Theses and Dissertations

2014

Understanding the Structural Basis for Transmembrane Signaling in the CLR-RAMP1 Heterodimer

Pin-Chuan Su
Lehigh University

Understanding the Structural Basis for Transmembrane Signaling in the CLR-RAMP1 Heterodimer

By

Pin-Chuan (Daniel) Su

A Dissertation
Presented to the Graduate and Research Committee
of Lehigh University
in Candidacy for the Degree of
Doctor of Philosophy

in
Chemical Engineering

Lehigh University
May 2014

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemical Engineering.

"Understanding the Structural Basis for Transmembrane Signaling in the CLR-RAMP1 Heterodimer"

Pin-Chuan (Daniel) Su

_____

April 15th, 2014

_____

Accepted Date

_____

Bryan W. Berger, Ph.D., Dissertation Advisor

_____

Anand Jagota, Ph.D., Committee Member

_____

Mark A. Snyder, Ph.D., Committee Member

_____

Kerney Jebrell Glover, Ph.D., Committee Member

iii

Copyright Permissions

The material contained in Chapters 2, 3, and 4 has been published prior to submittal of this dissertation. The proper copyright permission was obtained from each journal prior to submittal of this document, with the publication cited at the beginning of each chapter.

Acknowledgments

I wish to thank many people for their support throughout my Ph.D. studies. First and foremost, I would like to express my most sincere gratitude to my advisor, Dr. Bryan W. Berger, for his patience and guidance in my Ph.D. studies as well as personal life. Being the first Ph.D. student for a new faculty is not easy but it is certainly harder to train the very first Ph.D. student of his own. Dr. Berger has done an excellent job in training me how to think, tackle problems, design and execute experiments, and write as a professional and independent scientist. Although there were some disagreement and argument along the way (mostly due to my stubbornness, and I am very thankful for Dr. Berger's patient and forgiveness), after everything, it was a memorable experience and I accomplished and learned a lot.

I would also like to thank my committee members Dr. Anand Jagota, Dr. Mark A. Snyder, and Dr. Kerney Jebrell Glover for their great support and valuable guidance in my Ph.D. work. I am grateful to the knowledge I learned from the Chemical Engineering faculty. Also, special thanks to the department staffs Tracey Lopez, Janine Jekels, Barbara Kessler, Elaine Correll, John Caffrey, and Lee Graham for keeping my time here smooth and on top of things regarding paperwork and orders all the time.
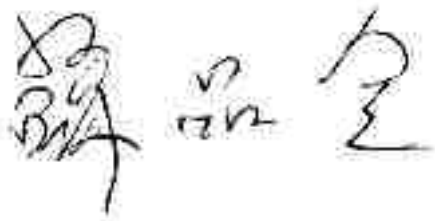
I would like to thank my colleagues Logan C. MacDonald, Sara MacDonald, Rachael E. Barton, Zhou Yang, and Benjamaporn Wonganu for their good company and making my days at Lehigh full of surprises and joyfulness and all the holidays feel like being at home. After all, I have seen you guys more than anyone else for the past 4 years. I would also like to thank the undergraduates who worked in the Berger lab for bringing in the energy and young hearts as well as homemade desserts (William Si,

Victoria Berard, Diego Liriano Jr., Danica Palacio, Alyssa Driscoll, Lindsey Stretz, Samuel E. Flores, Kelly Callahan, Mei-Fan Parnes, and Hanlon M. MacDonald). Also, I would like to thank all the graduate students who I have collaborated with.

My Ph.D. study is a mixture of sweet and bitter memories. I would like to thank all my friends for making my time at Lehigh meaningful and enjoyable. I would also like to thank the people who taught me becoming a more mature and discipline individual. Lastly, I would like to express my greatest appreciation for the endless support from my dad, my mom, and my sisters to carry me through all the tear and discouragement during my Ph.D. journey.

Lastly, I would like to conclude by sharing a quote from my all time favorite movie: "**the night is darkest just before the dawn**". Cherish every moment and significant ones in your life. Stay faithful because everything that meant to be will find a way. Cheers to all!

Pin-Chuan (Daniel) Su

Table of Contents

List of Figures

MBP antibody indicates proper expression of homo- and heterodimeric chimera, with the intensity of the bands for co-expression of both AraC- and AraC*-containing chimera twice that of the corresponding bands for single chimera expression. C, Chimera containing AraC* fused to wild-type and L980A integrin α_{IIb} TM-CYTO were expressed in MalE-deficient MM39 cells, streaked on a 0.4% maltose M9 plate and incubated for 40 h at 37°C. Each construct was properly integrated into the inner membrane of *E. coli*, as indicated by robust growth on the 0.4% maltose M9 plate similar to the positive control (pAraTM containing integrin αIIb TM-CYTO chimera fused to AraC and MBP-HA) As expected, no growth is observed on the negative control (AraCY, which expresses MBP without the N-terminal signal peptide sequence fused to AraC protein).

Abstract

  Membrane proteins comprise of 50% of all pharmaceutical targets in the human genome. These proteins reside in equilibrium between resting and active states that are responsible for various intracellular signaling. We are interested in studying a type of membrane protein belonging to the G-protein coupled receptor (GPCR) family named calcitonin receptor-like receptor (CLR). When associated with a single pass membrane protein called receptor activity modifying protein 1 (RAMP1), the CLR-RAMP1 complex forms a specific receptor for calcitonin gene-related peptide (CGRP). Dysregulation of this receptor complex is linked to various disease states including migraine and acute coronary syndrome (ACS). Small molecule antagonists targeting CLR-RAMP1 complex have been shown to be effective in treating acute migraines, however, they are only limited to intravenous delivery. Therefore, it is very important to obtain a detail understanding on how these receptors interact to form the functional receptor for superior drug design.

  Membrane proteins are extremely difficult to study due to their water insoluble nature. To overcome this limitation, we developed several tools to assist our research. One important mechanism that regulates membrane protein activation is receptor oligomerization. In order to gain an understanding and identify residues that govern receptor oligomerization in a straightforward and high throughput fashion, we developed *E. coli* transcription factor AraC-based methods named AraTM and DN-AraTM which look at receptor homo- and heterodimerization respectively. By using AraTM assay, we were able to identify a specific juxtamembrane region within the cytosolic domain (A375-P394) of receptor for advanced glycation endproducts (RAGE)

1

which mediates its homodimerization. Moreover, we also developed a *T7*-based expression vector (pOmpF) using an engineered fragment of outer membrane protein F (OmpF) as the fusion protein to direct full-length membrane protein overexpression in *E. coli* for high-resolution structure determination. Utilizing pOmpF vector, we successfully purified thermally stable RAMP1 protein in detergent Fos-choline 15 (FC15). By using circular dichroism, dynamic light scattering, and tryptophan fluorescence spectroscopy, we were able to show that the purified RAMP1 protein has a native tertiary structure consisting of disulfide bonds with 90% helical content.

Finally, using sequence-directed searches of transmembrane structural databases, we identified a P-x-x-x-T motif interface that is conserved in RAMP1 among different species as well as between different human RAMPs. By applying *cAMP* signaling assay, *in vivo* bioluminescent resonance energy transfer assay, and zebrafish RAMP1 phenotypic knockdown and rescue experiments, we were able to show that this predicted P-x-x-x-T motif plays a critical role in CLR-RAMP1 association and function. Altogether, this work not only provides innovative tools to improve membrane protein research but also sheds light on understanding the structural basis for CLR-RAMP1 receptor signaling.

Chapter 1
INTRODUCTION

The cells of virtually every living organism are surrounded by a phospholipid bilayer membrane, acting as a continuous hydrophobic barrier which retains various proteins, ions, and other biomolecules inside the cell and prevents unwanted foreign substances from entering the cell. However cells must be able to communicate and receive instructions from neighboring cells and the surrounding environment in order to function properly. Communication is achieved by membrane receptor proteins which span the cell membrane, recognize signals from the outside, and trigger intracellular changes through signal transduction. This process plays critical roles in various aspects of cellular activity such as wound healing, cell development, and generating an immune response. Membrane proteins comprise 30% of all proteins in the human genome and 50% of all pharmaceutical targets, however less than 0.3% of protein structures in the Protein Data Bank (PDB; www.rcsb.org) are membrane proteins (1). This is largely due to the inability to obtain sufficient quantities of membrane protein sample because of their insolubility in aqueous solutions and low abundance in their native membrane environment, therefore complicating efforts aimed at developing new pharmaceuticals (1-4).

In terms of pharmacology, G-protein coupled receptors (GPCRs) are an incredibly important type of membrane protein because they composed over 40% of clinically approved drug targets currently (5). GPCRs are a superfamily of seven transmembrane spanning proteins that are associated with a variety of intracellular signaling including hormones, neurotransmitters, light, odor, and taste via interaction

with guanine nucleotide binding proteins (G-proteins) (6). As shown in Figure 1.1, binding of GPCR ligands results in conformational changes in GPCRs that facilitate the exchange of guanosine diphosphate (GDP) to guanosine triphosphate (GTP) on the heterotrimeric G protein Gα subunit which leads to the dissociation of Gα from the Gβ and Gγ subunit for subsequence signaling (7, 8). Based on Gene Ontology and Celera's Panther Classification, GPCR can be categorized into three main classes: Rhodopsin-like (class A), Secretin-like (class B), and Glutamate-like (class C), where Rhodopsin-like class accounts for about 85% of the known GPCRs genomes (9). Although a number of active and inactive crystal structures of class A GPCR are available, they were limited usage for studying class B GPCR due to the lack of sequence conservation between these two classes (10). Up to date, there is no solved full-length class B GPCR crystal structure; therefore, there remains a lot of interest and opportunities in understanding how class B GPCRs function and signal.
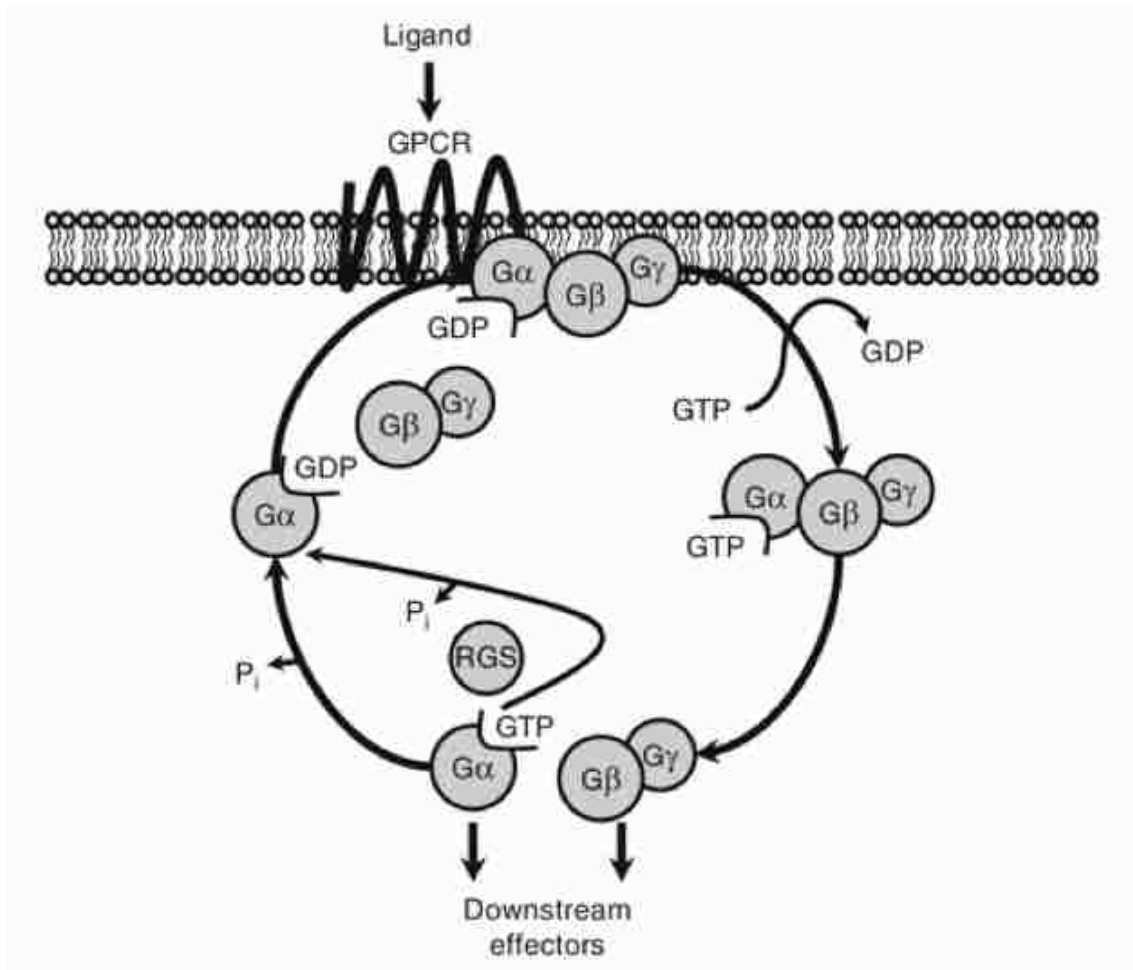
Figure 1. 1. The G protein cycle. Ligand binding to the G protein-coupled receptor (GPCR) leads to exchange of GDP on the Gα subunit and dissociation of the Gα and Gβγ heterodimer. Both Gα-GTP and Gβγ may regulate downstream effectors. GTP hydrolysis by the Gα subunit results in reassociation of GDP-bound Gα with Gβγ heterodimer and the GPCR, thus completing the cycle. RGS (regulator of G protein signaling) proteins accelerate the rate of GTP bydrolysis by the Gα subunit (7).

Calcitonin receptor-like receptor (CLR) is a class B GPCR with a 22 amino acid N-terminus signal peptide, a 124 amino acid extracellular domain with three N-linked glycosylation sites and three pairs of disulfide bonds, seven α-helical transmembrane domains, and a 73 amino acid cytosolic tail (11-13). CLR is linked with diverse disease conditions such as osteoporosis, diabetes, obesity, venous insufficiency, and

5

cardiovascular disorder (14). As shown in Figure 1.2, CLR alone does not recognize any ligand, thus the ligand-specificity of this receptor is dependent on its interaction with a group of accessory membrane proteins named receptor activity modifying proteins (RAMPs) (15-18). There are three types of RAMPs (RAMP1, RAMP2 and RAMP3), which share less than 30% sequence identity (18). The RAMPs are 148-175 amino acids in size with a cleavable N-terminal signaling peptide and similar topology: a large N-terminal extracellular domain (~100aa), a single transmembrane spanning domain (~20aa) and a short C-terminal intracellular domain (~10aa) (19). Associating with RAMP1 or RAMP2/3, CLR forms a calcitonin gene-related peptide (CGRP) receptor or an adrenomedullin (AM) receptor respectively (18). CGRP is one of the most potent vasodilators discovered thus far (20). Increased cranial circulation of CGRP has been associated with migraines therefore antagonism of CGRP receptors has become an important target for migraine treatment (21). Nevertheless, current antimigraine treatments such as sumatriptan cause vasoconstriction and other CLR antagonist chemical-based drugs developed by several pharmaceutical companies either exhibit poor oral bioavailability (olcegepant) or cause liver damage (telcagepant) (21, 22). One major challenge to design small molecule antagonist against class B GPCRs is the endogenous peptide-based ligand nature of these receptors, which involves competing a larger molecule which has more point of interaction with the receptor with a smaller molecule (23). Therefore, in order to design more effective migraine medications which possess high receptor selectivity, it is critical to understand how CLR and RAMP1 interact to form the ligand-binding pocket and how this receptor mediates signal transduction.

6
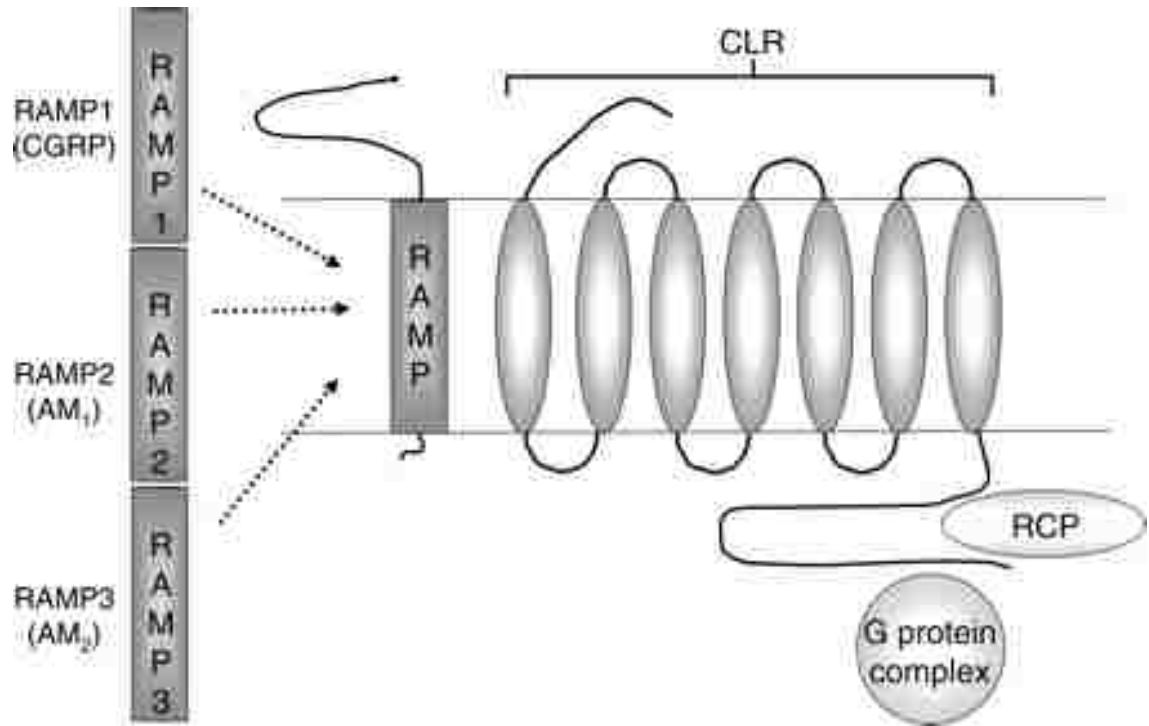
Figure 1. 2. The CGRP/adrenomedullin (AM) receptor model. The calcitonin receptor-like receptor (CLR) component is common to all three receptors and is a G protein-coupled 7-transmembrane receptor. The three RAMP components are single transmembrane domain proteins. The active receptor is a functional heterodimer of one CLR complexed with a RAMP, at the cell membrane. The interaction of RAMP1 with CL produces a CGRP receptor, RAMP2 with CLR an AM receptor, and RAMP3 with CLR a CGRP/AM receptor. The proposed receptor component protein (RCP), which is suggested to allow coupling to intracellular signaling pathways, is also included (15).

Efforts have been made to investigate the role different RAMP1 domains play in the formation of a functional CGRP receptor. In summary, RAMP1 extracellular domain alone can N-glycosylate CLR, is able to form a functional receptor unit with CLR and traffic the receptor to plasma membrane; however, the half maximum effective concentration is increased about 4000 fold compared to full length RAMP1 (24). This indicates that the remaining domains of RAMP1 play a significant role in CGRP binding and signaling. Steiner et al. showed that by deleting the short intracellular tail of RAMP1 had no effect in signaling (25). Moreover, in 2010, a high-

resolution structure of RAMP1-CLR extracellular domain in complex with bound chemical based antagonist was solved by ter Haar et al. (12). They demonstrated this complex was able to compete with the wild-type receptor binding to CGRP using radioligand displacement binding assay, but with a much lower binding affinity (12). This means that this structure does not fully recapitulated the binding of CGRP to CLR-RAMP1 complex. They also crystallized the soluble portion of the receptor in complex with antagonist olcegepant or telcagepant, but the size of the antagonist is much smaller than the size of the CGRP, so it does not provide a full picture for CGRP-receptor binding. All of the above results point toward the specific interaction of CLR-RAMP1 being governed by their TM domains. Therefore, in order to understand how CLR and RAMP1 assemble to form a functional receptor for CGRP, we need to study them from a perspective that is in context with their TM domains.

One means of understanding how external signals are transduced by membrane receptors is through the analysis of receptor homo- and heterodimerization. Receptor homo- and heterodimerization is a key mechanism that regulates transmembrane signal transduction through a coordinated process that involves ligand binding in the extracellular (EX) domain as well as conformational changes in the receptor transmembrane (TM) and cytoplasmic (CYTO) domains (26). For GPCRs, it had been proposed and proven in class A GPCRs that GPCR dimerization can play a role in receptor maturation, become a target for ligand induced dynamic regulation, lead to both positive and negative ligand binding cooperativity, change G-protein selectivity, and promote co-internalization of two receptors (27). Several classes of type I transmembrane receptors, including neuropilins and integrins, exist in a regulated

equilibrium between homo- and heterodimeric states which defines the inactive or activated forms of the co-receptor complex (28, 29). Within the TM domain of integrin α_{IIb} or β_3, mutations that promote preferential TM homodimer formation or destabilize the heterodimeric interface can result in receptor activation (28). Thus, in order to understand the basis for transmembrane signal transduction, it is essential to characterize the interplay between specific homo- and heterodimeric interactions involving TM and juxtamembrane (CYTO, EX) domains.

Numerous biological tools have been developed to study membrane receptor interactions and understand the basis for transmembrane signaling, including polyacrylamide gel-shift assay, analytical ultracentrifugation, and Forster resonance energy transfer (FRET). Among these methods, transcription factor-based selection systems have the advantage of straightforward implementation, their ability to be performed *in vivo*, high sensitivity, and rapid access to identify importance interfaces and domains within membrane receptors in defining homo- and heterodimeric states (30, 31). However, popular transcription factor-based methods such as TOXCAT and GALLEX rely on transcription factors (ToxR for TOXCAT, LexA for GALLEX) with type II orientation, in which the N-terminal DNA-binding domain is expressed as a fusion to a C-terminal TM domain of interest (30, 31). Thus, the C-terminal TM domain acts as a signal peptide to direct membrane trafficking and integration, and any modification of the TM domain, including addition of CYTO and EX domains, can prevent proper membrane integration. As a result, bacterial assays TOXCAT and GALLEX are limited to analysis of only isolated TM domains and these TM domain interactions were studied in a reverse orientation relative to their native states.

9

Another aspect of understanding how membrane receptors signal is through high-resolution protein structure determination by X-ray crystallography. X-ray crystallography is a technique in which beams of X-ray are bombarded at a protein crystal and the scattering pattern of the X-ray are collected and analyzed to generate a 3D protein structure (32). The most difficult part of X-ray crystallography is the formation of protein crystal. Due to their irregularly shaped surfaces and heterogeneity, it requires high purity (>97%) and large amounts of protein sample to empirically determine the ideal crystallization condition for different proteins (33, 34). However, obtaining high yields of homogeneous, soluble membrane protein is a major challenge due to their low abundance in their native membrane environment (3, 4, 35). Numerous approaches have been developed to improve the yield and recovery of integral membrane proteins. Among all the strategies to obtain high yield membrane proteins, *E. coli*-based strategies utilizing protein fusions are the most popular method for large-scale, recombinant membrane protein expression (36, 37). *E. coli* has several advantages for membrane protein overexpression: defined growth media is relatively inexpensive, numerous genetic tools are available for straightforward cell and target protein manipulation, and several expression conditions are established for large-scale synthesis (38, 39). However, in all cases, no single fusion protein was identified that gave consistently high expression independent of the particular target membrane protein. Furthermore, the effectiveness of many fusion protein vectors is only limited to expression of fragments of, rather than full-length membrane proteins.

To address these limitations of investigating how membrane receptors signal, we developed new approaches based on the *E. coli* AraC transcription factor to investigate

10

EX-TM-CYTO receptor domain homodimerization (Chapter 2) and heterodimeriztion (Chapter 3). We also developed a *T7*-based expression vector (pOmpF) using an engineered fragment of outer membrane protein F (OmpF) as the fusion protein to direct full-length membrane protein overexpression in *E. coli* for high-resolution structure determination (Chapter 4). Utilizing these novel tools and existing membrane receptor studying tools such as bioluminescence resonance energy transfer (BRET) and circular dichroism, we identified a specific sequence motif in RAMP1 TM domain that plays a significant role in CLR-RAMP1 association and signaling (Chapter 5). Finally, potential research directions are proposed in Chapter 6.

1.1   References

(1) Korepanova, A., Gao, F. P., Hua, Y., Qin, H., Nakamoto, R. K., and Cross, T. A. (2005) Cloning and expression of multiple integral membrane proteins from Mycobacterium tuberculosis in Escherichia coli. Protein Sci 14, 148-158.

(2) Flock, T., Venkatakrishnan, A., Vinothkumar, K., and Babu, M. M. (2012) Deciphering membrane protein structures from protein sequences. Genome Biol 13, 160.

(3) Seddon, A. M., Curnow, P., and Booth, P. J. (2004) Membrane proteins, lipids and detergents: not just a soap opera. Biochim Biophys Acta 1666, 105-117.

(4) Loll, P. J. (2003) Membrane protein structural biology: the high throughput challenge. Journal of Structural Biology 142, 144-153.

(5) Wise, A., Gearing, K., and Rees, S. (2002) Target validation of G-protein coupled receptors. Drug Discov Today 7, 235-246.

(6) Ferguson, S. S. (2001) Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling. Pharmacol Rev 53, 1-24.

(7) Servin, J. A., Campbell, A. J., and Borkovich, K. A. (2012). G protein signaling components in filamentous fungal genomes. In *Biocommunication of Fungi* (Witzany, G., ed.), pp. 21-38. Springer Netherlands.

(8) Neer, E. J. (1995) Heterotrimeric G proteins: organizers of transmembrane signals. Cell 80, 249-257.

(9) Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L.,

12

Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M. L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y. H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N. N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J. F., Guigo, R., Campbell, M. J., Sjolander, K. V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y. H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M.

Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., and Zhu, X. (2001) The sequence of the human genome. Science 291, 1304-1351.

(10) Katritch, V., Cherezov, V., and Stevens, R. C. (2012) Diversity and modularity of G protein-coupled receptor structures. Trends Pharmacol Sci 33, 17-27.

(11) Kusano, S., Kukimoto-Niino, M., Hino, N., Ohsawa, N., Okuda, K., Sakamoto, K., Shirouzu, M., Shindo, T., and Yokoyama, S. (2012) Structural basis for extracellular interactions between calcitonin receptor-like receptor and receptor activity-modifying protein 2 for adrenomedullin-specific binding. Protein Sci 21, 199-210.

(12) ter Haar, E., Koth, C. M., Abdul-Manan, N., Swenson, L., Coll, J. T., Lippke, J. A., Lepre, C. A., Garcia-Guzman, M., and Moore, J. M. (2010) Crystal structure of the ectodomain complex of the CGRP receptor, a class-B GPCR, reveals the site of drug antagonism. Structure 18, 1083-1093.

(13) Zhang, Z., and Henzel, W. J. (2004) Signal peptide prediction based on analysis of experimentally verified cleavage sites. Protein Sci 13, 2819-2824.

(14) Barwell, J., Gingell, J. J., Watkins, H. A., Archbold, J. K., Poyner, D. R., and Hay, D. L. (2012) Calcitonin and calcitonin receptor-like receptors: common themes with family B GPCRs?. Br J Pharmacol 166, 51-65.

(15) Brain, S. D. and Grant, A. D. (2004). Vascular actions of calcitonin gene-related peptide and adrenomedullin. *Physiol Rev* 84, 903-34.

(16) Wunder, F., Rebmann, A., Geerts, A., and Kalthof, B. (2008) Pharmacological and kinetic characterization of adrenomedullin 1 and calcitonin gene-related peptide 1 receptor reporter cell lines. Mol Pharmacol 73, 1235-1243.

(17) Roh, J., Chang, C. L., Bhalla, A., Klein, C., and Hsu, S. Y. (2004) Intermedin is a

calcitonin/calcitonin gene-related peptide family peptide acting through the calcitonin receptor-like receptor/receptor activity-modifying protein receptor complexes. J Biol Chem 279, 7264-7274.

(18) McLatchie, L. M., Fraser, N. J., Main, M. J., Wise, A., Brown, J., Thompson, N., Solari, R., Lee, M. G., and Foord, S. M. (1998) RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor. Nature 393, 333-339.

(19) Hilairet, S., Belanger, C., Bertrand, J., Laperriere, A., Foord, S. M., and Bouvier, M. (2001) Agonist-promoted internalization of a ternary complex between calcitonin receptor-like receptor, receptor activity-modifying protein 1 (RAMP1), and beta-arrestin. J Biol Chem 276, 42182-42190.

(20) Brain, S. D., Williams, T. J., Tippins, J. R., Morris, H. R., and MacIntyre, I. (1985) Calcitonin gene-related peptide is a potent vasodilator. Nature 313, 54-56.

(21) Ho, T. W., Ferrari, M. D., Dodick, D. W., Galet, V., Kost, J., Fan, X., Leibensperger, H., Froman, S., Assaid, C., Lines, C., Koppen, H., and Winner, P. K. (2008) Efficacy and tolerability of MK-0974 (telcagepant), a new oral antagonist of calcitonin gene-related peptide receptor, compared with zolmitriptan for acute migraine: a randomised, placebo-controlled, parallel-treatment trial. Lancet 372, 2115-2123.

(22) Iovino, M., Feifel, U., Yong, C. L., Wolters, J. M., and Wallenstein, G. (2004) Safety, tolerability and pharmacokinetics of BIBN 4096 BS, the first selective small molecule calcitonin gene-related peptide receptor antagonist, following single intravenous administration in healthy volunteers. Cephalalgia 24, 645-656.

(23) Moore, E. L., and Salvatore, C. A. (2012) Targeting a family B GPCR/RAMP receptor complex: CGRP receptor antagonists and migraine, Br J Pharmacol 166, 66-78.

15

(24) Fitzsimmons, T. J., Zhao, X., and Wank, S. A. (2003) The extracellular domain of receptor activity-modifying protein 1 is sufficient for calcitonin receptor-like receptor function. J Biol Chem 278, 14313-14320.

(25) Steiner, S., Muff, R., Gujer, R., Fischer, J. A., and Born, W. (2002) The transmembrane domain of receptor-activity-modifying protein 1 is essential for the functional expression of a calcitonin gene-related peptide receptor. Biochemistry 41, 11398-11404.

(26) Klemm, J. D., Schreiber, S. L., and Crabtree, G. R. (1998) Dimerization as a regulatory mechanism in signal transduction. Annu Rev Immunol 16, 569-592.

(27) Terrillon, S., and Bouvier, M. (2004) Roles of G-protein-coupled receptor dimerization. EMBO Rep 5, 30-34.

(28) Li, W., Metcalf, D. G., Gorelik, R., Li, R., Mitra, N., Nanda, V., Law, P. B., Lear, J. D., Degrado, W. F., and Bennett, J. S. (2005) A push-pull mechanism for regulating integrin function. Proc Natl Acad Sci U S A 102, 1424-1429.

(29) Roth, L., Nasarre, C., Dirrig-Grosch, S., Aunis, D., Cremel, G., Hubert, P., and Bagnard, D. (2008) Transmembrane domain interactions control biological functions of neuropilin-1. Mol Biol Cell 19, 646-654.

(30) Berger, B. W., Kulp, D. W., Span, L. M., DeGrado, J. L., Billings, P. C., Senes, A., Bennett, J. S., and DeGrado, W. F. (2010) Consensus motif for integrin transmembrane helix association. Proc Natl Acad Sci U S A 107, 703-708.

(31) Schneider, D., and Engelman, D. M. (2003) GALLEX, a measurement of heterologous association of transmembrane helices in a biological membrane. J Biol Chem 278, 3105-3111.

16

(32) Smyth, M. S., and Martin, J. H. (2000) x ray crystallography. Mol Pathol 53, 8-14.

(33) Tooze, C. B. a. J. (1998) Introduction to protein structure, 2nd Edition ed. Garland Publishing, Inc., New York.

(34) Rhodes, G. (1993) Crystallograpphy made crystal clear. A guide for users of macromolecular models. Academic Press, San Diego.

(35) Wiener, M. C. (2004) A pedestrian guide to membrane protein crystallization. Methods 34, 364-372.

(36) Zapun, A., Bardwell, J. C., and Creighton, T. E. (1993) The reactive and destabilizing disulfide bond of DsbA, a protein required for protein disulfide bond formation in vivo. Biochemistry 32, 5083-5092.

(37) Kuliopulos, A., Nelson, N. P., Yamada, M., Walsh, C. T., Furie, B., Furie, B. C., and Roth, D. A. (1994) Localization of the affinity peptide-substrate inactivator site on recombinant vitamin K-dependent carboxylase. J Biol Chem 269, 21364-21370.

(38) Bernaudat, F., Frelet-Barrand, A., Pochon, N., Dementin, S., Hivin, P., Boutigny, S., Rioux, J. B., Salvi, D., Seigneurin-Berny, D., Richaud, P., Joyard, J., Pignol, D., Sabaty, M., Desnos, T., Pebay-Peyroula, E., Darrouzet, E., Vernet, T., and Rolland, N. (2011) Heterologous expression of membrane proteins: choosing the appropriate host. PLoS One 6, e29191.

(39) Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Hogbom, M., van Wijk, K. J., Slotboom, D. J., Persson, J. O., and de Gier, J. W. (2008) Tuning Escherichia coli for membrane protein overexpression. Proc Natl Acad Sci U S A 105, 14371-14376.

Chapter 2

IDENTIFYING KEY JUXTAMEMBRANE INTERACTIONS IN CELL MEMBRANES USING ARAC-BASED TRANSCRIPTIONAL REPORTER ASSAY (ARATM)

Dimerization is a key regulatory mechanism in activation of transmembrane (TM) receptors during signal transduction. This process involves a coordinated interplay between extracellular (EX), TM, and cytoplasmic (CYTO) regions to form a specific interface required for both ligand binding and intracellular signaling to occur. While several transcriptional activator-based methods exist for investigating TM interactions in bacterial membranes, expression of TM chimera in these methods occurs in a reverse orientation, and are limited to only TM domains for proper membrane trafficking and integration. We therefore developed a new, AraC-based transcriptional reporter assay (AraTM) that expresses EX-TM-CYTO chimera in their native orientation, thereby enabling membrane trafficking to occur independent of the TM chimera used as well as permitting analysis of EX-TM-CYTO interactions in biological membranes. Using integrin α_{IIb} TM-CYTO as a model, we observe a large increase in homodimerization for the constitutively active TM mutant L980A relative to wild-type in the TM-CYTO construct (A963-E1008). We also characterized the receptor for advanced glycation endproducts (RAGE), whose homooligomeric state is critical in

ligand recognition, and find the specific juxtamembrane region within the CYTO (A375-P394) mediates homodimerization, and is dominant over effects observed when the extracellular C2 domain is included. Furthermore, we find good agreement between our AraTM measurements in bacterial membranes and BRET measurements made on corresponding RAGE constructs expressed in transfected HEK293 cells. Overall, the AraTM assay provides a new approach to identify specific interactions between receptor EX-TM-CYTO domains in biological membranes that are important in regulation of signal transduction.

2.1   Introduction

Transmembrane receptors play critical roles in regulating diverse biological processes, including cell-cell communication, adhesion, and proliferation. In particular, receptor dimerization is a key mechanism by which extracellular cues such as ligand binding are communicated across the cell membrane to activate intracellular signaling processes through formation of a multimeric, signaling complex (1). Recent work has illustrated the importance of specific interfaces involving transmembrane (TM) regions of single- and multi-pass membrane proteins in stabilizing receptor signaling complexes, as well as cooperative interactions between TM, cytoplasmic (CYTO) and extracellular (EX) regions that regulate ligand-dependent signal transduction across cell membranes (2). One such example are the integrins, which are a family of heterodimeric, cell surface receptors involved in cell adhesion and cell-cell communication (3). For platelet integrin $\alpha_{IIb}\beta_3$, heterodimeric complexes stabilized

through specific EX, TM and CYTO interactions occur in the inactive form, and both α_{IIb} and β_3 undergo homooligomeric clustering into larger focal adhesion complexes upon inside-out activation (3–7). Likewise, basic residues in the juxtamembrane (JM) region of glycoprotein Ib CYTO are important in stabilizing the glycoprotein Ib-IX complex (8, 9). In the case of platelet-derived growth factor β receptor (PDGFβ), while TM domain interactions involving a conserved, GX_3G interface drive homodimerization, purified peptide corresponding to a JM-TM-CYTO fragment of PDGFβ forms a obligate dimer, which suggests the CYTO rather than the TM enforces the stoichiometry of the signaling complex (10). Thus, the interplay between TM, CYTO and EX interactions is central to defining both the active and inactive forms of oligomeric receptor complexes during signal transduction.

Transcription factor-based selection systems such as TOXCAT, POSSYCAT, TOXluc and GALLEX are essential tools dimerization (11–15). In particular, TOXCAT and POSSYCAT are based on TM domain fusions to ToxR, which is a transmembrane transcription factor from V. cholera (16). Upon dimerization, ToxR binds the *ctx* promoter and activates gene transcription through its N-terminal DNA binding domain (16). Unlike most single-pass TM receptors, which contain an N-terminal signal peptide to direct membrane insertion (type I) as well as a TM domain (17), ToxR is a type II integral membrane protein, where the C-terminal TM domain functions to direct insertion and integration into the cytoplasmic membrane (18). For TOXCAT, TOXluc, and POSSYCAT, the TM domain of ToxR is replaced with a given TM domain of interest and the periplasmic domain is replaced with a truncated form of maltose-binding protein (MBP) that lacks its N-terminal signal peptide sequence (12–14). For

20

each assay, the heterologous TM domain of interest functions as a surrogate signal peptide to direct insertion of the TM-MBP fusion in an inverted orientation in the inner membrane of *E. coli* (12–14). The orientation of expressed chimera using ToxR is illustrated in Fig. 1. The extent of TM-mediated oligomerization is reflected in the level of chloramphenicol acetyltransferase (CAT), β-galactosidase, or luciferase reporter gene transcription, all of which are under control of the *ctx* promoter (12–14). Although ToxR-based methods are useful for studying the homodimeric interaction of TM domains, the N-terminal DNA-binding domain and type II orientation of ToxR requires TM constructs to be expressed as C-terminal fusions directly adjacent to ToxR. Thus, addition of EX or CYTO fragments (i.e. EX-TM-CYTO) to ToxR-TM-MBP chimera interfere with the ability of the TM to act as a signal peptide to properly traffic chimera to the bacterial inner membrane. Additionally, the LexA-based GALLEX method has been used with success to investigate TM domain homo- and heterodimerization (11). However, as with ToxR, the N-terminal orientation of the LexA DNA binding domain necessitates expression of TM-MBP fusions in a reverse orientation in order to insert into the bacterial inner membrane (Fig. 1), thereby permitting only TM domains for analysis (11). Furthermore, wild-type LexA is a soluble, cytosolic protein, so the heterologous TM domain of interest fused to LexA acts as a surrogate signal peptide sequence to the MBP fragment to promote insertion of LexA-TM-MBP chimera into the bacterial inner membrane in an inverted orientation. Therefore, two limitations of current LexA- and ToxR-based methods are that they can only be applied to TM domains, and they place TM domains in an inverted orientation in the bacterial inner membrane.

