2020

Comparative Analysis Of The Translocation Kinetics For Clpa'S Atpase Domains 1 & 2

Nathaniel Scull
University of Alabama at Birmingham

COMPARATIVE ANALYSIS OF THE TRANSLOCATION KINETICS FOR CLPA'S ATPASE DOMAINS 1 & 2

by

NATHANIEL W. SCULL

AARON L. LUICUS, COMMITTEE CHAIR
JAMES C. PATTERSON
PETER E. PREVELIGE
DAVID A. SCHNEIDER
JUN ZHANG

A DISSERTATITION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMGINHAM, ALABAMA

2020

COMPARATIVE ANALYSIS OF THE TRANSLOCATION KINETICS FOR CLPA'S ATPASE DOMAINS 1 & 2

NATHANIEL W. SCULL

CHEMISTRY

ABSTRACT

ClpA, a Class I AAA+ (ATPases Associated with various cellular Activities) motor, assembles into a tiered hexamer containing two AAA+ nucleotide binding domains (NDBs), D1 and D2. These NBDs couple nucleotide triphosphate binding and hydrolysis to mechanical work that drives polypeptide unfolding and translocation. Here we study the kinetic mechanisms of polypeptide translocation at the two domains to elucidate their contributions to the overall wildtype activity of ClpA. We subjected Walker B variants of ClpA that lack ATPase activity in one of the two domains to single-turnover stopped-flow techniques using a new fluorescence anisotropy and total fluorescence method. We found that the two domains translocate polypeptide substrate with unique rates and similar kinetic step-sizes at saturating ATP. D2 was found to support the majority of ClpA's translocation activity, with D1 constituting less than 1% of WT activity. Both D1 and D2 were found to traverse approximately 15 aa per repeating cycle of translocation, similar to WT. Moreover, we showed that both variants exhibited positive cooperativity with respect to ATP binding indicating the presence of inter-monomer interactions within hexameric ClpA.

Keywords: ClpA, AAA+ motor proteins, stopped-flow, transient-state kinetics, Walker B, fluorescence anisotropy

DEDICATION

To Susan, without whom this would never have been possible. You have supported me and made so many sacrifices to get me here. I can never thank you enough.

"Happy is the man who finds a true friend, and far happier is he who finds that true friend in his wife." -Franz Schubert

To Barrett, who kept my spirits afloat and helped me make it through until the end.

"Of all the titles I've been privileged to have, 'Dad' has always been the best."

-Ken Norton

To my family, who support me in everything I do.

"Families are the compass that guides us. They are the inspiration to reach great heights and our comfort when we occasionally falter." –Brad Henry

Acknowledgements

I would first and foremost like to express my thanks to my mentor Dr. Aaron Lucius. You have both pushed and challenged me, while always offering support and guidance. Thank you for helping to shape me into a professional thinker and scientist who seeks to ask and answer questions. The environment you have built in your lab has allowed all of us to strive to obtain what we really want out of our Ph.D. careers. Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Schneider, Dr. Prevelige, Dr. Zhang and Dr. Patterson, for their insightful comments and encouragement. To my fellow graduates, I would like to express my gratitude in all the support you have provided. It has truly been a pleasure to go through both the failures and successes of graduate school together, I have learned a lot from all of you. Zach, thank you for being there every day as we battled our way through this thing, and for teaching me so much. Clarissa and Elizabeth, you pushed and taught me so much, thank you for all the support and laughs. Kaila and Ethan, thanks for being someone I could always go to for discussion or just to talk. Finally, I would like to thank my family for their unwavering support in all aspects of my life; they made this possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAA+	ATPases Associated with various cellular Activities
ADP	Adenosine 5'- diphosphate
ATP	Adenosine 5'- triphosphate
ATPγS	Adenosine 5'- [gamma-thio]-triphosphate
Clp	Caseinolytic peptidase
Cy3	Cyanine 3 Maleimide
Cy5	Cyanine 5 Maleimide
D1	Nucleotide binding domain 1
D2	Nucleotide binding domain 2
EDTA	Ethylenediaminetetraacetic acid
EM	Electron Microscopy
Flu	Fluorescein
FRET	Förster resonance energy transfer
GA	Genetic Algorithm
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
Hsp	Heat shock protein
MENOTR	Multi-start Evolutionary Nonlinear OpTimizeR
2-ME	2-mercaptoethanol (β-mercaptoethanol)
NBD	Nucleotide binding domain
NLLS	Nonlinear least squares
Pi	Inorganic phosphate
PMT	Photomultiplier tube

Introduction

Caseinolytic peptidase A, ClpA, is a AAA+ motor found in *Escherichia coli* (*E. coli*) that couples repeated rounds of ATP binding and hydrolysis to the unfolding and translocation of polypeptide substrates (1). ATPases Associated with various cellular Activities (AAA+) play varied and crucial roles in DNA replication (2) and recombination (3), cellular proteostasis (4-6), membrane fusion (7, 8), cell-cycle control (9, 10), signal transduction (11) and many other pathways (12-14). ClpA specifically has been shown to be implicated in both protein remodeling (15, 16) and proteolysis reactions (17) as part of cellular proteostasis.

ClpA Structure and Function

AAA+ motors are subcategorized based on the number of nucleotide binding domains (NBDs) they contain, as either Class I or II (13). Examples of Class I AAA+ motors include ClpA (18), ClpB (19), p97 (20), and NSF (8), which contain two NBDs per monomer. While Class II motors contain a single NBD per monomer and include the motors ClpX (21), HslU (22), Lon (23), and FtsH (24).

As a Class I AAA+ chaperone, ClpA monomers consist of an N-terminal domain, shown in purple in **Fig.1 D**, followed by NBD 1 and 2 (D1 and D2), shown in **Fig. 1 D** as blue and green, respectively. In the presence of ATP, or ATP analogs, ClpA oligomerizes into a two-tiered homo-hexameric ring with a hollow central channel as seen in **Fig. 1 A -**

C (25-28). Each tier consists of six identical domains of either D1 or D2 that both contain canonical Walker A and B motifs. These motifs are ubiquitous across ATP-dependent enzymes responsible for ATP binding and hydrolysis (29, 30). Cryo-EM structures of hexameric ClpA and a crystal structure of ClpA monomer are shown in **Fig. 1 A-C & D**, respectively.

Binding and hydrolysis of ATP at each ClpA NBD induces conformational changes in pore-loops located between the Walker A & B motifs in each domain (31, 32). Through structural and cross-linking studies it has been shown that these pore-loops extend into the central channel of hexameric ClpA, and contact the polypeptide substrate (18, 33, 34). Evidence from the investigations of ClpA and other AAA+ motors suggests that these loops alternate between up and down conformations during rounds of ATP binding and hydrolysis. Thus, exerting tugging and pulling forces to the substrate that result in unfolding of the substrate and translocation through the central channel (28, 33, 35-37).

We have previously proposed a mechanistic model of ClpA catalyzed translocation (38). In that investigation, we concluded that translocation is being catalyzed by both NBDs, and that each domain has a unique mechanistic role in overall translocation. Our hypothesis for protease-independent ClpA catalyzed polypeptide translocation is as follows: The pore-loops of both NBDs start prebound to the polypeptide substrate, with ATP bound to each domain. D1 starts by hydrolyzing ATP and translocating the substrate into the central cavity of ClpA. The hydrolysis of ATP to ADP and P_i at D1 reduces its

affinity for the substrate (32). This allows D1 to release the peptide, reset, and subsequently rebinds both the substrate at a new position and a new ATP. D2 would then hydrolyze ATP and translocate the substrate out of the central cavity in a similar manner.

From our previous work, we have hypothesized that D1 translocates ~14 aa per translocation step while D2 translocates ~ 5 aa per repeated step (38, 39). The ratio of these step-sizes would cause the substrate to initially crimp within the ClpA cavity as D1 translocates. Then, D2 would need to go through multiple rounds of translocation to remove the substrate. The proposed mechanism is the result of the combined analysis of ClpA and ClpAP catalyzed translocation, substrate crosslinking, synchrotron footprinting (40), and structural studies (38). It, however, does not represent a direct quantitative measurement of the kinetic mechanisms of translocation at each domain.

Specific Aim of Dissertation

This dissertation sets out to test our previous hypotheses by directly interrogating how translocation is catalyzed by each NBD of ClpA. Mutations in the Walker motifs of other AAA+ motors have been successfully used to investigate activities of a single NBD at a time (41-43). In all AAA+ NBDs the Walker B motif contacts ATP creating an enzymatic pocket that coordinates Mg^{2+} and activates water for ATP hydrolysis. It has been shown that the mutation of the conserved Walker B glutamate, which activates the water in hydrolysis, eliminates ATP hydrolysis but not binding (14, 44, 45). These mutations have been successfully used in ClpA (44, 46) as well as other analogous AAA+ proteins including ClpX (42), ClpB (43), and Hsp104 (41). Moreover, we have previously characterized the effects these mutations have on ClpA oligomerization (47, 48).

Kress et al. have previously designed and overexpressed the Walker B variants $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$ that lack hydrolysis activity in D1, D2, or both domains, respectively (44). They exploited these variants to investigate the rates of ATP hydrolysis at each domain. In summary, they found that D1 exhibited lower steady-state ATPase activity relative to D2. D2 exhibited ATPase activity that was ~85% of that of WT, while D1 constituted only ~10% of the activity of WT. Additionally, Baytshtok et al. used similar ClpA variants, $ClpA_{E286Q}$ and $ClpA_{E565Q}$, to investigate unfolding and degradation of model substrates. They found that when activity was abolished at D1, the D2 active variant produced unfolding rates eightfold slower than that of WT. Strikingly, when D2 activity is abolished, D1 active variants lost all activity (46).

While there have been attempts to measure the activity of each domain, a method has yet to be applied that can be used to directly describe their translocation kinetics. Here, in **Chapter 2** we will present a new method of directly monitoring translocation catalyzed by AAA+ motors using fluorescence anisotropy and total fluorescence simultaneously. The strength of this method is its ability to monitor the residence time of motors on a substrate lattice that reflects only a single round of catalysis. Thus, we are sensitive to only the molecular events that occur in the active site of the motor during translocation. Additionally, in **Chapter 1** we outline the development of a novel approach to quantitatively examine the resultant fluorescence anisotropy and total fluorescence time courses that can overcome many of the obstacles faced in the rigorous analysis of biochemical reactions.

Fluorescence Anisotropy and Total Fluorescence Technique for Examination of Translocation Kinetics

We have previously developed a novel single-turnover fluorescence stopped-flow technique that allowed for the examination of $ClpA_{WT}$ in the absence of ClpP (39). This method relies on the fact that fluorescently modified polypeptide substrates experience a quenching of fluorescence when bound to ClpA. Thus, upon translocation, there is a transition of bound to unbound substrate resulting in an observed increase in fluorescence as a function of time. Modeling of this data allowed for the development of a molecular mechanism for ClpA catalyzed translocation.

However, this method has inherent limitations. The changes observed in fluorescence are the consequence of differences in the quantum yield of the fluorophore on each species, rotational artifacts on the time scale of fluorescence (49), or a combination of the two. Thus, time dependent changes in the rotational dynamics of the fluorescent substrate have the potential to create fluorescence artifacts that convolute the observed time courses. Moreover, this method assumes an all or none scenario, where the substrate is only either in a bound or unbound fluorescent state. If translocation intermediates have unique quantum yields, the analysis of these data sets will fail to accurately represent the molecular mechanism of translocation.

In **Chapter 2,** we present the development of a single-turnover fluorescence anisotropy and total fluorescence stopped-flow technique that can overcome these obstacles. These methods can account for unique intermediate species, as well as distinguish the contribution of changes in quantum yield and rotational mobility to the overall signal. In this setup, translocation is monitored concurrently in two ways, as total fluorescence and

fluorescence anisotropy. Total fluorescence is collected such that rotational artifacts are eliminated from the signal (49, 50). Therefore, total fluorescence is defined by changes only in quantum yield. Fluorescence anisotropy reports on the rotational mobility of the complex, which in turn depends on the size of complex. Thus, anisotropy is sensitive to the presence of the motor on its substrate (50). With these techniques, we have been successful in reproducing kinetic parameters from our original analysis of $ClpA_{WT}$ translocation of polypeptides. Moreover, we have been able to gain quantitative interpretations of the ClpA catalyzed translocation of polypeptide substrates longer than 100 aa and estimate the kinetic parameters that describe translocation catalyzed by each NBD of ClpA.

Analysis of Biochemical Data Using Hybrid Algorithms

The analysis of biochemical data requires the use of parameter optimization techniques to determine quantities such as kinetic rates, equilibrium binding and dissociation constants, rate constants, step-sizes, catalytic efficiencies and turn overs, and many other sought-after parameters. Nonlinear least squares (NLLS) methods of parameter optimization are one of the most common types of analysis strategies applied to biochemical problems (51, 52). These NLLS algorithms are well described and are used across all fields of study (53-56).

In summary, NLLS algorithms attempt to minimize the weighted sum of the squares, or the error, between the data and the model equation; here the model equation is the mathematical representation of the chemical system under investigation. All NLLS algorithms require an initial guess, provided by the user, that is then refined until each iteration no longer returns an improved solution to the problem. This approach is deterministic, meaning that given the same initial guess and stop criteria it will always produce the

same results (57). In this manner, NLLS methods quickly converge on a solution that has an error minimum often near the initial guess.

When optimizing highly correlated parameters it often becomes the case that multiple solutions can be found that yield good results for the problem (55, 58-62). This is a consequence of parameters compensating for one another in the optimization, creating multiple local minima in the error space being evaluated. The final parameter values determined in NLLS methods are assumed to have the maximum likelihood of being the best and lowest error. However, this assumption does not always hold in the presence of multiple local minima, as the optimization routine can become trapped. How can these problems be overcome to increase the likelihood of arriving at the true global minimum?

Metaheuristics have been shown to yield good results in solving high-level optimization problems across a variety of fields. Metaheuristics are "upper level methodologies" (meta) that work "to discover" (heuristic) solutions to a variety of fitting problems across multiple disciplines. The underlying characteristic that lends itself to solving problems like the ones imposed by NLLS is the stochastic nature of these metaheuristic algorithms.

In **Chapter 1** we outline the development and use of a novel hybrid genetic and nonlinear least squares algorithm, MENOTR, for optimizing kinetic parameters that describe translocation of a lattice substrate by RecBCD and ClpA. We compare the determined parameters and goodness of fit to previously published results and found that our algorithm can reproduce previous results, but moreover, determine parameter values that are statistically superior to those previously published. Additionally, we outline the general advantages that this method has in analyzing data over conventional nonlinear least

squares methods, and how it can take advantage of modern computing techniques and technologies.

In totality, this dissertation sets out to demonstrate the use of a single-turnover fluorescence anisotropy and total fluorescence technique paired with a novel parameter optimization algorithm to obtain mechanistic details on the translocation of polypeptide catalyzed by each NBD of ClpA. We were able to investigate and determine quantitative measurements of the rate constants, kinetic step-sizes and overall rates of translocation governing both domains.

References

1. Katayama, Y., S. Gottesman, J. Pumphrey, S. Rudikoff, W. P. Clark, and M. R. Maurizi. 1988. The two-component, ATP-dependent Clp protease of Escherichia coli. Purification, cloning, and mutational analysis of the ATP-binding component. J Biol Chem 263(29):15226-15236.

2. Kelman, L. M., and Z. Kelman. 2014. Archaeal DNA Replication. Annual review of genetics 48(1):71-97.

3. Yamada, K., M. Ariyoshi, and K. Morikawa. 2004. Three-dimensional structural views of branch migration and resolution in DNA homologous recombination. Current opinion in structural biology 14(2):130-137.

4. Sauer, R. T., and T. A. Baker. 2011. AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. Annu Rev Biochem 80:587-612.

5. Desantis, M. E., and J. Shorter. 2012. The elusive middle domain of Hsp104 and ClpB: location and function. Biochim Biophys Acta 1823(1):29-39.

6. Sousa, R. 2014. Structural mechanisms of chaperone mediated protein disaggregation. Front Mol Biosci 1:12.

7. Block, M. R., B. S. Glick, C. A. Wilcox, F. T. Wieland, and J. E. Rothman. 1988. Purification of an N-ethylmaleimide-sensitive protein catalyzing vesicular transport. Proc Natl Acad Sci U S A 85(21):7852-7856.

8. Fleming, K. G., T. M. Hohl, R. C. Yu, S. A. Muller, B. Wolpensinger, A. Engel, H. Engelhardt, A. T. Brunger, T. H. Sollner, and P. I. Hanson. 1998. A revised model for the oligomeric state of the N-ethylmaleimide-sensitive fusion protein, NSF. J Biol Chem 273(25):15675-15681.

9. Ingmer, H., C. Miller, and S. N. Cohen. 2001. The RepA protein of plasmid pSC101 controls Escherichia coli cell division through the SOS response. Molecular microbiology 42(2):519-526.

10. Wickner, S. H. 1990. Three Escherichia coli heat shock proteins are required for P1 plasmid DNA replication: formation of an active complex between E. coli DnaJ protein and the P1 initiator protein. Proc Natl Acad Sci U S A 87(7):2690-2694.

11. Confalonieri, F., and M. Duguet. 1995. A 200-amino acid ATPase module in search of a basic function. Bioessays 17(7):639-650.

12. Miller, J. M., and E. J. Enemark. 2016. Fundamental Characteristics of AAA+ Protein Family Structure and Function. Archaea (Vancouver, B.C.) 2016:9294307.

13. Neuwald, A. F., L. Aravind, J. L. Spouge, and E. V. Koonin. 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome research 9(1):27-43.

14. Hanson, P. I., and S. W. Whiteheart. 2005. AAA+ proteins: have engine, will work. Nat Rev Mol Cell Biol 6(7):519-529. Review.

15. Pak, M., and S. Wickner. 1997. Mechanism of protein remodeling by ClpA chaperone. Proc Natl Acad Sci U S A 94(10):4901-4906.

16. Wickner, S., S. Gottesman, D. Skowyra, J. Hoskins, K. McKenney, and M. R. Maurizi. 1994. A molecular chaperone, ClpA, functions like DnaK and DnaJ. Proc Natl Acad Sci U S A 91(25):12218-12222.

17. Maurizi, M. R. 1992. Proteases and protein degradation in Escherichia coli. Experientia 48(2):178-201.

18. Guo, F., M. R. Maurizi, L. Esser, and D. Xia. 2002. Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease. J Biol Chem 277(48):46743-46752.

19. Lee, S., M. E. Sowa, Y. H. Watanabe, P. B. Sigler, W. Chiu, M. Yoshida, and F. T. Tsai. 2003. The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. Cell 115(2):229-240.

20. Banerjee, S., A. Bartesaghi, A. Merk, P. Rao, S. L. Bulfer, Y. Yan, N. Green, B. Mroczkowski, R. J. Neitz, P. Wipf, V. Falconieri, R. J. Deshaies, J. L. Milne, D. Huryn, M. Arkin, and S. Subramaniam. 2016. 2.3 A resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. Science 351(6275):871-875.

21. Glynn, S. E., A. Martin, A. R. Nager, T. A. Baker, and R. T. Sauer. 2009. Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. Cell 139(4):744-756.

22. Sousa, M. C., C. B. Trame, H. Tsuruta, S. M. Wilbanks, V. S. Reddy, and D. B. McKay. 2000. Crystal and solution structures of an HslUV protease-chaperone complex. Cell 103(4):633-643.

23. Lin, C.-C., S.-C. Su, M.-Y. Su, P.-H. Liang, C.-C. Feng, S.-H. Wu, and C.-I. Chang. 2016. Structural Insights into the Allosteric Operation of the Lon AAA+ Protease. Structure 24(5):667-675.

24. Langklotz, S., U. Baumann, and F. Narberhaus. 2012. Structure and function of the bacterial AAA protease FtsH. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1823(1):40-48.

25. Maurizi, M. R., S. K. Singh, M. W. Thompson, M. Kessel, and A. Ginsburg. 1998. Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and clpP. Biochemistry 37(21):7778-7786.

26. Grimaud, R., M. Kessel, F. Beuron, A. C. Steven, and M. R. Maurizi. 1998. Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP. J Biol Chem 273(20):12476-12481.

27. Kessel, M., M. R. Maurizi, B. Kim, E. Kocsis, B. L. Trus, S. K. Singh, and A. C. Steven. 1995. Homology in structural organization between E. coli ClpAP protease and the eukaryotic 26 S proteasome. J Mol Biol 250(5):587-594.

28. Lopez, K. E., A. N. Rizo, E. Tse, J. Lin, N. W. Scull, A. C. Thwin, A. L. Lucius, J. Shorter, and D. R. Southworth. 2019. Conformational Plasticity of the ClpAP AAA+ Protease Couples Protein Unfolding and Proteolysis. bioRxiv:820209.

29. Gottesman, S., C. Squires, E. Pichersky, M. Carrington, M. Hobbs, J. S. Mattick, B. Dalrymple, H. Kuramitsu, T. Shiroza, T. Foster, and et al. 1990. Conservation of the regulatory subunit for the Clp ATP-dependent protease in prokaryotes and eukaryotes. Proc Natl Acad Sci U S A 87(9):3513-3517.

30. Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. Embo J 1(8):945-951.

31. Gates, S. N., and A. Martin. 2020. Stairway to translocation: AAA+ motor structures reveal the mechanisms of ATP-dependent substrate translocation. Protein Sci 29(2):407-419.

32. Farbman, M. E., A. Gershenson, and S. Licht. 2007. Single-Molecule Analysis of Nucleotide-Dependent Substrate Binding by the Protein Unfoldase ClpA. J Am Chem Soc.

33. Hinnerwisch, J., W. A. Fenton, K. J. Furtak, G. W. Farr, and A. L. Horwich. 2005. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. Cell 121(7):1029-1041.

34. Farbman, M. E., A. Gershenson, and S. Licht. 2008. Role of a conserved pore residue in the formation of a prehydrolytic high substrate affinity state in the AAA+ chaperone ClpA. Biochemistry 47(51):13497-13505.

35. Martin, A., T. A. Baker, and R. T. Sauer. 2008. Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. Nat Struct Mol Biol 15(11):1147-1151.

36. Weibezahn, J., P. Tessarz, C. Schlieker, R. Zahn, Z. Maglica, S. Lee, H. Zentgraf, E. U. Weber-Ban, D. A. Dougan, F. T. Tsai, A. Mogk, and B. Bukau. 2004. Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. Cell 119(5):653-665.

37. Biter, A. B., S. Lee, N. Sung, and F. T. Tsai. 2012. Structural basis for intersubunit signaling in a protein disaggregating machine. Proc Natl Acad Sci U S A 109(31):12515-12520.

38. Miller, J. M., J. Lin, T. Li, and A. L. Lucius. 2013. E. coli ClpA Catalyzed Polypeptide Translocation is Allosterically Controlled by the Protease ClpP. Journal of Molecular Biology 425(15):2795-2812.

39. Rajendar, B., and A. L. Lucius. 2010. Molecular mechanism of polypeptide translocation catalyzed by the Escherichia coli ClpA protein translocase. J Mol Biol 399(5):665-679.

40. Bohon, J., L. D. Jennings, C. M. Phillips, S. Licht, and M. R. Chance. 2008. Synchrotron protein footprinting supports substrate translocation by ClpA via ATP-induced movements of the D2 loop. Structure 16(8):1157-1165.

41. Schaupp, A., M. Marcinowski, V. Grimminger, B. Bosl, and S. Walter. 2007. Processing of proteins by the molecular chaperone Hsp104. J Mol Biol 370(4):674-686.

42. Hersch, G. L., R. E. Burton, D. N. Bolon, T. A. Baker, and R. T. Sauer. 2005. Asymmetric interactions of ATP with the AAA+ ClpX6 unfoldase: allosteric control of a protein machine. Cell 121(7):1017-1027.

43. Werbeck, N. D., S. Schlee, and J. Reinstein. 2008. Coupling and dynamics of subunits in the hexameric AAA+ chaperone ClpB. J Mol Biol 378(1):178-190.

44. Kress, W., H. Mutschler, and E. Weber-Ban. 2009. Both ATPase domains of ClpA are critical for processing of stable protein structures. J Biol Chem 284(45):31441-31452.

45. Singh, S. K., and M. R. Maurizi. 1994. Mutational analysis demonstrates different functional roles for the two ATP-binding sites in ClpAP protease from Escherichia coli. J Biol Chem 269(47):29537-29545.

46. Baytshtok, V., T. A. Baker, and R. T. Sauer. 2015. Assaying the kinetics of protein denaturation catalyzed by AAA+ unfolding machines and proteases. Proc Natl Acad Sci U S A 112(17):5377-5382.

47. Duran, E. C., and A. L. Lucius. 2018. ATP hydrolysis inactivating Walker B mutation perturbs E. coli ClpA self-assembly energetics in the absence of nucleotide. Biophys Chem 242:6-14.

48. Duran, E. C., and A. L. Lucius. 2019. Examination of the nucleotide linked assembly mechanism of E. coli ClpA. Protein Sci.

49. Otto, M. R., M. P. Lillo, and J. M. Beechem. 1994. Resolution of multiphasic reactions by the combination of fluorescence total-intensity and anisotropy stopped-flow kinetic experiments. Biophys J 67(6):2511-2521.

50. Lakowicz, J. R. 1999. Principles of fluorescence spectroscopy. Kluwer Academic/Plenum, New York.

51. Johnson, M. L. 1992. Why, when, and how biochemists should use least squares. Analytical biochemistry 206(2):215-225.

52. Johnson, M. L., and L. M. Faunt. 1992. [1] Parameter estimation by least-squares methods. Methods in Enzymology. Academic Press, pp. 1-37.

53. Gavin, H. P. 2019. The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems

54. Bevington, P. R. a. R., D Keith. 2003. Data reduction and error analysis for the physical sciences; 3rd ed. McGraw-Hill, New York, NY.

55. Johnson, M. L., and S. G. Frasier. 1985. [16] Nonlinear least-squares analysis. Methods in Enzymology. Academic Press, pp. 301-342.

56. Caceci, M., and W. Cacheris. 1984. Fitting curves to data: The simplex algorithm is the answer. Byte 9.

57. Black, P. E. 2009. deterministic algorithm. In Dictionary of Algorithms and Data Structures. NIST.

58. Straume, M., and M. L. Johnson. 1992. Monte Carlo method for determining complete confidence probability distributions of estimated model parameters. Methods in enzymology 210:117-129.

59. Williams, D. J., and K. B. Hall. 2000. Monte Carlo applications to thermal and chemical denaturation experiments of nucleic acids and proteins. Methods in enzymology 321:330-352.

60. Xu, C., and G. Z. Gertner. 2008. Uncertainty and sensitivity analysis for models with correlated parameters. Reliability Engineering & System Safety 93(10):1563-1573.

61. Johnson, M. 2000. Parameter correlations while curve fitting. Methods in enzymology 321:424-446.

62. Johnson, M. L. 1983. Evaluation and propagation of confidence intervals in nonlinear, asymmetrical variance spaces. Analysis of ligand-binding data. Biophys J 44(1):101-106.

63. Xia, D., L. Esser, S. K. Singh, F. Guo, and M. R. Maurizi. 2004. Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone. Journal of structural biology 146(1-2):166-179.

64. Goddard, T. D., C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, and T. E. Ferrin. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci 27(1):14-25.

65. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. Journal of molecular graphics 14(1):33-38, 27-38.

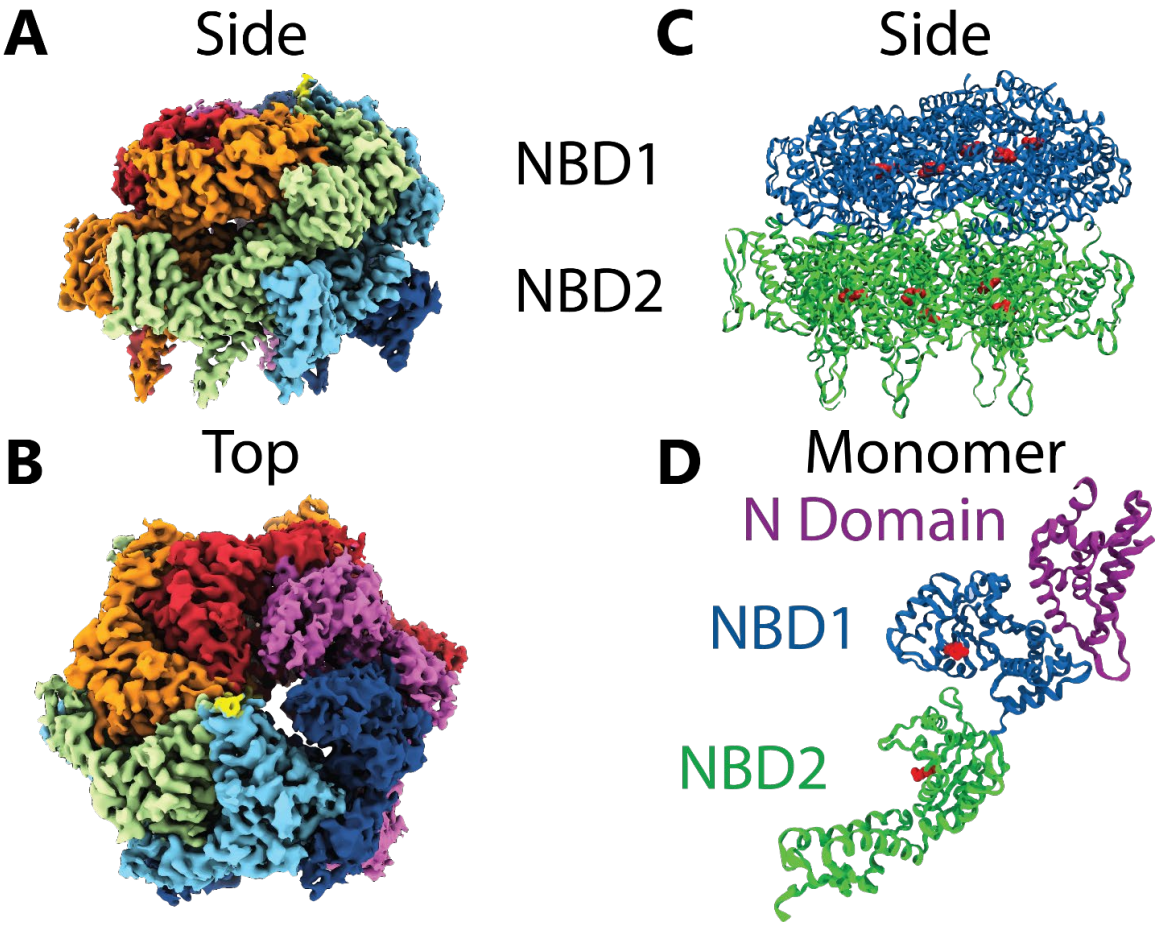


FIGURE 1 ClpA hexameric and monomeric structures. (**A, B, & C**) ClpA structure maps from the determination of a ~3.0 Å resolution cryo-EM structures of *E. coli* ClpAP in the presence of ATPγS and RepA-tagged GFP substrate (28). Surface rendering of the (**A**) side and (**B**) top down views of nucleotide binding domains 1 and 2. Due to the high mobility of the N domain, it could not be resolved, and was not modeled. These images were prepared using UCSF Chimera (Computer Graphics Laboratory, University of California, San Francisco. (**C**) Hexameric ribbon representation with NB1 (blue) and NB2 (green) domains. (**D**) Monomer ribbon representation of crystal structure of ClpA with N (purple), NB1 (blue), and NB2 (green) domains (63). In both crystal structures, the red residues indicate the location of the Walker B motifs in each domain. These are the positions that have mutated in the D1 and D2 Walker B variants, $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$. These structures were prepared using UCSF ChimeraX (64) and Visual Molecular Dynamics 1.9.3 (University of Illinois) (65).

CHAPTER 1

MENOTR: A HYBRID NLLS & GENETIC ALGORITHM FOR BIOCHEMICAL PROBLEMS

NATHANIEL W. SCULL, ZACHARIAH M. INGRAM
AARON L. LUCIUS

Introduction

Nonlinear least squares (NLLS) analysis is one of the most common forms of parameter optimization used to model biochemical data (1). In NLLS, the method attempts to minimize the weighted sum of the squared error between the data and the model function. This is accomplished by iteratively adjusting parameter values though Gauss-Newton (2), Nelder-Mead (3), Levenberg-Marquardt (4), or other similar methods until a best-fit set of values are determined.

The algorithms used to perform NLLS analysis are well described and will not be detailed here (3, 5-8). However, we will point out some of the common characteristics of NLLS algorithms that are relevant to this discussion. All NLLS algorithms require an initial user input as a guess of the parameter values. The algorithm will iteratively improve the initial set of parameters until there is no longer a difference between the preceding set of parameters and the resultant improved parameters to some tolerance. This approach is deterministic, meaning that given the same initial guesses and stopping criteria it will always produce the same results (9). In this manner, NLLS methods quickly and predictively converge on the closest error minimum in relation to the initial guess.

When trying to determine a set of parameters that best describe a set of data, it is often the case that more than one set of parameters can be found that yield good results (10-12). This is often encountered when a model contains a set of highly correlated parameters (13-15). In such cases, the parameters are able to compensate for one another, creating multiple local minima in the error space being evaluated. Meaning that there are

multiple possible solutions to the fitting problem. The complex models used in biochemical analysis tend to have many and varied local minima in their error space. Thus, the optimization method needs to determine which of these local minima is globally, or overall, the lowest, as it corresponds to the set of parameters that best describe the experimental data. The question then becomes, how does the optimization determine which set of parameters is the best, and further, when does the method stop looking for better solutions?

In the context of general NLLS methods, the answer to these questions is straightforward. Once the algorithm has reached a point in which the values of the parameters being optimized do not change within a given tolerance, the algorithm stops. Those final parameter values are assumed to have the maximum likelihood of being the best set and are reported. However, this assumption does not always hold in the presence of the local minima often found in biochemical data as the optimization can become trapped locally (13, 14). So how can we overcome these problems to increase the likelihood of arriving at the true global minimum?

Metaheuristics have been an avenue of active research for about four decades (16) and have yielded good results in solving high-level optimization problems across a variety of fields. Metaheuristics are “upper level methodologies” (meta) that work “to discover” (heuristic) solutions to a problem. Examples include simulated annealing (17), ant colony optimization (18), particle swarm optimization (19), bees algorithm (20), and genetic algorithms (21). As might be obvious, these methods are often nature-inspired and utilize biological ideas like mutation, fitness, gene crossover, and natural selection to overcome optimization problems. The underlying characteristic of these algorithms that

lends itself to overcoming problems like the ones imposed by NLLS are their stochastic nature.

In our strategy, a multi-start genetic (evolutionary) algorithm is used in conjunction with an NLLS algorithm. In general, a genetic algorithm (GA) works to find suitable solutions to optimization problems by pseudo-randomly generating a population of possible solutions over a given range. Each solution is considered an individual and those individuals are then subjected to genetic operators that both refine and diversify the solutions. Common genetic operators include mutation, crossover, and elitism. Compared to NLLS methods, GAs can overcome local minima but lack the ability to converge on a final solution in a reasonable amount of time (22, 23).

Here we report the development of MENOTR, pronounced “minotaur”, a hybrid algorithm that balances the strengths of genetic and NLLS algorithms to offset their corresponding limitations. MENOTR, Multi-start Evolutionary Nonlinear OpTimizeR, was written as a MATLAB toolbox (The MathWorks, Inc., Natick MA) in MATLAB 2018b. It contains a set of scripts and functions that can be edited to match the data and model function being evaluated. Within the files is the base algorithm that will optimize a set of parameters for a given data set and model function, as well as a set of statistical, visualization, and secondary analysis tools.

Two benchmarks were developed to demonstrate MENOTR’s parameter optimization capabilities. The benchmarks are reevaluations of results from previously published manuscripts.

Methods

MENOTR is defined by two general stages, which consist of diversification and refinement. Stage one involves the creation of a set of diverse initial solutions by the multi-start and genetic portions of MENOTR. Stage two consists of solution refinement using both genetic operators and the NLLS algorithm. The best solutions of stage two are passed back to stage one in an iterative manner until stage two is no longer producing improved solutions.

Multi-Start

Multi-start is a diversification strategy that is used to generate an initial set of possible solutions, or initial guesses (24). With every iteration of stage 1 the multi-start algorithm creates multiple unique initial guess based on the previous best solution or a user suppled input. Each guess generated by multi-start is run through MENOTR's optimization routines independently with no information being passed between the optimizations. The optimization routines will be outlined in the genetic and NLLS algorithm sections below.

After the optimizations of all initial guesses are complete, the results are pooled, and a set of user defined stopping criteria is applied to the pooled results. The first stopping criteria determines the coefficient of variation, mathematically equivalent to the standard deviation divided by the mean, for each of the parameters across the optimizations. If the mean of the coefficient of variation for all parameters is below the user designated threshold, then the first stopping criteria is satisfied. This criterion is a measurement of whether the parameters determined in each optimization are similar to the other

optimizations. In other words, are the optimizations converging on the same set of best-fit parameters within the current iteration?

The second criterion compares the chi-squared value generated by the parameters at the beginning of the current iteration to the final chi-squared of the current iteration. Therefore, if the percent difference between the starting and final chi-squared is lower than the user-defined threshold the second stopping criteria is satisfied. Thus, this criterion is a measurement of whether the optimization is still finding better solutions with each iteration of the algorithms, or if the best set of parameters has been reached.

