CHAPTER 5

CRISPR ACTIVATION SCREENING OF ADHESION GPCRS IDENTIFIES KNOWN AND NOVEL INTERACTIONS

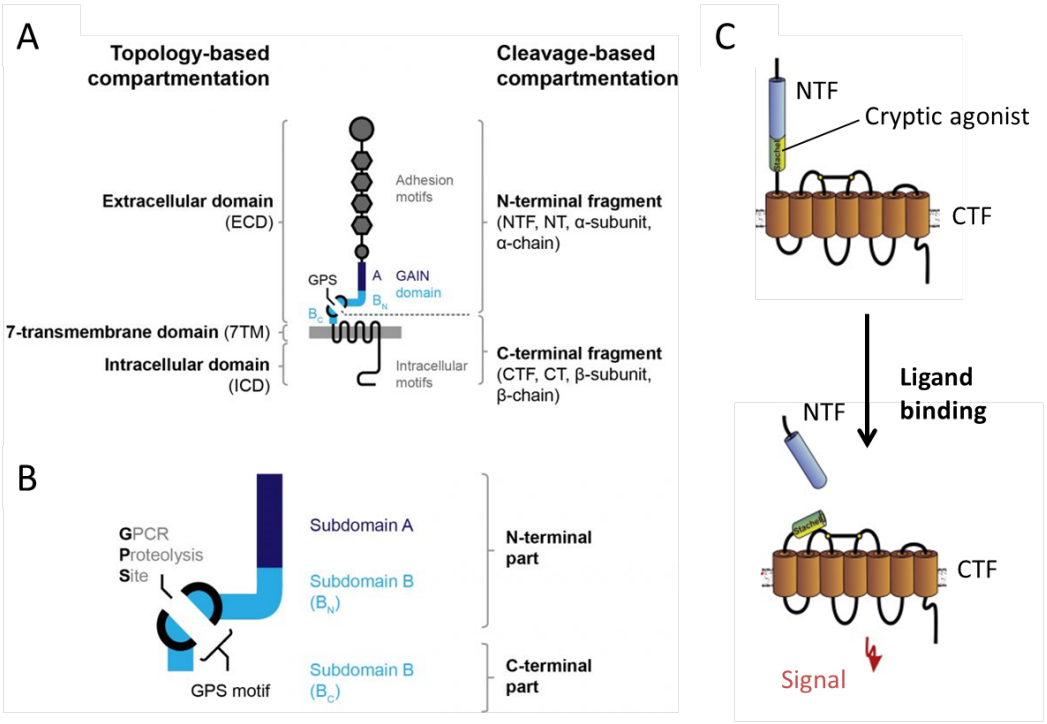
5.1 Introduction

The G-protein coupled receptor (GPCR) superfamily is one of the largest groups of transmembrane proteins, and is extensively targeted by therapeutic drugs. An estimated 34% of FDA-approved drugs target just 108 GPCRs (Hauser et al., 2018). GPCRs are characterised by seven transmembrane domains and cytosolic association with heterotrimeric G-proteins, which participate in a wide variety of downstream signalling cascades. However, many GPCRs do not have known endogenous ligands and as a result, there is great interest in understanding the biology of such GPCRs and identifying their ligands, a process called 'GPCR deorphanisation'. Adhesion GPCRs, a subfamily characterised by large N-terminal domains containing multiple adhesion-related motifs, contain the highest number of 'orphan' receptors which have no known endogenous ligand. Members of the adhesion GPCR family play important roles in immune regulation, central nervous system development, and angiogenesis (Bjarnadóttir et al., 2004). Knowledge of endogenous ligands would therefore shed some light on the molecular mechanisms of these receptors.

As previously mentioned, the large extracellular N-terminal regions of adhesion contain multiple protein domains involved in cell adhesion, as well as a conserved GPCR Proteolysis Site (GPS) located in the ectodomain almost adjacent to the first transmembrane domain (Figure 5.1A). During translation, the GPS is cleaved, forming a C-terminal fragment containing the seven transmembrane domains and an N-terminal fragment consisting most of the ectodomain (Figure 5.1B). Both fragments remain non-covalently attached during trafficking to the plasma membrane, and the prevailing theory of adhesion GPCR activation is that ligand binding to the N-terminal ectodomain causes a conformational change or complete dissociation of the N-terminal fragment revealing a cryptic tethered peptide agonist (Figure 5.1C) (Araç et al., 2012; Stoveken et al., 2015). As such, adhesion GPCRs are thought to bind cell surface or extracellular matrix (ECM) proteins to cause N-terminal fragment dissociation. This is supported by the fact that known binding partners include GPI-anchored, transmembrane, and fibrous proteins like collagen or laminin. The propensity for binding cell surface receptors along with large N-terminal ectodomains that can be produced in soluble recombinant form makes this family of proteins ideal for CRISPRa interaction screening.

