Elucidating Biophysical and Structural Mechanisms of Multivalent Lectin-Glycan Binding using Glyconanoparticle Probes

James Isaac Hooper

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Food Science and Nutrition

May 2023

The candidate confirms that the work submitted is their own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 1 contains some text from the following book chapter:

Hooper, J.; Guo, Y.; Zhou, D. Polyvalent Glycan-Quantum Dots as Multifunctional Structural Probes for Multivalent Lectin-Carbohydrate Interactions. In *Glycan-Based Cellular Communication: Techniques for Carbohydrate-Protein Interactions*, ACS Symposium Series, Vol. 1346; American Chemical Society, 2020; pp 47-66.

James Hooper performed the literature review and wrote the manuscript. Dr Yuan Guo and Professor Dejian Zhou provided comments and edits towards the final manuscript.

Chapter 3 is based on work from the following publication:

Hooper, J.; Liu, Y.; Budhadev, D.; Ainaga, D. F.; Hondow, N.; Zhou, D.; Guo, Y., Polyvalent Glycan Quantum Dots as a Multifunctional Tool for Revealing Thermodynamic, Kinetic, and Structural Details of Multivalent Lectin–Glycan Interactions. *ACS Appl. Mater. Interfaces.* **2022**. 14 (42), 47385–47396.

James Hooper wrote the manuscript and conducted all experiments and data analysis, excluding the STEM data which was performed and analysed by Dario Fernandez Ainaga and Dr Nicole Hondow. Ligand precursors were synthesised by Dr Yuanyuan Liu and Dr Darshita Budhadev. Dr Yuan Guo and Professor Dejian Zhou supervised the project and provided comments and edits towards the final manuscript.

Chapter 4 is based on work from the following publication:

Hooper, J.; Budhadev, D.; Fernandez Ainaga, D. L.; Hondow, N.; Zhou, D.; Guo, Y. Polyvalent Glycan Functionalized Quantum Nanorods as Mechanistic Probes for Shape-Selective Multivalent Lectin-Glycan Recognition. *ACS Appl. Nano Mater.* **2023**. 6 (6), 4201-4213.

James Hooper wrote the manuscript and conducted all experiments and data analysis, excluding the STEM imaging which was performed by Dario Fernandez Ainaga and Dr Nicole Hondow. Ligand precursors were synthesised by Dr Darshita Budhadev. Dr Yuan Guo and Professor Dejian Zhou supervised the project and provided comments and edits towards the final manuscript.

Acknowledgements

Firstly, I would like to thank my primary supervisor, Dr Yuan Guo, for their wonderful guidance, teaching and support throughout the course of my PhD, and for providing and overseeing this excellent project. Also, thank you to my co-supervisor, Professor Dejian Zhou, who has overseen my progress for the entirety of my higher education. I would also like to the thank the School of Food Science and Nutrition at the University of Leeds for offering me this scholarship to allow me to conduct this project, and the RSC and BBSRC for their additional contributions to funding this research.

In addition, I would also like to thank the people who have assisted and collaborated with me and the group to aid the research presented in this thesis: Dr Darshita Budhadev, Dr Yuanyuan Liu, Dr Emma Poole and Dr Chadamas Sakonsinsiri (Zhou Group, University of Leeds) for synthesising many of the glycans, ligands and linker molecules. Dario Fernandez Ainaga and Dr Nicole Hondow (School of Chemical and Process Engineering, University of Leeds) who performed all cryoEM and S/TEM imaging. Dr Ralf Richter (School of Biomedical Sciences and the School of Physics and Astronomy, University of Leeds) who co-supervised the project investigating surface affinity enhancing mechanisms for multivalent lectin-glycan interactions using QCM-D. Alexander St John, Martin Lepšík and Dr Emanuele Paci (FBS, University of Leeds) who performed the MD simulations of DC-SIGN/R informed by our smFRET data. Inga Nehlmeier, Amy Madeleine Kempf, Markus Hoffmann Nadine Krüger and Professor Dr Stefan Pöhlmann (German Primate Center, Göttingen, and Faculty of Biology and Psychology, Georg-August-University Göttingen) who prepared the pseudotyped viruses and performed viral inhibition assays of our lectin-nanoparticles against pseudotyped and authentic SARS-CoV-2. Dr Yue Ma and Dr Christos Pliotas (FBS, University of Leeds) who guided me in performing the cwEPR measurements of spin-labelled lectins. Dr Bogachan Tahirbegi and Dr Liming Ying (Imperial College London) who guided me in performing the smFRET analysis of FRET pair labelled DC-SIGN and DC-SIGNR. I would also like to thank those who majorly helped in training, facilities managing and general availability for advice on various instruments throughout my PhD: Dr Stewart Warriner (HRMS); Dr Martin Huscroft and Dr Jeanine Williams (LCMS); Dr Mark Howard (NMR); Dr Joanna Sier and Miles Ratcliffe (FSN facilities); Itzel Garcia Monge and Dr Xiaoli Zhang (SUV preparation and QCM-D); Abigail Roberts and Dr Ruth Hughes (confocal