21

To address these limitations, we developed a new approach based on the Escherichia coli AraC transcription factor to investigate EX-TM-CYTO receptor domain homodimerization (AraTM). Unlike ToxR or LexA, AraC contains a C-terminal DNA binding domain, and is therefore compatible with expressing EX-TM-CYTO fusions in their native orientation as an N-terminal fusion to full-length MBP that includes its native signal peptide (Fig. 2.1). The native signal peptide present in MBP promotes membrane trafficking, eliminating the need for the TM domain to act as a surrogate signal peptide (19). To quantifying dimerization, we developed an eGFP reporter under control of the AraC-regulated *araBAD* promoter that enables measurements on whole cells directly from culture. We used integrin α_{IIb} TM-CYTO as a model system to benchmark the sensitivity of our assay to specific TM domain mutations known to affect homodimerization, and extended this approach to identify domains involved in the homodimerization of the receptor for advanced glycation endproducts (RAGE). Our results demonstrate the sensitivity of the AraTM method to specific point mutations or domain deletions within EX-TM-CYTO constructs, and illustrates the utility of AraTM as a complementary approach to other mammalian cell-based assays to investigate the key domains and motifs responsible for mediating receptor dimerization.
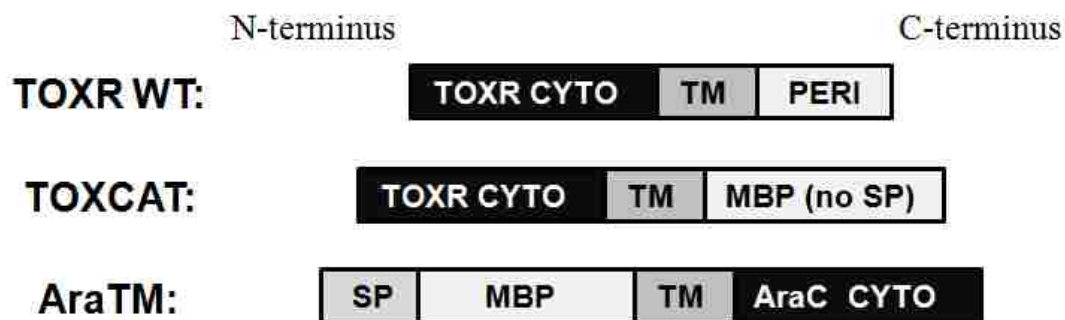
Figure 2. 1. Organization of AraC and ToxR Fusions Used in TM Interaction Assays. In ToxR-based assays (12–14), constructs are configured with ToxR as an N-terminal fusion to a TM domain of interest, with mature MBP lacking its signal peptide (SP) fused at the C-terminus. ToxR is a type II integral membrane protein (18), and therefore the TM domain of interest acts both as a signal peptide to direct membrane trafficking as well as membrane integration with the TM domain in a reverse orientation. In the AraTM assay, the C-terminal orientation of the DNA-binding domain within AraC enables constructs to be expressed with a N-terminal MBP fusion, which includes its native signal peptide to direct membrane integration. Thus, AraTM constructs are expressed in their native orientation, and membrane integration is decoupled from the specific sequence fused to MBP and AraC.

2.2   Materials and Methods

2.2.1   Subcloning

Unless otherwise stated, standard molecular biology techniques were used. All constructs used were verified by DNA sequencing (Department of Genetics Core Facility, University of Pennsylvania School of Medicine).

Plasmid pTrcRSF was generated by amplifying the *tac* promoter region and multiple cloning site from plasmid pTrc99a as a SpeI fragment and RSF origin of replication and kanamycin resistance marker from plasmid pRSF1b as a SpeI fragment, and ligating these two fragments after SpeI digestion and dephosphorylation of the RSF

fragment using CIP. AraC was amplified from plasmid pDS439 and subcloned into pTrcRSF as a KpnI/HindIII fragment (20). Full-length MBP, including the N-terminal signal peptide sequence for direct trafficking to the periplasmic membrane, was amplified from pMAL-p2e (NEB) with the reverse primer containing an HA epitope sequence (YPYDVPDYA) and subcloned into pTrcRSF as a NcoI/SacI fragment. A fragment of the MCS from pET28a with additional, unique restriction enzyme sites was subcloned between SacI and KpnI sites. The resultant plasmid (pAraTM) was used for all subsequent experiments, with RAGE and integrin α_{IIb} domains cloned in-frame as fusions with MBP and AraC (amino acids 168–293) as SacI/KpnI fragments. Integrin α_{IIb} inserts were generated using overlap-extension PCR with synthetic oligonucleotides corresponding to the TM and JM regions) and RAGE inserts were amplified from plasmid (OriGene) containing the full-length receptor.

The negative control plasmid pAraCY used for the maltose complementation test, which lacks both a TM domain and N-terminal signal peptide sequence, was generated by amplifying the MBP fragment lacking N-terminal signal peptide from plasmid pToxR1, and ligating after NcoI/SacI digestion into plasmid pTrcRSF containing AraC as described above. The negative control plasmid used for all other experiments was pTrcRSF.

The reporter plasmid pAraGFP (Sequence 2.S3) was derived from plasmid pDS439 by amplifying a fragment containing the *araBAD* promoter, MCS, ampicillin resistance marker and pBR322 origin of replication. This fragment also contains GFP under control of the *araBAD* promoter, and removes the AraC transcriptional activator present on plasmid pDS439. The reporter plasmid pAraGFP was generated by digestion

with SpeI and self-ligation with T4 DNA ligase.

For BRET, full-length human RAGE receptor and RAGE receptor fragments were amplified from plasmid (OriGene) and cloned as fragments into plasmidspGFP2 - N3 as NheI/SacI fragments and pRluc-N2 as BglII/KpnI fragments (BioSignal Packard).

2.2.2   AraTM dimerization assay

Plasmids pAraTM and pAra-GFP were co-transformed into the AraC-deficient *E. coli* strain SB1676 (The E. coli Genetic Stock Center at Yale University) and streaked on selective LB plates (100 μg/ml ampicillin, 50 μg/ml kanamycin). 8 colonies were picked for each construct and grown in selective LB media for 16 h at 37 °C. After 16 h, 1 μl of saturated culture was added to 1 ml of fresh, selective LB media and grown for an additional 16 h. 400 μl of cultures from each sample was transferred to a black 96-well, clear bottom plate (Greiner) and a series of 2-fold serial dilutions was prepared using selective LB media. A_{600} measurements and GFP fluorescence emissions spectra (excitation maximum at 485 nm and emissions maximum at 530 nm) were collected using a M200 Infinity plate reader (Tecan). Results are reported as the ratio of fluorescence emission at 530 nm to absorbance at 560 nm.

For FACS measurements, 1 ml of each AraTM construct was grown in selective LB (100 μg/ml ampicillin, 50 μg/ml kanamycin) for 16 h at 37 °C and 200 rpm overnight. After 16 h, 1 μl of saturated culture was added to 1 ml of selective LB media and grown for an additional 16 h. Afterward, cultures were analyzed by flow cytometry

(BC FACSCanto II) using 515–545 nm emissions filter for GFP. 100,000 events were collected for each construct and the gate was set so that 0% of the negative control cells were GFP-positive.

2.2.3 Maltose complementation test

Plasmid pAraTM containingTMinserts were transformed into the MBP-deficient *E. coli* strain MM39, streaked onto selective LB plates (50 μg/ml kanamycin) and incubated at 37 °C. The following day, colonies from each plate were grown in selective LB media (50 μg/ml kanamycin) at 37 °C and 200 rpm overnight. 5 μl of saturated culture from each construct was streaked on selective M9 minimal media plates (50 μg/ml kanamycin) containing 0.4% (w/v) maltose and incubated at 37 °C for 2 days to assess growth (12).

2.2.4 Spheroplast protection assay and immunoblotting

For spheroplast immunoblotting, pAraTM plasmids were transformed into MBP-deficient MM39 cells and cultures grown to A_{600} of 0.6 at 37 °C and 200 rpm before collecting cells. For immunoblotting to detect AraC-chimera expression from whole-cell extracts, plasmid pAraTM containing TM inserts and pAraGFP were transformed into SB1676 cells and cultures were grown for 16 h at 37 °C and 200 rpm. The following morning, cell cultures were diluted back down to A_{600} of 0.6 before processing cell lysates.

20 μl of culture was mixed with 5 μl of 5X Lammeli sample buffer, heated briefly at 90 °C, then loaded onto a 12% acrylamide gel. Samples were run for 1 h at 200 V using Lammeli running buffer, then transferred to a nitrocellulose membrane (Amersham Biosciences Hybond ECL) for 90 min, blocked for 1 h at room temperature using 5% milk in TBST, then incubated with 1:10,000 dilution of HRP-conjugated anti-MBP monoclonal antibody (NEB) for 1 h at room temperature. Membranes were developed using a chemiluminescent substrate (GE) and imaged using a Typhoon imager.

We used a previously described spheroplast protection assay to assess membrane insertion and integration of MBP-containing chimeras (12). Briefly, cultures were processed using osmotic lysis (Epicenter Biotechnologies) to remove the outer membrane and isolate periplasts and spheroplasts. Spheroplasts were separated from periplasts by centrifugation (12,000 x g) for 5 min, and treated with proteinase K (to determine chimera orientation in theperiplasmic membrane) and Nonidet P-40 (to dissolve spheroplasts and release cytosolic proteins). Chimera expression levels in periplasts, spheroplasts, and cytosolic fractions were detected by immunoblotting as described above.

2.2.5   Bioluminescence resonance energy transfer (BRET)

8 μg of RAGE-GFP and -Rluc fusions were transfected into HEK293 cells by electroporation (GenePulser, Bio-Rad) using the HEK293 preset protocol in HEBS buffer (pH 7.05). Immediately after electroporation, cells were transferred to a white,

roundbottom 96 well plate (100 µl of cells per well) in DMEM containing 10% FBS with L -glutamine/VitaMax supplements and 1% Penn/Strep, and incubated at 37 °C and 5% CO_2 for 48 h. Media was removed, cells were washed once with 100 µl of PBS per well before 100 µl of BRET buffer (0.1 g/liter $CaCl_2$, 0.1 g/liter $MgCl_2$, and 1 g/liter D -glucose in PBS) was added to each well. 5 µl of Deep Blue C (GoldBio) was added to each sample well, and luminescence measurements (Filter 1: Green and Filter 2: Magenta) were collected using a M200 Infinity plate reader (Tecan) over the course of the 10 s. The energy transfer efficiency is calculated by dividing the intensity of the signal for the green channel by the intensity of the signal for the magenta channel (21) in Equation 2.1.

Energy Transfer Efficiency = (GFP Signal at 515 nm) / (Magenta Signal at 410 nm) (Eq. 2.1)

The expression level of each eGFP2 -fused and Rluc-fused RAGE truncation was evaluated through Western blotting against anti-GFP and anti-Rluc antibody, respectively. Loading control was evaluated through Western blotting against antitubulin antibody and HEK293 cells expressing empty eGFP2 and Rluc vectors were used as negative control for BRET (Fig. 2.S1).

2.3   Results

2.3.1   Design of AraTM assay

The AraTM assay is based on transcriptional activity of the *E. coli* AraC transcription factor, which is active at the *araBAD* promoter as a homodimer (22). Wild-type AraC includes an N-terminal dimerization and ligand-binding domain (amino acids 1–87), which undergoes a conformational change upon L-arabinose binding to bind the I1 and I2 half-sites in the *araBAD* promoter, thereby driving gene expression (19). Genetic interaction studies indicate that replacement of the wild-type AraC dimerization domain with homo- and heterodimeric coiled-coils such as GCN4 and Fos-Jun causes constitutive activation at the *araBAD* promoter, leading to the view that AraC is a modular transcriptional activator whose activity can be modified through controlling N-terminal domain dimerization (23, 24). We therefore modified wild-type AraC by replacing the N-terminal dimerization domain with receptor EX-TM-CYTO domains, and utilize GFP expression under control of the *araBAD* promoter as a measure of receptor domain dimerization. To ensure that EX-TM-CYTO fusions are properly integrated into the bacterial inner membrane, we generated chimeras that also contain an N-terminal fusion to full-length maltose-binding protein (MBP), including its native signal peptide sequence. MBP is a soluble, periplasmic protein responsible for high affinity binding and transport of maltose across the periplasmic membrane (25). Thus, full-length MBP directs the correct expression and integration of receptor EX-TM-CYTO fusions into the periplasmic membrane through its native signal peptide sequence, thereby eliminating the need for the TM domain to act as a surrogate signal

peptide. Proper integration of MBP-EX-TM-CYTO-AraC chimera can be confirmed using both complementation experiments on growth media containing maltose as the sole carbon source as well as immunoblotting with anti-MBP antibodies (12). An overview of the AraTM assay is given in Fig. 2.2, and nucleic acid sequences for plasmids and primers used to develop the AraTM assay are provided in supplemental materials Section 2.6.
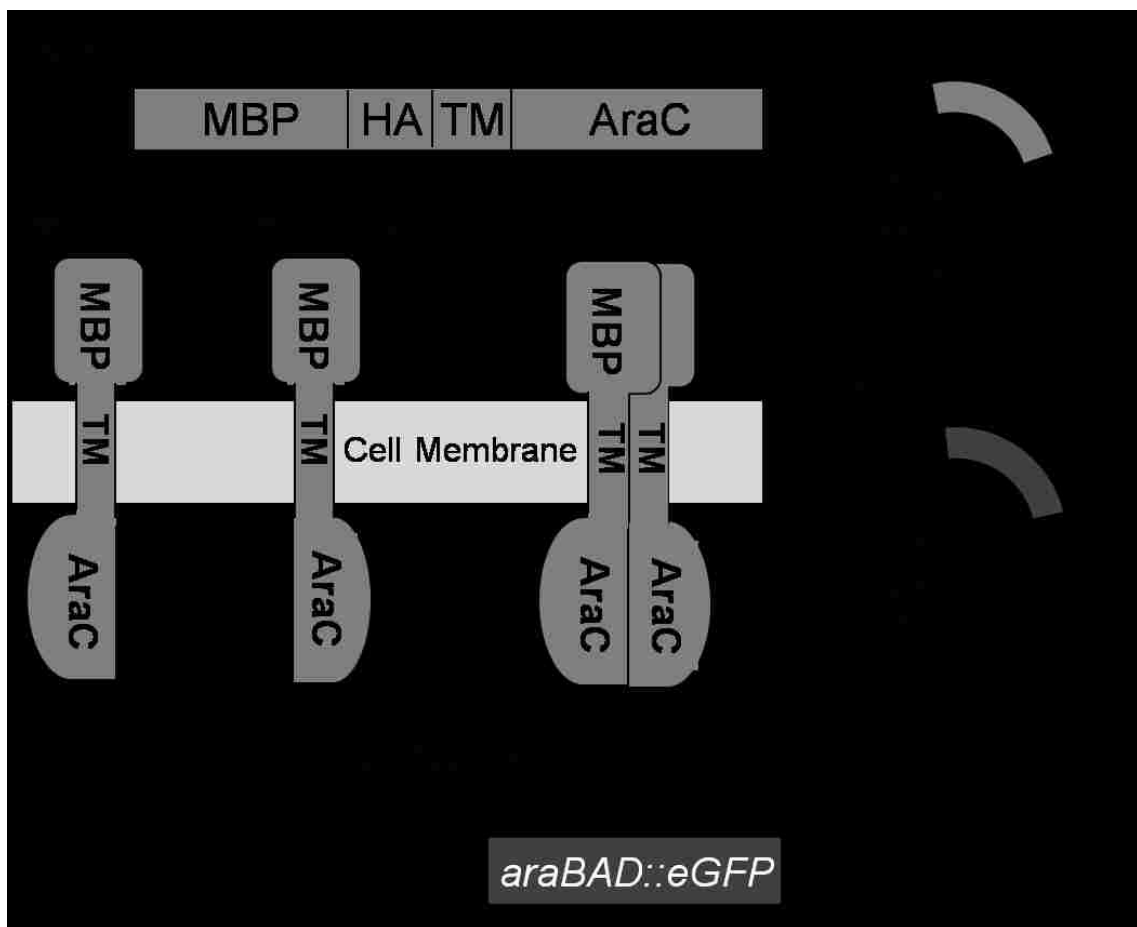
Figure 2. 2. Overview of AraTM Assay. Chimeric proteins containing N-terminal MBPand C-terminal AraC domains fused with an in-frame receptor fragment are expressed by the regulator plasmid (pAraTM; kanamycin resistant). Once expressed, MBP directs expression of the chimera to the inner membrane of *E. coli*. Homodimerization brings AraC transcriptional factors in close proximity, enabling binding to the *araBAD* promoter on the reporter plasmid (pAraGFP; ampicillin resistance) and activating transcription of the reporter gene GFP.

2.3.2   Homodimerization of integrin α_{IIb} TM-CYTO domains

To confirm the AraTM assay is sensitive to specific residues known to be responsible for driving TM-CYTO domain dimerization, we used the TM-CYTO of human integrin αIIb as a model system (A963-E1008; Fig. 2.3A). Integrin $\alpha_{IIb}\beta_3$ is the major platelet receptor for fibrinogen, and is comprised of 2 type I integral membrane proteins that form a non-covalent heterodimer in the inactive state stabilized by both specific TM and CYTO interactions (3). Specifically, the integrin α_{IIb} TM domain contains a conserved, GX_3G motif that is important for homodimerization, which has led to a proposed push-pull mechanism for regulation of the inactive, active and clustered states in which the interplay between TM and CYTO interactions are important (4, 26, 27). For the wild-type integrin α_{IIb} TM-CYTO (Figs. 2.3 and 2.4), significant homodimer formation occurs, with a GFP signal more than twice that of the negative control. Furthermore, when we introduce the specific TM mutation L980A into the TM-CYTO (A963-E1008), which was shown to substantially increase homooligomer formation measured using TOXCAT on a truncated TM construct (W968-K989) as well as by analytical ultracentrifugation and gel-shift assays on a purified TM-CYTO construct (A958-E1008), we observe a 1.5-fold increase in GFP signal relative to wild-type integrin α_{IIb} TM-CYTO (Fig. 2.3, B and C) (6, 26). To correct for variation in homodimerization signal with cell density, results are reported as the slope of fluorescent intensity for GFP emission versus cell density for a series of cell dilutions. Both wild-type and L980A integrin α_{IIb} TM-CYTO gives consistent results across several independent replicates at varying cell densities, with the majority

31

of individual measurements falling within the estimated 95% confidence interval for the entire dataset (Fig. 2.3B). We also quantified homodimerization of integrin α_{IIb} TM-CYTO constructs in individual cells using flow cytometry (Fig. 2.4). Compared with the wild-type integrin α_{IIb} TM-CYTO, the constitutively active L980A mutant has a nearly 2-fold increase in GFP-positive cell population (Fig. 2.4, A and B), with a clear shift in the overall cell population to higher GFP expression levels rather than skew caused by a small subpopulation of highly fluorescent cells. Thus, the increase in dimerization observed for the L980A mutant reflects a true population average throughout the entire cell suspension. Furthermore, the magnitude of the increase in GFP fluorescence for wt versus L980A is similar for measurements using a micro- plate reader (Fig. 2.3) and flow cytometer (Fig. 2.4), which indicates the robustness of the AraTM method to multiple, independent methods for quantitation.

Figure 2. 3. AraTM Assay rRsults for Wild-type Integrin α_{IIb} TM-CYTO and Its Mutations. A, integrin α_{IIb} TM-CYTO sequence subcloned into pAraTM. The L980A mutation is also highlighted. B, fluorescence intensity at 530 nm of wild-type integrin α_{IIb} TM-CYTO and mutant L980A from serial dilutions of bacterial cultures plotted against the corresponding cell density (OD600). Solid lines represent the best-fit line through the experimental data, and the dotted lines represent the upper and lower bound of the 95% confidence interval for each data set. C, average slopes from fluorescence intensity versus OD600 for each construct are compared from eight independent replicates, with error bars representing standard error.

33

Figure 2. 4. Population Distribution of GFP-positive Cells for Various Integrin α_{IIb} AraTM Constructs by Flow Cytometry. A, compared with wild-type, the mutant L980A shows an increase in the % of total cell population (100,000 events) that is GFP-positive as well as a shift in the overall cell population measured in terms of FSC to a higher average GFP signal, which is in good agreement with whole-cell fluorescence measurements from cell suspension (Fig. 2.3). B, integrin α_{IIb} L980A mutant shows a 2-fold increase in the GFP-positive population as compared with wild-type.

Integrin α_{IIb} wild-type and L980A TM-CYTO were both properly integrated in the *E. coli* inner membrane, as evidenced by growth on maltose M9 minimal plates (Fig. 2.5A), and expressed at similar levels, as confirmed by immunoblotting from whole-cell lysates (Fig. 2.5B). For integrin α_{IIb} TM-CYTO mutant L980A, we also confirmed proper membrane integration and orientation using a modified spheroplast digestion assay described previously (12, 26). AraTM constructs were detected using an HRP-conjugated anti-MBP antibody. In particular, we observe intact chimera in the spheroplast fraction, but not the periplasmic or extracellular space, consistent with an inner membrane-integrated TM-CYTO chimera (Fig. 2.5C) rather than a secreted, soluble periplasmic TM-CYTO chimera. Furthermore, degradation of MBP by proteinase K treatment occurs in the spheroplast fraction in the absence of Nonidet P-40 (Fig. 2.5C), which indicates the correct, N-terminal orientation of MBP in the periplasmic space accessible to proteinase K rather than accumulation of improperly trafficked chimera in the cytoplasmic space. If the MBP-TM-CYTO-AraC chimera was unable to integrate into the bacterial inner membrane, the membrane would act as a barrier to prevent degradation by proteinase K, resulting in detection of full-length chimera after proteinase K treatment of intact spheroplasts. Whereas in TOXCAT, the spheroplast prevents complete cleavage of MBP by proteinase K for properly integrated chimera (26), the addition of the flexible linker, including an HA epitope, between the N-terminal MBP and integrin α_{IIb} TM-CYTO domain in AraTM increases proteinase K accessibility, which explains the loss of detectable chimera observed in intact spheroplasts by anti-MBP immunoblotting. In addition, for integrin α_{IIb} TM-CYTO

35

L980A mutant, chimera expression begins after 3 h and remaining constant overnight as detected by immunoblotting from whole-cell lysates with anti-MBP antibody (Fig. 2.6). Thus, the AraTM assay is able to capture relative changes in homodimerization between wild-type integrin α_{IIb} TM-CYTO (A963-E1008) and the key mutant L980A that has been show previously to cause constitutive activation in the full-length receptor (26). Furthermore, the AraTM assay is capable of properly integrating larger TM-CYTO fragments (A963-E1008; Fig. 2.3) into the bacterial inner membrane, in contrast to truncated TM domain fragments (W968-K989) used previously in TOXCAT (26).
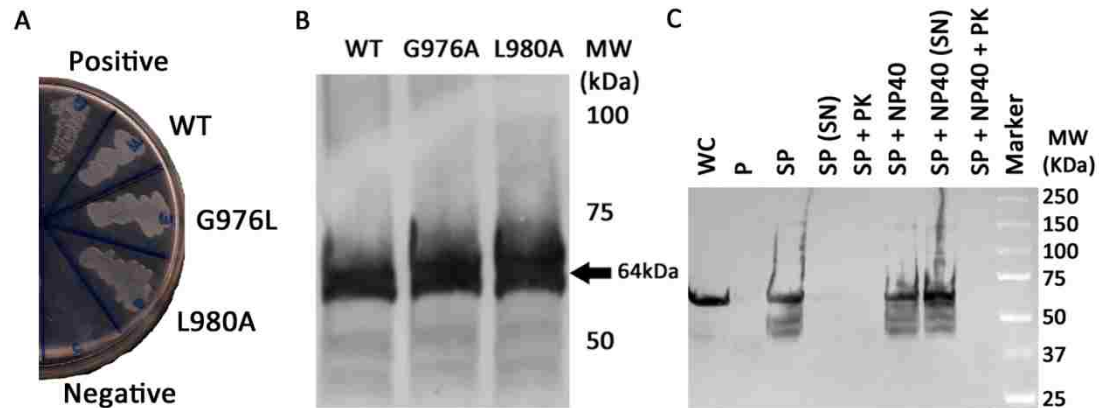
Figure 2. 5. The Integrin α_{IIb} TM-CYTO Chimera Express at Similar Levels and Are Properly Integrated into the Inner Membrane of *E. coli*. A, AraTM chimeras containing wild-type and mutant integrin α_{IIb} TM-CYTO expressed in MalE-deficient MM39 cells were streaked on a 0.4% maltose M9 plate and incubated for 48 h at 37 °C. Each construct is properly integrated into the inner membrane of *E. coli*, as indicated by robust growth on the 0.4% maltose M9 plates similar to the positive control (pTrcRSF containing MBP-AraC chimera). As expected, no growth is observed on the negative control (AraCY). B, wild-type and mutant integrin α_{IIb} TM-CYTO chimera were expressed at equal levels as determined by immunoblotting with HRP-conjugated anti-MBP antibody, and the observed chimera MWs were consistent with the expected MWs. C, periplasts and spheroplasts were prepared for mutant integrin α_{IIb} L980A TM-CYTO, treated with and without Nonidet P-40 (1% v/v) and proteinase K (50 μg/ml), and blotted against anti-MBP antibody (WC: whole cell, P: periplast, SP: spheroplast, SN: supernatant, PK: proteinase K, and Nonidet P-40: Detergent Nonidet P-40). No chimera is detected in intact spheroplasts treated with proteinase K (SP+PK) nor in the periplasmic fraction of the cell, consistent with the expected periplasmic orientation and membrane integration of the MBP fusion.

36
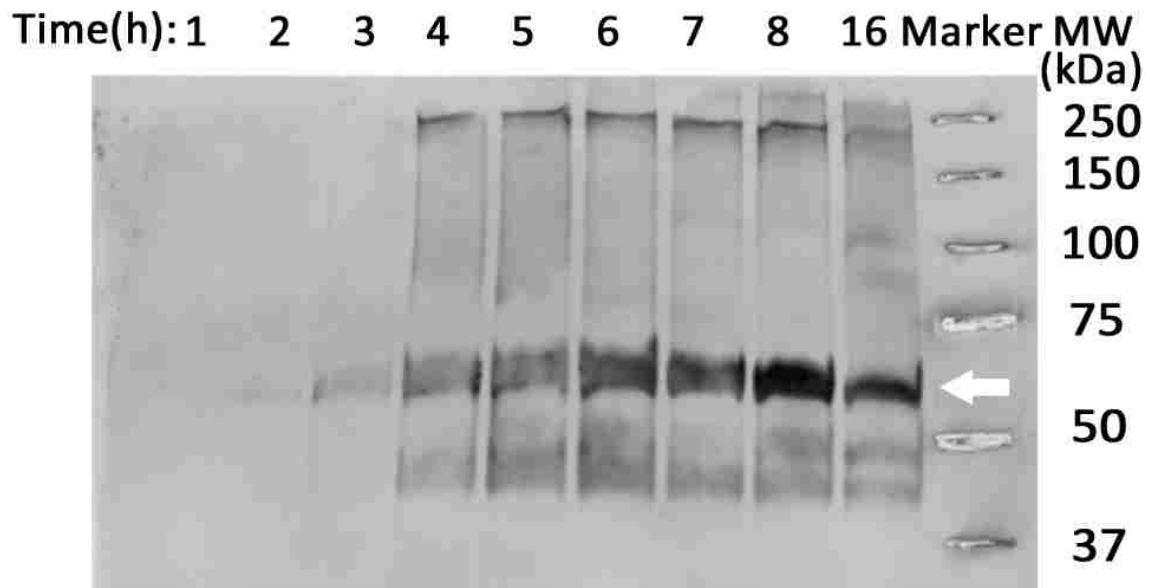
Figure 2. 6. Monitoring Expression of AraTM Chimera in Cell Culture. Cells expressing integrin α_{IIb} L980A chimera were collected hourly to monitor expression level and degradation products by immunoblotting with an anti-MBP antibody. Chimera expression began 3 h post-induction at 37 °C and remained constant overnight.

2.3.3   Identifying domains critical for homodimerization of RAGE

RAGE is a 43 kDa type I TM receptor that belongs to the immunoglobulin superfamily (28). RAGE is expressed in a wide range of tissues, including brain, liver, and heart, where it activates pro-inflammatory signaling as part of the innate immune system in response to external stress (29). RAGEis capable of binding several ligands such as S100-family proteins, amyloid-β peptide and advanced glycation endproducts (AGE), and the binding of each ligand as well as subsequent signaling depends on the oligomeric state of the receptor (28, 30). RAGE exists in 2 major forms (Fig. 2.7A): a full-length, TM receptor (RAGE) as well as a soluble, extracellular domain (sRAGE) that acts as a dominant-negative to inhibit ligand-dependent signaling (28). Most work

has focused on determining the structural basis for dimerization of sRAGE, which includes the V-domain (Fig. 2.7) that is responsible for binding AGE, as well as the V-C1-C2 domain fragment (Fig. 7) that is responsible for binding multimeric S100A-family ligands (31, 32). However, much less is known about specific TM-CYTO interactions that occur in oligomeric states of RAGE.
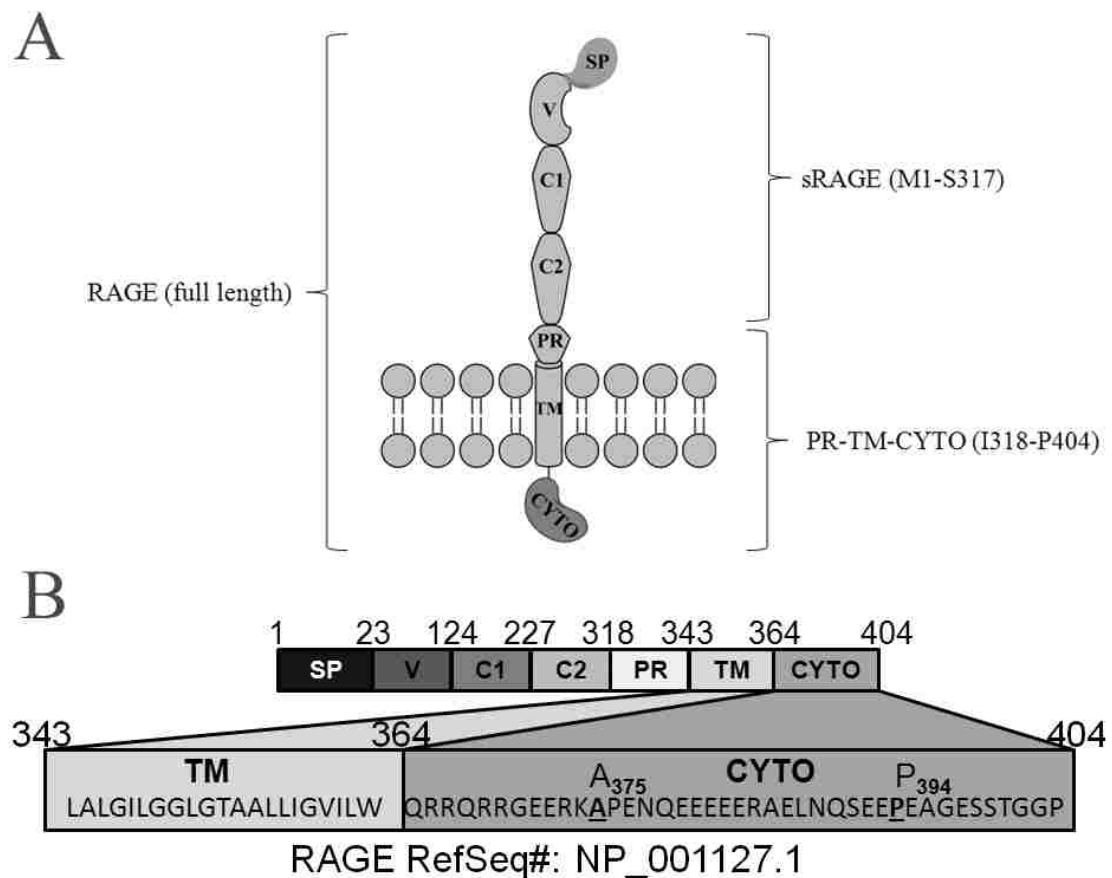
Figure 2. 7. RAGE. A, graphic illustration of domain structure in the full-length receptor, including the TM and CYTO regions. B, annotation of amino acid sequences in full-length RAGE corresponding to specific domains within the receptor. Sequences for the TM and CYTO region of RAGE are given, including positions (A375, P394) for specific truncations in the CYTO region (SP: signal peptide, V: V domain, C1: C1 domain, C2: C2 domain, PR: proximal domain).

We generated a series of RAGE domain deletions, including deletions of extracellular C2 and CYTO domains, and quantified their homodimerization using the AraTM assay. Domains and numbering used to delineate the specific domains with RAGE used are given in Fig. 2.7B. Full-length RAGE (V-C1-C2-PR-TM-CYTOfull; A23-P404) and RAGE chimera containing C1-domains (C1-C2-PR-TM-CYTOfull; P124-P404) were unable to complement grow on maltose minimal media, and were therefore excluded from further analysis (Fig. 2.8A). However, all other RAGE constructs were properly integrated into the bacterial inner membrane, as indicated by robust growth on maltose M9 minimal plates (Fig. 2.8A), and each was expressed at similar levels as indicated by whole-cell immunoblotting with anti-MBP (Fig. 2.8B). Isolation of spheroplasts from cells expressing RAGE PR-TM-CYTOfull (R314-P404) indicated they were properly oriented into the inner membrane (Fig. 2.8C), as degradation of MBP in intact spheroplasts occurred only after proteinase K treatment in the absence of Nonidet P-40 (12). We find that deletion of the C2 domain has little impact on homodimerization in constructs containing the CYTO domain (C2-PR-TM-CYTOfull; P224-P404 and PR-TM-CYTOfull; R314-P404), whereas deletion of the CYTO domain significantly reduces homodimerization largely independent of the C2 domain (C2-PR-TM; P224-R365 and PR-TM; R314-R365) (Fig. 2.9). We made additional deletions within RAGE CYTO to identify which regions were most important for homodimerization. We find that removal of the C-terminal distal region (PR-TM-CYTOP394; R314-P394) has no measurable impact on homodimerization, whereas deletion of the glutamic acid-rich region of the CYTO (PR-TM-CYTOA375; R314-A375) causes a modest decrease in homodimerization. However, deletion of the

39

entire CYTO domain (PR-TM; R314-R365) causes a greater than 2-fold decrease in dimerization (Fig. 2.9), which indicates the juxtamembrane region (R365-A375) present within the CYTO is key for dimerization of TM RAGE constructs. Thus, we conclude that the RAGE CYTO exhibits strong homodimerization in the context of C2-TM-CYTO domains expressed in cell membranes, particularly the juxtamembrane region (R365-A375) present within the CYTO.
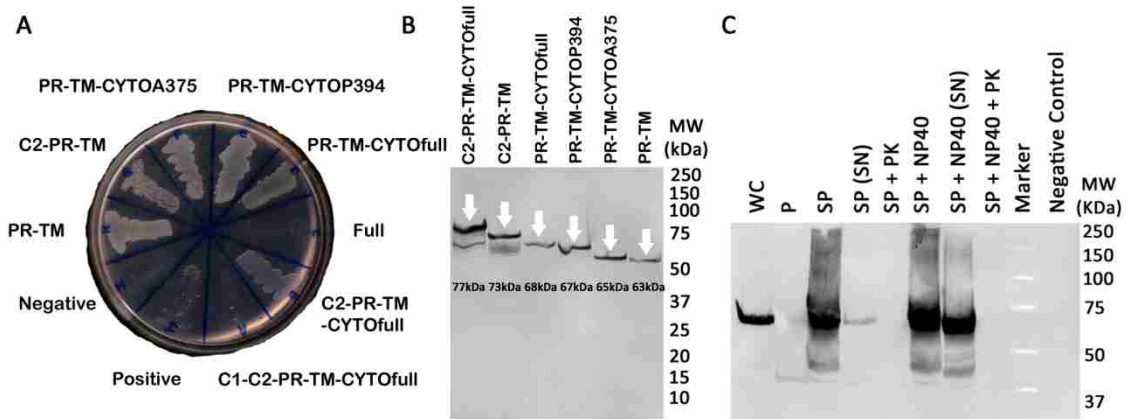
Figure 2. 8. Expression of RAGE constructs in AraTM assay. A, RAGE C2-PR-TM-CYTOfull AraTM chimera as well as additional truncations of RAGE C2-PR-TM-CYTOfull are able to complement growth on maltose M9 minimal plates. B, cells expressing RAGE chimera were expressed at similar levels and the chimera MWs are consistent with the expected MWs as determined by immunoblotting from whole-cell lysates with anti-MBP antibody. C, periplasts and spheroplasts of the RAGE PR-TM-CYTOfull chimera were prepared, treated with/without Nonidet P-40 (1% v/v) and proteinase K (50 μg/ml), and blotted against anti-MBP antibody (WC: whole cell, P: periplast, SP: spheroplast, SN: supernatant, PK: proteinase K, and Nonidet P-40: detergent Nonidet P-40). No chimera is detected in intact spheroplasts treated with proteinase K (SP+PK) nor in the periplasmic fraction of the cell, consistent with the expected periplasmic orientation and membrane integration of the MBP-RAGE-AraC fusion.