If both criteria are met, then the algorithm concludes, and the results of the current iteration are presented as the best-fit results. If one or both criteria are not met, the best solutions for that iteration are passed forward to the next iteration of the code as a new initial guess. Additionally, a third criterion exists where if a maximum number of iterations have occurred, the optimization is halted, and the current solutions are presented as the best fit. This becomes necessary if the other stopping criteria are too strict and cannot be achieved. In some cases, this becomes an issue and the optimization will become stuck in an infinite loop, circling about the global minimum.

Genetic Algorithm

In stage 2, MENOTR's optimization routines contain two main algorithms, the first of which is a genetic algorithm. This is a metaheuristic technique that generates solutions by utilizing stochastic sampling, genetic operators, and fitness evaluations.

Stochastic sampling is a method by which a population of solutions is created by pseudo-randomly choosing parameter values for each solution in the population. The algorithm will choose parameter values stochastically from a normal Gaussian distribution

about the initial guess provided by the multi-start. This type of sampling has the two-fold advantage of surveying large diverse error spaces for possible solutions, while also being able to rescue the routine from local optima.

Genetic operators act to both diversify and refine candidate solutions. The operators used in this method include mutation, crossover, and elitism. Crossover and mutation are the most common genetic operators and work to diversify the solution population (21). Mutation involves indiscriminately choosing parameters within a candidate solution to be altered. The new value that is generated for each mutated parameter is based on its corresponding initial value prior to mutation. In contrast, crossover creates entirely new candidate solutions by stochastically selecting parameter values from one or multiple current solutions and combining them. These operators are applied to a selection of candidate solutions with both high and low fitness.

In the context of these types of algorithms, the fitness of a candidate solution is based on how well that solution satisfies the problem, goodness of fit. Here we use chi-squared as an indication of fitness; higher chi-squared values indicate lower fitness and vice versa. In this way crossover and mutation work to diversify the population, but also lead to increases in fitness of solutions and overall population.

In contrast to the first two operators, elitism works only to improve the overall fitness of the population and opposes diversification. It does so by preserving a set of the high fitness candidate solutions to be passed forward without perturbation. Thus, maintaining a selection of good candidates at all times avoiding increases in the overall fitness of the population

Due to low convergence rates, GAs are considered computationally expensive despite being mathematically inexpensive and easily parallelizable. In the early stages of optimization convergence is fast, but will slow in the later stages of searching for an optimized solution (22). This is a consequence of the algorithm containing little in the way of direct selective fitness pressures. It is lacking what would be akin to directed evolution or artificial selection mechanisms, thinking about this from the biological metaphor. However, GAs can function as computationally cheap ways of finding improved and diverse initial guesses for a partner deterministic method like NLLS.

NLLS Algorithm

The second algorithm used within the MENOTR optimization follows nonlinear least square methods. The toolbox presented here has been designed such that the user can select between two algorithms: Levenberg-Marquardt or trust-region-reflective. Each handles solving nonlinear least squares problems differently and the user must decide which is more appropriate for their system. We will not go into how each of these algorithms work, as they are well described elsewhere, but we recommend empirical testing of each method for new optimization problems (4, 25, 26). Additionally, MATLAB provides documentation for each method with suggestions on choosing an algorithm: "Least-Squares (Model Fitting) Algorithms" and "Choosing the Algorithm."

While these two algorithms work differently, the deterministic nature of both allows for quick convergence on optima, but neither guarantees that the solution represents a global optimum. However, when partnered with more heuristic methods the reliability of converging on the global optima is increased (22, 23, 27, 28). It is also important to note that unlike the GA, NLLS algorithms cannot be intrinsically parallelized due to each

internal NLLS iteration depending on the previous. To minimize this disadvantage, the toolbox was designed to run multiple NLLS algorithms in parallel when possible.

Results & Discussion

Discussion on the usefulness of MENOTR over conventional fitting methods

The optimization strategies used in MENOTR fall into one of two categories: stochastic or deterministic. Stochastic methods are semi-random optimization strategies that work to diversify and refine the pool of possible solutions. A key feature of stochastic methods is that when provided identical starting points, a stochastic method can arrive at a different set of outputs over separate runs. This leads to a diversification in the sampling pool and aids in overcoming local minima to ensure that the final optimization solutions are global. In the algorithm presented here, the genetic and multi-start portions are both stochastic.

Deterministic methods contrast the randomness of the stochastic methods and will always arrive at the same set of solutions when given the same starting points. A deterministic method works by refining the starting points such that the goodness of fit increases. Deterministic approaches improve the goodness of fit iteratively until some stopping criteria is reached; thus, ensuring convergence to a minimum. However, this method is subject to becoming trapped in local minima. Temporary decreases in goodness of fit are often necessary to overcome local minima and to find the global solution. Deterministic methods do not allow for this. The deterministic algorithms used here are nonlinear least squares algorithms.

MENOTR is a hybrid fitting method that utilizes not only the intrinsic strengths of both stochastic and deterministic techniques but also balances the strengths of each

method to offset their corresponding limitations. The GA is able to survey large and diverse error spaces to find possible solutions. However, due to lack of convergence, the GA becomes computationally expensive and underperforms in the later stages of optimization. In contrast, the NLLS algorithm converges quickly on a local minimum, but cannot guarantee a global minimum due to its inability to escape local minimums or survey sufficient solution space. In this hybrid method, the GA will quickly find good starting points for the slower running NLLS algorithm and simultaneously assist in the escape from local minima. The synergy between the approaches yields an overall more robust and superior method of generating high-quality solutions to fitting problems.

Escape local minima

Local minima are a common obstacle encountered in fitting data with NLLS routines. This is especially apparent when dealing with models that have correlated parameters. As a result, it is often difficult to identify the 'best' global solution to the optimization problem.

One strategy to overcome this obstacle is the manual searching of the error space using NLLS routines with varying starting points. The goal of this strategy is to test as many and as diverse starting solutions as possible. At each starting point, the NLLS routine will be given a different set of parameters to begin refinement. In doing so, the user can manually escape local minima and probe different sections of the error space.

Once the search has been conducted, the results of each optimization can be compared to find the overall best set of solutions. This method is inherently tedious and requires the user to keep a detailed record of the NLLS results at different starting points.

MENOTR was designed in part to automate the user's manual procedure detailed above; allowing for the error space to be more extensively probed.

MENOTR uses the stochastic nature of both the multi-start and GA to automate the selection process for each iteration of the search. A description of how each of these components works is given in the **Methods**, but we will outline the application here. In a single iteration of the fit, MENOTR uses a multi-start procedure to initialize multiple and separate runs of the GA each with unique starting points. Each GA will generate a set of solutions that will be handed off to separate parallel NLLS optimizations. The independent results of each are then compared for goodness of fit to evaluate which is the best solution. In this way, the code follows the same logic detailed above in the manual method. However, the automation of this method allows for the ability to search many more solutions in a more efficient manner. Automation of this process is both more efficient and more robust at tackling the issue of local minima compared to traditional NLLS methods alone.

Minimal user intervention

Data optimization carried out using NLLS algorithms require an initial starting guess of the parameter solutions. There are strengths and weaknesses to this approach. Computational run times can be greatly reduced if the initial starting point is near the global minimum. Often this is achieved by inspecting the data or predicting values based on previous investigations. A downside to this approach is that the initial starting point can result in the algorithm arriving at a local minimum rather than the global minimum. As such, the initial starting point is particularly sensitive to user bias. A central goal of this software package was to minimize user intervention during parameter optimization in

order to both lessen time requirements on the user and eliminate user bias in the fitting process.

As already discussed, in order to ensure that the results of the optimization are a global minimum, strategies must be employed to survey many different possible solutions. The question then becomes how many minimums must be surveyed to conclude that the best solution is indeed the global minimum. The simplest answer is as many as possible. However, what is reasonable or even feasible to do in a finite amount of time? The strategy that we have employed here is to set stopping criteria for the automated fitting process.

The stopping criteria used in MENOTR are a set of tolerances based on the change in chi-squared between iterations, the magnitude of variation in the fit parameters within an iteration, and the total number of iterations already performed. Only when the change in chi-squared and the variance in parameter values has dropped below the user-supplied tolerance will the fitting routine stop. This acts to stop the fit once a minimum is found that the optimization is not escaping to a lower minimum. However, the limitation of this strategy is that if only minimal changes are made in the starting points of each NLLS fit then escaping some local minima would not be possible.

In order to create a set of potential starting points that surveys enough solutions to guarantee the fit is not stuck in a local minimum, large changes in the starting points must be used. To do this manually becomes unrealistic and would result in many person-hours. The way the fitting has been automated in MENOTR allows the user to establish a set of settings and tolerances at the beginning of the fitting process and be able to walk away until the fit is complete.

Another issue that arises in the process of fitting from a pool of varied starting points is how to choose the points. We know that it is necessary to have a large and diverse set of starting points, but what dictates a diverse and large enough difference between points? To do this manually requires the user to choose points based on what they might know about the system. As such, the user inserts some level of bias into the fit, because the points were specifically chosen. We chose to use statistical tools to approach this issue in an automated and semi-random manner that lowers bias. Given one initial guess, the code will create a pool of starting points by randomly selecting points that follow a normal Gaussian distribution about the initial point. Each multi-start iteration of the code will generate its own unique pool of points. The automation of this process relieves the user of having to do this manually and minimizes any type of user-induced bias.

Modern computational techniques

The majority of NLLS algorithms are not able to perform calculations in parallel. Consequently, NLLS algorithms are unable to capitalize on the multiple cores found in modern computers or the hundreds of cores found in high performance computer clusters. MENOTR was designed to run as many steps in parallel as possible, resulting in faster run times.

The stochastic nature of the GA directly benefits from this approach. All the genetic operators, chi-squared calculations, and population manipulations can be set up to run across all available processors simultaneously as they are independent of one another. Additionally, when the set of solutions to be refined using NLLS methods are found they are split across processors to parallelize multiple NLLS optimizations. This allows the algorithm to maximize the amount of processing power used at any given time.

We have found that MENOTR is able to take advantage of the computational power of high-performance computing clusters. We recommend exploring your access to these types of systems like the NSF-funded XSEDE, the NIH's Biowulf, or any of the university run computer clusters. At the University of Alabama at Birmingham we used the Cheaha cluster for most of the work presented here. Most of these systems will allow for the batching of scripts making it possible to load multiple variations of a given optimization problem simultaneously. This is especially useful when trying to determine the optimization settings that work best for a given problem.

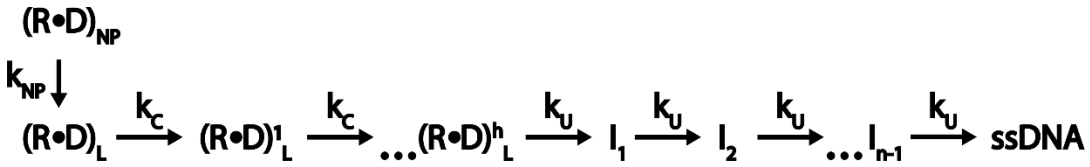
Benchmarks

To demonstrate the abilities of MENOTR, we established a series of benchmarks where optimization routines were performed on previously published data. The goal of the benchmarks was to reproduce published fit results with minimal user intervention. Each of the benchmarks were chosen to demonstrate the capabilities of MENOTR to fit data sets with correlated parameters. The first benchmark involved the application of MENOTR to the fitting of published time courses collected from the RecBCD catalyzed unwinding of duplex DNA. The second benchmark applied MENOTR to published time courses of polypeptide translocation catalyzed by ClpA.

Benchmark I: Duplex DNA unwinding catalyzed by RecBCD helicase

Benchmark 1 involved the application of MENOTR to analyze a set of DNA unwinding time courses that were previously published in 2004 by Lucius et al. (29). These time courses, shown as solid traces in **Fig. 1**, were collected in the investigation of DNA unwinding catalyzed by *E. coli* RecBCD using a FRET based stopped-flow assay. This assay monitors the FRET signal between a Cy3 and Cy5 pair attached on either side of a

nick in duplex DNA. At time zero, the signal from Cy3 is low and the signal for Cy5 is high. Upon DNA unwinding, the two dyes are separated resulting in an increase in the signal for Cy3 and a decrease in the signal for Cy5. The time courses shown in **Fig. 1** reflect the Cy3 signal in the experiment. Weighted global nonlinear least squares analysis of unwinding by RecBCD was performed using **Scheme 1** and **Eq. 1**.



Scheme 1

$$f_{\mathrm{Rec}BCD}(t) = A \cdot \mathcal{L}^{-1}\left(\frac{k_C^h k_U^n (k_{NP} + s \cdot x)}{s(k_C + s)^h (k_{NP} + s)(k_U + s)^n}\right) \quad (1)$$

Scheme 1 depicts the reaction mechanism used to describe RecBCD catalyzed unwinding of duplex DNA into ssDNA. Prior to unwinding, the enzyme-substrate complex exists in two forms: nonproductive complex $(R•D)_{NP}$ and productive complex $(R•D)_L$. Upon mixing with ATP, $(R•D)_{NP}$ slowly isomerizes with the rate constant k_{NP} to form $(R•D)_L$. Since these two states are present before mixing with ATP, the relative populations of the two states is defined by the fraction of productively bound complexes, x, where x is the ratio of productive complex to total complex. There is an additional slow conformational step to $(R•D)^h_L$ described by the rate constant k_C that repeats h number of times. Unwinding begins and some number, n-1, of intermediate steps are taken, I, governed by the rate constant k_U. With the last translocation step, the enzyme-ssDNA complex dissociates. The system of coupled differential equations that result from **Scheme 1**

are solved as detailed in Lucius et al 2004 to arrive at **Eq. 1** that describes formation of ssDNA (29). It should be noted that we report the kinetic step-size, m, in place of number of steps, n. The two parameters are related by the relationship m = n/L, where L is the length of the DNA being unwound.

The resulting published fit parameters are shown in **Table 1.1**. The fit was performed on eight different lengths of duplex substrate using the global fitting strategy where k_U, k_C, k_{NP}, m, and h are constrained as global fitting parameters, meaning they have the same value for all duplex lengths, while A and x are local parameters with unique values for each duplex length. This fit was also subjected to Monte Carlo analysis to generate uncertainties within 68% confidence; these correspond to the error on the fit parameters in **Table 1.1**.

The goal of this benchmark is to determine whether MENOTR can reproduce the previously published kinetic parameters given in **Table 1.1**. Therefore, MENOTR was applied to the same eight RecBCD time courses using **Scheme 1** and **Eq. 1**; as well as the same global fitting strategy outlined in above section. Best-fit simulations of the data are shown in **Fig. 1** as dashed black traces for three of the substrate lengths; 24, 43, and 60 bp. Inspection of the fits shows good overlap of the best-fit simulation and the data. The chi-squared value from the MENOTR fit was found to be ~71% smaller than the previously published fit result. Thus, the MENOTR fit was found to be statistically better at a 68% confidence interval.

The fit parameters determined by MENOTR are provided in **Table 1.1** underneath their corresponding published values. While the fit from MENOTR was determined to be

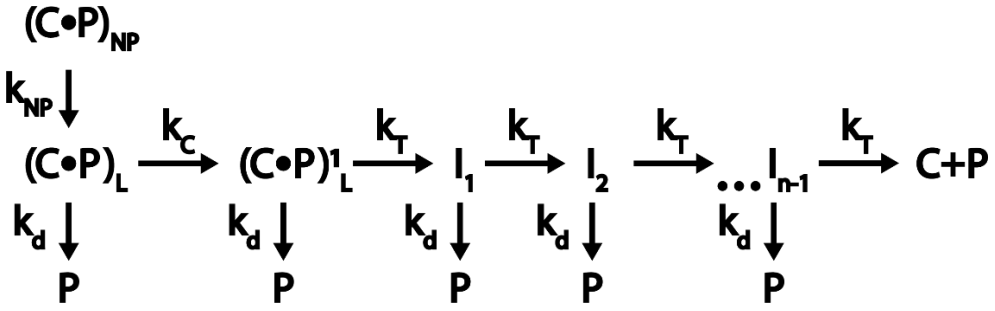
statically better, the resulting kinetic parameters did not vary dramatically from the previous results. The fit was subjected to Monte Carlo analysis to generate uncertainties within 68% confidence, which are also found in **Table 1.1** with their corresponding parameter values. With this, all but one of the kinetic parameters were found to be within error of the previous results. The parameter k_{NP} was found to have a value of $(6.49 \pm 0.08)\ s^{-1}$ from the MENOTR analysis while its published value is $(6.0 \pm 0.3)\ s^{-1}$. While k_{NP} is outside of error, we still conclude there is good agreement between the methods.

In this benchmark, MENOTR was able to reproduce fits comparable to those previously published using only nonlinear least squares fitting strategies. The simulated best fit and the values of the kinetic parameters determined from the MENOTR analysis agreed with both the data and the previously published results.

Benchmark II: Polypeptide translocation catalyzed by ClpA

The second benchmark involves the application of MENOTR to fitting time courses collected from polypeptide translocation catalyzed by *E. coli* ClpA. In 2010, Rajendar et al. developed a fluorescence stopped-flow method of studying ClpA catalyzed translocation of polypeptide substrate. The assay monitors the change in fluorescence signal of fluorecein as translocation occurs. When ClpA is bound to the polypeptide substrate, the fluorescence is quenched. During translocation, ClpA resides on the polypeptide substrate and the fluorescence remains quenched. Upon completion of translocation, the ClpA dissociates, and fluorescence is restored. This assay allows for quantitative measurements of the ClpA residence time on the polypeptide and the corresponding translocation kinetics. Translocation time courses were collected for three polypeptide

lengths shown in **Fig. 2** as solid traces; 30, 40, and 50 aa in length. Weighted global non-linear least squares analysis of polypeptide translocation catalyzed by ClpA was performed using **Scheme 2** and **Eq. 2** to find the set of best-fit parameters shown in **Table 1.2**.



Scheme 2

$$S(s)=\frac{1}{s}\left(\sum_{j=1}^{h}\frac{k_d k_C^{j-1}\left(k_{NP}+sx\right)}{\left(k_C+k_d+s\right)^{j}\left(k_{NP}+s\right)}+\sum_{i=1}^{n-1}\frac{k_d k_T^{h} k_T^{i-1}\left(k_{NP}+sx\right)}{\left(k_C+k_d+s\right)^{h}\left(k_{NP}+s\right)\left(k_d+k_T+s\right)^{i}}+\frac{\left(k_T+k_d\right)k_C^{h}k_T^{n-1}\left(k_{NP}+sx\right)}{\left(k_C+k_d+s\right)^{h}\left(k_{NP}+s\right)\left(k_d+k_T+s\right)^{n}}\right) \quad (2)$$

In **Scheme 2**, before mixing with ATP, ClpA is bound to substrate in two forms: nonproductive ClpA complex $(C•P)_{NP}$ and productive ClpA complex $(C•P)_L$. Upon mixing with ATP, $(C•P)_{NP}$ slowly isomerizes with rate constant k_{NP} to form $(C•P)_L$. As seen previously, x is the ratio of productive complex to total complex. Translocation begins, and some number, n-1, of intermediate steps, I, are taken governed by the rate constant k_T. With the last translocation step, the ClpA peptide complex will dissociate. Throughout the scheme it is possible for peptide, P, to dissociate from ClpA prior to the completion of translocation, these steps are described by the rate constant k_d. Here the slow conformational step to $(C•P)^h_L$ is described by only one k_C step, h = 1. The system of coupled

differential equations that result from **Scheme 2** are solved, as detailed in Rajendar et al. 2009, to arrive at **Eq. 2,** which describes appearance of P (29). Again, we report the kinetic step-size, m, in place of number of steps, n.

The fit was performed using the global fitting strategy of k_T, k_d, k_{NP}, k_C, m and h being constrained as global fitting parameters, while A and x were established as local parameters for each peptide length. It should be noted that it was previously determined that k_d was never detectable in these experiments at this [ATP], and it therefore falls out of the fit. This observation was taken into consideration when fitting with MENOTR. The uncertainties reported in **Table 1.2** were generated using Monte Carlo simulations and correspond to 68% confidence.

MENOTR was initialized with the same model and global strategy as used in the published fit detailed above. Best-fit simulations of the data, generated using **Eq. 2** and the MENOTR fit parameters in **Table 1.2**, are shown in **Fig. 2** as dashed black traces for the substrate lengths. Inspection of the data showed good overlap between the best-fit simulation and the experimental data. MENOTR was able to achieve a 50% reduction in the fit chi-squared, which resulted in a statistically better fit at 68% confidence.

As in benchmark 1, the fit was statically better, but there were not dramatic differences in the kinetic parameter values determined from MENOTR compared to the published results. This fit was also subjected to Monte Carlo analysis to generate uncertainties within 68% confidence, found in **Table 1.2** with their corresponding parameter values. Even though we will later show that this type of error is not appropriate for these parameters, we will use it to keep consistency with the published results. The values determined for k_T and m were both within error of their published values. The other two rate

constants, k_C and k_{NP}, were just outside of error of their published values, $k_{C\text{-}NLLS} = (0.20 \pm 0.003)\ s^{-1}$ compared to $k_{C\text{-}MENOTR} = (0.165 \pm 0.008)\ s^{-1}$ and $k_{NP\text{-}NLLS} = (0.045 \pm 0.0005)\ s^{-1}$ compared to $k_{NP\text{-}MENOTR} = (0.040 \pm 0.003)\ s^{-1}$. After performing grid search analysis on each of the kinetic parameters it was concluded that MENOTR reports $k_{NP} = (0.037 – 0.043)\ s^{-1}$ and $k_C = (0.15 – 0.18)\ s^{-1}$ with a 68% confidence. A description of this type of analysis is provided below. While these parameter values are still just outside of error of one another, they are similar enough to conclude agreement between the methods.

In this benchmark, MENOTR was able to demonstrate its ability to reproduce quality fits comparable to previous methods using strictly nonlinear least squares. The simulated best-fit and the values of the kinetic parameters determined from the MENOTR analysis agreed with both the data and the previously published results.

Monte Carlo

In addition to data fitting, MENOTR, has two built in secondary analysis tools. The first tool is a Monte Carlo simulator. Monte Carlo analysis is able to generate probability distributions for each parameter using the best fit and the experimental error on the data (11, 12).

A set of simulated time courses are created and processed to produce population distributions for each investigated parameter. In this type of analysis, some number of data simulations are generated by adding random Gaussian noise, scaled by the experimental error, to the data. This essentially creates a large number of simulated experimental replicates that can be subsequently fit. The best-fit parameters from the fitting of all the simulated data sets are synthesized into probability distribution histograms.

The difficulty of performing Monte Carlo analysis arises in the large number of fits that must be performed on the simulated time courses. To get good distributions, hundreds, if not thousands, of fits need to be performed. This can become tedious and time consuming as the number of fits is increased to increase rigor. However, due to MENOTR's ability to parallelize its work, we can perform this type of analysis more rigorously than is reasonable with other non-parallelized methods. Each Monte Carlo simulation can be fit in parallel as they are completely independent of one another. MENOTR can generate and fit thousands of simulations in reasonable time frames, allowing for large populations to be created for use in generating the probability distributions.

Monte Carlo simulations have been used to determine the degree and type of parameter correlation in a number of publications (30, 31). Plotting two parameters as a function of one another will yield a correlation scatter plot. Trends in these scatter plots can be used to reveal information about parameter correlation. An example of a correlation scatter plot for k_T and m for ClpA catalyzed translocation from the data found in **Fig. 1** can be found in **Fig. 3 A**. It can be seen that k_T and m are negatively correlated based on the spread of the data. As step-size is increased there is a decrease in the rate constant and vice versa. The degree of correlation is reflected in the steepness of a best fit line through the points, the correlation coefficient. This coefficient can yield information on the constraint of each parameter. For example, in **Fig. 3 A** the best fit line has a slope of -0.088. If both k_T and m had the same degree of correlation the slope would be predicted to be -1, since they are negatively correlated. However, the correlation coefficient is

greater than -1. Thus, we can predict that the kinetic step-size is less constrained compared to the translocation rate constant, and that small changes in the rate constant result in relatively large changes in step-size.

Grid Search

The second tool contained in MENOTR involves grid search analysis coupled with F statistics. This tool is ideal for further investigating parameter constraint. Grid search analysis generates error contours for each parameter as a function of the parameter's value (1, 14, 29, 32). Evaluation of these contours with F statistics allows for the determination of the confidence interval on each parameter. The confidence interval for a parameter corresponds to the statistical upper and lower bounds on the determined best-fit answer. Any two parameter values that fall between the two bounds are statistically equivalent at that confidence level.

This type of analysis is more rigorous than the traditional methods of reporting standard deviation. When reporting standard deviations, the assumption is that the fit error on the parameter follows a normal distribution. However, for many parameters, this is not the case and such reporting is inappropriate. With confidence intervals generated from MENOTR there is no assumption of symmetric error and this type of error can be reported appropriately.

Each contour is generated by plotting the error of the best fit at fixed values of the parameter of interest. Contours of k_T and m for ClpA catalyzed translocation from the data found in **Fig. 1** are shown in **Fig. 3 B & C**, respectively. Values of the parameter being evaluated are chosen to the right and left of the overall best-fit value. The data is then

analyzed to find the set of best-fit parameters while holding the parameter of interest constant at one of the chosen values. This is repeated for all the chosen values for the parameter of interest and the corresponding errors are tabulated. Once the contour has been generated, the errors can be converted into F calculated values by normalizing each by the overall best-fit error. An F critical value can be calculated from the number of data points in the fit, the number of fit parameters and a confidence level. Parameter values with F calculated values above the F critical value are considered statistically different. The confidence interval is found by extrapolating which values of the parameter generate an F calculated that value crosses the F critical value.

Error contours are also helpful in identifying constrained and unconstrained parameters. A constrained parameter exhibits a concave contour with two clear intersections with the F-critical line. In contrast, less constrained or unconstrained parameters deviate from this structure. One example is when there is only one intersection between the error contour and the F-critical line indicating either a lower or upper bound but not both. Inspection of error contours in MENOTR allows an observer to quickly identify if a parameter is constrained and provide insight for subsequent fitting strategies.

In **Fig. 3 B & C** it was found that the confidence intervals of k_T = (1.01 – 2.12) s^{-1} and m = (9.00 – 15.09) aa $step^{-1}$. Compared to the Monte Carlo standard deviation which estimates k_T = (1.3 - 1.7) s^{-1} and m = (13 - 15) aa $step^{-1}$, it can be seen that Monte Carlo produces underestimations in error. Additionally, the shapes of the error contours and the differences between the two types of errors show that the error on both parameters are not symmetrical. A simple standard deviation would unlikely convey the nuances of parameter error for these fits. The translocation rate constant is more constrained to the left,

while the step-size is more constrained to the right. This inverse relationship in constraint is a consequence of the inverse correlation between the parameters themselves, further supporting the conclusions drawn from the Monte Carlo correlation plots.

Conclusions

Here we present a novel MATLAB toolbox for parameters optimization that can overcome common issues that arise in the analysis of biochemical data. The hybrid algorithm was designed to take advantage of the refinement capabilities of classical nonlinear least squares fitting algorithms coupled with the stochastic nature of a genetic algorithm, which can overcome local minima and provide sufficient diversification of the solution pool. Thus, increasing the likelihood of the final solution being a true global minimum and moreover being the best set of parameters. Built into the toolbox are additional statistical tools for secondary data analysis and visualization methods.

MENOTR was able to determine a set of kinetic parameters for two biochemical problems that agree with previously published results and improved the overall fits of that data. Both data sets were previously analyzed using the strategy of manually probing possible initial parameter guesses. As such, each were manually optimized multiple times until the researcher was satisfied that a global minimum was found. Here, due to the automation capabilities of MENOTR, these optimizations were performed with minimal user intervention thus freeing the researcher to move their efforts elsewhere. Additionally, MENOTR's ability to be parallelized and automated allowed for efficient usage of high-performance computing; bringing fit times down dramatically.

MENOTR will be made available online with an accompanying user guide. The guide will outline how to set up optimizations, change the model being fit to, edit and personalize the optimization settings, and give general suggestions.

References

1. Johnson, M. L. 1992. Why, when, and how biochemists should use least squares. Analytical biochemistry 206(2):215-225.
2. Schaback, R. 1985. Convergence analysis of the general Gauss-Newton algorithm. Numerische Mathematik 46(2):281-309.
3. Nelder, J. A., and R. Mead. 1965. A Simplex Method for Function Minimization. The Computer Journal 7(4):308-313.
4. Gavin, H. P. 2019. The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems
5. Caceci, M., and W. Cacheris. 1984. Fitting curves to data: The simplex algorithm is the answer. Byte 9.
6. Marquardt, D. W. 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. Journal of the Society for Industrial and Applied Mathematics 11(2):431-441.
7. Bevington, P. R. a. R., D Keith. 2003. Data reduction and error analysis for the physical sciences; 3rd ed. McGraw-Hill, New York, NY.
8. Johnson, M. L., and L. M. Faunt. 1992. [1] Parameter estimation by least-squares methods. Methods in Enzymology. Academic Press, pp. 1-37.
9. Black, P. E. 2009. deterministic algorithm. In Dictionary of Algorithms and Data Structures. NIST.
10. Johnson, M. L., and S. G. Frasier. 1985. [16] Nonlinear least-squares analysis. Methods in Enzymology. Academic Press, pp. 301-342.
11. Straume, M., and M. L. Johnson. 1992. Monte Carlo method for determining complete confidence probability distributions of estimated model parameters. Methods in enzymology 210:117-129.
12. Williams, D. J., and K. B. Hall. 2000. Monte Carlo applications to thermal and chemical denaturation experiments of nucleic acids and proteins. Methods in enzymology 321:330-352.
13. Xu, C., and G. Z. Gertner. 2008. Uncertainty and sensitivity analysis for models with correlated parameters. Reliability Engineering & System Safety 93(10):1563-1573.
14. Johnson, M. 2000. Parameter correlations while curve fitting. Methods in enzymology 321:424-446.

15. Johnson, M. L. 1983. Evaluation and propagation of confidence intervals in nonlinear, asymmetrical variance spaces. Analysis of ligand-binding data. Biophys J 44(1):101-106.

16. Glover, F. 1986. Future paths for integer programming and links to artificial intelligence. Computers & Operations Research 13(5):533-549.

17. Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi. 1983. Optimization by Simulated Annealing. Science 220(4598):671.

18. Dorigo, M., and C. Blum. 2005. Ant colony optimization theory: A survey. Theoretical Computer Science 344(2):243-278.

19. Kennedy, J. 2006. Swarm Intelligence. Handbook of Nature-Inspired and Innovative Computing: Integrating Classical Models with Emerging Technologies. A. Y. Zomaya, editor. Springer US, Boston, MA, pp. 187-219.

20. Yuce, B., M. S. Packianather, E. Mastrocinque, D. T. Pham, and A. Lambiase. 2013. Honey Bees Inspired Optimization Method: The Bees Algorithm. Insects 4(4):646-662.

21. Reeves, C. 2010. Genetic Algorithms. pp. 109-139.

22. Liu, G. R., X. Han, and K. Y. Lam. 2002. A Combined Genetic Algorithm and Nonlinear Least Squares Method for Material Characterization using Elastic Waves. Computer Methods in Applied Mechanics and Engineering 191:1909-1921.

23. Liu, Y., and P. Tian. 2015. A multi-start central force optimization for global optimization. Applied Soft Computing 27:92-98.

24. Martí, R., J. M. Moreno-Vega, and A. Duarte. 2010. Advanced Multi-start Methods. Handbook of Metaheuristics. M. Gendreau and J.-Y. Potvin, editors. Springer US, Boston, MA, pp. 265-281.

25. Byrd, R. H., R. B. Schnabel, and G. A. Shultz. 1987. A Trust Region Algorithm for Nonlinearly Constrained Optimization. SIAM Journal on Numerical Analysis 24(5):1152-1170.

26. Yuan, Y.-x. 1999. A Review of Trust Region Algorithms for Optimization. ICM99: Proceedings of the Fourth International Congress on Industrial and Applied Mathematics.

27. Martí, R., M. G. C. Resende, and C. C. Ribeiro. 2013. Multi-start methods for combinatorial optimization. European Journal of Operational Research 226(1):1-8.

28. Ribas, P. C., L. Yamamoto, H. L. Polli, L. V. R. Arruda, and F. Neves-Jr. 2013. A micro-genetic algorithm for multi-objective scheduling of a real world pipeline network. Engineering Applications of Artificial Intelligence 26(1):302-313.

29. Lucius, A. L., C. Jason Wong, and T. M. Lohman. 2004. Fluorescence stopped-flow studies of single turnover kinetics of E.coli RecBCD helicase-catalyzed DNA unwinding. J Mol Biol 339(4):731-750.

30. Rajendar, B., and A. L. Lucius. 2010. Molecular mechanism of polypeptide translocation catalyzed by the Escherichia coli ClpA protein translocase. J Mol Biol 399(5):665-679.

31. Miller, J. M., and A. L. Lucius. 2014. ATP-gamma-S Competes with ATP for Binding at Domain 1 but not Domain 2 during ClpA Catalyzed Polypeptide Translocation. Biophys Chem 185:58-69.

32. Lucius, A. L., A. Vindigni, R. Gregorian, J. A. Ali, A. F. Taylor, G. R. Smith, and T. M. Lohman. 2002. DNA unwinding step-size of E. coli RecBCD helicase determined from single turnover chemical quenched-flow kinetic studies. J Mol Biol 324(3):409-428.

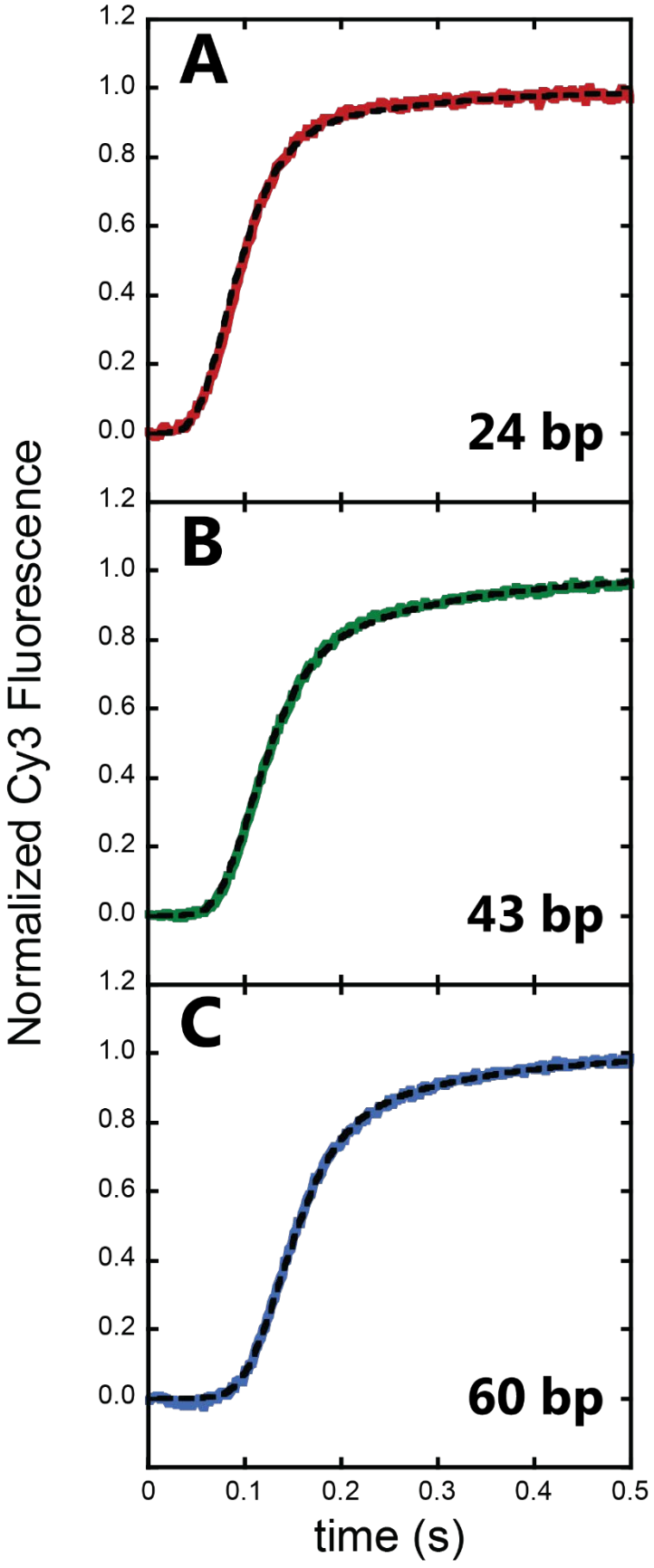


FIGURE 1 MENOTR analysis of the single-turnover RecBCD catalyzed DNA unwinding for a set of duplex lengths. Displayed are three representative time courses for (**A**) 24 bp, (**B**) 46 bp, and (**C**) 60 bp in the solid colored traces. The signals reported here come from monitoring Cy3 fluorescence as a function of reaction time. The time courses were globally analyzed using **Scheme 1** and **Eq. 1** across all eight DNA duplex lengths. Parameters k_U, k_C, k_{NP}, m, and h were set as global parameter for all time courses, while allowing A and x to float locally for each duplex length. The best fit parameters determined from this can be found in **Table 1.1** and the corresponding best-fit lines are show here as broken black traces. A corresponding figure for the original fits and analysis of this data can be found in Lucius et al. 2004, **Figure 8** (29).