microscopy); and the rest of the staff within the School of Chemistry and the School of Food Science and Nutrition.

Furthermore, I would like to say a big thank you to my lovely colleagues within the group: Darshita, who has been a fantastic friend, lab buddy and genuinely wonderful person throughout the entirety of my PhD. Liz Kalverda for training and assisting me with protein production, and for being a great teacher and friend. Emma, Anjo and Yuanyuan for teaching me how to synthesise various ligands and nanoparticles and making me feel welcomed into the group. Zeyang, Rahman, Akshath, Xinyu, Qian, Xinran and Zhiming, for their continual friendship and support throughout the various stages of my PhD. I will deeply cherish my time working in this group.

Lastly, I wish to thank my amazing wife, Stephanie, who has supported me throughout it all (amidst my complaining and my celebrating), and who has ensured that I have been continually topped up with tea during the writing of this thesis. You are truly wonderful and I could not have done this without you.

Abstract

Multivalent lectin-glycan interactions (MLGIs) are widespread and vital for biology, and also hold the key to many therapeutic applications. However, the underlying structural and biophysical mechanisms for many MLGIs remain poorly understood, limiting our ability to design glycoconjugates that can potently target specific MLGIs for therapeutic intervention. Glycosylated nanoparticles (NPs) have recently emerged as powerful biophysical probes for studying MLGIs, revealing key information regarding the binding behaviours of different lectins, which can subsequently inform their therapeutic potential. Despite these advances, however, the structural and biophysical mechanisms behind the specific binding behaviours of different lectins towards glycan-NPs is largely unknown. This research focuses on furthering our understanding of these mechanisms by exploiting the unique multifunctional properties of glycan-conjugated quantum dots (glycan-QDs) and gold nanoparticles (glycan-GNPs), as well as other nanoscale tools, to improve our understanding of the structural, thermodynamic and kinetic rationale behind the MLGIs of two closely related and immunologically interesting tetrameric lectins, DC-SIGN and DC-SIGNR.

DC-SIGN and DC-SIGNR (collectively, DC-SIGN/R) display near-identical mannose-binding motifs and have both been identified to facilitate viral infection. Despite this, these lectins have been shown to have very different affinities to some of the same viral glycoproteins or other glycoconjugates, both in solution and on cell surfaces. This has been attributed to their differences in multivalent spatial specificity, induced by differences in their tetrameric structures. For example, our group previously demonstrated that upon incubation with the same glycan-QDs, DC-SIGN bound simultaneously to individual QDs, whereas DC-SIGNR formed much weaker crosslinking interactions. Here, a QD-FRET technique for determining the thermodynamic and kinetic contributions of MLGIs in solution has been developed, and has revealed that, though both DC-SIGN/R display similar enthalpically driven MLGIs with mannosylated QDs (4× that of monovalent binding), DC-SIGNR incurs a greater entropic penalty and slower kinetics which can be attributed to its less favourable crosslinking binding mode (Chapter 3). Furthermore, a 16 amino acid C-terminal (not present in DC-SIGNR) has been shown to contribute to the specificity of DC-SIGN MLGIs; the removal of which completely alters the thermodynamics towards entropically driven binding, and enables the possibility of crosslinking. In addition to these assays, DC-SIGN/R-functionalised supported lipid bilayers

have been developed to better replicate the native environment of lectins on cell membranes (Chapter 5). Here, QCM-D and cryo-EM studies with mannosylated GNPs show that these different binding modes can still be observed on membrane surfaces, where DC-SIGNR·glycan-GNP binding is highly dependent upon the lectin surface density, whereas DC-SIGN binding shows no such dependency. This may therefore contribute to the observed differences between DC-SIGN/R MLGIs in both solution-phase and cell surface assays.