2.3.4 Homodimerization of RAGE receptor expressed in mammalian membranes

To confirm that our results from the AraTM assay are relevant in the context of full-length RAGE, we generated full-length RAGE and RAGE CYTO deletions as fusions with both eGFP2 and Rluc, co-expressed both eGFP2 and Rluc fusions in HEK293 cells, and measured the apparent interaction between each CYTO domain deletion using BRET (33, 34). The extent of energy transfer was calculated by dividing the green signal (515 nm) by the magenta signal (410 nm) as described previously (33). Expression levels for each construct as eGFP2 and Rluc fusions were similar among different truncations as determined by immunoblotting with specific antibodies (supplemental materials Fig. 2.10). Consistent with previous FRET measurements in transfected HEK293 cells (35), we find that full-length RAGE receptor exhibits a strong homodimeric signal in the absence of ligand (Fig. 2.11). Consistent with our AraTM results, removal of the C-terminal distal region (V-C1-C2-PR-TM-CYTOP394; A23-P394) has no impact on homodimerization, but removal of the glutamic acid-rich CYTO region (V-C1-C2-PR-TM-CYTOA375; A23-A375) causes a significant decrease of nearly 25% in the measured BRET signal relative to wild-type RAGE (Fig. 2.11). It should be noted that the decrease in observed homodimerization for the V-C1-C2-PR-TM-CYTOA375 construct measured using BRET (Fig. 2.11) is greater than the intermediate reduction in homodimerization signal measured using AraTM (Fig. 2.9), which indicates a significant, but intermediate effect on dimerization. Deletion of the CYTO reduces the measured BRET signal by 50% to near background levels. Thus, the trend in terms of reduction in homodimerization for the PR-TM-CYTO deletions

41

observed in AraTM (Fig. 2.9) are consistent with RAGE truncations expressed in HEK293 cells (Fig. 2.11), and the unliganded receptor is stabilized through interactions involving the juxtamembrane region of RAGE (R365-A375) present in the CYTO.

Figure 2. 9. Cytoplasmic Truncations of RAGE Reduce Homodimerization in Cell Membranes. Removing the last 10 amino acids in the cytosolic domain of RAGE (PR-TM-CTYOP394) had minimal affect on RAGE dimerization, whereas removal of the central domain (PR-TM-CYTOA375) reduced dimerization, but not to background levels observed in the cytoplasmic domain deletion construct. Results shown are from three independent replicates and the error bars represent standard error.
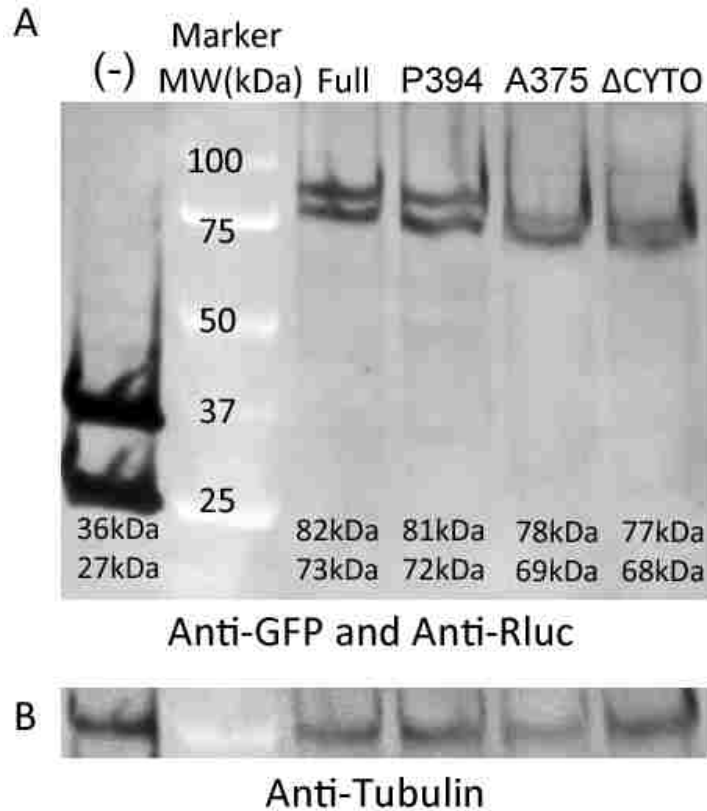
Figure 2. 10. Molecular weights for the expected products in each lane are provided at the bottom of the gel. A, HEK293 cells expressing various RAGE truncations fused to GFP were blotted against anti-GFP and anti-Rluc antibodies, and similar ratios of receptor expression are observed for both constructs. B, The same membrane probed with anti-GFP and anti-Rluc from A was blotted with an anti-tubulin antibody as a loading control to confirm similar levels of total sample per well.
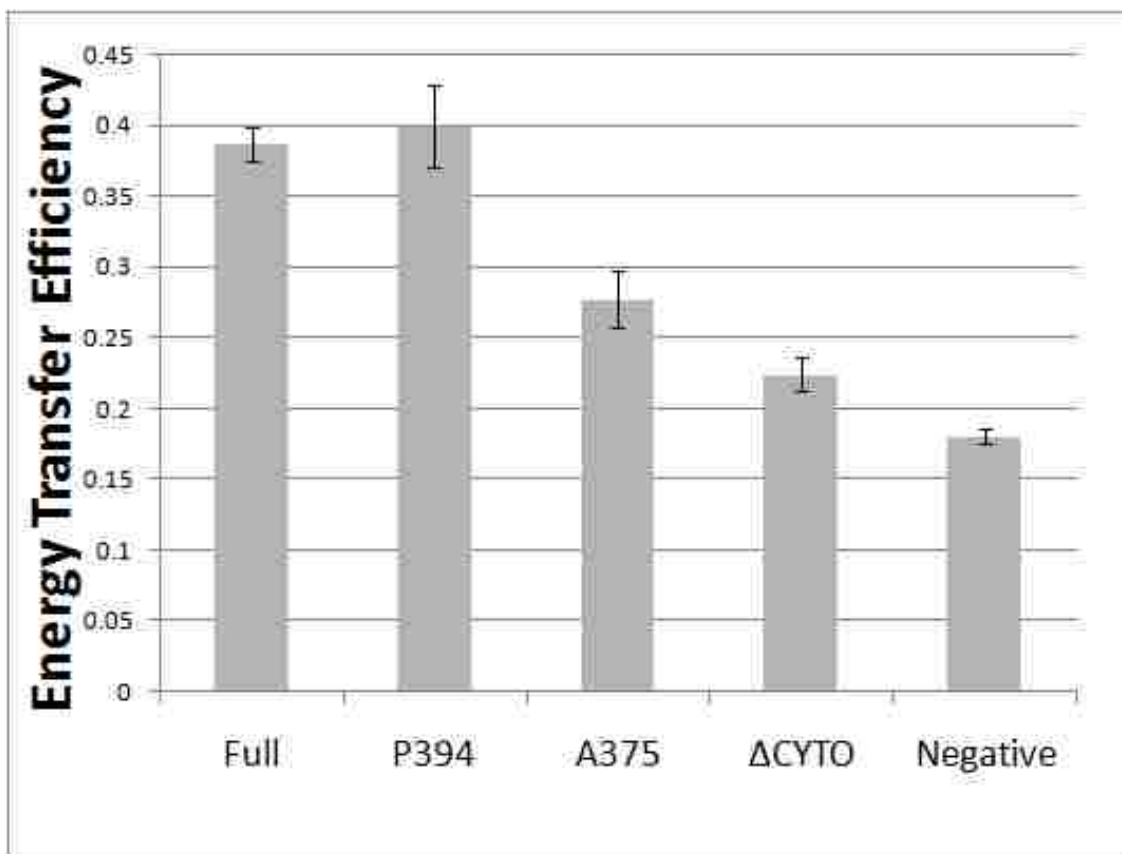
Figure 2. 11. RAGE Domain Interactions in Mammalian Membranes Correlate with AraTM Results, and Highlight the Importance of Cytoplasmic Domain Interactions in Ligand-independent Dimerization. Significant dimerization of RAGE is observed in the absence of the ligand. Removal of the distal C-terminal region (P394) has no impact on homodimerization, but removal of the central CYTO region (A375) causes a significant decrease in homodimerization. Experiments were repeated three times in triplicate, and error bars represent stand error of the mean (Full: full-length, P394: CYTO truncation at P394, A375: CYTO truncation at A375, ΔCYTO: CYTO truncation, and Negative: negative control).

2.4   Discussion

The AraTM method provides several advantages in investigating the importance of extracellular (EX), transmembrane (TM), and cytoplasmic (CYTO) domains in defining the dimeric state of TM receptors. In particular, the C-terminal orientation of the DNA-binding domain in AraC enables expression of chimera, including both

44

soluble and transmembrane domains, as a fusion to full-length MBP, including its native signal peptide, to direct membrane insertion. As mentioned previously, the ToxR transcription factor used in TOXCAT, POSSYCAT and TOXluc assays is a type II integral membrane protein containing an N-terminal DNA-binding domain, in which the ToxR TM domain functions as a signal peptide to direct trafficking to the bacterial inner membrane as well as membrane integration of ToxR (17, 18). Thus, modifications to ToxR-TM-MBP chimera that include both soluble EX or CYTO domains fused to TM domains, or specific mutations within a given TM domain, can interfere with membrane trafficking and lead to large differences in chimera expression levels between different constructs (36). In contrast, the C-terminal orientation of the DNA-binding domain in AraC enables expression of constructs in their native orientation, and utilizes the native signal peptide of MBP to direct insertion into the bacterial inner membrane (Figs. 2.1 and 2.2). Thus, the AraTM method decouples the dependence of a given TM construct on membrane trafficking from its dimerization in the bacterial inner membrane. This is particularly important given the large impact on extent of dimerization that can occur with cell-to-cell and construct-to-construct variability in ToxR-TM-MBP expression observed using TOXCAT, as well as the need to perform serial N- and C-terminal deletions of specific amino acids within a given TM domain to maximize membrane integration of expressed ToxR-TM-MBP chimera (12, 36). As illustrated for both integrin α_{IIb} (Fig. 2.5) and RAGE (Fig. 2.8), the AraTM system is robust to expression and proper membrane integration of a wide range of EX, TM, and CYTO constructs, including RAGE C2-PR-TM-CYTOfull (P224-P404, 21.3 kDa) and integrin α_{IIb} TM-CYTO (A963-E1008; 5.2 kDa), versus truncated TM domains such as

that used in TOXCAT for integrin α_{IIb} TM (W968-K989; 2.5 kDa) (26). Thus, the AraTM method (Fig. 2.2) enables analysis of dimerization in cell membranes of multi-domain receptor chimera that include EX, TM2 and CYTO, which is of particular use in investigating juxtamembrane interactions important in defining active and inactive states during signal transduction (6, 8–10).

The use of eGFP as a reporter gene in AraTM also allows for assays to be performed directly from cell culture using a microplate reader or flow cytometer, without a need for cell lysis, sample preparation or addition of exogenous substrates required for other common reporters such as CAT or luciferase. As mentioned before, cell-to-cell variability in TOXCAT chimera expression levels has a major impact on the observed dimerization signal (36), which we can measure directly from whole cells using flow cytometry to determine average dimerization per cell for a statistically significant sample size ($>10^6$ cells; Fig. 2.4). As illustrated by our measurements using integrin α_{IIb} TM-CYTO (A693-E1008), the AraTM method is sensitive to specific point mutations in the context of multidomain constructs, with a significant increase in homodimerization observed in the case of the TM L980A mutant relative to wt TM-CYTO (Fig. 2.4) (4, 26). Furthermore, the quantitative agreement between the relative increase in dimerization for wt integrin α_{IIb} TM-CYTO and L980A mutant measured using a microplate fluorescence plate reader (Fig. 2.3) and flow cytometer (Fig. 2.4), as well as the shift in total population mean observed between wt and L980A (Fig. 2.4), indicate the increase in observed GFP signal for L980A is due to a uniform increase in dimerization across the entire cell population rather than skew in the distribution caused by a specific cell subpopulation. Thus, the AraTM method is able to capture effects of

specific mutations on homodimerization in the context of TM-CYTO receptor fragments that are consistent with results from previously described in vitro and cell-based assays for integrin α_{IIb}.

In addition, our results point to a direct role for the RAGE CYTO in stabilizing the homooligomeric, unliganded form of the receptor. Previous studies have focused on the soluble, extracellular form of RAGE (sRAGE) containing the V-, C1-, and C2-domains as the primary driving force for homodimerization (31, 32). Purified sRAGE and C1-C2 domains assemble to form stable tetramers in vitro, and upon hexameric calgranulin binding, sRAGE undergoes significant conformational rearrangement in the ectodomain to form higher-order oligomers in the soluble, liganded state (32). Overexpression of full-length RAGE in transfected mammalian cells as mCFP and mYFP fusions result in significant co-localization and FRET transfer efficiency at the surface of transfected HEK293 cells, which supports the idea of RAGE exists in a homooligomeric state in the absence of ligand (35). Using AraTM, we find that the C2 domain does not play a key role in EX-TM-CYTO homodimerization, whereas the CYTO domain plays a key role in stabilizing the unliganded, homodimeric state of RAGE (Fig. 2.9). Specifically, deletion of the distal C-terminal region (PR-TM-CYTOP394; R314-P394) has no impact on dimerization, whereas deletion of the central cytoplasmic domain (PR-TM-CYTOA375; R314-A375) significantly reduces RAGE homodimerization, and the cytoplasmic domain deletion eliminated homodimerization. Previous work has indicated the extracellular V- and C1-domains as critical for dimerization of sRAGE in the absence of ligand (31, 32), whereas our BRET results for full-length RAGE indicate the strong self-association propensity of the CYTO domain

in stabilizing the resting state of the receptor, specifically the juxtamembrane region (A375-P394) within the CYTO important for stabilization of the unliganded, homodimeric state (Fig. 10). Recent NMR data investigating RAGE CYTO binding to the cytosolic signaling protein mDia1 suggests that the TM-proximal region is important for protein-protein interactions involving cytosolic signaling molecules, with the R365-P376 region of the CYTO forming a stable α-turn in solution (37). Our AraTM (Fig. 2.9) and BRET (Fig. 2.10) results indicate this region may also play a key role in homodimerization in the unliganded state. Interestingly, there are multiple, conserved families of primarily soluble proteins containing glutamic acid rich regions, in which the glutamic acid-rich region is found to have important roles in protein structure and complex formation. Specifically, in the case of the soluble protein SH3BGR, the C-terminal, glutamic acid-rich region is predicted to adopt coiled-coil like structures capable of mediating protein-protein interactions (38). Other studies have suggested glutamic acid-rich regions may act as low-affinity, calcium-binding domains important in regulating local calcium concentrations (39). Thus, the disordered, glutamic acid region of the CYTO observed previously by NMR may reflect one of several conformations for the RAGE CYTO during signal transduction (37). Furthermore, the overall correlation between the magnitude of homooligomerization measured using AraTM (Fig. 2.9) and BRET (Fig. 2.10) emphasizes the utility of the AraTM method to investigate the role of receptor dimerization in the biologically relevant context of cell membranes.

2.5   Conclusion

Overall, given the good agreement between our results for integrin α_{IIb} and previously published results investigating homodimerization *in vivo* and *in vitro*, as well as the correlation between our results investigating RAGE domain homodimerization in bacterial membranes and BRET signal in transfected mammalian cells, the AraTM provides a useful, complementary method to other mammalian cell-based measurement techniques such as BRET and bacterial 2-hybrid assays such as BACTH to rapidly assess the importance of specific interfaces and domains within transmembrane receptors in defining their dimeric states (33, 40).

2.6   References

(1) Klemm, J. D., Schreiber, S. L., and Crabtree, G. R. (1998) Dimerization as a regulatory mechanism in signal transduction. Annu. Rev. Immunol. 16, 569–592

(2) Moore, D. T., Berger, B. W., and DeGrado, W. F. (2008) Protein-protein interactions in the membrane: sequence, structural, and biological motifs. Structure 16, 991–1001

(3) Bennett, J. S. (2005) Structure and function of the platelet integrin $\alpha_{IIb}\beta_3$. J. Clin. Investig. 115, 3363–3369

(4) Li, W., Metcalf, D. G., Gorelik, R., Li, R., Mitra, N., Nanda, V., Law, P. B., Lear, J. D., Degrado, W. F., and Bennett, J. S. (2005) A push-pull mechanism for regulating integrin function. Proc. Natl. Acad. Sci. U.S.A. 102, 1424–1429

(5) Metcalf, D. G., Kulp, D. W., Bennett, J. S., and DeGrado, W. F. (2009) Multiple approaches converge on the structure of the integrin $\alpha_{IIb}\beta_3$ transmembrane heterodimer. J. Mol. Biol. 392, 1087–1101

(6) Li, R., Babu, C. R., Lear, J. D., Wand, A. J., Bennett, J. S., and DeGrado, W. F. (2001) Oligomerization of the integrin $\alpha_{IIb}\beta_3$: roles of the transmembrane and cytoplasmic domains. Proc. Natl. Acad. Sci. U.S.A. 98, 12462–12467

(7) Berger, B. W., Kulp, D. W., Span, L. M., DeGrado, J. L., Billings, P. C., Senes, A., Bennett, J. S., and DeGrado, W. F. (2010) Consensus motif for integrin transmembrane helix association. Proc. Natl. Acad. Sci. U.S.A. 107, 703–708

(8) Mo, X., Luo, S. Z., Lo´pez, J. A., and Li, R. (2008) Juxtamembrane basic residues in glycoprotein Ibβ cytoplasmic domain are required for assembly and surface expression of glycoprotein Ib-IX complex. FEBS Lett. 582, 3270–3274

(9) Mo, X., Luo, S. Z., Munday, A. D., Sun, W., Berndt, M. C., Lo´pez, J. A., Dong, J. F., and Li, R. (2008) The membrane-proximal intermolecular disulfide bonds in glycoprotein Ib influence receptor binding to von Willebrand factor. J. Thrombosis Haemostasis 6, 1789–1795

(10) Oates, J., King, G., and Dixon, A. M. (2010) Strong oligomerization behavior of PDGFβ receptor transmembrane domain and its regulation by the juxtamembrane regions. Biochim. Biophys. Acta 1798, 605–615

(11) Schneider, D., and Engelman, D. M. (2003) GALLEX, a measurement of heterologous association of transmembrane helices in a biological membrane. J. Biol. Chem. 278, 3105–3111

(12) Russ, W. P., and Engelman, D. M. (1999) TOXCAT: a measure of transmembrane

helix association in a biological membrane. Proc. Natl. Acad. Sci. U.S.A. 96, 863–868

(13) Gurezka, R., and Langosch, D. (2001) In vitro selection of membrane-spanning leucine zipper protein-protein interaction motifs using POSSYCCAT. J. Biol. Chem. 276, 45580–45587

(14) Roth, L., Nasarre, C., Dirrig-Grosch, S., Aunis, D., Cre´mel, G., Hubert, P., and Bagnard, D. (2008) Transmembrane domain interactions control biological functions of neuropilin-1. Mol. Biol. Cell 19, 646–654

(15) Ladant, D., and Karimova, G. (2000) Genetic systems for analyzing protein-protein interactions in bacteria. Res. Microbiol. 151, 711–720

(16) DiRita, V. J. (1992) Co-ordinate expression of virulence genes by ToxR in V. cholerae. Mol. Microbiol. 6, 451–458

(17) Ott, C. M., and Lingappa, V. R. (2002) Integral membrane protein biosynthesis: why topology is hard to predict. J. Cell Sci. 115, 2003–2009

(18) Kolmar, H., Hennecke, F., Go¨tze, K., Janzer, B., Vogt, B., Mayer, F., and Fritz, H. J. (1995) Membrane insertion of the bacterial signal transduction protein ToxR and requirements of transcription activation studied by modular replacement of different protein substructures. EMBO J. 14, 3895–3904

(19) Bustos, S. A., and Schleif, R. F. (1993) Functional domains of the AraC protein. Proc. Natl. Acad. Sci. U.S.A. 90, 5638–5642

(20) Siegele, D. A., and Hu, J. C. (1997) Gene expression from plasmids containing the *araBAD* promoter at subsaturating inducer concentrations represents mixed populations. Proc. Natl. Acad. Sci. U.S.A. 94, 8168–8172

(21) Bonde, M. M., Hansen, J. T., Sanni, S. J., Hauns, S., Gammeltoft, S., Lyngs, C.,

and Hansen, J. L. (2010) Biased signaling of the angiotensin II type 1 receptor can be mediated through distinct mechanisms. PLoS ONE 5

(22) Schleif, R. (2003) AraC protein: a love-hate relationship. BioEssays 25, 274–282

(23) Weldon, J. E., and Schleif, R. F. (2006) Specific interactions by the N-terminal arm inhibit self-association of the AraC dimerization domain. Protein Science 15, 2828–2835

(24) Eustance, R., and Schleif, R. (1996) In vivo association of protein fragments giving active AraC. Proteins-Structure Function Bioinformatics 25, 501–505

(25) Nikaido, H. (1994) Maltose transport system of *E. coli*: an ABC-type transporter. FEBS Lett. 346, 55–58 26. Li, R., Gorelik, R., Nanda, V., Law, P. B., Lear, J. D., DeGrado, W. F., and Bennett, J. S. (2004) Dimerization of the transmembrane domain of Integrin α_{IIb} subunit in cell membranes. J. Biol. Chem. 279, 26666–26673

(27) Schneider, D., and Engelman, D. (2004) Involvement of transmembrane domain interactions in signal transduction by α/β integrins. J. Biol. Chem. 279, 9840–9846

(28) Schmidt, A. M., Yan, S. D., Yan, S. F., and Stern, D. M. (2001) The multiligand receptor RAGE as a progression factor amplifying immune and inflammatory responses. J. Clin. Investig. 108, 949–955

(29) Hofmann, M. A., Drury, S., Fu, C., Qu, W., Taguchi, A., Lu, Y., Avila, C., Kambham, N., Bierhaus, A., Nawroth, P., Neurath, M. F., Slattery, T., Beach, D., McClary, J., Nagashima, M., Morser, J., Stern, D., and Schmidt, A. M. (1999) RAGE mediates a novel proinflammatory axis: a central cell surface receptor for S100/calgranulin polypeptides. Cell 97, 889–901

(30) Ostendorp, T., Leclerc, E., Galichet, A., Koch, M., Demling, N., Weigle, B.,

Heizmann, C. W., Kroneck, P. M., and Fritz, G. (2007) Structural and functional insights into RAGE activation by multimeric S100B. EMBO J. 26, 3868–3878

(31) Zong, H., Madden, A., Ward, M., Mooney, M. H., Elliott, C. T., and Stitt, A. W. (2010) Homodimerization is essential for the receptor for advanced glycation end products (RAGE)-mediated signal transduction. J. Biol. Chem. 285, 23137

(32) Xie, J., Burz, D. S., He, W., Bronstein, I. B., Lednev, I., and Shekhtman, A. (2007) Hexameric calgranulin C (S100A12) binds to the receptor for advanced glycated end products (RAGE) using symmetric hydrophobic target- binding patches. J. Biol. Chem. 282, 4218–4231

(33) Ramsay, D., Kellett, E., McVey, M., Rees, S., and Milligan, G. (2002) Homoand hetero-oligomeric interactions between G-protein-coupled receptors in living cells monitored by two variants of bioluminescence resonance energy transfer (BRET): hetero-oligomers between receptor subtypes form more efficiently than between less closely related sequences.

(34) Buensuceso, C. (2003) Detection of integrin $\alpha_{IIb}\beta_3$ clustering in living cells. J Biol. Chem. 278, 15217–15224

(35) Xie, J., Reverdatto, S., Frolov, A., Hoffmann, R., Burz, D. S., and Shekhtman, A. (2008) Structural basis for pattern recognition by the receptor for advanced glycation end products (RAGE). J. Biol. Chem. 283, 27255–27269

(36) Duong, M. T., Jaszewski, T. M., Fleming, K. G., and MacKenzie, K. R. (2007) Changes in apparent free energy of helix-helix dimerization in a biological membrane due to point mutations. J. Mol. Biol. 371, 422– 434

(37) Rai, V., Maldonado, A. Y., Burz, D. S., Reverdatto, S., Yan, S. F., Schmidt, A. M.,

and Shekhtman, A. (2012) Signal transduction in receptor for advanced glycation end products (RAGE): solution structure of C-terminal rage (ctRAGE) and its binding to mDia1. J. Biol. Chem. 287, 5133–5144

(38) Scartezzini, P., Egeo, A., Colella, S., Fumagalli, P., Arrigo, P., Nizetic, D., Taramelli, R., and Rasore-Quartino, A. (1997) Cloning a new human gene from chromosome 21q22.3 encoding a glutamic acid-rich protein expressed in heart and skeletal muscle. Human Genetics 99, 387–392

(39) Haber-Pohlmeier, S., Abarca-Heidemann, K., Ko¨rschen, H. G., Dhiman, H. K., Heberle, J., Schwalbe, H., Klein-Seetharaman, J., Kaupp, U. B., and Pohlmeier, A. (2007) Binding of Ca2_ to glutamic acid-rich polypeptides from the rod outer segment. Biophys. J. 92, 3207–3214

(40) Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proc. Natl. Acad. Sci. U.S.A. 95, 5752–5756

Chapter 3

A NOVEL ASSAY FOR ASSESSING JUXTAMEMBRANE AND TRANSMEMBRANE DOMAIN INTERACTIONS IMPORTANT FOR RECEPTOR HETERODIMERIZATION

Understanding the basis of specificity in receptor homodimerization versus heterodimerization is essential in determining the role receptor plays in signal transduction. Specificity in each of the interfaces formed during signal transduction involves cooperative interactions between receptor extracellular, transmembrane (TM), and cytoplasmic domains. While methods exist for studying receptor heterodimerization in cell membranes, they are limited to either TM domains expressed in an inverted orientation or capture only heterodimerization in a single assay. To address this limitation, we have developed an assay (DN-AraTM) that enables simultaneous measurement of homodimerization and heterodimerization of type I receptor domains in their native orientation, including both soluble and TM domains. Using integrin α_{IIb} and RAGE (receptor for a dvanced glycation endproducts) as model type I receptor systems, we demonstrate both specificity and sensitivity of our approach, which will provide a novel tool to identify specific domain interactions that are important in regulating signal transduction.

3.1   Introduction

Receptor homodimerization and heterodimerization is a key mechanism that regulates transmembrane (TM) signal transduction (1). Several classes of type I TM receptors, including neuropilins and integrins, all exist in a regulated equilibrium between homodimeric and heterodimeric states, with each state defining the inactive or activated forms of the co-receptor complex (2,3). Furthermore, for both integrins and neuropilins, specific, cooperative interactions involving the TM, cytosolic [cytoplasmic (CYTO)], and juxtamembrane [extracellular (EX)] domains are critical in stabilizing the homodimeric and the heterodimeric states of the co-receptor complex. In the case of integrin $\alpha_{IIb}\beta_3$, the inactive, heterodimeric state is stabilized through a TM reciprocal " large-small" helical packing interface and a cytosolic clasp formed by an Arg-Asp salt bridge (4,5). Disruption of the cytosolic clasp by intracellular binding of talin to β_3 or mutations to residues in the TM heterodimer interface triggers a series of conformational changes that lead to integrin activation and fibrinogen binding (6,7). Within the TM domain, activation can involve mutations to integrin α_{IIb} or β_3 that promote preferential TM homodimer formation, as well as mutations to either α_{IIb} or β_3 that destabilize the heterodimeric interface (3). Thus, characterizing the interplay between specific, homodimeric and heterodimeric interactions involving TM and juxtamembrane (CYTO and EX) domains is essential in understanding the basis for TM signal transduction.

Among the methods available to characterize type I receptor heterodimerization in cell membranes, *Escherichia coli*-based transcription factor assays GALLEX and

DN-ToxRed have the advantage of straightforward implementation, high sensitivity, and rapid quantification of homodimerization and heterodimerization in terms of reporter gene expression levels such as red fluorescent protein (4,8,9). In the case of TOXCAT, extension of this method to polytopic membrane proteins has also been described (10). However, both assays rely on transcription factors (LexA for GALLEX and ToxR for DN-ToxRed) that have a type II orientation, in which the N-terminal DNA-binding domain is expressed as a fusion to a C-terminal TM domain of interest. Thus, the C-terminal TM domain acts as a signal peptide to direct membrane trafficking and integration, and any modification of the TM domain, including addition of CYTO and EX domains, can prevent proper membrane integration. As a result, bacterial heterodimerization assays GALLEX and DN-ToxRed are limited to analysis of only isolated TM domains. Likewise, while split-enzyme assays such as BACTH based on reconstitution of split adenylyl cyclase are effective in determining heterodimerization of a given receptor domain, they do not capture the simultaneous competition between homodimerization and heterodimerization often observed for full-length receptors during signal transduction (11).

To address the limitation of current bacterial heterodimerization assays, we developed a dominant-negative assay (DN-AraTM) to investigate type I receptor heterodimerization (Fig. 3.1). In DN-AraTM, a chimera containing wild-type AraC is co-expressed with a chimera that contains a non-functional, mutant AraC (AraC*) that is unable to activate gene transcription. Preferential heterodimerization (AraC-AraC*) reduces the level of green fluorescent protein (GFP) reporter gene transcription (under control of the *araBAD* promoter) relative to homodimerization (AraC-AraC), thereby

57

giving a quantitative measure of heterodimerization relative to homodimerization that can be measured directly in whole cells. As with AraTM, each chimera contains an N-terminal maltose-binding protein (MBP) fusion that includes a signal peptide to direct membrane insertion, thereby enabling expression of receptor domains that include EX, TM, and CYTO domains (12). To confirm that both chimera are expressed in cell membranes, we introduce a unique HA or myc epitope into each chimera, allowing for quantification of expression levels by Western blotting using either unique HA or myc epitopes or MBP. We illustrate the effectiveness of DN-AraTM using integrin α_{IIb} and RAGE (receptor for a dvanced glycation endproducts) as model type I receptor systems. Collectively, our results indicate that DN-AraTM is robust to systems that exhibit preferential homodimerization and heterodimerization and provides a unique method to investigate the structural basis for domain interactions important in the regulation of type I receptor signal transduction.
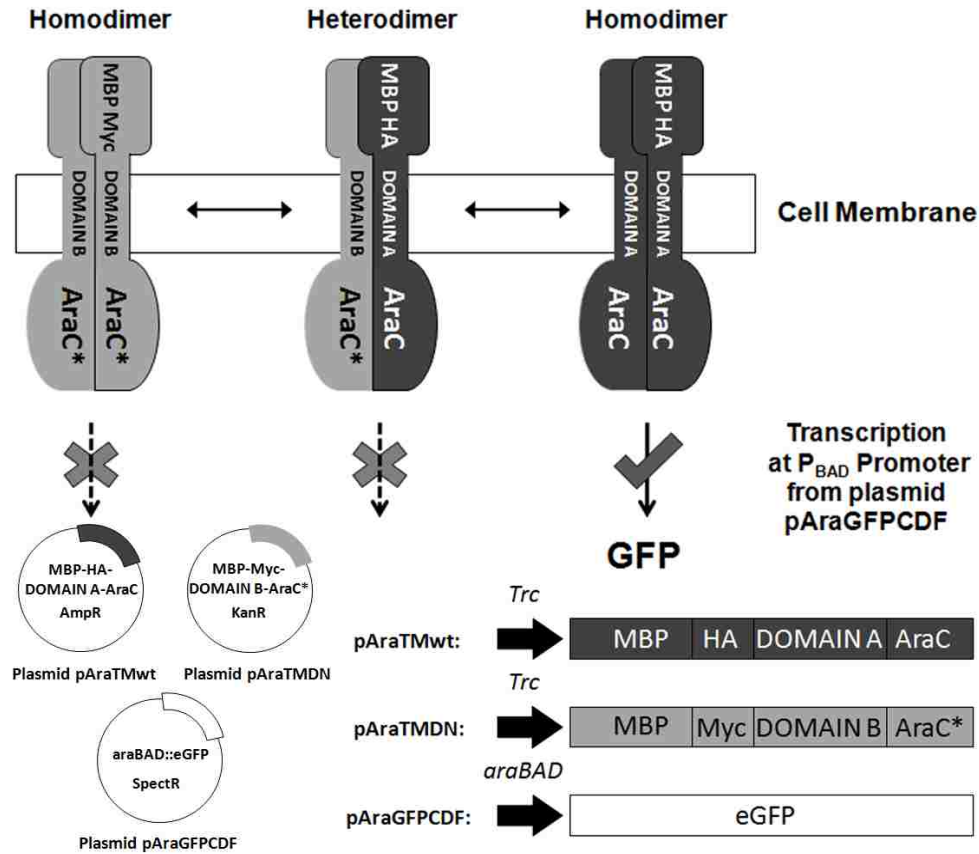
Figure 3. 1. Overview of DN-AraTM Assay. Chimera containing N-terminal MBP fused to either an in-frame receptor A fragment (DOMAIN A) and C-terminal AraC or an in-frame receptor B fragment (DOMAIN B) and C-terminal disabled AraC unable to activate transcription at the *araBAD* promoter (AraC*) are expressed by the regulator plasmids (pAraTMwt, ampicillin resistant; pAraTMDN, kanamycin resistant). Once expressed, MBP directs integration of chimera in the inner membrane of *E. coli*. Homodimerization of receptor A (AraC-AraC) brings the AraC transcription factors in close proximity and activates the *araBAD* promoter to produce GFP. If receptor A has a higher affinity to heterodimerize with receptor B versus homodimerize with receptor A, a reduction in GFP will be observed (AraC-AraC*) due to the inability of heterodimers containing receptor B-fused AraC* to activate transcription at the *araBAD* promoter.

3.2 Materials and Methods

3.2.1 Subcloning

Unless otherwise stated, standard molecular biology techniques were used. All constructs used were verified by DNA sequencing. An overview of the DN-AraTM assay is given in Fig. 3.1. Plasmid pAraTM was generated as described previously (12). Plasmid pAraTMmyc was generated by amplifying the MBP sequence from the pAraTM plasmid with primers that introduce a C-terminal myc-tag sequence (EQKLISEEDL). The resulting PCR product was subcloned into the pAraTM plasmid as an NcoI/SacI fragment.

pAraTMDN plasmid was generated by making a point mutation R210A in the truncated AraCamino acid sequence (amino acids 168–293) in the pAraTMmyc plasmid using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies) to generate a dominant-negative form of AraC (AraC*). The truncated AraC amino acid sequence (amino acids 168–293) lacks its N-terminal dimerization domain, which is instead replaced with the receptor domain of interested to promote AraC dimerization. The pAraTMDN plasmid was used for all subsequent experiments, with RAGE and integrin αIIb domains cloned in-frame as fusions with MBP and AraC* as SacI/KpnI fragments.

Plasmid pAraTMwt was generated by amplifying the MBP-HA-MCS-AraC (amino acids 168–293) from the pAraTM plasmid and subcloning into pTrc99a as an NcoI/HindIII fragment. The pAraTMwt plasmid was used for all subsequent experiments, with RAGE and integrin α_{IIb} domains cloned in-frame as fusions with

MBP and AraC as SacI/KpnI fragments. Note that the LacI gene on the pAraTMwt allows IPTG-regulated expression of both AraC and AraC* chimera; for all experiments, 1 mM IPTG was added to culture.

The reporter plasmid pAraGFPCDF was derived from plasmid pAraGFP by amplifying a fragment containing the *araBAD* promoter and GFP coding sequence and subcloning into a pCDFDuet-1 plasmid (Novagen) as an EcoRI/PstI fragment.

3.2.2   DN-AraTM assay

Plasmids pAraTMDN, pAraTMwt, and pAraGFPCDF were co-transformed into the AraC-deficient *E. coli* strain SB1676 (The *E. coli* Genetic Stock Center at Yale University) and streaked on selective LB plates (100 μg/mL ampicillin, 50 μg/mL kanamycin, and 100 μg/mL spectinomycin). One colony was picked from each construct and grown in 4 mL selective LB media for 16 h at 37 °C and 200 rpm. After 16 h, the cultures were diluted to A560 of 0.5 in three wells on a 2.0-mL-deep, 96-well PP plate (PlateOne) with each well containing 400 μL of selective LB media (100 μg/mL ampicillin, 50 μg/mL kanamycin, 100 μg/mL spectinomycin, and 1 mM IPTG) and grown for an additional 6 h at 37 °C and 300 rpm. We transferred 200 μL of cultures from each sample to a black 96-well, clear bottom plate (Greiner) and a series of 2-fold dilutions was prepared using selective LB media. A560 measurements and GFP fluorescence emissions spectra (excitation maximumat 485 nm and emissions maximum at 530 nm) were collected using an M200 Infinity Plate Reader (Tecan). Results are reported as the ratio of fluorescence emission at 530 nm to absorbance at

560 nm and normalized to the negative control (pAraGFPCDF, pTrc99a, and pTrcRSF transformed cells), where the results were first divided by the negative control ratio and rescaled so that the negative control is equal to zero.

3.2.3 Proper integration of AraC chimera into the inner membrane of E. coli and immunoblotting

The maltose complementation test was carried out as described previously to demonstrate the proper integration of AraC chimera into the inner membrane of *E. coli* (12). Immunoblotting was carried out as described previously to demonstrate proper expression of each full-length AraC chimera using anti-myc (Cell Signaling) and anti-HA (Cell Signaling) monoclonal antibodies (12).