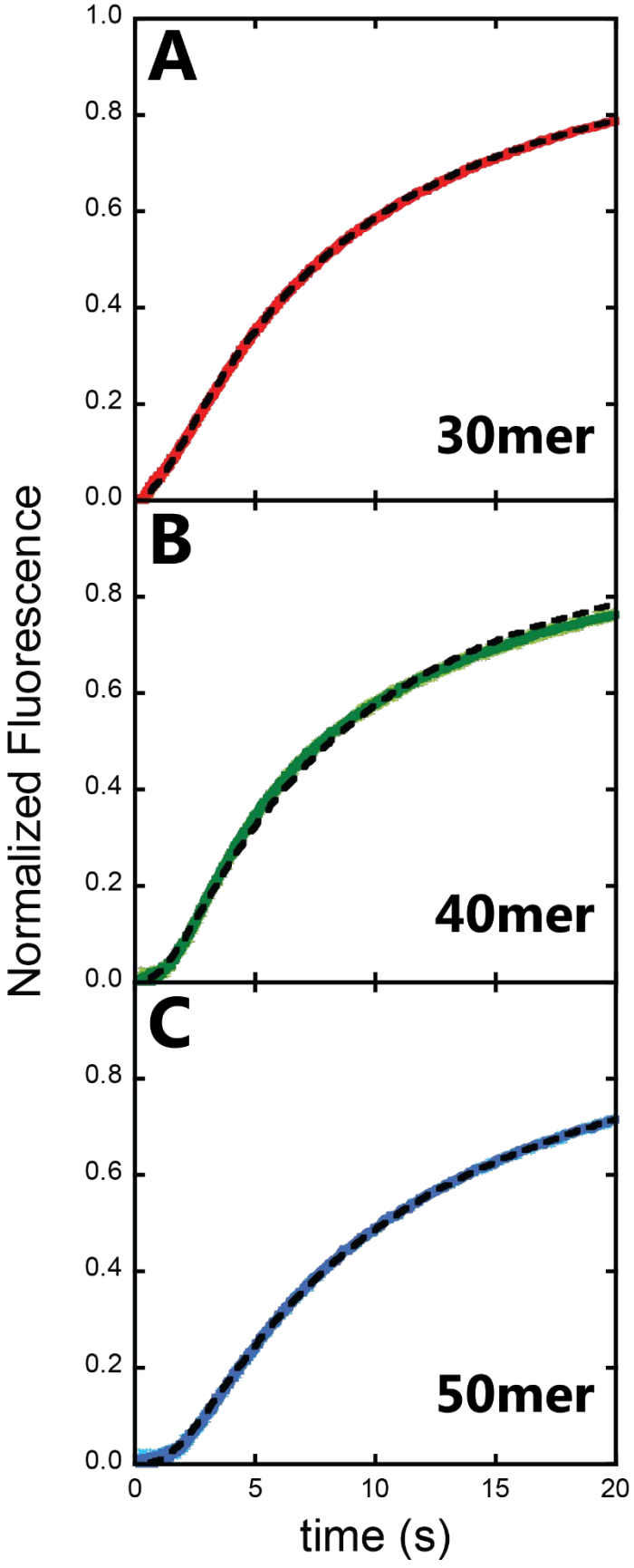


FIGURE 2 MENOTR analysis of single-turnover ClpA catalyzed translocation of SsrA (**A**) 30, (**B**) 40, and (**C**) 50 aa polypeptide substrates. Representative time courses for the translocation of each substrate length are shown in the solid colored traces. Fluorescence of fluorescein was monitored as a function of time for each substrate. Global analysis was performed using **Scheme 2** and **Eq. 2** across all three substrates lengths. The rate constants k_T, k_C, k_{NP} as well as h and m were set as global parameters, while the amplitudes A and x were designated as local parameters for each substrate. The best-fit parameters determined from this are presented in **Table 1.2** and the corresponding best-fits are plotted here as broken black traces. A corresponding figure for the original fits and analysis of this data can be found in Rajendar et al. 2009, Figure 3 (30).

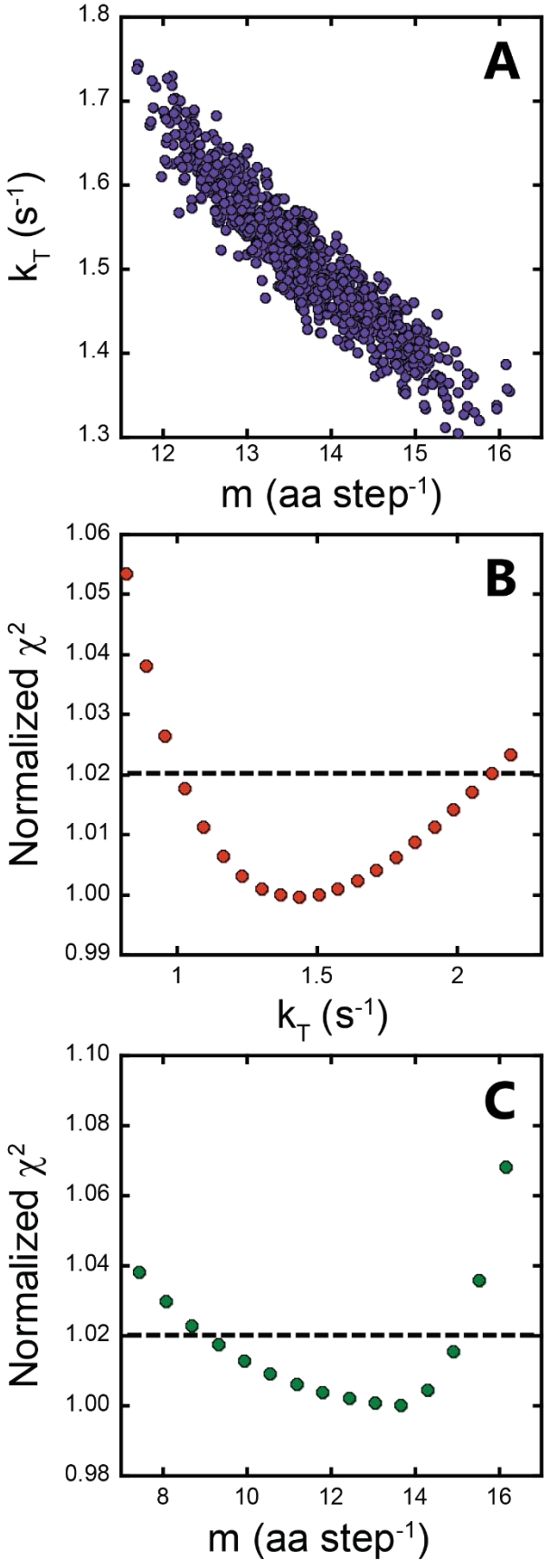


FIGURE 3 Monte Carlo and grid search analysis of parameter correlation and confidence intervals for ClpA catalyzed translocation of polypeptide substrate. (**A**) Plot of the translocation rate constant, k_T, as a function of kinetic step-size, m. This analysis was performed on the time courses found in **Fig. 2**. The k_T and m pairs are generated from 1000 Monte Carlo simulations as described in the **Methods**. Plots of normalized chi-squared as a function of (**B**) the translocation rate constant and (**C**) kinetic step-size from the time courses found in **Fig. 2**. The F critical for this data was determined at a 68% confidence to be 1.0201, which is plotted as a horizontal broken black trace. Comparing the error contour to the F critical line yields confidence intervals of k_T = (1.01 – 2.12) s^{-1} and m = (9.00 – 15.09) aa $step^{-1}$. A corresponding figure for the original analysis of this data with Monte Carlo and grid searches can be found in Rajendar 2009, **Figure 5** (30).

TABLE 1 Optimized parameter comparison for kinetic benchmarks I and II

(1.1) RecBCD catalyzed DNA unwinding parameters

Parameter	k_U (s^{-1})	k_C (s^{-1})	k_{NP} (s^{-1})	m (bp $step^{-1}$)	h (steps)
Published Fits	200 ± 40	51 ± 5	6.0 ± 0.3	3.4 ± 0.6	3.2 ± 0.3
MENOTR Fit Results	185.5 ± 0.1	54.9 ± 0.6	6.49 ± 0.08	3.68 ± 0.02	3.31 ± 0.04

(1.2) ClpA catalyzed polypeptide translocation parameters

Parameter	k_T (s^{-1})	k_C (s^{-1})	k_{NP} (s^{-1})	m (aa $step^{-1}$)	h (steps)
Published Fits	1.46 ± 0.05	0.20 ± 0.003	0.045 ± 0.0005	13.4 ± 0.5	1
MENOTR Fit Results	1.5 ± 0.2	0.165 ± 0.008	0.040 ± 0.003	14 ± 1	1

CHAPTER 2

APPLICATION OF COMBINED TOTAL FLUORESCENCE AND FLUORESCENCE ANISOTROPY TECHNIQUES TO THE TRANSIENT-STATE KINETICS OF ATP-DEPENDENT MOTOR PROTEINS

NATHANIEL W. SCULL
AARON L. LUCIUS

Introduction

ATPases Associated with various cellular Activity (AAA+) is a large superfamily of protein translocases. Some of these translocases function biologically by associating with a protease partner, and act as both a motor and regulatory component. Examples include the eukaryotic 19S cap of the 26S proteasome (1), members of the bacterial Clp/Hsp100 (2) and HslU (3) families, and PAN/20S in archaea (4, 5). The investigation of the mechanisms of translocation catalyzed by such systems has been accomplished through steady-state degradation of model substrates (6), like green fluorescent protein (GFP), and Förester resonance energy transfer (FRET) using dyes located on the substrate and inside the protease (7, 8). However, other translocases such as NSF (9), katanin/spastin (10, 11), VPS4 (12), ClpB/Hsp104 (13, 14) and p97 (15) perform their functions without a proteolytic partner. In these examples, there is no covalent modification of the substrates they translocate. This makes monitoring translocation catalyzed by these motors more challenging. Consequently, it remains unclear if these motors fully translocate substrate through their axial channels, or only partially translocate (16, 17).

Recent advances in Cryo-EM techniques has led to a wealth of important new structural insights on AAA+ polypeptide translocases (18-24). With new structures have come many structural hypotheses on the mechanism of polypeptide translocation catalyzed by AAA+ molecular motors. However, static structures are not able to reveal dynamic parameters such as step-size (distance travelled per turnover), coupling efficiency

(ATP molecules hydrolyzed per step), elementary rate constants, and processivity. Consequently, there is an urgent need for techniques that can test these structural hypotheses in solution.

We have previously developed a transient-state kinetics technique to investigate motors that do not associate with, or have been separated from, their proteolytic partner (25). This technique employed fluorescently modified synthetic polypeptide substrates (17, 25, 26). When these fluorescent substrates are bound to the motor proteins, there is a marked decrease in the observed fluorescence intensity. Upon translocation, there is a transition from a population of bound to unbound substrate as the enzyme and substrate dissociate. This results in an observed increase in the fluorescence as a function of time. The difference between the bound and the unbound substrate fluorescence can arise from differences in the quantum yield of the fluorophore on each species, rotational artifacts of each species on the time scale of fluorescence (27), or a combination of the two. When using these techniques, it assumes an all or none scenario, where the substrate is either in a bound or unbound fluorescent state.

The investigation of translocation catalyzed by ClpA on unstructured polypeptide substrates indicated that ClpA is taking ~13 amino acid (aa) sized steps on its substrate during translocation (25). Given that the substrates used were 30-50 aa in length, only 2-4 steps were taken per polypeptide translocated. While we have been able to determine a set of kinetic parameters describing the mechanism of translocation, some properties of translocation could not be resolved. One important parameter is processivity, or the likelihood of the translocase to continue translocating versus dissociating from its lattice. ClpA

has been reported as a highly processive motor that is able to translocate entire target substrates into its proteolytic partner ClpP for degradation (28-30). However, a quantitative estimate of processivity has not been reported. We were only able to report a lower limit of processivity, at limiting ATP of P = ~0.88 (25). Longer polypeptide substrates will be required to quantitatively determine processivity at saturating ATP concentrations where translocation is fast. However, longer substrates in our previously developed stopped-flow strategy did not yield interpretable signal changes (unpublished results).

In principle, anisotropy stopped-flow techniques would be better suited for an examination of polypeptide translocation because it is sensitive to the presence of the motor on the lattice. That is to say, the polypeptide substrate being translocated is much smaller than hexameric ClpA (~500,000 Daltons). Thus, a polypeptide substrate bound by ClpA should exhibit a larger anisotropy than substrate alone, and the anisotropy should decrease as ClpA dissociates from the peptide after complete translocation. In fact, we and others have reported stopped-flow anisotropy time courses that exhibit exactly that behavior for ClpA (7, 25), ClpB (16), and Hsp104 (17). However, only a qualitative description of those time courses was reported. The reason for this is that anisotropy time courses do not follow typical exponential behavior that can be described by sums of exponentials. Thus, the analysis is uniquely complex because anisotropy time courses are described by a ratio of sums of exponentials that are sensitive to both changes in anisotropy and quantum yield.

Here we define a fluorescence anisotropy method for observing and modeling translocation of polypeptide substrates by protein translocases that do not covalently modify the polypeptide on which they translocate. These methods can account for unique

intermediate species as well as distinguish the contribution of changes in quantum yield and rotational mobility to the overall signal. In this setup, translocation is monitored concurrently in two ways. First, as total fluorescence, which is collected such that rotational artifacts are eliminated from the signal (31). As such, total fluorescence is defined singularly by quantum yield. Second, as fluorescence anisotropy that reports on the rotational mobility of the complex, which in turn depends on the size of the complex and is thereby sensitive to the presence of the motor on the lattice (31). Here we report the simultaneous analysis of both total fluorescence and anisotropy time courses using the techniques described by Otto et al. and our previous application of n-step sequential mechanisms (25, 27, 32, 33). With these techniques, we have been successful in both reproducing kinetic parameters from the analysis of the original polypeptides (25) and, for the first time, gaining quantitative interpretations of the translocation of polypeptide substrates longer than 100 aa using stopped-flow anisotropy.

Materials and Methods

Materials

Solutions

All solutions were prepared in double distilled water from a Siemens Water Technology, Purelab Ultra Genetic system (Alpharetta, GA) with commercially available reagent grade chemicals. All experiments were carried out in buffer H300: 25 mM HEPES pH 7.5 at 25 °C, 300 mM NaCl, 10 mM $MgCl_2$, 2 mM 2-mercaptoethanol, and 10% v/v glycerol.

Polypeptide substrates

SsrA polypeptides were synthesized and fluorescently labeled by CPC Scientific (Sunnyvale, CA) and were certified as >90% pure by LC-MS analysis. The peptides consist of the C-terminal SsrA tag flanked by truncations of the Titin I27 domain. Each peptide has an N-terminal cysteine that has been labeled with a fluorescein dye. These peptides have been previously analyzed using CD spectroscopy to show that they lack any significant structure and bind to ClpA (34).

αS1-casein truncations were constructed as detailed in our previously published work. Each substrate consists of a C-terminal truncation of αS1-casein that is labeled with fluorescein-5-maleimide on an N-terminal cysteine. These peptides have been previously analyzed using CD spectroscopy to show that they lack any significant structure and bind to ClpA (34).

ClpA

ClpA was purified in a manner similar to that as described in Veronese 2009 Biochemistry (35). The only significant difference was the omission of the final Blue Sepharose FF column. The reported concentrations of ClpA were determined in H300 buffer using the molar extinction coefficient 31,000 $M^{-1}cm^{-1}$ and reported as monomers.

<u>*Methods*</u>

Polypeptide translocation experiments

Polypeptide translocation experiments were carried out on an Applied Photophysics SX20 stopped-flow spectrometer (Leatherhead, UK). First, 4 μM ClpA was incubated with 300 μM ATPγS in H300 for 30 min at 25 °C. Then 200 nM fluorescent SsrA substrate was added and incubated for 15 min to achieve binding equilibrium. Equilibrium

was assured by monitoring fluorescence change in a HORIBA Jobin Yvon Fluorolog-3 spectrofluorometer (Albany, NY) over two hours. No change was observed after 15 minutes. Additionally, 6 mM ATP and 20 μM α-casein were incubated in H300 for 45 min at 25 °C.

Raw fluorescence stopped-flow experiments were collected in an L configuration using the longer 10 mm path length**.** Prior to each set of experimental acquisitions, the stopped-flow system is prepared by adjusting the PMT voltages. A single acquisition is taken, and the chemistry is allowed to react. Once the steady state of the final reaction conditions is met, the PMT voltages are set to 6 V and auto adjusted by the instrument. The stopped-flow system is washed with buffer matching that of reaction conditions between each experiment. Each data set gathered represent the mean and standard deviation of at least 4 acquisitions.

Fluorescence anisotropy data for translocation was collected using a T-format. Prior to each set of experimental acquisitions, the stopped flow is prepared by adjusting the PMT voltages and determining the G factor. A single acquisition is taken, and the chemistry is allowed to react. Once the steady state of the final reaction conditions is met, the PMT voltages are set to 15% and the G-factor is obtained to a 10% level of precision as described in the instrument manual. The stopped-flow system is washed with buffer matching that of reaction conditions between each experiment. Each data set gathered represents the mean and standard deviation of at least 4 acquisitions.

MENOTR

All the data analysis performed here was carried out in the MATLAB (The MathWorks Inc., Natick, MA) toolbox MENOTR. MENOTR, Multi-start Evolutionary Nonlinear OpTimizeR, is a hybrid multi-start genetic and NLLS algorithm. This hybrid fitting method utilizes not only the intrinsic strengths of both stochastic, genetic, and deterministic, NLLS, techniques but also balances the strengths of each method to offset their corresponding limitations. The genetic algorithm surveys large and diverse error spaces to find possible solutions. However, due to lack of convergence, the genetic algorithm becomes computationally expensive and underperforms in the later stages of optimization. In contrast, the NLLS algorithm converges quickly on a local minimum, but cannot guarantee a global minimum due to its inability to escape local minimums or survey sufficient solution space. In this hybrid method, the genetic algorithm will quickly find good starting points for the slower running NLLS algorithm and simultaneously assist in the escape from local minima. The collaboration between the approaches yields an overall more robust and superior method of generating high-quality solutions to fitting problems. More information on this algorithm can be found in (Ingram, Scull. 2020, In Preparation). MENOTR uses the model equations discussed in this manuscript to find a set of optimized parameters that best describe each data set. This is discussed in detail in the below sections.

Results and Theory

Application of raw fluorescence stopped-flow methods to examine polypeptide translocation

We have developed a single-turnover transient-state kinetics technique to examine the mechanism of polypeptide translocation catalyzed by ClpA, ClpAP, ClpB and Hsp104 (16, 17, 25, 26). The experimental design of the technique is schematized in **Fig. 1 A & B**. Briefly, in the context of ClpA, we pre-mix ClpA with the slowly hydrolysable ATP analog, ATPγS, and polypeptide substrate that contains a fluorescein at the N-terminus and a known binding sequence for ClpA at the C-terminus. The ATPγS is included because ClpA requires nucleotide binding to form hexamers active in polypeptide substrate binding. This sample, which contains hexameric ClpA bound to a fluorescently modified polypeptide substrate, is loaded into syringe 1 of the stopped-flow system. Syringe 2 is loaded with a solution containing ATP and non-fluorescent polypeptide substrate. The non-fluorescent substrate is included at large excess over the prebound ClpA-substrate complex to serve as a trap for any ClpA that is free in solution or that dissociates from the polypeptide during or after translocation, thereby maintaining single-turnover conditions.

The contents of the two syringes are rapidly mixed, and fluorescence emission from fluorescein is monitored, see **Fig. 2 A** for two representative time courses. As we have previously reported and show in the supplemental (see **Fig. S1 A - C**), fluorescein fluorescence is quenched when ClpA is bound relative to fluorescence of the free substrate (25). Consequently, the kinetic time courses shown in **Fig. 2 A** exhibit low fluorescence at time zero followed by an increase in fluorescence over time indicating that ClpA

is dissociating upon mixing with ATP and protein trap. These experiments are single turnover with respect to the polypeptide substrate, which means that ClpA cannot rebind the fluorescently modified substrate after the first round of translocation. Therefore, the kinetic time courses represent a single round of substrate translocation followed by dissociation. Consistent with our previous observations, the kinetic time courses shown in **Fig. 2 A** exhibit a constant fluorescence signal, or lag, between 0 and ~2 s prior to the increase in fluorescence. The duration of the lag is observed to increase with increasing substrate length as seen in **Fig. 2 A** for representative kinetic time courses collected with two different substrate lengths of 30 and 50 aa. We have interpreted this increase in the lag with increasing substrate length to indicate that ClpA is taking more kinetic steps before dissociation on a longer substrate compared to a shorter substrate. This observation is consistent with translocation from the binding site at the C-terminus to the N-terminus.

We developed this method using short synthetic polypeptides substrates ranging in length between 30 and 50 aa (see **Table 1**). Synthetic polypeptides longer than 50 aa are both difficult to synthesize and cost prohibitive. Thus, our attempts to examine substrates longer than 50 aa were performed with fluorescently labeled truncations of the protein αS1-casein, which were expressed recombinantly. This substrate was chosen for two reasons. First, αS1-casein is a natural substrate for ClpA. Second, αS1-casein is an intrinsically disordered protein. Although we are interested in examining translocation on folded proteins, we have shown that the short synthetic polypeptides are unfolded and we have extensively characterized translocation on these substrates (34). Thus, for consistency, we are seeking to examine translocation on substrates with minimal structure.

To test translocation on substrates longer than 50 aa we performed single-turnover stopped-flow experiments with 2 μM ClpA, 300 μM ATPγS, and 200 nM fluorescent αS1-casein. The αS1-casein substrates are 102 and 127 aa in length (See **Table 1**). Syringe 2 was loaded with 6 mM ATP and 20 μM unlabeled α-casein to serve as a trap. **Fig. 2 B** shows representative time courses collected with the αS1-casein substrates. The observed signal does not show the same clear trend that was observed with the shorter synthetic substrates shown in **Fig. 2 A**.

There are multiple potential explanations for the signal observed with αS1-casein. For example, although casein is a natural substrate for ClpA, there is no clear binding site for ClpA. Therefore, we do not know if ClpA binds the N-terminus, C-terminus, or somewhere in between. Meaning, ClpA may bind close or far away from the fluorescein dye. Consequently, the kinetic time courses indicate that the effect of the protein on the dye is unclear.

To overcome these unknowns, we sought to test translocation of the longer polypeptides using anisotropy stopped-flow methods. Anisotropy is sensitive to the size of the fluorescently modified species. Thus, fluorescently modified αS1-casein bound to hexameric ClpA (~500 kDa) will exhibit a higher anisotropy value relative to free αS1-casein (~12 kDa). We have previously shown this to be the case in steady-state anisotropy titrations and single-turnover translocation experiments (17, 34, 36). Thus, we hypothesized that an anisotropy time course collected with ClpA prebound to polypeptide substrate should exhibit a high anisotropy at time zero followed by a decrease towards an anisotropy value consistent with free substrate (See **Fig. S2**) as ClpA dissociates. Thus, the

anisotropy time course will reflect the residence time of ClpA on the polypeptide substrate regardless of where ClpA binds relative to the position of the fluorophore.

Replication of results for ClpA catalyzed translocation of SsrA substrates using fluorescence anisotropy

To determine if anisotropy stopped-flow is a viable method for examining polypeptide translocation catalyzed by ClpA we first sought to determine if we could acquire the same kinetic parameters for our previously reported synthetic polypeptide substrates of 30 – 50 aa (see **Table 1**). Anisotropy stopped-flow experiments were carried out as schematized in **Fig. 1 A** and the signal was monitored as schematized in **Fig. 1 C**. Briefly, the sample is excited with vertically polarized light at 494 nm. The emission path is set up in a T-format with two photomultiplier tube (PMT) detectors set 90° to the incident light and 180° to one another. One detector collects vertically polarized emitted light ($I_{\parallel}$) while the other collects horizontal emissions ($I_{\perp}$). Each PMT is fit with a 515+ nm long pass filter to block any excitation light.

In this experiment, total fluorescence **Eq. 1** and fluorescence anisotropy **Eq. 2** are gathered simultaneously, meaning two time courses are collected for each acquisition. The time courses for ClpA catalyzed translocation of SsrA 30, 40, and 50mer are shown in **Fig. 3 A & B** for total fluorescence and anisotropy, respectively. In the total fluorescence time courses, we see an initially quenched fluorescence that proceeds through a lag phase followed by a second phase that rises to a final plateau. These data follow the same trends as described in the raw fluorescence time courses shown in **Fig. 2 A** and previously reported. As predicted, anisotropy starts at an initial high anisotropy value. It then

proceeds through a lag followed by a decrease in signal eventually plateauing at an anisotropy value consistent with free peptide (the anisotropy value for free peptide is shown in **Fig. S2**).

The question is, can we extract mechanistic information from the anisotropy time courses since the time courses cannot be described by a simple sum of exponentials (27). As stated, when collecting fluorescence anisotropy experiments, two types of signals are simultaneously gathered: total fluorescence (*TF*) and fluorescence anisotropy (*r*). When exciting the sample with vertically polarized light each signal is described by **Eq. 1 & 2**, respectively.

$$TF = G \cdot I_{\parallel} + 2 \cdot I_{\perp} \quad (1)$$

$$r = \frac{(G \cdot I_{\parallel} + I_{\perp})}{(G \cdot I_{\parallel} + 2 \cdot I_{\perp})} \quad (2)$$

Where $I_{\parallel}$ and $I_{\perp}$ are the measured fluorescence intensities detected parallel and perpendicular to the incident polarization (See **Fig. 1 C**). The G-factor, G, is the ratio of the sensitivity of the two detectors. The G factor accounts for any differences in the optical paths of the two detectors.

In order to apply **Eq. 1 & 2** to the kinetic time courses collected for ClpA catalyzed polypeptide translocation we need them in a form that can be used to describe all of the species that have unique fluorescence values (quantum yields) and unique anisotropy values as a function of time. Total fluorescence as a function of time is the summation of all the states, y_i, multiplied by their respective quantum yield, q_i, (See **Eq. 3**). Similarly, anisotropy is the summation of all the states multiplied by both their respective quantum

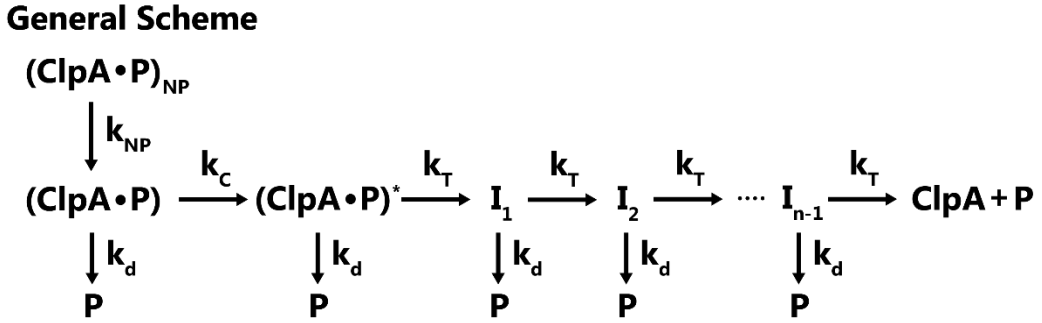
yield and the anisotropy, r_i. However, as can be seen in **Eq. 2**, anisotropy is normalized by total fluorescence. Therefore, an additional term in the time dependent anisotropy function that accounts for this normalization is required. By dividing the anisotropy summation by the total fluorescence summation, the same normalization can be achieved in the time dependent function (See **Eq. 4**).

$$TF(t) = \sum y_i(t) \cdot q_i \quad 3$$

$$r(t) = \frac{\sum y_i(t) \cdot q_i \cdot r_i}{\sum y_i(t) \cdot q_i} \quad 4$$

The task at hand is to determine time dependent functions, $y_i(t)$, for each species in our reaction. We have previously reported that **Scheme 1** describes ClpA catalyzed polypeptide translocation (25). From this scheme, we can construct a system of coupled differential equations that describe each species in the mechanism. It has been demonstrated that Laplace transforms can be applied to solve these systems of differential equations (33). The transformation of the system of differential equations into Laplace space results in a much simpler system of coupled algebraic equations, $y_i(s)$. This not only simplifies the system but more importantly allows us to construct a model within Laplace space where the number of steps, n, is a floating parameter. This is essential because it is not initially known how many steps ClpA is taking on a given substrate lattice. Thus, n, is a fitting parameter of interest. Once we acquire each $y_i(s)$ the task is to find the inverse Laplace transform to yield $y_i(t)$ given by **Eq. 5**.

$$y_i(t) = \mathcal{L}^{-1} y_i(s) \quad 5$$



(**Scheme 1**)

Scheme 1 depicts the reaction mechanism used to describe ClpA catalyzed translocation of polypeptide substrate. Before mixing with ATP, ClpA is bound to fluorescent substrate in two forms: nonproductive ClpA complex $(ClpA•P)_{NP}$ and productive ClpA complex (ClpA•P). Upon mixing with ATP $(ClpA•P)_{NP}$ slowly isomerizes with rate constant k_{NP} to form (ClpA•P). Since these two states are present before rapid mixing with ATP, we define the relative populations of the two states by the fraction of productively bound complexes, x, where x is the ratio of productive complex to total complex. Translocation begins and some number, n-1, of intermediate steps, I, are taken governed by the rate constant k_T. With the last translocation step, the ClpA peptide complex will dissociate. Throughout the scheme it is possible for peptide, P, to dissociate from ClpA prior to the completion of translocation; these steps are described by the rate constant k_d. Additionally, there is a slow conformational step to (ClpA•P)* described by the rate constant k_C that repeats h number of times. Applying **Eq. 5** to **Eq. 3 & 4** yields Eq. **6 & 7**.

$$TF(t) = \mathcal{L}^{-1}\sum y_i(s)\cdot q_i \qquad 6$$

$$r(t) = \frac{\mathcal{L}^{-1}\left(\sum y_i(s)\cdot q_i \cdot r_i\right)_i}{\mathcal{L}^{-1}\left(\sum y_i(s)\cdot q_i\right)} \quad 7$$

With a strategy for finding the time dependent functions, $y_i(t)$ for each intermediate, the next step is to approximate the quantum yield values, q_i, and the anisotropy values, r_i. In our experimental design, the signal is sensitive to only fluorescently modified substrate. In the simplest case, two states of species exist during translocation. The time courses in **Fig. 3 A & B** indicate, at a minimum, that bound and free substrate have different anisotropy and total fluorescence values. Thus, the first level of analysis is to assume that the quantum yield and anisotropy of all of the substrate bound states are given by q_1 and r_1, respectively, and all free substrate is given by q_2 and r_2 (see **Eq. 8 & 9**). To apply **Eq. 6 & 7** to the analysis of the time courses in **Fig. 3 A & B** the summation is expanded and is given by **Eq. 8 & 9.**

$$TF(t) = \mathcal{L}^{-1}\sum y_i(s)\cdot q_i = \mathcal{L}^{-1}\Big(q_1\big([ClpA\bullet S]_{NP}(s) + [ClpA\bullet S](s) + [ClpA\bullet S]^*(s) + I_1(s) + I_2(s) + I_{n-1}(s)\big) + q_2\big(P(s)\big)\Big) \quad 8$$

$$r(t) = \frac{\mathcal{L}^{-1}\Big(r_1 q_1\big([ClpA\bullet S]_{NP}(s) + [ClpA\bullet S](s) + [ClpA\bullet S]^*(s) + I_1(s) + I_2(s) + I_{n-1}(s)\big) + r_2 q_2\big(P(s)\big)\Big)}{\mathcal{L}^{-1}\Big(q_1\big([ClpA\bullet S]_{NP}(s) + [ClpA\bullet S](s) + [ClpA\bullet S]^*(s) + I_1(s) + I_2(s) + I_{n-1}(s)\big) + q_2\big(P(s)\big)\Big)} \quad 9$$

During analysis, MENOTR solves **Eq. 8 & 9** by numerically approximating the inverse Laplace transform of both equations. Note that for anisotropy, inverse Laplace

transforms must be performed on the numerator and denominator separately prior to normalization. The transforms yield solutions in the time domain that represent the total fluorescence and anisotropy signals. For full expansions of the fitting functions, see supplemental equations **Eq. S1 - S4.**

The time courses from **Fig. 3 A & B** were subjected to global MENOTR analysis to determine a set of best-optimized kinetic parameters using **Scheme 1** and the derived two-state model (**Eq. S1 - S4**). A description of the analysis can be found in the **Methods**. The resulting best fits of the data are plotted as broken black traces in **Fig. 3 A & B**. All the determined parameters agreed with those determined using raw fluorescence techniques. The translocation rate constant was found to be $k_T = (1.8 \pm 0.6)\ s^{-1}$ compared to $k_T = (1.5 \pm 0.2)\ s^{-1}$, while the kinetic step-size was found to be m = (14.4 ± 0.1) aa compared to m = (14 ± 1) aa. The slow pre-translocation transition rate constant was determined to be $k_{NP} = (0.06 \pm 0.01)\ s^{-1}$ compared to $k_{NP} = (0.040 \pm 0.003)\ s^{-1}$. The slow non-translocation rate constant was found to be $k_C = (0.2 \pm 0.1)\ s^{-1}$ compared to $k_C = (0.168 \pm 0.008)\ s^{-1}$.

It can be seen in **Fig. 3 A** that the fit lines do not describe the total fluorescence time courses as well as the anisotropy time courses in **Fig. 3 B**. This is the consequence of the fact that the total fluorescence time courses exhibit ‘drift’ from acquisition to acquisition. Recall, that the time courses represent the average of multiple acquisitions, where each acquisition represents a time course collected with the same sample contained within the two syringes of the stopped flow (see **Fig. 1 A**). Although the shape of each time course from acquisition to acquisition is identical, its total fluorescence value is variable (See **Fig. S3**). In our previously published raw fluorescence results we also observe

this phenomenon. However, previously the raw fluorescence time courses were offset so that they overlay each other. Since the shape of the curves are all the same, this procedure of overlaying the curves does not influence the determination of the kinetic parameters. However, in the analysis of the total fluorescence time courses we cannot offset the time courses because the absolute value of the fluorescence is required to analyze the anisotropy time courses, see **Eq. 3 & 4**. Consequently, when we average multiple acquisitions there is a standard deviation on each time point that accounts for the drift in the signal. Strikingly, the anisotropy time courses from acquisition to acquisition overlay near perfectly (See **Fig. S3**). Upon averaging the acquisitions, as expected due to drift, we see that there is a larger standard deviation on each total fluorescence data point compared to each anisotropy time point (See **Fig. S4**). The consequence of this is that the anisotropy time points are weighted heavier in the global analysis than the total fluorescence time points because the analysis uses the standard deviation to weight each data point when finding the minimum of the fitting function. However, all the fit lines shown in **Fig. 3 A & B** fall within the standard deviation of the data points and thus we conclude that the combined fit of both total fluorescence and anisotropy is a good description of the data.

Single-turnover time courses of extended length polypeptide substrates

The analysis of the total fluorescence and the anisotropy time courses collected with synthetic polypeptide substrates yielded results consistent with our raw fluorescence time courses. Therefore, we next sought to test the anisotropy technique with the αS1-casein substrates. Recall, the αS1-casein substrates yielded time courses in raw fluorescence that we deemed uninterpretable, see **Fig. 2 B**. For these experiments, fluorescently modified αS1-casein 102mer and 127mer (see **Table 1**) were pre-bound to ClpA and

loaded in Syringe 1 of the stopped-flow fluorometer and, in syringe 2, was loaded with 6 mM ATP and protein trap, as shown in **Fig. 1 A**. The contents of the two syringes were rapidly mixed and total fluorescence and anisotropy time courses were collected. The resultant time courses are shown in **Fig. 4 A & B**.

The total fluorescence time course for 127mer exhibits an initial lag followed by a clearly defined decrease, while anisotropy exhibited an initial rise followed by the expected decay. The trends in the 102mer total fluorescence time courses are not the same. The total fluorescence initially rises over the first 100 s and then decays back towards the initial total fluorescence value. On the other hand, the anisotropy time courses displayed an initial lag followed by the expected decay. It should also be noted that, in addition to being different from each other, the signals for the 102mer and 127mer are not the same as what we observe with the SsrA substrates.

To try and better understand the observed signal changes between bound and unbound peptide, steady-state measurements of the αS1-casein and SsrA substrates alone as well as incubated with 15 μM ClpA and 300 μM ATPγS were collected (**Fig. S1**). The synthetic SsrA substrates (**Fig. S1 A - C**) all exhibit steady-state fluorescence signals that are consistent with the time courses collected here (**Fig. 3 A**) and our previous results. Where unbound substrate has a higher fluorescence intensity compared to the quenched fluorescence of bound substrate. In contrast to this, the αS1-casein steady-state fluorescence measurements showed an opposite effect upon binding to ClpA. Unbound substrate had lower fluorescent intensities than that of bound, i.e. fluorescein fluorescence is enhanced upon binding. While the aS1-casein substrates did not exhibit a fluorescence con-

sistent with the SsrA polypeptide substrates, the two aS1-casein substrates were consistent with each other and consistent with the observations of total fluorescence in the stopped flow shown in **Fig. 4 A**.

We have previously reported steady-state anisotropy values of bound and unbound αS1-casein 127mer and 102mer. We found that 127mer and 102mer had steady-state anisotropy values of 0.11 and 0.08, respectively. When bound to another Hsp/Clp chaperone, ClpB, their steady-state anisotropy values increased to 0.23 and 0.18 respectively (36). With ClpA, we reported that 127mer increased to an anisotropy of ~0.25 (34). We did not previously determine a value for 102mer bound to ClpA; however, we expect to see an increase with this substrate as well. Based on the results of both the steady-state fluorescence and anisotropy for bound and unbound αS1-casein substrates we predict that there will need to be at least two values of quantum yield and anisotropy used in the model to describe that data in **Fig. 4**.

