Characterising the molecular interaction between CLEC14A and Multimerin 2

By

Aleen Baber

UNIVERSITY OF BIRMINGHAM

A thesis submitted to the University of Birmingham for the degree

DOCTOR OF PHILOSOPHY

Institute of Cardiovascular Sciences

College of Medical and Dental Sciences

The University of Birmingham

July 2024

Abstract

The C-type lectin domain group 14A (CLEC14A) protein, a single pass transmembrane spanning glycoprotein, is an endothelial expressed protein found to be highly and selectively expressed on tumour vessels. CLEC14A interacts with various ligands including the extracellular matrix protein Multimerin 2 (MMRN2) and heparan sulfate polysaccharides. CLEC14A's interaction with MMRN2 elicits a proangiogenic phenotype and upon blocking this interaction there is a reduction in tube formation and cell migration *in vitro* and tumour size *in vivo*. The aim of this project was to further characterise the interactions of CLEC14A with MMRN2 and heparin as well as explore the structure of CLEC14A and MMRN2. To this end an alanine-scanning mutagenesis approach, coupled with AlphaFold predictive modelling, determined that residues S137, T139 and R141 of human CLEC14A directly interact with MMRN2. The R161 residue of CLEC14A was also shown to partially contribute to the interaction with heparin. It was revealed that the R100 and R141 residues of CLEC14A form part of the binding epitope for the CRT4 antibody which blocks the interaction between CLEC14A and MMRN2. The CRT4 antibody, known to modulate tumour vessel formation, also blocked the interaction between CLEC14A and heparin, thus expanding our understanding of how this reagent is acting to mediate its biological effects. Additionally, analytical ultracentrifugation revealed that MMRN2 forms a trimer in solution, providing new insights into how CLEC14A may bind to MMRN2. Taken together, the data presented in this thesis expands on the current understanding of the interaction between CLEC14A and MMRN2 and heparan sulfate, providing valuable strategies that could be used to better determine how different ligands affect the biology of CLEC14A. This may ultimately give rise to novel therapeutic strategies to target CLEC14A in cancer.

In the name of God, the most gracious and most merciful

ربِّ زِدْنِي عِلْماً

Acknowledgments

I would like to begin by thanking my primary supervisor, Dr. Victoria Heath, for all of her support and supervision throughout my PhD. Her scientific integrity and subject knowledge have been an inspiration. I will look back fondly on our many conversations about the chaos and immense joys of family life. I would also like to extend my thanks to my secondary supervisor, Professor Roy Bicknell, for his support and constant willingness to impart his knowledge to myself and others.

Second only to my supervisors, I would like to thank Alex Slater, an excellent mentor and friend throughout my PhD, a phenomenal scientist and teacher, who always delivered advice, knowledge and sarcastic comments with perfect comedic timing. I would also like to extend my thanks to Abs and Kabir Khan, who have both taken time out of their lives to provide advice, guidance, and expertise. My gratitude extends to Shannon O'Brian who was very kind and patient when helping me with the NanoBRET experiments.

I would like to acknowledge all the friends I have made in ICVS throughout my PhD, an incredible set of people with whom I was lucky enough to cross paths with in my lifetime. Dylan Marques, who dominated 'joke of the day', I appreciate all the laughs. Leena Patel who I have had the pleasure of accompanying during my PhD, your hard work and humour has been an inspiration. Benjamin Keenleyside, who in a very short time become a good friend and has provided the right amount of chaos while maintaining a healthy HR-minded workplace. Stanimir Tashev, who brought me a lot of comfort and joy. Thank you for being a great friend … and sometimes a pain in the neck. I would also like to acknowledge the junior PhD students Liv, Abbie, Jess, Katie and Arnon for their exceptional sense of humour and optimism. I know each of you will grow to be amazing scientists. I also want to acknowledge

all the friends who have since left the university Amar Azad, Caitlin Hall and Max Cumberland, thank you for all the laughs and great memories. Finally, I would like to extend my thanks to Ashleigh Renelt for being so helpful during my PhD particularly during experiment emergencies but also for being a great friend. I will look back fondly on our conversations which brought me tears of laughter and the artistic posters you would make of me.

Next, I want to acknowledge Simren, one of the most important people to me—my sister, my comfort, my confidant, and my personal comedian. Growing with you and going through degree after degree with you by my side has been an immense pleasure. My thanks extend to Chaggar, Niki, and Ciara. I couldn't have asked for a more amazing, inspiring group of girls to call my best friends. I am so grateful to have you in my life, thank you for everything.

Finally, I would like to thank my family for their company and love. To my mum, for all her love and support—I am grateful for all the "Letts" books and time you spent on me. Thank you for believing in me when nobody else did and for wiping my tears away when I felt alone. To my sisters Iqra and Hira- thank you for being the kindest, most patient and loving people in my life. I am indebted you each of you for shaping me to be the woman I am (for better or worse).

Table of Contents

List of figures

List of tables

CHAPTER 1: Introduction

CHAPTER 1: Introduction

1.1. The Endothelium

The Endothelium is a monolayer of endothelial cells which line blood vessels and exert a range of homeostatic functions including the regulation of vascular tone, angiogenesis, inflammation, leukocyte trafficking, platelet activation and vascular permeability (Rajendran *et al.,* 2013). Originally delineated from the mesoderm during the early stages of embryo development, endothelial cells are one of the first cell populations to be formed as part of the cardiovascular system (Liu and Gotlieb, 2010).

The endothelium forms the interface between blood vessels and organs and functions to regulate the movement of substances between the blood and tissue, a key process during inflammation. During inflammation immune cells within the circulation move to the site of tissue damage in a process called "leukocyte trafficking". Upon inflammation pro-inflammatory cytokines and chemokines such as TNF-alpha, IL-1 and IL-8 are released by immune cells and activate endothelial cells to express adhesion molecule such as ICAM-1 and VCAM-1. Such adhesion molecules promote the contact of immune cells from the circulation onto the luminal side of the blood vessel by binding to integrins on the surface of circulating immune cells. Upon inflammatory activation endothelial cells also modify their tight junctions to become more permeable creating space for immune cells to extravasate into the tissue (Panes and Granger).

The structure of the endothelium is specialised and differs based on location and function. High endothelial venules found in lymphoid tissues have cuboidal endothelial cells which facilitate movement of lymphocytes from the blood stream to the lymphatic system (Vella *et al.,* 2021). In areas where rapid substance exchange is required such as the kidney, intestines

and endocrine glands fenestrated endothelial cells are present (Finch *et al.,* 2023). In the case of the blood brain barrier endothelial cells are tightly connected by tight junctions to form a continuous high-resistance barrier to aid highly selective permeability (Kadry *et al.,* 2020).

Changes in vascular tone are driven by a whole host of factors such as release of hormones from the nervous system, physical factors such as temperature, myogenic regulators and mechanical forces such as shear stress (Kvietys, 2010). Shear stress is sensed by mechanosensors associated with the membrane of endothelial cells and triggers endothelial-derived factors to regulate vascular tone such as nitric oxide. The endothelium also plays roles in coagulation, thrombosis and leukocyte trafficking which is essential for protecting tissues against infections and injuries (Granger and Senchenkova, 2010). This cell type also orchestrates the generation of new vessels from existing vessels in the process of angiogenesis.

1.2. Angiogenesis

Angiogenesis is the formation of blood vessels from pre-existing ones and plays a critical role in embryonic development and physiological processes including the regeneration of the endometrium in the menstrual cycle and wound healing. There are two modes of angiogenesis: intussusception (also known as splitting angiogenesis) and sprouting angiogenesis. Both are regulated by a careful balance of angiogenic and angiostatic factors (Carmeliet, 2003).

Sprouting angiogenesis is initiated when parenchymal cells detect a hypoxic microenvironment. Hypoxia-inducible Factor 1α (HIF-1α) is a key transcription factor subunit which, in physiological conditions, is produced and targeted for degradation by the ubiquitin proteasome pathway. Under normoxic conditions prolyl hydroxylases hydroxylate HIF-1α in

an oxygen-dependent manner. This post-translational modification is recognised by the Von Hippel-Lindau protein, the recognition subunit of the E3 ubiquitin ligase, thus enabling the polyubiquitination and subsequent degradation of HIF-1α. In hypoxic conditions the low levels of oxygen are insufficient to drive hydroxylation of HIF-1α, subsequently HIF-1α forms a complex with HIF-1β. This complex translocates into the nucleus and initiates the transcription of up to 2000 genes including *VEGFA* and *ANG1* (Strowitzki *et al.,* 2019).

Vascular endothelial growth factor A (VEGF-A) and Angiopoietin 1 (Ang1) are key angiogenic factors which are upregulated in the context of angiogenesis and bind to their cognate receptors VEGFR2 and Tie2 respectively on endothelial cells. Activation of VEGR2 by VEGFA drives sprouting angiogenesis in the early stages by playing a critical role in inducing quiescent ECs to become tip cells as well as causing endothelial release of enzymes such as matrix metalloproteinases (MMPs) which degrade the basement membrane, facilitating the formation of a sprout. Pericytes detach and ECs expressing the most activated VEGFR2 compete to become tip cells. Tip cells induce filopodia which are heavily saturated with VEGFRs to sense the gradient of angiogenic growth factor and guide migration. Tip cells signal to adjacent endothelial cells promoting their development of a stalk cell phenotype; stalk cells proliferate causing the capillary sprout to elongate and vacuoles within stalk cells form to create a new lumen. Once two or more sprouts reach the site of growth factor secretion, the tip cells fuse together to make a closed circuit for oxygenated blood flow. The vessel is then stabilised by recruitment of pericytes, a process driven by Ang1-Tie2 signalling. This involves deposition of a novel basement membrane and the recruitment of pericytes (Figure 1.1) (Udan *et al.,* 2014; Suri *et al.,* 1996). The resulting vascular plexus produced is a very well-defined and efficient network of vessels secured with an intact basement membrane and full pericyte coverage.

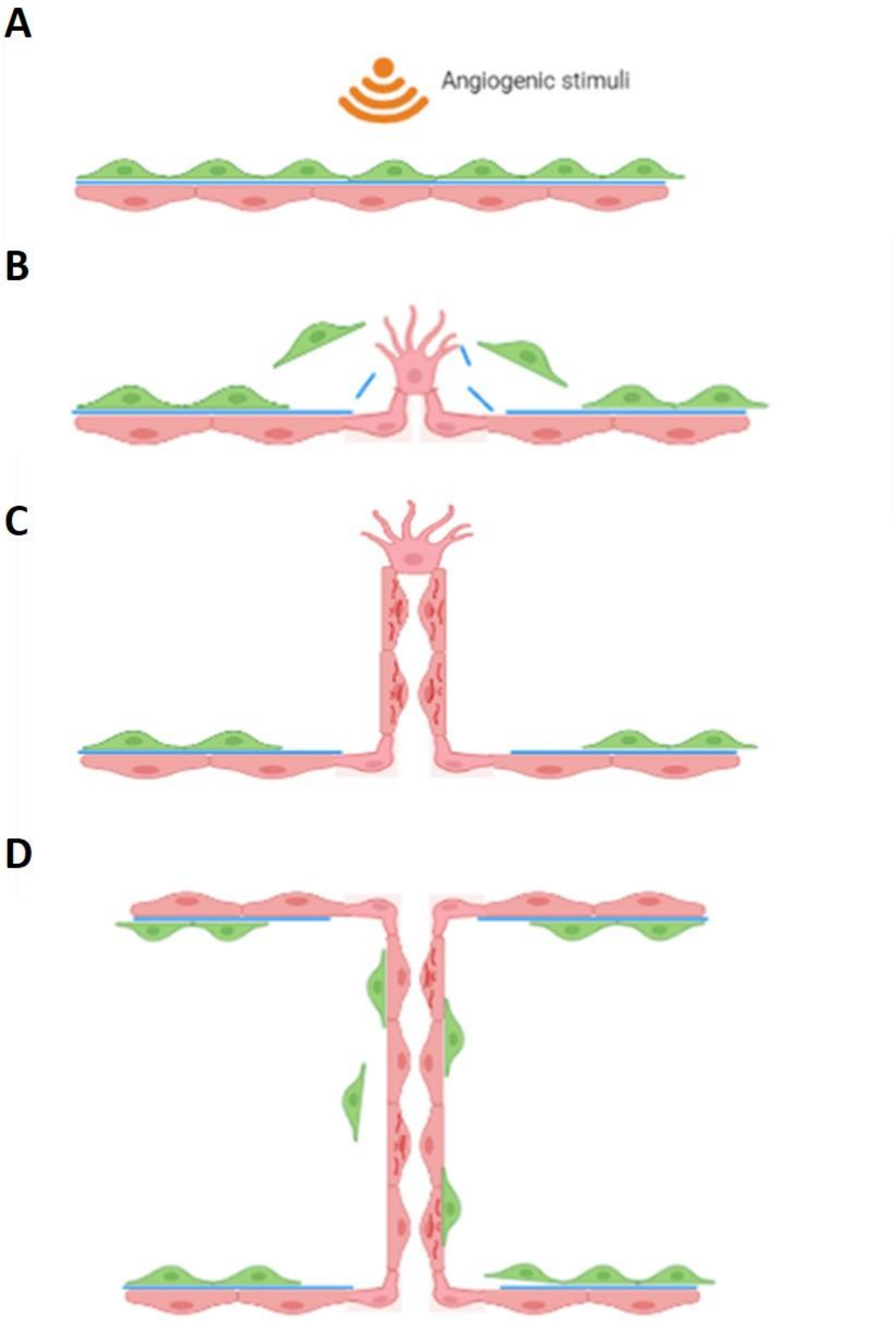


Figure 1.1. Schematic diagram of sprouting angiogenesis . (A) Endothelial cells (pink) become activated following angiogenic stimuli (Orange). (B) Pericytes cells (green) on the vessel detach and the basement membrane (blue) breaks down exposing the endothelial cell. (C) Endothelial cells facing the stimuli become tips cells expressing filopodia projections or stalk cells and reach out of the vessel. (D) The angiogenic sprout reaches another sprout and fuses to form a new vessel. The basement membrane reforms, and pericytes cells are recruited. (Adapted from Bergers and Benjamin, 2003) (Created in BioRender).

1.2.1. Tumour endothelium and angiogenesis

Angiogenesis is a hallmark of cancer and is required for tumour growth and stability (Hanahan and Weinberg, 2011). Whilst tumour angiogenesis is mechanistically similar to physiological angiogenesis, the resulting vascular network is quite different due to excessive expression of angiogenic factors such as VEGFA (Duffy *et al.,* 2013). The tumour vasculature is a disorganised network of leaky vessels due to reduced intercellular endothelial adhesion and incomplete pericyte coverage. Tumour vessels vary in diameter and length and do not follow the normal hierarchy of vessel organisation. This disorganised network is often poorly perfused resulting in a hypoxic and acidic tumour microenvironment, this persistent hypoxia results in chronic production of angiogenic factors (Lugano *et al.,* 2020). Furthermore, incomplete pericyte coverage promotes metastasis as tumour cells gain easy access to the blood stream (Hida *et al.,* 2022).

Given the dysfunctional nature of tumour blood vessels and the distinct tumour microenvironment, it is unsurprising that the tumour endothelium expresses a particular angiogenic gene signature. Tumour endothelial markers (TEMs) are defined as proteins expressed on tumour endothelium but absent from the normal vasculature. Such markers are of particular interest because they offer an opportunity to specifically target the tumour endothelium with anti-cancer agents such as antibody-drug conjugates or chimeric antigen receptor (CAR) T cells (Liang *et al.,* 2021). Pioneering studies by Croix *et al.* identified 46 genes elevated in the tumour associated endothelium in malignant colorectal cancer samples, some of which were expressed in other tumour types (Croix *et al.,* 2000). More recently, Wragg *et al*. used microarray analyses to discover the upregulation of melanoma cell adhesion molecule (MCAM) and its partner protein, laminin alpha 4 (LAMA4) in a number of tumour-associated endothelial tissues including renal cell carcinoma (RCC), colorectal carcinoma and

colorectal liver metastasis. Moreover, it was shown this upregulation predicted poor survival in cases such as RCC (Wragg *et al.,* 2016). Another key tumour endothelial marker is CLEC14A and this is discussed in detail in the next sections of this introduction (Robinson *et al.,* 2020).

The concept of targeting tumour angiogenesis was first proposed by Judah Folkman in 1971 (Folkman, 1971). Since the early 2000s there has been a whole host of antiangiogenic therapies targeting angiogenic factors. The first FDA approved drug used to target VEGF signalling was bevacizumab (Avastin) in 2004. Bevacizumab is a humanized monoclonal antibody which targets VEGF-A and was approved for treatment of metastatic colorectal cancer (Willett *et al.,* 2004). However, these antibody-based therapies came with a range of side effects such as bleeding, haemorrhaging and hypertension (Elice and Rodeghiero, 2012). Another major drawback is drug resistance, despite initial success tumours have been shown to initiate alternative angiogenic pathways and increased hypoxia following treatment has been shown to lead to a more aggressive and invasive phenotype for some tumours (Martin *et al.,* 2013). To address this, antiangiogenic agents were used in combination with other antiangiogenic agents or chemotherapeutic agents or immunotherapy, which in the case of Bevacizumab is routinely administered alongside Paclitaxel (An *et al.,* 2021; Fountzilas *et al.,* 2011). An important aspect of antiangiogenic therapy is vascular normalisation, a process by which limiting excessive availability of VEGFA causes the vasculature to become more normal and hence better able to perfuse the tumour. This allows better delivery of chemotherapeutic drugs into the tumour and is why bevacizumab is most often combined with chemotherapy, and is rarely effective as a monotherapy (Jain, 2001, Yang *et al.,* 2021).

1.3. The C-type lectin domain group 14 family

The C-type lectin domain (CTLD) superfamily, comprising 17 subfamilies, is a large group of extracellular proteins with diverse functions which all feature a CTLD domain. The CTLD is usually characterised by a unique “loop-in-loop” structure consisting of a short loop and a longer flexible loop fixed in place by a disulphide bridge, which mediates interactions with proteins and carbohydrates alike (Zelensky and Gready 2005). CLEC14A, CD93, Thrombomodulin (THBD) and CD248 are all members of the CTLD group 14 subfamily; they are single pass transmembrane-spanning glycoproteins which share structural domains. These include an N-terminal CTLD followed by a sushi domain, one or more EGF-like domains, a mucin-like region, a transmembrane spanning domain and a short C-terminal intracellular tail (Khan *et al.,* 2019) (Figure 1.2). In the case of THBD, the structure C-terminal to the CTLD does not qualify as a sushi domain as it lacks the critical two disulphide bonds which characterise this type of domain (Norman *et al.,* 1991). The mucin-like regions of each protein vary in length and are subjected to O-linked glycosylation to varying degrees with the predicted site as follows; CLEC14A (17), CD93 (23), THBD (7) and CD248 (51) (NetOGlyc 4.0 server) (Steentoft *et al.*,2013). CD248 is expressed in the surface of pericytes and is found to be upregulated in tumour-associated pericytes and stromal cells as well as malignant cells of mesenchymal origin (Teicher, 2019). THBD is expressed on vascular cells such as endothelial cells, mesothelial cells and epidermal keratinocytes (Lager *et al.,* 1995). CD93 is expressed on a whole host of cells including immune cells such as dendritic cells, neutrophils and macrophages as well as endothelial cells and is upregulated on tumour endothelial cells (Tossetta *et al.,* 2023). CLEC14A expression is restricted to tumour endothelial cells as well as endothelial cells experiencing low shear stress (Mura *et al.,* 2012; Robinson *et al.,* 2020).

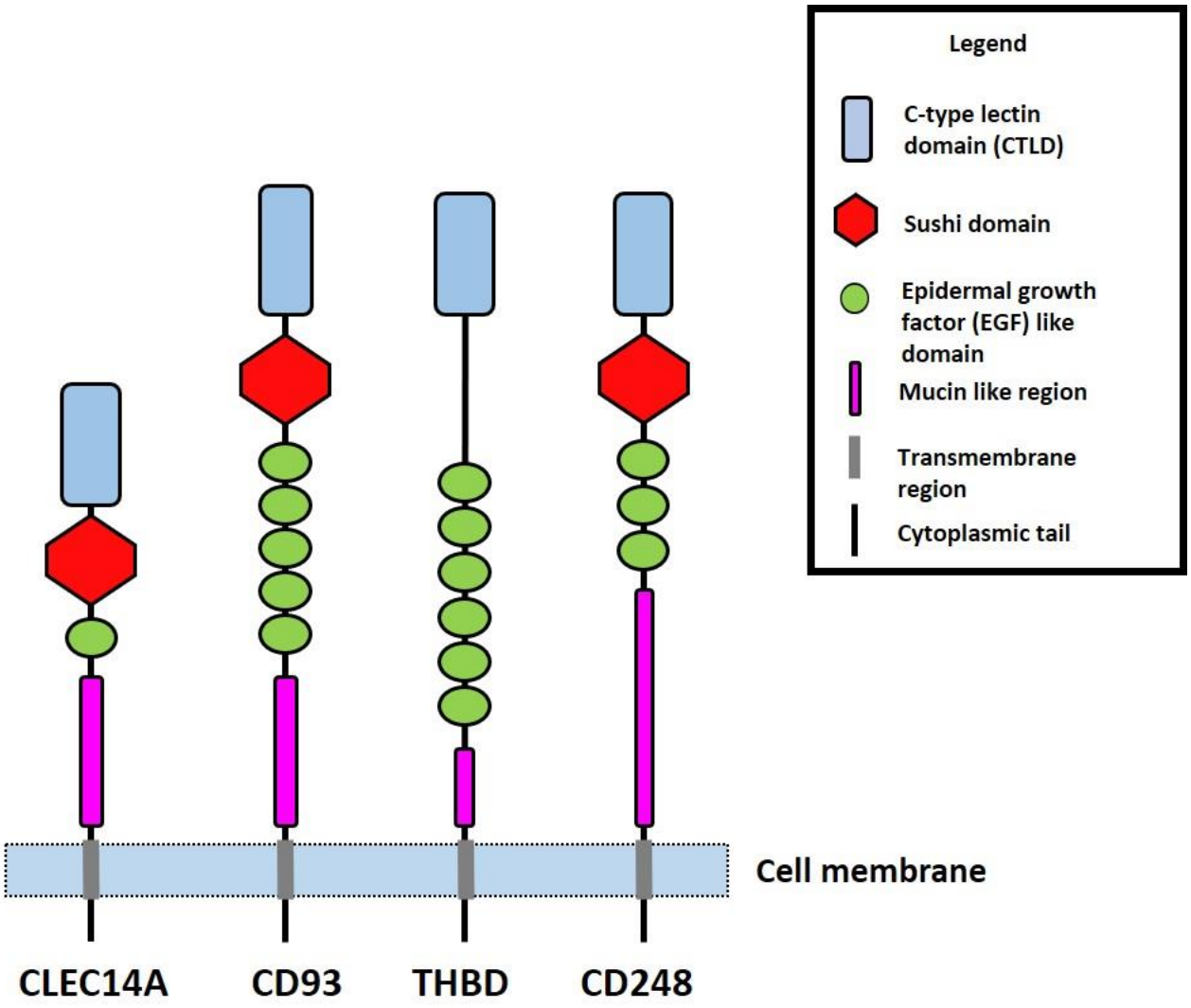


Figure 1.2. Schematic illustration of members of the C-type lectin domain (CTLD) group 14 family alongside a legend indicating specific domains. Each member of the CTLD group 14 family are expressed on the cell membrane. They each contain an N-terminal CTLD. In the case of CLEC14A, CD93 and CD248 the next domain is a sushi domain which is not present for Thrombomodulin. This is then followed by one or more EGF-like domains and then a mucin domain which is often glycosylated. The transmembrane portion of the protein follows and is completed with a short intracellular tail.

1.3.1. CLEC14A expression

CLEC14A was first identified from database-mining and microarray analyses which aimed to identify novel differentially expressed genes in a range of different types of endothelial cells (Ho *et al.,* 2003). CLEC14A has since been identified as a tumour endothelial marker, analysis of both RNA levels and immunohistochemical staining of tissue sections revealed increased expression of CLEC14A on vessels in kidney, bladder, oesophagus, colon, prostate, rectum, stomach, endometrium, liver, breast, lung and skin tumours compared to the corresponding healthy tissues (Robinson *et al.,* 2020). CLEC14A mRNA levels were also increased in hepatocellular carcinoma (HCC) compared to adjacent non-tumour tissue and this expression correlated with increased tumour size (Yan *et al.,* 2022). CLEC14A expression is also found on Type H vessels, a specialised subtype of blood vessels involved in angiogenesis in the context of bone formation (Neag *et al.,* 2024).

The expression of *CLEC14A* is increased in areas of low shear stress (Rho *et al.,* 2012; Robinson *et al.,* 2020). In terms of CLEC14A localisation at a cellular level, immunofluorescent staining revealed expression of CLEC14A at the cell membrane, particularly at the leading edge of migrating HUVECs as well as areas of cell-cell contacts and this localisation was disrupted upon removal of the N-terminal CTLD (Rho *et al.,* 2011; Mura *et al.,* 2012). Beyond pathological conditions, CLEC14A is expressed in both mouse embryos at day E10.5 and vessels in zebrafish embryos between 12-24 hours post fertilisation (Rho *et al.,* 2011; Mura *et al.,* 2012).

1.3.2. CLEC14A function

At the organismal level the consequence of the loss or reduction of CLEC14A has been studied in mice and zebrafish. In the latter, morpholino knockdown of CLEC14A resulted in defective development of intersomitic and dorsal longitudinal anastomotic vessels (Mura *et al.,* 2012). Though the mice were viable, the hind brain of E13.5 *clec14a*$^{-/-}$ mice showed increased microvessel density, haemorrhages and dilated vessels compared to wild type (WT) littermates (Lee *et al.,* 2017). Similarly, another study corroborated that *clec14a*$^{-/-}$ mice were viable but displayed defected tumour growth, suggesting that CLEC14A helped promote tumour growth however decreased microvessel density was observed for *clec14a*$^{-/-}$. Defective ex-vivo angiogenic sprouting was observed in isolated *clec14a*$^{-/-}$ aortas, with both fewer and shorter sprouts compared to WT littermates (Noy *et al.,* 2015).

Given the endothelial-specific expression of CLEC14A, the functional significance of the protein has been broadly studied in a range of angiogenesis assays. Following siRNA mediated knockdown of CLEC14A, tube formation and cell migration were decreased in Matrigel and scratch wound assays, respectively. Similar results were observed in these assays when HUVECs were treated with polyclonal antiserum targeting CLEC14A (Mura *et al.,* 2012). Ectopic expression of CLEC14A in HeLa and HEK293T cells induced filopodia, a key characteristic of tip cells in sprouting angiogenesis (Mura *et al.,* 2012).

Whilst CLEC14A is expressed on the membrane of endothelial cells, its extracellular domain (ECD) is also subject to cleavage by rhomboid-like protein 2 (RHBDL2). RHBDL2 is a protease which cleaves the membrane-proximal mucin-like domain of CLEC14A resulting in the extracellular portion of the protein being deposited into the circulation. HUVEC spheroids were used as a model for sprouting angiogenesis and were treated with a recombinant

CLEC14A ECD in order to mimic the RHBDL2-processed CLEC14A. The results indicated that spheroids treated with recombinant CLEC14A ECD developed fewer and shorter sprouts compared to control treated spheroids (Noy *et al.,* 2016). These data suggest that cleaved CLEC14A could antagonise the function of membrane-bound CLEC14A.

CLEC14A function was also studied in the context of ischemic stroke and brain injury. Silencing of CLEC14A in HUVECs resulted in increased vascular permeability as well as decreased expression of tight junction proteins. This was studied further in *clec14*$^{-/-}$ mice in the context of the blood brain barrier (BBB). *clec14a*$^{-/-}$ mice did not exhibit significant differences in BBB permeability compared to WT mice. However, using the middle cerebral artery occlusion (MCAO) mouse model used to mimic ischemic stroke it was shown that *clec14a*$^{-/-}$ mice exhibited increased BBB permeability compared to WT MCAO mice. Taken together, this suggests a significant role for CLEC14A on the BBB in the context of ischemic stroke and brain injury (Kim *et al.,* 2020).

More recently a study conducted by Neag *et al* highlighted a role for CLEC14A in the context of bone formation. Premature condensation of the type-H vasculature was observed in *clec14a*$^{-/-}$ mice knockout mice which resulted in accelerated skeletal development, increased bone length and expanded distribution of osteoblasts. This phenotype was consistence when mice were treated with the CRT4 blocking antibody (Neag *et al.,* 2024).

1.3.3. CART cells against CLEC14A

Five monoclonal antibodies were generated against the extracellular domains of CLEC14A which cross react with human and mouse CLEC14A; CRT1-5 and chimeric antigen receptor (CAR) T-cells were generated based on CRT3 and CRT5. CAR T-cells are cytotoxic T-cells engineered to express the antigen-specific portion of antibodies fused to a cassette

comprising the T cell receptor as well as co-stimulatory domains (Lin *et al.,* 2021). When CAR-T cells bind their cognate ligand a T cell response is elicited. CAR T-cells were engineered to express a CAR comprised of the antigen binding elements of CRT3 or CRT5. Both sets of CAR T-cells produced IFN-γ, TNF-α and IL-2 following binding with CLEC14A on HUVECs. CRT5 engineered CAR T-cells were tested in three different tumour mouse models; Rip-Tag2 mice which develop pancreatic β-cell tumours, FVB/n mice which were injected with murine pancreatic ductal adenocarcinoma cells in the pancreas and C57BL/6J mice subcutaneously implanted with Lewis lung carcinoma cells. In each model a significant reduction in tumour size was observed following treatment with CAR engineered T-cells versus non-engineered T-cells (Zhuang *et al.,* 2020). This data suggests targeting CLEC14A may be a potential avenue for anti-cancer therapy.

1.3.4. CLEC14A binding partners

As described in section 1.3 CLEC14A contains an N-terminal CTLD domain, this domain mediates interaction with a number of partners: Heat shock protein 70-1A (HSP70-1A), heparin and Multimerin 2 (MMRN2). CLEC14A's cytoplasmic tail also interacts with VEGFR3 (Figure 1.3).

1.3.4.1. CLEC14A-HSP70-1A

The interaction between HSP70-1A and CLEC14A was first identified following a pull-down using CLEC14A-CTLD fused with an Fc tag coupled to mass spectrometry. CLEC14A-CTLD-Fc bound to HSP70-1A at the surface of HUVECs with a binding affinity of ~8nM via amino acids 43-69 of the CTLD (Jang *et al.,* 2017). HSP70-1A is a relatively ubiquitously expressed protein which functions in response to stress as a molecular chaperone to assist proper folding of proteins (Daugaard *et al.,* 2007). SiRNA-mediated knockdown of HSP70-1A disrupted both

tube formation and cell migration of HUVECs, suggesting that potentially its interaction with CLEC14A promotes an angiogenic phenotype. In addition, cell-cell contact in HUVECs increased following treatment with HSP70-1A in a dose-dependent manner. This was abrogated by the addition of recombinant CLEC14A CTLD, which acted to prevent the recombinant HSP70-1A from interacting with endogenous membrane bound CLEC14A. How HSP70-1A modulates endothelial behaviour is not known, however its binding to CLEC14A promoted ERK phosphorylation. The signalling pathway connecting CLEC14A with the MAP kinase pathway has not yet been characterised (Kim *et al.,* 2016; Jang *et al.,* 2017).

1.3.4.2. CLEC14A-Heparin

More recently, the interaction between CLEC14A and heparan sulphate was identified using liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) in a study aimed to discover novel interactions between this glycoform and membrane proteins cleaved from the surface of HUVECs. The binding affinity between heparin and CLEC14A was ~25nM, and subsequent mutagenesis of CLEC14A demonstrated that R161 was a key residue involved in this interaction. The CLEC14A R161A was eluted from the heparin resin at a much lower salt concentration compared to WT CLEC14A and the other mutants, indicating that it bound to heparin more weakly (Sandoval *et al.,* 2020). The functional significance of the interaction has not yet been explored.

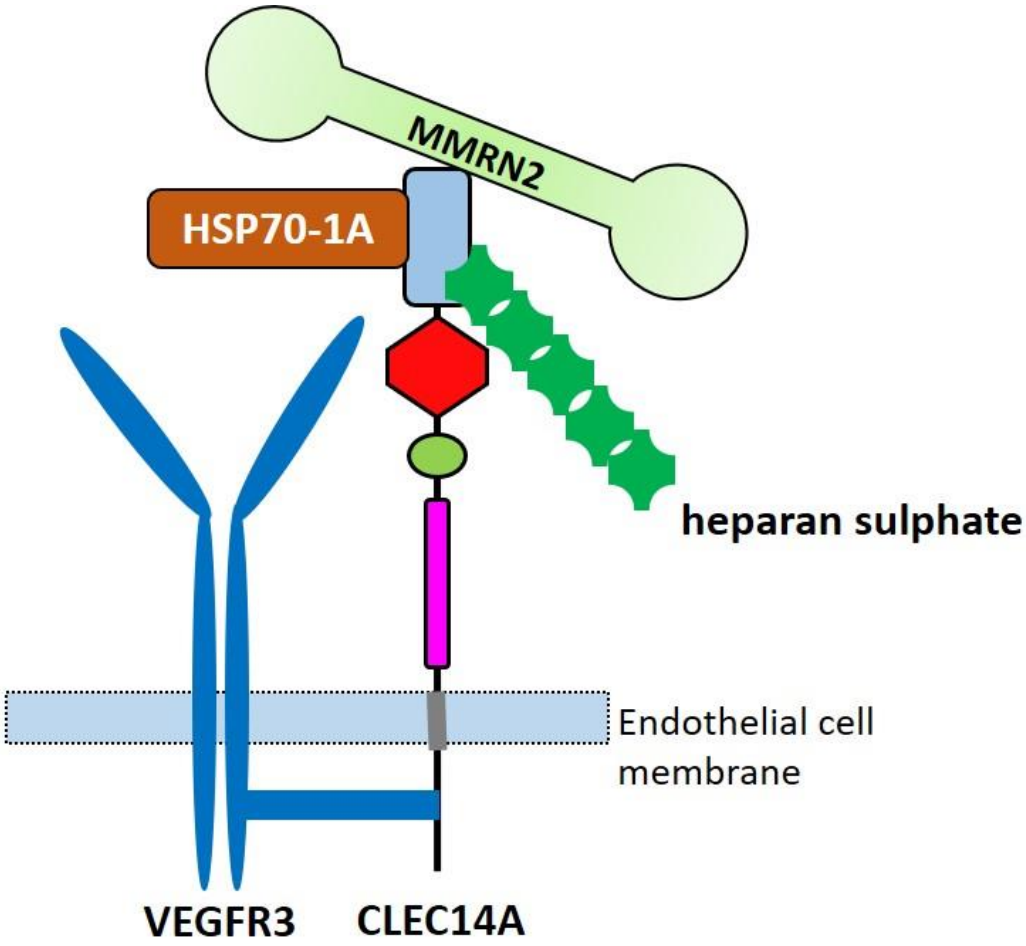


Figure 1.3. Schematic illustration of CLEC14A and its binding partners. CLEC14A interacts with MMRN2, heparan sulfate proteoglycans and Heat shock protein 70-1A via its extracellular N-terminal CTLD. VEGFR3 interacts with CLEC14A intracellularly. Each of these interactions except that with heparan sulfate proteoglycans have been shown to modulate angiogenesis.

1.3.4.3. CLEC14A-VEGF

Almost all the current known binding partners of CLEC14A interact with the extracellular portion of the protein, however, immunoprecipitation studies revealed the intracellular domain of CLEC14A interacts with VEGFR3 and this was confirmed using GFP-tagged CLEC14A. (Lee *et al.,* 2017). VEGFR3 and VEGFR2 are both expressed on the surface of endothelial cells and knockdown of VEGFR3, potentiates VEGFR2 activation and phosphorylation from VEGFA stimuli (Tammela *et al.,* 2011). Silencing of CLEC14A in HUVECs and VEGFA treatment also resulted in phosphorylation and activation of VEGFR2 however, since CLEC14A does not directly interact with VEGFR2 it is plausible that CLEC14A indirectly regulates VEGFR2 through its interaction with VEGFR3. Study of *clec14a*$^{-/-}$ mouse embryo retina showed decreased expression of VEGFR3 and increased expression of VEGFR2. Increased vascular density was also observed in *clec14a*$^{-/-}$ mice in the retina and lymphatic vessels (Lee *et al.,* 2017). This is the first data to suggest an antiangiogenic role for CLEC14A by attenuating VEGFR2 signalling.

1.3.4.4. CLEC14A-MMRN2

Another key binding partner of CLEC14A which has been widely studied is Multimerin 2 (MMRN2). MMRN2, also referred to as EndoGlyx-1, is a large extracellular matrix protein which is produced by ECs and deposited within the interface between ECs and pericytes. It was first identified as an antigen for the mAbH572 antibody which was generated from mice immunized with HUVECs. MAbH572 recognised MMRN2 within the extracellular matrix of blood vessels ranging from capillaries to arteries (Sanz-Moncasi *et al.,* 1994; Christian *et al.,* 2001). Due to its structure MMRN2 belongs to the EMILIN family of glycoproteins alongside EMILIN 1, EMILIN 2 and MMRN1. MMRN2 is comprised of 3 domains; an N-terminal EMI domain followed by a coiled-coil domain which is often subjected to N-linked and O-linked

glycosylation, and finally a C-terminal globular C1q domain (Christian *et al.,* 2001; Colombatti *et al.,* 2012).

The interaction between CLEC14A and MMRN2 was first identified by probing HUVEC whole cell lysate using Fc-tagged recombinant CLEC14A ECD coupled with mass spectrometry (Noy *et al.,* 2015). Mapping the interaction determined that the N-terminal CTLD of CLEC14A bound to a non-glycosylated portion of the coiled-coil domain of MMRN2 between amino acid 530-624 (Khan *et al.,* 2017). The significance of this interaction has been studied using inhibitors in the form of antibodies and peptides. Of the CRT antibodies introduced in section 1.3.3; CRT1, 4 and 5 each block the CLEC14A-MMRN2 interaction whilst CRT2 and CRT3 do not (Khan *et al.,* 2017). Treatment with the CRT4 antibody, which cross-reacts with both mouse and human forms of CLEC14A, reduced HUVECs migration and tube formation in scratch wound assays and matrigel assays, respectively, and sprout formation in aortic ring cultures. *In vivo*, 50ug of the CRT4 antibody or the isotype control was administered every 2 weeks to mice harbouring subcutaneously implanted Lewis lung carcinomas (LCC). A significant reduction in tumour size of CRT4 treated mice was observed compared to isotype control treated mice (Noy *et al.,* 2016).

Another inhibitor used to characterise the interaction between CLEC14A and MMRN2 is the MMRN2 495-674 fragment, the region of MMRN2 which binds CLEC14A, this inhibitor was compared with another MMRN2 fragment between amino acids 495-603 which does not abrogate the interaction with CLEC14A. HUVECs have been shown in adhesion assays to adhere to full length MMRN2 coated on a plate (Lorenzon *et al.,* 2012). This adherence of HUVECs was mirrored with plates coated with MMRN2 495-674 but was lost when plates were coated with MMRN2 495-603 (Khan *et al.,* 2017). Similar to that observed with CRT4,

blocking the CLEC14A-MMRN2 interaction using the mouse MMRN2 495-678 fragment fused to Fc inhibited the growth of subcutaneously implanted LLC tumour cells, compared with Fc control treatment (Khan *et al.,* 2017; Noy *et al.,* 2015). Taken together, this data implicates the CLEC14A-MMRN2 interaction to have a significant function in sprouting angiogenesis particularly in a tumour setting and is therefore an attractive target for anti-angiogenic therapy.

1.5. Multimerin 2

Multimerin 2 (MMRN2) is a key extracellular matrix protein which plays a significant role in maintaining vascular homeostasis and regulating angiogenesis by interacting with a range of binding partners to exert different functions. Alongside CLEC14A, MMRN2 interacts with VEGFA, matrix metalloproteinases and CD93 as discussed below (Figure 1.4).

1.5.1. MMRN2 expression

Using the anti-MMRN2 mAbH572 antibody it was shown that MMRN2 expression is restricted to the extracellular matrix of the endothelium both on the luminal and abluminal side of almost all healthy endothelium with the exception of hepatic and splenic sinusoids (Sanz-Moncasi *et al.,* 1994; Christian *et al.,* 2001). More notably, MMRN2 expression extended to the tumour endothelium and various pathologies, such as human skin cancer particularly at the invasive front in areas of angiogenesis (Huber *et al.,* 2006). Various studies have shown MMRN2 peptides to be present in malignant pleural effusion of advanced lung adenocarcinoma patients; urine of patients with early invasive breast cancer; plasma of symptomatic ovarian cancer patients; and spinal fluid of HIV patients. Moreover, increased plasma levels of MMRN2 peptides were also highly associated with myocardial infarctions (Soltermann *et al.,* 2008; Beretov *et al.,* 2015; Sheild-Artin *et al.,* 2012; Bora *et al.,* 2014).

MMRN2 shows promise as a potential biomarker for disease. Whilst MMRN2 expression was shown to increase in a range of pathologies, decreased expression is also apparent in cases such as gastric tumour samples where healthy gastric vessels express high levels of MMRN2 compared to gastric cancer samples (Andreuzzi *et al.,* 2018). Similarly, biopsies of colorectal cancer were also shown to exhibit decreased MMRN2 expression compared to normal samples at the protein level. The same study revealed decreased mRNA levels of MMRN2 in tumour-derived ECs from prostate, breast and kidney tumours compared to HUVECs (Andreuzzi *et al.,* 2017).

MMRN2 expression at a cellular level has been analysed using immunofluorescence staining in multiple studies, aortic ring assays showed MMRN2 expression was reduced in the leading tip and instead was deposited along the vessel (Andreuzzi *et al.,* 2017). Conversely, staining of the developing mouse retina revealed MMRN2 expression on the leading edge of sprouts and surrounding filopodia (Lugano *et al.,* 2018) suggesting MMRN2 expression may vary according to the context of angiogenesis.

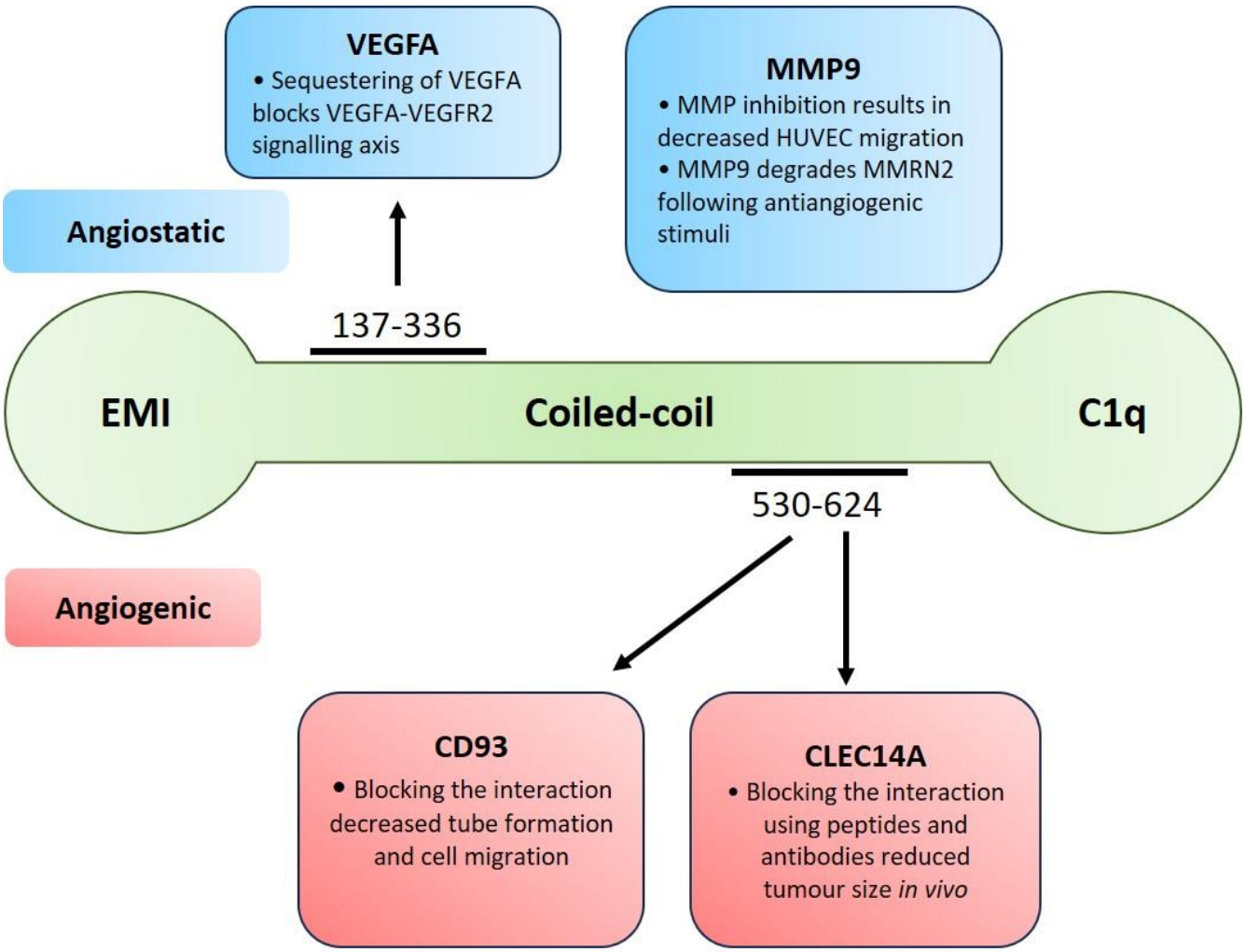


Figure 1.4. Illustration of the key binding partners of MMRN2 which elicit either angiostatic or angiogenic phenotypes. MMRN2 is comprised of an EMI domain following by a coiled-coil domain and then a C1q domain. The angiostatic function of MMRN2 is driven by binding with VEGFA or processing by MMP9. The angiogenic roles are mediated by interaction with heparin, and endothelial membrane proteins including CLEC14A and CD93.

1.5.2. MMRN2 function

MMRN2 has both pro and anti-angiogenic functions. As described below, while anti-angiogenic properties are associated with sequestration of VEGFR2 and MMP-9 processing, its pro-angiogenic properties are driven by interactions with CLEC14A and CD93 (Figure 1.4).

VEGFA interacts with MMRN2 between amino acids 137-336 of the coiled coil domain, an interaction promoted by MMRN2 glycosylation, evidenced by heparin's ability to block this interaction (Colladel *et al.,* 2015). MMRN2 can sequester VEGFA thus perturbing the pro-angiogenic VEGFA/VEGFR2 signalling axis (Lorenzon *et al.,* 2012). The angiogenic stimuli VEGFA and FGF-2 were shown to inhibit MMRN2 expression and to promote its degradation primarily via MMP-9 and to a lesser extent MMP-2. This degradation of MMRN2 by MMP-9 was shown to correlate with reduced HUVEC migration (Andreuzzi *et al.,* 2017). These studies suggest that MMRN2 with its ability to sequester VEGF has an antiangiogenic role, and that during angiogenesis its expression is reduced thus potentiating the function of VEGF.

Conversely, there is evidence to suggest an angiogenic function of MMRN2. A study conducted by Fejza *et al*, revealed that MMRN2 functions in pericyte adhesion, a key process in the final stages of sprouting angiogenesis as described in section 1.2. MMRN2 expression was found to upregulate in HUVECs in the presence of but not in direct contact with pericytes. Furthermore, presence of recombinant MMRN2 increased production of chemokines and cytokines such as PDGF and Heparin binding epidermal growth factor from both endothelial cells and pericytes which function to promote pericyte motility and adhesion. In this context MMRN2 does not initiate angiogenesis but instead plays a critical role in the later stages by exerting an adhesive function for both endothelial cells and pericytes (Fejza *et al.,* 2021).

These findings are consistent with *in vivo* studies of *mmrn2*$^{-/-}$ mice which had impaired pericyte recruitment leading to defective, collapsed vessels (Pellicani *et al.,* 2020).

MMRN2 has also been shown in the literature to exert a pro-angiogenic function predominantly through its interactions with CD93 and CLEC14A. As previously reviewed in section 1.3.3.4, the interaction between CLEC14A and MMRN2 appears to elicit a pro-angiogenic phenotype. MMRN2 also interacts with another member of the CLTD group 14 family, CD93 which is expressed on both healthy and tumour endothelial cells and binds to the coiled-coil domain of MMRN2 between amino acids 530-624 (Khan *et al.,* 2017). CD93 functions to promote cell-cell and cell-matrix adhesion as well as promote angiogenesis and tumour growth *in vivo* (Langenkamp *et al.,* 2015). Expression studies of CD93 and MMRN2 were performed using immunofluorescence and revealed that MMRN2 and CD93 colocalised in a range of different tumour types including retinoblastoma, malignant choroidal melanoma, endometrioid adenocarcinoma and others. Colocalization occurred on the abluminal side of blood vessels and MMRN2 was often found deposited to cells expressing high levels of CD93 (Lugano *et al.,* 2018; Galvagni *et al.,* 2017). A MMRN2-CD93 inhibitor in the form of a recombinant CD93 was shown to decrease HUVEC migration, tube formation, sprouting and transmigration. The study went on to generate a CD93 mutant, F238A, which could not bind to MMRN2. HUVECs silenced for CD93 had reduced capacity to migrate. Using lentiviral constructs resistant to CD93 silencing, the disturbed phenotype was rescued with WT CD93. F238A CD93 however maintained the phenotype of reduced cell migration. Moreover, the phenotype observed from HUVECs expressing F238A CD93 mirrored the phenotype from total knockdown of CD93 (Galvagni *et al.,* 2017). Blocking the interaction between CD93 and MMRN2 using monoclonal anti-CD93 antibody also resulted in decreased vascular sprouting in choroid sprouting assays (Tosi *et al.,* 2020). SiRNA-mediated knockdown

of MMRN2 (siMMRN2) resulted in decreased CD93 at the protein level but not at the mRNA level. MMRN2-dependent CD93 stabilisation was confirmed by blocking protein synthesis in both control and siMMRN2 HUVECs, the results indicated accelerated decline of CD93 in siMMRN2 HUVECs (Lugano *et al.,* 2018). Overall, it appears that CD93 and MMRN2 work synergistically to promote an angiogenic phenotype in the early stages of sprouting angiogenesis.

It is clear that MMRN2 functions in a context-dependent manner in order to regulate and maintain vascular stability and function. MMRN2 expression is often dysregulated in a range of pathologies which makes it a potential biomarker or drug target in certain pathologies. However, much is yet to be discovered about the mechanisms driving MMRN2 function.

1.6. Heparin and Heparan Sulfate

Heparan sulfate (HS) is a type of glycosaminoglycan (GAG) which is often expressed in conjunction with a protein core to form a heparan sulfate proteoglycan (HSPG). HSPGs are generally expressed as membrane proteins or extracellular matrix proteins and exert their functions by interacting with heparin-binding proteins (Olczyk *et al.,* 2015). They induce a plethora of functions varying from activation and inhibition of enzyme activity, regulating growth factor levels within the extracellular space, promoting oligomerization of proteins as well as regulating functions such as angiogenesis (Xu and Esko, 2014). Heparin is another GAG which is structurally similar to HS but contains more sulfate groups and therefore has an overall negative charge (Shriver *et al.,* 2012). It is regularly used as a lab-based alternative to HS and used to study the HS interactome. The significance of HS has been studied using mouse models which have certain genes involved in the biosynthesis of HS knocked out. For the most part these mutants generally display defects in growth however the specific phenotype

depends on which genes in the biosynthesis pathway of HS are knocked out (Fuster and Wang, 2010).

1.6.1. HSPGs in the vascular system

HSPGs are expressed on the surface of ECs or within the extracellular space between ECs and pericytes and regulate aspects of the vascular system including angiogenesis and inflammation (Olczyk *et al.,* 2015). The VEGF family of protein and receptors play a key role in angiogenesis and are modulated by HS (Zhao *et al.,* 2012). The interaction between heparan sulfate and VEGFA heavily relies on the sulfate groups of HS and upon removal of the sulfate groups from HS, HUVECs exhibit reduced tube formation (Ashikara-Hada *et al.,* 2005). A study conducted by Wijelath *et al,* involved the generation of mutant mouse endothelial cells which lacked cell surface HS, the cells displayed reduced VEGF signalling despite presence of VEGFA, this was indicated by decreased ERK phosphorylation. Surface plasmon resonance (SPR) techniques used to measure VEGFA-VEGFR2 binding revealed enhanced binding between VEGFA and VEGFR2 in the presence of unfractionated heparin. The mechanism by which heparin supported the interaction is yet unknown however it was hypothesized that cell surface HSPG bind to and stabilised VEGFA at the cell surface to activate VEGFR2 (Wijelath *et al.,* 2012). To this end the HS-VEGFA interaction is not absolutely required for the VEGFA/VEGFR signalling axis but certainly supports the signalling by stabilising VEGFA.

1.7. Hypothesis and aims

The work in this thesis follows on from extensive studies of CLEC14A, from its identification as a TEM, to the discovery of its interaction with MMRN2, and determining its function and that of MMRN2.

The following hypotheses and aims were explored in this project.

1. *The loop-in-loop structure within the CTLD of CLEC14A directly interacts with MMRN2 and the sequence involved can be used in targeting the interaction.*

   Aim 1: To further characterise the interaction and determine key residues of CLEC14A and MMRN2 which interact with one another by using AlphaFold predictive modelling and molecular biology techniques such as far western blotting and flow cytometry.

2. *The interaction between CLEC14A and heparin contributes to the function of CLEC14A.*

   Aim 2: To identify if the heparin binding sites converged between CLEC14A and MMRN2 as well as evaluate the effect, if any, of the CRT antibodies on the CLEC14A-heparin interaction using pull down assays with heparin beads.

3. *The interaction between CLEC14A and MMRN2 has a strong binding affinity.*

   Aim 3: To quantify the binding affinity between CLEC14A and MMRN2 using the NanoBRET assay.

4. *MMRN2 likely forms a trimer and binds to the CTLD of CLEC14A.*

   Aim 4: To determine the structure of the interaction using biophysical techniques such as analytical ultracentrifugation and Cryo-electron microscopy.

CHAPTER 2: Materials and Methods

Chapter 2: Materials and Methods

All chemicals used were from Sigma-Aldrich unless stated otherwise.

2.1. Reagents

2.1.1. Commonly used buffers

Table 2.1. Commonly used buffers and the concentration of each component.

Name	Components
Phosphate buffered Saline solution (PBS) (Gibco)	40 mM NaCl, 10 mM Na_2HPO_4, 2.7 mM KCl and 1.76 mM KH_2PO_4, pH 7.4
Tris-acetate-EDTA (TAE)	40 mM Tris (Millipore), 20 mM Acetic acid, 1 mM EDTA
2 x SDS-PAGE sample buffer (reducing)	50 mM Tris pH 6.8 (Millipore), 2% (w/v) SDS (fisher scientific), 10% (v/v) Glycerol (Millipore), 10% (v/v) β-mercaptoethanol, 0.03% (w/v) Bromophenol blue
2 x SDS-PAGE sample buffer (non-reducing)	50 mM Tris pH 6.8 (Millipore), 2% (w/v) SDS (fisher scientific), 10% (v/v) Glycerol (Millipore), 0.03% (w/v) Bromophenol blue
SDS-PAGE running buffer	25 mM Tris (Millipore), 250 mM Glycine, 0.1% (w/v) SDS (fisher scientific), pH 8.3
Western blot transfer buffer	25 mM Tris (Millipore), 187.2 mM Glycine
Tris Buffered Saline with Tween (TBST)	150 mM NaCl, 20 mM Tris pH 7.5 (Millipore), 0.1% (v/v) Tween
NP40 lysis buffer	1% (v/v) NP40, 10 mM Tris pH 7.5 (Millipore), 150mM NaCl and 1 mM EDTA with 1% (v/v) protease inhibitor
Flow cytometry buffer	PBS with 0.2% (w/v) BSA (VWR life science), 0.02% (w/v) Sodium Azide
Hanks buffered Salt solution (HBSS)	2 mM Sodium Pyruvate (Gibco), 145 mM NaCl, 10 mM HEPES, 5 mM KCl, 1 mM $MgSO_4.7H_2O$, 1.3 mM $CaCl_2.2H_2O$, 1.5 mM $NaHCO_3$, 10 mM D-Glucose, 0.1% (w/v) BSA (VWR Life Science)

2.1.2. Antibodies

Table 2.2. Primary and Secondary antibodies. Listed with manufacturers details, species raised in and application. Western blot (WB), Flow cytometry (FC), enzyme-linked immunosorbent assay (ELISA), Immunoprecipitation (IP).

Antibody	**Manufacturer and catalogue number**	**Species**	**Application (working concentration/dilution)**
Human MMRN2	Sigma-Aldrich (020241)	Rabbit	WB (1 µg/mL)
Human CLEC14A polyclonal	R&D system (AF4968)	Sheep	WB (0.1 µg/mL)
Myc tag monoclonal	Cell signalling technology (9B11)	Mouse	WB (1:5000)
Human IgG Fc-HRP conjugated	Sigma-Aldrich (AP113P)	Goat	WB and ELISA (1:2500)
6-His tag HRP conjugated	Cambridge Bioscience (A190-114P)	Rabbit	WB and ELISA (1:5000)
Snap tag polyclonal	Invitrogen (CAB4255)	Rabbit	WB (1 µg/mL)
Alexa Fluor 488	Invitrogen (A11094)	Rabbit	WB (1:2500)
Anti-His conjugated to AF647	BioLegend (J095G46)	Mouse	FC (0.2 µg/mL)
Anti-mouse IgG conjugated to Alexa Fluor 674	Invitrogen (A21237)	Goat	FC (5 µg/mL)
CLEC14A monoclonal (Clone CRT2)	N/A produced in laboratory	Mouse	FC (10 µL/mL), IP (1 µg) Binding assay (10 µg)
CLEC14A monoclonal (Clone CRT3)	N/A produced in laboratory	Mouse	FC (10 µL/mL), IP (1 µg), Binding assay (10 µg)

CLEC14A monoclonal (Clone CRT4)	N/A produced in laboratory	Mouse	FC (10 µL/mL), IP (1 µg), binding assay (10 µg)
Normal mouse IgG	Invitrogen (10400C)	Mouse	IP (1 µg), binding assay (10 µg)
Anti- mouse HRP	CAYMAN chemical company (0565692-2)	Goat	WB (1:5000)
Anti- rabbit HRP	Invitrogen (A16124)	Goat	WB (2 µg/mL)
Anti- sheep HRP	Millipore (AP147P)	Rabbit	WB (1:5000)
Anti- goat HRP	Dako (P0160)	Rabbit	WB (1:5000)

2.1.3. Beads, Recombinant proteins and conditioned media

Table 2.3. Beads, Recombinant proteins and conditioned media. Listed with manufacturers details or method of production.