The crosslinking behaviour of DC-SIGNR can be attributed to the inability of its four binding sites to bridge the glycan display of an individual glycan-NP. Here, a single molecule FRET technique has been developed, using a new tetramer FRET pair labelling strategy, in conjugation with MD simulations to provide an estimate of the inter-binding site distances of DC-SIGN/R (Chapter 6). These results demonstrate a broader tetrameric model for DC-SIGNR which may explain its inability to tetravalently bind to glycan-NPs of these sizes. However, this has been demonstrated to be overcome using glycosylated quantum rods (glycan-QRs), whereby the reduced surface curvature of the cylindrical middle section of these nanomaterials is able to encourage tetravalent binding of all four binding sites of DC-SIGNR, reducing the amount of crosslinking and thus resulting in stronger MLGIs (Chapter 4).

Additionally, understanding of these interactions has also paved the foundation of other innovative nano-therapeutics which apply these principles. Here, DC-SIGN-conjugated NPs have been developed to target the glycosylation sites of the SARS-CoV-2 S protein, showing to effectively neutralise a range of pseudotyped and authentic SARS-CoV-2 variants (Chapter 8).

Overall, the findings presented in this thesis reveal crucial information regarding the binding of key MLGIs which can be used to help inform the rational design of more potent and specific lectin or glycan targeting agents for potential therapeutic applications.

Table of Contents

Acknowledgements ..... **iv**
Abstract ..... **vi**
Table of Contents ..... **viii**
List of Tables ..... **xiii**
List of Figures ..... **xvi**
List of Abbreviations ..... **xxxv**
Chapter 1 Introduction **39**
1.1 Overview 39
1.2 Multivalent Lectin-Glycan Interactions 40
1.2.1 Protein-Carbohydrate Interactions in the Immune System 40
1.2.2 Lectins, LGIs and MLGIs 41
1.2.3 DC-SIGN 47
1.2.4 DC-SIGNR 53
1.2.5 Limitations and Gaps in Current Research 54
1.3 Glyconanotechnology Strategy and Application in the Study of MLGIs 55
1.3.1 Introduction to Glyconanotechnology 55
1.3.2 Glyconanoparticle Design 57
1.3.3 Spectroscopic Methods for Affinity Determination 63
1.3.4 Electron Microscopy and Dynamic Light Scattering 70
1.3.5 MLGIs on Surfaces 74
1.3.6 Spectroscopic Rulers 76
1.4 Summary and Research Aims 79
Chapter 2 General Materials and Methodology **89**
2.1 Materials and Methods 89
2.1.1 Materials 89
2.1.2 General instrumentation and computation 90
2.1.3 DHLA-EG_n-Glycan 91
2.1.4 Production and characterisation of proteins 92
2.2 Experimental methodology and instrumentation 96
2.2.1 Experimental Theory 96
2.2.2 General Experimental Method 102

Chapter 3 Polyvalent Glycan-Quantum Dots as a Multifunctional Tool for Revealing Thermodynamic, Kinetic and Structural Details of Multivalent Lectin-Glycan Interactions 107

3.1 Introduction 107

3.2 Experimental 110

3.2.1 QD-EG_n-Glycan 110

3.2.2 DC-SIGN, DC-SIGNR and DC-SIGN-C 111

3.2.3 QD-FRET thermodynamic assays 113

3.2.4 QD-FRET kinetic assays 114

3.2.5 QD STEM imaging 115

3.3 Results and Discussion 115

3.3.1 Materials synthesis and characterisation. 115

3.3.2 Quantifying binding affinity and thermodynamics *via* QD-FRET. 117

3.3.3 Role of C-terminal segment in DC-SIGN multivalent binding. 124

3.3.4 Investigating MLGI kinetics by QD-FRET. 127

3.4 Conclusions 132

Chapter 4 Polyvalent Glycan-Quantum Rods as Multifunctional Mechanistic Probes for Shape-Selective Multivalent Lectin-Glycan Recognition 136

4.1 Introduction 136

4.2 Experimental 139

4.2.1 Preparation of QR-DiMan. 139

4.2.2 QR-FRET assays 140

4.2.3 Lectin Titration Analysis via DLS 141

4.2.4 S/TEM QR images 142

4.3 Results and Discussion 142

4.3.1 QR-DiMan Preparation and Characterisation 142

4.3.2 Binding Affinity Determination *via* FRET 143

4.3.3 Binding Mode Determination *via* FRET and Hydrodynamic Size Analysis 148

4.3.4 Binding Mode Rationale *via* S/TEM Imaging 153

4.4 Conclusion 161

Chapter 5 Investigating the binding of glycan-gold nanoparticles upon DC-SIGN/R conjugated supported lipid bilayers using QCM-D and cryoTEM 164

