12-2022

# Using Molecular Dynamics Simulations to Decipher Mechanistic Details of Biomolecular Processes of Biology and Biotechnology Oriented Applications

Adithya Polasa
*University of Arkansas, Fayetteville*

Using Molecular Dynamics Simulations to Decipher Mechanistic Details of Biomolecular Processes of Biology and Biotechnology Oriented Applications

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Cell and Molecular Biology

by

Adithya Polasa
Anurag group of Institutions
Bachelor of Pharmacy, 2015
Long Island University
Master of Science in Biomedical Science, 2017

December 2022
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Mahmoud Moradi, Ph.D.
Dissertation Director

Bob Beitle Jr., Ph.D.
Committee member

Yong Wang, Ph.D.
Committee member

Suresh Thallapuranam, Ph.D.
Committee member

**Abstract**

Researchers in chemistry and biology often utilize computer simulations, in conjunction with experimental data, to model and predict the structures, energies, kinetics, processes, and functions of the systems that are their focus of study, ranging from single molecules to whole viruses. Here, we use molecular dynamics (MD) techniques to gain a deeper understanding of biomolecular processes in biology and biotechnology-oriented applications. Using a mixture of equilibrium and non-equilibrium MD simulations, this work describes the insertion process of YidC at the atomic level. In order to better comprehend the insertion process, several docking models of YidC-Pf3 in the lipid bilayer were created. We employed conventional molecular dynamics simulations combined with a non-equilibrium approach to undertake a thorough analysis of the conformational difference between the two docking models produced. Our results show that this high-affinity association requires the YidC transmembrane (TM) groove, and that the hydrophilic YidC groove is crucial for protein trafficking through the cytoplasmic membrane bilayer to the periplasmic side. The Pf3 coat protein is mechanistically affected by conformational changes in the YidC TM domain and membrane core at various phases of the insertion process. The hydration and dehydration of the YidC's hydrophilic groove are also crucial throughout the insertion phase. These findings show that during insertion, Pf3's coat protein interacts with the membrane and YidC in a variety of conformational states. Finally, the thorough investigation directly supports YidC's role as a independent insertase. We have performed the first comprehensive analysis of the gram-negative bacterial YidC protein using microsecond-level all-atom MD simulations. By using equilibrium MD simulations, this study clarifies the relevance of many domains in the YidC structure at the atomic level. The purpose of this study was to describe the crucial function of the C2 loop and the periplasmic domain found in gram-negative YidC, which is lacking in its gram-positive counterpart. Various models of YidC embedded in the lipid bilayer were created. According to our findings, the C2 loop stabilizes the protein overall, especially in the transmembrane region, and it also has an allosteric effect on the periplasmic domain. Important intra- and inter-domain interactions that support the protein's stability and functionality have been identified. Lastly, we also

provide a unique computational method for predicting the trend of the size and activity of peptide-directed nanoparticles by calculating the peptide's binding affinity to a single ion. To examine how the solitary, Green Fluorescent Protein (GFP)-fused peptides behave differently from one another, we used MD simulations. The usual nanoparticle sizes observed from images taken with a transmission electron microscope are in line with the binding free energies we estimated for palladium (Pd). Additionally, computationally estimated Pd binding affinities match Stille coupling and Suzuki-Miyaura reaction turnover frequencies (TOF). The findings demonstrate that although employing Pd4 and its two known variations (A6 and A11) alone results in nanoparticles of various sizes, fusing these peptides to the GFPuv protein results in nanoparticles with comparable sizes and activity. To put it another way, the GFPuv makes the nanoparticles less sensitive to the peptide sequence. In this study, a computational framework for making peptides that are both free and attached to proteins is shown. This makes it easier to make nanoparticles with well-defined properties.

**Acknowledgement**

First and foremost, I would want to express my gratitude to Dr. Mahmoud Moradi, my adviser, for providing me with the chance to work in his lab and for always being there for me during good and difficult times. His exceptional character and dedication to serious scientific inquiry really impressed me. With his methodical and patient direction, I learnt how to plan and carry out investigations. I also had access to a wide range of cutting-edge technology in his lab. He has been a true mentor to me during my PhD studies, and it is a pleasure for me to work with him. For their assistance and support, I also like to thank Dr. Bob Beitle Jr., Dr. Suresh Kumar Thallapuranam, and Dr. Yong Wang. from my committee.

Second, I would like to express my gratitude to the University of Arkansas's Cell and Molecular Biology program as well as its Chemistry and Biochemistry department for providing excellent settings for both academic study and scientific investigation.

Thirdly, I would like to express my gratitude to my lab mates, Dylan Ogden, Kalyan Immadisetty, Vivek Govind Kumar, Ugochi Isu, Shadi A Badiee, Matthew Brownd, Mortaza Derakhshani, James Losey, Ehsaneh Khodadadi, Stephanie Sauve, and Joseph Williamson all of whom were very helpful to me, not just in my lab work but also in other aspects of my life. I would like to thank Imann Mosleh for providing experimental results for the nanoparticle project.

In addition, I also want to express my deep affection for the pleasant city of Fayetteville, Arkansas.

I would like to extend my deepest and most sincere gratitude to my family, which includes my parents, Ashok Polasa and Rama Polasa, as well as my brothers, Shiva Krishna Polasa and Mani Krishna Polasa. If it weren't for their unwavering love and support, I wouldn't be able to list all of the things I've accomplished to this point. I would want to express my gratitude to my uncle Laxmi Veggalam and my aunt Sowmya Veggalam for all of the assistance and advice they have given me throughout my professional life.

Many Thanks to Vikram, Naveen, and Vinay

**Table of contents**

**List of published/submitted papers**

1. Chapter 1 – **Polasa, A.**, Hettige, J. J., Immadisetty, K., & Moradi, M. (2022). An investigation of the YidC-mediated membrane insertion of Pf3 coat protein using molecular dynamics simulations. Frontiers in Molecular Biosciences, 9, 2296-889X. https://doi.org/10.3389/fmolb.2022.954262

2. Chapter 2 – **Polasa**, **A.,** and Moradi, M. (2022) Deciphering the Inter-domain Decou-pling in the Gram-negative Bacterial Membrane Insertase, bioRxiv 2022.08.09.503346. https://doi.org/10.1101/2022.08.09.503346.

3. Chapter 3 – **A. Polasa**, I. Mosleh, J. Losey, A. Abbaspourrad, R. Beitle, and M. Moradi. (2022). "Developing a Rational Approach to Designing Recombinant Proteins for Peptide-Directed Nanoparticle Synthesis," Nanoscale Advances. 4, 3161. https://doi.org/10.1039/D2NA00212D

## Introduction

Researchers in the fields of chemistry and biology now routinely combine computer simulations with a wide range of experimental information to learn and make predictions about the structures, energies, kinetics, mechanisms, and functions of many fascinating and important systems, which can range in size from single molecules to entire viruses [1, 2]. The steady development of new technologies and their applications, such as the use of large simulation programs and state-of-the-art methodologies, now allow us to observe the molecules of life in action, resulting in a scientific revolution just as impactful as the introduction of light microscopes and X-ray techniques in the seventeenth and nineteenth centuries [1]. The ability to use bio-molecular modeling and simulation in both basic and applied research has allowed us to propose and answer new questions and explore tough challenges. Examples of challenges that need to be addressed include elucidating the pathways by which proteins fold or function, identifying new therapeutic targets for common human diseases, and creating novel materials and pharmaceuticals [1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12].

Atomistic molecular dynamics (MD) simulation have become a standard technique for investigating the behavior of proteins in explicit aqueous and membrane environments [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. This is one of the reasons why all-atom MD is so effective. The vast majority of MD-based studies begin with the presumption that local conformational changes which are observed in brief simulations at the nanosecond level can be used to accurately describe functionally relevant large-scale conformational transitions that take place over much longer timescales [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. This study focuses on the use of diverse computational methodologies in decoding the mechanistic intricacies of bio-molecular processes in biology and biotechnology related applications.

This study used microsecond-level MD simulations to examine the conformational changes relevant to the membrane insertion process of YidC in gram-positive and gram-negative bacteria. During the process of membrane protein co-translation, about 33 percent of all membrane proteins are inserted and embedded in the plasma membrane bilayer [13, 14, 15, 16, 17, 18, 19]. YidC, Oxa1, and Alb3 are membrane proteins that collaborate to fold incoming peptides into the

membrane in the most time and energy-effective manner feasible [20, 21, 22, 23, 24, 25, 26]. In the course of a laboratory investigation, it was shown that the membrane insertion and folding processes of over sixty different cytoplasmic membrane proteins were significantly inserted when YidC was not present [20, 21, 22, 23, 24, 25, 26]. Gram-negative and gram-positive bacterial YidC are believed to promote membrane insertion simply by binding nascent chains and promoting their insertion into the lipid bilayer using cytoplasmic loop interactions, hydrophobic force, and interactions in the groove [27, 28, 29, 30] with the last one being of particular interest to this work. A large number of previous researchers have suggested a variety of possible explanations for the YidC independent process [31, 32, 33, 34, 35, 36, 30]. However, the global and local structural changes that take place in YidC as a result of the process are not entirely described at this time. It is not known how the cytoplasmic hairpin region of YidC and the water molecule that is located within the groove area of YidC contribute to the insertion process in any way. How would the structure and conformation of the peptide or protein that was being processed change throughout this process? In this study, we focused on examining the Sec-independent insertion of gram-negative and gram-positive YidCs. According to the findings of our study, the C2 loop of gram-positive and gram-negative YidC is important for the overall stability of the protein, most notably in the transmembrane region, and has an allosteric effect on the periplasmic domain of gram-negative YidC. We have identified important interactions that contribute to the protein's stability and functionality. Lastly, our research proposes a SecY-independent insertion mechanism for both gram-positive and gram-negative bacterial YidC.

Recent developments in the field of nanotechnology have led to the creation of a wide variety of effective processes for the production and characterization of nanoparticles. Within the past two decades in particular, many peptides have been introduced to identify inorganic metal surfaces [37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50]. Some of these peptides have been utilized to produce nanoparticles. The production of nanoparticles of variable sizes, shapes, or aggregation stabilities by employing peptide immobilization has applications in several sub-fields of biotechnology, including sensors and bioanalytical techniques [51, 52, 53]. The configuration

of the nanoparticle assemblies produced by this method have the potential to be sensitive to even minute changes in the content and sequence of the peptide conjugate. It has recently come to light that the composition of the nanoparticles can be regulated by carefully manipulating the peptide sequence [54, 55, 56].

The size and shape of nanoparticles have a significant impact on their physical and chemical characteristics. Nanoparticles (NPs) may take on a variety of shapes and sizes [57, 58, 59]. It is probable that the design of the spatial arrangement of NPs and the structure of their interfaces will help in detecting correlations between structure and property as well as boosting performance. In the field of nanotechnology, atomistic simulation techniques such as MD, which offer a physical insight into understanding phenomena on an atomic scale and make it possible to predict some properties of nanomaterials, have emerged as powerful tools. These techniques provide a physical insight into understanding the phenomena on an atomic scale. This is due to the fact that these methodologies make it possible to comprehend the occurrences on a scale that is even smaller than atoms. It is anticipated that atomistic simulations might increasingly play a vital role as an alternative to experimental approaches in the field of nanomaterials. This is due to the fact that the experimental investigation of nanomaterials would unavoidably face numerous technological challenges, not to mention concerns regarding the cost.

The primary objective of this section of our research is to construct a computational model that will allow us to investigate the peptide-nanoparticle interaction in the context of free peptides as well as peptides that are bound to proteins. We do not claim to have a technique that completely explains all elements of the interaction between peptides and nanoparticles: indeed, the interaction with a single Pd ion is the only subject of our attention in this body of work. It is interesting to note that the Pd binding free energies that were measured for several systems, including isolated and GFPuv-fused Pd4 as well as its two known mutants (H6A and H11A, denoted A6 and A11 respectively), provided a good predictor of the behavior of the resulting nanoparticles, both in terms of their sizes and their activities.

## References


[1] T. Schlick and S. Portillo-Ledesma, "Biomolecular modeling thrives in the age of technology," *Nature Computational Science*, vol. 1, no. 5, pp. 321–331, 2021.

[2] L. Casalino, A. C. Dommer, Z. Gaieb, E. P. Barros, T. Sztain, S.-H. Ahn, A. Trifan, A. Brace, A. T. Bogetti, A. Clyde, H. Ma, H. Lee, M. Turilli, S. Khalid, L. T. Chong, C. Simmerling, D. J. Hardy, J. D. C. Maia, J. C. Phillips, T. Kurth, A. C. Stern, L. Huang, J. D. McCalpin, M. Tatineni, T. Gibbs, J. E. Stone, S. Jha, A. Ramanathan, and R. E. Amaro, "AI-driven multiscale simulations illuminate mechanisms of SARS-CoV-2 spike dynamics," *The International Journal of High Performance Computing Applications*, vol. 35, no. 5, pp. 432–451, 2021.

[3] V. Govind Kumar, D. S. Ogden, U. H. Isu, A. Polasa, J. Losey, and M. Moradi, "Prefusion Spike Protein Conformational Changes Are Slower in SARS-CoV-2 than in SARS-CoV-1," *Journal of Biological Chemistry*, vol. 298, 2022.

[4] K. Immadisetty, A. Polasa, R. Shelton, and M. Moradi, "Elucidating the Molecular Basis of pH Activation of an Engineered Mechanosensitive Channel," *bioRxiv*, 2021.

[5] A. Polasa, I. Mosleh, J. Losey, A. Abbaspourrad, R. Beitle, and M. Moradi, "Developing a Rational Approach to Designing Recombinant Proteins for Peptide-Directed Nanoparticle Synthesis," *Nanoscale Adv.*, vol. 4, pp. 3161–3171, 2022.

[6] M. Moradi, G. Enkavi, and E. Tajkhorshid, "Atomic-level characterization of transport cycle thermodynamics in the glycerol-3-phosphate:phosphate transporter," *Nat. Commun.*, vol. 6, p. 8393, 2015.

[7] M. Moradi, C. Sagui, and C. Roland, "Calculating relative transition rates with driven nonequilibrium simulations," *Chem. Phys. Lett.*, vol. 518, pp. 109–113, 2011.

[8] S. C. Moradi, Mahmoud and C. Roland, "Investigating rare events with nonequilibrium work measurements: I. nonequilibrium transition path probabilities," *J. Chem. Phys.*, vol. 140, no. 3, p. 034114, 2014.

[9] M. Moradi and E. Tajkhorshid, "Mechanistic picture for conformational transition of a membrane transporter at atomic resolution," *Proc. Natl. Acad. Sci. USA*, vol. 110, no. 47, pp. 18916–18921, 2013.

[10] D. Ogden and M. Moradi, "Molecular Dynamics–Based Thermodynamic and Kinetic Characterization of Membrane Protein Conformational Transitions," *Structure and Function of Membrane Proteins*, pp. 289–309, 2021.

[11] H. Chen, D. Ogden, S. Pant, W. Cai, E. Tajkhorshid, M. Moradi, B. Roux, and C. Chipot, "A Companion Guide to the String Method with Swarms of Trajectories: Characterization, Performance, and Pitfalls," *Journal of Chemical Theory and Computation*, vol. 18, pp. 1406–1422, 2022.

[12] K. Immadisetty, A. Polasa, R. Shelton, and M. Moradi, "Elucidating the Molecular Basis of Spontaneous Activation in an Engineered Mechanosensitive Channel," *Computational and Structural Biotechnology Journal*, vol. 20, pp. 2539–2550, 2022.

[13] S. Nagamori, I. N. Smirnova, and H. R. Kaback, "Role of YidC in folding of polytopic membrane proteins," *Journal of Cell Biology*, vol. 165, no. 1, pp. 53–62, 2004.

[14] N. E. Lewis and L. J. Brady, "Breaking the bacterial protein targeting and translocation model: Oral organisms as a case in point," *Molecular Oral Microbiology*, vol. 30, no. 3, pp. 186–197, 2015.

[15] T. Serdiuk, D. Balasubramaniam, J. Sugihara, S. A. Mari, H. R. Kaback, and D. J. Müller, "YidC assists the stepwise and stochastic folding of membrane proteins," *Nature Chemical Biology*, vol. 12, no. 11, pp. 911–917, 2016.

[16] S. Kol, B. R. Turrell, J. De Keyzer, M. Van Der Laan, N. Nouwen, and A. J. Driessen, "YidC-mediated membrane insertion of assembly mutants of subunit c of the F1F0 ATPase," *Journal of Biological Chemistry*, vol. 281, no. 40, pp. 29762–29768, 2006.

[17] M. Van der Laan, M. L. Urbanus, C. M. Ten Hagen-Jongman, N. Nouwen, B. Oudega, N. Harms, A. J. Driessen, and J. Luirink, "A conserved function of YidC in the biogenesis of respiratory chain complexes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 10, pp. 5801–5806, 2003.

[18] L. Yi and R. E. Dalbey, "Oxa1/Alb3/YidC system for insertion of membrane proteins in mitochondria, chloroplasts and bacteria," *Molecular Membrane Biology*, vol. 22, no. 1-2, pp. 101–111, 2005.

[19] E. Van Bloois, G. Jan Haan, J. W. De Gier, B. Oudega, and J. Luirink, "F1F0 ATP synthase subunit c is targeted by the SRP to YidC in the E. coli inner membrane," *FEBS Letters*, vol. 576, no. 1-2, pp. 97–100, 2004.

[20] R. E. Dalbey and A. Kuhn, "Membrane Insertases Are Present in All Three Domains of Life," *Structure*, vol. 23, no. 9, pp. 1559–1560, 2015.

[21] M. A. McDowell, M. Heimes, and I. Sinning, "Structural and molecular mechanisms for membrane protein biogenesis by the Oxa1 superfamily," *Nature Structural & Molecular Biology*, vol. 28, no. 3, pp. 234–239, 2021.

[22] F. Jiang, M. Chen, L. Yi, J. W. De Gier, A. Kuhn, and R. E. Dalbey, "Defining the regions of Escherichia coli YidC that contribute to activity," *Journal of Biological Chemistry*, vol. 278, no. 49, pp. 48965–48972, 2003.

[23] K. J. Nass, I. M. Ilie, M. J. Saller, A. J. M. Driessen, A. Caflisch, R. A. Kammerer, and X. Li, "The role of the N-terminal amphipathic helix in bacterial YidC: Insights from functional studies, the crystal structure and molecular dynamics simulations," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1864, no. 3, p. 183825, 2022.

[24] B. Güngör, T. Flohr, S. G. Garg, and J. M. Herrmann, "The ER membrane complex (EMC) can functionally replace the Oxa1 insertase in mitochondria," *PLOS Biology*, vol. 20, p. e3001380, mar 2022.

[25] T. A. Rapoport, "Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes," *Nature*, vol. 450, no. 7170, pp. 663–669, 2007.

[26] A. Krogh, B. Larsson, G. Von Heijne, and E. L. Sonnhammer, "Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes," *Journal of Molecular Biology*, vol. 305, no. 3, pp. 567–580, 2001.

[27] R. E. Dalbey, A. Kuhn, L. Zhu, and D. Kiefer, "The membrane insertase YidC," *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1843, no. 8, pp. 1489–1496, 2014.

[28] D. Kiefer and A. B. T. I. R. o. C. Kuhn, "YidC as an Essential and Multifunctional Component in Membrane Protein Assembly," *Academic Press*, vol. 259, pp. 113–138, 2007.

[29] A. Gallusser and A. Kuhn, "Initial steps in protein membrane insertion. Bacteriophage M13 procoat protein binds to the membrane surface by electrostatic interaction," *The EMBO journal*, vol. 9, no. 9, pp. 2723–2729, 1990.

[30] A. Polasa, J. Hettige, K. Immadisetty, and M. Moradi, "An Investigation of the YidC-Mediated Membrane Insertion of Pf3 Coat Protein Using Molecular Dynamics Simulations," *bioRxiv*, 2022.

[31] Y. Chen, S. Capponi, L. Zhu, P. Gellenbeck, J. A. Freites, S. H. White, and R. E. Dalbey, "YidC Insertase of Escherichia coli: Water Accessibility and Membrane Shaping," *Structure*, vol. 25, no. 9, pp. 1403–1414.e3, 2017.

[32] R. J. Schulze, J. Komar, M. Botte, W. J. Allen, S. Whitehouse, V. A. M. Gold, J. A. L. a Nijeholt, K. Huard, I. Berger, C. Schaffitzel, and I. Collinson, "Membrane protein insertion and proton-motive-force-dependent secretion through the bacterial holo-translocon SecYEG-SecDF-YajC-YidC," *Proceedings of the National Academy of Sciences*, vol. 111, no. 13, pp. 4844–4849, 2014.

[33] Y. Chen, R. Soman, S. K. Shanmugam, A. Kuhn, and R. E. Dalbey, "The role of the strictly conserved positively charged residue differs among the gram-positive, gram-negative, and chloroplast YidC homologs," *Journal of Biological Chemistry*, vol. 289, no. 51, pp. 35656–35667, 2014.

[34] S. Funes, F. Kauff, E. O. Van Der Sluis, M. Ott, and J. M. Herrmann, "Evolution of YidC/Oxa1/Alb3 insertases: Three independent gene duplications followed by functional specialization in bacteria, mitochondria and chloroplasts," *Biological Chemistry*, vol. 392, no. 1-2, pp. 13–19, 2011.

[35] S. Funes, A. Hasona, H. Bauerschmitt, C. Grubbauer, F. Kauff, R. Collins, P. J. Crowley, S. R. Palmer, L. J. Brady, and J. M. Herrmann, "Independent gene duplications of the YidC/Oxa/Alb3 family enabled a specialized cotranslational function," *Proceedings of the*

*National Academy of Sciences of the United States of America*, vol. 106, no. 16, pp. 6656–6661, 2009.

[36] N. Shimokawa-Chiba, K. Kumazaki, T. Tsukazaki, O. Nureki, K. Ito, and S. Chiba, "Hydrophilic microenvironment required for the channel-independent insertase function of YidC protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 16, pp. 5063–5068, 2015.

[37] D. J. Gaskin, K. Starck, and E. N. Vulfson, "Identification of inorganic crystal-specific sequences using phage display combinatorial library of short peptides: A feasibility study," *Biotechnology Letters*, vol. 22, no. 15, pp. 1211–1216, 2000.

[38] S. R. Whaley, D. S. English, E. L. Hu, P. F. Barbara, and A. M. Belcher, "Selection of peptides with semiconductor binding specificity for directed nanocrystal assembly," *Nature*, vol. 405, no. 6787, pp. 665–668, 2000.

[39] R. R. Naik, S. E. Jones, C. J. Murray, J. C. McAuliffe, R. A. Vaia, and M. O. Stone, "Peptide templates for nanoparticle synthesis derived from polymerase chain reaction-driven phage display," *Advanced Functional Materials*, vol. 14, no. 1, pp. 25–30, 2004.

[40] R. R. Naik, S. J. Stringer, G. Agarwal, S. E. Jones, and M. O. Stone, "Biomimetic synthesis and patterning of silver nanoparticles," *Nature Materials*, vol. 1, no. 3, pp. 169–172, 2002.

[41] K. I. Sano and K. Shiba, "A Hexapeptide Motif that Electrostatically Binds to the Surface of Titanium," *Journal of the American Chemical Society*, vol. 125, no. 47, pp. 14234–14235, 2003.

[42] S. Wang, E. S. Humphreys, S. Y. Chung, D. F. Delduco, S. R. Lustig, H. Wang, K. N. Parker, N. W. Rizzo, S. Subramoney, Y. M. Chiang, and A. Jagota, "Peptides with selective affinity for carbon nanotubes," *Nature Materials*, vol. 2, no. 3, pp. 196–200, 2003.

[43] M. B. Dickerson, R. R. Naik, M. O. Stone, Y. Cai, and K. H. Sandhage, "Identification of peptides that promote the rapid precipitation of germania nanoparticle networks via use of a peptide display library," *Chemical Communications*, vol. 4, no. 15, pp. 1776–1777, 2004.

[44] C. K. Thai, H. Dai, M. S. Sastry, M. Sarikaya, D. T. Schwartz, and F. Baneyx, "Identification and characterization of Cu2O- and ZnO-binding polypeptides by escherichia coli cell surface display: Toward an understanding of metal oxide binding," *Biotechnology and Bioengineering*, vol. 87, no. 2, pp. 129–137, 2004.

[45] C. Mao, D. J. Solis, B. D. Reiss, S. T. Kottmann, R. Y. Sweeney, A. Hayhurst, G. Georgiou, B. Iverson, and A. M. Belcher, "Virus-Based Toolkit for the Directed Synthesis of Magnetic and Semiconducting Nanowires," *Science*, vol. 303, no. 5655, pp. 213–217, 2004.

[46] B. D. Reiss, C. Mao, D. J. Solis, K. S. Ryan, T. Thomson, and A. M. Belcher, "Biological routes to metal alloy ferromagnetic nanostructures," *Nano Letters*, vol. 4, no. 6, pp. 1127–1132, 2004.

[47] R. Zuo, D. Örnek, and T. K. Wood, "Aluminum- and mild steel-binding peptides from phage display," *Applied Microbiology and Biotechnology*, vol. 68, no. 4, pp. 505–509, 2005.

[48] M. Umetsu, M. Mizuta, K. Tsumoto, S. Ohara, S. Takami, H. Watanabe, I. Kumagai, and T. Adschiri, "Bioassisted room-temperature immobilization and mineralization of zinc oxide - The structural ordering of ZnO nanoparticles into a flower-type morphology," *Advanced Materials*, vol. 17, no. 21, pp. 2571–2575, 2005.

[49] M. Sarikaya, C. Tamerler, A. K. Jen, K. Schulten, and F. Baneyx, "Molecular biomimetics: Nanotechnology through biology," *Nature Materials*, vol. 2, no. 9, pp. 577–585, 2003.

[50] J. M. Slocik and R. R. Naik, "Biologically programmed synthesis of bimetallic nanostructures," *Advanced Materials*, vol. 18, no. 15, pp. 1988–1992, 2006.

[51] J. J. Storhoff, R. Elghanian, R. C. Mucic, C. A. Mirkin, and R. L. Letsinger, "One-pot colorimetric differentiation of polynucleotides with single base imperfections using gold nanoparticle probes," *Journal of the American Chemical Society*, vol. 120, no. 9, pp. 1959–1964, 1998.

[52] S. Liang, D. T. Pierce, C. Amiot, and X. Zhao, "Photoactive nanomaterials for sensing trace analytes in biological samples," *Synthesis and Reactivity in Inorganic, Metal-Organic and Nano-Metal Chemistry*, vol. 35, no. 9, pp. 661–668, 2005.

[53] S. G. Penn, L. He, and M. J. Natan, "Nanoparticles for bioanalysis," *Current Opinion in Chemical Biology*, vol. 7, no. 5, pp. 609–615, 2003.

[54] P. R. Selvakannan, S. Mandal, S. Phadtare, A. Gole, R. Pasricha, S. D. Adyanthaya, and M. Sastry, "Water-dispersible tryptophan-protected gold nanoparticles prepared by the spontaneous reduction of aqueous chloroaurate ions by the amino acid," *Journal of Colloid and Interface Science*, vol. 269, no. 1, pp. 97–102, 2004.

[55] S. Si and T. K. Mandal, "Tryptophan-based peptides to synthesize gold and silver nanoparticles: A mechanistic and kinetic study," *Chemistry - A European Journal*, vol. 13, no. 11, pp. 3160–3168, 2007.

[56] R. R. Bhattacharjee, A. K. Das, D. Haldar, S. Si, A. Banerjee, and T. K. Mandal, "Peptide-assisted synthesis of gold nanoparticles and their self-assembly," *Journal of Nanoscience and Nanotechnology*, vol. 5, no. 7, pp. 1141–1147, 2005.

[57] I. Mosleh, M. Benamara, L. Greenlee, M. H. Beyzavi, and R. Beitle, "Recombinant peptide fusion proteins enable palladium nanoparticle growth," *Materials Letters*, vol. 252, pp. 68–71, 2019.

[58] I. Mosleh, H. R. Shahsavari, R. Beitle, and M. H. Beyzavi, "Recombinant Peptide Fusion Protein-Templated Palladium Nanoparticles for Suzuki-Miyaura and Stille Coupling Reactions," *ChemCatChem*, vol. 12, no. 11, pp. 2942–2946, 2020.

[59] Tejada-Vaprio, Rita and Mosleh, Imann and Mukherjee, Rudra Palash and Aljewari, Hazim and Fruchtl, McKinzie and Elmasheiti, Ahmed and Bedford, Nicholas and Greenlee, Lauren and Beyzavi, M. Hassan and Beitle, Robert, "Recombinant peptide fusion construction for protein-templated catalytic palladium nanoparticles," *Biotechnology Progress*, vol. 36, no. 3, p. 3, 2020.

# Chapter I: An Investigation of the YidC-Mediated Membrane Insertion of Pf3 Coat Protein Using Molecular Dynamics Simulations

Adithya Polasa[1], Jeevapani Hettige[1], Kalyan Immadisetty[1] and Mahmoud Moradi[1] *

[1]Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701

## Abstract

YidC is a membrane protein that facilitates the insertion of newly synthesized proteins into lipid membranes. Through YidC, proteins are inserted into the lipid bilayer via the SecYEG-dependent complex. Additionally, YidC functions as a chaperone in protein folding processes. Several studies have provided evidence of its independent insertion mechanism. However, the mechanistic details of the YidC SecY-independent protein insertion mechanism remain elusive at the molecular level. This study elucidates the insertion mechanism of YidC at an atomic level through a combination of equilibrium and non-equilibrium molecular dynamics (MD) simulations. Different docking models of YidC-Pf3 in the lipid bilayer were built in this study to better understand the insertion mechanism. To conduct a complete investigation of the conformational difference between the two docking models developed, we used classical molecular dynamics simulations supplemented with a non-equilibrium technique. Our findings indicate that the YidC transmembrane (TM) groove is essential for this high-affinity interaction and that the hydrophilic nature of the YidC groove plays an important role in protein transport across the cytoplasmic membrane bilayer to the periplasmic side. At different stages of the insertion process, conformational changes in YidC's TM domain and membrane core have a mechanistic effect on the Pf3 coat protein. Furthermore, during the insertion phase, the hydration and dehydration of the YidC's hydrophilic groove are critical. These results demonstrate that Pf3 coat protein interactions with the membrane and YidC vary in different conformational states during the insertion process. Finally, this extensive study directly confirms that YidC functions as an independent insertase.

## Introduction

Approximately 33% of all membrane proteins are inserted and embedded in the plasma membrane bilayer during co-translation [1, 2]. The membrane proteins YidC/Oxa1/Alb3 work to fold incoming peptides into the membrane as efficiently as possible [3, 4, 5, 6, 7]. In an experimental study, over sixty cytoplasmic membrane proteins were found whose membrane insertion/folding is substantially hindered in the absence of YidC [8]. YidC catalyzes the transmembrane insertion of newly synthesized membrane proteins in the absence of an energy supply domain, such as an ATPase [9], and is also involved in the insertion and placement of membrane proteins in microbes. The extent to which insertase proteins are required for inserting proteins into the membrane has been thoroughly investigated. They can be found in all kingdoms of life and are necessary for cell viability [10, 11, 12, 13]. They are adaptable proteins and can function along with the SecYEG pathway to insert peptides into the membrane through the Signal Recognition Particle (SRP) mechanism. They can fold and insert proteins independently of the Sec pathway [10, 14, 15, 16, 17, 18, 19, 20]. This study primarily focuses on the conformational dynamics of YidC, including both local and global conformational changes involved in the insertion process of the Pf3 coat protein.

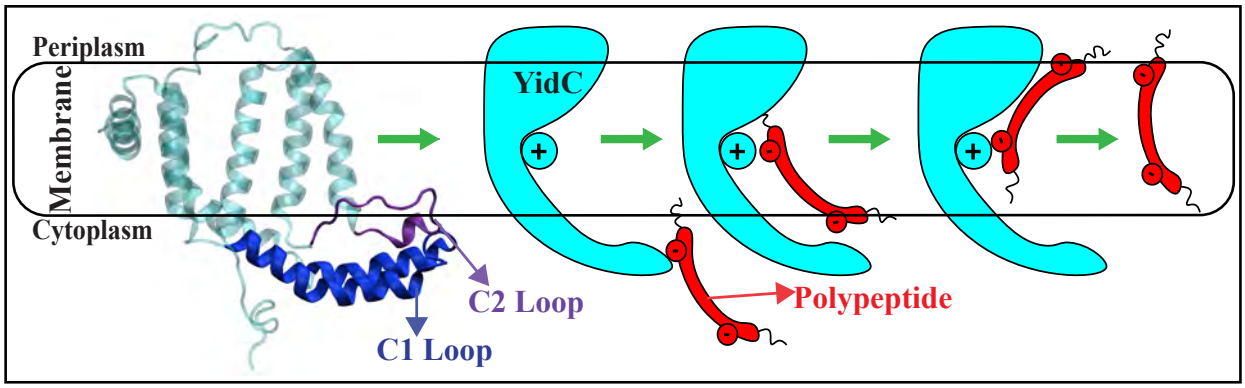


Figure 1: The cartoon representation of YidC and the schematic illustration of the SecY-independent insertion mechanism. A cartoon representation of YidC's C1 and C2 loops on the cytoplasmic side (left). Schematic illustration of the YidC Sec-independent insertion of polypeptide (right).

YidC completes its function either independently as a membrane insertase or as a chaperone in the SecYEG complex mechanism. In an experimental study, the deletion of YidC resulted in a

conformational change of LacY during the insertion process by the SecYEG complex [21]. Hence, YidC plays a critical role in the insertion of the LacY lactose permease membrane layer protein [22, 16, 23, 24, 25]. Also, YidC is involved in the incorporation of subunit II of cytochrome o oxidase in *E.Coli* [26, 27, 28]. Initially, the Sec-autonomous pathway was thought to function without the contribution of an insertase. However, many studies have demonstrated that YidC is fundamental for the addition of small phage coat proteins like Pf3 and M13 in a Sec-free pathway [10, 25, 29, 30, 31, 32, 33, 34].

A few experimental studies have explored the role of YidC in various microbial organisms. The genomes of most gram-positive microscopic organisms encode two YidC proteins: YidC1 and YidC2 [35, 36]. Although YidC typically exists as a dimer or tetramer [37] under physiological conditions, it was discovered that YidC can also function as a monomer in lipid bilayers [9, 25, 38, 39].

The protein is firmly anchored in the lipid bilayer by interfacial aromatic residues, a cytoplasmic salt-bridge group, and a periplasmic helix enhanced with aromatic residues. The aromatic residues above the R72 amino acid in YidC from Bacillus halodurans may offer a polar hydrophobic surface for the insertion of peptides into the lipid bilayer [40, 41].