Beads	**Source**
Protein A Sepharose beads	Sigma-Aldrich (MKCL1655)
Protein G beads	Sigma-Aldrich (P3296)
Glutathione agarose beads	Thermo Scientific (16100)
Heparin-Sepharose beads	Abcam (ab193268)
Heparin- Agarose beads	Merck life sciences (H6508-5ML)
Ni-NTA Agarose beads	Qiagen (1018244)
CLEC14A ECD purified protein	Produced in lab using HEK293T cells (section 2.5.1)
CLEC14A ECD- Fc conditioned media	Produced in lab using HEK293T cells (section 2.5.2)
MMRN2 495-674- His purified protein and conditioned media	Produced in lab using HEK293T cells (section 2.5.3)
MMRN2 coiled-coil- His conditioned media	Produced in lab using HEK293T cells (section 2.5.3)

2.2. Cell culture

Cell culture techniques were carried out in a laminar flow hood under sterile conditions. Human embryonic kidney cells (HEK293T) were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% v/v Foetal Calf Serum (FCS) (Thermo Fisher), 2 mM L-glutamine (Thermo Fisher) and 20 µg/mL penicillin/streptomycin (Thermo Fisher) (cDMEM). Cells were grown in a humidified incubator at 37°C, 5% v/v CO_2. Upon 80% confluency, the cDMEM was removed and the cells were washed with sterile PBS to remove any excess cDMEM. The cells were then incubated with 2X Trypsin-EDTA (0.5% (v/v) Trypsin 5.3 Mm EDTA.4Na; Thermo Fisher) at room temperature for 3 mins or until the cells were visibly detached from the base of the dish. Detached cells were resuspended in cDMEM to neutralise the Trypsin-EDTA and replated onto new dishes, diluted appropriately with cDMEM.

2.2.1. Plasmid transfection

HEK293T cells were transfected with plasmid DNA using polyethylenimine (PEI). For 10 cm dishes, cells were seeded at a density of 3 x 10^6 HEK293T in cDMEM 24 hrs before transfection. The next day a transfection mix was made using 1 mL of sterile Opti-MEM media and 9 µg of plasmid DNA, this was then vortexed well before adding 100 µL of PEI (1 mg/mL). The sample was then vortexed well and incubated at room temperature for 10 mins to allow the DNA and PEI to form complexes. The transfection mix was then pipetted onto the cells drop-wise around the dish and mixed by tilting the plate in a north to south and then east to west movement to ensure the transfection mix covered the full area of the plate. Volumes of each component were scaled according the assay being performed (Table 2.4).

Table 2.4. Reagent quantities required for different vessel sizes.

Vessel type	Cells plate	Media volume	OptiMEM	DNA	PEI
6-well plate	3 x 10^5	2 mL	150 µL	2 µg	15 µL
6 cm dish	1 x 10^6	3 mL	300 µL	3 µg	30 µL
10 cm dish	3 x 10^6	10 mL	1 mL	9 µg	100 µL
15 cm dish/ T175 flask	6 x 10^6	20 mL	2.25 mL	36 µg	225 µL

2.3. Molecular biology

2.3.1. Restriction enzyme digest

Restriction enzyme digests were carried out using reagents from New England Biolabs and each restriction enzyme used was compatible in the CutSmart buffer. Restriction digests were set up to ensure an excess of restriction enzyme per µg of DNA where 0.5 µL (10 units) of restriction enzyme were used per µg of DNA. Small scale digests used for diagnostic purposes were processed in a total volume of 10 µL. For this, 1 µg of DNA was used, followed by 1 µL of 10x CutSmart buffer and 0.5 µL of each restriction enzyme of interest, the final volume was made up using dH_2O. Large scale digestions involved the use of 5 µg of DNA and reagents were scaled up.

2.3.2. Polymerase Chain reaction (PCR)

Polymerase chain reaction (PCR) to generate fragments for cloning was performed using primers described in Table 2.2 using the high-fidelity enzyme Q5 (NEB). The basic PCR reaction was made up using 1X QC buffer, 1X QC enhancer buffer, 200 µM dNTPs, 0.3 µM forward primer, 0.3 µM reverse primer, 50 ng DNA template, dH_20 up to 99 µL and 1-unit Q5 polymerase. The PCR cycle parameters were as follows; initial denaturation at 98°C for 30 seconds, further denaturation at 98°C for 30 seconds, annealing at 55°C for 20 seconds, extension at 72°C for 30 seconds/ kb (32 cycles) and final extension at 72°C for 7 mins.

Colony PCRs were performed using 2x RedTaq ready mix (Sigma-Aldrich), with forward and reverse primers to amplify the area of interest, dH_2O was used to dilute the master mix to 1X. The final volume varied based on the number of colonies being screened where each colony requires a 15µL reaction mix. Individual colonies were picked from the parent plate, dotted onto a fresh agar plate enriched with the same antibiotic as the parent plate and resuspended into the 15µL reaction mix. The PCR reaction cycle was the same as stated above and the PCR product was run on an agarose gel and visualised using blue light as described in section 2.3.3.

2.3.3. DNA agarose gel electrophoresis

Agarose gels were made up at different percentages based on the size of the DNA fragments being run, for example 1% (w/v) gels were used for larger constructs at ~10k bases and 2% (w/v) gels were used for ~500bp fragments. Gels were made up in 1X TAE buffer and SYBR safe DNA gel stain (Invitrogen) was added at a dilution of 1 µL per 10 mL of agarose solution, this allowed for visualisation of the DNA under blue light. DNA samples were diluted in loading dye (New England Biolabs) and loaded into the wells of the gel. Gels were run at either 100V for 30 mins or 80V for 40 mins depending on the size of the gel and separation required. Gels were visualised using the UV transilluminator (Cleaver scientific).

2.3.4. DNA Gel extraction

DNA bands of interest were excised out using a clean scalpel and subjected to gel extraction using the gel extraction kit (Thermo scientific) as per the manufacturer's instructions. Concentrations of fragments were measured using the NanoDrop (Geneflow ND-1000).

2.3.5. Molecular cloning

In order to generate and manipulate plasmids two methods of molecular cloning were employed. These were conventional cloning (Figure 2.1A) and Gibson assembly cloning

(Figure 2.1B). Both of which generally involved amplification of an area of interest being inserted into digested plasmids. Corresponding primers used are displayed in Table 2.5.

2.3.5.1. Conventional cloning

Following PCR amplification of the inserts according to section 2.3.2, the PCR product was purified using the QIAquick PCR purification kit as per the manufacturers instruction with the exception of eluting the PCR product in 40 µL dH_2O (Qiagen). The sample was then subjected to restriction digest by addition of 5 µL of 10X Cutsmart buffer followed by 2.5 µL of the required restriction enzymes. The digestion mix was incubated for 1 hr at 37°C. The plasmid of interest was also digested as described in section 2.3.1. The samples were then subjected to gel electrophoresis and extracted as described in sections 2.3.3 and 2.3.4, respectively. Ligation reactions were made up to a final volume of 10 µL. This included the PCR inserts and the cut vectors with complimentary ends using a molar ratio of 3:1 respectively. 1 µL of T4 DNA ligase and 1 µL of 10X T4 DNA ligase buffer where then added. To make up the final volume, dH_2O was added. The ligation reactions were incubated at room temperature for 1 hr and then 5 µL of the mixture was transformed into competent DH5α competent *E. coli* (New England Biolabs) as described in section 2.3.6 (Figure 2.1A).

2.3.5.2 Gibson Assembly cloning

The NEBuilder program was recruited to design the primers used to amplify the sequence of interest of which at least 18 nucleotides were gene specific, the inserts were then amplified by PCR as described in section 2.3.2. using the recommended annealing temperature from the NEBuilder program. The vector of interest was linearised with the appropriate restriction enzymes as described in section 2.3.1. The ratio of linearised vector to insert was dependent on the number of inserts. In the case of 1-3 inserts a molar ration of 1:2 was used. An equal

volume of 2X Gibson reaction master mix (New England Biolabs) was added generally aiming for a final volume of 10 µL. The mixture was incubated at 50°C for 1 hr. 5 µL of the reaction mix was then transformed into DH5α competent *E. coli* (New England Biolabs) as described in section 2.3.6 (Figure 2.1B).

2.3.5.3 Linker addition

Short linkers were added to some pHLAvitag3 constructs to improve flexibility. Linker DNA was ordered from Eurogentec as two complimentary single strands, whereby the 5' end contained the cut sequence of KpnI and the 3' end contained the cut sequence for EcoRV. Between the end of the linker sequence and the EcoRV site, the sequence for the restriction enzyme PmeI was inserted as well as 6 extra bases to allow space between the restriction sites so both the PmeI and EcoRV sites could be cut simultaneously. Oligomers were bound together by heating at 95^{o}C at equimolar concentrations and allowing to cool to room temperature. Linkers were inserted into the KpnI and EcoRV digested pHLAvitag3 vector as described in section 2.3.5.1.

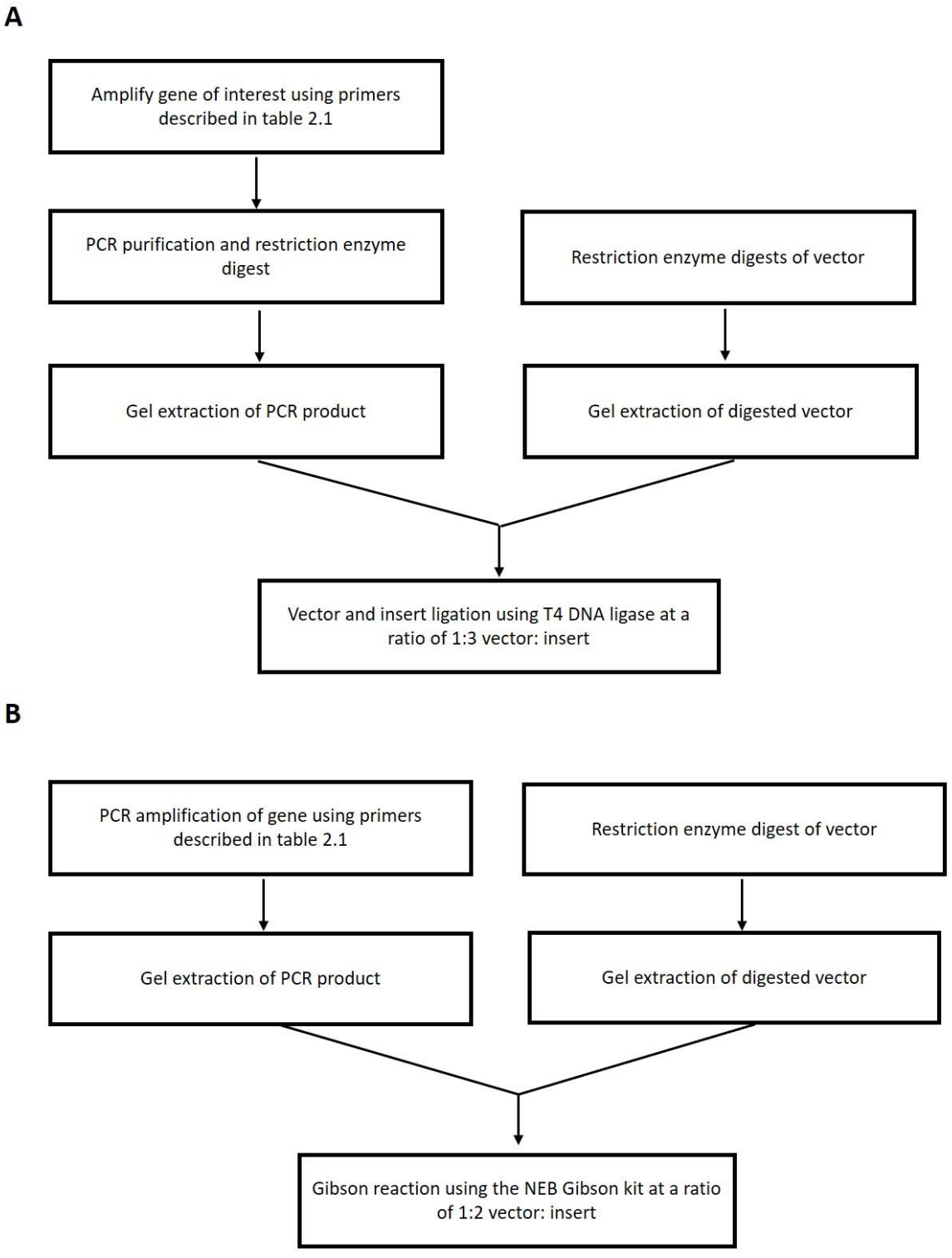


Figure 2.1. Flow diagram describing construction of plasmids. (A) Construction of plasmid using traditional restriction digest and ligation. (B) Construction of plasmids using Gibson assembly reactions.

2.3.6. Heat shock Transformations

Competent strains of DH5α *E. coli* (New England Biolabs) were thawed on ice in 20 µL aliquots. 5 µL of a ligation mix was added to the cells and the tube was gently flicked to ensure the sample was mixed and the cells were incubated on ice for a further 30 mins. The cells were then heat-shocked on a 42°C hot plate for 30 seconds and immediately placed back on ice to incubate for a further 2 mins. 200 µL of SOC outgrowth media (New England Biolabs) was then added to the cells and left to incubate at 37°C in a shaking incubator for 40 min. Finally, the full transformation mix was spread onto pre-warmed LB agar plates containing Ampicillin (Sigma-Aldrich) at 100 µg/mL. The plates were left to grow colonies overnight at 37°C.

2.3.7. Plasmid DNA isolation

Individual colonies carrying the plasmid of interest were cultured in LB Broth containing 100 µg/mL Ampicillin overnight (approximately 16 hrs) at 37°C with orbital shaking (180rpm). Small scale protein preps used 1.5 mL of cultured LB broth and a mini prep kit (Thermo Scientific), midi preps used 50 mL of cultured LB broth (Sigma-Aldrich). Large scale preps were performed using 500 mL of cultured LB broth and a maxi prep kit (Qiagen). Each kit was used as per the manufacturer's instructions and eluted in elution buffer (10 mM Tris, pH 8.5, 0.1 mM EDTA). Plasmid concentrations were measured using a NanoDrop (Geneflow ND-1000).

2.3.8. Sequencing

Plasmid DNA was sequenced using sanger sequencing (Source Bioscience). DNA was prepared at a concentration of 100 ng/µL using elution buffer. 5 µL of DNA was required for each sequence run. The primers used were either from Source bioscience stock primers or specially designed sequencing primers. The primers were prepared at a concentration of 3.5 pM using dH_2O and 5 µL were required per reaction.

2.3.9. Primers

Primers were ordered from Eurogentec and reconstituted in nuclease free water at a concentration of 100 µM. Table 2.4 displays the primers used in this project including their sequence and individual uses.

Table 2.5. List of primers used for PCR reactions, cloning and sequencing. Sequences are displayed 5' to 3'.

Name	Description	Sequence
VH609	Fc Reverse for pIgPP Fc (sequencing)	CAGTTGAACTTGACCTCAGG
VH206	T7 Forward for pIgPP Fc (Sequencing)	TAATACGACTCACTATAGGG
EGFP-N	GFP rev (sequencing)	CGTCGCCGTCCAGCTCGACCAG
VH575	CLEC14A 5' 1st half forward for pWPXL	CGAGACTAGCCTCGAGGTTTAAACATGAGGCCGGCGTTCGCC
VH588	CLEC14A 3' 2nd half reverse for pWPXL	ATCATATGACTAGTCCCGGGAATTCTCACAAAGCTTGAGCTCG
647	CLEC14A reverse (654-674) (sequencing)	GTAACTGAGATCGGGAGCTGT
AB139	CLEC14A 5' 1st half forward for pIgPP	AACGGCCGCCAGTGTGCTGGAATTCAGAATGAGGCCGGCGTTC
AB142	CLEC14A 3' 2nd half reverse for pIgPP	CTCACTCGATGAAGAACCGCGGCCGCCCACGGCAGAGGAGGAGTC
AB140	R98A CLEC14A 1st half 3' reverse	AACGCCTGGCCTCCAGTGCGACCCAGAAC

AB141	R98A CLEC14A 2nd half 5' forward	CGCACTGGAGGCCAGGCGTTCCCACTGC
AB143	R100A CLEC14A 1st half 3' reverse	AGTGGGAAGCCCTGCGCTCCAGTGCGAC
AB144	R100A CLEC14A 2nd half 5' forward	GGAGCGCAGGGCTTCCCACTGCACCCTG
AB145	R98,99,100A CLEC14A 1st half 3' reverse	AAGCCGCGGCCTCCAGTGCGACCCAGAAC
AB146	R98,99,100A CLEC14A 2nd half 5' forward	CGCACTGGAGGCCGCGGCTTCCCACTGC
AB019	S101A CLEC14A 1st half reverse	TGCAGTGAGCACGCCTGCGCTCCAGTGC
AB020	S101A CLEC14A 2nd half forward	GCGCAGGCGTGCTCACTGCACCCTGGAG
AB021	H102A CLEC14A 1st half reverse	GGGTGCAAGCGGAACGCCTGCGCTCCAG
AB022	H102A CLEC14A 2nd half forward	CAGGCGTTCCGCTTGCACCCTGGAGAAC
AB023	L105A CLEC14A 1st half reverse	CGTTCTCAGCGGTGCAGTGGGAACGCCTG
AB024	L105A CLEC14A 2nd half forward	CCACTGCACCGCTGAGAACGAGCCTTTG
AB025	E106A CLEC14A 1st half reverse	GCTCGTTAGCCAGGGTGCAGTGGGAACGC
AB026	E106A CLEC14A 2nd half forward	CTGCACCCTGGCTAACGAGCCTTTGCGG
AB027	N107A CLEC14A 1st half reverse	AAGGCTCAGCCTCCAGGGTGCAGTGGGAAC

AB028	N107A CLEC14A 2nd half forward	CACCCTGGAGGCTGAGCCTTTGCGGGGT
AB029	E108A CLEC14A 1st half reverse	GCAAAGGAGCGTTCTCCAGGGTGCAGTGGG
AB030	E108A CLEC14A 2nd half forward	CCTGGAGAACGCTCCTTTGCGGGGTTTC
AB031	E106K CLEC14A 1st half reverse	GCTCGTTCTTCAGGGTGCAGTGGGAACGC
AB032	E106K CLEC14A 2nd half forward	CTGCACCCTGAAGAACGAGCCTTTGCGG
AB033	E108K CLEC14A 1st half reverse	GCAAAGGCTTGTTCTCCAGGGTGCAGTGGG
AB034	E108K CLEC14A 2nd half forward	CCTGGAGAACAAGCCTTTGCGGGGTTTC
AB037	E106,108K CLEC14A 1st half reverse	GCTTGTTCTTCAGGGTGCAGTGGGAACGC
AB038	E106,108K CLEC14A 2nd half forward	CTGCACCCTGAAGAACAAGCCTTTGCGG
AB039	E124K CLEC14A 1st half reverse	TGTCGCTCTTGAGACCGCCGGGGTCGGA
AB040	E124K CLEC14A 2nd half forward	CGGCGGTCTCAAGAGCGACACGCTGCAG
AB041	E124A CLEC14A 1st half reverse	TGTCGCTCGCGAGACCGCCGGGGTCGGA
AB042	E124A CLEC14A 2nd half forward	CGGCGGTCTCGCGAGCGACACGCTGCAG
AB043	E153K CLEC14A 1st half reverse	CTGCGGGCTTGACCCCACCGGTGGCCTG
AB044	E153K CLEC14A 2nd half forward	CGGTGGGGTCAAGCCCGCAGGCTGGAAG

AB045	E153A CLEC14A 1st half reverse	CTGCGGGCGCGACCCCACCGGTGGCCTG
AB046	E153A CLEC14A 2nd half forward	CGGTGGGGTCGCGCCCGCAGGCTGGAAG
AB047	S101A, H102A CLEC14A 1st half reverse	TGCACGCTGCACGCCTGCGCTCCAGTGC
AB048	S101A, H102A CLEC14A 2nd half forward	GCGCAGGCGTGCAGCGTGCACCCTGGAG
AB049	L105A, N107A CLEC14A 1st half reverse	CCGCCTCCGCGGTGCAGTGGGAACGCCTG
AB050	L105A, N107A CLEC14A 2nd half forward	CCACTGCACCGCGGAGGCGGAGCCTTTG
AB051	S137A, T139A and R141 CLEC14A 1st half reverse	CAGCGCATGCGCGTTGGGGCTCCTCCAC
AB052	S137A, T139A and R141 CLEC14A 2nd half forward	GCCCCAACGCGCATGCGCTGCGGCCAGA
AB071	S137A CLEC14A 1st half reverse	CGGTGCATGCGCGTTGGGGCTCCTCCAC
AB072	S137A CLEC14A 2nd half forward	GCCCCAACGCGCATGCACCGCGCGGAGA
AB073	T139A CLEC14A 1st half reverse	TCCGCGCTGCGCAGGAGCGTTGGGGCTC
AB074	T139A CLEC14A 2nd half forward	ACGCTCCTGCGCAGCGCGGAGATGCGCG

AB075	R141A CLEC14A 1st half reverse	CGCATCTTGCCGCGGTGCAGGAGCGTTG
AB076	R141A CLEC14A 2nd half forward	CTGCACCGCGGCAAGATGCGCGGTACTC
AB077	S137,T139A CLEC14A 1st half reverse	CAGCGCATGCGCGTTGGGGCTCCTCCAC
AB078	S137,T139A CLEC14A 2nd half forward	GCCCCAACGCGCATGCGCTGCGCGGAGA
AB079	S137,R141A CLEC14A 1st half reverse	CGGTGCATGCGCGTTGGGGCTCCTCCAC
AB080	S137,R141A CLEC14A 2nd half forward	GCCCCAACGCGCATGCACCGCGGCTAGA
AB081	T139,R141A CLEC14A 1st half reverse	TAGCCGCTGCGCAGGAGCGTTGGGGCTC
AB082	T139,R141A CLEC14A 2nd half forward	ACGCTCCTGCGCAGCGGCTAGATGCGCG
AB083	R142A CLEC14A 1st half reverse	CCGCGCATGCCCGCGCGGTGCAGGAGCG
AB084	R142A CLEC14A 2nd half forward	CACCGCGCGGGCATGCGCGGTACTCCAG
AB117	MMRN2 495-674 5' 1st half forward	GATGGGTTGCGTAGCTGAAACCGGTCAGAAGCTCTATTTAGAC
AB118	MMRN2 495-674 3' 2nd half reverse	GGAACCACCGGAACCTCCGGTACCCGGCCGCGGGGG
AB099	E617A MMRN2 1st half reverse	CCAGCACTGCCTCCCCGAAGAGCGCGGC
AB100	E617A MMRN2 2nd half forward	CTTCGGGGAGGCAGTGCTGGAGGAGATG
AB101	E620A MMRN2 1st half reverse	ACATCTCTGCCAGCACCTCCTCCCCGAAGAGC

AB102	E620A MMRN2 2nd half forward	GGAGGTGCTGGCAGAGATGTCTGAGCAG
AB103	S623A MMRN2 1st half reverse	TCTGCTCTGCCATCTCCTCCAGCACCTCCTCCCC
AB104	S623A MMRN2 2nd half forward	GGAGGAGATGGCAGAGCAGACGCCGGGA
AB105	E624A MMRN2 1st half reverse	GCGTCTGTGCAGACATCTCCTCCAGCACCTCCTCC
AB106	E624A MMRN2 2nd half forward	GGAGATGTCTGCACAGACGCCGGGACCG
AB119	E617,E620,S623,E624 MMRN2 1st half reverse	GTGCTGCCATCTCTGCCAGCACTGCCTC
AB120	E617,E620,S623,E624 MMRN2 2nd half forward	GCTGGCAGAGATGGCAGCACAGACGCCGGGACC
AB013	BamHI CLEC14A forward w/o signal peptide	TAGTAGGGATCCGAACACCCCACTGCCGAC
AB014	XhoI reverse CLEC14A	TAGTAGCTCGAGCTATGCATCACTAGAGCC
AB015	XbaI reverse CLEC14A	TAGTAGTCTAGACTATGCATCACTAGAGCC
AB065	MMRN2 495-674 BamHI	TAGTAGGGATCCACCGGTCAGAAGCTCTATTTAGACCTG G
AB066	XbaI reverse His-tag	TAGTAGTCTAGATCATTAGTGATGGTGATGGTGGTGCTTGGT
AB067	XbaI reverse His-tag	TAGTAGCTCGAGTCATTAGTGATGGTGATGGTGGTGCTTGGT
AB093	Signal peptide pcDNA3.1 forward	TTTAAACTTAAGCTTGGTACCGCCACCATGCGGCTCTGC
AB094	Signal peptide pcDNA3.1 forward	AGCCGGTCACCAGGGTAAGCTTCCGGCTG

AB095	SmBiT amplification forward	GCTTACCCTGGTGACCGGCTACCGGCTG
AB096	SmBiT amplification reverse	AGCGGGTTTAAACGGGCCCTCTAGACTCGAGCTATGCATCAC TAGAGCC
AB131	Linker sequence forward	CGGAGGTGGCGGGTCTGGAGGTGGCGGGAGTGGAGGTGG CGG GTCAAGTTTAAACGTCGCAGAT
AB132	Linker sequence reverse	ATCTGCGACGTTTAAACTTGACCCGCCACCTCCACTCCCGCCA CC TCCAGACCCGCCACCTCCGGTAC
AB133	Nluc into pAvitag linker forward	GGAGGTGGCGGGTCAAGTTTAAACGTCTTCACACTCGAAGA TTTCGTTG
AB134	Nluc into pAvitag linker reverse	ATCTTCTGAGCTTCAAAGATATCCGCCAGAATGCGTTCGCAC
AB135	SNAP into pAvitag linker forward	GGAGGTGGCGGGTCAAGTTTAAACGACAAAGACTGCGAAAT GAAGCGC
AB136	SNAP into pAvitag linker reverse	ATCTTCTGAGCTTCAAAGATATCCAGCCCAGGCTTGCCCAG
AB137	LgBiT into pAvitag linker forward	GGAGGTGGCGGGTCAAGTTTAAACGTCTTCACACTCGAAGA TTTC
AB138	LgBiT into pAvitag linker reverse	ATCTTCTGAGCTTCAAAGATATCGTTGATGGTTACTCGGAAC
AB107	Nluc into pAvitag C-term forward	GTGGTTCCGGTCTGAATGATATCGTCTTCACACTCGAAGATT TCGTTG
AB108	Nluc into pAvitag C-term reverse	ATCTTCTGAGCTTCAAAGATATCCGCCAGAATGCGTTCGCAC

AB109	SNAP into pAvitag C-term forward	GTGGTTCCGGTCTGAATGATATCGGCTGTCGAGGCGCGGCG
AB110	SNAP into pAvitag C-term reverse	ATCTTCTGAGCTTCAAAGATATCTAACCTCAAGTCA AGGCTTTTCTATGGAATAAGGAATGGACAGC
AB121	pHLAvitag3 5' forward (sequencing)	AAAGAATTGCGGCCGTCTCAGGC
AB122	pHLAvitag3 3' reverse (sequencing)	TTCTGATAGGCAGCCTGCACCTG
AB063	MMRN2 forward (1495-1515) (sequencing)	TTAGACCTGGACGTCATCCGG
AB064	MMRN2 reverse (1700-1718) (sequencing)	TCATCCAGCGCCTGCACTT

2.3.10. Plasmids

The plasmids used in this project were either constructed through Gibson assembly (Table 2.6) or conventional cloning (Table 2.7) or provided by others (Table 2.8). Details of each construct are present in Tables 2.6 through to 2.8.

Table 2.6. List of plasmids constructed through Gibson assembly. Displaying plasmid names, the cloning method, primers used and tags or notable features within the construct. Fragments were generated from WT templates.

Plasmid	Cloning method	Primers used	Tags/features
pWPXL- R98A CLEC14A	Gibson assembly of 2 PCR product into linearised (PmeI and EcoRI) pWPXL	VH575-2 3-VH588	Myc tag, GFP
pWPXL- R100A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-5 6-VH588	Myc tag, GFP
pWPXL- R98,99,100A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-7 8-VH588	Myc tag, GFP
pWPXL- S101A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB019 AB020-VH588	Myc tag, GFP
pWPXL- H102A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB021 AB022-VH588	Myc tag, GFP
pWPXL- S101A, H102A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB047 AB048-VH588	Myc tag, GFP
pWPXL- L105A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB023 AB024-VH588	Myc tag, GFP
pWPXL- N107A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB049 AB050-VH588	Myc tag, GFP
pWPXL- L105A, N107A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB023 AB024-VH588	Myc tag, GFP

pWPXL- E106A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB025 AB026-VH588	Myc tag, GFP
pWPXL- E106K CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB031 AB032-VH588	Myc tag, GFP
pWPXL- E108A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB029 AB030-VH588	Myc tag, GFP
pWPXL- E108K CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB033 AB034-VH588	Myc tag, GFP
pWPXL- E106, E108K	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB037 AB038-VH588	Myc tag, GFP
pWPXL- S137, T139, R141A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB051 AB052-VH588	Myc tag, GFP
pWPXL- S137, T139A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB077 AB078-VH588	Myc tag, GFP
pWPXL- S137, R141A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB079 AB080-VH588	Myc tag, GFP
pWPXL- T139, R141A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB081 AB082-VH588	Myc tag, GFP
pWPXL- S137A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB071 AB072-VH588	Myc tag, GFP
pWPXL- T139A CLEC14A	Gibson assembly of 2 PCR products into linearised (PmeI and EcoRI) pWPXL	VH575-AB073 AB074-VH588	Myc tag, GFP
pWPXL- R141A CLEC14A	Gibson assembly of 2 PCR products into	VH575-AB075 AB076-VH588	Myc tag, GFP

	linearised (PmeI and EcoRI) pWPXL		
pIgPP- R98A CLEC14A	Gibson assembly of 1 PCR product into linearised (EcoRI and NotI) pIgPP	AB139-AB140 AB141-AB142	Fc tag, protease cleavage site
pIgPP- R100A CLEC14A	Gibson assembly of 1 PCR product into linearised (EcoRI and NotI) pIgPP	AB139-AB143 AB144-AB142	Fc tag, protease cleavage site
PIgPP-R98,99,100A CLEC14A	Gibson assembly of 1 PCR product into linearised (EcoRI and NotI) pIgPP	AB139-AB145 AB146-AB142	Fc tag, protease cleavage site
pIgPP- S137, T139,R141A CLEC14A	Gibson assembly of 1 PCR product into linearised (EcoRI and NotI) pIgPP	AB139-AB142	Fc tag, protease cleavage site
pIgPP- S137,T139A CLEC14A	Gibson assembly of 1 PCR product into linearised (EcoRI and NotI) pIgPP	AB139-AB142	Fc tag, protease cleavage site
pIgPP- S137,R141A CLEC14A	Gibson assembly of 1 PCR product into linearised (EcoRI and NotI) pIgPP	AB139-AB142	Fc tag, protease cleavage site
pIgPP- T137,R141A CLEC14A	Gibson assembly of 1 PCR product into linearised (EcoRI and NotI) pIgPP	AB139-AB142	Fc tag, protease cleavage site
pIgPP- S137A CLEC14A	Gibson assembly of 1 PCR product into linearised (EcoRI and NotI) pIgPP	AB139-AB142	Fc tag, protease cleavage site
pIgPP- T139A CLEC14A	Gibson assembly of 1 PCR product into linearised (EcoRI and NotI) pIgPP	AB139-AB142	Fc tag, protease cleavage site
pIgPP- R141A CLEC14A	Gibson assembly of 1 PCR product into linearised (EcoRI and NotI) pIgPP	AB139-AB142	Fc tag, protease cleavage site
pAvitag Nluc MMRN2 495-674 C-term linker	Gibson assembly of amplified 1 PCR product linearised (PmeI and EcoRV) pAvitag3	AB133-AB134	Nluc, His tag

pAvitag SNAP MMRN2 495-674 C-term linker	Gibson assembly of amplified 1 PCR product linearised (PmeI and EcoRV) pAvitag3	AB134-AB136	SNAP, His tag
pAvitag LgBiT MMRN2 495-674 C-term linker	Gibson assembly of amplified 1 PCR product linearised (PmeI and EcoRV) pAvitag3	AB137-AB138	LgBiT, His tag
pAvitag Nluc MMRN2 495-674 C-term	Gibson assembly of 1 PCR product into linearised (EcoRV) pAvitag3	AB107-AB108	Nluc, His tag
pAvitag SNAP MMRN2 495-674 C-term	Gibson assembly of 1 PCR product into linearised (EcoRV) pAvitag3	AB109- AB110	SNAP, His tag
pHLAvitag3- E617A MMRN2 495-674	Gibson assembly of 2 PCR products into linearised (AgeI and KpnI) pHLAvitag3 constructed by Sowmiya Navamanirajah	AB117-AB099 AB100-AB118	His tag AgeI and KpnI
pHLAvitag3- E620A MMRN2 495-674	Gibson assembly of 2 PCR products into linearised (AgeI and KpnI) pHLAvitag3 constructed by Sowmiya Navamanirajah	AB117-AB101 AB102-AB118	His tag
pHLAvitag3- S623A MMRN2 495-674	Gibson assembly of 2 PCR products into linearised (AgeI and KpnI) pHLAvitag3 constructed by Sowmiya Navamanirajah	AB117-AB103 AB104-AB118	His tag
pHLAvitag3- E624A MMRN2 495-674	Gibson assembly of 2 PCR products into linearised (AgeI and KpnI) pHLAvitag3 constructed by Sowmiya Navamanirajah	AB117-AB105 AB106-AB118	His tag

Table 2.7. List of plasmids constructed through conventional cloning methods. Displaying plasmid name, methods of construction, primers used and tags or notable features within the construct.

Plasmid	Cloning method	Primers used	Tags/features
pcDNA3.1- Nluc MMRN2 495-674 N-term	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB065-AB067	Nluc
pcDNA3.1- SNAP MMRN2 495-674 N-term	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB065-AB067	SNAP
pcDNA3.1- LgBiT MMRN2 495-674 N-term	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB065-AB067	LgBiT
pcDNA3.1- SmBiT MMRN2 495-674 N-term	BamHI and XbaI digested PCR product ligated into complimentary cut pcDNA3.1	AB065-AB067	SmBiT
pcDNA3.1- Nluc WT CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	Nluc
pcDNA3.1- Nluc S137T139R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	Nluc
pcDNA3.1- Nluc S137T139A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	Nluc
pcDNA3.1- Nluc S139R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	Nluc
pcDNA3.1- Nluc T139R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	Nluc
pcDNA3.1- Nluc S137A CLEC14A	BamHI and XhoI digested PCR product ligated into	AB013-AB014	Nluc

	complimentary cut pcDNA3.1		
pcDNA3.1- Nluc T139A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	Nluc
pcDNA3.1- Nluc R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	Nluc
pcDNA3.1- Nluc R161A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	Nluc
pcDNA3.1- SNAP WT CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	SNAP
pcDNA3.1- SNAP S137T139R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	SNAP
pcDNA3.1- SNAP S137T139A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	SNAP
pcDNA3.1- SNAP S139R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	SNAP
pcDNA3.1- SNAP T139R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	SNAP
pcDNA3.1- SNAP S137A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	SNAP
pcDNA3.1- SNAP T139A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	SNAP
pcDNA3.1- SNAP R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	SNAP

pcDNA3.1- SNAP R161A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	SNAP
pcDNA3.1- LgBiT WT CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	LgBiT
pcDNA3.1- LgBiT S137T139R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	LgBiT
pcDNA3.1- LgBiT S137T139A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	LgBiT
pcDNA3.1- LgBiT S139R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	LgBiT
pcDNA3.1- LgBiT T139R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	LgBiT
pcDNA3.1- LgBiT S137A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	LgBiT
pcDNA3.1- LgBiT T139A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	LgBiT
pcDNA3.1- LgBiT R141A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	LgBiT
pcDNA3.1- LgBiT R161A CLEC14A	BamHI and XhoI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB014	LgBiT
pcDNA3.1- SmBiT WT CLEC14A	BamHI and XbaI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB015	SmBiT
pcDNA3.1- SmBiT S137T139R141A CLEC14A	BamHI and XbaI digested PCR product ligated into	AB013-AB015	SmBiT

	complimentary cut pcDNA3.1		
pcDNA3.1- SmBiT S137T139A CLEC14A	BamHI and XbaI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB015	SmBiT
pcDNA3.1- SmBiT S139R141A CLEC14A	BamHI and XbaI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB015	SmBiT
pcDNA3.1- SmBiT T139R141A CLEC14A	BamHI and XbaI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB015	SmBiT
pcDNA3.1- SmBiT S137A CLEC14A	BamHI and XbaI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB015	SmBiT
pcDNA3.1- SmBiT T139A CLEC14A	BamHI and XbaI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB015	SmBiT
pcDNA3.1- SmBiT R141A CLEC14A	BamHI and XbaI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB015	SmBiT
pcDNA3.1- SmBiT R161A CLEC14A	BamHI and XbaI digested PCR product ligated into complimentary cut pcDNA3.1	AB013-AB015	SmBiT

Table 2.8. List of plasmids provided through the project. Displaying plasmid name, source and notable features within the construct.

Plasmid	**Source**	**Tag/features**
pWPXL- WT CLEC14A	Constructed by Victoria Heath	Myc tag, GFP
pWPXL- THBD switch	Constructed by Victoria Heath	Myc tag, GFP
pWPXL- R161A CLEC14A	Constructed by Victoria Heath	Myc tag, GFP
pWPXL- no ICD CLEC14A	Constructed by Victoria Heath	Myc tag, GFP
pIgPP- WT CLEC14A	Constructed by Victoria Heath	Fc tag, protease cleavage site
pIgPP- R161A CLEC14A	Constructed by Victoria Heath	Fc tag, protease cleavage site
pWPI- FL MMRN2 Avitag	Constructed by Kabir Khan	

pHLAvitag3- MMRN2cc	Constructed by Kabir Khan	His tag
pHLAvitag3- MMRN2cc first half (133-487)	Constructed by Kabir Khan	His tag
pHLAvitag3- MMRN2cc second half (487-820)	Constructed by Kabir Khan	His tag
pHLAvitag3- MMRN2 487-674	Constructed by Kabir Khan	His tag
pHLAvitag3- MMRN2 675-820	Constructed by Kabir Khan	His tag
pHLAvitag3- MMRN2 495-674	Constructed by Kabir Khan	His tag

2.4. Protein Biochemistry

2.4.1. Cell lysis

HEK293T cells grown for biochemistry-based experiments were harvested by scraping the cells into a small volume of PBS. The cells were pelleted at 1844g for 2 mins. The pellet was then resuspended in 5 pellet volumes of NP40 lysis buffer. Pellet size varied based on the dish used, in the case of a 6 cm dish 300µL of NP40 lysis buffer was used. The mixture was incubated on ice for 15 min. Insoluble material was pelleted at 16,602g for 15 mins at 4°C. Supernatants were collected in clean centrifuge tubes, equal volumes of 2X sample buffer was added and samples were stored at -20°C.

2.4.2. SDS-PAGE

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was used as a method to separate protein samples. SDS polyacrylamide resolving gels were made up of H_2O, 30% (v/v) acrylamide mix, 1.5 M Tris pH 8.8, 10% (w/v) SDS (fisher scientific), 10% (w/v) ammonium persulfate and TEMED in disposable cassettes (Life Technologies). Volumes of H_2O, acrylamide and TEMED were scaled accordingly for varying percentage acrylamide gels. Stacking gels were made up of H_2O, 5% (v/v) acrylamide, 1M Tris pH 6.8, 10% (w/v) SDS (fisher scientific), 10% (w/v) ammonium persulfate and TEMED. Protein samples were run at 100V

through the stacker and 130V for the resolving portion of the gel. Gels were either stained for protein using a Coomassie based stain (Sigma-Aldrich) or transferred onto PVDF membranes in transfer buffer for 2 hrs at 30V. Membranes were then blocked in 5% (w/v) milk in TBST (150 mM NaCl, 20 mM Tris pH 7.5, 0.1% (v/v) Tween) for 1 hr at room temp.

2.4.3. Western blotting

PVDF membranes with the proteins transferred and blocked were then incubated overnight with appropriately diluted primary antibodies (Table 2.5) in either in 5% (w/v) milk or 5% (w/v) BSA in TBST at 4°C. The membrane was then washed in TBST for 30 mins, replacing the TBST every 6 mins, to remove non-specific binding of antibodies before being incubated with appropriately diluted secondary antibodies conjugated with HRP in 5% (w/v) milk for 2 hrs at room temperature. The membrane was then washed in TBST to remove non-specific binding of secondary antibodies. The presence of HRP conjugated antibodies was detected using enhanced chemi-luminescence (ECL). Light and dark ECL was mixed in equal volumes and immediately placed on the membrane. The membrane was then inserted into the Odyssey (LiCore) using the chemi and 700nm channels to visualise the proteins.

2.4.4. Far western blotting

Proteins being studied using far western blotting were prepared in non-reducing conditions and separated by SDS-PAGE then transferred and blocked as described above. The membranes were incubated overnight at 4°C with the protein of interest, either CLEC14A-Fc in conditioned media or from lysates of HEK293T cells transfected with FL MMRN2 (diluted 1:50 in 5% (w/v) milk). The membranes were then washed in TBST and incubated with either the Anti-Fc-HRP antibody for the CLEC14A-Fc conditioned media blots or the MMRN2 primary antibody for the FL MMRN2 lysate treated blots. Membranes were incubated for 3 hrs at

room temp. Non-specific binding antibodies were washed off using TBST for 30 mins wherein fresh TBST was used every 6 mins, the membrane was incubated with the appropriately diluted secondary antibody conjugated to HRP if required. The membrane was washed in TBST for 30 mins wherein TBST was replaced every 6 mins, and the presence of HRP-conjugated antibodies were detected using enhanced chemi-luminescence (ECL) (Figure 2.2).

2.4.5. Statistical analysis

Band intensities from blots were quantified using the LI-COR and analysed in the Image Studio Lite version 5.2 software (LI-COR). Band intensity values were tested for normality using the Shapiro-Wilks test. The values were analysed using an ANOVA test for significance. This was then followed by the Dunnett's test where experimental groups were compared to the WT control group. Statistical analyses were performed in GraphPad Prism version 9.

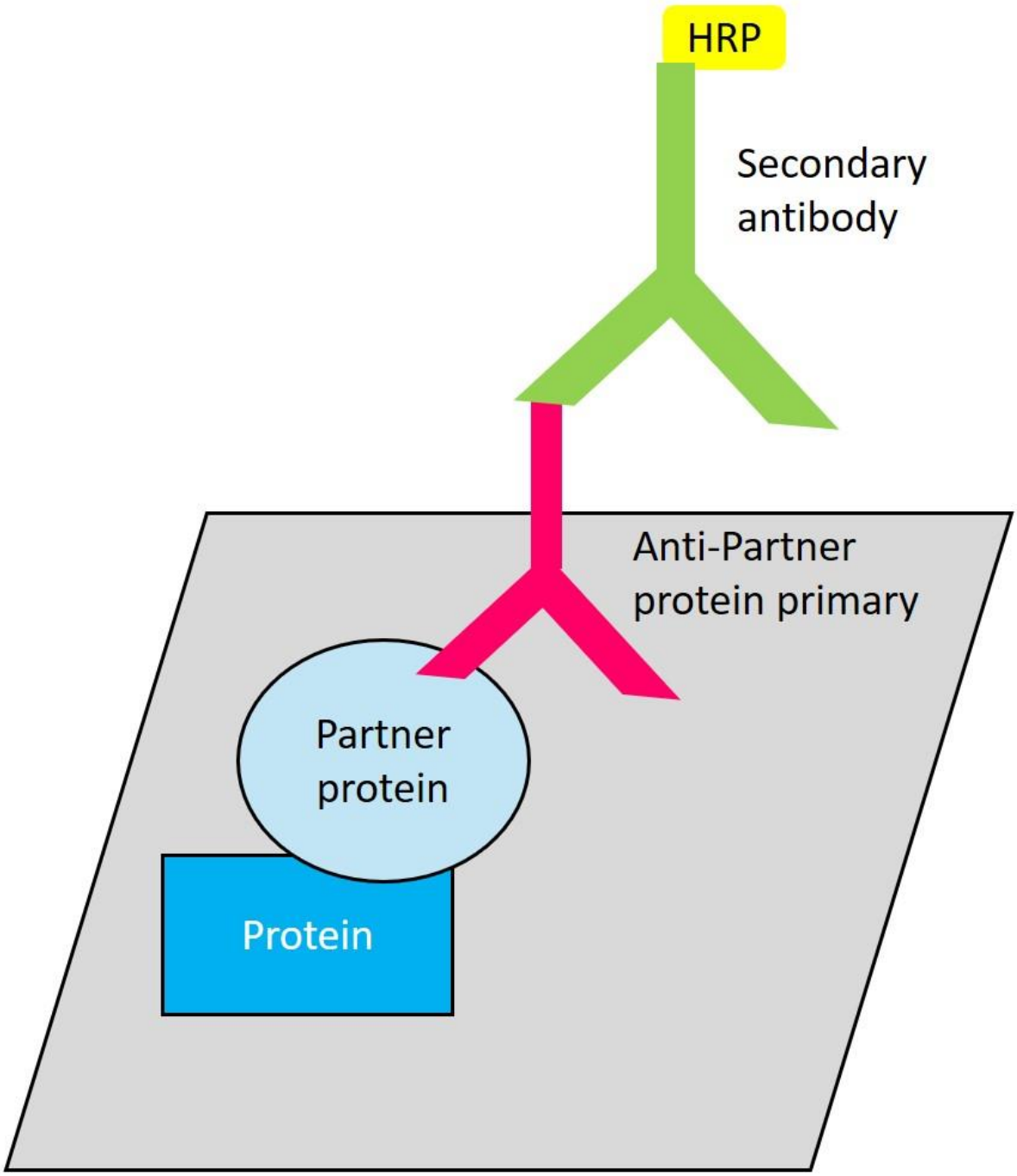


Figure 2.2. Schematic illustration of a Far western blot. A far western blot involves subjecting a protein sample in non-reducing conditions to SDS-PAGE. The proteins in the gel are then transferred onto a PVDF membrane which is incubated with cellular lysates containing the partner protein on interest. The membrane is then treated with an antibody to target the partner protein of interest, followed by the appropriate secondary antibody conjugated to HRP.

2.4.6. Enzyme-link immunosorbent assay (ELISA)

Protein interactions were studied using ELISAs, this involved coating wells of a 96-well plate (Bio-Techne DY990) with 100 ng of the protein of interest in a 50 µL volume of PBS overnight at 4°C. The wells were then washed three times with 100 µL PBS. The wells were then blocked using 100 µL of PBS containing 3% (w/v) BSA for 90 mins at room temperature. The blocking buffer was removed, and wells were washed three times with 100 µL of PBS, the wells were then treated with a binding protein in the form of 50% diluted conditioned media and/or a peptide of interest (Cambridge Research Biochemicals) in a 100 µL volume and incubated for 90 mins at room temperature. The sample was removed, and wells were washed three times with 100 µL PBS, the wells were then incubated with the appropriate HRP-conjugated antibody for 90 mins at room temperature. The antibody was removed, and wells were washed three times with 100 µL PBS. The wells were treated with BM Blue POD substrate (Sigma-Aldrich) and incubated in the dark for 3-5 mins before being neutralized with 1 mM HCl. The plate was read at 450nm using the plate reader (Molecular devices, VersaMax) on the SoftMax Pro software.

2.4.7. Flow cytometry

Transfected HEK293T cells expressing the protein of interest were detached from the plate using 2X Trypsin, the volume of trypsin used was scaled depending on the dish used in the case of a 6 well dish each well required 1 mL of 2X Trypsin. Once detached, the cells were transferred into a tube and pelleted at 190g 3 mins. The cells were then kept on ice for the full sample preparation procedure. Cells were then resuspended in either 100 µL of conditioned media or the primary antibody of interest at 10 µg/mL diluted in FACs buffer and incubated for 30 mins on ice. After the incubation, the cells were washed twice in 1 mL FACS buffer (PBS with 0.2% (w/v) BSA (VWR Life Science) and 0.02% (v/v) sodium azide). The cells

were then resuspended in 100 µL of the appropriate primary or secondary antibody conjugated to the relevant fluorophore appropriately diluted in FACs buffer (Table 2.2) and incubated for 30 mins on ice. 700 µL of cDMEM was added to each tube and the sample was passed through a 70 µm filter. The cells were studied using the CytoFLEX S (Beckman Coulter) and the data was analysed using the FlowJo software (BD Biosciences). Mean florescent intensities were documented and normalised by dividing the value with the isotype control reading and statistical analyses were performed as described in section 2.4.5.

2.5. Protein expression and purification

2.5.1. Producing and purifying Fc-tagged protein

HEK293T cells were transfected with the CLEC14A pIgPP plasmid containing the protein of interest fused to a human IgG1 Fc tag. The cells produced and secreted the protein into the OptiMEM media in which they grew. The media was collected every 2 days for 2 weeks and flowed through PBS-washed Protein A beads twice at a flow rate of 1 mL/min (Figure 2.3B). In in the case of a 1 L prep 500 µL of packed Protein A beads were used which were washed 3X in 3 mL PBS. Post media flow-through the beads were collected and washed 3X in 3 mL of precision protease buffer (25mM Tris pH7.5, 150mM NaCl, 5mM $MgCl_2$, 0.1% (v/v) Triton x100, 1mM EDTA, 1mM DTT). After the final wash the beads were resuspended in 500 µL of precision protease buffer and treated with GST-tagged PreScission Protease (Sigma-Aldrich), which was scaled based on the prep size, in the case of a 1 L prep 4 µL of protease was used (Figure 2.3C). This was incubated at 4°C overnight by rotation. The next day the beads were spun down at 664g for 1 min at 4°C. To remove the GST-tagged PreScission Protease the supernatant was collected and incubated with 100 µL of Glutathione-agarose beads washed 3X in 1 mL of precision protease buffer without DTT, after the final wash the beads were

resuspended in 500 µL of the precision protease buffer without DTT and the supernatant was added and incubated at 4°C for 2 hrs by rotation (Figure 2.3E). The beads were spun down at 664g for 3 mins at 4°C and the incubation was repeated with another batch of Glutathione beads washed in precision protease buffer without DTT 3X in 1 mL. The beads were spun down at 664g for 3 mins at 4°C and the supernatant containing the CLEC14A was collected. The precision protease buffer was removed using a membrane dialysis tube (GeBa), the supernatant was placed in the tube and the tube was submerged and stirred in a 10x volume of PBS overnight at 4°C. The next day the protein was collected and treated with 0.02% (v/v) Sodium Azide and stored at 4°C. Protein concentrations were measured using the NanoDrop (Geneflow ND-1000) and protein integrity was studied using SDS-PAGE.

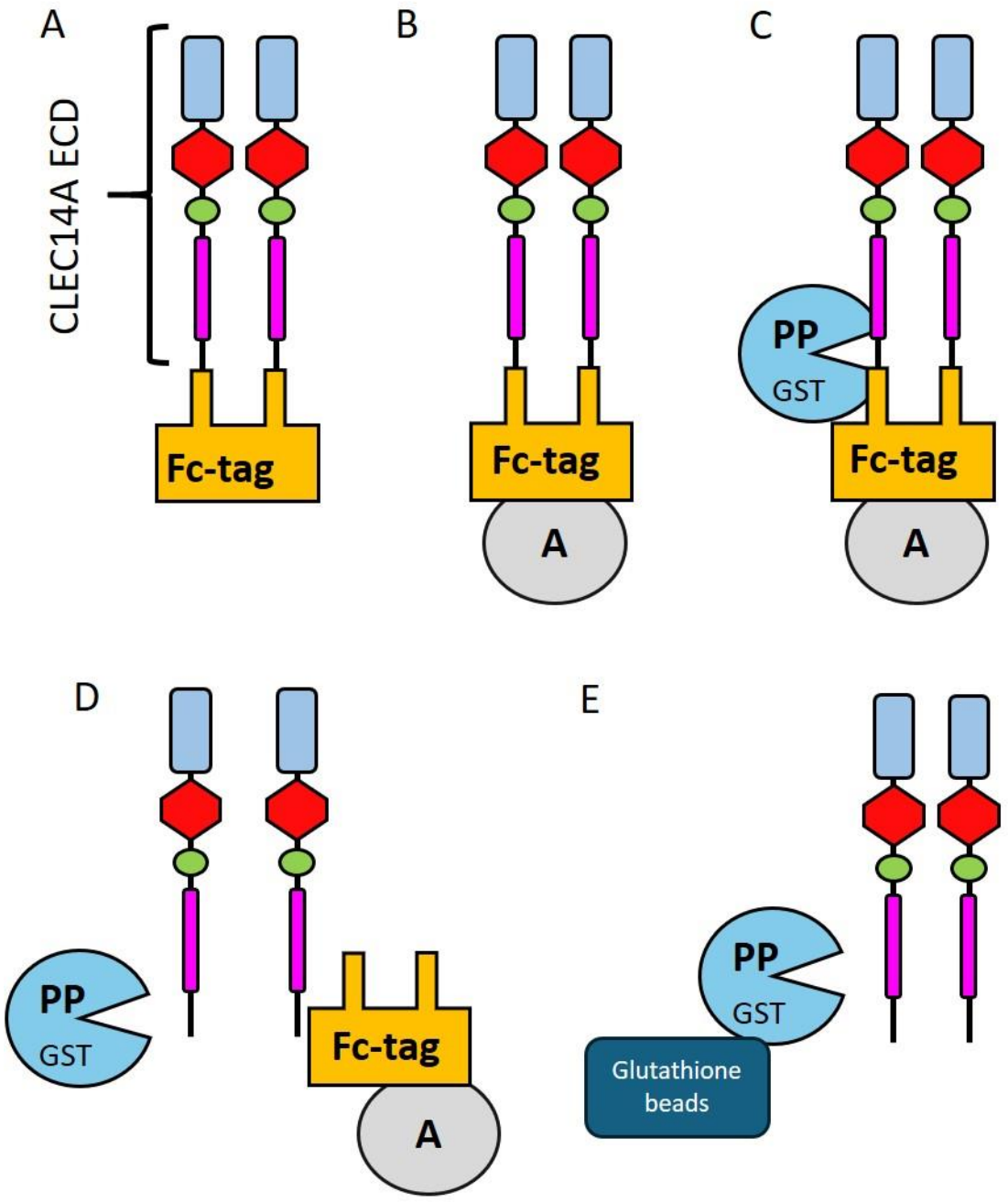


Figure 2.3. Schematic illustration of Fc-tagged purification using the precision protease cleavage system. (A) The extracellular domain (ECD) of CLEC14A was expressed as an Fc-tag fusion using HEK293T cells. (B) The protein was separated from the conditioned media using protein A beads. (C) The Fc-tag was cleaved off using precision protease (PP) which contained a GST tag. (D) The Fc-tag conjugated to the protein A bead was removed. (E) Glutathione beads were used to remove the GST-tagged PP and this was performed twice.

2.5.2. Eluting Fc-tagged protein

HEK293T cells were transfected with the pIgPP plasmid containing the protein of interest fused to a human IgG1 Fc tag. Cells produced and secreted the Fc-tagged protein of interest into the OptiMEM media which was collected every 2 days for 2 weeks and an equal volume of PBS was added to the collected media. The media was then flowed through PBS-washed Protein A beads twice at a flow rate of 1mL/min, in the case of a 1 L prep 500 µL of packed beads were used which were washed 3X in 3mL PBS. Post flow through the beads were washed in PBS and the protein was eluted by incubating the beads for 20 mins at room temp in 1 mL of PBS containing 10mM Tris at pH 2.3. The eluates were immediately neutralised in 1M Tris pH 9.0. As eluates were collected and neutralised, protein levels were detected using a colorimetric assay. 5 µL of the eluate was taken and added to a well of a 96 well plate, the colorimetric assay dye (BioRad) was diluted 1:5 with dH_2O and 100 µL was added to the well with the protein. The colour change was then visibly monitored. Once no colour change was detected eluates were no longer collected. The protein was then buffer exchanged into PBS, eluates were pooled and placed into a membrane dialysis tube (GeBa), the tube was submerged into 10X volume of PBS overnight at 4°C. The next day the protein was collected and treated with 0.02% (v/v) Sodium Azide and stored at 4°C.

2.5.3. Producing and purifying His-tagged protein

HEK293T cells were transfected with pHLAvitag3 plasmids containing the gene encoding the protein of interest conjugated to a short sequence of Histidine amino acids. Cells produced and secreted the protein of interest into the OptiMEM media. The media was collected every 2 days for 2 weeks and an equal volume of PBS was added. The media was flowed through a column containing washed Nickel beads at a flow rate of 1 mL/min. For a 1 L prep, 1 mL of packed Nickel beads were washed 3X in 3 mL of PBS containing 10mM imidazole (VWR Life

Science). Once the media had flowed through the beads twice, the beads were washed with one column volume of PBS containing 10 mM Imidazole. The protein was then eluted off in 1 mL fractions using a buffer containing 50 mM Na_2HPO_4 pH7.4, 400 mM NaCl, 10% (v/v) Glycerol (Millipore), 500 mM imidazole and 0.5 mM TCEP. The beads were incubated in 1 mL of the buffer for 20 mins at room temperature and fractions were eluted. As fractions were being collected protein levels were being detected using a colorimetric assay as described in section 2.5.2. The eluates were collected, and the protein was buffer exchanged into PBS using a membrane dialysis tube (GeBa), the tube was submerged into 10X volume of PBS overnight at 4°C. The next day the protein was collected and treated with 0.02% (v/v) Sodium Azide and stored at 4°C.

