CHAPTER 6

GENERAL DISCUSSION

6.1 Summary of results

Extracellular interactions form the basis of how cells sense and respond to their environment. As such, these interactions are involved in a variety of biological processes including development, immune regulation and pathogen invasion. Such interactions also make attractive drug targets as they are readily accessible to systemically delivered drugs, such as therapeutic monoclonal antibodies. However, the identification of key receptor-ligand pairs, especially low-affinity cell adhesion interactions, can be technically challenging to perform at scale. CRISPR activation (CRISPRa) provides an attractive potential strategy for genome-wide extracellular interaction screening as it enables in principle upregulation of virtually any cell surface receptor in the genome and investigation of extracellular interactions in the context of a cell membrane. CRISPRa screening also circumvents the need for the production of large recombinant protein or cDNA libraries which can be costly and resource-intensive.

In this thesis I have investigated some parameters affecting cell surface protein upregulation using CRISPRa, and established a workflow for extracellular interaction screening using a gRNA library targeting the promoter regions of genes encoding all putative membrane proteins in the human genome. Using both antibodies and endogenous ligands, I show that CRISPRa screening can detect interactors with high confidence, even those which

bind with low affinity. Finally, I applied the CRISPRa screening approach to members of the adhesion GPCR family and identified novel cell surface ligands for ADGRB1.

6.2 Evaluation of CRISPR activation screening as an approach for receptor identification

Currently available strategies for identifying extracellular interactions include affinity purification with mass spectrometry (AP-MS), plate-based interaction assays using soluble recombinant ectodomains, arrayed cDNA overexpression screening, and CRISPR/Cas9 knockout screening. Each approach has its advantages and limitations, but some key advantages and limitations of CRISPRa screening are discussed below.

6.2.1 CRISPR activation uses gRNA libraries that are cost-effective to produce and maintain

Improvements in oligonucleotide synthesis methods have greatly reduced the cost of producing complex pools of oligonucleotides with good accuracy. As such, gRNA libraries capable of targeting thousands of genes can be synthesised at a fraction of the cost of a comparably-sized library of full-length cDNA or recombinant protein expression constructs. Furthermore, a large plasmid or lentiviral library preparation can be used for numerous screens, reducing the need for maintenance of cDNA stocks or repeated protein production. In this study I designed and cloned a gRNA library targeting 6,213 genes encoding all putative membrane proteins. By contrast, the largest plate-based recombinant protein screen tested pairwise interactions of 250 proteins (Martin et al., 2010), whilst the largest available membrane protein cDNA library contains clones encoding 4,493 membrane proteins (Mullican et al., 2017), or an estimated 75% of the human surfaceome. One caveat with the membrane protein gRNA library is that I was unable to obtain an accurate estimate of the fraction of proteins that could be successfully upregulated. However, this is difficult to assess without having access to a large number of antibodies and cloning hundreds of individual gRNAs, or performing single-cell experiments which can be costly and technically challenging.

6.2.2 CRISPR activation allows genome-scale interaction screening

Aside from using cost-effective gRNA libraries to screen a large number of cell surface proteins at once, endogenous overexpression using CRISPRa also circumvents the restrictions of maximum insert lengths associated with conventional cloning into plasmid or virus-based expression vectors. By contrast, cDNA libraries tend to be biased towards smaller transcripts, with most plasmid-based expression vectors exhibiting reduced cloning efficiencies for inserts exceeding 7,000 - 8,000 bp. In this study, CRISPRa screening of ADGRL1 resulted in the enrichment of two members of the Teneurin family, TENM3 and TENM4, which are large proteins with a coding region of at least 8,097 and 7,209 bp respectively. This highlights the utility of endogenous overexpression for studying receptors with large domains. Furthermore, CRISPRa allows screening of multi-pass membrane proteins, which is difficult to achieve using recombinant protein approaches. This because non-contiguous ectodomains may not be able to fold independently of the transmembrane domains when produced recombinantly. In this study I demonstrated that CRISPRa can upregulate other multi-pass membrane receptors such as P2RX7 and ADGRE5. Thus, compared to currently available methods, CRISPRa is arguably the closest to achieving genome-scale extracellular interaction screening.