3.3   Results

3.3.1 Design of DN-araTM heterodimer assay

The DN-AraTM assay is a dominant-negative system based on the *E. coli* transcription factor AraC, which is active at the *araBAD* promoter as a homodimer (13). The overall process is illustrated in Fig. 3.1, where a wild-type AraC chimera (AraC) is co-expressed with a AraC chimera in which a point mutation in the DNA-binding domain of AraC (R210A) was introduced to inactivate transcription (AraC*). We screened a series of previously described inactivating mutations in the DNA-binding domain of AraC (H203A, R210A, and R216A) and found that, of these, R210A

exhibited reproducibly minimumactivity at the *araBAD* promoter when expressed as part of a TM domain-containing chimera (Fig. 3.2) (14). As with AraTM, both chimera also contain an N-terminal MBP fusion with signal peptide to direct integration into the inner *E. coli* membrane, thereby enabling expression of type I receptors including EX, TM, and CYTO domains in their native orientation (12). When chimera containing AraC and AraC* are each co-expressed from unique plasmids (pAraTMwt and pAraTMDN) and properly integrated into the inner membrane, the AraC*-containing chimera acts as a dominant negative (DN) to inhibit the homodimerization of the AraC-containing chimera, thereby reducing the transcriptional activity at the *araBAD* promoter due to formation of AraC-AraC* heterodimeric complexes. In this way, the relative decrease in expression of the reporter gene GFP for a heterodimeric competitor involving two unique receptor sequences (A-B*) relative to a homodimeric sequence (A-A) provides a quantitative measure of preferential receptor heterodimerization. If the reduction in signal for the heterodimer (A-B*) is more than that for the respective homodimer (A-A), then the two receptor domains (A and B) form a preferential heterodimer, with the strength of the relative heterodimer quantified in terms of difference in signal between A-A and A-B*.

A

B

WT R210A

← 64kDa
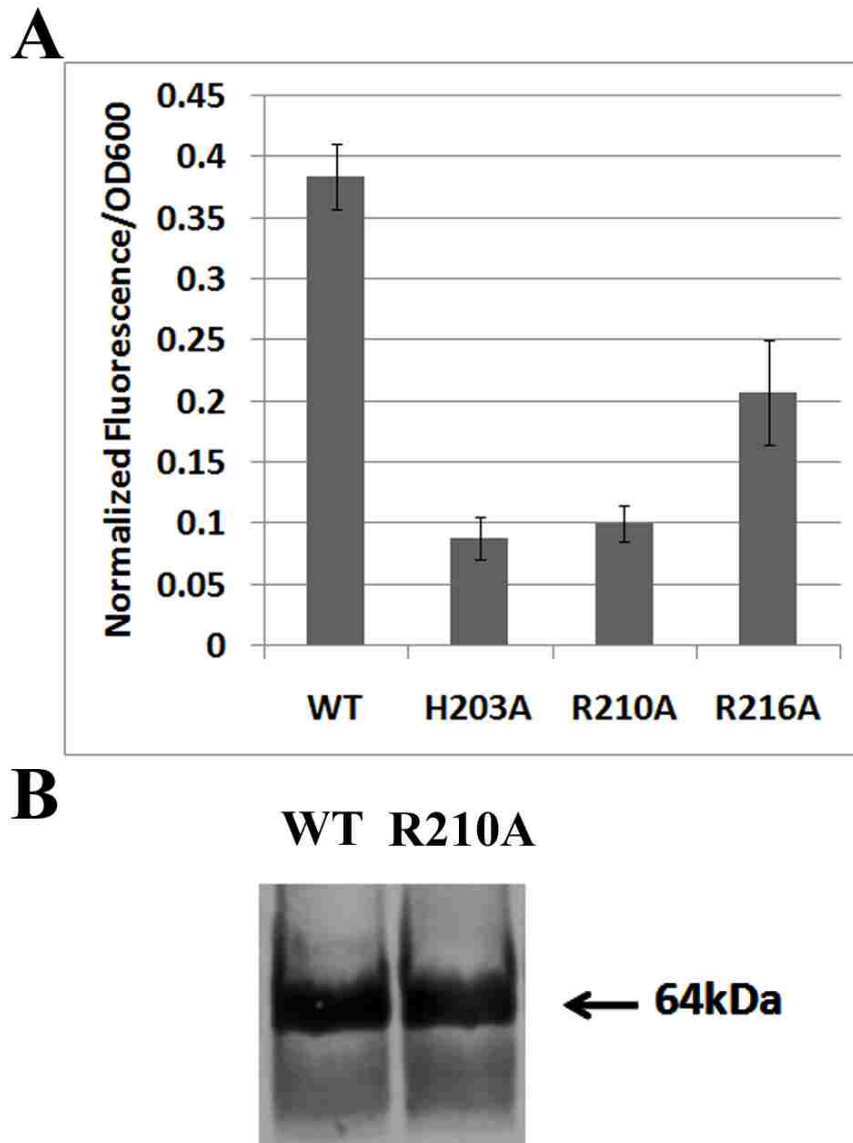
Figure 3. 2. Dominant-negative Form of AraC (R210A) Disrupts Its Ability to Turn on *araBAD* Promoter. A, The R210A mutation in the DNA binding domain of AraC eliminates the ability of chimera containing wild-type integrin α_{IIb} TM-CYTO AraC to activate GFP transcription at the *araBAD* promoter. Experiments were repeated three times in triplicate, and error bars represent standard deviation. B, The expression level of AraC and AraC* chimera containing wild-type integrin αIIb TM-CYTO are similar as shown by immunoblotting with anti-MBP antibody from whole-cell lysates (WT: wild-type AraC chimera, H203A: H203A mutant AraC chimera, R210A: R210A mutant AraC chimera, and R216A: R216A mutant AraC chimera).

3.3.2   Multiple epitope tags to confirm heterodimer expression

Unlike AraTM, in which a single receptor is expressed to measure homodimer formation, in DN-AraTM, two independent receptor domains are co-expressed, with the relative competition between homodimer and heterodimer formation measured (Fig. 3.1). Thus, it is important to ensure that expression levels of both AraC and competitor AraC* fusions are similar in order to have an accurate measure of homodimer versus heterodimer formation. To address this issue, we replaced the HA-tag on the pAraTMDN plasmid with a myc-tag; thus, AraC chimera expressed by the pAraTMwt can be detected with an anti-HA antibody and AraC* chimera expressed by pAraTMDN can be detected with an anti-myc antibody. Additionally, we can utilize an anti-MBP antibody to compare expression levels for co-expression of both AraC and AraC* chimera versus expression of either AraC or AraC* chimera alone.

First, to confirm that the epitope used has no effect on chimera orientation or expression in the *E. coli* inner membrane, we replaced the HA-tag on the pAraTM plasmid by a myc-tag to make pAraTMmyc plasmid and we compared GFP reporter gene expression from integrin α_{IIb} L980A TM-CYTO for both chimera; this construct was used previously as a model homodimer for AraTM (12). After expression in *E. coli*, both constructs could activate transcription at the *araBAD* promoter, resulting in similar GFP reporter gene expression levels (Fig. 3.3B), indicating that the myc-tag did not significantly alter AraC transcriptional activity relative to the HA-tag used previously in AraTM (12). Both chimera expressed by pAraTMmyc and pAraTM can properly integrate into the *E. coli* inner membrane as demonstrated by growth on maltose

minimal media plates (Fig. 3.3C). Moreover, immunoblotting whole-cell lysates from integrin α_{IIb} L980A TM-CYTO expressed from pAra-TMmyc and pAraTM with anti-myc and anti-HA, respectively, indicate that both fusion tags were expressed and can be recognized by the antibodies. Independent of HA and myc epitopes, immunoblotting whole-cell lysates against anti-MBP antibody confirmed both constructs expressed at similar levels (Fig. 3.3D). Thus, we have two independent methods to confirm expression of AraC- and AraC*-containing chimera: (1) intensity of total anti-MBP signal for total AraC and AraC* chimera and (2) comparison of anti-myc and anti-HA signal specific to either AraC or AraC* chimera.
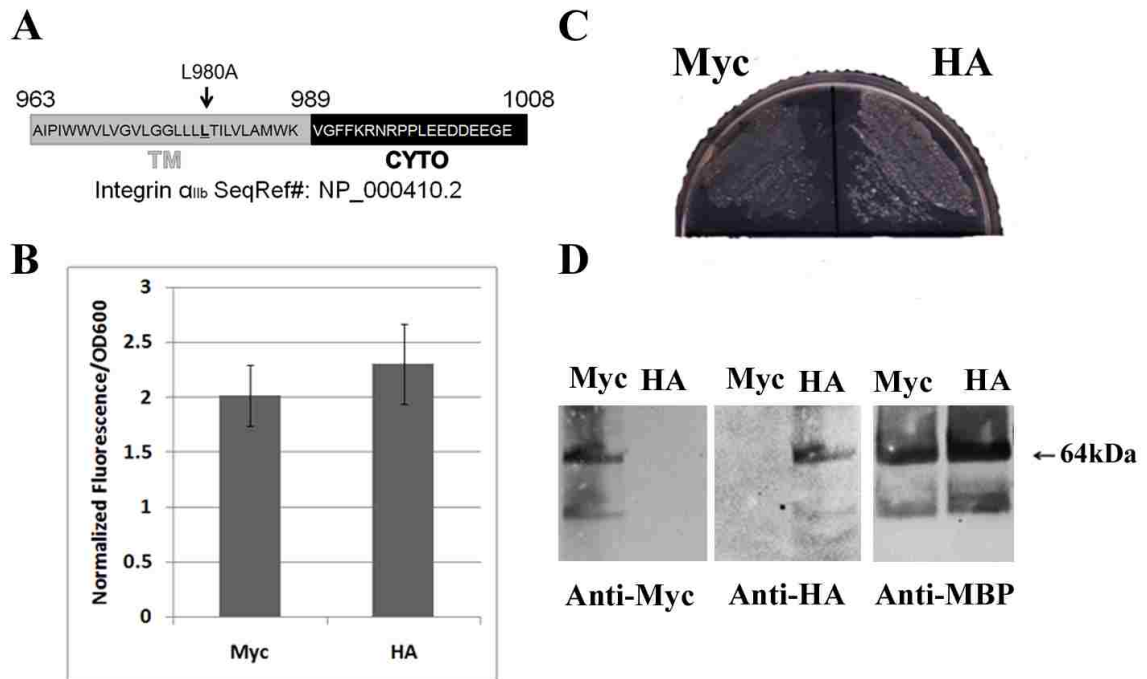
Figure 3. 3. AraTM Chimera Containing Either an HA-tag and a Myc-tag Are Expressed and Integrated Correctly into *E. coli* Inner Membrane, and Both Chimera Retain Proper AraC Transcriptional Activity. A, Amino acid sequence of integrin α_{IIb} TM-CYTO and the position of L980A mutation. B, AraTM chimera with the homodimeric integrin αIIb L980A TM-CYTO construct that contain either an HA-tag or a myc-tag exhibit comparable AraC transcriptional activity. C, AraTM chimera with the homodimeric integrin αIIb L980A TM-CYTO construct that contain either an HA-tag or a myc-tag were expressed from plasmid pAraTM or pAraTMmyc in MBP-deficient MM39 cells, streaked onto a 0.4% maltose M9 plate and incubated for 40 h at 37 °C. Both HA-tagged and myc-tagged constructs are properly integrated into the inner membrane of *E. coli* as evidenced by robust growth similar to the positive control (pTrcRSF containing MBP-AraC chimera). As expected, no growth is observed on the negative control (AraCY, which expresses MBP without the N-terminal signal peptide sequence fused to AraC protein). D, Immunoblotting of whole-cell lysate with anti-myc, anti-HA, and anti-MBP antibodies indicate that both chimera are expressed at their expected molecular masses (64 kDa).

3.3.3 Heterodimerization of integrin α_{IIb} and RAGE TM-CYTO domains

To confirm DN-AraTM is sensitive to heterodimer formation, we used the TM-CYTO of human integrin α_{IIb} (residues A963-E1008) and its dimer-forming TM mutant

L980A as a model system (15). Previous studies have shown that human integrin α_{IIb} TM domain forms a homodimer through a conserved GX_3G motif, and the L980A mutant enhances homodimer formation by placing an additional small residue, alanine, on the same face of the TM helix as the conserved GX_3G (16). When α_{IIb} L980A is expressed as a fusion to wild-type AraC, a strong homodimer signal is observed, with a GFP signal 4-fold higher than the negative control, which expresses the reporter plasmid pAraGFPCDF with empty pTrc99a and pTrcRSF (Fig. 3.4). Likewise, when α_{IIb} L980A-AraC* is co-expressed as a competitor to α_{IIb} L980A-AraC, a significant decrease in GFP expression is observed, again consistent with the strong homodimer forming tendency of α_{IIb} L980A (Fig. 3.4). However, when α_{IIb} wild-type AraC* is co-expressed as a competitor to α_{IIb} L980A-AraC, no significant decrease in GFP expression is observed relative to the α_{IIb} L980A-AraC homodimer, again consistent with the expected result that α_{IIb} L980A forms a strong, preferential homodimer (large decrease in GFP signal in DN-AraTM) versus an α_{IIb} L980A-α_{IIb} wild-type heterodimer (no decrease in GFP signal in DN-AraTM). Integrin α_{IIb} wild-type and α_{IIb} L980A TM-CYTO constructs were both expressed and properly integrated in the *E. coli* membrane, as evidenced by growth on maltose M9 minimal plates (Fig. 3.5) for both AraC- and AraC*-containing chimera. When expressed individually, pAraTMwt and pAraTMDN chimera had similar immunoblotting band intensities (Fig. 3.5), and the intensity doubled when both chimera were co-expressed (Fig. 3.5), indicating that both chimera are expressed at similar levels.

Figure 3. 4. Quantifying Receptor Homodimerization and Heterodimerization using the DN-AraTM Assay. Significant heterodimerization is observed for AraC*-integrin α_{IIb}L980A TM-CYTO as competitor to AraC-integrin αIIb L980A TM-CYTO, but not with AraC*-wild-type integrin α_{IIb} TM-CYTO as competitor. The sequences for integrin α_{IIb} wild-type and L980A TM-CYTO are given for reference. Thus, the strong AraC-AraC* heterodimer formed from integrin α_{IIb} L980A TM-CYTO gives rise to a reduction in overall GFP expression indicative of heterodimerization, whereas the strong AraC-AraC homodimer formed from integrin α_{IIb} L980A TM-CYTO is not destabilized by the AraC*-integrin α_{IIb} wild-type TM-CYTO, therefore indicative of preferential homodimerization by retaining high GFP expression. Each measurement is from three independent replicates, with error bars representing standard deviation.

69

A

Construct	1	2	3	4	5
pAraTMwt Insert: α_{IIb} L980A TM-CYTO	X	X	X	-	-
pAraTMDN Insert: α_{IIb} L980A TM-CYTO	-	X	-	X	-
pAraTMDN Insert: α_{IIb} wt TM-CYTO	-	-	X	-	X
pTrc99a	-	-	-	X	X
pTrcRSF	X	-	-	-	-
pAraGFPCDF	X	X	X	X	X

Figure 3. 5. AraC* chimera were properly expressed and integrated into the bacterial inner membrane. A, Summary of plasmid combinations used in co-expression experiments. B, Immunoblotting of whole cell lysate from each sample with anti-MBP antibody indicates proper expression of homo- and heterodimeric chimera, with the intensity of the bands for co-expression of both AraC- and AraC*-containing chimera twice that of the corresponding bands for single chimera expression. C, Chimera containing AraC* fused to wild-type and L980A integrin α_{IIb} TM-CYTO were expressed in MalE-deficient MM39 cells, streaked on a 0.4% maltose M9 plate and incubated for 40 h at 37°C. Each construct was properly integrated into the inner membrane of *E. coli*, as indicated by robust growth on the 0.4% maltose M9 plate similar to the positive control (pAraTM containing integrin αIIb TM-CYTO chimera fused to AraC and MBP-HA) As expected, no growth is observed on the negative control (AraCY, which expresses MBP without the N-terminal signal peptide sequence fused to AraC protein).

In addition to integrin α_{IIb} TM-CYTO, we also investigated whether co-expression of RAGE TM-CYTO as an AraC* fusion could inhibit RAGE homodimer formation (Fig. 3.6). RAGE is a type I TM receptor whose dimeric and higher-order oligomeric state is important in determining the specificity of EX ligand binding (17). In particular, RAGE exists in an unliganded state as a dimer, and the EX domain forms tetramers and octamers when bound to the ligands S100A12 and S100B (18,19). Our previous work indicates that the unliganded RAGE homodimer also depends on the cytosolic region of the receptor, specifically the juxtamembrane region enriched in glutamic acids to promote homodimer formation (12). Thus, the RAGE TM-CYTO provides a useful, complementary system to integrin α_{IIb} to assess the sensitivity of the DN-AraTM assay. When RAGE TM-CYTO-AraC* is co-expressed as a competitor to RAGE TM-CYTO-AraC, a significant decrease in GFP expression is observed relative to the corresponding RAGE TM-CYTO-AraC homodimer (Fig. 3.6). This further demonstrates the ability of DN-AraTM assay to assess receptor domain heterodimerization.

71

A

Construct	A	B	C	(+)
pAraTMwt Insert: RAGE PR-TM-CYTOfull	X	X	-	-
pAraTMDN Insert: RAGE PR-TM-CYTOfull	-	X	X	-
pTrcRSF	X	-	-	-
pAraGFPCDF	X	X	-	-
pAraTM Insert: α_{IIb} L980A TM-CYTO	-	-	-	X

B — Normalized Fluorescence/OD600

C — ← 68kDa

D — (+) / C
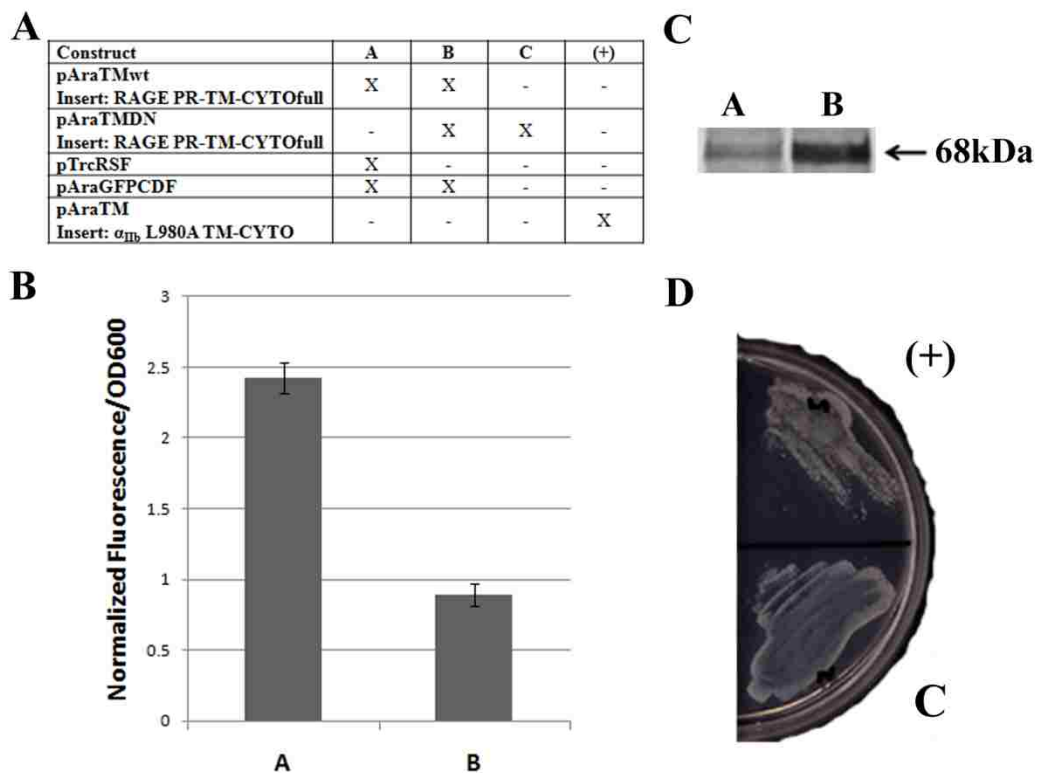
Figure 3. 6. Effectiveness of Heterodimeric AraTM Assay was Demonstrated by Competing Wild-type and R210A AraC Fused RAGE PR-TM-CYTOfull Chimera. A, Summary of plasmid combinations used in co-expression experiments. B, The average slopes from 2 fluorescence intensity versus OD_{600} for each construct are compared from three independent replicates, with error bars representing standard deviation. When chimera containing AraC fused to RAGE PR-TM-CYTO are co-expressed with AraC* chimera containing RAGE PR-TM-CYTO, a significant decrease in GFP was observed as a result of strong homo- and heterodimerization, consistent with previous results indicating RAGE PR-TM-CYTO forms a strong homodimer (11). C, Immunoblotting of whole cell lysate from each sample with anti-MBP antibody confirm co-expression of chimera, as indicated by the increase in intensity of the anti-MBP signal for AraC- and AraC* heterodimer co-expression relative to the AraC homodimer alone sample. D, AraTM chimera containing AraC* fused to RAGE PR-TM-CYTO were expressed in MalE-deficient MM39 cells, streaked on a 0.4% maltose M9 plate and incubated for 40 h at 37°C. AraC* fused RAGE PR-TM-CYTO chimera is properly integrated into the inner membrane of *E. coli*, as indicated by robust growth on the 0.4% maltose M9 plate. As expected, no growth is observed on the negative control (AraCY, which expresses MBP without the Nterminal signal peptide sequence fused to AraC protein).

72

3.4 Discussion

The DN-AraTM assay provides several advantages in investigating the importance of EX, TM, and CYTO domains in defining the heterodimeric state of type I TM receptors. First, the C-terminal orientation of the DNA-binding domain in AraC enables expression of receptor chimera, including both soluble (EX and CYTO) and TM domains, as a fusion to full-length MBP including its native signal peptide. The N-terminal MBP with signal peptide enables expression of type I receptors in their native orientation and proper trafficking to the bacterial inner membrane, unlike ToxR-based (DN-ToxRed) or LexA-based (GALLEX) transcriptional assays in which a type II orientation is required (4,8,9). For type II orientations, the TM domain acts both as the dimerization domain of interest and as the signal peptide to direct protein trafficking to the membrane (20). Thus, modifications to the TM domain of interest, including addition of EX or CYTO domains, can impair membrane trafficking for both DN-ToxRed and GALLEX, unlike the DN-AraTM assay. Given the importance of cooperative interactions between TM and soluble (EX and CYTO) regions in stabilizing receptor heterodimers such as integrins, this is a significant advantage when understanding the underlying structural basis for heterodimerization (5,6,15,16).

In addition, the DN-AraTM assay provides a method to quantify the equilibrium between homodimerization and heterodimerization. Most receptors exist in a regulated equilibrium between homodimeric and heterodimeric states, with each corresponding to the biologically relevant resting or activated state during signal transduction (1). The DN-AraTM assay allows monitoring of the homodimeric state of a given receptor as an

73

AraC fusion, and through competition with a non-functional AraC fusion (AraC*), the relative decrease in homodimerization due to heterodimer formation can be compared (Fig. 3.1). Since the DN-AraTM assay depends on the reduction in reporter gene signal to quantify heterodimer formation, it is also important to have a sensitive reporter such as GFP that can be detected directly with a high degree of sensitivity from whole cells. As illustrated for both integrin α_{IIb} and RAGE receptor (Fig. 3.4 and Fig. 3.6), the DN-AraTM assay is robust to specific interactions between TM domains that exhibit preferential homodimers and heterodimers.

3.5   Conclusion

Overall, given the good agreement between the observed and expected results for the integrin α_{IIb} heterodimer competition experiments (Fig. 3.4) and the ability to assay expression level of each AraC chimera (Fig. 3.2), the AraTM heterodimeric assay provides a useful, complementary method to access important specific interfaces and domains within TM domains of proteins in defining receptor heterodimerization. We feel that DN-AraTM will provide a useful, complementary technique to mammalian cell-based BRET and bacterial two-hybrid BACTH system in investigating the importance of receptor TM and juxtamembrane domain interactions in regulating signal transduction (11,21).

3.6   References

(1) Moore DT, Berger BW, DeGrado WF. Protein–protein interactions in the membrane: sequence, structural, and biological motifs. Structure 2008;16:991–1001.

(2) Roth L, Nasarre C, Dirrig-Grosch S, Aunis D, CremelG, Hubert P, et al. Transmembrane domain interactions control biological functions of neuropilin-1. Mol Biol Cell 2008;19:646–54.

(3) Li W, Metcalf DG, Gorelik R, Li R, Mitra N, Nanda V, et al. A push-pull mechanism for regulating integrin function. Proc Natl Acad Sci USA 2005;102:1424–9.

(4) Berger BW, Kulp DW, Span LM, DeGrado JL, Billings PC, Senes A, et al. Consensus motif for integrin transmembrane helix association. Proc Natl Acad Sci USA 2010;107:703–8.

(5) Hughes PE, Diaz-Gonzalez F, Leong L, Wu C, McDonald JA, Shattil SJ, et al. Breaking the integrin hinge. A defined structural constraint regulates integrin signaling. J Biol Chem 1996;271:6571–4.

(6) Li R, Mitra N, Gratkowski H, Vilaire G, Litvinov R, Nagasami C, et al. Activation of integrin $\alpha_{IIb}\beta_3$ by modulation of transmembrane helix associations. Science 2003;300: 795–8.

(7) Moser M, Legate KR, Zent R, Fassler R. The tail of integrins, talin, and kindlins. Science 2009;324:895–9.

(8) Schneider D, Engelman DM. GALLEX, a measurement of heterologous association of transmembrane helices in a biological membrane. J Biol Chem 2003;278:3105–11.

(9) Lindner E, Langosch D. A ToxR-based dominant-negative system to investigate

heterotypic transmembrane domain interactions. Proteins 2006;65:803–7.

(10) Joce C, Wiener AA, Yin H. Multi-Tox: application of the ToxR-transcriptional reporter assay to the study of multi-pass protein transmembrane domain oligomerization. Biochim Biophys Acta 2011;1808:2948–53.

(11) Karimova G, Pidoux J, Ullmann A, Ladant D. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proc Natl Acad Sci USA 1998;95:5752–6.

(12) Su PC, Berger BW. Identifying key juxtamembrane interactions in cell membranes using AraC-based transcriptional reporter assay (AraTM). J Biol Chem 2012;287:31515–26.

(13) Bustos SA, Schleif RF. Functional domains of the AraC protein. Proc Natl Acad Sci USA 1993;90:5638–42.

(14) Cass LG, Wilcox G. Novel activation of araC expression and a DNA site required for araC autoregulation in Escherichia coli B/r. J Bacteriol 1988;170:4174–80.

(15) Li R, Babu CR, Lear JD, Wand AJ, Bennett JS, DeGrado WF. Oligomerization of the integrin $\alpha_{IIb}\beta_3$: roles of the transmembrane and cytoplasmic domains. Proc Natl Acad Sci USA 2001;98:12462–7.

(16) Li R, Gorelik R, Nanda V, Law PB, Lear JD, DeGrado WF, et al. Dimerization of the transmembrane domain of Integrin $\alpha_{IIb}\beta_3$ subunit in cell membranes. J Biol Chem 2004;279:26666–73.

(17) Zong H, Madden A, Ward M, Mooney MH, Elliott CT, Stitt AW. Homodimerization is essential for the receptor for advanced glycation end products (RAGE)-mediated signal transduction. J Biol Chem 2010;285:23137–46.

(18) Xie J, Reverdatto S, Frolov A, Hoffmann R, Burz DS, Shekhtman A. Structural basis for pattern recognition by the receptor for advanced glycation end products (RAGE). J Biol Chem 2008;283:27255–69.

(19) Ostendorp T, Leclerc E, Galichet A, Koch M, Demling N, Weigle B, et al. Structural and functional insights into RAGE activation by multimeric S100B. EMBO J 2007;26:3868–78.

(20) Ott CM, Lingappa VR. Integral membrane protein biosynthesis: why topology is hard to predict. J Cell Sci 2002;115: 2003–9.

(21) Ramsay D, Kellett E, McVey M, Rees S, Milligan G. Homoand hetero-oligomeric interactions between G-proteincoupled receptors in living cells monitored by two variants of bioluminescence resonance energy transfer (BRET): hetero-oligomers between receptor subtypes form more efficiently than between less closely related sequences. Biochem J 2002;365:429–40.

Chapter 4

HIGH-YIELD MEMBRANE PROTEIN EXPRESSION FROM *E. COLI* USING AN ENGINEERED OUTER MEMBRANE PROTEIN F FUSION

The work described in this chapter has been published in "High-yield Membrane Protein Expression from E. coli Using an Engineered Outer Membrane Protein F Fusion" by Pin-Chuan Su, William Si, Deidre L. Baker, and Bryan W. Berger, Protein Science 2013 22 435-443.

Obtaining high yields of membrane proteins necessary to perform detailed structural study is difficult due to poor solubility and variability in yields from heterologous expression systems. To address this issue, an Escherichia coli-based membrane protein overexpression system utilizing an engineered bacterial outer membrane protein F (pOmpF) fusion has been developed. Full-length human receptor activity-modifying protein 1 (RAMP1) was expressed using pOmpF, solubilized in FC15 and purified to homogeneity. Using circular dichroism and fluorescence spectroscopy, purified full-length RAMP1 is composed of approximately 90% a-helix, and retains its solubility and structure in FC15 over a wide range of temperatures (20–60℃). Thus, our approach provides a useful, complementary approach to achieve high-yield, full-length membrane protein overexpression for biophysical studies.

4.1   Introduction

Membrane receptors play a significant role in regulating numerous fundamental biological processes, including membrane trafficking, signal transduction, and cell–cell

78

communication (1). Thus, it is estimated that 20–30% of all predicted proteins in sequenced genomes encode for integral membrane proteins (2). Despite their prevalence and numerous important biological roles, the number of integral membrane protein high-resolution structures available in the protein data bank (PDB) is <1% (3). The challenges in solving membrane protein structures are many, including their insolubility in aqueous solutions and low abundance in their native membrane environment. Thus, obtaining the high yields of homogeneous, soluble membrane protein required for detailed structural studies is a major challenge (4–6).

Numerous approaches have been developed to improve the yield and recovery of integral membrane proteins. Among the most popular methods for largescale, Recombinant membrane protein expression are *E. coli*-based strategies utilizing protein fusions to improve yield (7–14). *E. coli* has several advantages for membrane protein overexpression: defined growth media is relatively inexpensive, numerous genetic tools are available for straightforward cell and target protein manipulation, and several expression conditions are established for large-scale synthesis (10,15). In particular, fusions that direct hydrophobic polypeptides into inclusion bodies have been particularly successful for cell-based membrane protein overexpression by minimizing toxicity associated with disruption of bacterial membranes (11,13,14) Both soluble and insoluble fusions have also been effective in improving membrane protein yield from cell-free expression systems (12). However, in all cases, no single fusion protein was identified that gave consistently high expression independent of the particular target membrane protein. Furthermore, the effectiveness of many fusion protein vectors is limited to expressing fragments of, rather than full-length membrane proteins. Thus, a

range of vectors is typically screened initially in order to identify specific fusions that maximize membrane protein production for a new target.

To address these issues, we developed a *T7*-based expression vector (pOmpF) using an engineered fragment of outer membrane protein F (OmpF) as a fusion protein to direct full-length membrane protein overexpression in *E. coli*. OmpF is the major outer membrane porin in *E. coli*, and exhibits high stability due to its b-barrel structure, including eight short periplasmic β-hairpins and eight extracellular antiparallel β-strands (16,17). Highyield expression, purification, and refolding protocols have been established for OmpF, suggesting it may be an effective fusion partner to promote high-yield expression of full-length integral membrane proteins (18–20). OmpF as well as other b-barrel membrane proteins are also considerably more polar than a-helical membrane proteins, and thus the improved solubility of OmpF in aqueous solution enhances its refolding using detergents (19–21) We utilized pOmpF to overexpress human receptor activity-modifying protein 1 (RAMP1), a type I integral membrane protein co-receptor for calcitonin with calcitonin-like receptor (CLR) (22). Prior work concerning RAMP1 structure has focused mainly on the extracellular domain, whereas, relatively little is known concerning the structure of the transmembrane (TM) and cytoplasmic domains (23). We demonstrate that pOmpF is able to generate high-yield expression of full-length, human RAMP1, in contrast to other fusion proteins commonly used for membrane protein overexpression, and identified conditions that promote the stability and structure of the purified co-receptor. Thus, pOmpF provides a useful, complementary tool to enable overexpression and purification of full-length integral membrane proteins such as RAMP1 for biophysical studies.

4.2   Materials and Methods

4.2.1   Subcloning and plasmid design

The engineered OmpF sequence (amino acids A23-S164; Uniprot ID P02931) was designed to include strands 1–7 and remove the N-terminal signal peptide sequence. An optimized nucleotide sequence (Genscript) corresponding to the polyhistidine tag, fusion protein sequence, polyglycine linker, thrombin cleavage site, and MCS were subcloned into the kanamycin-resistant pET-28a(+) vector as a NcoI/XhoI fragment to construct plasmid pOmpF (Supplemental Materials Fig. 4.S1). Plasmid pTrpLE containing an optimized TrpLE fragment was kindly provided by Dr. Jebrell Glover (Department of Chemistry, Lehigh University) (13). Plasmid pBCL99 was generated by amplifying the corresponding 99-amino acid BclXL fragment described previously for high-level membrane protein expression from plasmid pEF6-BclXL using specific primers, and subcloned as an NcoI-XhoI fragment into plasmid pET28a(+) (14,24). The human RAMP1 CDS (amino acids M1-V148; Uniprot ID O60894) was generated using overlap extension PCR with synthetic, *E. coli* codon-optimized oligonucleotides. The gene was inserted as a BamHI/XhoI fragment in pOmpF.

4.2.2   Protein expression

Auto-induction

The pOmpF-RAMP1 plasmid was transformed into BL21 cells and allowed to grow for 1 h in SOC media before 10 μl of cells were plated on a LB agar plate

containing kanamycin (50 µg/ml) and grown overnight at 37°C. The following day, individual colonies were isolated and grown in selective LB media (50 µg/ml kanamycin) overnight at 37°C. The next day, saturated cell culture was diluted 100-fold into fresh, selective ZYP media and grown at 37°C for 24 h (25). Cell pellets from induced cultures was harvested by centrifugation and stored at –20°C until further use.

<u>IPTG induction</u>

The pOmpF-RAMP1 plasmid was transformed into BL21 cells and grown for 1 h in SOC media. 10 µl of transformed cells were then plated on a kanamycin selective (50 µg/ml) LB agar plate and grown overnight at 37°C. Colonies were then isolated and allowed to grow in selective LB media (50 µg/ml kanamycin) overnight at 37°C. The next day, saturated cell culture was centrifuged, spent media decanted, and the cell pellet resuspended in fresh M9 minimal media. The cell suspension was then diluted to OD600 of 0.4 in fresh M9 minimal media containing 0.5% v/v glucose and 50 µg/ml kanamycin, and grown at 20°C for 16 h to equilibrate growth temperature. Equilibrated cells were harvested by centrifugation, spent media decanted, then resuspended in fresh M9 minimal media. Equilibrated cultures were then diluted to OD_{600} of 0.4 in M9 minimal media containing 0.5% v/v glycerol and 50 µg/ml kanamycin, grown at 20°C for an hour, then 1 mM IPTG added to induce protein expression. After 16–24 h, cells were collected from induced cultures by centrifugation and stored at –20°C until further use.

4.2.3   Protein purification

Surfactant extraction and solubility test

Induced cultures were centrifuged (12,000 x g) to remove spent media, the pellet resuspended in a 1% (w/v) test surfactant solution, then mixed for 20 min. Samples were then centrifuged and the supernatant analyzed by PAGE gel electrophoresis using 12% acrylamide gels with 1x MES running buffer. To measure the percentage recovery of RAMP1 from the whole-cell extract (% soluble), the intensity of the band corresponding to RAMP1 in the whole-cell extract was compared to that of RAMP1 solubilized using a given test surfactant solution using the densiometry function in ImageJ. The total volume for each sample was kept constant, and the result reported as the ratio of test detergent band intensity relative to the whole-cell extract.

Inclusion body preparation

Fifty milliliter of induced culture was centrifuged (12,000 x g) for 15 min and the supernatant removed. The resulting pellet was resuspended in 50 mL of deionized water and sonicated (15 W) for 20 min to disrupt cell membranes. Next, cultures were mixed using a stir plate with agitation (200 rpm) for 20 min to release soluble proteins, and then centrifuged (12,000 x g) for 15 min. The wash procedure using deionized water was repeated a minimum of three times. After three washes, the resulting pellet containing primarily OmpF-RAMP1 was resuspended in 5 mL of FC15 buffer (1% w/v FC15, 20 mM Tris, pH 8.3), mixed using a stir plate with agitation (200 rpm) for 20 min, and centrifuged (12,000 x g) for 15 min. The supernatant, containing primarily residual contaminating proteins, was decanted, the pellet collected, and the entire FC15

wash procedure repeated at least once. To solubilize OmpF-RAMP1, the FC15-washed pellet was then resuspended in an equal volume of FC15 buffer supernatant.

<u>Immobilized metal ion affinity chromatography (IMAC)</u>

Five-hundred microliter of chelating sepharose Ni-NTA resin (GE Healthcare) was charged using 0.2 M $NiCl_2$ and equilibrated with 1 ml of FC15 buffer prior to use. An equal volume of FC15-solubilized OmpF-RAMP1 solution and settled Ni-NTA resin was mixed in a 2 ml microcentrifuge tube for 20 min on a rotator mixer. The sample was centrifuged briefly to sediment the resin, and the supernatant containing unbound impurities was removed. The resin was then sequentially washed with 1 bed volume of FC15 buffer, 4 bed volumes of FC15 IMAC wash buffer 1 (FC15 buffer. 10 mM imidazole), 1 bed volume of FC15 IMAC wash buffer 2 (FC15 buffer. 20 mM imidazole), 1 bed volume of FC15 IMAC wash buffer 3 (FC15 buffer. 30 mM imidazole), and 1 bed volume of FC15 IMAC wash buffer 4 (FC15 buffer. 50 mM imidazole) to remove bound impurities. Finally, the target protein was eluted with 1 bed volume of FC15 IMAC elution buffer 1 (FC15 buffer. 100 mM imidazole) and 1 bed volume of FC15 IMAC elution buffer 2 (FC15 buffer. 300 mMimidazole).