2.6. Gel filtration

Gel filtration data was collected by Dr Alex Slater at the University of Birmingham. Some purified proteins were subjected to size exclusion chromatography gel filtration using the AKTA pure S200 column. Approximately 500 mL of PBS buffer was degassed using a vacuum and the column was equilibrated using two column volumes of the degassed PBS buffer. 500 µL of the purified protein was then injected into the AKTA and a flow rate of 0.3 mL/min was used. 0.5 mL Fractions were taken and a gel filtration trace was produced. Peaks from the trace were compared to a standard curve which was generated using the gel filtration markers kit for protein molecular weights (Sigma-Aldrich). Collected fractions were subjected to western blotting.

2.7. Fluorescently labelling SNAP-tagged constructs

The SNAP-Surface Alexfluor488 substrate was dissolved in 50 µL of DMSO to make a stock solution of 1 mM SNAP-tag substrate. The stock solution was stored at -20°C and diluted in

PBS to a working concentration of 10 µM per 5 µM of SNAP-tagged protein when required. SNAP-tagged proteins were either expressed as purified soluble proteins or on the cell surface of transfected HEK293T cells.

2.7.1. Labelling SNAP-tagged purified protein

Purified SNAP-tagged proteins were labelled as using a 2X molar excess of the SNAP substrate. For 5 µM of purified SNAP-tagged protein, 10 µM of the SNAP-surface Alexafluor 488 substrate was used and 1 mM of DTT. This was scaled up accordingly. This mixture was incubated for 1 hr at room temperature in the dark and excess free SNAP substrate was removed using a dialysis membrane (GeBa).

2.7.2. Labelling cell surface SNAP-tagged protein

In the case of cell surface SNAP-tag labelling, the media was removed from the dish and was replaced with HBSS media containing 0.25 µM of the SNAP-surface Alex Fluor 488 substrate. Volumes varied based on the size of the dish, 6 well plates required 3 mL per well. This was incubated for 1 hr at 37°C. The free unlabelled dye was removed from the dish and the cells were washed twice in HBSS buffer.

2.8. NanoBRET

HEK293T cells were seeded onto 6cm dishes at a density of 700,000 cells in 3 mL of cDMEM. After 24 hrs the cells were transfected with 3 µg of the plasmid of interest and left for 24 hrs. The cells were then removed using 2X Trypsin and re-seeded onto a solid white-bottom 96-well plate pre-coated with Poly-L-lysine at a density of 100,000 cells per well in 100 µL of cDMEM. After 24 hrs the cells were fluorescently labelled if required (as described in section 2.8.2). The cDMEM was removed and replaced with 45 µL HBSS buffer containing the partner protein of interest at concentrations increasing from 0 nM to 1000 nM. This was incubated

for 1 hr at 37°C. A fluorescence reading was then taken using the PHERAstar FSX (BMG-LABTECH) using the fluorescence intensity module and the gain was adjusted as required. The cells were then treated with 5 µL of furimazine diluted 1:500 in HBSS buffer for a final volume of 50 µL. This was incubated for 5 mins at room temperature before the fluorescence and luminescence reading was taken using both the BRET1 and BRET2 module, gain adjusted as required. Data was analysed in excel where BRET ratios were calculated by dividing the acceptor (fluorescence) value by the donor (luminescence) value.

2.9. Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) was performed by Dr Gemma Harris at the Research Complex at Harwell. Purified proteins in PBS buffer were subjected to ultracentrifugation at 50,000rpm using either the Beckman XLI or Beckman Optima analytical ultracentrifuge with an An-50Ti rotor (Beckman Coulter). Protein samples in PBS were placed into sample wells with standard two-sector open centrepieces and sapphire windows. Data were recorded at 20°C using the absorbance (measured at 280 nm) and interference optical detection systems. The density and viscosity of the buffer was measured experimentally using a DMA 5000M densitometer (Anton Paar) equipped with a Lovis 2000ME viscometer module. The partial specific volume of the proteins was calculated using SEDFIT from the amino acid sequences and data analysis was performed using SEDFIT, fitting to the c(s) model. Figures were made using GUSSI (UT southwestern).

2.10. Intact Mass spectrometry

Mass spectrometry was performed by Dr Jinglei Yu at the University of Birmingham. Purified proteins were prepared in a buffer containing 5% (v/v) methanol (VWR) and 1% (v/v) formic acid (Fisher scientific) in 50 µL of liquid chromatography mass spec grade water (Fisher

scientific) for a final concentration of 1-10mg/mL. Samples were subjected to mass spectrometry using the Orbitrap Q-Exactive HF mass spectrometer (Thermo Fisher Scientific, Bremen, Germany).

CHAPTER 3: Mapping the interaction of CLEC14A with MMRN2

CHAPTER 3: Mapping the interaction of CLEC14A with MMRN2

3.1. Introduction

CLEC14A is single pass transmembrane spanning glycoprotein expressed at high levels of the tumour endothelium with little to no expression on normal blood vessels (Robinson *et al.,* 2020) . It is part of a wider family called the C-type lectin domain group 14 family alongside CD93, Thrombomodulin (THBD) and CD248. Each member of this family contains similar domains including an N-terminal C-type lectin domain (CTLD), a sushi domain, at least one EGFR-like domain, a glycosylated mucin-like domain followed by the transmembrane domain and a short intracellular tail. MMRN2 is a common binding partner of CLEC14A, CD93 and CD248 but not THBD and interacts with the N-terminal CTLD of each protein (Khan *et al.,* 2017). CLEC14A plays a role in angiogenesis by promoting tube formation and cell migration as shown in a number of *in vitro* assays (Mura *et al.,* 2012). The functional significance of the interaction between CLEC14A and MMRN2 is not yet well understood however blocking the interaction *in vivo* using blocking antibodies and peptides, results in a marked reduction in tumour size (Noy *et al.,* 2015; Khan *et al.,* 2017).

It was hypothesized that the CLEC14A-MMRN2 interaction promotes tumour angiogenesis. The aim of this part of the project was to identify the key amino acids on CLEC14A and MMRN2 involved in the interaction. The next step was to test the blocking capacity of peptides which mimic the interacting region of CLEC14A. This would provide valuable insight into the mechanism of action of the interaction.

3.1.1. Mutagenesis strategy

CLEC14A and other CTLD-containing proteins contain a loop in loop structure as indicated in the AlphaFold predicted model highlighted in blue (Figure 3.1A). This loop in loop structure

involves a single flexible strand protruding from the core of the CTLD and forming a loop before folding into a short beta sheet. From the beta sheet another flexible loop folds into a second loop which folds around the first loop before standing adjacent to the first loop and terminating into a beta sheet. Previously it has been shown that by replacing the sequence of amino acids from the first loop of CLEC14A, between amino acid E97- E108, with the aligning sequence of THBD which does not bind to MMRN2, CLEC14A can no longer bind to MMRN2. This mutant was coined the THBD-switch mutant (Khan *et al.,* 2017)

Since CLEC14A and CD93 are both known to bind to MMRN2 within the same region and THBD does not, the rationale used was that amino acids conserved between CLEC14A and CD93 but not THBD could potentially play a role in the interaction and were targeted for mutagenesis. A site-directed mutagenesis approach often referred to as an alanine screen was employed. This is a common strategy employed to study protein-protein interaction. Alanine is a short, uncharged amino acids and was therefore used as a replacement for candidate residues targeted for mutagenesis. The mutants were constructed into a pWPXL vector including a Myc tag followed by the P2A self-cleaving peptide and the sequence for GFP. The Clustal Omega protein alignment tool was used to align the amino acid sequence from the loop in loop structures of CLEC14A, CD93 and THBD (Figure 3.1B). This revealed a small positively charged patch of amino acids at the N-terminal of the first loop of CLEC14A which came in the form of 3 consecutive arginine residues: R98, R99 and R100. The R98 residue aligned with an arginine residue on CD93 whilst the aligned amino acid on THBD was a hydrophobic leucine. The R99 residues of CLEC14A aligned with a negative glutamic acid residue on CD93 and a hydrophobic proline on THBD and therefore did not fit the criteria. Lastly, the R100 residue of CLEC14A aligned with another positively charged residue on CD93; lysine but again a hydrophobic proline on THBD.

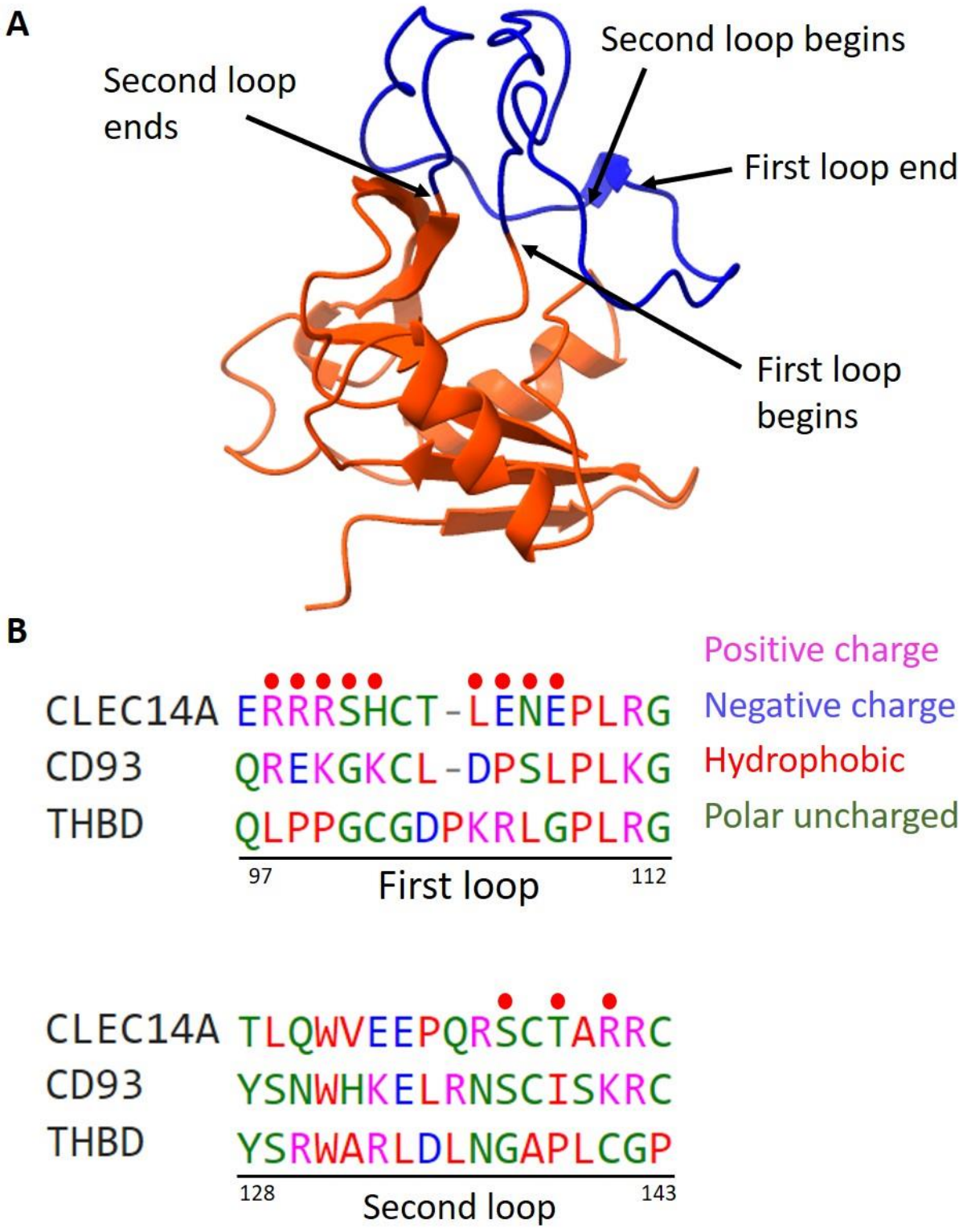


Figure 3.1. Protein sequence alignments of members of the C-type lectin domain group 14A family alongside the AlphaFold predicted structure of CLEC14A CTLD viewed on Chimera software. (A) The predicted tertiary structure of the CTLD of CLEC14A containing a 'loop in loop' structure highlighted in Blue (Accession number Q86T13). (B) PRofile ALIgNEment (PRALINE) generated alignments of CLEC14A, CD93 and Thrombomodulin (THBD) showing amino acids which form parts of the 'loop in loop'. Residues targeted for mutagenesis are pinpointed with a red dot.

3.2. Residues within the first loop of CLEC14A do not interact with MMRN2

Mutants of CLEC14A were created through PCR to generate fragments and ligated using Gibson assembly. To test whether the first and last arginine residues within this 3-amino acid stretch could interact with MMRN2, point mutations of R98A and R100A were generated. Another mutant where all three arginine residues were mutated to alanine was also generated in order to completely abolish the positive charge. This was generated in the instance that the CLEC14A-MMRN2 interaction was based on opposing charges.

It has been well documented that WT CLEC14A binds to MMRN2 in a far western setting, this involves CLEC14A being processed in non-reducing conditions to conserve the disulphide bridges which appear to be critical in its binding to MMRN2 (Khan *et al.,* 2017). The THBD switch has also been shown previously to not bind to MMRN2 and was also used to to ensure results were consistent. The R161A CLEC14A mutant has been demostrated in the literature to abrogate the interaction between CLEC14A and heparin (Sandoval et al., 2020). This mutant was also included in the study to see if it was involved in the interaction with MMRN2. . Cellular lysates from HEK293T cells transfected with the empty pWPXL vector was used as the negative control. Lysates containing full length MMRN2 were run alongside to confirm expression of the protein used to probe the membrane in the far western blot analysis. Far western blotting revealed WT, R161A and R100A CLEC14A each bound to FL MMRN2. Consistent with previous findings, the THBD switch did not bind to MMRN2 nor did the R98A or R98-99-100A triple mutant. FL MMRN2 was present in the lysate (Figure 3.2A). Band intensities were quantified and statistically analysed to confirm significantly reduced binding between the THBD switch, the R98A and R98-99-100A mutants to MMRN2 (Figure 3.2B).

Corresponding western blots probed with an anti-Myc antibody were also studied in order to quantify the expression of each mutant form of CLEC14A. The blot showed differential expression patterns of mutants compared to WT CLEC14A. WT CLEC14A revealed two bands, one at approximately 100kDa and another lower band are approximately 60kDa with similar band intensities. The R161A mutant followed the same pattern of two bands at similar intensities. In the case of the THBD switch and R98A mutant, the higher molecular weight band is barely visible whilst the lower bands are much more intense. The R100A and R98-99-100A mutants show both bands clearly, however in both instances the lower molecular weight band is slightly more intense than the higher. The negative controls did not show any banding (Figure 3.3A).

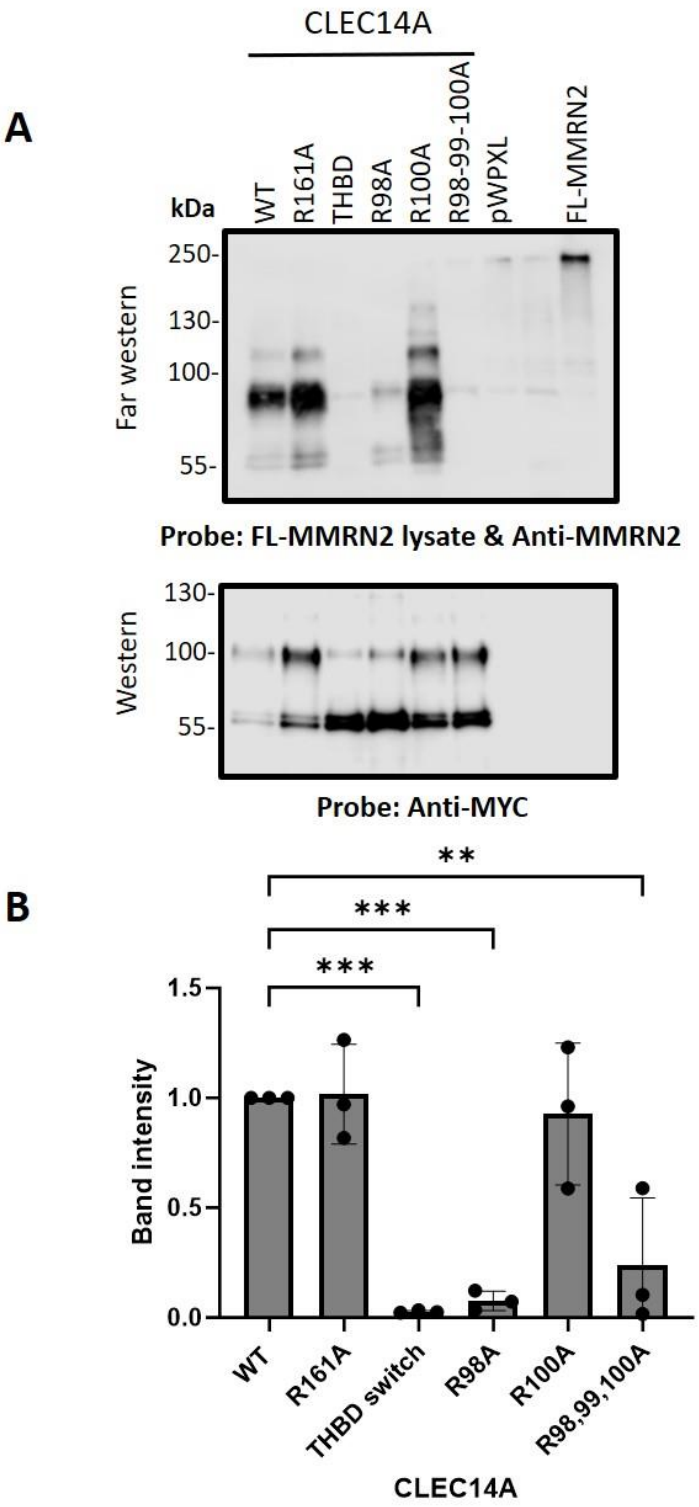


Figure 3.2. The R98A and R98,99,100A mutants blocked the CLEC14A-MMRN2 interaction. (A) HEK293T cells were transfected with plasmids to express either WT or mutant CLEC14A, lysates were collected and either run in reducing condition for western blot analysis or non-reducing conditions for far western blot analysis. Far western blots were probed with cellular lysates prepared from cells expressing full length MMRN2 followed by the anti-MMRN2 antibody. Western blots were probed with the Anti-myc antibody to determine expression levels of the various forms of CLEC14A. (B) Band intensities were quantified using the LI-COR and intensities of mutant bands were compared to WT CLEC14A bands. The results from the comparison are shown in the graph alongside error bars and plots for each experimental repeat. Statistical analysis of band intensities was performed using an ANOVA test followed by a Dunnett's test to compare significance to the WT control, mean ± S.D. (n=3). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

Given the alternative banding patterns of expression between WT CLEC14A it was suspected that some mutant proteins had altered glycosylation which suggest potential defects in folding and trafficking to the cell surface. Therefore, a flow cytometry assay was recruited to further study how the mutational profile of the protein effected cell surface expression. Monoclonal antibodies have been generated against the extracellular domains of CLEC14A and utilised in a range of experiments to reveal substantial information on CLEC14A. Five antibodies were generated by Cancer research technologies (CRT) these were CRT1 through to CRT5. It was revealed that the antibodies CRT1, CRT4 and CRT5 blocked the interaction between CLEC14A and MMRN2 whilst CRT2 and CRT3 did not (Noy *et al.,* 2015; Khan *et al.,* 2017). The CRT antibodies recognise CLEC14A in its native state and were therefore used in the flow cytometry assays. The CRT2 antibody was used as a control to confirm if the protein localised to the cell surface, this is because the CRT2 binding epitope is slightly further from the CTLD. As the CRT4 blocks the interaction between CLEC14A and MMRN2, it was of interest to identify any residues that may be part of the binding epitope of this antibody.

The gating strategy employed selected live, single cells which expressed GFP. Mean fluorescence intensities (MFI), were also documented and normalised. WT CLEC14A was used as a positive control as it has previously been shown to be recognised by both the CRT2 and CRT4 antibodies, a result which was consistent in this study. The MFIs of each mutant form of CLEC14A was compared with that of WT CLEC14A.

The R161A mutant was recognised by both the CRT2 and CRT4 antibodies with MFIs comparable to that of WT CLEC14A indicating it was expressed on the cell surface. The R100A mutant exhibited an MFI significantly higher than WT indicating enhanced cell surface expression and the MFI of the R98,99,100A mutant was similar to that of CRT4. Both

R98,99,100A and R100A both also exhibited significantly reduced MFI for CRT4 binding compared to that of WT. In the case of R100A the CRT4 signal was similar to that of the negative control staining indicating a loss of the CRT4 binding epitope. The THBD-switch and R98A mutants both exhibited significant reductions in MFI signal from both antibodies compared to WT indicating negative effects on protein expression and trafficking to the cell surface. The pWPXL negative control did not display any signal above the negative control (Figure 3.3).

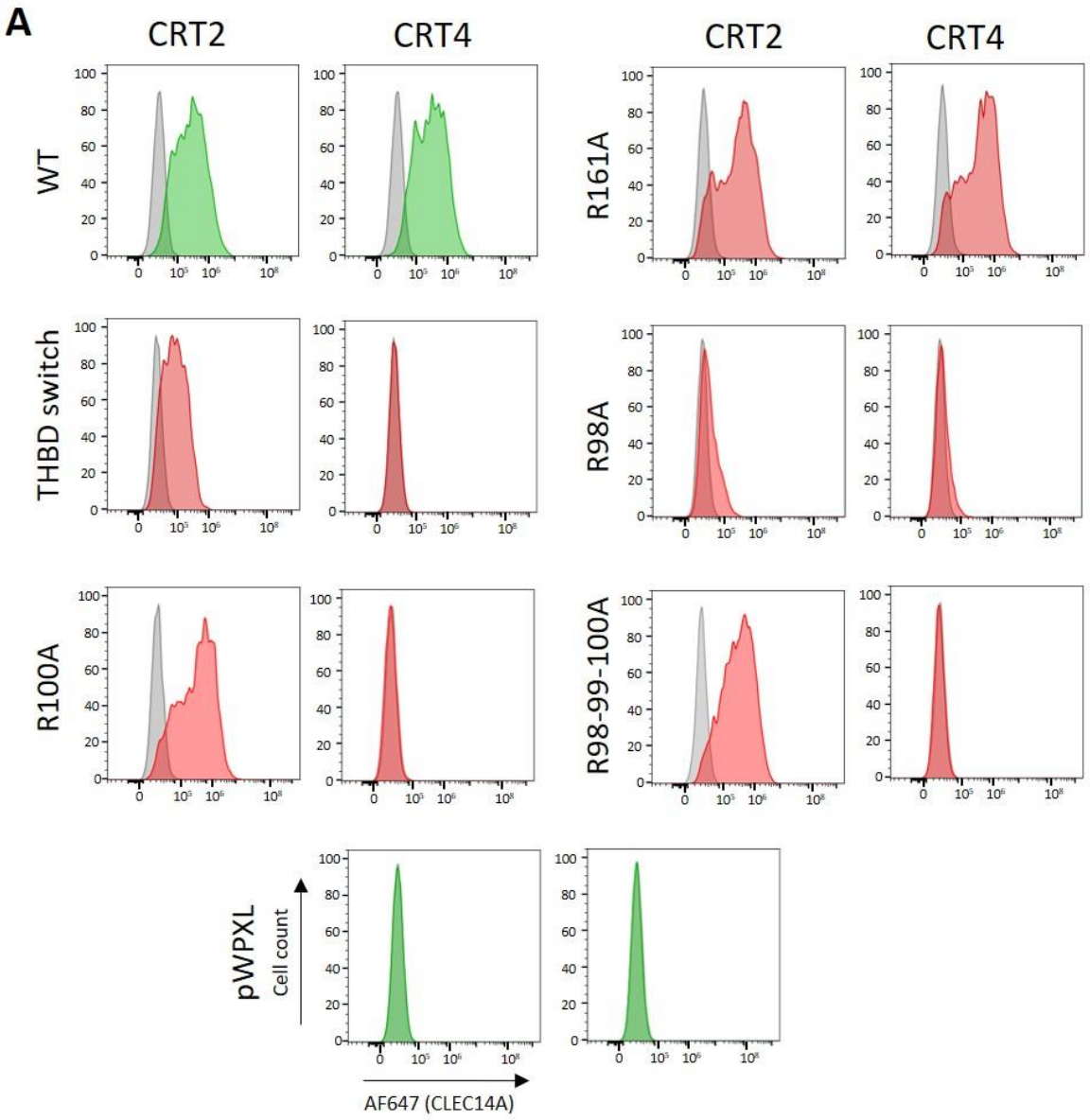


Figure Legend on next page

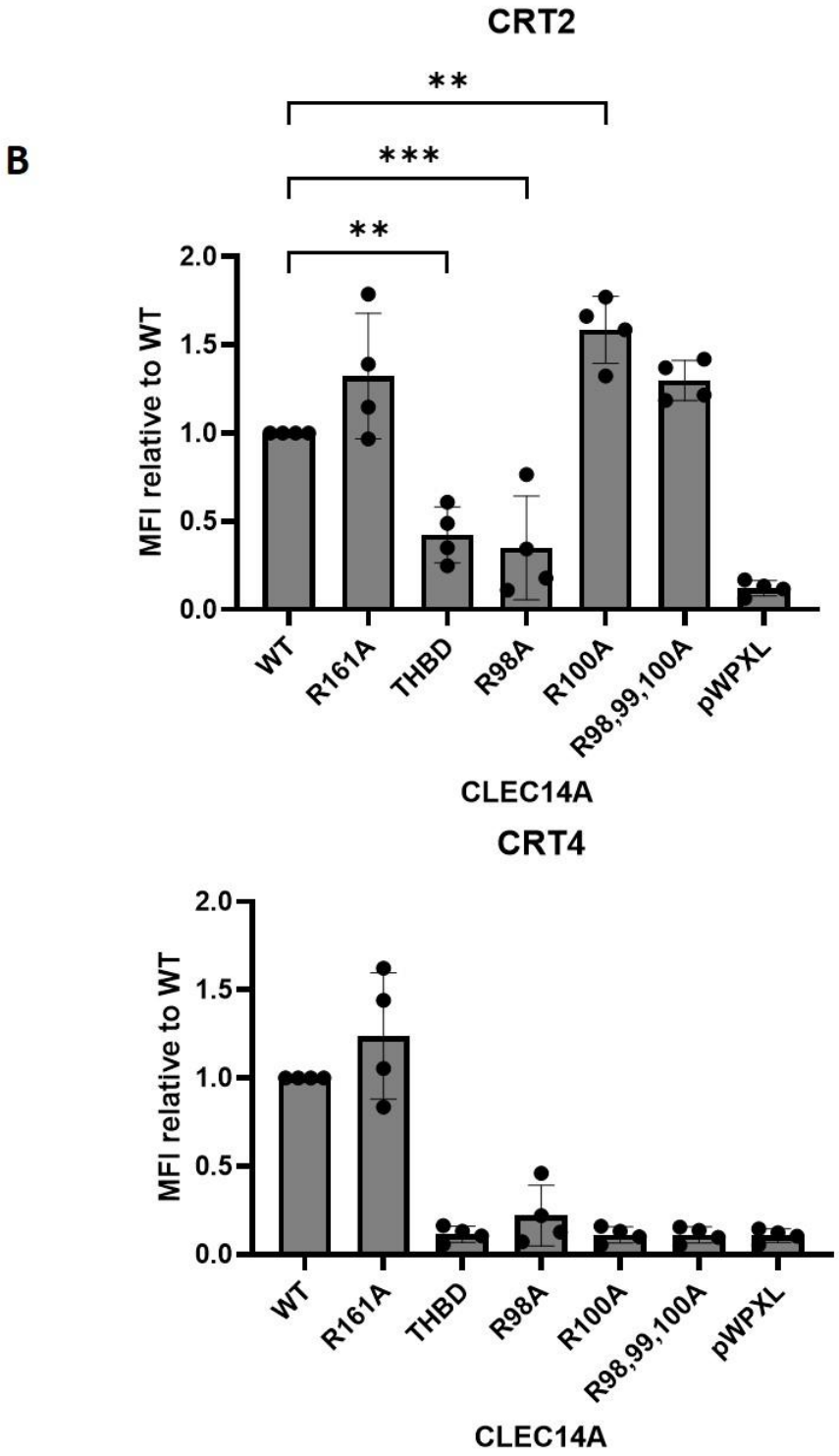


Figure 3.3. The R100 residue of CLEC14A forms part of the binding epitope for the CRT4 antibody. HEK293T cells were transfected with plasmids to express either WT or mutant CLEC14A or the pWPXL empty vector. Cells were incubated with CRT2 or CRT4 as primary antibodies followed by incubation with an anti-mouse IgG Alexafluor647 conjugated antibody. (A) Flow cytometry was performed, and the histograms presented show fluorescence in the APC channel gated on live cells expressing GFP. The histograms for the cells transfected with either pWPXL control or expressing WT CLEC14A are highlighted in green, whereas those of mutant forms of CLEC14A are in red. Signal from the control antibody is shown in grey. This figure is representative of data produced from three experimental replicates. (B) MFIs for each mutant and control were identified and analysed relative to WT CLEC14A for both the signal from CRT2 and CRT4 treatment. The result from a triplicate set of experiments were statistically analysed using an ANOVA test followed by a Dunnett's test to compare significance to the WT control, mean ± S.D. (n=3). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

Further computational predictive modelling revealed the orientation of the side chains within the CTLD of CLEC14A. The orientation of the positive arginine patch revealed that the R98 residue which, when mutated to alanine, was shown to not bind to MMRN2 and not express at the cell surface, has an interior-facing amino acid side chain. The R100 residue which was shown to express at the cell surface and interact with MMRN2 when mutated was solvent-facing. The middle R99 residue was planar facing which means the side chain was along the surface of the domain. The R98 interior-facing residue was potentially significant in the folding and structural integrity of the protein. This could explain the lack of cell surface expression and MMRN2 binding.

Given the probability that interiors facing proteins are likely involved in protein structure and stability and would not be expressed as the cell surface, the next iteration of mutagenesis strategically targeted only solvent-facing residues within the first loop. The AlphaFold predicted model indicated that the other solvent-facing residues included S101, H102, L105, N107, E106 and E108 each of which were targeted for mutagenesis. Stable protein-protein interactions often involve more than one amino acid, therefore both single and double mutants were generated. Furthermore, protein-protein interactions may be charge-based therefore the negatively charge glutamic acid residues were also mutated to a positively charged lysine both individually and in combination. MMRN2 binding capacity for both single and double mutants were studied using far western blotting and a corresponding western blot was run to study levels of protein in the lysate. WT CLEC14A was used as a positive control for MMRN2 binding and the pWPXL alone was used as the negative control. The results indicate that WT CLEC14A and each mutant within the first loop of CLEC14A bound to MMRN2 (Figure 3.4). The corresponding western blots revealed two bands, a higher molecular weight band at approximately 100kDa and a lower molecular weight band at approximately 55kDa.

The higher molecular weight band was more intense than the lower band which was consistent for each mutant (Figure 3.4A and C). Band intensities were statistically analysed to confirm no significant difference in MMRN2 binding between WT CLEC14A and each solvent facing mutant (Figure 3.4B and D).

Flow cytometry was also performed on the solvent-facing mutants of the first loop using the CRT2 and CRT4 antibodies. The gating strategy described for Figure 3.3 was employed and the MFIs were documented. WT CLEC14A was employed as positive controls which were recognised by both antibodies. Each single and double mutant expressed at the cell surface and was recognised by both the CRT2 and CRT4 antibody (Figures 3.5 and 3.6).

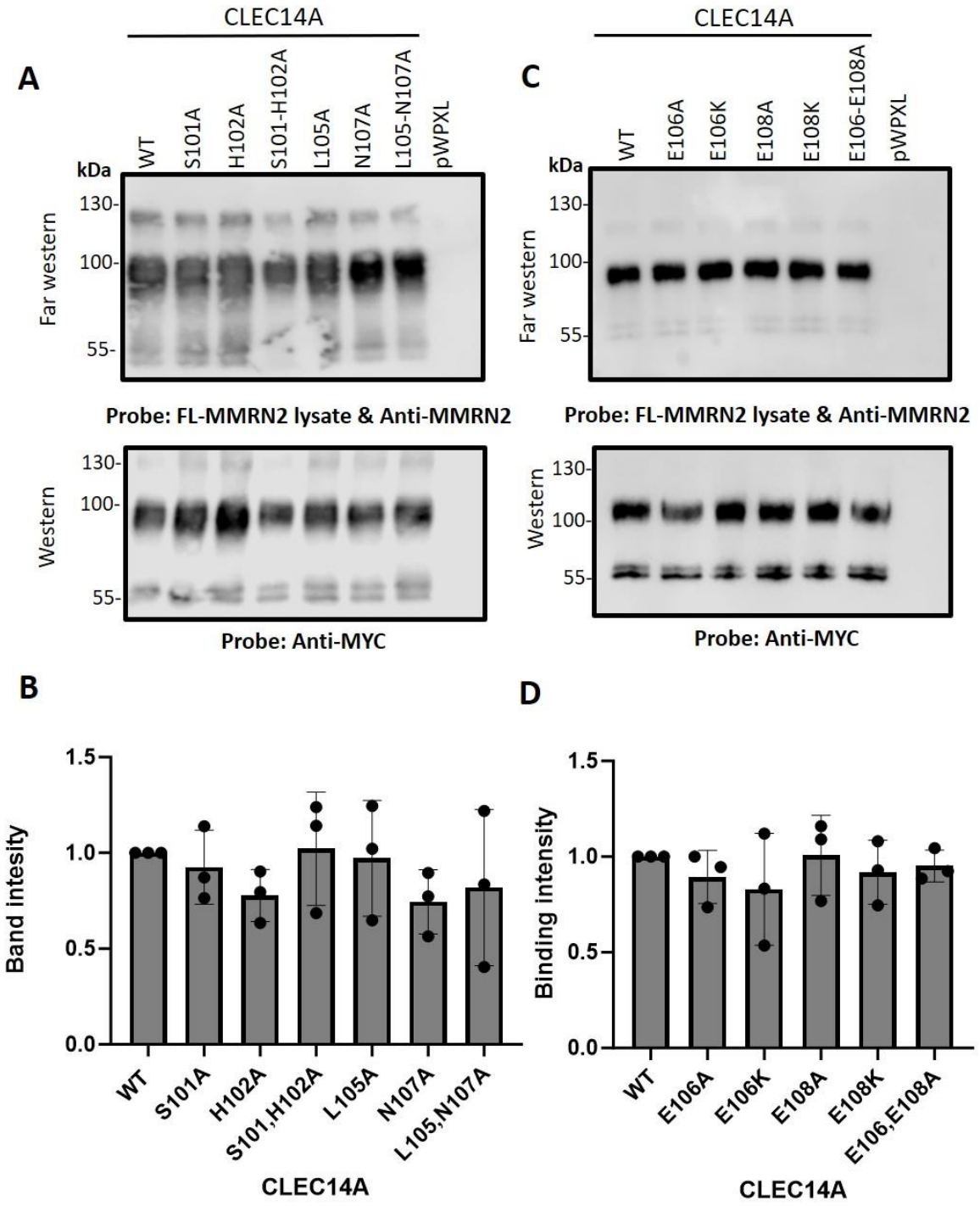


Figure 3.4. Solvent-facing residue mutants within the first loop bind to MMRN2. (A and C) HEK293T cells were transfected with plasmids to express either WT or mutant CLEC14A, lysates were collected and either run in reducing condition for western blot analysis or non-reducing conditions for far western blot analysis. Far western blots were probed with cellular lysates prepared from cells expressing full length MMRN2 followed by the anti-MMRN2 antibody. Western blots were probed with the Anti-myc antibody to determine expression levels of the various forms of CLEC14A. (B and D) Band intensities were quantified using the LI-COR and intensities of mutant bands were compared to WT CLEC14A bands. The results from the comparison are shown in the graphs alongside error bars and plots for each experimental repeat. Statistical analysis of band intensities was performed using an ANOVA test followed by a Dunnett's test to compare significance to the WT control, mean ± S.D. (n=3). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

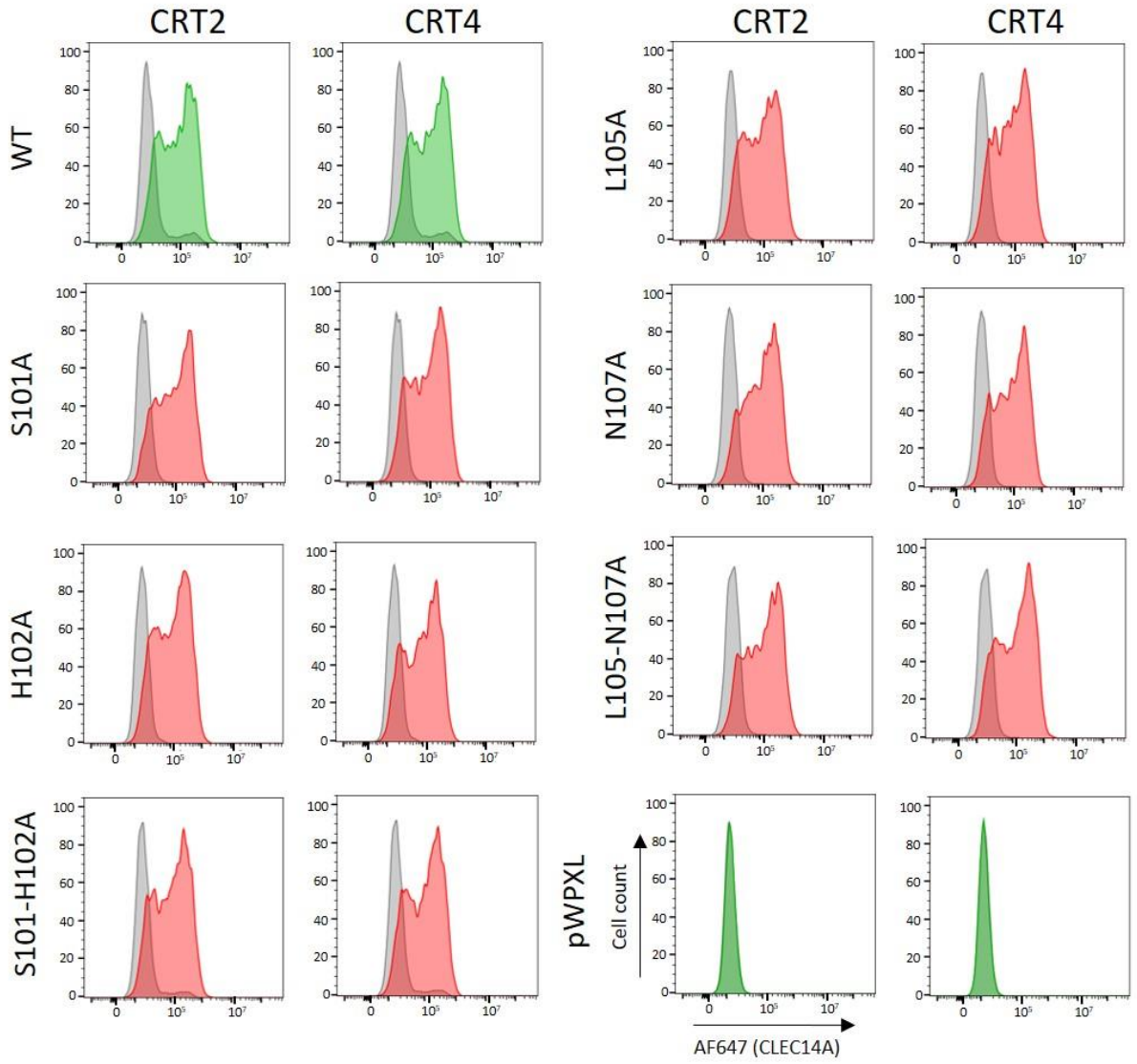


Figure 3.5. Solvent-facing mutants within the first loop expressed at the cell surface. HEK293T cells were transfected with plasmids to express either WT or mutant CLEC14A or the pWPXL empty vector. Cells were incubated with CRT2 or CRT4 as primary antibodies followed by incubation with an anti-mouse IgG Alexafluor647 conjugated antibody. Flow cytometry was performed, and the histograms presented show fluorescence in the APC channel gated on live cells expressing GFP. The histograms for the cells transfected with either pWPXL control or expressing WT CLEC14A are highlighted in green, whereas those of mutant forms of CLEC14A are in red. Signal from the control antibody is shown in grey. This figure is representative of data produced from three experimental replicates.

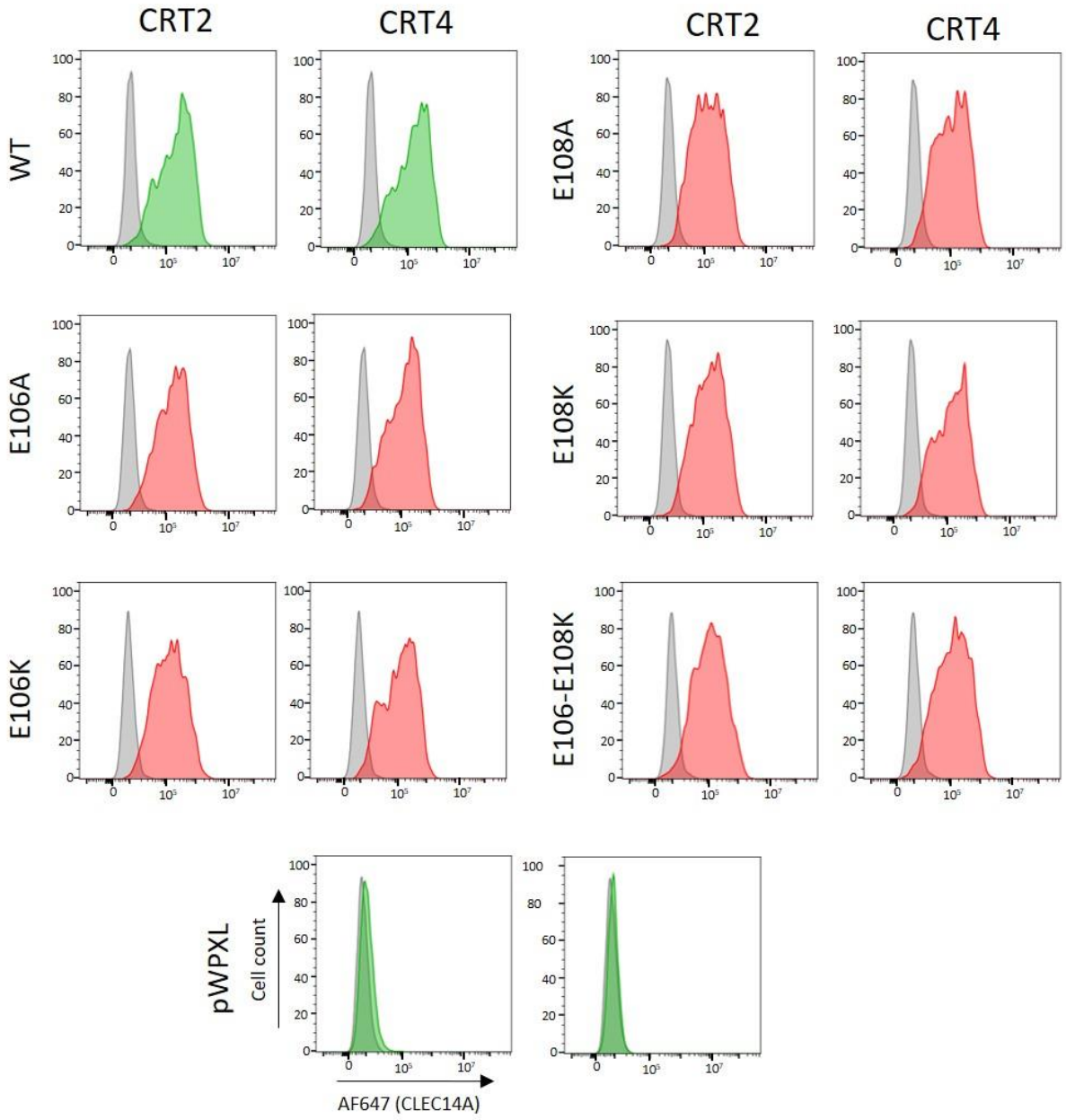


Figure 3.6. Glutamic acid solvent-facing mutants within the first loop expressed at the cell surface. HEK293T cells were transfected with plasmids to express either WT or mutant CLEC14A or the pWPXL empty vector. Cells were incubated with CRT2 or CRT4 as primary antibodies followed by incubation with an anti-mouse IgG Alexafluor647 conjugated antibody. Flow cytometry was performed, and the histograms presented show fluorescence in the APC channel gated on live cells expressing GFP. The histograms for the cells transfected with either pWPXL control or expressing WT CLEC14A are highlighted in green, whereas those of mutant forms of CLEC14A are in red. Signal from the control antibody is shown in grey. This figure is representative of data produced from three experimental replicates.

3.3. S137, T139 and R141 of CLEC14A directly interact with MMRN2

The next strategy employed was to target solvent-facing residues within the second loop of the loop in loop structure. . The S137, T139 and R141 residues were mutated as triple, double and single mutants where each residue was mutated to alanine. When tested in a far western, the triple mutant S137-T139-R141A and each double mutant; S137-T139A, S137-R141A and T139-R141A did not show any visible banding indicating each did not bind to MMRN2 when WT CLEC14A did (Figure 3.7A). Each single mutant: S137A, T139A and R141A bound to MMRN2 to varying degrees and to a much lesser extent when compared to the binding of WT CLEC14A to MMRN2 (Figure 3.7A). Statistical analysis confirmed each mutant had a significantly reduced capacity to bind to MMRN2 to varying degrees. These results reached statistical significance in three individual experiments (Figure 3.7B). The corresponding western blot revealed the two bands pattern seen previously which was present for WT and each mutant form of CLEC14A. This indicates relatively equal protein expression in the cellular lysate compared to WT CLEC14A. Interestingly, in each case the lower molecular weight band was more intense compared with the higher molecular weight band (Figure 3.7A).

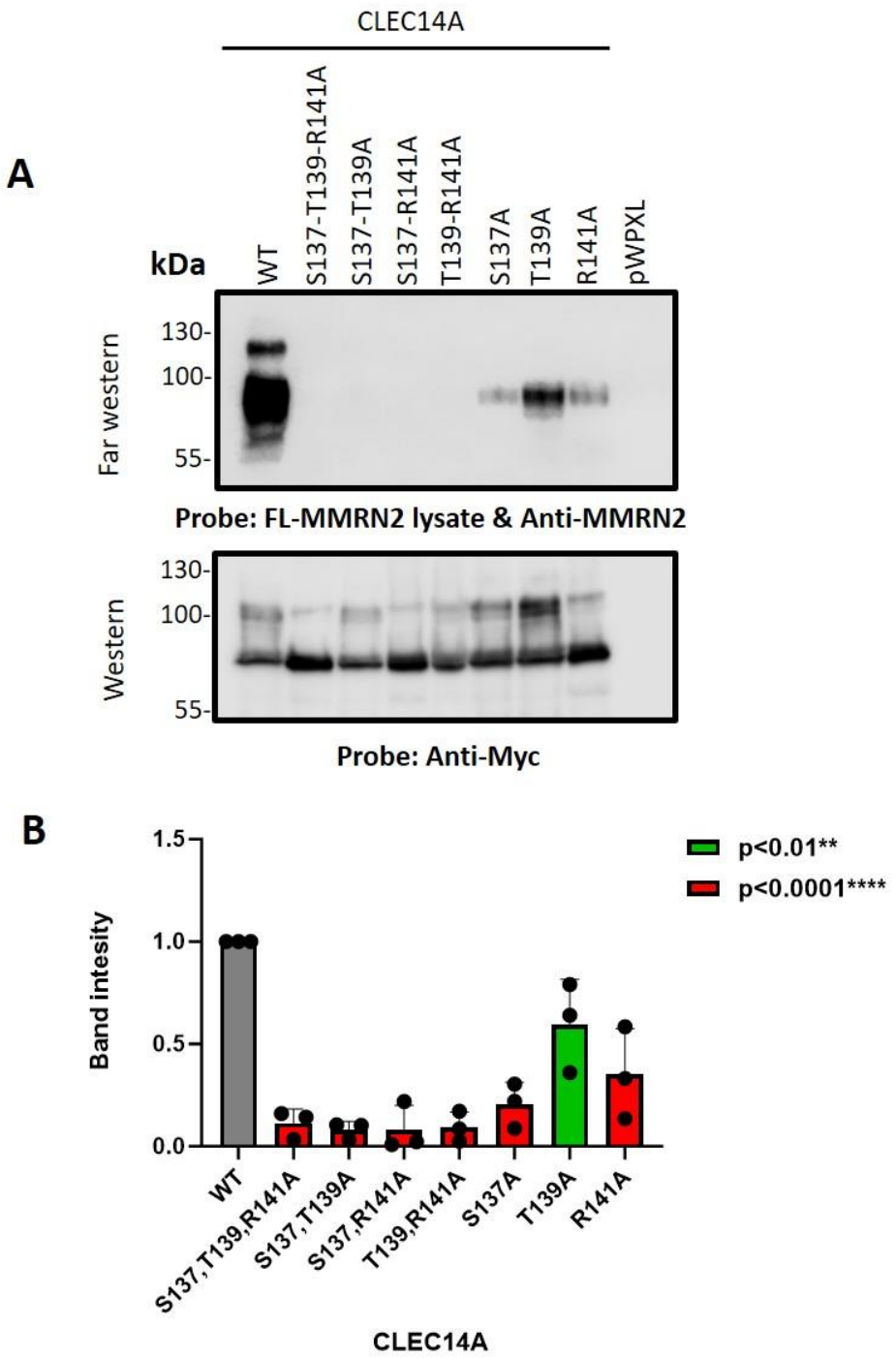


Figure 3.7. The S137, T137, R141 residues of CLEC14A bind to MMRN2. (A) HEK293T cells were transfected with plasmids to express either WT or solvent-facing mutant CLEC14A within the second loop, lysates were collected and either run in reducing conditions for western blot analysis or non-reducing condition for far western blot analysis. Far western blots were probed with cellular lysates prepared from cells expressing full length MMRN2 followed by the anti-MMRN2 antibody. Western blots were probed with the Anti-myc antibody to determine expression levels of the various forms of CLEC14A. (B) Band intensities were quantified using the LI-COR and intensities of mutant bands were compared to WT CLEC14A bands. The results from the comparison are shown in the graphs alongside error bars and plots for each experimental repeat. Statistical analysis of band intensities was performed using an ANOVA test followed by a Dunnett's test to compare significance to the WT control, mean ± S.D. (n=3). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

The S137-T39-R141A triple mutant and each double and single mutant generated was also studied by flow cytometry using the CRT2 and CRT4 antibodies, the same gating strategy was employed and MFIs were documented. WT CLEC14A was used as a positive control which expressed at the cell surface and was recognised by both CRT2 and CRT4 (Figure 3.8A). The triple mutant and each double and single mutant involving S137, T139 or R141 expressed at the cell surface as they were each recognised by the CRT2 antibody and exhibited generally similar MFIs compared to WT. Conversely signal from the CRT4 antibody varied between mutants. Mutants S137-T139-R141A, S137-R141A, T139-R141A and R141A each exhibited reductions in MFI compared to WT CLEC14A, particularly R141A where the CRT4 signal was reduced to levels similar to negative staining. Other mutants which did not include a mutated R141 residue appeared to have higher MFIs for both CRT2 and CRT4 staining compared to WT (Figure 3.8B). Taken together, this data set confirms that the R141 residue is an integral part of the binding epitope of the CRT4 antibody. None of the mutants containing a mutated form of R141 were recognised by the CRT4 antibody as well as WT CLEC14A (Figure 3.8). This is the first data to suggest the CRT4 binding epitope and the MMRN2 binding site on CLEC14A overlap.

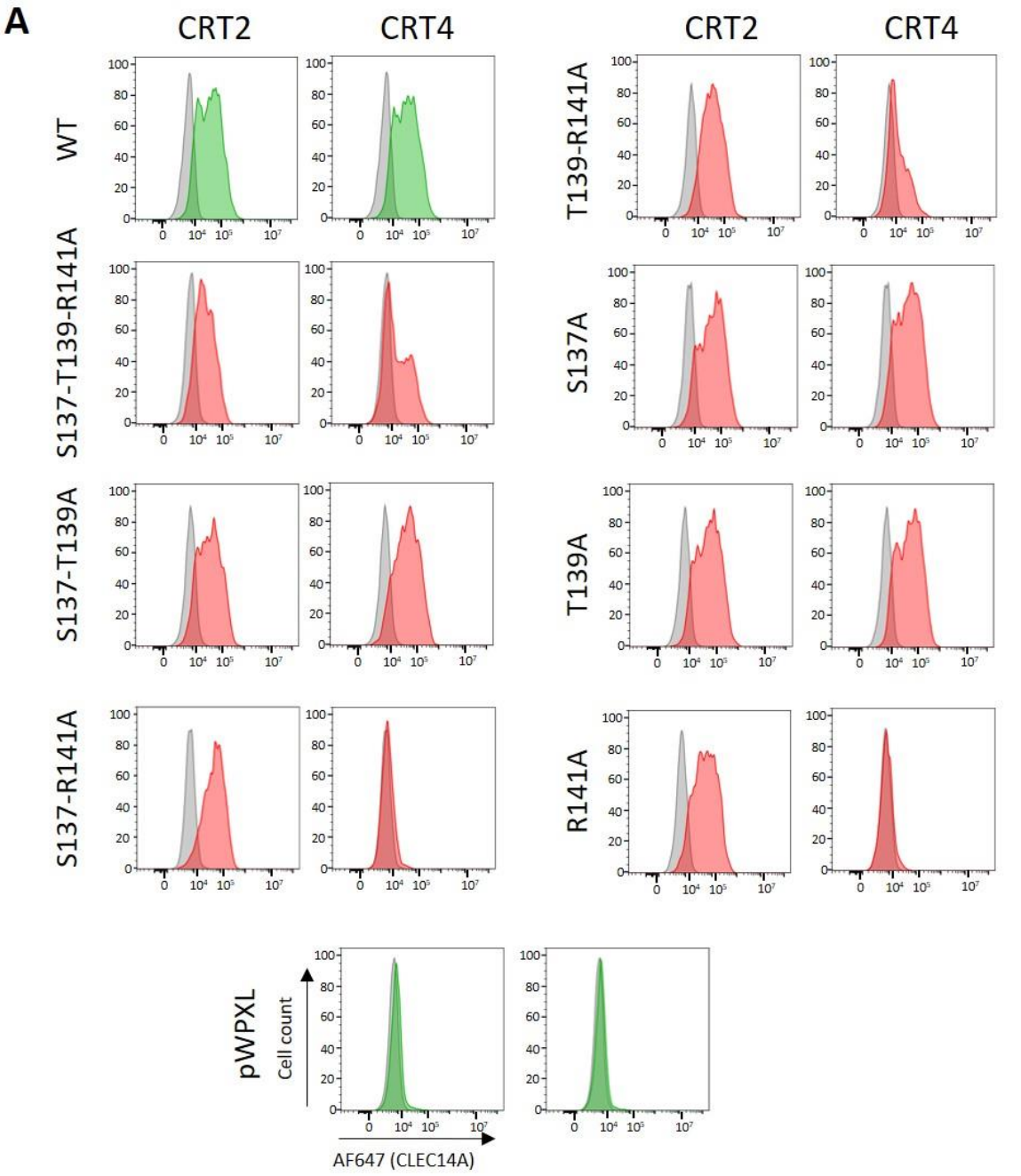


Figure Legend on next page

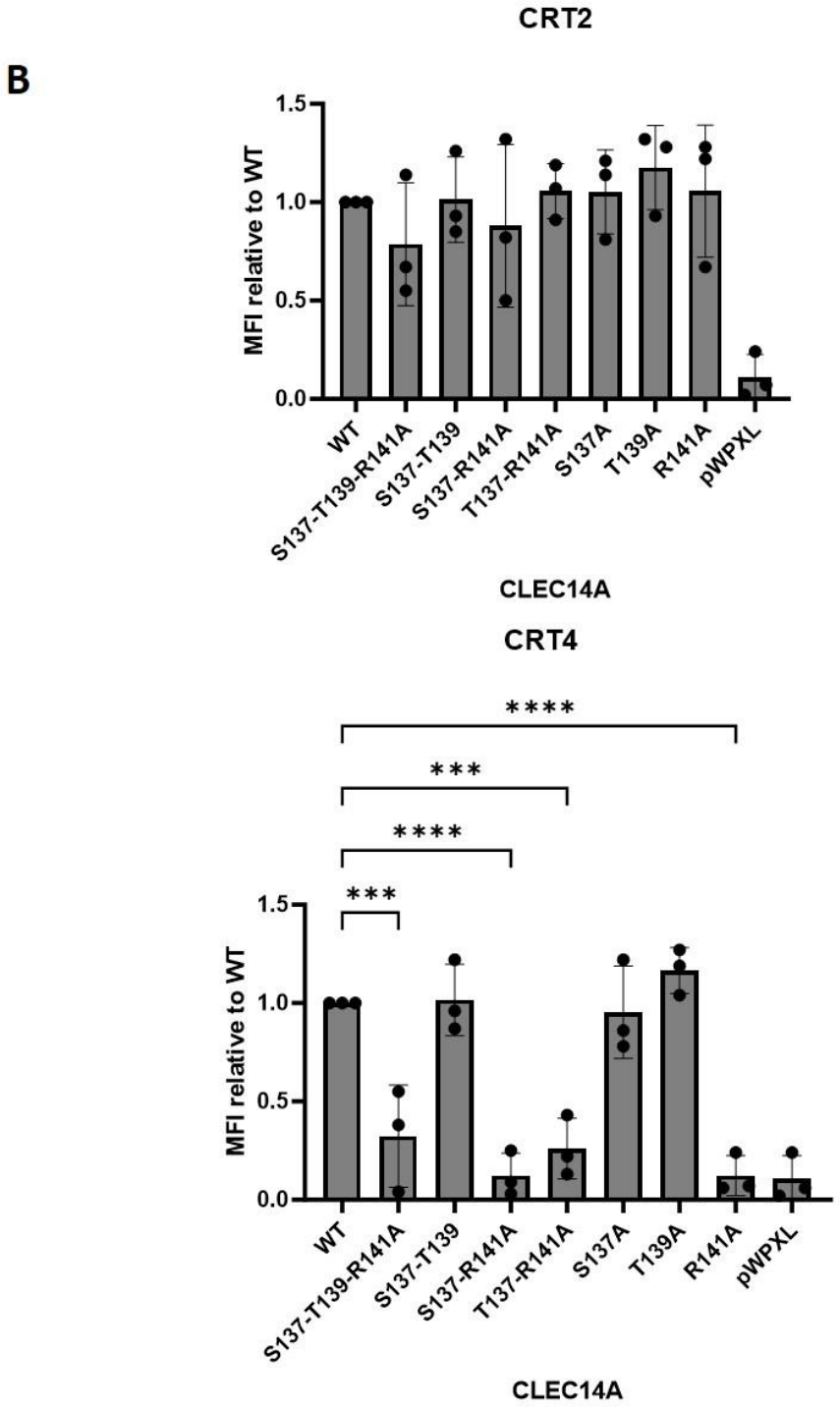


Figure 3.8. The R141 residue of CLEC14A forms part of the binding epitope for the CRT4 antibody. HEK293T cells were transfected with plasmids to express either WT or mutant CLEC14A or the pWPXL empty vector. Cells were incubated with CRT2 or CRT4 as primary antibodies followed by incubation with an anti-mouse IgG Alexafluor647 conjugated antibody. (A) Flow cytometry was performed, and the histograms presented show fluorescence in the APC channel gated on live cells expressing GFP. The histograms for the cells transfected with either pWPXL control or expressing WT CLEC14A are highlighted in green, whereas those of mutant forms of CLEC14A are in red. Signal from the control antibody is shown in grey. Normalised mean fluorescent intensity values are depicted in the corner of each histogram. This figure is representative of data produced from three experimental replicates. (B) MFIs for each mutant and control were identified and analysed relative to WT CLEC14A for both the signal from CRT2 and CRT4 treatment. The result from a triplicate set of experiments were statistically analysed using an ANOVA test followed by a Dunnett's test to compare significance to the WT control, mean ± S.D. (n=3). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

3.4. The study of the CLEC14A mimicking peptide.

Given the specific expression of CLEC14A on the tumour endothelium as well as its role in angiogenic processes such as tube formation and cell migration, CLEC14A is an attractive target for anti-cancer therapy. Previously it has been shown that the CRT4 antibody which blocks the interaction between CLEC14A and MMRN2, results in decreased tumour size *in vivo* (Noy *et al*., 2015; Khan *et al.,* 2017*)*. Antibody treatment comes with a range of drawbacks including reduced tissue perfusion due to strong binding affinity as well as mode of administration. Conversely, small molecule inhibitors can be administered orally as oppose to intravenously and have high tissue perfusion. Furthermore, due to the large size of antibodies it is possible that binding of antibodies can block binding of other binding partners.