6.2.3 CRISPR activation screening is not restricted to receptors that are already expressed by a cell line

Furthermore, gain-of-function studies enable systematic testing of interactions without being restricted by endogenous expression in the screening cell line. In contrast, AP-MS and CRISPR knockout studies rely on pre-existing expression of a receptor candidate, possibly missing other interactors which are expressed in different tissues or under different contexts. In addition, CRISPR knockout screening might have problems identifying multiple, co-expressed receptors if knocking out one receptor on the cell does not reduce ligand binding (Sharma et al., 2018). On the other hand, CRISPRa can identify multiple interaction partners provided they can be expressed on the cell surface. For example, CRISPRa screening using EFNA1 identified EPHA2, EPHA4 and EPHA7.

6.2.4 CRISPRa screening is able to detect low affinity interactions

Importantly, I show that CRISPRa screening can identify endogenous interactions of medium to low affinity, particularly the CD55-ADGRE5 interaction. This is important as low-affinity interactions are often understudied, and can be difficult to detect with certain methods such as AP-MS, where stringent wash steps and the use of detergents for solubilisation of protein complexes can cause low-affinity binders to be lost (Wright, 2009). I was unable to detect the other low-affinity interaction tested (CTLA4-CD86) using CRISPRa screening, despite showing that the CTLA4 ectodomain probe was active and able to bind to a second receptor, CD80. However, the failure to detect the CTLA4-CD86 interaction may not be only due to low affinity and other explanations such as alternative TSSs are explored in Section 6.2.6. Ideally, more screens should be performed to gain a more accurate estimate of the sensitivity of this approach. However, as the cost of running each screen is not trivial, I limited the number of proteins tested to demonstrate that CRISPRa screening works.

6.2.5 CRISPRa screening is unable to detect certain types of extracellular interactions

One limitation of CRISPRa screening is that it cannot detect interactions between secreted proteins, or those requiring heteromeric receptors, unless the other subunits are already expressed in HEK293 cells. An example of the latter is integrins, which are formed from one α and one β subunit and exhibit specific binding patterns depending on the combination of subunits. CRISPRa screening using an antibody against integrin αvβ3 detected enrichment of the gene encoding the β-subunit, ITGB3, but not the α subunit, suggesting that the α-subunit is already expressed in HEK293 cells and able to form additional heterodimers with upregulated ITGB3. Additionally, interactions between soluble secreted factors or with extracellular matrix proteins tend to be of higher affinity than that between cell surface molecules and hence might be more amenable to mass spectrometry-based approaches.

6.2.6 Potential explanations for false negatives arising from CRISPRa screening and suggestions for improvement

One issue with CRISPRa screening is that it has a high false negative rate of an estimated 33.3%. This estimate is based on a relatively low number of interactions (nine). In addition, some known adhesion GPCR interactions like ADGRL1-TENM2 and ADGRL3-

FLRT1 were not identified with CRISPRa screening. Similarly, the ADGRB1-RTN4R interaction was not detected in the initial screen even though it was later shown using cDNA overexpression assays and AVEXIS. In the above examples of known false negatives, the selection probe was active as other cell surface binding partners were identified.