The C–terminus of monomeric YidC cooperates with the ribosomes, and the short interhelical loops come into contact with the ribosomal proteins [42]. YidC is believed to promote membrane insertion simply by binding nascent chains and promoting their insertion into the lipid bilayer using hydrophobic force [9]. The hydrophilic groove inside the membrane core of the YidC will increase the rate of accepting the hydrophilic moieties of a substrate into the membrane [43, 44]. The YidC hydrophilic region traverses the inner side of the membrane and is closed to the periplasmic side of the bilayer. This decreases the hydrophobicity of the membrane towards the external side of the lipid bilayer. This hypothesis of the YidC mechanism provides excellent opportunity to study the conformational dynamics of YidC. In the first step, it interacts with a hydrophilic protein region temporarily in its groove, and in the second step, this peptide is translocated to the periplasmic side [18].

Many prior studies have reported various explanations of the YidC independent mechanism. However, the global and local structural changes that occur in YidC during the process are not completely defined. It's unknown how the cytoplasmic hairpin region of YidC and the water molecule inside the groove area of YidC take action during the insertion process. How would the incoming peptide or protein's structure and conformation change during the process? We examined this topic using a combination of docking, classical molecular dynamics, and non-equilibrium simulations to analyze Sec-independent YidC's (Fig. 1) insertion of the Pf3 coat protein into the membrane. In this study, we looked at the local and global conformational changes of YidC associated with Pf3 coat protein insertion into the hydrophilic groove, Pf3 coat protein interactions with YidC and the membrane, and conformational changes in Pf3 coat protein that occurred during the insertion process.

## Methods

The crystal structure of YidC (PDB:3WO7 [45]) was downloaded from the Protein Data Bank. Initially, the system was prepared using the Molecular Operating Environment (MOE) software [46] by removing the water molecules from the crystal structure and assigning the appropriate protonation states for the residues using the protonate3D facility.

We used MOE software to create two docking structures of the Pf3 coat protein interacting with YidC based on the previously hypothesized stages involved in the YidC sec-independent insertion process [9, 47, 48]. In pose1, Pf3 coat protein is docked in the YidC's hydrophilic groove (Fig. 2A) to evaluate probable interactions and conformational changes in the mechanism's initial phase. The Pf3 coat protein is docked near to the periplasmic side (Fig. 2A) of the protein in pose2 to identify the interactions and conformational changes involved in the mechanism's final phase. Biased and unbiased all-atom molecular dynamics (MD) simulations were performed to characterize the conformational differences of the two bacterial YidC2 - Pf3 docking models pose1 and pose2 (Fig. 2A) in a membrane environment.

All simulations were performed with the NAMD 2.13 [49] using the CHARMM36m [50] force

field [51]. TIP3P [52] waters were used to solvate the protein. YidC was inserted into the lipid bilayer, solvated, and ionized using the membrane builder on CHARMM-GUI [53]. In these MD studies, we used palmitoyloleoyl phosphatidylethanolamine (POPE) lipids to build a lipid bilayer. A membrane layer surface of 110 Å $\times$ 110 Å was built along the XY plane. The protein lipid-assembly was solvated in water with 25 Å thick layers of water on top and bottom. 0.15 M of $Na^+$ and $Cl^-$ ions were added to the solution with a slight modification in the number of ions to neutralize the system. The solvated system contained $\approx$142000 atoms. Before the equilibrium simulation, the structure was energy minimized using the conjugate gradient algorithm [54]. Following that, we used the standard CHARMM-GUI [53] protocol to progressively relax the systems using restrained MD simulations. In the NPT ensemble at 310 K, 550 ns of equilibrium MD simulations were performed under periodic boundary conditions for each system. In the simulations, a Langevin integrator with a damping coefficient of $\gamma$=0.5 $ps^{-1}$ and 1 atm pressure was maintained using the Nose-Hoover Langevin piston method [55, 56].

Trajectories were visualized and analyzed using VMD software [57]. Salt bridge interaction analysis was conducted via VMD plugins. For salt bridge analysis, the cut-off distance was set at 4 Å and the distance between the oxygen atoms of the acidic residues and the nitrogen atoms of basic residues were calculated. The interhelical angles were calculated as the angle between the third principal axes of the corresponding helices [58, 59]. The TM helices and other subdomains were defined for analysis as follows: TM1a (58$-$78), TM1b (79$-$104), TM2 (134$-$155), TM3 (175$-$190), TM4 (219$-$233), TM5 (233$-$258), C1 region (84$-$133), C2 loop (195$-$216), and modified C-terminal region (256$-$272) respectively. The number of contacts within 3 Å of selection was measured for contact analysis. We counted the number of water molecules within 5 Å of R72 for water analysis. For the Pf3 coat protein bending angle, we chose two pairs of residues selection for the top (ASP7-ASP18) and bottom (ASP18-LEU29) regions of Pf3 coat protein and measured their third principal axes, denoted by $v_1$ and $v_2$, respectively. The angle between the two vectors was calculated as $180^\circ - \arccos(\frac{v_1 \cdot v_2}{|v_1||v_2|})$. Principal component analysis (PCA) was performed on each trajectory using PRODY [60] software. Only $C_\alpha$ atoms of the

peptide were considered in the PCA calculations of both docking simulations. The VMD plugin MEMBPLUGIN was used to calculate the area per lipid and the deterium order parameter of equilibrated POPE membranes [61].

The combination of equilibrium and non-equilibrium MD simulations has proven effective for investigating biological challenges [62, 63, 64, 65, 66, 67, 68, 69, 13]. In this work, the YidC independent insertion mechanism was studied using non-equilibrium targeted MD (TMD) as implemented within the colvars module of NAMD [70].

A TMD simulation was performed on the final conformation of the pose 1 system obtained from the 550 ns equilibrium trajectory in order to transfer the Pf3 coat protein to the periplasmic side of the membrane, as seen in pose 2. The RMSD collective variable was used in the TMD simulation. As a collective variable, we used the RMSD of Pf3 coat protein backbone atoms from the last frame of pose2's equilibrium simulation trajectory. The TMD simulation was run for 100 ns with a force constant of 44 $kcal/mol/\text{Å}^2$. To ensure conformational accuracy, the final frame of the targeted MD simulation was equilibrated for 20 ns without any restraints.

## Results and Discussion

### YidC Undergoes Major Conformational Changes in Sec-independent Insertion Process.

A protein must undergo various local and global conformational changes in a mechanism. A set of approximate docking models (Fig. 2A) of YidC and Pf3 coat protein were developed to represent the insertion process.The docking postures created for this study were constructed based on hypotheses previously proposed in the literature [71, 9, 47, 48], that C1 and C2 loop regions could detect the substrate initially during YidC SecY-independent insertion. Because of electrostatic and hydrophilic interactions between the substrate and YidC, the substrate is then momentarily trapped within the YidC groove. Following that, the captured substrate protein is transferred from the cavity into the membrane through hydrophobic interactions between membrane lipids and the protein. We utilized MOE docking software to produce 20 distinct docking positions that were sorted according to docking scores. For the simulations, we chose the top posture (i.e., pose1) and the

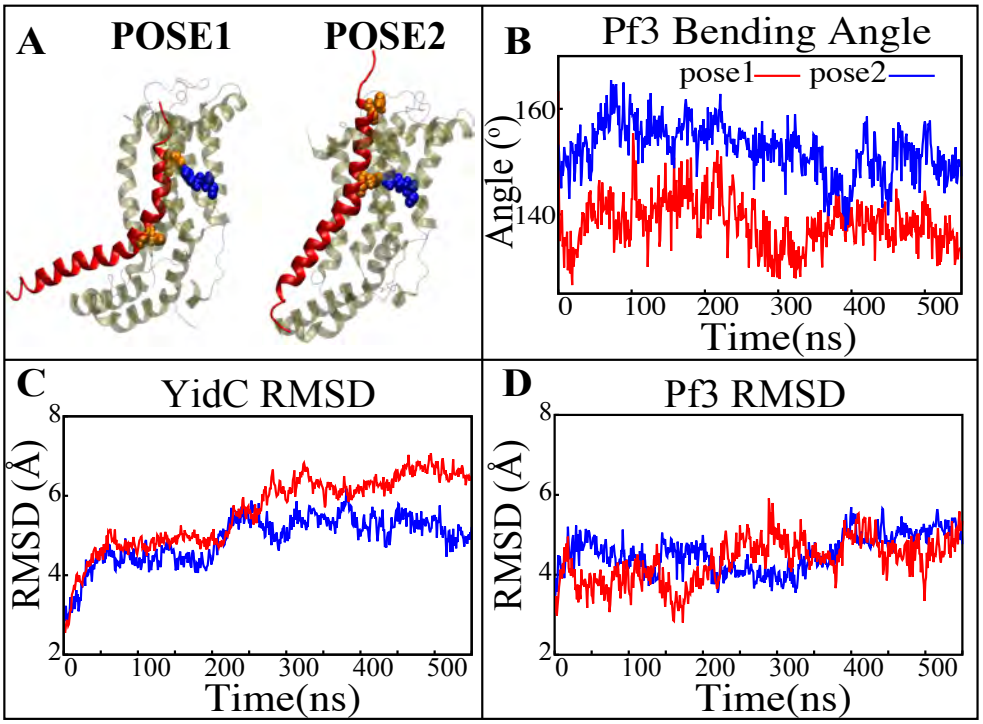


Figure 2: Structural stability evaluation of YidC and Pf3 coat protein in the insertion process. (A) Cartoon representation of pose1 & pose2 docking models generated based on the YidC (PDB:3WO7) Pf3 coat protein insertion process, as described previously in methods. (B) The bending angle of the Pf3 coat protein helix in pose1 (red) and pose2 (blue) models shown as a function of time. (C & D) Root mean square deviation of the YidC & Pf3 coat protein in pose1 (red) and pose2 (blue) models. Based on RMSD results, we have observed that in pose2 stage the YidC is significantly more stable compared to the pose1 state.

eighth pose (i.e., pose2) from the 20 docked positions. The top posture was chosen since it was rated first by MOE and represented an intermediate stage of insertion. On the other hand, the 8th pose was selected since it was the highest ranked pose among the ones representing a late stage of insertion. MD simulations of these docking models were performed to investigate various conformational characteristics to see how YidC and Pf3 coat protein altered conformational properties during the insertion process. Several measures or quantities linked to local and global conformational properties were evaluated and monitored for the two conformational poses developed in this study. The C$\alpha$ root mean square deviation (RMSD) of the systems were evaluated independently of the framework to test the impact of the Pf3 coat protein on YidC's global structure. It was found that the Pf3 coat protein had a relatively comparable RMSD in the two models simulated in this

study (Fig. 2D). However, the YidC protein fluctuated more in pose 1 (Fig. 2C) than in pose 2 (Fig. 2C).

At the start of a process or mechanism, a protein is anticipated to undergo substantial conformational changes. The fact that YidC's RMSD in pose 1 (Fig. 2C) is 2 Å greater than that in pose 2 (Fig. 2C) suggests that YidC goes through significant conformational changes at the start of the process. In a recent computational analysis reported by us, the RMSD of the YidC without Pf3 coat protein was determined to be less than 4 Å, which is lower than what we observed in the presence of Pf3 coat protein [72]. The lower RMSD in the absence of Pf3 coat protein supports our hypothesis that YidC protein undergoes major conformational changes in the presence of Pf3 coat protein. This demonstrates that the effect of Pf3 coat protein insertion into the membrane differs depending on the stage of the process. Although we see comparable RMSD of Pf3 coat protein in both poses, the Pf3 coat protein bending angle analysis (Fig. 2A) more clearly suggests a structural difference between the two states in support of our hypothesis.

The bending angle indicates that Pf3 coat protein has a lower bend at the start of the insertion process and changes its conformation inside the YidC groove (Fig. 2A) as it progresses deeper into the groove. This brings us to the conclusion, that Pf3 coat protein undergoes major conformational changes to adapt to the YidC groove environment during the insertion process. In the next investigation, additional local components of YidC were rigorously investigated to elucidate more details of the insertion process.

**Widening of the Transmembrane Domain is Essential for Incorporation of Proteins in Membrane During the Insertion Process.**

Previous studies have revealed that the YidC transmembrane (TM) region is crucial for membrane protein insertion into the lipid bilayer [73, 74]. The helical angle between each TM pair was determined in this study. In the pose2 docking simulation, the transmembrane helices (TM1a, TM2, TM3, TM4, and TM5) seem more slanted than in the pose1 docking simulation, which has a difference in the angle of over 10 degrees (Fig. 3B-G). This implies that the central TM

groove of YidC is substantially enlarged during the insertion of Pf3 coat protein into the membrane bilayer. The critical helices TM1a (Fig. 3E,F,&G) and TM2 (Fig. 3B,C,&D) undergo significant modifications following peptide insertion because they are stretched onto the cytoplasmic side of the membrane, which is the entrance point of Pf3 coat protein. Based on this, one may assume that throughout the insertion process, YidC experiences a gradual and tranquil conformational shift, organically adjusting to the incoming peptide. In this scenario, Yidc progressively expands its transmembrane groove to make room for the incoming peptide and then returns to a normal state once the peptide is fully incorporated into the membrane.

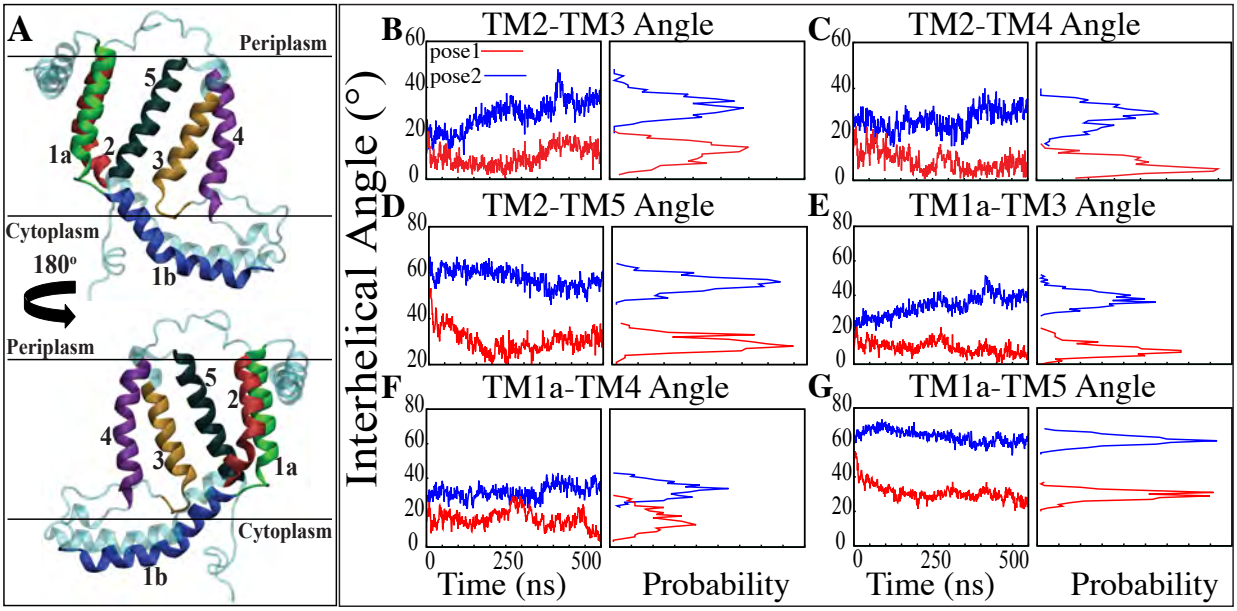


Figure 3: Inter-helical angles between trans-membrane helices of YidC in both docking model simulations. (A) Graphical representation of YidC protein on periplasmic, cytoplasmic and transmembrane regions labeled with helical numbers in the transmembrane region. (B-D) Overall inter-helical angle between transmembrane helix 2 and other helices of the protein in pose 1 (red) and pose 2 (blue) simulations. (E-G) Overall inter-helical angle between transmembrane helix 1a and other helices of the protein in pose 1 and pose 2 simulations. Also, the probability density distribution is shown for all graphs.

The interactions of the membrane and YidC with the Pf3 coat protein were studied to learn more about the insertion process. The number of interactions of Pf3 coat protein with YidC and the membrane within 3 Å were taken into account for the interaction analysis (Fig. 4A&B). The contact of the Pf3 coat protein with the membrane determines its position in the bilayer system; the

Pf3 coat protein positioned inside YidC's hydrophilic groove has a slightly higher lipid interaction distribution than the Pf3 coat protein positioned just outside the groove area (Fig. 4B). Because the Pf3 coat protein is now ready to be incorporated into the membrane, Pf3 coat protein has a high degree of contact with the membrane in pose2. Previous experiments have shown that the hydrophobic interaction between the substrate and the lipid aliphatic chains can make it easier for a substrate to get inside the membrane [75, 48, 76, 77]. Here, our analysis shows that the N-terminus of the Pf3 coat protein interacts with lipids better in pose2 than in pose1 (Fig. S3A). As the Pf3 coat protein progresses through the insertion mechanism, it establishes hydrophobic interactions with aliphatic chains, and these hydrophobic interactions could aid in the insertion of the protein. The divergence of Pf3 coat protein lipid interactions supports the suggested mechanistic models

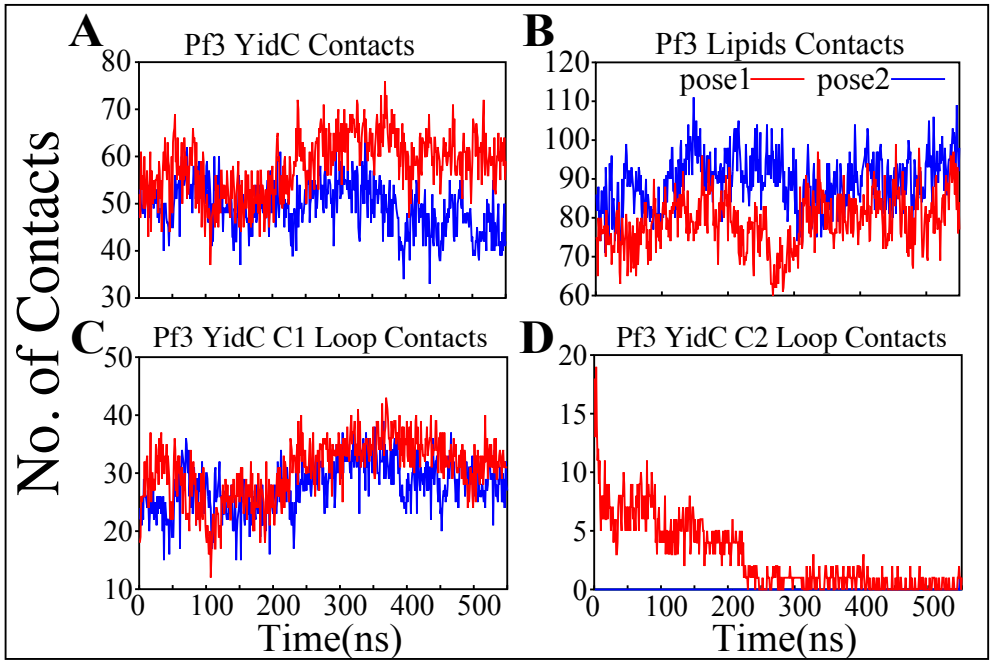


Figure 4: Pf3 coat protein overall interaction with YidC and POPE lipid tails in both the docking model simulations. (A & B) Respective number of YidC and lipid interactions with the Pf3 coat protein in pose1 (red) and pose2 (blue), shown as a function of time. (C & D) Number of contacts between Pf3 coat protein and the C1 and C2 loops of YidC, shown as a function of time.

for the YidC independent insertion pathway in this study. In addition to lipid interactions, the interactions between YidC and Pf3 coat protein are also important in this process. The distribution of such interactions should confirm the outcomes of the lipid interactions. Because of the greater

distribution of lipid contacts in the pose2 model, Pf3 coat protein decreases its interaction with the YidC protein by shifting closer to the lipid bilayer (Fig. 4B). However, at the pose1 stage of the insertion process, the association of YidC and Pf3 coat protein should be significantly strengthened before establishing the peptide in YidC's hydrophilic groove. This would explain the YidC higher interactions with the Pf3 coat protein that were observed in the pose1 model (Fig. 4A).

**Interaction of C1 and C2 Loops with Pf3 Coat Protein Stabilizes the Insertion Process.**

Cytoplasmic loops C1 and C2 (Fig. 1A) are important components in YidC's independent insertion mechanism. Previous studies on YidC with or without the C2 loop found that its presence stabilizes YidC in the membrane [72]. In both the pose1 and pose2 models, the YidC cytoplasmic loop C1 interacts with the Pf3 coat protein. This interaction aids in the stability of the Pf3 coat protein inside YidC's hydrophilic groove. Furthermore, these loops establish a strong interaction to retain the incoming proteins inside YidC's U-shaped groove (Fig. 4C&D). At the beginning of the insertion mechanism, the cytoplasmic C1 loop, which is deeply expanded into the cytoplasmic side, creates extremely strong contacts with the Pf3 coat protein (Fig. 4C). These interactions between the C1 loop and the peptide are critical for keeping the peptide under control during the insertion process. According to our contact analysis results, the C2 loop contacts (Fig. 4D) are formed only in the pose1 model, since it is the starting point of the insertion process, and a high number of protein interactions are necessary to stabilize such a long peptide. As the process progresses, the C2 loop loses its interactions (Fig. 4D) with the incoming peptide once the peptide is settled inside the U-shaped groove of YidC, as seen in pose2. Thus far, we have shown that YidC undergoes significant global and local conformational changes, such as TM domain expansion, and its interactions with Pf3 coat protein, specifically through contacts in the cytoplasmic loop region. All the findings presented above confirm the major hypothesis of the study on YidC conformational changes throughout the independent insertion process.

Principal component analysis (PCA) was used to identify the key differences between the pose1 and pose2 models. Pose1 and pose2 systems were clearly differentiated by projections onto prin-

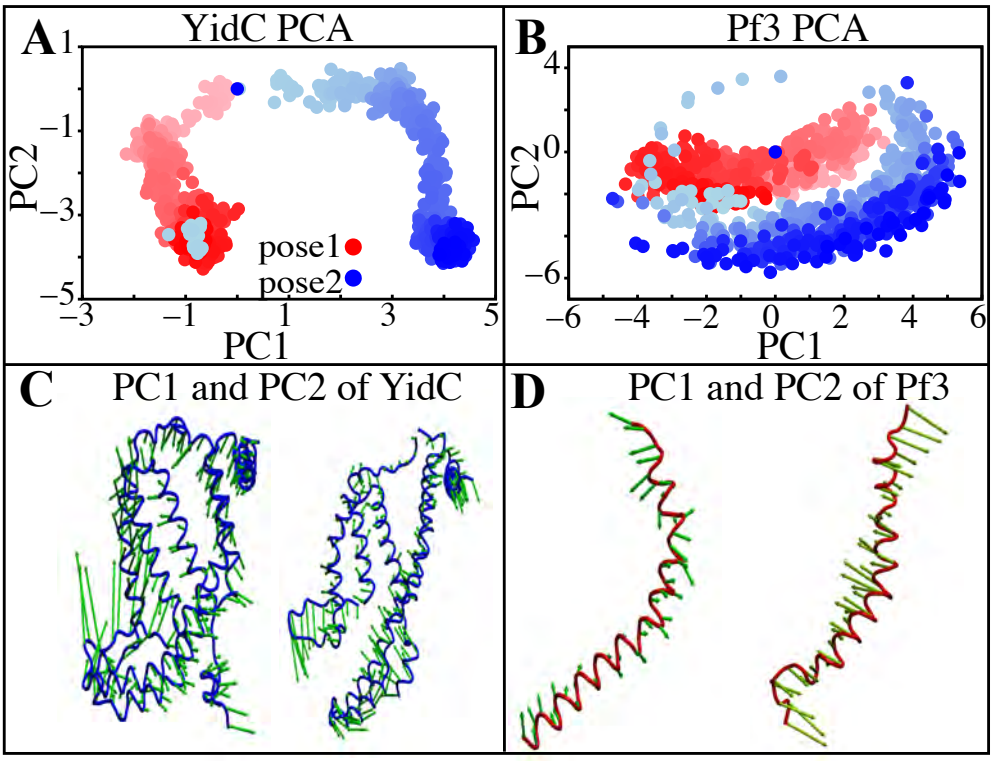


Figure 5: Principal component projections along PCs 1 and 2. (A & B) PC1 vs PC2 for pose1(red) and pose2(blue) models of YidC & Pf3 coat protein. (C) Structural variation in PC1 and PC2 of YidC, respectively. (D) Structural variations in PC1 and PC2 of Pf3, respectively. The bidirectional arrow shows the direction of the fluctuation of the structure and the length of the arrow reflects the magnitude of the fluctuation. The color shading in the picture indicates a timeline, with light and dark shades representing the beginning and end of the simulation, respectively.

cipal components (PCs) 1 and 2. Only YidC $C_\alpha$ atoms are considered in this study. PC1 and PC2 contributed 49.9 and 18.8 percent of the total variance, respectively (Fig. 5A). As expected, the structural analysis of pose1 and pose2 models contradicts each other in PC1 and PC2, which is logical given the significant conformational differences (Fig. 5A) observed previously. The Pf3 coat protein, on the other hand, has clustered similarly along PC1 (Fig. 5B). However, the fluctuation of Pf3 coat protein is different around PC2 (Fig. 5B), which may be the result of a shift in Pf3 coat protein interactions and conformational changes (PC1 and PC2 contributed 45.9 and 25.4 percent of the total variance). To demonstrate this visually, square displacements of PC residues were projected onto the structure, as seen in Fig. 5C-D. Overall, the major finding of the PC analy-

sis was that the behavior of the pose1 and pose2 proteins differed considerably. This confirms our previous notion that YidC conformational dynamics play an important role in the insertion process. The PCA results are consistent with the early evidence for global and local structural changes.

**YidC's Hydrophilic Groove Hydration and Dehydration are Critical in the Insertion Mechanism.**

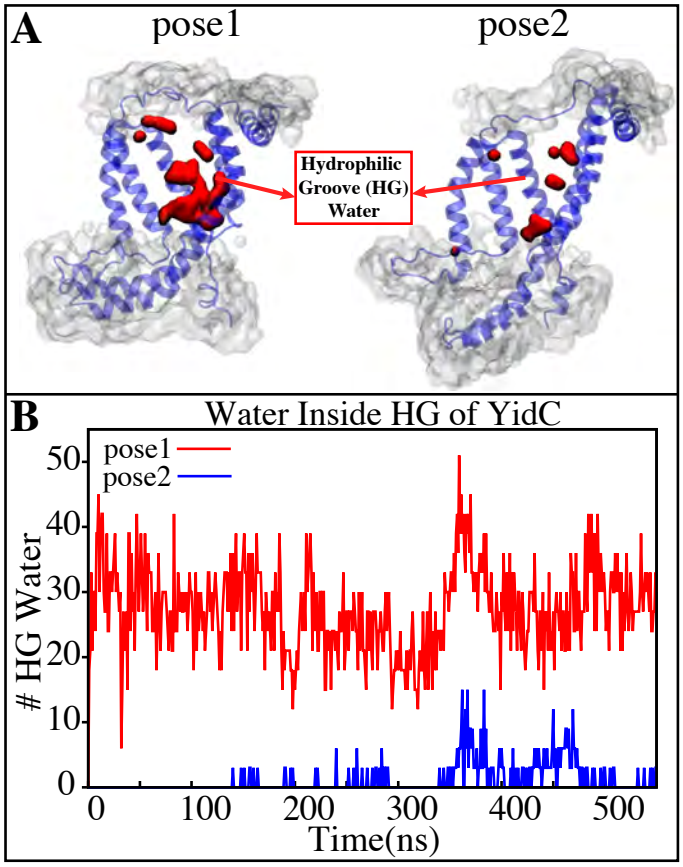


Figure 6: (A) Graphical representation of the docking models showing the average water molecule count in the hydrophilic groove (HG) region of YidC. (B) Number of water molecules inside the hydrophilic groove (HG) region of YidC in docking poses 1 (red) & pose 2 (blue).Pf3 coat protein entry into the TM area is aided by the water in the groove, which creates a water slide which aids in further insertion of Pf3 coat protein into the groove.

YidC has a U-shaped hydrophilic groove that is closed on the periplasmic side but exposed to the cytoplasmic side of the membrane bilayer. To examine the water content of the groove within helices TM1-TM5 (Fig. 6A), the number of water molecules inside the groove region of the YidC

protein was measured and plotted over the simulation time. The water analysis results reveal that the number of water molecules within the groove region is higher in pose1, which is considered the starting state of the insertion process. Whereas in the docking model pose2, the water content is close to zero throughout the simulation (Fig. 6B). This confirms the previous hypothesis that a water slide motion is important in the initial positioning of the Pf3 coat protein [43, 44]. The peptide enters the YidC groove via the cytoplasmic side of the membrane bilayer; the central TM helices are then widened to form a water slide [9, 78, 79] and the YidC groove region is filled with water to provide a smooth sliding motion for the entering protein. As Pf3 coat protein progresses through the insertion processes, the cytoplasmic groove of YidC becomes more compact and water molecules are pushed out of the TM groove. These two factors combine to cause a hydrophobic shift in the region, making it more susceptible to membrane insertion. Previous experimental studies have reported that the hydrophilic cavity of YidC reduces the energy barrier associated with the insertion of the substrate by shortening the hydrophobic core of the membrane [40, 80]. Based on our results, we hypothesize that the Pf3 coat protein is initially stabilized in the groove by hydrophilic interactions [40, 80], and dehydration of the groove, later in the process, will aid in breaking interactions with YidC (Fig. 4A) to facilitate the translocation of Pf3 coat protein from the groove into the membrane.

**The Saltbridge Interaction of Pf3 Coat Protein with YidC R72 in the Hydrophilic Groove is a Significant Event in the Insertion Process.**

The YidC residue Arginine 72 (R72) is in the core cavity of the YidC transmembrane region and forms a salt-bridge with incoming protein chains. It has been suggested that before translocation, a YidC protein's hydrophilic groove is forced into the hydrophobic cavity, implying that peptides may reach R72 for bond formation [81]. According to salt bridge analysis results, R72 is available for interactions with the incoming Pf3 coat protein. During the insertion process, the R72 residue of YidC forms a stable salt-bridge with D7 and D18 of Pf3 coat protein in the pose1 and pose2 simulations, respectively (Fig. 7). These two residues were experimentally shown to have an

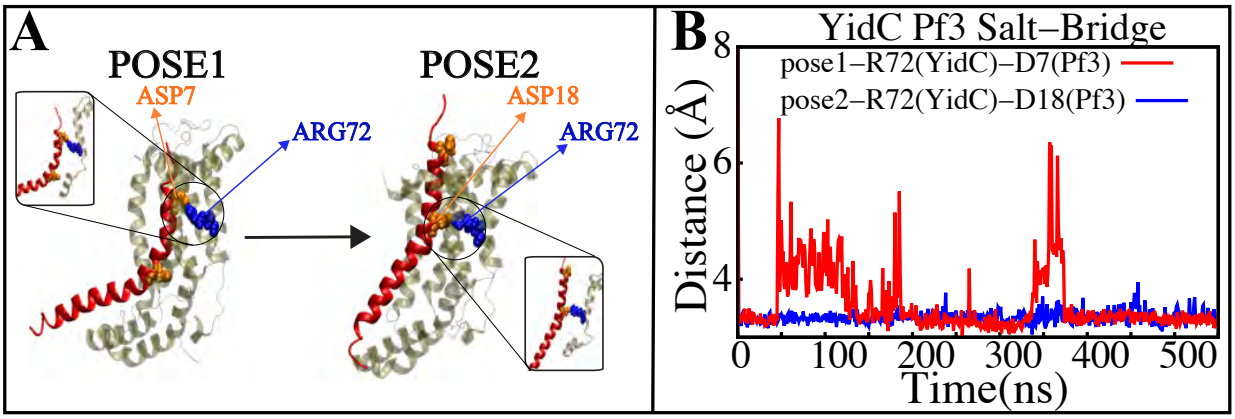


Figure 7: Salt−bridge connectivity of R72 (YidC) located in the groove. (A) Graphical representation of significant salt-bridge interactions between the Pf3 coat protein and YidC involved in the insertion process. (B) Distance between salt-bridge Arg72(YidC)−Asp18(Pf3 coat protein) and Arg72(YidC)−Asp7(Pf3 coat protein) (labeled with blue and red lines, respectively) in YidC and Pf3 coat protein docking models.

important function in the translocation of Pf3 coat protein into the membrane in an experimental research [79]. During the first phase of YidC insertion, the salt-bridge interaction between YidC's R72 and Pf3's D7 stabilizes the Pf3 coat proteins in the TM helical groove as soon as it enters the TM groove. As the Pf3 coat proteins move towards the periplasmic side of the protein, salt-bridge residue interactions with the Pf3 coat proteins are sequentially moved from D7 to D18 (Fig. 8A).