<u>Thrombin cleavage</u>

One milliliter of IMAC eluate was cleaved with bovine thrombin (BioPharm Laboratories) by incubation for 40 min (5 U/ml thrombin, 2.5 mM $CaCl_2$) with mixing at room temperature on a rotisserie.

<u>Cation exchange chromatography (CEC)</u>

Fivehundred microliter of SP Sepharose (GE Healthcare) was equilibrated with 1 ml of FC15 buffer, then an equal volume of thrombin-cleaved OmpF-RAMP1

solution and settled SP Sepharose was mixed for 10 min. The mixture was centrifuged briefly to sediment resin and the supernatant containing unbound protein removed. The resin was then sequentially washed with 4 bed volumes of FC15 buffer, 1 bed volume of FC15 IEX wash buffer 1 (FC15 buffer, 10 mM NaCl), 1 bed volume of FC15 IEX wash buffer 2 (FC15 buffer. 20 mM NaCl), 1 bed volume of FC15 IEX wash buffer 3 (FC15 buffer. 30 mM NaCl), and 1 bed volume of FC15 IEX wash buffer 4 (FC15 buffer . 50 mM NaCl) to remove any weakly bound material. Solublilized RAMP1 was eluted with 1 bed volume of FC15 IEX elution buffer 2 (FC15 buffer. 500 mM NaCl). NaCl was removed by dialysis against a 1000-fold excess of low-salt FC15 buffer using a 3500 MWCO dialysis membrane (Pierce).

4.2.4   Immunoblotting

Twenty microliter of protein sample was mixed with 5 µl of 5x Lammeli sample buffer, heated briefly at 90°C, then loaded onto a 12% acrylamide gel with MES running buffer. Samples were run for 1 h at 200 V using Lammeli running buffer, then transferred to a nitrocellulose membrane (Amersham Hybond ECL) for 90 min, blocked for 1 h at room temperature using 5% milk in TBST, then incubated with a 1000x dilution of anti-RAMP1 rabbit polyclonal antibody (ab96125, abcam) for 16 h at 4°C or a 3000x dilution of anti-His antibody (Cell Signaling) for 1 h at room temperature. The membrane was washed with TBST, exchanging solution every 5 min, followed by incubation with a 3000x dilution of goat polyclonal anti-rabbit IgG antibody (ab6721, abcam) for 1 h at room temperature or a 15,000x dilution of anti-mouse IgG antibody

for 1 h at room temperature. Membranes were developed using a chemiluminescent substrate (GE) and imaged using a Typhoon imager.

4.2.5 Mass spectrometry

The MW of purified human RAMP1 protein was verified by MALDI-TOF mass spectrometry. Protein samples in FC15 buffer were mixed with sinapinic acid matrix (acetonitrile/water/TFA: 50/50/0.1) at a 1:1 ratio, spotted on a MSP 96 target ground steel plate (Bruker) and allowed to air dry before analysis using a Microflex mass spectrometer (Bruker Daltonics).

4.2.6 Circular dichroism (CD)

Spectra were collected using 400 μl of dialyzed protein samples (0.14 mg/ml) in 0.1 cm quartz cuvettes (Starna) using a J-815 CD spectrometer (JASCO) with heated sample chamber. The scanning speed was set to 500 nm/min and measurements collected from 180 to 260 nm. All CD spectra were collected in nanopure water (Millipore) containing 1% w/v FC15; this solution was also used for background subtraction. For thermal stability measurements, samples were first heated to a specific temperature (in 10° increments) from 20 to 60°C and then cooled down to 20°C. The spectrum of buffer containing 1% w/v FC15 at each temperature was used for background subtraction. Raw spectra (in mdeg) were converted to molar ellipticity (in degree cm_2/dmol), and percent secondary structure estimated from the converted spectra

using the K2D3 online server (26).

4.2.7 Fluorescence measurements

Fluorescence emission spectrums (310–410 nm) were measured at 25°C using a Cary Eclipse Spectrophotometer (Agilent Technologies) with an excitation wavelength of 295 nm. Buffer conditions were identical to those used for CD measurements. For free tryptophan, a concentration of 0.01 mg/ml was used. For the disulfide bond reducing experiment, a DTT concentration of 10 mM was used.

4.2.8 Dynamic light scattering (DLS)

Measurements were performed using a Brookhaven Instruments spectrometer and the intensity–intensity time correlation function was measured by means of a BIO-9000 AT multichannel digital correlator. The scattering angle 90°, 75°, and 60° were used. The CONTIN method was used to analyze the normalized electric field time correlation function to determine the particle's apparent diffusion coefficient (D) (27). The mean hydrodynamic radius (Rh) was calculated from the Stokes–Einstein equation:

$$R_h = (k_b T)/(6\pi\eta D)$$

where k_b is the Boltzmann constant, T the temperature, and η the shear viscosity of the solvent. Buffer conditions were identical to those used for CD measurements.

4.3 Results

4.3.1 Design of pOmpF expression vector

Expression plasmid pOmpF was constructed using a *T7* promoter to drive
overexpression in *E. coli.* Details of the plasmid design and construction are provided in
''Materials and Methods Section 4.2.1''. The main features of the expression system are
illustrated in Figure 4.1: N-terminal polyhistidine tag, an engineered OmpF fragment
(amino acids A23-S164; Uniprot ID P02931), a polyglycine linker, a thrombin cleavage
site, and multiple cloning sites (MCS). In the engineered OmpF fusion, the signal
peptide of OmpF was removed to prevent protein trafficking to the outer membrane and
direct inclusion body formation.
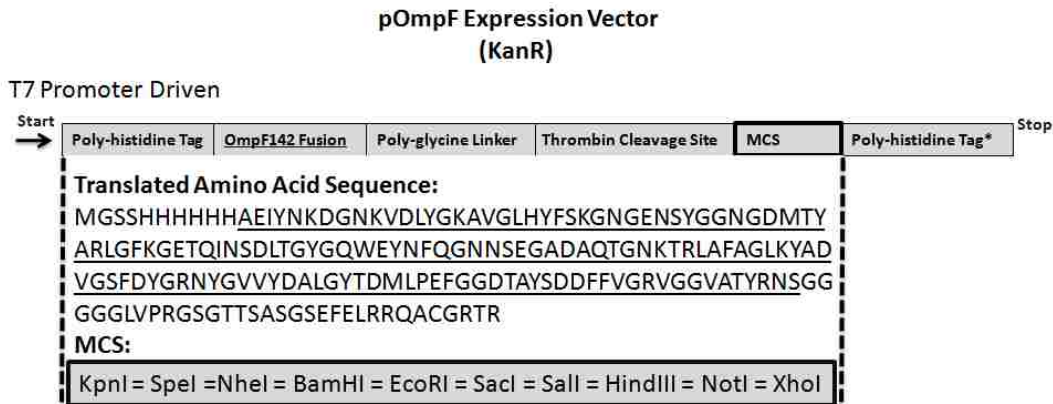
Figure 4. 1. pOmpF Expression Vector Plasmid Map. Key features of pOmpF
(kanamycin resistant) are illustrated above, including N-terminal poly-histidine
purification tag, engineered OmpF fragment, poly-glycine linker with thrombin
cleavage site and multiple cloning site. Fusion protein expression is driven by a *T7*
promoter. The translated protein sequence for the engineered OmpF fusion is also
included.

88

4.3.2   Expression of RAMP1 protein as an OmpF fusion

We generated an in-frame fusion with our engineered OmpF fragment to full-length human RAMP1 (amino acids M1-V148; Uniprot ID O60894), and compared expression to three other commonly used fusions for membrane protein overexpression: trpLE, BclXL, and glutathione-S-transferase [pET-42a (+) expression plasmid] (13,14). Fusions were expressed in LB broth using BL21 (DE3) cells for 12–16 h at a growth temperature of 20°C as described in ''Materials and Methods Section 4.2.2''. As illustrated in Figure 4.2, RAMP1 expression was only observed as a fusion with pOmpF, with a prominent band visible by PAGE at the expected fusion MW (35.1 kDa) present in the whole-cell extract and confirmed by western blotting using anti-His antibody. Furthermore, high-yield OmpF-RAMP1 expression is observed for several different media formulations (Fig. 4.3), including autoinduction (ZYP growth media) at 20°C and 37°C, LB media with 1 mM IPTG at 20°C, and M9 minimal media with IPTG at 20°C (25). Thus, the engineered OmpF fragment is able to facilitate overexpression of full-length RAMP1, in contrast to other fusions used previously for membrane protein overexpression. Furthermore, OmpF fusions can be expressed in minimal media at a scale suitable for metabolic labeling and structure determination as well as in rich media to increase protein expression levels for biophysical analysis. We also successfully expressed a 2 TM fragment of the cognate co-receptor for RAMP1, CLR, using pOmpF (Fig. 4.4) under similar conditions to those described for RAMP1, further demonstrating the potential of pOmpF to produce single and multi-pass TM proteins at a scale appropriate for biophysical studies.
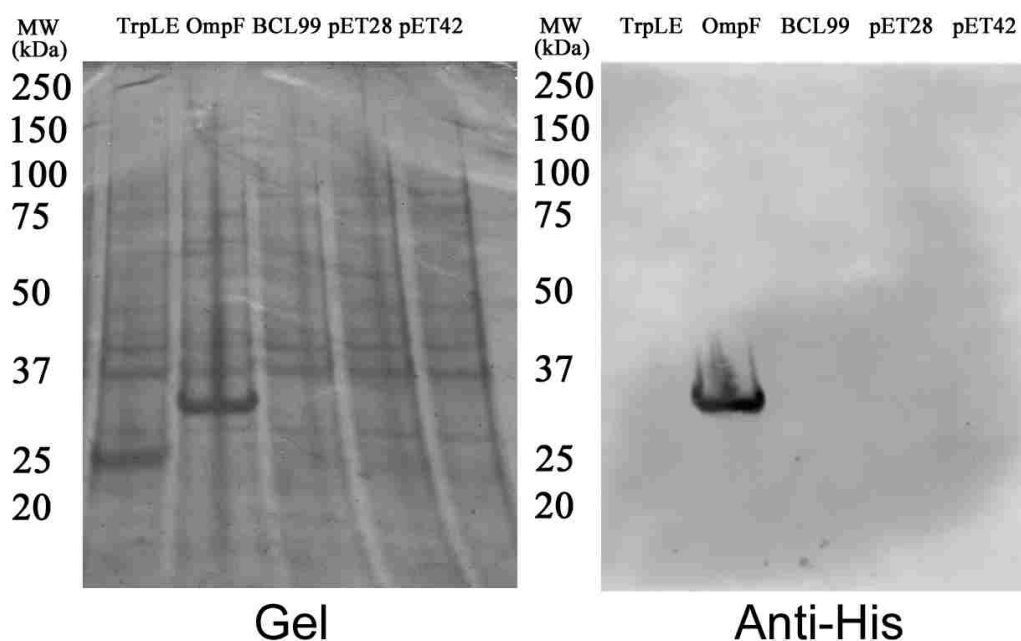
Figure 4. 2. Comparison of RAMP1 Expression with Different Fusion. Full-length RAMP1 CDS was subcloned into three different expression vectors (TrpLE, BCL99, and pET42) containing commonly used fusions for membrane protein overexpression, as well as pOmpF, which contains an engineered OmpF fragment, and pET28, which does not contain a fusion tag. RAMP1 fusions were expressed using IPTG induction from BL21 (DE3) cells for 12–16 h at a growth temperature of 20°C. A prominent band at the expected size (35.1 kDa) for the OmpF-RAMP1 fusion is observed in whole-cell lysates from pOmpF, whereas no prominent bands are observed for any of the other constructs. Immunoblotting using an anti-His antibody confirms the band at the expected size for OmpF-RAMP1 is specific, whereas no bands are observed for any of the other constructs. [TrpLE: pET-TrpLE, OmpF: OmpF, BCL99: pBCL99, pET28: pET-28a (+), and pET42: pET-42a (+)]
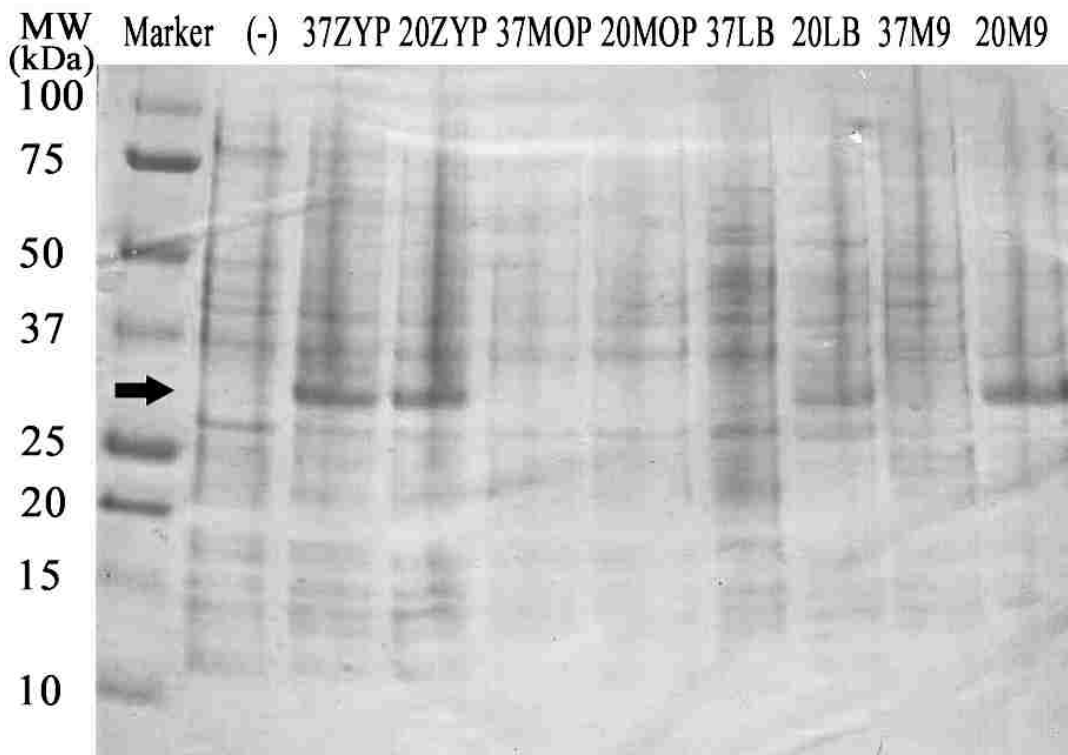
Figure 4. 3. Different Induction Conditions for pOmpF Expression Vector. pOmpF containing full-length RAMP1 was transformed into BL21 (DE3) cells and expression

induced using different growth medias, growth temperatures and inducers. Prominent bands at the expected size for OmpF-RAMP1 were observed from whole-cell lysates expressed using ZYP media at 37°C, ZYP media at 20°C, LB media. IPTG at 20°C, and M9 minimal media . IPTG at 20°C (band between 25kDa and 37kDa). [(-): Uninduced culture, 37ZYP: ZYP media at 37°C, 20ZYP: ZYP media at 20°C, 37MOP: MOPS media at 37°C, 20MOP: MOPS media at 20°C, 37LB: LB media at 37°C, 20LB: LB media at 20°C, 37M9: M9 minimal media at 37°C, and 20M9: M9 minimal media at 20°C].
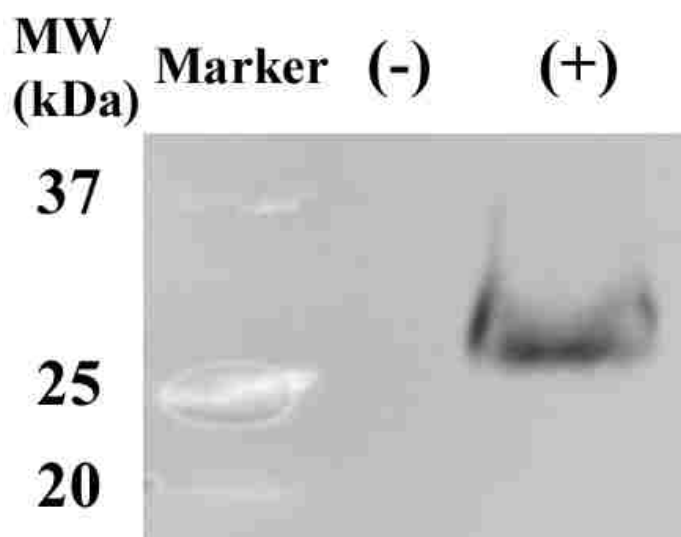
Figure 4. 4. Anti-His Western Blot on OmpF Fused Human Calcitonin Gene-related Peptide Type 1 Receptor (CLR) Fragment (30kDa). The anti-His western blot demonstrates the effectiveness of OmpF expression vector expressing a human calcitonin gene-related peptide type 1 receptor protein fragment (SeqRef#: NP_005786.1) that includes dual transmembrane domains (L237-R336) ((-): uninduced culture and (+): 20°C IPTG induced culture).

4.3.3 Purification of RAMP1 protein

The overexpressed OmpF-RAMP1 fusions were retained in the insoluble fraction of the cell after lysis using 1% Triton X-100 and sonication (Fig. 4.5 and Fig. 4.6). An advantage of insoluble membrane protein fusions is the ability to use differential extraction with detergents, salts and other solutes to selectively enrich for the fusion in the insoluble fraction prior to solubilization (28,29). For OmpF-RAMP1, we were able to achieve >90% pure OmpF-RAMP1 fusion protein as assessed by PAGE (Fig. 4.4A lane WS2) using a series of washes followed by resolubilizing the protein in zwitterionic detergent fos-choline 15 (FC15). Further details of the wash and solubilization procedure are provided in ''Materials and Methods Section 4.2.3".
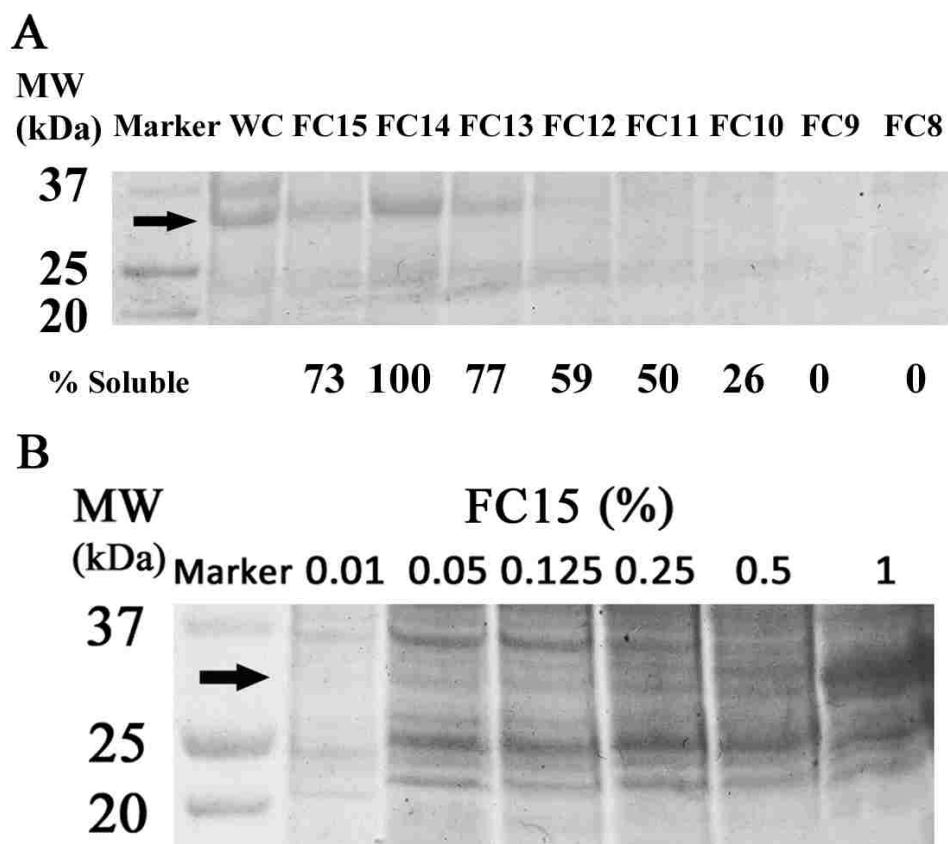
Figure 4. 5. Extraction and Solubilization of OmpF-RAMP1 as a Function of Surfactant Type. A, Induced cultures were lysed, centrifuged and the insoluble fraction mixed with a test detergent solution at 1% (w/v) concentration. After incubation, the soluble fraction was collected by centrifugation and analyzed by SDS-PAGE. OmpF-RAMP1 was highly soluble in FC15, FC14, and FC13 and moderately soluble in FC12 and FC11 (WC: whole cell lysate, FC15: FOS-CHOLINE-15, FC14: FOS-CHOLINE-14, FC13: FOS-CHOLINE-13, FC12: FOS-CHOLINE-12, FC11: FOS-CHOLINE-11, FC10: FOS-CHOLINE-10, FC9: FOS-CHOLINE-9, and FC8: FOS-CHOLINE-8). B, OmpF-RAMP1 was solubilized in different concentration of FC15, and the minimum FC15 concentration necessary to solubilize OmpF-RAMP1 is 1% w/v. Percent solubilization of RAMP1 by a given test detergent is reported in terms of the relative ratio of RAMP1 present in the whole-cell extract versus that in a given test detergent solution.
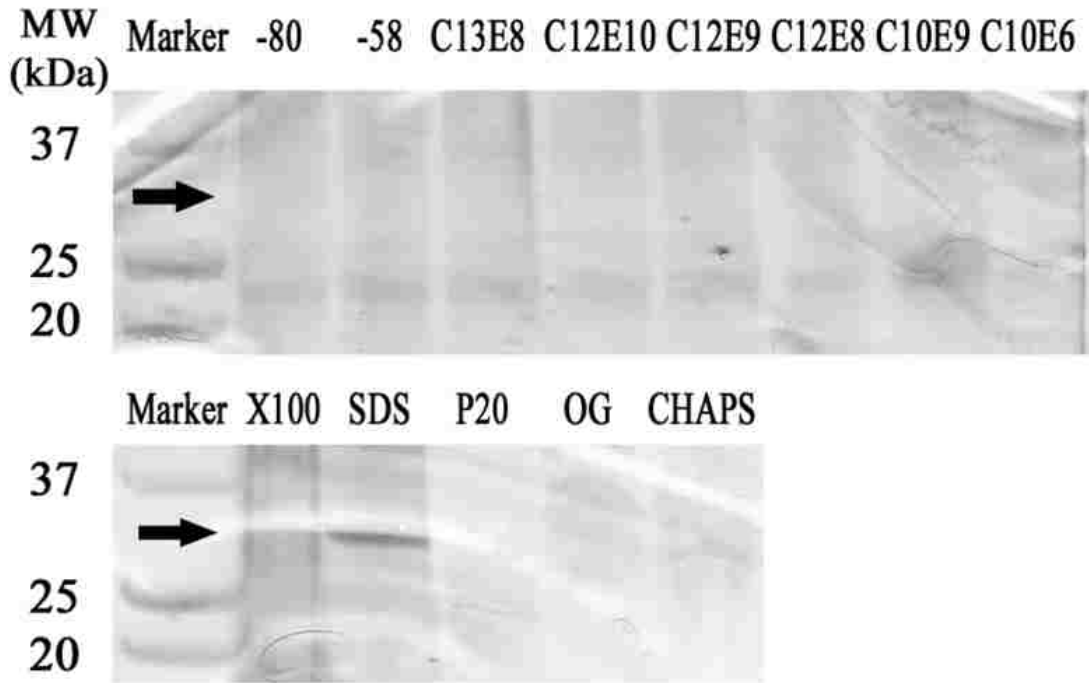
Figure 4. 6. Comparing Extraction and Solubilization of OmpF-RAMP1 as a Function of Surfactant Type. Induced cultures were lysed, centrifuged and the insoluble fraction mixed with a test detergent solution at 1% (w/v) concentration. After incubation, the soluble fraction was collected by centrifugation and analyzed by SDS-PAGE. OmpF-RAMP1 was only fully soluble in SDS (-80: ANAPOE-80, -58: ANAPOE-58, C13E8: ANAPOE-$C_{13}E_8$, C12E10: ANAPOE-$C_{12}E_{10}$, C12E9: ANAPOE-$C_{12}E_9$, C12E8: ANAPOE-$C_{12}E_8$, C10E7: ANAPOE-$C_{10}E_7$, C10E6: ANAPOE-$C_{10}E_6$, X100: Triton X-100, SDS: Sodium dodecyl sulfate, P20: Polysorbate-20, OG: N-oxytyl-β-D-glucoside, and CHAPS: CHAPS).

We tested the ability of a series of non-ionic, ionic, and zwitterionic detergents to effectively solubilize OmpF-RAMP1, with the expectation that mild detergents will be effective at solubilizing the fusion. Of those tested, OmpF-RAMP1 was highly soluble in the denaturing detergent SDS as well as the zwitterionic FC detergents FC13 (77%), FC14 (100%) and FC15 (73%), and moderately soluble in FC11 and FC12 (Fig. 4.5); percent solubilization for each sample is reported as the ratio of RAMP1 present in the detergent extract relative to the whole-cell lysate. FC detergents have been used

successfully for solubilization, chromatographic purification, and biophysical characterization of a wide range of integral membrane proteins overexpressed in *E. coli*, including G-protein coupled receptors and OmpF (16,30) For FC15, the minimum detergent concentration required to solubilize the fusion protein was determined to be 1% (w/v) (Fig. 4.5B). Therefore the mild zwitterionic detergent FC15 is a suitable choice to solubilize OmpF-RAMP1 directly, without a need for harsh denaturants or high detergent concentrations (5,31).
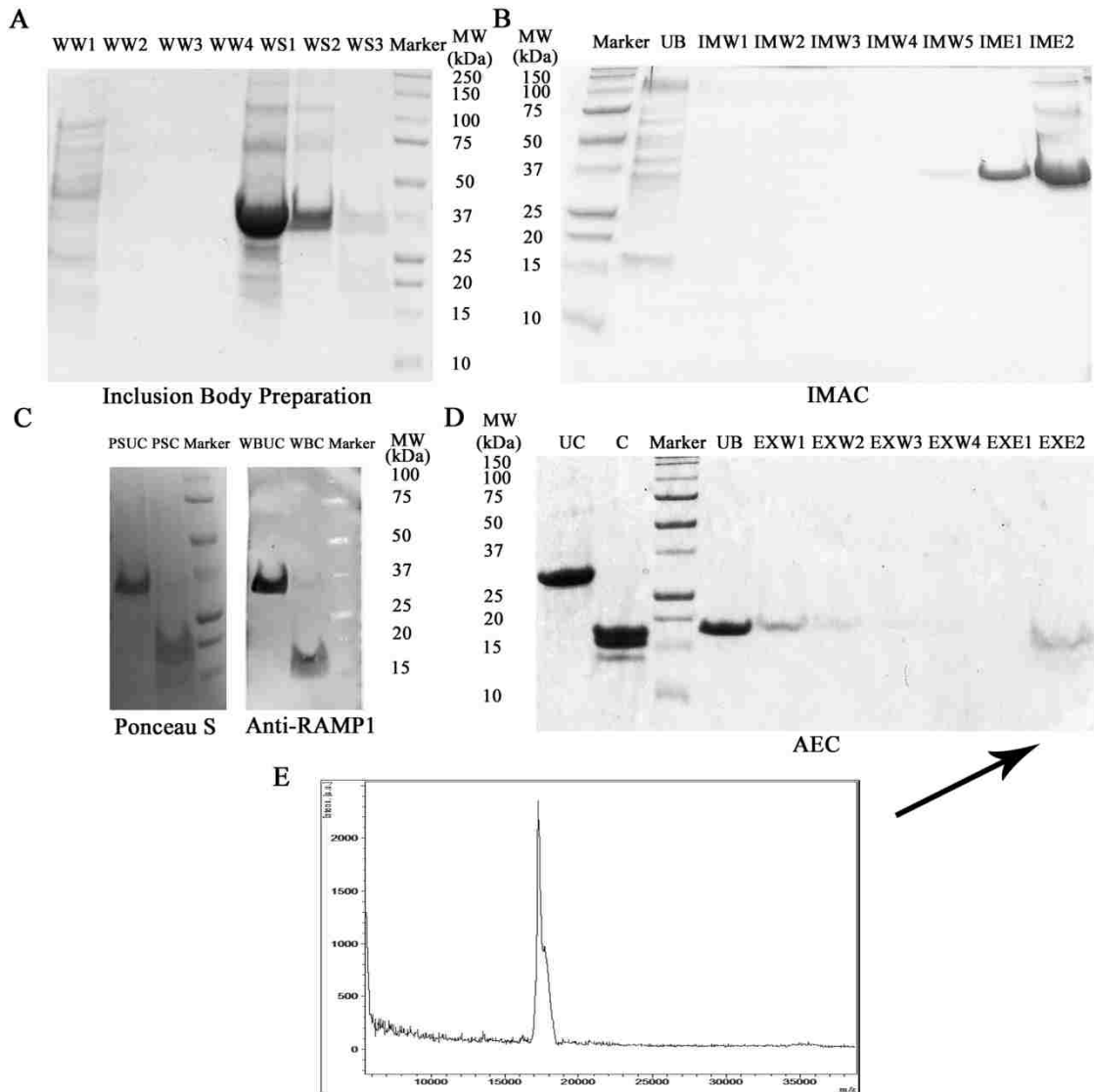
Figure 4. 7. Purification of RAM1 from OmpF-RAMP1 Fusion. A, Cellular proteins from whole-cell extracts are removed by successive washes of Insoluble fraction using water (WW) and FC15 (WS) (WW1: water wash 1, WW2: water wash 2, WW3: water wash 3, WW4, water wash 4, WS1: FOS-15 buffer solubilization 1, WS2: FOS-15 buffer solubilization 2, and WS3: FOS-15 buffer solubilization 3). B, Immobilized metal ion affinity chromatography (IMAC) was used to isolate His-tagged OmpF-RAMP1 fusion protein using a series of low imidazole washes (IMW; 10–30 mM) to remove impurities followed by high imidazole wash (IME; 100 mM) to elute bound OmpF-RAMP1 (UB: unbound, IMW1: IMAC wash 1, IMW2: IMAC wash 2, IMW3: IMAC wash 3, IMW4: IMAC wash 4, IMW5: IMAC wash 5, IME1: IMAC elution 1, and IME2: IMAC elution 2). C, Purified protein samples were transferred to a nitrocellulose membrane, transferred proteins were visualized using Ponceau S stain and RAMP1 identified in visualized bands by immunoblotting with anti-RAMP1

96

antibody (Anti-RAMP1). The anti-RAMP1 immunoblot confirms RAMP1 expression in the OmpF-RAMP1 fusion and in the thrombin-cleaved sample to remove OmpF (PSUC: Ponceau S stained uncleaved OmpF-RAMP1 fusion, PSC: Ponceau S stained thrombin cleaved OmpF-RAMP1 fusion, WBUC: anti-RAMP1 blotted uncleaved OmpF-RAMP1 fusion, and WBC: anti-RAMP1 blotted thrombin cleaved OmpF-RAMP1 fusion). D, Cation exchange chromatography (CEC) at pH 8.3 was used to isolate RAMP1 from OmpF and other proteins present in cleavage solution (UC: uncleaved, C: cleaved, UB: unbound, EXW1: IEX wash 1, EXW2: IEX wash 2, EXW3: IEX wash 3, EXW4: IEX wash 4, EXE1: IEX elution 1, and EXE2: IEX elution 2). E, MALDI MS of purified RAMP1 from lane EXE1 in Figure 4D indicates a single peak at 17 kDa, which corresponds to MW of RAMP1 (Theoretical pI and MW: OmpF-RAMP1 [6.59 and 35.1kDa), OmpF (5.73 and 17.4kDa), RAMP1 (8.56 and 17.8kDa), Bovine Thrombin (7.05)].

For purification, IMAC affinity chromatography was used as an initial capture step for the OmpFRAMP1 fusion protein, followed by ion exchange chromatography to separate OmpF from RAMP1 after treatment with thrombin. Having pre-washed the OmpF-RAMP1 fusion present in the insoluble fraction to improve purity to >90% (Fig. 4.7A), the fusion was immobilized onto IMAC, washed with low concentrations of imidazole to remove residual impurities and eluted using 100 mM imidazole (Fig. 4.7B). After IMAC, the OmpF-RAMP1 fusion was essentially pure as assessed by PAGE. Thrombin treatment of the eluate was robust to the presence of FC15 and imidazole, and >90% of RAMP1 was released from OmpF after a 40-min digestion. Importantly, OmpF-RAMP1 and RAMP1 both remained soluble and stable in FC15 throughout the digest. Due to the similar sizes of RAMP1 (17.8 kDa) and OmpF (17.4 kDa), an anti-RAMP1 western blot was used to confirm the presence of full-length RAMP1 before and after thrombin treatment. As observed on nitrocellulose membranes visualized by Ponceau S stain (Fig. 4.7C), two prominent species between 15 and 20 kDa are present in the thrombin-treated sample. However, only one of the two species

present in the thrombin-treated sample were reactive against anti-RAMP1, confirming RAMP1 was successfully removed from the OmpF fusion in a soluble, stable form. Interestingly, RAMP1 appeared to migrate faster than OmpF on a PAGE gel, which may be a result of altered detergent binding to TM helices resistant to unfolding by SDS (32). Thus, FC15 is compatible with chromatographic capture of OmpF-RAMP1 using IMAC, thrombin cleavage and stabilization of full-length RAMP1 in solution.

Cation exchange chromatography (CEC) was used to remove residual OmpF and thrombin from FC15-solubilized RAMP1. CEC was preferred based on the predicted isoelectric point (pI) of RAMP1 (8.6), which is considerably higher than that of OmpF-RAMP1 (6.6), OmpF (5.7) and thrombin (7.0). After loading the thrombin digestion mixture onto the CEC at pH 8.3, an initial isocratic wash with low NaCl (50 mM) at pH 8.3 was used to remove OmpF-RAMP1, OmpF, and thrombin. A subsequent isocratic wash at high NaCl concentration (0.5M) then enabled purification of bound, FC15-solubilized RAMP1 (Fig. 4.7D). Purity was assessed MALDI MS (Fig. 4.7E), which indicates a single species at the expected MW for full-length, human RAMP1.

4.3.4   *Characterization of purified RAMP1 protein from pOmpF expression vector*

Using dynamic light scattering (DLS), we observe a single, monodisperse population for FC-solubilized, purified RAMP1 with an estimated hydrodynamic radius (3 nm) consistent with a RAMP1-FC15 complex (Fig. 4.8). Moreover, in order to confirm purified RAMP1 is soluble and stable after purification from OmpF, we used

circular dichroism (CD) to measure the secondary structure of purified RAMP1 solubilized in FC15. At 20°C, the two local minima at 208 nm and 222 nm and maximum between 190 and 195 nm indicative of α-helical secondary structure (Fig. 4.9) (33). We also examined the thermal stability of RAMP1 secondary structure between 20 and 60°C in FC15, and observed only a modest (5%) decrease in the molar ellipticity at 195 nm and increase in the molar ellipticity at 208 nm at 60°C which was reversible upon cooling to 20°C. Importantly, no precipitation of RAMP1 was observed over the entire 20–60°C temperature range using either FC15. Thus, we conclude RAMP1 is stable in FC15 at the level of secondary structure, and the observed secondary structure is insensitive to temperature. We also measured the secondary structure and thermal stability of purified RAMP1 in 1:1 dodecyl β-D-maltoside:FC15 (Fig. 4.10) and found essentially identical results to FC15 alone, indicating the secondary structure and thermal stability of purified RAMP1 is independent of detergent type.

