The dynamic action of SecA during the initiation of protein translocation. *The Biochemical journal* **449**, 695–705 (2012).

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

9. W. J. Allen, *et al.*, Two-way communication between SecY and SecA suggests a Brownian ratchet mechanism for protein translocation. *eLife* **5** (2016).

10. A. Economou, J. A. Pogliano, J. Beckwith, D. B. Oliver, W. Wickner, SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF. *Cell* **83**, 1171–1181 (1995).

11. V. A. M. Gold, A. Robson, A. R. Clarke, I. Collinson, Allosteric regulation of SecA: magnesium-mediated control of conformation and activity. *The Journal of biological chemistry* **282**, 17424–17432 (2007).

12. R. Lill, *et al.*, SecA protein hydrolyzes ATP and is an essential component of the protein translocation ATPase of Escherichia coli. *The EMBO journal* **8**, 961–966 (1989).

13. A. Robson, V. A. M. Gold, S. Hodson, A. R. Clarke, I. Collinson, Energy transduction in protein transport and the ATP hydrolytic cycle of SecA. *Proceedings*

of the National Academy of Sciences of the United States of America **106**, 5111–5116 (2009).

14. Z. Ahdash, *et al.*, HDX-MS reveals nucleotide-dependent, anti-correlated opening and closure of SecA and SecY channels of the bacterial translocon. *eLife* **8**, E1786 (2019).

15. B. W. Bauer, T. Shemesh, Y. Chen, T. A. Rapoport, A “Push and Slide” Mechanism Allows Sequence-Insensitive Translocation of Secretory Proteins by the SecA ATPase. *Cell* **157**, 1416–1429 (2014).

16. G. C. Pereira, *et al.*, A High-Resolution Luminescent Assay for Rapid and Continuous Monitoring of Protein Translocation across Biological Membranes. *Journal of molecular biology* **431**, 1689–1699 (2019).

17. A. L. Lucius, N. K. Maluf, C. J. Fischer, T. M. Lohman, General methods for analysis of sequential “n-step” kinetic mechanisms: application to single turnover kinetics of helicase-catalyzed DNA unwinding. *Biophysical journal* **85**, 2224–2239 (2003).

18. J. A. Ali, T. M. Lohman, Kinetic Measurement of the Step Size of DNA Unwinding by Escherichia coli UvrD Helicase. *Science* **275**, 377–380 (1997).

19. S. E. McClelland, D. T. F. Dryden, M. D. Szczelkun, Continuous Assays for DNA Translocation Using Fluorescent Triplex Dissociation: Application to Type I Restriction Endonucleases. *J Mol Biol* **348**, 895–915 (2005).

20. R. A. Corey, *et al.*, Unlocking the Bacterial SecY Translocon. *Structure (London, England : 1993)* **24**, 518–527 (2016).

21. T. Fessl, *et al.*, Dynamic action of the Sec machinery during initiation, protein translocation and termination. *eLife* **7** (2018).

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

23. D. Hizlan, *et al.*, Structure of the SecY complex unlocked by a preprotein mimic. *Cell reports* **1**, 21–28 (2012).

24. L. Brundage, J. P. Hendrick, E. Schiebel, A. J. Driessen, W. Wickner, The purified E. coli integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. *Cell* **62**, 649–657 (1990).

25. E. Schiebel, A. J. Driessen, F. U. Hartl, W. Wickner, Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell* **64**, 927–939 (1991).

26. K. E. Chatzi, *et al.*, Preprotein mature domains contain translocase targeting signals that are essential for secretion. *The Journal of cell biology* **216**, 1357–1369 (2017).

27. P. Bariya, L. L. Randall, Coassembly of SecYEG and SecA Fully Restores the Properties of the Native Translocon. *J Bacteriol* **201**, JB.00493-18 (2018).

28. C. Mao, P. Bariya, Y. Suo, L. L. Randall, Comparison of Single and Multiple Turnovers of SecYEG in E. coli. *J Bacteriol* (2020) https:/doi.org/10.1128/jb.00462-20.

29. C. Mao, *et al.*, Stoichiometry of SecYEG in the active translocase of Escherichia coli varies with precursor species. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11815–11820 (2013).