**Non-Equilibrium Simulation of YidC's Sec-independent Mechanism of Pf3 Coat Proteins Insertion in the Membrane Bilayer.**

The insertion process was further investigated using the above-mentioned non-equilibrium (NE) simulation (Fig. 8A) approach. Many of the key factors discussed above, such as Pf3 coat protein bending angle (Fig. 8B), radius of gyration (Fig. 8C), Pf3 coat protein lipid interactions (Fig. 8D), the presence of water in the groove (Fig. 9C), and Pf3 coat protein contacts with YidC (Fig. 9D), are evaluated for the NE simulation trajectory. Our NE simulation results are totally in agreement with results produced in equilibrium simulations. The bending of Pf3 coat protein is observed in the NE simulations, where Pf3 coat protein has gone from a lower to a greater bending angle (Fig. 8B) to adapt to the groove environment. The radius of gyration analysis also confirms our hy-

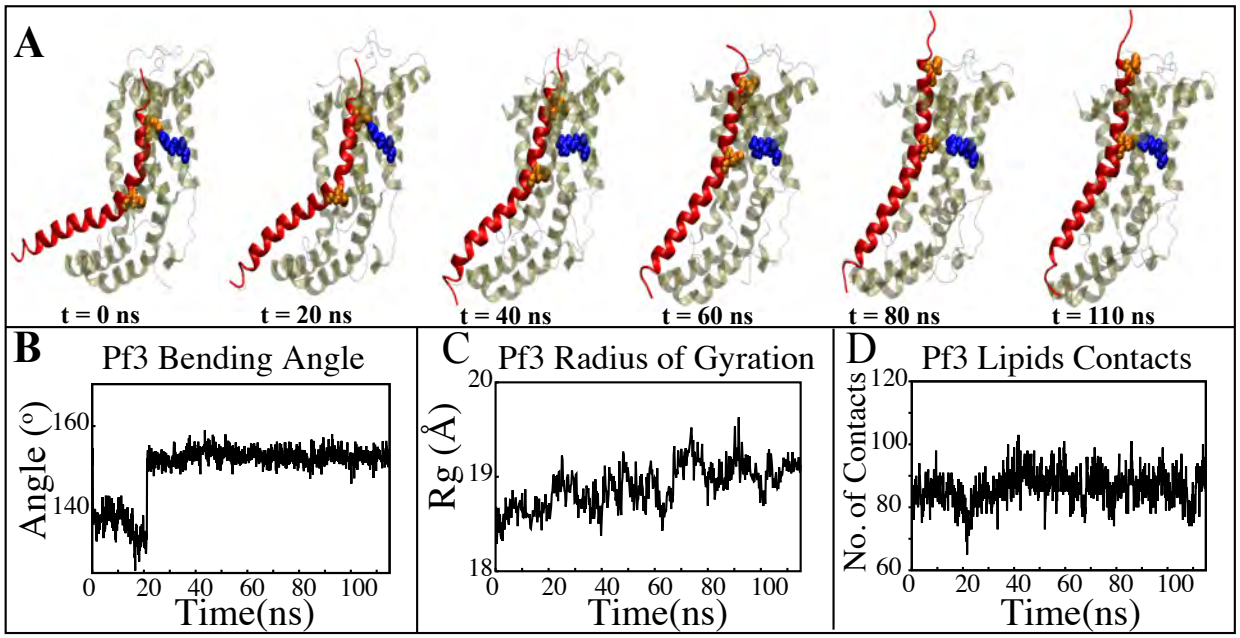


Figure 8: Characterizing the insertion process using targeted MD simulations. (A) Graphical representation of a series of targeted MD snapshots taken at different stages of the simulation. (B) Bending angle analysis of the Pf3 coat protein helix. (C) Radius of gyration of Pf3 coat protein peptide in the insertion process. (D) Interactions of Pf3 coat protein with lipid tails in the NE simulation process.

pothesis about Pf3 coat protein conformational changes during the insertion process (Fig. 8C). The increase and decrease in the amount of water inside the groove significantly supports the hydration and dehydration hypothesis (Fig. 9C). Robust interactions of Pf3 coat protein with lipid tails (Fig. 8D) play a significant role in the insertion process. As previously stated, YidC loses connections with the Pf3 coat protein as the insertion process progresses, as seen in the NE simulations, where the number of YidC-Pf3 contacts decreases during the targeted MD simulation (Fig. 9D). During the insertion of Pf3 coat protein inside the membrane, YidC undergoes significant conformational changes, which we observed previously in our analysis. As expected, Yidc underwent substantial conformational changes from the beginning to the completion of the insertion process as indicated by the overall RMSD (Fig. 9A) and radius of gyration (Fig. 9B) analyses. Overall, based on equilibrium and NE simulations, the following mechanism for YidC's Sec-independent insertion mechanism is proposed in this study: the incoming Pf3 coat protein first interacts with the cytoplasmic loops and gradually moves into the hydrophilic groove located in the transmembrane

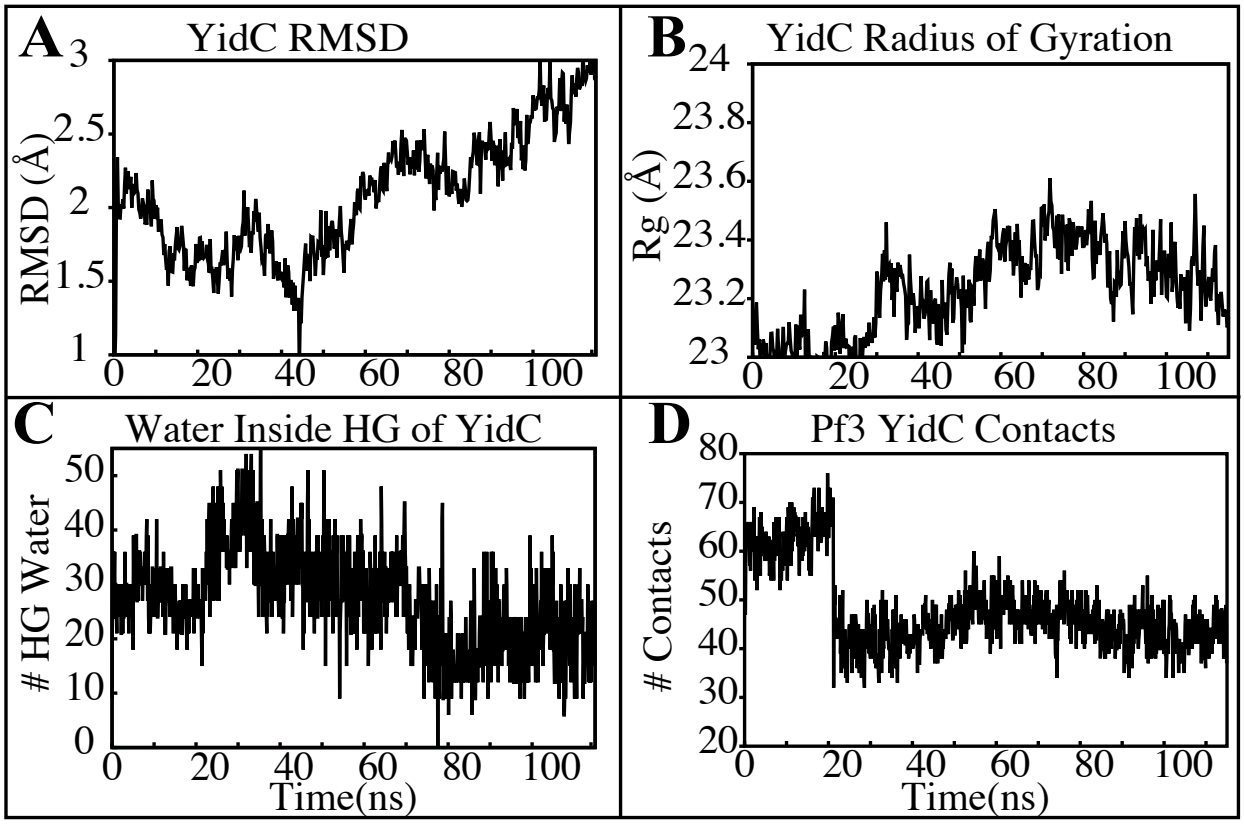


Figure 9: YidC conformational changes observed in the targeted MD simulations. (A & B) Root mean square deviation and radius of gyration of YidC in the insertion process of Pf3. (C) The water content in the hydrophilic groove of the protein during a 100ns NE simulation followed by a 15 ns equilibrium simulation. (D) Interaction of YidC with Pf3 coat protein in the NE simulation process.

region, forming a salt bridge with R72. The Pf3 coat protein's negatively charged D7 residue forms a salt bridge with the positively charged R72, which is critical to the insertion mechanism. The hydrophilic interactions within the groove (Fig. 4A) and salt-bridge interactions between the negatively charged D18 residue of Pf3 coat protein and positively charged R72 of YidC would drive the Pf3 coat protein to break the initial salt-bridge and move further into the groove. The N-terminal then moves into the deep groove and dehydration of the groove takes place. The Pf3 coat protein then migrates towards the periplasmic side of the membrane, assisted by the hydrophobic force, i.e., the hydrophobic interactions of the hydrophobic regions of the Pf3 coat protein with lipid tails out of the YidC hydrophilic groove.

It is important to note that in this study, we did not attempt to investigate the entire insertion

process from the initial stages of the binding to the full dissociation of the Pf3 coat protein. Instead, we only focused to look at a crucial stage of the process where the bound substrate moves up within the YidC-membrane environment. We particularly looked at the protein conformational dynamics within this part of the process. To simplify, particularly since we did not intend to investigate initial stages of binding that is likely to involve lipid headgroups, we employed a homogeneous POPE membrane instead of the anionic (phosphatidyl-glycerol (POPG) and cardiolipin (CL)) lipid-rich bacterial inner membrane. YidC's behavior in a pure POPE membrane and a heterogeneous POPE/POPG/CL membrane was compared in a recent MD study in its *apo* state [72]. The YidC conformation and protein/lipid interactions have been shown to be unaffected by the presence or absence of the anionic lipids POPG and CL in MD simulations [72]. However, it is highly likely for the anionic lipids to play a crucial role in the initial stages of binding and insertion process [75, 48, 76, 77, 9, 79]. Further in-depth computational and experimental studies are needed to have a better grasp on lipid specificity in the insertion process. More specifically, the proton motive force of the membrane, which is not the focus of our current study, may aid protein insertion in a lipid-specific manner. The proton motive force facilitates the YidC mediated membrane insertion by electrostatically attracting the negatively charged extracellular residues of the single-spanning membrane protein from the YidC groove, in addition to the hydrophobic interaction of the lipid tails with Pf3 coat protein [73, 48, 82, 77, 9, 79]. The functional group attached to the phosphate moiety determines the charge of the phospholipid. Compared to zwitterionic POPE, anionic POPG and CL have a greater likelihood of protons binding to their negatively charged headgroups [83]. It is likely that Pf3 coat protein in a lipid bilayer containing POPG and CL will experience stronger electrostatic attractions with enhanced proton binding compared to a pure POPE bilayer. However, the current study does not focus on the lipid-specific behavior of YidC-mediated membrane insertion, nor does it focus on the direction of the insertion, which is influenced by the proton motive force. Here we have only focused on a specific part of the insertion process, which is less dependent on proton motive force and lipid specificity and more dependent on YidC-Pf3 interactions coupled with conformational dynamics of YidC.

## Conclusions

Based on our equilibrium and non-equilibrium MD simulation results, YidC must undergo major conformational changes during the SecY-independent insertion process. The incoming Pf3 coat protein would first come into contact with the cytoplasmic loops and then penetrate into the hydrophilic groove, forming a salt bridge with R72. The YidC loops on the cytoplasmic side of the bilayer are critical for moving Pf3 coat protein into YidC's hydrophilic groove. At first, these cytoplasmic loops make contact with the Pf3 coat protein. The negatively charged D7 residue of Pf3 coat protein interacts with the positively charged R72 of YidC to form a stable salt bridge. The formation of this salt bridge is crucial in the insertion process to stabilize the Pf3 coat protein in the YidC's TM groove. The hydrophilic interactions within the groove also aid in the passage of the protein towards the periplasmic side, which is also supported by the salt bridge between D18 of Pf3 coat protein and R72 of YidC; this combination stabilizes the position of Pf3 coat protein inside the groove. Finally, when the Pf3 coat protein is completely inside the YidC's hydrophilic groove, it will come into contact with lipid tails. The Pf3 coat protein then travels towards the periplasmic side of the membrane, helped by the proton motive force and hydrophobic interaction with the membrane. The protein then moves into the membrane through the groove.

Despite the field's stunning advancements in recent years and the widespread use of docking techniques, there are still a few drawbacks. The fact that model quality and docking accuracy have a substantial impact on simulation results is one of these limitations. Additional studies using more docking models, including a range of substrate proteins in various conformational states, are required to fully understand the process. Results from this study would help in creating a plan both for experimental and computational scientists to study YidC SecY-independent mechanism for deeper understanding.

## References


[1] T. A. Rapoport, "Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes," *Nature*, vol. 450, no. 7170, pp. 663–669, 2007.

[2] A. Krogh, B. Larsson, G. Von Heijne, and E. L. Sonnhammer, "Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes," *Journal of Molecular Biology*, vol. 305, no. 3, pp. 567–580, 2001.

[3] R. E. Dalbey and A. Kuhn, "Membrane Insertases Are Present in All Three Domains of Life," *Structure*, vol. 23, no. 9, pp. 1559–1560, 2015.

[4] M. A. McDowell, M. Heimes, and I. Sinning, "Structural and molecular mechanisms for membrane protein biogenesis by the Oxa1 superfamily," *Nature Structural & Molecular Biology*, vol. 28, no. 3, pp. 234–239, 2021.

[5] F. Jiang, M. Chen, L. Yi, J. W. De Gier, A. Kuhn, and R. E. Dalbey, "Defining the regions of Escherichia coli YidC that contribute to activity," *Journal of Biological Chemistry*, vol. 278, no. 49, pp. 48965–48972, 2003.

[6] K. J. Nass, I. M. Ilie, M. J. Saller, A. J. M. Driessen, A. Caflisch, R. A. Kammerer, and X. Li, "The role of the N-terminal amphipathic helix in bacterial YidC: Insights from functional studies, the crystal structure and molecular dynamics simulations," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1864, no. 3, p. 183825, 2022.

[7] B. Güngör, T. Flohr, S. G. Garg, and J. M. Herrmann, "The ER membrane complex (EMC) can functionally replace the Oxa1 insertase in mitochondria," *PLOS Biology*, vol. 20, p. e3001380, mar 2022.

[8] A. N. Gray, J. M. Henderson-Frost, D. Boyd, S. Shirafi, H. Niki, M. B. Goldberg, and S. Maloy, "Unbalanced Charge Distribution as a Determinant for Dependence of a Subset of Escherichia coli Membrane Proteins on the Membrane Insertase YidC," *mBio*, vol. 2, no. 6, pp. e00238–11, 2011.

[9] R. E. Dalbey, A. Kuhn, L. Zhu, and D. Kiefer, "The membrane insertase YidC," *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1843, no. 8, pp. 1489–1496, 2014.

[10] R. E. Dalbey and M. Chen, "Sec-translocase mediated membrane protein biogenesis," *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1694, no. 1-3 SPEC.ISS., pp. 37–53, 2004.

[11] M. T. Borowska, P. K. Dominik, S. A. Anghel, A. A. Kossiakoff, and R. J. Keenan, "A YidC-like Protein in the Archaeal Plasma Membrane," *Structure*, vol. 23, no. 9, pp. 1715–1724, 2015.

[12] A. Kuhn and D. Kiefer, "Membrane protein insertase YidC in bacteria and archaea," *Molecular Microbiology*, vol. 103, no. 4, pp. 590–594, 2017.

[13] H. Chen, D. Ogden, S. Pant, W. Cai, E. Tajkhorshid, M. Moradi, B. Roux, and C. Chipot, "A Companion Guide to the String Method with Swarms of Trajectories: Characterization, Performance, and Pitfalls," *Journal of Chemical Theory and Computation*, vol. 18, pp. 1406–1422, 2022.

[14] P. A. Scotti, M. L. Urbanus, J. Brunner, J. W. L. De Gier, G. Von Heijne, C. Van Der Does, A. J. Driessen, B. Oudega, and J. Luirink, "YidC, the Escherichia coli homologue of mitochondrial Oxa1p, is a component of the Sec translocase," *EMBO Journal*, vol. 19, no. 4, pp. 542–549, 2000.

[15] S. J. Facey and A. Kuhn, "Membrane integration of E. coli model membrane proteins," *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1694, no. 1-3 SPEC.ISS., pp. 55–66, 2004.

[16] N. E. Lewis and L. J. Brady, "Breaking the bacterial protein targeting and translocation model: Oral organisms as a case in point," *Molecular Oral Microbiology*, vol. 30, no. 3, pp. 186–197, 2015.

[17] J. C. Samuelson, M. Chen, F. Jiang, I. Möller, M. Wiedmann, A. Kuhn, G. J. Phillips, and R. E. Dalbey, "YidC mediates membrane protein insertion in bacteria," *Nature*, vol. 406, no. 6796, pp. 637–641, 2000.

[18] D. Kiefer and A. Kuhn, "Yidc-mediated membrane insertion," *FEMS Microbiology Letters*, vol. 365, 2018.

[19] P. R. Laskowski, K. Pluhackova, M. Haase, B. M. Lang, G. Nagler, A. Kuhn, and D. J. Müller, "Monitoring the binding and insertion of a single transmembrane protein by an insertase," *Nature Communications*, vol. 12, no. 1, p. 7082, 2021.

[20] A. J. O. Lewis and R. S. Hegde, "A unified evolutionary origin for the ubiquitous protein transporters SecY and YidC," *BMC Biology*, vol. 19, no. 1, p. 266, 2021.

[21] L. Zhu, H. R. Kaback, and R. E. Dalbey, "YidC Protein, a Molecular Chaperone for LacY Protein Folding via the SecYEG Protein Machinery*," *Journal of Biological Chemistry*, vol. 288, no. 39, pp. 28180–28194, 2013.

[22] S. Nagamori, I. N. Smirnova, and H. R. Kaback, "Role of YidC in folding of polytopic membrane proteins," *Journal of Cell Biology*, vol. 165, no. 1, pp. 53–62, 2004.

[23] T. Serdiuk, D. Balasubramaniam, J. Sugihara, S. A. Mari, H. R. Kaback, and D. J. Müller, "YidC assists the stepwise and stochastic folding of membrane proteins," *Nature Chemical Biology*, vol. 12, no. 11, pp. 911–917, 2016.

[24] S. Kol, B. R. Turrell, J. De Keyzer, M. Van Der Laan, N. Nouwen, and A. J. Driessen, "YidC-mediated membrane insertion of assembly mutants of subunit c of the F1F0 ATPase," *Journal of Biological Chemistry*, vol. 281, no. 40, pp. 29762–29768, 2006.

[25] D. Spann, E. Pross, Y. Chen, R. E. Dalbey, and A. Kuhn, "Each protomer of a dimeric yidc functions as a single membrane insertase," *Scientific Reports*, vol. 8, no. 1, p. 589, 2018.

[26] M. Van der Laan, M. L. Urbanus, C. M. Ten Hagen-Jongman, N. Nouwen, B. Oudega, N. Harms, A. J. Driessen, and J. Luirink, "A conserved function of YidC in the biogenesis of respiratory chain complexes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 10, pp. 5801–5806, 2003.

[27] L. Yi and R. E. Dalbey, "Oxa1/Alb3/YidC system for insertion of membrane proteins in mitochondria, chloroplasts and bacteria," *Molecular Membrane Biology*, vol. 22, no. 1-2, pp. 101–111, 2005.

[28] E. Van Bloois, G. Jan Haan, J. W. De Gier, B. Oudega, and J. Luirink, "F1F0 ATP synthase subunit c is targeted by the SRP to YidC in the E. coli inner membrane," *FEBS Letters*, vol. 576, no. 1-2, pp. 97–100, 2004.

[29] Y. Endo, Y. Shimizu, H. Nishikawa, K. Sawasato, and K.-i. Nishiyama, "Interplay between MPIase, YidC, and PMF during Sec-independent insertion of membrane proteins," *Life Science Alliance*, vol. 5, no. 1, 2022.

[30] Y. Xin, Y. Zhao, J. Zheng, H. Zhou, X. C. Zhang, C. Tian, and Y. Huang, "Structure of YidC from Thermotoga maritima and its implications for YidC-mediated membrane protein insertion," *FASEB Journal*, vol. 32, no. 5, pp. 2411–2421, 2018.

[31] J. Yuan, G. J. Phillips, and R. E. Dalbey, "Isolation of cold-sensitive yidC mutants provides insights into the substrate profile of the YidC insertase and the importance of transmembrane 3 in YidC function," *Journal of Bacteriology*, vol. 189, no. 24, pp. 8961–8972, 2007.

[32] C. Klenner and A. Kuhn, "Dynamic disulfide scanning of the membrane-inserting Pf3 coat protein reveals multiple YidC substrate contacts," *Journal of Biological Chemistry*, vol. 287, no. 6, pp. 3769–3776, 2012.

[33] S. Kol, N. Nouwen, and A. J. Driessen, "Mechanisms of YidC-mediated insertion and assembly of multimeric membrane protein complexes," *Journal of Biological Chemistry*, vol. 283, no. 46, pp. 31269–31273, 2008.

[34] S. Ernst, A. K. Schönbauer, G. Bär, M. Börsch, and A. Kuhn, "YidC-driven membrane insertion of single fluorescent Pf3 coat proteins," *Journal of Molecular Biology*, vol. 412, no. 2, pp. 165–175, 2011.

[35] S. Funes, F. Kauff, E. O. Van Der Sluis, M. Ott, and J. M. Herrmann, "Evolution of YidC/Oxa1/Alb3 insertases: Three independent gene duplications followed by functional specialization in bacteria, mitochondria and chloroplasts," *Biological Chemistry*, vol. 392, no. 1-2, pp. 13–19, 2011.

[36] S. Funes, A. Hasona, H. Bauerschmitt, C. Grubbauer, F. Kauff, R. Collins, P. J. Crowley, S. R. Palmer, L. J. Brady, and J. M. Herrmann, "Independent gene duplications of the YidC/Oxa/Alb3 family enabled a specialized cotranslational function," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 16, pp. 6656–6661, 2009.

[37] R. Kohler, D. Boehringer, B. Greber, R. Bingel-Erlenmeyer, I. Collinson, C. Schaffitzel, and N. Ban, "YidC and Oxa1 Form Dimeric Insertion Pores on the Translating Ribosome," *Molecular Cell*, vol. 34, no. 3, pp. 344–353, 2009.

[38] A. Kedrov, M. Sustarsic, J. De Keyzer, J. J. Caumanns, Z. C. Wu, and A. J. Driessen, "Elucidating the native architecture of the YidC: Ribosome complex," *Journal of Molecular Biology*, vol. 425, no. 22, pp. 4112–4124, 2013.

[39] K. Kumazaki, T. Kishimoto, A. Furukawa, H. Mori, Y. Tanaka, N. Dohmae, R. Ishitani, T. Tsukazaki, and O. Nureki, "Crystal structure of Escherichia coli YidC, a membrane protein chaperone and insertase," *Scientific Reports*, vol. 4, p. 7299, 2014.

[40] Y. Chen, S. Capponi, L. Zhu, P. Gellenbeck, J. A. Freites, S. H. White, and R. E. Dalbey, "YidC Insertase of Escherichia coli: Water Accessibility and Membrane Shaping," *Structure*, vol. 25, no. 9, pp. 1403–1414.e3, 2017.

[41] S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers, and Y. Zhang, "Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification," vol. 466, pp. 1129–1133, 2010.

[42] A. Kedrov, S. Wickles, A. H. Crevenna, E. O. van der Sluis, R. Buschauer, O. Berninghausen, D. C. Lamb, and R. Beckmann, "Structural Dynamics of the YidC:Ribosome Complex during Membrane Protein Biogenesis," *Cell Reports*, vol. 17, no. 11, pp. 2943–2954, 2016.

[43] S. Wickles, A. Singharoy, J. Andreani, S. Seemayer, L. Bischoff, O. Berninghausen, J. Soeding, K. Schulten, E. O. van der Sluis, and R. Beckmann, "A structural model of the active ribosome-bound membrane protein insertase YidC," *eLife*, vol. 3, no. July2014, pp. 1–17, 2014.

[44] K. Ito, N. Shimokawa-Chiba, and S. Chiba, "Sec translocon has an insertase-like function in addition to polypeptide conduction through the channel," *F1000Research*, vol. 8, pp. F1000 Faculty Rev–2126, 12 2019.

[45] M. Fujihashi, K. Mito, E. F. Pai, and K. Miki, "Atomic resolution structure of the orotidine 5'-monophosphate decarboxylase product complex combined with surface plasmon resonance analysis: implications for the catalytic mechanism," *The Journal of biological chemistry*, vol. 288, pp. 9011–9016, 03 2013.

[46] . Molecular Operating Environment (MOE), "Chemical Computing Group Inc. Molecular Operating Environment (MOE); Chemical Computing Group Inc. 1010 Sherbooke St. West, Suite# 910: Montreal, QC, Canada,," *Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2013.*, 2015.

[47] T. Tsukazaki, "Structural Basis of the Sec Translocon and YidC Revealed Through X-ray Crystallography," *The Protein Journal*, vol. 38, no. 3, pp. 249–261, 2019.

[48] K. Kumazaki, S. Chiba, M. Takemoto, A. Furukawa, K. I. Nishiyama, Y. Sugano, T. Mori, N. Dohmae, K. Hirata, Y. Nakada-Nakura, A. D. Maturana, Y. Tanaka, H. Mori, Y. Sugita, F. Arisaka, K. Ito, R. Ishitani, T. Tsukazaki, and O. Nureki, "Structural basis of Sec-independent membrane protein insertion by YidC," *Nature*, vol. 509, no. 7501, pp. 516–519, 2014.

[49] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with NAMD," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1781–1802, 2005.

[50] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and A. D. MacKerell Jr, "CHARMM36m: an improved force field for folded and intrinsically disordered proteins," *Nature methods*, vol. 14, pp. 71–73, jan 2017.

[51] J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, and R. W. Pastor, "Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types," *Journal of Physical Chemistry B*, vol. 114, no. 23, pp. 7830–7843, 2010.

[52] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *The Journal of Chemical Physics*, vol. 79, no. 2, pp. 926–935, 1983.

[53] S. Jo, T. Kim, and W. Im, "Automated builder and database of protein/membrane complexes for molecular dynamics simulations," *PLoS ONE*, vol. 2, no. 9, 2007.

[54] J. K. Reid, "On the method of conjugate gradients for the solution of large sparse systems of linear equations," pp. 231–254, 1971.

[55] G. J. Martyna, D. J. Tobias, and M. L. Klein, "Constant pressure molecular dynamics algorithms," *J. Chem. Phys.*, vol. 101, no. 5, pp. 4177–4189, 1994.

[56] S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, "Constant pressure molecular dynamics simulation: The Langevin piston method," *J. Chem. Phys.*, vol. 103, no. 11, pp. 4613–4621, 1995.

[57] W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual molecular dynamics," *Journal of Molecular Graphics*, vol. 14, no. 1, pp. 33–38, 1996.

[58] K. Immadisetty, A. Polasa, R. Shelton, and M. Moradi, "Elucidating the Molecular Basis of Spontaneous Activation in an Engineered Mechanosensitive Channel," *Computational and Structural Biotechnology Journal*, vol. 20, pp. 2539–2550, 2022.

[59] K. Immadisetty, J. Hettige, and M. Moradi, "What Can and Cannot Be Learned from Molecular Dynamics Simulations of Bacterial Proton-Coupled Oligopeptide Transporter GkPOT?," *J. Phys. Chem. B*, vol. 121, pp. 3644–3656, 2017.

[60] A. Bakan, L. M. Meireles, and I. Bahar, "ProDy: Protein dynamics inferred from theory and experiments," *Bioinformatics*, vol. 27, no. 11, pp. 1575–1577, 2011.

[61] R. Guixà-González, I. Rodriguez-Espigares, J. M. Ramírez-Anguita, P. Carrió-Gaspar, H. Martinez-Seara, T. Giorgino, and J. Selent, "Membplugin: studying membrane complexity in vmd," *Bioinformatics*, vol. 30, no. 10, pp. 1478–1480, 2014.

[62] V. Govind Kumar, D. S. Ogden, U. H. Isu, A. Polasa, J. Losey, and M. Moradi, "Prefusion Spike Protein Conformational Changes Are Slower in SARS-CoV-2 than in SARS-CoV-1," *Journal of Biological Chemistry*, vol. 298, 2022.

[63] K. Immadisetty, A. Polasa, R. Shelton, and M. Moradi, "Elucidating the Molecular Basis of pH Activation of an Engineered Mechanosensitive Channel," *bioRxiv*, 2021.

[64] A. Polasa, I. Mosleh, J. Losey, A. Abbaspourrad, R. Beitle, and M. Moradi, "Developing a Rational Approach to Designing Recombinant Proteins for Peptide-Directed Nanoparticle Synthesis," *Nanoscale Adv.*, vol. 4, pp. 3161–3171, 2022.

[65] M. Moradi, G. Enkavi, and E. Tajkhorshid, "Atomic-level characterization of transport cycle thermodynamics in the glycerol-3-phosphate:phosphate transporter," *Nat. Commun.*, vol. 6, p. 8393, 2015.

[66] M. Moradi, C. Sagui, and C. Roland, "Calculating relative transition rates with driven nonequilibrium simulations," *Chem. Phys. Lett.*, vol. 518, pp. 109–113, 2011.

[67] S. C. Moradi, Mahmoud and C. Roland, "Investigating rare events with nonequilibrium work measurements: I. nonequilibrium transition path probabilities," *J. Chem. Phys.*, vol. 140, no. 3, p. 034114, 2014.

[68] M. Moradi and E. Tajkhorshid, "Mechanistic picture for conformational transition of a membrane transporter at atomic resolution," *Proc. Natl. Acad. Sci. USA*, vol. 110, no. 47, pp. 18916–18921, 2013.

[69] D. Ogden and M. Moradi, "Molecular Dynamics–Based Thermodynamic and Kinetic Characterization of Membrane Protein Conformational Transitions," *Structure and Function of Membrane Proteins*, pp. 289–309, 2021.

[70] G. Fiorin, M. L. Klein, and J. Hénin, "Using collective variables to drive molecular dynamics simulations," *Mol. Phys.*, vol. 111, no. 22–23, pp. 3345–3362, 2013.

[71] Y. Chen, R. Soman, S. K. Shanmugam, A. Kuhn, and R. E. Dalbey, "The role of the strictly conserved positively charged residue differs among the gram-positive, gram-negative, and chloroplast YidC homologs," *Journal of Biological Chemistry*, vol. 289, no. 51, pp. 35656–35667, 2014.

[72] T. Harkey, V. Govind Kumar, J. Hettige, S. H. Tabari, K. Immadisetty, and M. Moradi, "The role of a crystallographically unresolved cytoplasmic loop in stabilizing the bacterial membrane insertase yidc2," *Scientific Reports*, vol. 9, no. 1, p. 14451, 2019.

[73] M. Chen, J. C. Samuelson, F. Jiang, M. Muller, A. Kuhn, and R. E. Dalbey, "Direct interaction of YidC with the Sec-independent Pf3 coat protein during its membrane protein insertion," *Journal of Biological Chemistry*, vol. 277, no. 10, pp. 7670–7675, 2002.

[74] Z. Yu, G. Koningstein, A. Pop, and J. Luirink, "The conserved third transmembrane segment of YidC contacts nascent Escherichia coli inner membrane proteins," *Journal of Biological Chemistry*, vol. 283, no. 50, pp. 34635–34642, 2008.

[75] N. Shimokawa-Chiba, K. Kumazaki, T. Tsukazaki, O. Nureki, K. Ito, and S. Chiba, "Hydrophilic microenvironment required for the channel-independent insertase function of YidC protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 16, pp. 5063–5068, 2015.

[76] A. Gallusser and A. Kuhn, "Initial steps in protein membrane insertion. Bacteriophage M13 procoat protein binds to the membrane surface by electrostatic interaction," *The EMBO journal*, vol. 9, no. 9, pp. 2723–2729, 1990.

[77] D. Kiefer and A. B. T. I. R. o. C. Kuhn, "YidC as an Essential and Multifunctional Component in Membrane Protein Assembly," *Academic Press*, vol. 259, pp. 113–138, 2007.

[78] H. He, A. Kuhn, and R. E. Dalbey, "Tracking the Stepwise Movement of a Membrane-inserting Protein In Vivo," *Journal of Molecular Biology*, vol. 432, no. 2, pp. 484–496, 2020.

[79] A. Steudle, D. Spann, E. Pross, S. K. Shanmugam, R. E. Dalbey, and A. Kuhn, "Molecular communication of the membrane insertase yidc with translocase secyeg affects client proteins," *Scientific Reports*, vol. 11, no. 1, p. 3940, 2021.

[80] N. Shimokawa-Chiba, K. Kumazaki, T. Tsukazaki, O. Nureki, K. Ito, and S. Chiba, "Hydrophilic microenvironment required for the channel-independent insertase function of YidC protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 16, pp. 5063–5068, 2015.

[81] K. Kumazaki, T. Tsukazaki, T. Nishizawa, Y. Tanaka, H. E. Kato, Y. Nakada-Nakura, K. Hirata, Y. Mori, H. Suga, N. Dohmae, R. Ishitani, and O. Nureki, "Crystallization and preliminary X-ray diffraction analysis of YidC, a membrane-protein chaperone and insertase from Bacillus halodurans," *Acta Crystallographica Section F:Structural Biology Communications*, vol. 70, no. 8, pp. 1056–1060, 2014.

[82] J. C. Samuelson, F. Jiang, L. Yi, M. Chen, J.-W. de Gier, A. Kuhn, and R. E. Dalbey, "Function of YidC for the Insertion of M13 Procoat Protein inEscherichia coli: TRANSLOCATION OF MUTANTS THAT SHOW DIFFERENCES IN THEIR MEMBRANE POTENTIAL DEPENDENCE AND Sec REQUIREMENT*," *Journal of Biological Chemistry*, vol. 276, no. 37, pp. 34847–34852, 2001.

[83] M. Y. Yoshinaga, M. Y. Kellermann, D. L. Valentine, and R. C. Valentine, "Phospholipids and glycolipids mediate proton containment and circulation along the surface of energy-transducing membranes," *Progress in Lipid Research*, vol. 64, pp. 1–15, 2016.

## Chapter II: Deciphering the Inter-domain Decoupling in the Gram-negative Bacterial Membrane Insertase

Adithya Polasa[1], and Mahmoud Moradi[1] *

[1]Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701

### Abstract

YidC is a protein found in membranes that plays an important role in the process of inserting newly generated proteins into lipid membranes. The SecYEG-dependent complex is responsible for inserting proteins into the lipid bilayer, and this process is facilitated by YidC. In addition, YidC acts as a chaperone during the folding processes of proteins. Multiple investigations have conclusively shown that the gram-positive bacterium YidC has SecY-independent insertion mechanisms. Through the use of microsecond level all-atom molecular dynamics simulations, we have carried out the first in-depth investigation of the YidC protein originating from gram-negative bacteria. This research sheds light on the significance of several structural areas related to YidC at an atomic level by utilizing equilibrium molecular dynamics (MD) simulations. In this research, multiple models of YidC inside the lipid bilayer were constructed in order to achieve a deeper understanding of the critical role of the C2 loop and the extra periplasmic domain present in gram-negative YidC. According to the results of our research, the C2 loop is responsible for the overall stabilization of the protein, most notably in the transmembrane region, and it also has an allosteric influence on the periplasmic domain. We have found critical interactions that contribute to the stability of the protein as well as its functional aspect. Finally, our study provides a hypothetical SecY-independent insertion mechanism for gram-negative bacterial YidC.

### Introduction

Membrane proteins participate in fundamentally crucial biological processes such as signaling, transcriptional regulation, ion and macro-molecule transport, proteolysis, motility, metabolism, energy creation, and energy transfer. Specialized cellular machinery enables proper membrane

protein folding and insertion into the lipid bilayer. A group of membrane proteins may be inserted through the Sec apparatus in bacteria. YidC is one of the proteins working with Sec in order to introduce client proteins into the membrane [1, 2, 3, 4, 5, 6, 7].