As the specific region of CLEC14A where MMRN2 binds has been identified (Figure 3.7), a 20 amino acid-long peptide was commissioned to mimic this region. This peptide spanned between W130 and T149 of CLEC14A, a scrambled version of the peptide was used as a negative control. The capacity of the peptide to inhibit the CLEC14A-MMRN2 interaction was studied using an enzyme-linked immunosorbent assay (ELISA) (section 2.4.5). The ELISA set up was validated prior to addition of peptides. This involved coating a 96-well plate with 100ng of purified MMRN2 495-674 protein. The wells were blocked and treated with conditioned media containing the CLEC14A-Fc construct. A dilution series was generated using Opti-MEM media. Wells were treated with each dilution of the conditioned media and then with anti-Fc HRP. The results indicate that the ELISA set up successfully detected the interaction between MMRN2 495-674 and CLEC14A-Fc and the signal increased as the concentration of CLEC14A-Fc increased (Figure 3.9).

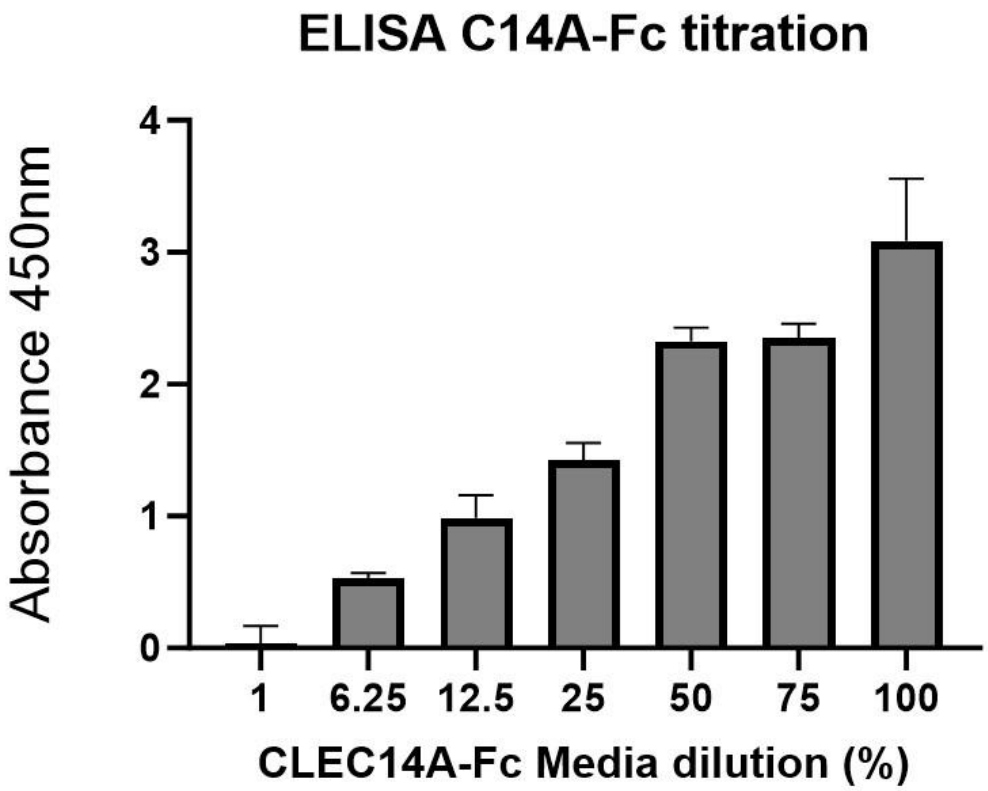


Figure 3.9. Conditioned media containing CLEC14A-Fc binds to MMRN2 495-674 in an ELISA. Wells of an ELISA plate were coated with 100ng of MMRN2 495-674 purified protein for 3 technical repeats, the wells were then blocked and incubated with a dilution series of CLEC14A-Fc-containing conditioned media. The wells were then treated with the Anti-human Fc antibody conjugated to HRP. Following substrate addition, the plate was read at 450nm, and the absorbance was plotted in a bar chart. Values shown are mean ± S.D. (n=3).

3.5. The CLEC14A-mimicking peptide does not block the CLEC14A- MMRN2 interaction

The ELISA was set up, which involved coating the plate with 100 ng of MMRN2 495-674 and treating with CLEC14A-Fc conditioned media preincubated with increasing concentrations of either the CLEC14A-mimicking peptide or scrambled control. The results indicated that both the CLEC14A-mimicking peptide and the scrambled control produced similar readings at each concentration level. The CLEC14A-mimicking peptide did not block the interaction between CLEC14A and MMRN2 in the ELISA nor was there any significant change as the amount of peptide increased. The MMRN2 495-674 purified protein was used as an inhibitor control which was preincubated with the CLEC14A-Fc containing conditioned media at a 2X excess molar concentration, this blocked CLEC14A from binding to the MMRN2 on the plate (Figure 3.10).

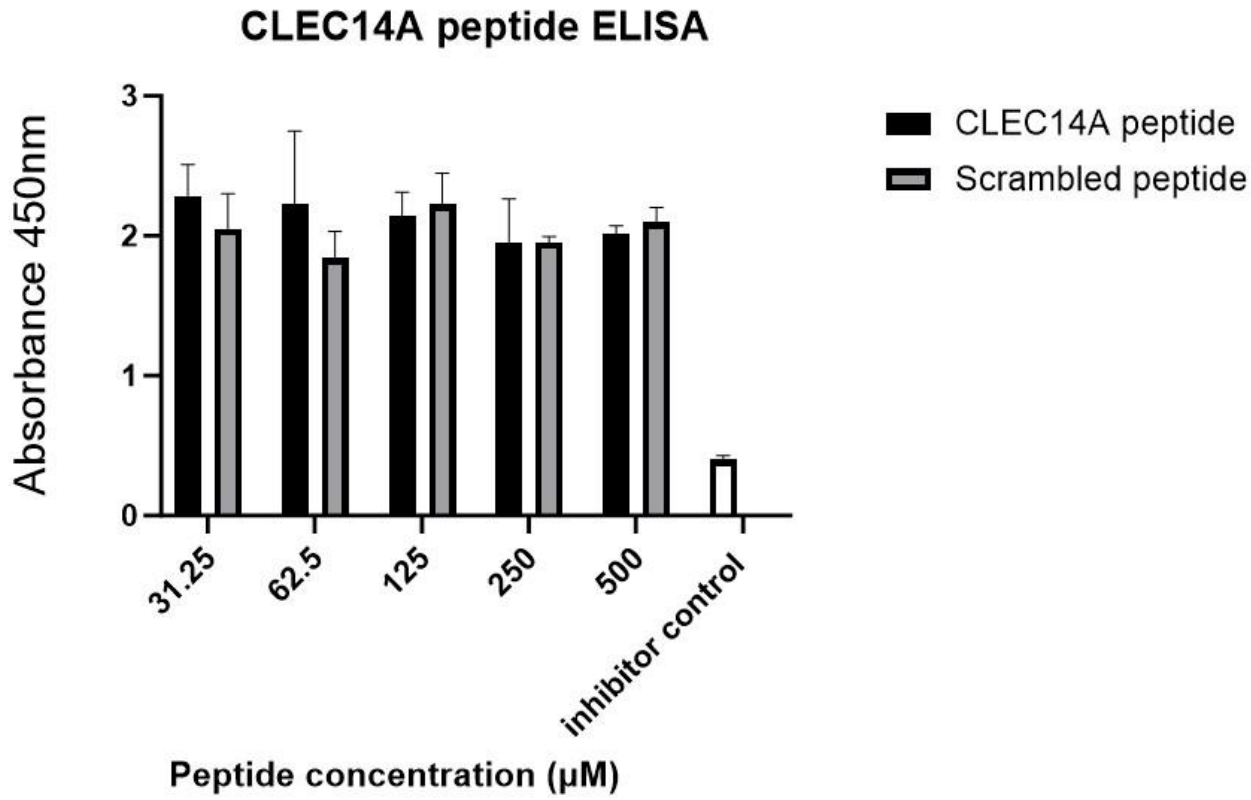


Figure 3.10. Neither the CLEC14A-mimicking peptide or scrambled peptide inhibit the interaction between CLEC14A-Fc and MMRN2 495-674. Wells of an ELISA plate were coated with MMRN 495-674 purified protein and blocked. The wells were then treated with CLEC14A-Fc conditioned media pre-incubated with increasing concentrations of either the CLEC14A-mimicking peptide or the scrambled peptide. The inhibitor control was CLEC14A-Fc conditioned media preincubated with an excess of MMRN2 495-674 purified protein. Wells were then incubated with the Anti-Fc antibody conjugated to HRP and following substrate addition the plate was read at 450nm. Values shown are mean ± S.D. (n=3).

3.6. MMRN2 residues E617, E620, S623 and E624 do not individually bind to CLEC14A

Predictive docking modelling of the CTLD of CLEC14A and MMRN2 495-674 fragment suggest a short coil of MMRN2 in close proximity to the 2nd loop of CLEC14A where S137, T139 and R141A are situated (Philip Morrison PhD thesis, 2023). The whole fragment of MMRN2 495-674 is folded into a coil which slightly kinks out between amino acids 617-624, within this short coil four amino acids project out in the direction of the second loop of CLEC14A. These are E617, E620, S623 and E624 (Figure 3.11).

Each of these residues were mutated to alanine using primers to amplify fragments which were assembled through Gibson assembly. Attempts to mutate all 4 residues simultaneously were unsuccessful. Mutants were constructed into the pHLAvitag3 vector which included a C-terminal His-tag. The plasmids were transfected into HEK293T cells which produced and secreted the protein into the growth media. To study protein expression and binding to CLEC14A, the western and far western blotting techniques were used again. Flow cytometry was also used to detect MMRN2 binding to WT CLEC14A on the surface of cells.

WT MMRN2 495-674 was used as a positive control for expression and CLEC14A binding. WT MMRN2 was expressed in the cellular lysate as indicated by a double band at approximately 20kDa, this expression level was mirrored with each mutant tested (Figure 3.11). The corresponding far western blot indicated the same blotting pattern to the western blot, this indicated that WT MMRN2 495-674 bound to WT CLEC14A. Since each mutant revealed a band in the far western blot, it can be concluded that none of the MMRN2 mutants generated affected binding with CLEC14A. This result was confirmed in three individual experiments. Statistical analysis confirmed no decrease in binding between MMRN2 mutants and WT CLEC14A (Figure 3.12).

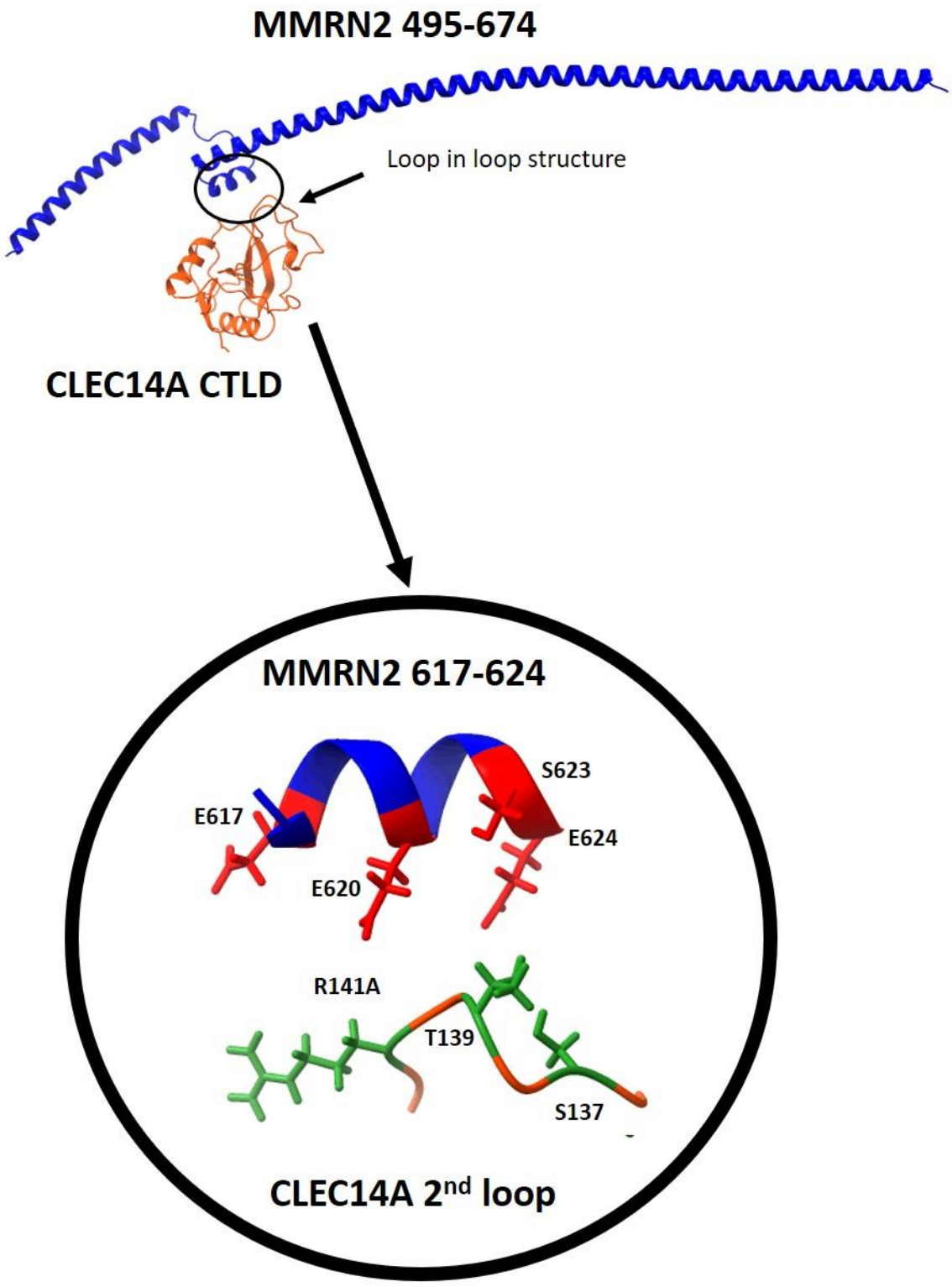


Figure 3.11. Predicted binding model of the CTLD of CLEC14A with the Coiled-coil domain of MMRN2 495-674. The CTLD of CLEC14A (green) with the 495-674 fragment of MMRN2 (Blue). A detailed view of the key amino acid side chains on MMRN2 616-624 predicted to be in proximity to the MMRN2 binding site of CLEC14A.

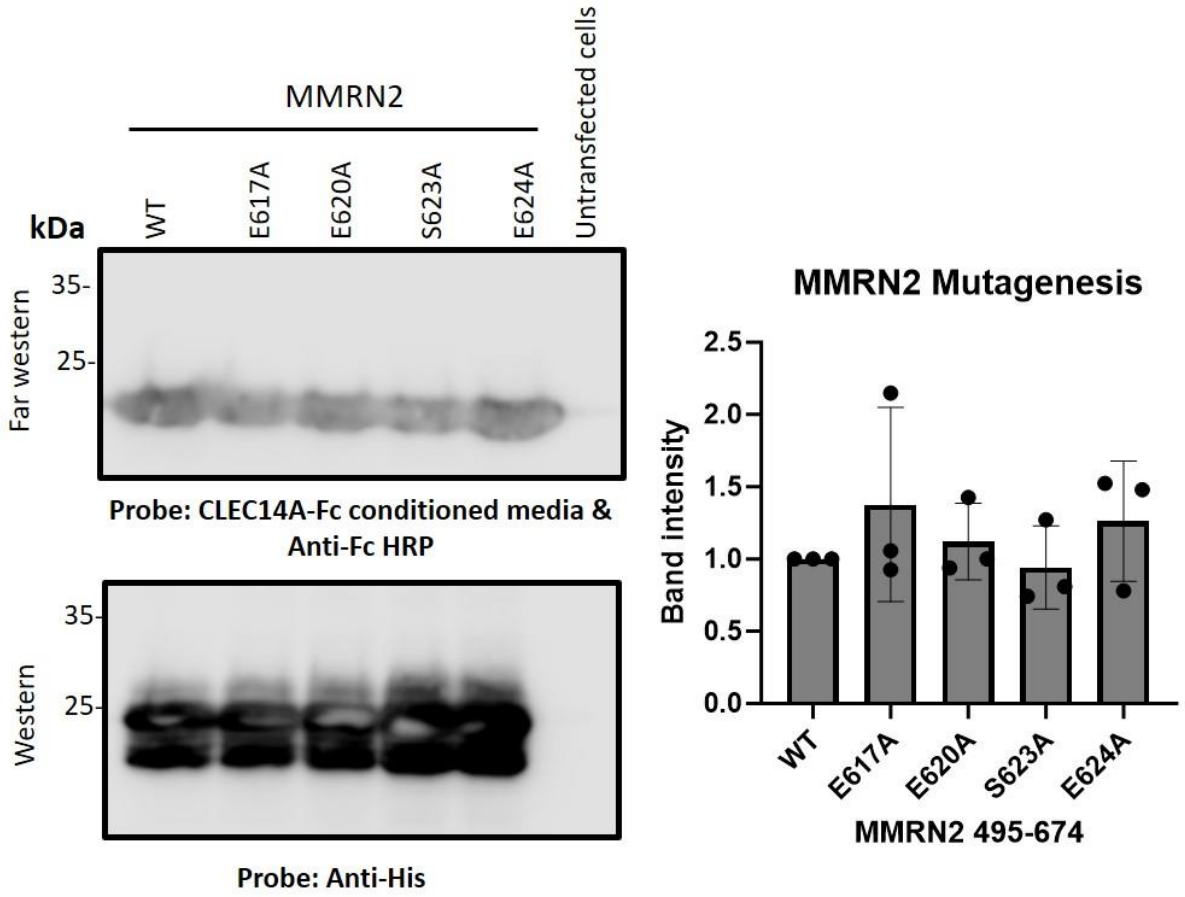


Figure 3.12. No MMRN2 495-674 mutant blocked the interaction with WT CLEC14A. HEK293T cells were transfected with plasmids to express the MMRN2 495-674 fragment. Lysates were collected from the cells and run in either non-reducing conditions for far western analysis or reducing conditions for western blot analysis. Far westerns blots were probed with conditioned media containing a CLEC14A-Fc construct followed by an anti-Fc HRP conjugated antibody. Western blots were probed with an anti-His tag antibody to determine expression of levels of the various forms of MMRN2 495-674. (B) Band intensities were quantified using the LI-COR and intensity values of mutant MMRN2 were compared to WT MMRN2. The results from the comparison are presented in the graph alongside errors bars for each experimental repeat. Statistical analysis of band intensities was performed using an ANOVA test followed by a Dunnett's test to compare significance to the WT control, mean ± S.D. (n=3). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

Flow cytometry analysis was also performed to study the binding of WT and mutant MMRN2 with WT CLEC14A using the gating strategy described for Figure 3.3. Flow cytometry analysis revealed that WT CLEC14A expressed at the cell surface, as it was recognised by the CRT4 antibody. It was also shown that WT MMRN2 bound to CLEC14A on the cell surface as the protein was recognised by the anti-His tag antibody. These results were mirrored for each MMRN2 mutant confirming that every mutant bound to CLEC14A, the pWPXL negative control did not display any peaks above the isotype control (Figure 3.13).

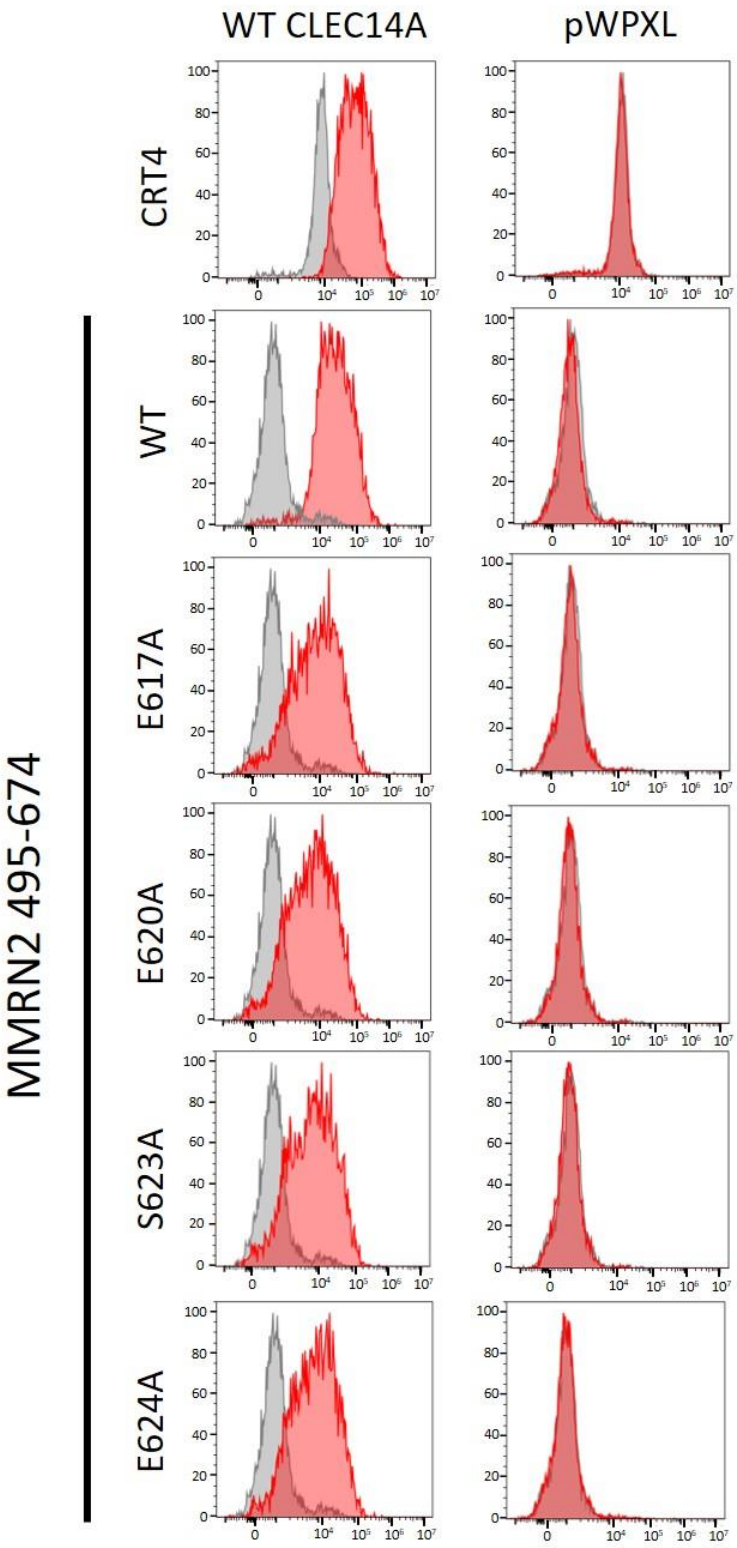


Figure 3.13. WT and mutant forms of MMRN2 495-674 bound to CLEC14A in flow cytometry. HEK293T cells were transfected with the plasmid to express WT CLEC14A, or the pWPXL empty vector. Cells were then incubated with either the CRT4 antibody or conditioned media contain WT or mutant MMRN2. This was obtained from HEK293T cells transfected with plasmid to express and secrete either WT or mutant MMRN2. The CRT4 treated cells were then incubated with an anti-mouse IgG Alexafluor647 conjugated antibody whilst the MMRN2-containing conditioned media treated cells were incubated with an anti-His tag Alexafluor647 conjugated antibody. Flow cytometry was performed, and the histograms presented show fluorescence in the APC channel gated on live single cells. Signal from negative controls are shown in grey. This figure is representative of data produced from three experimental replicates.

3.7 Discussion

In this portion of the project the residues of CLEC14A which directly interact with MMRN2 were identified as well as the residues which form part of the binding epitope of the CRT4 blocking antibody. The peptide which mimicked the MMRN2 binding region of CLEC14A did not inhibit the interaction.

The residues of CLEC14A which directly engage with MMRN2 were found to be S137, T139 and R141 which are situated on the second loop of the "loop-in-loop" structure of the CTLD of CLEC14A (Figure 3.14A). Previously in the literature, it was suggested that the 97-ERRRSHCTLENE-108 sequence within the first loop played a role in MMRN2 binding because upon replacement of that sequence with the analogous sequence from THBD the interaction was abrogated (Khan *et al.,* 2017). However, data presented within this chapter show that this mutant form is not efficiently trafficked to the cell surface and has an altered banding pattern indicating differential glycosylation. This study instead demonstrates that it is the second loop of the CTLD that is involved in MMRN2 binding.

This study highlights the importance of using accurate predictive modelling tools such as AlphaFold (Jumper *et al.,* 2021) to guide decision making in using mutagenesis to map interactions. When residues predicted to be solvent-facing were altered, protein expression and trafficking were not affected thus enabling a clearer interpretation of the resulting interaction data.

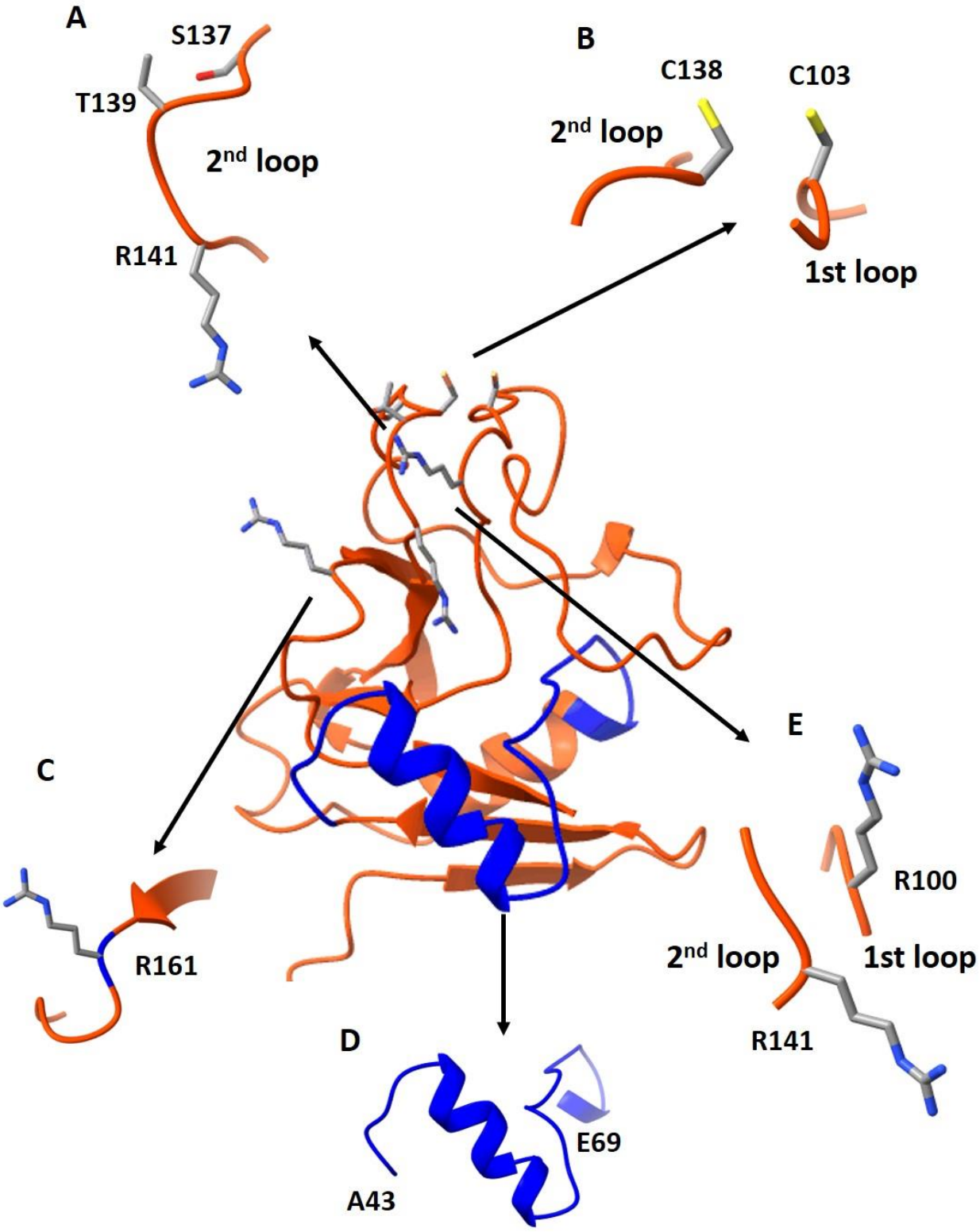


Figure 3.14. The AlphaFold predicted structure of the CTLD viewed on chimera software. The figure highlights key regions within the domain. (A) Highlights the MMRN2 binding site on CLEC14A. (B) Highlights C103 and C138 which form a disulphide link. (C) Highlights a heparin binding residue (Sandoval *et al.,* 2020). (D) Highlights the binding region of the heat shock protein 70-1A (Jang *et al.,* 2017) (E) Highlights the residues which form the binding epitope for the CRT4 antibody.

Another key finding in this project was that the R100 and R141 residues form part of the binding epitope of the CRT4 antibody. These residues which are separated by 40 amino acids in the primary structure are predicated to be co-located at the base of both loops (Figure 3.14E). The CRT4 antibody has been used extensively as a reagent to block the CLEC14A and MMRN2 interaction (Noy *et al.,* 2015; Khan *et al.,* 2017). It was known prior to this study that CRT4 was able to block the binding of MMRN2 to CLEC14A, and that both CRT4's binding epitope and the region to which MMRN2 binds is within the CTLD (Khan *et al.,* 2017). This study shows that the same residue R141 is part of both the CRT4 and MMRN2 binding site.

A similar binding experiment was performed by Galvagni *et al*, which involved studying the CD93-MMRN2 interaction using different techniques to our study. To investigate the CD93-MMRN2 interaction, a solid phase binding assay with recombinant CD93 and MMRN2 was employed alongside western blots. This approach allowed for the study of protein binding and glycosylation status. However, the major drawback in this study was the lack of evaluation of cell surface expression and folding. At the time of the study, AlphaFold was not available. Therefore, homology docking models were generated using the Phyre2 server. These models were based on the crystal structures of distantly related proteins and were less accurate than those generated by AlphaFold. The study concluded that the F238 residue in the sushi domain of CD93 was important for interaction with MMRN2, however due to the lack of folding confirmation, this could be an indirect effect due to protein misfolding (Galvagni *et al.,* 2017). More recently, the same lab resolved the crystal structure of CD93, revealing that CD93 forms a homodimer with a higher binding capability to MMRN2 compared to monomeric CD93 (Barbera *et al.,* 2023). The homodimer interface spans both the CTLD and sushi domain, suggesting that the F238 residue within the sushi domain could play a critical role in

dimerization and subsequent MMRN2 binding. These findings highlight the importance of using accurate structural models as well as comprehensive analyses to study interactions.

The analogous residues on CD93 that align with the MMRN2-binding residues on CLEC14A are relatively conserved. Specifically, S137 of CLEC14A aligns with S135 of CD93, T139 of CLEC14A aligns with I137 of CD93 (both uncharged), and R141 of CLEC14A corresponds to K139 of CD93 (both positively charged)

CLEC14A is known to mediate interactions with other proteins and carbohydrates through its CTLD. It has been shown in the literature that the R161 residue engages with heparin resin (Sandoval *et al.,* 2020) and that the fragment between A43 and E69 binds to Heat shock protein70-1A (HSP70-1A) (Jang *et al.,* 2017). In the context of the literature the work conducted in this study expands the currently mapped interactions of the CTLD. Each CTLD binding interaction appears to be spatially separated from one another and neither heparan sulfate nor HSP70-1A interact with the "loop-in-loop" structure (Figure 3.14C and D).

Subsequently, a CLEC14A mimicking peptide was generated which failed to block the interaction between CLEC14A and MMRN2. S137, T139 and R141 are presented on a loop of the CTLD which is stabilized in place by a disulfide bond at the apex of both loops (Figure 3.14B) (Khan *et al.,* 2017). The peptide used in this study was linear and may not have presented the MMRN2 binding residues in the same orientation of native CLEC14A. Since the residues are present on a loop it may be the case that a cyclical peptide would better present the residues for MMRN2 binding.

The choice of peptides over antibodies for blocking protein-protein interactions offers several advantages. Peptides have increased tissue perfusion compared to antibodies due to their

small size and are less likely to trigger an immune response, peptides are also more amenable to modification which may support absorption of peptides when taken orally as oppose to antibodies which are highly sensitive to digestive enzymes and the acidic environment (Vora *et al.,* 2022).

These novel findings open up the possibility of studying the CLEC14A-MMRN2 interaction in greater detail *in vivo* using genetically engineered mice that express the S137-T139-R141A mutant. This method would provide benefits over the using of CLEC14A knockout mice as only the CLEC14A-MMRN2 interaction would be affected and not the CD93-MMRN2 interaction nor the CLEC14A-heparan sulfate interaction. Further binding experiments using HSP70-1A would be required to ensure the mutants do not block the CLEC14A-HSP70-1A interaction. Other potential avenues involve studying the effect of the MMRN2-blocking mutants *in vitro* using induced pluripotent stem cell (iPSC) vascular models where CRISPR-Cas9 is used to introduce the S137, T139, R141A mutants to CLEC14A. These models can then be used in a range of angiogenesis-based assays such as sprouting assays, tube formation assays, permeability assays to study barrier function of the endothelium, cell migration assays as well as angiogenic factor response assays (Esparza *et al.,* 2024).

Overall, the data presented in this chapter expands on our current understanding of the CLEC14A-MMRN2 interaction and provides valuable insights into potential therapeutic interventions, as well as future avenues for studying this interaction in greater depth. This work also underscores the significant benefits of integrating computational predictive modeling, such as AlphaFold, into the research strategy. This predictive modeling approach facilitated informed decision-making, allowing for the efficient identification of the next steps in the project.

CHAPTER 4: Investigating the interactions of heparin with MMRN2 and CLEC14A

CHAPTER 4: Investigating the interactions of heparin with MMRN2 and CLEC14A

4.1. Introduction

Heparin sulphate proteoglycans (HSPG) are an abundant class of glycosaminoglycans present on the cell surface and in the extracellular matrix space as either transmembrane spanning proteoglycans or chains of glycosaminoglycans. HSPGs exert a wide range of functions including but not limited to cell adhesion to the extracellular matrix and angiogenesis (Olczyk *et al.,* 2015; Chiodelli *et al.,* 2015). CLEC14A has been shown in the literature to elicit a proangiogenic phenotype which can be disturbed via treatment with the CRT4 antibody. CLEC14A also binds directly to heparin with nanomolar affinity and upon mutagenesis of the R161 residue to alanine the interaction with heparin is abrogated (Noy *et al.,* 2015; Sandoval *et al.,* 2020). The functional significance of this interaction has not yet been fully explored however given the proangiogenic properties of both CLEC14A and heparin sulfate (Mura *et al.,* 2012; Chiodelli *et al.,* 2015), it may be possible that CLEC14A and HSPGs work synergistically to promote angiogenesis. Furthermore, MMRN2 has also been shown in the literature to interact with heparin sulfate, an interaction which blocks the MMRN2-VEGFA interaction to illicit pro angiogenic phenotype (Colladel *et al.,* 2015).

The overall aim of this part of the project was to study the effect of CRT antibody treatments on the interaction between CLEC14A and heparin. The study also aimed to expand on the current knowledge of the heparin interactions with CLEC14A and MMRN2 in terms of interacting regions and residues.

4.2. The R161A CLEC14A mutant reduces binding to Heparin.

It has previously been shown in the literature that CLEC14A interacts with the heparin sulfate proteoglycans and that when mutating the R161 residue to alanine the interaction is abrogated (Sandoval *et al.,* 2020). To confirm these results the R161A CLEC14A extracellular domain was cloned into the pIgPP vector to include a C-terminal Fc-tag (constructed by Victoria Heath). The construct was then tested for expression using HEK293T cells and protein A beads. In order to test the capacity of WT and R161A mutant CLEC14A to bind to heparin, heparin coated agarose beads were used. Both WT and R161A CLEC14A were expressed and secreted into the media at similar levels as indicated in the protein A bead pull-down (Figure 4.1A). WT CLEC14A bound to heparin beads and the band was comparable to that of the protein A bead pull down of WT CLEC14A thus indicating no particular reduction or increase when binding to heparin. Conversely, the R161A CLEC14A mutant also bound to heparin however the amount of CLEC14A pulled down with heparin beads was reduced resulting in a lower band intensity (Figure 4.1A). Statistical analyses on the band intensities revealed a significant decrease in binding capacity of the R161A CLCEC14A mutant to heparin beads compared with binding to protein A beads (Figure 4.1A).

Given that the R161A CLEC14A mutant did not completely abrogate binding with heparin, the next step was to identify other mutants on CLEC14A which may contribute to the interaction. To this end the mutant forms of CLEC14A which impeded the interaction with MMRN2 (Figure 3.4C and D) were studied. Each variation of the CLEC14A mutants which reduces binding to MMRN2 were subcloned into the pIgPP vector which included a C-terminal Fc-tag. Again, the proteins were expressed using HEK293T cells and the media was collected. The proteins were then incubated with either Protein A beads or heparin coated agarose beads to study expression and heparin binding respectively. The CLEC14A mutants; S137-T139-R141A, S139-

T139A, S137-R141A, T139-R141A, S137A, T139A and R141A which reduced binding to MMRN2 appeared to express to different degrees within the conditioned media. Therefore, the media was diluted prior to heparin and protein A bead pull down and equal amounts were loaded (Figure 4.1A). The heparin bead pull-down blot revealed that each mutant bound to heparin to similar degrees when compared to WT CLEC14A. This indicated that the distinct residues involved in the CLEC14A-MMRN2 interaction do not interact with heparin (Figure 4.1A). Statistical analysis of band intensities was performed to confirm no significant decrease or increase in heparin binding of each MMRN2-blocking mutant when compared to the protein A bead pull-down (Figure 4.1B).

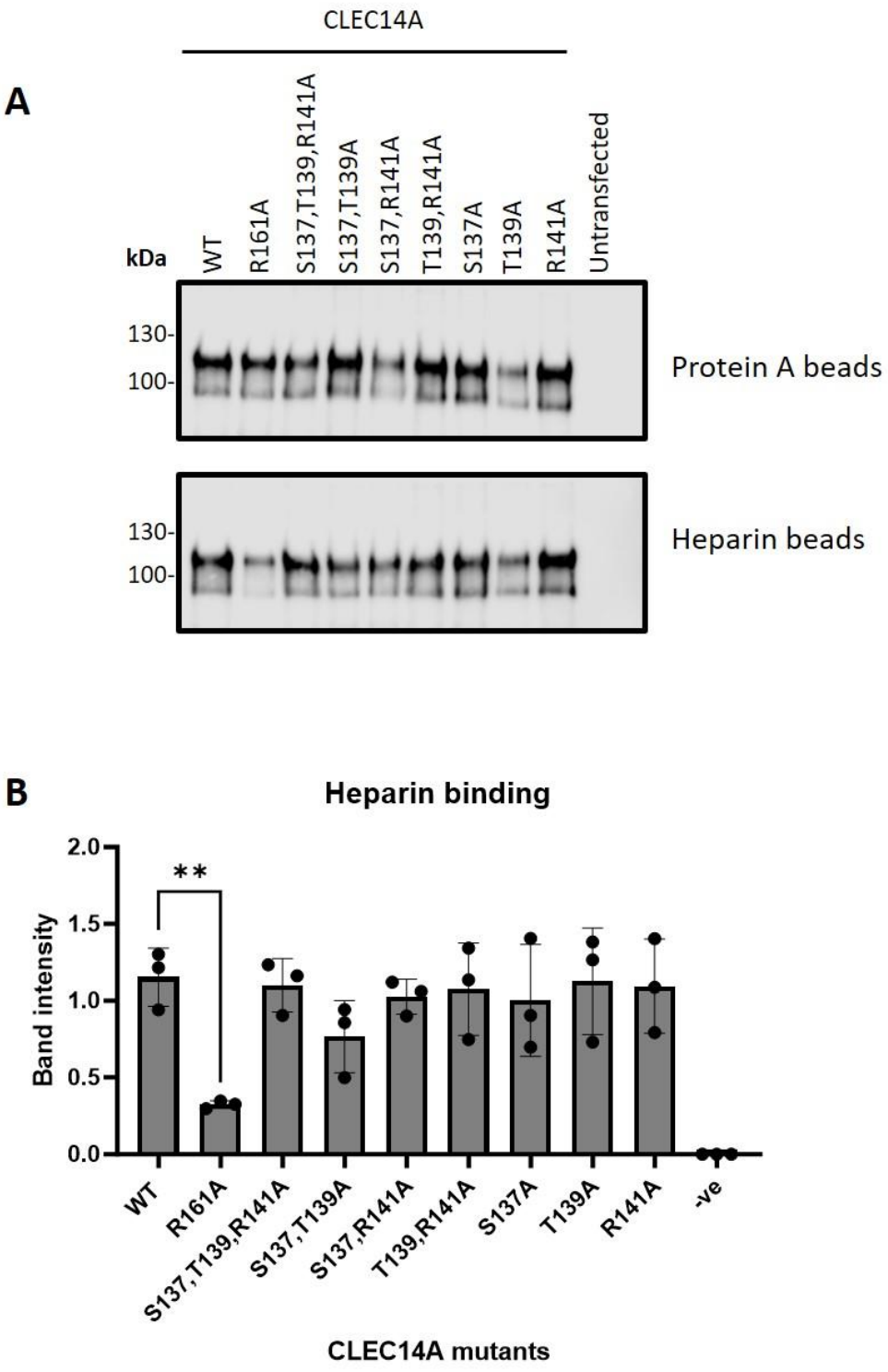


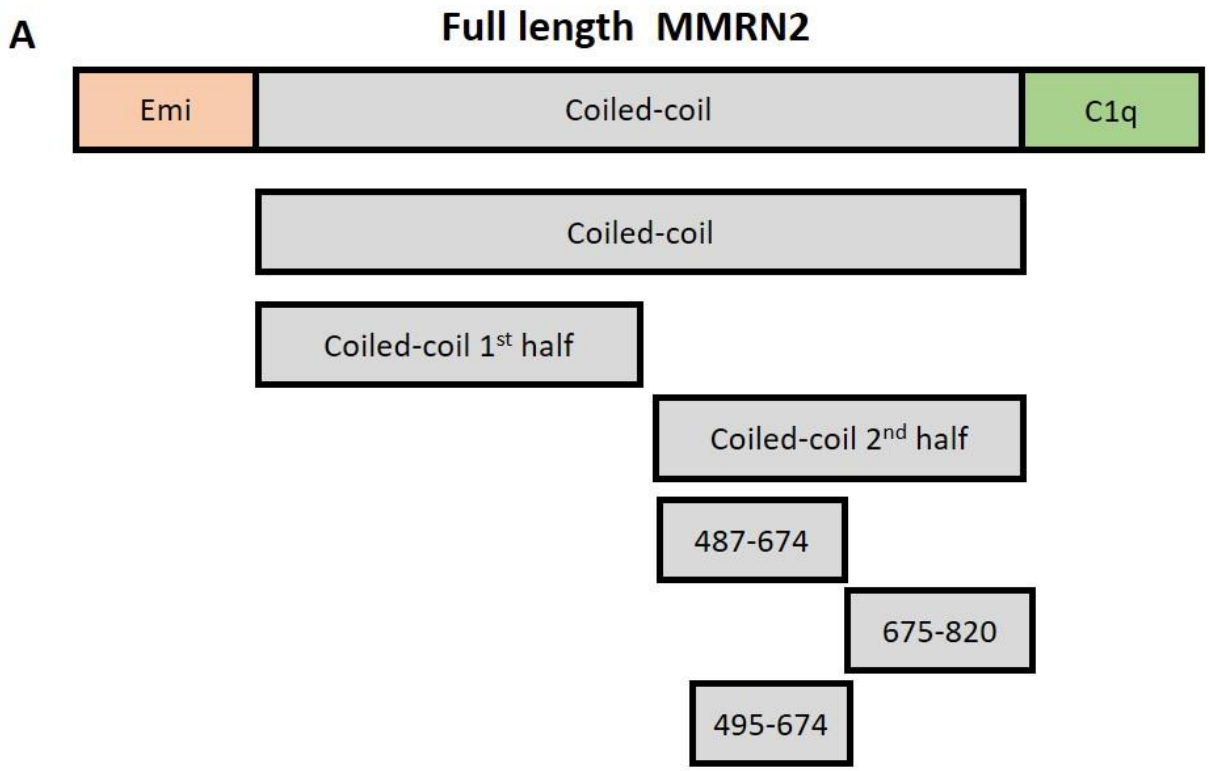
Figure 4.1. CLEC14A mutants which do not bind to MMRN2 can bind to heparin. HEK293T cells were transfected with plasmids to produce and secrete either WT or mutant forms of CLEC14A which were fused to a human Fc tag into the surrounding growth media. The conditioned media containing the proteins were collected and incubated with either protein A beads to study protein expression or heparin beads to study heparin binding. (A) Beads were treated with reducing sample buffer, boiled and subjected to western blotting, where membranes were probed with anti-Fc antibodies conjugated to HRP. (B) Band intensities were then quantified using the LiCore and band intensity values of each heparin binding band was compared to the corresponding expression bands. The results from the comparison are presented in the graph alongside errors bars for each experimental repeat. Statistical analysis of band intensities was performed using Dunnett's test, mean ± S.D. (n=3). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

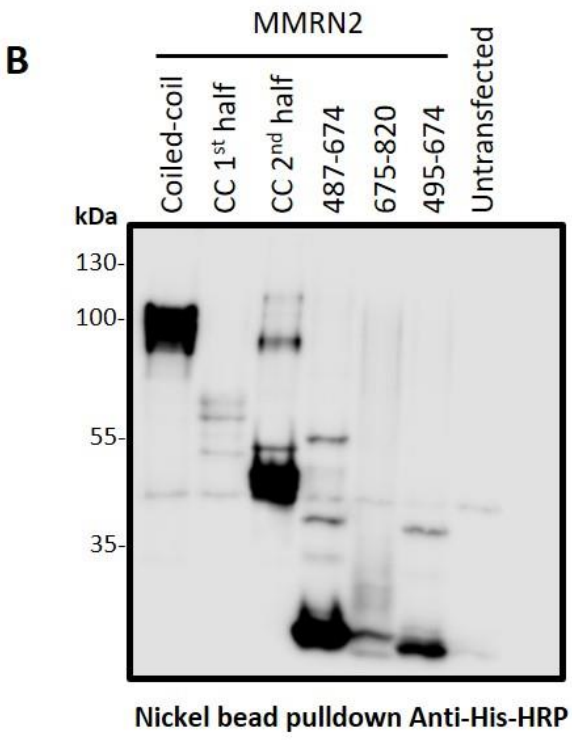
4.3. Heparin does not have a distinct binding site on MMRN2.

In the same study which showed that CLEC14A interacted with HSPGS, MMRN2 was identified in a liquid chromatography tandem mass spec screen. Furthermore, it has been documented that the interaction between MMRN2 and heparin sulphate inhibits the interactions between MMRN2 and VEGFA (Sandoval *et al.,* 2020; Colladel *et al.,* 2015). VEGFA binds to MMRN2 between the amino acids 137-336 which is located within the first half of the coiled-coil domain. Since heparin sulfate abrogates this interaction, it is likely that heparin binds to MMRN2 within this region. However, the possibility of heparin sulfate having more than one binding site on MMRN2 and potentially effecting the interaction between CLEC14A and MMRN2 cannot be ruled out.

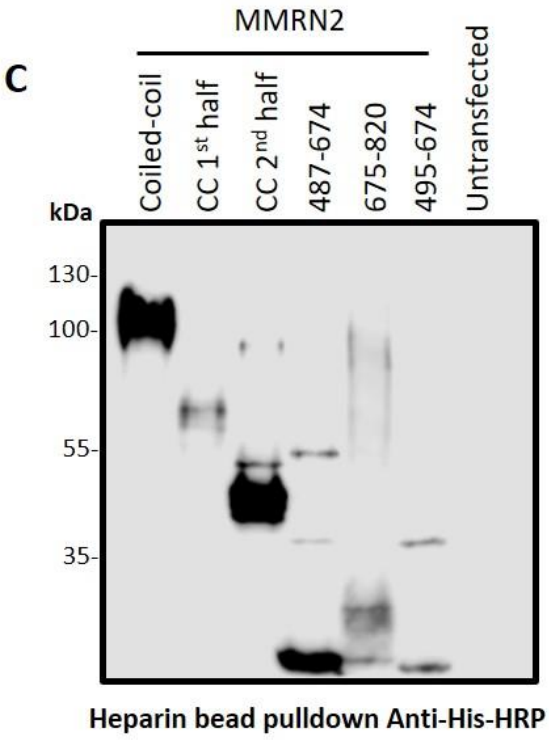
Constructs containing various fragmented forms of MMRN2 spanning the coiled-coil domain of MMRN2 were generated by Kabir Khan in the pHLAvitag3 vector which included a C-terminal His-tag. The fragments used in this study included the full length coiled-coil domain, the first half of the coiled-coil domain, the second half of the coiled-coil domain and three fragments which spanned the second half of the coiled-coil domain including 487-674, 675-820 and lastly the 495-674 fragment (Figure 4.2A). HEK293T cells were transfected with each construct to secreted the various MMRN2 fragments into the growth media. The conditioned media was collected and each fragment was subjected to Ni-NTA agarose bead pull-downs to study protein expression. Heparin bead pull-downs were also performed to study heparin binding. The results indicate that each fragment of MMRN2 was expressed as indicated by the clear bands at the correct predicted sizes for each MMRN2 fragment. The fragments were, however, differentially expressed (Figure 4.2B). The western blot from the heparin bead pull-down almost identically mirrored the loading blot thus indicating that each fragment bound to heparin and neither fragment appeared to have more or less binding compared to the

loading control (Figure 4.2C). Taken together, this data suggests that there is not one distinct binding region on MMRN2 which heparin binds.





Nickel bead pulldown Anti-His-HRP



Heparin bead pulldown Anti-His-HRP

Figure Legend on next page

C

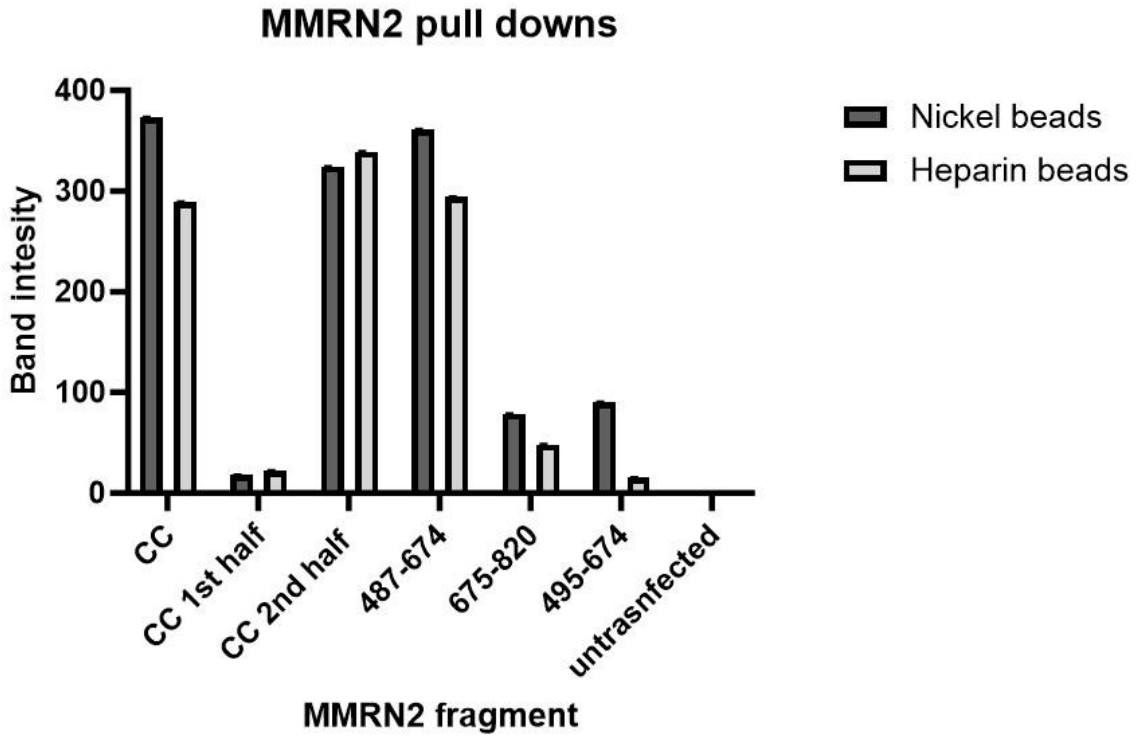


Figure 4.2. The full coiled-coil domain as well as fragments of the coiled-coil domain of MMRN2 bind to heparin beads. (A) The fragments of MMRN2 being tested included the coiled-coil domain as well as short fragments which cover this region. HEK293T cells were transfected with plasmids to produce and secrete various fragments of MMRN2 into the surrounding media. The conditioned media was then subjected to pull downs using either Ni-NTA beads or heparin beads. The beads were treated with reducing sample buffer, boiled and subjected to western blotting. (B) Membranes were probed with anti-His tag antibodies conjugated to HRP. (C) Band intensities from both blots were then quantified using LiCore and are presented side by side graph.

4.4. The CRT4 antibody blocks the interaction between CLEC14A and heparin

Five monoclonal antibodies raised in mice against CLEC14A (Cancer research technologies) have been extensively studied in the lab to reveal the antibodies CRT1, 3, 4 and 5 each bind to the CTLD of CLEC14A whilst the CRT2 antibody interacts with the SUSHI domain of CLEC14A. Both CRT4 and CRT5 appear to block the interaction between CLEC14A and MMRN2. Experimental studies using CRT4 inhibits tube formation and cell migration *in vitro* as well as decrease tumour size in mice harbouring Lewis lung carcinoma (LLC) tumours (Noy *et al.,* 2015).