30. J. Young, F. Duong, Investigating the stability of the SecA–SecYEG complex during protein translocation across the bacterial membrane. *J Biol Chem* **294**, 3577–3587 (2019).

31. K. Morita, H. Tokuda, K.-I. Nishiyama, Multiple SecA Molecules Drive Protein Translocation across a Single Translocon with SecG Inversion. *Journal of Biological Chemistry* **287**, 455–464 (2012).

32. A. Robson, A. E. G. Booth, V. A. M. Gold, A. R. Clarke, I. Collinson, A large conformational change couples the ATP binding site of SecA to the SecY protein channel. *Journal of molecular biology* **374**, 965–976 (2007).

33. J. Park, *et al.*, PcrA helicase dismantles RecA filaments by reeling in DNA in uniform steps. *Cell* **142**, 544–55 (2010).

34. P. R. Bianco, *et al.*, Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature* **409**, 374–378 (2001).

35. S. Whitehouse, *et al.*, Mobility of the SecA 2-helix-finger is not essential for polypeptide translocation via the SecYEG complex. *The Journal of cell biology* **199**, 919–929 (2012).

36. D. Tomkiewicz, N. Nouwen, R. van Leeuwen, S. Tans, A. J. M. Driessen, SecA supports a constant rate of preprotein translocation. *The Journal of biological chemistry* **281**, 15709–15713 (2006).

37. Q. Zhang, Y. Li, R. Olson, I. Mukerji, D. Oliver, Conserved SecA Signal Peptide-Binding Site Revealed by Engineered Protein Chimeras and Förster Resonance Energy Transfer. *Biochemistry* **55**, 1291–1300 (2016).

38. P. MITCHELL, A General Theory of Membrane Transport From Studies of Bacteria. *Nature* **180**, 134–136 (1957).

39. I. Collinson, R. A. Corey, W. J. Allen, Channel crossing: how are proteins shipped across the bacterial plasma membrane? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **370** (2015).

40. T. Cranford-Smith, D. Huber, The way is the goal: how SecA transports proteins across the cytoplasmic membrane in bacteria. *FEMS Microbiology Letters* **365** (2018).

41. J. A. Pogliano, J. Beckwith, SecD and SecF facilitate protein export in Escherichia coli. *The EMBO journal* **13**, 554–561 (1994).