YidC is a member of the Oxa/Alb3/YidC family of insertases found in mitochondria, chloroplasts, and bacteria [8, 9, 10, 11, 12, 13, 14]. YidC catalyzes the transmembrane insertion of freshly produced membrane proteins in the absence of an energy supply domain, such as an ATPase [15]. It also plays a vital role in insertion in the insertion and positioning of membrane proteins in bacteria [16, 17, 18]. Insertase proteins, such as YidC, have been exhaustively investigated to determine their importance for the insertion of proteins into membranes [19]. They are versatile proteins that, in conjunction with the SecYEG pathway [1, 2, 3, 4, 5, 6, 7], act to insert peptides into the membrane through the Signal Recognition Particle (SRP) mechanism. YidC may fold and insert polypeptides independently of the SecY pathway [20, 21, 22, 23, 2, 24, 25, 26, 27]. YidC is a crucial protein for the insertion of small phage coat proteins like Pf3 coat and M13 in a SecY-independent pathway [20, 28, 29, 30, 31, 32, 33].

A few experimental studies have explored the role of YidC in various microbial organisms. The genome of most gram-positive microscopic organisms encode two YidC proteins: YidC1 and YidC2 [34, 35]. Although YidC typically exists as a dimer or tetramer [36] under physiological conditions, it is discovered that YidC can also exist as a monomer in lipid bilayers [15, 37]. The gram-negative YidC protein possesses an additional transmembrane (TM) segment at the N-terminus and a large periplasmic domain (PD) [38, 39]. Although the PD region and the additional N-terminus TM segment are not required for YidC activity, PD region inteacts with SecY machinery and helps creating stable complex. [38]. The area with the C-terminal five TM segments are vital for the membrane insertase activity of gram-negative YidC [40]. In both gram-negative and gram-positive strains of bacteria, the protein is firmly anchored in the lipid bilayer by interfacial aromatic residues, a cytoplasmic salt-bridge group, and a periplasmic helix enhanced with aromatic residues. The highly conserved arginine residue (R366) in the hydrophilic groove was found in the same locations as in gram-positive (R72), implying that the arginine residue is as

important for gram-negative function as it is for gram-positive function [38]. A group of aromatic residues around R72/R366 may bind with incoming peptide during insertion into the lipid bilayer [41, 42, 40, 34, 35, 43, 19]. The C–terminus of monomeric YidC cooperates with the ribosomes, and the short interhelical loops come into contact with the ribosomal proteins [44].

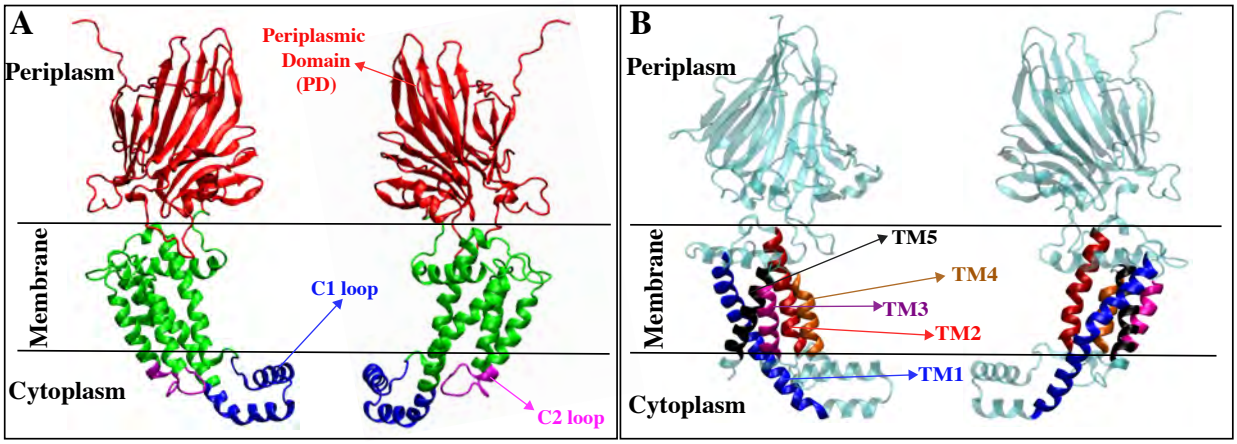


Figure 1: The cartoon representation of YidC (PDB:6AL2) (A) includes periplasmic domain (PD), transmembrane (TM) domain, C1, and C2 loops on the cytoplasmic side. (B) The cartoon representation of YidC's individual TM helices TM1 (blue), TM2 (Red), TM3 (purple), TM4 (orange), and TM5 (black).

Both in gram-negative and gram-positive bacteria, YidC is believed to promote membrane insertion simply by binding nascent chains and promoting their insertion into the lipid bilayer using cytoplasmic loop interactions, hydrophobic force, and interactions in the groove [15, 45, 46, 19]. The hydrophilic groove inside the membrane core of YidC increases the rate of accepting the hydrophilic moieties of a substrate into the membrane [47, 48]. During the process of its independent insertion mechanism, YidC also goes through a number of conformational changes, including widening of the TM region and hydration and dehydration of the hydrophilic groove. Additionally, a broad range of interactions with the incoming protein are engaged at each step of the insertion process, for example the salt bridge interaction with R72 aromatic residue [19].

In the past, gram-positive bacterial YidC insertase has been widely researched [15, 45, 46, 19]. The significance of the extra PD region in gram-negative YidC is not fully understood, yet. We don't know if the PD area of YidC affects protein stability and function, nor do we know if the cy-

toplasmic loops serve the same purpose in gram-positive and gram-negative bacterial YidC. Here, gram-negative YidC's structure was investigated using microsecond level molecular dynamics simulations. We examined the regional and global conformational changes in YidC brought on by the loss of the periplasmic domain and the cytoplasmic loop.

## Methods

From the Protein Data Bank, the crystal structure of the gram-negative bacterial YidC (PDB:6al2 [49]) was downloaded. The CHARMM36m [50] force field [51], together with the NAMD 2.14 [52] software package were used for all molecular dynamics simulations. The protein was solvated in TIP3P [53] water. Using the membrane builder on CHARMM-GUI [54], YidC was introduced into the lipid bilayer, solvated, and ionized. In these MD investigations, YidC was placed in a lipid bilayer of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPC) lipids. A 90 Å $\times$ 90 Å membrane layer surface was constructed along the XY plane. The protein-lipid assembly was solvated in TIP3 [53] water with 18 Å thick layers of water on top and bottom. To neutralize the system, 0.15 M of $Na^+$ and $Cl^-$ ions were added to the solution with a slight modification in the number of ions to neutralize the system. There were about $\approx$143000 atoms in the final solvated system. Utilizing the conjugate gradient technique [55], each system was energy minimized before the equilibrium simulation. The systems were then gradually relaxed using constrained MD simulations in accordance with the standard CHARMM-GUI [54] procedure. In the NPT ensemble at 310 K, 1 $\mu$s of equilibrium MD simulations were performed under periodic boundary conditions for each system. In the simulations, a Langevin integrator with a damping coefficient of $\gamma$ =0.5 $ps^{-1}$ and 1 atm pressure was maintained using the Nose-Hoover Langevin piston method [56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 18, 66].

We are interested in the significance of the cytoplasmic C2 loop and extracellular periplasmic domain (PD). It has been shown in the past that the C2 loop is crucial to YidC's overall conformation and function in gram-positive bacteria [67, 19]. We are interested in learning more about the functions of the PD (Fig. 1A), which is absent in gram-positive bacterial YidC and present

in gram-negative bacterial YidC. Accordingly, we created four systems: YidC with PD and C2 loop (YidC), YidC without C2 loop (YidC ΔC2), YidC without PD region (YidC ΔPD ), and YidC without PD region and C2 loop (YidC ΔPD ΔC2).

The preliminary simulations were run on TACC Stampede supercomputer. Subsequently, the production run for each model was increased to 400 ns with a timestep of 2.5 fs on Anton2 [58]. Every 240 picoseconds, conformations were gathered on Anton2. Initial processing of the Anton2 simulation trajectories was carried out on Kollman [58]. Later, the simulations were extended to 1 microsecond on TACC Stampede and one additional set of 1 $\mu$s simulation was performed for all systems.

All trajectories were visualized and examined using the VMD software [68]. A VMD plugin was used to analyze salt bridge interactions by measuring the distance between the oxygen atoms of acidic residues and the nitrogen atoms of basic residues with a cut-off distance of 4 Å. The interhelical angles were determined as the angle between the third main axes of the respective helices [66, 69, 60, 19]. The selection of the TM helices and other sub-domains are as indicated: TM1 (355-388); TM2 (423-442); TM3 (466-479); TM4 (497-508); TM5 (511-528); C1 area (380-420); C2 loop (480-492); and PD (49-326) (Fig. 1). To analyze the water inside the groove region, we counted the number of water molecules within 5 Å of R366. Principle component analysis (PCA) was performed for each trajectory using PRODY [70, 60, 19], taking only protein $C_\alpha$ atoms into account.

More quantitative data on the coordinated movements of the $C_\alpha$ atoms was obtained using Dynamic Network Analysis (DNA) of the associated motions of the protein [71, 58]. The correlation coefficient for the motion of each $C_\alpha$ atom with respect to the other $C_\alpha$ atoms was determined using MD-TASK [72], a software package of MD analysis tools. For each of the TM regions in all the simulated trajectories, a correlation matrix $M$ was created.

To quantify the differences in correlation between a system and some reference, a difference

matrix Δ was calculated,

$$\Delta = |M_i - M_{Ref.}| \tag{1}$$

where $M_i$ is the correlation matrix of interest, and $M_{Ref}$ is the correlation matrix of a reference conformation. In this work, the difference between a TM region in complete YidC structure conformation and other YidC structures was our point of interest. For this reason, the TM region in the wild type YidC simulations were compared with TM region in the other YidC simulation system described above.

## Results and Discussion

### The Overall Protein Conformation is Stabilized by the Presence of the C2 loop and the PD Domain.

The root mean square deviation (RMSD) of the protein was first calculated in order to assess the stability during equilibrium simulations. The RMSD of the YidC TM region was computed, and the results are shown as a function of the simulation run-time (Fig. 2). According to the $C_\alpha$ RMSD of the four different systems, the presence of the PD and C2 loop helps to stabilize the protein when it is in its native state (Fig. 2A). The system without the C2 loop is, beyond a shadow of a doubt, less stable than the system with the loop (Fig. 2B & D). The combined effect of removing the PD and the C2 loop has a higher impact on RMSD (Fig. 2D), which suggests that the PD could be responsible for preserving the stability of the protein. However, the effect of just removing the C2 loop (Fig. 2B) on protein RMSD is slightly higher than the native system (Fig. S1), but not as high as the system without PD and the C2 loop (Fig. 2D). We believe in between the C2 loop and the PD region, the C2 loop is more important for the stability of the protein compared to the PD region, as you can see in system YidC ΔPD (Fig. 2C) the RMSD of the TM region is very similar to the wild type YidC structure (Fig. 2A). Therefore, we have arrived at the conclusion that the impact of eliminating the C2 loop on gram-negative bacterial YidC is far more substantial than the

effect of removing just PD regions (Fig. S1).

The principal component analysis (PCA) was used in order to determine the most important distinctions between the systems could be determined. The projections onto principle components (PCs) 1 and 2 allowed for a straightforward differentiation between the YidC wild type and the other systems (Fig. 3). In this particular analysis, only the YidC $C_\alpha$ atoms that are located in the TM area are taken into account. PC1 contributed 27.5% of the total variance, while PC2 contributed 19.1% of the total variance. It was anticipated that the structural analysis of YidC ΔC2 and YidC ΔC2 ΔPD models would contradict that of YidC and YidC ΔPD in PC1 and PC2 (Fig. 3). This contradiction is rational, considering the large conformational discrepancies that were found earlier in the RMSD analysis (Fig. S1). In general, the most important thing that came out of the principal components analysis was the realization that the behavior of the YidC ΔC2 ΔPD (Fig. 3 D & H) protein system was quite different from the other systems (Fig. 3). The results of this study lend credence to the theory that the C2 loop play a significant part in the conformational dynamics of YidC.

Previous studies have revealed that the YidC transmembrane (TM) region is crucial for membrane protein insertion mechanism [73, 74, 19]. In order to examine the impact of deleting the PD and C2 loop on the TM helices, the helical angle between each pair of TMs was measured in this work (Fig. 4). When compared to the wild type, we found that the local shape of the TM helices was altered in all of the other systems. Similar to the analysis above, the local conformation was more affected by the removal of both the PD and the C2 loop (Fig. 4D & H). However, when just the PD is removed, we do see some changes in the helical angle (Fig. 4C & G), despite the fact that the impact is not as large as what is seen in systems that do not have a C2 loop (Fig. 4B, F, D, & H). This demonstrates the critical role that the C2 loop plays in maintaining the structural stability of transmembrane region of YidC. We do observe a similar trend in other transmembrane helices combinations (Fig. S2).

Furthermore, we used dynamic network analysis (DNA), which finds the linear connection between various residue pairs, in order to conduct a comprehensive study on the allosteric interactions

of the C2 loop and the PD region with the various protein domains. The correlation coefficient of each residue pair is shown in Figure 5. This coefficient was calculated from the trajectory with the C2 loop and PD region, then subtracted from the same quantity that was calculated from the trajectory without the loop and PD region systems, and the result was reported as its absolute value. The amount that is presented for each pair of residues measures the size of the difference in the correlation behavior of the two residues that is brought on by the presence of the C2 loop and the PD domain in the TM region. The presence or absence of the C2 loop has been demonstrated to create significant variations in the correlations between YidC's distinct domains. The inter-domain correlations, notably between the TM/C1 loop region, vary greatly between the YidC and other systems (Fig. S3). This difference in the cross-correlation is more pronounced, especially between TM1 and TM4 (Fig. 5). Based on this, we believe that the C2 loop does play a crucial role in the YidC conformational dynamics and that the conformational dynamics of the TM region are affected in its absence. The DNA findings (Fig. S3) are consistent with the early evidence for global and regional structural alterations. Overall, the results show that the C2 loop affects the behavior of the functionally essential areas of gram-negative YidC.

**C2 loop and the Periplasmic Domain Allosterically Influence YidC's Other Functionally Important Regions**

YidC's U-shaped hydrophilic groove, which is exposed on the cytoplasmic side of the membrane bilayer, is essential for the insertion process [47, 48]. The membrane proteins enter the YidC groove through the cytoplasmic side of the membrane bilayer during the insertion process. The YidC groove area is filled with water to provide a smooth sliding motion for the entering protein. As a membrane protein advances through the insertion processes, the groove water molecules are expelled from the TM groove. These two variables produce a change in the region's hydrophobicity, making it more vulnerable to membrane insertion [41, 75, 19].

To analyze the water content of the groove inside the TM region, the amount of water molecules within the groove region of the YidC protein was determined and plotted against simulation time.

The lack of the C2 loop had a significant impact on the amount of water that was present inside the groove area (Fig. 6B & D). Because hydrophilic contacts initially serve to keep membrane proteins in place inside the groove whilst the insertion process is being carried out [19]. This provides us with strong evidence that the C2 loop not only contributes to the conformational dynamics of the protein but also plays a significant role in the function of the protein. The water quantity is significantly reduced in system YidC ΔC2 ΔPD loop compared to system YidC. Our findings lead us to infer that the removal of just PD does have a marginal impact on the conformational dynamics of YidC. However, when this modification is coupled with the removal of the C2 loop, the effect is significantly amplified, which could ultimately effect the insertion process.

We also found an intradomain hydrogen bond in the TM region between Y516 and G429 that is only stable in the wild type system compared to other systems (Fig. 7). Especially in the YidC ΔC2 ΔPD system, this bond is completely broken. We think that this hydrogen bond is unstable in the system YidC ΔC2 ΔPD because fluctuation of the TM region caused due to absence of C2 loop (Fig 7B & D). These results clearly show that there is a link between the functionally important C2 loop and the PD region on the TM side of YidC. We also think this hydrogen bond (Fig 7) could be important to the insertion process. As the incoming membrane protein moves along the groove and toward the periplasmic side, it breaks this link (Fig 7), which causes widening of the TM region and leads the water to leave the hydrophilic groove, resulting in a hydrophobic shift and increasing the likelihood of membrane insertion. To this point, we have shown that the C2 loop and PD directly alter the structural behavior of YidC, while the PD region also slightly affects the conformational dynamics of the protein. However, the absence of both the PD and the C2 loop has an allosteric influence on the behavior of the YidC conformational dynamics, although the effect of the C2 loop absence is substantially larger than the PD absence. This leads us to the conclusion that the C2 loop is much more important to the structure and function of YidC than the PD region, which is in support of previous research [38, 39]. On the other hand, the removal of just PD did affect the conformational dynamics of the TM region. To determine what caused this effect, we analyzed interactions between the PD and the TM region that play a significant role in the stability

of the protein.

**Inter-domain Amino Acids Interactions Play a Key Role in the Stabilization of the Transmembrane Domain**

Table 1: Occupancy (%) of Inter-domain H-Bonds between PD and TM region

| | **Y259-M441 (%)** | | **A214-Y437 (%)** | | **T327-E312 (%)** | |
|---|---|---|---|---|---|---|
| System | **set1** | **set2** | **set1** | **set2** | **set1** | **set2** |
| **YidC** | 40 | 40 | 66 | 42 | 81 | 43 |
| **YidC ΔC2** | 0 | 0 | 0 | 0 | 0 | 0 |
| **YidC ΔPD** | NA | NA | NA | NA | NA | NA |
| **YidC ΔC2 ΔPD** | NA | NA | NA | NA | NA | NA |

Previous experimental findings lead researchers to the hypothesis that the interactions between the PD region of YidC and the TM region of YidC are crucial for maintaining the protein's stable state in the membrane [38]. Through this research, we were able to identify critical hydrogen bonds and salt-bridle interactions that take place between the PD region and the TM region. It is interesting to note that the hydrogen bonds that had been formed between the PD region and the TM region were only observed in wild type YidC system and entirely disrupted in the YidC ΔC2 system. However, there is no direct interaction between the PD area and the C2 loop region which does have an effect on stability of TM region. We also identified two salt bridge interaction between PD and TM which are contributing to stability of TM region. The salt bridge between D315 and K345 is stably formed in both systems of YidC even without C2 loop this salt bridge is unaffected (Fig. 8A & B). However the salt bridge between K232 and D329 (Fig. 9) is disrupted in YidC ΔC2 (Fig. 9B).

We believe that the removal of the C2 loop has caused an effect on the hydrogen bond (Fig. 7) in the TM core region, causing instability of the TM region and ultimately causing disruption of hydrogen bonds and salt bridge interactions between the PD and the TM region. Based on the analysis presented above, we postulate that the the hydrogen bond (Fig. 7) in the protein's groove region is essential for preserving the structural stability of the protein. The C2 loop is crucial for this hydrogen bond to remain stable in the structure. Although the YidC PD region does have some

influence on the protein's structure, its contribution to the protein's overall function is noticeably less substantial than that of the C2 loop. Even without the PD region, it is still possible to have a normal Sec-independent insertase mechanism; however, the absence of the C2 loop may have a detrimental influence on the protein function.

**Proposed Independent Insertion Mechanism of Gram-negative Bacterial YidC**

According to the findings and earlier hypotheses, gram-negative bacterial YidC must likewise go through significant conformational changes during the SecY-independent insertion procedure, much like gram-positive bacterial YidC. During the SecY-independent insertion process, the entering membrane protein would first make contact with the cytoplasmic loops and then squeak into the hydrophilic groove, where it would join forces with R366 to create a salt bridge. Incoming protein must be moved into YidC's hydrophilic groove by the YidC loops on the cytoplasmic side of the bilayer. The salt bridge between the incoming protein and R366 of YidC also contributes to the passage of the protein towards the periplasmic side, stabilizing its position within the groove. As the incoming protein moves through the groove and approaches the periplasmic side, it breaks the hydrogen bond between Y516 and G429, causing widening of the TM region that results in a hydrophobic shift through dehydration of the groove. The protein will finally make contact with lipid tails. After that, with the aid of the proton motive force and the membrane's hydrophobic interaction, the protein moves in the direction of the membrane. Then, via the groove, the protein enters the membrane.

**Conclusion**

We have come to the conclusion that the C2 cytoplasmic loop of YidC could be an important part of the protein. This could be achieved, for example, by the C2 cytoplasmic loop of YidC helping to stabilize the protein structure via its indirect effects on interactions in the transmembrane core region, as well as its indirect effects on other periplasmic domains, in particular on interactions between the PD and the TM region. It also indicates that the existence of the C2 loop has an

influence on the functional features of YidC, such as the hydration of the groove. In order to further understand the C2 loop's role in the sec-independent insertion process of small single-spanning membrane proteins like the pf3 coat protein, whose interactions with the cytoplasmic region are thought to be essential, more study is required. In the context of molecular dynamics simulations, the results of our research show that C2 loops may be required for the structural stability and SecY-independent function of the gram-negative YidC membrane protein dynamics.

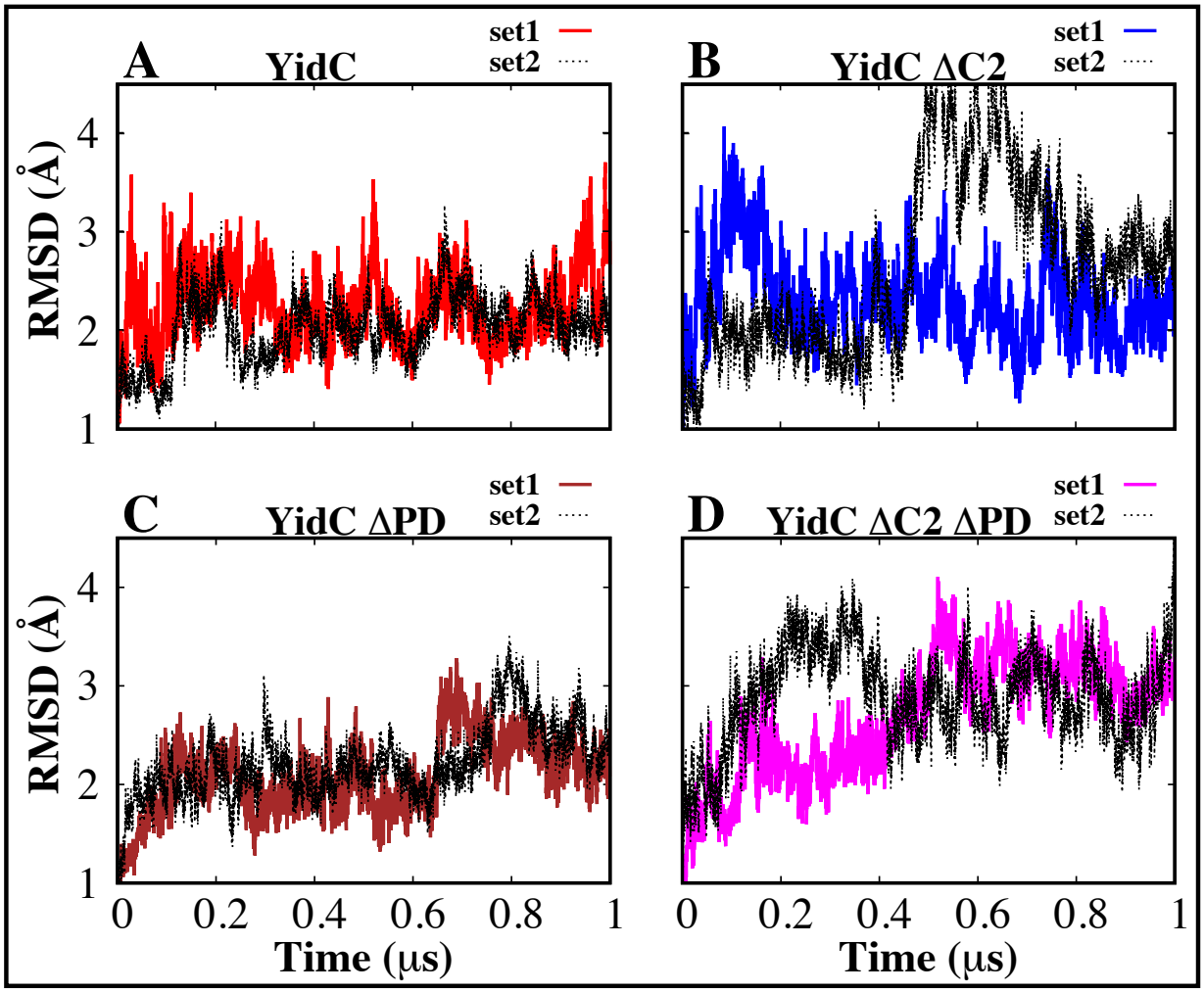


Figure 2: Analysis of YidC's structural stability in presence and absence of PD and C2 loop. (A–D) The root mean square deviation of the YidC in different systems. Based on RMSD data, we have observed that YidC fluctuates more in the system with the C2 loop removed compare to systems with C2 loop. The simulations for each system were run twice, and the dashed lines in the graphs reflect the second run of those simulations.

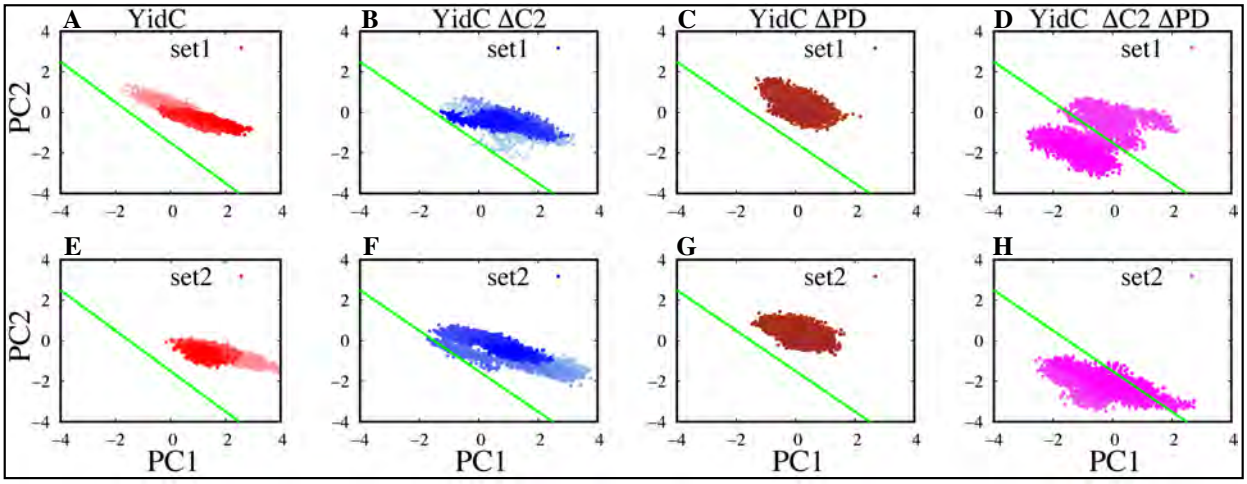


Figure 3: Projections of the principal components (PC's) 1 and 2. (A–H) PCA findings of PC1 vs PC2 for YidC systems from set 1 simulations are shown in the top row, while the results of set 2 simulations are displayed in the bottom row. The gradation of colors in the image denotes a timeline, with lighter shades reflecting earlier points in the simulation and darker hues denoting later points. Only the PCA analysis of the TM area that is present in all systems is shown here for the sake of consistency. For easier comparison, the green line in the plot is to show the difference in projections of PC's.

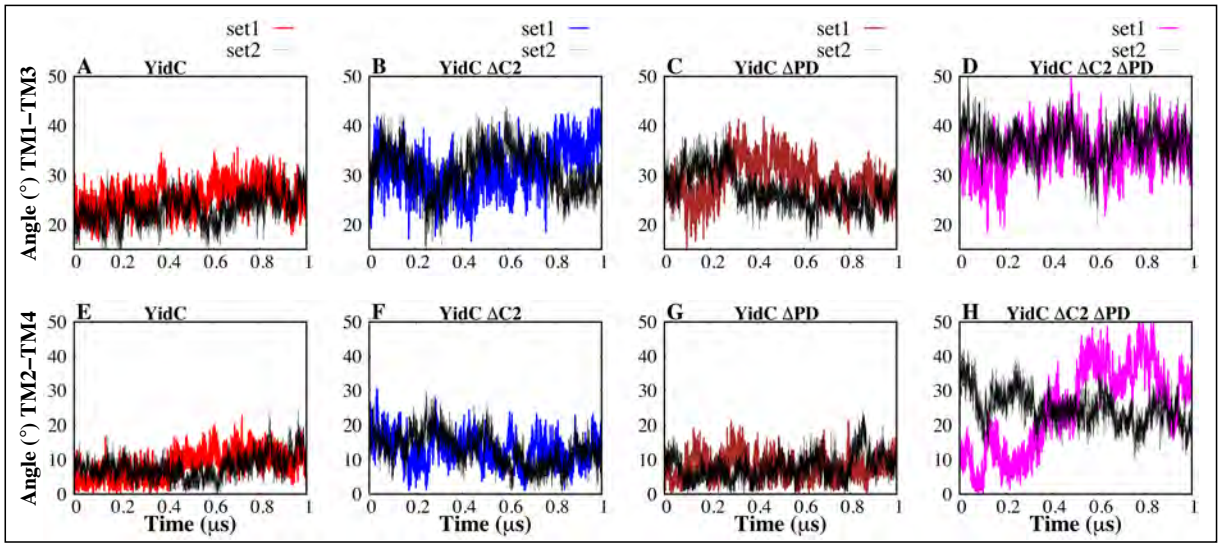


Figure 4: Inter-helical angles between transmembrane helices of YidC. (A–D) The inter-helical angle between the transmembrane helix 1 and helix 3 of the protein. (E–H) The overall inter-helical angle between helix 2 region and helix 4 of the protein.

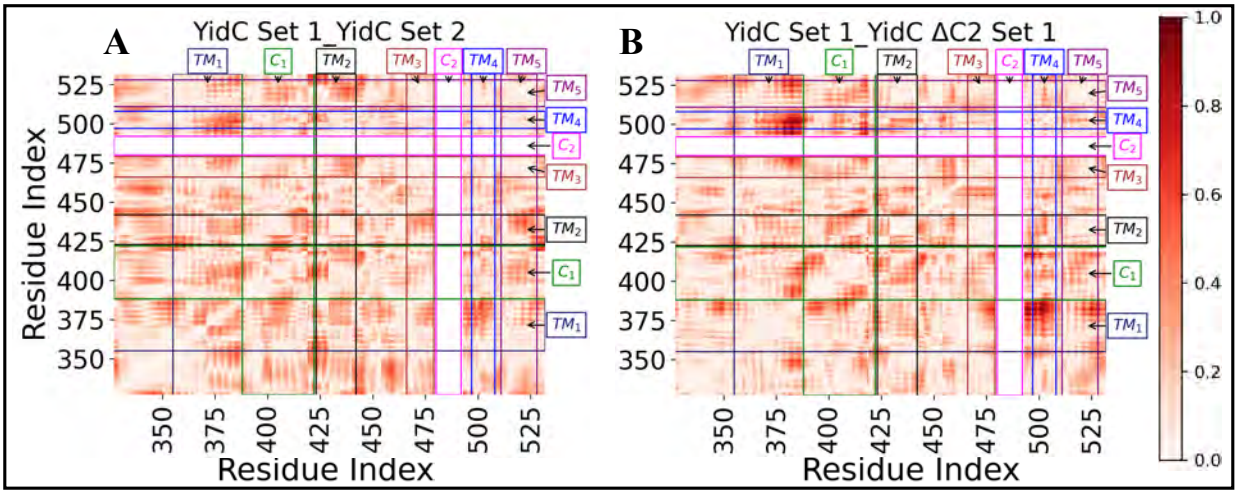


Figure 5: The DNA analysis revealed differences in correlation between the YidC set 1 control system and YidC ΔC2 ΔPD system examined. The theoretical maximum for correlation difference is 2, but the observed maximum was less than 1. (A–B) Differences in correlation are shown as a red gradient, with darker red indicating a larger difference.