5.1 Introduction 164

5.2 Experimental 168
5.2.1 G5-Ct 168
5.2.2 G5-Man 169
5.2.3 Protein Production and Non-Specific Labelling 170
5.2.4 DOPC and tris-NTA DODA SUVs 171
5.2.5 QCM-D 172
5.2.6 Confocal Microscopy 172
5.2.7 CryoTEM 173
5.3 Results and Discussion 174
5.3.1 Materials Synthesis and Characterisation 174
5.3.2 SLB-lectin formation 176
5.3.3 SLB-lectin·glycan-GNP binding 181
5.3.4 SLB-lectin·glycan-GNP affinity quantification 183
5.3.5 SLB-lectin·glycan-GNP kinetics 189
5.3.6 Imaging glycan-GNP bound SLB-lectin surfaces 192
5.4 Conclusion 198
5.5 Future Direction 199
Chapter 6 Evaluation of the inter-carbohydrate binding domain distances of DC-SIGN and DC-SIGNR using single molecule FRET 202
6.1 Introduction 202
6.2 Experimental 204
6.2.1 Donor and Acceptor Labelled DC-SIGN/R 204
6.2.2 GBDC-SIGN 206
6.2.3 HTDC-SIGN/R 206
6.2.4 Site-Directed Spin-Labelling of DC-SIGN/R 209
6.2.5 FRET Assays 210
6.2.6 SmFRET 210
6.2.7 Cw-EPR 211
6.2.8 Computational Methods 211
6.3 Results and Discussion 212
6.3.1 Double purification strategy for FRET pair labelling DC-SIGN/R 212
6.3.2 Inter-CRD Distance Determination *via* smFRET 216
6.3.3 Comparison of Inter-CRD Distances *via* Molecular Dynamics 219
6.4 Conclusions 223

Chapter 7 A Polyvalent Nano-Lectin Potently Neutralises SARS-CoV-2 by Targeting Glycans on the Viral Spike Protein 226