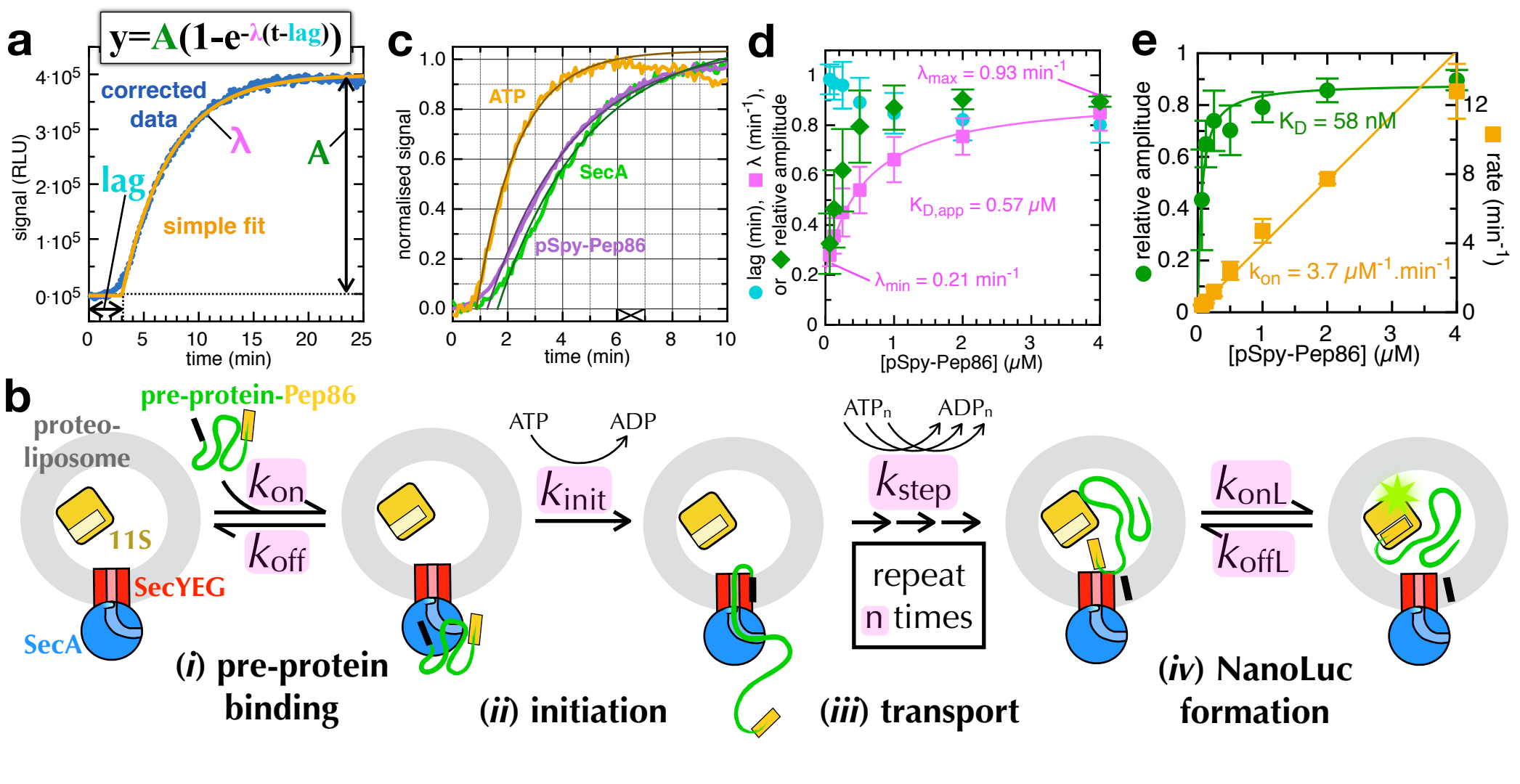
42. B. Jauss, *et al.*, Noncompetitive binding of PpiD and YidC to the SecYEG translocon expands the global view on the SecYEG interactome in Escherichia coli. *J Biol Chem* **294**, 19167–19183 (2019).