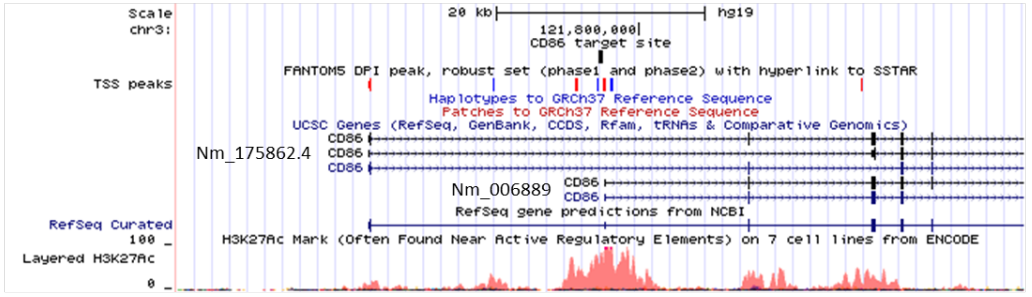
Potential reasons for the occurrence of false negatives including inactive guides, mis-annotation of transcription start sites (TSSs), targeting of alternative TSSs, or the lack of specific chaperones for surface transport and other post-transcriptional contextual effects. For instance, screening with CTLA4 ectodomains was able to identify a known binding partner CD80, but not a lower affinity interaction with CD86. Mapping of the CD86 gRNA targeting region alongside predicted gene models shows that gRNA in the library target a non-cannonical isoform of CD86 with a shorter signal peptide (Figure 6.1). This could result in a lack of expression of CD86 during pooled screening.

Whilst the inability to upregulate surface expression due to inactive guides or mis-targeting of TSSs is unique to CRISPRa, a lack of chaperones or other context-specific effects should be shared with cDNA-based overexpression approaches and might be remedied by using a different cell line for screening. Regarding gRNA design, there is ongoing research to study the parameters affecting guide effectiveness specifically for CRISPRa, as well as looking at more sophisticated ways to determine canonical TSSs. In fact, new genome-wide libraries have been published with improved selection algorithms that consider nucleosome positioning (Horlbeck et al., 2016), or with improved TSS predictions (Sanson et al., 2018). Although experimental comparisons are needed to determine whether more sophisticated design algorithms do indeed improve CRISPRa efficiency, future gRNA libraries might reduce the number of false negatives from CRISPRa screening that are due to the failure of inefficient or mistargeted gRNAs to upregulate receptors at the cell surface.

6.3 Implications of ADGRB1-RTN4R interactions and suggestions for further investigation

In this study, CRISPRa screening identified and validated a set of interactions between ADGRB1 and all three members of the Nogo receptor family (RTN4R, RTN4RL1, and RTN4RL2). All four proteins are implicated in synaptogenesis and neurite outgrowth, although no neuronal ligands for ADGRB1 have yet been identified. Nogo receptors on the other hand are known to regulate synaptogenesis through interactions between RTN4R

A



B

```
NM_006889       1   ------MGLSNILFVMAFLLSGAAPLKIQAYFNETADLPCQFANSQNQSLSELVVFWQDQ   54
NM_175862.4     1   MDPQCTMGLSNILFVMAFLLSGAAPLKIQAYFNETADLPCQFANSQNQSLSELVVFWQDQ   60
                    ************************************************************

NM_006889      55   ENLVLNEVYLGKEKFDSVHSKYMGRTSFDSDSWTLRLHNLQIKDKGLYQCIIHHKKPTGM  114
NM_175862.4    61   ENLVLNEVYLGKEKFDSVHSKYMGRTSFDSDSWTLRLHNLQIKDKGLYQCIIHHKKPTGM  120
                    ************************************************************

NM_006889     115   IRIHQMNSELSVLANFSQPEIVPISNITENVYINLTCSSIHGYPEPKKMSVLLRTKNSTI  174
NM_175862.4   121   IRIHQMNSELSVLANFSQPEIVPISNITENVYINLTCSSIHGYPEPKKMSVLLRTKNSTI  180
                    ************************************************************

NM_006889     175   EYDGIMQKSQDNVTELYDVSISLSVSFPDVTSNMTIFCILETDKTRLLSSPFSIELEDPQ  234
NM_175862.4   181   EYDGIMQKSQDNVTELYDVSISLSVSFPDVTSNMTIFCILETDKTRLLSSPFSIELEDPQ  240
                    ************************************************************

NM_006889     235   PPPDHIPWITAVLPTVIICVMVFCLILWKWKKKRPRNSYKCGTNTMEREESEQTKKREK  294
NM_175862.4   241   PPPDHIPWITAVLPTVIICVMVFCLILWKWKKKRPRNSYKCGTNTMEREESEQTKKREK  300
                    ************************************************************

NM_006889     295   IHIPERSDEAQRVFKSSKTSSCDKSDTCF                               323
NM_175862.4   301   IHIPERSDEAQRVFKSSKTSSCDKSDTCF                               329
                    *****************************
```