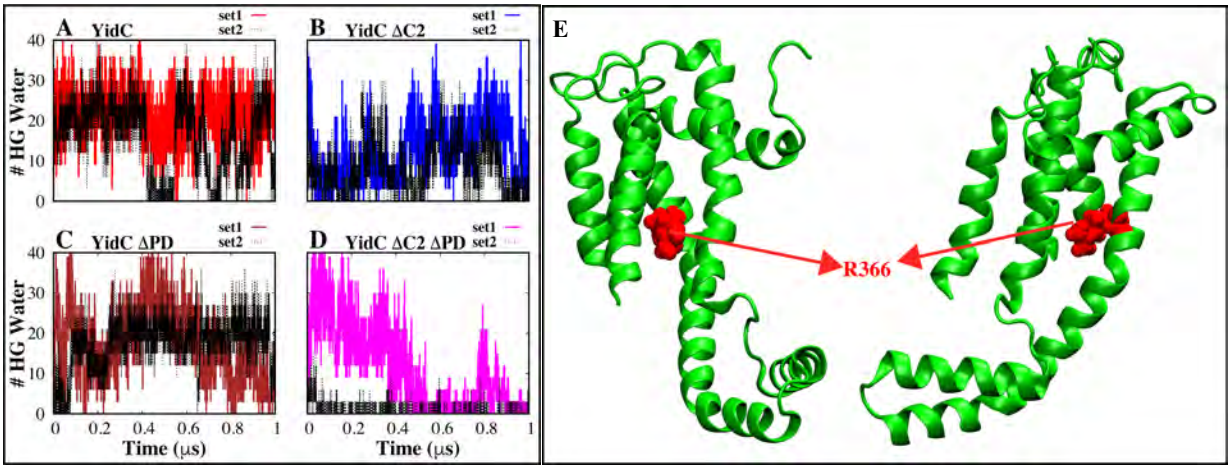


Figure 6: Analyses of the water inside the YidC groove. (A–D) The number of water molecules that are located within 5 Å of the R366 residue inside the hydrophilic groove (HG) region of YidC in each and every system. (E) The graphical illustration of the residue R366, which may be found in the central part of the hydrophilic groove (HG) of YidC

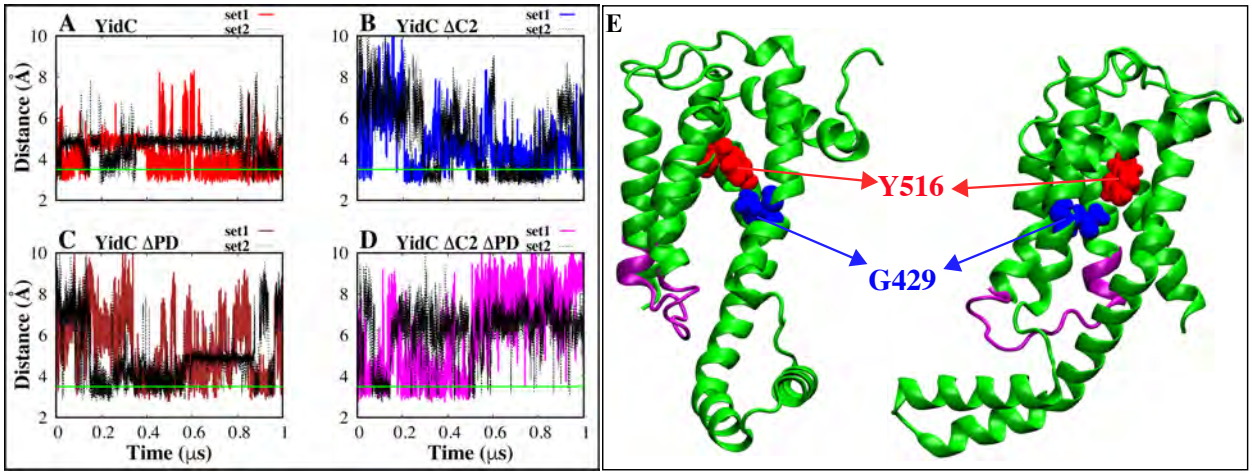


Figure 7: Interaction analysis of a hydrogen bond between Y516 and G429 (YidC), which is situated inside the groove area. (A–D) Analysis of the distance between the hydrogen bonds formed by Y516 and G429. (E) The graphical representation of the residues that participate in the interaction involving hydrogen bonds

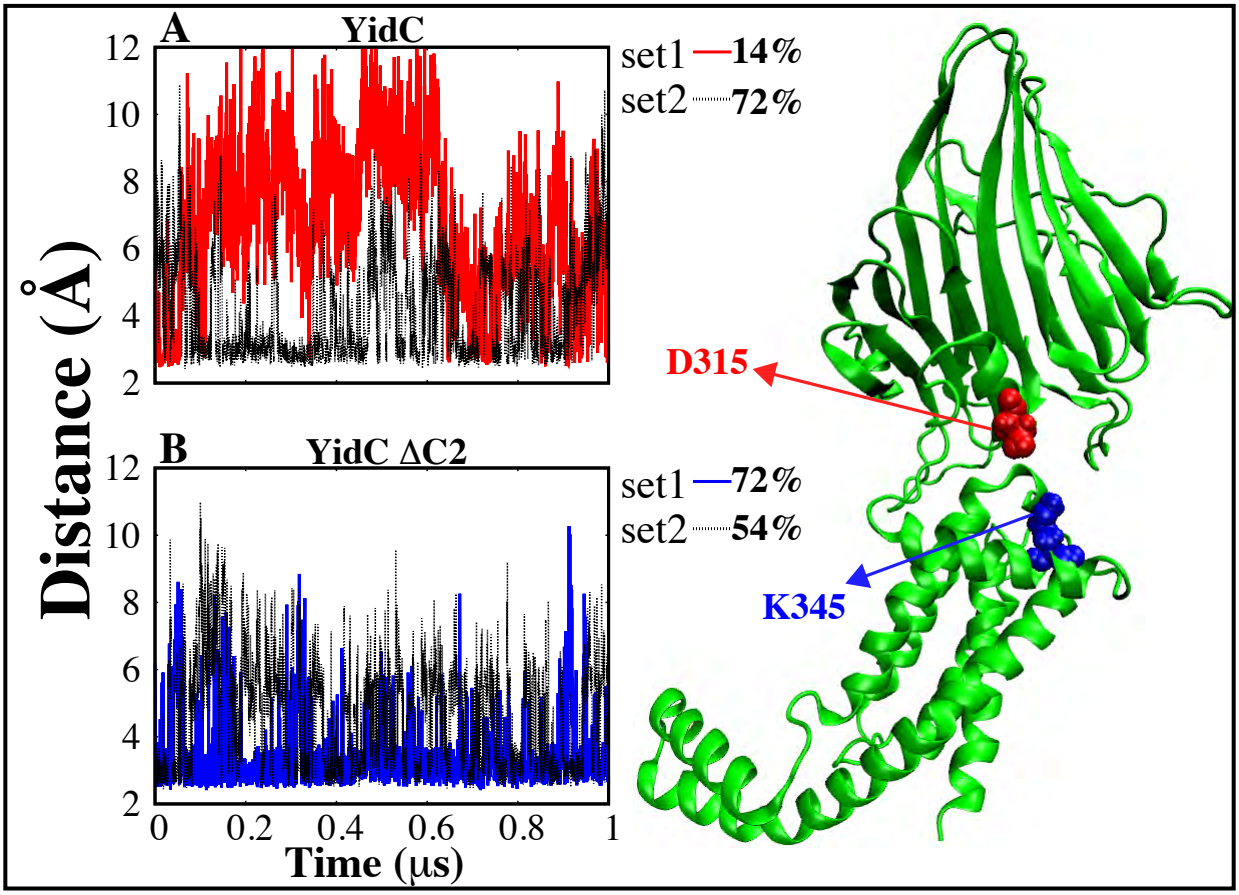


Figure 8: Salt–bridge interaction between D315 and K345 (YidC), which is located between the PD and TM regions. Graphical illustration of the important salt-bridge interactions that take place between the PD region and the TM area (right). (A–B) Distance analysis between D315–K345 salt-bridge and we also reported the occupancy of salt bridge interaction.

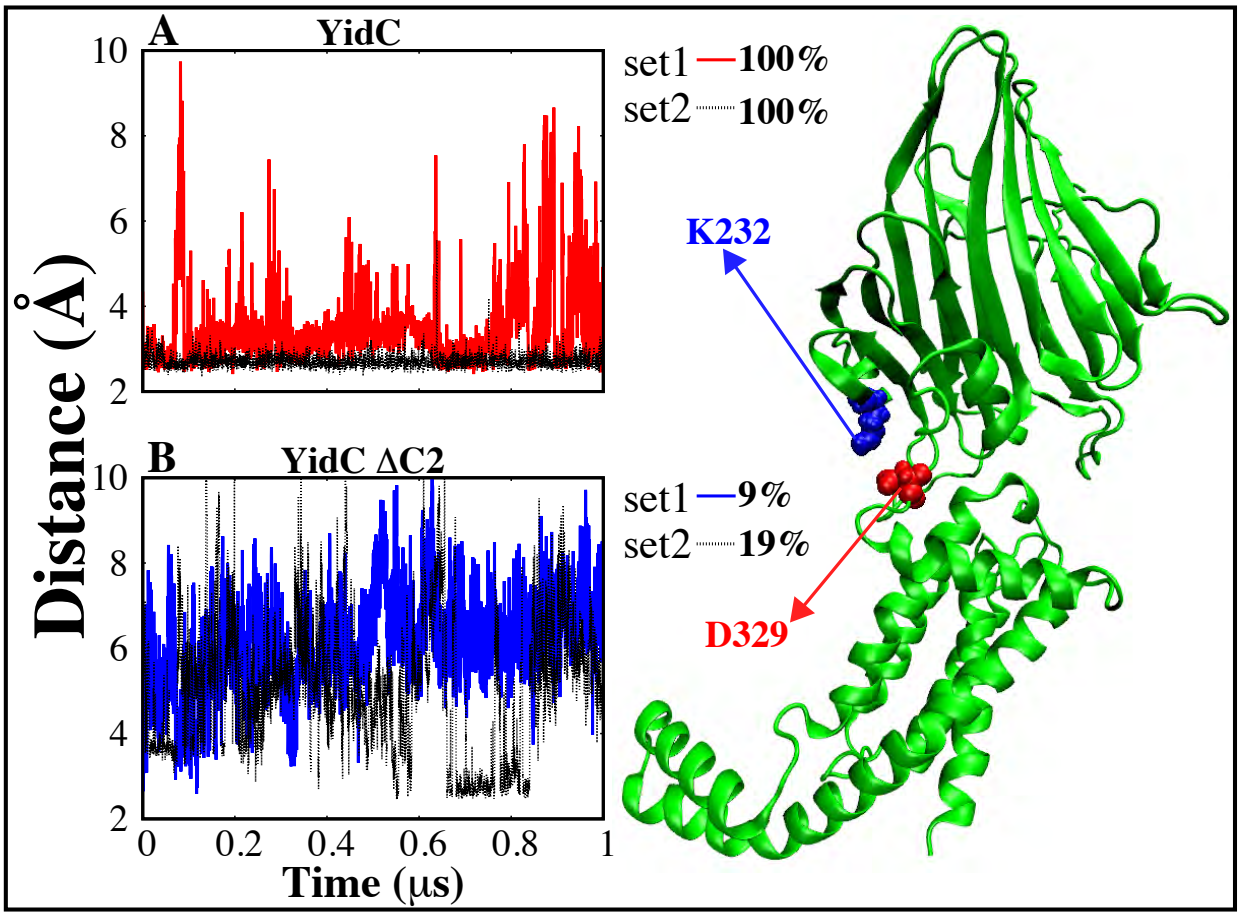


Figure 9: Salt–bridge formed between D329 and K232 (YidC), which is located between the PD and TM regions. The cartoon representation of the salt-bridge interactions that take place between the PD region and the TM area (right). (A–B) Distance analysis between D329–K232 salt-bridge and we also reported the occupancy of salt-bridge interaction.

## References


[1] S. Nagamori, I. N. Smirnova, and H. R. Kaback, "Role of YidC in folding of polytopic membrane proteins," *Journal of Cell Biology*, vol. 165, no. 1, pp. 53–62, 2004.

[2] N. E. Lewis and L. J. Brady, "Breaking the bacterial protein targeting and translocation model: Oral organisms as a case in point," *Molecular Oral Microbiology*, vol. 30, no. 3, pp. 186–197, 2015.

[3] T. Serdiuk, D. Balasubramaniam, J. Sugihara, S. A. Mari, H. R. Kaback, and D. J. Müller, "YidC assists the stepwise and stochastic folding of membrane proteins," *Nature Chemical Biology*, vol. 12, no. 11, pp. 911–917, 2016.

[4] S. Kol, B. R. Turrell, J. De Keyzer, M. Van Der Laan, N. Nouwen, and A. J. Driessen, "YidC-mediated membrane insertion of assembly mutants of subunit c of the F1F0 ATPase," *Journal of Biological Chemistry*, vol. 281, no. 40, pp. 29762–29768, 2006.

[5] M. Van der Laan, M. L. Urbanus, C. M. Ten Hagen-Jongman, N. Nouwen, B. Oudega, N. Harms, A. J. Driessen, and J. Luirink, "A conserved function of YidC in the biogenesis of respiratory chain complexes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 10, pp. 5801–5806, 2003.

[6] L. Yi and R. E. Dalbey, "Oxa1/Alb3/YidC system for insertion of membrane proteins in mitochondria, chloroplasts and bacteria," *Molecular Membrane Biology*, vol. 22, no. 1-2, pp. 101–111, 2005.

[7] E. Van Bloois, G. Jan Haan, J. W. De Gier, B. Oudega, and J. Luirink, "F1F0 ATP synthase subunit c is targeted by the SRP to YidC in the E. coli inner membrane," *FEBS Letters*, vol. 576, no. 1-2, pp. 97–100, 2004.

[8] R. E. Dalbey and A. Kuhn, "Membrane Insertases Are Present in All Three Domains of Life," *Structure*, vol. 23, no. 9, pp. 1559–1560, 2015.

[9] M. A. McDowell, M. Heimes, and I. Sinning, "Structural and molecular mechanisms for membrane protein biogenesis by the Oxa1 superfamily," *Nature Structural & Molecular Biology*, vol. 28, no. 3, pp. 234–239, 2021.

[10] F. Jiang, M. Chen, L. Yi, J. W. De Gier, A. Kuhn, and R. E. Dalbey, "Defining the regions of Escherichia coli YidC that contribute to activity," *Journal of Biological Chemistry*, vol. 278, no. 49, pp. 48965–48972, 2003.

[11] K. J. Nass, I. M. Ilie, M. J. Saller, A. J. M. Driessen, A. Caflisch, R. A. Kammerer, and X. Li, "The role of the N-terminal amphipathic helix in bacterial YidC: Insights from functional studies, the crystal structure and molecular dynamics simulations," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1864, no. 3, p. 183825, 2022.

[12] B. Güngör, T. Flohr, S. G. Garg, and J. M. Herrmann, "The ER membrane complex (EMC) can functionally replace the Oxa1 insertase in mitochondria," *PLOS Biology*, vol. 20, p. e3001380, mar 2022.

[13] T. A. Rapoport, "Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes," *Nature*, vol. 450, no. 7170, pp. 663–669, 2007.

[14] A. Krogh, B. Larsson, G. Von Heijne, and E. L. Sonnhammer, "Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes," *Journal of Molecular Biology*, vol. 305, no. 3, pp. 567–580, 2001.

[15] R. E. Dalbey, A. Kuhn, L. Zhu, and D. Kiefer, "The membrane insertase YidC," *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1843, no. 8, pp. 1489–1496, 2014.

[16] M. T. Borowska, P. K. Dominik, S. A. Anghel, A. A. Kossiakoff, and R. J. Keenan, "A YidC-like Protein in the Archaeal Plasma Membrane," *Structure*, vol. 23, no. 9, pp. 1715–1724, 2015.

[17] A. Kuhn and D. Kiefer, "Membrane protein insertase YidC in bacteria and archaea," *Molecular Microbiology*, vol. 103, no. 4, pp. 590–594, 2017.

[18] H. Chen, D. Ogden, S. Pant, W. Cai, E. Tajkhorshid, M. Moradi, B. Roux, and C. Chipot, "A Companion Guide to the String Method with Swarms of Trajectories: Characterization, Performance, and Pitfalls," *Journal of Chemical Theory and Computation*, vol. 18, pp. 1406–1422, 2022.

[19] A. Polasa, J. Hettige, K. Immadisetty, and M. Moradi, "An Investigation of the YidC-Mediated Membrane Insertion of Pf3 Coat Protein Using Molecular Dynamics Simulations," *bioRxiv*, 2022.

[20] R. E. Dalbey and M. Chen, "Sec-translocase mediated membrane protein biogenesis," *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1694, no. 1-3 SPEC.ISS., pp. 37–53, 2004.

[21] D. Spann, E. Pross, Y. Chen, R. E. Dalbey, and A. Kuhn, "Each protomer of a dimeric yidc functions as a single membrane insertase," *Scientific Reports*, vol. 8, no. 1, p. 589, 2018.

[22] P. A. Scotti, M. L. Urbanus, J. Brunner, J. W. L. De Gier, G. Von Heijne, C. Van Der Does, A. J. Driessen, B. Oudega, and J. Luirink, "YidC, the Escherichia coli homologue of mitochondrial Oxa1p, is a component of the Sec translocase," *EMBO Journal*, vol. 19, no. 4, pp. 542–549, 2000.

[23] S. J. Facey and A. Kuhn, "Membrane integration of E. coli model membrane proteins," *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1694, no. 1-3 SPEC.ISS., pp. 55–66, 2004.

[24] J. C. Samuelson, M. Chen, F. Jiang, I. Möller, M. Wiedmann, A. Kuhn, G. J. Phillips, and R. E. Dalbey, "YidC mediates membrane protein insertion in bacteria," *Nature*, vol. 406, no. 6796, pp. 637–641, 2000.

[25] D. Kiefer and A. Kuhn, "Yidc-mediated membrane insertion," *FEMS Microbiology Letters*, vol. 365, 2018.

[26] P. R. Laskowski, K. Pluhackova, M. Haase, B. M. Lang, G. Nagler, A. Kuhn, and D. J. Müller, "Monitoring the binding and insertion of a single transmembrane protein by an insertase," *Nature Communications*, vol. 12, no. 1, p. 7082, 2021.

[27] A. J. O. Lewis and R. S. Hegde, "A unified evolutionary origin for the ubiquitous protein transporters SecY and YidC," *BMC Biology*, vol. 19, no. 1, p. 266, 2021.

[28] Y. Endo, Y. Shimizu, H. Nishikawa, K. Sawasato, and K.-i. Nishiyama, "Interplay between MPIase, YidC, and PMF during Sec-independent insertion of membrane proteins," *Life Science Alliance*, vol. 5, no. 1, 2022.

[29] Y. Xin, Y. Zhao, J. Zheng, H. Zhou, X. C. Zhang, C. Tian, and Y. Huang, "Structure of YidC from Thermotoga maritima and its implications for YidC-mediated membrane protein insertion," *FASEB Journal*, vol. 32, no. 5, pp. 2411–2421, 2018.

[30] J. Yuan, G. J. Phillips, and R. E. Dalbey, "Isolation of cold-sensitive yidC mutants provides insights into the substrate profile of the YidC insertase and the importance of transmembrane 3 in YidC function," *Journal of Bacteriology*, vol. 189, no. 24, pp. 8961–8972, 2007.

[31] C. Klenner and A. Kuhn, "Dynamic disulfide scanning of the membrane-inserting Pf3 coat protein reveals multiple YidC substrate contacts," *Journal of Biological Chemistry*, vol. 287, no. 6, pp. 3769–3776, 2012.

[32] S. Kol, N. Nouwen, and A. J. Driessen, "Mechanisms of YidC-mediated insertion and assembly of multimeric membrane protein complexes," *Journal of Biological Chemistry*, vol. 283, no. 46, pp. 31269–31273, 2008.

[33] S. Ernst, A. K. Schönbauer, G. Bär, M. Börsch, and A. Kuhn, "YidC-driven membrane insertion of single fluorescent Pf3 coat proteins," *Journal of Molecular Biology*, vol. 412, no. 2, pp. 165–175, 2011.

[34] S. Funes, F. Kauff, E. O. Van Der Sluis, M. Ott, and J. M. Herrmann, "Evolution of YidC/Oxa1/Alb3 insertases: Three independent gene duplications followed by functional specialization in bacteria, mitochondria and chloroplasts," *Biological Chemistry*, vol. 392, no. 1-2, pp. 13–19, 2011.

[35] S. Funes, A. Hasona, H. Bauerschmitt, C. Grubbauer, F. Kauff, R. Collins, P. J. Crowley, S. R. Palmer, L. J. Brady, and J. M. Herrmann, "Independent gene duplications of the YidC/Oxa/Alb3 family enabled a specialized cotranslational function," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 16, pp. 6656–6661, 2009.

[36] R. Kohler, D. Boehringer, B. Greber, R. Bingel-Erlenmeyer, I. Collinson, C. Schaffitzel, and N. Ban, "YidC and Oxa1 Form Dimeric Insertion Pores on the Translating Ribosome," *Molecular Cell*, vol. 34, no. 3, pp. 344–353, 2009.

[37] H. He, A. Kuhn, and R. E. Dalbey, "Tracking the Stepwise Movement of a Membrane-inserting Protein In Vivo," *Journal of Molecular Biology*, vol. 432, no. 2, pp. 484–496, 2020.

[38] K. Kumazaki, T. Kishimoto, A. Furukawa, H. Mori, Y. Tanaka, N. Dohmae, R. Ishitani, T. Tsukazaki, and O. Nureki, "Crystal structure of Escherichia coli YidC, a membrane protein chaperone and insertase," *Scientific Reports*, vol. 4, p. 7299, 2014.

[39] K. Kumazaki, S. Chiba, M. Takemoto, A. Furukawa, K. I. Nishiyama, Y. Sugano, T. Mori, N. Dohmae, K. Hirata, Y. Nakada-Nakura, A. D. Maturana, Y. Tanaka, H. Mori, Y. Sugita, F. Arisaka, K. Ito, R. Ishitani, T. Tsukazaki, and O. Nureki, "Structural basis of Sec-independent membrane protein insertion by YidC," *Nature*, vol. 509, no. 7501, pp. 516–519, 2014.

[40] Y. Chen, R. Soman, S. K. Shanmugam, A. Kuhn, and R. E. Dalbey, "The role of the strictly conserved positively charged residue differs among the gram-positive, gram-negative, and chloroplast YidC homologs," *Journal of Biological Chemistry*, vol. 289, no. 51, pp. 35656–35667, 2014.

[41] Y. Chen, S. Capponi, L. Zhu, P. Gellenbeck, J. A. Freites, S. H. White, and R. E. Dalbey, "YidC Insertase of Escherichia coli: Water Accessibility and Membrane Shaping," *Structure*, vol. 25, no. 9, pp. 1403–1414.e3, 2017.

[42] R. J. Schulze, J. Komar, M. Botte, W. J. Allen, S. Whitehouse, V. A. M. Gold, J. A. L. a Nijeholt, K. Huard, I. Berger, C. Schaffitzel, and I. Collinson, "Membrane protein insertion and proton-motive-force-dependent secretion through the bacterial holo-translocon SecYEG-SecDF-YajC-YidC," *Proceedings of the National Academy of Sciences*, vol. 111, no. 13, pp. 4844–4849, 2014.

[43] N. Shimokawa-Chiba, K. Kumazaki, T. Tsukazaki, O. Nureki, K. Ito, and S. Chiba, "Hydrophilic microenvironment required for the channel-independent insertase function of YidC protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 16, pp. 5063–5068, 2015.

[44] A. Kedrov, S. Wickles, A. H. Crevenna, E. O. van der Sluis, R. Buschauer, O. Berninghausen, D. C. Lamb, and R. Beckmann, "Structural Dynamics of the YidC:Ribosome Complex during Membrane Protein Biogenesis," *Cell Reports*, vol. 17, no. 11, pp. 2943–2954, 2016.

[45] D. Kiefer and A. B. T. I. R. o. C. Kuhn, "YidC as an Essential and Multifunctional Component in Membrane Protein Assembly," *Academic Press*, vol. 259, pp. 113–138, 2007.

[46] A. Gallusser and A. Kuhn, "Initial steps in protein membrane insertion. Bacteriophage M13 procoat protein binds to the membrane surface by electrostatic interaction," *The EMBO journal*, vol. 9, no. 9, pp. 2723–2729, 1990.

[47] S. Wickles, A. Singharoy, J. Andreani, S. Seemayer, L. Bischoff, O. Berninghausen, J. Soeding, K. Schulten, E. O. van der Sluis, and R. Beckmann, "A structural model of the active ribosome-bound membrane protein insertase YidC," *eLife*, vol. 3, no. July2014, pp. 1–17, 2014.

[48] K. Ito, N. Shimokawa-Chiba, and S. Chiba, "Sec translocon has an insertase-like function in addition to polypeptide conduction through the channel," *F1000Research*, vol. 8, pp. F1000 Faculty Rev–2126, 12 2019.

[49] Y. Tanaka, A. Izumioka, A. Abdul Hamid, A. Fujii, T. Haruyama, A. Furukawa, and T. Tsukazaki, "2.8-Å crystal structure of Escherichia coli YidC revealing all core regions, including flexible C2 loop," *Biochemical and Biophysical Research Communications*, vol. 505, no. 1, pp. 141–145, 2018.

[50] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and A. D. MacKerell Jr, "CHARMM36m: an improved force field for folded and intrinsically disordered proteins," *Nature methods*, vol. 14, pp. 71–73, jan 2017.

[51] J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, and R. W. Pastor, "Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types," *Journal of Physical Chemistry B*, vol. 114, no. 23, pp. 7830–7843, 2010.

[52] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with NAMD," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1781–1802, 2005.

[53] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *The Journal of Chemical Physics*, vol. 79, no. 2, pp. 926–935, 1983.

[54] S. Jo, T. Kim, and W. Im, "Automated builder and database of protein/membrane complexes for molecular dynamics simulations," *PLoS ONE*, vol. 2, no. 9, 2007.

[55] J. K. Reid, "On the method of conjugate gradients for the solution of large sparse systems of linear equations," pp. 231–254, 1971.

[56] G. J. Martyna, D. J. Tobias, and M. L. Klein, "Constant pressure molecular dynamics algorithms," *J. Chem. Phys.*, vol. 101, no. 5, pp. 4177–4189, 1994.

[57] S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, "Constant pressure molecular dynamics simulation: The Langevin piston method," *J. Chem. Phys.*, vol. 103, no. 11, pp. 4613–4621, 1995.

[58] V. Govind Kumar, D. S. Ogden, U. H. Isu, A. Polasa, J. Losey, and M. Moradi, "Prefusion Spike Protein Conformational Changes Are Slower in SARS-CoV-2 than in SARS-CoV-1," *Journal of Biological Chemistry*, vol. 298, 2022.

[59] K. Immadisetty, A. Polasa, R. Shelton, and M. Moradi, "Elucidating the Molecular Basis of pH Activation of an Engineered Mechanosensitive Channel," *bioRxiv*, 2021.

[60] A. Polasa, I. Mosleh, J. Losey, A. Abbaspourrad, R. Beitle, and M. Moradi, "Developing a Rational Approach to Designing Recombinant Proteins for Peptide-Directed Nanoparticle Synthesis," *Nanoscale Adv.*, vol. 4, pp. 3161–3171, 2022.

[61] M. Moradi, G. Enkavi, and E. Tajkhorshid, "Atomic-level characterization of transport cycle thermodynamics in the glycerol-3-phosphate:phosphate transporter," *Nat. Commun.*, vol. 6, p. 8393, 2015.

[62] M. Moradi, C. Sagui, and C. Roland, "Calculating relative transition rates with driven nonequilibrium simulations," *Chem. Phys. Lett.*, vol. 518, pp. 109–113, 2011.

[63] S. C. Moradi, Mahmoud and C. Roland, "Investigating rare events with nonequilibrium work measurements: I. nonequilibrium transition path probabilities," *J. Chem. Phys.*, vol. 140, no. 3, p. 034114, 2014.

[64] M. Moradi and E. Tajkhorshid, "Mechanistic picture for conformational transition of a membrane transporter at atomic resolution," *Proc. Natl. Acad. Sci. USA*, vol. 110, no. 47, pp. 18916–18921, 2013.

[65] D. Ogden and M. Moradi, "Molecular Dynamics–Based Thermodynamic and Kinetic Characterization of Membrane Protein Conformational Transitions," *Structure and Function of Membrane Proteins*, pp. 289–309, 2021.

[66] K. Immadisetty, A. Polasa, R. Shelton, and M. Moradi, "Elucidating the Molecular Basis of Spontaneous Activation in an Engineered Mechanosensitive Channel," *Computational and Structural Biotechnology Journal*, vol. 20, pp. 2539–2550, 2022.

[67] T. Harkey, V. Govind Kumar, J. Hettige, S. H. Tabari, K. Immadisetty, and M. Moradi, "The role of a crystallographically unresolved cytoplasmic loop in stabilizing the bacterial membrane insertase yidc2," *Scientific Reports*, vol. 9, no. 1, p. 14451, 2019.

[68] W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual molecular dynamics," *Journal of Molecular Graphics*, vol. 14, no. 1, pp. 33–38, 1996.

[69] K. Immadisetty, J. Hettige, and M. Moradi, "What Can and Cannot Be Learned from Molecular Dynamics Simulations of Bacterial Proton-Coupled Oligopeptide Transporter GkPOT?," *J. Phys. Chem. B*, vol. 121, pp. 3644–3656, 2017.

[70] A. Bakan, L. M. Meireles, and I. Bahar, "ProDy: Protein dynamics inferred from theory and experiments," *Bioinformatics*, vol. 27, no. 11, pp. 1575–1577, 2011.

[71] A. Sethi, J. Eargle, A. A. Black, and Z. Luthey-Schulten, "Dynamical networks in tRNA:protein complexes," *Proc. Natl. Acad. Sci. USA*, vol. 106, pp. 6620–6625, 2009.

[72] D. K. Brown, D. L. Penkler, O. S. Amamuddy, C. Ross, A. R. Atilgan, C. Atilgan, and Ö. T. Bishop, "MD-TASK: a software suite for analyzing molecular dynamics trajectories," *Bioinformatics*, vol. 33, pp. 2768–2771, 2017.

[73] M. Chen, J. C. Samuelson, F. Jiang, M. Muller, A. Kuhn, and R. E. Dalbey, "Direct interaction of YidC with the Sec-independent Pf3 coat protein during its membrane protein insertion," *Journal of Biological Chemistry*, vol. 277, no. 10, pp. 7670–7675, 2002.

[74] Z. Yu, G. Koningstein, A. Pop, and J. Luirink, "The conserved third transmembrane segment of YidC contacts nascent Escherichia coli inner membrane proteins," *Journal of Biological Chemistry*, vol. 283, no. 50, pp. 34635–34642, 2008.

[75] N. Shimokawa-Chiba, K. Kumazaki, T. Tsukazaki, O. Nureki, K. Ito, and S. Chiba, "Hydrophilic microenvironment required for the channel-independent insertase function of YidC protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 16, pp. 5063–5068, 2015.

# Chapter III: Developing a Rational Approach to Designing Recombinant Proteins for Peptide-Directed Nanoparticle Synthesis

Adithya Polasa[1], Imann Mosleh[2,3], James Losey[1], Alireza Abbaspourrad [2],
Robert Beitle[3], and Mahmoud Moradi[1] *

[1]Department of Chemistry and Biochemistry, University of Arkansas,
Fayetteville, AR 72701, USA.

[2]Department of Food Science, College of Agriculture and Life Sciences, Cornell University,
Ithaca, NY, 14853, USA

[3]Ralph E. Martin Department of Chemical Engineering, University of Arkansas,
Fayetteville, AR 72701, USA

## Abstract

The controlled formation of nanoparticles with optimum characteristics and functional aspects has proven successful via peptide-mediated nanoparticle synthesis. However, the effects of peptide sequence and binding motif on surface features and physicochemical properties of nanoparticles are not well-understood. In this study, we investigate in a comparative manner how a specific peptide known as the Pd4 and its two known variants may form nanoparticles both in an isolated state and when attached to the green fluorescent protein (GFPuv). More importantly, we introduce a novel computational approach to predict the trend of the size and activity of the peptide-directed nanoparticles by estimating the binding affinity of the peptide to a single ion. We used molecular dynamics (MD) simulations to explore the differential behavior of the isolated and GFP-fused peptides and their mutants. Our computed palladium (Pd) binding free energies match the typical nanoparticle sizes reported from transmission electron microscope pictures. Stille coupling and Suzuki-Miyaura reaction turnover frequencies (TOF) also correspond with computationally predicted Pd binding affinities. The results show that while using Pd4 and its two known variants (A6 and A11) in isolation produces nanoparticles of varying sizes, fusing these peptides to the GFPuv protein produces nanoparticles of similar sizes and activity. In other words, the GFPuv

reduces the sensitivity of the nanoparticles to the peptide sequence. This study provides a computational framework for designing free and protein-attached peptides that helps the synthesis of nanoparticles with well-regulated properties.

## Introduction

Recent advances in nanotechnology have resulted in the development of various efficient synthesis and characterization procedures for nanoparticles. Within the last two decades, in particular, many peptides have been introduced to identify inorganic metal surfaces [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14], some of which have been utilized to produce nanomaterials [15, 13, 14]. Production of nanoparticles with varying size, shape, or aggregation stability using peptide immobilization is useful in sub-fields of biotechnology [16], sensors [17, 18, 19], and bioanalytical procedures [20, 21, 22]. Small variations in the composition and sequence of the peptide conjugate may have a significant effect on the configuration of the resulting nanoparticle assembly. It has become apparent that by carefully tailoring the peptide sequence, one can regulate the composition of the nanoparticles [23].

Self-assembly motif peptides guide nanoparticle assembly into specific architectures [23, 24]. Peptides may produce well-defined nanostructures such as nanotubes, nanofibers, nanoparticles, nanotapes, gels, and nanorods by self-assembly [25]. Changes to the amino acid sequence of a peptide used in the synthesis of a nanoparticle can alter the nanoparticle properties. Studies have found evidence that oligopeptides with tryptophan and tyrosine in their sequence are potentially involved in reducing metal ions into their respective metals thus forming nanoparticles [26, 27, 28]. In contrast, histidine (His) containing oligopeptides, bound to the material surfaces, may increase the interaction between the solvent and metallic surface [29]. Recent computational studies have demonstrated that aromatic residues His10 and His12 for Pd2 and His6 and His11 for Pd4 have lower surface interaction and mobility on the surface of palladium nanoparticles compared to other peptides in the study [29]. Substitution of these His residues with an alanine (Ala) affects the reactivity and nanoparticle fabrication capability of the peptide, resulting in varying turnover fre-

quencies (TOFs) for the Stille coupling reaction [29]. Furthermore, minor mutations in the amino acid sequence, such as cysteine and alanine modifications, will significantly reduce the surface structure of nanocatalysts, lowering the peptide absorption energy of the palladium nanoparticle [30, 31, 32]. The difference between His6 and His11 in the Pd4 peptide has been the focus of some studies, as mutation of either of these residues is known to completely modify the structural and functional abilities of the peptide in nanoparticle synthesis[24, 30]. A computational study has proposed that His6 has a slightly lower surface interaction energy and mobility on the surface of the palladium substrate than His11 [29]. However, His11 has more interaction sites than His6, which explains the difference in the free energy profiles of these residues in many studies [24, 30, 29]. Attaching the peptides to a green fluorescent protein (GFPuv) using recombinant fusion proteins can offer an alternative strategy to using chemically synthesized, isolated peptides that is more cost effective [33, 34, 35]. Additionally, studies have shown that the use of this GFPuv protein helps in monitoring the synthesis of the nanoparticle in a single-step process [36]. Despite the success of this approach to successfully produce functioning nanoparticles similar to common peptide-mediated nanoparticles, relying only on experimental techniques does not provide the insight needed to understand the phenomenon at the molecular level. Here the focus of our study is to develop a computational framework for investigating the peptide-nanoparticle interaction in both isolated peptides and protein-attached peptides. We do not claim to have a methodology that fully explains all aspects of the peptide-nanoparticle interaction but we only focus, in this work, on the interaction with a single Pd ion. Interestingly, the Pd binding free energies measured for several systems including isolated and GFPuv-fused Pd4 and its two known mutants (H6A and H11A denoted by A6 and A11, respectively) provided a good predictor of the behavior of the resulting nanoparticles both in terms of their sizes and activities. Employing all-atom molecular dynamics (MD) simulations and enhanced sampling techniques, we have investigated the differential behavior of isolated and protein-fused Pd4 peptides and its mutants at the atomic level. The peptides studied in this MD simulation study are the peptide models Pd4 (TSNAV**H**PTLR**H**L) and alanine substitution mutants called A6 (TSNAV**A**PTLR**H**L) and A11 (TSNAV**H**PTLR**A**L) (Fig. 1A-C).