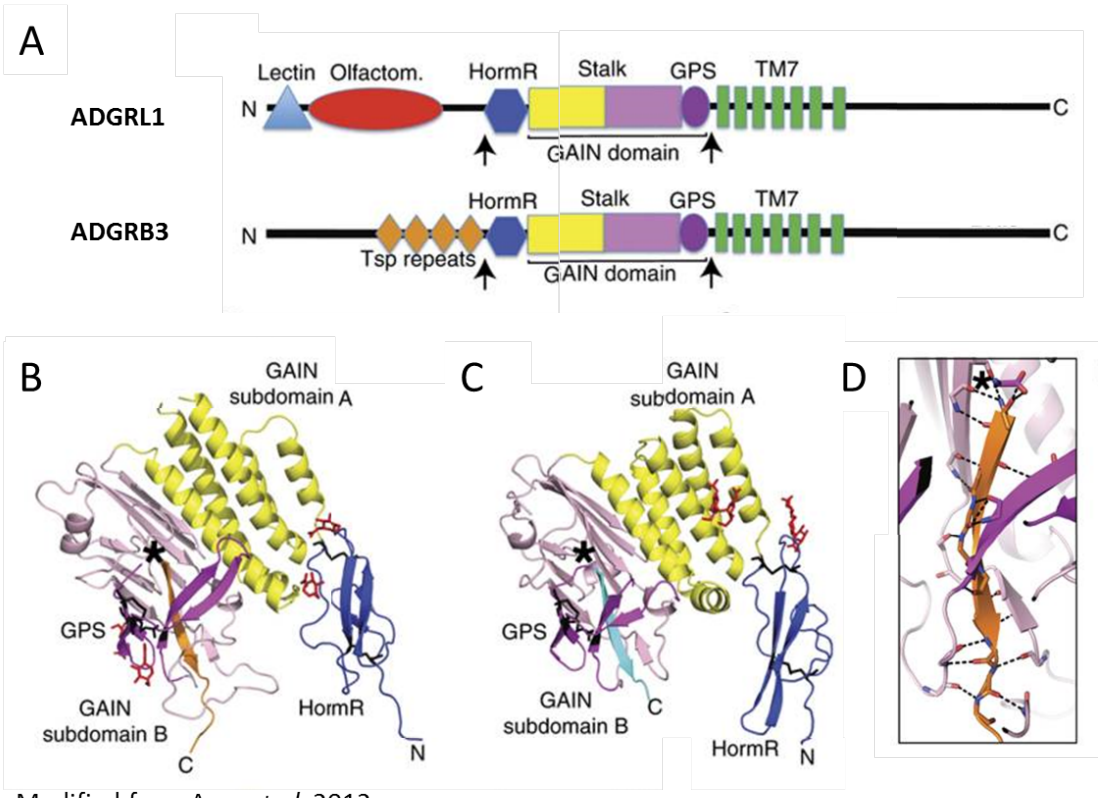
All 33 members of the adhesion GPCR family in humans harbour a conserved GPCR autoproteolysis-inducing (GAIN) domain which includes the GPS motif and is minimally required for proteolysis. The GPS motif consists of a conserved histidine, leucine and threonine or serine, with proteolysis occurring between the leucine and threonine/serine residues (H↓T/S). Structural studies indicate that the C-terminal region of the GAIN domain consists of a twisted β-sandwich including 13 β-strands and two small α-helices (Figure 5.2A-C) (Araç et al., 2012). The GPS is located between the last two β-strands and cleavage results in the separation of the last β-strand, which is then kept in place by numerous hydrophobic interactions with the surrounding β-strands (Figure 5.2D). Mutagenesis of the GPS motif indicated that autoproteolysis is not required for surface transport, although certain mutations resulted in cytosolic retention of the receptor, possibly due to the steric changes affecting the structure of the GPS/GAIN domain (Araç et al., 2012).

In this chapter, I applied the CRISPRa screening approach to identify cell surface binding partners of adhesion GPCRs. To do so, I expressed the entire ectodomain of 13 adhesion GPCRs as soluble recombinant proteins by mutating the GPS site to prevent proteolysis. Screening identified several previously described interactions, as well as novel associations between brain angiogenesis inhibitor 1 (ADGRB1) and members of the myelin-associated Nogo receptor family. Using cDNA overexpression and Avidity-based extracellular interaction screening (AVEXIS), I confirmed that ADGRB1 binds all three



Modified from Langenhan *et al*, 2013 and Liebscher *et al*, 2014

Figure 5.1 Structure of a typical adhesion GPCR A) Adhesion GPCR structure can be compartmentalised with reference to topology or cleavage at the GPCR proteolytic site (GPS). All adhesion GPCRs consist of a tripartite structure consisting of an extracellular domain (ECD), a seven transmembrane domain (7TM), and an intracellular (ICD). Some adhesion GPCRs undergo autoproteolysis at the GPS to produce an N-terminal (NTF) and C-terminal fragment (CTF). B) The GAIN domain is a complex fold that mediates autoproteolysis and subsequent attachment of cleaved NTF and CTFs. It is divided into two subdomains, A and B. Subdomain B contains and is cleaved at a conserved sequence of residues (HL↓T/S) located within the GPS motif. C) Ligand binding to the NTF is thought to induce intracellular signalling by causing structural changes or complete dissociation of the NTF to reveal a cryptic tethered agonist which then binds to and activates the receptor.



