

Genome-scale identification of cellular pathways required for cell surface recognition



Sumana Sharma

Wellcome Trust Sanger Institute
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

DECLARATION

I hereby declare that the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other, University. This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. This thesis does not exceed the word limit set by the Faculty of Biology.

Sumana Sharma
December 2017

ACKNOWLEDGEMENTS

First, I would like to express my sincere gratitude to my advisor Dr. Gavin Wright for the guidance, motivation, patience and the encouraging words all through my PhD. Thank you mostly for encouraging me to appreciate technology development related research! Thank you also to the members of my thesis committee, Dr. Julian Rayner and Dr. Kosuke Yusa for their insightful comments and encouragement, and also for providing a different perspective to my research.

Deepest gratitude to the members of the Wright Lab who have been a cherished combination of colleagues and friends. Special thank you to Josefin, who has provided instrumental support to my research throughout the PhD; Nicole, Shan, Laura and Kirsty for readily sharing their reagents and technical abilities; and Cecile and Enrica whose openness and scientific insight I always admire. I would also like to thank the members of the flow cytometry core-facility (Bee Ling Ng, Jennifer Grahm, Chistopher Hall and Sam Thompson) for their flow cytometry training and advice; Mandy Sanders for help with sequencing sample submission; and Sanger Institute Core Sequencing facility for sequencing.

Most of all, I would like to thank my parents, whose encouragement and love has made possible my journey from the hills of Rohote to Cambridge! Also, thanks to my beloved siblings, Rashmi and Prashant, for their guidance and unwavering support. Thank you to James- you have been amazingly patient and always been there to support me. Thanks to my friends, John, Katie and Tom who have made my time during my PhD a very enjoyable one.

ABSTRACT

A range of biochemically diverse molecules located in the plasma membrane—such as proteins, glycans, and lipids—mediate cellular recognition events, initiation of signalling pathways, and the regulation of processes important for the normal development and function of multicellular organisms. Interactions mediated by cell surface receptors can be challenging to detect in biochemical assays, because they are often highly transient, and membrane-embedded receptors are difficult to solubilise in their native conformation. The biochemical features of low-affinity extracellular protein interactions have therefore necessitated the development of bespoke methods to detect them.

Here, I develop a genome-scale cell-based genetic screening approach using CRISPR-Cas9 knockout technology that reveals cellular pathways required for specific cell surface recognition events. Using a panel of high-affinity monoclonal antibodies, I first establish a method from which I identify not only the direct receptor but also other required gene products, such as co-receptors, post-translational modifications, and transcription factors contributing to antigen expression and subsequent antibody-antigen recognition on the surface of cells. I next adapt this method to identify cellular factors required for receptor interactions for a panel of recombinant proteins corresponding to the ectodomains of cell surface proteins to the endogenous surface receptors present on a range of cell lines. In addition to finding general cellular features recognised by many ectodomains, I also identify direct interaction partners of recombinant protein probes on cell surfaces together with intracellular genes required for such associations.

Using this method, I identify IGF2R as a binding partner for the R2 subunit of GABAB receptors, providing a mechanism for the internalisation and regulation of GABAB receptor signalling. The results here demonstrate that this single approach can identify the molecular nature and cell biology of surface receptors without the need to make any prior assumptions regarding their biochemical properties.

TABLE OF CONTENTS

List of figures	xiii
List of tables	xvii
1 Introduction	1
1.1 Molecules mediating cell surface recognition	1
1.1.1 Proteins	2
1.1.2 Glycans	3
1.1.3 Lipids	3
1.2 The challenges of studying extracellular ligand-receptor interactions	5
1.3 Methods to study extracellular protein-protein interactions	6
1.3.1 General overview	6
1.3.2 Cell-based binding assays	7
1.3.3 Cell-free protein interaction approaches	9
1.3.4 Mass-spectrometry-based methods	11
1.4 Methods to study extracellular protein-glycan interactions	12
1.4.1 General overview	12
1.4.2 Binding inhibition approaches	13
1.4.3 cDNA expression methods	14
1.4.4 Glycan arrays	14
1.5 Limitations of the existing methods	15
1.6 Loss-of-function genetic approaches to study cellular recognition process	17
1.6.1 General overview	17
1.6.2 Study of naturally occurring mutants	17
1.6.3 Genetic screening approaches	18
1.6.4 CRISPR-Cas9 approach	21