HEK293T cells were transfected with pIgPP constructs containing either WT or R161A CLEC14A. The cells produced and secreted protein into the growth media which was collected and studied in the binding assay. Protein A and heparin bead pull downs were performed for WT and R161A CLEC14A to confirm protein expression and binding to the relevant beads. For the binding experiment, conditioned media from cells transfected with either WT or R161A CLEC14A-Fc or untransfected were treated with a 10-fold molar excess of either the CRT2, CRT3, CRT4 or mouse IgG isotype control (10 μg). Following this incubation, the proteins within the media were then tested for their capacity to bind to heparin by performing a heparin bead pull-down. The beads were treated with sample buffer and the contents were subjected to western blotting. The results indicated that following incubation with the CRT2 and CRT3 antibodies, the protein could still bind to heparin as indicated by the clear bands showing CLEC14A bound to heparin beads, this is the case for both the WT and R161A CLEC14A-Fc proteins. The CRT4 antibody however, which has been shown to block the interaction between CLEC14A and MMRN2, also blocks heparin from binding to CLEC14A (Figure 4.3A). This is the case in both WT and R161A CLEC14A and suggests that the specific binding site on CLEC14A where heparin binds is blocked by the CRT4 antibody.

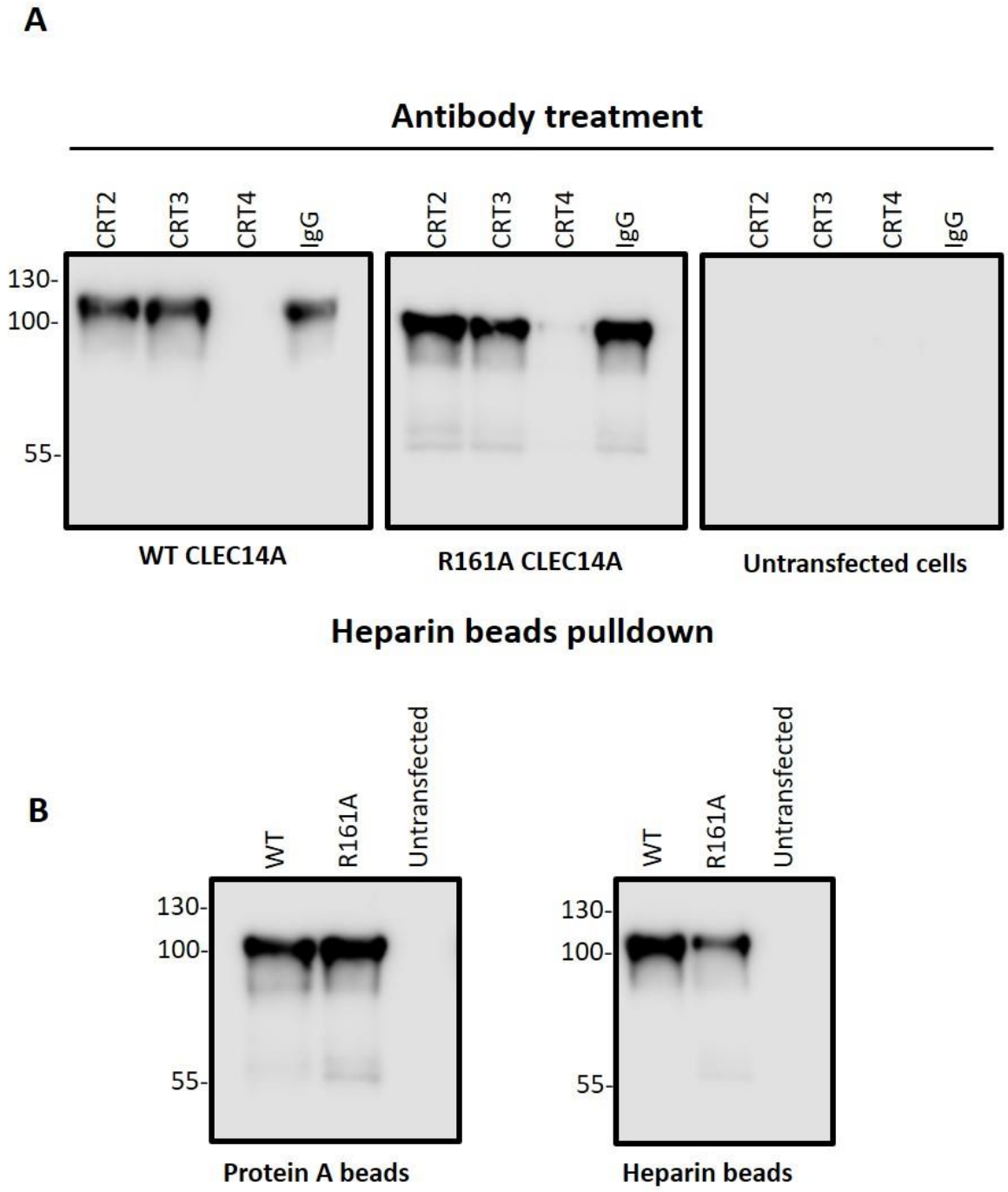


Figure 4.3. The CRT4 antibody blocks CLEC14A from binding to heparin beads. HEK293T cells were transfected with either WT or R161A CLEC14A to produce and secrete the protein into the surrounding media. The conditioned media was then collected and processed. (A) Conditioned media was incubated with either the CRT2, CRT3 or CRT4 antibodies followed by incubation with heparin beads. The bead were then treated with sample buffer and subjected to western blotting. (B) Conditioned media was also incubated with either protein A beads or heparin beads which were then treated with sample buffer and subjected to western blotting. Each membrane was incubated with the anti- human Fc antibody conjugated to HRP.

4.5 Discussion

The R161 residue of CLEC14A reduces binding to heparin beads and none of the CLEC14A mutants which blocked binding to MMRN2 effected the interaction with heparin. It was also found that the CRT4 antibody which has been shown to block the interaction between CLEC14A and MMRN2 also blocked binding between CLEC14A and heparin beads.

This study revealed that the R161A CLEC14A mutant significantly reduced binding to heparin-coated agarose and Sepharose beads but did not completely eliminate the interaction. This finding supports the literature as it confirms the R161 residue to be involved in the interaction however the finding also opposes the same study from Sandoval *et al.,* which indicated in an ELISA that the R161A CLEC14A mutant completely eliminates the interaction. Our finding suggest it is more likely that other residues are also involved in the interaction although the other positively charged residues within this region of the CTLD (R141, K138 and R165) have already been studied in the literature and have shown to not be involved the CLEC14A-heparin interaction. The carbohydrate recognition domain often also includes parts of the "loop-in-loop" structure which, in the case of CLEC14A, is close to where the R161 residue is situated (Veldhuizen *et al.,* 2011).

The key distinction between both studies is the type of heparin-based reagent used: our study employed heparin-coated agarose or Sepharose beads. whereas Sandoval *et al.,* used porcine mucosal heparin in their ELISA. This difference in heparin reagent may account for the discrepancies in CLEC14A binding between each study. In both instances heparan sulfate was substituted with heparin. Due to structural complexity and purification difficulties heparan sulfate is expensive and may lead to inconsistency in experimental outcomes due to batch

variability (Shriver *et al.,* 2012). The Sandova*l et al.,* study confirmed that CLEC14A does bind to endothelial heparan sulfate however the subsequent experiments used heparin.

The S137A, T139A and R141A mutants which decrease MMRN2 binding did not affect heparin binding nor does the R161A mutant effect MMRN2 binding as shown in Chapter 3 (Figure 3.2). Taken together these data confirm that the binding sites for both these binding partners on the CTLD are separate. It has previously been shown in the literature that R141A mutant did not weaken heparin binding, a result which was confirmed in this study (Sandoval *et al.,* 2020). As of yet it is not clear if MMRN2 and heparin can bind to CLEC14A simultaneously, a binding experiment using heparin coated beads incubated with MMRN2 495-674 and CLEC14A-Fc would provide insights in this regard.

The CRT4 antibody blocks the interaction between heparin and CLEC14A. Previously it has been shown that the CRT4 antibody blocks the interaction between CLEC14A and MMRN2 and this resulted in a significant phenotype of reduced tube formation and cell migration *in vitro* as well as decreased tumour size *in vivo*. When these experiments were performed it was deduced that these results were due to blockage of the CLEC14A and MMRN2 interaction (Noy *et al.,* 2015; Khan *et al.,* 2017). The new data presented in this part of the project regarding the CRT4 antibody raises new questions that the observed phenotypes following CRT4 treatment may be due to blockage of the interaction between CLEC14A and heparan sulfate. It also raises the question of if the CRT4 antibody can block other CLEC14A interactions.

Of each CRT antibody developed CRT1, 4 and 5 blocked the interaction between CLEC14A and MRMN2. Antibodies CRT2 and CRT3 were shown to not effect this interaction and the binding epitope were vaguely mapped to the sushi domain for CRT2 and CTLD for CRT3. Whilst the

CRT3 antibody does not block the interaction between CLEC14A and MMRN2, it has been shown that reduce tube formation and cell migration following treatment on HUVECs (Puja Lodhia PhD thesis, 2016). The CRT3 antibody did not affect the interaction between CLEC14A and heparin nor did it appear that the R161 residue formed part of the binding epitope. This was also the case for the CRT2 antibody. This indicates that the phenotype observed following CRT3 treatment is due to another reason, potentially blocking the interaction of CLEC14A with Heat shock protein 70-1A.

The data presented in this chapter expands upon current literature concerning CLEC14A and heparin, while also identifying gaps in our understanding. A logical next step would be to investigate the impact of CRT antibodies, particularly CRT4 and CRT3, on the interaction between CLEC14A and heat shock protein 70-1A. This investigation would clarify whether the observed effects of CRT3 treatment on HUVECs are mediated by this interaction and could enhance our comprehension of CRT4 treatment effects. Another significant gap in the literature pertains to the lack of functional studies evaluating the importance of the CLEC14A-heparan sulfate interaction. Both CLEC14A and heparan sulfate have been shown to play a critical role in angiogenesis. The functional significance of the interaction between CLEC14A and heparan sulfate is yet to be explored, one avenue would involve the use of HUVECs knocked down for WT CLEC14A and expressing the R161A mutant. The HUVECs can then be assessed in a range of *in vitro* assays such as tube formation and cell migration assays. Differences in phenotypes between WT CLEC14A, R161A CLEC14A and CLEC14A knock down HUVECs would provide valuable insights into the function of this interaction.

CHAPTER 5: Development of a NanoBRET assay to interrogate the CLEC14A-MMRN2 interaction

CHAPTER 5: Development of a NanoBRET assay to interrogate the CLEC14A-MMRN2 interaction

5.1 Introduction

The CLEC14A-MMRN2 interaction has significant potential as a target for anti-angiogenic therapy as abrogation of the interaction has been illustrated to affect angiogenesis *in vitro* as well as decrease tumour size *in vivo* (Noy *et al.,* 2015). When considering a protein interaction as a drug target, the molecular pharmacology of the interaction must be understood; this includes binding affinity, avidity and kinetics parameters such as receptor internalisation. Previous work performed in the Bicknell lab has interrogated the binding kinetics of CLEC14A and MMRN2 using surface plasmon resonance (SPR) and revealed the CLEC14A-MMRN2 interaction to be relatively stable. However, the MMRN2 fragment was so strongly bound to CLEC14A that the off rate was too slow to enable an accurate measurement of the binding affinity (Kabir Khan PhD thesis, 2016). In light of this, the aim of this part of the study was to quantify the binding affinity between CLEC14A and MMRN2 using Nano Bioluminescence resonance energy transfer (NanoBRET) and better characterise the kinetics of the interaction.

5.1.1 The NanoBRET assay

NanoBRET is an assay used to study the binding and kinetics of protein interactions in real-time using live cells by employing a donor luciferase and a complementary acceptor fluorophore. Proteins of interest were either tagged with a luciferase enzyme such as NanoLuciferase (NLuc), a small 19.1kDa protein which acted as the energy donor, or fused to an acceptor fluorophore. Initially derived from deep sea shrimp, the luciferase enzyme has been mutated and modified to optimise luminescent output. Moreover, NLuc has increased

stability and is smaller in size when compared to previous luciferase-based enzymes (England *et al.,* 2016). Quantitative data from the NanoBRET assay can inform key binding characteristics such as binding affinity and the effects of drugs or mutants on a given interaction. The NanoBRET assay involves two proteins which will be referred to as protein A and protein B for clarity in the subsequent explanation. Protein A is tagged with NLuc which oxidises its substrate furimazine which produces luminescence at a wavelength of 470nm. Protein B is tagged with an energy acceptor molecule in the form of a fluorescently labelled tag (i.e., SNAP- CLIP- or HALO-tag) or a fluorophore. When protein A and protein B are within close proximity (<10nm), the luciferase energy donor, NLuc will transfer energy to the complementary donor fluorophore, resulting in a detectable acceptor emission signal at a longer wavelength. Within the assay, both the light emission of the NLuc and fluorophore can be detected, when in close proximity at their respective wavelengths. Energy is absorbed by the acceptor from the donor, the signal of luminescence from the donor protein reduces and the fluorescent signal from the acceptor protein increases. The ratio between both signals is the BRET ratio (Figure 5.1A).

NanoLuc Binary Technology (NanoBiT) is whereby the NanoLuciferase is split into two NanoBiT fragments coined SmBiT (small bit) and LgBiT (large bit) (Dixon *et al.,* 2016). As purified proteins, these subunits have a relatively weak affinity for one another of 190μM, however, when the fragment re-complements, the luciferase retains a bright signal (Cooley *et al.,* 2020). In this instance protein A is tagged with LgBiT and protein B is tagged with SmBiT. When the two fragments bind, the full NLuc reforms which, in the presence of furimazine, results in detectable luminescence peaking at 470nm (Figure 5.1B). The NanoBiT set-up does not involve a fluorescently labelled protein and therefore there is no bioluminescent energy transfer, however, it is theoretically possible to have a three-way system to study three

proteins in close proximity by adding of a third fluorescently labelled protein and measuring subsequent BRET.

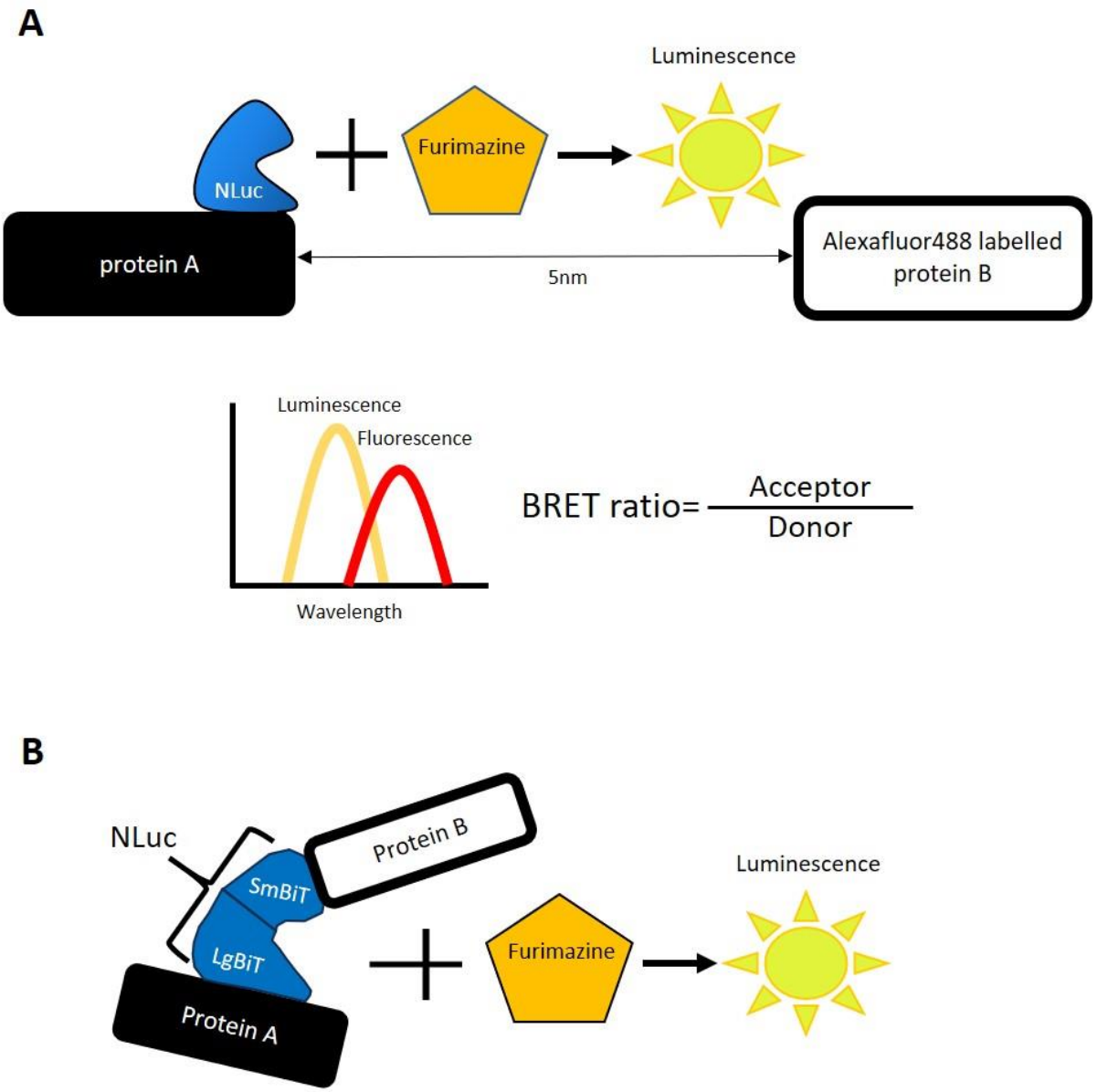


Figure 5.1. Schematic illustrations of a NanoBRET experiment as well as the NanoBiT experiment. (A) The two proteins being studied are tagged with Nanoluciferase (Nluc) or fluorescently labelled with Alexafluor 488. Following addition of the substrate furimazine, the Nluc tagged protein emits luminescence. When the 2 proteins are in 5nm proximity, the luminescence is absorbed by the fluorescently labelled protein and reemitted at a higher wavelength. The wavelength of the acceptor (fluorescent protein) is divided by the wavelength of the luminescence to give the BRET ratio. (B) The NanoBiT experiment involves 2 proteins tagged with either LgBiT or SmBiT each of which are fragments of Nluc and when bound together form Nluc. When the proteins are close enough for both tags to bind and form Nluc, following addition of the substrate furimazine, luminescence is emitted.

5.2. Construct design

In order to investigate CLEC14A and MMRN2 using NanoBRET and NanoBiT assays, a range of constructs were designed to include the NanoBRET tags (NLuc & SNAP) and NanoBiT tags (LgBiT and SmBiT) fused to either CLEC14A or MMRN2. Plasmids containing NanoBRET and NanoBiT tags were provided by Laura Kilpatrick (University of Nottingham).

Initially, DNA fragments encoding the proteins of interest were cloned into the pcDNA3.1 plasmid to include the various NanoBRET and NanoBiT tags fused to the N-terminus of either CLEC14A or MMRN2. CLEC14A localises to the cell surface, therefore, full length CLEC14A was used, the plasmids included a signal peptide therefore the native signal peptide of CLEC14A was excluded. MMRN2 was produced as a purified protein, the MMRN2 495-674 fragment was used as this is the shortest stably expressed fragment which binds CLEC14A. MMRN2 495-674 contained a C-terminal His-tag which was used for purification. Following the initial NanoBRET attempts several additional constructs were designed to optimise the set up by adjusting the placement of the NanoBRET tags (Figure 5.2).

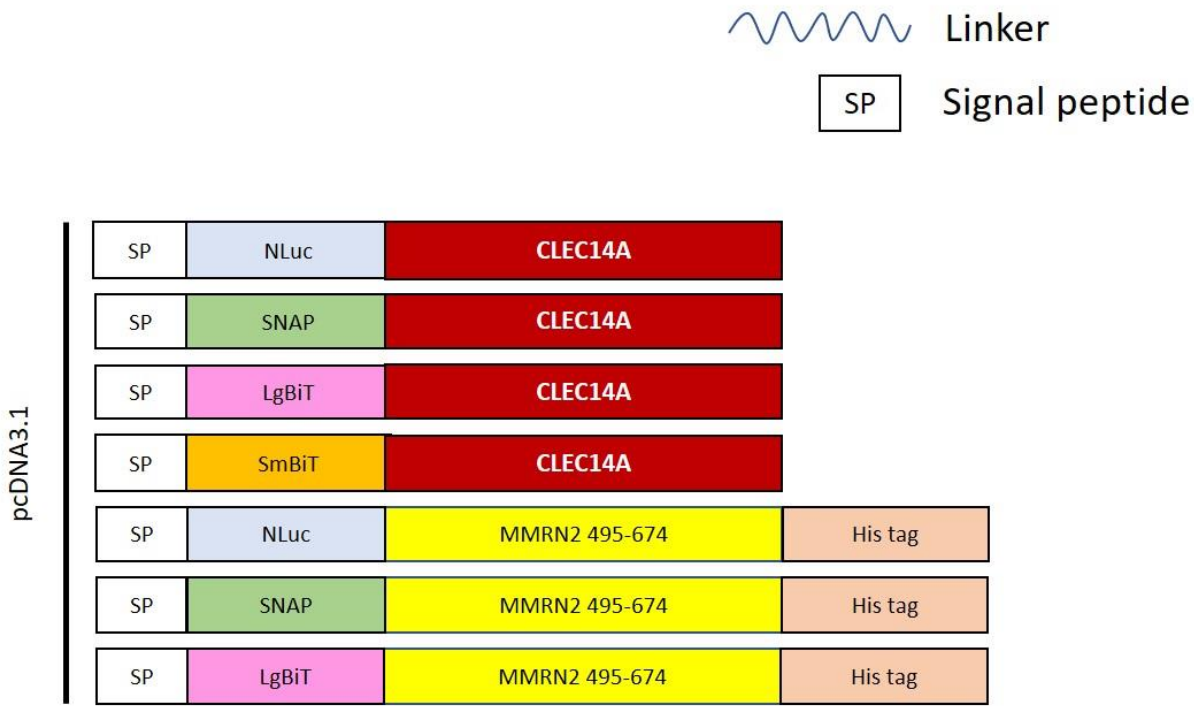


Figure 5.2. Schematic illustration of the contructs designed for this NanoBRET study. Plasmids were constructed wherein either WT or mutant CLEC14A or MMRN2 495-674 were tagged with the various NanoBRET or NanoBiT tags. Each plasmid included a signal peptide followed by the NanoBRET and NanoBiT tagged fused to the N-terminus of the protein. These were constructed in the pcDNA3.1 vector.

5.3. Expressing NanoBRET and NanoBiT tagged CLEC14A and MMRN2

5.3.1 Expressing NanoBRET and NanoBiT tagged WT CLEC14A

Once all plasmids encoding the different CLEC14A or MMRN2 constructs, as detailed in Figure 5.2, were cloned, their expression and binding capacities were tested. This was to ensure the BRET tags did not disturb either MMRN2 or CLEC14A folding and/or trafficking. To verify this, lysates from HEK293T cells transfected with plasmids containing tagged CLEC14A were subjected to western and far western blotting. Untagged CLEC14A which has previously been shown to bind to MMRN2 was used as a positive control, meanwhile lysates from untransfected HEK293T cells were used as a negative control. The results indicate that WT CLEC14A was expressed, as indicated by two bands present in the western blot (Figure 5.3A). This two-band pattern previously observed indicates two alternative glycoforms of CLEC14A. Each NanoBRET and NanoBiT tagged form of CLEC14A also expressed with varying banding patterns and intensities. Both NLuc and SmBiT-tagged CLEC14A indicated faint double bands close together at approximately 120kDa and 90kDa respectively, whilst SNAP and LgBiT-tagged CLEC14A indicated two more intense bands indicating the two alternative glycoforms of the protein. In each case the bands for each tagged-CLEC14A ran at a higher molecular weight than untagged CLEC14A (Figure 5.3A).

The corresponding far western revealed that the untagged CLEC14A positive control bound to full length MMRN2 (Figure 5.3B). The NLuc, SNAP and SmBiT-tagged CLEC14A each also bound to MMRN2 confirming addition of each tag did not affect the interaction. However, the LgBiT-tagged CLEC14A did not bind to MMRN2 indicating that the tag may have caused steric hindrance or disrupted the structure of the CTLD (Figure 5.3B). In both blots the untransfected negative control did not reveal any banding.

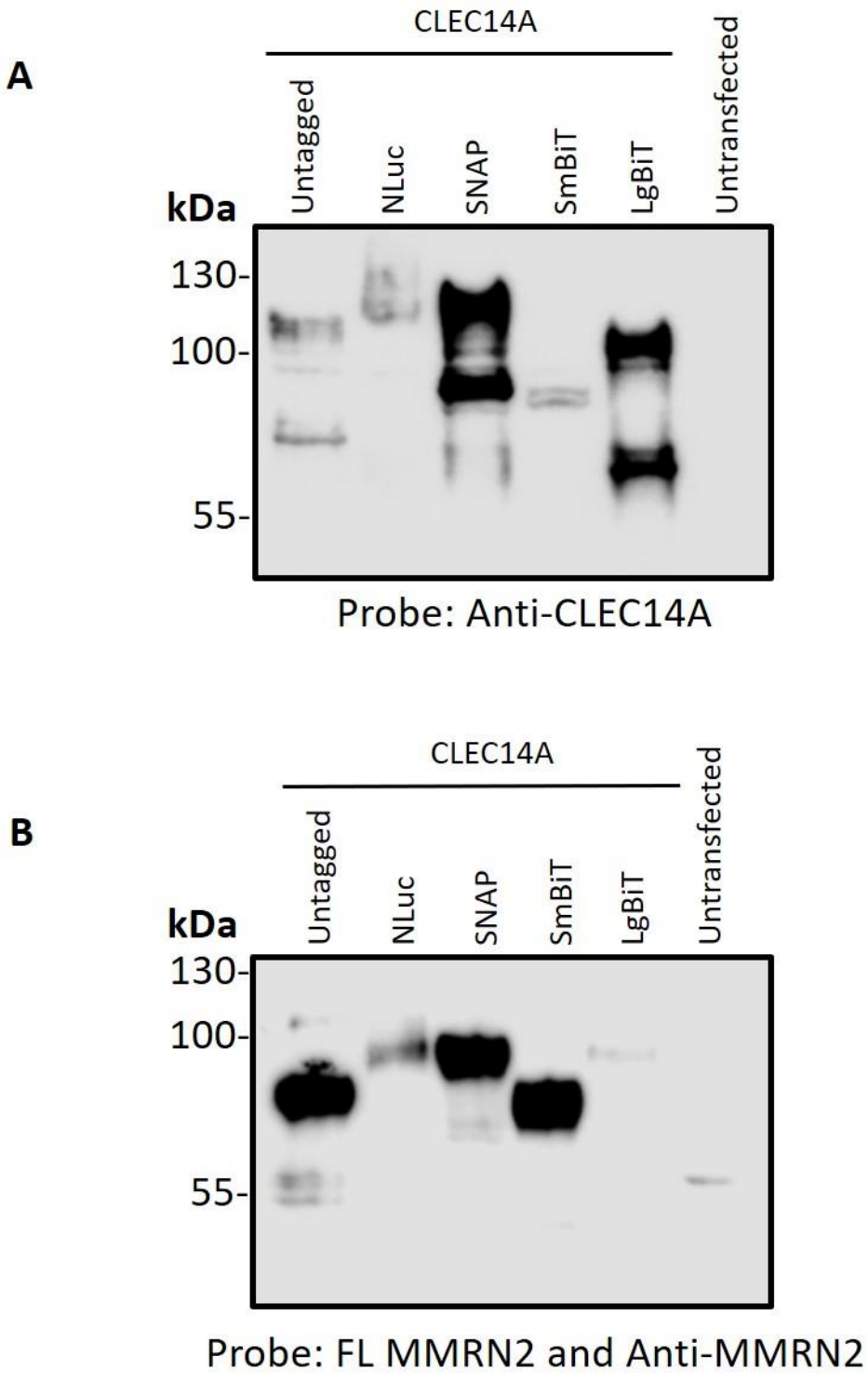


Figure 5.3. NLuc, SNAP and SmBiT CLEC14A each bound to MMRN2 in a far western. HEK293T cells were transfected with the plasmids constructed to express CLEC14A tagged with either NLuc, SNAP, LgBiT or SmBiT. Lysates were then taken and subjected to either western blotting or far western blotting. (A) Western blots were run in reducing conditions and then probed with anti-CLEC14A antibodies followed by incubation with an anti-sheep antibody conjugated to HRP. (B) Far western blots were run in non-reducing conditions and probed with cellular lysate containing FL MMRN2 followed by an anti-MMRN2 antibody before incubation with an anti-rabbit antibody conjugated to HRP. Lysates taken from untransfected cells were used as a negative control.

Since the NanoBRET and NanoBiT experiments would involve studying the CLEC14A-MMRN2 interaction at the cell surface, it was important to ensure the addition of each tag did not affect the capacity of tagged-CLEC14A to traffic to the cell surface. Therefore, flow cytometry was performed to study cell surface expression with untagged CLEC14A used as a positive control. Antibodies against CLEC14A, CRT2 and CRT4, were employed for this study. The CRT4 antibody epitope partly converges with the MMRN2 binding site on CLEC14A via the R141 residue as shown in chapter 3. Therefore, if CRT4 recognises the tagged forms of CLEC14A, this would confirm that the tags do not block the CLEC14A-MMRN2 interaction. Conversely, if the CRT4 antibody did not recognise tagged-CLEC14A this might not solely be down to the tag covering the binding site, but could potentially be due to protein misfolding following tag addition. Therefore, the CRT2 antibody which has a binding epitope distinct from the CTLD was also used.

Flow cytometry analyses revealed that the untagged CLEC14A positive control was expressed at the cell surface and recognised by both the CRT2 and CRT4 antibodies (Figure 5.4A). The results also indicated that NLuc-, SNAP- and SmBiT-tagged CLEC14A were also expressed at the cell surface and were recognised by both the CRT2 and CRT4 antibodies. The LgBiT-tagged CLEC14A had reduced expression compared to the other proteins and the untransfected negative control did not indicate any peaks above the isotype control (figure 5.4A). The mean fluorescence intensities (MFIs) were documented and normalised to untagged CLEC14A as each protein appeared to express to varying degrees. For both the CRT2 and CRT4 signals, each tagged form of CLEC14A had decreased MFIs compared to untagged CLEC14A. In the case of LgBiT-tagged CLEC14A the MFI decreased to levels similar to that of the untransfected negative control indicating that little to no protein was expressed at the cell surface (figure 5.4B).

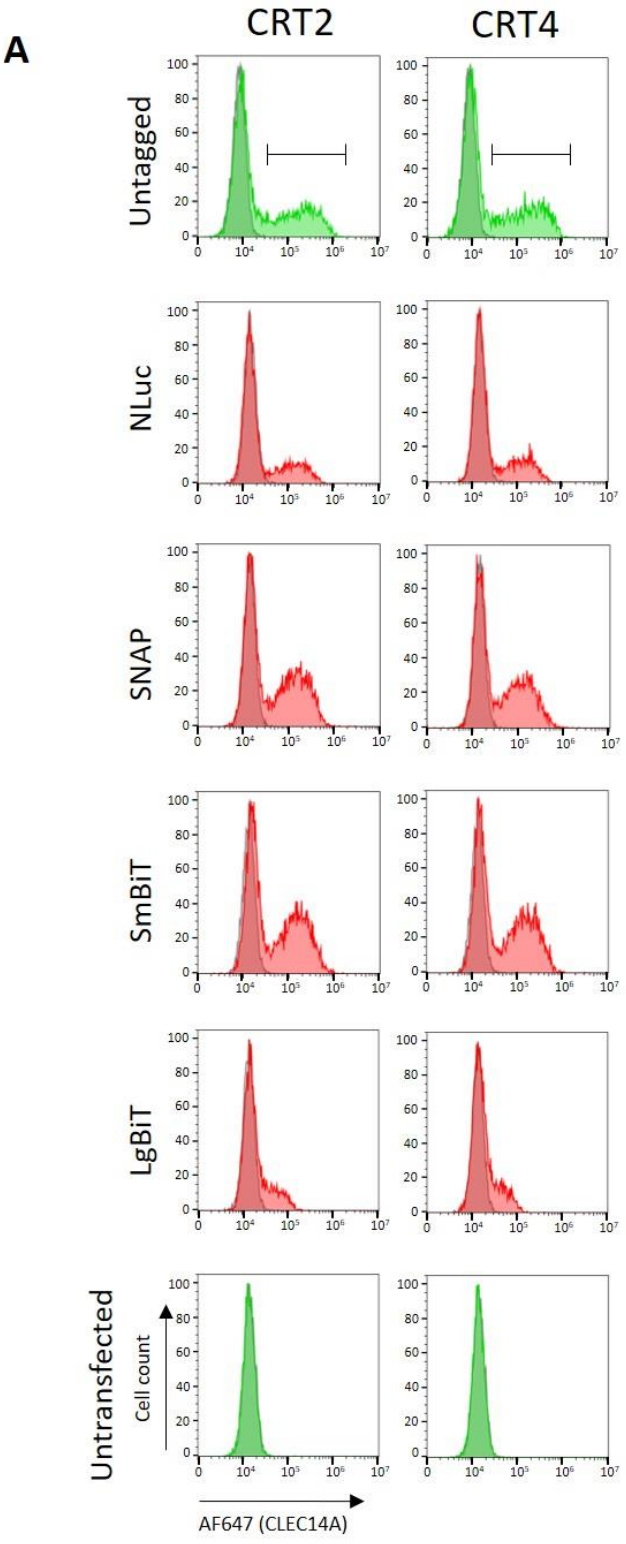


Figure Legend on next page

B

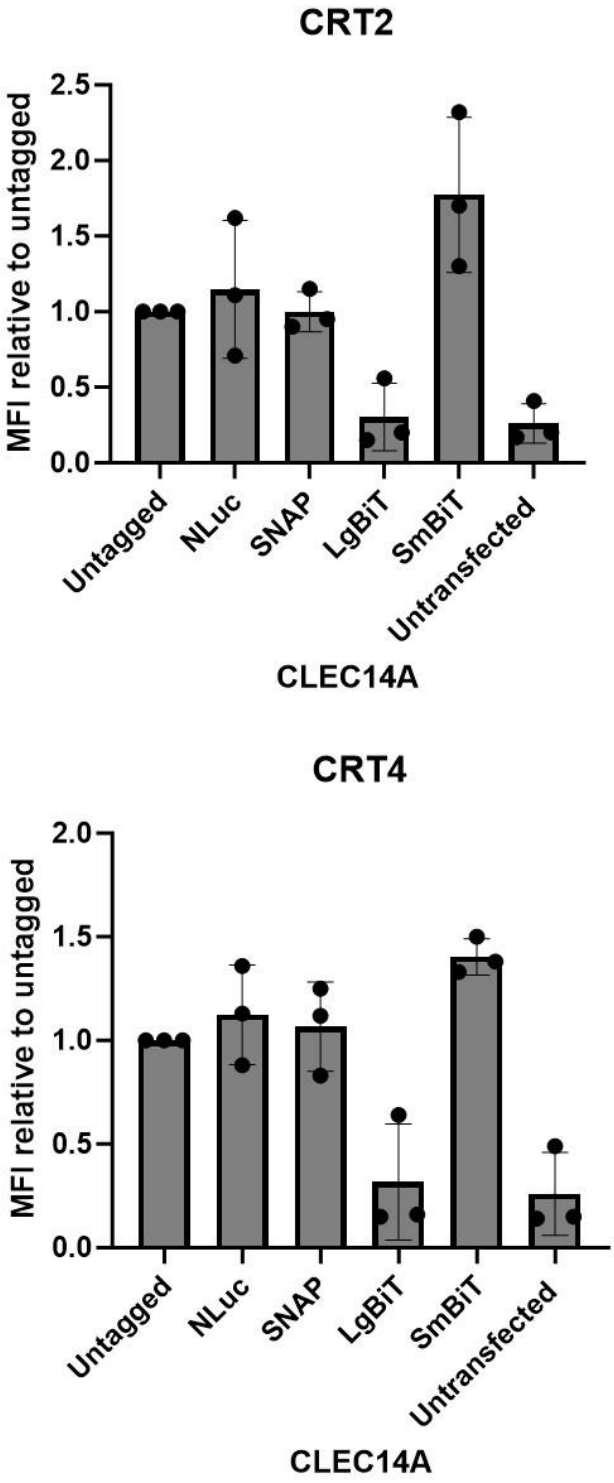


Figure 5.4. Nluc, SNAP and SmBiT tagged CLEC14A expressed at the cell surface. HEK293T cells were transfected with plasmids to express either untagged or various tagged forms of CLEC14A. Cells were incubated with either the CRT2 or CRT4 primary antibodies followed by incubation with anti-mouse IgG Alexafluor647 conjugated antibody. (A) Flow cytometry was performed, and the histograms presented show fluorescence in the APC channel gated on live cells. The histograms for untagged CLEC14A and the untransfected cells negative controls are highlighted in green, whereas the tagged CLEC14A are shown in red. Signal from the control antibody is shown in grey. Below each histogram is shown the percentage of cells expressing the protein of interest. This figure is representative of data produced from three experimental replicates. (B) The MFIs from the three experimental replicates were normalised to the untagged CLEC14 and are presented on the graphs. Values shown are mean ± S.D. (n=3).

5.3.2 Expressing NanoBRET and NanoBiT tagged mutant CLEC14A

Previously in this project it was shown that the S137, T139 and R141 residues of CLEC14A directly interact with MMRN2 (section 3.3). By assessing the changes in binding affinity between each mutant form of CLEC14A and MMRN2, the relative contribution of each of these residues to the interaction could be determined. To do so, each CLEC14A mutant (where these residues were switched to alanine) was cloned into the validated NanoBRET and NanoBiT constructs. These mutants included the triple S137-T139-R141A mutant, the double S137-T139A, S137-R141A and T139-R141A mutants and the single S137A, T139A and R141A mutants. Alongside this, the R161A point mutant was also cloned with the NanoBRET tags as it had previously been shown to decrease binding to heparin, but did not affect binding to MMRN2 (section 3.2). Given that no detectable protein expression of the LgBiT-tagged WT CLEC14A was observed, each mutant was tagged at the N-terminus with NLuc, SNAP or SmBiT. To test the expression of the proteins, western blotting was performed using cellular lysates from HEK293T cells transfected with the various constructs. The positive control used for each tagged mutant was the corresponding tagged WT CLEC14A as these constructs were shown to express previously in a western blot (Figure 5.3). Far western blotting was not performed as most mutants would not be expected to bind to MMRN2.

The NLuc-tagged CLEC14A blots revealed that the WT positive control was expressed. Additionally, each MMRN2-blocking mutant was also expressed. Each protein revealed a similar banding pattern when compared with NLuc-tagged WT CLEC14A, however, the mutants appeared to be differentially expressed. The R161A, T139-R141A and R141A mutants expressed at relatively similar levels when compared to WT. Conversely, the S137-T139A, S137A and T139A mutants each expressed to a higher degree compared to WT. The S137-T139-R141A and S137-R141A mutants expressed at a lesser extent compared to WT. The

SNAP-tagged constructs revealed a more defined two band pattern compared with the NLuc blot. SNAP-tagged WT CLEC14A was expressed and each SNAP-tagged mutant followed the same pattern but were differentially expressed. The R161A and S137-T139A mutants displayed only the single lower molecular weight band which previously indicated altered glycosylation. The T139A mutant displayed a faint higher molecular weight band and a more defined lower band. Each of the other mutants; S137-T139-R141A, S137-R141A, T139-R141A, S137A and R141A expressed both a higher and lower molecular weight band at relatively similar intensities. SmBiT-tagged WT CLEC14A also displayed a two-band pattern which was followed by the other SmBiT-tagged mutants. Both the S137-T139-R141A and S137-R141A mutants indicated faint higher molecular weight bands whilst the other mutants were expressed at relatively the same level and expressed both the higher and lower molecular weight bands at similar intensities. In each blot the untransfected negative controls did not show any bands (Figure 5.5). Between each mutant there does not appear to be a clear pattern regarding expression although the S137-R141A mutant consistently displayed decrease band intensities of the higher molecular weight band.

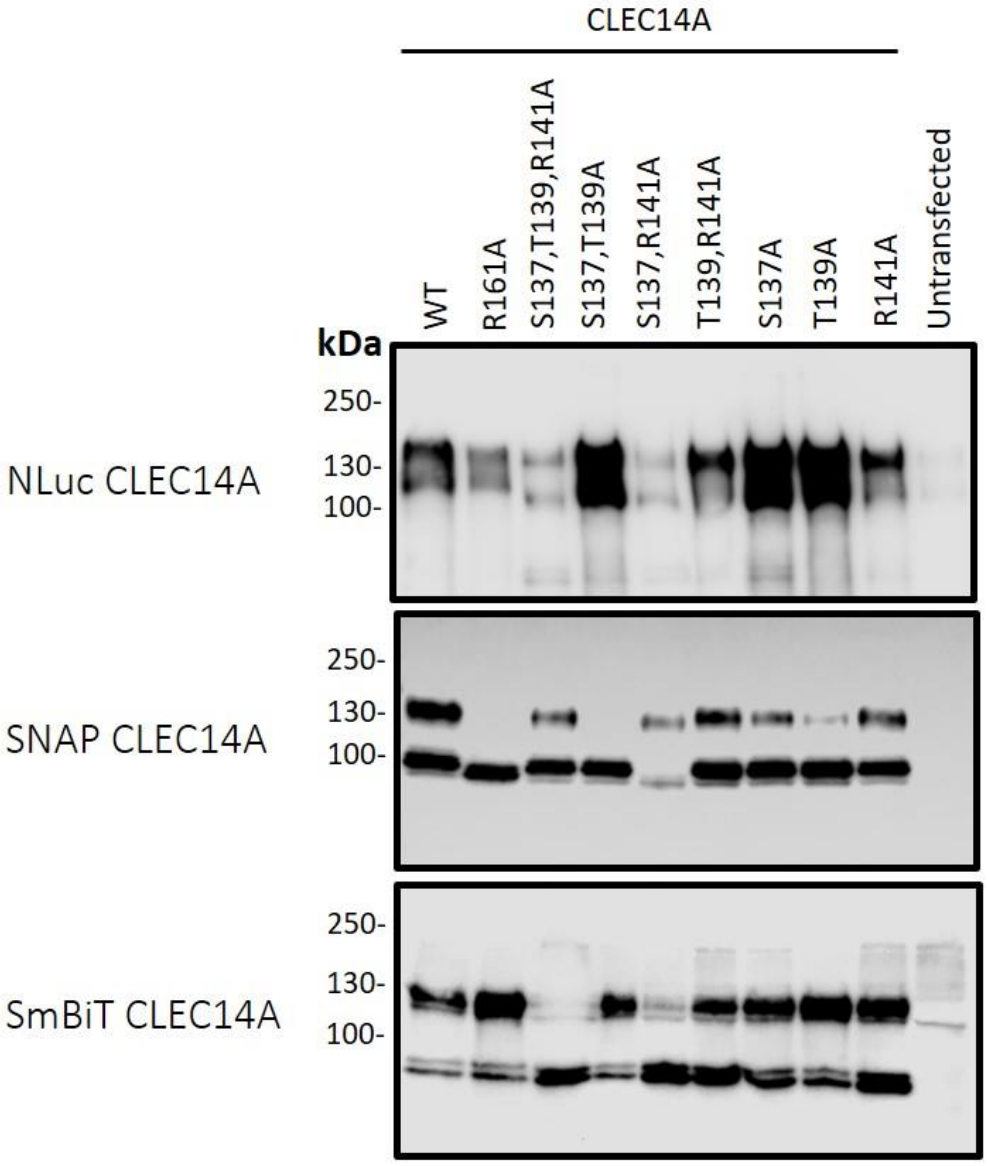


Figure 5.5. NanoBRET and NanoBiT tagged mutants differentially express. HEK293T cells were transfected with plasmids to express mutant forms of CLEC14A tagged with either NLuc, SNAP or SmBiT. Lysates were collected from the cells and subjected to western blotting. The membranes were each probed with an anti-CLEC14A antibody followed by incubation with a secondary anti-sheep antibody conjugated to HRP.

Given the alternative expression patterns of each mutant it was important to determine whether this affected cell surface expression as this is necessary for the NanoBRET and NanoBiT assays. Therefore, flow cytometry analysis was performed on HEK293T cells transfected with the various mutant plasmids. The CRT2 antibody alone was used to detect the proteins at the cell surface due to its binding epitope distinct from the MMRN2-binding CTLD. Additionally, the CRT4 antibody was not used as the R141 residue, which is mutated in most of the mutants, is an integral part of its binding epitope. For each flow cytometry experiment the WT-tagged form of the protein was used as a positive control, as each tagged form had previously been shown to express at the cell surface (Figure 5.4). For each analysis performed, the MFIs of the positive peak (as determined by the isotype control) were determined and normalised relative to the tagged WT CLEC14A positive control. This was to ensure the MFIs calculated represented signal from cells expressing the protein of interest only.

The Nluc-, SNAP- and SmBiT- tagged CLEC14A mutants flow cytometry analysis revealed that each protein was recognised by the CRT2 antibody and was therefore successfully expressed at the cell surface albeit at varying degrees (Figures 5.6, 5.7 and 5.8). The varying MFI reading for each tagged protein did not completely correspond to the band intensities of the western blot (Figure 5.5). Overall, each CLEC14A mutant with the various tags were detected at the cell surface which is sufficient for this assay as very little amounts of donor (NLuc) or acceptor (fluorescent molecule) is required to elicit a detectable and robust BRET signal. Therefore, regarding CLEC14A, no further optimisation would be required to increase expression.

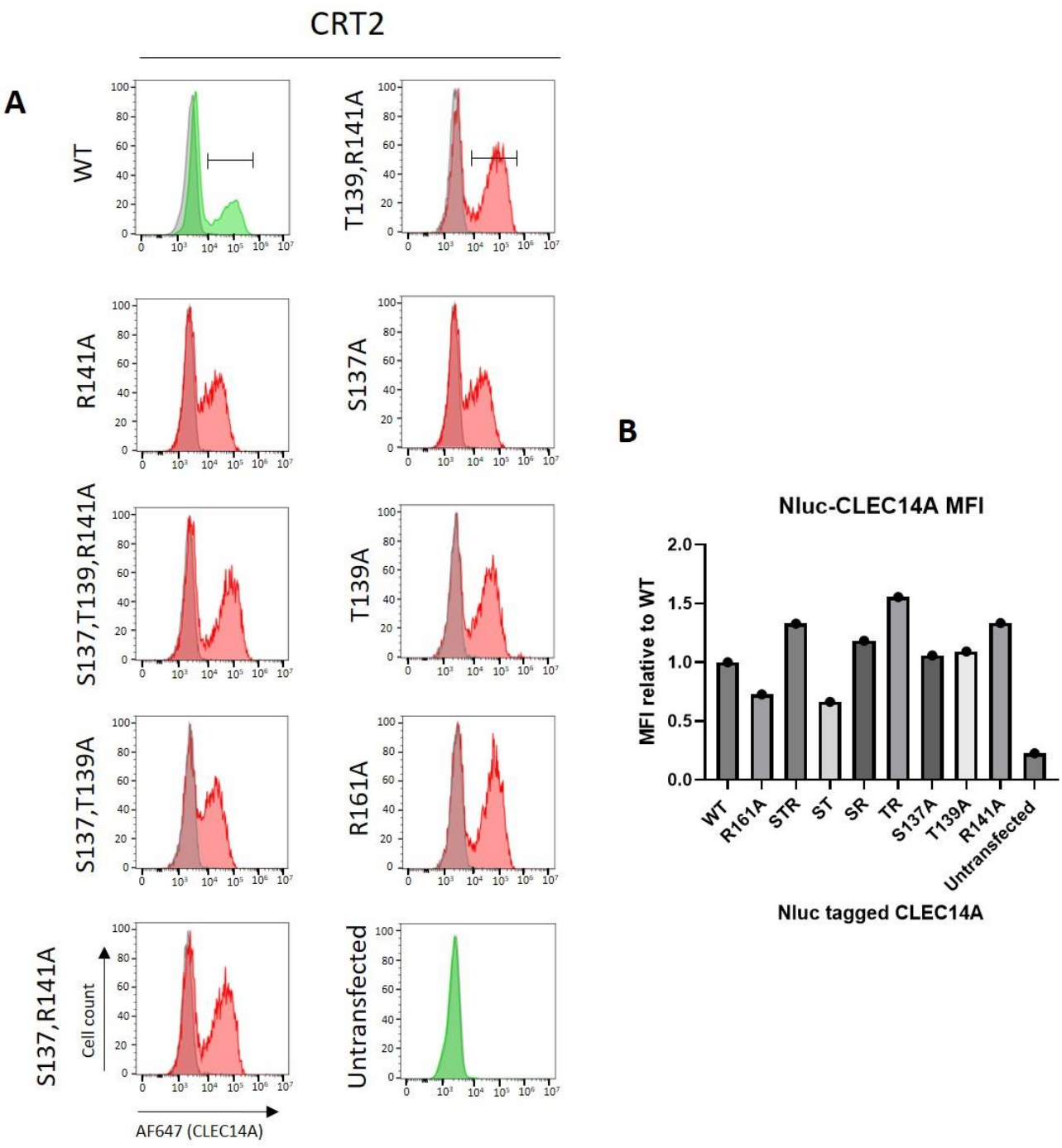


Figure 5.6. NLuc-tagged CLEC14A mutants were expressed on the cell surface. HEK293T cells were transfected with the CLEC14A mutants fused with the NLuc tag. Cells were incubated with the CRT2 primary antibody followed by incubation with an anti-mouse IgG Alexafluor647 conjugated antibody. (A) Flow cytometry was performed, and the histograms presented show fluorescence in the APC channel gated on live cells. The histograms for both WT CLEC14A and untransfected cells are highlighted in green, whereas each mutant is shown in red. Signal from the control antibody is shown in grey. (B) MFIs were normalised to the NLuc-tagged WT CLEC14A and is presented in a graph. This experiment was performed once.

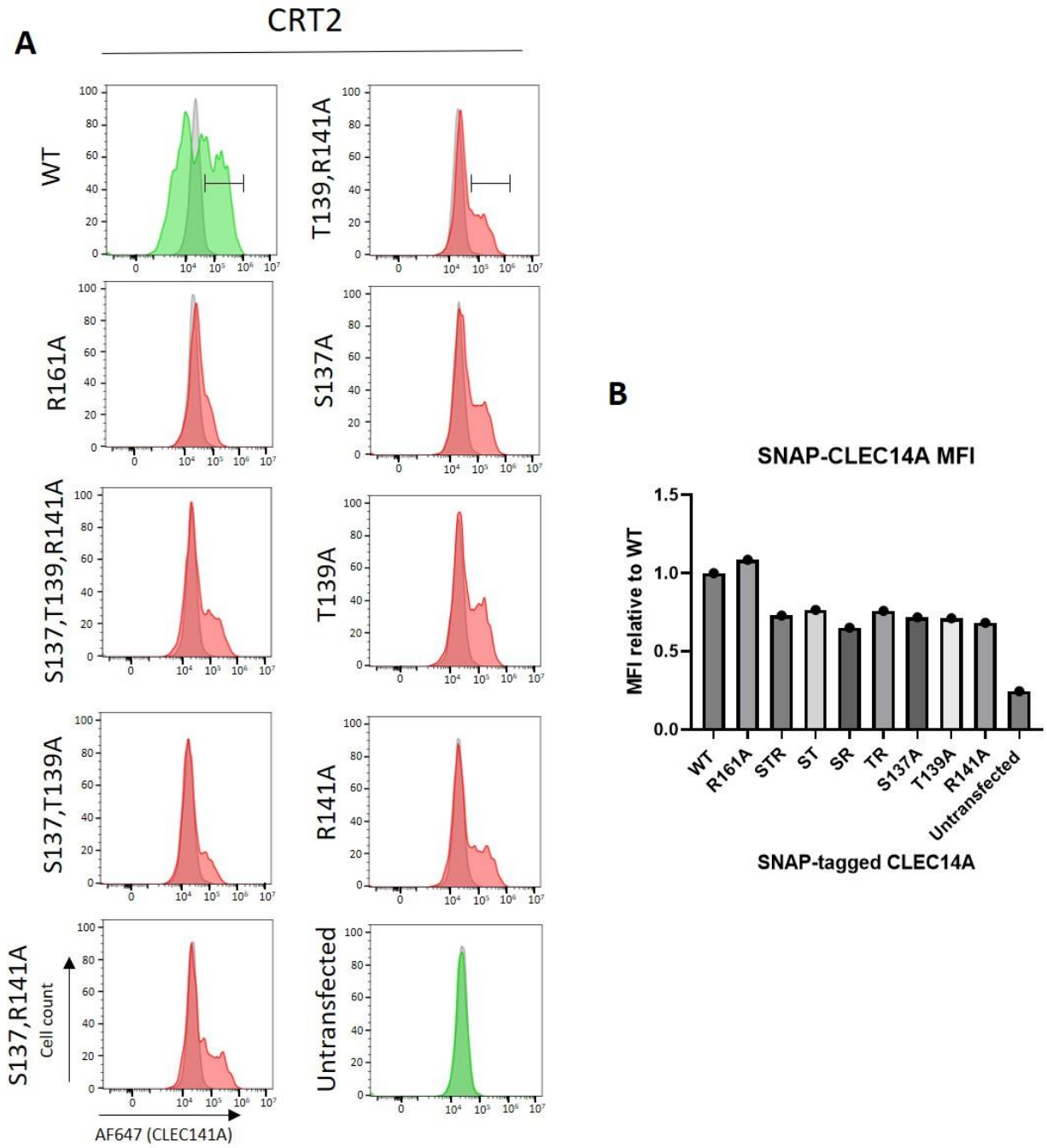


Figure 5.7. SNAP-tagged CLEC14A mutants were expressed on the cell surface. HEK293T cells were transfected with the CLEC14A mutants fused with the SNAP tag. Cells were incubated with the CRT2 primary antibody followed by incubation with an anti-mouse IgG Alexafluor647 conjugated antibody. (A) Flow cytometry was performed, and the histograms presented show fluorescence in the APC channel gated on live cells. The histograms for both WT CLEC14A and untransfected cells are highlighted in green, whereas each mutant is shown in red. Signal from the control antibody is shown in grey. (B) MFIs were normalised to the SNAP-tagged WT CLEC14A and is presented in a graph. This experiment was performed once.

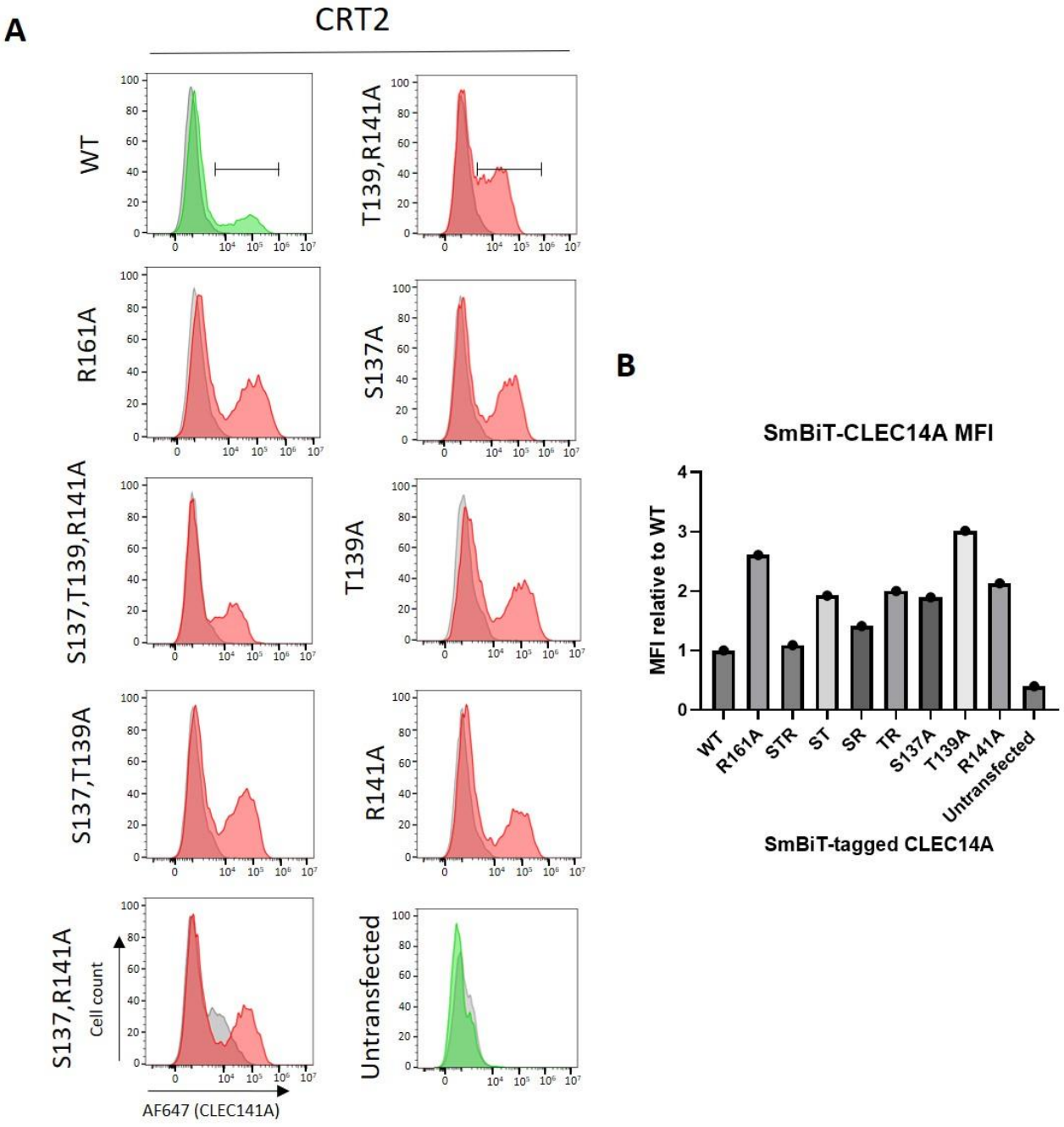


Figure 5.8. SmBiT-tagged CLEC14A mutants were expressed on the cell surface. HEK293T cells were transfected with the CLEC14A mutants fused with the SmBiT tag. Cells were incubated with the CRT2 primary antibody followed by incubation with an anti-mouse IgG Alexafluor647 conjugated antibody. (A) Flow cytometry was performed, and the histograms presented show fluorescence in the APC channel gated on live cells. The histograms for both WT CLEC14A and untransfected cells are highlighted in green, whereas each mutant is shown in red. Signal from the control antibody is shown in grey. (B) MFIs were normalised to the SmBiT-tagged WT CLEC14A and is presented in a graph. This experiment was performed once.