Modified from Araç *et al*, 2012

Figure 5.2 The GPS is not an autonomously folded domain but is part of a larger domain. A) Diagram of ADGRL1 and ADGRB3 showing the domains suggested by the SMART protein domain prediction server. The GPS is defined as a separate domain in the Pfam database (dark purple). B) Structures of the GPCR autoproteolysis-inducing (GAIN) domain of ADGRL1 and C) ADGRB3 by Araç et al. (2012) show that the GPS motif is part of a more complex fold comprising 13 β sheets and 2 α helices. D) After cleavage, NTF and CTF remain attached by numerous hydrogen bonds shown between the cleaved β-strand (orange) and the surrounding β-strands (purple) in ADGRL1. The cleavage site is indicated with a black star in B), C) and D)

members of the Nogo receptor family (RTN4R, RTN4RL1, RTN4RL2) and showed that the first three thrombospondin repeats on ADGRB1 are sufficient for binding.

5.2 Results

5.2.1 A T/S→G mutation at the GPS site enables high level of recombinant ectodomain production

To identify binding partners using CRISPRa screening, I sought to produce the ectodomains of members of the adhesion GPCR family in soluble, recombinant form. Using constructs encoding recombinant ADGRL4 and ADGRG1 that were already available in the lab, I transfected HEK293-6E cells and harvested culture supernatants six days post transfection. Western blotting of culture supernatant showed faint bands that did not correspond to the expected molecular weights, indicating a lack of proper protein expression (Figure 5.3A). These constructs contained the ectodomains of both receptors truncated at the GPS cleavage motif (Figure 5.3B). Given that the N- and C-terminal fragments of the receptors remain associated by numerous hydrophobic bonds (Figure 5.2D), I hypothesised that the GAIN domain might be unable to undergo proper folding without the last β-strand, which was not included in these constructs. Thus, I designed new constructs that consisted of the full length ectodomains of ADGRL4 and ADGRG1 up to the start of the first transmembrane domain, and included a T→G mutation at the GPS motif which had been previously shown to abolish proteolysis but not trafficking to the cell surface (Araç et al., 2012). Expression of the new constructs yielded a much higher level of recombinant protein that was expressed at the expected molecular weights (Figure 5.3C). Hence, I applied this strategy to design recombinant ectodomain constructs for expressing other members of the adhesion GPCR family.

After eliminating adhesion GPCRs with ectodomains exceeding the maximum length for gene synthesis, I designed a total of 21 constructs, 18 of which were successfully synthesised and 13 produced sufficient protein for CRISPRa screening (Figure 5.3D). To determine whether constructs harbouring the T/S→G mutation retained the ability to bind to their endogenous ligands, I performed plate-based interaction assays with recombinant CD55 and ADGRE5, and showed that recombinant ADGRE5 was able to bind to CD55 (Figure 5.3E).