7.1 Introduction 226

7.2 Experimental 228

7.2.1 DC-SIGN EDC and CRD 228

7.2.2 LA N-terminal labelling of ECD and CRD 229

7.2.3 G13-Ct 229

7.2.4 G13-(EG_7-OH)$_X$ 230

7.2.5 G13-ECD and G13-CRD. 231

7.2.6 QD-DiMan and QD-Gal 233

7.2.7 QD Fluorescence Quenching assays 234

7.2.8 Pseudo-SARS CoV-2 virus preparation and inhibition studies. 234

7.2.9 Inhibition of authentic SARS-CoV-2 infection by G13-CRD. 235

7.3 Results 235

7.3.1 Lectin-GNP Conjugation 235

7.3.2 Lectin-GNP Glycan Binding 239

7.3.3 Inhibition of entry of SARS-CoV-2 S protein bearing pseudotypes into Vero76 cells by lectin-GNPs 243

7.3.4 Inhibition of authentic SARS-COV-2 entry into Vero76 cells by lectin-GNPs 247

7.4 Discussion and conclusions 248

Chapter 8 Concluding Remarks and Future Directions 255

8.1 Determinations of Biophysical Contributions Influencing MLGIs 256

8.2 Determination of Structural Contributions Influencing MLGIs 257

8.3 Determination of Surface-Phase Influences on MLGIs 259

8.4 Therapeutic Application 260

Appendix A General Information 263

A.1 Amino acid sequences 263

A.2 SARS-CoV-2 S protein mutations 264

Appendix B Glycan-NP Characterisation 266

B.1 NP Glycan Display 266

B.2 Spectral overlap and Förster radius of QD-Atto594 FRET pair 267

B.3 QD and Atto594 labelled lectin fluorescence-concentration relationships 269

B.4 QD-DiMan·Lectin Dissociation Tests 269

Appendix C Lectin-Glycan Binding Assays 271
C.1 Thermodynamic QD-FRET studies 271
C.2 Kinetic QD-FRET studies 272
C.3 QR Hydrodynamic Size Studies 272
C.3.1 Lectin Titration Analysis *via* DLS 272
C.3.2 QR-DiMan·Lectin Dissociation with Mannose *via* DLS 275
C.4 QR STEM Images 279
C.5 G5-Man bound DC-SIGNR-SLB 281
C.6 Inhibition of pseudo-typed VSV entry into Vero76 cells by lectin-GNPs 282
Appendix D Lectin Inter-CRD Measurements 284
D.1 SmFRET FRET Efficiency Distributions 284
D.1.1 DC-SIGN-Atto488 Control 284
D.1.2 HTDC-SIGN/R smFRET signal 287
D.2 Reweighted MD Distances 292
Appendix E Equation Derivations 295
E.1.1 Derivation of Second Order FRET Rate Equation for Association 295
E.1.2 Pseudo-First Order FRET Rate Equation for Dissociation 296
E.2 Model of QR Glycan Display and Lectin Binding 297
Appendix F Preliminary Experiments 301
F.1 Investigating QD-Man Glycan Density on MLGIs 301
F.2 QD-EG_3-DiMan·lectin Binding Thermodynamics 304

List of Tables

Table 3.1 Fitting parameters for the FRET curves of QD-DiMan binding with labelled DC-SIGN, DC-SIGN-C and DC-SIGNR at varying temperatures (R^2 >0.99 for all fits, SDs represent fitting errors). 120

Table 3.2 Summary of the binding thermodynamic parameters for QD-DiMan binding with DC-SIGN, DC-SIGN-C, and DC-SIGNR (SDs are propagated from the fitting errors obtained in Figure 3.9, according to Equation 3.4). 123

Table 3.3 Summary Kinetic Parameters for QD-DiMan Binding with DC-SIGN, DC-SIGN-C, and DC-SIGNR (SDs Represent Fitting Errors). 129

Table 4.1 Summary of fitting parameters obtained from the FRET ratio-concentration relationship for QR-DiMan binding with DC-SIGN and DC-SIGNR fitted with the Hill equation (Equation 4.1). Where, PQR is the protein:QR molar ratio, $F\mathrm{max}$ is the FRET ratio at saturation, $K\mathrm{d}$ is the dissociation constant, β is the enhancement factor over the monovalent interaction and N is the number of glycans per QR (R^2 > 0.99 for all fits). 147

Table 4.2 Summary of fitting parameters obtained from the Gaussian fits of the nearest neighbour distances (NND) of QR-DiMan particles binding with DC-SIGN, DC-SIGNR or DC-SIGNR with 2 mM Man. Where ppNND is the perimeter-perimeter NND and ccNND is the centre-centre NND (where 2.9 < ppNND < 7.7 nm). 156

Table 5.1 Summary of parameters for the binding of DC-SIGN/R with different glycan-GNPs, derived from the quenching efficiency, $E = \mathrm{Pn}/(K\mathrm{d}n + \mathrm{P}n)$; and the inhibition of EBOV-GP into DC-SIGN/R^+ 293T cells with different glycan-GNPs, derived from the luciferase activity, $A = \mathrm{IC50}m/(\mathrm{IC50}m + \mathrm{Gm})$. Where [L] is the concentration of lectin, $K\mathrm{d}$ is the dissociation constant, n is the Hill coefficient, IC_{50} is the concentration of inhibitor required to achieve 50% luciferase activity, m is the inhibition coefficient and [G] is the concentration of glycan-GNP inhibitor. 165

Table 5.2 Summary of parameters for the Langmuir fits of the association of G5-Man to SLBs bearing DC-SIGN or DC-SIGNR of varying densities (Equation 5.2; where $-\Delta f\mathrm{max}$ is the maximum $-\Delta f$ value, $K\mathrm{d}$ is the dissociation constant and n is the cooperativity; n = 1 for DC-SIGN; $R^2 \geq 0.98$). 186

Table 5.3 Summary of the kinetic parameters for the dissociation of G5-Man from SLB-DC-SIGN/R of different densities fitted with Equation 5.3, where $k\mathrm{off}$ is the dissociation coefficient, a is the linear decay term and kon is the association coefficient derived by $K\mathrm{d} = k\mathrm{off}/k\mathrm{on}$. 190

Table 6.1 Summary of average FRET efficiencies and distances obtained by comparing the apparent FRET efficiencies derived from smFRET of HTDC-SIGN and HTDC-SIGNR against a donor only control. 219

Table 6.2 Average inter-dye and inter-domain distances calculated from reweighted MD snapshots. Reweighting was performed to fit snapshots to the experimental acceptor threshold of 7 (RSS = 1.52 and 1.2 for DC-SIGN and DC-SIGNR, respectively). 221