5.3.3. Expressing NanoBRET and NanoBiT tagged MMRN2 495-674

The MMRN2 495-674 fragment is routinely used when studying the CLEC14A-MMRN2 interaction as it is the shortest stable fragment of MMRN2 that binds to CLEC14A (Khan *et al.,* 2017). Therefore, constructs designed for NanoBRET used this fragment. As already illustrated in Figure 5.2, NanoBRET and NanoBiT tags were constructed onto the N-terminus of the MMRN2 495-674 fragment. Since the LgBiT-tagged CLEC14A constructs did not express at the cell surface (Figure 5.9), its corresponding fragment, SmBiT, was omitted. Instead NLuc-, SNAP- and LgBiT-tagged MMRN2 495-674 were produced. Plasmids were cloned using conventional cloning methods described in section 2.3.5.1. HEK293T cells were transfected with plasmids encoding tagged MMRN2. The constructs were designed to allow for production and secretion into the surround growth media. The conditioned media was collected and since the MMRN2 constructs included a His-tag, the protein was purified using Ni-NTA coated agarose beads. The protein was then buffer exchanged to remove imidazole (section 2.5.3).

To test for expression, a small-scale Ni-NTA bead pull-down was performed, the beads were treated with reducing sample buffer and subjected to SDS-PAGE and the gel was stained with Coomassie. Untagged MMRN2 495-674 was run alongside as a positive control. The gel indicates that the untagged MMRN2 495-674 construct expressed well, as indicated by the band at approximately 25kDa (Figure 5.9). The NLuc-, SNAP- and LgBiT- tag each increased the size of the protein with the respective molecular weights being 19.1kDa, 19.4kDa and 18kDa. The gel indicated that each tagged form of MMRN2 expressed and the corresponding bands for each of these fusions ran at a higher molecular weight. 1µg and 5µg of BSA were also run alongside each protein in order to estimate the amount of protein run on the gel and thus extrapolate the approximate expression efficiency. The gel indicated that each protein

expressed at relatively the same level and was comparable to the 1µg band of BSA (Figure 5.9).

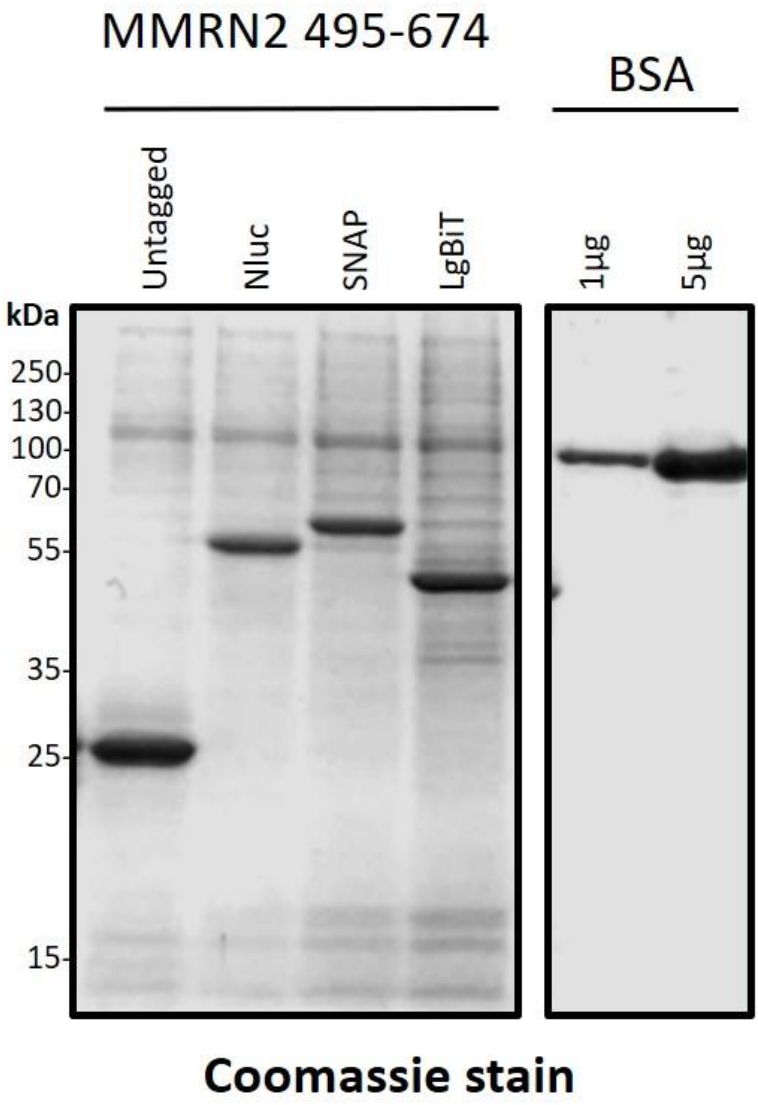


Figure 5.9. Successful expression of the NanoBRET and NanoBiT tagged MMRN2 fusion proteins. HEK293T cells were transfected with plasmids to produce and secrete untagged MMRN2 as well as the tagged MMRN2 variation. The proteins were secreted into the surrounding growth media and collected. Conditioned media was incubated with Ni-NTA beads. Beads were then treated with sample buffer, boiled and subjected to SDS-PAGE. The proteins were run on the gel alongside BSA controls. The gel was then stained with a Coomassie stain then destained with water.

5.4. NLuc-MMRN2 and SNAP-CLEC14A did not elicit a BRET signal

The initial NanoBRET experiment was performed using SNAP-tagged WT CLEC14A labelled with AlexaFluor 488 (AF488), and N-terminal tagged NLuc MMRN2 495-674 purified protein. The experiment was designed to study the interaction between CLEC14A and MMRN2 by monitoring the BRET interactions over time. To do so, two different concentrations of purified NLuc-MMRN2 495-674 protein was tested to gage the approximate binding affinity. HEK293T cells were transfected with 100 ng of SNAP-CLEC14A which were labelled using a membrane impermeable AF488 dye as described in section 2.8.2. The labelled cells were treated with either 1 nM or 100 nM of the NLuc-MMRN2 purified protein followed by the NLuc substrate Furimazine. Alongside this, a series of controls were run concurrently; untransfected cells treated with 100 nM of NLuc-MMRN2, unlabelled SNAP-CLEC14A transfected cells, and a negative control of untreated untransfected cells.

Wells were treated with Furimazine for 5 mins and then luminescence and fluorescence emissions were detected every 20 seconds over the course of one hour, and the BRET ratios were calculated. At the initial time point, 0 seconds, the BRET values between the two experimental samples and the negative controls were not vastly different. It appeared that the negative controls indicated a higher BRET ratio than both the 1 nM and 100 nM NLuc-MMRN2 treated samples (Figure 5.10A). The BRET ratios over time for the 100 nM NLuc-MMRN2 treated cells remained consistent at a value of 1 until 240 seconds.

This result indicated that the luminescence signal from the NLuc-MMRN2 protein was from completely saturated NLuc and the data upto 240 seconds was not an accurate representation of any changes in luminescence emissions or interaction with SNAP-CLEC14A (as measured by BRET). After the 12^{th} cycle, furimazine was oxidised to an appropriate level

which resulted in a slight increase in BRET signal. This increased BRET signal was higher than the 1 nM NLuc-MMRN2 reading at 240 seconds but also higher than the 1 nM NLuc-MMRN2 reading at 0 seconds. For the 1nM Nluc-MMRN2 sample, the BRET ratio decreased over time and the 100nM Nluc-MMRN2 sample BRET ratio also decreased over time at the point which the luminescence signal had not saturated. These data suggested no significant difference in BRET ratio between the two samples (Figure 5.10B). Moreover, higher BRET ratios were observed for the negative controls which consistently remained higher than both the 1nM and 100nM NLuc-MMRN2 treated wells (figure 5.10C). Taken together, these data suggested that there was no energy transfer between SNAP-CLEC14A and NLuc-MMRN2 indicative of either no interaction between these tagged proteins or orientation issues effecting the proximity of the Nluc and SNAP-tagged proteins resulting in suboptimal energy transfer.

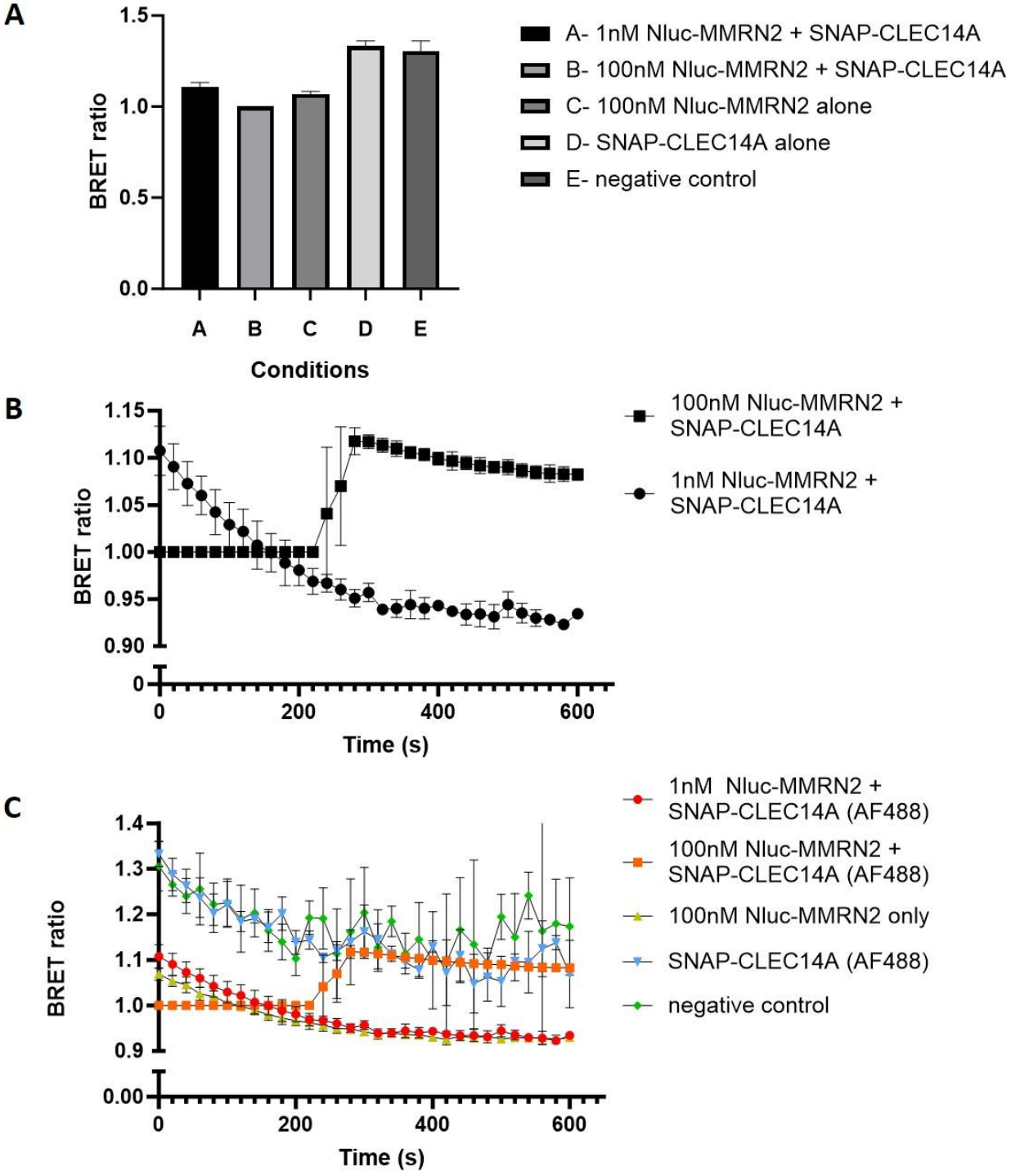


Figure 5.10. The interaction between Nluc-MMRN2 and SNAP-CLEC14A did not elicit a significant BRET signal. HEK293T cells were transfected with 100ng of plasmids to express SNAP-CLEC14A. The proteins were labelled with Alexafluore488 (AF488). The cells were treated with 1nM Nluc-MMRN2 or 100nM Nluc-MMRN2 purified protein. Controls were run alongside which negating one or other protein as well as a negative control of untransfected and untreated HEK293T cells. Following incubation cells were treated with the substrate furimazine and incubated for 5 minute at room temperature. The luminescent and fluorescent reading were then taken using the PHERAstar BRET1 plus module. (A) The BRET ratios for each condition was calculated at the initial time point of 0s. The BRET ratios are presented on the bar chart. (B) The BRET ratios were then calculated for the 1nM MMRN2 condition and the 100nM MMRN2 condition and plotted over time between 0s-600s which is presented on the graph. (C) The BRET ratios for each condition including negative controls was calculated and plotted over time between 0s-600s, this is presented on the graph. This was a single experiment with 3 technical repeats.

5.5. NanoBRET using GPVI as a positive control

The drawback when monitoring the BRET signal between NLuc-MMRN2 and SNAP-CLEC14A was the lack of a positive control. In order to ensure that the set up in terms of experimental design and procedure as well as analysis was correct, a positive control in the form of the Glycoprotein VI (GPVI) protein was utilised. The GPVI protein is a membrane protein expressed on platelets which can dimerise. The dimerization of GPVI has been previously studied using NanoBRET by co-transfecting HEK293T cells with NLuc-GPVI and Halo tag-GPVI (Clark *et al.,* 2021). The plasmids containing each construct were provided by Dr Joanne Clark, University of Birmingham.

The experiment was designed to study the BRET signal of GPVI dimerization by co-transfecting HEK293T cells with 100ng of NLuc-GPVI and increasing concentrations of HaloTag-GPVI ranging from 0 to 200ng. The cells were then labelled with membrane impermeable AF488. The controls were the same co-transfections except Halo-tagged GPVI was not labelled. The fluorescent intensity was read prior to addition of the substrate furimazine. This was to ensure the experimental wells were successfully labelled with AF488. The results indicate an increase in fluorescence of labelled wells compared to the unlabelled. Furthermore, the fluorescence increased as the concentration of HaloTag GPVI increased whilst the unlabelled well remained the same indicating that the cells were successfully labelled (Figure 5.11A). Wells were then treated with Furimazine, and the luminescent and fluorescent readings were documented. The BRET ratios were calculated and the results revealed that as the concentration of AF488-labelled HaloTag GPVI increased, the BRET ratio also increased indicating successful energy transfer as the GPVI monomers formed dimers. The BRET ratio for the unlabelled wells remained at a baseline level (Figure 5.11B).

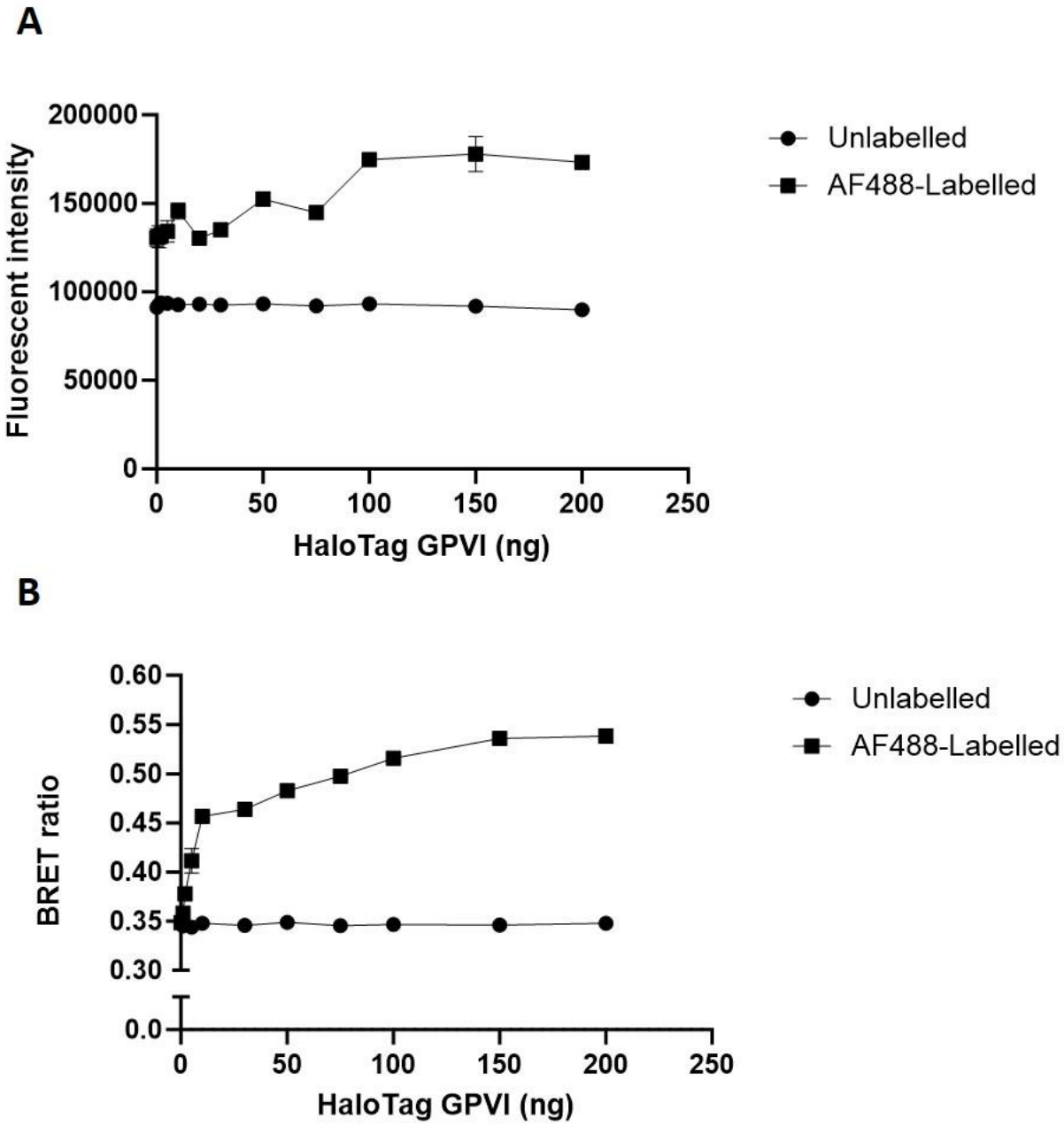


Figure 5.11. The GPVI dimerization successfully elicited a BRET signal. HEK293T cells were co-transfected with 100ng of plasmid to express Nluc-tagged GPVI alongside increasing amounts of HaloTag-GPVI plasmid ranging from 0-200ng. The HaloTag protein was labelled with Alexafluor488, and some wells were left an unlabelled controls. (A) Fluorescence was measured on the PHERAstar using the Fluorescence intensity (485 520) module and the gain was adjusted to 459. The fluorescent intensity values of the labelled and unlabelled wells are displayed in the graph. (B) The wells were treated with furimazine and allowed to incubate for 5mins at room temperature before fluorescence and luminescent readings were taken using the PHERAstar BRET1 plus module. The BRET ratios were calculated and are presented on the graph for both labelled and unlabelled HaloTag GPVI. This was a single experiment with 3 technical repeats, values shown are mean ± S.D.

5.6. Optimising NanoBRET tagged MMRN2 495-674

The successful detection of GPVI dimers using NanoBRET, suggested that the physical set up was not the issue with the MMRN2/CLEC14A NanoBRET experiments. To explore why there was no BRET signal when studying the CLEC14A-MMRN2 interaction the predictive docking model was employed. The model reveals 32 individual alpha helical turns between the N-terminus of MMRN2 495-674 and the CLEC14A binding site (Figure 5.12). Since there are 3.6 amino acids and a distance of 5.6 Å per turn (Mazzone *et al.,* 2011), then 32 turns equate to a distance of 172.8 Å or 17.3 nm. Based on this predicted model, the energy transfer between Nluc-MMRN2 and SNAP-CLEC14A would likely not occur as protein tags are required to be within 10nm of each other to elicit a meaningful BRET signal (Dale *et al.,* 2019). However, the distance between the predicted CLEC14A binding site and the C-terminus of MMRN2 495-674 has 9 alpha helical turns. This translates to 4.86nm, which theoretically should allow for efficient energy transfer. Therefore, the constructs were re-engineered to add Nluc, SNAP and LgBiT tags to the C-terminus of MMRN2 495-674 (Figure 5.13). It has been well documented in the literature that addition of linkers can also optimise recombinant proteins by improving structural stability, flexibility, function as well as decrease potential steric hindrance (Chen *et al.,* 2013). In this instance the aim of the linker was to increase flexibility of the NanoBRET or LgBiT tag. To ensure the linker would be flexible, not fold into a secondary structure nor affect the folding and expression of MMRN2 495-674 a Glycine and Serine rich linker was generated as depicted in Figure 5.14. The linker designed followed a (GGGGS) X3 sequence which has previously been shown in the literature to enhance protein expression as well as not interfere with secondary structure (Trinh *et al.,* 2004). The sequence of the original pHLAvitag3 plasmid containing MMRN2 495-674 included restriction sites KpnI and EcoRV at the 3' end of the

MMRN2 495-674 sequence which allowed for insertion of the linker as described in section 2.3.5.3.

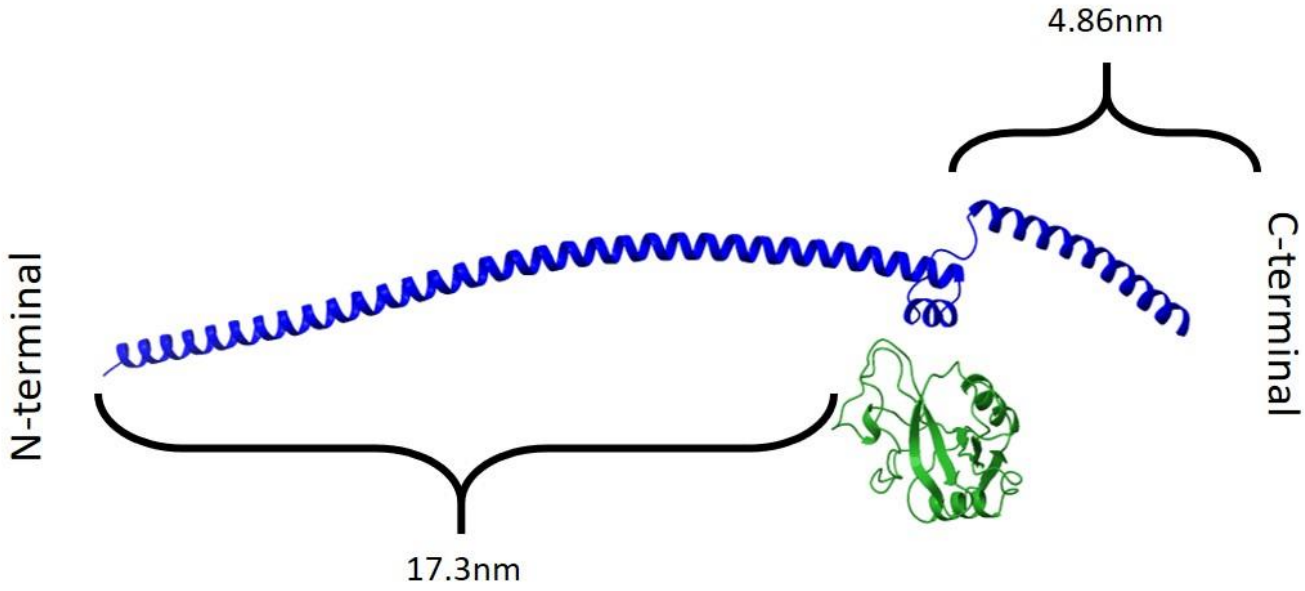


Figure 5.12. Predicted model of the CLEC14A-MMRN2 interaction. The AlphaFold predicted model was viewed in the Chimera software and depicts the CTLD of CLEC14A (green) and MMRN2 495-674 fragment (Blue). Figure is labelled with the calculated distance in nanometres between the predicted binding interface and either the N-terminal or the C-terminal of MMRN2 495-674.

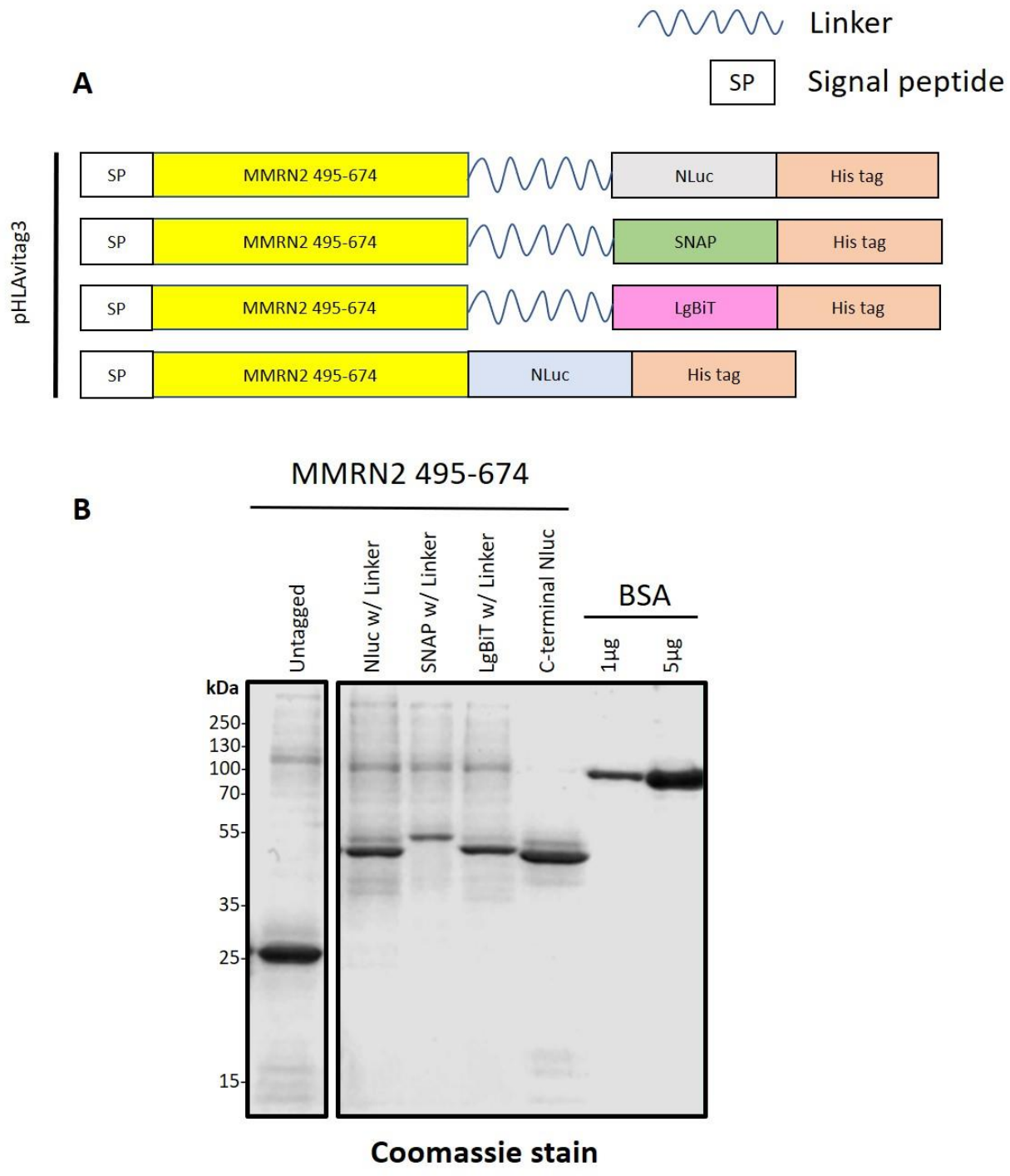


Figure 5.13. Successful expression of the NanoBRET and NanoBiT tagged MMRN2 fusion proteins. (A) Schematic illustration of newly constructed C-terminal fused NanoBRET and NanoBiT tagged MMRN2 495-674. Plasmids include a signal peptide for expression and were constructed in the pHLAvitag3 vector (B) HEK293T cells were transfected with plasmids to produce and secrete untagged MMRN2 as well as the tagged MMRN2 variation. The proteins were secreted into the surrounding growth media and collected. Conditioned media was incubated with Ni-NTA beads. Beads were then treated with sample buffer, boiled and subjected to SDS-PAGE. The proteins were run on the gel alongside BSA controls. The gel was then stained with a Coomassie stain then destained with water.

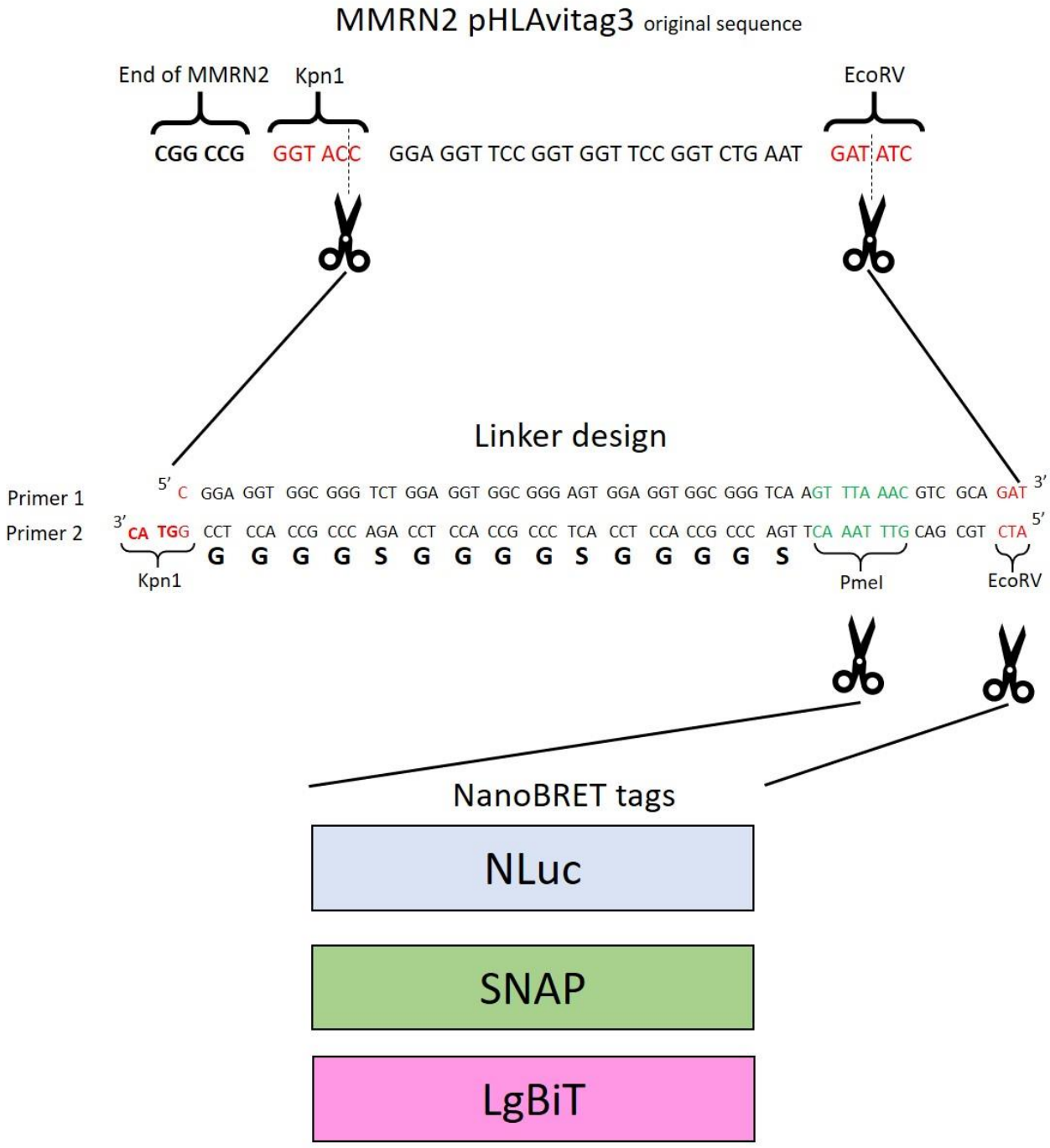


Figure 5.14. Schematic illustration of linker design. The pHLAvitag3 MMRN2 495-674 plasmid was modified to include a 15/20 amino acid-long linker 3' of the MMRN2 sequence and 5' of the His tag between the restriction sites KpnI and EcoRV. The linker included a PmeI restriction site. NanoBRET tags were inserted between the PmeI restriction site and the EcoRV site of the newly formed plasmid.

The C-terminal SNAP-tagged MMRN2 fragment with a linker was expressed as a purified protein for the second NanoBRET attempt, therefore, the protein was tested in a flow cytometry assay as well as a western blot to ensure binding to NLuc-CLEC14A at the cell surface (Figure 5.15 and 5.16). Binding of the CRT2 antibody confirmed that both untagged and Nluc-tagged CLEC14A were expressed at the cell surface as indicated by the small peak shifted to the right of the isotype control. This result was mirrored when studying SNAP-MMRN2-linker binding for both untagged and NLuc-tagged CLEC14A. The percentage of cells within the sample which were recognised by either CRT2 or SNAP-MMRN2-linker were monitored for both untagged CLEC14A and NLuc-tagged CLEC14A and were found to be broadly similar (Figure 5.15), confirming that the SNAP-MMRN2-linker can bind to NLuc-tagged CLEC14A. Untransfected cells were used as a negative control and did not indicate any peaks above the isotype control.

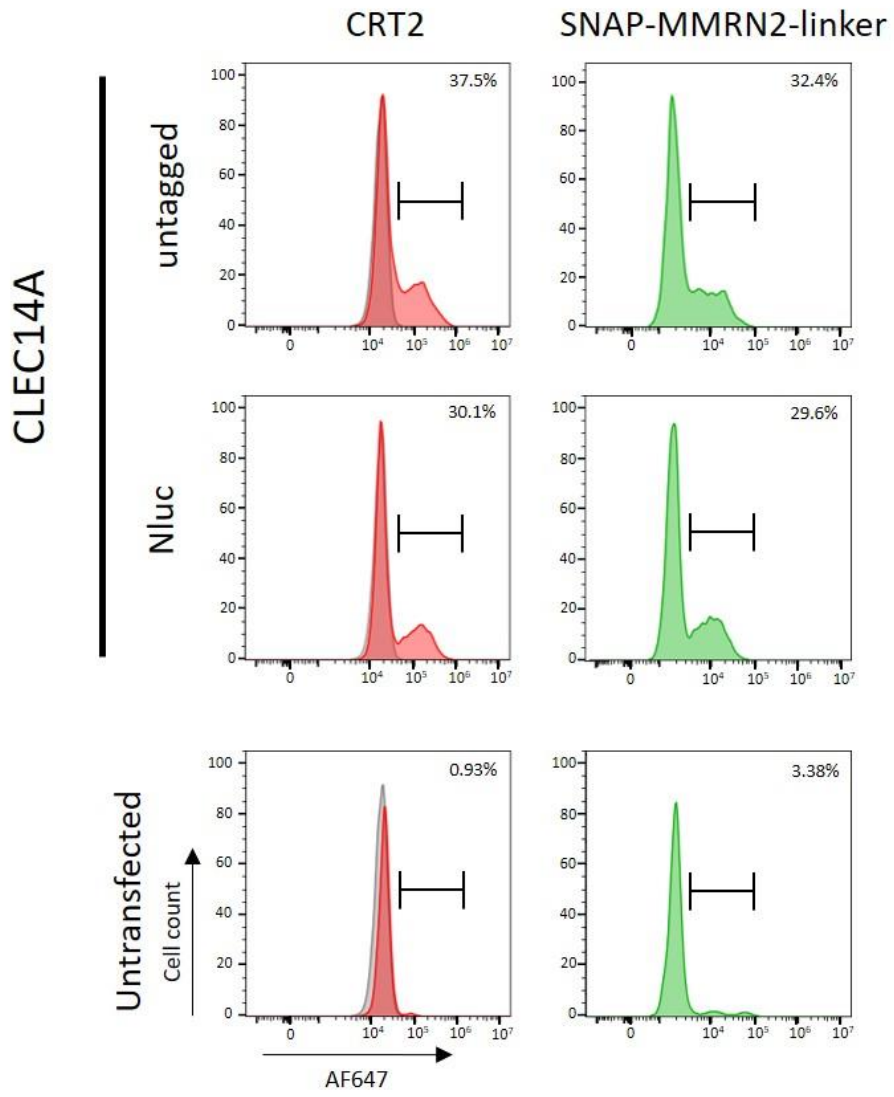


Figure 5.15. SNAP-tagged MMRN2 with a linker binds to Nluc-tagged CLEC14A. HEK293T cells were transfected with plasmids to express either untagged CLEC14A or Nluc tagged CLEC14A. One set of cells was treated with the CRT2 antibody as a primary antibody followed by incubation with an anti-mouse IgG Alexafluor647 conjugated antibody. Another set of cells was incubated with the MMRN2-SNAP-linker purified protein followed by incubation with an anti-his antibody conjugated to Alexafluor647. Flow cytometry was performed, and the histograms presenting the CRT2 antibody incubations are highlighted in red including the isotype control (grey). The histograms presenting the MMRN2-SNAP-linker purified protein incubation are highlighted in green. The percentage of cells within the sample which elicited a signal above the isotype control are displayed in the corner of each histogram. Untransfected cells were used as a negative control.

The purified protein was also studied by western blotting using the anti-SNAP tag antibody. The results indicated two bands for both labelled and unlabelled MMRN2-SNAP; a higher molecular weight band at around 45kDa, which is the approximate predicted size of the MMRN2-SNAP, and another lower molecular weight band at approximately 20kDa (Figure 5.16A). Since both bands appear on the anti-SNAP western blot it was expected that that the lower molecular weight band represents the C-terminal portion of the protein. To test if the smaller fragment was labelled with AF488, a second western blot was performed using an anti-AF488 antibody. The results indicated that the labelling of the purified protein was successful, and the unlabelled protein was not labelled. Interestingly, the lower molecular weight protein was also labelled and was more intense than the higher band (Figure 5.16B).

It is likely that the lower molecular weight band does not bind to NLuc-CLEC14A, and the flow cytometry results represent the binding of the higher molecular weight species to NLuc-CLEC14A. However, when determining the concentration of the sample both species are considered therefore to separate both species and correctly determine the amount of higher molecular weight species in the sample, size exclusion gel filtration was performed. Gel filtration was performed using the AKTA S200 column as described in section 2.6. The gel filtration trace did not appear to separate two specific species within the sample, instead the trace revealed a peak which is slightly broadened by a “shoulder” peak detected at 8.86 mL. This peak corresponded to ~680kDa which is much higher than expected. The trace also revealed excessively low concentrations of protein post filtration as the highest absorption value was 1.5. There were also shorter peaks detected at 19.17 mL and 21.22 mL which correspond to ~0.929kDa and ~0.55kDa respectively (Figure 5.17A). The fractions representing each peak in the gel filtration trace were collected and subjected to western blotting using the Anti-His antibody. The protein pre-filtration was also run alongside as a

reference which revealed the two bands as seen previously. There were no visible bands when the fractions from each peak within the trace were run most likely due to the extremely low concentration of the protein post-gel filtration (Figure 5.17B).

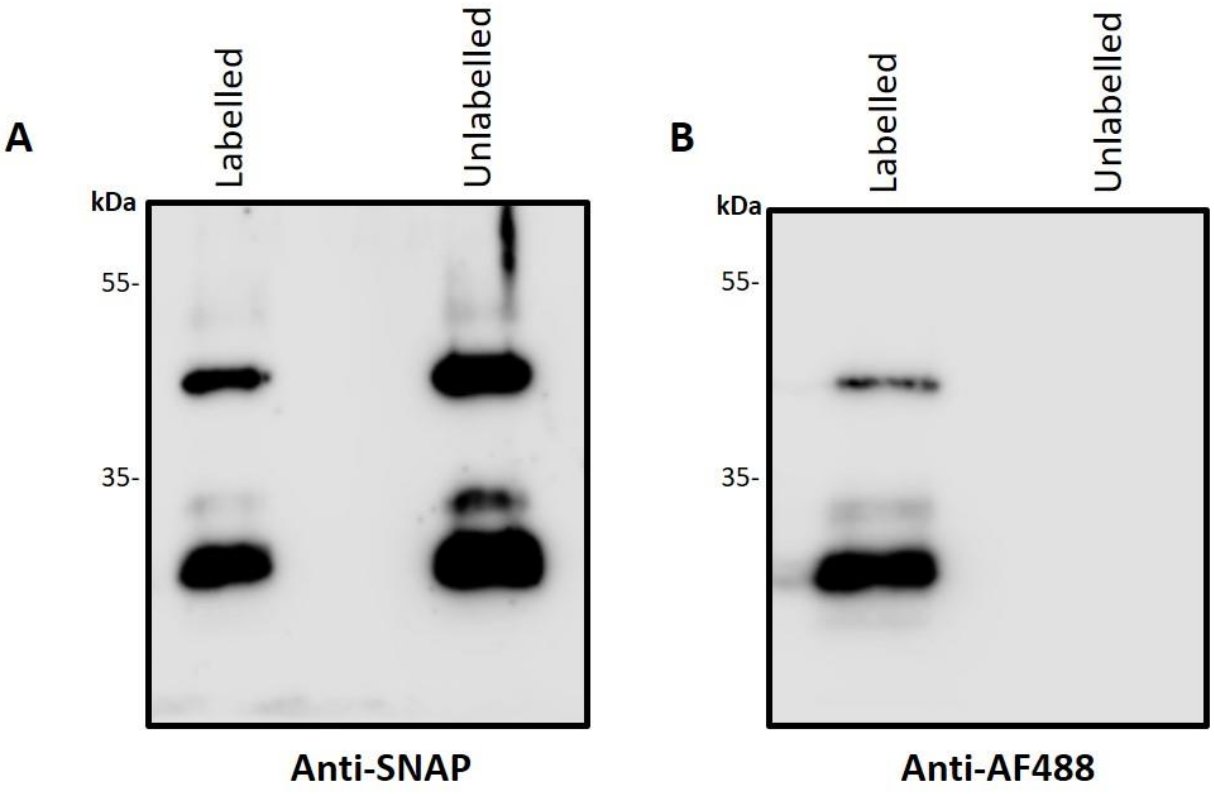


Figure 5.16. SNAP-MMRN2-linker purified protein displays two bands, both of which are labelled with AF488. HEK293T cells were transfected with plasmids to produce and secrete SNAP-MMRN2-linker purified protein into the surrounding media. The protein was then purified from the conditioned media using Ni-NTA beads which were then eluted off the beads using imidazole. The purified protein was labelled with Alexafluor488. Both labelled and unlabelled protein was subjected to western blotting. (A) The membrane was probed with an anti-SNAP tag antibody followed by an incubation with a secondary anti-rabbit antibody conjugated to HRP. (B) Another membrane was probed with an anti-AF488 antibody followed by incubation with a secondary anti-rabbit antibody conjugated to HRP.

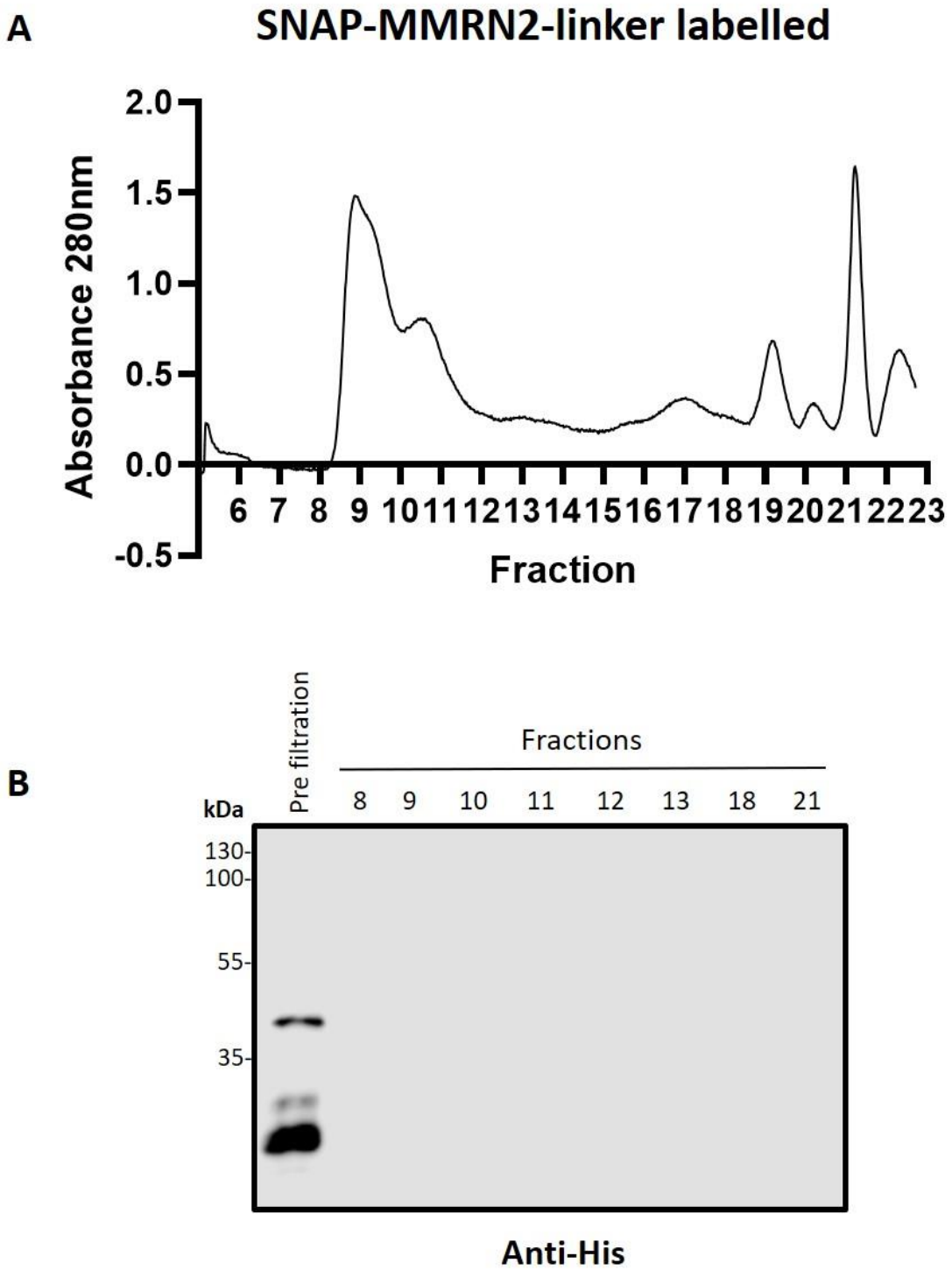


Figure 5.17. Purified MMRN2-SNAP-linker aggregates to form large molecular complexes in solution. The purified MMRN2-SNAP-linker protein was subjected to gel filtration using the AKTA S200 column. The sample was run at a rate of 0.3mL/min. (A) The absorbance was taken at 280nm and presented on the graph indicating the various fractions taken. (B) The fractions covering the peak was collected and subjected to western blotting alongside the purified protein pre-gel filtration. The membrane was probed with an anti-His antibody conjugated to HRP.

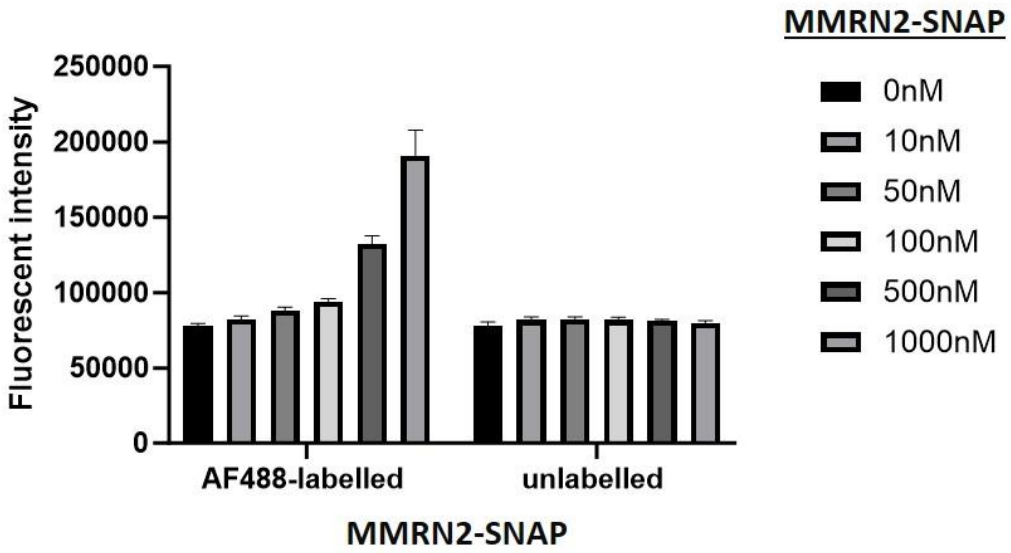
5.7. SNAP-MMRN2-linker and NLuc-CLEC14A did not elicit a BRET signal

Since the purified SNAP-MMRN2-linker protein was expressed, successfully labelled and able to bind to NLuc-CLEC14A, the NanoBRET assay was performed. Given that the two molecular weight species within the protein sample could not be separated without loss of protein, the heterogenous sample was used in the BRET experiment and the concentration range of purified protein used in the assay was increased up to 1000nM.

A fixed amount of 50ng NLuc-CLEC14A was transfected into HEK293 cells and further treated with increasing concentrations of purified and labelled SNAP-MMRN2-linker ranging from 1-1000nM. The corresponding negative control samples were 50ng NLuc-CLEC14A transfected cells treated with unlabelled MMRN2-SNAP.

Prior to Furimazine addition, the fluorescence was read using the fluorescence intensity module. The results confirmed that the SNAP-MMRN2-linker protein was fluorescently labelled, and fluorescence signal increased as the concentration of protein added increased, moreover, the unlabelled wells remained at a baseline value (Figure 5.18A). Wells were then treated with Furimazine and the luminescence and fluorescence readings were taken. The results indicated no significant difference in BRET signal as the amount of labelled SNAP-MMRN2-linker increased. There was no significant difference in BRET signal between the experimental and negative control wells indicating that this combination of proteins did not elicit a meaningful BRET signal (Figure 5.18B).

A



B

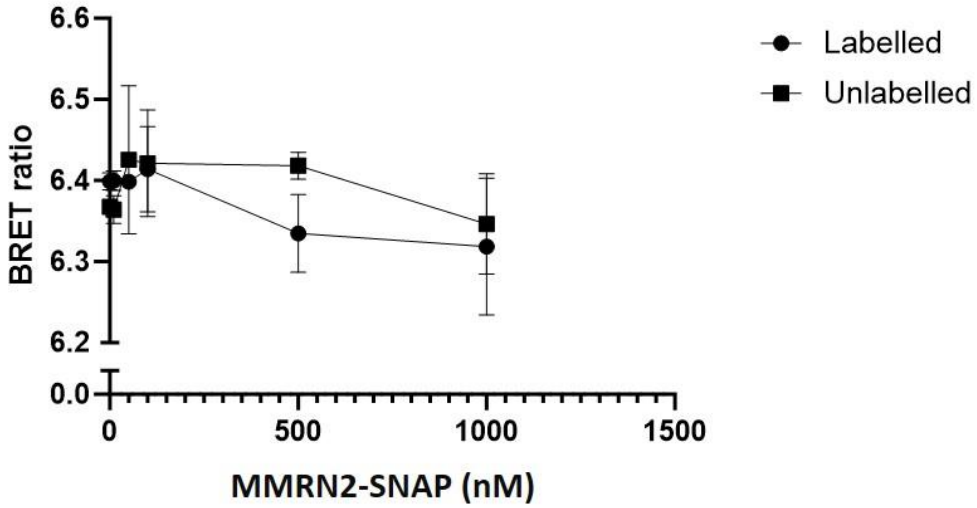


Figure 5.18. The interaction between N-terminal tagged MMRN2-SNAP-linker did not elicit a significant BRET signal. HEK293T cells were transfected with 50ng Nluc-CLEC14A and then treated with increasing amounts of either AF488-labelled or unlabelled MMRN2-SNAP-linker ranging between 0nM to 1000nM. (A) Fluorescence was then read using the PHERAstar Fluorescence intensity (485 520) module and the gain was adjusted to 171. Fluorescent intensity is presented in the bar chart. (B) the wells were then incubated with furimazine for 5mins at room temp and the fluorescence and luminescence was read using the PHERAstar BRET1 plus module. The BRET ratio was calculated and presented on the graph. This was a single experiment with three technical repeats.

5.8. Expression of C-terminal tagged MMRN2-NLuc

It is clear that the SNAP-MMRN2-linker construct did not elicit a BRET signal, therefore a further construct was designed wherein the MMRN2 495-674 fragment was fused at the C-terminal with the NLuc tag without a linker. This was designed in the instance that the linker added too much flexibility and, by extension, too much distance between the tags once CLEC14A and MMRN2 had bound. This construct was cloned and purified using Ni-NTA beads and imidazole. Throughout the purification process, samples were taken and subjected to SDS-PAGE. The gel was then Coomassie stained to reveal that the protein had successfully been captured onto the beads and was purified. The beads were left empty, and the imidazole was removed leaving a single band representing purified MMRN2-NLuc. 1µg and 5µg of BSA were run alongside for comparison (Figure 5.19).

Once the protein was purified it was important to ensure that the placement of the tag did not affect binding to SNAP-tagged CLEC14A. Therefore, flow cytometry was performed using HEK293T cells expressing either untagged and SNAP-tagged CLEC14A which were treated with CRT2 or MMRN2-NLuc. The results indicate that both untagged and SNAP-CLEC14A were expressed at the cell surface as both were recognised by CRT2 (Figure 5.20). Untagged CLEC14A bound to both CRT2 and MMRN2-NLuc however in both instances the peak was very small and not very well defined from the isotype control peak. The percentage of cells recognised by the CRT2 antibody were 25.6% and similarly MMRN2-NLuc bound to 22.8% of cells. CRT2 and MMRN2-NLuc also bound to SNAP-CLEC14A however in these instances the peak indicating the signals were well defined and distinct compared to the isotype control peak. The percentage of cells recognised by CRT2 and MMRN2-NLuc were 38.4% and 32.4% respectively. This indicated that the transfection efficiency was higher for SNAP-CLEC14A than untagged CLEC14A and binding of CRT2 and MMRN2-NLuc were similar. Untransfected cells

were used as the negative control and did not indicate any signal above the isotype control peak (Figure 5.20).

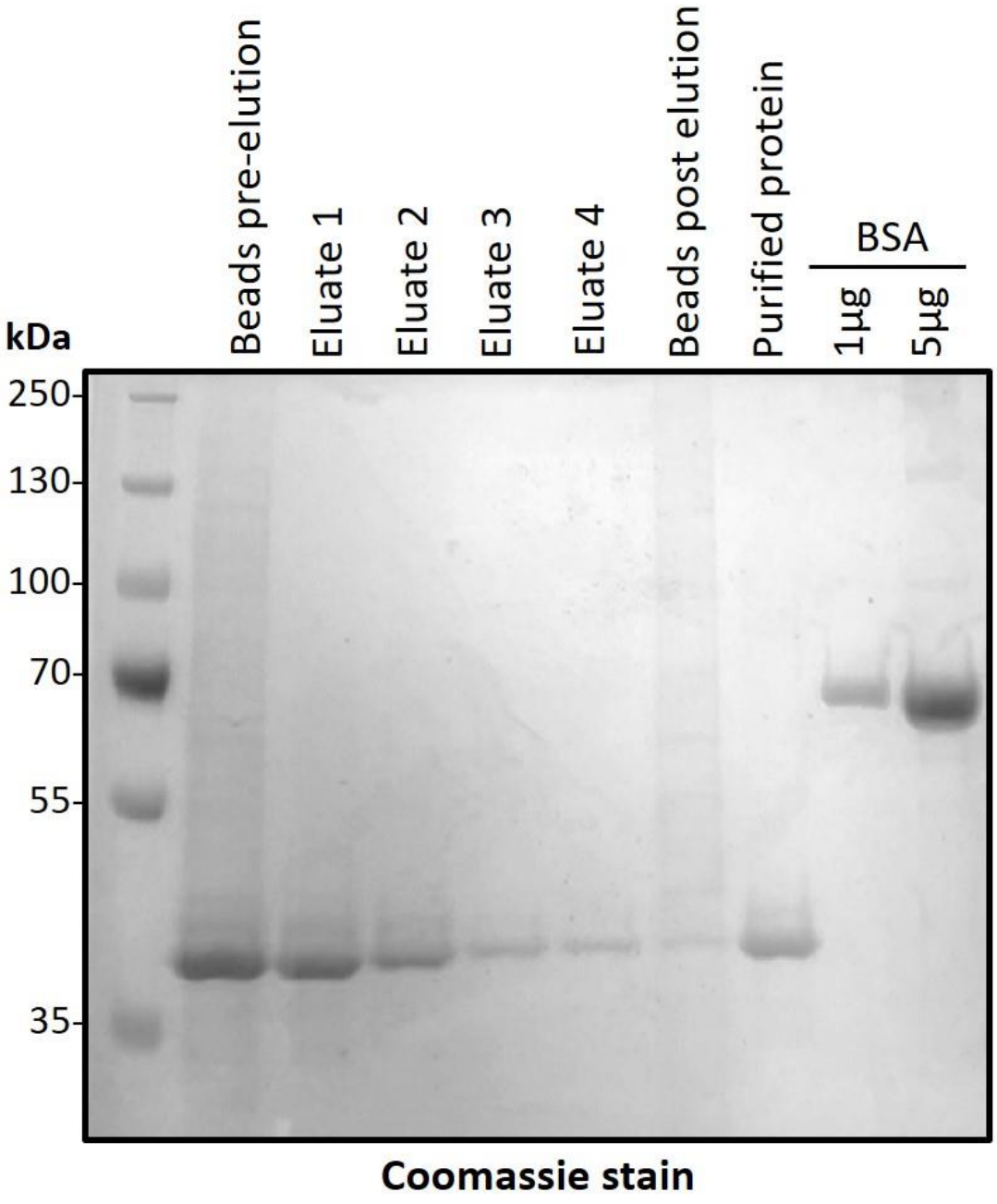


Figure 5.19. Coomassie stained gel of C-terminal Nluc-tagged MMRN2 495-674 purification. HEK293T cells were transfected with plasmid to produce and secrete MMRN2-Nluc into the surrounding media. The conditioned media was then collected, and the protein was purified using Ni-NTA beads. The beads were then treated with PBS with 500mM Imidazole to elute off the protein. Imidazole was removed by buffer exchange. Each purification step was run on a gel alongside BSA controls and stained with Coomassie and de-stained with water.