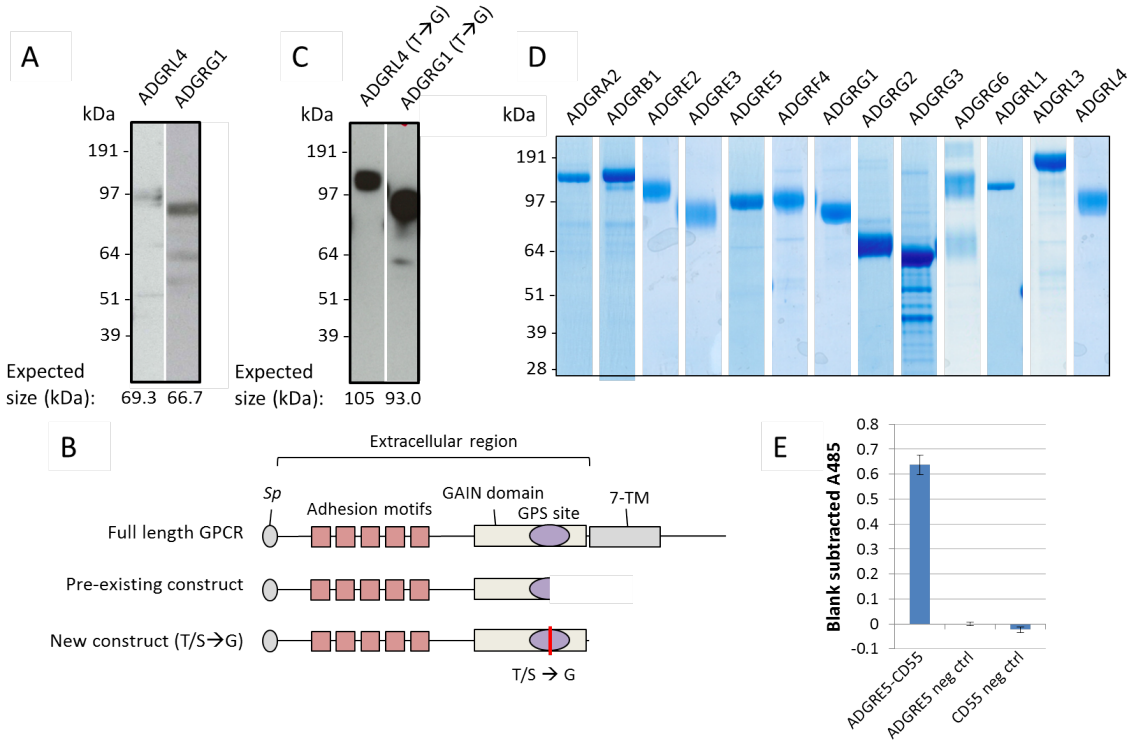


Figure 5.3 T/S→G mutation at the GPS site enables the production of soluble recombinant adhesion GPCR ectodomains for CRISPRa screening. A) Constructs expressing truncated ectodomains of ADGRL4 and ADGRG1 do not produce biotinylated proteins at the expected sizes as observed by western blotting. 10 μL of culture supernatant was loaded in each well. Detection of biotinylated proteins was performed by incubating blot with streptavidin conjuated to HRP and visualised with chemiluminescent peroxidase substrate. B) Diagram of truncated ectodomain encoded by the original constructs and full length ectodomain encoded in the new constructs. Full length ectodomains are resistent to cleavage at the GPS by a T/S→G mutation (red line) adjacent to the cleavage site. C) Constructs expressing full length ectodomains produce higher levels of biotinylated recombinant protein at the expected sizes. The same amount of culture supernatant as in A) was loaded in all wells and masses listed include predicted glycosylation. D) 13 adhesion GPCRs ectodomains were produced as biotinylated, His-tagged recombinant proteins and purified using Ni^{2+} affinity beads. Purified protein corresponded to their expected sizes as determined by SDS-PAGE and Coomassie staining. E) Recombinant ADGRE5 interacts with its endogenous ligand CD55. Increased absorbance at 485 nm indicate retention of β-lactamase-tagged CD55 prey in wells coated with recombinant ADGRE5 bait. Negative controls were performed with an unrelated protein, rCd200, which did not interact with either recombinant ADGRE5 or CD55. Bars represent blank subtracted mean $\pm$ s.d.; n=3.

5.2.2 CRISPRa screening of adhesion GPCR ectodomains identifies known interactions

For CRISPRa screening, I prepared highly avid tetramers from biotinylated ectodomains of 13 adhesion GPCRs according to the strategy described in (Chapter 4.2.4). Fluorescently labelled tetramers were screened using the improved CRISPRa screening protocol at a 5% sort threshold, and any genes enriched with an FDR of < 0.1 were considered 'hits'. CRISPRa screening using ADGRL1 and ADGRL3 tetramers detected previously reported interactions with members of the FLRT and TENM families (Figure 5.4A). ADGRL1 and ADGRL3 belong to the Latrophilin subfamily and are neuronal receptors for α-latrotoxin, controlling neurotransmitter release and presynaptic calcium levels. In separate mass spectrometry based studies, ADGRLs have been found to interact with FLRTs (fibronectin leucine-rich transmembrane proteins) and TENMs (Teneurins) that are expressed on the neuronal cell surface and are implicated in controlling neurite outgrowth and patterning (O'Sullivan et al., 2012; Silva et al., 2011). Importantly, the detection of the ADGRL1-TENM3/4 interactions would likely not be possible using plate-based assays as TENMs are very large, Type II transmembrane proteins which might be difficult to produce in soluble recombinant form.

Additionally, a screen using ADGRA2 tetramers detected an enrichment of guides targeting syndecans (SDC1 and SDC2), a major family of heparan sulfate proteoglycans (Figure 5.4B). ADGRA2 is known to bind sulfated glycosaminoglycans (GAG) such as heparan and chondroitin sulfates (Vallon and Essler, 2006), and syndecans are composed mainly of such GAGs attached to a core protein. To determine if binding is dependent purely on sulfated glycosaminoglycans, I performed cell binding assays with a cell line lacking SLC35B2, a transporter of 3'-phosphoadenosine-5'-phosphosulfate (PAPS), which acts as a sulfate donor during GAG sulfation. SLC35B2 knockout lines retain expression of syndecans but are unable to produce sulfated GAGs like heparan sulfate. A complete loss of binding of ADGRA2 tetramers was observed with SLC35B2 knockout cells (Figure 5.4C), suggesting that GAGs form the major determinant of ADGRA2 binding rather than binding being specific to syndecans. Taken together, the detection of known interactions provide further validation of the CRISPRa screening approach for the identification of extracellular receptor-ligand pairs.