Table 7.1 Summary of inhibition data by G13-CRD_{100} and G13-CRD_{115} against infection of Vero76 cells by pseudo-typed SARS-CoV-2 bearing S proteins from various variants (Figure 7.13; fitted with Equation 7.2; n = 1 for all fits). A higher CRD valency and larger size of G13-CRD appears to be beneficial for boosting antiviral potency. 247

Table A.1 Summary of SARS-CoV-2 protein mutation sites in different variants (data obtained from Ref. [10]). 264

Table B.1 Summary of physical parameters derived from the hydrodynamic diameter ($D\mathrm{h}$) and glycan valency (N) of glycan-NPs. Where, $L\mathrm{h}$ is the hydrodynamic length, k is the average surface area of each ligand, θ is the average deflection angle, and X is the inter-glycan distance. 267

Table B.2 Summary of parameters for obtaining the Förster radius ($R0$) of each NP-Atto594 FRET pair, derived from Equation B.6-B.7 from their corresponding quantum yields (QY) and spectral overlap J-values. 268

Table C.1 Summary of parameters for the multimodal Gaussian fits of the hydrodynamic size distribution histograms for the binding of QR-DiMan with DC-SIGN/R. Where PQR is the protein to QR molar ratio; $D\mathrm{h}$ is the mean hydrodynamic size; FWHM is the full width at half maximum; and A is the integrated area. (All errors represent the standard deviation as observed by the fitting). 274

Table C.2 Summary of parameters for the multimodal Gaussian fits of the Dh distribution histograms for the dissociation of QR-DiMan·DC-SIGN/R with free D-mannose. Where [Man] is the concentration of D-mannose; $D\mathrm{h}$ is the mean hydrodynamic size; FWHM is the full width at half maximum; and A is the integrated area (All errors represent the standard deviation as observed by the fitting). 278

Table D.1 Summary of parameters for the unimodal Gaussian fits of the FRET efficiency distributions recorded for DC-SIGN-Atto488 donor only control sample with a fixed threshold for the sum of acceptor and donor intensities of 15 a.u.. FWHM is the full width at half maximum and A is the integrated area under the fit. *These values do not represent "real" FRET events but demonstrate the how much of the control is detected using these thresholds. 286

Table D.2 Summary of parameters for the multimodal Gaussian fits of the FRET efficiency distributions recorded for HTDC-SIGN and HTDC-SIGNR with a fixed threshold for the sum of acceptor and donor intensities of 15 a.u.. FWHM is the full width at half maximum and A is the integrated area under the fit..... 289

Table D.3 Distances between free C274 (*i.e.* Cα…Cα) and dye molecules (*i.e.* O8…O8) derived from the MD sampling reweighted towards the smFRET histograms obtained at each $\boldsymbol{TA}$ for each of the 4 side (s) and 2 diagonal (d) inter-CRD distances of DC-SIGN. 293

Table D.4 Distances between free C274 (*i.e.* Cα…Cα) and dye molecules (*i.e.* O8…O8) derived from the MD sampling reweighted towards the smFRET histograms obtained at each $\boldsymbol{TA}$ for each of the 4 side (s) and 2 diagonal (d) inter-CRD distances of DC-SIGN. 293

List of Figures

Figure 1.1 Schematic depicting three possible immunological routes involving specific protein-carbohydrate interactions on a cell membrane. 40

Figure 1.2 Schematic depicting the equilibrium tetravalent binding affinity ($\Delta Gtetra$) between a tetravalent lectin L with a tetravalent glycan G. 42

Figure 1.3 Annotated schematic of DC-SIGN. 49

Figure 1.4 DC-SIGN-mediated trans-infection of T cells by HIV-1, either *via* endocytosis and internalisation into multivesicular bodies which then release the virus at the infectious synapse in the presence of exosomes; or *via* DC-SIGN and CD4 binding, both associated with the infectious synapse. Additionally, transient DC-SIGN-ICAM-3 interactions facilitate DC-T cell interaction.[50, 51] 50