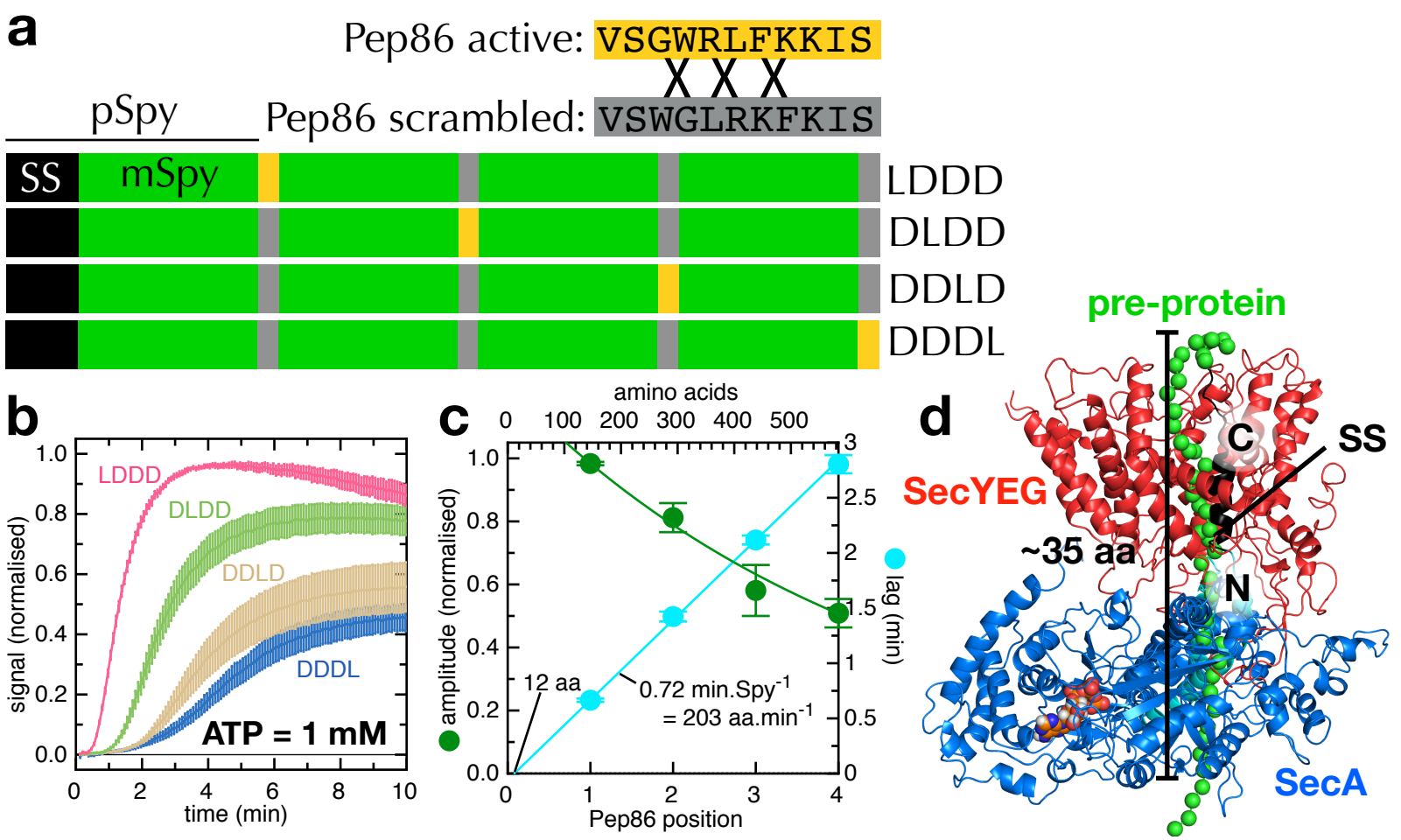
43. A. Tsirigotaki, *et al.*, Long-Lived Folding Intermediates Predominate the Targeting-Competent Secretome. *Structure* **26**, 695-707.e5 (2018).

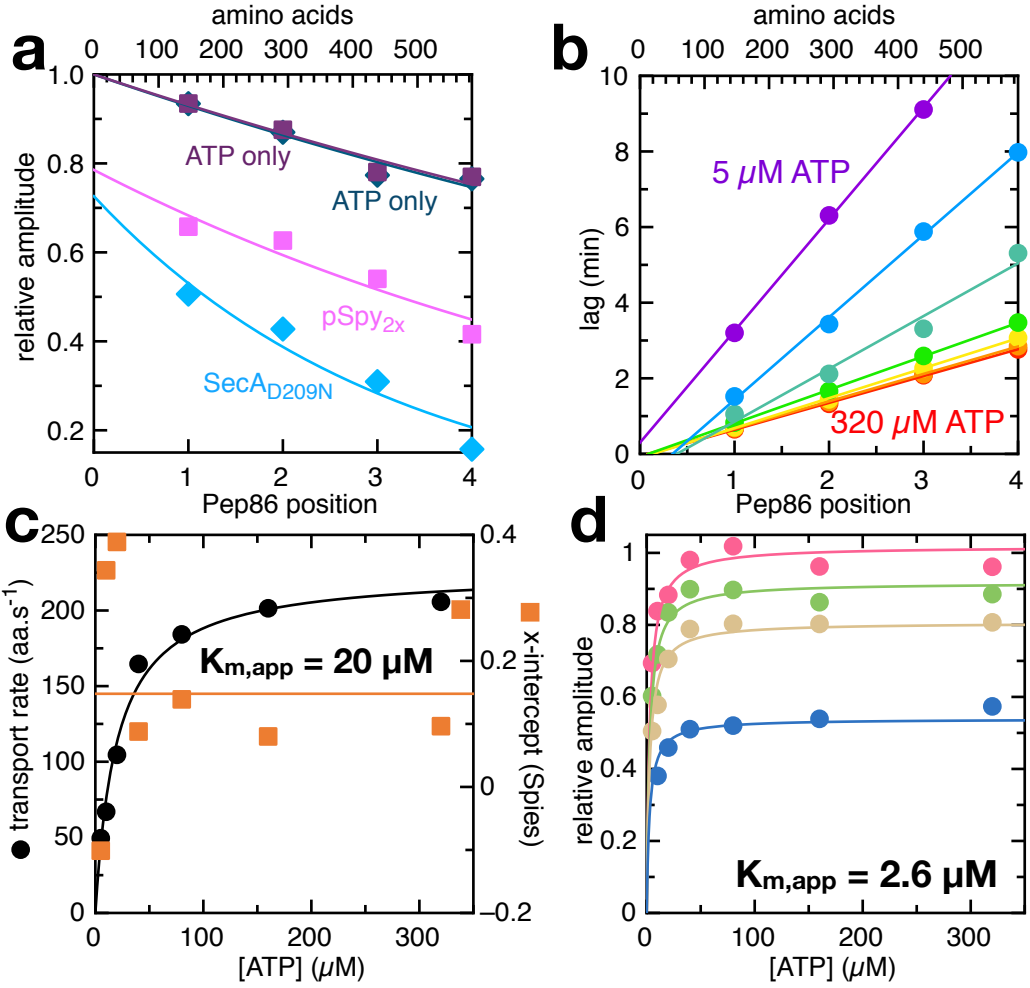
44. D. Huber, *et al.*, SecA Cotranslationally Interacts with Nascent Substrate Proteins In Vivo. *J Bacteriol* **199**, e00622-16 (2017).