Recall, the time courses collected with the SsrA substrates were described assuming that all substrate bound states had the same quantum yield and anisotropy value and free substrate has a different quantum yield and anisotropy. We applied the same model used to describe ClpA translocation of the SsrA substrates here with the αS1-casein substrates. However, the model did not adequately describe the time courses based both on visual inspection and poor chi-squared values (fits not shown). It is likely still the case that free αS1-casein substrate can be described by a single quantum yield and a single anisotropy value. Thus the differences in the observed signal compared to the SsrA substrates is likely due either to the pre-bound states ($(ClpA•P)_{NP}$ and (ClpA•P)) or the various translocation intermediates, I_i, exhibiting different quantum yields and anisotropies.

To test this idea, we derived a model where we assume that the initial pre-bound ClpA states exhibits a quantum yield, q_1, and anisotropy, r_1 whereas the summation of all the translocation intermediates exhibit a different quantum yield, q_2, and anisotropy, r_2. Finally, free peptide exhibits a quantum yield and anisotropy, q_3 and r_3, respectively. Other 3 state and 4 state models were tested, none of which produce statistically superior results than this simpler three state model (results not shown). Applying this to **Eq. 6 & 7** we can expand the summations to **Eq. 10 & 11**.

$$TF(t) = \mathcal{L}^{-1}\left(\sum y_i(s)\cdot q_i\right) = \mathcal{L}^{-1}\left(q_1\left([ClpA\bullet S]_{NP}(s) + [ClpA\bullet S](s)\right) + q_2\left([ClpA\bullet S]^*(s) + I_1(s) + I_2(s) + I_{n-1}(s)\right) + q_3\left(P(s)\right)\right) \quad (10)$$

$$r(t) = \frac{\mathcal{L}^{-1}\left(\sum y_i(s)\cdot q_i\cdot r_i\right)}{\mathcal{L}^{-1}\left(\sum y_i(s)\cdot q_i\right)} = \frac{\mathcal{L}^{-1}\left(r_1q_1\left([ClpA\bullet S]_{NP}(s) + [ClpA\bullet S](s)\right) + r_2q_2\left([ClpA\bullet S]^*(s) + I_1(s) + I_2(s) + I_{n-1}(s)\right) + r_3q_3\left(P(s)\right)\right)}{\mathcal{L}^{-1}\left({}_1q_1\left([ClpA\bullet S]_{NP}(s) + [ClpA\bullet S](s)\right) + q_2\left([ClpA\bullet S]^*(s) + I_1(s) + I_2(s) + I_{n-1}(s)\right) + q_3\left(P(s)\right)\right)} \quad (11)$$

Three-state models were derived (See **Eq. S5 - S8**) and used to perform global MENOTR analysis on the αS1-casein substrate data. The resulting translocation rate constant was found to be $k_T = (1.18 \pm 0.03)\ s^{-1}$, similar to what was determined for the SsrA substrates, $k_T = (1.8 \pm 0.6)\ s^{-1}$. The kinetic step-size found for the αS1-casein substrates, m = (11.9 ± 0.2) aa, was just under that of the SsrA substrates, m = (14.4 ± 0.1) aa. The other kinetic parameters determined for the αS1-casein substrates were also similar to the previous results, with $k_{NP} = (0.014 \pm 0.003)\ s^{-1}$ and $k_C = (0.08 \pm 0.02)\ s^{-1}$.

Discussion

Here, we have presented a fluorescence anisotropy stopped-flow technique for use in examining the mechanism of single-turnover polypeptide translocation catalyzed by ClpA in the absence of its protease. Equally as important, we developed and report methods to quantitatively evaluate the resultant time courses to determine kinetic parameters that describe the translocation mechanism. With these techniques and methods, we show that we are able to overcome obstacles previously encountered in obtaining mechanistic information about the translocation of substrates greater than 50 aa in length. Moreover, we have shown that these techniques are suitable for monitoring the residence time of a translocase on its lattice independent of a partner protease.

Monitoring ClpA catalyzed translocation using a novel fluorescence anisotropy technique

Prior to applying florescence anisotropy to any novel investigations, we first sought to test whether this technique is sensitive to ClpA catalyzed translocation. We have previously developed a single-turnover fluorescence stopped-flow technique sensitive to translocation catalyzed by ClpA (25). In that assay a monochromator is used to select the excitation wavelength and fluorescence emission is detected by a photomultiplier tube containing a long pass filter (See **Fig. 1 B**). Thus, the fluorescence signal detected in that technique represents the integrated area under the emission spectra for all wavelengths above the cut off of the filter being used. Additionally, neither the excitation nor emission light passes through a polarizer, leaving all the light circularly polarized. We refer to this as ‘raw fluorescence’ to distinguish it from the total fluorescence collected in fluorescence anisotropy.

In our previous work, raw fluorescence time courses for the fluorescently modified polypeptides exhibited a substrate length dependent lag followed by an increase to a final plateau. An increase in the extent of lag with increasing lattice length is evidence that ClpA is taking more kinetic steps as the lattice length is increased. This observation is consistent with directional translocation along the polypeptide chain (25). Thus, the analysis of those data allowed for the determination of several kinetic parameters describing the elementary steps of ClpA catalyzed translocation.

Global analysis of time courses collected for both raw fluorescence and combined total fluorescence and fluorescence anisotropy produced kinetic parameters that agreed with one another. We showed here that the translocation rate constant, k_T, and the kinetic step-size, m, determined using the combined signals from anisotropy and total fluorescence were within error of those determined from raw fluorescence alone (25). This observation suggests that we are monitoring the same kinetic mechanism of polypeptide translocation in fluorescence anisotropy as previously observed using raw fluorescence. Thus, the application of these fluorescence and anisotropy-based methods to novel questions regarding ClpA and other potential translocases is appropriate and warranted.

Fluorescence anisotropy measurements allow for the observation of either partial or complete disassociation of the motor from substrate. We showed in **Fig. S2** that ClpA catalyzed translocation of all three polypeptide substrate lengths resulted in an anisotropy change consistent with complete dissociation of ClpA from its fluorescently modified substrates. Thus, we know that ClpA is not stalled on the lattice partially translocated or at the end of the substrate. This information is important in deciding on the mechanistic meanings behind the derived kinetic parameters obtained from analysis of this data. If some

subpopulation of the motors are stalled on the lattice without dissociating, the schemes describing translocation must be reevaluated to account for this type of phenomenon.

Examination of ClpA catalyzed translocation of substrates greater than 100 aa in length

Processivity, P, represents the probability of a translocase taking another step forward along its lattice versus disassociating. Processivity can be quantitatively defined as the translocation rate constant, k_T, divided by the summation of the translocation rate constant and the dissociation rate constant, k_d, see **Eq. 13** and **Scheme 1** (37). Where P = 1 indicates that a motor completely translocates its lattice with every binding event, and P < 1 indicates the motor has some probability of dissociating prior to complete translocation of its lattice. Using this probability, it is possible to predict the average number of aa translocated, N, by a motor before dissociating from the polypeptide, see **Eq. 14** (33). As a reminder, m, is the kinetic step-size of the motor, which represents the average number of aa translocated between two rate limiting steps.

$$P = \frac{k_T}{k_T + k_d} \qquad 12$$

$$N = \frac{-m}{\ln(P)} \qquad 13$$

We have assumed that ClpA is a processive motor under saturating ATP concentrations because dissociation during translocation was not detected. However, a quantitative estimate of processivity was not possible (25, 38). In those investigations k_d was only detectable at [ATP] below 300 μM, resulting in P = 0.876 at sub-saturating [ATP] using Eq. 13. The average number of amino acids translocated per binding event was calculated

to be N = ~100 AA Using Eq. 14 and P = 0.876. However, that estimate is two-fold larger than the longest substrate used in the study. Thus, we conclude that this estimate of processivity represents a lower limit and a more precise estimate is needed.

Here we have shown that fluorescence anisotropy and total fluorescence techniques can be used to collect time courses that can be quantitatively evaluated for substrates of lengths up to at least 127 aa. Strikingly, the results of the anisotropy translocation experiments performed on the fluorescently modified αS1-casein substrates, 102 and 127 aa in length resulted in the determination of kinetic parameters that agreed with both the fluorescence anisotropy and raw fluorescence results for the shorter 30-50 aa substrates. However, the analysis of both the 102 and 127 aa αS1-casein substrates did not lead to a detectible k_d precluding us from calculating a processivity based on Eq. 13. Thus, leading us to conclude that ClpA is translocating at least 127 aa per binding event and updating our estimate of processivity to P > 0.893 using Eq. 14.

We have previously constructed unstructured polypeptide substrates as long as 177 aa for use in examining ClpA binding specificity (34). In that work we found that at chain lengths greater than 127 aa multiple ClpA hexamers can bind a single substrate. When using either raw fluorescence or the combined Anisotropy and total fluorescence approach it is imperative to have only one ClpA hexamer bound per polypeptide substrate. Consequently, when using fluorescence anisotropy and total fluorescence to examine substrates between 150 and 200 aa in length we will have to perform these experiments under conditions that favor one-to-one binding of ClpA hexamer to substrate. Thus, there are still challenges to overcome to examine substrates long enough to report on processivity. On the

other hand, we are now poised to begin examining folded proteins with specific ClpA binding sequences. In this case, the binding sequence is sufficiently short to accommodate only one ClpA hexamer and ClpA will not bind to the folded region. Thus, the combined total fluorescence and anisotropy technique will allow us to determine the impact of folded structures on the polypeptide translocation mechanism, which was also not possible with raw fluorescence.

Fluorescence anisotropy as an in-solution method for evaluating translocation mechanisms

Gates et al. recently published a review of 41 cryo-EM structures that revealed comparable spiral-staircase arrangements in multiple AAA+ motors (18, 21-23, 39-44). From these structures, multiple mechanistic hypotheses for AAA+ catalyzed translocation have been proposed. These hypotheses propose that translocation involves conformational shifts in motor protomers as a consequence of ATP binding and hydrolysis. These conformational changes create a rotating spiral-staircase arrangement in the motor that is suggested to support a hand-over-hand or ratchet-like mechanism for translocation. These mechanisms are based on structural snapshots of the mutant motor proteins bound to model substrates in the presence of the slowly hydrolysable ATP analog, ATPγS.

The Class I AAA+ motor proteins ClpB (24, 42, 45) and Hsp104 (23), and just recently ClpAP (18), have been shown to adopt spiral staircase architecture in both nucleotide binding domains. Multiple motor states determined in these studies are suggested to indicate successive hydrolysis events at domains 1 and 2 that catalyze a 2 aa translocation steps per ATP hydrolysis cycle. This model is in contrast to the traditional power stroke model that has been established using in-solution biochemical assays (39). Additionally,

the 2 aa step-size is in direct contrast to all biochemical reports of ClpAP catalyzed translocation that conclude a step-size of ~ 5 aa per repeated rate limiting step (26, 46).

We propose that in-solution techniques, like the fluorescence anisotropy and total fluorescence techniques presented here, are necessary to bridge the gap between what is being observed structurally compared to previous kinetic assays. Structural studies lack the ability to directly monitor reaction pathway information because they are only sensitive to path-independent states. Moreover, it is unknown where these states exist along the reaction pathway and if they even represent reaction intermediates. In contrast, in-solution kinetic methods have the ability to shed light on path-dependent parameters like kinetic step-sizes, ATP coupling efficiencies, elementary rate constants, and processivity. While these recent structural studies provide new hypotheses for AAA+ catalyzed translocation, there is a need to have path-dependent experiments that can test these hypotheses.

Fluorescence anisotropy as a method to investigate protease-independent translocation

Many kinetic techniques applied to the investigation of AAA+ translocases utilize the degradation of substrate by proteolytic partners to monitor the reaction. These methods take advantage of proteolysis to monitor translocation by relating it to disappearance of substrate bands on an SDS-PAGE gel (47-49), loss of fluorescence as labeled lattice substrate is cleaved (6, 50), or FRET signal as labeled lattice approaches a partner dye inside the protease (7, 51). While these coupled methods are excellent for monitoring AAA+ motors in the presence of their proteases, they assume that the mechanism of translocation is the same in both the presence and absence of the proteolytic partner. However, there are

many examples of AAA+ motors that do not interact with a proteolytic partner nor do they covalently modify the polypeptide chain on which the operate.

It has been shown that some AAA+ motors, including ClpA and ClpX (26, 52), have translocation activity that is allosterically controlled by the presence of their proteases, ClpP. In the cases of ClpA and ClpX, the presence of ClpP has been shown to induce increases in rate and lead to changes in kinetic parameters like step-size. Baytshtok et al. showed that ClpXP exhibited three fold higher unfoldase activity compared to ClpX, and that ClpAP unfolds substrates sevenfold faster than ClpA (52). Additionally, Miller et al. showed that ClpAP translocates polypeptide with an almost twofold faster rate and translocates almost half as many aa per rate limiting step compared to ClpA. In order to draw conclusions about these types of allosteric interactions between the partner and its chaperone it is necessary to have methods that can measure activity of the motor in the presence and absence of the protease.

AAA+ motors such as NSF, katanin, spastin, VPS4, ClpB, and Hsp104, do not natively associate with a partner protease. If these motors processively translocate a polypeptide through their axial channel, then no covalent modification of the polypeptide chain would occur. Consequently, determining if these motors are processive translocases has proven to be a difficult question to answer. One way this problem has been overcome is through protein engineering strategies. For example, ClpB and Hsp104 have been mutated such that they include the helix-loop-helix IGL loop responsible for ClpA's association with ClpP (47, 48). As such, the corresponding variants BAP and HAP, named as either ClpB or Hsp104 that contains a ClpA-P-loop that makes contact with ClpP, are competent in forming complex with ClpP. The idea being that if proteolytic degradation is observed

in the presence of BAP or HAP then they must be threading the polypeptide chain through the axial channel and into the proteolytic chamber of ClpP the same way that ClpA and ClpX do naturally. In that work they showed that when BAP-ClpP or HAP-ClpP was provided polypeptide substrate and ATP, disappearance of substrate bands corresponding to a-casein on an SDS-PAGE gels was observed. They concluded that this is indicative of complete and processive threading of polypeptide substrate through the axial channel of both motors, similar to ClpA. It should be noted that only a single cleavage event is needed to observe this disappearance of substrate bands on a gel.

Recent transient-state kinetic analysis of ClpB and Hsp104 revealed both to be non-processive translocases in direct contrast to the BAP and HAP results (16, 17). Both $Hsp104_{WT}$ and $ClpB_{WT}$ were shown to only take 2 kinetic steps on their substrates per binding event on all substrate lengths tested. In addition, Li et al showed that the previous degradation assays performed on BAP discussed above (48) yielded substrate degradation in the absence of ATP. Thus, it is unclear whether any of the observed degradation in the BAP studies is the consequence of ATP-dependent translocation rather than another mechanism that is enhanced by the presence of ATP. Li et al. proposed that the ATP-independent degradation could be the consequence of dysregulation of ClpP activity (16, 53). Similar dysregulation is seen in ADEPs studies; where ADEPS are small molecules that have been shown to bind to ClpP in a similar manner to the ClpA-P-loops. Their binding induces opening of the ClpP axial pores that allows for unregulated degradation of substrate without a translocase (53).

We propose that the fluorescence anisotropy techniques presented here overcome some of these types of problems. The fluorescence anisotropy methods do not rely on the

presence of a partner protein and thus have the potential to be used in both protease-dependent and protease-independent investigations. Further, the only modification necessary to monitor fluorescence anisotropy is the addition of a fluorescent dye to the substrates being translocated. Thus, the motor remains unmodified and wildtype actives of AAA+ motors that do not natively associate with a protease can be directly observed.

Fluorescence anisotropy as a robust method for investigating AAA+ translocases

Here we propose that the global analysis of total fluorescence and fluorescence anisotropy represent a robust method for monitoring translocation catalyzed by AAA+ translocases. Otto et al. have shown that kinetic experiments with combined total fluorescence and anisotropy data sets can provide more insight into the reaction mechanisms compared to either alone (27).

While stopped-flow techniques involving fluorescence or anisotropy have been extensively used to observe biochemical reactions, little has been done in the way of combining both data types. ClpA (7, 25), ClpB/Hsp104 (16, 17, 43), p97 (54), 26S proteasome (51), Hsp70/60 (55), and HslU (56) systems have all been analyzed using anisotropy based techniques. However, all have only supplied qualitative descriptions of the kinetic time courses, and none have attempted to couple fluorescence anisotropy to total fluorescence in the manner described in this manuscript. Here we will discuss the drawbacks of anisotropy or fluorescence alone and explain how the combination of the two methods can overcome the limitations of each individually; thus, supporting the need for a combined strategy.

Fluorescence is a common method of monitoring reaction kinetics. However, raw fluorescence can suffer from rotational artifacts that do not represent steps in the reaction

(27, 31). In the raw fluorescence experiments detailed in **Fig. 2 B**, for the αS1-casein substrates, it was found that the resultant ClpA catalyzed translocation time courses did not exhibit the same expected trends that were observed when translocating on the shorter substrates given in **Fig. 2 A**. The differences observed between total and raw fluorescence in **Fig. 2 A** and **Fig. 2 B** could be the result of rotational artifacts that arise from changes in the rotational mobility of ClpA when translocating on the αS1-casein substrates.

Cryo-EM structures show that ~25 amino acids traverse the axial channel of ClpA. This would mean that 77 and 102 amino acids would be outside the ClpA channel for the 102 and 127 aa aS1-casein substrates, respectively (18). As ClpA translocates the substrates, the amount of overhang is decreasing as a function of time, and thus could induce changes in the rotational mobility of the ClpA-peptide complex on the time scale of translocation. This could explain what was observed in the total fluorescence time courses when the a-casein substrates were used.

When the αS1-casein substrates were examined here by simultaneously collecting total fluorescence and anisotropy the resultant total fluorescence time courses exhibit clear trends. The signal reflects only changes in quantum yield since the polarizers are set such that rotational correlation is removed when observing total fluorescence. This contrasts with what was observed in raw fluorescence, where the emitted light is circularly polarized and reflects both changes in quantum yield and rotational correlation. Similarly, the anisotropy time courses exhibit predictable behavior. The total fluorescence and anisotropy time courses were simultaneously analyzed, and the resultant kinetic parameters gave good agreement with those determined on the short synthetic polypeptide substrates, which was not possible with raw fluorescence.

The analysis of fluorescence anisotropy time courses in the absence of total fluorescence often leads to over or underestimations of the kinetic parameters determined (27). This is a consequence of fluorescence anisotropy being sensitive to quantum yield changes, see **Eq. 4**. Therefore, large changes in total fluorescence on the time scale of the reaction cause anisotropy time courses to not directly represent the changes in rotational mobility of the population during the reaction. Otto et al showed that by globally analyzing both total fluorescence and fluorescence anisotropy simultaneously the changes in rotational mobility and quantum yield could be decoupled. In this combined method, the changes in total fluorescence and fluorescence anisotropy contain direct information on the residence time of the motor on its substrate free of artifacts. Thus, providing more overall information on translocation than either of the methods individually.

References

1. Baumeister, W., and A. Lupas. 1997. The proteasome. Current opinion in structural biology 7(2):273-278.

2. Katayama-Fujimura, Y., S. Gottesman, and M. R. Maurizi. 1987. A multiple-component, ATP-dependent protease from Escherichia coli. J Biol Chem 262(10):4477-4485.

3. Kwon, A. R., C. B. Trame, and D. B. McKay. 2004. Kinetics of protein substrate degradation by HslUV. Journal of structural biology 146(1-2):141-147.

4. Sauer, R. T., and T. A. Baker. 2011. AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. Annu Rev Biochem 80:587-612.

5. Horwitz, A. A., A. Navon, M. Groll, D. M. Smith, C. Reis, and A. L. Goldberg. 2007. ATP-induced Structural Transitions in PAN, the Proteasome-regulatory ATPase Complex in Archaea. Journal of Biological Chemistry 282(31):22921-22929.

6. Weber-Ban, E. U., B. G. Reid, A. D. Miranker, and A. L. Horwich. 1999. Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. Nature 401(6748):90-93.

7. Reid, B. G., W. A. Fenton, A. L. Horwich, and E. U. Weber-Ban. 2001. ClpA mediates directional translocation of substrate proteins into the ClpP protease. Proc Natl Acad Sci U S A 98(7):3768-3772.

8. Kolygo, K., N. Ranjan, W. Kress, F. Striebel, K. Hollenstein, K. Neelsen, M. Steiner, H. Summer, and E. Weber-Ban. 2009. Studying chaperone-proteases using a real-time approach based on FRET. Journal of structural biology 168(2):267-277.

9. Fleming, K. G., T. M. Hohl, R. C. Yu, S. A. Muller, B. Wolpensinger, A. Engel, H. Engelhardt, A. T. Brunger, T. H. Sollner, and P. I. Hanson. 1998. A revised model for the oligomeric state of the N-ethylmaleimide-sensitive fusion protein, NSF. J Biol Chem 273(25):15675-15681.

10. Hartman, J. J., J. Mahr, K. McNally, K. Okawa, A. Iwamatsu, S. Thomas, S. Cheesman, J. Heuser, R. D. Vale, and F. J. McNally. 1998. Katanin, a microtubule-severing protein, is a novel AAA ATPase that targets to the centrosome using a WD40-containing subunit. Cell 93(2):277-287.

11. Eckert, T., D. T. Le, S. Link, L. Friedmann, and G. Woehlke. 2012. Spastin's microtubule-binding properties and comparison to katanin. PLoS One 7(12):e50161.

12. Monroe, N., H. Han, P. S. Shen, W. I. Sundquist, and C. P. Hill. 2017. Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. eLife 6:e24487.

13. Mogk, A., E. Kummer, and B. Bukau. 2015. Cooperation of Hsp70 and Hsp100 chaperone machines in protein disaggregation. Front Mol Biosci 2:22.

14. Doyle, S. M., and S. Wickner. 2009. Hsp104 and ClpB: protein disaggregating machines. Trends Biochem Sci 34(1):40-48.

15. Dalal, S., M. F. N. Rosser, D. M. Cyr, and P. I. Hanson. 2004. Distinct roles for the AAA ATPases NSF and p97 in the secretory pathway. Mol Biol Cell 15(2):637-648.

16. Li, T., C. L. Weaver, J. Lin, E. C. Duran, J. M. Miller, and A. L. Lucius. 2015. Escherichia coli ClpB is a non-processive polypeptide translocase. Biochem J 470(1):39-52.

17. Durie, C. L., J. Lin, N. W. Scull, K. L. Mack, M. E. Jackrel, E. A. Sweeny, L. M. Castellano, J. Shorter, and A. L. Lucius. 2019. Hsp104 and Potentiated Variants Can Operate as Distinct Nonprocessive Translocases. Biophys J 116(10):1856-1872.

18. Lopez, K. E., A. N. Rizo, E. Tse, J. Lin, N. W. Scull, A. C. Thwin, A. L. Lucius, J. Shorter, and D. R. Southworth. 2019. Conformational Plasticity of the ClpAP AAA+ Protease Couples Protein Unfolding and Proteolysis. bioRxiv:820209.

19. Matyskiela, M. E., G. C. Lander, and A. Martin. 2013. Conformational switching of the 26S proteasome enables substrate degradation. Nat Struct Mol Biol 20(7):781-788.

20. Han, H., J. M. Fulcher, V. P. Dandey, J. H. Iwasa, W. I. Sundquist, M. S. Kay, P. S. Shen, and C. P. Hill. 2019. Structure of Vps4 with circular peptides and implications for translocation of two polypeptide chains by AAA+ ATPases. Elife 8.

21. Sandate, C. R., A. Szyk, E. A. Zehr, G. C. Lander, and A. Roll-Mecak. 2019. An allosteric network in spastin couples multiple activities required for microtubule severing. Nat Struct Mol Biol 26(8):671-678.

22. White, K. I., M. Zhao, U. B. Choi, R. A. Pfuetzner, and A. T. Brunger. 2018. Structural principles of SNARE complex recognition by the AAA+ protein NSF. Elife 7.

23. Gates, S. N., A. L. Yokom, J. Lin, M. E. Jackrel, A. N. Rizo, N. M. Kendsersky, C. E. Buell, E. A. Sweeny, K. L. Mack, E. Chuang, M. P. Torrente, M. Su, J. Shorter, and D. R. Southworth. 2017. Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. Science 357(6348):273-279.

24. Rizo, A. N., J. Lin, S. N. Gates, E. Tse, S. M. Bart, L. M. Castellano, F. DiMaio, J. Shorter, and D. R. Southworth. 2019. Structural basis for substrate gripping and translocation by the ClpB AAA+ disaggregase. Nat Commun 10(1):2393.

25. Rajendar, B., and A. L. Lucius. 2010. Molecular mechanism of polypeptide translocation catalyzed by the Escherichia coli ClpA protein translocase. J Mol Biol 399(5):665-679.

26. Miller, J. M., J. Lin, T. Li, and A. L. Lucius. 2013. E. coli ClpA Catalyzed Polypeptide Translocation is Allosterically Controlled by the Protease ClpP. Journal of Molecular Biology 425(15):2795-2812.

27. Otto, M. R., M. P. Lillo, and J. M. Beechem. 1994. Resolution of multiphasic reactions by the combination of fluorescence total-intensity and anisotropy stopped-flow kinetic experiments. Biophys J 67(6):2511-2521.

28. Lee, C., M. P. Schwartz, S. Prakash, M. Iwakura, and A. Matouschek. 2001. ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. Mol Cell 7(3):627-637.

29. Thompson, M. W., S. K. Singh, and M. R. Maurizi. 1994. Processive degradation of proteins by the ATP-dependent Clp protease from Escherichia coli. Requirement for the multiple array of active sites in ClpP but not ATP hydrolysis. J Biol Chem 269(27):18209-18215.

30. Hoskins, J. R., M. Pak, M. R. Maurizi, and S. Wickner. 1998. The role of the ClpA chaperone in proteolysis by ClpAP. Proc Natl Acad Sci U S A 95(21):12135-12140.

31. Lakowicz, J. R. 1999. Principles of fluorescence spectroscopy. Kluwer Academic/Plenum, New York.

32. Lucius, A. L., J. M. Miller, and B. Rajendar. 2011. Application of the Sequential n-Step Kinetic Mechanism to Polypeptide Translocases. Methods in enzymology 488:239-264.

33. Lucius, A. L., N. K. Maluf, C. J. Fischer, and T. M. Lohman. 2003. General methods for analysis of sequential "n-step" kinetic mechanisms: application to single turnover kinetics of helicase-catalyzed DNA unwinding. Biophys J 85(4):2224-2239.

34. Li, T., and A. L. Lucius. 2013. Examination of Polypeptide Substrate Specificity for E. coli ClpA. Biochemistry 52:4941-4954.

35. Veronese, P. K., R. P. Stafford, and A. L. Lucius. 2009. The Escherichia coli ClpA Molecular Chaperone Self-Assembles into Tetramers. Biochemistry 48(39):9221-9233.

36. Li, T., J. Lin, and A. L. Lucius. 2015. Examination of polypeptide substrate specificity for Escherichia coli ClpB. Proteins 83(1):117-134.

37. McClure, W. R., and Y. Chow. 1980. The kinetics and processivity of nucleic acid polymerases. Methods in enzymology 64:277-297.

38. Duran, E. C., C. L. Weaver, and A. L. Lucius. 2017. Comparative Analysis of the Structure and Function of AAA+ Motors ClpA, ClpB, and Hsp104: Common Threads and Disparate Functions. Front Mol Biosci 4:54.

39. Gates, S. N., and A. Martin. 2020. Stairway to translocation: AAA+ motor structures reveal the mechanisms of ATP-dependent substrate translocation. Protein Sci 29(2):407-419.

40. Cooney, I., H. Han, M. G. Stewart, R. H. Carson, D. T. Hansen, J. H. Iwasa, J. C. Price, C. P. Hill, and P. S. Shen. 2019. Structure of the Cdc48 segregase in the act of unfolding an authentic substrate. Science 365(6452):502-505.

41. Twomey, E. C., Z. Ji, T. E. Wales, N. O. Bodnar, S. B. Ficarro, J. A. Marto, J. R. Engen, and T. A. Rapoport. 2019. Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. Science 365(6452).

42. Yu, H., T. J. Lupoli, A. Kovach, X. Meng, G. Zhao, C. F. Nathan, and H. Li. 2018. ATP hydrolysis-coupled peptide translocation mechanism of Mycobacterium tuberculosis ClpB. Proc Natl Acad Sci U S A 115(41):E9560-e9569.

43. Deville, C., M. Carroni, K. B. Franke, M. Topf, B. Bukau, A. Mogk, and H. R. Saibil. 2017. Structural pathway of regulated substrate transfer and threading through an Hsp100 disaggregase. Sci Adv 3(8):e1701726.

44. Ho, C. M., J. R. Beck, M. Lai, Y. Cui, D. E. Goldberg, P. F. Egea, and Z. H. Zhou. 2018. Malaria parasite translocon structure and mechanism of effector export. Nature 561(7721):70-75.

45. Deville, C., K. Franke, A. Mogk, B. Bukau, and H. R. Saibil. 2019. Two-Step Activation Mechanism of the ClpB Disaggregase for Sequential Substrate Threading by the Main ATPase Motor. Cell Rep 27(12):3433-3446 e3434.

46. Olivares, A. O., A. R. Nager, O. Iosefson, R. T. Sauer, and T. A. Baker. 2014. Mechanochemical basis of protein degradation by a double-ring AAA+ machine. Nat Struct Mol Biol.

47. Tessarz, P., A. Mogk, and B. Bukau. 2008. Substrate threading through the central pore of the Hsp104 chaperone as a common mechanism for protein disaggregation and prion propagation. Molecular microbiology 68(1):87-97.

48. Weibezahn, J., P. Tessarz, C. Schlieker, R. Zahn, Z. Maglica, S. Lee, H. Zentgraf, E. U. Weber-Ban, D. A. Dougan, F. T. Tsai, A. Mogk, and B. Bukau. 2004. Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. Cell 119(5):653-665.

49. Gottesman, S., E. Roche, Y. Zhou, and R. T. Sauer. 1998. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. Genes Dev 12(9):1338-1347.

50. Maglica, Z., K. Kolygo, and E. Weber-Ban. 2009. Optimal efficiency of ClpAP and ClpXP chaperone-proteases is achieved by architectural symmetry. Structure 17(4):508-516.

51. Bhattacharyya, S., J. P. Renn, H. Yu, J. F. Marko, and A. Matouschek. 2016. An assay for 26S proteasome activity based on fluorescence anisotropy measurements of dye-labeled protein substrates. Analytical biochemistry 509:50-59.

52. Baytshtok, V., T. A. Baker, and R. T. Sauer. 2015. Assaying the kinetics of protein denaturation catalyzed by AAA+ unfolding machines and proteases. Proc Natl Acad Sci U S A 112(17):5377-5382.

53. Brotz-Oesterhelt, H., D. Beyer, H. P. Kroll, R. Endermann, C. Ladel, W. Schroeder, B. Hinzen, S. Raddatz, H. Paulsen, K. Henninger, J. E. Bandow, H. G. Sahl, and H. Labischinski. 2005. Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. Nat Med 11(10):1082-1087. Research Support, Non-U.S. Gov't.

54. Chou, T. F., S. L. Bulfer, C. C. Weihl, K. Li, L. G. Lis, M. A. Walters, F. J. Schoenen, H. J. Lin, R. J. Deshaies, and M. R. Arkin. 2014. Specific inhibition of p97/VCP ATPase and kinetic analysis demonstrate interaction between D1 and D2 ATPase domains. J Mol Biol 426(15):2886-2899.

55. Bukau, B., and A. L. Horwich. 1998. The Hsp70 and Hsp60 chaperone machines. Cell 92(3):351-366.

56. Burton, R. E., T. A. Baker, and R. T. Sauer. 2005. Nucleotide-dependent substrate recognition by the AAA+ HslUV protease. Nat Struct Mol Biol 12(3):245-251.

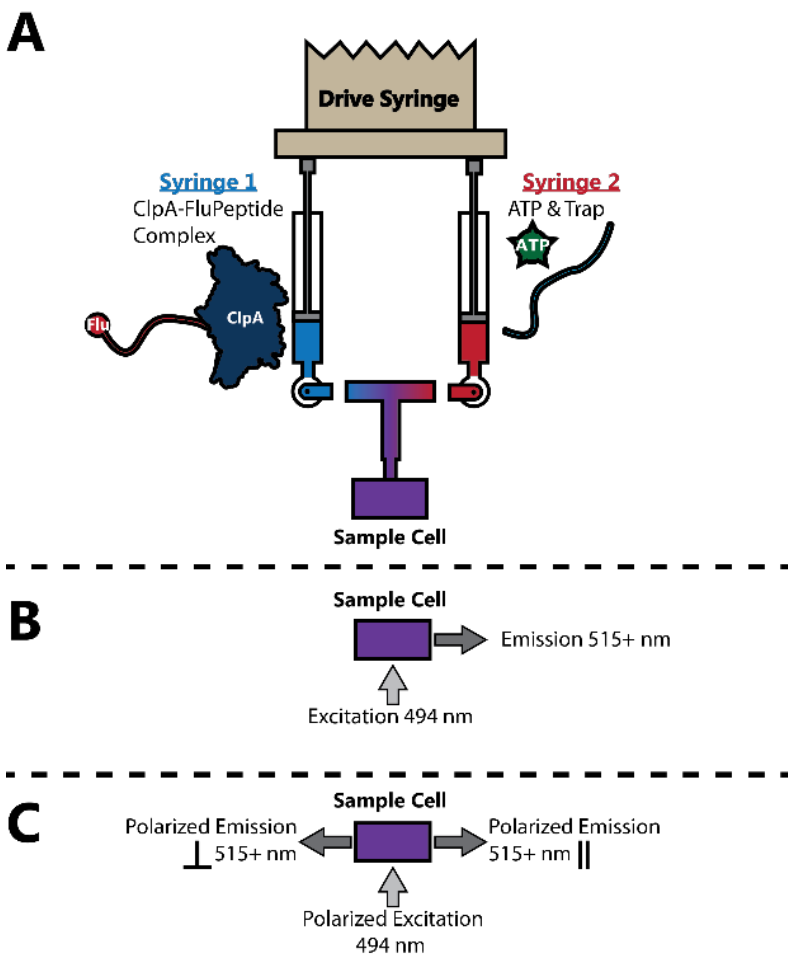


FIGURE 1 Schematic outlining single-turnover translocation experiments in a SX20 stopped-flow spectrometer (Applied Photophysics). (**A**) Syringe 1, shown in blue, contains the pre-assembled ClpA/flu-polypeptide complex. This consists of 2 μM ClpA, 300 μM ATPγS, and 200 nM flu-polypeptide. Syringe 2, shown in red, contains ATP to fuel translocation and 20 μM α-casein to serve as a trap for unbound ClpA, maintaining single-turnover conditions. The two syringes are rapidly mixed into the sample chamber, shown in purple, where the sample is excited, and emission is observed as detailed below. Upon mixing, the final concentration of each reagent in the sample chamber is half of its original in the premixing syringes. (**B**) L-format setup for raw fluorescence. The sample is excited with light at 494 nm. Fluorescence emission of the sample is collected 90° to the incident light through a 515 nm long pass filter with a single photomultiplier tube (PMT) detector. (**C**) T-format setup for fluorescence anisotropy. In this method, the fluorescein sample is excited with vertically polarized light at 494 nm. The fluorescence emission is measured at 90° to the incident light with two PMTs set 180° to one another. One PMT is fitted with a vertical polarizer, detecting emitted light parallel to excitation (∥), and the other a horizontal polarizer that detects light perpendicular to excitation (⊥). Each PMT detector is fitted with a 515 nm long pass filter to observe fluorescein fluorescence.