Scattering Angle (Degree)	90	75	60
Hydrodynamic Radius (nm)	2.589	2.814	3
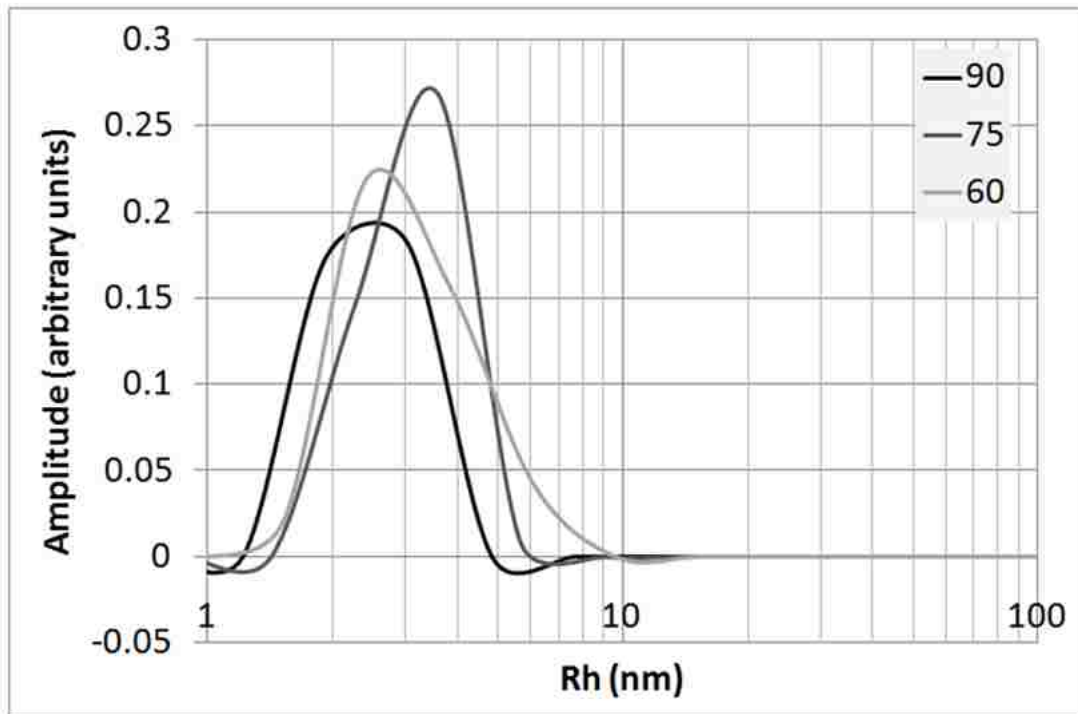
Figure 4 8. Dynamic Light Scattering Analysis of RAMP1-FC15 Complex. A homogeneous population of RAMP1 solubilized in FC15 is observed, with an estimated hydrodynamic radius of about 3 nm, which is consistent with a RAMP1 monomer solubilized in FC15.
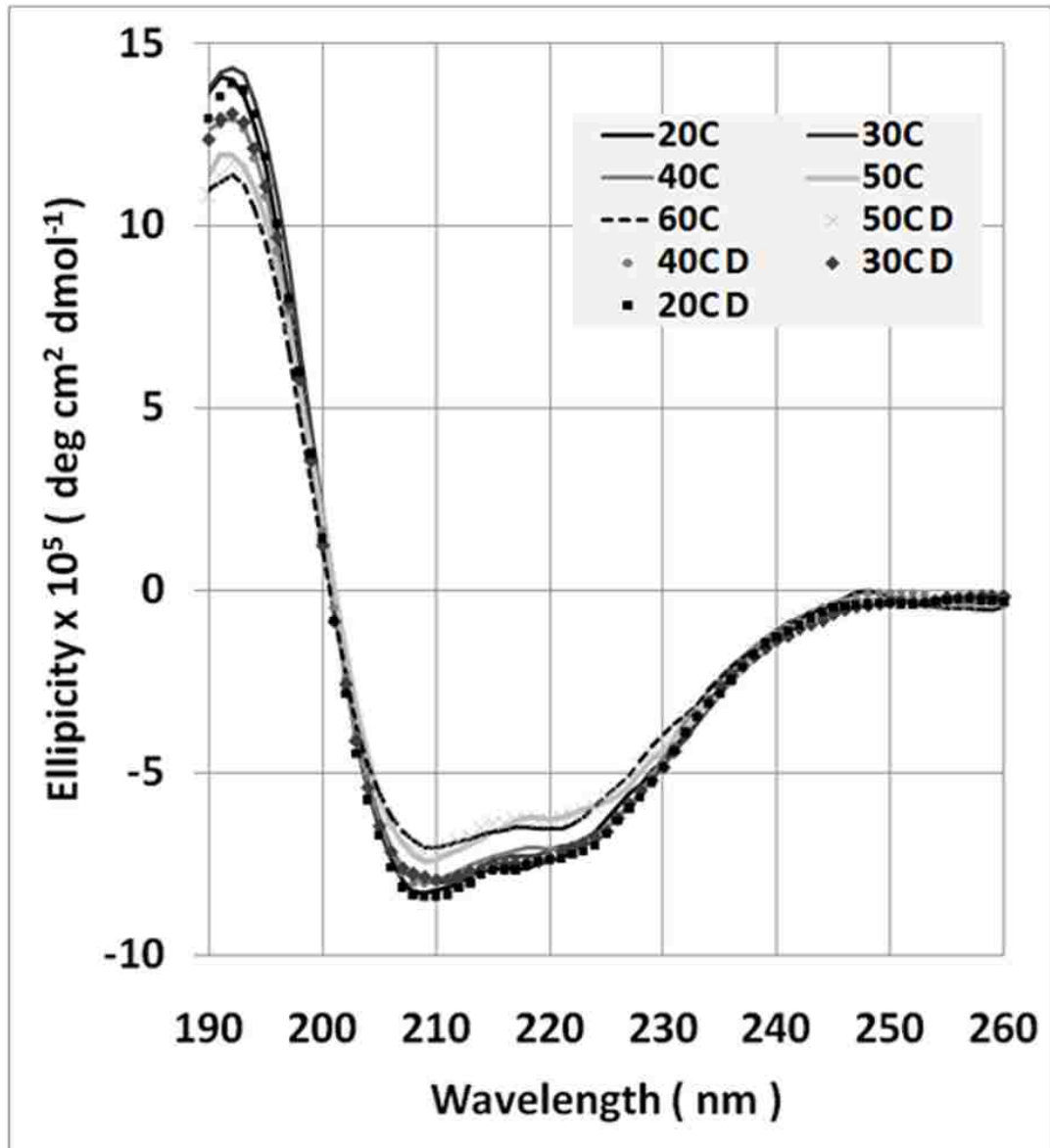
Figure 4. 9. Circular Dichroism Spectra of Purified RAMP1. All spectra represent an average of three scans. As temperature was increased to 60°C, the minima at 208 nm and 222 nm increased slightly (5%), but the overall spectra remained α-helical. The change in minima was completely reversible upon cooling samples to 20°C. (D: reducing temperature).
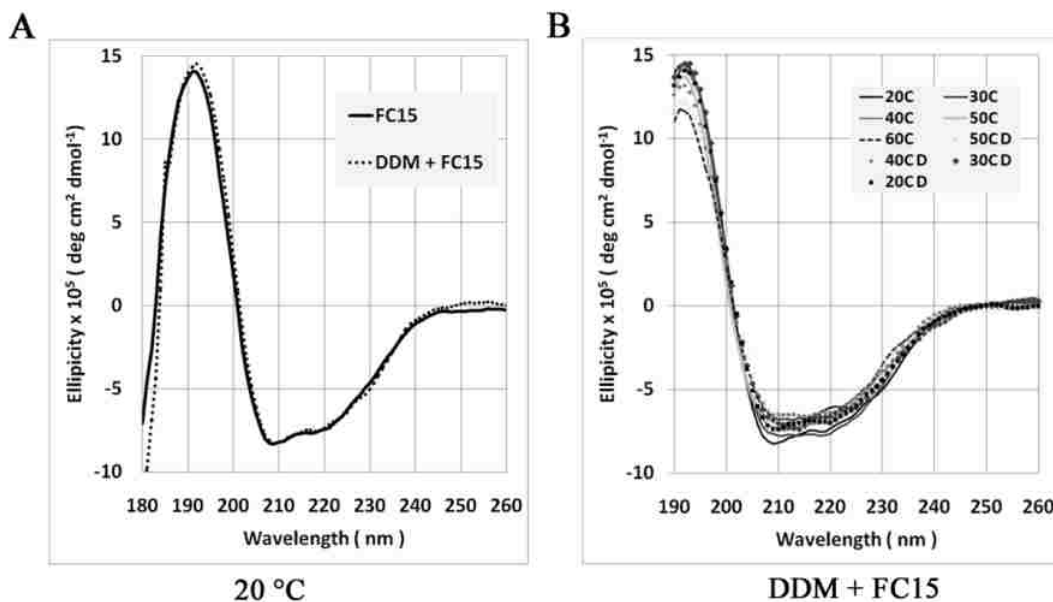
Figure 4. 10. Circular Dichroism Spectra of Purified RAMP1. All spectra represent an average of three scans. A, Circular dichroism spectra of RAMP1 at 20°C showed that RAMP1 is primarily α-helical (local minima at 208 and 222nm) in both detergent conditions (1% FC15 and 1% FC15 + 1% DDM). B, CD spectra of RAMP1 solubilized in 1% FC15 + 1% DDM detergent mixture is essentially identical to that measure for 1% FC15, indicating α-helical secondary structure, and remained constant over a wide range of temperatures (20-60°C).

As further confirmation that the FC15-RAMP1 complex forms a stable tertiary structure, we compared tryptophan emissions spectra for wild-type RAMP1 with and without reducing agent dithiothreitol (DTT). Compared to free tryptophan, which has an emissions maximum at 359 nm, wild-type RAMP1 is significantly blue-shifted, with an emissions maximum at 340 nm (Fig. 4.11), consistent with burial of extracellular tryptophan residues. Addition of DTT to wild-type RAMP1 causes a nearly 25% increase in the intensity of the tryptophan emission spectrum within 10 min; the increase in fluorescence under reducing conditions is indicative of tryptophan fluorescence quenching in the non-reduced, purified RAMP1-FC15 sample due to

102

disulfide bond formation (34).
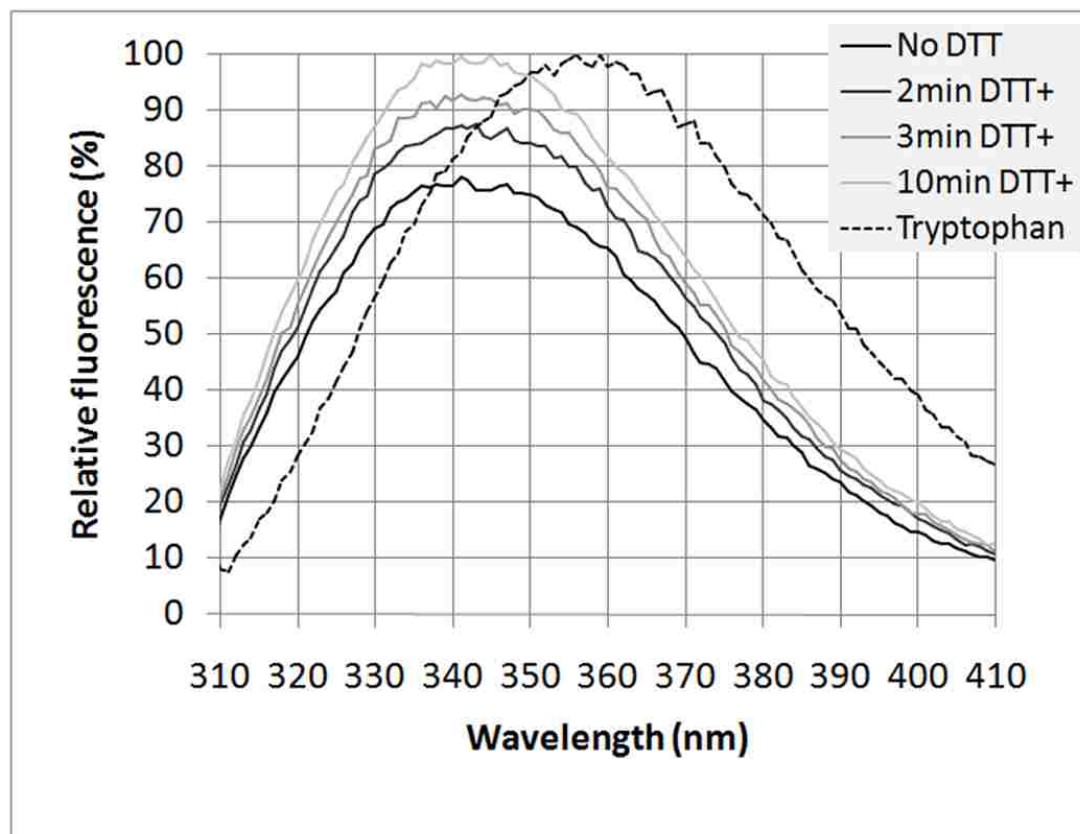
Figure 4. 11. Fluorescence Spectrum of Wild-type and Reduced RAMP1. An excitation wavelength of 295 nm was used to selectively observe the effect of tryptophan residues. Compared to the fully solvent exposed tryptophan sample (maximum at 359 nm), the fluorescence spectrum of unreduced RAMP1 is blue shifted (maximum at 341 nm), indicating burial of the extracellular tryptophan residues. Furthermore, the 25% increase in spectrum maximum in the reduced RAMP1 sample relative to unreduced RAMP1 is consistent with the presence of disulfide bonds, which when reduced, lead to increased tryptophan exposure (No DTT: no DTT added, 2min DTT+: 2 min after DTT was added, 3min DTT+: 3 min after DTT was added, 10min DTT+: 10 min after DTT was added, and Tryptophan: free tryptophan).

4.3   Discussion

A major challenge in working with membrane proteins is identifying conditions that promote both solubility and stability of the fusion and purified target membrane

protein. In general, membrane proteins are only soluble in strong denaturants such as urea and SDS or organic solvents such as formic acid due to their hydrophobicity. However, harsh conditions typically used for solubilization, such as 70% formic acid, 1–5% SDS or 8M urea, not only unfold the membrane protein, but may also lead to irreversible denaturation and aggregation (28,35). Thus, choice of a polar fusion protein, such as OmpF, can often enhance the overall solubility of the fusion, thereby reducing the amount of denaturant required or in some cases, enable solubilization using non-denaturing detergents.

When designing the pOmpF Expression Vector, we found that the truncated 15.4 kDa fragment (A23-S164) of OmpF including strand 7 (G157-S163) was most robust to consistent, high-yield expression, and was therefore chosen for further study. We also included several features in the pOmpF Expression Vector to facilitate the purification process of the target protein: the addition of an N-terminal poly-histidine tag in-frame with the expressed protein enables purification via IMAC and immunoblotting to confirm expression from cell lysate and the incorporation of a thrombin cleavage site allows the option of cleaving the target protein from aqueous solutions in a detergent-solubilized, native-like state. Previous work has indicated that thrombin is a suitable protease for cleavage of membrane protein fusions, and is largely insensitive to the type of detergent present when solubilizing fusion proteins (36).

To confirm the protein produced by pOmpF Expression Vector has the native structure is extremely critical for further protein structural determination purpose. With that in mind, we need to prove that our purified RAMP1 protein constitutes the native RAMP1 structure and this is achieved through several experimental techniques

including DLS, CD, and Tryptophan fluorescent quenching. First, DLS results showed that purified RAMP1 protein is in a monodisperse form with an estimated hydrodynamic radius of 3nm, indicating the peptides were not in large aggregates. Then, we used K2D3 to extract secondary structure from the experimental CD spectra of purified RAMP1 (26) and found full-length RAMP1 is composed of 90% α-helix, with negligible β-strand. For comparison, the previously published RAMP1 extracellular domain crystal structure (PDB 2YX8) is comprised of a three-helix bundle with approximately 81% α-helix, which is less than the estimated 90% helical content by CD for full-length RAMP1 (23). Thus, we conclude the RAMP1 TM and likely cytosolic domains are also contributing to the additional α-helical secondary structure estimated by CD for full-length, TM RAMP1. Given that RAMP1 is a type I integral membrane protein with a predicted single, TM α-helix, the experimentally measured increase in helicity for full-length, TM RAMP1 versus soluble, extracellular RAMP1 is not surprising. Thus, we conclude that FC15 is an appropriate detergent for stabilization of purified, full-length RAMP1, with the measured secondary structure (90% α-helix) consistent with the expected result based on the previous crystal structure for the RAMP1 extracellular domain (23).

Additionally, we compared tryptophan emissions spectra for wild-type RAMP1 with and without reducing agent dithiothreitol (DTT) to confirm that the FC15-RAMP1 complex forms a stable tertiary structure. RAMP1 is predicted to have three extracellular disulfide bonds that contribute to tertiary structure, and multiple studies of the effects of disulfide bond formation on protein tertiary structure have demonstrated that an increase in the intensity of tryptophan fluorescence upon disulfide bond

105

reduction occurs concomitant with an increase in the partial molar volume due to loss of tertiary structure (34,37). RAMP1 has also been shown to regulate the glycosylation of its cognate co-receptor CLR, which could potentially influence the tertiary structure of RAMP1 in a CLR-RAMP1 heterodimeric complex (38). However, previous studies of *E. coli* expressed CLR-RAMP1 extracellular domain fusions have demonstrated comparable ligand binding affinity to full-length receptors in mammalian membranes despite not being glycosylated (39). Thus, we conclude that RAMP1 assumes a tertiary structure consistent with a stable, folded state, particularly when considered along with the expected secondary structure (Figs. 4.9 and 4.10), monodispersity (Fig. 4.8), and thermal stability (Figs. 4.9 and 4.10) measured for RAMP1-FC15.

Moreover, pOmpF Expression Vector has really high membrane protein expression efficiency. We found that from an initial culture volume of 25 mL, we are able to purify approximately 1 mg of RAMP1 solubilized in FC15. Thus, using RAMP1 as a representative integral membrane protein, we anticipate yields of purified, target membrane protein on the order of 1–10 mg purified receptor per liter culture from pOmpF, which enables membrane protein purification on a scale necessary for biophysical and structural studies.

4.5   Conclusion

The engineered OmpF fusion (pOmpF) provides a novel, complementary vector to others used for high-yield expression and purification of full-length, integral membrane proteins from *E. coli*. In particular, we find that the engineered OmpF

truncation enables overexpression of full-length integral membrane proteins such as the 17 kDa human RAMP1 in a form suitable for isolation and enrichment from whole-cell lysates, yet provides adequate solubility to enable extraction directly into zwitterionic detergents such as FC15 without a need for harsh denaturants. High-yield expression of human RAMP1 using pOmpF is observed from a variety of expression conditions, including minimal media suitable for metabolic labeling and enriched autoinduction broth for high-density protein expression. The ability to extract soluble, stable membrane protein fusion directly from cell lysates into FC15 simplifies subsequent purification, with a single IMAC capture step capable of purifying the fusion to near homogeneity. Furthermore, the solubility and stability of the FC15-solubilised OmpF-RAMP1 fusion enables removal of the OmpF fusion and recovery of the purified, full-length target membrane protein in yields sufficient for biophysical and structural characterization. In the case of purified, full-length RAMP1, the experimentally measured secondary structure by CD was consistent with the previous structure of full-length human RAMP1 extracellular domain, exhibited high thermal stability and tryptophan fluorescence emissions spectra consistent with a folded, stable RAMP1 tertiary structure (23). Overall, the engineered OmpF provides a useful method for robust, high-yield membrane protein expression and purification, particularly in instances where other commonly-used fusions tags are not effective at improving expression yield or conferring solubility.

4.6   References

(1) Klemm, JD, Schreiber SL, Crabtree GR (1998) Dimerization as a regulartory mechanism in signal transduction. Annu Rev Immunol 16:569–592.

(2) Wallin E, Von Heijne G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci 7:1029 1038.

(3) Heijne AE (2007) Membrane protein structure: prediction versus reality. Annu Rev Biochem 76:125–140.

(4) Seddon AM, Curnow P, Booth PJ (2004) Membrane proteins, lipids and detergents: not just a soap opera. Biochim Biophys Acta 1666:105–117.

(5) Loll PJ (2003) Membrane protein structural biology: the high throughput challenge. J Struct Biol 142:144–153.

(6) Wiener MC (2004) A pedestrian guide to membrane protein crystallization. Methods 34:364–372.

(7) Zapun A, Bardwell JCA, Creighton TE (1993) The reactive and destabilizing disulfide bond of DsbA, a protein required for protein disulfide bond formation in vivo. Biochemistry 32:5083–5092.

(8) Kuliopulos A, Walsh CT (1994) Production, purification, and cleavage of tandem repeats of recombinant peptides. J Am Chem Soc 116:4599–4607.

(9) Leviatan S, Sawada K, Moriyama Y, Nelson N (2010) Combinatorial method for overexpression of membrane proteins in Escherichia coli. J Biol Chem 285: 23548–23556.

(10) Bernaudat F, Frelet-Barrand A, Pochon N, Dementin S, Hivin P, Boutigny S,

Rioux JB, Salvi D, Seigneurin- Berny D, Richaud P, Joyard J, Pignol D, Sabaty M, Desnos T, Pebay-Peyroula E, Darrouzet E, Vernet T, Rolland N (2011) Heterologous expression of membrane proteins: choosing the appropriate host. PLoS ONE 6: e29191.

(11) Laage R, Langosch D (2001) Strategies for prokaryotic expression of eukaryotic membrane proteins. Traffic 2: 99–104.

(12) Ishihara G, Goto M, Saeki M, Ito K, Hori T, Kigawa T, Shirouzu M, Yokoyama S (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. Protein Expr Purif 41:27–37.

(13) Diefenderfer C, Lee J, Mlyanarski S, Guo Y, Glover KJ (2009) Reliable expression and purification of highly insoluble transmembrane domains. Anal Biochem 384: 274 278.

(14) Thai K, Choi J, Franzin CM, Marassi FM (2005) Bcl-XL as a fusion protein for the high-level expression of membrane-associated proteins. Protein Sci 14:948–955.

(15) Wagner S, Klepsch MM, Schlegel S, Appel A, Draheim R, Tarry M, H€ogbom M, van Wijk KJ, Slotboom DJ, Persson JO, de Gier J-W (2008) Tuning *Escherichia coli* for membrane protein overexpression. Proc Natl Acad Sci USA 105:14371–14376.

(16) Kefala G, Ahn C, Krupa M, Esquivies L, Maslennikov I, Kwiatkowski W, Choe S (2010) Structures of the OmpF porin crystallized in the presence of foscholine- 12. Protein Sci 19:1117–1125.

(17) Cowan SW, Schirmer T, Rummel G, Steiert M, Ghosh R, Pauptit RA, Jansonius JN, Rosenbusch JP (1992) Crystal structures explain functional properties of two *E.coli* porins. Nature 358:727–733.

(18) Eisele JLJ, Rosenbusch JPJ (1990) In vitro folding and oligomerization of a

membrane protein. Transition of bacterial porin from random coil to native conformation. J Biol Chem 265:10217–10220.

(19) Visudtiphole V, Thomas MB, Chalton DA, Lakey JH (2005) Refolding of *Escherichia coli* outer membrane protein F in detergent creates LPS-free trimers and asymmetric dimers. Biochem J 392:375.

(20) Watanabe Y (2002) Characterization of the refolding and reassembly of an integral membrane protein OmpF porin by low-angle laser light scattering photometry coupled with high-performance gel chromatography. J Chromat A 961:137–146.

(21) Surrey T, Schmid A, Jahnig F (1996) Folding and membrane insertion of the trimeric β-barrel protein OmpF. Biochemistry 35:2283–2288.

(22) McLatchie LM, Fraser NJ, Main MJ, Wise A, Brown J, Thompson N, Solari R, Lee MG, Foord SM (1998) RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor. Nature 393: 333–339.

(23) Kusano S, Kukimoto-Niino M, Akasaka R, Toyama M, Terada T, Shirouzu M, Shindo T, Yokoyama S (2008) Crystal structure of the human receptor activity-modifying protein 1 extracellular domain. Protein Sci 17: 1907–1914.

(24) Boise LH, Gonz_alez-Garcı´a M, Postema CE, Ding L, Lindsten T, Turka LA, Mao X, Nu~nez G, Thompson CB (1993) bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. Cell 74: 597–608.

(25) Studier FW (2005) Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41:207–234.

(26) Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C (2012) Prediction of protein secondary structure from circular dichroism using theoretically derived spectra.

Proteins 80:374–381.

(27) Provencher SW (1976) A Fourier method for the analysis of exponential decay curves. Biophys J 16:27–41.

(28) Rogl H (1998) Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilised by nickel chelating chromatography. FEBS Lett 432:21–26.

(29) Mouillac BBJ (2010) Mammalian membrane receptors expression as inclusion bodies in Escherichia coli. Methods Mol Biol 601:39–48.

(30) Ren H, Yu D, Ge B, Cook B, Xu Z, Zhang S (2009) High-level production, solubilization and purification of synthetic human GPCR chemokine receptors CCR5, CCR3, CXCR4 and CX3CR1. PLoS One 4:e4509.

(31) Sanders CR, Kuhn Hoffmann A, Gray DN, Keyes MH, Ellis CD (2004) French swimwear for membrane proteins. ChemBioChem 5:423–426.

(32) Rath A, Glibowicka M, Nadeau VG, Chen G, Deber CM (2009) Detergent binding explains anomalous SDSPAGE migration of membrane proteins. Proc Natl Acad Sci USA 106:1760–1765.

(33) Norden B, Rodger A, Dafforn T (2010) Linear dichroism and circular dichroism : a textbook on polarized-light spectroscopy. Cambridge: Royal Society of Chemistry, p.293.

(34) Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C (2012) Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. Proteins 80:374–381.

(35) Yang Z, Zhang L, Zhang Y, Zhang T, Feng Y, Lu X, Lan W, Wang J, Wu H, Cao C, Wang X (2011) Highly efficient production of soluble proteins from insoluble

inclusion bodies by a two-step-denaturing and refolding method. PLoS One 6:e22981.

(36) Vergus JM, Wiener MC (2011) The variable detergent sensitivity of proteases that are utilized for recombinant protein affinity tag removal. Protein Expr Purif 78:139 142.

(37) Zapun A, Bardwell JCA, Creighton TE (1993) The reactive and destabilizing disulfide bond of DsbA,a protein required for protein disulfide bond formation in vivo. Biochemistry 32:5083–5083.

(38) Gekko K, Kimoto A, Kamiyama T (2003) Effects of disulfide bonds on compactness of protein molecules revealed by volume, compressibility, and expansibility changes during reduction. Biochemistry 42:13746–13753.

(39) Hilairet S, Foord SM, Marshall FH, Bouvier M (2001) Protein-protein interaction and not glycosylation determines the binding selectivity of heterodimers between the calcitonin receptor-like receptor and the receptor activity-modifying proteins. J Biol Chem 276: 29575–29581.

Chapter 5

THE ASSOCIATION BETWEEN CALCITONIN RECEPTOR-LIKE RECEPTOR (CLR) AND RECEPTOR ACTIVITY MODIFYING PROTEIN 1 (RAMP1) TRANSMEMBRANE DOMAINS IS CRITICAL FOR FORMING A FULLY FUNCTIONAL CALCITONIN GENE-RELATED PEPTIDE RECEPTOR

Calciton receptor-like receptor (CLR), a class B G-protein coupled receptor (GPCR), when bound with receptor activity modifying protein 1 (RAMP1), forms a specific calcitonin gene-related peptide (CGRP) receptor. Many complications had been associated with this co-receptor complex and obtaining detail understanding on how CLR and RAMP1 associate is critical for superior antagonist design. In this study, we identified a highly conserved RAMP1 TM P-x-x-x-T motif through genetic sequence alignments. By applying site-directed mutagenesis to substitute the motif residues and various biochemical experimental techniques, we were able to demonstrate the importance of this motif in CLR-RAMP1 association and signaling. Furthermore, our DN-AraTM results on individual CLR and RAMP1 transmembrane (TM) domains suggested CLR TM3, TM6, and TM7 might play a significant role in associating with RAMP1.

5.1   Introduction

Calcitonin receptor-like receptor (CLR) is a class B G-protein coupled receptor (GPCR), a family of proteins which pass through the plasma membrane seven times and represent the primary receptor-type targets for prescription drugs. CLR is associated with numerous disease conditions such as osteoporosis, diabetes, obesity, venous

insufficiency, and various cardiovascular disorders (1). The ligand-specificity of CLR is dependent on its interaction with other accessory membrane proteins (2, 3). CLR alone does not recognize any ligand. Ligand binding occurs only after association with a group of co-receptors called receptor activity modifying proteins (RAMPs) (4). There are three types of RAMPs (RAMP1, RAMP2 and RAMP3), which share less than 30% sequence identity (4). The RAMPs are 148-175 amino acids in size with a cleavable N-terminal signaling peptide and similar topology: a large N-terminal extracellular domain (~100 aa), one transmembrane spanning domain (~20 aa) and a short C-terminal intracellular domain (~10 aa) (5). When associating with RAMP1 in a 1:1 ratio, CLR forms a specific receptor for calcitonin gene-related peptide (CGRP), which is one of the most potent vasodilators discovered thus far (6). Increased cranial circulation of CGRP has been associated with migraines and antagonism of CGRP receptors has become an important target for migraine treatment (7). Nevertheless, current antimigraine treatments such as sumatriptan cause vasoconstriction and other CLR antagonist chemical-based drugs either exhibit poor oral bioavailability (BIBN 4096 BS) or cause liver damage (MK-0974) (7, 8). In order to design more effective migraine medications, it is critical to understand how CLR and RAMP1 interact and how this receptor mediates signal transduction.

While there is substantial information regarding the domains of RAMP1 which are critical for ligand binding, little is known about how it associates with CLR and forms a functional CGRP receptor (9). Previous studies have shown that the RAMP1 extracellular domain alone is sufficient for N-glycosylation, trafficking, and signaling of CLR; however, the half minimum effective concentration (EC_{50}) of CGRP ligand was

increased ~4000 fold compared to the full-length RAMP1 receptor (10). Other studies have shown that the CLR-RAMP1 oligomer is no longer observed on the membrane surface when RAMP1 transmembrane (TM) domain was deleted, whereas there is no effect in signaling when the short RAMP1 intracellular domain is deleted (11). All evidence points toward the direct functional significance of the RAMP1 TM domain in orienting RAMP1 to form the CLR-RAMP1 complex in a ligand specific manner.

In the current study, we identified a highly conserved P-x-x-x-T motif within RAMPs TM domain. Through BRET analysis, a cylic AMP (cAMP) functional assay, TM peptide circular dichroism analysis, and zebrafish gene knockdown and rescue experiments, we were able to demonstrate the significant role this motif plays in CLR-RAMP1 association and signaling. We also examined individual CLR TM and RAMP1 TM domain interaction using DN-AraTM assay and identified CLR TM4, TM6 and TM7 as potential RAMP1 TM association partners.

5.2   Materials and Methods

5.2.1   Subcloning

Unless otherwise stated, standard molecular biology techniques were used. All constructs used were verified by DNA sequencing (Genewiz).

For BRET assay, full-length human RAMP1 receptor sequence was amplified from plasmid (OriGene) and cloned into pGFP^2-N2 as an NheI/KpnI fragment. Full-length human CLR receptor sequence was amplified from plasmid (OriGene) and cloned into pRluc-N2 as an XhoI/KpnI fragment. RAMP1 mutants T126A, T128I,

T130I, and T134I were generated by site-directed mutagenesis using QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies).

For cAMP functional assay, full-length RAMP1 receptor sequence was amplified including a stop codon at the end from plasmid (OriGene) and cloned into pGFP^2-N2 as an NheI/KpnI fragment.

For DN-AraTM assay, RAMP1 TM sequence (I106-V138) was amplified from plasmid (OriGene) and cloned into pAraTMwt as a SacI/KpnI fragment. A flexible poly-proline linker sequence (RQLPTAAPEPAKV) was inserted between RAMP1 and AraC protein to allow proper functioning of the DN-AraTM assay. Individual CLR TM sequence (TM1: Y143-S170, TM2: S170-A197, TM3: V210-L240, TM4: Q250-Y277, TM5: S286-I317, TM6: K333-G358, and TM7: E363-Q392) was amplified from plasmid (OriGene) and cloned into pAraTMDN as an SacI/KpnI fragment. A flexible poly-proline linker sequence (RQLPTAAPEPAKV) was inserted between CLR TM domains and AraC* protein (transcription inactivated form of AraC protein) to allow proper functioning of the DN-AraTM assay.

5.2.2 RAMP TM sequence alignment

RAMP TM sequences were obtained from Protein Data Bank (PDB) and the highly conserved amino acids were identified by Weblogo tool. The statistical significance of the identified conserved motif was analyzed using TMSTAT.

5.2.3 Bioluminescent resonance energy transfer (BRET) assay

8 μg of RAMP1 or mutant RAMP1-GFP and CLR-Rluc fusions were transfected into HEK293 cells by electroporation (GenePulser, Bio-Rad) using the HEK293 preset protocol in HEBS buffer (pH 7.05). Immediately after electroporation, cells were transferred to a white, round-bottom 96 well plate (100 μL of cells per well) in DMEM containing 10% FBS with L-glutamine/VitaMax supplements and 1% Penn/Strep, and incubated at 37°C and 5% CO_2 for 48 h. Media was removed, cells were washed once with 100μL of PBS per well before 100μL of BRET buffer (0.1 g/L $CaCl_2$, 0.1 g/L $MgCl_2$, and 1 g/L D-glucose in PBS) was added to each well. 5μL of Deep Blue C (GoldBio) was added to each sample well, and luminescence measurements (Filter 1: Green and Filter 2: Magenta) were collected using a M200 Infinity plate reader (Tecan) over the course of 10 seconds. The energy transfer efficiency is calculated by dividing the intensity of the signal for the green channel by the intensity of the signal for the magenta channel (12):

Energy Transfer Efficiency = (GFP Signal at 515nm) / (Magenta Signal at 410nm)

The expression level of each eGFP[2]-fused and Rluc-fused RAMP1 and CLR proteins was evaluated through western blotting against anti-GFP and anti-Rluc antibody respectively. Loading control was evaluated through western blotting against anti-tubulin antibody and HEK293 cells expressing empty eGFP[2] and Rluc vectors were used as negative control for BRET.

5.2.4 Cyclic adenosine monophosphate (cAMP) functional assay

8 μg of RAMP1 or mutant RAMP1 and pGloSensor™ -20F cAMP Plasmid (Promega) were transfected into HEK293 cells by electroporation (GenePulser, Bio-Rad) using the HEK293 preset protocol in HEBS buffer (pH 7.05). Immediately after electroporation, cells were transferred to a white, round-bottom 96 well plate (100 μL of cells per well) in DMEM containing 10% FBS with L-glutamine/VitaMax supplements and 1% Penn/Strep, and incubated at 37°C and 5% CO_2 for 24 h. LAR II (firefly luciferase substrate) and Stop & Glo Reagent (rennila luciferase substrate) were prepared as described by Dual-Luciferase Reporter Assay System protocol (Promega). A final concentration of 10^{-5} M αCGRP was added to the cells and allows the cells to incubate for 1h. After 1h, the media was removed. 50μL of LAR II was added to each well and luminescence measurements were collected using a M200 Infinity plate reader (Tecan) over the course of 20 minutes. Following that, 50μL of Stop & Glo Reagent was added to each well and measurements were collected until maximum luminescence is observed. The cAMP signal was calculated by dividing the maximum LAR II luminescent signal by maximum Stop & Glo Reagent luminescent signal.

5.2.5 Circular dichroism analysis of RAMP1 TM peptides

RAMP1 TM wt (KGSILYPFIVVPITVTLLVTALVVWK), RAMP1 TM T128I, and RAMP1 TM T130I peptides were chemically synthesized by GenScript. The peptides were solubilized in 1% w/v fos-choline 15 detergent using a bath sonicator.

Spectra were collected using 200 µL of peptide samples (~1 mg/mL) in 0.1 cm quartz cuvettes (Starna) using a J-815 circular dichroism spectrometer (JASCO). The scanning speed was set to 500 nm/min and measurements collected from 180-260 nm. The spectrum for buffer containing 1% w/v fos-choline 15 was used for background subtraction (13). Raw spectra (mdeg) were converted to molar ellipticity (degree cm^2 dmol^{-1})

5.2.6   Zebrafish RAMP1 gene knockdown and rescue

For RAMP1 gene knockdown, zebrafish embryos were injected with 10 ng of each complementary RAMP1 MO (UTR1 and UTR2), both of which target the 5' UTR region of RAMP1. As controls, we used 2 MOs (scr) with scrambled sequences. For RAMP1 phenotype rescue, zebrafish embryos were injected with RAMP1-MOs followed by RAMP1 mRNA synthesized from plasmid pCS2 to rescue RAMP1 expression. Full-length zebrafish RAMP1 coding sequence was amplified from zebrafish cDNA using specific primers and subcloned into plasmid pCS2 as an EcoRI/XhoI fragment. Using site-directed mutagenesis, we also generated point mutations P126A and T130I in RAMP1-pCS2.

5.2.7   DN-AraTM assay

Plasmids pAraTMDN, pAraTMwt, and pAraGFPCDF were co-transformed into the AraC- deficient *E. coli* strain SB1676 (The *E. coli* Genetic Stock Center at Yale University) and streaked on selective LB plates (100 µg/mL ampicillin, 50 µg/mL

kanamycin, and 100 μg/mL spectinomycin). 1 colony was picked from each construct and grown in 4 mL selective LB media for 16 h at 37°C and 200rpm. After 16 h, the cultures were diluted to A_{600} of 0.5 in three wells on a 2.0 mL 96 deep well PP plate (PlateOne) with each well containing 400 μL of selective LB media (100 μg/mL ampicillin, 50 μg/mL kanamycin, 100 μg/mL spectinomycin, and 1 mM IPTG) and grown for an additional 16 h at 37°C and 300 rpm. 200 μL of cultures from each sample was transferred to a black 96-well, clear bottom plate (Greiner) and a series of 2-fold dilutions was prepared using selective LB media. A_{600} measurements and GFP fluorescence emissions spectra (excitation maximum at 485nm and emissions maximum at 530nm) were collected using a M200 Infinity plate reader (Tecan). Results are reported as the ratio of fluorescence emission at 530 nm to absorbance at 560 nm and normalized to the negative control (pAraGFPCDF transformed cells) (14).

5.3   Results

5.3.1   *Identification of a conserved RAMP1 TM motif through sequence alignments*

Through a cross-species TM sequence analysis (Figure 5.1), we were able to identify a highly conserved TM domain P-x-x-x-T motif. TMSTAT, an online program developed by Engelman lab which compares a provided TM sequence to a membrane protein database in order to determine the statistical significance and frequency of the given TM sequence, found that the P-x-x-x-T motif was over-represented in known TM helices (Odds ratio=1.07), indicating that this motif may play a significant role in TM receptor function (15).
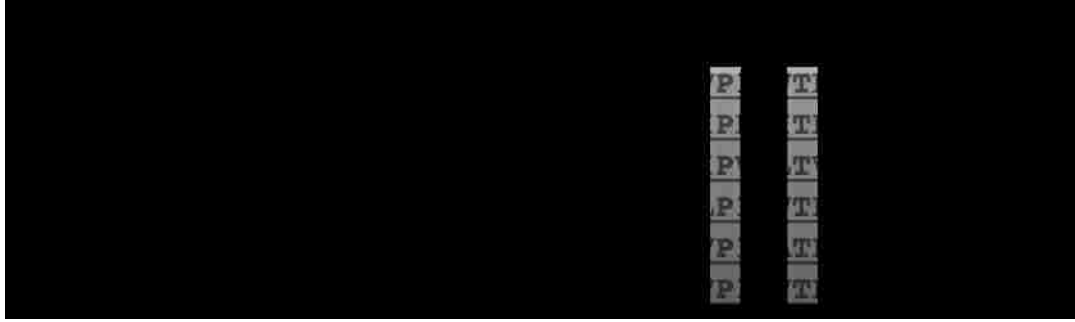
Figure 5. 1. RAMP TM domain sequence alignments. The P-x-x-x-T motif is highly conserved among human RAMPs and across species, with the registry of the motif within the transmembrane region varying among RAMP family members.