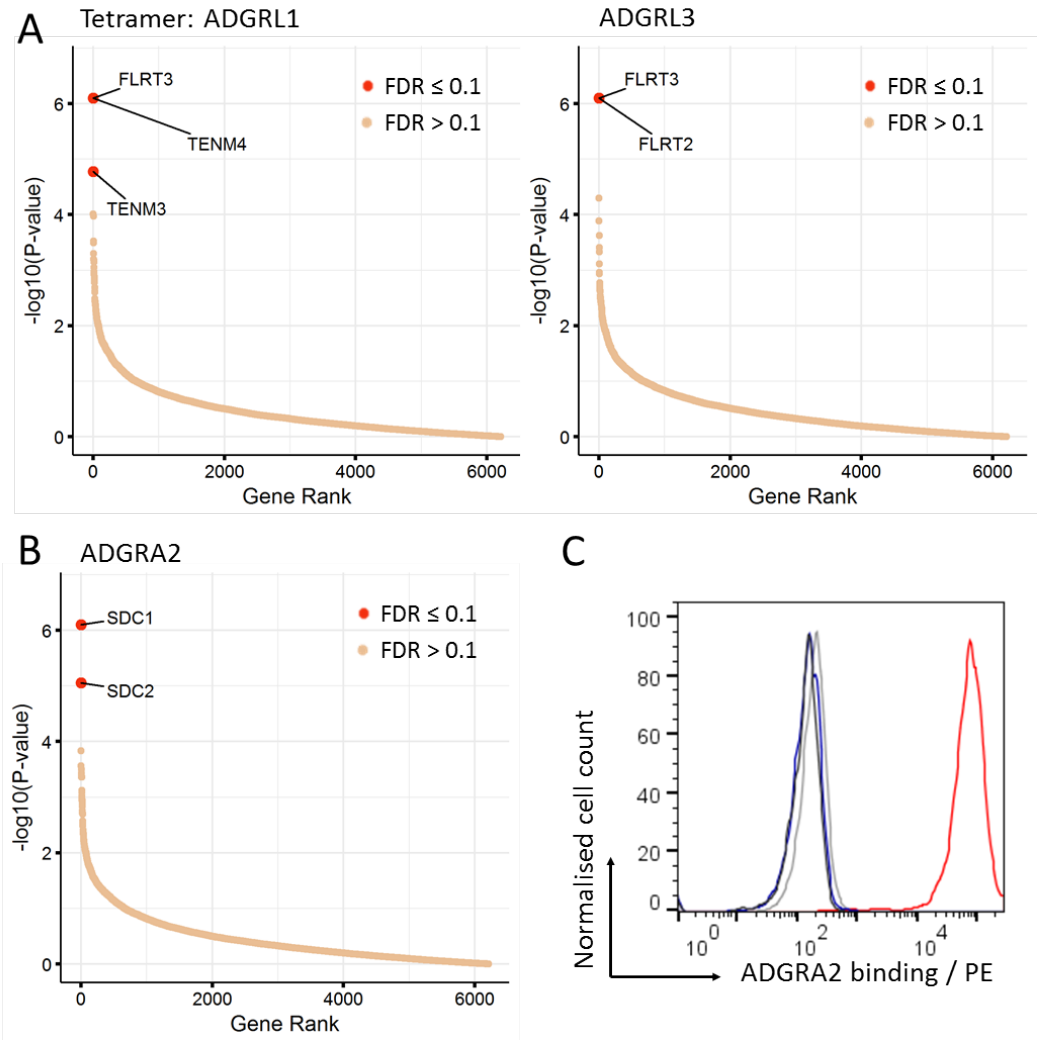


Figure 5.4 CRISPRa screening identifies known interactions of ADGRL1 and ADGRL3, as well as glycosaminoglycan (GAG)-binding properties of ADGRA2. A) Transformed gene enrichment P-values are plotted against a rank-ordered gene list for CRISPRa enrichment screens with cells selected using recombinant tetramers for ADGRL1 (left), ADGRL3 (right) and B) ADGRA2. An FDR cut-off of 0.1 was used to determine which genes were considered significantly enriched (red dots). C) Cell surface binding assays with SLC35B2 knockout (KO) HEK293 cells suggest that ADGRA2 binding is GAG-dependent. Fluorescently labelled ADGRA2 tetramers bound to wildtype HEK293 (red trace) but not SLC35B2 KO cells (blue trace). Unstained wildtype HEK293 cells (black trace) or cells incubated with streptavidin-PE (grey trace) were used as negative controls. A representative of three independent experiments is shown.

5.2.3 CRISPRa screening with ADGRB1 uncovers novel interactions with Nogo receptors

In addition to known interactions, CRISPRa screening identified previously unreported binding partners of ADGRB1. Guides targeting RTN4RL1 and RTN4RL2 were significantly enriched in a screen performed with ADGRB1 tetramers (Figure 5.5A). ADGRB1 (Brain Angiogenesis Inhibitor 1) is a phosphatidylserine receptor on professional phagocytes (Park et al., 2007), and is enriched in the postsynaptic density in neurons where it regulates excitatory synapse formation in hippocampal and cortical cultures (Duman et al., 2013) but has no documented ligands in the nervous system. On the other hand, RTN4RL1 and RTN4RL2 belong to the Nogo receptor family which are GPI-linked membrane proteins and are known to be involved in regulating axon growth and synapse formation, most notably through interactions between Nogo receptor 1 (RTN4R) and the myelin-associated inhibitor, Nogo-66 (Liu et al., 2002).

Given that guides targeting RTN4RL1 and RTN4RL2 were highly enriched in the sorted population, along with the fact that all three proteins were expressed in the brain, I decided to validate this interaction using cDNA overexpression in HEK293 cells. Transfection of HEK293 cells with constructs expressing the full-length cDNAs of RTN4RL1 and RTN4RL2 conferred an increased affinity for ADGRB1 tetramer binding relative to an untransfected control (Figure 5.5B). Surprisingly, overexpression of RTN4R also resulted in increased binding, even though guides targeting RTN4R were not significantly enriched in the initial screen (Figure 5.5C). The gain of binding phenotype was not due to recombinant protein tags as a control protein bearing the same tags did not show any increase in binding to transfected cells. To determine if the increase in binding was due to an indirect effect of upregulating extracellular phosphatidylserine, I performed Annexin V staining and did not see a significant increase in levels of extracellular phosphatidylserine in cells overexpressing Nogo receptors relative to an untransfected control (Figure 5.5D). These data suggest that ADGRB1 binds to all three members of the Nogo receptor family.