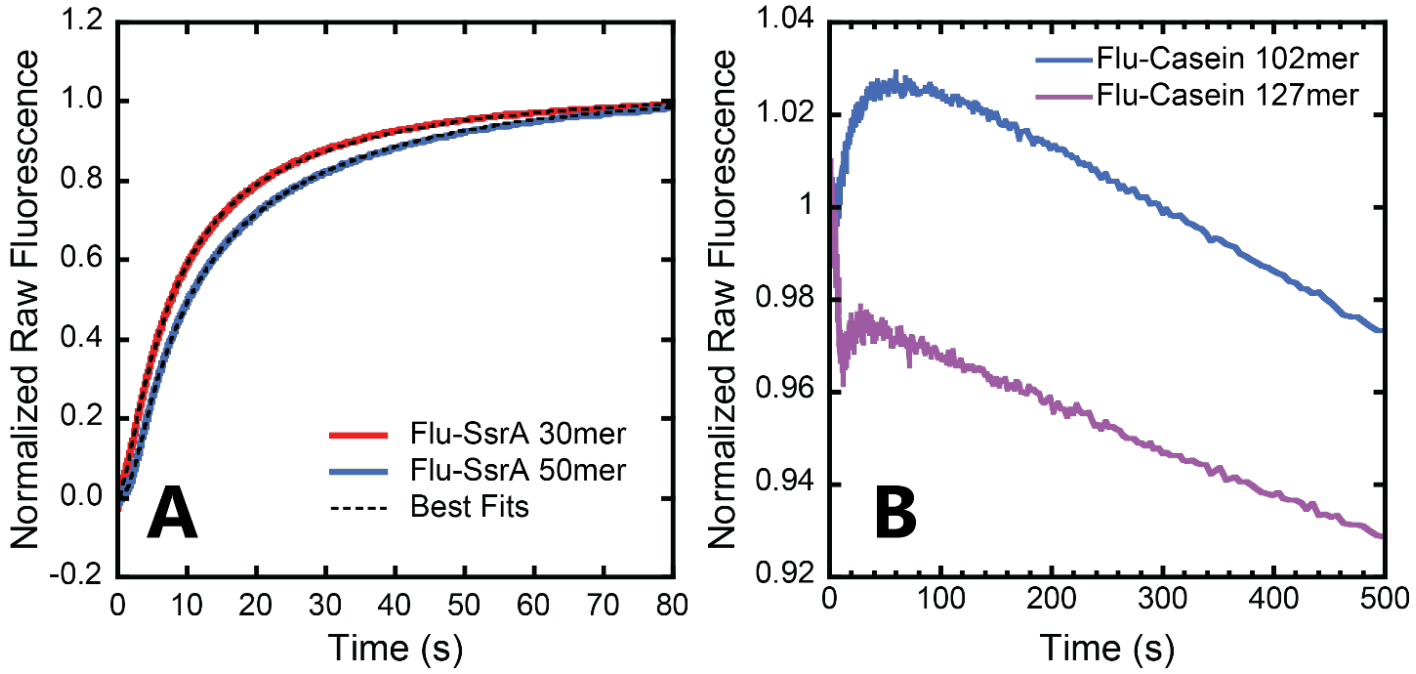


FIGURE 2 Raw fluorescence time courses of ClpA catalyzed translocation of (A) SsrA and (B) αS1-casein substrates collected as described in **Fig. 1**. (A) Single-turnover translocation time courses for SsrA 30mer (red) and 50mer (blue) shown in solid traces with best fits shown in black broken traces. Best fits were generated using MENOTR parameter optimization applied to **Scheme 1** as described in the **Methods**. Analysis was performed globally across the 30, 40, and 50mer SsrA peptides (40mer data not shown). The translocation rate constant (k_T) was found to be (1.5 ± 0.2) s^{-1} and the kinetic step-size (m) was determined to be (14 ± 1) aa step-1. (B) Single-turnover translocation time courses for αS1-casein 102mer (blue) and 127mer (purple) are shown in solid traces. Parameter optimization using Scheme 1 failed to produce any viable best-fit results (not shown).

TABLE 1: Fluorescent Polypeptide Substrates

Name	Length (aa)	Sequence or Source
Flu-SsrA 30mer	30	Flu-CTKSAANLKVKELRSKKKL**AANDENYALAA**
Flu-SsrA 40mer	40	Flu-CTGEVSFQAANTKSAANLKVKELRSKKKL**AANDENYALAA**
Flu-SsrA 50mer	50	Flu-CLILHNKQLGMTGEVSFQAANTKSAANLKVKELRSKKKL**AANDENYALAA**
Flu-Casein 102mer	102	C-terminal 102 amino acids of as 1 casein
Flu-Casein 127mer	127	C-terminal 127 amino acids of as 1 casein

Bolded sequences represent the 11 aa SsrA tag. All sequences are listed N to C with the fluorophore being placed on the N-terminus of each substrate.

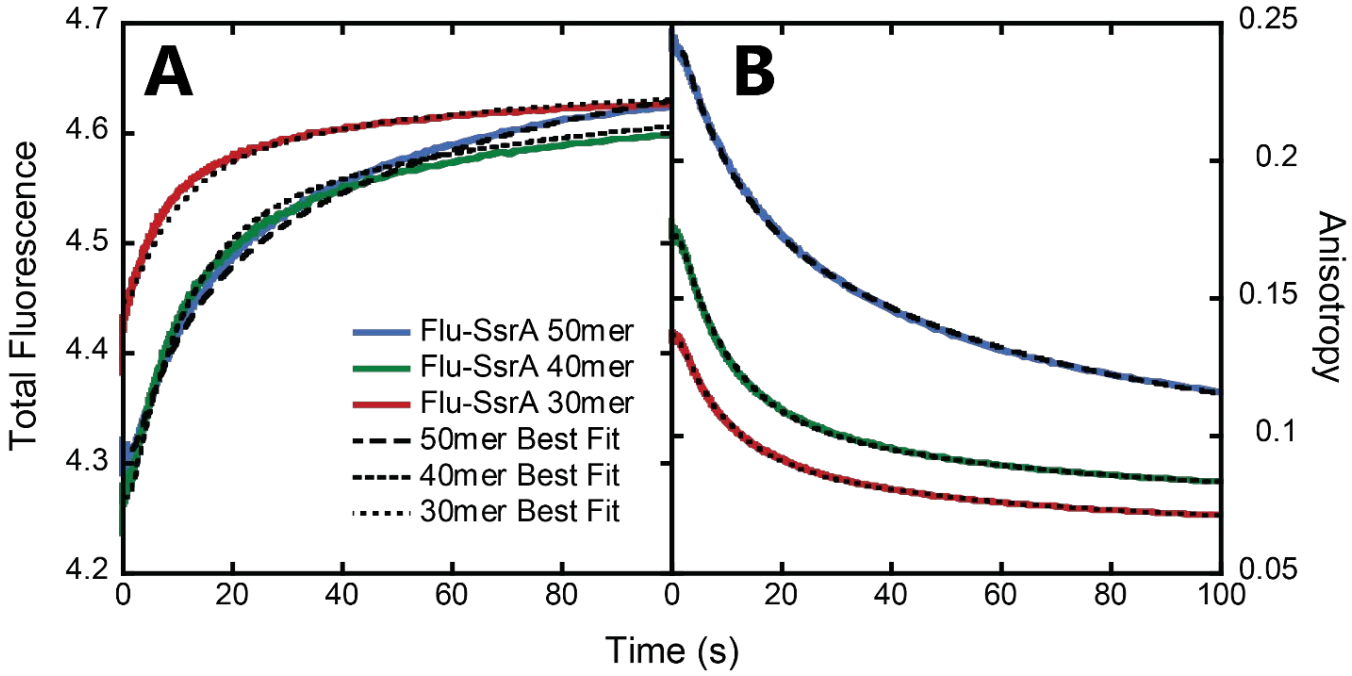


FIGURE 3 (**A**) Total fluorescence and (**B**) fluorescence anisotropy time courses of ClpA catalyzed translocation of 30, 40, and 50mer substrates (See **Table 1**). Solid traces show representative time courses of 30mer (red), 40mer (green), and 50mer (blue) that were collected at 3 mM [ATP] as described in **Fig. 1**. The anisotropy and total fluorescence data were subjected to MENOTR global analysis described in the **Methods** section using **Scheme 1**. The resulting best fits are shown in the overlaid black broken traces. The optimized parameters for this fit are as follows. The translocation rate constant $k_T = (1.8 \pm 0.6)\ s^{-1}$; nonproductive rate constant $k_{NP} = (0.06 \pm 0.01)\ s^{-1}$; slow step rate constant $k_C = (0.2 \pm 0.1)\ s^{-1}$; kinetic step-size m = (14.4 ± 0.1) aa. The time courses were fit globally across total fluorescence and fluorescence anisotropy for all three lengths. The best-fit simulations better describe the anisotropy data compared to that of the total fluorescence. This is due to the relative error on the respective data with total fluorescence having an error, ± 0.1, two orders of magnitude higher than that of anisotropy, ± 0.001 (see **Supplemental Fig. S3-4**).

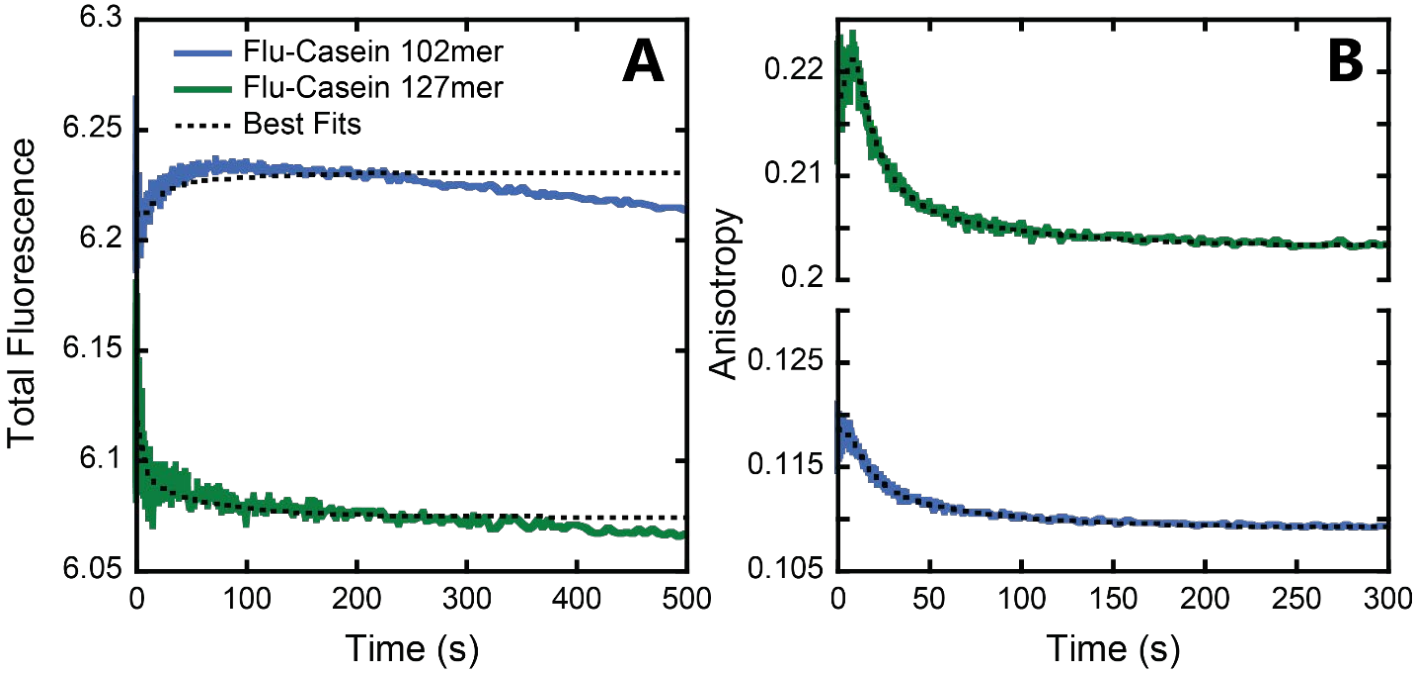


FIGURE 4 (**A**) Total fluorescence and (**B**) fluorescence anisotropy time courses of ClpA catalyzed translocation of αS1-casein 102 and 127mer (See **Table 1**). Solid traces show representative time courses of 102mer (blue) and 127mer (green) that were collected at 3 mM [ATP] as described in **Fig. 1**. The anisotropy and total fluorescence data were subjected to MENOTR global analysis described in the **Methods** section using **Scheme 1**. The resulting best fits are shown in the overlaid black broken traces. The optimized parameters for this fit are as follows. The translocation rate constant k_T = (1.18 ± 0.03) s^{-1}; dissociation rate constant k_d, was not detected; nonproductive rate constant k_{NP} = (0.014 ± 0.003) s^{-1}; slow step rate constant k_C = (0.08 ± 0.02) s^{-1}; kinetic step-size m = (11.9 ± 0.2) aa. The time courses were fit globally across total fluorescence and fluorescence anisotropy and all substrate lengths. The best-fit simulations better describe the anisotropy data compared to that of the total fluorescence. This is due to the relative error on the respective data with total fluorescence having an error, ± 0.1, two orders of magnitude higher than that of anisotropy, ± 0.001 (see **Supplemental Fig. S5-6**).

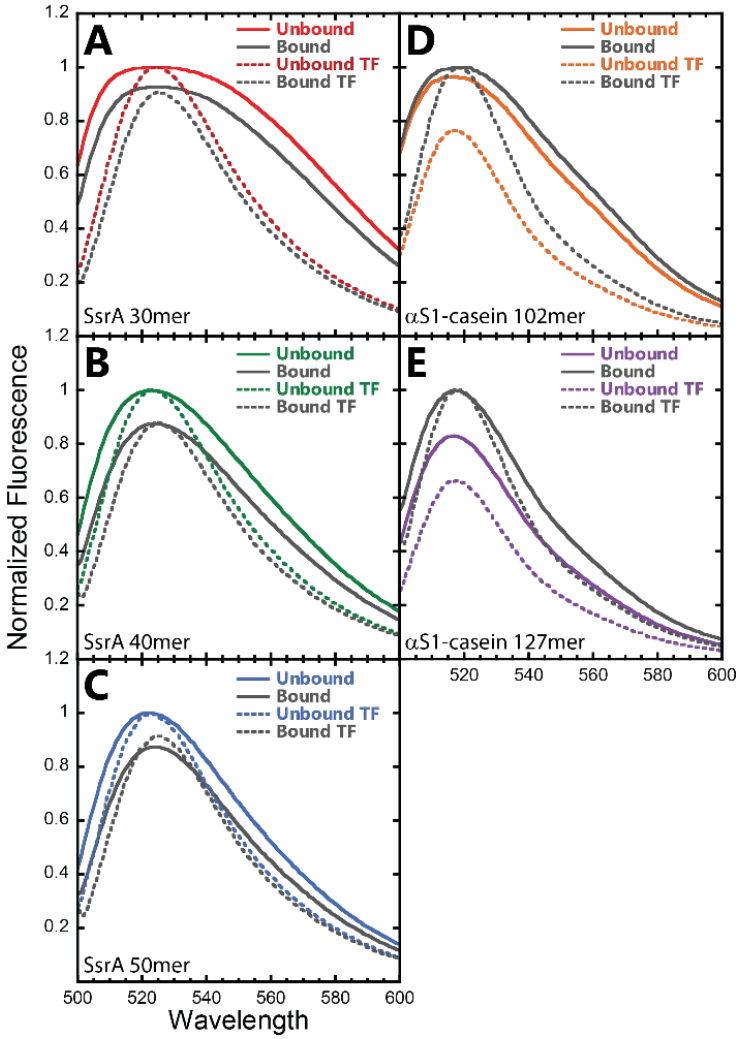


FIGURE S1 Steady-state fluorescence spectra for a set of fluorescein labeled polypeptide substrates. (**A**) SsrA 30mer, (**B**) SsrA 40mer, (**C**) SsrA 50mer, (**D**) αS1-casein 102mer, and (**E**) αS1-casein 127mer; see **Table 1** for the sequences. For each panel, the solid traces represent the steady-state fluorescence collected under raw fluorescence conditions while the broken traces represent the total fluorescence spectra. Total fluorescence was collected by using magic angle conditions in the fluorimeter. This was achieved by polarizing the incident light at 0° and polarizing the detector at 55°. In each panel, the colored traces represent spectra of the 200 nM peptide and 300 μM ATPγS in H300. The gray traces were collected after 200 nM peptide was allowed to form bound complex with 14 μM ClpA in the presence of 300 μM ATPγS in H300. All measurements were collected in a Fluorolog-3 spectrofluorometer (HORIBA Jobin Yvon) by exciting at 494 nm and observing signal at 515 nm with a 5 nm slit width.

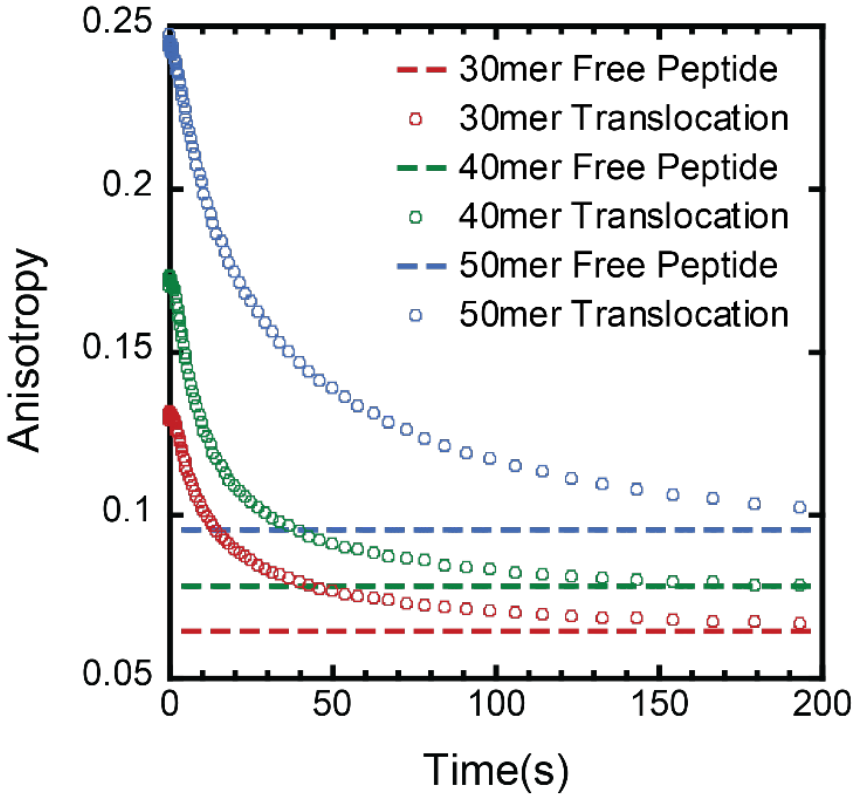


FIGURE S2 Overlay of anisotropy time courses and steady-state anisotropy measurements collected on the 30mer, 40mer, and 50mer polypeptide substrates described in **Table 1**. Time courses shown here are the same as those found in **Fig. 3 B** and collected as described in **Fig. 1 A & C**. Throughout the time courses, anisotropy approaches that of steady-state free peptide, suggesting complete dissociation of peptide from ClpA. Steady-state measurements were collected in the stopped-flow by filling syringe 1 with 200 nM peptide in H300 and rapid mixing it against H300 buffer in syringe 2. The anisotropy value reported in the figure represents the average signal collected over 50 s. All the steady-state measurements exhibited flat time courses over this period.

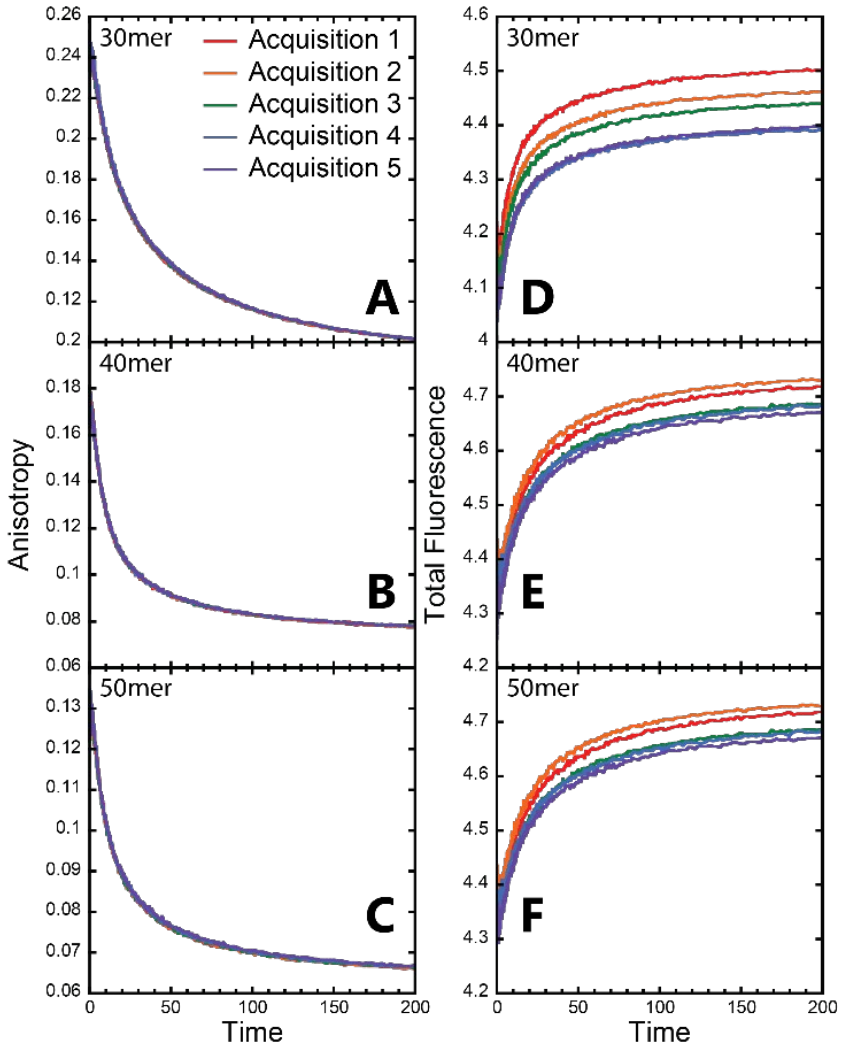


FIGURE S3 Anisotropy (**A - C**) and total fluorescence (**D - F**) time courses of five stopped-flow acquisitions collected on the 30mer, 40mer, and 50mer polypeptide substrates described in **Table 1**. The time courses were collected as described in **Fig. 1 A & C**, where enough sample is loaded in syringe 1 and 2 such that at least 5 acquisitions can be taken on the same sample. The acquisitions were collected back-to-back starting with acquisition 1, and finishing with acquisition 5. The anisotropy time courses for all SsrA substrates overlay well from acquisition to acquisition, while the total fluorescence time courses do not. However, all the total fluorescence time courses exhibit the same shape, and, if the drifts are offset, all the time courses overlay. When the data is averaged, prior to analysis, this drift results in a larger standard deviation in the total fluorescence relative to anisotropy as seen in **Fig. S4**.

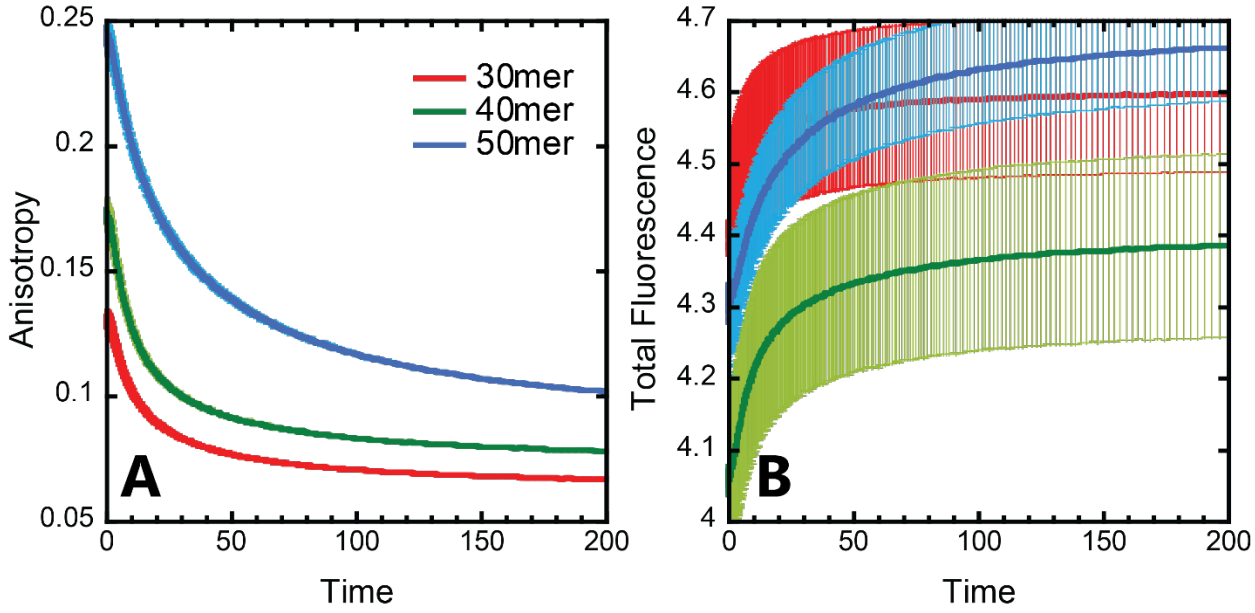


FIGURE S4 (**A**) Anisotropy and (**B**) total fluorescence stopped-flow time courses for 30, 40, and 50mer. These represent the average and standard deviation of all five acquisitions gathered in **Fig. S3** and the data analyzed in **Fig. 3**. Total fluorescence has an average error of ~± 0.1, which is two orders of magnitude higher than that of anisotropy at ~± 0.001.

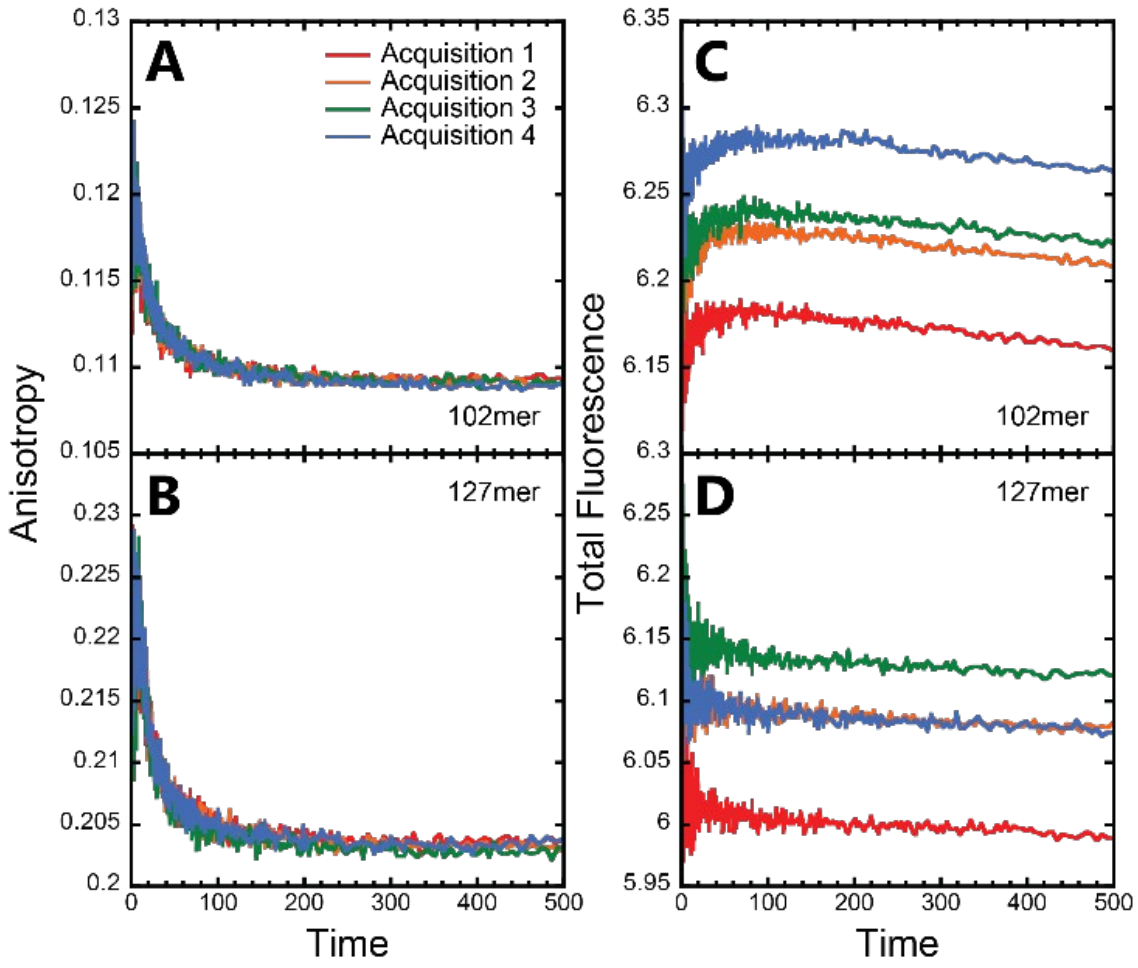


FIGURE S5 Anisotropy (**A & B**) and total fluorescence (**C & D**) time courses of four stopped-flow acquisitions collected on the αS1-casein substrates, **A & C** being 102mer and **B & D** being 127mer, described in **Table 1**. The time courses were collected as described in **Fig. 1A & C** where enough sample is loaded in syringe 1 and 2 so that at least 4 acquisitions can be taken on the same sample. The acquisitions were collected back-to-back starting with acquisition 1, and finishing with acquisition 4. The anisotropy time courses for all αS1-casein peptides overlay well from acquisition to acquisition, while the total fluorescence time courses again do not. However, all the total fluorescence time courses exhibit the same shape, and, if the drifts are offset, all the time courses overlay. When the data is averaged, prior to analysis, this drift results in a larger standard deviation in the total fluorescence relative to anisotropy, as seen in **Fig. S6**.

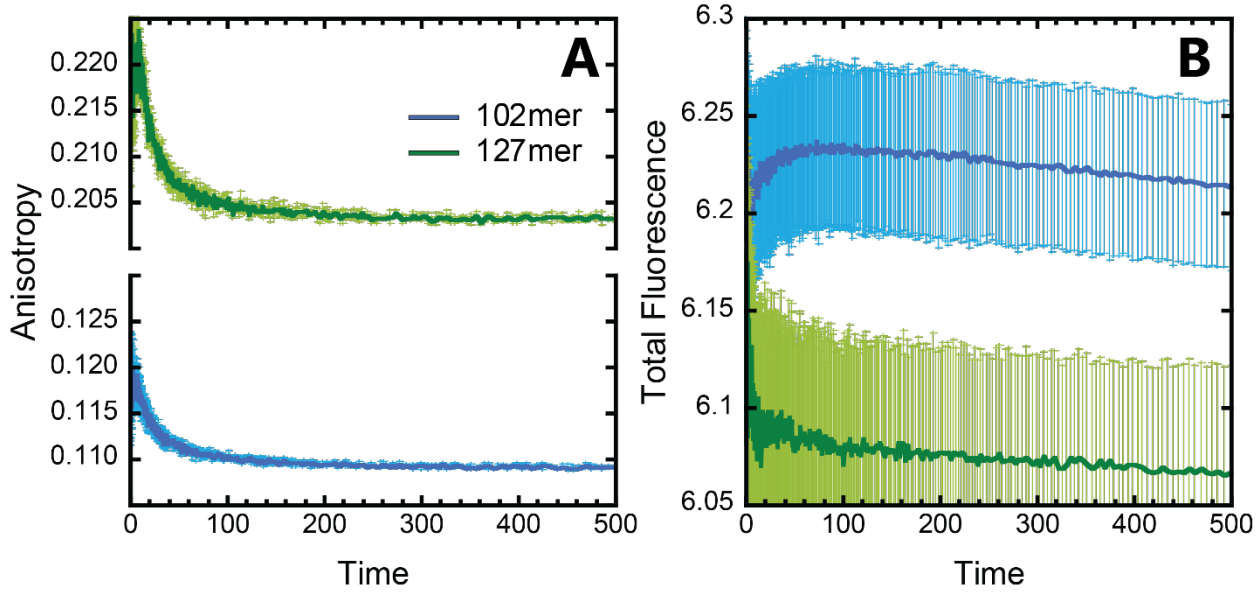


FIGURE S6 (**A**) Anisotropy and (**B**) total fluorescence stopped-flow time courses for αS1-casein 102mer and 127mer. These represent the average and standard deviation of all the acquisitions gathered in **Fig. S5** and the data fit in **Fig. 4**. Total fluorescence has an average error of ~± 0.1, which is two orders of magnitude higher than that of anisotropy at ~± 0.001.

Supplemental Equations

Two-State Model

$$TF(t) = \mathcal{L}^{-1}\left(q_1\left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i\left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^i} + \frac{1-x}{k_{NP}+s}\right) + q_2\left(\frac{1}{s}\left(\frac{k_T^{n-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^n}k_T\right)\right)\right) \quad \text{S1}$$

$$r(t) = \frac{r(t)_{num}}{r(t)_{den}} \quad \text{S2}$$

$$r(t)_{num} = \mathcal{L}^{-1}\left(r_1 q_1\left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i\left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^i} + \frac{1-x}{k_{NP}+s}\right) + r_2 q_2\left(\frac{1}{s}\left(\frac{k_T^{n-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^n}k_T\right)\right)\right) \quad \text{S3}$$

$$r(t)_{den} = \mathcal{L}^{-1}\left(q_1\left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i\left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^i} + \frac{1-x}{k_{NP}+s}\right) + q_2\left(\frac{1}{s}\left(\frac{k_T^{n-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^n}k_T\right)\right)\right) \quad \text{S4}$$

Three-State Model

$$TF(t) = \mathcal{L}^{-1}\left(q_1\left(\frac{1-x}{k_{NP}+s}\right) + q_2\left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i\left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^i}\right) + q_3\left(\frac{1}{s}\left(\frac{k_T^{n-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^n}k_T\right)\right)\right) \quad \text{S5}$$

$$r(t) = \frac{r(t)_{num}}{r(t)_{den}} \tag{S6}$$

$$r(t)_{num} = \mathcal{L}^{-1}\left(r_1 q_1 \left(\frac{1-x}{k_{NP}+s} \right) + r_2 q_2 \left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i \left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1} k_C^h \left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h \left(k_{NP}+s\right)\left(k_T+s\right)^i} \right) + r_3 q_3 \left(\frac{1}{s}\left(\frac{k_T^{n-1} k_C^h \left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h \left(k_{NP}+s\right)\left(k_T+s\right)^n} k_T \right) \right) \right) \tag{S7}$$

$$r(t)_{den} = \mathcal{L}^{-1}\left(q_1 \left(\frac{1-x}{k_{NP}+s} \right) + q_2 \left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i \left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1} k_C^h \left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h \left(k_{NP}+s\right)\left(k_T+s\right)^i} \right) + q_3 \left(\frac{1}{s}\left(\frac{k_T^{n-1} k_C^h \left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h \left(k_{NP}+s\right)\left(k_T+s\right)^n} k_T \right) \right) \right) \tag{S8}$$

CHAPTER 3

MECHANISTIC EXAMINATION OF POLYPEPTIDE TRANSLOCATION CATALYZED BY CLPA'S ATPASE DOMAINS 1 & 2

NATHANIEL W. SCULL
AARON L. LUCIUS

Introduction

ClpA is an Hsp100/Clp protein and member of the AAA+ (ATPases Associated with various cellular Activities) superfamily of molecular motors. ATP-dependent molecular motors like ClpA are crucial for many cellular pathways including DNA replication (1), cellular proteostasis (2-4), cell-cycle control (5, 6), signal transduction (7), and many others (2, 8, 9). In these pathways, the AAA+ proteins couple the energy released from nucleotide triphosphate binding and hydrolysis to physical movements of substrate lattices. In particular, ClpA has been shown to be implicated in both protein remodeling (10) and proteolysis reactions in *Escherichia coli* (*E. coli*) (11).

AAA+ motors can be subclassified as either Class I or Class II based on the number of nucleotide binding domains (NBDs) they contain (2). Class I AAA+ proteins contain two NBDs per monomer and include motors such as ClpA (12), ClpB (13), and Hsp104 (14). In contrast, Class II AAA+ proteins contain a single NBD per monomer and include the motors ClpX (15), HslU (16), FtsH (17), and others. Regardless of Class, all AAA+ NDBs contain conserved Walker A and B motifs that are responsible for nucleotide triphosphate binding and hydrolysis, respectively (18). It is known that ATP binding to the NBDs drives oligomerization of the motors and that repeated rounds of ATP binding and hydrolysis drives their mechanical movements.

Structurally, ClpA contains three domains: an N-terminal domain (colored purple in **Fig. 1 D**) and two AAA+ NBDs (D1 and D2) shown in **Fig. 1 C & D**, colored in blue

and green, respectively. In ClpA, the primary sequence and structure of D1 and D2 are generally non-homologous except for the conserved Walker motifs (12). In the presence of ATP, or ATP analogs, ClpA oligomerizes into a two-tiered hexameric ring with a hollow central channel as illustrated in **Fig. 1 A-C**. Each tier consists of six identical D1 or D2 domains, with each monomer contributing one D1 and D2 per tier (19, 20). Cryo-EM structures of hexameric ClpA and a crystal structure of ClpA monomer are shown in **Fig. 1 A-C & D**, respectively.

Hexameric ClpA couples the binding and hydrolysis of ATP to induce conformational changes in a flexible Helix-2 insertion, known as the pore-loop, located between the Walker A and B motifs of each domain (21, 22). These 12 pore-loops extend into the central channel of the hexameric ring, and have been shown to contact ClpA's polypeptide substrate (12, 23, 24). Evidence from the investigations of ClpA and other AAA+ motors suggests that these loops alternate between up and down conformations, thus exerting tugging and pulling forces on the substrate, resulting in translocation through the central channel and unfolding of the substrate (23, 25-28).

In the absence of other partner proteins, ClpA acts as a protein remodeling system that unfolds targeted protein substrates. However, ClpA also functions as a molecular chaperone in the ATP-dependent protease system, ClpAP. ClpAP translocates substrates through ClpA's central channel directly into the sequestered proteolytic chamber of ClpP for degradation. Despite the fact that both activities involve the same general polypeptide unfolding and translocation by ClpA, multiple groups have shown that ClpAP translo-

cates with an overall faster rate compared to ClpA alone (29-32). The prevailing hypothesis is that each NBD of ClpA has a unique translocation mechanism and that the presence of ClpP allosterically regulates those activities.