5.3.2 Examination of the conserved RAMP1 TM motif in CLR-RAMP1 association through BRET assay

BRET is an *in vivo* mammalian cell-based method used to study receptor dimerization. We applied BRET as a mean to observe how mutations within this conserved P-x-x-x-T motif affect CLR-RAMP1 heterodimerization. In the case of studying CLR-RAMP1, the CLR receptor is fused to RLuc and the RAMP1 or mutant RAMP1 (single amino acid substitution within the conserved P-x-x-x-T motif) receptor is fused to GFP[2]. Conservative amino acid mutations were made such that the size variation between the native and the substituted side-chains was minimized, for example proline was mutated to alanine and threonine to isoleucine.

121
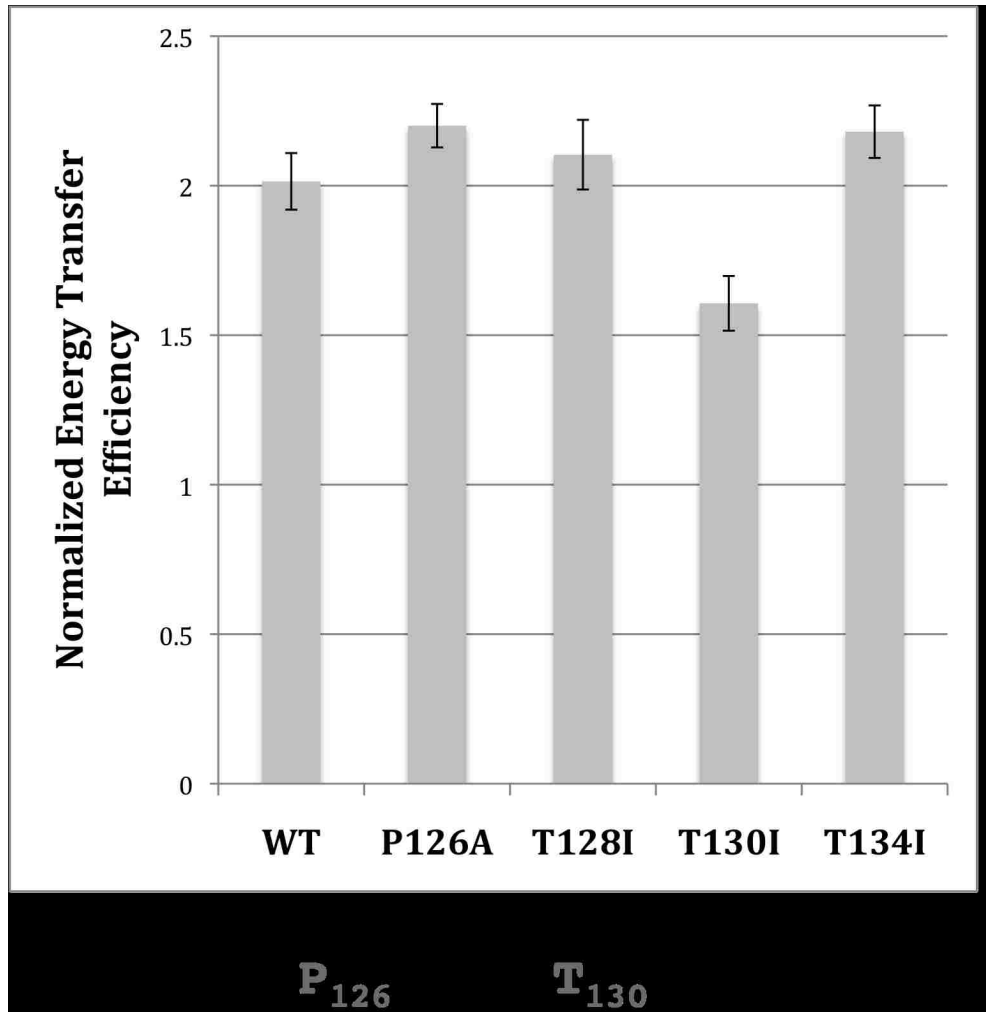
Figure 5. 2. RAMP1 T130I mutant disrupts CLR-RAMP1 association in HEK cells as shown by BRET. All RAMP1 mutations includin P126A, T128I, and T134I had similar BRET energy transfer efficiency as the RAMP1 wt. However, a ~50% reduction in energy transfer efficiency was observed in RAMP1 T130I mutant. Experiments were repeated three times in triplicate, and error bars represent standard deviation (wt: RAMP1 wt, P126A: RAMP1 P126A, T128I: RAMP1 T128I, T130I: RAMP1 T130I, and T134I: RAMP1 T134I).

By analyzing the resulting BRET energy transfer efficiency with different RAMP1 mutants (Figure 5.2), we found that only the T130I mutation (~50% reduction) but not the control mutants (T128I and T134I) showed a significant decrease in the energy transfer ratio indicative of the disruption of this mutation to CLR-RAMP1 association which agrees well with our prediction. Interesting enough, the P126A had no affect on

122

energy transfer ratio. However, BRET only indicates how close the two receptors are, it does not provide any information on whether the receptor is functional.

5.3.3 *Examination of the conserved RAMP1 TM motif in CLR-RAMP1 signaling through cAMP functional assay*

cAMP is a second messenger used by CLR-RAMP1 for intracellular signal transduction (10). Studies by Mishima's group discovered increased levels of cAMP in patients under migraine attack (16). Therefore, variation of cAMP levels in the cell could be used as an indication of *in vivo* CLR-RAMP1 activity and possibly migraine states. To determine the functional significant of the P-x-x-x-T motif in CLR-RAMP1 signaling, we transiently expressed RAMP1 or mutant RAMP1 in HEK293 cells, and as well as a plasmid that produces a firefly luciferase fusion protein (FLuc). Binding of this FLuc with cAMP results in an increase in luminescence and by adding a FLuc substrate and inducing with CGRP, the cAMP signal being transduced by the functional receptor formed by RAMP1 or mutant RAMP1 with endogenously expressed CLR can be monitored in the system. As shown in Figure 5.3, before induced by CGRP, all the samples have similar basal cAMP level. After addition of 10^{-5} M CGRP agonist, we observed a significant increase (~1.8 fold increase as compared to the non-ligand induced state) in cAMP response for RAMP1 wt, P126A, and T128I samples, but a similar cAMP response profile as the negative control (~1.3 fold increase as compared to the non-ligand induced state) for RAMP1 T130I sample was observed. The reduction in RAMP1 T130I cAMP response is consistent with our prediction of the significance

of this threonine within the conserved RAMPs TM motif; however, we see no

significant penalty in signaling by mutating the RAMP1 P126 residue.
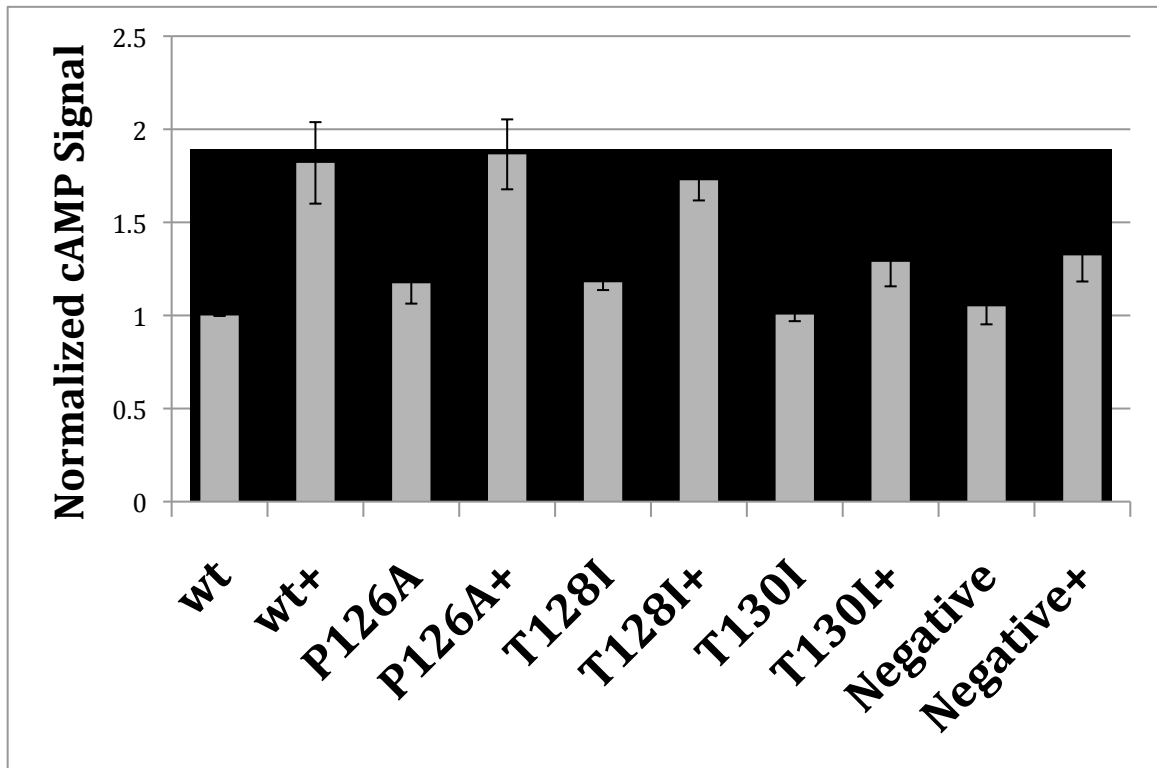
Figure 5. 3. RAMP1 T130I exhibits a significant disruption in receptor signaling when induced with peptide agonist αCGRP. Experiments were repeated three times in quadruplicate, and error bars represent stand error of the mean (wt: RAMP1 wt, P126A: RAMP1 P126A, T128I: RAMP1 T128I, T130I: RAMP1 T130I, and +: induced with 10^{-5} αCGRP).

5.3.4 *Zebrafish RAMP1 gene knockdown and rescue*

The expression of RAMP1 receptor in zebrafish was repressed during embryo

development using the anti-sense morpholinos (MOs) where the morpholino oligos

bound to the 5'-untranslated region of RAMP1 messenger RNA and interfered

ribosomal initiation complex from progressing through the 5' cap to the start codon. As

shown in Figure 5.4, 24 hours post fertilization (hpf), we observed numerous cardiovascular defects in zebrafish embryos injected with specific MOs to knockdown RAMP1 expression compared to the scrambled randomized MOs (scr) injected zebrafish embryos. These defects include cardiac edema, blood pooling through the heart chamber, and lack of heart chamber formation. Moreover, the tubule is still able to function by pumping blood through the vasculature, but at a significantly reduced rate. By 48 hpf, the larvae have severe malformations and gross morphological defects, and do not remain viable beyond 48-72 hpf. Our next step was to rescue the phenotype by introducing RAMP1 mRNA into the embryos. As shown in Figure 5.5, when we introduced the wt RAMP1 mRNA to the RAMP1-MO knockdown zebrafish embryo, the zebrafish was able to have fully developed healthy cardiovascular system similar to the zebrafish embryos injected with scr MOs. In contrast, when TM mutant RAMP1 mRNAs (P126A, T130I) were introduced to the embryos, neither of them was able to rescue the RAMP1-MO phenotype. Both RAMP1 mutants resulted in enlarged and unhealthy cardiovascular system zebrafish phenotype (Figure 5.5).

10ng ctrl scr MO   **10ng UTR1 MO + UTR2 MO**
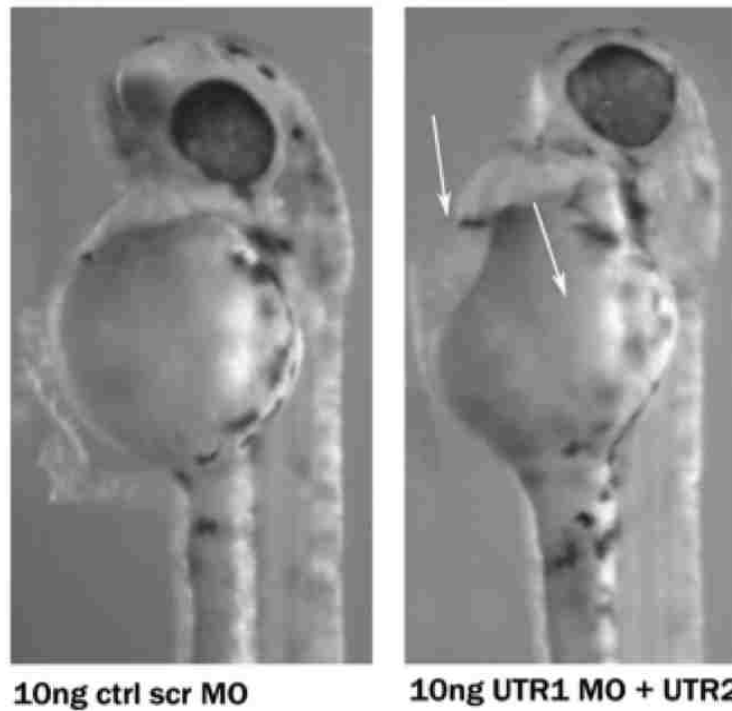
Figure 5. 4. Cardiovascular defects observed in RAMP1-MO treated zebrafish at 24 hpf. Embryos were injected with 2 complementary MOs targeting the 5' UTR (untranslated region) of RAMP1 and compared to controls using scrambled MO sequences (scr). Larvae exhibited numerous cardiovascular defects at 24 hpf, including cardiac edema, blood pooling throughout the heart chamber and lack of heart chamber formation.
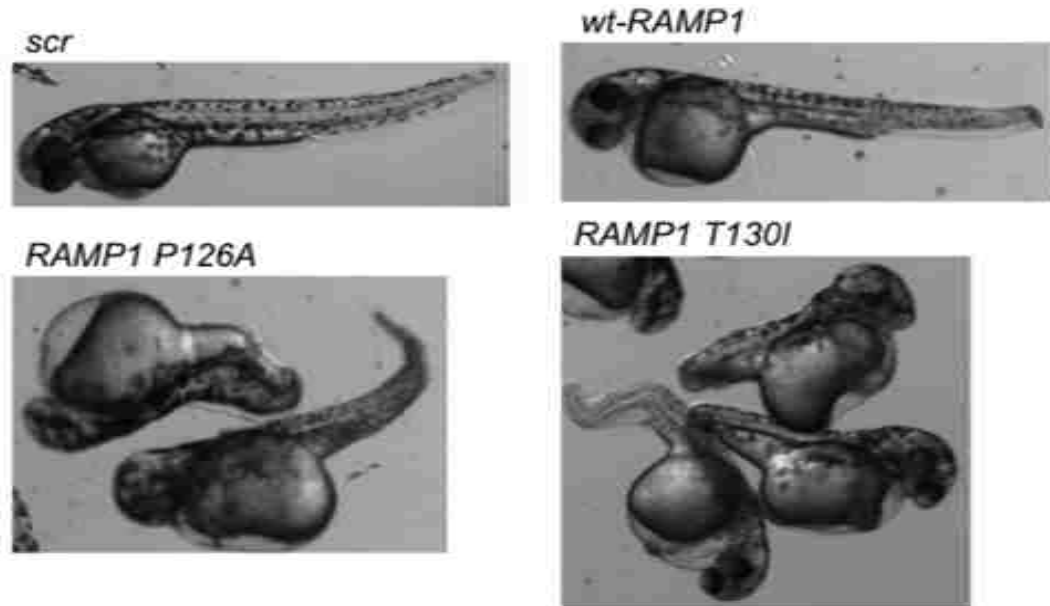
RAMP1: VRDPP⌈GSILYPFIVV**P**ITV**T**LLVTALVVW⌉QSKRTEGIV

Figure 5. 5. Rescue of cardiovascular defects observed in RAMP1-MO larvae with wild-type RAMP1 mRNA: Wild-type RAMP1 and point mutations P126A and T130I were generated, injected and imaged at 24 hpf. Both mutations occur in the TM region of RAMP1, and are predicted to disrupt CLR-RAMP1 complex formation. Strikingly, both mutations are unable to rescue the RAMP1-MO phenotype, in contrast to wt-RAMP1.

5.3.5   Circular dichroism analysis of RAMP1 TM peptides

To examine the contribution of threonine mutation to overall RAMP1 TM secondary structure, we chemically synthesized RAMP1 TM peptides (wt, T128I, and T130I) and solubilized the peptides in 1% fos-choline w/v 15 micelles, which was previously shown to solubilize full-length RAMP1 protein and retain its native folded secondary and tertiary structure (13). The secondary structure of each RAMP1 TM peptide in fos-choline 15 micelles was determined via circular dichroism. RAMP1 is a type I single-span membrane protein which means its TM domain is expected to have an α-helical secondary structure. As shown in Figure 5.6, RAMP1 TM wt resembles a

semi-α-helical structure. Whereas, both RAMP1 TM T128I and T130I appear to have

majorly β-sheet secondary structure with a single minimum at around 218nm (T128I at

217nm and T130I at 215nm).
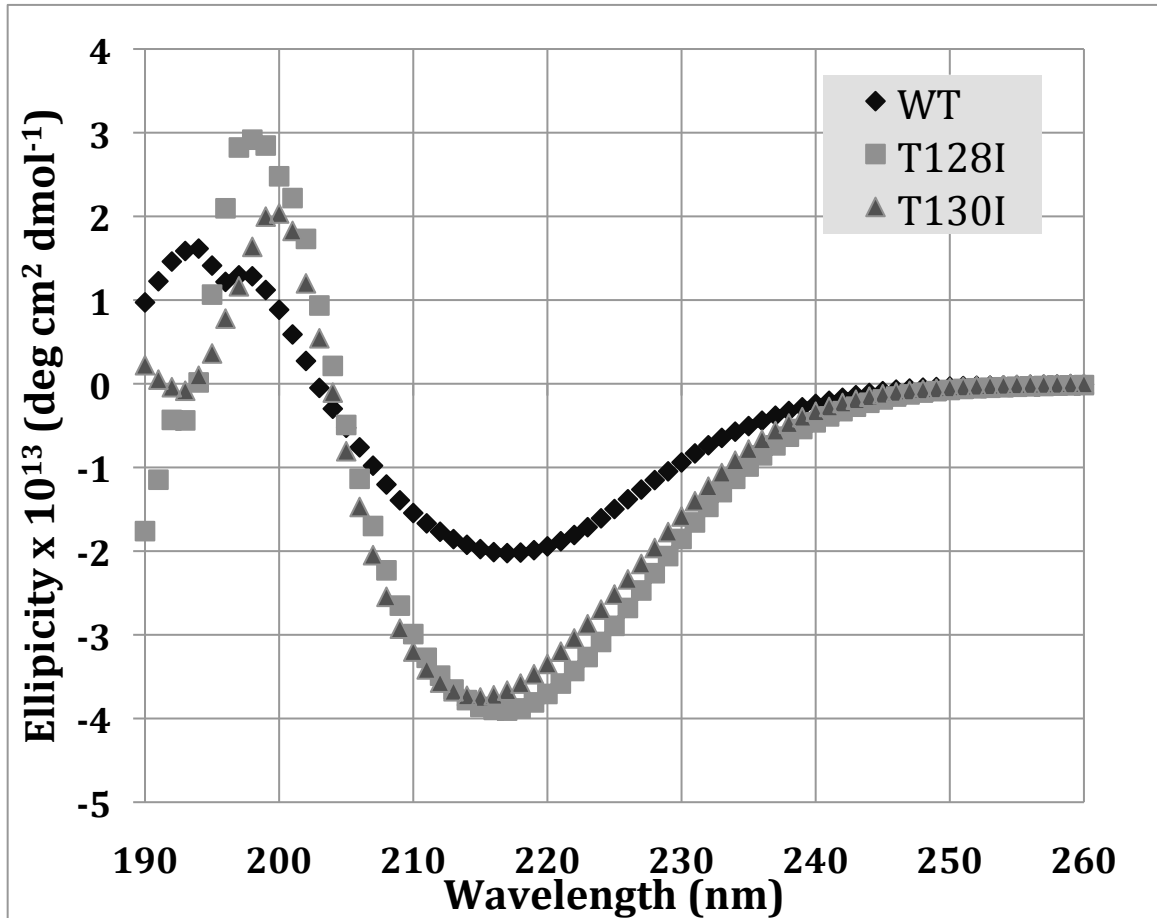
Figure 5. 6. Circular dichroism spectrum for RAMP1 TM peptides. Both RAMP1 threonine mutant peptides had a more defined β-sheet secondary structure than the RAMP1 wt peptide (wt: RAMP1 wt, T128I: RAMP1 T128I, and T130I: RAMP1 T130I).

5.3.6   DN-AraTM assay analysis on CLR-RAMP1 interacting TM domains

DN-AraTM assay is an *E. coli* transcription factor AraC based dominant-negative system that looks at isolated protein domain interaction and correlates amount of heterodimer being formed to observed GFP signal. Homodimerization of the reference protein domain gives rise to baseline GFP protein. When the competitor protein domain interacts with the reference protein domain, this results in a decrease in GFP signal as a consequence of reduction in homodimerization of the reference protein domain. We used RAMP1 TM-CYTO domain as the reference protein and individual CLR TM domains (TM1-TM7) as the competitor proteins in this experiment. First, we observed a strong baseline GFP when we expressed RAMP1 TM-CYTO domain indicative of strong homodimer formation (~2 times GFP signal of the negative control). Then we introduced individual CLR TM domain as competitor into the system. As shown in Figure 5.7, we observed decrease in GFP signals in all samples. Among the CLR TM competitors, TM3, TM6, and TM7 had the most significant GFP signal drop, indicative of these CLR domains having a higher tendency to form a heterodimer with the RAMP1 TM-CYTO domain.
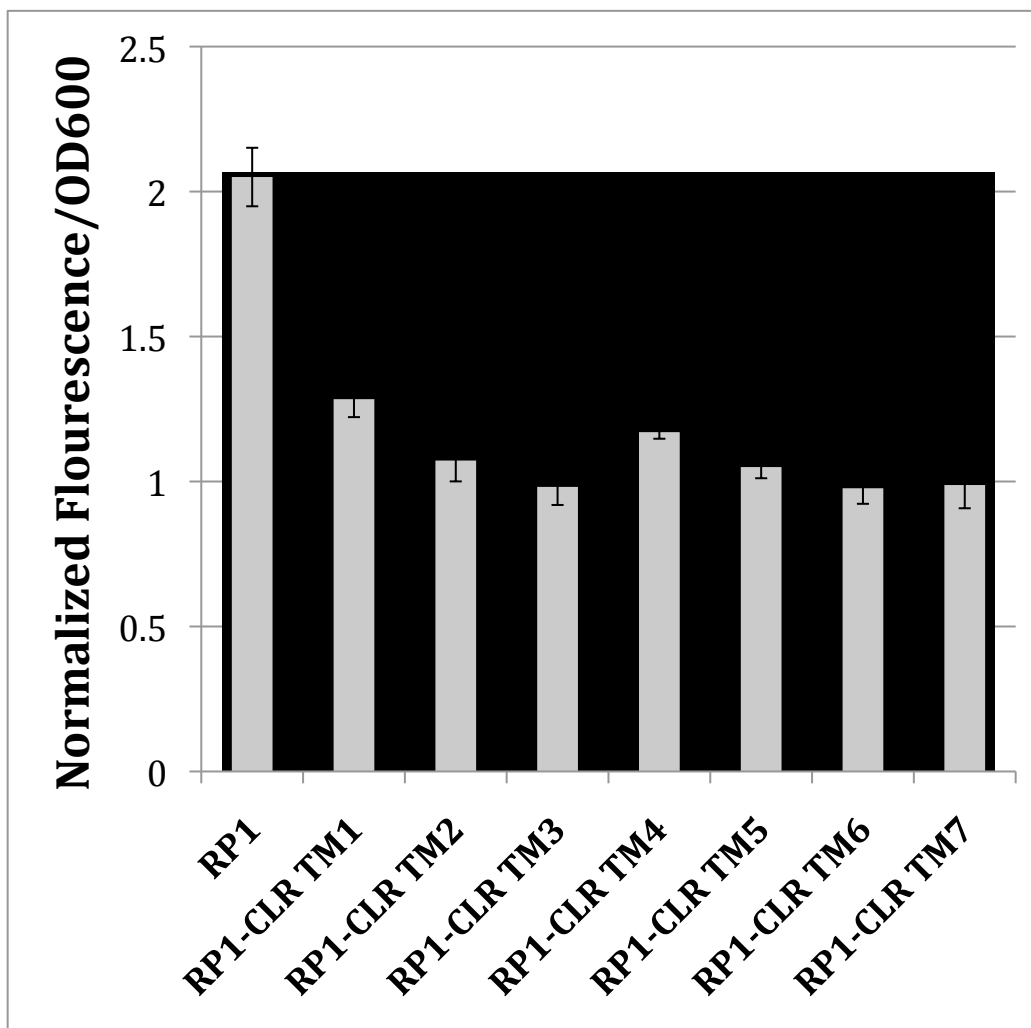
Figure 5. 7. DN-AraTM results for RAMP1 and CLR TM domains. Although all CLR TM domains forms heterodimer with RAMP1 TM (reduction in fluorescence compared to expressing RAMP1 TM alone), CLR TM3, TM6, and TM7 had the most reduction. Experiments were repeated three times in quintuplicate, and error bars represent stand error of the mean.

5.4   Discussion

Observing a proline residue in a primarily α-helix TM domain of protein is interesting because proline has a ring structure side chain which lacks the backbone amine group necessary for creating a hydrogen bond with the backbone carbonyl group of a residue one turn (or four residues) later, resulting in an unstable kink in the helix.

Hydrogen bonding is the primary force required to stabilize the membrane protein's α-helical structure, therefore, there must be some kind of force to compensate for the lost in structural stability caused by the proline residue. One possibility is the contribution of stability from the surrounding residues. Studies had shown that threonine side-chain hydroxyl group often forms an additional hydrogen bond to the peptide backbone (17). α-helices exhibit a repeated i+4 hydrogen bonding, resulting in every 5th amino acid residue residing on the same side of the helix, as such, proline and threonine in the P-x-x-x-T motif of RAMP1 should be on the same face of the helix. The ability of threonine to form an additional hydrogen bond could provide structural stability that compensates for the resulting helical instability introduced by the proline residue. As a consequence, both proline and threonine residues in the P-x-x-x-T motif reinforce the unique kinked structure that could a play significant role in CLR-RAMP1 interaciton. In addition, studies by the Fleming group indicate that both proline and threonine have a high tendency to face the protein interior with a higher packing value, which is a measure of occluded surface area, in TM helices; hence, these residues may contribute to formation of a tight interface, which may be critical for CLR-RAMP1 interaction (18).

The results of BRET analysis of RAMP1 P-x-x-x-T mutations revealed that only the T130I mutation, and not P126A, affected CLR-RAMP1 association. This is surprising due to the fact that the proline side chain is the cause for the kinked structure. If the mutation of threonine resulted in disruption of CLR-RAMP1 association is real, we should see consequences in CLR-RAMP1 signaling when we introduce the mutations. Therefore, we employed the cAMP functional assay to examine the effect of the RAMP1 mutation on signaling. Interesting enough, there was no effect in signaling

131

with the proline substitution, but a significant signaling disruption was observed with the threonine substitution, which is consistent with our BRET results. The considerably moderate change in magnitude of the cAMP signal may be due to the fact that we were using endogenously expressed CLR and receptor component protein in pairing with our wt and mutant RAMP1 for cAMP activation experiments. Therefore, the signaling response we observed was limited to the amount of endogenous receptors found in HEK293 cells.CGPR receptor is widely distributed in mammalian cardiovascular system and expression of RAMP1 and CLR were observed during the development of rat heart (19, 20). These studies suggested RAMP1 could play a critical role in early stage cardiovascular system development in the vertebrates. Therefore, if the function of RAMP1 receptor was disrupted, we should observe a defect in cardiovascular system. We used zebrafish as an *in vivo* model system to monitor the role RAMP1 receptor plays in cardiovascular development due to its transparent and quick developing nature as well as its ease to manipulate genetically. Our results showed that only the wt RAMP1 but not the TM P-x-x-x-T motif residue mutated RAMP1s (P126A, T130I) were able to rescue the cardiovascular defect phenotype in the RAMP1-MO knockdown zebrafish. Again, threonine residue within the conserved motif was important for proper functioning of RAMP1. In contrary to our BRET and cAMP, the proline residue actually plays a role in RAMP1 function in this case. The study done by Sexson's group has shown that human RAMP1 can also interact with other receptors such as the most common variant of the human calcitonin receptor (hCTR2) (21). The inability to rescue cardiovascular defect we observed by the RAMP1 proline mutation mRNA could be a result of interference with the function of other co-receptors rather

than the CLR receptor, considering zebrafish embryo development is an incredibly complex system.

Our next step was to investigate whether this threonine residue contributes to RAMP1 TM domain structure. We solubilized the synthesized RAMP1 TM peptides in 1% w/v fos-choline and analyzed their secondary structure using circular dichroism. Compared to wt RAMP1 TM peptide, both RAMP1 TM T128I and T130I peptides resembled much more β-sheet like structure. Though we could not fully reconstitute the α-helical structure of RAMP1 TM wt peptide, it is more α-helical-like compared to the mutant peptides which implies the threonine plays a important role in stabilizing the secondary structure of the proline kink.

All evidence indicates that the P-x-x-x-T motif is important in maintaining the specific structure necessary for CLR-RAMP1 signaling. However, the proline residue substitution had no influence in CLR-RAMP1 signaling. This could be explained by the evolution hypothesis of the TM proline kinks in GPCRs which states that because this kinked structure is so important for function, the surrounding amino acids acclimate to this specific structure such that even without the proline residue, the structure is still retained (22). This can also explains why CLR-RAMP1 receptor is very sensitive to threonine substitution.

Various studies have shown that TM domain swapping between human RAMP1, RAMP2, and RAMP3 had minimum to moderate effect in ligand binding affinity and signaling (9, 23). This suggests the interaction between RAMPs and CLR through their TM domains is similar. The highly conserved P-x-x-x-T motif exists in all human RAMPs and mutagenesis experiments demonstrated the consequence in CLR-

RAMP1 signaling when this particular structure is disrupted. Moreover, the two-domain binding model of CGRP receptor proposed high-affinity binding of CGRP C-terminal with N-terminal regions of CLR and RAMP1 (24). This implies that the shape complementation resulting from the proline kink may be the driving force for CLR and RAMP1 TM domain association. Therefore, the existence of a common structural feature between TM domains of RAMPs that dominates CLR-RAMPs interaction is highly possible.

5.5 Conclusion

Though the current study does not provide a complete picture, it is a start for understanding how CLR and RAMP1 assembles and forms a functional CGRP receptor through their TM domains. Future work such as cysteine mutagenesis and cross-linking between RAMP1 and CLR could be used to map out specific TM interaction domains and interfaces.

5.6 References

(1) Barwell, J., Gingell, J. J., Watkins, H. A., Archbold, J. K., Poyner, D. R., and Hay, D. L. (2012) Calcitonin and calcitonin receptor-like receptors: common themes with family B GPCRs?, *Br J Pharmacol 166*, 51-65.

(2) Wunder, F., Rebmann, A., Geerts, A., and Kalthof, B. (2008) Pharmacological and kinetic characterization of adrenomedullin 1 and calcitonin gene-related peptide 1 receptor reporter cell lines, *Mol Pharmacol 73*, 1235-1243.

(3) Roh, J., Chang, C. L., Bhalla, A., Klein, C., and Hsu, S. Y. (2004) Intermedin is a

calcitonin/calcitonin gene-related peptide family peptide acting through the calcitonin receptor-like receptor/receptor activity-modifying protein receptor complexes, *J Biol Chem 279*, 7264-7274.

(4) McLatchie, L. M., Fraser, N. J., Main, M. J., Wise, A., Brown, J., Thompson, N., Solari, R., Lee, M. G., and Foord, S. M. (1998) RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor, *Nature 393*, 333-339.

(5) Hilairet, S., Belanger, C., Bertrand, J., Laperriere, A., Foord, S. M., and Bouvier, M. (2001) Agonist-promoted internalization of a ternary complex between calcitonin receptor-like receptor, receptor activity-modifying protein 1 (RAMP1), and beta-arrestin, *J Biol Chem 276*, 42182-42190.

(6) Brain, S. D., Williams, T. J., Tippins, J. R., Morris, H. R., and MacIntyre, I. (1985) Calcitonin gene-related peptide is a potent vasodilator, *Nature 313*, 54-56.

(7) Ho, T. W., Ferrari, M. D., Dodick, D. W., Galet, V., Kost, J., Fan, X., Leibensperger, H., Froman, S., Assaid, C., Lines, C., Koppen, H., and Winner, P. K. (2008) Efficacy and tolerability of MK-0974 (telcagepant), a new oral antagonist of calcitonin gene-related peptide receptor, compared with zolmitriptan for acute migraine: a randomised, placebo-controlled, parallel-treatment trial, *Lancet 372*, 2115-2123.

(8) Iovino, M., Feifel, U., Yong, C. L., Wolters, J. M., and Wallenstein, G. (2004) Safety, tolerability and pharmacokinetics of BIBN 4096 BS, the first selective small molecule calcitonin gene-related peptide receptor antagonist, following single intravenous administration in healthy volunteers, *Cephalalgia 24*, 645-656.

(9) Qi, T., Simms, J., Bailey, R. J., Wheatley, M., Rathbone, D. L., Hay, D. L., and Poyner, D. R. (2009) Structure-function analysis of RAMP1-RAMP3 chimeras, *Biochemistry 49*, 522-531.

(10) Fitzsimmons, T. J., Zhao, X., and Wank, S. A. (2003) The extracellular domain of receptor activity-modifying protein 1 is sufficient for calcitonin receptor-like receptor function, *J Biol Chem 278*, 14313-14320.

(11) Steiner, S., Muff, R., Gujer, R., Fischer, J. A., and Born, W. (2002) The transmembrane domain of receptor-activity-modifying protein 1 is essential for the functional expression of a calcitonin gene-related peptide receptor, *Biochemistry 41*, 11398-11404.

(12) Bonde, M. M., Hansen, J. T., Sanni, S. J., Hauns, S., Gammeltoft, S., Lyngs, C., and Hansen, J. L. (2010) Biased Signaling of the Angiotensin II Type 1 Receptor Can Be Mediated through Distinct Mechanisms, *PLoS ONE 5*.

(13) Su, P. C., Si, W., Baker, D. L., and Berger, B. W. (2013) High-yield membrane protein expression from E. coli using an engineered outer membrane protein F fusion, *Protein Sci 22*, 434-443.

(14) Su, P. C., and Berger, B. W. (2013) A novel assay for assessing juxtamembrane and transmembrane domain interactions important for receptor heterodimerization, *J Mol Biol 425*, 4652-4658.

(15) Russ, W. P., and Engelman, D. M. (2000) The GxxxG motif: a framework for transmembrane helix-helix association, *J Mol Biol 296*, 911-919.

(16) Mishima, K., Takeshima, T., Shimomura, T., Okada, H., Kitano, A., Takahashi, K., and Nakashima, K. (1997) Platelet ionized magnesium, cyclic AMP, and cyclic GMP levels in migraine and tension-type headache, *Headache 37*, 561-564.

(17) Orzaez, M., Salgado, J., Gimenez-Giner, A., Perez-Paya, E., and Mingarro, I. (2004) Influence of proline residues in transmembrane helix packing, *J Mol Biol 335*, 631-640.

(18) Eilers, M., Shekar, S. C., Shieh, T., Smith, S. O., and Fleming, P. J. (2000) Internal packing of helical membrane proteins, *Proc Natl Acad Sci U S A 97*, 5796-5801.

(19) Li, J., and Wang, D. H. (2005) Development of angiotensin II-induced hypertension: role of CGRP and its receptor, *J Hypertens 23*, 113-118.

(20) Dvorakova, M., Haberberger, R. V., Hagner, S., McGregor, G. P., Slavikova, J., and Kummer, W. (2003) Expression and distribution of the calcitonin receptor-like receptor in the developing rat heart, *Anat Embryol (Berl) 207*, 307-315.

(21) Christopoulos, G., Perry, K. J., Morfis, M., Tilakaratne, N., Gao, Y., Fraser, N. J., Main, M. J., Foord, S. M., and Sexton, P. M. (1999) Multiple amylin receptors arise from receptor activity-modifying protein interaction with the calcitonin receptor gene product, *Mol Pharmacol 56*, 235-242.

(22) Yohannan, S., Faham, S., Yang, D., Whitelegge, J. P., and Bowie, J. U. (2004) The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors, *Proc Natl Acad Sci U S A 101*, 959-963.

(23) Zumpe, E. T., Tilakaratne, N., Fraser, N. J., Christopoulos, G., Foord, S. M., and Sexton, P. M. (2000) Multiple ramp domains are required for generation of amylin

receptor phenotype from the calcitonin receptor gene product, *Biochem Biophys Res Commun 267*, 368-372.

(24) Moore, E. L., and Salvatore, C. A. (2011) Targeting a family B GPCR/RAMP receptor complex: CGRP receptor antagonists and migraine, *Br J Pharmacol 166*, 66-78.

Chapter 6
CONCLUSIONS AND FUTURE PERSPECTIVES

The focus of this thesis is developing novel tools to resolve the missing gap in membrane protein research and investigate the structural basis for membrane receptor signaling for a class B G-protein coupled receptor (GPCR) system: calcitonin receptor-like receptor (CLR) and receptor activity modifying protein 1 (RAMP1). Overall, we developed novel assays for straightforward cooperative membrane protein domain interaction analysis (AraTM and DN-AraTM assays) and large quantity membrane protein sample preparation for high-resolution structural studies (pOmpF). Furthermore, by applying these novel tools along with classical methods to CLR and RAMP1, we gained new insights into how these receptors heterodimerize and form a functional receptor for calcitonin gene-related peptide (CGRP). Obtaining a full picture on how membrane proteins interact with other proteins, are activated by their ligands, and signal through conformational changes is a rapidly developing field due to the significant roles they play in various biological processes and as important pharmaceutical targets. Therefore, there are plenty of opportunities and future work worth exploring.