Figure 1.5 DC-SIGN-mediated pathogen recognition and immune regulation. Left: recognition of mannose-containing glycans, in conjunction with activation of other PRRs such as TLR4, enables transcription of genes encoding for IL-12 and IL-6 cytokines, which polarise T_H1 cells and T_H17 that in turn activate cellular immune response (involving macrophages, neurophils and NK cells). Right: recognition of fucose-containing glycans displaces the RAF-1 signalosome (associated with the mannose-associated pathway, thus downregulating IL-12, IL-6 and IL-23) and instead allows the assembly of the fucose signalosome which, in conjugation with other PRRs such as TLR4, enables transcription of anti-inflammatory cytokine IL-10 and chemokines CCL17 and CCL22, which attract T_H2 cells that in turn activate a humoral immune response (involving IgE^+ B cells, eosinophils and basophils), as well as IL-27 (*via* crosstalk with the interferon-α/β receptor (IFNAR)-associated pathway *via* the upregulation interferon-β (IFNβ)) which activates T_{FH} cell differentiation (involved in IgG^+ memory B cell activation). 52

Figure 1.6 Comparison of the amino acid sequences of DC-SIGN and DC-SIGNR (asterisks identify identical residues, blue bold shows the transmembrane domain, grey highlights show the 7½ neck repeat units and red bold shows the mannose binding motifs). 54

Figure 1.7 Schematic of glyconanoparticle design consisting of carbohydrate bearing glycan ligands polyvalently conjugated onto a nanoparticle core. 57

Figure 1.8 Elements of glyconanoparticle design such as the type of functional groups, the core material, the flexibility and density of glycans, and the size and shape of the glyconanoparticle can have a profound effect on its multivalent lectin binding behaviour. However, study of how many of these attributes effect the biophysical mechanisms of lectin binding remains in its infancy. 59

Figure 1.9 Conceptual schematic of different glycan-NPs binding to a trimeric lectin with varying valencies; demonstrating that by changing the size, shape, glycan density and/or deformability, a glycan-NP may be able to bridge a larger number of carbohydrate binding sites in order to improve its overall avidity (green: occupied binding site, pink: unoccupied binding site). 61

Figure 1.10 Schematic depicting the FRET process between an excited glycan-QD and dye-labelled DC-SIGN. 64

Figure 1.11 Schematic depicting the quenching process between excited dye-labelled DC-SIGN and glycan-GNP. 66

Figure 1.12 Multifunctional anchor-linker-glycan ligand design (top: lipoic acid-tetra(ethylene glycol)-mannose, or LA-EG_4-Man, with an yne-azide linkage; middle: lipoic acid-undeca(ethylene glycol)-mannose-α-1,2-mannose, or LA-EG_{11}-DiMan, with a cycoloctyne-azide linkage; bottom: thiol-hepta(ethylene glycol), or HS-EG_7-OH). 67

Figure 1.13 Hydrodynamic diameter (Dh) size distribution histograms of (A) QD-EG_{11}-DiMan in HEPES binding buffer; (B) QD-EG_{11}-DiMan incubated with DC-SIGN at a PQR of 12.5; (C) QD-EG_{11}-DiMan incubated with DC-SIGNR at a PQR of 12.5. (D, E, F) show the corresponding STEM images of the cryogenic prepared QD and QD-protein assembly samples. (Figure copied from Reference [104]). 72

Figure 1.14 Schematic of the binding modes of DC-SIGN (*i.e.* simultaneous binding of all CRDs to an individual NP) and DC-SIGNR (*i.e.* crosslinking with multiple NPs) upon binding to the polyvalent glycan-NPs highlighted in References [99, 104]. 73

Figure 1.15 Schematic depiction of a single molecule FRET approach to measure inter-binding site distances, demonstrating the passing of single protein molecules through the confocal volume of a focus excitation beam $h\nu$ex which excites the donor fluorophore (emission $h\nu$D) which transfers energy to the acceptor fluorophore (emission $h\nu$A) with a FRET efficiency of E (left). By then obtaining E for all inter-binding site distances, for a tetrameric lectin, values for the 4 side and 2 diagonal distances may be obtained. 78

Figure 2.1. Mass spectrometry data and chemical structures of in-house synthesized ligands for functionalize the QD, (A) DHLA-EG_{11}-DiMan and (B) DHLA-EG_3-OH.[a] 92

Figure 2.2. Deconvoluted HRMS spectra of the monomeric extracellular domains of (A) DC-SIGN (wild type; WT) and (B) DC-SIGNR (WT). 94

Figure 2.3. Volume population hydrodynamic size distribution histograms fitted with lognormal Gaussian distribution for extracellular segments of wide-type lectins: (A) DC-SIGN (WT) (B) DC-SIGNR (WT), at 40 nM in binding buffer (Dh values given as mean ± ½ FWHM). 94