Here we set out to distinguish and quantify the unique contributions each NBD makes to ClpA's translocation activity in the absence of ClpP. Specifically; we will define the kinetic mechanism of polypeptide translocation catalyzed by each NBD. In our previous work, we have proposed a mechanistic model of ClpA catalyzed translocation in the absence of ClpP (31, 33). In our model, both NBDs start prebound to the polypeptide substrate, with ATP bound to each domain. D1 starts by hydrolyzing ATP and translocating the substrate into the central cavity of ClpA. The hydrolysis of ATP to ADP and P_i at D1 reduces its affinity for substrate (22), thus allowing D1 to release the substrate, reset, and subsequently rebind both the substrate at a new position and a new ATP. D2 would then hydrolyze ATP and translocate the substrate out of the central cavity in a similar manner. From our previous work, we have proposed that D1 translocates polypeptide with a step size of ~14 amino acids (aa) while D2 has a step size of ~ 5 aa (31, 34). The ratio of these step sizes would cause the peptide to initially crimp within the ClpA cavity as D1 translocates. Thus, D2 would need to go through multiple rounds of translocation to remove the substrate prior to D1 taking another step. While this hypothesis is based on known structural data and various biophysical studies, no direct quantitative measurement of the kinetic mechanisms of translocation at each domain has been determined.

In order to test these hypotheses, ClpA variants that are sensitive to translocation at only one of the two NBDs will be examined using transient kinetic techniques. Muta-

tions in the Walker motifs of other AAA+ motors have been successfully used to investigate activities of a single NBD at a time (35-37). The P-loop of the Walker A motif contacts the phosphates of bound ATP, and its mutation has been shown to eliminate ATP binding, rendering the AAA+ protein inactive and perturbing oligomerization of the motor (8, 20, 38). The Walker B motif also contacts ATP creating an enzymatic pocket that coordinates Mg^{2+} and activates water for ATP hydrolysis. It has been shown that the mutation of the conserved Walker B glutamate, which activates the water in hydrolysis, arrests ATP hydrolysis but not binding (8, 20, 39). Walker B mutations have been used by others in steady-state ClpA ATPase (39) and unfoldase (32) assays, and has been shown to have minimal and known effects on ClpA hexamerization (40, 41).

Kress et al. have designed and overexpressed the Walker B variants $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$ that lack hydrolysis activity in D1, D2, or both domains, respectively (39). The variants were used to investigate the ATPase activity of each NBD. They reported, and we have also confirmed, that $ClpA_{E286A/E565A}$ binds substrate but is unable to hydrolyze ATP (40). The ATPase activity for D2 was reported as (410 ± 33) min^{-1} while D1 was reported as (49 ± 9) min^{-1}. This can be compared to the wildtype (WT) activity reported at (479 ± 38) min^{-1}. They concluded that the observed ATPase activity of D1 and D2 adds up to (96 ± 7) percent of wildtype activity; suggesting that ATP hydrolysis in each domain is independent. Moreover, they demonstrated that D2 contributes the most to overall ATPase activity, at 86 % of WT, while D1 only makes up 10 % of observed WT activity.

Another ClpA study conducted by Baytshtok et al. reported on the steady-state unfolding activity of model substrates catalyzed by similar Walker B mutants, $ClpA_{E286Q}$

and $ClpA_{E565Q}$ (32). Their results indicated that only the D2 active variant, $ClpA_{E286Q}$, exhibited unfolding activity greater than their control reactions. They reported an eightfold reduction in steady-state unfolding for D2, which is dramatically lower than the predicted 1.2-fold decrease estimated from the ATPase activity reported by Kress et al. From this, Baytshtok et al. concluded that D1 makes important contributions to substrate unfolding, but its activity is not required. Thus, D2 is predicted to contribute more to the overall unfolding activity of ClpA.

Here we report the results of single-turnover stopped-flow assays monitoring the translocation of polypeptide catalyzed by D1 and D2 of ClpA using new fluorescence anisotropy and total fluorescence methods. Specifically, we were able to determine quantitative measurements of the elementary rate constants, kinetic step-sizes, and macroscopic rates of translocation governing both domains. From this, we were able to conclude that D2 translocates polypeptide substrate with a rate of ~16 aa s^{-1}, while D1 translocates with a rate of ~0.35 aa s^{-1} at saturating ATP and 500 mM NaCl conditions. This is in contrast to the rate of ~23 aa s^{-1} determined for WT. We show that changes in the rate for both variants are the consequence of decreases in the kinetic rate constant while the kinetic step-size were found to remain approximately the same at ~10-20 aa. Additionally, we concluded that the dependence of the kinetic rate constant on [ATP] determined for both variants suggests inter-monomer cooperativity between adjacent (D1 & D1) or (D2 & D2) domains in the hexamer. With these results, we propose a revised molecular model for polypeptide translocation catalyzed by ClpA.

Materials and Methods

Solutions

All solutions were prepared in double distilled water from a Siemens Water Technology, Purelab Ultra Genetic system (Alpharetta, GA) with commercially available reagent grade chemicals. All experiments were carried out in buffer H500: 25 mM HEPES pH 7.5 at 25 °C, 500 mM NaCl, 20 mM $MgCl_2$, 2 mM 2-mercaptoethanol, and 10 % v/v glycerol.

Peptide substrates

Polypeptide substrates were synthesized and fluorescently labeled by CPC Scientific (Sunnyvale, CA) and were certified as >90 % pure by LC-MS analysis. The substrates consist of a C-terminal SsrA tag followed by truncations of the Titin I 27 domain. Each substrate has an N-terminal cysteine that has been labeled with a fluorescein dye. The substrates have been previously analyzed using CD spectroscopy to show that they lack any significant structure and bind to ClpA (42). Full sequences of each substrate can be found in **Table 1**.

ClpA

Genes for the ClpA variants ($clpA_{E286A}$, $clpA_{E565A,}$ and $clpA_{E286A/E565A}$) were cloned into pET30a vectors under a T7 promoter system, containing kanamycin resistance, as described in Duran et al. 2018 (40). Prior to expression of each variant, the vectors were transformed into $ClpA_{WT}$ knockout BL21(DE3) *E.coli* cells, $\Delta ClpA_{WT}$-BL21(DE3), that were constructed using recombineering methods (43). The genomic $ClpA_{WT}$ gene was replaced with an ampicillin cassette, allowing for final cell selection using ampicillin and

kanamycin in LB media. All ClpA variants were purified in a manner similar to that described in Veronese 2009 Biochemistry (44). The only significant difference was the omission of the final Blue Sepharose FF column. The reported concentrations of ClpA were determined in H500 buffer using the molar extinction coefficient 31,000 (M monomer)$^{-1}$cm^{-1} and as all concentrations are in monomer units.

Polypeptide translocation experiments

Polypeptide translocation experiments were carried out on an Applied Photophysics SX20 stopped-flow spectrometer (Leatherhead, UK). First, 4 μM ClpA was incubated with 300 μM ATPγS in H500 for 30 min at 25 °C. Then 200 nM fluorescent SsrA substrate was added and incubated for 15 min to achieve binding equilibrium. Equilibrium was assessed by monitoring fluorescence change in a HORIBA Jobin Yvon Fluorolog-3 spectrofluorometer (Albany, NY) over two hours. No change was observed after 15 minutes. Additionally, a given [ATP] and 20 μM α-casein were incubated in H500 for 45 min at 25 °C. ATP concentrations were varied across experiments and the final concentration is indicated in the text.

Fluorescence anisotropy data for polypeptide translocation was collected using a T-format. Prior to each set of experimental acquisitions, the stopped flow is prepared by adjusting the PMT voltages and determining the G factor. A single acquisition is taken, and the chemistry is allowed to react. Once the steady state of the final reaction conditions are met, the PMT voltages are set to 15 % and the G-factor is obtained to a 10 % level of precision as described by the instrument manual. The stopped-flow sampled chamber, flow lines and syringes are washed with buffer matching that of reaction conditions between each experiment.

MENOTR

All data analysis was carried out in the MATLAB (MathWorks, Natick, MA) toolbox MENOTR. MENOTR, Multi-start Evolutionary Nonlinear OpTimizeR, is a hybrid multi-start genetic and NLLS algorithm. This hybrid fitting method utilizes both a genetic and nonlinear least squares (NLLS) algorithm. Genetic algorithms can survey large and diverse error spaces to find quality solutions. However, due to lack of convergence, the genetic algorithm becomes computationally expensive and underperforms in the final stages of optimization. In contrast, the NLLS algorithm converges quickly on a local minimum, but cannot guarantee a global solution due to its inability to escape local minimums or survey sufficient solution space. In this hybrid method, the genetic algorithm will quickly find initial solutions for the slower running NLLS algorithm and simultaneously assist in the escape from local minima, while the NLLS algorithm guarantees convergence on a final solution. The collaboration between the approaches yields an overall more robust and superior method of generating high-quality solutions to fitting problems compared to either method alone. More information on this algorithm can be found in (Ingram, Scull. 2020, In Preparation). MENOTR can use the model equations discussed in this manuscript to find a set of optimized parameters that best describe each data set. This is discussed in detail in the below sections.

Analysis of total fluorescence and fluorescence anisotropy

When collecting fluorescence anisotropy experiments, two types of signals are simultaneously gathered: total fluorescence (*TF*) and fluorescence anisotropy (*r*). When exciting the sample with vertically polarized light each signal is described by **Eq. 1 & 2**.

$$TF = G \cdot I_{\parallel} + 2 \cdot I_{\perp} \quad (1)$$

$$r = \frac{(G \cdot I_{\parallel} + I_{\perp})}{(G \cdot I_{\parallel} + 2 \cdot I_{\perp})} \quad (2)$$

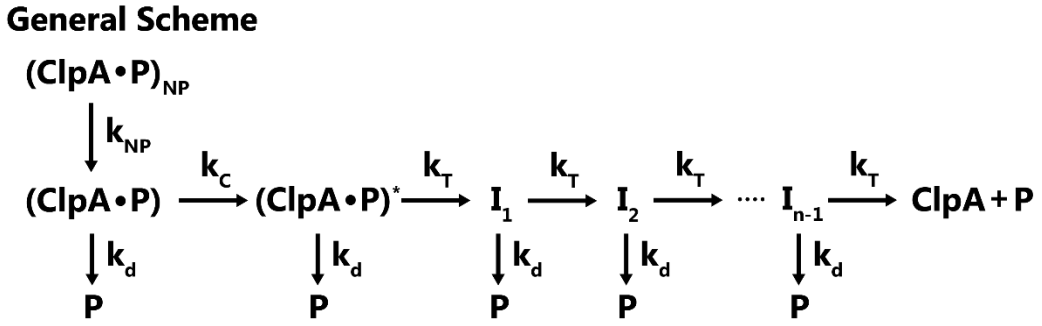
Where $I_{\parallel}$ and $I_{\perp}$ are the measured fluorescence intensities detected parallel and perpendicular to the incident polarization (See **Fig. 2**). The G-factor, G, is the ratio of the sensitivity of the two detectors. The G factor accounts for any differences in the optical paths of the two detectors.

In order to apply **Eq. 1 & 2** to the kinetic time courses collected for ClpA catalyzed polypeptide translocation they need to be in a form that can be used to describe all the kinetic states that have unique total fluorescence values (quantum yields) and unique anisotropy values as a function of time. Total fluorescence as a function of time can be represented by the summation of all of the states, y_i, multiplied by their respective quantum yield, q_i, given by **Eq. 3**. Similarly, anisotropy is the summation of all the states multiplied by both their respective quantum yield and the anisotropy, r_i. However, as can be seen in **Eq. 2**, anisotropy is normalized by total fluorescence. Meaning, an additional term in the time dependent anisotropy function that accounts for this normalization is required. By dividing the anisotropy summation by the total fluorescence summation, the same normalization can be achieved in the time dependent function as shown in **Eq. 4**.

$$TF(t) = \sum y_i(t) \cdot q_i \quad (3)$$

$$TF(t) = \frac{\sum y_i(t) \cdot q_i \cdot r_i}{\sum y_i(t) \cdot q_i} \quad (4)$$

The task at hand is to determine time dependent functions, $y_i(t)$, for each species in the reaction. We have previously reported that **Scheme 1** describes ClpA catalyzed polypeptide translocation (34). From this scheme, a system of coupled differential equations can be constructed that describe each species in the mechanism. Previously we have demonstrated that Laplace transforms can be applied to solve these systems of differential equations (45). The transformation of the system of differential equations into Laplace space results in a much simpler system of coupled algebraic equations, $y_i(s)$. This not only simplifies the system but more importantly allows us to construct a model within Laplace space where the number of steps, n, is a fitting parameter of interest. This is essential because it is not initially known how many steps ClpA is taking on a given substrate lattice. During data analysis, it becomes necessary to find the inverse Laplace transform of $y_i(s)$, which can be numerically solved by MENOTR to yield a solution in the time domain.



(**Scheme 1**)

Scheme 1 depicts the reaction mechanism used to describe ClpA catalyzed translocation of polypeptide substrate. Before mixing with ATP, ClpA is bound to fluorescent substrate in two forms: nonproductive ClpA complex $(ClpA•P)_{NP}$ and productive ClpA

complex (ClpA•P). Upon mixing with ATP, $(ClpA•P)_{NP}$ slowly isomerizes with rate constant k_{NP} to form (ClpA•P). Since these two states are present before rapid mixing with ATP, the relative populations of the two states is defined by the fraction of productively bound complexes, x, where x is the ratio of productive complex to total complex (**Eq. 5**).

$$x = \frac{[ClpA•P]}{[ClpA•P] + [ClpA•P]_{NP}} \quad 5$$

Translocation begins and some number, n-1, of intermediate steps, I, are taken governed by the rate constant k_T. With the n^{th} translocation step, the ClpA polypeptide complex will dissociate from its substrate. Throughout the scheme it is possible for polypeptide, P, to dissociate from ClpA prior to the completion of translocation, these steps are described by the rate constant k_d. Additionally, there can be a slow conformational step to $(\mathrm{ClpA} \bullet \mathrm{S})^*$ described by the rate constant k_C that repeats h number of times. The need for this additional k_C step is discussed in the **Results** for each ClpA variant.

When fitting translocation data globally across multiple peptide substrates the kinetic parameter m, step-size, is used in place of the number of steps, n. This is done by relating the number of translocation steps in each experiment to their respective substrate length, L. This relationship is described by **Eq. 6**.

$$m\ (\text{aa step}^{-1}) = \frac{L\ (\text{aa})}{n\ (\text{step})} \quad 6$$

Model equations used to describe $ClpA_{WT}$ and $ClpA_{E286A}$ catalyzed translocation time courses

With a set of equations that describe all the reaction species, $y_i(t)$, the question now becomes, how many species exist within the reaction pathway that have unique quantum yield and anisotropy values? In the simplest case, two states of species exist during translocation. Here the fluorescence anisotropy and total fluorescence measurements are sensitive to only fluorescently modified peptide. Additionally, it is known that bound and free peptide have unique anisotropy and total fluorescence values. Thus, it is assumed that all bound species share the same q and r values while free peptide has its own. The system of coupled differential equations that result from **Scheme 1** were solved to obtain expressions describing each species as a function of the Laplace variable, S(s). The expressions were grouped together as either bound or unbound and multiplied by their respective quantum yield and anisotropy values creating the two-state model equations for total fluorescence (**Eq. 7**) and anisotropy (**Eq. 8**):

$$TF(t) = \mathcal{L}^{-1}\sum y_i(s)\cdot q_i = \mathcal{L}^{-1}\Big(q_1\big([ClpA{\bullet}P]_{NP}(s) + [ClpA{\bullet}P](s) + [ClpA{\bullet}P]^*(s) + I_1(s) + I_2(s) + I_{n-1}(s)\big) + q_2\big(P(s)\big)\Big) \quad (7)$$

$$r(t) = \frac{\mathcal{L}^{-1}\left(\sum y_i(s)\cdot q_i\cdot r_i\right)_i}{\mathcal{L}^{-1}\left(\sum y_i(s)\cdot q_i\right)} = \frac{\mathcal{L}^{-1}\Big(r_1q_1\big([ClpA{\bullet}P]_{NP}(s) + [ClpA{\bullet}P](s) + [ClpA{\bullet}P]^*(s) + I_1(s) + I_2(s) + I_{n-1}(s)\big) + r_2q_2\big(P(s)\big)\Big)}{\mathcal{L}^{-1}\Big(q_1\big([ClpA{\bullet}P]_{NP}(s) + [ClpA{\bullet}P](s) + [ClpA{\bullet}P]^*(s) + I_1(s) + I_2(s) + I_{n-1}(s)\big) + q_2\big(P(s)\big)\Big)} \quad (8)$$

During analysis, MENOTR solves the model equations by numerically approximating the inverse Laplace transform of both equations. Note that for anisotropy, inverse Laplace transforms must be performed on the numerator and denominator separately prior to normalization. The transforms yield solutions in the time domain that represent the total fluorescence and anisotropy signals. For full expansions of the time dependent fitting functions, see supplemental equations **Eq. S1-S4.**

Model equations used to describe $ClpA_{E565A}$ catalyzed translocation time courses

The time courses of $ClpA_{E565A}$ catalyzed translocation exhibited an additional phase that could not be described by the two-state model used in the analysis of $ClpA_{WT}$ and $ClpA_{E286A}$ and would require at least three states. It is likely still the case that free substrate can be described by a single quantum yield and a single anisotropy value. Thus the differences in the observed signal is likely due either to the pre-bound states ($(ClpA•P)_{NP}$ and (ClpA•P)) or the various translocation intermediates, I_i, exhibiting different quantum yields and anisotropies. To test this idea, we derived a model where we assume that the initial pre-bound ClpA states exhibits a quantum yield, q_1, and anisotropy, r_1 whereas the summation of all the translocation intermediates exhibit a different quantum yield, q_2, and anisotropy, r_2. Finally, free peptide exhibits a quantum yield and anisotropy, q_3 and r_3, respectively. Applying this to **Scheme 1** we were able to derive a three-state model for total fluorescence (**Eq. 9**) and anisotropy (**Eq. 10**). Full expansions of **Eq. 9 & 10** in the time domain can be found in **Eq. S5 - S8**.

$$TF(t) = \mathcal{L}^{-1}\left(\sum y_i(s)\cdot q_i\right) = \mathcal{L}^{-1}\left(q_1\left([ClpA•S]_{NP}(s) + [ClpA•S](s)\right) + q_2\left([ClpA•S]^*(s) + I_1(s) + I_2(s) + I_{n-1}(s)\right) + q_3\left(P(s)\right)\right) \quad (9)$$

$$r(t) = \frac{\mathcal{L}^{-1}\left(\sum y_i(s)\cdot q_i \cdot r_i\right)}{\mathcal{L}^{-1}\left(\sum y_i(s)\cdot q_i\right)} = \frac{\begin{array}{l}\mathcal{L}^{-1}\Big(r_1 q_1\left([ClpA{\bullet}S]_{NP}(s)+[ClpA{\bullet}S](s)\right)+ \\ \quad r_2 q_2\left([ClpA{\bullet}S]^{*}(s)+I_1(s)+I_2(s)+I_{n-1}(s)\right)+r_3 q_3\left(P(s)\right)\end{array}}{\begin{array}{l}\mathcal{L}^{-1}\Big({}_1 q_1\left([ClpA{\bullet}S]_{NP}(s)+[ClpA{\bullet}S](s)\right)+ \\ \quad q_2\left([ClpA{\bullet}S]^{*}(s)+I_1(s)+I_2(s)+I_{n-1}(s)\right)+q_3\left(P(s)\right)\end{array}} \quad (10)$$

Analysis strategy

The concurrent analysis of both total fluorescence and anisotropy requires that the two data sets be linked. This is accomplished by globally constraining the parameters with respect to the total fluorescence and anisotropy. Meaning, that within an experiment all parameters are forced to be the same for both the total fluorescence and anisotropy data sets. Using **Fig. 3** as an example, the kinetic parameters used to describe the 30mer data (red) in total fluorescence (**Fig. 3 A**) are the same as those used to describe 30mer (red) in anisotropy (**Fig. 3 B**). However, while parameters like the translocation rate constant, kinetic step size, and other rate constants are global across all three substrate experiments in **Fig. 3**, parameters like quantum yield and steady state anisotropy values are local within each substrate experiment. Meaning that each substrate has a unique set of q and r values associated with it. To that end, the parameter optimization strategy was designed to constrain the quantum yield and steady state anisotropy parameters to be local with respect to the substrate experiment while being global with respect to the two data types. On the other hand, the kinetic parameters are global with respect to both the substrate experiments and the two data sets.

Analysis of the [ATP] dependence of the translocation rate and rate constant

The ATP dependence of the kinetic parameters k_T and mk_T for $ClpA_{WT}$ (**Fig. 3 C & D**), $ClpA_{E286A}$ (**Fig. 4 C & D**), and $ClpA_{E565A}$ (**Fig. 5 C & D**) were subjected to NLLS analysis using Synergy Software KaleidaGraph (Reading, PA) and a Hill model containing a y intercept, b, **Eq. 11**.

$$y = \frac{(y_{\max})(K_a[ATP])^H}{1+(K_a[ATP])^H} + b \qquad 11$$

Where y is either k_T or mk_T and y_{max} is the maximum translocation rate constant or translocation rate, respectively. K_a is the association constant, and H is the hill coefficient.

Results

To interrogate the kinetic mechanism of polypeptide translocation catalyzed by $ClpA_{WT}$, $ClpA_{E286A}$, and $ClpA_{E565A}$, we performed single-turnover fluorescence anisotropy stopped-flow experiments as schematized in **Fig. 1**. To determine the kinetic step-size for each variant, the polypeptide chain length is varied. The three substrate lengths used in this study are shown in **Table 1**. Each substrate contains a C-terminal 11 aa SsrA binding motif that ClpA is known to recognize, as well as an N-terminal cysteine that has been labeled with fluorescein (42, 46, 47).

In these experiments, 4 μM ClpA monomer, 300 μM ATPγS, and 200 nM fluorescent substrate are combined and loaded into syringe 1 as described in the **Methods, Fig. 2,** and previously (31, 34). Here the presence of ATPγS, a slowly hydrolysable ATP analogue, is required to form ClpA hexamers that are competent in binding polypeptide

substrate. Hexameric ClpA binds the C-terminal SsrA sequence of the polypeptide, forming a prebound complex, which is poised to initiate translocation upon introduction of ATP. Syringe 2 is loaded with a solution containing ATP and 20 μM non-fluorescent polypeptide, α-casein. The non-fluorescent polypeptide is included in large excess over the fluorescently modified substrate to serve as a trap for any ClpA that is free in solution or that dissociates from the polypeptide chain during or after translocation, thereby maintaining single-turnover conditions.

Upon rapid mixing of the two syringes, fluorescein is excited with vertically polarized light at 494 nm. During the time course, emission is collected in a T-format with two photomultiplier tube (PMT) detectors set 90° to the incident light and 180° to one another as illustrated in **Fig. 2**. One detector collects vertically polarized emitted light while the other collects horizontal emissions. Each PMT is fit with a 515 nm long pass filter to block any excitation light.

Examination of polypeptide translocation catalyzed by $ClpA_{WT}$ in H500.

Here we seek to compare the polypeptide translocation mechanisms for $ClpA_{WT}$, $ClpA_{E286A}$, and $ClpA_{E565A}$ ClpA to determine the roles of the D1 and D2 ATPase sites in translocation. It was previously found that the $ClpA_{E286A}$, $ClpA_{E565A}$, and the double mutant, $ClpA_{E286A/E565A}$, all exhibited reduced solubility in the 300 mM NaCl buffer conditions previously used to examine WT. All three variants exhibit higher solubility in buffer containing 500 mM NaCl compared to the previous concentration. An analysis of the assembly state of these variants in the higher salt buffer conditions has been previously conducted (40, 41). However, all previous kinetic examinations of WT catalyzed polypeptide translocation were performed in 300 mM NaCl (31, 34). Thus, for comparison

purposes, single-turnover translocation experiments using WT in H500 as described in the **Methods** and **Fig. 1** were performed. Here we used fluorescence anisotropy, which includes the simultaneous observation of total fluorescence and anisotropy, to monitor translocation. This strategy overcomes some of the limitations of the previous raw fluorescence strategy (Scull 2020, In Preparation).

Representative total fluorescence and fluorescence anisotropy time courses for the 30, 40, and 50 aa substrates are shown in **Fig. 3 A & B**, respectively. Fluorescein fluorescence is quenched when ClpA binds the fluorescently modified polypeptide substrate (31, 34). Consequently, the kinetic time courses shown in **Fig. 3 A** exhibit low initial total fluorescence followed by an increase over time as ClpA dissociates. Conversely, the anisotropy of the ClpA-substrate complex is expected to exhibit a higher anisotropy compared to free substrate (42). As predicted, the time courses in **Fig. 3 B** exhibit an initial high anisotropy followed by a decrease as ClpA dissociates over time.

These experiments are single turnover with respect to the polypeptide substrate, which means that ClpA cannot rebind the fluorescently modified substrate after the first round of translocation. Thus, the kinetic time courses represent a single round of substrate translocation followed by dissociation. All the kinetic time courses shown in **Fig. 3 A & B** exhibit a constant signal, or lag, prior to the observed change in signal. This extent of lag is observed to increase with increasing substrate length as seen across all three polypeptide substrates. We have previously interpreted this dependence on substrate length in single-turnover experiments to indicate that ClpA is taking more kinetic steps before dissociation on longer substrates compared to shorter substrates. This is consistent with

ATP-driven translocation, and suggests that we are monitoring ClpA catalyzed translocation with these techniques (31, 34, 48).

To determine if the number of steps translocated scales linearly with substrate length, L, we fit the data to the simplest n-step sequential model with local n, the number of steps, and k_T, the translocation rate constant, for each substrate length (34). To evaluate the relationship between the number of steps and substrate length, a plot of n as a function of L was constructed; see the black traces in **Fig S1**. As found previously, a linear relationship was observed in the n versus L plot and additionally a positive y-intercept was found. Intuitively this predicts that at L = 0 these are some number of steps that have already been taken (49, 50). This is indicative of additional steps outside of translocation existing in the mechanism (45, 51). The way we have modeled this previously has been to include a slow step, k_C, that can account for these non-translocation steps (34). Inclusions of k_C into our model now results in an n versus L plot, see **Fig. S1** red traces, with a positive x-intercept, or negative y-intercept, that we have interpreted as ClpA's contact site size. Therefore, the final scheme used to model the data for WT will include a k_C step.

The system of coupled differential equations that results from **Scheme 1** were solved using Laplace transforms to obtain a set of algebraic equations. These equations were used to generate two-state model equations for analyzing the anisotropy and total fluorescence data. A full description of how this is accomplished is outlined in the **Methods**. As such, the total fluorescence and anisotropy time courses were subjected to global MENOTR analysis using the model equations, **Eq. S1-S4,** derived from **Scheme 1**. Best

fits of the representative 3 mM translocation data are plotted in **Fig. 3 A & B** as broken black traces and the corresponding kinetic parameters can be found in **Table 2**.

The time courses in **Fig. 3 A & B** were collected in H500 in the presence of 3 mM ATP. The determined kinetic parameters for these data vary slightly compared to those of WT parameters previously found in H300 (34). To distinguish between the two sets of data, the NaCl concentration is denoted in the subscript of the corresponding parameters. The translocation rate constant k_{T500} = (2.4 ± 0.3) s^{-1} was found to be ~1.5 times faster than that of k_{T300} = (1.46 ± 0.05) s^{-1}. The nonproductive rate constant k_{NP500} = (0.015 ± 0.001) s^{-1} was found to be 3 times slower than that of k_{NP300} = (0.045± 0.0005) s^{-1}. The slow rate constant k_{C500} = (0.091 ± 0.005) s^{-1} was found to be 2 times slower than that of k_{C300} = (0.20± 0.003) s^{-1}. The kinetic step-size m_{500} = (10 ± 3) aa was found to be within error of m_{300} = (13.4 ± 0.5) aa. The macroscopic rate constant mk_T was comparable for each, with mk_{T500} = (23 ± 3) aa s^{-1} and mk_{T300} = (19.5 ± 0.3) aa s^{-1}.

The kinetic step-size measured here represents the average number of aa translocated between two rate limiting steps. Consequently, if two rate limiting steps with the same rate constant occur for each mechanical movement of the enzyme then this kinetic step-size would be two-fold smaller than the mechanical step-size (50). One way to test this possibility is to examine the ATP concentration dependence of the translocation observed across a range of sub to saturating ATP concentrations. The idea being that if one is observing multiple steps between each mechanical translocation step then when the ATP concentration is reduced the observation will be reduced to observing only a single step. Under such conditions, one would predict that the observed step-size would increase by two-fold.

Translocation experiments were collected as described in **Fig. 2** across a range of [ATP]. Here the experiments were performed from 0.1 – 7 mM [ATP] keeping all other conditions the same. Each experiment was collected in triplicate, and all data sets were subjected to global MENOTR analysis using the same model equations discussed above (see **Eq. S1-S4**). **Table 2** contains the resulting kinetic parameters for each [ATP], with each representing the mean and standard deviation of the three replicates. The translocation rate and rate constant were plotted as a function of [ATP], shown in **Fig. 3 C.** The rate constant and rate both exhibit steep hyperbolic characteristics with respect to [ATP], indicating cooperative binding of ATP to ClpA. Thus, each isotherm was subjected to analysis using an infinity cooperative binding model, the Hill equation **Eq. 11**. The analysis yielded Hill coefficients of (2.5 ± 1.2) and (2.0 ± 0.6), respectively. Both are within error of what has been previously reported for WT in 300 mM NaCl (51). The analysis of the k_T isotherm produced a k_T max = (2.4 ± 0.5) s^{-1}, K_a = (0.8 ± 0.1) mM^{-1}, and b = (0.3 ± 0.1) s^{-1}; while mk_T produced a mk_T max = (17.6 ± 4.3) aa s^{-1}, K_a = (0.7 ± 0.2) mM^{-1}, and b = (7.1 ± 2.7) s^{-1}. **Fig. 3 D** is a plot of the kinetic step-size as a function of [ATP]. At low [ATP], ranging from 0.1 to 0.3 mM, the step-size is ~20 aa. Upon transitioning to higher [ATP], the step-size drops to ~10 aa after reaching 2 mM ATP. This is unlike WT in H300, which we have previously reported to have no ATP dependence with respect to step-size.

Fig. 3 D suggests that there is a two-fold increase in the kinetic step-size that occurs below 1 mM ATP. However, the uncertainty on the kinetic step-size in this range is such that nearly all the measurements are within error. It is important to note that the translocation rate constant, k_T, and the kinetic step-size, m, are negatively correlated. So,

the question remains, is the kinetic steps-size changing with decreasing [ATP] or is this an artifact of the negative parameter correlation? To test this, grid search analyses was performed to determine the confidence intervals on the kinetic step-sizes at 0.1 and 7 mM ATP.

In this grid search, the value of the kinetic step-size is fixed at values between 1 and 35 aa and the other parameters are optimized using multiple rounds of NLLS analysis. A goodness of fit, χ^2, is generated for each of the optimizations, which is then used to calculate a corresponding F calculated value. By comparing the F calculated value of each fit to the overall F critical value it is possible to determine the range over which there is certainty on the value of m. The F calculated values were evaluated at a 68 % confidence level for each [ATP] resulting in confidence intervals of (1.8 - 35.2) aa at 0.1 mM ATP and (3.6 to 19.7) aa at 7 mM ATP. The size of these ranges indicates that there is reduced resolution on m compared to previous results in H300, especially at low [ATP]. The values of all m reported in **Fig. 3 D** fall within these confidence intervals, thus suggesting that no significance can be given to the trends seen in the isotherm.

Examination of polypeptide translocation catalyzed by D2, $ClpA_{E286A}$, in H500

Single-turnover fluorescence anisotropy and total fluorescence time courses for D2 active $ClpA_{E286A}$ were collected in a manner identical to $ClpA_{WT}$, and as described in **Fig. 2**. Representative time courses of D2 catalyzed translocation are shown in **Fig. 4 A & B** as solid traces for the 30mer, 40mer, and 50mer polypeptides in the presence of 3 mM ATP and H500. These time courses exhibit a length dependent lag and the same trends in signal change as observed for WT.

The D2 time courses were first evaluated for the presence of the slow step, k_C, as described above for WT. In the analysis of D2, a positive y-intercept was found for the dependence of n on L when fitting without k_C, (See **Fig. S2 A**). This indicates the need for the inclusion of k_C in the fitting scheme. As such, the D2 time courses were subjected to the same global MENOTR analysis as WT, using the model equations derived from **Scheme 1**, See **Eq. S1-S4**. The resulting best fits are shown as black broken traces in **Fig. 4 A & B** and the determined kinetic parameters for D2 can be found in **Table 3.** The translocation rate constant was found to be (1.7 ± 06) s^{-1}, which is ~71 % of that of WT_{500}. Both k_{NP} and k_C were within error of WT_{500} at (0.01 ± 0.004) s^{-1} and (0.06 ± 0.04) s^{-1}, respectively. The kinetic step-size was found to be m = (11 ± 6) aa while the translocation rate was determined at mk_T = (16 ± 3) aa s^{-1}. While the step-size found matched well with WT_{500}, the rate was slightly slower than that of WT_{500}. Similar to the translocation rate constant, the translocation rate was ~70 % of that of WT_{500}.

D2 catalyzed translocation experiments were collected across a range of [ATP]. The experiments were performed from (0.1 – 7) mM [ATP] keeping all other conditions the same. Each experiment was collected in triplicate and were globally analyzed with MENOTR and **Eq. S1-S4**. **Table 3** contains the resulting kinetic parameters, with each representing the mean and standard deviation of the three replicates. Plotting the translocation rate and rate constant as a function of [ATP], we created the isotherms shown in **Fig. 4E**, gold and green respectively. Both the translocation rate constant and rate show steep hyperbolic characteristics with respect to [ATP], again consistent with positive binding cooperativity of ATP to ClpA. As such, both isotherms were subjected to NLLS

analyzed using the Hill equation, **Eq. 11**,. The analysis of k_T as a function of ATP produced a k_{TD2} max = (1.9 ± 0.3) s^{-1}, K_{aD2} = (1.7 ± 0.4) mM^{-1}, H_{D2} = (2.2 ± 0.8), and b_{D2} = (0.2 ± 0.2) s^{-1}; while mk_T produced a mk_{TD2} max = (12 ± 6) aa s^{-1}, K_{aD2} = (1.5 ± 0.7) mM^{-1}, H_{D2} = (2 ± 1), and b_{D2} = (4 ± 3) s^{-1}. The Hill coefficients found for both the rate and rate constant are within error of that of WT_{500} and are indicative of positive cooperativity. Similar to the WT_{500}data, the D2 step-size shows an apparent [ATP] dependence in **Fig. 4D**. Where at low [ATP] m = ~20 aa and as the [ATP] transitions to higher concentrations m goes to ~7 aa around 0.3 mM ATP.

The trend in the kinetic step-size was tested to determine if it reflects an actual transition in m with [ATP] or an artifact of correlation between m and k_T. Therefore, confidence intervals of m at low [ATP] were evaluated over the range of (1 – 35) aa using grid search strategies. At 0.1 mM ATP the confidence interval was only found on one sided at 2.6 aa, meaning that at low [ATP] m can fall anywhere in the range of (2.6 to 35) aa. We predict that the observed increase in m at low [ATP] is a consequence of parameter correlation and lowered resolution on the step-size at low [ATP]. Thus, the step size is independent of [ATP] and is predicted to be approximately 11 aa for all [ATP].

Examination of polypeptide translocation catalyzed by D1, $ClpA_{E565A}$, in H500

Single-turnover translocation experiments for D1 active $ClpA_{E565A}$ were collected using the experimental design described in **Fig. 2**. Representative time courses of D1 catalyzed translocation are shown in **Fig. 5 A & B** as solid colored traces for the 30mer, 40mer, and 50mer SsrA tagged substrates in the presence of 3 mM ATP and H500. In stark contrast to both $ClpA_{WT}$ and Clp_{AE286A}, the Clp_{AE565A} time courses exhibited much slower kinetics. The longest time courses for WT and D2 were collected over 700 s,

while the shortest time courses for D1 were collected over 2000 s. Moreover, the lower [ATP] D1 time courses were collected out to 5000 s to capture the entire reaction.

The total fluorescence time courses collected for D1 catalyzed translocation exhibited a substrate length dependence in the extent of lag and a biphasic shape similar to D2 and WT. However, in the anisotropy time courses for D1, it was seen that the 30mer substrate progresses through an initial rise phase prior to the expected decrease as the enzyme dissociates from the substrate. We attempted to analyze the D1 time courses using the previous two-state model given in **Eq. S1-S4**. However, the model was not able to adequately describe the time courses based both on visual inspection of the fit and the resultant high chi-squared values (fits not shown). In order to account for this new rise phase in the time courses we derived three-state models, **Eq. S5 - S8**, as described in the **Methods**. These three-state models were used in the analysis of the time courses for $ClpA_{E565A}$ catalyzed translocation.

The time courses of D1 catalyzed translocation were evaluated for the presence of the slow step, k_C, as described here previously for WT and D2. However, unlike WT and D2, no positive y-intercept was found in the plot of n as a function of L when fitting without k_C, (See **Fig. S2 B**). Thus, the D1 time courses were subjected to global MEN-OTR analysis using the model equations **Eq. S5 - S8** derived from **Scheme 1** without the inclusion of k_C. The resulting best-fit traces are shown as black broken overlays in **Fig. 5 A & B** and the determined kinetic parameters can be found in **Table 4**. The translocation rate constant for D1 catalyzed translocation at 3 mM ATP was found to be (0.017 ± 0.007) s^{-1}, which represents less than 1 % of the WT translocation rate constant. The de-

termined step-size, m = (16 ± 4) aa, was within error of that of both WT and D2. Comparing the macroscopic rate constant of (0.24 ± 0.002) aa s^{-1} to WT we see that the rate is also ~1 % of the WT magnitude. Additionally, all the values determined for k_{NP} were found to be one order of magnitude lower than those of WT.