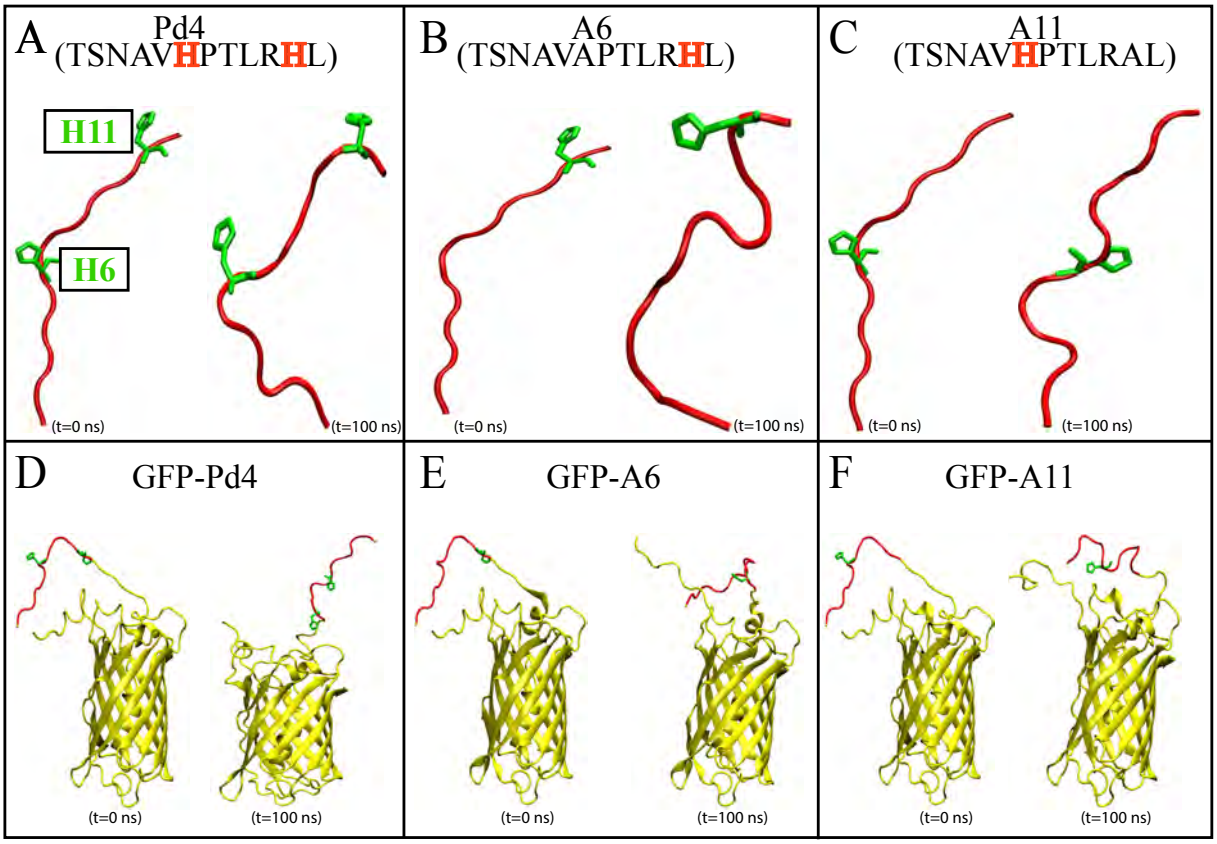


Figure 1: The initial and final MD snapshots of Pd4, A6, A11 (A-C), GFP-Pd4, GFP-A6 and GFP-A11 (D-F) in cartoon representation. The peptide in each protein is colored red and GFPuv is colored yellow (D-F). Histidines in the peptide are colored green and shown in licorice representation.

Additionally, for each peptide described above, computationally generated GFPuv bound peptide models (Fig. 1D-F) were created. Following structural modeling of the peptides, MD simulations were performed in an aqueous environment, and free energy calculations were done to measure the absolute binding free energy between each His and Pd ion. Experimental recombinant peptide fusion GFPuv-mediated palladium nanoparticle synthesis results for particle size and turnover frequencies (TOF) were used to validate our computational results, and provide a comparison to existing experimental data for free peptides. These simulations provide valuable insights into the sequence-dependent structural and functional dynamics of the peptides at the atomic level. The results of this study could potentially be used to guide the design or selection of peptides for isolated or protein-fused peptide-directed nanoparticle synthesis.

## Methods

All-atom MD simulations were performed to obtain relaxed structures of peptide models to investigate conformational dynamics and calculate binding energies of His residues on the surface of Pd ion. For the fusion peptides, the crystal structure of GFPuv (PDB: 1W7S) [37] was obtained from protein data bank with a resolution of 1.85 Å. Initially, Modeller [38] was used for the construction of all peptide and GFPuv fusions (Fig. 1). Next, CHARMM-GUI [39, 40] web-server was used to build the MD simulation models of peptides and GFPuv fusions in aqueous solution of TIP3P [41] water. $Na^+$ and $Cl^-$ ions were used to both neutralize and add 0.15 M of salt to each system. The total number of atoms for peptide and GFPuv fusion peptide systems was ≈22,300 and ≈68,000, respectively. NAMD 2.13 [42] was utilized to run the MD simulations with periodic boundary conditions (PBC) at 310K in the NPT ensemble, and 1 atm pressure was maintained using the Nosé-Hoover Langevin piston method [43, 44]. Initially, we used the conjugate gradient [45] technique to energy-minimize each system for 10,000 steps. Following that, we used the standard CHARMM-GUI [46] protocol to progressively relax the systems using restricted MD simulations. An NVT ensemble was used for the initial relaxation, whereas an NPT ensemble was used for all production runs. Simulations were performed using a Langevin integrator with a damping value of $\gamma = 0.5$ ps$^{-1}$ and a 2 fs time step at 310 K temperature. Every system was equilibrated for 100 nanoseconds using CHARMM36m all-atom force field [47, 48] parameters. VMD [49] was used to visualize and analyze trajectory data. Principal component analysis (PCA) was performed on each trajectory using PRODY [50] software. Only $C_\alpha$ atoms of peptide were considered in the PCA calculations of both free and fused peptide simulations. Secondary structure analysis of each residue was done using the $\phi - \psi$ angles [51] to identify various areas in the Ramachandran plot. The definitions of different regions in the Ramachandran plot (F, $\beta$, $\alpha_R$, and $\alpha_L$ regions) labeled in Fig. 2 and Fig. 4 are based on boundaries defined by Moradi et al. [51]. Further, we calculated the binding free energy of palladium ion and histidines in the peptide sequence using following methodology.

**Binding free energy calculations**

The Pd binding free energy of histidines in the peptide was calculated for all systems mentioned above (Fig. 1) using the free energy perturbation (FEP) method [52]. All simulations were carried out using NAMD 2.13 [42]. For FEP calculations, 20 snapshot from a 100 ns equilibrated system trajectory were obtained, and the $Pd^{2+}$ ion was positioned within 2 Å of the His aromatic ring structure's center of mass. For Pd ion, the force field parameters developed by Heinz et al. [53] were used, which is already available within CHARMM-METAL [47]. The CHARMM-METAL [53] parameters provide reliable results for calculating values such as the surface binding energy of metals (e.g., palladium). This force field model, on the other hand, ignores the contribution of metal atom polarizability to dynamic forces as well as a realistic representation of the spatial and orientational structure of liquid water at the metal interface [54]. These characteristics are expected to have a significant role in biomolecule adsorption behavior [54]. However, in previous computational investigations of peptide binding for bio-based nanomaterials [55, 56], using the CHARMM-METAL force field, significantly demonstrated the relationship between peptide surface binding and nanoparticle size and structure. Thus, we used the CHARMM-METAL force field to demonstrate the qualitative difference between free peptide and GFPuv fused peptide binding to nanoparticles and show a phenomenological relationship between histidine binding affinity and nanoparticle size and catalytic activity. Employing polarizable force fields would be beneficial in future studied to gain a deeper understanding of the relationship between nanoparticle synthesis and binding free energies. Solvation free energy calculations for palladium ion was carried out under PBC conditions at constant pressure in aqueous TIP3P [41] water solution. The systems were simulated with a 2 fs time step using Langevin dynamics at a temperature of 310K. Initially, we utilized the conjugate gradient [45] approach to minimize energy in each system for 10,000 steps. Every system was equilibrated for 3 nanoseconds (ns). To constrain the distance between the Pd ion and the His aromatic ring structure in all FEP simulations, we employed a harmonic restraint [57] in terms of the center of mass distance between Pd ion and His aromatic ring structure. To restrain the protein conformational dynamics, we also used the root mean square deviation (RMSD) of the

protein as a collective variable [57]. FEP simulations were performed by varying $\lambda$ from zero to one (forward) or from one to zero (backward) in 32 stages (i.e., $\Delta\lambda$ =0.03125). Each stage had a simulation time of 100 *ps* including 2 *ps* of equilibration and 98 *ps* of data collection. A soft-core potential was used to avoid the endpoint problem for van der Waals interactions [58, 59] and the electrostatic interactions were gradually decoupled from $\lambda = 0$ to 0.9 in the forward direction and they were coupled from $\lambda = 0.9$ to 0 in the backward direction.

The free energy values were computed using Bennette Acceptance Ratio method [60] within the parsefep plugin [61] from forward and backward simulations. The thermodynamic cycle illustrated in Figure S1 was used for final calculation of the binding free energies. This cycle connects the binding of palladium ion with the histidine in peptide individually for each system from unbound state to bound state, i.e., $\Delta G_{unbound.\rightarrow bound}$= $\Delta G_2$ with solvation of palladium ion in aqueous solution $\Delta G_{vacuum.\rightarrow aqueous}$= $\Delta G_1$ ($\Delta G_1$ = 282.5 kcal/mol). Overall, the binding free energy of palladium was calculated based on the equation $\Delta G_{binding}$ = $\Delta G_1$ - $\Delta G_2$ derived from the thermodynamic cycle (Fig. S1). The mean value and standard deviation of the mean were estimated from 20 independent FEP calculations based on 20 different snapshots of each system obtained as described above.

**Fusion Protein Preparation**

According to the codon preference of *E.coli*, plasmids encoding Pd4, A6, and A11 peptides fused to GFPuv were constructed. The fragment containing codons of peptides were introduced to $5'$ end of the GFPuv gene using primer forward primers through polymerase chain reaction (PCR). Synthetic genes containing desired peptides and addgene-plasmid-51559 were double digested with EcoRI and XbaI restriction enzymes before the ligation. Finally, the designed DNA containing desired peptides were ligated to the DNA to construct plasmids containing GFPuv fusion proteins. Bacterial lysates from arabinose induced cells were obtained, followed a method adapted from our previous research [33]. Protein concentrations were determined using the detergent compatible (DC) protein assay (Bio-Rad, Hercules, CA).

**Nanoparticle synthesis**

To synthesize palladium nanoparticles (Pd NPs) at room temperature, 0.16 mg $K_2PdCl_4$ was added to synthesis mixtures (one milliliter total volume) containing 0.23 mg fusion protein. These amounts result in a 2:1 ratio of $Pd^{+2}$ to Pd4. After mixing for 0.5 hour, 1.5 mg $NaBH_4$ was added to the mixture to reduce $Pd^{+2}$ ions to metallic Pd and NPs were formed rapidly after reduction indicated by a color change (yellow to light brown). The NP shape and size distribution were analyzed using transmission electron microscopy (TEM). A droplet containing ten microliter samples of the reaction mixture were placed on a 300 mesh standard lacey carbon grid. FEI Titan $80-300$ instrument was used to producing nanoparticle size distribution and morphology images. TEM images were analyzed using ImageJ software [33] in order to measure particle sizes. Experimental data on nanoparticle size for the three systems were limited. Using bootstrapping, the particle size data was re-sampled 100 times for each system for statistics calculations. The bootstrap method involves iteratively resampling a set of data to generate a sub-sample that is smaller than or equal to the size of the given data set. This sub-sample is produced using replacement, which allows each data point to be sampled multiple times or not at all. The bootstrap particle size data was also processed to generate probability density functions using the kernel density estimator. Corresponding histograms (Fig. 3A) were generated using the 100 probability density functions derived from kernel density estimate.

**Screening of reaction parameters for Suzuki-Miyaura coupling reaction**

For optimizing the reaction condition, a method developed by Mosleh *et al.,* was adopted [33]. The different reaction conditions including base, temperature, and solvent were evaluated using the model coupling reaction (Table S1). The reaction did not result in high yield when KOtBu, $K_2HPO_4$, and $KH_2PO_4$ (entries 1-3) were used as base while the reaction proceeded with excellent catalytic activity in the presence of $K_2CO_3$ (entry 4). The catalytic performance of reaction under different temperatures indicated that by increasing the temperature, higher yields could be obtained. Indeed, the presence of EtOH in water-contained solvents as a green solvent with the

ratio of 1:1 was found to be the best solvent for Suzuki-Miyaura coupling reaction (entries 6-8).

### Screening of reaction parameters for Stille coupling reaction

To obtain the optimized condition for Stille coupling reaction, the reaction of iodobenzene and phenyltin trichloride was used as the model reaction and other reaction conditions including base, temperature, and catalyst loading were evaluated (Table S2). For the base study, a series of bases were explored and $K_3PO_4$ was found to be the best base as the yield of biphenyl production was 97% (entries 1-4). Higher yields could be obtained when higher temperatures were used. Although 70% yield of biphenyl was proceeded at 60 °C, the reaction was performed for 20 hrs. Employing 80 °C resulted in biphenyl production with 96% yield after 6 hrs. Furthermore, increasing the amount of catalyst did not alter the yield of biphenyl while lower catalytic activities were observed when 2 mmol% and 1 mmol% of Pd was present in the reaction (entries 7-9).

## Results and Discussion

### Histidine-palladium binding free energies for free peptides

The Pd4, A6, and A11 peptides described above were first computationally modeled and equilibrated using MD as described in the Methods section. The binding free energy of Pd with histidine residues at positions 6 and 11 was then calculated using the free energy perturbation (FEP) method [52] to determine the site-specific binding strengths of the peptide (see Methods section - Binding free energy calculations). The Pd binding free energies of the histidines are reported in Table 2 along with average particle size and turnover frequency (TOF) from Ref. 24 and Ref. 30.

Table 2: Binding free energy of Pd nanoparticles with free peptide

| **Peptide** | **Binding Free Energy(kcal/mol)** | | | **TOF($h^{-1}$)** | | **Average Particle Size (nm)** | |
|---|---|---|---|---|---|---|---|
| | His6 | His11 | Effective | Ref. 24 | Ref. 30 | Ref. 24 | Ref. 30 |
| **Pd4** | -89±5 | -157±8 | -157±8 | 2234±99 | 2200±100 | 1.9±0.3 | 2.1±0.4 |
| **A6** | | -145±7 | -145±7 | 5224±381 | 5200±400 | 2.2±0.4 | 2.2±0.7 |
| **A11** | -118±6 | | -118±6 | 1298±107 | 1300±10 | 2.4±0.5 | 2.6±0.4 |

The differential binding free energy values highlight the importance of the histidine residues

in these free peptides. As the peptide sequence is modified, the binding free energy value changes significantly. For systems with His11 present, the free energy of Pd ion binding with His11 was -157 $\pm$ 8 and -145 $\pm$ 7 kcal/mol for Pd4 and A6, respectively. On the other hand, the free energy for binding with His6, was -89 $\pm$ 5 and -118 $\pm$ 6 kcal/mol, in Pd4 and A11, respectively. The lower free energy for binding with His11 made it the preferred Pd binding site when it was available, which is directly related to its high affinity and TOF value in palladium nanoparticle synthesis. Therefore, the His11 has a much higher affinity to Pd than His6. His6 is also potentially involved in the production of nanoparticles, even in the absence and presence of His11 in A11 and Pd4 peptides, respectively. However, it is unlikely to be the dominant binding site due to its low affinity for the palladium ion, possibly resulting in a lower TOF value.

The binding free energies from Table 2 provided more insight into existing catalytic rate data of TOF values from Ref. 24 and Ref. 30. Even though the Pd binding free energy of His11 in A6 is slightly higher than that of His11 in Pd4, the TOF for A6 was doubled, from 2200 $\pm$ 100 $h^{-1}$ (Pd4) to 5200 $\pm$ 400 $h^{-1}$ (A6). Conversely, when only the His6 is present, as in A11, the TOF decreases to 1298 $\pm$ 107 $h^{-1}$. With both His6 and His11 present in Pd4, the catalytic activity was slowed down. The lower TOF in Pd4 as compared to A6 has been attributed to the shared interaction activity of the two histidines in Pd4, where a larger area of nanoparticle is covered by the peptide when two histidines can bind to the nanoparticle 30. It is apparent from the binding free energy data that both histidine residues can be involved in the nanoparticle interaction. In prior studies, it was discovered that when the peptide was altered at residue His6, the experimentally estimated binding free energy of the peptide had a greater affinity for palladium nanoparticle than when modified at residue His11 [55, 56].

Other synthesis measures showed similar relationship with the free energy calculations. Existing nanoparticle size data from transmission electron microscopy (TEM) measurements [24, 30] displayed a linear relationship with the minimum binding free energy of the histidines (Fig. 3A). Higher binding free energy of the palladium with the peptide during synthesis correlated with larger particle size. These free energy calculations shed light on the competitive relationships be-

tween histidine binding sites that determine catalytic activity of the peptide and the size of the nanoparticle produced. To investigate these differences, the peptide simulation trajectories were subjected to further in-depth examination.

### Secondary structure propensity for free peptides

To evaluate the influence of the mutations on the secondary structure of peptide, the $\phi - \psi$ angles of the peptides' residues 6 and 11 were analyzed on Ramachandran plots [62, 51] (Fig. 2).

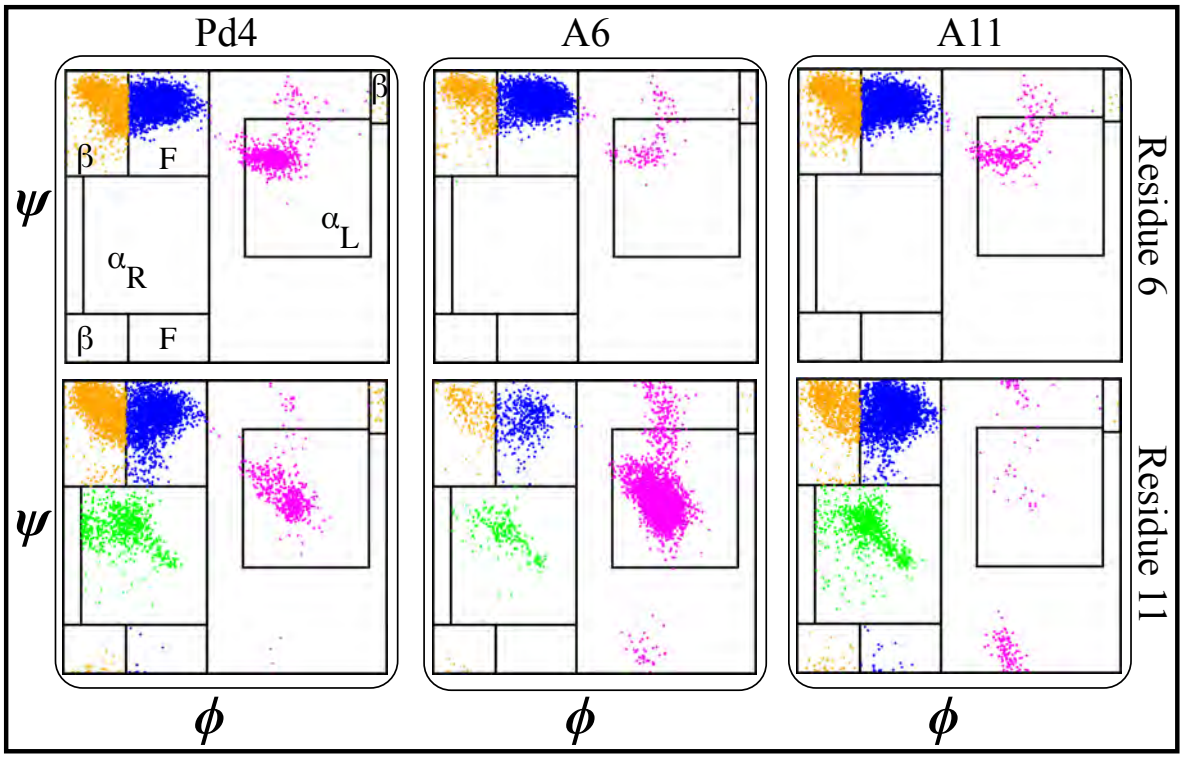


Figure 2: The structural propensities of peptide Pd4, A6, and A11 shown as Ramachandran plots (x-axis $\phi$ and y-axis $\psi$ angles) of residue 6 (top) and residue 11 (bottom). Secondary structures are colored as follows: $\alpha_R$-green, $\alpha_L$-magenta, $\beta$-orange, and F-blue.

This analysis revealed that the $\alpha_L$ propensity of residues was significantly affected by the mutation of His6 and His11 to alanine. The secondary structure of the residues was primarily acquired within the F area of the Ramachandran plots in wild type Pd4 peptides (Fig. 2). However, the mutation H6A in A6 changed the secondary structure of His11, which features a significantly higher

$\alpha_L$ propensity value giving it more $\alpha$ helical structure compared to the residue in peptides Pd4 and A11 (Fig. 2). In A11, the mutation of the His11 to an alanine changed the propensity of the $\alpha_L$ secondary structure of residue 11 but not as prominently as in the A6 and Pd4 peptide structure (Fig. 2). This explain the low TOF of A11 than Pd4 and A6 (Table 2). The $\alpha_L$ propensities of both His6 and His11 in the Pd4 peptide are comparable, but not as high as His11 in the A6 peptide, As a result, we believe that Pd4 has lower binding free energy and TOF than A6 (Table 2). This secondary structure difference could be a reason for the difference in the TOF and free energy results, as the $\alpha$ structure of the amino acid is directly proportional to the reactivity because of the rigid structure of the residue. Although similar results have been reported in previous study [55], our method for calculating this is unique and distinct, and the major purpose of this study is to examine the differences between peptide and GFP-fused recombinant peptides (Fig. S2 & S3). The following section of our findings show how the fusing of GFPuv protein to these peptides affects the structure of the peptide and the production of nanoparticles.

**GFPuv fusion peptide simulations and experiments**

Table 3: Binding free energy of Pd nanoparticles with GFPuv fused peptides

| **Peptide** | **Binding Free Energy (kcal/mol)** | | | **TOF ($h^{-1}$)** | | **Average Particle Size (nm)** |
|---|---|---|---|---|---|---|
| | His6 | His11 | Effective | Stille | Suzuki-Miyaura | |
| **GFP-Pd4** | -118±6 | -122±6 | -122±6 | 2945±103 | 11731±839 | 2.6±0.5 |
| **GFP-A6** | | -122±6 | -122±6 | 2912±101 | 11093±481 | 2.7±0.7 |
| **GFP-A11** | -118±6 | | -118±6 | 2942±65 | 10867±443 | 2.6±0.4 |

To study the ability of the GFPuv fusion peptide framework to produce Pd nanoparticles, MD and FEP simulations of peptides bound to the GFPuv protein were conducted. GFPuv fused peptide directed palladium nanoparticles were generated experimentally and used as catalysts for the Stille and Suzuki-Miyaura coupling reactions for characterization. A schematic of the different reactions and a plot of the catalytic data is presented in Fig. S4. Turn over frequency (TOF) observed in these catalytic rate results are reported in Table 3. The nanoparticles directed by three GFPuv fused peptide (Pd4, A6 and A11) systems were analyzed by Transmission electron microscopy

(TEM) to measure particle size (Fig. S5). Regardless of mutation, the in silico binding free energy of palladium with histidines in GFPuv fused peptides are consistent with each other (Table 3). Interestingly, the binding free energy estimates for His6 are identical (-118$\pm$6 kcal/mol) in GFP-Pd4 and GFP-A6. Similarly, the binding free energy estimates for His11 are identical (-122$\pm$6 kcal/mol) in GFP-Pd4 and GFP-A11. Even His6 and His11 have quite similar binding free energies with only 4 kcal/mol difference, which is within the uncertainty of the calculations. Note that these binding free energies are estimated from 20 distinct sets of independent FEP simulations each with a different initial conformation, so our findings are statistically significant. TOF and nanoparticle size estimates, on the other hand, follow the same pattern as the free energy values in that the mutations do not have a significant impact on the size and function of the nanoparticles. Previous studies [63, 29] of nanoparticle synthesis using free peptides, produced nanoparticles with a varying particle sizes. This could be because of difference in the effective binding free energy between the histidine and Pd ion in the free peptides.

To test this hypothesis, we investigated the particle size of nanoparticles made with recombinant fusion peptides to see whether there was a relationship between free energy values and the size of the particles produced. Average particle sizes of $2.6 \pm 0.5$ (nm), $2.7 \pm 0.7$ (nm), and $2.6 \pm 0.4$ (nm) were observed when the nanoparticles were prepared from GFPuv fused Pd4, A6 and A11 peptides, respectively. The particle sizes of the three systems are not only identical on average, but also have similar density distributions (Fig. 3A). The TEM particle size findings for every GFPuv fusion peptide-mediated nanoparticle production were re-sampled 100 times using bootstrapping to estimate kernel density distribution (Fig. 3A). These nanoparticle particle size (Table 3) and distribution results (Fig. 3A) are significantly similar and have a consistent trend like binding free energies calculated in our study (Table 3). In Still coupling processes, similarities between TOF values were also observed for all fusion peptides, which follows the same pattern as the binding free energies. The Suzuki-Miyaura reactions provided a different and higher TOF values than Stille coupling method, but the uniform catalytic behavior of the fusion peptides persisted across these two different reaction types. When compared to free peptide simulations results

(Table 2), the TOF of the fusion peptides in nanoparticle synthesis are different for each peptide, which is expected because the binding free energy of all the His combinations in free state are significantly different. Whereas in the fused form, the TOF observations from both tests are considerably comparable; as a result, the binding free energies of all His combinations in the fused state are quite similar. This confirms that the binding free energy value of His residues with Pd ion are proportional to the catalytic rate (Table 3).

As previously discussed the TOF is varied when either histidine in the free Pd4 peptide is mutated to alanine with A6 and A11 having the greatest and lowest TOF, respectively (Table 1). The TOF variability is not observed in GFP fused peptides (Table 2). While the lower TOF in A11 is justifiable by assuming that His11 is more reactive than His6, which is also consistent with our free energy estimates (Table 1), the reason for higher TOF of the A6 peptide is less clear. This has been attributed to the presence of two histidines in Pd4 and their ability to cover more of the nanoparticle surface as compared to one histidine in A6 [30]. However, this justification does not seem to be consistent with what we have observed with the GFP fused peptides. Here, we observe very similar TOF values for A6, A11, and Pd4 peptides when attached to the GFP protein. Although we agree with Bedford et al. [30] that the presence of two histidines rather than one causes the lower TOF in Pd4 as compared to A6, we hypothesize that the intramolecular interaction of the two histidines may be the more important factor in lowering the activity as compared to the shared interaction with the nanoparticle. This argument is consistent with the fact that we observe significantly more frequent histidine-histidine interactions in free Pd4 than GFP-Pd4. To quantify the interactions more accurately, we have measured His6-His11 interaction energies in Pd4 and GFP-Pd4. We specifically measured the frequency of observing His6-His11 interaction energies that are larger than 1 kcal/mol in magnitude (as a measure of non-negligible interactions). In GFP-Pd4, this frequency is only 7% as compared to 23% in the free Pd4, indicating there is a 3-fold decrease in the frequency of His6-His11 interactions when Pd4 is attached to GFP as compared to when it is free. Alternatively, one may consider the interaction energy distributions, which similarly indicates that larger interaction energies are observed more frequently in free Pd4 than

in GFP-Pd4 (Fig. S6). We therefore hypothesize that the histidine-histidine interactions may be behind the lower TOF of Pd4 relative to A6 and since these interactions are less significant in GFP fused peptide, we no longer observe a lower TOF for GFP-Pd4 relative to GFP-A6.

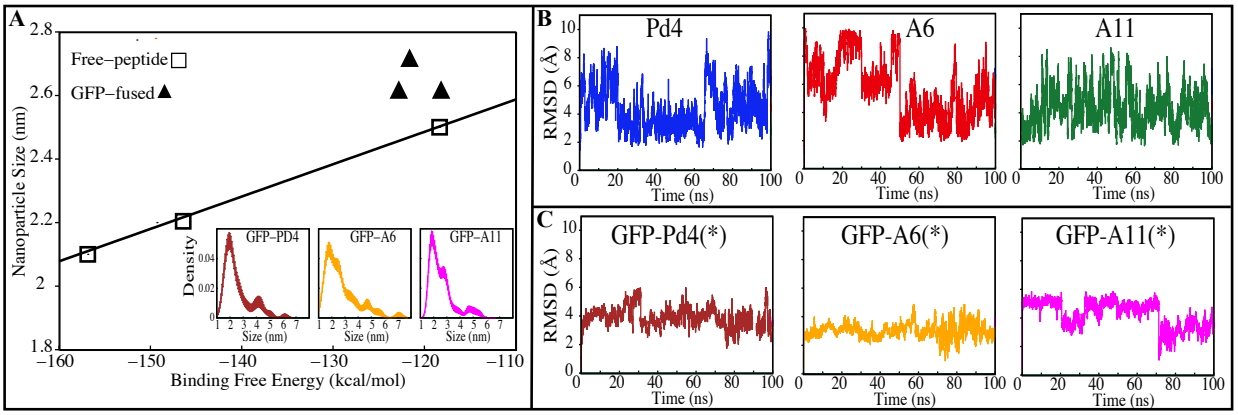


Figure 3: (A) Nanoparticle size as a function of binding free energies (line represents linear regression of free peptide results). The inset shows the kernel density estimation of 100 bootstrap resamples of experimentally measured nanoparticle sizes. Size distribution for nanoparticles synthesized using GFPuv fusion peptides. (B-C) RMSD profile of free peptides and GFPuv fused peptides, respectively. (*) Only the peptide region was used for RMSD calculation in (C).

All GFPuv fusion peptides show a similar trend in binding free energy, TOF, average nanoparticle size and distribution, supporting our above assumption. Root mean square deviation (RMSD) of the peptides were stable and unvarying in all GFPuv fused systems (Fig. 3C). This implies that GFPuv might be stabilizing the peptide structure during nanoparticle production. In contrast to fused peptides, the free peptide backbone RMSD (Fig. 3B) fluctuated more. Based on the RMSD data, we believe that the conformational stability of the peptide is important in the production of uniform nanoparticles. In support of this claim, all GFPuv fusion peptide experiments produced consistent nanoparticles. The GFPuv controlled the peptide structural variations throughout nanoparticle synthesis, resulting in a quite similar nanoparticle size and distribution (Fig. 3A) across all GFPuv fused peptide-mediated nanoparticle synthesis.

**Secondary structural propensity for GFPuv fusion peptides**

A secondary structure analysis of the simulated GFPuv fusion peptides revealed further differences between free peptide simulations and the fusion peptide simulations. The Ramachandran plots of peptides Pd4, A6, and A11 bound to GFPuv are shown in Fig. 4.

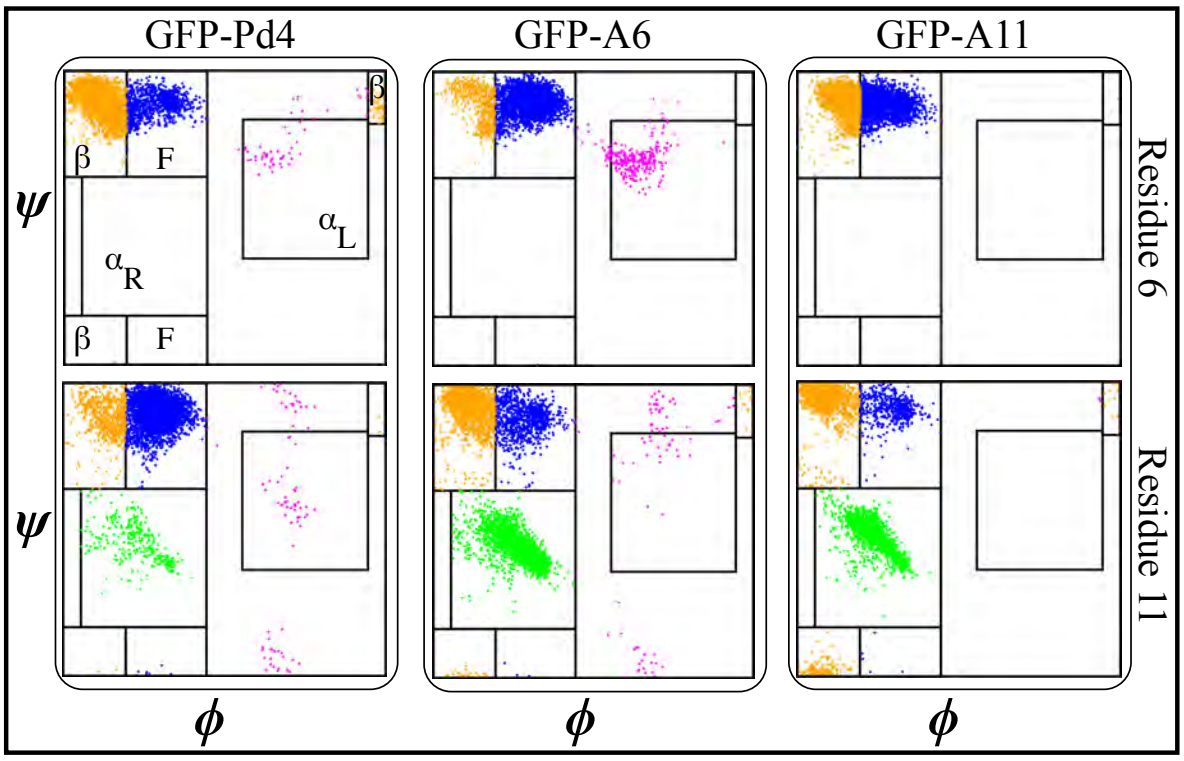


Figure 4: Structural propensity of the GFPuv fused peptide Pd4, A6, and A11 shown as Ramachandran plots (x-axis $\phi$ and y-axis $\psi$ angles) of residue 6 (top row) and residue 11 (bottom row). Secondary structures are colored with $\alpha_R$ as green, $\alpha_L$ as magenta, $\beta$ as orange, and F as blue.

The coloring cluster of the plots are consistent with Fig. 2. The propensity for $\alpha_L$ states was lower in the all the bound peptide simulations. Also, the relationship between histidine and $\alpha_L$ propensity observed in the free peptide simulations did not exist for the GFPuv fusion peptides. The presence of histidine at position 6 and 11 had either no correlation or inverse correlation to the $\alpha_L$ propensity. For GFP-Pd4, the $\alpha_L$ propensity was almost identical for residue 6 and 11. In the GFP-A6 simulations, the histidine at residue 6 showed a very slight increase in $\alpha_L$ propensity,

while residue 11 had a a secondary structural propensity similar to GFP-Pd4. Finally, GFP-A11 showed near zero $\alpha_L$ states for both residues, regardless of the presence of a histidine. Unlike the free peptide secondary structure properties discussed previously, the secondary structure of amino acid residues at positions 6 and 11 remains consistent across all GFP-bound states. Furthermore, in contrast to the free peptide-mediated method, the lack of high $\alpha_L$ propensity differences resulted in identical binding free energy and similar size nanoparticles in GFPuv fusion peptides approach. Hence, secondary structure of the protein ultimately altering the peptide behaviour, which directly implicated on the nanoparticle size (Fig. 3), catalytic reactivity (Table. 3) and binding free energy (Table. 3) calculations in our examination. The experimental TOF calculations and TEM (Table 3) outcomes for GFPuv fused peptides strongly supports the computationally determined free energy (Table 3) results and structural properties of peptides (Fig. 4), thus showing the effectiveness of the approach. Peptides are generated chemically, which makes them expensive. GFP-fused recombinant peptides, on the other hand, can be made more quickly and at a lower cost.