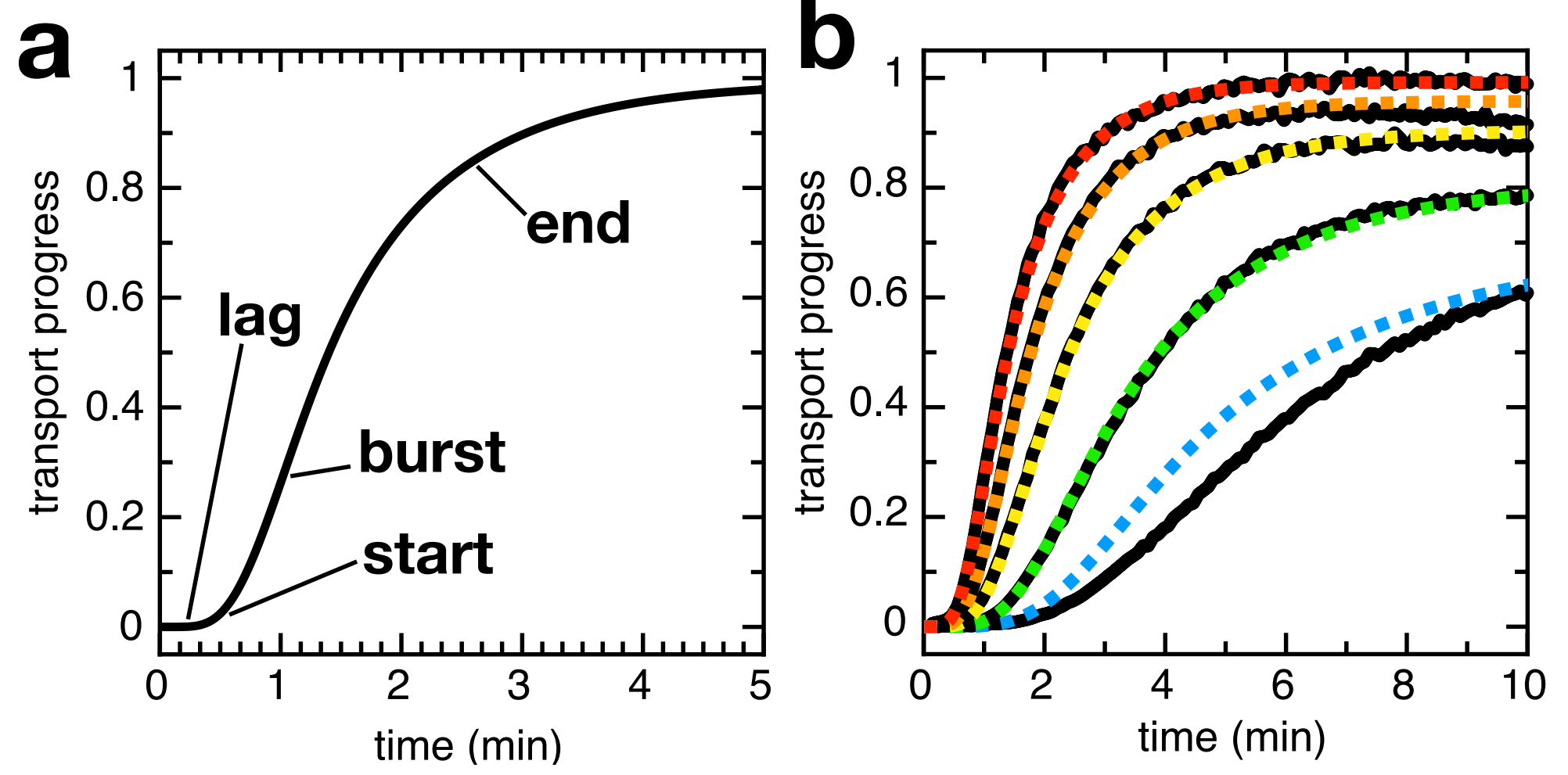
45. R. Singh, *et al.*, Cryo-electron microscopic structure of SecA protein bound to the 70S ribosome. *Journal of Biological Chemistry* **289**, 7190–7199 (2014).