D1 catalyzed translocation experiments were collected across a range of [ATP] as detailed in **Fig. 2**. Identical to WT and D2, the experiments were performed from (0.1 – 7) mM [ATP] keeping all other conditions the same. Each experiment was collected in triplicate and all data sets were analyzed globally with MENOTR using **Eq. S5 - S8**. However, it was found that at [ATP] below 0.7 mM the three-state models were failing to describe the D1 translocation time courses being analyzed. This was determined based on visual inspection of the fits and comparison of the subsequent chi-squared values. At these concentrations, the resulting D1 catalyzed translocation time courses had no apparent lag phase, and were better described by fits to a single exponential. Thus, indicating that the time courses are no longer sensitive to the repeated rounds of translocation observed previously. **Table 4** contains the resulting kinetic parameters for the analysis of experiments carried out over (0.7 – 7) mM ATP, with each representing the mean and standard deviation of three replicates.

The D1 translocation rate and rate constant were plotted as a function of [ATP] from (0.7 – 7) mM, shown in **Fig. 5 C** as green and gold respectively. Both isotherms exhibited steep hyperbolic characteristics similar to WT and D2, again suggesting cooperativity. Thus, each was subjected to NLLS analysis using the Hill equation, **Eq. 11**. The analysis of the k_T isotherm resulted in a k_{TD1} max = (0.023 ± 0.007) s^{-1}, K_{aD1} = (0.3 ± 0.1) mM^{-1}, H_{D1} = (3 ± 2), and b_{D1} = (0 ± 0) s^{-1}; while the mk_T isotherm analysis produced a

mk_{TD1} max = (0.39 ± 0.03) aa s^{-1}, K_{aD1} = (0.38 ± 0.2) mM^{-1}, H_{D1} = (4.0 ± 0.8), and b_{D1} = (0.02 ± 0.02) s^{-1}. The D1 Hill coefficient for k_T is within error of both WT and D2, while mk_T was just outside of error of both.

In **Fig. 5 D**, we plotted the kinetic step-size of D1 as a function of [ATP] from (0.7 – 7) mM ATP. At lower [ATP], ranging from 0.7 to 1 mM, the step-size is estimated to be ~20 aa. Upon transitioning to higher [ATP], the step-size drops to ~10 aa after reaching 2 mM ATP. Grid search analysis of 7 mM ATP from m = 1 to 35 aa yielded a one-sided confidence interval of 13.9. Meaning that at 7 mM ATP, step-size values from (13.9 – 35) aa produce the same goodness of fit in the analysis of the data. Moreover, at lower [ATP] the grid search analysis failed to find confidence intervals for the step-size with all values of m between 1 and 35 aa providing the same goodness of fit in the analysis. The large error and lack of grid search confidence intervals on 0.7 and 1 mM ATP suggest that parameter correlation is most likely causing the apparent shift in kinetic step-size. Thus, we predict that the step-size for D1 is approximately 20 aa across all the ATP concentrations.

Discussion

Previous work done by our lab has resulted in the determination of the elementary rate constants, overall translocation rates, and the kinetic step-sizes for the translocation of polypeptide by ClpA and ClpAP. These kinetic parameters were elucidated in an attempt to characterize the transient-state molecular mechanism of translocation by both systems. That work led to the proposal of the most encompassing mechanistic model to date, but the proposal lacked direct information on the mechanisms that describe translocation at each NBD of ClpA (31). We present here the first quantitative description of the

molecular mechanisms of polypeptide translocation catalyzed by D1 and D2 of *E. coli* ClpA in the absence of ClpP.

Dependence of translocation mechanism on [NaCl]

In order to determine how each NBD contributes to the overall translocation activity, the kinetics of translocation catalyzed by $ClpA_{E286A}$ and $ClpA_{E565A}$ would need to be compared to that of $ClpA_{WT}$. Although we have already established a mechanism for WT, our previous investigations were performed in the presence of 300 mM NaCl. It was found that all ClpA Walker B variants displayed reduced stability at this salt concentration compared to WT. As such, the NaCl concentration was raised to accommodate the solubility of the mutants. Thus, the activity of WT was ascertained under these new conditions to rule out the contribution of salt on the activities of the mutants.

Fluorescence anisotropy and total fluorescence stopped-flow assays were performed on WT in buffer containing 500 mM NaCl at [ATP] from (0.1 – 7) mM. Global quantitative analysis of the resulting time courses produced kinetic parameters that were comparable to those previously determined for translocation under 300 mM conditions. The overall rate of polypeptide translocation at saturating ATP was found to (23 ± 3) aa s^{-1} which is within error of the previously determined rate of ~20 aa s^{-1}. The slight increase in the rate was found to be the consequence of an increase in the translocation rate constant.

These results are not surprising as an increase in NaCl is likely to affect ATP binding, which we have shown to be kinetically coupled to the translocation rate constant (34). From the ATP dependence of the rate and rate constant, we were able to predict that the K_a of ATP binding is ~ 1 mM^{-1} in 500 mM NaCl; while the previously determined K_a

was ~2 mM^{-1} in 300 mM NaCl. While there is a twofold difference in the equilibrium constant for ATP binding, the observed dependence in the overall rate on ATP is small if any. To probe this farther would require these experiments to be carried out across a larger range of ATP concentrations and was not explored further in this work.

Rate of D1 & D2 catalyzed polypeptide translocation

The analysis of the resultant time courses for both $ClpA_{E286A}$ and $ClpA_{E565A}$ catalyzed translocation revealed that the majority of the translocation activity of WT is catalyzed by D2. D2 was found to translocate polypeptide at saturating ATP with an apparent rate ~ 70 % of that of WT, compared to D1 at only ~1 %. Interestingly, Kress et al. showed that for the same variants the steady-state ATPase activities of D2 and D1 were ~86 % and ~10 % of WT activity, respectively (39). It should be noted that since ATP hydrolysis and translocation are coupled events, the rates determined for each should represent the same rate-limiting step in the reaction.

The findings presented by Kress et al. match well with our reported rates of translocation, and the small discrepancy between the reported values could be the consequence of their ATPase activity being measured in the steady state in the absence of polypeptide substrate. Meaning, their ATP turnover numbers represent ATP hydrolysis involved in not only translocation by hexameric ClpA, but also hydrolysis catalyzed by lower order ClpA oligomers (41). Nevertheless, both findings are consistent with the prevailing theory of ATP hydrolysis at D2 providing the majority of the energy and mechanical work for substrate unfolding by WT (20, 23, 31, 38).

A similar D2 active ClpA variant, $ClpA_{E286Q}$, was reported to have an unfolding rate ninefold lower than that determined for WT (32). This constitutes only ~11 % the activity of WT. In those studies, $ClpA_{E286Q}$ catalyzed the steady-state unfolding of an Arc-Gcn-SsrA dimer. This substrate was specifically designed to induce a high level of stability by adding a parallel coiled coil sequence adjacent to the SsrA recognition tag to slow down the rate of substrate unfolding. However, it has been determined that hydrolysis of ATP at D1 becomes important when ClpA encounters a highly stable protein structure near the terminal ClpA recognition tag (39, 52, 53). Kress et al. proposed that D1 is involved in the initial capture and partial unraveling of target substrate prior to facilitating handover to D2 (23, 39). Thus, it is possible that abolition of ATP hydrolysis at D1 disrupts ClpA's ability to initially grip folded substrates without them refolding and dissociating prior to translocation.

In the $ClpA_{E286Q}$ variant studied by Baytshtok et al., the loss of ATP hydrolysis at D1 has the potential to perturb the domain's ability to properly handover folded polypeptide substrate to D2. Consequently, the steps preceding D2 catalyzed unfolding could become rate limiting rather than unfolding at D2 itself. This is consistent with the lower unfolding activity reported. In contrast, our translocation experiments on $ClpA_{E286A}$ are not sensitive to unfolding or initial binding as all the substrates have been shown to be unstructured and are prebound to ClpA prior to experimentation (42). Additionally, Kress et al. showed that degradation of unstructured peptide is less affected by removal of hydrolysis at D1 in ClpAP. Thus, the discrepancies between our results and those found by Baytshtok et al. could be the consequence of how the NBDs play differing roles in unfolding of structured substrates versus translocation of unstructured substrates. The examination

of D1 and D2 catalyzed translocation of folded substrates using the methods described here is warranted to test this hypothesis.

Interpretation of the ATP dependence of the translocation mechanism at D1 & D2

By preassembling the ClpA-polypeptide complex and maintaining saturating levels of non-fluorescent polypeptide trap, the experiments presented here are single-turnover. Meaning, the method is only sensitive to a single round of polypeptide translocation, and is independent of affects from ClpA oligomerization or polypeptide binding. Thus, we are sensitive to only the molecular events that occur in the active site of the motor during translocation. However, within each cycle of polypeptide translocation, the motor must go through at a minimum ATP binding and hydrolysis, mechanical movement, various conformational changes, and ADP release. The rate constant, k_T, reported here represents the slowest repeating step or steps in the translocation cycle. This step could represent any of the above-mentioned processes or a combination of two that have similar rate constants.

In an attempt to reveal which of these steps is rate limiting we have examined the ATP dependence of the kinetic parameters that describe translocation by ClpA (31, 34). If more than a single step in the translocation cycle is rate limiting at saturating ATP, then a reduction in ATP concentration will cause a change in both the apparent number of steps and step-size. As the concentration of ATP is reduced, either ATP binding or a step coupled to ATP binding will become rate limiting resulting in only a single step being observed. A transition from two repeating steps per cycle to a single step will result in a two-fold reduction in the number of steps or a two-fold increase in the step-size observed

(45, 50). For WT we have previously determined that we are observing the step that is kinetically coupled to and immediately follows ATP binding (34). We have predicted this step to be either ATP hydrolysis or physical movement of the pore loop (31).

Here we found that in all cases as ATP is reduced from 7 mM to 0.1 mM there is an apparent effect on the overall rate and translocation rate constant. Upon initial inspection the rate and rate constant of both D1 and D2 appeared to have a hyperbolic dependence on ATP indicating that ATP binding is not rate limiting. If ATP binding was rate limiting, we would predict a linear dependence of ATP on the rate. After analysis, it was found that the isotherms for both the rate and rate constant were too steep to be described by a simple one-to-one binding model, and required an infinity cooperative model, the Hill equation. This degree of steepness in the ATP dependence indicates that ATP binding is cooperative, that the rate-limiting step is kinetically coupled to ATP binding, and that the rate-limiting step immediately follows ATP binding. Thus, we conclude that like WT, translocation at D1 and D2 is predicted to be either ATP hydrolysis or physical movement of the pore loop.

Analysis of ClpA and ClpAP translocation coupled to the ATPase activities of D1 and D2 has led to the proposal of cooperative interactions between NBDs either on the same monomer or between NBDs of adjacent monomers (31, 34, 39). Here we found that positive cooperatively in ATP binding for both D1 and D2 was equivalent to WT. Thus, suggesting that cooperativity is not lost when hydrolysis is abolished in one of the domains. Additionally, it has been reported elsewhere that ATP hydrolysis in D1 is not coupled to hydrolysis in D2 (39). Both of these observations are consistent with inter-monomer interactions driving the observed cooperativity in ATP binding.

Unlike WT in 300 mM NaCl, here we found that all three ClpA variants exhibited an apparent change in kinetic step-size as ATP concentration was lowered. However, the experimental uncertainty on the kinetic step-size is such that nearly all the measurements are within error at low [ATP]. Additionally, F statistics and grid search analysis revealed that at low ATP concentrations the confidence intervals for step-size is increased such that resolution on the parameter is lowered. This coupled with the previous observation of the ClpA step-size being independent of [ATP] indicates that the observed decrease is most likely a consequence of negative parameter correlation between the rate constant and step-size and large uncertainty on the values.

Updated molecular model of ClpA catalyzed translocation

In our previous model of ClpA catalyzed translocation we concluded that in the absence of ClpP the rate limiting step of translocation occurs at D1 with a rate constant of ~1.39 s^{-1} and step-size of 14 aa (31). This was based on the analysis of translocation catalyzed by ClpA and ClpAP (31, 34), the steady-state ATPase hydrolysis rates of D1 and D2 (39), and crosslinking studies (23), but it lacked direct measurements of translocation catalyzed by each domain.

Thus, the question becomes do the results presented here support a mechanism in which D1 is rate limiting? It was determined that D1 catalyzed substrate translocation with an overall rate of 0.35 aa s^{-1} at saturating ATP. If the rate-limiting step of WT catalyzed translocation occurred at D1 then it would be predicted that the activity of WT would be similar to that determined for D1 under the same conditions. However, WT was found to translocation substrate with an overall rate of 24 aa s^{-1} under saturating ATP, an

order of magnitude higher than that of D1. In contrast, D2 was found to translocate substrate with a rate of 15 aa s-1, suggesting that translocation at D2 is more likely to contain the rate limiting step.

While the translocation activity of D2 accounts for a large percentage of WT's translocation activity it has been seen elsewhere that full wildtype activity requires both NBDs (23, 32, 39). As discussed earlier, ATP hydrolysis at D2 accounted for ~86 % of $ClpA_{WT}$ hydrolysis, with D1 making up the other ~10 %. Additionally, Hinnerwisch et al. showed via crosslinking studies that ClpA engaged with its substrate through the pore loops located on D2 and D1, and that mutations in the D1 pore loops abolished translocation activity. Thus, our results and findings from the field support a mechanistic model of WT catalyzed translocation where D2 drives substrate translocation and D1 provides additional activity that is not directly involved in the repeating cycles of translocation.

If D1 is not directly implicated in the repeating translocation cycle, then what purpose does it serve? Recent work by Kotamarthi et al. has reported on the analysis of ClpAP translocation where ClpA contains the E286Q mutation abolishing ATP hydrolysis at D1 (54). The analysis of single-molecule optical-tweezers assays provided evidence of excessive stalling of the motor during D2 catalyzed translocation, especially at low [ATP]. They observe an average translocation velocity that was 60 % slower than that of $ClpAP_{WT}$, but when the stalls were removed from the analysis, the average velocity and step-size of translocation were similar to that of $ClpAP_{WT}$. Indicating that the slower translocation is almost entirely caused by the observed pausing, and that D2 has a decreased ability to grip substrates properly in the absence of D1. This is consistent with the

proposals made by Kress, Hinnerwisch, and Bohon et al., as well as what we have reported here. Thus, we conclude that D1 likely functions as an auxiliary/regulator motor that stabilizes the substrate during translocation and may assist in initial substrate handover to D2 prior to translocation.

In such a model, ClpA would start prebound with polypeptide substrate making contact with the pore loops of D2 in an up conformation (23, 55). Translocation would initiate with D2 hydrolyzing ATP causing the pore loop to move down thereby translocating the substrate by ~11 aa. The hydrolysis of ATP to ADP and P_i at D2 reduces its affinity for substrate (22), thus allowing D2 to release the substrate, reset, and subsequently rebinds both the substrate and a new ATP. This cycle of repeated translocation steps at D2 would occur approximately every 0.3 s, based on a rate constant of 3.3 s^{-1}. During reset of the D2 pore-loop, D1 stabilizes the substrate preventing back slipping and facilitating proper rebinding of the D2 pore-loops, which further prevents stalling of the motor.

References

1. Kelman, L. M., and Z. Kelman. 2014. Archaeal DNA Replication. Annual review of genetics 48(1):71-97.

2. Neuwald, A. F., L. Aravind, J. L. Spouge, and E. V. Koonin. 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome research 9(1):27-43.

3. Gottesman, S. 1996. Proteases and their targets in Escherichia coli. Annu Rev Genet 30:465-506.

4. Sauer, R. T., and T. A. Baker. 2011. AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. Annu Rev Biochem 80:587-612.

5. Ingmer, H., C. Miller, and S. N. Cohen. 2001. The RepA protein of plasmid pSC101 controls Escherichia coli cell division through the SOS response. Molecular microbiology 42(2):519-526.

6. Wickner, S. H. 1990. Three Escherichia coli heat shock proteins are required for P1 plasmid DNA replication: formation of an active complex between E. coli DnaJ protein and the P1 initiator protein. Proc Natl Acad Sci U S A 87(7):2690-2694.

7. Confalonieri, F., and M. Duguet. 1995. A 200-amino acid ATPase module in search of a basic function. Bioessays 17(7):639-650.

8. Hanson, P. I., and S. W. Whiteheart. 2005. AAA+ proteins: have engine, will work. Nat Rev Mol Cell Biol 6(7):519-529. Review.

9. Miller, J. M., and E. J. Enemark. 2016. Fundamental Characteristics of AAA+ Protein Family Structure and Function. Archaea (Vancouver, B.C.) 2016:9294307.

10. Pak, M., and S. Wickner. 1997. Mechanism of protein remodeling by ClpA chaperone. Proc Natl Acad Sci U S A 94(10):4901-4906.

11. Maurizi, M. R. 1992. Proteases and protein degradation in Escherichia coli. Experientia 48(2):178-201.

12. Guo, F., M. R. Maurizi, L. Esser, and D. Xia. 2002. Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease. J Biol Chem 277(48):46743-46752.

13. Lee, S., M. E. Sowa, Y. H. Watanabe, P. B. Sigler, W. Chiu, M. Yoshida, and F. T. Tsai. 2003. The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. Cell 115(2):229-240.

14. Parsell, D. A., Y. Sanchez, J. D. Stitzel, and S. Lindquist. 1991. Hsp104 is a highly conserved protein with two essential nucleotide-binding sites. Nature 353(6341):270-273.

15. Glynn, S. E., A. Martin, A. R. Nager, T. A. Baker, and R. T. Sauer. 2009. Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. Cell 139(4):744-756.

16. Sousa, M. C., C. B. Trame, H. Tsuruta, S. M. Wilbanks, V. S. Reddy, and D. B. McKay. 2000. Crystal and solution structures of an HslUV protease-chaperone complex. Cell 103(4):633-643.

17. Langklotz, S., U. Baumann, and F. Narberhaus. 2012. Structure and function of the bacterial AAA protease FtsH. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1823(1):40-48.

18. Gottesman, S., C. Squires, E. Pichersky, M. Carrington, M. Hobbs, J. S. Mattick, B. Dalrymple, H. Kuramitsu, T. Shiroza, T. Foster, and et al. 1990. Conservation of the regulatory subunit for the Clp ATP-dependent protease in prokaryotes and eukaryotes. Proc Natl Acad Sci U S A 87(9):3513-3517.

19. Maurizi, M. R., S. K. Singh, M. W. Thompson, M. Kessel, and A. Ginsburg. 1998. Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and clpP. Biochemistry 37(21):7778-7786.

20. Singh, S. K., and M. R. Maurizi. 1994. Mutational analysis demonstrates different functional roles for the two ATP-binding sites in ClpAP protease from Escherichia coli. J Biol Chem 269(47):29537-29545.

21. Gates, S. N., and A. Martin. 2020. Stairway to translocation: AAA+ motor structures reveal the mechanisms of ATP-dependent substrate translocation. Protein Sci 29(2):407-419.

22. Farbman, M. E., A. Gershenson, and S. Licht. 2007. Single-Molecule Analysis of Nucleotide-Dependent Substrate Binding by the Protein Unfoldase ClpA. J Am Chem Soc.

23. Hinnerwisch, J., W. A. Fenton, K. J. Furtak, G. W. Farr, and A. L. Horwich. 2005. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. Cell 121(7):1029-1041.

24. Farbman, M. E., A. Gershenson, and S. Licht. 2008. Role of a conserved pore residue in the formation of a prehydrolytic high substrate affinity state in the AAA+ chaperone ClpA. Biochemistry 47(51):13497-13505.

25. Martin, A., T. A. Baker, and R. T. Sauer. 2008. Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. Nat Struct Mol Biol 15(11):1147-1151.

26. Weibezahn, J., P. Tessarz, C. Schlieker, R. Zahn, Z. Maglica, S. Lee, H. Zentgraf, E. U. Weber-Ban, D. A. Dougan, F. T. Tsai, A. Mogk, and B. Bukau. 2004. Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. Cell 119(5):653-665.

27. Biter, A. B., S. Lee, N. Sung, and F. T. Tsai. 2012. Structural basis for intersubunit signaling in a protein disaggregating machine. Proc Natl Acad Sci U S A 109(31):12515-12520.

28. Lopez, K. E., A. N. Rizo, E. Tse, J. Lin, N. W. Scull, A. C. Thwin, A. L. Lucius, J. Shorter, and D. R. Southworth. 2019. Conformational Plasticity of the ClpAP AAA+ Protease Couples Protein Unfolding and Proteolysis. bioRxiv:820209.

29. Olivares, A. O., A. R. Nager, O. Iosefson, R. T. Sauer, and T. A. Baker. 2014. Mechanochemical basis of protein degradation by a double-ring AAA+ machine. Nat Struct Mol Biol.

30. Kolygo, K., N. Ranjan, W. Kress, F. Striebel, K. Hollenstein, K. Neelsen, M. Steiner, H. Summer, and E. Weber-Ban. 2009. Studying chaperone-proteases using a real-time approach based on FRET. Journal of structural biology 168(2):267-277.

31. Miller, J. M., J. Lin, T. Li, and A. L. Lucius. 2013. E. coli ClpA Catalyzed Polypeptide Translocation is Allosterically Controlled by the Protease ClpP. Journal of Molecular Biology 425(15):2795-2812.

32. Baytshtok, V., T. A. Baker, and R. T. Sauer. 2015. Assaying the kinetics of protein denaturation catalyzed by AAA+ unfolding machines and proteases. Proc Natl Acad Sci U S A 112(17):5377-5382.

33. Duran, E. C., C. L. Weaver, and A. L. Lucius. 2017. Comparative Analysis of the Structure and Function of AAA+ Motors ClpA, ClpB, and Hsp104: Common Threads and Disparate Functions. Front Mol Biosci 4:54.

34. Rajendar, B., and A. L. Lucius. 2010. Molecular mechanism of polypeptide translocation catalyzed by the Escherichia coli ClpA protein translocase. J Mol Biol 399(5):665-679.

35. Schaupp, A., M. Marcinowski, V. Grimminger, B. Bosl, and S. Walter. 2007. Processing of proteins by the molecular chaperone Hsp104. J Mol Biol 370(4):674-686.

36. Hersch, G. L., R. E. Burton, D. N. Bolon, T. A. Baker, and R. T. Sauer. 2005. Asymmetric interactions of ATP with the AAA+ ClpX6 unfoldase: allosteric control of a protein machine. Cell 121(7):1017-1027.

37. Werbeck, N. D., S. Schlee, and J. Reinstein. 2008. Coupling and dynamics of subunits in the hexameric AAA+ chaperone ClpB. J Mol Biol 378(1):178-190.

38. Seol, J. H., S. H. Baek, M. S. Kang, D. B. Ha, and C. H. Chung. 1995. Distinctive roles of the two ATP-binding sites in ClpA, the ATPase component of protease Ti in Escherichia coli. J Biol Chem 270(14):8087-8092.

39. Kress, W., H. Mutschler, and E. Weber-Ban. 2009. Both ATPase domains of ClpA are critical for processing of stable protein structures. J Biol Chem 284(45):31441-31452.

40. Duran, E. C., and A. L. Lucius. 2018. ATP hydrolysis inactivating Walker B mutation perturbs E. coli ClpA self-assembly energetics in the absence of nucleotide. Biophys Chem 242:6-14.

41. Duran, E. C., and A. L. Lucius. 2019. Examination of the nucleotide linked assembly mechanism of E. coli ClpA. Protein Sci.

42. Li, T., and A. L. Lucius. 2013. Examination of Polypeptide Substrate Specificity for E. coli ClpA. Biochemistry 52:4941-4954.

43. Sharan, S. K., L. C. Thomason, S. G. Kuznetsov, and D. L. Court. 2009. Recombineering: a homologous recombination-based method of genetic engineering. Nature protocols 4(2):206-223.

44. Veronese, P. K., R. P. Stafford, and A. L. Lucius. 2009. The Escherichia coli ClpA Molecular Chaperone Self-Assembles into Tetramers. Biochemistry 48(39):9221-9233.

45. Lucius, A. L., N. K. Maluf, C. J. Fischer, and T. M. Lohman. 2003. General methods for analysis of sequential "n-step" kinetic mechanisms: application to single turnover kinetics of helicase-catalyzed DNA unwinding. Biophys J 85(4):2224-2239.

46. Farrell, C. M., A. D. Grossman, and R. T. Sauer. 2005. Cytoplasmic degradation of ssrA-tagged proteins. Molecular microbiology 57(6):1750-1761.

47. Karzai, A. W., E. D. Roche, and R. T. Sauer. 2000. The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. Nature structural biology 7(6):449-455.

48. Durie, C. L., J. Lin, N. W. Scull, K. L. Mack, M. E. Jackrel, E. A. Sweeny, L. M. Castellano, J. Shorter, and A. L. Lucius. 2019. Hsp104 and Potentiated Variants Can Operate as Distinct Nonprocessive Translocases. Biophys J 116(10):1856-1872.

49. Lucius, A. L., C. Jason Wong, and T. M. Lohman. 2004. Fluorescence stopped-flow studies of single turnover kinetics of E.coli RecBCD helicase-catalyzed DNA unwinding. J Mol Biol 339(4):731-750.

50. Lucius, A. L., and T. M. Lohman. 2004. Effects of temperature and ATP on the kinetic mechanism and kinetic step-size for E.coli RecBCD helicase-catalyzed DNA unwinding. J Mol Biol 339(4):751-771.

51. Lucius, A. L., J. M. Miller, and B. Rajendar. 2011. Application of the Sequential n-Step Kinetic Mechanism to Polypeptide Translocases. Methods in enzymology 488:239-264.

52. Lee, C., M. P. Schwartz, S. Prakash, M. Iwakura, and A. Matouschek. 2001. ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. Mol Cell 7(3):627-637.

53. Kenniston, J. A., T. A. Baker, J. M. Fernandez, and R. T. Sauer. 2003. Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. Cell 114(4):511-520.

54. Kotamarthi, H. C., R. T. Sauer, and T. A. Baker. 2020. The Non-dominant AAA+ Ring in the ClpAP Protease Functions as an Anti-stalling Motor to Accelerate Protein Unfolding and Translocation. Cell Rep 30(8):2644-2654.e2643.

55. Bohon, J., L. D. Jennings, C. M. Phillips, S. Licht, and M. R. Chance. 2008. Synchrotron protein footprinting supports substrate translocation by ClpA via ATP-induced movements of the D2 loop. Structure 16(8):1157-1165.

56. Goddard, T. D., C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, and T. E. Ferrin. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci 27(1):14-25.

57. Xia, D., L. Esser, S. K. Singh, F. Guo, and M. R. Maurizi. 2004. Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone. Journal of structural biology 146(1-2):166-179.

58. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. Journal of molecular graphics 14(1):33-38, 27-38.

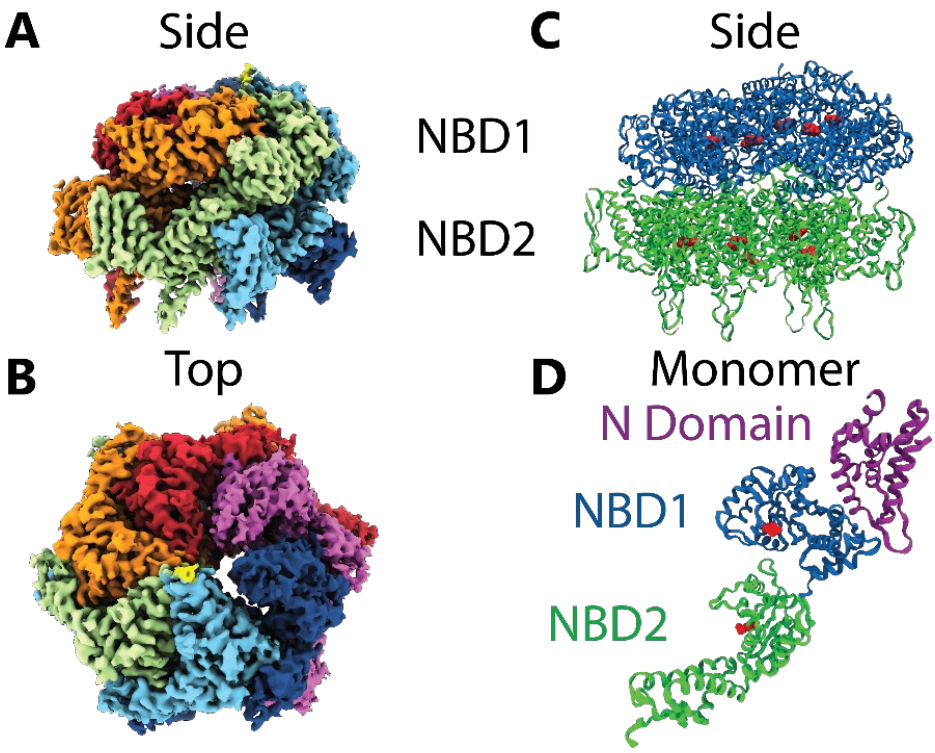


FIGURE 1 ClpA hexamer and monomer structures. (**A, B, and C**) ClpA structure maps coming from the determination of a ~3.0 Å resolution cryo-EM structure of E. coli ClpAP in the presence of ATPγS and RepA-tagged GFP substrate (28). Surface rendering of the (**A**) side and (**B**) top down views of nucleotide binding domains 1 and 2. Due to the high mobility of the N domain, it cannot be resolved, and was not modeled in the cryo-EM structure. These images were prepared using UCSF Chimera (Computer Graphics Laboratory, University of California, San Francisco (56). (**C**) Hexameric ClpA ribbon representation of the cryo-EM ClpA structure with NB1 in blue and NB2 in green. (D) Monomer ribbon representation of the crystal structure of ClpA with N (purple), NB1 (blue), and NB2 (green) domains (57). In both **C & D**, the red residues indicate the positions that have mutated in the D1 and D2 Walker B variants, $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$. These structures were prepared using Visual Molecular Dynamics 1.9.3 (University of Illinois) (58).

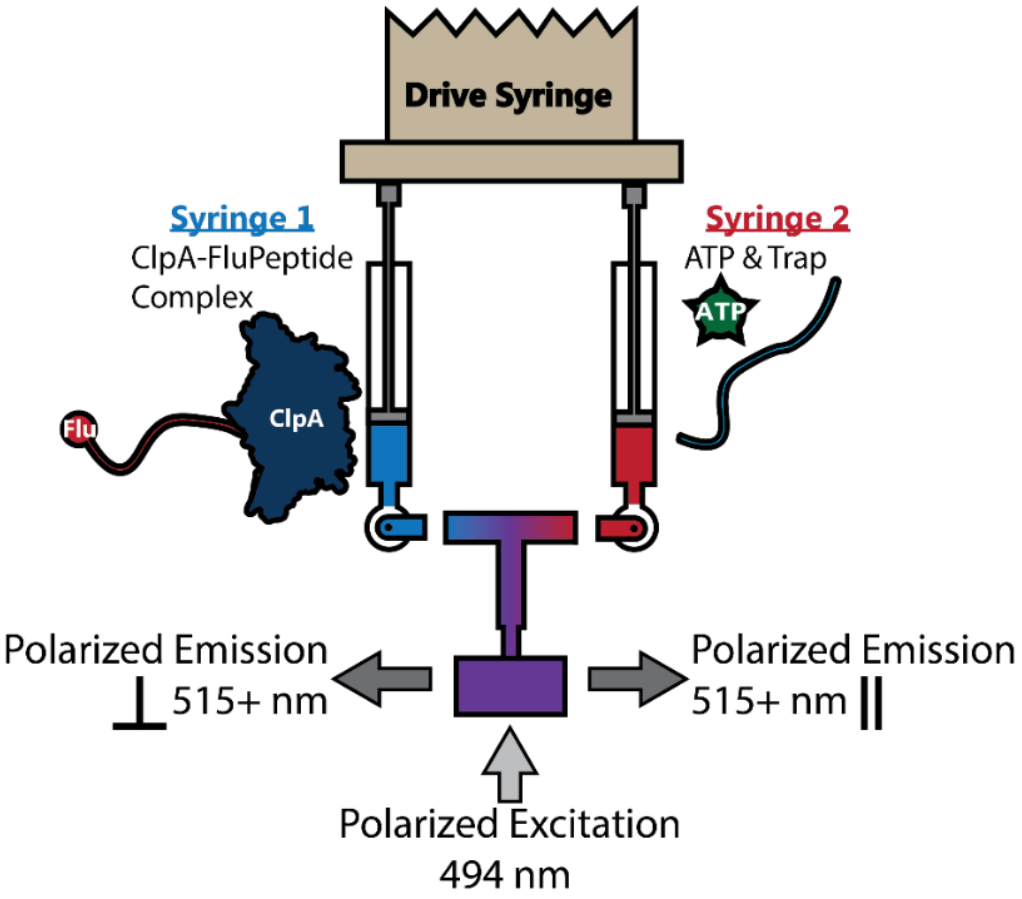


FIGURE 2 Schematic outlining single-turnover translocation experiments in a SX20 stopped-flow spectrometer (Applied Photophysics). Syringe 1, shown in blue, contains 4 μM ClpA, 300 μM ATPγS, and 200 nM flu-SsrA polypeptide. These reagents are mixed as outlined in **Methods** to form a prebound ClpA peptide complex that is poised to start translocation when mixed with ATP. Syringe 2, shown in red, contains ATP at a concentration indicated in the text and 20 μM α-casein to serve as a trap for unbound ClpA, thereby maintaining single-turnover conditions. The contents of the two syringes are rapidly mixed and flow into the sample chamber shown in purple, where the sample is excited, and emission is observed as detailed below. The stopped-flow is arranged in a T-format for fluorescence anisotropy. Here, the fluorescein sample is excited with vertically polarized light at 494 nm. The fluorescence emission is measured at 90° to the incident light with two PMTs set 180° to one another. One PMT is fit with a vertical polarizer, detecting emitted light parallel to excitation, and the other a horizontal polarizer that detects light perpendicular to excitation. Each PMT is also fit with a 515+ nm long pass filter to block any excitation light. Upon mixing, the final concentration of each reagent in the sample chamber is half of its original in the premixing syringes.

TABLE 1 Fluorescent Polypeptide Substrates

Name	Length (aa)	Sequence or Source
Flu-SsrA 30mer	30	Flu-CTKSAANLKVKELRSKKKL**AANDENYALAA**
Flu-SsrA 40mer	40	Flu-CTGEVSFQAANTKSAANLKVKELRSKKKL**AANDENYALAA**
Flu-SsrA 50mer	50	Flu-CLILHNKQLGMTGEVSFQAANTKSAANLKVKELRSKKKL**AANDENYALAA**

Bolded sequences indicate the 11 aa SsrA tag

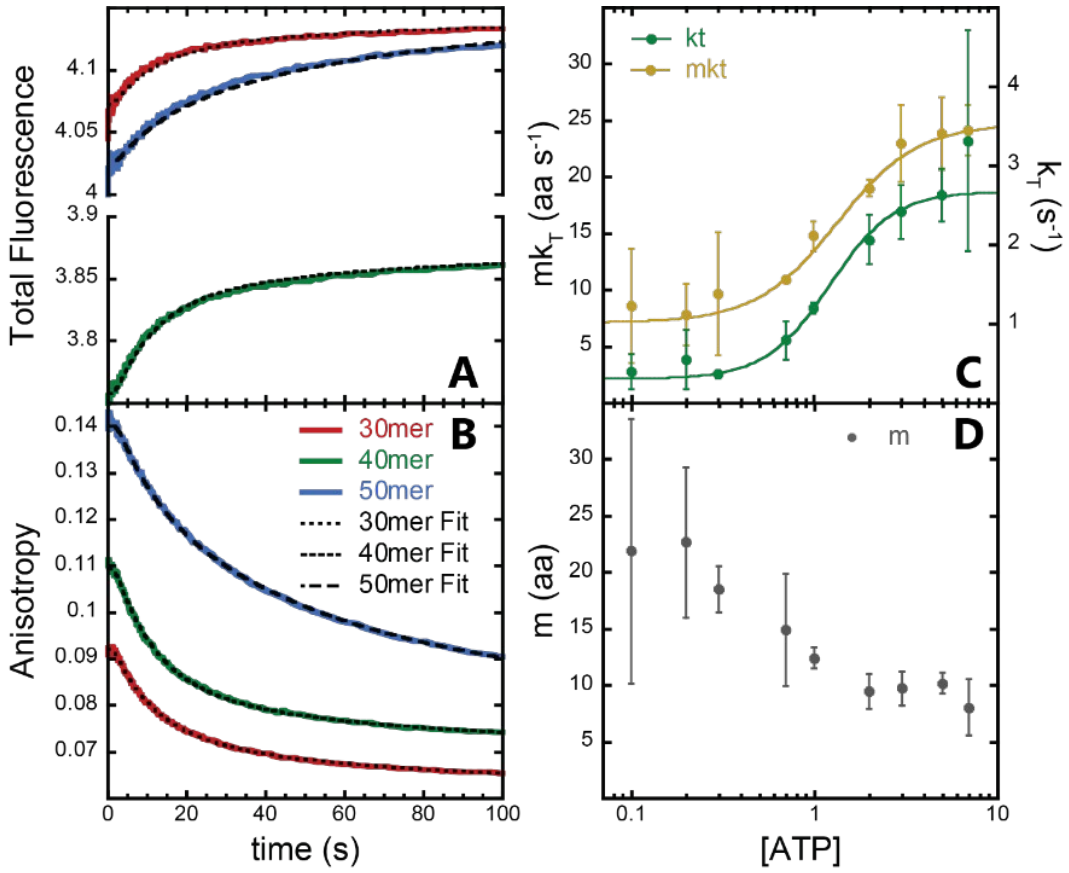


FIGURE 3 Fluorescence anisotropy and total fluorescence time courses for ClpA$_{WT}$ catalyzed translocation of SsrA tagged polypeptide substrates. Representative time courses (**A & B**) were collected at 3 mM [ATP] as described in **Fig. 2**. Translocation was monitored concurrently using (**A**) total fluorescence and (**B**) anisotropy for each time course. The solid colored traces correspond to the 30, 40, 50 amino acid SsrA substrates (See **Table 1**). The anisotropy and total fluorescence data were subjected to MENOTR global analysis described in the methods section using **Scheme 1** and **Eq. S1-S4**. The resulting best fits for the 3 mM ATP data are shown as overlaid black broken traces and the optimized parameters can be found in **Table 2**. Translocation time courses were collected over a range of [ATP], from 0.1 - 7 mM. The kinetic parameters (**C**) k_T, mk_T, and (**D**) m are shown plotted as a function of [ATP]. Each [ATP] experiment was collected in triplicate, with the figure representing the average and standard deviation of the three. Both the mk_T and k_T isotherms were subjected to NLLS analysis using the Hill equation (**Eq. 11**). The best fits are shown as a solid trace of the respective color. The fit of k_T produced a k_T max = (2.4 ± 0.5) s^{-1}, K_a = (0.8 ± 0.1) mM^{-1}, H = (2.5 ± 1.2), and b = (0.3 ± 0.1) s^{-1}; while mk_T produced a mk_T max = (17.6 ± 4.3) aa s^{-1}, K_a = (0.7 ± 0.2) mM^{-1}, H = (2.0 ± 0.6), and b = (7.1 ± 2.7) s^{-1}. The [ATP] dependence of m shows that at low [ATP] m approached a value of ~22 aa, and as [ATP] increased there is a transition to an m of ~10 aa.