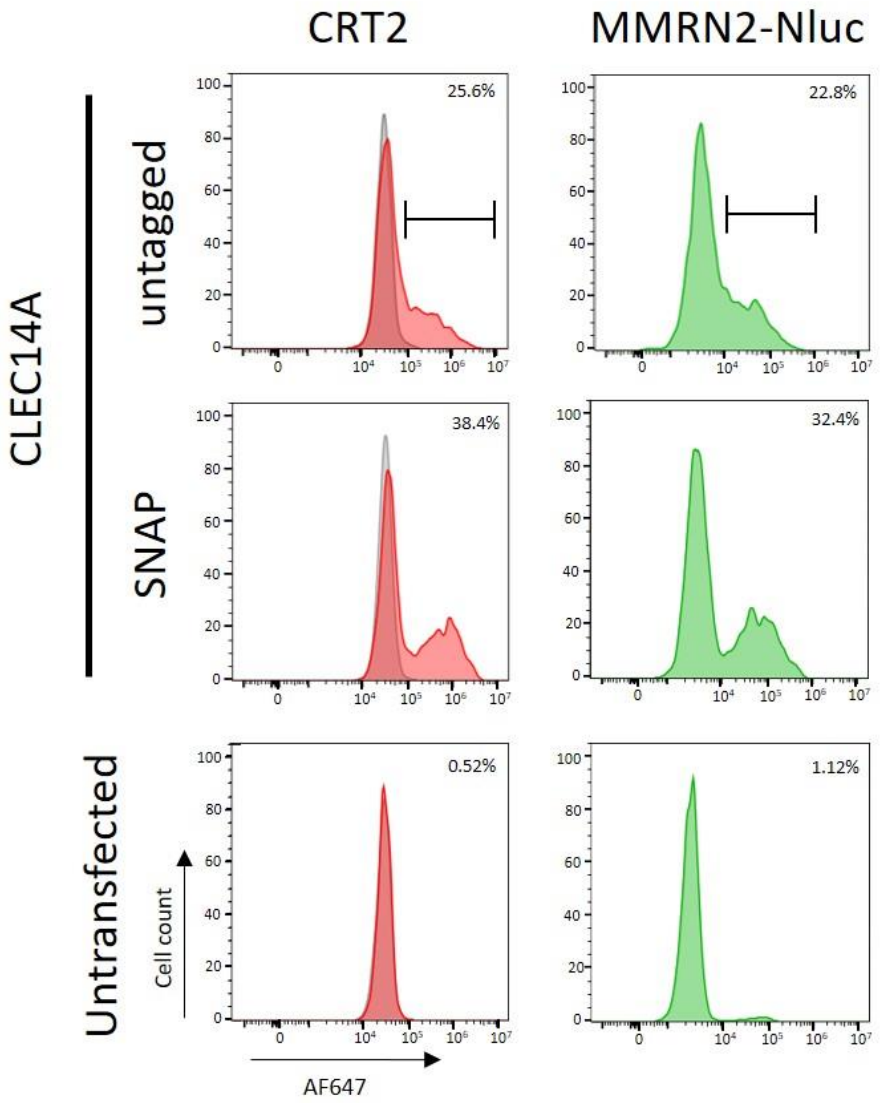


Figure 5.20. C-terminal Nluc tagged MMRN2 binds to SNAP-tagged CLEC14A. HEK293T cells were transfected with plasmids to express either untagged CLEC14A or SNAP tagged CLEC14A. One set of cells was treated with the CRT2 antibody as a primary antibody followed by incubation with an anti-mouse IgG Alexafluor647 conjugated antibody. Another set of cells was incubated with purified MMRN2-Nluc followed by incubation with an anti-his antibody conjugated to Alexafluor647. Flow cytometry was performed, and the histograms presenting the CRT2 antibody incubations are highlighted in red including the isotype control (grey). The histograms presenting the MMRN2-Nluc purified protein incubation are highlighted in green. The percentage of cells within the sample which elicited a signal above the isotype control are displayed in the corner of each histogram. Untransfected cells were used as a negative control.

5.9. MMRN2-NLuc and SNAP-CLEC14A did not elicit a BRET signal

Given that the MMRN2-NLuc protein was expressed and bound to SNAP-CLEC14A a further attempt at the NanoBRET assay was performed. Since SNAP-CLEC14A appears to have a higher transfection efficiency than untagged CLEC14A the protein was studied at three different concentrations. HEK293T cells were either left untransfected or transfected with 50 ng or 100 ng of the SNAP-CLEC14A plasmid. For each condition 10 wells were left unlabelled and 10 were labelled with the membrane impermeable AF488 as described in section 2.8. The fluorescence intensity was read, and the results indicate that the untransfected cells had a baseline fluorescence which increased as higher concentrations of fluorescently labelled MMRN2 increased. It appears that the 50 ng transfected cells indicated a much higher level of fluorescence intensity compared to the 100 ng control (Figure 5.21A). Therefore, BRET experiments were performed on 50 ng treated cells.

NanoBRET was then performed with the addition of furimazine and luminescence and fluorescence readings were taken. In this instance, wells were treated with increasing amounts of MMNR2-NLuc purified protein ranging from 0-500nM. Again, unlabelled cells were used as the negative control. The results indicated very little increase in the BRET ratio as the amount of MMRN2-NLuc purified protein was added. However, the unlabelled negative control indicated higher BRET ratios than the experimental results. At the maximum protein treatment condition of 500nM of MMRN2-NLuc, both BRET ratios for labelled and unlabelled wells were equal. This indicates that no meaningful BRET signal was observed (Figure 5.21B).

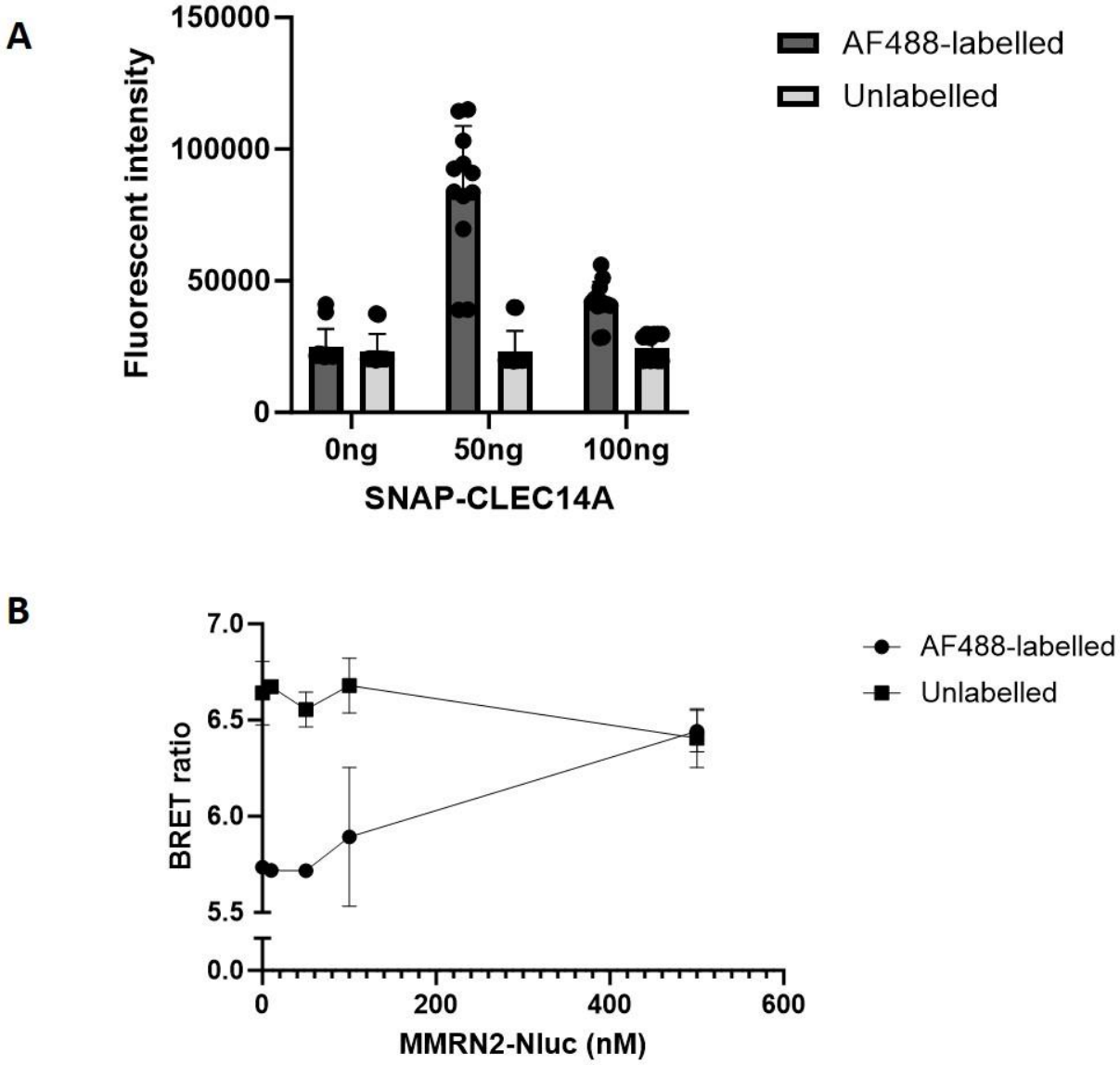


Figure 5.21. The interaction between SNAP-CLEC14A and N-terminal Nluc-tagged MMRN2 did not elicit a significant BRET ratio. HEK293T cells were either transfected with 50ng or 100ng SNAP-CLEC14A or left untransfected. Wells were either labelled with Alexafluor488 or left as controls. (A) Fluorescence was then read using the PHERAstar Fluorescence intensity (485 520) module. The wells were then incubated with furimazine for 5mins at room temperature and the fluorescence and luminescence was read using the PHERAstar BRET 1 plus module. (B) The BRET ratio was then calculated and is presented on the graph for both labelled and unlabelled wells. This was a single experiment with 3 technical repeats.

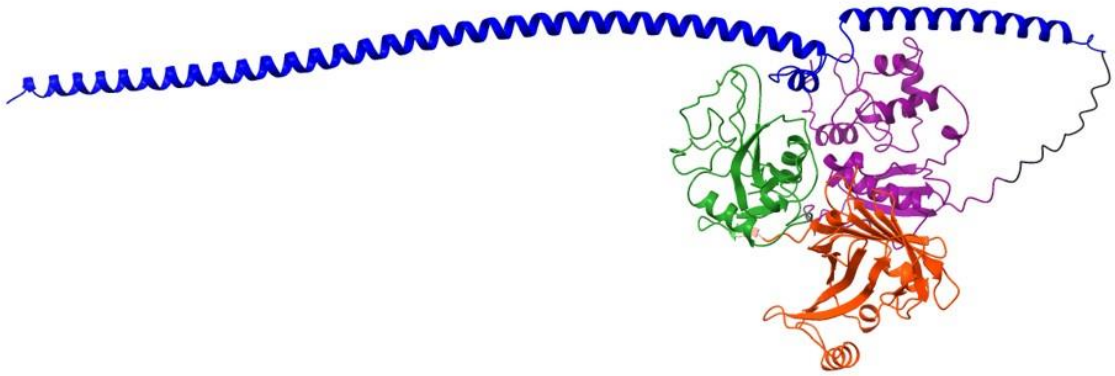
5.10. Discussion

The aim was to study the binding kinetics between CLEC14A and MMRN2 this would involve identifying the binding affinity which, based on previous studies using SPR, is predicted to be in the nanomolar range (Kabir Khan PhD thesis, 2015). Throughout this portion of the project numerous attempts and cloning strategies were performed in order to elicit a BRET signal between CLEC14A and MMRN2 none of which were successful.

Addition of each tag to either CLEC14A or MMRN2 495-674 did not diminish the capacity of the protein to express or bind to its partner protein. Despite this, no BRET signal was detected across all constructs tested. This suggests that while the constructs likely bound to CLEC14A, they were not in sufficiently close proximity (<10nm) to facilitate energy transfer. C-terminal tagged MMRN2 (495-674) proteins which were produced to optimize the assay also did not elicit a BRET signal. According to the predicted docking models of CLEC14A and MMRN2 which were used in the second phase of the study it appeared that the addition of the linker would increase the proximity of the Nluc and SNAP tags compared the non-linker variations (Figure 5.22). Based on the docking model of SNAP-MMRN2 with a linker both the SNAP and Nluc tag would be in close enough proximity to elicit a BRET signal (Figure 5.22A). The C-terminus Nluc tagged MMRN2 495-674 which did not contain a linker did appear to be further away from the SNAP tagged CLEC14A in the docking model which may provide an explanation as to why the assay did not work (Figure 5.22B).

In the instance that the docking model may have some inaccuracies particularly when modelling the SNAP-MMRN2 495-674 linker protein, addition of the SNAP or NLuc tag closer to the predicted binding site of CLEC14A might have been an option. However, this would involve cloning the tags half way into the MMRN2 fragment, the major concern is that this

A



B

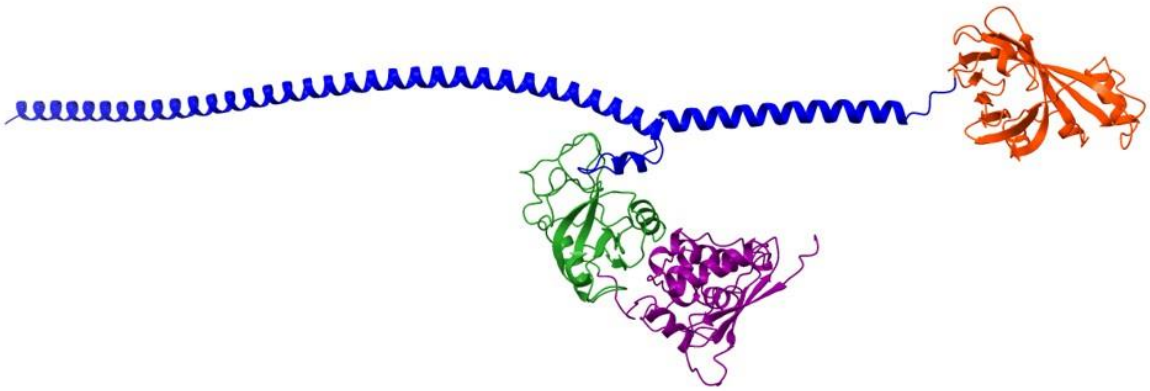


Figure 5.22. AlphaFold predicted models of CLEC14A and MMRN2 constructs used in the study viewed in the Chimera software. The CTLD of CLEC14A is depicted in green, MMRN2 495-674 is depicted in blue, the Nluc tag is depicted in red and the SNAP tag is depicted in purple. (A) Docking model of Nluc tagged CLEC14A CTLD with C terminal tagged SNAP-MMRN2 with a linker (black). (B) Docking model of SNAP-tagged CLEC14A CTLD docked with C-terminus fused Nluc-MMRN2 without a linker.

could potentially disrupt the coiled structure of MMRN2 and negatively impact its interaction with CLEC14A. It is generally recommended to place tags at either the N- or C-terminus, or between two domains, to minimize structural obstruction (Malhotra, 2009).

NanoBRET attempts using CLEC14A and MMRN2 (495-674) involved the use of the NLuc-tag and the SNAP-tag labelled with AF488. Although NanoBiT tags (LgBiT and SmBiT) were cloned into CLEC14A and MMRN2, they were not utilized in the project due to the time-consuming and expensive nature of protein production. Therefore, only one set of proteins was produced at a time, and the NLuc and SNAP tags have been extensively shown in the literature to elicit a BRET signal (Kilpatrick *et al.,* 2017). The use of NanoBiT tags typically involves the LgBiT-tag and the HiBiT-tag. Both SmBiT and HiBiT are smaller fragments of NLuc; however, HiBiT has a higher binding affinity for LgBiT compared to SmBiT (Liu and Guo, 2022), making HiBiT more preferable as a tag for this experiment. Numerous attempts to clone the HiBiT-tag onto CLEC14A and MMRN2 (495-674) failed during the early stages of the cloning process. To overcome this issue, an alternative strain of E. coli and Gibson assembly primers were used, but bacterial colonies failed to grow. We considered that the tertiary structure of the resulting DNA or certain sequences might be toxic to the bacteria.

A BRET signal was successfully elicited using a positive control involving GPVI, a membrane protein previously shown to produce a BRET signal upon dimerization (Clark *et al.,* 2021). However, the main limitation of this positive control is that it involves the interaction between two membrane proteins dimerizing at the cell surface, unlike the interaction between CLEC14A and MMRN2, which occurs between a membrane protein and an extracellular matrix protein. Membrane protein interactions are typically stronger and more stable compared to extracellular interactions, which tend to be inherently weaker and more transient (Hynes,

2009). For a more accurate comparison, a verified BRET interaction between a membrane protein and an extracellular matrix protein, such as Nluc-tagged VEGFR2 with fluorescently labelled $VEGF_{165}$a, would have been more suitable (Kilpatrick *et al.,* 2017). Additionally, it is possible that the protein concentration range used in this project was not optimal to elicit a BRET signal.

Previous studies have shown that CLEC14A is cleaved by RHBDL2 (Noy *et al.,* 2016) and undergoes internalization following antibody treatment (Puja Lodhia, PhD thesis, 2016). Whether these events can be triggered by MMRN2 binding remains unknown and could have been investigated using the NanoBRET setup. Alternatively, microscopy-based imaging techniques using SNAP-tagged CLEC14A and SNAP-tagged MMRN2 (495-674) labelled with different fluorophores could also be employed to monitor the interactions between these proteins. This approach would provide valuable data regarding the rate of receptor internalization and potential recycling to the cell surface.

CHAPTER 6: Biophysical analysis of CLEC14A and MMRN2

CHAPTER 6: Structural analysis of CLEC14A and MMRN2

6.1. Introduction

Understanding the structure of individual proteins and protein complexes often reveal insights into their function and regulation (Jisna and Jayaraj, 2021). CLEC14A and MMRN2 interact with each other at the interface between the luminal and abluminal side of tumour endothelial cells and pericytes, with both proteins being involved in regulating angiogenesis (Noy *et al.,* 2015).

MMRN2 belongs to the EMILIN-like family of proteins and contains three domains, an N-terminal EMI domain followed by a coiled-coil domain and a C-terminal C1q domain (Christian *et al.,* 2001). CLEC14A is a type I integral membrane protein comprising an N-terminal CTLD, a sushi domain, one epidermal growth factor-like domain, a highly glycosylated mucin-like region, a transmembrane region and a short intracellular tail (Khan *et al.,* 2019). As indicated in previous chapters the CTLD of CLEC14A engages with the fragment of MMRN2 between amino acids 495-674 (Figure 6.1A). At present, neither CLEC14A nor MMRN2 have any experimentally resolved structures, instead AlphaFold predictive modelling was used to investigate the structural interface between these two proteins (Figure 6.1B and C). The aim of this section was to study the structure of CLEC14A and MMRN2 using a range of biophysical techniques such as analytical ultracentrifugation (AUC), gel filtration and intact mass spectrometry to study how the protein exists in solution, including the size, folding, and binding to its partner protein of interest. This would provide valuable insights into the structure of CLEC14A and MMRN2 which would aid in the development of novel therapies. Determining the structure of the complex using Cryo-EM was also a key aim in this study.

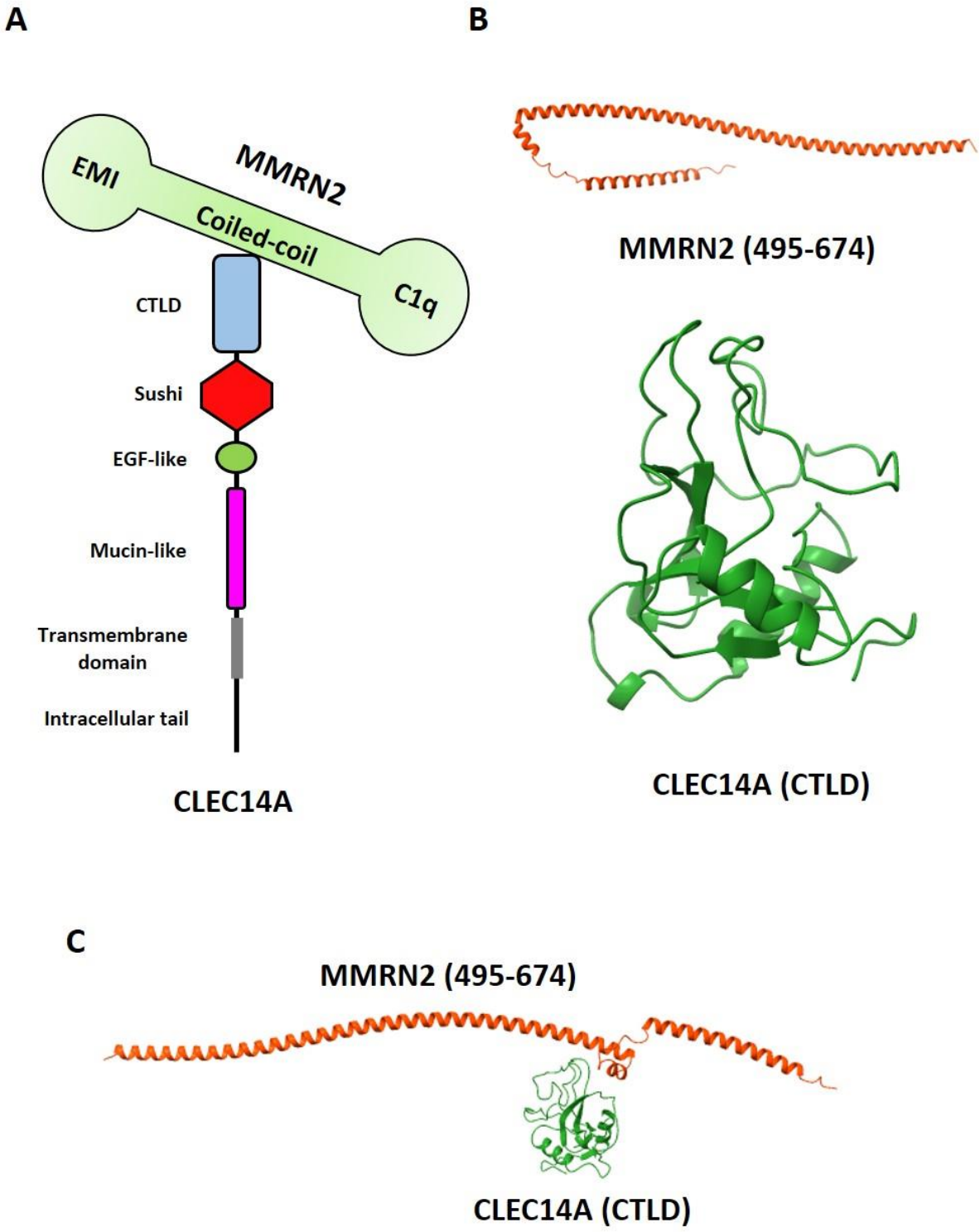


Figure 6.1. Schematic illustrations and AlphaFold models of CLEC14A and MMRN2. (A) Schematic illustrations of full length CLEC14A and MMRN2 with individually highlighted domains. (B) The AlphaFold predicted models of the CTLD of CLEC14A depicted in green and the MMRN2 495-674 region of the coiled-coil domain depicted in red. (C) The AlphaFold predicted docking model of CLEC14A CTLD in green with MMRN2 495-674 in red.

6.2. Protein production and purification

In order to study the structures of CLEC14A and MMRN2, soluble forms of each protein were generated. Expression of full length MMRN2 was problematic and so instead two fragments of MMRN2 were generally used: the MMRN2 coiled-coil domain (MMRN2cc) and MMRN2 495-674. Both fragments were previously cloned into the pHLAvitag3 plasmid which provided a C-terminal His-tag and this recombinant protein was expressed in mammalian HEK293T cells. A secretable form of CLEC14A extracellular domain (ECD) was also previously constructed into a pIgPP vector which included a C-terminal human Fc-tag and had been modified to include a cleavable peptide between the protein and the Fc-tag, such that the CLEC14A ECD could be released from the Protein A-bound Fc domain. Purification techniques for both Fc-tagged and His-tagged proteins are described in sections 2.5.1 and 2.5.3 respectively.

In the instance of His-tagged purified proteins, samples were collected at each step of the purification and subjected to SDS-PAGE and Coomassie stained. The gel corresponding to the purification of the coiled-coil domain of MMRN2 (MMRN2cc) indicated that the protein successfully bound to the Nickel beads from the conditioned media (Figure 6.2A). The protein was then eluted off the beads in three elution steps. No bands were evident from the nickel beads post-elution indicating completed removal of the protein. The protein was subjected to buffer exchange via dialysis to remove imidazole from the sample and was subsequently analysed on the gel. The purified MMRN2cc ran as two bands between 100 and 130kDa, most likely being different glycoforms of MMRN2CC.

The MMRN2 495-674 fragment was purified in the same manner as MMRN2cc and samples were collected through the purification process and subjected to SDS-PAGE followed by

Coomassie staining. The results confirm that the MMRN2 495-674 protein was successfully collected on the nickel beads and eluted from the beads in three steps. The protein was then dialysed to remove the imidazole and remained intact post-dialysis as indicated by the similarities in the banding pre and post-dialysis. when comparing bands pre and post-dialysis (Figure 6.2B). In both instances, 1 µg and 5 µg of BSA were also run alongside to gauge the approximate protein concentration and overall yield. In the case of MMRN2 495-674 the purified protein was comparable to the band indicating 1µg of BSA, whilst the purified protein band of MMRN2cc was much fainter than the 1µg BSA band (Figure 6.2).

In the case of CLEC14A ECD-Fc protein purification (section 2.5.1), samples were collected at each step of the process and subjected to SDS-PAGE and Coomassie stained. Protein A beads bound CLEC14A ECD-Fc from conditioned media as yielding a band of ~120kDa (Figure 6.3). The protein was then cleaved from the beads using GST-tagged PreScission Protease (Sigma-Aldrich) and the Fc tag remained on the Protein A beads as indicated by the band at ~35kDa. The PreScission Protease was removed using glutathione-agarose. The resulting purified protein ran as a smeared band at ~90kDa, with variably sized protein likely due to different levels of glycosylation (Figure 6.2).

The purified CLEC14A ECD was then subjected to SDS-PAGE however the proteins were run in reducing and non-reducing conditions (Figure 6.4). Both proteins ran at 70kDa indicating no major discrepancies.

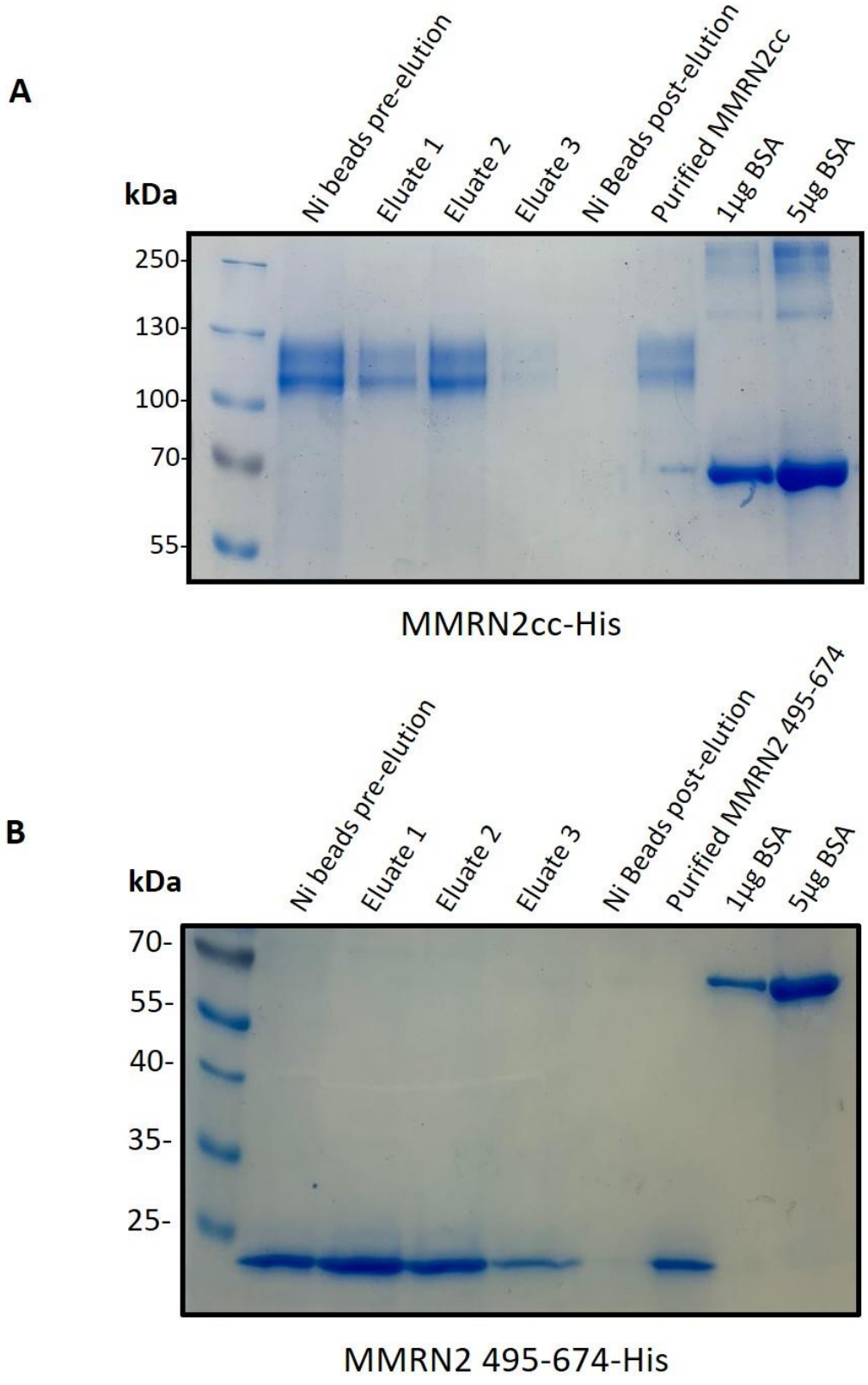


Figure 6.2. Coomassie stain gel showing purification steps of His-tagged MMRN2 proteins. (A) His-tagged MMRN2cc and (B) His-tagged MMRN2 495-674 proteins were purified from conditioned media using Ni-NTA beads. Each step of the purification process was run on a gel and Coomassie stained. Both proteins were eluted off the beads using imidazole in three elution steps and the remaining beads were stripped of protein. The purified protein post-dialysis was then run alongside 1µg and 5µg BSA controls.

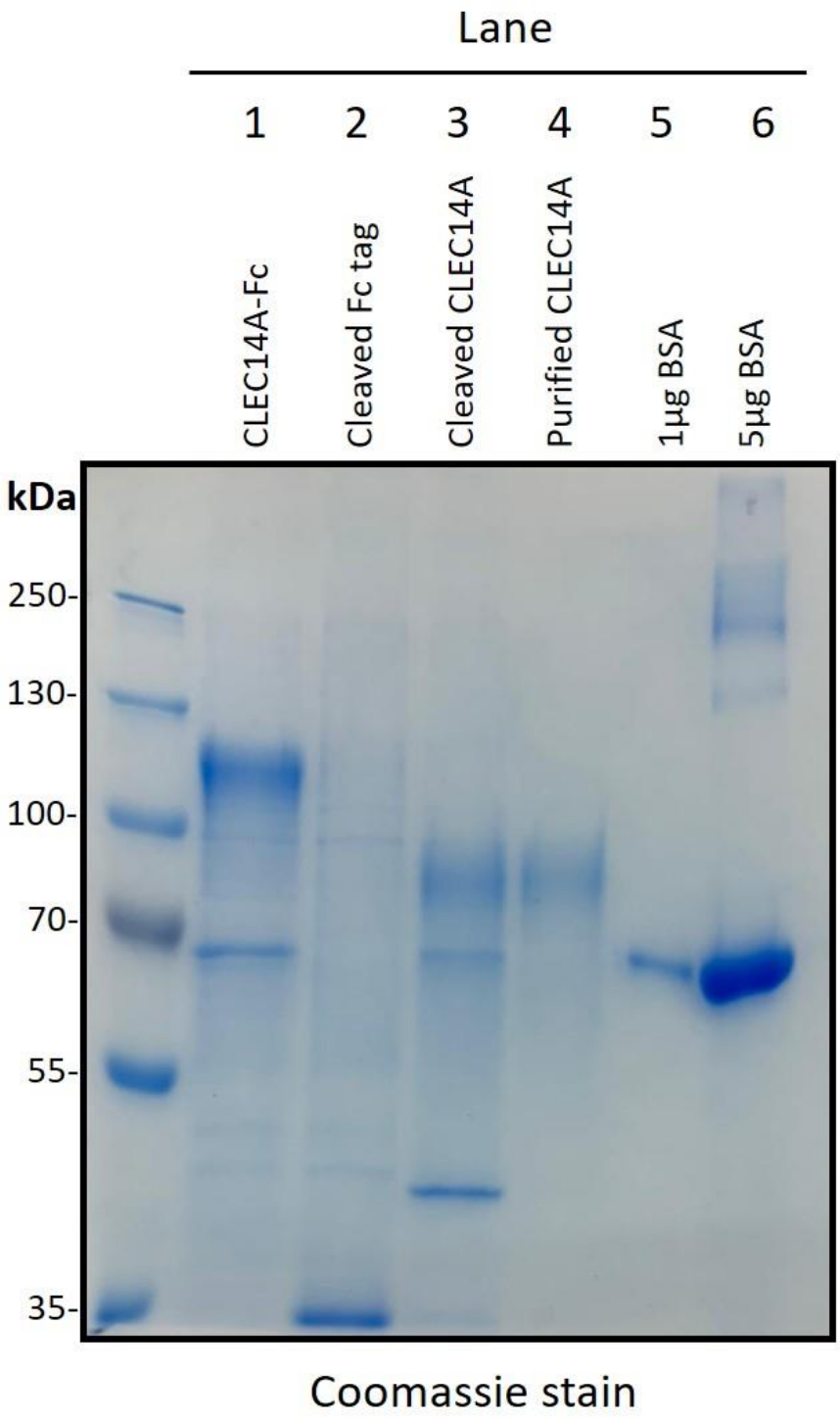


Figure 6.3. Coomassie stained gel indicating method of purification of CLEC14A ECD. CLEC14A ECD was expressed with an Fc tag and isolated from conditioned media using protein A beads (lane 1). The Fc tag was then cleaved from the protein (lane 2) using precision protease and CLEC14A ECD remained with the protease (lane 3). The precision protease was removed using Glutathione beads and CLEC14A ECD was then buffer exchanged into PBS (lane 4). 1µg and 5µg of BSA were run alongside (lanes 5 and 6).

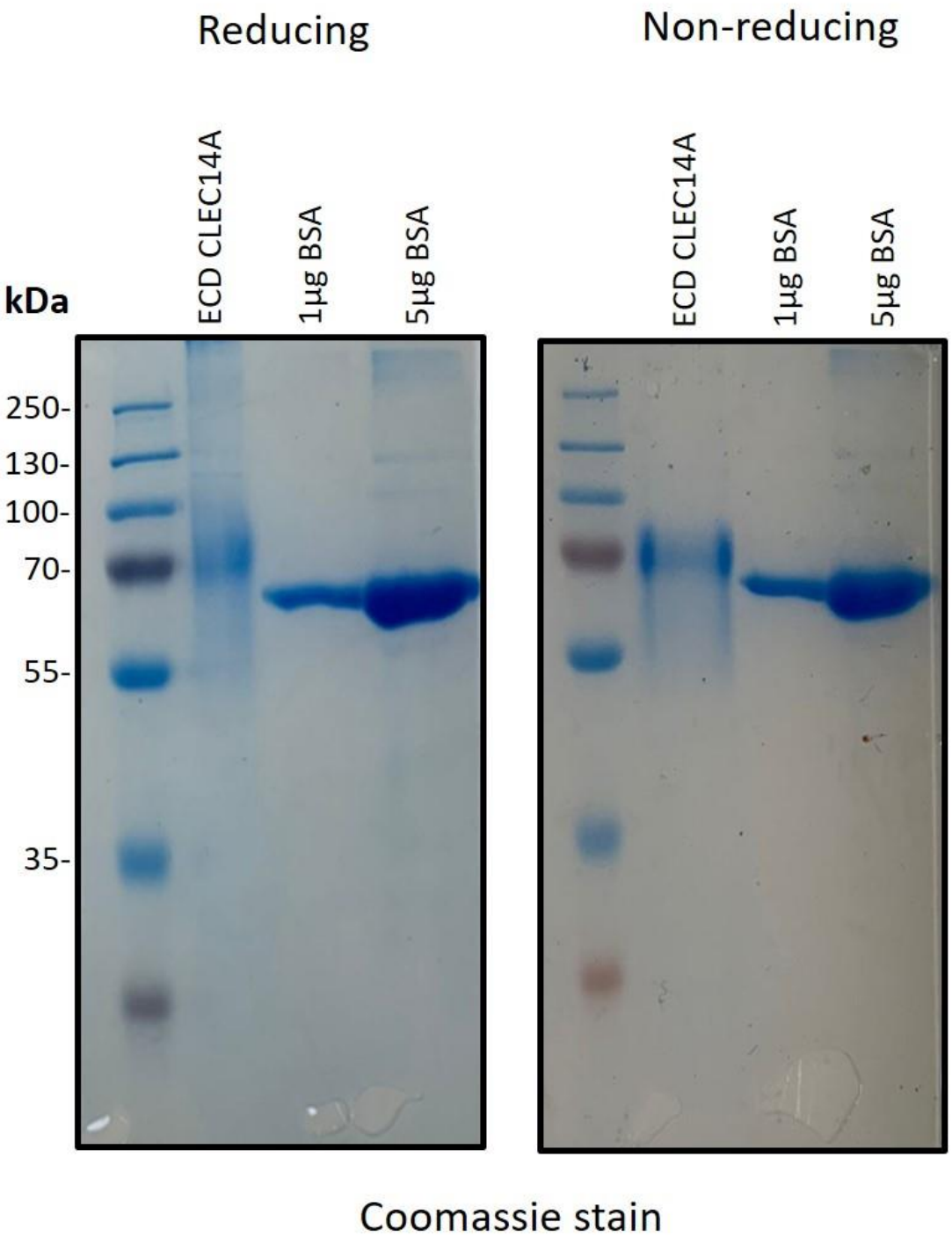


Figure 6.4. Coomassie stained gel indicating CLEC14A ECD cleaved from Fc tag. The cleaved CLEC14A ECD was treated with either reducing or non-reducing sample buffer. The proteins were run on separate gels alongside 1µg and 5µg of BSA.

6.3. Analytical Ultracentrifugation

The first biophysical technique employed was analytical Ultracentrifugation (AUC), a method used to study the aggregation, polymerization, stoichiometry and stability of the CLEC14A-MMRN2 complex in solution (Unzai, 2018). This method involves spinning samples at high speeds and monitoring protein sedimentation to yield insights into how the protein is folded, the oligomeric state of the protein as well as its molecular weight.

In this instance MMRN2 495-674 and CLEC14A ECD were studied individually and in combination. Each protein was purified in PBS therefore PBS was used in the reference cell. Two optical systems were recruited: the absorbance optic and the interference optic. The absorbance optic is less sensitive and involved monitoring the absorbance reading at 280nm of each section of the cell over the course of several minutes. Each section of the cell refers to the distance across the radius of the cell from the central rotating point, therefore as the protein sediments to the bottom of the cell the absorbance reading within each section changes over time. The rate of sedimentation is identified and the sedimentation coefficient is calculated which directly relates to molecular weight using the Svedberg equation. The smaller the sedimentation coefficient the lower the molecular weight. S was plotted on the X axis against protein concentration on the Y axis, which indicates which species are more abundant within the samples. The Interference optic was also used to studying the shape and size of a protein by studying the light paths generated after shining white light onto the sample. The light and dark spots generated from this image are called the fringes and are plotted on the Y axis against the sedimentation coefficient. Inference readings are more sensitive than absorbance readings therefore very little protein is required however contaminants are more likely detected and effect the data.

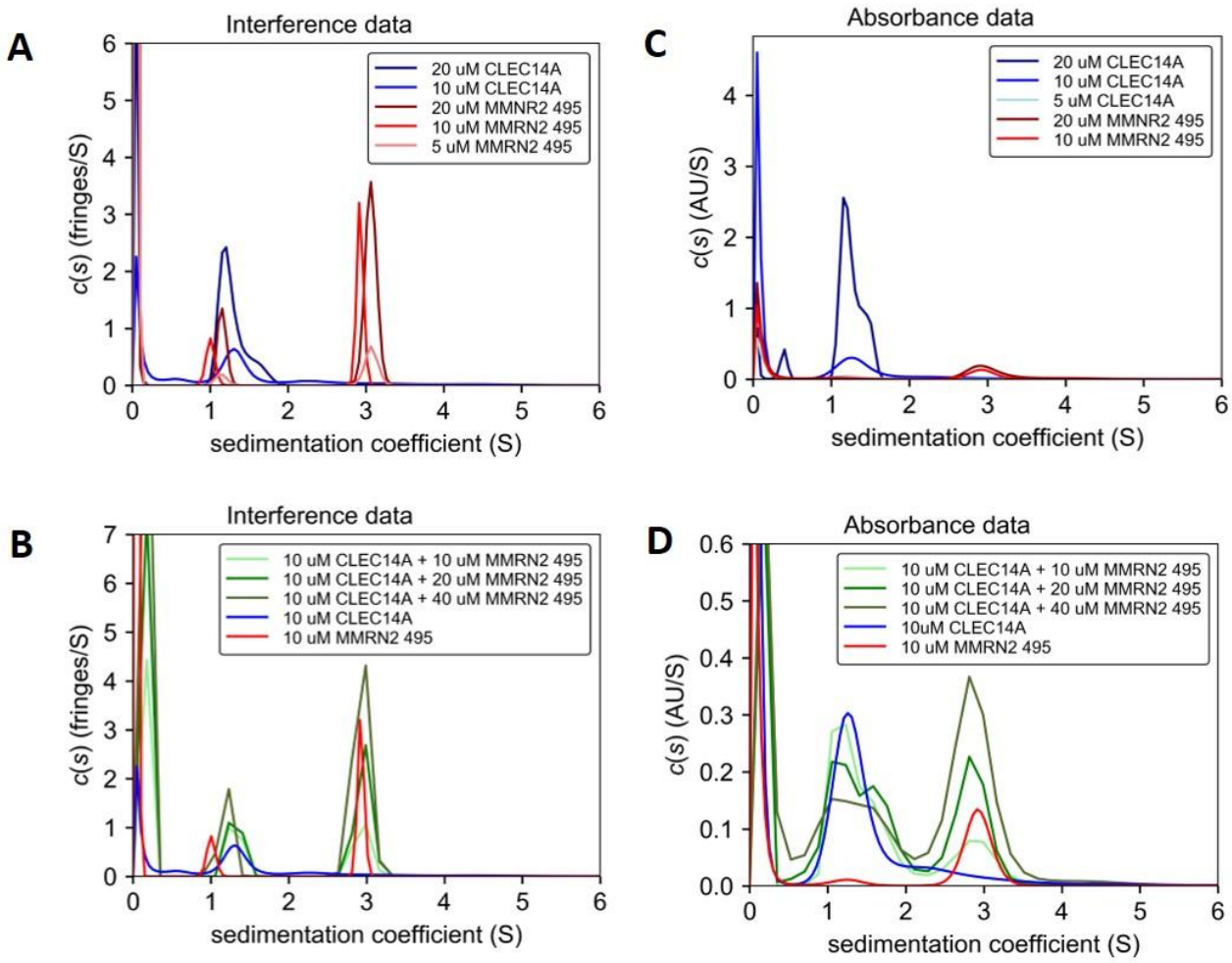


Figure 6.5. Analytical ultracentrifugation graphs indicating peaks from absorbance and interference optics of increasing amounts of CLEC14A ECD, MMRN2 495-674-his and a mixture of both. (A) Data from interference optics indicate 2 sharp peaks for CLEC14A with a sedimentation coefficient of approximately 0 and 1. MMRN2 showed sharp peaks present at 3. (B) The mixed sample did not differ nor were there major discrepancies when the amount of protein analysed increased. (C and D) Data using the absorbance optics was taken at 280nm and showed broad peaks present at similar sedimentation coefficients. Data presented was collect by Dr Gemma Harris at the Research complex at Harwell.

Initial readings were taken of each individual protein at different concentrations using both the absorbance and interference optics. The CLEC14A ECD protein was studied at 5 µM, 10 µM and 20 µM using the absorbance optics and 10 µM and 20 µM using the interference optics. Using the interference optics, a relatively broad peak at 1.3S was observed for the CLEC14A sample. This corresponded to a molecular weight of 17.4kDa (Figure 6.5A). The absorbance readings indicated that the 5 µM sample was not detected however both 10 µM and 20 µM were and the mirrored the peaks in the interference data (Figure 6.5B). This molecular weight does not represent what was visualised on the gel of CLEC14A ECD (Figure 6.4) and may indicate degradation or cleavage of CLEC14A which would need to be studied further.

MMRN2 495-674 was also studied using the interference and absorbance optics. 5 µM, 10 µM and 20 µM amounts were studied using the interference optics, whilst only 10 µM and 20 µM amounts were used for absorbance optics. The interference data indicated two peaks a larger one at 3s and a smaller on at 1S. In each instance the peak size also decreased as the concentration of the protein decreased. The shorter of the two peaks were present at approximately 1S which corresponds to 19.6kDa, this was concordant with the molecular weight of the band representing monomeric MMRN2 495-674 on the gel (Figure 6.2). The taller of two peaks were present at 3S which corresponds to a molecular weight of 59.2kDa which is almost triple the molecular weight of the monomer. This indicates that the MMRN2 fragment most likely trimerizes in solution and this form of MMRN2 is present at a higher proportion than the monomeric form. The detection from the absorbance optics was much lower and only the 10 µM and 20 µM samples were detected as very short peaks at 3S which represent trimeric MMRN2 (Figure 6.5A and C).

The proteins were then studied in combination wherein 10 µM of CLEC14A ECD was treated with increasing amounts of MMRN2 495-674 including 10 µM, 20 µM and 40 µM using both the absorbance and interference optics. 10 µM of CLEC14A ECD and 10 µM of MMRN2 495-674 were used as controls. The controls repeated what was previously seen where CLEC14A ECD observed a single peak at 1.3S. The MMRN2 control also indicated what was shown before with two peaks at 1S and 3S. The mixed samples indicated peaks at the same S values as the control for both absorbance and interference readings which indicate that CLEC14A and MMRN2 proteins did not form a complex. If a complex had formed a peak at a much higher sedimentation coefficient would have been detected (Figure 6.5B and D).

6.4. ELISA to study binding between CLEC14A and MMRN2

It has been previously shown in section 3.4 that the MMRN2 495-674 fragment can bind to CLEC14A-Fc in conditioned media (Figure 3.9) however in the AUC experiment MMRN2 495-674 did not bind to cleaved CLEC14A ECD (Figure 6.5B and D). In order to test the capacity of cleaved CLEC14A to bind to MMRN2 495-674 an ELISA was performed. This ELISA was set up using the same stock of MMRN2 495-674 that was studied in the AUC experiment and the purified CLEC14A ECD and conditioned media using CLEC14A-Fc were both generated from the same plasmid which was used to generate the proteins for the AUC experiment. If cleaved CLEC14A binds to MMRN2 495-674 coated on the plate, it would block binding between the incoming CLEC14A-Fc. MMRN2 495-674-His was coated on an ELISA plate and treated with a 2X molar excess of purified CLEC14A ECD or PBS. This was followed by treatment with conditioned media containing CLEC14A-Fc. The wells were then treated with an anti-Fc antibody conjugated to HRP followed by treatment with the substrate and read at 450nm. The negative control wells were as follows: wells coated with MMRN2 495-674 and treated

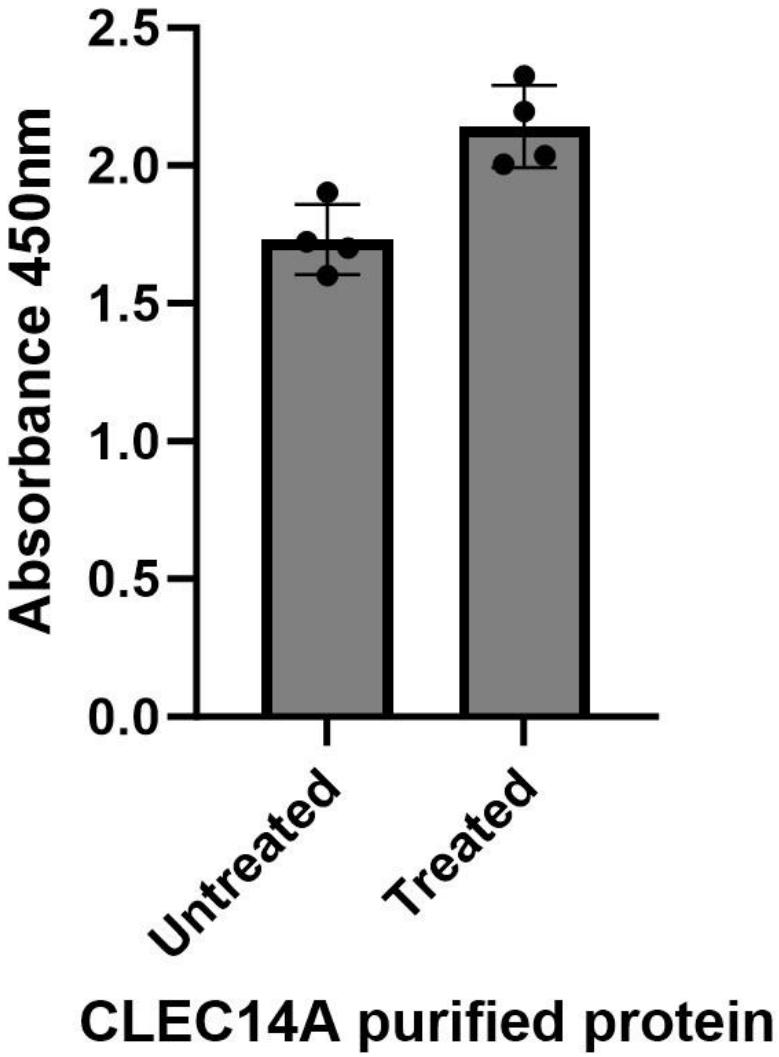


Figure 6.6. Purified ECD CLEC14A does not bind to MMRN2 495-674 in an ELISA. An ELISA plate was coated with 100ng MMRN2 495-674 and either treated with purified CLEC14A ECD or PBS. The wells were then treated with CLEC14A-Fc conditioned media. There was no difference in absorbance values.

with only PBS, uncoated wells treated with CLEC14A ECD alone, uncoated wells treated with CLEC14A-Fc alone and finally uncoated wells treated with only PBS. The results indicated no difference in absorbance between the wells treated with the purified CLEC14A-ECD versus untreated. Had the purified CLEC14A ECD bound to MMRN2 there would be a decrease in absorbance reading, instead there was a slight increased which means the proteins did not interact (Figure 6.6).

6.5. Further study of CLEC14A ECD

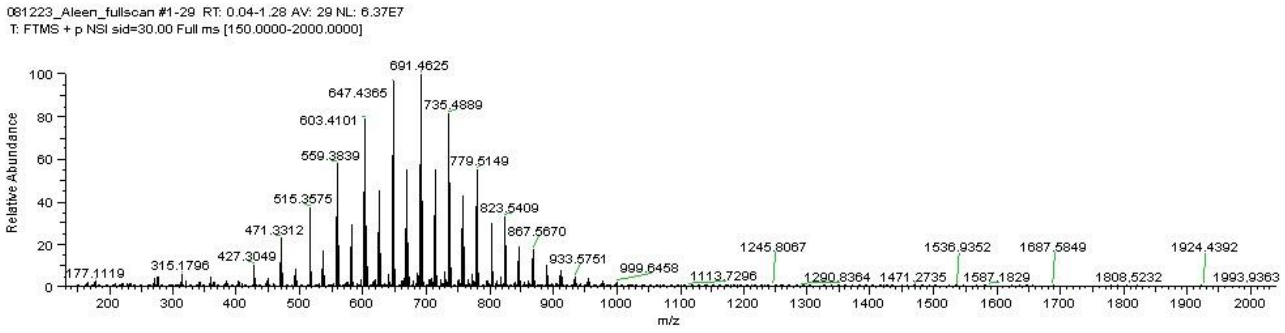
Due to the discrepancies in molecular weight of CLEC14A ECD on an SDS-gel compared to the AUC experiment (Figure 6.4 and 6.5), alternative methods of analysis were performed. Mass spectrometry analysis and gel filtration were used to further study the CLEC14A ECD protein in solution similarly to AUC.

6.5.1. Tandem intact mass spectrometry

Intact (top-down) tandem mass spectrometry was performed wherein the protein was kept intact and the heterogeneity of the sample can be studied. Within the resulting trace there is a low mass range and a high mass range which refers to the detectable range of the mass-to-charge (m/z) ratios. The high mass range generally refers to proteins and larger molecular structures whilst the lower mass range refers to peptides and small molecules (Donnelly *et al.,* 2019). CLEC14A, being a 90kDa protein would likely be present within the high mass range with peaks between 4000-6000m/z. The trace indicated a range of peaks within the low mass range between 440-900m/z which does not correspond to proteins and instead represents polymer peaks of contaminants. In the high mass range, there are no clear peaks which equate to CLEC14A, instead there is a broad signal between 2000-6000m/z. It may be the case

that this signal represents alternative glycoforms of CLEC14A or degraded CLEC14A however it may also be background signal (Figure 6.7).

Mass spectrum in low mass range



Mass spectrum in high mass range

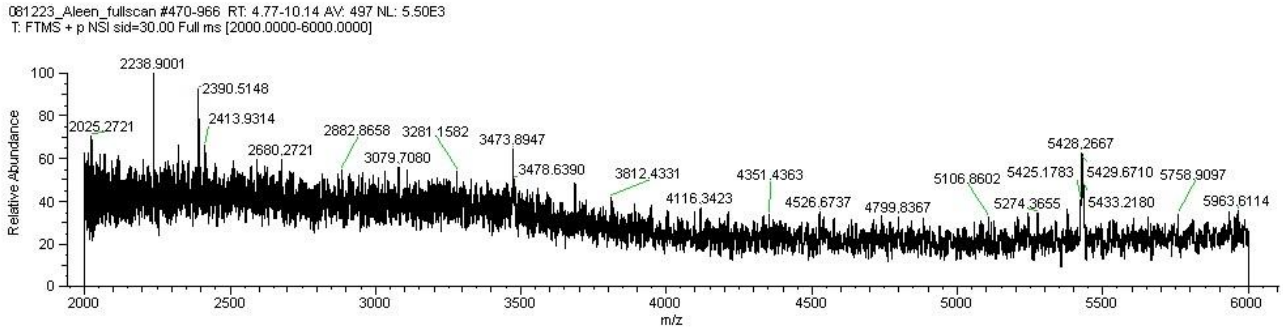


Figure 6.7. Top-down mass spec analysis of CLEC14A. Mass spec trace indicates some low molecular weight peaks within the low mass range. The high mass range revealed numerous peaks ranging from 2000-6000 m/z. Data presented was collect by Dr Jinglei Yu at the University of Birmingham.

6.5.2 Gel filtration

Since the analyses from the AUC, mass spectrometry analysis and SDS-PAGE each indicated different molecular weights, gel filtration was performed as another method to study CLEC14A ECD in solution. In the case of gel filtration, molecules are separated using porous beads based on size where larger molecules are unable to enter the pores of the beads and pass through the column quicker than smaller molecules which pass through the beads (O Fagain *et al.,* 2010). Gel filtration was performed on CLEC14A-ECD using the superdex200 column generally used for proteins between 10-600kDa. In this experiment, absorbance at 280 nm is plotted against elution volume and samples were collected in 0.5 mL fractions. Two peaks were detected in the trace, a peak single peak between 7.5 mL and 10 mL and another broader peak detected between 17 mL and 25 mL (Figure 6.8A). The first, more well-defined peak, according to a standard curve was approximately 677.9kDa which is much higher than the expected molecular weight of 90kDa and is within the void volume of the column and out of the resolution range.

The second broad peak was at the lower end of the calibration indicating a species less than ~2.84kDa in size and may be degraded protein or contaminants which absorb at 280nm. In this instance neither peak reflected the predicted molecular weight of CLEC14A ECD nor does it corroborate with the AUC result. Fractions 8-10 were subjected to western blotting and run alongside purified CLEC14A ECD. The results indicate that in each fraction CLEC14A was present and comparable to purified CLEC14A (Figure 6.8B).