1.6.5	The scope of CRISPR-Cas9 knockout screening system in the context of cellular interactions	25
1.7	Thesis aims	26
2	Materials and Methods	27
2.1	Buffers/Media/Solutions	27
2.2	Generation of expression plasmids	28
2.2.1	Cloning of IGF2R expression construct	30
2.3	Recombinant protein production using HEK-293-6E cells	31
2.4	Protein purification and quantification	32
2.4.1	Immobilized metal ion affinity chromatography	32
2.4.2	Determination of protein expression and quality	33
2.4.3	Plate-based direct protein interaction assay	34
2.5	Human Cell line culture	35
2.6	Flow cytometry based cell binding assay	36
2.7	Genome-wide screening and validation	38
2.7.1	Construction of gRNA expression vector	38
2.7.2	Lentiviral production	38
2.7.3	Lentivirus transduction	39
2.7.4	Cell surface phenotyping, selection and amplification of selected gRNAs	41
2.7.5	Data analysis	42
3	Development of a CRISPR-Cas9 based knockout screen for cellular recognition	45
3.1	Introduction	45
3.1.1	Monoclonal antibodies are ideal probes for the study of cell surface recognition	45
3.1.2	Genome-scale knockout screening approach has the potential to determine the specificity of mAbs.	46
3.1.3	Considerations for knockout screening approach to identify directly interacting receptors	46
3.1.4	Knockout approach used in this study	47
3.1.5	Scope of this chapter	48
3.2	Results	49
3.2.1	Generation of stable cell lines expressing Cas9	49
3.2.2	Quality control of the genome-scale mutant cell library	54

3.2.3	Genome-scale screens using monoclonal antibodies . . .	58
3.3	Discussion	74
4	Application of the genetic screening approach to identify interactions mediated by recombinant proteins	79
4.1	Introduction	79
4.1.1	<i>P. falciparum</i> ‘merozoite’ protein library	80
4.1.2	The platelet receptor library	82
4.1.3	Scope of this chapter	82
4.2	Results	83
4.2.1	Proof-of-concept study: BSG and RH5 interaction	83
4.2.2	Heparan sulphates serve as common factors for cellular recognition	87
4.2.3	Investigating extracellular interactions mediated by human proteins	92
4.3	Discussion	99
5	Identification and characterisation of IGF2R as an endosomal receptor for GABBR2	103
5.1	Introduction	103
5.1.1	Introduction to GABA-B receptors	103
5.1.2	Introduction to IGF2R	106
5.1.3	Scope of this chapter	108
5.2	Results	108
5.2.1	IGF2R is required for the binding of GABBR2 to HEK-293-E cells	108
5.2.2	Plasma membrane expression of IGF2R is required for the binding of GABBR2 ectodomain	110
5.2.3	IGF2R ectodomain and GABBR2 ectodomain directly interact	111
5.2.4	GABBR2 interacts with IGF2R in a M6P-dependent manner	112
5.3	Discussion	114
6	Discussion and future directions	117
6.1	Overview for screening parameters	120
6.1.1	Sorting strategy	120
6.1.2	The timing of phenotypic selections	122

6.2	Potential of genome-scale KO screens using mAbs for the study of receptor biology	123
6.2.1	Potential for the study of receptor biology in a high-throughput manner	124
6.3	Assessment of the approach to identify receptors of soluble protein ectodomains	124
6.4	Interaction between IGF2R and GABBR2	127
6.5	Concluding remarks	129
	Bibliography	131
	Appendix A	163

LIST OF FIGURES

1.1	Overview of type II CRISPR-Cas9 mediated gene editing . . .	22
1.2	Outline of the major approaches that are utilised for the study of extracellular interactions.	26
2.1	Schematic diagram representing the plasmids used for protein production using the HEK-293-6E protein expression system .	29
3.1	Time-dependent decrease in cell surface expression of membrane receptors is observed in Cas9-expressing cells that are transduced with gene specific gRNAs	50
3.2	Selecting clonal cell lines with high Cas9 activity for efficient genome-scale genetic screening	52
3.3	Cas9-expressing human cell lines generated in this study demonstrate high <i>GFP</i> and <i>BSG</i> cleavage efficiency	53
3.4	Mutant cell libraries can be created reproducibly by transduction with the lentiviral gRNA library	55
3.5	gRNAs targeting genes involved in essential biological processes were depleted during extended culture of the mutant cell library.	57
3.6	Schematic of the genetic screening approach.	58
3.7	Cell sorting strategy for a proof-of-principle genome-scale screening for recognition of BSG in HEK-293-E cells by an anti-BSG mAb	59
3.8	A positive selection screen using anti-BSG mAb demonstrates the successful application of a CRISPR-based KO screen to identify factors required for epitope recognition by a mAb . . .	61
3.9	Cell sorting profile for flow-cytometry based CRISPR-Cas9 screen in HEK-293-E cells with an anti-CD59 antibody	63