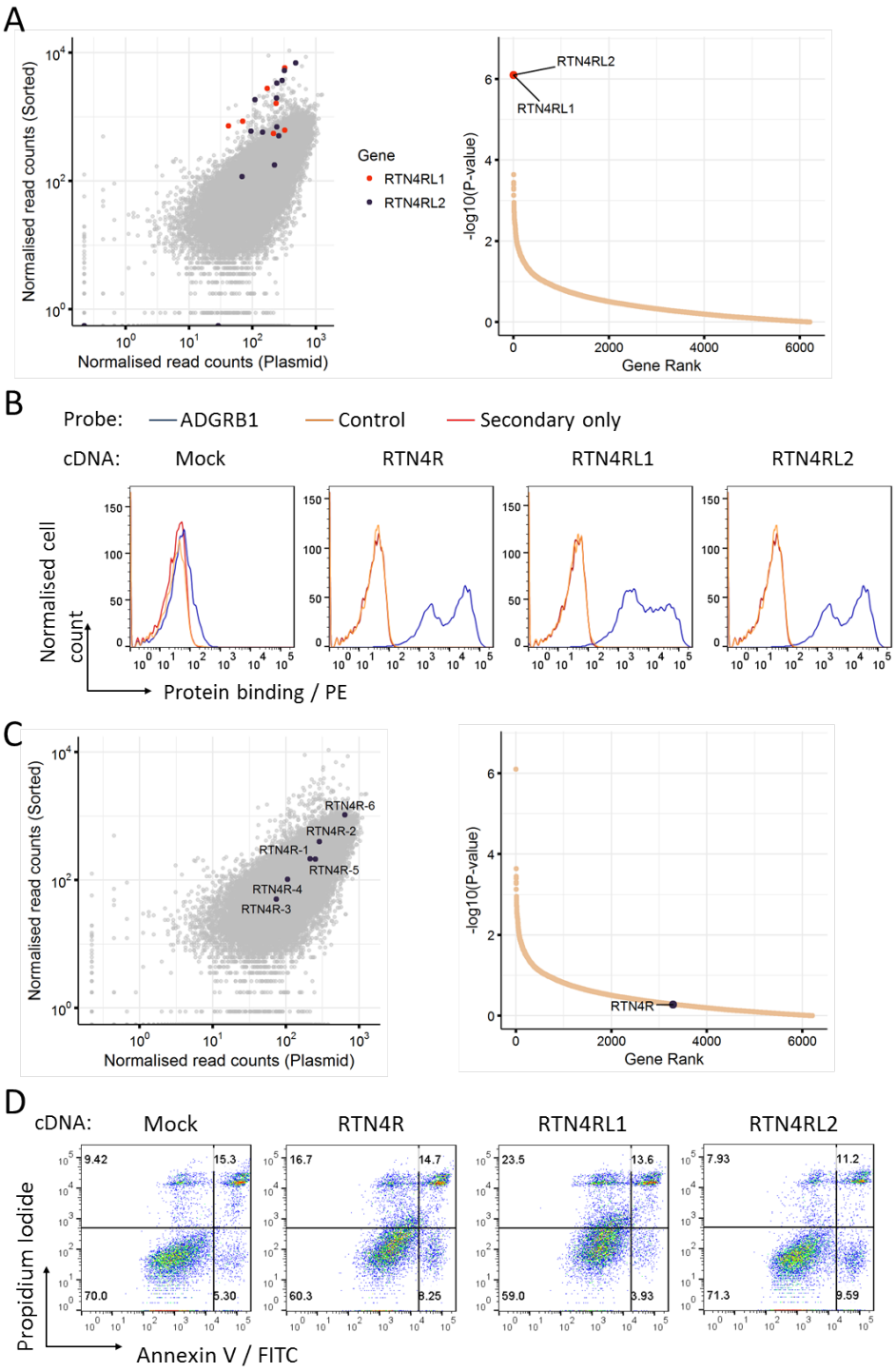

A
Normalised read counts (Sorted)
Normalised read counts (Plasmid)
Gene
RTN4RL1
RTN4RL2
-log10(P-value)
Gene Rank
B
Probe:
ADGRB1
Control
Secondary only
cDNA:
Mock
RTN4R
RTN4RL1
RTN4RL2
Normalised cell count
Protein binding / PE
C
RTN4R-6
RTN4R-2
RTN4R-1
RTN4R-5
RTN4R-4
RTN4R-3
D
Propidium Iodide
Annexin V / FITC
9.42
15.3
70.0
5.30
16.7
14.7
60.3
8.25
23.5
13.6
59.0
3.93
7.93
11.2
71.3
9.59