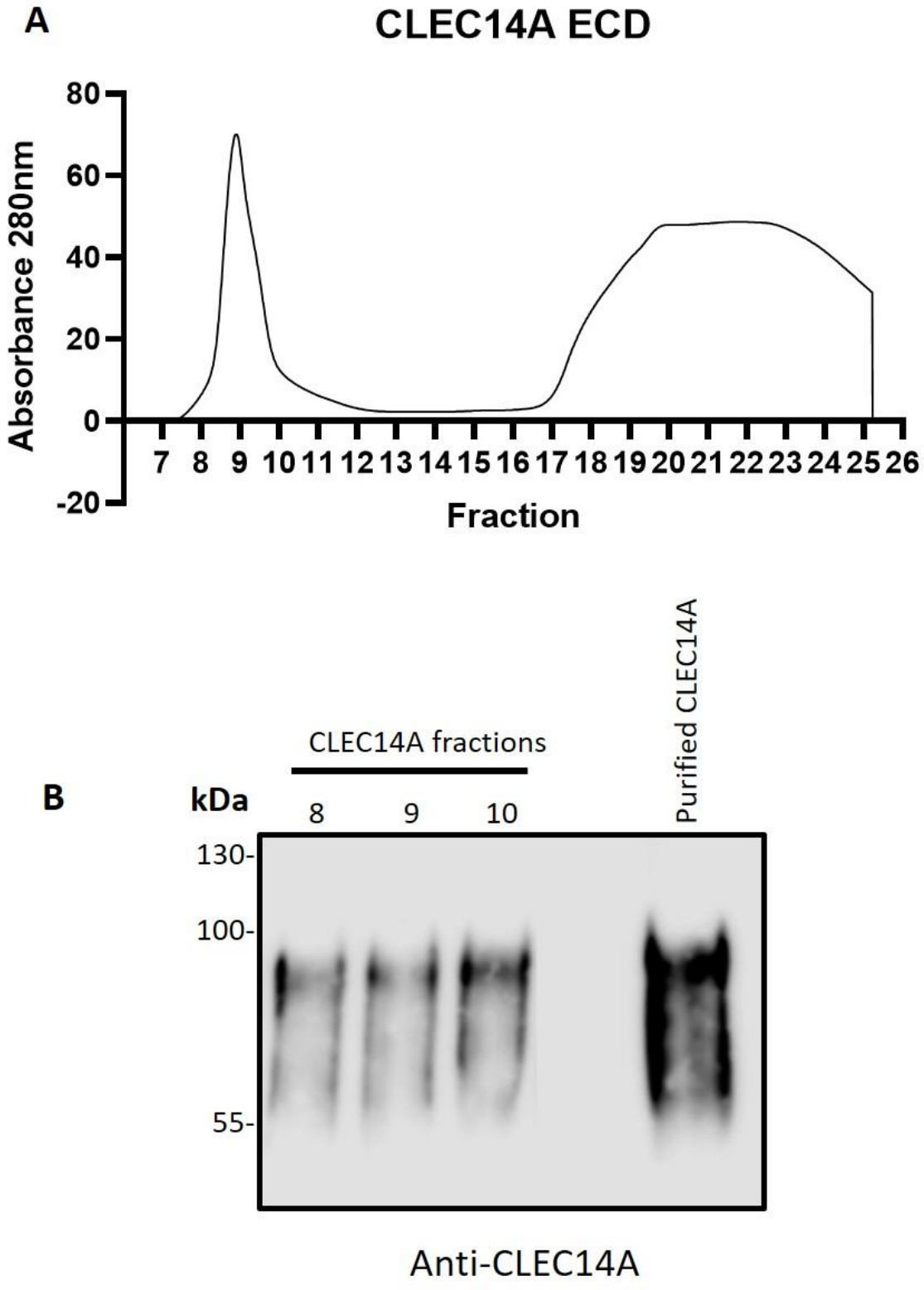


Figure 6.8. Gel filtration analysis of purified CLEC14A ECD and corresponding western blot. (A) Gel filtration was performed using the AKTA pure S200 column indicating a sharp peak between fractions 8-11 and a broad peak between 17-25. (B) Western blot of CLEC14A fraction 8-11 alongside purified CLEC14A probed using anti-CLEC14A antibody.

6.6. Discussion

In this section of the project biophysical techniques were employed to study CLEC14A and MMRN2. The CLEC14A extracellular domains (ECD), MMRN2 coiled-coil (CC) and 495-674 fragments were produced as purified proteins. MMRN2 495-674 and CLEC14A ECD were studied using Analytical ultracentrifugation to reveal MMRN2 495-674 trimerization. The CLEC14A ECD was shown to have a lower molecular weight than presented on an SDS-PAGE gel, therefore further analysis in the form of mass spectrometry and gel filtration were performed each of which indicated either no clear reference for CLEC14A ECD or excessively high molecular weights. The overall aim was to obtain an experimentally derived structure for the complex using Cryo-EM. However, CLEC14A ECD post-cleavage from the Fc-tag lost its binding function to MMRN2 495-674.

Purified proteins were produced using mammalian HEK293T cells. A common alternative method of protein production is the use of *E. coli* which serves as a cheaper, less time-consuming and higher yield method of production. The major drawback in this method and the reason it was not used in this study was because bacterial hosts lack endoplasmic reticulum and Golgi apparatus wherein key folding event and post translational modification occur such as formation of disulphide bonds and glycosylation (Overton, 2014; Ma *et al.,* 2020). The CTLD of CLEC14A which mediates interaction with most binding partners contains a disulphide bond which is crucial for the structural integrity of CLEC14A. Furthermore, both MMRN2CC and CLEC14A ECD are glycosylated (Khan *et al.,* 2017) thus posing an issue regarding bacterial expression of our constructs.

This study revealed through AUC that the MMRN2 495-674 fragment forms a homotrimer in solution. Whilst various algorithms have predicted MMRN2 trimerization through its coiled-

coil domain (Colombatti *et al.,* 2012) this is the first experimental data to confirm trimeric organisation of MMRN2. Currently in the literature there is experimental evidence showing closely related family members of the EMILIN-like family also form trimers such as EMILIN-1 and MMRN1 (Mongiat *et al.,* 2000; Hayward *et al.,* 1995). This expands our understanding of MMRN2 but moreover sheds some light onto its interaction with CLEC14A.

There were discrepancies in the molecular weight of purified CLEC14A ECD protein indicated by the different assays used in the study. The CLEC14A ECD purified protein on an SDS-PAGE gel revealed a band present at ~70-100kDa as previously seen in the literature (Khan *et al.,* 2017) and this was the same between reduced and non-reduced treated samples. The AUC data indicated one peak for the CLEC14A ECD calculated to 17.4kDa, which is a considerably lower molecular weight than predicted. Potentially the protein could have been degraded however, in our hands, the protein remains intact over a long period of time. The proteins were also transported to the Research Complex at Harwell and stored in a different facility prior to AUC analysis. Intact mass spec did not display a clear reference to CLEC14A. The multiple peaks in the high mass range most likely represent background noise and these signals may supress the signal of CLEC14A. Alternatively the protein signal may have been too low to be detected by the instrument. The final method employed to study CLEC14A ECD was gel filtration which revealed an excessively high molecular weight species at 677.9kDa representing CLEC14A. This was within the void volume of the column meaning it was too large to be resolved by the S200 column which is typically used to resolve complexes between 10kDa to 600kDa. It is important to note that the standard curve was produced using globular proteins and therefore assumes the protein to be spherical and compact in shape. This may not be the case for CLEC14A which may be elongated or irregular. Furthermore, glycosylation alters the hydrodynamic volume of a protein and may

indicate a larger apparent size resulting in discrepancies in molecular weight prediction (Kilgore *et al.,* 2020). It is plausible that CLEC14A formed higher-order polymers as it has previously been shown in the literature using size exclusion chromatography that CLEC14A ECD formed a trimer, and upon removal of the mucin-like region this trimerization was disturbed and CLEC14A was presented as a monomer (Sandoval *et al.,* 2020). The data shown in this study did not agree with the literature as the high molecular weight species shown to be CLEC14A ECD was excessively higher than a trimeric form of CLEC14A. However, the study by Sandoval *et al,* does provide an explanation for the higher order polymerization which may have been the case for the protein subjected to gel filtration.

The native organisation of soluble CLEC14A ECD is not yet well understood. Based on the data collected and the study from Sandoval *et al,* its most likely that in solution CLEC14A ECD forms a disorganised polymer and this is potentially due to the mucin-like region. A way to overcome this would be to produce recombinant CLEC14A which does not include the mucin-like region and therefore does not include glycosylation.

CLEC14A ECD-Fc could bind to MMRN2 495-674 however upon cleavage of the protein from the Fc-tag the protein lost its functional capacity to bind to MMRN2 495-674. The MMRN2 binding site is on the N-terminal CTLD of CLEC14A (Khan *et al.,* 2017) whilst the cleavage site is situated at the C-terminal of the protein, therefore, it was surprising that MMRN2 binding was affected following cleavage. The impact of cleavage on the folding of the CTLD has not been explored, though is unlikely that cleavage would affect the structure of CLEC14A, since the non-cleaved protein is fully able to bind MMRN2 in an ELISA. A pull-down assay of CLEC14A ECD post-cleavage using CRT4 could provide useful insights as the antibody only recognises the protein in its native, correctly folded form. If the CLEC14A ECD forms

polymers in solution, then possibly there could be steric hindrance preventing MMRN2 binding.

Further work in this project would involve confirming the cleaved form of CLEC14A does not bind to MMRN2, this could be done by performing a far western. The ELISA could also be designed differently to test for direct interactions between CLEC14A ECD and MMRN2 495-674 as oppose to studying any blocking activity of CLEC14A ECD. The fundamental drawback within this part of the project is that the monomeric form of CLEC14A does not bind to MMRN2, consequently an experimentally resolved structure of both proteins interacting could not be resolved. The MMRN2 495-674 protein was found to be trimeric in solution and this provides functional insights as well as structural understanding. Previously predicted models depicted MMRN2 as a monomeric protein and were therefore not accurate.

CHAPTER 7: Final discussion

CHAPTER 7: Final discussion

7.1. Brief overview

CLEC14A is a tumour endothelial marker known to interact with the large extracellular matrix protein MMRN2 (Khan et al., 2017). Blocking this interaction has been shown to inhibit tube formation and cell migration in vitro, and to reduce tumour size in vivo (Noy et al., 2015). This project aimed to expand on the current understanding of this interaction in order to understand the biology of CLEC14A and MMRN2 and target the interaction for therapeutic intervention.

7.2. Potential role of CLEC14A-MMRN2 interaction

The site directed mutagenesis approach recruited in this project identified S137, T139 and R141 of CLEC14A be critical in interacting with MMRN2. To date, the current understanding of the role of this interaction has been studied by the use of a MMRN2 protein fragment or the CRT4 CLEC14A-MMRN2 blocking antibody (Khan *et al.,* 2017; Noy *et al.,* 2015). The main drawback of using recombinant MMRN2 fragments is that this is also able to bind to CD93, since both CD93 and CLEC14A compete for binding the same region of MMRN2. It was also found in this project that the CRT4 antibody also blocks the interaction between CLEC14A and heparin as well as the interaction between CLEC14A and MMRN2. Therefore, the phenotypes observed *in vitro* and *in vivo* following CRT4 treatment could be due to blocking the interaction between CLEC14A and heparan sulfate interactions. Therefore, a more tailored method of targeting this interaction is required.

The use of iPSC vascular models expressing mutant CLEC14A or mouse models knocked in with mutant CLEC14A provide the best route of studying the effect of the interaction with

MMRN2 in a more tailored way compared to complete knockdown of CLEC14A or the use of the CRT4 antibody. The S137,T139,R141A CLEC14A mutant does however need to be tested for its capacity to bind to heat shock protein 70-1A. Recent studies have used induced pluripotent stem cell (iPSC) vascular models to investigate the effects of specific genetic mutations on angiogenesis. One notable study introduced the L914F TIE2 mutant into iPSCs, resulting in a line of cells that, when differentiated into endothelial cells, exhibited a constitutively activated TIE2 pathway. This mutation provided insights into the role of TIE2 in angiogenesis as the mutant displayed increased cell migration but decreased cell proliferation. There was also a decrease in alignment of cells to the direction of flow (Lazovic *et al.,* 2024).

In terms of the binding region and function a comprehensive overview of the most up to date information regarding CLEC14A binding proteins and antibodies is presented in Table 7.1.

Table 7.1. Overview of CLEC14A binding interactions. Displaying binding partners, binding domains and regions and binding effects following CLEC14A binding including references.

CLEC14Abinding partner or antibody	**CLEC14A binding region/epitope**	**Binding effect**
MMRN2	S137, T139 and R141 within CTLD (Current project, Chapter 3)	Promotes tube formation and cell migration in vitro Potentially promotes tumour growth in vivo Noy *et al.,* 2015; Khan *et al.,* 2017
Heparin	Partially R161 within CTLD. (Current project, Chapter 4)	Potentially promotes tumour growth *in vivo* (Noy *et al.,* 2017)
CRT1	CTLD (Puja Lodhia PhD thesis, 2017)	Blocks CLEC14A-MMRN2 (Kabir Khan PhD thesis, 2016)
CRT2	Between CTLD and SUSHI domain (Puja Lodhia PhD thesis, 2017)	Unknown
CRT3	CTLD (Puja Lodhia PhD thesis, 2017)	Decreased tube formation and cell migration. Triggers receptor internalisation (Puja Lodhia PhD thesis, 2017)
CRT4	R100 and R141 within CTLD (Current project, Chapter 4)	Blocks CLEC14A-MMRN2 interaction and CLEC14A-heparin interaction Reduces tumour size *in vivo* (Noy *et al.,* 2015)
CRT5	CTLD (Puja Lodhia PhD thesis, 2017)	Blocks CLEC14A-MMRN2 interaction (Kabir Khan PhD thesis, 2016)
Heat shock protein 70-1A	CTLD between A43-E69 (Jang *et al.,* 2017)	Promoted cell-cell contacts and tube formation *in vitro* (Jang *et al.,* 2017)
VEGFR3	Intracellular tail (Lee *et al.,* 2016)	Decreases VEGFR2 signalling (Lee *et al.,* 2016)
Facin	Intracellular tail (Puja Lodhia PhD thesis, 2017)	Unknown

7.3. Predictive modelling of CLEC14A and MMRN2

The AlphaFold predicted model of the CTLD of CLEC14A accurately displayed the loop-in-loop structure and predicted the solvent-facing residues which were confirmed with mutagenesis. In the case of MMRN2 however the docking model used indicated MMRN2 as a monomer, later in this project it was revealed that the MMRN2 495-674 fragment folds as a trimer. Therefore, the predictive docking model used was not completely accurate and a more accurate model using the AlphaFold3 server is displayed in Figure 7.1. Residues E617, E620, S623 and E624 of MMRN2 were mutated and did not block the interaction however when studying the new predicted CLEC14A-MMRN2 interface it appears that the three MMRN2 coils between amino acids A598 and H606 were also in close proximity to the CTLD. These coils are denoted in the figure as Coil A, Coil B and Coil C of which coils A and B were closest to the second loop of the CTLD (Figure 7.1B). Based on this model, residues L600 and L604 of Coil A and D602 and R605 of Coil B may directly interact with MMRN2.

Furthermore, it has previously been shown in the literature using gel filtration that CLEC14A ECD may potentially form a trimer in solution (Sandoval *et al.,* 2020). In the instance that CLEC14A ECD does form a trimer in solution an AlphaFold model was developed to reflect this using the CTLD for simplicity (Figure 7.2A). The model suggested that, for each CLEC14A molecule, the MMRN2 binding sites face in opposite directions. It is unlikely that membrane-bound CLEC14A exists as a trimer in this conformation however post cleavage by RHBDL2 (Noy *et al.,* 2016), the resulting circulating CLEC14A ECD could and may exhibit a higher binding efficacy than membrane-bound CLEC14A. In the instance that a trimeric form of CLEC14A binds to trimeric MMRN2 an AlphaFold predicted model was generated (Figure 7.2B) which indicated each CLEC14A molecule independently interacted with each coil of the MMRN2

trimer. In this model CLEC14A did not display as a trimer and instead as three separate molecules bound to trimeric MMRN2.

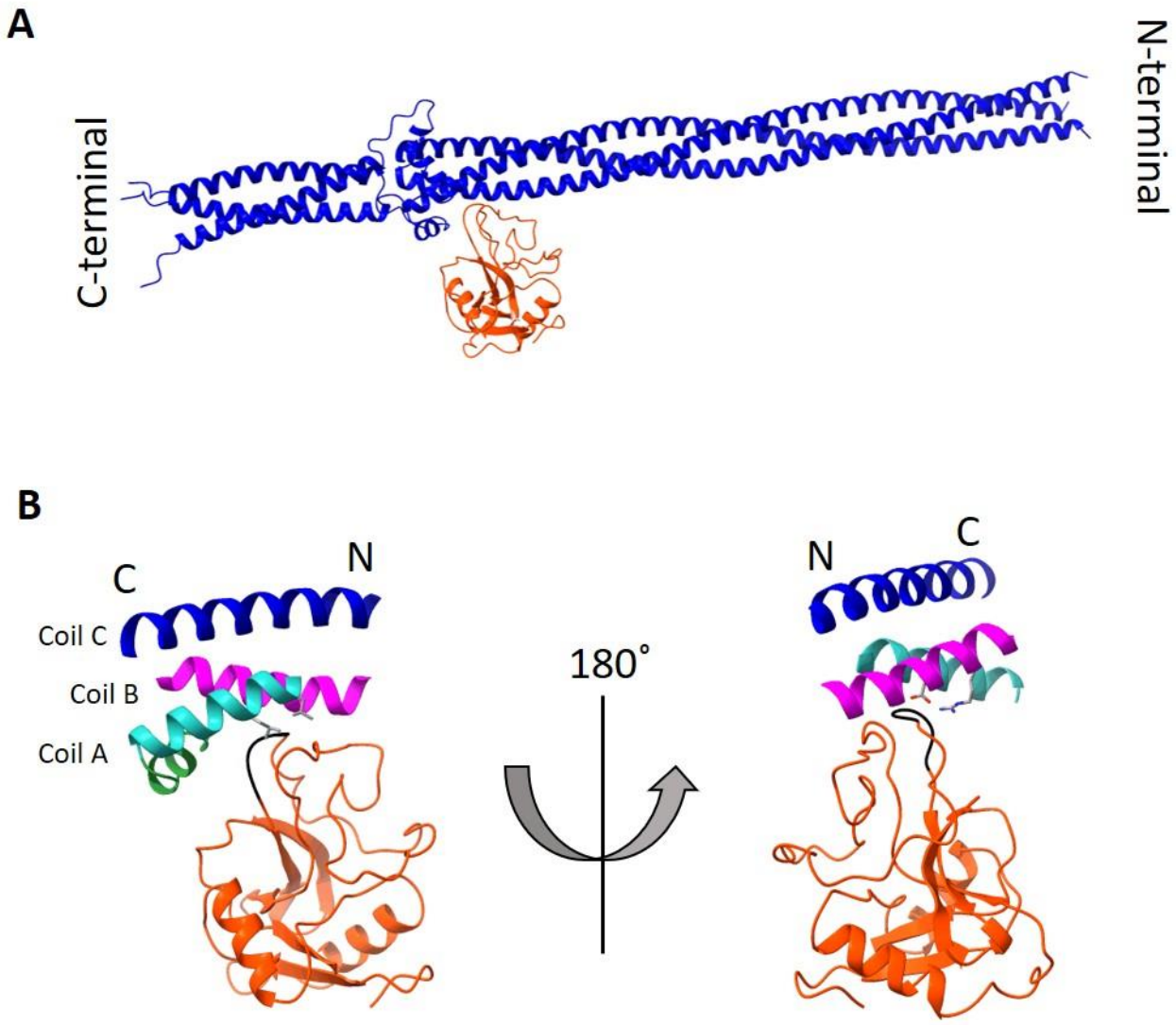


Figure 7.1. The AlphaFold predicted docking model of the CTLD of CLEC14A with MMRN2 fragments displayed as trimers viewed in Chimera. (A) The model displays the CTLD of CLEC14A (orange) and MMRN2 495-674 trimer (dark blue) labelled to indicate the C and N terminal. (B) Highlights the CTLD of CLEC14A (orange) and second loop highlighted (black) with shorter fragments of the three coils of MMRN2 between amino acids A598 and H606. The models indicate the three coils labelled Coil A (light blue and green to indicate the previously mutated region), Coil B (magenta) and Coil C (dark blue). The model is displayed in front view where residues L600 and L604 are highlighted shown on Coil A. The model is also displayed in back view where residues D602 and R605 are displayed on Coil B (magenta).

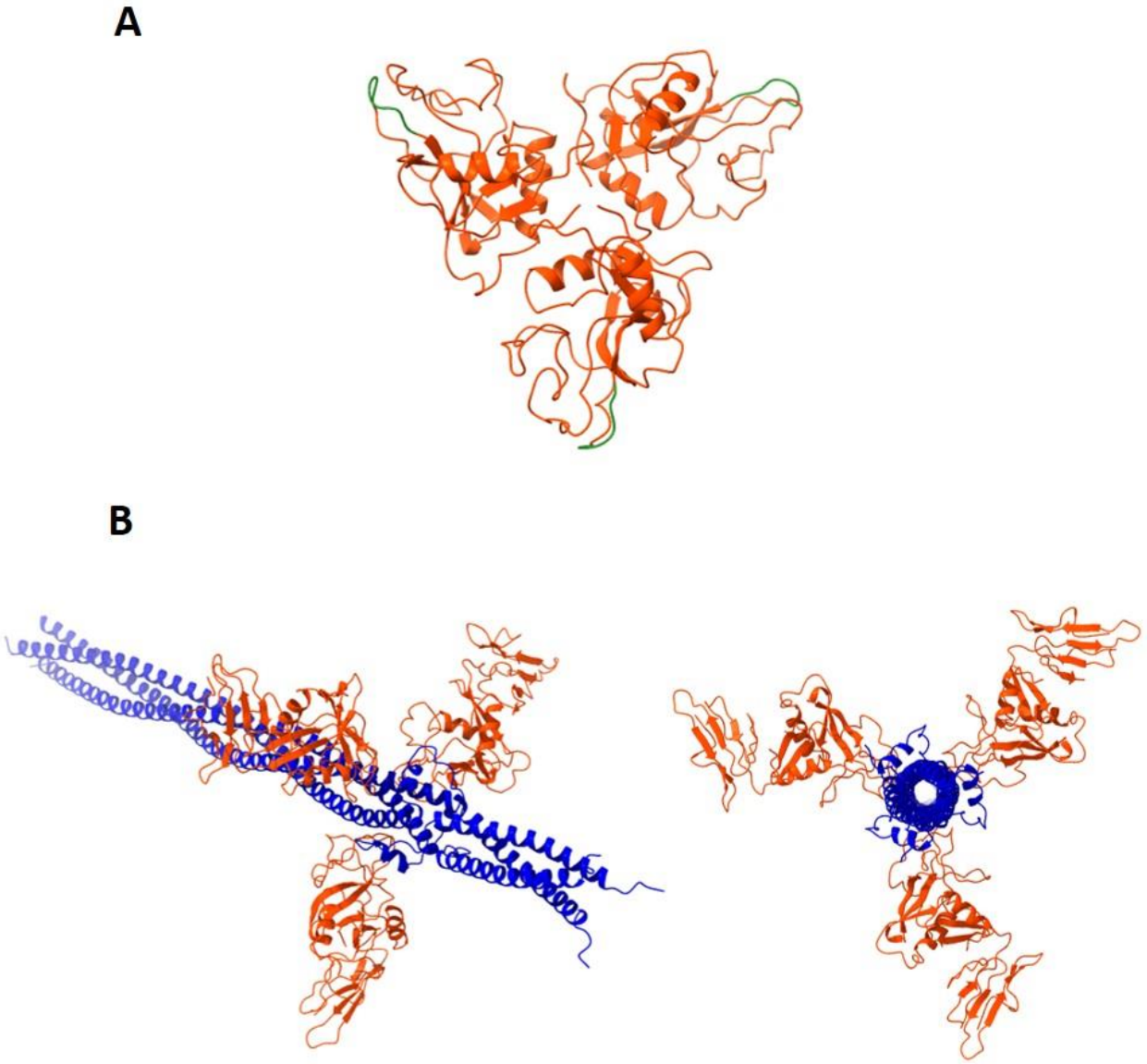


Figure 7.2. The AlphaFold predicted models of trimeric forms of CLEC14A and MMRN2. (A) The model displays the CTLD of CLEC14A in a trimeric form where the MMRN2 binding site is highlighted in green. (B) The docking model of trimeric MMRN2 495-674 (blue) and three molecules of the CTLD and sushi domains of CLEC14A.

It is also important to note that since the biophysical assays performed on CLEC14A were unsuccessful, the native physical orientation of CLEC14A is as of yet unknown. It has been shown in the literature that CD93 forms a homodimer which increases its capacity to bind to MMRN2 (Barbera *et al.,* 2023). The CTLD and SUSHI domain of CD93 drives this anti-parallel dimer which forms an arginine-rich positive core which houses a sulfate ion. Whilst we did not study CLEC14A dimerization in this project we found CLEC14A, fused to a human Fc-tag, was functional in binding to MMRN2 and upon cleavage from the tag, binding was lost. It may be the case that the Fc-tag promoted CLEC14A dimerization which may be critical for its function and this capacity to dimerize is lost when the protein is free in solution. It also may be the case that CLEC14A could bind to ligands more effectively when fused to an Fc-tag or on the cell membrane due to increased avidity. It has been shown in the literature that fragmented forms of CLEC14A as short as the CTLD and SUSHI domain are functional (Khan *et al.,* 2017). Therefore, an AlphaFold model was generated which depicts dimerization of these two domains (Figure 7.3). The model differs from the resolved structure of CD93 as it does not form an anti-parallel dimer nor does it form a positive arginine-rich core. The aligning residues from the positive core are present on opposite sides of the structure and are solvent-facing. However, within the model it is depicted that this dimerization involved “domain switching”. Domain switching is a phenomenon by which sequences or part of two identical proteins are exchanged or fuse together to exert or optimise their function (Mascarenhas and Gosavi, 2017). If this is the case this would provide an alternative method for targeting CLEC14A, by inhibiting domain switching. Domain switching upon to promote dimerization and decreased function following Fc-tag cleavage due to loss of capacity to dimerise has been shown in the literature by other membrane proteins such as GPVI (Slater *et al.,* 2021)

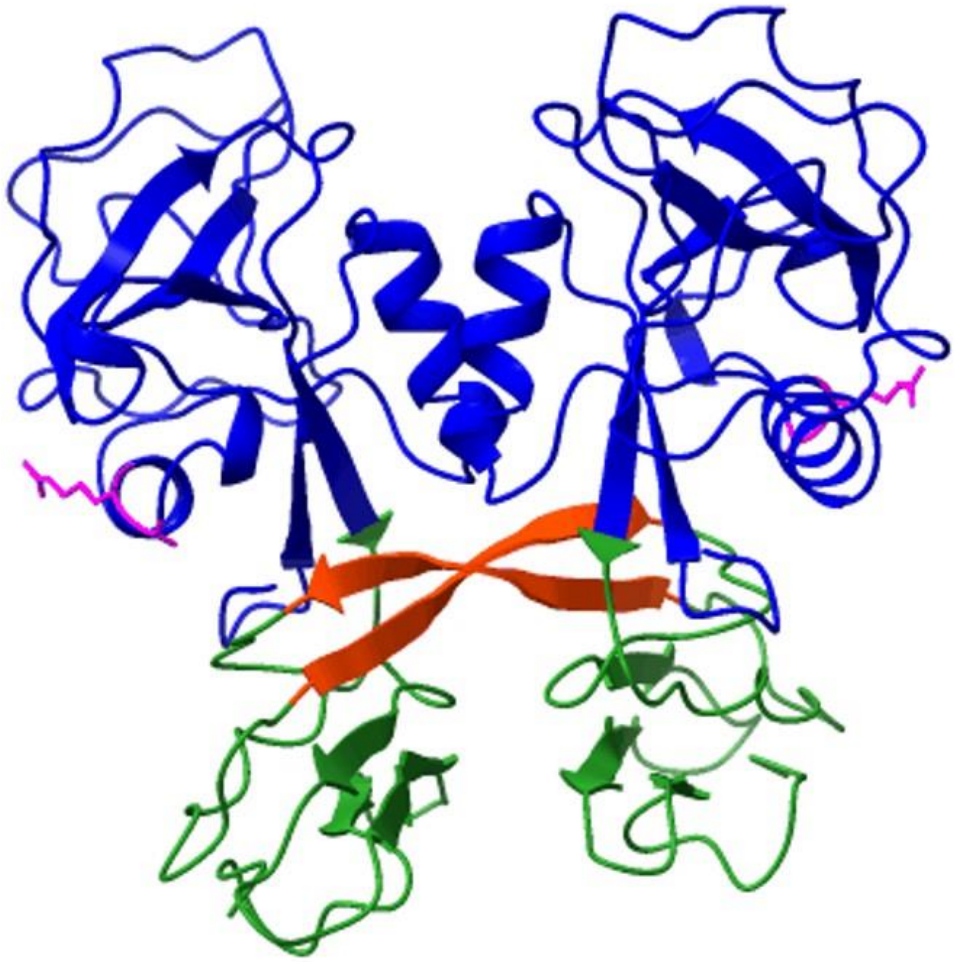


Figure 7.3. The AlphaFold predicted model of CLEC14A CTLD and SUSHI domain dimerization. The model displays the CTLD (blue) and SUHSI domain (green) of CLEC14A dimerising with part of the SUHSI domain engaging in domain switching (orange). The model also highlights residues R78 and A79 of each protein presented on opposite sides of the dimer facing outwards.

7.4. Future plans

Moving forward from this project, several important areas of research are needed to deepen our understanding of CLEC14A and its interactions. First, it will be essential to identify the specific residues of CLEC14A that directly interact with heparin, as this may reveal critical insights into the molecular dynamics of these interactions. Additionally, investigating whether the MMRN2 495-674 fragment can block the interaction between CLEC14A and heparin will help clarify the role of this region in modulating CLEC14A function. Another priority will be to examine the impact of CRT antibodies on the interaction between CLEC14A and HSP70-1A, which could provide further information about regulatory mechanisms. Finally, studying the effects of CLEC14A mutants that do not bind to MMRN2 in functional assays, such as tube formation, cell migration, and in vivo studies (e.g., using CRISPR genetically modified mice), will shed light on the functional significance of these interactions and their potential implications for vascular biology.

7.5. Concluding remarks

In summary, the data presented in this thesis expands on our understanding of the CLEC14A-MMRN2 interaction in terms of location and folding. The CTR4 antibody was also shown to be multifunctional which opens up a range of possibilities when studying CLEC14A. The use of AI modelling in the form of AlphaFold to predict protein structure has been instrumental in this project and the data presented reflects both the benefits and inaccuracies of computational modelling. Following on from this project there are many avenues which can be taken regarding the study of this interaction. Most importantly the CLEC14A mutants which block the interaction with MMRN2 must be studied for their capacities to bind to other CLEC14A

binding partners. Dimerization studies of CLEC14A would also be useful in expanding our understanding of CLEC14A orientation and folding in physiology.

References

An, D., Banerjee, S. and Lee, J.-M. (2021) Recent advancements of antiangiogenic combination therapies in ovarian cancer. *Cancer treatment reviews*, 98: 102224. doi:10.1016/j.ctrv.2021.102224.

Andreuzzi, E., Capuano, A., Pellicani, R., et al. (2018) Loss of Multimerin-2 and EMILIN-2 Expression in Gastric Cancer Associate with Altered Angiogenesis. *International Journal of Molecular Sciences*, 19 (12): E3983. doi:10.3390/ijms19123983.

Andreuzzi, E., Colladel, R., Pellicani, R., et al. (2017) The angiostatic molecule Multimerin 2 is processed by MMP-9 to allow sprouting angiogenesis. *Matrix Biology: Journal of the International Society for Matrix Biology*, 64: 40–53. doi:10.1016/j.matbio.2017.04.002.

Ashikari-Hada, S., Habuchi, H., Kariya, Y., et al. (2005) Heparin Regulates Vascular Endothelial Growth Factor165-dependent Mitogenic Activity, Tube Formation, and Its Receptor Phosphorylation of Human Endothelial Cells. *Journal of Biological Chemistry*, 280 (36): 31508–31515. doi:10.1074/jbc.M414581200.

Barbera, S., Raucci, L., Tassone, G., et al. (2023) Dimerization of the C-type lectin-like receptor CD93 promotes its binding to Multimerin-2 in endothelial cells. *International Journal of Biological Macromolecules*, 224: 453–464. doi:10.1016/j.ijbiomac.2022.10.136.

Beretov, J., Wasinger, V.C., Millar, E.K.A., et al. (2015) Proteomic Analysis of Urine to Identify Breast Cancer Biomarker Candidates Using a Label-Free LC-MS/MS Approach. *PloS One*, 10 (11): e0141876. doi:10.1371/journal.pone.0141876.

Bora, A., Ubaida Mohien, C., Chaerkady, R., et al. (2014) Identification of putative biomarkers for HIV-associated neurocognitive impairment in the CSF of HIV-infected patients under cART therapy determined by mass spectrometry. *Journal of Neurovirology*, 20 (5): 457–465. doi:10.1007/s13365-014-0263-5.

Carmeliet, P., 2003. Angiogenesis in health and disease. *Nature medicine*, *9*(6), pp.653-660.

Chen, L., Deng, H., Cui, H., et al. (2017) Inflammatory responses and inflammation-associated diseases in organs. *Oncotarget*, 9 (6): 7204–7218. doi:10.18632/oncotarget.23208.

Chen, X., Zaro, J. and Shen, W.-C. (2013) Fusion Protein Linkers: Property, Design and Functionality. *Advanced drug delivery reviews*, 65 (10): 1357–1369. doi:10.1016/j.addr.2012.09.039.

Chiodelli, P., Bugatti, A., Urbinati, C., et al. (2015) Heparin/Heparan Sulfate Proteoglycans Glycomic Interactome in Angiogenesis: Biological Implications and Therapeutical Use. *Molecules*, 20 (4): 6342–6388. doi:10.3390/molecules20046342.

Christian, S., Ahorn, H., Novatchkova, M., et al. (2001) Molecular cloning and characterization of EndoGlyx-1, an EMILIN-like multisubunit glycoprotein of vascular

endothelium. *The Journal of Biological Chemistry*, 276 (51): 48588–48595. doi:10.1074/jbc.M106152200.

Clark, J., Neagoe, R.A.I., Zuidscherwoude, M., et al. (2021) Evidence that GPVI is expressed as a mixture of monomers and dimers, and that the D2 domain is not essential for GPVI activation: GPVI dimerisation is not critical for ligand binding. *Thrombosis and Haemostasis*, 2021 (00): 1–26. doi:10.1055/a-1401-5014.

Colladel, R., Pellicani, R., Andreuzzi, E., et al. (2016) MULTIMERIN2 binds VEGF-A primarily via the carbohydrate chains exerting an angiostatic function and impairing tumor growth. *Oncotarget*, 7 (2): 2022–2037. doi:10.18632/oncotarget.6515.

Colombatti, A., Spessotto, P., Doliana, R., et al. (2012) The EMILIN/Multimerin Family. *Frontiers in Immunology*, 2: 93. doi:10.3389/fimmu.2011.00093.

Cooley, R., Kara, N., Hui, N.S., et al. (2020) Development of a cell-free split-luciferase biochemical assay as a tool for screening for inhibitors of challenging protein-protein interaction targets. *Wellcome Open Research*, 5: 20. doi:10.12688/wellcomeopenres.15675.1.

Croix, B.St., Rago, C., Velculescu, V., et al. (2000) Genes Expressed in Human Tumor Endothelium. *Science*, 289 (5482): 1197–1202. doi:10.1126/science.289.5482.1197.

Dale, N.C., Johnstone, E.K.M., White, C.W., et al. (2019) NanoBRET: The Bright Future of Proximity-Based Assays. *Frontiers in Bioengineering and Biotechnology*, 7: 56. doi:10.3389/fbioe.2019.00056.

Daugaard, M., Rohde, M. and Jäättelä, M. (2007) The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions. *FEBS Letters*, 581 (19): 3702–3710. doi:10.1016/j.febslet.2007.05.039.

Dixon, A.S., Schwinn, M.K., Hall, M.P., et al. (2016) NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chemical Biology*, 11 (2): 400–408. doi:10.1021/acschembio.5b00753.

Duffy, A.M., Bouchier-Hayes, D.J. and Harmey, J.H. (2013) "Vascular Endothelial Growth Factor (VEGF) and Its Role in Non-Endothelial Cells: Autocrine Signalling by VEGF." In *Madame Curie Bioscience Database [Internet]*. Landes Bioscience. Available at: https://www.ncbi.nlm.nih.gov/books/NBK6482/.

Elice, F. and Rodeghiero, F. (2012) Side effects of anti-angiogenic drugs. *Thrombosis Research*, 129 Suppl 1: S50-53. doi:10.1016/S0049-3848(12)70016-6.
England, C.G., Ehlerding, E.B. and Cai, W. (2016) NanoLuc: A Small Luciferase is Brightening up the Field of Bioluminescence. *Bioconjugate chemistry*, 27 (5): 1175–1187. doi:10.1021/acs.bioconjchem.6b00112.

Esparza, A., Jimenez, N., Borrego, E.A., et al. (2024) Review: Human stem cell-based 3D in vitro angiogenesis models for preclinical drug screening applications. *Molecular Biology Reports*, 51 (1): 260. doi:10.1007/s11033-023-09048-2.

Fejza, A., Poletto, E., Carobolante, G., et al. (2021) Multimerin-2 orchestrates the cross-talk between endothelial cells and pericytes: A mechanism to maintain vascular stability. *Matrix Biology Plus*, 11: 100068. doi:10.1016/j.mbplus.2021.100068.

Finch, N.C., Neal, C.R., Welsh, G.I., et al. (2023) The unique structural and functional characteristics of glomerular endothelial cell fenestrations and their potential as a therapeutic target in kidney disease. *American Journal of Physiology-Renal Physiology*, 325 (4): F465–F478. doi:10.1152/ajprenal.00036.2023.

Folkman, J. (1971) Tumor angiogenesis: therapeutic implications. *The New England Journal of Medicine*, 285 (21): 1182–1186. doi:10.1056/NEJM197111182852108.

Fountzilas, G., Kourea, H.P., Bobos, M., et al. (2011) Paclitaxel and Bevacizumab as First Line Combined Treatment in Patients with Metastatic Breast Cancer: The Hellenic Cooperative Oncology Group Experience with Biological Marker Evaluation. *Anticancer Research*, 31 (9): 3007–3018.

Fuster, M.M. and Wang, L. (2010) Endothelial Heparan Sulfate in Angiogenesis. *Progress in molecular biology and translational science*, 93: 179–212. doi:10.1016/S1877-1173(10)93009-3.

Galvagni, F., Nardi, F., Spiga, O., et al. (2017) Dissecting the CD93-Multimerin 2 interaction involved in cell adhesion and migration of the activated endothelium. *Matrix Biology*, 64: 112–127. doi:10.1016/j.matbio.2017.08.003.

Granger, D.N. and Senchenkova, E. (2010) "Leukocyte–Endothelial Cell Adhesion." In *Inflammation and the Microcirculation*. Morgan & Claypool Life Sciences. Available at: https://www.ncbi.nlm.nih.gov/books/NBK53380/.

Hanahan, D. and Weinberg, R.A. (2011) Hallmarks of Cancer: The Next Generation. *Cell*, 144 (5): 646–674. doi:10.1016/j.cell.2011.02.013.

Hayward, C.P., Hassell, J.A., Denomme, G.A., et al. (1995) The cDNA sequence of human endothelial cell multimerin. A unique protein with RGDS, coiled-coil, and epidermal growth factor-like domains and a carboxyl terminus similar to the globular domain of complement C1q and collagens type VIII and X. *The Journal of Biological Chemistry*, 270 (31): 18246–18251. doi:10.1074/jbc.270.31.18246.

Hida, K., Maishi, N., Takeda, R., et al. (2022) "The Roles of Tumor Endothelial Cells in Cancer Metastasis." In Sergi, C.M. (ed.) *Metastasis*. Brisbane (AU): Exon Publications. Available at: http://www.ncbi.nlm.nih.gov/books/NBK580883/.

Ho, M., Yang, E., Matcuk, G., et al. (2003) Identification of endothelial cell genes by combined database mining and microarray analysis. *Physiological Genomics*, 13 (3): 249–262. doi:10.1152/physiolgenomics.00186.2002.
Huber, M.A., Kraut, N., Schweifer, N., et al. (2006) Expression of stromal cell markers in distinct compartments of human skin cancers. *Journal of Cutaneous Pathology*, 33 (2): 145–155. doi:10.1111/j.0303-6987.2006.00446.x.

Hynes, R.O. (2009) Extracellular matrix: not just pretty fibrils. *Science (New York, N.Y.)*, 326 (5957): 1216–1219. doi:10.1126/science.1176009.

Jain, R.K. (2001) Normalizing tumor vasculature with anti-angiogenic therapy: A new paradigm for combination therapy. *Nature Medicine*, 7 (9): 987–989. doi:10.1038/nm0901-987.

Jang, J., Kim, M.R., Kim, T.-K., et al. (2017) CLEC14a-HSP70-1A interaction regulates HSP70-1A-induced angiogenesis. *Scientific Reports*, 7: 10666. doi:10.1038/s41598-017-11118-y.
Jisna, V.A. and Jayaraj, P.B. (2021) Protein Structure Prediction: Conventional and Deep Learning Perspectives. *The Protein Journal*, 40 (4): 522–544. doi:10.1007/s10930-021-10003-y.

Jumper, J., Evans, R., Pritzel, A., et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, 596 (7873): 583–589. doi:10.1038/s41586-021-03819-2.

Kadry, H., Noorani, B. and Cucullo, L. (2020) A blood–brain barrier overview on structure, function, impairment, and biomarkers of integrity. *Fluids and Barriers of the CNS*, 17 (1): 69. doi:10.1186/s12987-020-00230-3.

Khan, K.A. (n.d.) *Kabir- Investigating the extracellular interactions of the tumour endothelial marker*., p. 316.

Khan, K.A., McMurray, J.L., Mohammed, F., et al. (2019) C-type lectin domain group 14 proteins in vascular biology, cancer and inflammation. *The FEBS Journal*, 286 (17): 3299–3332. doi:10.1111/febs.14985.

Khan, K.A., Naylor, A.J., Khan, A., et al. (2017) Multimerin-2 is a ligand for group 14 family C-type lectins CLEC14A, CD93 and CD248 spanning the endothelial pericyte interface. *Oncogene*, 36 (44): 6097–6108. doi:10.1038/onc.2017.214.

Kilgore, H.R., Latham, A.P., Ressler, V.T., et al. (2020) Structure and Dynamics of N-Glycosylated Human Ribonuclease 1. *Biochemistry*, 59 (34): 3148–3156. doi:10.1021/acs.biochem.0c00191.

Kilpatrick, L.E., Friedman-Ohana, R., Alcobia, D.C., et al. (2017) Real-time analysis of the binding of fluorescent VEGF 165 a to VEGFR2 in living cells: Effect of receptor tyrosine kinase inhibitors and fate of internalized agonist-receptor complexes. *Biochemical Pharmacology*, 136: 62–75. doi:10.1016/j.bcp.2017.04.006.

Kim, T.-K., Na, H.J., Lee, W.R., et al. (2016) Heat shock protein 70-1A is a novel angiogenic regulator. *Biochemical and Biophysical Research Communications*, 469 (2): 222–228. doi:10.1016/j.bbrc.2015.11.125.

Kim, Y., Lee, S., Zhang, H., et al. (2020) CLEC14A deficiency exacerbates neuronal loss by increasing blood-brain barrier permeability and inflammation. *Journal of Neuroinflammation*, 17 (1): 48. doi:10.1186/s12974-020-1727-6.

Kvietys, P.R. (2010) "Regulation of Vascular Tone and Oxygenation." In *The Gastrointestinal Circulation*. Morgan & Claypool Life Sciences. https://www.ncbi.nlm.nih.gov/books/NBK53101/.

Lager, D.J., Callaghan, E.J., Worth, S.F., et al. (1995) Cellular localization of thrombomodulin in human epithelium and squamous malignancies. *The American Journal of Pathology*, 146 (4): 933–943.

Langenkamp, E., Zhang, L., Lugano, R., et al. (2015) Elevated expression of the C-Type Lectin CD93 in the glioblastoma vasculature regulates cytoskeletal rearrangements that enhance vessel function and reduce host survival. *Cancer Research*, 75 (21): 4504–4516. doi:10.1158/0008-5472.CAN-14-3636.

Lazovic, B., Nguyen, H.-T., Ansarizadeh, M., et al. (2024) Human iPSC and CRISPR targeted gene knock-in strategy for studying the somatic TIE2L914F mutation in endothelial cells. *Angiogenesis*. doi:10.1007/s10456-024-09925-9.

Lee, S., Rho, S.-S., Park, H., et al. (2017) Carbohydrate-binding protein CLEC14A regulates VEGFR-2– and VEGFR-3–dependent signals during angiogenesis and lymphangiogenesis. *The Journal of Clinical Investigation*, 127 (2): 457–471. doi:10.1172/JCI85145.

Liang, J., Wang, S., Zhang, G., et al. (2021) A New Antitumor Direction: Tumor-Specific Endothelial Cells. *Frontiers in Oncology*, 11: 756334. doi:10.3389/fonc.2021.756334.

Lin, H., Cheng, J., Mu, W., et al. (2021) Advances in Universal CAR-T Cell Therapy. *Frontiers in Immunology*, 12: 744823. doi:10.3389/fimmu.2021.744823.

Liu, A.C. and Gotlieb, A.I. (n.d.) *Molecular Basis of Cardiovascular Disease*.

Liu, Y.-L. and Guo, Z.-Y. (2022) "The NanoBiT-Based Homogenous Ligand–Receptor Binding Assay." In Kim, S.-B. (ed.) *Bioluminescence: Methods and Protocols, Volume 2*. New York, NY: Springer US. pp. 139–153. doi:10.1007/978-1-0716-2473-9_10.

Lodhia, P. (n.d.) *Puja-Production and Characterisation of Anti-CLEC14A Antibody Drug Conjugates*., p. 172.

Lorenzon, E., Colladel, R., Andreuzzi, E., et al. (2012) MULTIMERIN2 impairs tumor angiogenesis and growth by interfering with VEGF-A/VEGFR2 pathway. *Oncogene*, 31 (26): 3136–3147. doi:10.1038/onc.2011.487.

Lugano, R., Vemuri, K., Yu, D., et al. (2018) CD93 promotes β_1 integrin activation and fibronectin fibrillogenesis during tumor angiogenesis. *The Journal of Clinical Investigation*, 128 (8): 3280–3297. doi:10.1172/JCI97459.

Ma, Y., Lee, C.-J. and Park, J.-S. (2020) Strategies for Optimizing the Production of Proteins and Peptides with Multiple Disulfide Bonds. *Antibiotics*, 9 (9): 541. doi:10.3390/antibiotics9090541.

Malhotra, A. (2009) Tagging for protein expression. *Methods in Enzymology*, 463: 239–258. doi:10.1016/S0076-6879(09)63016-0.

Martin, T.A., Ye, L., Sanders, A.J., et al. (2013) "Cancer Invasion and Metastasis: Molecular and Cellular Perspective." In *Madame Curie Bioscience Database [Internet]*. Landes Bioscience. https://www.ncbi.nlm.nih.gov/books/NBK164700/.

Mascarenhas, N.M. and Gosavi, S. (2017) Understanding protein domain-swapping using structure-based models of protein folding. *Progress in Biophysics and Molecular Biology*, 128: 113–120. doi:10.1016/j.pbiomolbio.2016.09.013.

Mazzone, A., Dibenedetto, D., Nico, G., et al. (2011) "Automatic identification of alpha-helices in X-ray diffraction patterns." In *2011 IEEE International Symposium on Medical Measurements and Applications*. May 2011. pp. 437–441. doi:10.1109/MeMeA.2011.5966729.

Mongiat, M., Mungiguerra, G., Bot, S., et al. (2000) Self-assembly and supramolecular organization of EMILIN. *The Journal of Biological Chemistry*, 275 (33): 25471–25480. doi:10.1074/jbc.M001426200.

Mura, M., Swain, R.K., Zhuang, X., et al. (2012) Identification and angiogenic role of the novel tumor endothelial marker CLEC14A. *Oncogene*, 31 (3): 293–305. doi:10.1038/onc.2011.233.

Norman, D.G., Barlow, P.N., Baron, M., et al. (1991) Three-dimensional structure of a complement control protein module in solution. *Journal of Molecular Biology*, 219 (4): 717–725. doi:10.1016/0022-2836(91)90666-T.

Noy, P.J., Lodhia, P., Khan, K., et al. (2015) Blocking CLEC14A-MMRN2 binding inhibits sprouting angiogenesis and tumour growth. *Oncogene*, 34 (47): 5821–5831. doi:10.1038/onc.2015.34.

Noy, P.J., Swain, R.K., Khan, K., et al. (2016) Sprouting angiogenesis is regulated by shedding of the C-type lectin family 14, member A (CLEC14A) ectodomain, catalyzed by rhomboid-like 2 protein (RHBDL2). *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 30 (6): 2311–2323. doi:10.1096/fj.201500122R.

Olczyk, P., Mencner, Ł. and Komosinska-Vassev, K. (2015) Diverse Roles of Heparan Sulfate and Heparin in Wound Repair. *BioMed Research International*, 2015: 549417. doi:10.1155/2015/549417.

Overton, T.W. (2014) Recombinant protein production in bacterial hosts. *Drug Discovery Today*, 19 (5): 590–601. doi:10.1016/j.drudis.2013.11.008.
Panés, J. and Granger, D.N. (1998) Leukocyte-endothelial cell interactions: molecular mechanisms and implications in gastrointestinal disease. *Gastroenterology*, 114 (5): 1066–1090. doi:10.1016/s0016-5085(98)70328-2.

Pellicani, R., Poletto, E., Andreuzzi, E., et al. (2020) Multimerin-2 maintains vascular stability and permeability. *Matrix Biology*, 87: 11–25. doi:10.1016/j.matbio.2019.08.002.

Rajendran, P., Rengarajan, T., Thangavel, J., et al. (2013) The Vascular Endothelium and Human Diseases. *International Journal of Biological Sciences*, 9 (10): 1057–1069. doi:10.7150/ijbs.7502.

Rho, S.-S., Choi, H.-J., Min, J.-K., et al. (2011) Clec14a is specifically expressed in endothelial cells and mediates cell to cell adhesion. *Biochemical and Biophysical Research Communications*, 404 (1): 103–108. doi:10.1016/j.bbrc.2010.11.075.

Robinson, J., Whitworth, K., Jinks, E., et al. (2020) An evaluation of the tumour endothelial marker CLEC14A as a therapeutic target in solid tumours. *The Journal of Pathology: Clinical Research*, 6 (4): 308–319. doi:10.1002/cjp2.176.

Sandoval, D.R., Gomez Toledo, A., Painter, C.D., et al. (2020) Proteomics-based screening of the endothelial heparan sulfate interactome reveals that C-type lectin 14a (CLEC14A) is a heparin-binding protein. *The Journal of Biological Chemistry*, 295 (9): 2804–2821. doi:10.1074/jbc.RA119.011639.

Sanz-Moncasi, M.P., Garin-Chesa, P., Stockert, E., et al. (1994) Identification of a high molecular weight endothelial cell surface glycoprotein, endoGlyx-1, in normal and tumor blood vessels. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, 71 (3): 366–373.

Shield-Artin, K.L., Bailey, M.J., Oliva, K., et al. (2012) Identification of ovarian cancer-associated proteins in symptomatic women: A novel method for semi-quantitative plasma proteomics. *Proteomics. Clinical Applications*, 6 (3–4): 170–181. doi:10.1002/prca.201100008.

Shriver, Z., Capila, I., Venkataraman, G., et al. (2012) Heparin and Heparan Sulfate: Analyzing Structure and Microheterogeneity. *Handbook of experimental pharmacology*, (207): 159–176. doi:10.1007/978-3-642-23056-1_8.

Slater, A., Di, Y., Clark, J.C., et al. (2021) Structural characterization of a novel GPVI-nanobody complex reveals a biologically active domain-swapped GPVI dimer. *Blood*, 137 (24): 3443–3453. doi:10.1182/blood.2020009440.

Soltermann, A., Ossola, R., Kilgus-Hawelski, S., et al. (2008) N-glycoprotein profiling of lung adenocarcinoma pleural effusions by shotgun proteomics. *Cancer*, 114 (2): 124–133. doi:10.1002/cncr.23349.

Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., et al. (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *The EMBO journal*, 32 (10): 1478–1488. doi:10.1038/emboj.2013.79.

Strowitzki, M.J., Cummins, E.P. and Taylor, C.T. (2019) Protein Hydroxylation by Hypoxia-Inducible Factor (HIF) Hydroxylases: Unique or Ubiquitous? *Cells*, 8 (5): 384. doi:10.3390/cells8050384.

Suri, C., Jones, P.F., Patan, S., et al. (1996) Requisite Role of Angiopoietin-1, a Ligand for the TIE2 Receptor, during Embryonic Angiogenesis. *Cell*, 87 (7): 1171–1180. doi:10.1016/S0092-8674(00)81813-9.

Tammela, T., Zarkada, G., Nurmi, H., et al. (2011) VEGFR-3 controls tip to stalk conversion at vessel fusion sites by reinforcing Notch signalling. *Nature Cell Biology*, 13 (10): 1202–1213. doi:10.1038/ncb2331.

Teicher, B.A. (2019) CD248: A therapeutic target in cancer and fibrotic diseases. *Oncotarget*, 10 (9): 993–1009. doi:10.18632/oncotarget.26590.

Tosi, G.M., Neri, G., Barbera, S., et al. (2020) The Binding of CD93 to Multimerin-2 Promotes Choroidal Neovascularization. *Investigative Ophthalmology & Visual Science*, 61 (8): 30. doi:10.1167/iovs.61.8.30.

Tossetta, G., Piani, F., Borghi, C., et al. (2023) Role of CD93 in Health and Disease. *Cells*, 12 (13): 1778. doi:10.3390/cells12131778.

Trinh, R., Gurbaxani, B., Morrison, S.L., et al. (2004) Optimization of codon pair use within the (GGGGS)3 linker sequence results in enhanced protein expression. *Molecular Immunology*, 40 (10): 717–722. doi:10.1016/j.molimm.2003.08.006.

Udan, R.S., Culver, J.C. and Dickinson, M.E. (2013) Understanding vascular development. *Wiley interdisciplinary reviews. Developmental biology*, 2 (3): 327–346. doi:10.1002/wdev.91.

Unzai, S. (2017) Analytical ultracentrifugation in structural biology. *Biophysical Reviews*, 10 (2): 229–233. doi:10.1007/s12551-017-0340-0.

Veldhuizen, E.J.A., van Eijk, M. and Haagsman, H.P. (2011) The carbohydrate recognition domain of collectins. *The FEBS Journal*, 278 (20): 3930–3941. doi:10.1111/j.1742-4658.2011.08206.x.

Vella, G., Guelfi, S. and Bergers, G. (2021) High Endothelial Venules: A Vascular Perspective on Tertiary Lymphoid Structures in Cancer. *Frontiers in Immunology*, 12. doi:10.3389/fimmu.2021.736670.

Vora, D., Dandekar, A.A. and Banga, A.K. (2022) "Therapeutic Peptide Delivery: Fundamentals, Formulations, and Recent Advances." In Jois, S.D. (ed.) *Peptide Therapeutics: Fundamentals of Design, Development, and Delivery*. Cham: Springer International Publishing. pp. 183–201. doi:10.1007/978-3-031-04544-8_5.

Wijelath, E., Namekata, M., Murray, J., et al. (2010) Multiple Mechanisms for Exogenous Heparin Modulation of Vascular Endothelial Growth Factor Activity. *Journal of cellular biochemistry*, 111 (2): 461–468. doi:10.1002/jcb.22727.

Willett, C.G., Boucher, Y., di Tomaso, E., et al. (2004) Direct evidence that the VEGF-specific antibody bevacizumab has antivascular effects in human rectal cancer. *Nature Medicine*, 10 (2): 145–147. doi:10.1038/nm988.

Wragg, J.W., Finnity, J.P., Anderson, J.A., et al. (2016) MCAM and LAMA4 are highly enriched in tumor blood vessels of renal cell carcinoma and predict patient outcome. *Cancer research*, 76 (8): 2314–2326. doi:10.1158/0008-5472.CAN-15-1364.

Xu, D. and Esko, J.D. (2014) Demystifying Heparan Sulfate–Protein Interactions. *Annual review of biochemistry*, 83: 129–157. doi:10.1146/annurev-biochem-060713-035314.

Yan, L., Li, X. and Yuan, Y. (2022) CLEC14A was up-regulated in hepatocellular carcinoma and may function as a potential diagnostic biomarker. *Clinics*, 77. doi:10.1016/j.clinsp.2022.100029.

Yang, T., Xiao, H., Liu, X., et al. (2021) Vascular Normalization: A New Window Opened for Cancer Therapies. *Frontiers in Oncology*, 11: 719836. doi:10.3389/fonc.2021.719836.

Zelensky, A.N. and Gready, J.E. (2005) The C-type lectin-like domain superfamily. *The FEBS Journal*, 272 (24): 6179–6217. doi:10.1111/j.1742-4658.2005.05031.x.

Zhao, W., McCallum, S.A., Xiao, Z., et al. (2012) Binding affinities of vascular endothelial growth factor (VEGF) for heparin-derived oligosaccharides. *Bioscience reports*, 32 (1): 71–81. doi:10.1042/BSR20110077.

Zhuang, X., Maione, F., Robinson, J., et al. (2020) CAR T cells targeting tumor endothelial marker CLEC14A inhibit tumor growth. *JCI Insight*, 5 (19): e138808. doi:10.1172/jci.insight.138808.