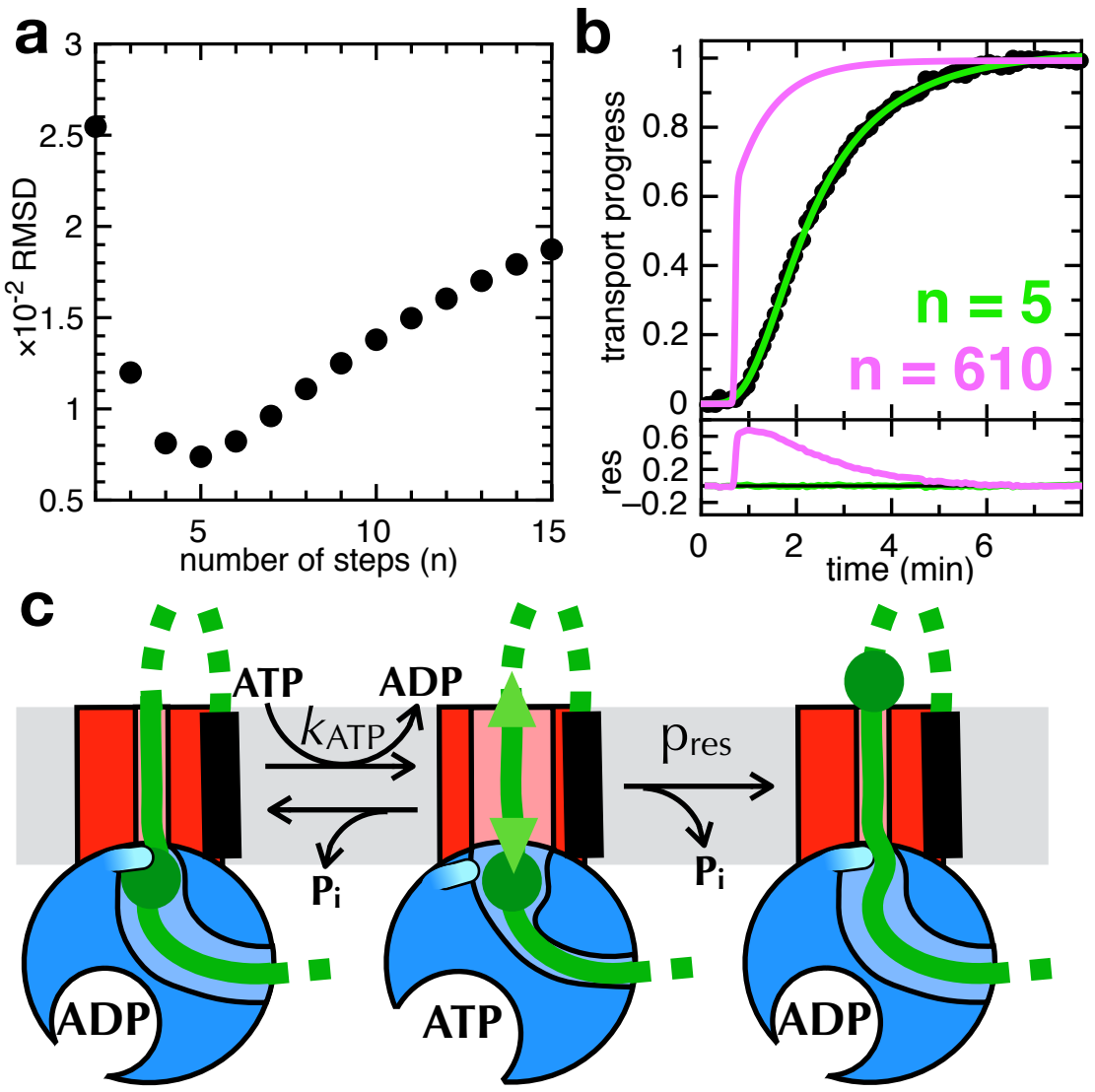
46. J. Chiu, P. E. March, R. Lee, D. Tillett, Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h. *Nucleic Acids Res* **32**, e174–e174 (2004).