Figure 5.5 CRISPRa screening identifies novel interactions between ADGRB1 and Nogo receptors. A) Guides targeting RTN4RL1 and 2 are enriched in a population of cells sorted for gain-of-function binding to ADGRB1 tetramers. A plot of normalised gRNA read counts in the sorted population against that of the plasmid library show increased abundance of RTN4RL1 and 2-targeting guides (left). Transformed gene enrichment P-values plotted against a rank-ordered gene list for a screen performed with ADGRB1 tetramers show that RTN4RL1 and 2 are the only genes found to be significantly enriched under an FDR of 0.1 (right). B) ADGRB1 tetramers stained cells transfected with cDNAs encoding full-length RTN4R, RTN4RL1, RTN4RL2 (blue lines) but not mock-transfected cells compared to a control ADGRL1 tetramer (orange line), or streptavidin-PE alone (red line). A representative of four independent experiments is shown. C) RTN4R-targeting guides were not enriched in the CRISPRa screen using ADGRB1 tetramers. Normalised read counts of all 5 RTN4R-targeting guides were similar between the ADGRB1-sorted population and the plasmid gRNA library (left). Transformed gene enrichment P-values plotted against gene rank also show that RTN4R was not highly ranked in terms of enrichment. D) Transfection of cells with cDNAs encoding full-length RTN4R, RTN4RL1, RTN4RL2 did not cause an increase in the levels of cell surface phosphatidylserine, a known ligand of ADGRB1, as determined by Annexin V staining of cells in comparison to mock-transfected cells.

5.2.4 The first three thrombospondin repeats on ADGRB1 is sufficient for RTN4R binding

To determine direct binding between ADGRB1 and RTN4Rs, I performed plate-based interaction assays using recombinant ectodomains of all four receptors as previously described in Bushell et al. (2008). Briefly, the full length ectodomains of ADGRB1 and RTN4R, RTN4RL1 and RTN4RL2 were produced either as biotinylated monomers ('baits') or as pentameric, β-lactamase tagged proteins ('preys'). Biotinylated baits were captured on streptavidin coated plates and overlaid with pentameric preys before washing to remove unbound preys. A nitrocefin hydrolysis assay was then used to detect the presence of remaining preys. Colourimetric readouts indicated that ADGRB1 interacted with all three Nogo receptors (Figure 5.6A). This interaction was shown in both orientations, with ADGRB1 as either a bait or prey. RTN4RL2 could not be expressed in pentameric form, but when produced as a bait showed binding to ADGRB1. Importantly, none of the Nogo receptors interacted with ADGRB2, a closely related receptor to ADGRB1. Both ADGRB1 and ADGRB2 ectodomains contain several thrombospondin repeats (TSRs) as well as a hormone binding domain (HRM), but clearly exhibit different binding properties. This suggests that Nogo receptors interact specifically with ADGRB1.

To investigate the minimal requirements for ADGRB1-RTN4R binding, I produced truncated versions of the ADGRB1 ectodomain and tested their ability to interact with RTN4R. I expressed just the first three TSRs (TSR1-3), all five TSRs (TSR1-5) or just the HRM and GAIN domains (HRM+GAIN) of ADGRB1. Plate-based interaction assays using these constructs indicated that the first three TSRs were sufficient for binding of RTN4R (Figure 5.6B). Accordingly, the TSR1-5 fragment was also able to interact with RTN4R, but not the HRM+GAIN fragment. This indicates that the HRM and GAIN domains do not play a role in the interaction between ADGRB1 and RTN4Rs. In summary, ADGRB1 interacts directly and specifically with all three Nogo receptors through the first three TSRs on its ectodomain.

5.3 Discussion

In this chapter, I applied CRISPRa extracellular interaction screening to identify cell surface binding partners of adhesion GPCRs. To do so I designed expression constructs to produce full length ectodomains with a mutation at the GPS site to prevent cleavage. The