**Hydrogen bond analysis**

To further explore the structural differences observed in simulation, a hydrogen bond analysis was performed. Hydrogen bonds are significant interactions in proteins and peptides, contributing to backbone conformational stability differences [64]. To quantify hydrogen bonds in the peptide backbones, the hydrogen bond occupancies were calculated with bond length and angle cutoffs of 4.0 Å and 40°, respectively, for the simulated trajectories of the free and the fusion peptides. To account for the weak hydrogen bond interactions, we used a relatively loose definition of hydrogen bonding; however, a more strict definition (distance and angle cutoffs of 3.5 Å and 30°, respectively) gives qualitatively similar results (Fig. S7). In the hydrogen bond analysis of the free peptides (Fig. 5A), the A6 peptide formed more stable backbone hydrogen bond interactions throughout the course of the simulation than Pd4 and A11, which both had very few hydrogen bonds in the occupancy percentage. We discovered the hydrogen bonds that resulted in significant occupancy in the A6 peptide in free peptide simulations. There are two major hydrogen bond

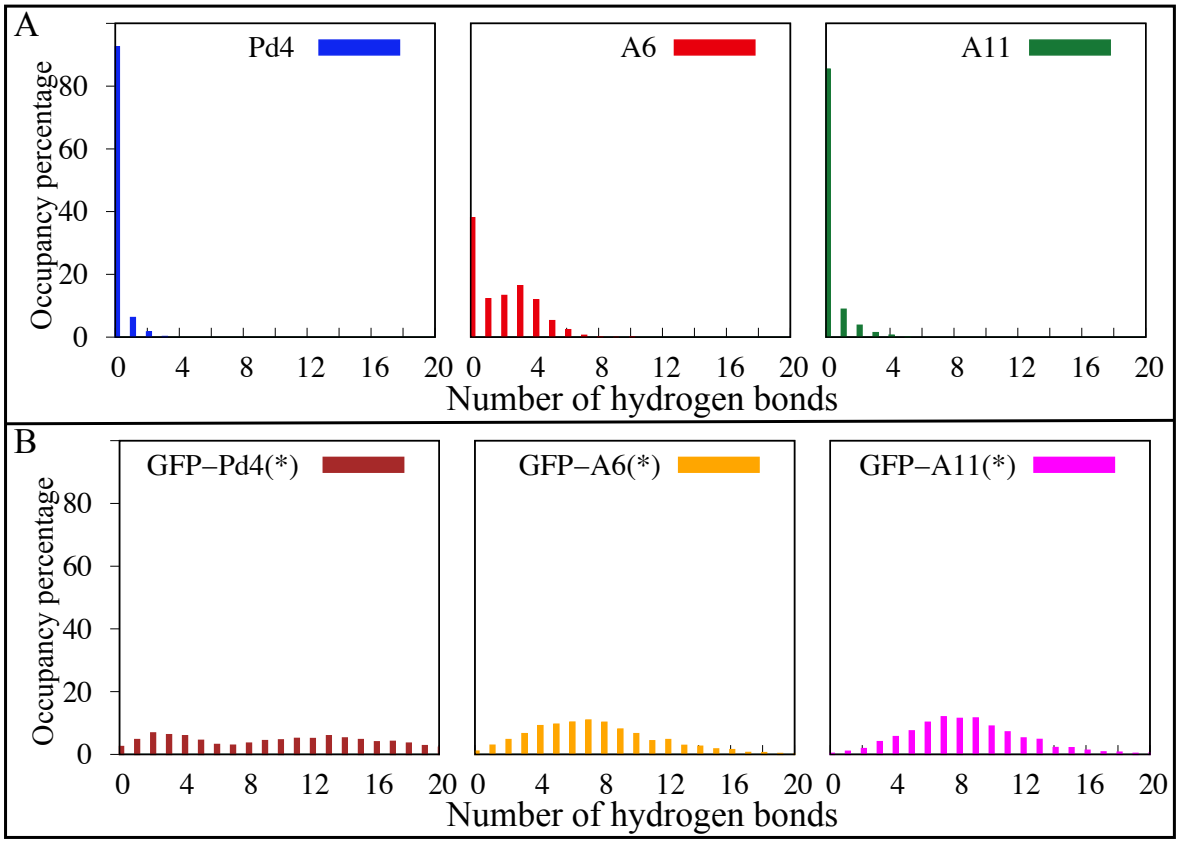


Figure 5: The occupancy percentage of hydrogen bonds calculated from the MD simulations for the (A) free peptides and (B) GFPuv fused peptide region. (*) Only the peptide region in the GFPuv fused peptide simulations was used for hydrogen bond analysis.

interactions (R10-P7 and A6-L9). The A6-L9 bond is particularly important, with a 38 percent occupancy. This hydrogen bond was formed only in the A6 peptide, and being associated with the mutated residue suggests that the H6A (A6) mutation may play a critical role in protein stability. In the A6 peptide, another hydrogen bond with a 46 percent occupancy was formed between R10 and P7 residues. Residue 11 had the greatest $\alpha_L$ propensity in the A6 free peptide simulations, which might be due to the hydrogen bond between R10 and P7, as they are adjacent residues. Since, hydrogen bond interactions are also responsible for backbone conformational stability and secondary structure differences [64]. As a result, we can see the difference in secondary structure between Pd4 and A11 with absence of these two hydrogen bonds.

In Fig. 5B, hydrogen bond interactions of just the peptide region for GFPuv fusion peptides are

reported. The distribution of hydrogen bonds was qualitatively similar for all the GFPuv fusion peptides, with a broad distribution in the number of hydrogen bonds observed. The inclusion of GFPuv improves the peptide stabilization by forming evenly distributed hydrogen bonds, resulting in decreased RMSD of the peptides and a comparable $\alpha_L$ propensity regardless of mutation. The principal component analysis (PCA) of $C_\alpha$ atoms in Fig. S8 also reveals substantial differences in the conformations of A6, Pd4, and A11. Pd4 and A11 are equally dispersed along PC1 and PC2, but A6 clustered differently. When peptides were fused to GFPuv, the PCA of the fusion peptides showed even higher stability, with thicker clusters than free peptides suggesting minimal variation during simulations. Our findings imply that the GFPuv fused peptide can govern the size and functioning of nanoparticles by stabilizing peptide fluctuations, secondary structure, and surface binding abilities, resulting in a uniform nanoparticle and TOF in experimental data. Mutations in GFPuv fused peptides had no effect on the binding free energies and secondary structure of the peptide, as well as the experimentally measured nanoparticle size and TOF values. Hence, all of our computational findings are in agreement with the experiment results.

Using a combination of simulations and experiments, we have shown a phenomenological relationship between histidine binding affinity, secondary structure, peptide size, and catalytic reactivity. Our simulation models, however, do not explain why a relationship exists between TOF and binding free energy. We believe that histidine binding affinity, peptide conformational stability, and secondary structure all influence catalytic reactivity and size of nanoparticle. However, understanding these relationships at a molecular level requires a more detailed study and a more complete picture of nanoparticle synthesis and activity, which may enable the development of more effective peptides for nanoparticle formation. Our findings may contribute to the development of other simulation studies investigating the observed relationships.

## Conclusions

Overall, our computational and experimental results have added molecular level detail to the existing sequence dependent different catalytic results in Pd nanoparticle production using free pep-

tides. This research demonstrates how a single amino acid substitution in the free Pd4 peptide sequence may alter the structure and catalytic properties of nanoparticle production. Our findings demonstrate that the H6A mutation increased the $\alpha_L$ propensity of residue 11 (Fig. 2) and hydrogen bond occupancy (Fig. 5) in the peptide backbone. This increased the affinity of palladium ions for residue 11 in A6 peptide. Peptide stability and secondary structure have a significant impact on these differences in nanoparticle production induced by single mutations in free peptide. The uniformity in experimental turn over frequency (TOF), particle size (Table 3) and distribution (Fig. 3A) results of nanoparticles using GFPuv fused peptides showed consistent results for all nanoparticle produced and without any aggregation, the nanoparticles remained stable. One of the advantages of our method is that the coupling process was carried out in a water/EtOH mixture, which resulted in a sufficient dipole moment, which prevented the nanoparticles from aggregating even at 80 degrees Celsius. The catalytic activity were attributed to Pd nanoparticles, which are crucial in coupling processes. Based on simulation predictions GFPuv fused peptides, are unaffected by peptide sequence changes which are reported in free peptide conditions. GFPuv acted as a stabilizer when linked to peptides. Since the peptide was stabilized in a fused state, the palladium binding free energy with the histidines (Table 3), RMSD (Fig. 3C), secondary structure (Fig. 4), and hydrogen bond occupancy (Fig. 5) are all extremely comparable between all systems.

This paper describes a novel approach for generating multi-functional peptides with distinct amino acid domains for the cost-effective production of nanoparticles. Using FEP/MD simulations, the approach used in this study might be used to successfully manufacture additional nano catalysts, as well as to explain and perhaps find peptide regions critical to nanoparticle production.

## References


[1] D. J. Gaskin, K. Starck, and E. N. Vulfson, "Identification of inorganic crystal-specific sequences using phage display combinatorial library of short peptides: A feasibility study," *Biotechnology Letters*, vol. 22, no. 15, pp. 1211–1216, 2000.

[2] S. R. Whaley, D. S. English, E. L. Hu, P. F. Barbara, and A. M. Belcher, "Selection of peptides with semiconductor binding specificity for directed nanocrystal assembly," *Nature*, vol. 405, no. 6787, pp. 665–668, 2000.

[3] R. R. Naik, S. E. Jones, C. J. Murray, J. C. McAuliffe, R. A. Vaia, and M. O. Stone, "Peptide templates for nanoparticle synthesis derived from polymerase chain reaction-driven phage display," *Advanced Functional Materials*, vol. 14, no. 1, pp. 25–30, 2004.

[4] R. R. Naik, S. J. Stringer, G. Agarwal, S. E. Jones, and M. O. Stone, "Biomimetic synthesis and patterning of silver nanoparticles," *Nature Materials*, vol. 1, no. 3, pp. 169–172, 2002.

[5] K. I. Sano and K. Shiba, "A Hexapeptide Motif that Electrostatically Binds to the Surface of Titanium," *Journal of the American Chemical Society*, vol. 125, no. 47, pp. 14234–14235, 2003.

[6] S. Wang, E. S. Humphreys, S. Y. Chung, D. F. Delduco, S. R. Lustig, H. Wang, K. N. Parker, N. W. Rizzo, S. Subramoney, Y. M. Chiang, and A. Jagota, "Peptides with selective affinity for carbon nanotubes," *Nature Materials*, vol. 2, no. 3, pp. 196–200, 2003.

[7] M. B. Dickerson, R. R. Naik, M. O. Stone, Y. Cai, and K. H. Sandhage, "Identification of peptides that promote the rapid precipitation of germania nanoparticle networks via use of a peptide display library," *Chemical Communications*, vol. 4, no. 15, pp. 1776–1777, 2004.

[8] C. K. Thai, H. Dai, M. S. Sastry, M. Sarikaya, D. T. Schwartz, and F. Baneyx, "Identification and characterization of Cu2O- and ZnO-binding polypeptides by escherichia coli cell surface display: Toward an understanding of metal oxide binding," *Biotechnology and Bioengineering*, vol. 87, no. 2, pp. 129–137, 2004.

[9] C. Mao, D. J. Solis, B. D. Reiss, S. T. Kottmann, R. Y. Sweeney, A. Hayhurst, G. Georgiou, B. Iverson, and A. M. Belcher, "Virus-Based Toolkit for the Directed Synthesis of Magnetic and Semiconducting Nanowires," *Science*, vol. 303, no. 5655, pp. 213–217, 2004.

[10] B. D. Reiss, C. Mao, D. J. Solis, K. S. Ryan, T. Thomson, and A. M. Belcher, "Biological routes to metal alloy ferromagnetic nanostructures," *Nano Letters*, vol. 4, no. 6, pp. 1127–1132, 2004.

[11] R. Zuo, D. Örnek, and T. K. Wood, "Aluminum- and mild steel-binding peptides from phage display," *Applied Microbiology and Biotechnology*, vol. 68, no. 4, pp. 505–509, 2005.

[12] M. Umetsu, M. Mizuta, K. Tsumoto, S. Ohara, S. Takami, H. Watanabe, I. Kumagai, and T. Adschiri, "Bioassisted room-temperature immobilization and mineralization of zinc oxide - The structural ordering of ZnO nanoparticles into a flower-type morphology," *Advanced Materials*, vol. 17, no. 21, pp. 2571–2575, 2005.

[13] M. Sarikaya, C. Tamerler, A. K. Jen, K. Schulten, and F. Baneyx, "Molecular biomimetics: Nanotechnology through biology," *Nature Materials*, vol. 2, no. 9, pp. 577–585, 2003.

[14] J. M. Slocik and R. R. Naik, "Biologically programmed synthesis of bimetallic nanostructures," *Advanced Materials*, vol. 18, no. 15, pp. 1988–1992, 2006.

[15] M. Sarikaya, C. Tamerler, D. T. Schwartz, and F. Baneyx, "Materials assembly and formation using engineered polypeptides," *Annual Review of Materials Research*, vol. 34, pp. 373–408, 2004.

[16] R. Narayanan and M. A. El-Sayed, "Catalysis with transition metal nanoparticles in colloidal solution: Nanoparticle shape dependence and stability," *Journal of Physical Chemistry B*, vol. 109, no. 26, pp. 12663–12676, 2005.

[17] C. K. Kim, R. R. Kalluru, J. P. Singh, A. Fortner, J. Griffin, G. K. Darbha, and P. C. Ray, "Gold-nanoparticle-based miniaturized laser-induced fluorescence probe for specific DNA hybridization detection: Studies on size-dependent optical properties," *Nanotechnology*, vol. 17, no. 13, pp. 3085–3093, 2006.

[18] J. Liu and Y. Lu, "Colorimetric biosensors based on DNAzyme-assembled gold nanoparticles," *Journal of Fluorescence*, vol. 14, no. 4, pp. 343–354, 2004.

[19] X. Luo, A. Morrin, A. J. Killard, and M. R. Smyth, "Application of nanoparticles in electrochemical sensors and biosensors," *Electroanalysis*, vol. 18, no. 4, pp. 319–326, 2006.

[20] J. J. Storhoff, R. Elghanian, R. C. Mucic, C. A. Mirkin, and R. L. Letsinger, "One-pot colorimetric differentiation of polynucleotides with single base imperfections using gold nanoparticle probes," *Journal of the American Chemical Society*, vol. 120, no. 9, pp. 1959–1964, 1998.

[21] S. Liang, D. T. Pierce, C. Amiot, and X. Zhao, "Photoactive nanomaterials for sensing trace analytes in biological samples," *Synthesis and Reactivity in Inorganic, Metal-Organic and Nano-Metal Chemistry*, vol. 35, no. 9, pp. 661–668, 2005.

[22] S. G. Penn, L. He, and M. J. Natan, "Nanoparticles for bioanalysis," *Current Opinion in Chemical Biology*, vol. 7, no. 5, pp. 609–615, 2003.

[23] L. Hwang, G. Zhao, P. Zhang, and N. L. Rosi, "Size-controlled peptide-directed synthesis of hollow spherical gold nanoparticle superstructures," *Small (Weinheim an der Bergstrasse, Germany)*, vol. 7, pp. 1938–1942, jul 2011.

[24] R. Coppage, J. M. Slocik, M. Sethi, D. B. Pacardo, R. R. Naik, and M. R. Knecht, "Elucidation of peptide effects that control the activity of nanoparticles," *Angewandte Chemie - International Edition*, vol. 49, no. 22, pp. 3767–3770, 2010.

[25] N. Habibi, N. Kamaly, A. Memic, and H. Shafiee, "Self-assembled peptide-based nanostructures: Smart nanomaterials toward targeted drug delivery," *Nano Today*, vol. 11, no. 1, pp. 41–60, 2016.

[26] P. R. Selvakannan, S. Mandal, S. Phadtare, A. Gole, R. Pasricha, S. D. Adyanthaya, and M. Sastry, "Water-dispersible tryptophan-protected gold nanoparticles prepared by the spontaneous reduction of aqueous chloroaurate ions by the amino acid," *Journal of Colloid and Interface Science*, vol. 269, no. 1, pp. 97–102, 2004.

[27] S. Si and T. K. Mandal, "Tryptophan-based peptides to synthesize gold and silver nanoparticles: A mechanistic and kinetic study," *Chemistry - A European Journal*, vol. 13, no. 11, pp. 3160–3168, 2007.

[28] R. R. Bhattacharjee, A. K. Das, D. Haldar, S. Si, A. Banerjee, and T. K. Mandal, "Peptide-assisted synthesis of gold nanoparticles and their self-assembly," *Journal of Nanoscience and Nanotechnology*, vol. 5, no. 7, pp. 1141–1147, 2005.

[29] R. B. Pandey, H. Heinz, J. Feng, B. L. Farmer, J. M. Slocik, L. F. Drummy, and R. R. Naik, "Adsorption of peptides (A3, Flg, Pd2, Pd4) on gold and palladium surfaces by a coarse-grained Monte Carlo simulation," *Physical Chemistry Chemical Physics*, vol. 11, no. 12, pp. 1989–2001, 2009.

[30] N. M. Bedford, H. Ramezani-Dakhel, J. M. Slocik, B. D. Briggs, Y. Ren, A. I. Frenkel, V. Petkov, H. Heinz, R. R. Naik, and M. R. Knecht, "Elucidation of peptide-directed palladium surface structure for biologically tunable nanocatalysts," *ACS Nano*, vol. 9, no. 5, pp. 5082–5092, 2015.

[31] N. M. Bedford, A. R. Showalter, T. J. Woehl, Z. E. Hughes, S. Lee, B. Reinhart, S. P. Ertem, E. B. Coughlin, Y. Ren, T. R. Walsh, and B. A. Bunker, "Peptide-Directed PdAu Nanoscale Surface Segregation: Toward Controlled Bimetallic Architecture for Catalytic Materials," *ACS Nano*, vol. 10, no. 9, pp. 8645–8659, 2016.

[32] H. Ramezani-Dakhel, N. M. Bedford, T. J. Woehl, M. R. Knecht, R. R. Naik, and H. Heinz, "Nature of peptide wrapping onto metal nanoparticle catalysts and driving forces for size control," *Nanoscale*, vol. 9, no. 24, pp. 8401–8409, 2017.

[33] I. Mosleh, M. Benamara, L. Greenlee, M. H. Beyzavi, and R. Beitle, "Recombinant peptide fusion proteins enable palladium nanoparticle growth," *Materials Letters*, vol. 252, pp. 68–71, 2019.

[34] I. Mosleh, H. R. Shahsavari, R. Beitle, and M. H. Beyzavi, "Recombinant Peptide Fusion Protein-Templated Palladium Nanoparticles for Suzuki-Miyaura and Stille Coupling Reactions," *ChemCatChem*, vol. 12, no. 11, pp. 2942–2946, 2020.

[35] Tejada-Vaprio, Rita and Mosleh, Imann and Mukherjee, Rudra Palash and Aljewari, Hazim and Fruchtl, McKinzie and Elmasheiti, Ahmed and Bedford, Nicholas and Greenlee, Lauren and Beyzavi, M. Hassan and Beitle, Robert, "Recombinant peptide fusion construction for protein-templated catalytic palladium nanoparticles," *Biotechnology Progress*, vol. 36, no. 3, p. 3, 2020.

[36] P. Sanpui, S. B. Pandey, S. S. Ghosh, and A. Chattopadhyay, "Green fluorescent protein for in situ synthesis of highly uniform Au nanoparticles and monitoring protein denaturation," *Journal of Colloid and Interface Science*, vol. 326, no. 1, pp. 129–137, 2008.

[37] J. J. Van Thor, G. Y. Georgiev, M. Towrie, and J. T. Sage, "Ultrafast and low barrier motions in the photoreactions of the green fluorescent protein," *Journal of Biological Chemistry*, vol. 280, no. 39, pp. 33652–33659, 2005.

[38] A. Ali, "Comparative protein modeling by satisfaction of spatial restraints," *Molecular Medicine Today*, vol. 1, no. 6, pp. 270–277, 1995.

[39] S. Jo, T. Kim, V. G. Iyer, and W. Im, "{CHARMM-GUI}: A Web-Based Graphical User Interface for {CHARMM}," vol. 29, pp. 1859–1865, 2008.

[40] J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks, A. D. MacKerell, J. B. Klauda, and W. Im, "CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field," *Journal of Chemical Theory and Computation*, vol. 12, no. 1, pp. 405–413, 2016.

[41] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *The Journal of Chemical Physics*, vol. 79, no. 2, pp. 926–935, 1983.

[42] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with NAMD," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1781–1802, 2005.

[43] G. J. Martyna, D. J. Tobias, and M. L. Klein, "Constant pressure molecular dynamics algorithms," *The Journal of Chemical Physics*, vol. 101, no. 5, pp. 4177–4189, 1994.

[44] S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, "Constant pressure molecular dynamics simulation: The Langevin piston method," *The Journal of Chemical Physics*, vol. 103, no. 11, pp. 4613–4621, 1995.

[45] B. N. Parlett, "Large Sparse Sets of Linear Equations (J. K. Reid, ed.); Sparse Matrices and Their Applications (Donald J. Rose and Ralph A. Willoughby, ed.)," *SIAM Review*, vol. 16, no. 3, pp. 396–398, 1974.

[46] S. Jo, T. Kim, and W. Im, "Automated builder and database of protein/membrane complexes for molecular dynamics simulations," *PLoS ONE*, vol. 2, no. 9, 2007.

[47] R. B. Best, X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, and A. D. MacKerell, "Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone $\phi$, $\psi$ and side-chain $\chi 1$ and $\chi 2$ Dihedral Angles," *Journal of Chemical Theory and Computation*, vol. 8, no. 9, pp. 3257–3273, 2012.

[48] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. De Groot, H. Grubmüller, and A. D. MacKerell, "Charmm36m: an improved force field for folded and intrinsically disordered proteins," *Nature methods*, vol. 14, no. 1, pp. 71–73, 2017.

[49] W. Humphrey, A. Dalke, and K. Schulten, "VMD: visual molecular dynamics," *J. Mol. Graphics*, vol. 14, pp. 33–38, 1996.

[50] A. Bakan, L. M. Meireles, and I. Bahar, "ProDy: Protein dynamics inferred from theory and experiments," *Bioinformatics*, vol. 27, no. 11, pp. 1575–1577, 2011.

[51] M. Moradi, C. Sagui, and C. Roland, "Calculating relative transition rates with driven nonequilibrium simulations," *Chem. Phys. Lett.*, vol. 518, pp. 109–113, 2011.

[52] R. W. Zwanzig, "High-temperature equation of state by a perturbation method. i. nonpolar gases," *The Journal of Chemical Physics*, vol. 22, no. 8, pp. 1420–1426, 1954.

[53] J. Henin, E. Tajkhorshid, K. Schulten, and C. Chipot, "Diffusion of glycerol through *Escherichia coli* aquaglyceroporin GlpF," *Biophys. J.*, vol. 94, pp. 832–839, 2008.

[54] Z. E. Hughes and T. R. Walsh, "Distinct Differences in Peptide Adsorption on Palladium and Gold: Introducing a Polarizable Model for Pd(111)," *The Journal of Physical Chemistry C*, vol. 122, pp. 19625–19638, aug 2018.

[55] R. Coppage, J. M. Slocik, B. D. Briggs, A. I. Frenkel, R. R. Naik, and M. R. Knecht, "Determining Peptide Sequence Effects That Control the Size, Structure, and Function of Nanoparticles," *ACS Nano*, vol. 6, pp. 1625–1636, feb 2012.

[56] R. Coppage, J. M. Slocik, H. Ramezani-Dakhel, N. M. Bedford, H. Heinz, R. R. Naik, and M. R. Knecht, "Exploiting Localized Surface Binding Effects to Enhance the Catalytic Reactivity of Peptide-Capped Nanoparticles," *Journal of the American Chemical Society*, vol. 135, pp. 11048–11054, jul 2013.

[57] G. Fiorin, M. L. Klein, and J. Hénin, "Using collective variables to drive molecular dynamics simulations," *Mol. Phys.*, vol. 111, no. 22–23, pp. 3345–3362, 2013.

[58] M. Zacharias, T. P. Straatsma, and J. A. McCammon, "Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration," *The Journal of Chemical Physics*, vol. 100, pp. 9025–9031, June 1994.

[59] T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, and W. F. van Gunsteren, "Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations," *Chemical Physics Letters*, vol. 222, no. 6, pp. 529–539, 1994.

[60] C. H. Bennett, "Efficient estimation of free energy differences from Monte Carlo data," *Journal of Computational Physics*, vol. 22, no. 2, pp. 245–268, 1976.

[61] P. Liu, F. Dehez, W. Cai, and C. Chipot, "A toolkit for the analysis of free-energy perturbation calculations," *Journal of Chemical Theory and Computation*, vol. 8, pp. 2606–2616, 08 2012.

[62] G. N. RAMACHANDRAN, C. RAMAKRISHNAN, and V. SASISEKHARAN, "Stereochemistry of polypeptide chain configurations," *Journal of molecular biology*, vol. 7, pp. 95–99, jul 1963.

[63] M. B. Dickerson, S. E. Jones, Y. Cai, G. Ahmad, R. R. Naik, N. Kröger, and K. H. Sandhage, "Identification and design of peptides for the rapid, high-yield formation of nanoparticulate TiO2 from aqueous solutions at room temperature," *Chemistry of Materials*, vol. 20, no. 4, pp. 1578–1584, 2008.

[64] C. Pace, H. Fu, K. Fryar, J. Landua, S. Trevino, D. Schell, R. Thurlkill, S. Iimura, J. M. Scholtz, K. Gajiwala, J. Sevcik, L. Urbániková, J. Myers, K. Takano, E. Hebert, B. Shirley, and G. Grimsley, "Contribution of hydrogen bonds to protein stability," *Protein science : a publication of the Protein Society*, vol. 23, p. 5, 05 2014.

## Conclusion

Through the use of MD simulations, we were able to construct an empirical framework for predicting the peptides that could be used in the effective synthesis of nanoparticles and investigate the insertion mechanisms of gram-negative and gram-positive YidC.

For the first time, we investigated the molecular aspects of YidC independent function in gram-positive and gram-negative bacteria using all-atom MD and biased MD simulation methodology (such as steered molecular dynamics).We have presented a comprehensive examination of the many different YidC domains that each play a crucial role in the stability and function of YidC. We have uncovered significant interactions that play a key role in the protein's stability as well as its function. In conclusion, the findings of our study point to a mechanism by which gram-positive and gram-negative YidC may insert without the involvement of SecY.

The computationally hypothesized Sec-independent insertion mechanism of polypeptides by YidC is in excellent accord with the findings from studies that were done in the past. According to the findings of this research, the cytoplasmic loops not only play an extremely important part in the stability of YidC but also serve as the primary contact locations at which insertion may take place. Incoming Pf3 coat protein contacts cytoplasmic loops and then penetrates the hydrophilic groove, forming a salt bridge with R72. This salt bridge stabilizes the Pf3 coat protein in the YidC TM groove at the initial stage of insertion. Hydrophilic interactions inside the groove facilitate the transit of the protein towards the periplasmic side. The salt bridge between D18 of the Pf3 coat protein and R72 of YidC stabilizes the Pf3 coat protein within the groove. When Pf3 is fully within the YidC hydrophilic groove, it contacts lipid tails. Proton motive force and hydrophobic contact with the membrane lipid tails let the Pf3 coat protein migrate to the periplasmic side of the membrane. In this mechanism, YidC undergoes large conformational changes during SecY-independent insertion, such as widening of the TM region. Despite recent advances and widespread docking, there are still several downsides. Model quality and docking accuracy affect simulation outcomes. To properly understand the process, further docking models are needed, including substrate proteins in different structural states. This research will assist experimental and computational scientists in

examining YidC SecY-independent mechanism.

The C2 cytoplasmic loop also has a significant role in YidC in gram-negative bacteria. Based on our analysis, the gram-negative YidC C2 cytoplasmic loop aids in stabilizing the protein structure by having indirect impacts on connections in the transmembrane core region and allosteric effects on other periplasmic domains, particularly on interactions between the PD and TM regions. Additionally, our analysis suggests that the C2 loop's absence affects YidC functional characteristics, such as the groove's hydration. Our study of gram-positive bacterial YidC using molecular dynamics simulations has shown that C2 loops may be required for the structural stability and functionality of gram-negative YidC membrane protein dynamics that are independent of SecY. However, more research is needed to better understand the function of the C2 loop in the insertion of tiny single-spanning membrane proteins, such as the pf3 coat protein, which are considered to interact with the cytoplasmic loops at the beginning of the insertion pathway of YidC in gram-positive bacteria.

In addition, we provide a unique method for generating peptides with multiple functions and different amino acid domains, which may be used in the inexpensive manufacture of nanoparticles. The methodology developed for this work, primarily involving FEP/MD simulations, has the potential to be applied to the efficient development of new nanocatalysts as well as the identification and explanation of peptide regions crucial to nanoparticle formation. The sequence-dependent, various catalytic outcomes in the production of Pd nanoparticles from free peptides have been enhanced by the molecular-level detail from our computational and experimental findings. This work demonstrates how altering the sequence of a single amino acid in the free Pd4 peptide may alter the structure and catalytic characteristics of nanoparticle production. We anticipate that this approach may be used in various peptide nanocatalyst development processes. To fully explore these theories, further experimental and computational research is required. However, the relationship between TOF and binding free energy is not explained by our simulation models. According to our theory, the size and catalytic activity of nanoparticles are influenced by the secondary structure, peptide conformational stability, and histidine binding affinity. However, a more thorough inves-

tigation and a fuller knowledge of nanoparticle synthesis and activity are needed to fully grasp these interactions at the molecular level. This computational approach will help scientists create more potent peptides for nanoparticle synthesis. Our results could aid in the development of more simulation studies that look into the relationship we proposed in our study.

## Appendix

### Supplementary Material for Chapter I: An Investigation of the YidC-Mediated Membrane Insertion of Pf3 Coat Protein Using Molecular Dynamics Simulations

Adithya Polasa[1], Jeevapani Hettige[1], Kalyan Immadisetty[1] and Mahmoud Moradi[1] *

[1]Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701

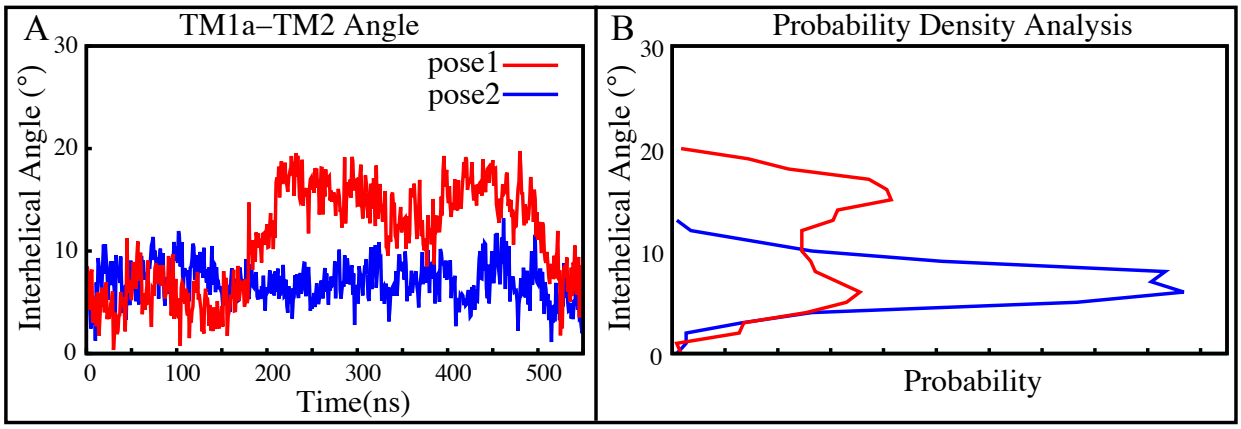


Figure S1: (A) Overall inter-helical angle between transmembrane helix 1a and 2 helices of the protein in pose 1 (red) and pose 2 (blue) simulations. (B) The probability density distribution for the angle between transmembrane helices 1a and 2.

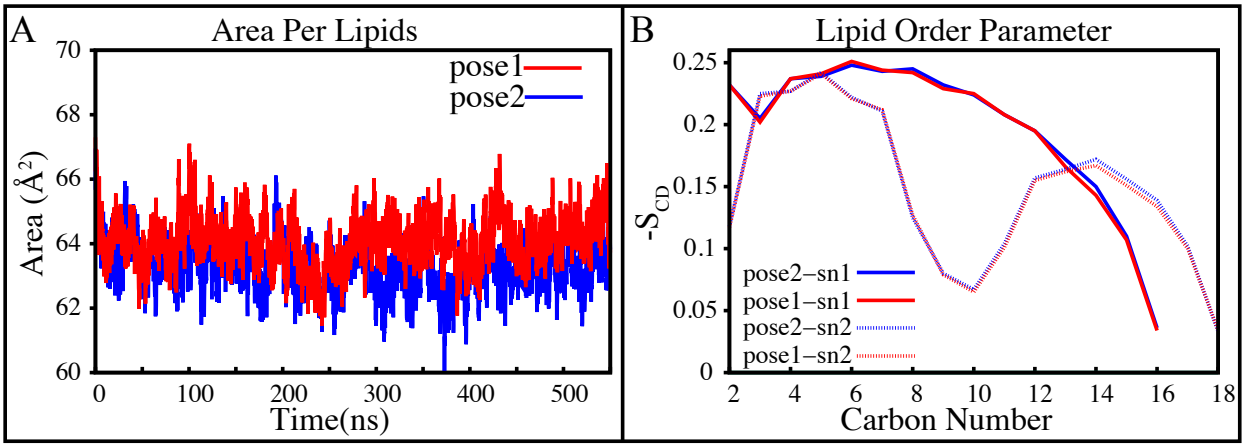


Figure S2: (A) Time-dependent area per lipid values of the equilibrated POPE membrane in pose1 and pose2. (B) Order parameters were calculated for the equilibrated lipids POPE simulations using MEMBPLUGIN analysis tool. Results are presented for (broad line) the sn-1 and (dotted line) sn-2 chains.