6.1 Application of Modified AraTM Assay on Influenza A Virus M2 Ion Channel Protein (AM2) Inhibitor Development

In Chapter 2, we designed an *E. coli* transcription factor-based assay, AraTM assay, which expresses a chimera protein (contains a maltose-binding protein (MBP) and a truncated AraC protein that is connected by a protein domain of interest) in which

when homodimerized leads to activation of *araBAD* promoter and production of green fluorescent protein (GFP). Thus, by quantifying the GFP signal, the AraTM assay allows for the investigation of coordinated interplay between different domains of membrane receptors in forming the specific interface required for functioning. In human, there are a large number of membrane proteins in which their function is governed by their homooligomeric states and how tightly the oligomer interfaces are packed. One particularly important example is the influenza A virus M2 ion channel protein (AM2). AM2 is a viral integral membrane protein that forms a homotetramer proton channel that plays an important role during the process of viral infection in humans (1). After the virus enters the human host by endocytosis, the AM2 proton channel initiates endosomal acidification by pumping H^+ ions into the cells. This acidification leads to the dissociation of matrix proteins from viral ribo nucleo proteins (RNP) which eventually results in viral replication in the host (2).

Blockage of AM2 proton channel by amantadine had been an effective treatment for influenza infection for the past couple decades (3). However, recent years the number of naturally occurring amantadine-resistant influenza A (H1N1 and H3N2) have increased drastically (4). Thus, it is essential to design new AM2 inhibitors that could block these variant influenza A viruses with amantadine-resistant AM2. Functional tests of AM2 such as electrophysiological readings using two-electrode voltage-clamp apparatus are not suitable for drug screening due to the complicity of the experiment (5). An ideal solution for drug screening would be a tool that is simple in implementation with a high throughput reporting ability. One possible future direction

for the application of AraTM is using the AraTM assay coupled with ratiometric pHluorin as a pH reporter to design a superior tool for AM2 inhibitor screening (6).

In the modified AraTM assay, there are two parts: the basic AraTM assay (described in Chapter 2) and the ratiometric pHluorin reporter. The basic AraTM assay serve two purposes: one is to investigate the contribution of instability or stability of different naturally occurring amantadine-resistant AM2 mutations to the AM2 homotetramer. The other is to verify the expressed AM2 protein are assembled and oligormized in their native states. As shown in Figure 6.1, various literature results on AM2 mutant homotetramer stability studies are in excellent agreement with our AM2 mutant AraC assay results (7). Particularly, both amantadine-resistant M2 mutations, L26F and S31N, showed stabilizing effects in homotetramer formation. This illustrates the AM2 proteins expressed in our AraTM assay behave as native-like structures. For the reporter, we would utilize ratiometric pHluorin as an *in vivo* approach that monitors *E. coli* intracellular pH which is directly linked to AM2 functions. We take advantage of the fact that ratiometric pHluorin has distinctive fluorescent spectrum properties dependent on the pH of the environment (Fig. 6.2). By adjusting the extra- and intracellular pH gradient that forces proton pumping by AM2, we could create a system that mimics the *in vivo* AM2 functional settings and is ideal for rapid AM2 inhibitor screening (6).
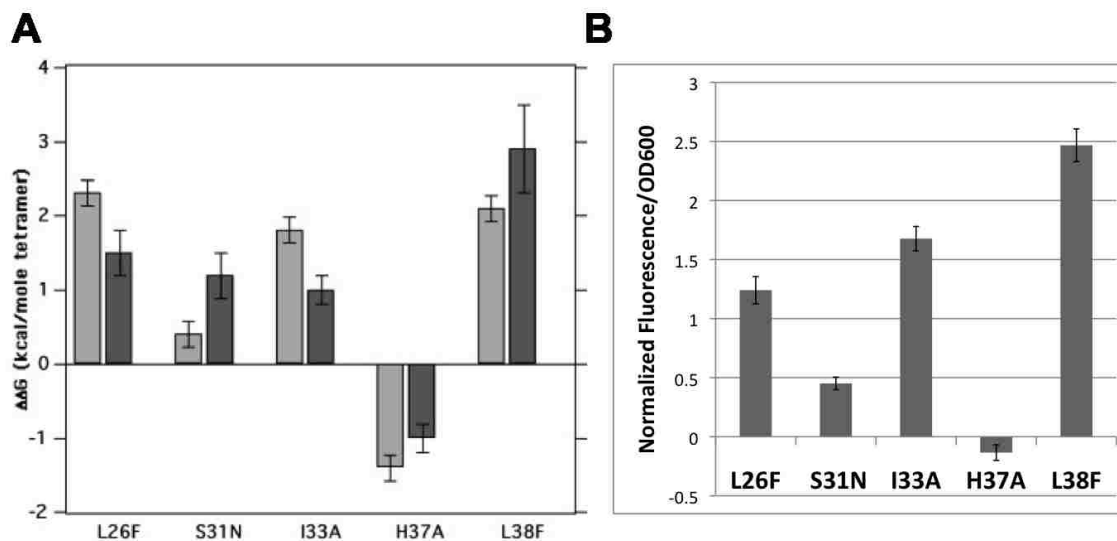
Figure 6. 1. The consequences of AM2 mutations in tetramer dissociation are similar between previous published data and our AraTM AM2 data. A, The change in free energy of tetramer dissociation upon mutation in micelles (gray) and vesicles (black) is shown. L26F, S31N, I33A, and L38F mutations stabilize tetramer formation in both environments, whereas H37A is destabilizing (7) B, The normalized AraTM results for AM2 mutations. Error bars represent standard deviations.
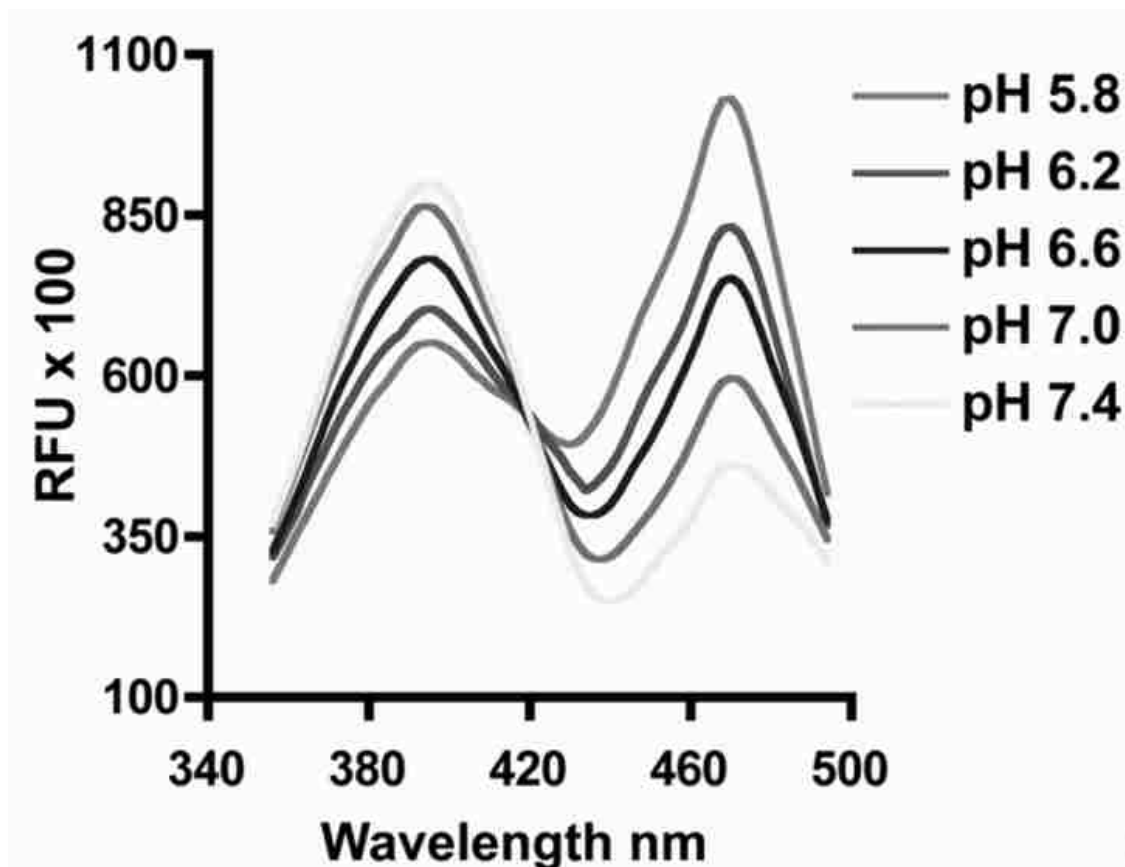
Figure 6 2. Excitation scans of cytoplasmic pHluorin2 expressed in HEK293 cells clamped at the indicated pH using K^+ ions and nigericin are shown. Data are representative of three independent experiments (6).

6.2 Identification of Critical Residues that Governs Receptor for Advanced Glycation Endproducts (RAGE) Homodimerization

RAGE is a multiple ligand receptor that is associated with various chronic pathologies, including cancer, neurodegeneration, atherosclerosis, and diabetic complications (8). Homodimerization of RAGE is essential for proper RAGE signaling and leads to various inflammation-related cell responses (9, 10). In order to treat diseases associated with RAGE, one option is to block RAGE mediated signaling via homodimerization disruption. In order to accomplish this goal, first we need to identify

143

important residues that drive homodimerization of RAGE. Therefore, one future direction we can pursue is to identify the residues that govern RAGE homodimerization which can be later used for the drug design target.

From our AraTM results on different RAGE truncations, we discovered that homodimerization of RAGE was mediated by the juxtamembrane region within the cytosolic tail (A375-P394). If we take a look at the sequence within this region (Fig. 2.7), there are a lot of charged amino acid residues including many glutamic acid residues. In Chapter 2, we mentioned that it is possible that this glutamic acid rich region could adopt coiled-coil like structures that mediates protein-protein interaction. Therefore, we thought this is a good starting point. We made several point mutations on these charged residues within RAGE cytosolic tail and tested with AraTM Assay. As shown in Figure 6.3, our preliminary results showed that none of these mutations were sufficient to disrupt RAGE homodimerization. It is possible that we haven't found the "hot spot" for RAGE homodimerization. Therefore, a comprehensive mutation screening could be done to identify the "hot spot".
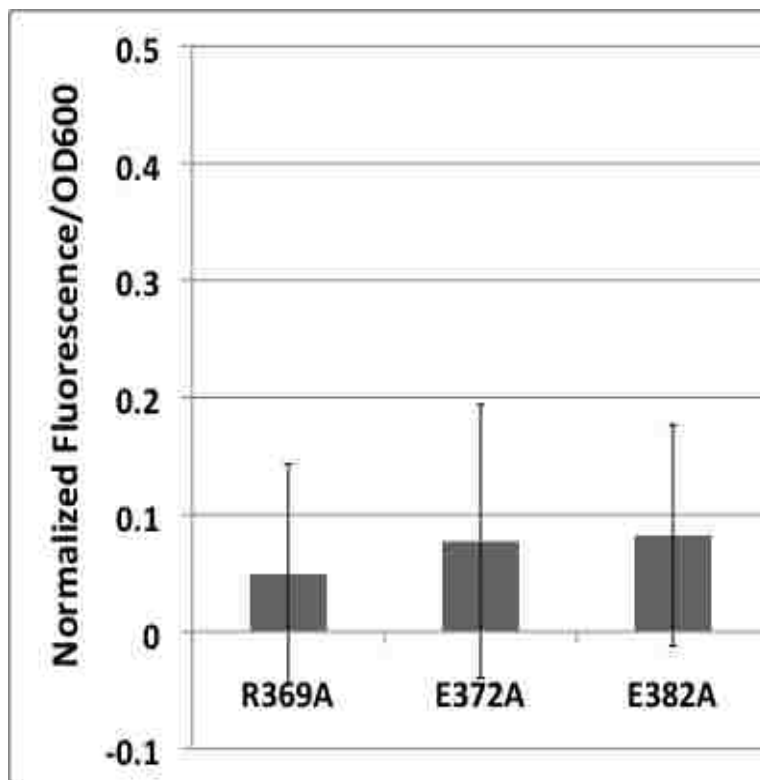
Figure 6. 3. AraTM results on point mutations within RAGE cytosolic tail. Point mutations were made so that each charged residue was replaced by an alanine residue. No single mutations were sufficient enough to disrupt RAGE homodimerization. The results were normalized to wild-type PR-TM-CYTOfull RAGE AraC fusion protein and error bars represent standard deviations.

6.3   Application of Modified DN-AraTM Assay as an Positive Selection Tool for

Peptide-based Protein Homooligomer Blocker Design

In Chapter 3, we developed an assay, named DN-AraTM, which produces a chimera containing wild-type AraC co-expressed with a chimera that contains a non-functional, mutant AraC (AraC*) that is unable to activate gene transcription. Preferential heterodimerization (AraC-AraC*) reduces the level of GFP reporter gene transcription (under control of the *araBAD* promoter) relative to homodimerization (AraC-AraC).   The   DN-AraTM   enables   simultaneous   measurement   of

homodimerization and heterodimerization of type I receptor domains, including both soluble and transmembrane (TM) domains, in their native orientation. Additionally, there is a very important potential application for this assay: a small peptide-based drug design system.

Numerous disease states are associated with oligmer formation of proteins, such as Alzheimer's disease. Oligomers of amyloid β-peptide (Aβ) are the earliest effectors that induce long-term potentiation (LTP) and lead to Alzheimer's disease (11-13). Small molecules that inhibit Aβ oligomer formation had been demonstrated by recent study to be effective in abrogation of LTP thus this shines light on a new mechanism of therapeutic intervention (14). Despite the fact that small molecule drugs often have the advantage of better oral bioavailability over peptide-based drugs, the peptide-based drugs are superior in terms of high potencies of action and few off-target side-effects (15). Therefore, designing peptide-based drugs that block oligomer formation is a rational approach for next generation drugs.

An ideal tool to assist the peptide-base drug design process needs to allow for the identification of the best target protein binding sequence among thousands of randomized peptide sequences in high efficiency. Therefore, one future direction for DN-AraTM assay is to use it as the foundation for the modified DN-AraTM assay, by replacing the reporter GFP with a lethal toxin gene for *E. coli*, we could build a positive selection system. The modified DN-AraTM assay would be designed as follows: first, the protein target is expressed as the AraC fusion protein and a library of randomized blocker peptide sequences will be expressed as AraC* fusion proteins. When homodimerization is preferred, the transcription of lethal toxin gene would kill the *E.*

146

coli cells with this genotype. Only when heterodimerization is preferred would one observe the survival of cells. By DNA sequencing the surviving cells, candidates that block target protein homooligomerization can be identified. Moreover, potential choices of lethal genes such as RelE (Endoribonuclease RelE) and ChpAK (Endoribonuclease MazF) of *E. coli* toxins have been shown to be effective in cell killing (16). As shown in Fig. 6.4, we replaced the GFP by RelE gene in the DN-AraTM, expressed both types of integrin α_{IIb} AraC chimeras (wild-type and obligated homodimer mutation L980A), and monitored cell growth by optical density at 600nm (OD_{600}). The results showed that, first, RelE gene expression leads to repression of cell grown could be governed by homodimerization of AraC chimeras. Second, we observed a more remarkable repression of cell growth for the L980A integrin α_{IIb} AraC chimera relative to wild-type integrin α_{IIb} AraC chimera, which fits well with the expectation that stronger homodimer of chimera construct results in higher expression of RelE gene, thus, more cell death. Both observations support the potential usage for modified DN-AraTM as a positive selection system for drug screening.
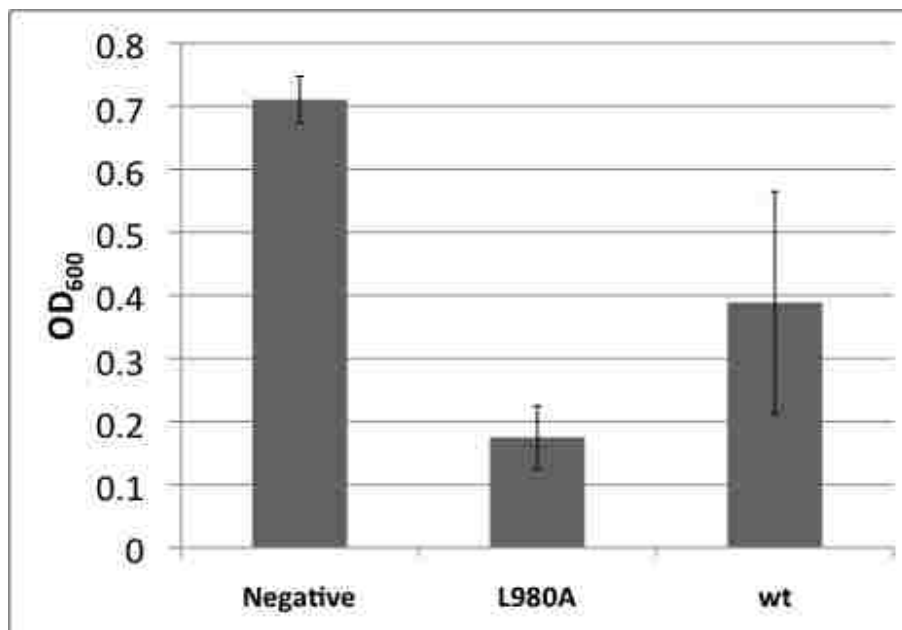
Figure 6. 4. Modified AraTM assay utilizing toxin RelE as the reporter for positive selection. Obligated homodimer-forming integrin α_{IIb} L980A mutant AraC chimera (L980A) had a stronger activation on *araBAD* promoter for RelE expression that leads to more cell growth repression compared to wild-type integrin α_{IIb} AraC chimera (wt) as well as cells not expressing any AraC chimera (Negative). Error bars represent standard errors (Negative: negative control, L980A: integrin α_{IIb} L980A mutant AraC chimera, and wt: wild-type integrin α_{IIb} AraC chimera).

6.4   Application of Modified DN-AraTM Assay that Reports Homodimer and Heterodimer Simultaneously for Drug Design Optimization

An ideal peptide-based drug should have two characteristics: first, it should have a high specificity against the receptor target. Second, it should not form aggregates that interfere with its therapeutic efficiency. Therefore, when designing a drug, it would be beneficial to have a system that could report heterodimer (specificity of the drug to its target) and homodimer (tendency to form aggregates) in an equilibrium system simultaneously. This leads us to another future direction for the DN-AraTM assay: modify the *araBAD* promoter so it requires two types of transcription factors for

148

transcription activation. As shown in Figure 6.5, binding of wild-type AraC to high-affinity O_2 and I_1 sites forms a DNA loop and represses gene transcription of *araBAD* promoter (17). In the presence of arabinose, the wild-type AraC protein homodimerize in another orientation and bind to I_1 and I_2 sites. This opens up the DNA loop thus RNA polymerase can bind and gene transcription is allowed. In our DN-AraTM assay (Chapter 3), we replaced the N-terminal domain of wild-type AraC protein (including arabinose binding pocket and N-terminal dimerization domain) with the protein domain of interest. Thus, the DNA loop formation is now arabinose independent but protein homo/heterdimerizaiton dependent.
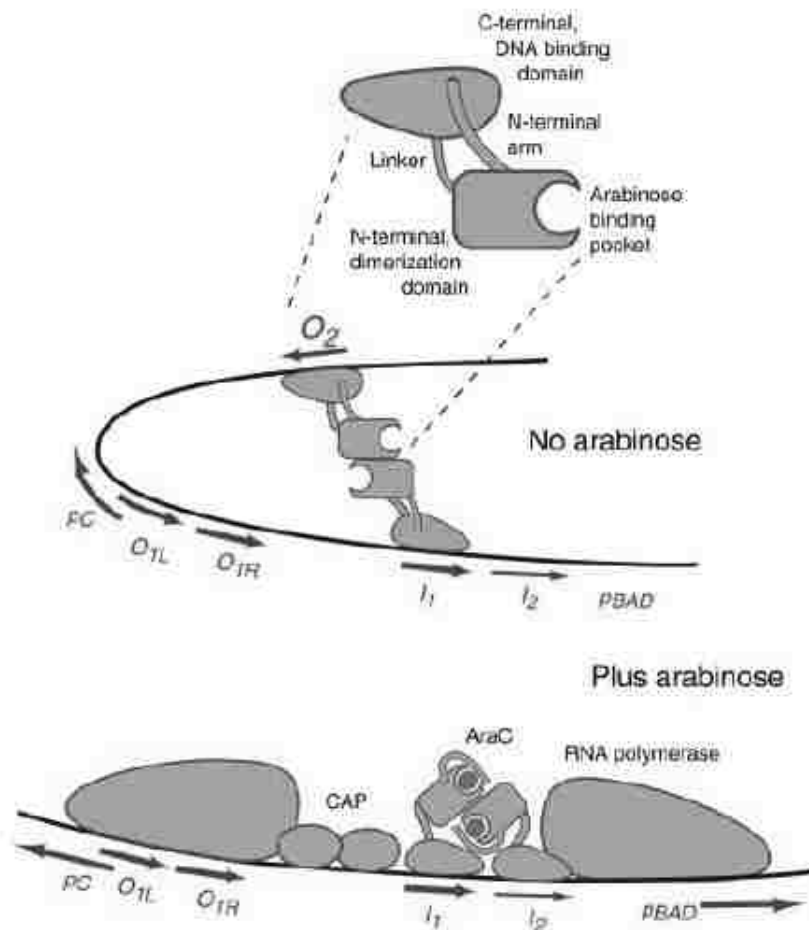
Figure 6. 5. Binding of AraC in *trans* to the O_2 and I_1 half-sites to form a DNA loop in the absence of arabinose and its binding *cis* to the I_1 and I_2 half-sites in the presence of arabinose that leads to unlooping and induction of p_{BAD} and transient depression of p_c and the light-switch mechanism (17).

In the modified DN-AraTM assay, we would like to replace the high-affinity I_1 site with another transcription factor binding site on the pAraGFPCDF plasmid (Chapter 3) and the AraC part of the chimera protein expressed by pAraTMDN (Chapter 3) with another transcription factor protein. We would also replace the reporter GFP gene on the modified pAraGFPCDF with RFP (red fluorescent protein) gene. Then, the modified reporter plasmid would have reporter gene controlled by two

150

transcription factor binding sites. In the experimental setup, we would co-express all four plasmids: pAraTMwt (Chapter 3), modified pAraTMDN, modified pAraGFPCDF, and pAraGFPCDF. As a result, heterodimerization of the two types of chimera proteins expressed by pAraTMwt and modified pAraTMDN could be monitored by RFP signal and homodimerization of chimera proteins expressed by pAraTMwt could be correlated to GFP signal.

Ideal replacement candidates would have a functional truncated protein (including its DNA-binding domain (DBD)) that is similar to truncated AraC (used in DN-AraTM assay) in size with well-studied DNA-binding site consensus sequence. One possible candidate is the *Streptomyces griseus* AdpA, which had been described by previous studies with a functional truncation that is 126aa in size (compared to truncated AraC that is 125aa) and known binding site consensus sequence $5'$-TGGCGGGTTC-$3'$ (18). As an initial attempt, we replaced the I_1 site sequence with the 10bp AdpA consensus sequence on the DN-AraTM assay reporter plasmid (pAraGFPCDF) and the AraC* protein gene on the pAraTMDN plasmid with DNA sequence that translates to truncated AdpA protein (residue 215-340). Then we expressed or co-expressed pAraTMwt, modified pAraTMDN, and modified pAraGFPCDF with each of pAraTMwt and modified pAraTMDN expressing integrin α_{IIb} L980A as the protein domain of interest. The results are shown in Figure 6.6. All three scenarios showed similar fluorescent signals (expressing modified pAraGFPCDF alone, co-expressed pAraTMwt with modified pAraGFPCDF, or co-expressed pAraTMwt, modified pAraTMDN, and modified pAraGFPCDF). Two conclusions could be made from these results. First, by replacing the native I_1 with the AdpA DNA-

151

binding sequence, homodimerization of AraC fusion proteins cannot activate the *AraC/AdpA dual* promoter. Second, heterodimerization of AraC and AdpA fusion proteins can not activate the *AraC/AdpA dual* promoter. The inability to turn on modified *AraC/AdpA dual* promoter may be due to incorrect spacing between the two transcription factor binding sites. Therefore, further optimization of the spacing between these binding sites is required.
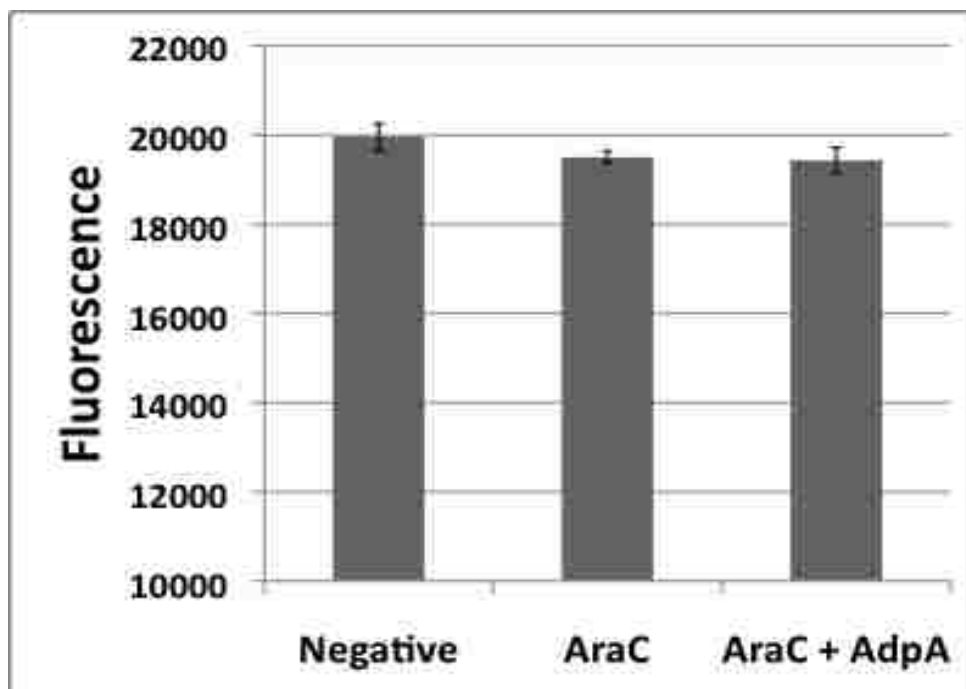
Figure 6. 6. Modified AraTM assay utilizing two transcription factor binding site as the mechanism to control GFP expression. Neither homodimeriztion of AraC fusion proteins nor heterodimerization of AraC/AdpA fusion proteins could activate the *AraC/AdpA dual* promoter on the modified pAraGFPCDF plasmid (Negative: negative control only expressing modified pAraGFPCDF plasmid, AraC: co-expresseing pAraTMwt with modified pAraGFPCDF plasmids, and AraC + AdpA: co-expressing pAraTMwt, modified pAraTMDN, and modified pAraGFPCDF plasmids).

6.5   Determination of High-resolution Human RAMP1 Structure by X-ray Crystallography

In Chapter 4, we designed an *E. coli*-based membrane protein overexpression system, pOmpF expression plasmid, which utilizes an engineered bacterial outer membrane protein F (OmpF) as the fusion to enhance target protein expression. Using this technique, we were able to make milligram quantities of pure full-length human RAMP1 protein. One future direction is to apply this method to purify enough human RAMP1 protein to set up crystallization tests for high-resolution structural studies. Even though structure of RAMP1 extracellular domain had already been solved, there is still no information on how the RAMP1 extracellular ligand-binding pocket is orientated respect to its TM domain (19). Considering the important role RAMP1 TM domain plays in signaling, including its role in forming heterooligomer with CLR, there is still great value in solving the full-length human RAMP1 structure (20). This structural information could not only provide insights in how RAMP1 is inserted into the cell membrane but also help in predicting which side of the RAMP1 TM domain is facing the CLR TM domains.

6.6   Identification of CLR-RAMP1 Association Interfaces through Cysteine Mutagenesis and Cross-linking

In Chapter 5, we identified a specific RAMP1 TM motif that governs the heterodimerization of CLR and RAMP1 and affects their signaling as a CGRP receptor. This finding is especially critical for anti-migraine drug design since it is often very difficult to develop small molecule antagonists against family B GPCR to compete with

their endogenous peptide-based ligand (21). One solution is to design allosteric regulator that does not inhibit receptor signaling by direct antagonist/agonist competition, instead, it inhibits receptor signaling by binding to locations on the receptor that modifies its ligand binding site and prevents ligand binding. However, in order to achieve this goal, we need to complete the other half of the picture, which is, how CLR TM domains interact with RAMP1. Therefore, one future direction that is worth exploring is to figure out what CLR TM domains are associating with RAMP1 and more specifically, which at which interfaces do these protein interact.

The experimental techniques that could be used are site-directed mutagenesis and cysteine cross-linking. This should be done by first performing a series of mutagenesis reaction to generate single cysteine residue substitution on each full-length CLR and RAMP1 protein TM domains. Then, the proteins would be expressed in HEK cells and cross-linked the free sulfhydryl group on the cysteine side-chain between mutant CLR and RAMP1 proteins using a maleimide crosslinker (Bismaleimidohexane) and immunoblotted with specific antibodies. Judging by the molecule size of the immunoblotted bands, we could determine which RAMP1 and CLR mutants were cross-linked thus their relative positions with respect to one another. In Figure 6.7, we successfully cross-linked HA-tagged RAMP1 I118C mutation with one of the CLR TM5 mutation (L290C) in HEK cells. This illustrates the plausibility of the experiments. Furthermore, by making single residue cysteine substitution on the RAMP1 TM and identified CLR TM domains at different faces of the helices and performing cross-linking experiments, CLR-RAMP1 association interfaces could be mapped out.
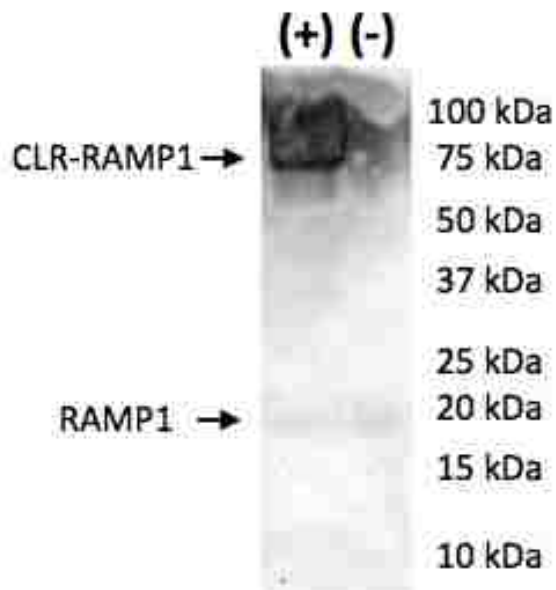
Figure 6. 7. CLR and RAMP1 cysteine cross-linking experiment. CLR and RAMP1 can only be cross-linked when cysteine point mutations when introduced to each receptor. The lower molecule weight band corresponds to HA-tagged RAMP1 (20.9 kDa) and the higher molecule weight band corresponds to HA-tagged RAMP1 cross-linked to CLR (70.9 kDa) ((+): RAMP1-HA I118C and CLR L290C and (-): RAMP1-HA wt and CLR wt).

6.7   Final Remarks

In summary, this work contributes to the field of membrane protein research. More specifically, the tools we developed can benefit the exploration of membrane proteins such as identifying residues that define protein oligomerization as well as high-resolution structural studies. Additionally, these tools can also be utilized as foundations for novel drug development methods. The findings regarding CLR and RAMP1 provide insights about their mode of association and signaling. However, this thesis only brings us a step closer in understanding many membrane protein-linked biological processes.

Thus, many future research directions were suggested. It's my hope that this work provides motivation for individuals who are enthusiastic and passionate to carry on.

6.8 References

(1) Pielak, R. M. & Chou, J. J. (2011). Influenza M2 proton channels. *Biochim Biophys Acta* **1808**, 522-9.

(2) Helenius, A. (1992). Unpacking the incoming influenza virus. *Cell* **69**, 577-8.

(3) Wang, J., Wu, Y., Ma, C., Fiorin, G., Pinto, L. H., Lamb, R. A., Klein, M. L. & Degrado, W. F. (2013). Structure and inhibition of the drug-resistant S31N mutant of the M2 ion channel of influenza A virus. *Proc Natl Acad Sci U S A* **110**, 1315-20.

(4) Boltz, D. A., Aldridge, J. R., Jr., Webster, R. G. & Govorkova, E. A. (2010). Drugs in development for influenza. *Drugs* **70**, 1349-62.

(5) Holsinger, L. J., Nichani, D., Pinto, L. H. & Lamb, R. A. (1994). Influenza A virus M2 ion channel protein: a structure-function analysis. *J Virol* **68**, 1551-63.

(6) Mahon, M. J. (2011). pHluorin2: an enhanced, ratiometric, pH-sensitive green florescent protein. *Adv Biosci Biotechnol* **2**, 132-137.

(7) Stouffer, A. L., Ma, C., Cristian, L., Ohigashi, Y., Lamb, R. A., Lear, J. D., Pinto, L. H. & DeGrado, W. F. (2008). The interplay of functional tuning, drug resistance, and thermodynamic stability in the evolution of the M2 proton channel from the influenza A virus. *Structure* **16**, 1067-76.

(8) Fritz, G. (2011). RAGE: a single receptor fits multiple ligands. *Trends Biochem Sci* **36**, 625-32.

(9) Zong, H., Madden, A., Ward, M., Mooney, M., Elliott, C. & Stitt, A. (2010).

Homodimerization is essential for the receptor for advanced glycation end products (RAGE)-mediated signal transduction. *J Biol Chem.* **285**, 23137-46.

(10) Chuah, Y. K., Basir, R., Talib, H., Tie, T. H. & Nordin, N. (2013). Receptor for Advanced Glycation End Products and Its Involvement in Inflammatory Diseases. *Int J Inflam* **2013**, 403460.

(11) Roher, A. E., Chaney, M. O., Kuo, Y. M., Webster, S. D., Stine, W. B., Haverkamp, L. J., Woods, A. S., Cotter, R. J., Tuohy, J. M., Krafft, G. A., Bonnell, B. S. & Emmerling, M. R. (1996). Morphology and toxicity of Abeta-(1-42) dimer derived from neuritic and vascular amyloid deposits of Alzheimer's disease. *J Biol Chem* **271**, 20631-5.

(12) Lambert MP, B. A., Chromy BA, Edwards C, Freed R, Liosatos M, Morgan TE, Rozovsky I, Trommer B, Viola KL, Wals P, Zhang C, Finch CE, Krafft GA, Klein WL. (1998). Diffusible, nonfribrillar ligands derived from Aβ1-42 are potent central nervous system neurotoxins. *Proc Natl Acad Sci U S A* **95**, 6448-53.

(13) Hartley, D. M., Walsh, D. M., Ye, C. P., Diehl, T., Vasquez, S., Vassilev, P. M., Teplow, D. B. & Selkoe, D. J. (1999). Protofibrillar intermediates of amyloid beta-protein induce acute electrophysiological changes and progressive neurotoxicity in cortical neurons. *J Neurosci* **19**, 8876-84.

(14) Walsh, D. M., Townsend, M., Podlisny, M. B., Shankar, G. M., Fadeeva, J. V., El Agnaf, O., Hartley, D. M. & Selkoe, D. J. (2005). Certain inhibitors of synthetic amyloid beta-peptide (Abeta) fibrillogenesis block oligomerization of natural Abeta and thereby rescue long-term potentiation. *J Neurosci* **25**, 2455-62.

(15) Craik, D. J., Fairlie, D. P., Liras, S. & Price, D. (2013). The future of peptide-based

drugs. *Chem Biol Drug Des* **81**, 136-47.

(16) Pedersen, K., Christensen, S. K. & Gerdes, K. (2002). Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Mol Microbiol* **45**, 501-10.

(17) Schleif, R. (2003). AraC protein: a love-hate relationship. *Bioessays* **25**, 274-82.

(18) Yao, M. D., Ohtsuka, J., Nagata, K., Miyazono, K., Zhi, Y., Ohnishi, Y. & Tanokura, M. (2013). Complex structure of the DNA-binding domain of AdpA, the global transcription factor in Streptomyces griseus, and a target duplex DNA reveals the structural basis of its tolerant DNA sequence specificity. *J Biol Chem* **288**, 31019-29.

(19) ter Haar, E., Koth, C. M., Abdul-Manan, N., Swenson, L., Coll, J. T., Lippke, J. A., Lepre, C. A., Garcia-Guzman, M. & Moore, J. M. (2010). Crystal structure of the ectodomain complex of the CGRP receptor, a class-B GPCR, reveals the site of drug antagonism. *Structure* **18**, 1083-93.

(20) Steiner, S., Muff, R., Gujer, R., Fischer, J. A. & Born, W. (2002). The transmembrane domain of receptor-activity-modifying protein 1 is essential for the functional expression of a calcitonin gene-related peptide receptor. *Biochemistry* **41**, 11398-404.

(21) Moore, E. L. & Salvatore, C. A. (2012). Targeting a family B GPCR/RAMP receptor complex: CGRP receptor antagonists and migraine. *Br J Pharmacol* **166**, 66-78.

VITA

Name: **Pin-Chuan (Daniel) Su**
Date of Birth: **November 15th, 1986**
Names of Parents: **Yin-Feng Lai & Chung-Yuan Su**

Education

Lehigh University, Bethlehem, PA
> Doctor of Philosophy, May 2014
> Chemical Engineering
> GPA 3.57/4.0

Carnegie Mellon University, Pittsburgh, PA
> Bachelor of Science, May 2009
> Double Majors in Chemical Engineering and Biomedical
> Engineering
> GPA 3.45/4.0

Publications

Su P.C., Berger B.W. Identifying Key Juxtamembrane Interactions in Cell Membranes Using AraC-based Transcriptional Reporter Assay (AraTM). *J Biol. Chem.* (2012)

Su P.C., Si W., Baker D. L., Berger, B. W. High-Yield Membrane Protein Expression from *E. coli* Using an Engineered Outer Membrane Protein F (OmpF) Fusion, *Protein Science* (2013)

Su P.C., Berger B.W. A Novel Assay for Assessing Juxtamembrane and Transmembrane Domain Interactions Important for Receptor Heterodimerization, *J Mol. Biol.* (2013)

Work Experience

National Taiwan University Hospital	Summer 2006
Doctor's research assistant	
Pro Broadband Inc. at Beijing	Summer 2008 & 2009
International sales assistant	