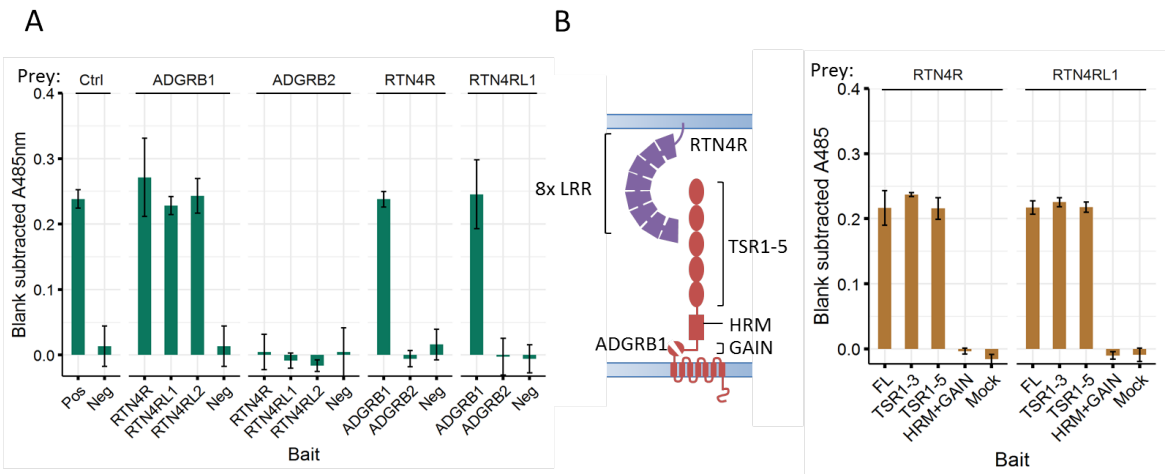


Figure 5.6 ADGRB1 specifically and directly interacts with Nogo receptors through the first three thrombospondin repeats (TSRs) in its ectodomain. A) The ectodomains of ADGRB1 and RTN4R family members directly interact. The extracellular regions of the named receptors were expressed as soluble biotinylated bait proteins, captured in individual wells of a streptavidin-coated plate and probed for interactions with pentameric β-lactamase-tagged prey proteins. Binding is quantified by absorbance at 485 nm of a hydrolysis product of the colourimetric β-lactamase substrate, nitrocefin. Bars represent blank-subtracted mean $\pm$ s.d; n=3. ADGRE5-CD55 interaction was used as a positive control; negative control bait was the CD55 ectodomain. B) The Nogo receptor binding interface on ADGRB1 is composed of the N-terminal three TSR domains. Schematic of the Nogo receptor family and ADGRB1 proteins showing their domain organization (left). Binding of RTN4R and RTN4RL1 preys to fragments of ADGRB1 encompassing the full-length ectodomain (FL), thrombospondin repeats 1-3 (TSR1-3), TSRs 1-5, or the hormone receptor motif and GAIN domain (HRM+GAIN) is shown (right). Bars represent blank subtracted mean $\pm$ s.d.; n=3.

success of this strategy over producing the N-terminal fragment of the ectodomain after proteolytic cleavage indicates that the presence of the extracellular portion of the C-terminal fragment is essential for proper folding of the receptor. Furthermore, the ability to produce fragments containing just the TSRs of ADGRB1 suggest that this dependence is specific to the GAIN domain, and that other adhesion motifs within the ectodomain are able to fold independently.

CRISPRa screening of adhesion GPCRs identified known interactions involving members of the Latrophilin subfamily and detected GAG-binding properties of ADGRA2. Interestingly, although CRISPRa screening identified TENM3 and TENM4 as binding partners for ADGRL1, these interactions were not observed in a cell binding assay using cDNA overexpression of members of the Teneurin family (Silva et al., 2011). This discrepancy could be due to the use of a specific isoform for cDNA overexpression which may be non-functional, whilst CRISPRa allows endogenous splicing decisions which might have resulted in the expression of an isoform capable of binding ADGRL1. This highlights an advantage of the CRISPRa platform over cDNA overexpression, particularly for receptors which are poorly annotated or have many isoforms. However, CRISPRa screening failed to identify several binding partners of both ADGRLs including FLRT1 and other members of the TENM family. Given that ADGRL tetramers were able to interact with their endogenous receptors, this is likely due to an inability of CRISPRa to upregulate FLRT1 and other TENMs. This could be due to inefficient guides, or targeting of alternative TSSs, and should be improved with better knowledge of guide design principles as well as TSS annotation.

For the remaining nine adhesion GPCRs, no hits were obtained at an FDR cut-off of 0.1. Reducing the threshold to FDR < 0.25 did not reveal any additional hits, suggesting that no genes were enriched in these screens. This could be due to a number of reasons, including the possibility that the tetramers for these adhesion GPCRs were unable to bind their endogenous partners, or that the adhesion GPCRs do not have cell surface ligands. Instead, some adhesion GPCRs may only bind extracellular matrix proteins. For instance, collagen type III is an activating ligand of ADGRG1 during cortical development and laminationLuo et al. (2011) whilst ADGRG6 interacts with laminin-211 and collage type IV to regulate Schwann cell development and peripheral nerve development respectively (Paavola et al., 2014; Petersen et al., 2015). This highlights a limitation of the CRISPRa screening platform, as this approach is unable to detect extracellular interactions with soluble secreted factors or extracellular matrix proteins.

CRISPRa screening of ADGRB1 tetramers detected novel interactions with members of the Nogo receptor family, RTN4RL1 and RTN4RL2. Although it did not appear as a hit in the initial screen, subsequent validation experiments showed that ADGRB1 was also able to bind to the third Nogo receptor, RTN4R. The expression of all four proteins are enriched in the brain, and all have documented functions in the regulation of neurite growth and synapse formation both *in vitro* and *in vivo* (Duman et al., 2013; Wills et al., 2012). Importantly, no neuronal ligands of ADGRB1 have been identified, and no common ligands have been identified which bind to all three RTN4Rs. The Nogo receptor family is known to function redundantly with regards to regulating neuronal growth *in vivo* (Wills et al., 2012), and therefore the discovery of common binding partners may provide an explanation for this functional redundancy.

In summary, I demonstrated the utility of CRISPRa screening by applying it to a family of GPCRs with few known ligands. CRISPRa screening of adhesion GPCRs identified previously reported interactions, which provides confidence that this screening strategy works, and also identified novel interactions between ADGRB1 and the Nogo receptor family. These interactions were validated using cDNA overexpression and plate-based interaction assays and were shown to be mediated by TSRs on ADGRB1.