3.10	An improved CRISPR-mediated forward genetic screen identifies the genes required for the trafficking of the receptor in addition to the gene encoding the antibody epitope.	64
3.11	Several genes involved in the early stages of N-glycan biosynthesis pathway were identified in the screen using an anti-CD59 mAb.	65
3.12	Genome-scale CRISPR-Cas9 screen using anti-CD47 mAbs targeting cell surface protein reveal members of SRP dependent ER protein translocation pathway	68
3.13	Genome-scale KO screen using an anti-GYPA antibody reveals mAb epitope and factors required for the cell surface GYPA display	70
3.14	Genome-scale loss-of-function screen using an anti-integrin $\alpha\text{IIb}\beta\text{1}$ mAb identifies the subunit encoding the antibody epitope and components of the cytoplasmic Arp2/3 complex	71
3.15	Genes required for protein export, N-glycan biosynthesis and general housekeeping functions were enriched in, and shared between, cells selected for the loss of cell surface mAb staining	73
4.1	Cellular organization and invasion process of <i>P. falciparum</i> merozoite.	81
4.2	RH5 binding to HEK293 was not completely dependent on BSG but was heat-labile.	84
4.3	Cell-based genetic screens identified BSG and heparan sulfate as independent receptors for <i>P. falciparum</i> RH5 on HEK293 cells.	85
4.4	The total observed binding of RH5 to HEK-293-E cell surface is the sum of independent contributions from BSG and HS.	86
4.5	All three genome-wide screens using merozoite proteins revealed the role of HS-biosynthesis pathway in mediating cell surface interactions	88
4.6	The binding of multiple merozoite proteins to the cell surface can be completely abrogated by inactivating <i>SLC35B2</i>	89
4.7	Schematics for the approach to determine the fractional contribution of HS towards binding.	91
4.8	The pre-screening strategy identifies human proteins the binding of which is not dependent solely on <i>SLC35B2</i>	92

4.9	Genome-scale screens identify HS-biosynthesis pathway when ligands that lose majority of binding upon targeting <i>SLC35B2</i> are used as screening probes	94
4.10	Cell surface receptors were identified using cellular genetic screens	95
4.11	Strategy for genetic screening using recombinant proteins . . .	96
4.12	Genome-scale screen using TNFRSF9 ectodomain as a sorting ligand identifies the interaction partner along with the p53 pathway.	98
5.1	Schematics of GABA _B receptors	104
5.2	Schematics of mannose-6-phosphate cellular transport pathway and structure of IGF2R	107
5.3	A genome-scale screen using the ectodomain of GABBR2 identifies IGF2R and genes involved in endosomal function and trafficking	109
5.4	Targeting <i>IGF2R</i> and <i>WDR7</i> on HEK-293-E cells leads to the loss of binding of GABBR2 ectodomain	110
5.5	Binding of GABBR2 can be conferred to NCI-SNU-1 cell line that do not display IGF2R on their surface	111
5.6	Interaction between IGF2R and GABBR2 is direct	112
5.7	Interaction between IGF2R and GABBR2 is dependent on the M6P-modified N-linked glycans of GABBR2	113
5.8	Proposed mechanism of IGF2R mediated GABAB receptor internalisation	115
A.1	Enlarged version of figure 3.10B for better clarity of gene names	167
A.2	Enlarged version of figure 3.12 for better clarity of gene names	168
A.3	Enlarged version of figure 3.14A for better clarity of gene names	169
A.4	Genome-scale KO screen using an anti-CD58 mAb (BRIC5) identifies <i>CD58</i> as the highest enriched gene in the sorted population	170
A.5	Primary screen for binding of 11 merozoite proteins to six cell lines	170
A.6	Binding of merozoite proteins RAMA and CyRPA to HEK-293-E cells can be completely blocked by soluble heparin	171
A.7	Summary of ‘overlapping factors’ identified in at least two out of six screen that identified the HS-biosynthesis pathway	171

LIST OF TABLES

2.1	Summary of reagent quantities for transfection of HEK-293-6E cells.	32
2.2	Growth condition for cell lines used in this study	36
3.1	Summary of mAbs used for genome-scale loss-of-function screening.	66
3.2	Genes involved in protein export (KEGG annotation) identified in screens carried out using anti-CD47 mAb on day nine post mutant library generation	67
3.3	FDR of identification of the genes encoding direct receptor in a genome-scale screening approach using monoclonal antibodies	72
4.1	Background of the ligands that demonstrated dependency on <i>SLC35B2</i> for binding to HEK-293-E cells.	93
4.2	Background of the identified receptor-ligand interaction partners.	96
5.1	Non-lysosomal interaction partners of IGF2R	108
6.1	Summary of genome-scale KO screens using the CRISPR-Cas9 approach to study cellular recognition events	118
A.1	Primers used for cloning IGF2R into protein expression vectors Fwd=Forward, Rev=Reverse	163
A.2	Sequences of guide RNAs used in this study	163
A.3	Summary of the screening parameters for all screens carried out in this study	164
A.4	Primer sequences for amplifying gRNA and NGS.	166
A.5	Summary of 'other factors' identified with FDR<0.05 in at least two out of the seven different screens carried out using mAbs .	172
A.6	Genes identified in KO screens (FDR<0.05) using recombinant proteins as screening probes	173