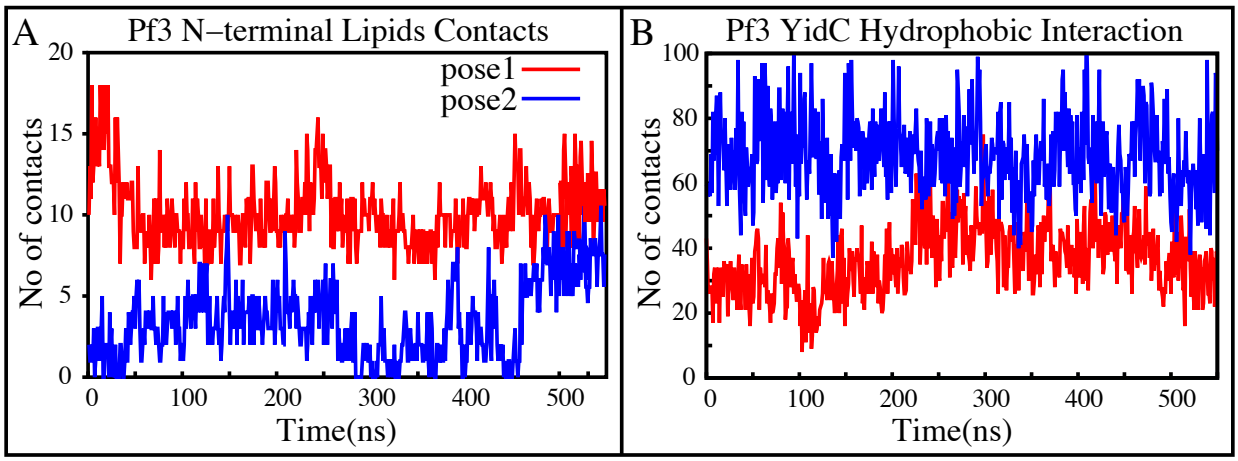


Figure S3: (A) Pf3 coat protein N-terminal end overall interaction with POPE lipid tails in both the docking model simulations. (B) Hydrophobic residue contacts between Pf3 coat protein and YidC protein in both the docking model simulations.

**Supplementary Material for Chapter II: Deciphering the Inter-domain Decoupling in the Gram-negative Bacterial Membrane Insertase**

Adithya Polasa[1], and Mahmoud Moradi[1] *

[1]Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701

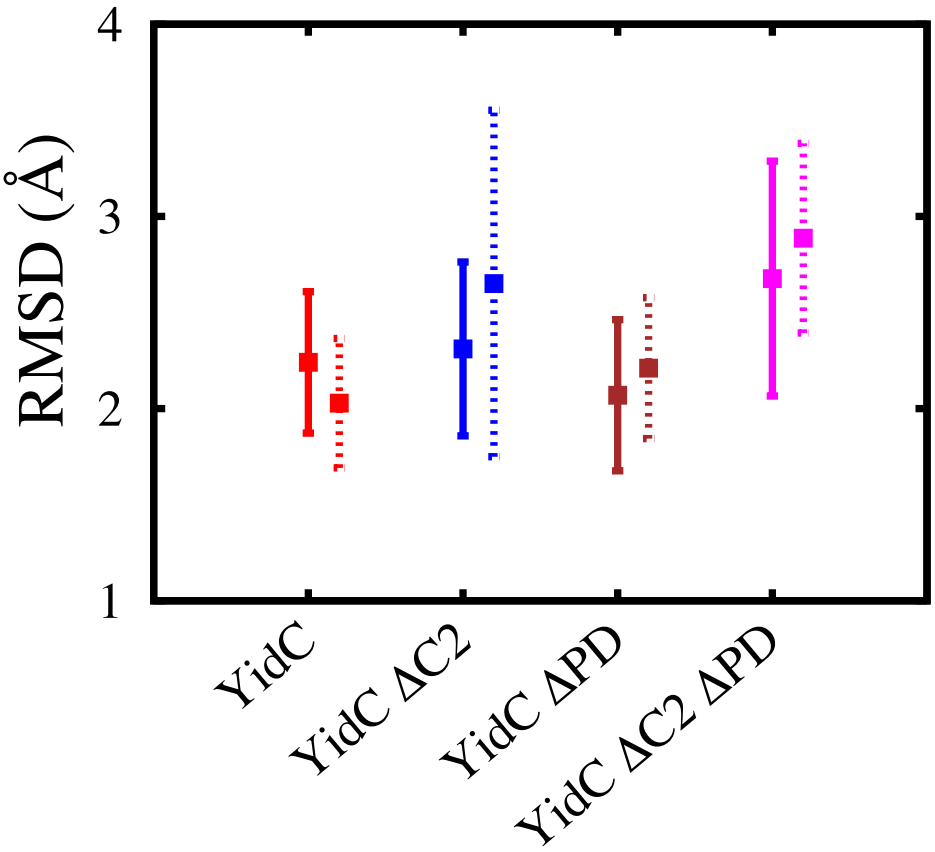


Figure S1: Analysis of the structural stability of YidC with and without the PD and C2 loop. The average and standard deviation of the YidC's root mean square deviation in various systems are shown in this figure. Based on RMSD data, we have shown that YidC is more fluctuating in the system without the C2 loop and PD than in the system with the C2 loop and PD. The dashed lines in the graphs represent the second simulation run for each individual system.

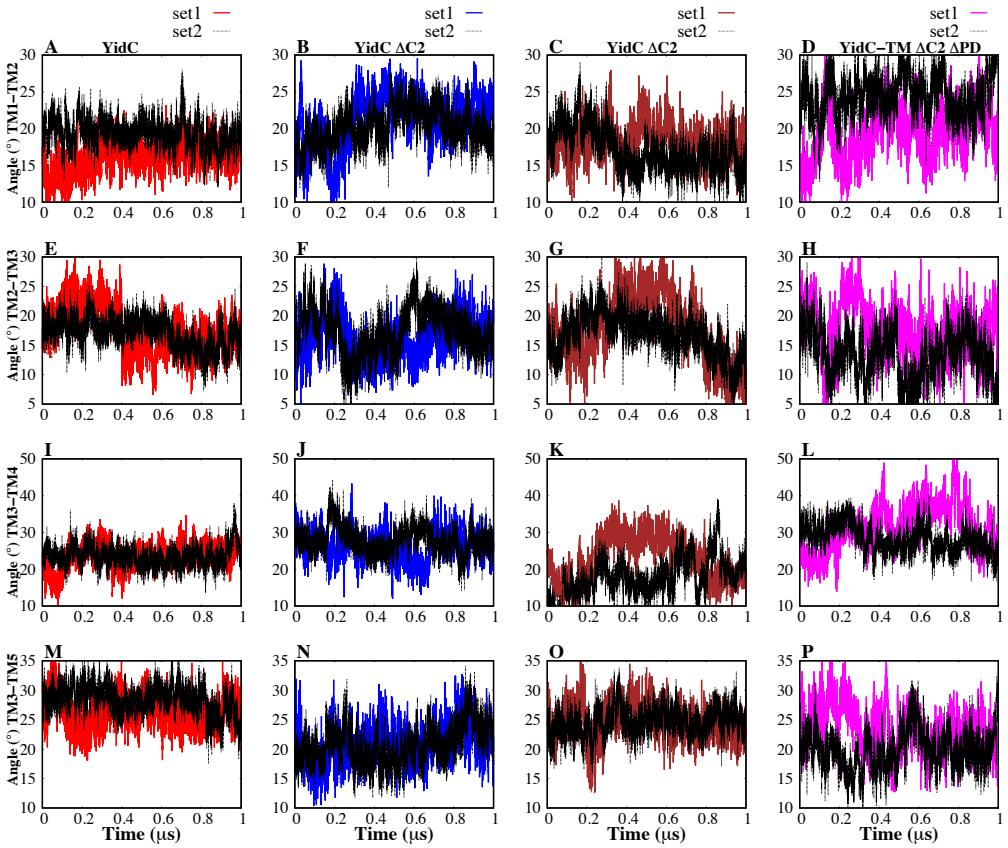


Figure S2: Inter-helical angles between transmembrane helices of YidC. (A–D) The inter-helical angle between the transmembrane helix 1 and helix 2 of the protein. (E–H) The overall inter-helical angle between helix 2 region and helix 3 of the protein. (I–L) The inter-helical angle between the transmembrane helix 3 and helix 4 of the protein. (M–P) The overall inter-helical angle between helix 3 region and helix 5 of the protein.

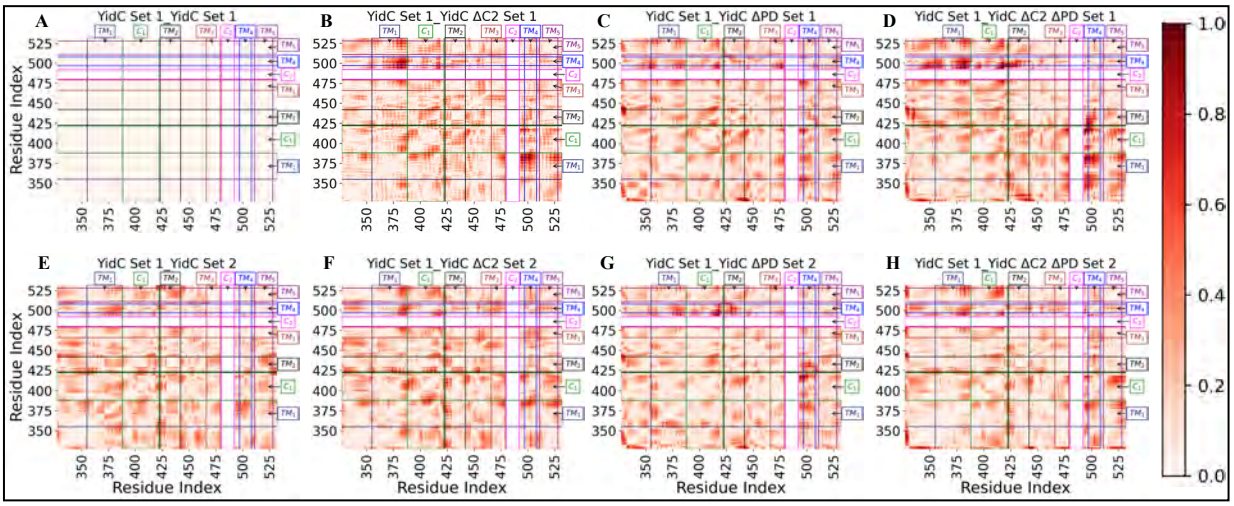


Figure S3: The DNA analysis showed that the YidC set 1 control system and the other YidC systems that were investigated for this research have distinct disparities in their correlations. Although the highest difference in correlation that may be detected in practice was less than one, the theoretical maximum is two. (A-H) Differences in correlation are shown as a gradient of red, with deeper red representing bigger differences.

**Supplementary Material for Chapter III: Developing a Rational Approach to Designing Recombinant Proteins for Peptide-Directed Nanoparticle Synthesis**

Adithya Polasa[1], Imann Mosleh[2,3], James Losey[1], Alireza Abbaspourrad [2],

Robert Beitle[3], and Mahmoud Moradi[1] *

[1]Department of Chemistry and Biochemistry, University of Arkansas,

Fayetteville, AR 72701, USA.

[2]Department of Food Science, College of Agriculture and Life Sciences, Cornell University,

Ithaca, NY, 14853, USA

[3]Ralph E. Martin Department of Chemical Engineering, University of Arkansas,

Fayetteville, AR 72701, USA

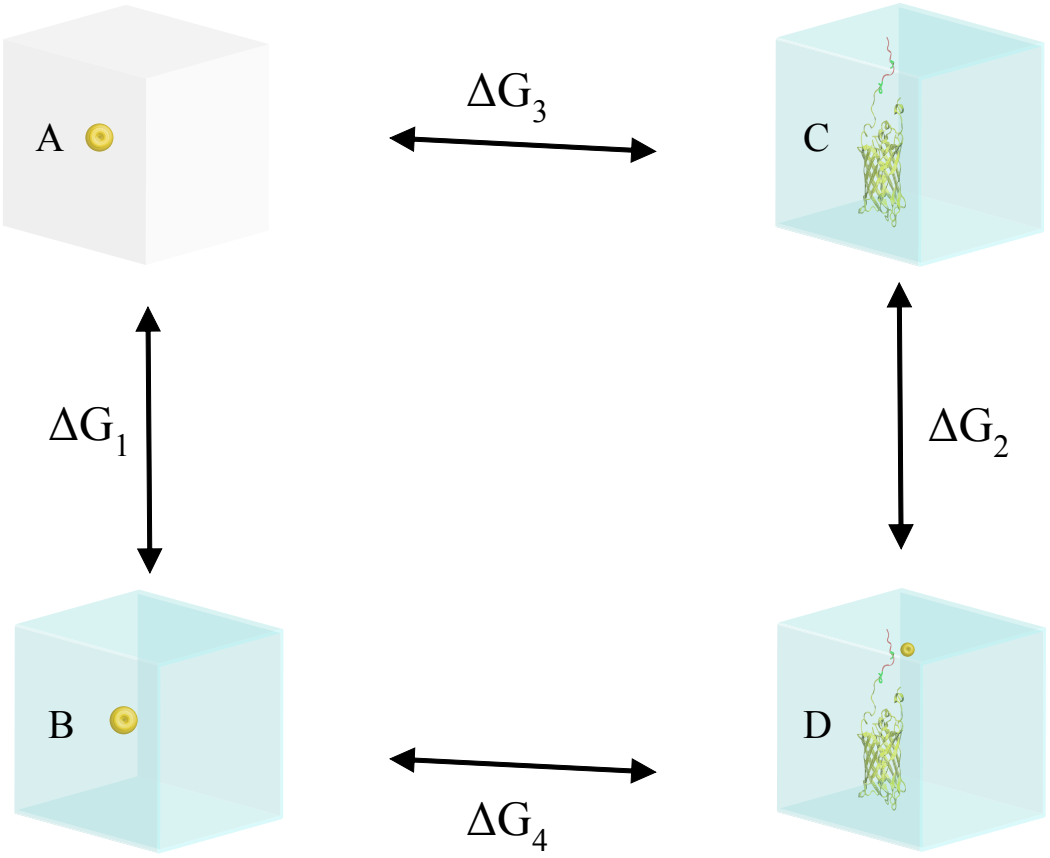


Figure S1: Schematic representation of thermodynamics cycle for calculating the relative binding free energy of palladium binding to the histidine of peptides.(A-B) Solvation free energy of palladium ion in the aqueous solution. (C-D) Relative free energy binding of the palladium with the histidines of the peptides. overall binding free energy is derived from $\Delta G_{binding} = \Delta G_2 - \Delta G_1$

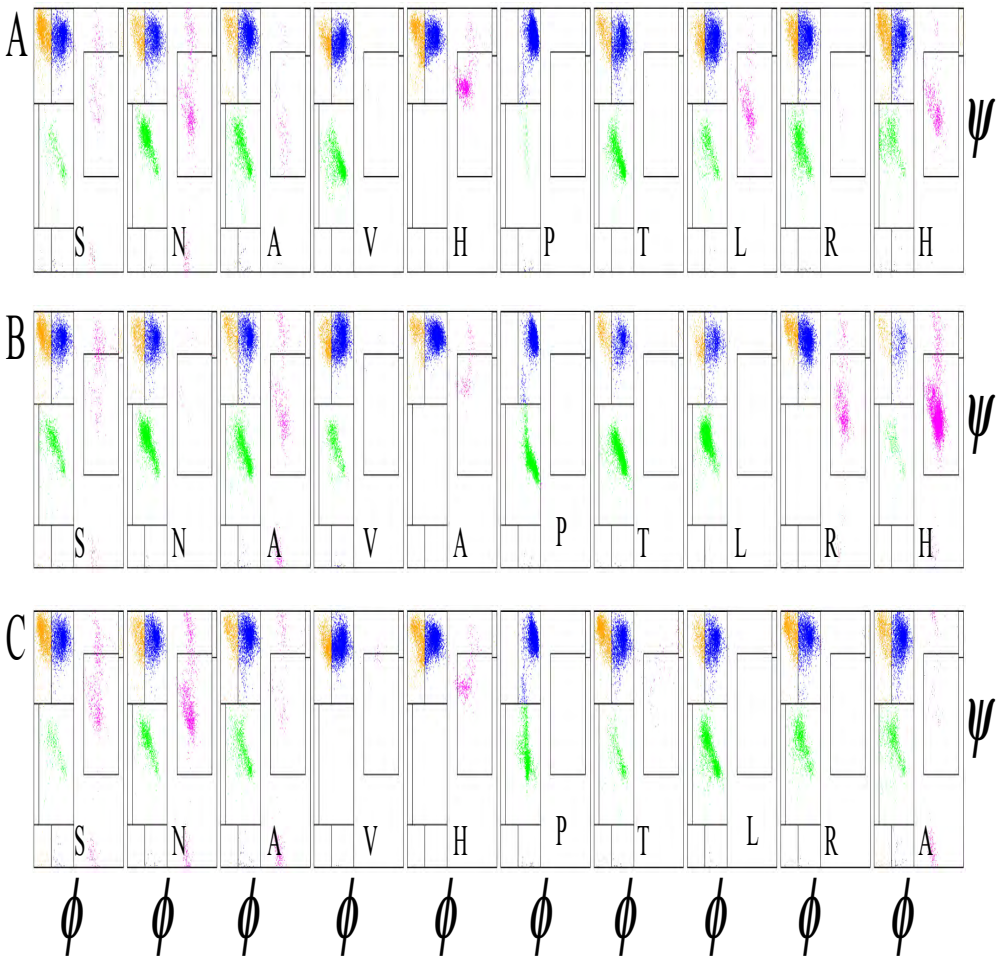


Figure S2: Ramachandran plots of residues in free peptides: (A-C) The backbone structure of the residue in peptides Pd4, A6, and A11 respectively. The region definitions are the same as Figure 1A. Orange, blue, pink, green, and gray clusters identify the $\beta$, F, $\alpha_L$, $\alpha_R$, and N regions, of Ramachandran plot respectively.

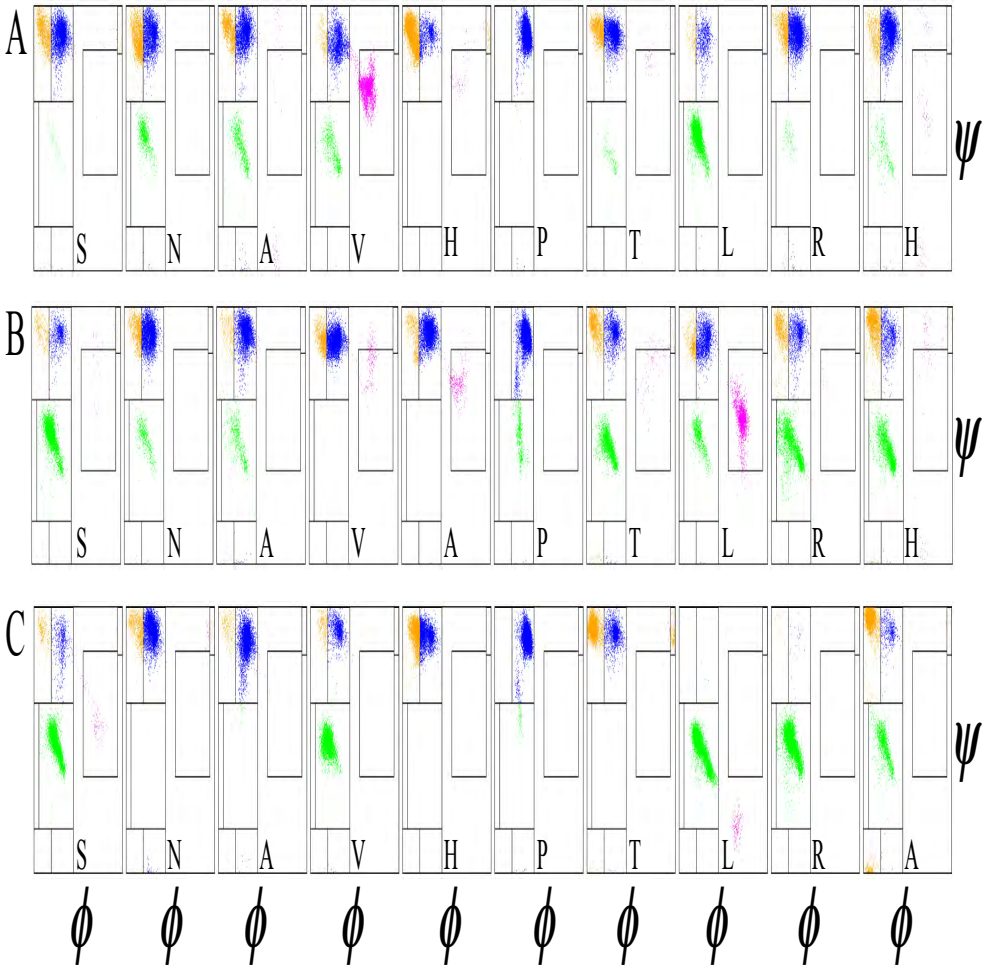


Figure S3: Ramachandran plots of residues in GFPuv fusion peptides: (A-C) The backbone structure of the residue in GFP bound peptides Pd4, A6, and A11 respectively. The region definitions are the same as Figure 1A. Orange, blue, pink, green, and gray clusters identify the $\beta$, F, $\alpha_L$, $\alpha_R$, and N regions, of Ramachandran plot respectively

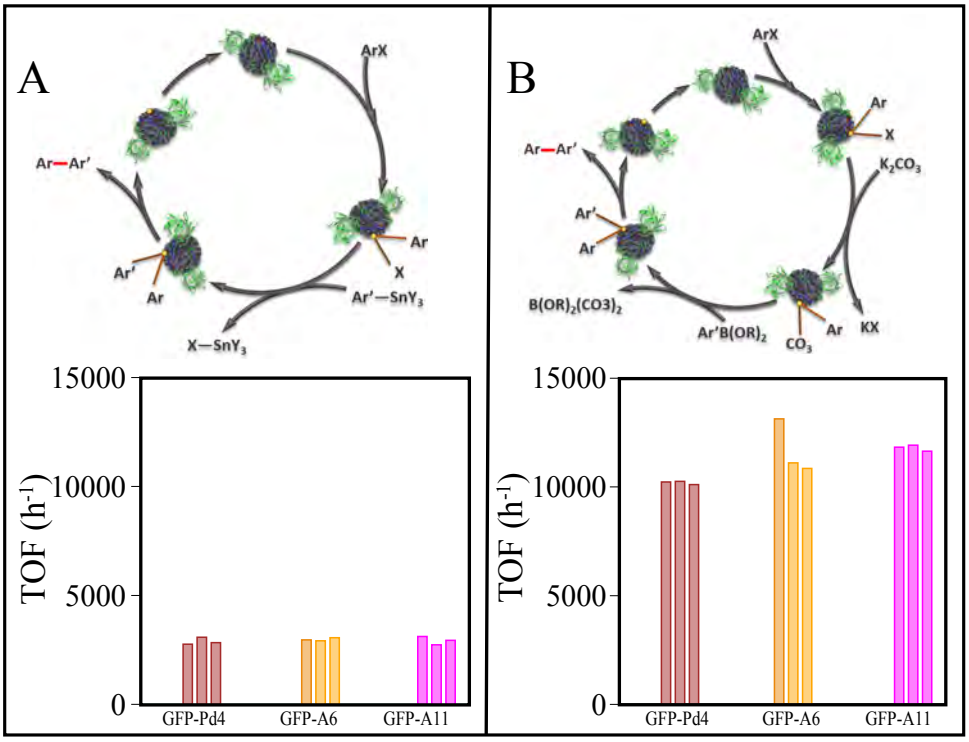


Figure S4: Schematic diagram and TOF of (A) Stille coupling and (B) Suzuki-Miyaura coupling reaction for GFP fused peptides Pd4, A6 and A11 respectively.

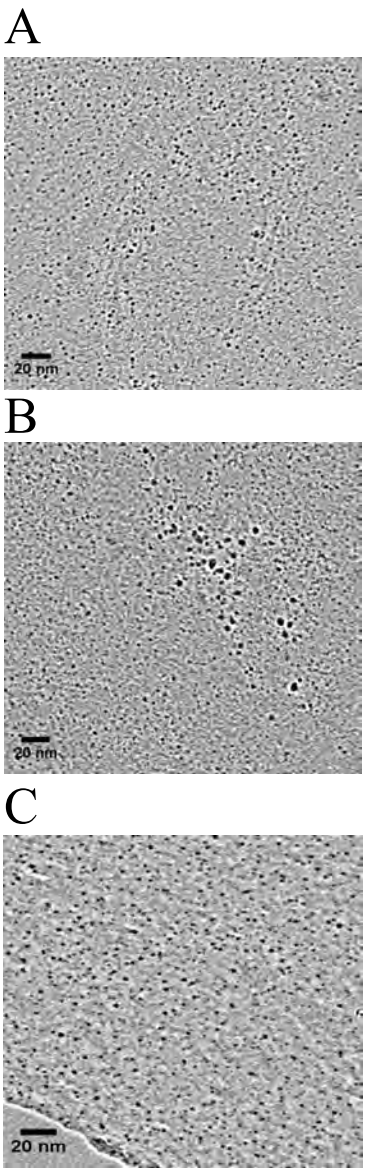


Figure S5: (A-C) TEM images of nanoparticles produced from GFPuv-attached peptides Pd4 (A), A6 (B) & A11 (C).

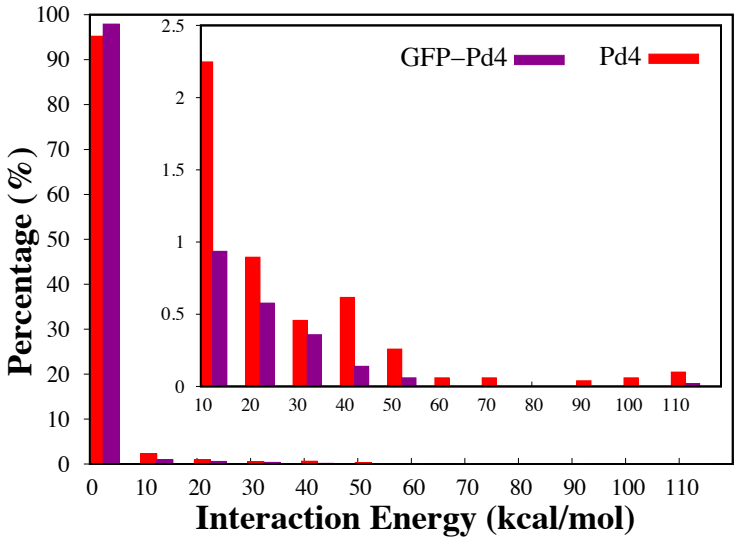


Figure S6: His6-His11 interaction energy distribution in the Pd4 and GFP-Pd4 systems. The normalized histograms are built with a bin size of 10 kcal/mol and based on the magnitude of interaction energies. In the inset, the distribution is shown, but the 0–10 bin is left out.

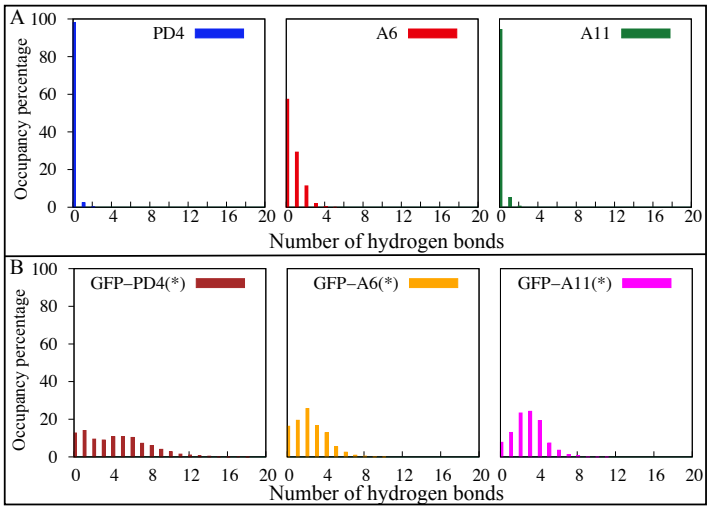


Figure S7: The occupancy percentage of hydrogen bonds calculated from the MD simulations for the (A) free peptides and (B) GFPuv fused peptide region. (*) Only the peptide region in the GFPuv fused peptide simulations was used for hydrogen bond analysis. This hydrogen bond analysis was done using a distance cutoff of 3.5 Å and an angle cutoff of 30°.

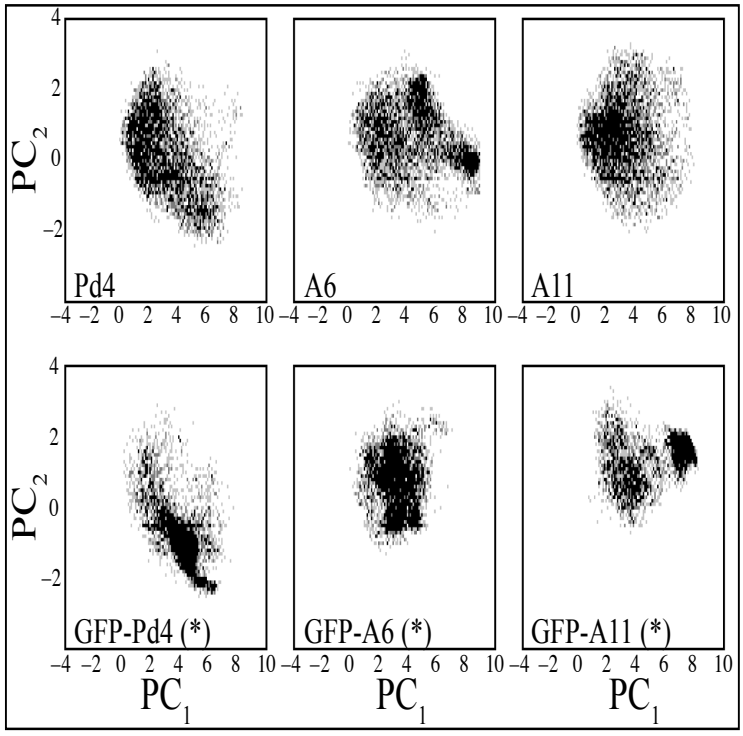


Figure S8: PCA of C-$\alpha$ of peptides in free state and GFP fused simulations. The intensity of the color in the plot represents the relative population of peptide PCA analysis.

Table S1: **Optimized condition via coupling reaction of iodobenzene and phenylboronic acid. Solvent, base, and temperature.**

$$\mathrm{Ar{-}X{+}}Ar' \xrightarrow[Solvent,Time,and Temperature]{Catalyst,Base} Ar-Ar' \tag{2}$$

| Entry | Solvent | Base (mmol) | Time (h) | Temperature (°C) | Yield[b] (%) |
|---|---|---|---|---|---|
| **1** | EtOH: $H_2O$ (1:1) | KOtBu | 3.0 | 80 | 47 |
| **2** | EtOH: $H_2O$ (1:1) | $K_2HPO_4$ | 3.0 | 80 | 65 |
| **3** | EtOH: $H_2O$ (1:1) | $K_2HPO_4$ | 3.0 | 80 | 62 |
| **4** | EtOH: $H_2O$ (1:1) | $K_2CO_3$ | 3.0 | 80 | 98 |
| **5** | EtOH: $H_2O$ (1:1) | $K_2CO_3$ | 8.0 | 80 | 82 |
| **6** | EtOH: $H_2O$ (3:1) | $K_2CO_3$ | 1.5 | 80 | 98 |
| **7** | EtOH: $H_2O$ (1:1) | $K_2CO_3$ | 1.5 | 80 | 97 |
| **8** | EtOH: $H_2O$ (1:3) | $K_2CO_3$ | 3.5 | 80 | 74 |

Reaction conditions:iodobenzene (0.1 mmol), phenylboronic acid (0.12 mmol), base (0.3 mmol) and Pd NPs (0.005 $\mu$mol) were mixed in solvent and refluxed under $N_2$. ([b])Yield were determined by HPLC.

Table S2: **Optimized condition via coupling reaction of iodobenzene and phenyltin trichloride. Solvent, base, and temperature.**

$$\mathrm{Ar{-}X} + Ar - SnCl_3 \xrightarrow[EtOH:H\_2O and Refulx]{Catalyst,Base} Ar-Ar' \tag{3}$$

| Entry | Solvent | Pd (mmol %) | Time (h) | Temperature (°C) | Yield[b] (%) |
|---|---|---|---|---|---|
| **1** | 5 | KOH | 6.0 | 80 | 72 |
| **2** | 5 | $K_2CO_3$ | 6.0 | 80 | 34 |
| **3** | 5 | $K_2HPO_4$ | 6.0 | 80 | 36 |
| **4** | 5 | CsF | 6.0 | 80 | 97 |
| **5** | 5 | $K_3PO_4$ | 48.0 | 40 | 58 |
| **6** | 5 | $K_3PO_4$ | 20.0 | 60 | 70 |
| **7** | 10 | $K_3PO_4$ | 6.0 | 80 | 96 |
| **8** | 2 | $K_3PO_4$ | 16 | 80 | 88 |
| **9** | 1 | $K_3PO_4$ | 16 | 80 | 64 |

Reaction conditions:iodobenzene (0.1 mmol), phenylboronic acid (0.12 mmol), base (0.3 mmol) and Pd NPs (0.005 $\mu$mol) were mixed in solvent and refluxed under $N_2$. ([b])Yield were determined by HPLC.

**List of other published papers**

1. K Immadisetty, **A Polasa**, R Shelton, and M Moradi, Elucidating the Molecular Basis of Spontaneous Activation in an Engineered Mechanosensitive Channel. Computational and Structural Biotechnology Journal.,

2. Govind Kumar, V., Ogden, D. S., Isu, U. H., **Polasa. A**, Losey, J., & Moradi, M. (2022). Pre-fusion Spike Protein Conformational Changes Are Slower in SARS-CoV-2 than in SARS-CoV-1 Journal of Biological Chemistry.,

3. James Losey, **Adithya Polasa**, Michael Jauch, Axel Cortes–Cubero, Haoxuan Wu, Roberto Rivera, David S. Matteson, and Mahmoud Moradi, Simulating Freely–diffusing Single–molecule smFRET Data with Consideration of Protein Conformational Dynamics. Biophysical Journal.,

4. Vivek Govind Kumar, **Adithya Polasa**, Shilpi Agrawal, Thallapuranam Krishnaswamy Suresh Kumar, and Mahmoud Moradi. Binding Affinity Estimation From Restrained Umbrella Sampling Simulations, Nature Computational Science.,