Figure 6.1 CRISPRa library guides target a non-cannonical isoform of CD86A) Guides for CD86 were designed to target a region upstream of the TSS of NM_006889 as denoted by a black rectangle (CD86 target site). Although this site is associated with a predicted CAGE-seq TSS peak (FANTOM5 DPI peak) as well as epigenetic marks commonly associated with promoter regions (H3K27ac), the longer isoform encoded by NM_175862.4 is annotated as the cannonical isoform. B) Amino acid alignment of coding sequences of transcripts NM_006889 and NM_175862.4 showing only a difference of 6 amino acids within the signal peptide.

and Nogo, an inhibitory molecule expressed on myelin (Fournier et al., 2001). Myelin-associated glycoprotein (MAG) is another inhibitory myelin-associated molecule that binds to RTN4R and RTN4RL2, but not RTN4RL1 (Robak et al., 2009). RTN4RL1 does not bind any members of the reticulon family, however all three receptors functionally compensate for each other *in vivo* (Wills et al., 2012). Thus, a common binding partner of all three RTN4Rs might help explain functional redundancy.

6.3.1 Knockout phenotypes complicate the assessment of ADGRB1-RTN4R function

The function of an interaction can be easily demonstrated when knocking out one binding partner results in a phenocopy of the other knockout (KO). However, knockout phenotypes of ADGRB1 and RTN4Rs do not exactly coincide. Whilst the knockout of either ADGRB1 and RTN4Rs have effects on synaptogenesis both *in vitro* and *in vivo*, ADGRB1 knockout mice show reduced synaptogenesis in the hippocampus whilst a triple KO of Nogo receptors in mice resulted in abnormally elevated synaptogenesis in the same region (Wills et al., 2012; Zhu et al., 2015). ADGRB1 deficient mice have additional deficits in synaptic plasticity and spatial learning (Zhu et al., 2015). The phenotype of Nogo receptor deficient mice has been attributed to the interaction between RTN4R and myelin-associated inhibitors like MAG and Nogo, but this does not explain the fact that individual knockdowns of RTN4RL1 and RTN4RL2 result in increased numbers of excitatory synapses in neuronal cultures. The difference in KO phenotypes between ADGRB1 and RTN4Rs suggest that the interaction could be inhibitory, with binding preventing normal signalling of one of the receptors, or that additional receptors are involved. This makes it more challenging to elucidate the exact function of these interactions *in vivo*.

6.3.2 Levels of downstream effectors can help determine if RTN4Rs are activating ligands of ADGRB1

ADGRB1 is known to signal through G-protein α-12/13, which is coupled to the activation of small GTPase RhoA (Stephenson et al., 2013). Additionally, ADGRB1 is known to signal through Rac1 GTPases independently of G-protein activation and couples with different Rac1-guanine nucleotide exchange factor modules during synaptogenesis and phagocytosis (Duman et al., 2013). To determine if Nogo receptors are activating (or inhibitory) ligands of ADGRB1, one could monitor levels of activated RhoA and Rac1

in cells expressing ADGRB1 and exposed to soluble ectodomains of RTN4Rs. Although I performed some preliminary experiments investigating RhoA activation in response to ADGRB1-RTN4R binding, I failed to see any activation of RhoA in HEK293 cells (data not shown). However, a more relevant cell line, such as the neuroblastoma-derived SHSY-5Y or primary hippocampal cultures, might provide a more accurate assessment of ADGRB1 activation/inhibition. Importantly, previous studies have shown that transfection of full length ADGRB1 into HEK293 cells results in a slight increase in RhoA activation above baseline (Stephenson et al., 2013), raising the possibility that inhibitory ligands for ADGRB1 could exist.