TABLE 2 Optimized fitting parameters for $ClpA_{WT}$ catalyzed translocation of polypeptide

[ATP] mM	k_T (s^{-1})	k_{NP} (s^{-1})	k_C (s^{-1})	m (aa)	mk_T (aa s^{-1})
0.1	0.4 ± 0.2	0.001 ± 0.001	0.006 ± 0.003	20 ± 10	9 ± 5
0.2	0.5 ± 0.3	0.03 ± 0.04	0.007 ± 0.005	23 ± 7	8 ± 3
0.3	0.36 ± 0.05	0.003 ± 0.001	0.0172 ± 0.0004	19 ± 2	10 ± 5
0.7	0.8 ± 0.2	0.0080 ± 0.0003	0.039 ± 0.001	15 ± 5	10.9 ± 0.2
1	1.20 ± 0.07	0.0098 ± 0.0006	0.050 ± 0.002	12.4 ± 0.9	15 ± 1
2	2.0 ± 0.3	0.0122 ±0.0007	0.066 ± 0.006	9 ±2	19.0 ± 0.7
3	2.4 ± 0.3	0.015 ± 0.001	0.091 ± 0.005	10 ± 3	23 ± 3
5	2.6 ± 0.3	0.022 ± 0.003	0.12 ± 0.01	10 ± 1	24 ± 3
7	3.3 ± 1.3	0.020 ± 0.003	0.11 ± 0.01	8 ± 3	24 ± 2

k_T, translocation rate constant; k_{NP}, nonproductive rate constant; k_C, slow step rate constant; m, step size; mk_T, macroscopic rate of translocation

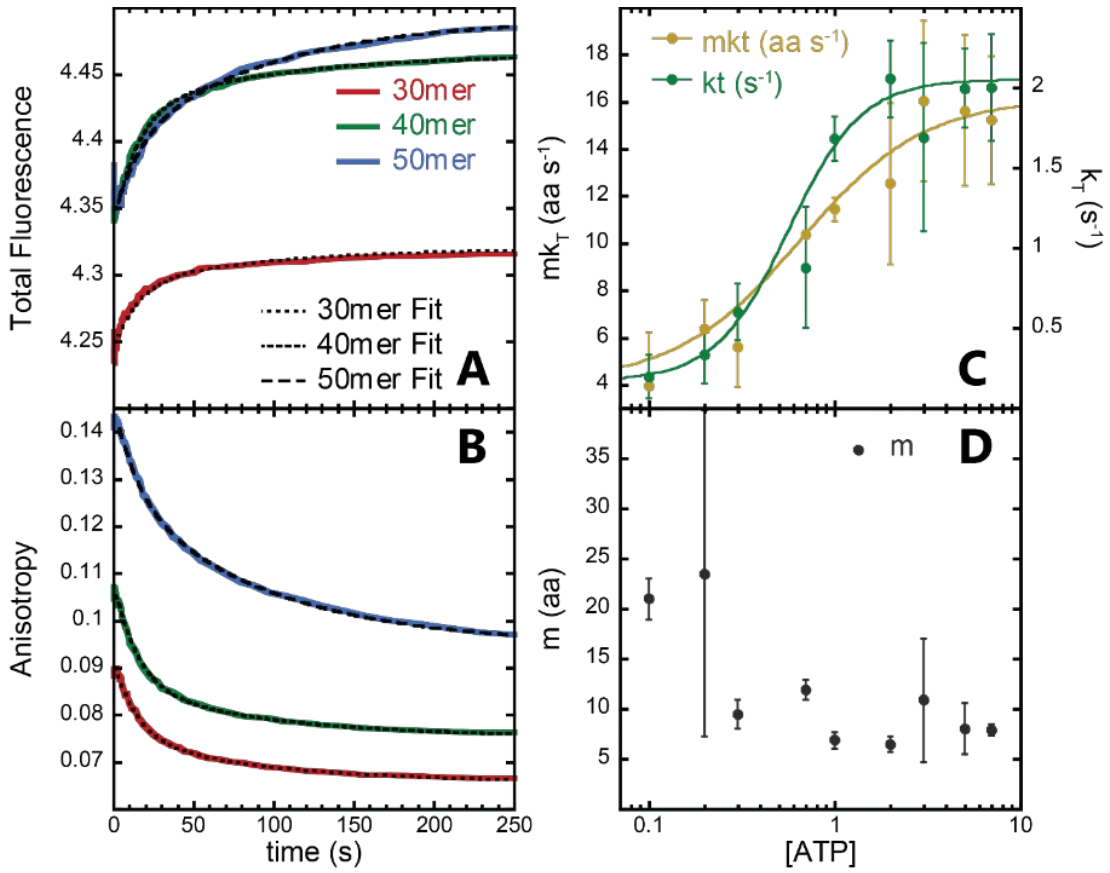


FIGURE 4 Fluorescence anisotropy and total fluorescence time courses for D2, $ClpA_{E286A}$, catalyzed translocation of polypeptide substrates. Representative time courses were collected at 3 mM [ATP] as described in **Fig.1**. Translocation was observed concurrently using (**A**) total fluorescence and (**B**) anisotropy for each time course. The solid colored traces represent the translocation of the 30, 40, and 50 aa substrates (See **Table 1**). The anisotropy and total fluorescence data were subjected to MENOTR global analysis as described in the **Methods** section using **Scheme 1** and **Eq. S1-S4**. The resulting best fits are shown as overlaid black broken traces and the corresponding optimized parameters can be found in **Table 3**. Translocation time courses were collected over a range of [ATP], from 0.1 - 7 mM. Each [ATP] experiment was evaluated using MENOTR global analysis as described in the **Methods** section using **Scheme 1** and **Eq. S1-S4**. The kinetic parameters (**C**) k_T, mk_T, and (**D**) m are shown plotted as a function of [ATP]. Each [ATP] experiment was collected in triplicate, with the figure representing the average and standard deviation of the three. Both the mk_T and k_T isotherms were subjected to NLLS analysis using the Hill equation (**Eq. 11**) and the resulting best fits are shown as a solid trace of the respective color. The fit of the k_T isotherm produced a k_{TD2} max = (1.9 ± 0.3) s^{-1}, K_{aD2} = (1.7 ± 0.4) mM^{-1}, H_{D2} = (2.2 ± 0.8), and b_{D2} = (0.2 ± 0.2) s^{-1}; while mk_T produced a mk_{TD2} max = (12 ± 6) aa s^{-1}, K_{aD2} = (1.5 ± 0.7) mM^{-1}, H_{D2} = (2 ± 1), and b_{D2} = (7.1 ± 2.7) s^{-1}. The [ATP] dependence of m shows that at low [ATP] m approached a value of ~20 aa and as [ATP] increased there is a transition to a value of m of ~7 aa above 0.3 mM ATP.

TABLE 3 Optimized fitting parameters for ClpA$_{E286A}$ catalyzed translocation of polypeptide

[ATP] mM	k_T (s^{-1})	k_{NP} (s^{-1})	k_C (s^{-1})	m (aa)	mkT (aa s^{-1})
0.1	0.2 ± 0.1	0.004 ± 0.006	0.004 ± 0.004	21 ± 2	4 ± 2
0.2	0.3 ± 0.2	0.0015 ± 0.0007	0.006 ± 0.002	20 ± 10	6 ± 1
0.3	0.6 ± 0.2	0.002 ± 0.002	0.013 ± 0.007	9 ± 1	6 ± 2
0.7	0.9 ± 0.4	0.00688 ± 0.00001	0.0378 ± 0.0001	12 ± 1	10.39 ± 0.01
1	1.7 ± 0.1	0.006 ± 0.002	0.036 ± 0.005	6.8 ± 0.8	11.4 ± 0.5
2	2.2 ± 0.3	0.007 ± 0.003	0.04 ± 0.02	6.4 ± 0.8	13 ± 3
3	1.7 ± 0.6	0.010 ± 0.004	0.06 ± 0.04	11 ± 6	16 ± 3
5	2.0 ± 0.2	0.0153 ± 0.0003	0.092 ± 0.003	8 ± 3	16 ± 3
7	2.0 ± 0.3	0.015 ± 0.001	0.08 ± 0.016	7.9 ± 0.6	15 ± 3

k_T, translocation rate constant; k_{NP}, nonproductive rate constant; k_C, slow step rate constant; m, step size; mk_T, macroscopic rate of translocation

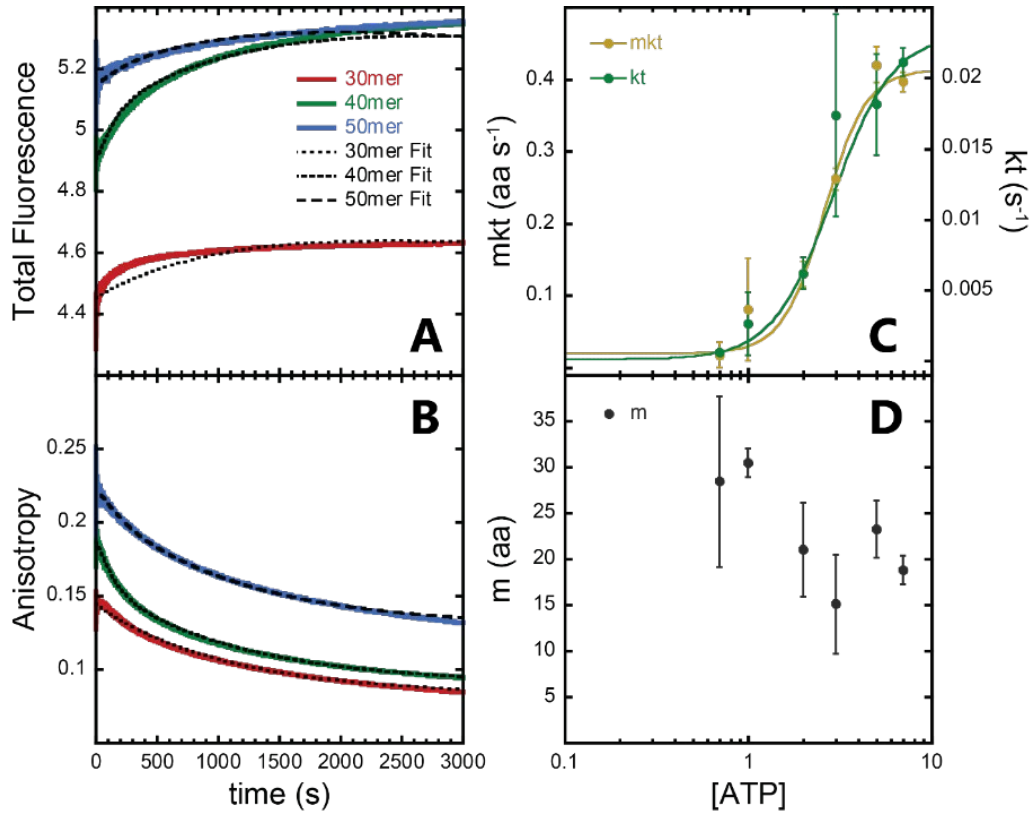


FIGURE 5 Fluorescence anisotropy and total fluorescence time courses for D1, $ClpA_{E565A}$, catalyzed translocation of polypeptide substrates. Representative time courses were collected at 3 mM [ATP] as described in **Fig.1**. Translocation was observed concurrently using (**A**) total fluorescence and (**B**) anisotropy for each time course. The solid colored traces represent the translocation of the 30, 40, 50 aa SsrA tagged substrates (See **Table 1**). The anisotropy and total fluorescence data were subjected to MENOTR global analysis as described in the **Methods** section using **Scheme 1** and **Eq. S5 - S8**. The resulting best fits are shown as overlaid black broken traces and the corresponding optimized parameters can be found in **Table 4**. Translocation time courses were collected over a range of [ATP], 0.7 - 7 mM. Each [ATP] was evaluated using MENOTR global analysis as described in the **Methods** section using **Scheme 1** and **Eq. S5 - S8**. Each [ATP] experiment was evaluated using MENOTR global analysis as described in the **Methods** section using **Scheme 1** and **Eq. S1-S4**. The kinetic parameters (**C**) k_T, mk_T, and (**D**) m are shown plotted as a function of [ATP]. Each [ATP] experiment was collected in triplicate, with the figure representing the average and standard deviation of the three. Both the mk_T and k_T isotherms were subjected to NLLS analysis using the Hill equation (**Eq. 11**) and the resulting best fits are shown as a solid trace of the respective color. The fit of the k_T isotherm produced a k_{TD1} max = (0.023 ± 0.007) s^{-1}, K_{aD1} = (0.3 ± 0.1) mM^{-1}, H_{D1} = (3 ± 2), and b_{D1} = (0 ± 0) s^{-1}; while mk_T produced a mk_{TD1} max = (0.39 ± 0.03) aa s^{-1}, K_{aD1} = (0.38 ± 0.2) mM^{-1}, H_{D1} = (4.0 ± 0.8), and b_{D1} = (0.02 ± 0.02) s^{-1}. The [ATP] dependence of m shows that at low [ATP] m approached a value of ~28 aa and as [ATP] increased there is a transition to a value of m of ~18 aa above 1 mM ATP.

Table 4: Optimized fitting parameters for ClpA$_{E565A}$ catalyzed translocation of polypeptide

[ATP] mM	k_T (s^{-1})	k_{NP} (s^{-1})	k_C (s^{-1})	m (aa)	mkT (aa s^{-1})
0.7	0.00061 ± 0.00010	0.00023 ±0.00004	NA	28 ± 9	0.02 ± 0.02
1	0.003 ± 0.002	0.0006 ± 0.0006	NA	30 ± 2	0.08 ± 0.07
2	0.006 ± 0.001	0.0004 ± 0.0004	NA	21 ± 5	0.12 ± 0.02
3	0.017 ± 0.007	0.0009 ± 0.0002	NA	16 ± 4	0.24 ± 0.02
5	0.018 ± 0.004	0.00127 ± 0.00002	NA	23 ± 3	0.36 ± 0.02
7	0.021 ± 0.001	0.0013 ± 0.0002	NA	19 ± 2	0.35 ± 0.01

k_T, translocation rate constant; k_{NP}, nonproductive rate constant; k_C, slow step rate constant; m, step size; mk_T, macroscopic rate of translocation

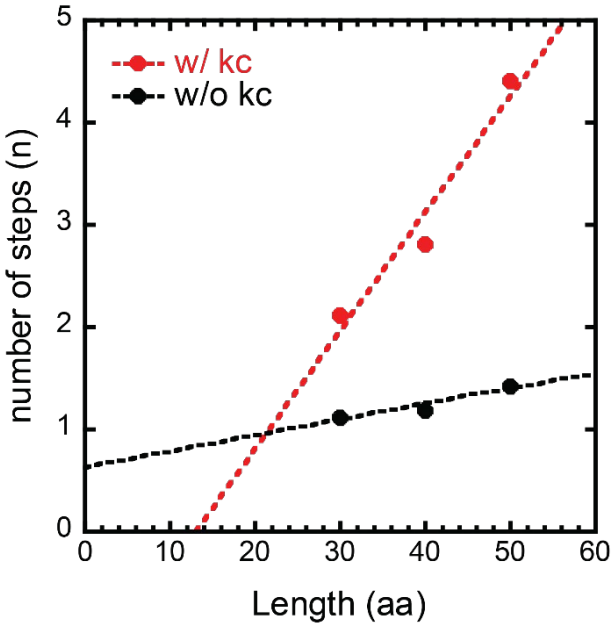
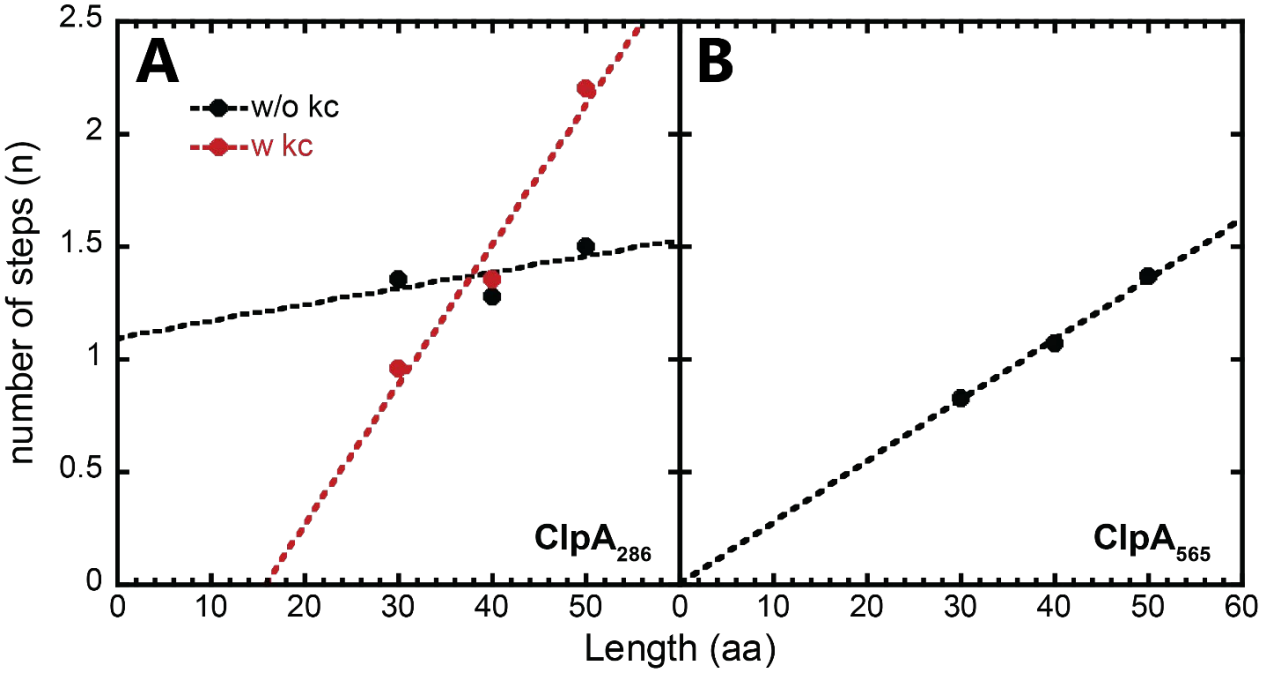


FIGURE S1 The dependence of the number of translocation steps, n, on the length of peptide translocated, L, for $ClpA_{WT}$. The data found in **Fig. 3** were analyzed using a fitting method that either included or excluded the slow step k_C. In this analysis, the kinetic rate constants were set as global parameters while the amplitudes and number of steps were allowed to float as local parameters for the three peptides lengths. The resulting number of translocation steps for each were plotted as a function of substrate length. The black filled circles represent the results of fitting while excluding a k_C step. The dashed black line represents a linear regression of the corresponding number of steps and produces a y-intercept of 0.6 steps. The red filled circles represent the results of fitting with a k_C step. The dashed red line represents a linear regression of the corresponding number of steps and produces a x-intercept of ~11 aa.



FIGU**RE S2** The dependence of the number of translocation steps, n, on the length of peptide translocated, L, for $ClpA_{E286A}$ and $ClpA_{E565A}$. The 3mM data found in **Fig. 4 & 5** were analyzed to evaluate the need for an additional step outside of translocation, k_C. In this analysis, the kinetic rate constants were established as global parameters, while the amplitudes and number of steps were allowed to float as local parameters for the three peptides lengths. The resulting number of translocation steps for each experiment were plotted as a function of substrate length. The black filled circles represent the results of fitting while excluding a k_C step. For 286, a linear regression of the plot produced a y-intercept of 1.1 number of steps, while the intercept for 565 was found to be ~0. Fitting 286 with a k_C step produced the red circles and the corresponding linear regression yielded a x-intercept of ~16 aa. Since 565 was not found to require a k_C step no second fit was performed.

Two-State Model

$$TF(t) = \mathcal{L}^{-1}\left(q_1\left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i\left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^i} + \frac{1-x}{k_{NP}+s}\right) + q_2\left(\frac{1}{s}\left(\frac{k_T^{n-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^n}k_T\right)\right)\right) \quad \text{S1}$$

$$r(t) = \frac{r(t)_{num}}{r(t)_{den}} \quad \text{S2}$$

$$r(t)_{num} = \mathcal{L}^{-1}\left(r_1 q_1\left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i\left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^i} + \frac{1-x}{k_{NP}+s}\right) + r_2 q_2\left(\frac{1}{s}\left(\frac{k_T^{n-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^n}k_T\right)\right)\right) \quad \text{S3}$$

$$r(t)_{den} = \mathcal{L}^{-1}\left(q_1\left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i\left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^i} + \frac{1-x}{k_{NP}+s}\right) + q_2\left(\frac{1}{s}\left(\frac{k_T^{n-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^n}k_T\right)\right)\right) \quad \text{S4}$$

Three-State Model

$$TF(t) = \mathcal{L}^{-1}\left(q_1\left(\frac{1-x}{k_{NP}+s}\right) + q_2\left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i\left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^i}\right) + q_3\left(\frac{1}{s}\left(\frac{k_T^{n-1}k_C^h\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h\left(k_{NP}+s\right)\left(k_T+s\right)^n}k_T\right)\right)\right) \quad \text{S5}$$

$$r(t) = \frac{r(t)_{num}}{r(t)_{den}} \tag{S6}$$

$$r(t)_{num} = \mathcal{L}^{-1}\left(r_1 q_1 \left(\frac{1-x}{k_{NP}+s} \right) + r_2 q_2 \left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i \left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1} k_C^h \left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h \left(k_{NP}+s\right)\left(k_T+s\right)^i} \right) + r_3 q_3 \left(\frac{1}{s}\left(\frac{k_T^{n-1} k_C^h \left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h \left(k_{NP}+s\right)\left(k_T+s\right)^n} k_T \right) \right) \right) \tag{S7}$$

$$r(t)_{den} = \mathcal{L}^{-1}\left(q_1 \left(\frac{1-x}{k_{NP}+s} \right) + q_2 \left(\sum_{i=1}^{h} \frac{k_C^{i-1}\left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^i \left(k_{NP}+s\right)} + \sum_{i=1}^{n} \frac{k_T^{i-1} k_C^h \left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h \left(k_{NP}+s\right)\left(k_T+s\right)^i} \right) + q_3 \left(\frac{1}{s}\left(\frac{k_T^{n-1} k_C^h \left(k_{NP}+s\cdot x\right)}{\left(k_C+s\right)^h \left(k_{NP}+s\right)\left(k_T+s\right)^n} k_T \right) \right) \right) \tag{S8}$$

Conclusions

We have previously reported on the translocation mechanisms of both ClpA and ClpAP using transient-state kinetic assays (1, 2). In those studies, we described the elementary rate constants, overall translocation rate, and the distance traveled per rate limiting step (step-size) in an attempt to characterize the molecular mechanism of translocation catalyzed by both systems. Elucidation of these parameters coupled to structural (3, 4), steady-state ATPase (5), and cross-linking (6) studies lead to the development of a proposed model for polypeptide translocation of both ClpA and ClpAP. In those models we proposed that in the absence of ClpP the rate-limiting translocation step occurs at Domain 1 (D1) (2). While those models represented the most encompassing hypothesis on the translocation activity of ClpA and ClpAP to date, they do not include direct quantitative measurements of translocation at each nucleotide binding domain (NBD).

We have presented in **Chapter 3** the investigation of ClpA Walker B variants that allow for monitoring translocation kinetics at either D1, $ClpA_{E565A}$, or D2, $ClpA_{E286A}$. In these variants, the Walker B glutamate of one of the two NBDs has been mutated such that it lacks the ability to hydrolyze ATP. Thus, we are only sensitive to translocation at the unmodified NBD. Importantly, it has been shown that these modifications do not inhibit ATP binding to the affected domain thereby still allowing for ClpA oligomerization (5, 7).

In **Chapter 2**, we detailed the development and utilization of a method that allows for the observation of ClpA catalyzed polypeptide translocation under transient-state conditions. In this method, ClpA is preassembled with fluorescently modified polypeptide

substrate and is rapidly mixed with ATP and unmodified polypeptide trap in a stopped-flow spectrometer. By preassembling the ClpA-polypeptide complex and maintaining saturating levels of unmodified polypeptide trap, we are able to achieve single-turnover conditions. Meaning, the method is only sensitive to a single round of polypeptide translocation and is independent of effects from ClpA oligomerization or polypeptide binding. Thus, we are sensitive to only the molecular events that occur in the active site of the motor during translocation.

The strength of this technique is its ability to monitor the residence time of motors on a substrate lattice using both fluorescence anisotropy and total fluorescence as detection methods. In this setup, total fluorescence is collected such that rotational artifacts are eliminated from the signal (8), and is consequently defined by only changes in quantum yield. When ClpA is bound to fluorescently modified substrate, there is a marked change in the total fluorescence of the complex. Therefore, changes in total fluorescence between that of bound and unbound substrate can be used to detect the presence of the motor. Fluorescence anisotropy is sensitive to the rotational mobility of the complex, and thus is sensitive to the presence of the motor on the lattice based on changes in the apparent size of the complex (8).

Application of single-turnover stopped-flow fluorescence anisotropy and total fluorescence methods to both Walker B variants allowed for the determination of elementary rate constants, overall translocation rate, and kinetic step-sizes for each NBD of ClpA in the absence of ClpP. It is important to note that all ClpA Walker B variants exhibited reduced solubility in the buffer conditions previously used to examine ClpA and

ClpAP (9). As such, $ClpA_{WT}$ was also analyzed in the final solution conditions under which $ClpA_{E286A}$ and $ClpA_{E565A}$ were found to be stable.

We reported that the kinetic parameters determined for $ClpA_{WT}$ in 500 mM NaCl varied only slightly in comparison to those previously determined in 300 mM NaCl using raw fluorescence techniques. ClpA in 500 mM NaCl was found to translocate polypeptide with an overall rate of ~23 aa s-1 compared to the previously determined rate of ~20 aa s^{-1}. The data suggests that the slight difference in rate is the consequence of the increase observed in the translocation rate constant for ClpA in 500 mM NaCl, from (1.46 ± 0.05) s^{-1} to (2.4 ± 0.3) s^{-1}, respectively. The determined step-size of ClpA catalyzed translocation under both conditions were found to be approximately the same with m = (10 ± 3) aa and (13.4 ± 0.5) aa, for 500 and 300 mM NaCl, respectively.

The investigation of polypeptide translocation catalyzed by $ClpA_{E286A}$ revealed that D2 is responsible for the majority of the translocation activity present in $ClpA_{WT}$. It was found D2 translocated polypeptide substrates with an overall rate of (16 ± 3) aa s^{-1}, which is ~70 % of that of ClpAWT in 500 mM NaCl. Interestingly, Kress et al. concluded that the ATPase activity assays for the same mutant revealed that D2 hydrolyzed ATP with a k_{cat} ~80 % of that of $ClpA_{WT}$ (5). The similarity between these numbers is not surprising given ATP binding, hydrolysis, and release is directly coupled to ClpA catalyzed translocation.

In stark contrast to the D2 findings, the results for the translocation experiments performed on ClpAE565A revealed that D1 has only minimal translocation activity. At 3 mM ATP the overall rate of translocation was found to be (0.24 ± 0.02) aa s^{-1} which is less than 1 % of the activity of $ClpA_{WT}$. The translocation rate constant and kinetic step-

size were determined at (0.017 ± 0.007) s^{-1} and (16 ± 4) aa, respectively. Thus, suggesting that the observed reduction in rate is a product of a reduced rate constant and not a change in step-size. Again, comparing our findings to those of Kress et al., those researchers reported that $ClpA_{E565A}$ turned over ATP at ~10 % to that of ClpWT (5). While D1 is still competent in polypeptide translocation our evidence in conjunction with ATPase (5), FRET unfolding (10), and single molecule molecular tweezer assay (11) suggests that D1 plays a secondary role in ClpA catalyzed translocation.

We have previously shown that by subjecting ClpA to translocation experiments over a range of [ATP] we can tease out mechanistic information regarding what step in the repeating translocation cycle is rate limiting. We found that for ClpA, the step immediately following ATP binding is rate limiting and that ATP binding is positively cooperative (1). This is determined by plotting the kinetic parameters determined as a function of [ATP] and analyzing the trends. The translocation rate constant and overall rate of translocation where both found to fit well to an infinity cooperative Hill equation suggesting cooperativity with respect to ATP binding (1, 2). Additionally, the plot of kinetic step-size was determined to have no ATP dependence indicating that only a single step is being monitored in each repeating cycle of translocation, and that the step being observed is coupled to ATP binding (12, 13).

Here we found that $ClpA_{WT}$, $ClpA_{E286A}$, and $ClpA_{E56A}$ all exhibited positive cooperativity in their rate and rate constants with respect to [ATP]. Fitting each of these two isotherms to a Hill equation produced Hill coefficients of ~2 for all the variants which is consistent with positive cooperativity between ~2-3 binding sites in the hexamer. The observation that cooperativity was not lost when deactivating hydrolysis in one of the two

NBD suggests that the cooperativity is inter-monomer as only domains on adjacent monomers were participating in translocation. This hypothesis of inter-monomer cooperativity was also suggested by Kress et al. (5).

In the development of total fluorescence and fluorescence anisotropy stopped-flow methods in **Chapter 2** we tested to see if we could further interrogate the processivity of ClpA. We have previously assumed that ClpA is a processive motor under saturating ATP concentrations because dissociation during translocation was not detected. However, a quantitative estimate of processivity was not possible (1, 14). We proposed that polypeptide substrates of lengths longer than 100 aa were necessary to probe the processivity of ClpA as its processivity was estimated at ~100 aa translocated per binding event. Using fluorescence anisotropy and total fluorescence we were able for the first time to monitor and analyze the kinetics on substrates longer than 100 aa. With those results we were able to conclude that ClpA is translocating at least 127 aa per binding event which updates our estimate of processivity up to $P > 0.893$.

These methods are applicable to the investigations of many other AAA+ motor proteins that reside on lattices like ClpA. Specifically, these techniques that can be used with motors that don't associate with partner proteins, don't covalently modify their substrate, and are unaltered. Meaning, the wildtype activities of traditionally hard to study motors can be directly monitored and quantified. Thus, combined fluorescence anisotropy and total fluorescence represent a new strategy of overcoming some of the limitations of the traditional kineticists toolkit.

All the analysis performed in the studies presented in **Chapters 2 & 3** was carried out using a custom-built MATLAB toolbox called MENOTR. MENOTR is a hybrid genetic and NLLS algorithm that can be used to analyze various biochemical data sets. In **Chapter 1** we showed that it was able to reproduce the kinetic results of two previously published analyses. Moreover, it generated kinetic parameters that produced statistically better fits of the data compared to the published results.

The toolbox was written to overcome the common issues associated with parameter correlation and user guess bias while also minimizing the amount of user intervention required to operate the analysis. Its application to the large data set presented in **Chapter 3** allowed for the coupling of high-performance computing, algorithm parallelization and minimal user intervention to make a previously tedious task more palatable by automating the process and fitting on supercomputer clusters. Meaning the experimentalist could apply a "set and forget" approach to data fitting allowing them to focus their attention elsewhere.

While this toolbox was intensively used in the investigation of n-step sequence kinetic mechanisms it is widely applicable to a variety of tasks. So far it has been applied at UAB in the Lucius and Schneider labs in investigating the ClpA, ClpB, Hsp104, Pol I, and Pol II systems. Additionally, it has been used in the Lohman and Millard labs at WUSL and GT to study RecBCD, ssBP, and Protein Kinase A. As such, it represents a robust tool that is able to overcome the obstacles facing many types of fitting problems and is designed in a way that it can be easily converted to meet the needs of different investigators and labs.

Future Directions

While we have made meaningful progress in defining the mechanism that describes translocation at each NBD of ClpA, questions remain. We have shown previously that ClpP allosterically affects that translocation mechanism of ClpA in the ClpAP system (2). Moreover, we presented hypotheses that suggested that these effects were the consequence of ClpP upregulating D2 of ClpA while simultaneously downregulating D1. With the initial investigations of the ClpA Walker B variants, it is now possible to apply the same kinetic experiments to $ClpA_{E286A}P$ and $ClpA_{E565A}P$ to directly test those hypotheses. This would include rigorous evaluations of the [ATP] dependence of the mechanisms that control translocation at each domain.

While we have presented methods that can be applied to polypeptide substrates greater than 100 aa, no detectable dissociation rate constant was determined on the 102 and 127 aa substrates interrogated in **Chapter 2**. The use of even longer substrates, between 150 and 200 aa, under conditions that favor one-to-one ClpA-substrate binding could be reasonably used to further probe processivity. Moreover, we are now poised to begin examining folded proteins with specific ClpA binding sequences. Thus, the combined total fluorescence and anisotropy technique will allow us to determine the impact of folded structures on the polypeptide translocation mechanism, which was also not possible with previous methods.

References

1. Rajendar, B., and A. L. Lucius. 2010. Molecular mechanism of polypeptide translocation catalyzed by the Escherichia coli ClpA protein translocase. J Mol Biol 399(5):665-679.

2. Miller, J. M., J. Lin, T. Li, and A. L. Lucius. 2013. E. coli ClpA Catalyzed Polypeptide Translocation is Allosterically Controlled by the Protease ClpP. Journal of Molecular Biology 425(15):2795-2812.

3. Bohon, J., L. D. Jennings, C. M. Phillips, S. Licht, and M. R. Chance. 2008. Synchrotron protein footprinting supports substrate translocation by ClpA via ATP-induced movements of the D2 loop. Structure 16(8):1157-1165.

4. Farbman, M. E., A. Gershenson, and S. Licht. 2007. Single-Molecule Analysis of Nucleotide-Dependent Substrate Binding by the Protein Unfoldase ClpA. J Am Chem Soc.

5. Kress, W., H. Mutschler, and E. Weber-Ban. 2009. Both ATPase domains of ClpA are critical for processing of stable protein structures. J Biol Chem 284(45):31441-31452.

6. Hinnerwisch, J., W. A. Fenton, K. J. Furtak, G. W. Farr, and A. L. Horwich. 2005. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. Cell 121(7):1029-1041.

7. Duran, E. C., and A. L. Lucius. 2019. Examination of the nucleotide linked assembly mechanism of E. coli ClpA. Protein Sci.

8. Lakowicz, J. R. 1999. Principles of fluorescence spectroscopy. Kluwer Academic/Plenum, New York.

9. Duran, E. C., and A. L. Lucius. 2018. ATP hydrolysis inactivating Walker B mutation perturbs E. coli ClpA self-assembly energetics in the absence of nucleotide. Biophys Chem 242:6-14.

10. Baytshtok, V., T. A. Baker, and R. T. Sauer. 2015. Assaying the kinetics of protein denaturation catalyzed by AAA+ unfolding machines and proteases. Proc Natl Acad Sci U S A 112(17):5377-5382.

11. Kotamarthi, H. C., R. T. Sauer, and T. A. Baker. 2020. The Non-dominant AAA+ Ring in the ClpAP Protease Functions as an Anti-stalling Motor to Accelerate Protein Unfolding and Translocation. Cell Rep 30(8):2644-2654.e2643.

12. Lucius, A. L., and T. M. Lohman. 2004. Effects of temperature and ATP on the kinetic mechanism and kinetic step-size for E.coli RecBCD helicase-catalyzed DNA unwinding. J Mol Biol 339(4):751-771.

13. Lucius, A. L., N. K. Maluf, C. J. Fischer, and T. M. Lohman. 2003. General methods for analysis of sequential "n-step" kinetic mechanisms: application to single turnover kinetics of helicase-catalyzed DNA unwinding. Biophys J 85(4):2224-2239.

14. Duran, E. C., C. L. Weaver, and A. L. Lucius. 2017. Comparative Analysis of the Structure and Function of AAA+ Motors ClpA, ClpB, and Hsp104: Common Threads and Disparate Functions. Front Mol Biosci 4:54.