a
y=A(1-e-λ(t-lag))
corrected data
λ
A
lag
simple fit
signal (RLU)
time (min)
c
ATP
SecA
pSpy-Pep86
normalised signal
time (min)
d
λmax = 0.93 min-1
KD,app = 0.57 μM
λmin = 0.21 min-1
lag (min), λ (min-1), or relative amplitude
[pSpy-Pep86] (μM)
e
KD = 58 nM
kon = 3.7 μM-1.min-1
relative amplitude
rate (min-1)
[pSpy-Pep86] (μM)
b
pre-protein-Pep86
proteo-liposome
11S
SecYEG
SecA
kon
koff
ATP
ADP
kinit
ATPn
ADPn
kstep
repeat n times
konL
koffL
(i) pre-protein binding
(ii) initiation
(iii) transport
(iv) NanoLuc formation


a
Pep86 active: VSGWRLFKKIS
X X X
pSpy
Pep86 scrambled: VSWGLRKFKIS
SS
mSpy
LDDD
DLDD
DDLD
DDDL
b
signal (normalised)
time (min)
LDDD
DLDD
DDLD
DDDL
ATP = 1 mM
c
amino acids
amplitude (normalised)
lag (min)
Pep86 position
12 aa
0.72 min.Spy-1
= 203 aa.min-1
d
pre-protein
SecYEG
C
SS
~35 aa
N
SecA


a
amino acids
0 100 200 300 400 500
relative amplitude
1.0
0.8
0.6
0.4
0.2
ATP only
ATP only
pSpy2x
SecAD209N
0 1 2 3 4
Pep86 position
b
amino acids
0 100 200 300 400 500
lag (min)
10
8
6
4
2
0
5 μM ATP
320 μM ATP
0 1 2 3 4
Pep86 position
c
transport rate (aa.s-1)
250
200
150
100
50
0
Km,app = 20 μM
x-intercept (Spies)
0.4
0.2
0
–0.2
0 100 200 300
[ATP] (μM)
d
relative amplitude
1
0.8
0.6
0.4
0.2
0
Km,app = 2.6 μM
0 100 200 300
[ATP] (μM)


a
b
transport progress
time (min)
lag
burst
start
end
0
0.2
0.4
0.6
0.8
1
0
1
2
3
4
5
0
2
4
6
8
10


a
3
2.5
2
1.5
1
0.5
×10-2 RMSD
5
10
15
number of steps (n)
b
1
0.8
0.6
0.4
0.2
0
transport progress
n = 5
n = 610
res
0.6
0.2
–0.2
0
2
4
6
time (min)
c
ATP
ADP
kATP
Pi
pres
Pi
ADP
ATP
ADP