One explanation for the lack of evidence for ADGRB1 activation could be that ADGRB1-RTN4R interactions are purely adhesive and do not trigger downstream signalling via either receptor. Non-activating ligands of adhesion GPCRs have been previously described (Safaee et al., 2013), although it is difficult to conclude with certainty if an interaction is non-activating, if the subsequent response was below the detection threshold, or if signalling is taking place through alternative pathways.

6.4 Other possible applications of CRISPRa extracellular interaction screening

CRISPRa extracellular interaction screening is not restricted to the testing of single defined ligands to identify cell surface receptors. The membrane protein gRNA library and pooled screening approach can be used together with more physiologically relevant selection assays or reagents to identify relevant cell surface interactions involved in a number of biological processes.

6.4.1 Uncovering novel viral receptors

Viruses are intracellular obligate parasites which rely on host machinery to replicate (Alkhatib, 2009). This involves breaching the cell membrane to insert viral DNA within a host cell. Many viruses show specific host cell tropism, suggesting that specific host factors play a part in enabling viruses to invade host cells. Well-established examples are the CD4 receptor which, along with CXCR4 and CCR5 co-receptors, interact with Human Immunodificiency Virus (HIV) coat protein gp120 to initiate HIV fusion with the host cell membrane (Alkhatib,

2009). HIV co-receptor CCR5 has subsequently become a key target for the development of antiviral drugs and immunotherapy (Lopalco, 2010). CRISPRa screening could be used to uncover novel viral receptors by selecting for cells which gain an ability to be infected by fluorescently-labelled viral or virus-like particles from a population of cells transduced with the membrane protein gRNA library. In fact, other studies have demonstrated the use of CRISPRa for identifying host cell factors involved in influenza infection using genome-wide gRNA libraries (Heaton et al., 2017). A focused gRNA library like the membrane protein gRNA library provides a smaller library size whilst focusing on identifying extracellular factors that can be easily targeted to prevent viral entry.

6.4.2 Identifying extracellular interactions underlying cancer metastasis

The pooled rather than arrayed format of CRISPRa screening also provides opportunities for use with other types of selection assays, such as *in vivo* models of cancer metastasis. Metastasis refers to the spread of cancer cells from the primary tumour to surrounding tissues and is estimated to be the primary cause of cancer mortality as well as relapse (Seyfried and Huysentruyt, 2013). As such, the discovery of the mechanisms of primary tumour cell extravasation, immune evasion, and subsequent establishment in different tissues is paramount to developing more effective anticancer therapy. Established metastatic models in mice involve the injection or engraftment of cancer cells in host mice and subsequent review of the resulting metastases. Such models could be used to select for receptors that confer an advantage or disadvantage for cancer cells to undergo metastasis by injecting mice with a pool of cancer cells transduced with gRNA targeting membrane proteins. However, a challenge of performing *in vivo* studies is that the maximum number of cells injected or engrafted is typically quite low, thus requiring a large number of animals to provide sufficient coverage for screening with pooled libraries. One solution is to use smaller libraries targeting subsets of genes likely to contribute to metastasis, and pre-select highly active gRNA. Another important consideration would be the pre-existing metastatic ability of the cancer cell line used for screening. A cell line that spontaneously forms many metastatic colonies might make it difficult to detect enrichment of gRNAs causing gain-of-function metastatic ability over baseline.

6.5 Conclusion

In conclusion, I have shown that CRISPRa is a feasible approach for overexpression of cell surface receptors and established a CRISPRa screening workflow to identify novel extracellular interactions. CRISPRa screening provides a complimentary approach to currently available techniques for identifying extracellular interactions, with the advantage of allowing cost-efficient genome-scale screening. However, it is limited to detecting cell surface binding partners and false negatives may occur from inactive gRNAs. Nonetheless, CRISPRa screening can be used to detect novel interactions and could serve as an approach for identifying multiple receptors to a defined ligand or pool of ligands in a single screen.