Figure 2.4. Deconvoluted HRMS spectra of the monomeric extracellular domains of (A) DC-SIGN Q274C labelled with maleimide-Atto594 (DC-SIGN-Atto594), (B) DC-SIGNR R287C labelled with maleimide-Atto594 (DC-SIGNR-Atto594). 96

Figure 2.5 Diagram depicting example plots related to Figure 1.10, for: the overlap between the adsorption spectra for donor (QD) and acceptor (dye) fluorophores (AD and AA, respectively) and their emission spectra (with fluorescence intensities ID and IA, respectively) (left); the fluorescence spectra of a FRET pair which increases in binding with increasing concentration (middle); and the change of FRET ratio (IA/ID) with concentration, showing the maximum FRET ratio (Fmax) and the dissociation constant (Kd) (right). 99

Figure 2.6 QCM-D monitoring of the formation of an SLB *via* the adsorption and subsequent rupturing of SUVs upon a hydrophilic support (reconstructed from Reference [6]). 100

Figure 2.7 QCM-D monitoring of the functionalisation of an SLB surface with protein, and subsequent association and dissociation of analyte. 101

Figure 2.8 (A) schematic of proxyl SDSL at a cysteine residue; (B) energy level diagram demonstrating the hyperfine splitting of electron spin upon coupling with a N^{14} nucleus (C) cw-EPR spectrum corresponding to the hyperfine splitting of nitroxide unpaired electrons. 102

Figure 2.9 (A) The background corrected absorption spectra of varying concentrations of N3-EG2-DiMan samples after reaction with 5 % phenol and sulphuric acid (1:5 v/v ratio) and (B) a plot of sample absorbance at 490 nm against N_3-EG_2-DiMan concentration, fitted by a linear relationship, giving $y = (0.0189 \pm 0.0003) \cdot x$; $R^2 = 0.998$. 105

Figure 3.1 Schematic showing the different binding modes for DC-SIGN/R leading to different QD assemblies. The simultaneous tetravalent DC-SIGN·glycan-QD binding leads to individual QD particles, even at high protein:QD ratios (PQRs), whereas the crosslinking between DC-SIGNR and glycan-QDs results in a multiple QDs being assembled together in large-scale assemblies. 109

Figure 3.2 Volume populations hydrodynamic size distributions fitted with lognormal Gaussian distribution curves for (A) QD capped with DHLA-EG$_{11}$-DiMan (QD-DiMan) and (B) QD capped with DHLA-EG$_3$-OH (QD-OH). 111

Figure 3.3. Deconvoluted HRMS spectra of the monomeric extracellular domains of (A) DC-SIGN-C, (B) DC-SIGN-C-Atto594. 113

Figure 3.4. Volume population hydrodynamic size distribution histograms fitted with a lognormal Gaussian distribution for (A) DC-SIGN-C at 40 nM in binding buffer (Dh values given as mean ± ½FWHM). 113

Figure 3.5 Schematic of DHLA-EG$_{11}$-DiMan coated CdSe/CdS/ZnS core/shell/shell QDs (QD-DiMan; left) with the chemical structure of deprotonated DHLA-EG$_{11}$-DiMan (LA-EG$_{11}$-DiMan; right). 116

Figure 3.6 Schematic depicting the FRET process between an excited glycan-QD (donor) and lectin labelled with Atto594 (acceptor), where $h\nu$ex is the excitation energy, $h\nu$D is the donor emission energy and $h\nu$A is the acceptor emission energy. 117

Figure 3.7 Background-corrected fluorescence spectra of different concentrations of a mixture of QD-DiMan with Atto594 labelled lectins at 20 ^{o}C for (A) DC-SIGN (PQR 1:1), (B) DC-SIGNR (PQR 10:1), and (C) DC-SIGN-C at a (PQR 1:1); and the corresponding apparent FRET ratio−protein concentration relationships at three different temperatures, fitted by Equation 3.1, for (D) 1:1 mixed QD and DC-SIGN, (E) 1:10 mixed QD and DC-SIGNR, and (F) 1:1 mixed QD and DC-SIGN-C. Error bars represent the standard deviations (SDs) of triplicate experiments at each concentration. 119

Figure 3.8 Plot of FRET ratio *vs* protein to QD molar ratio (PQR) obtained by titrating varying amounts of DC-SIGN (blue) or DC-SIGN-C (green) to a fixed concentration of QD-OH (20 nM). Neither plots showed significant FRET signals, confirming almost no non-specific interactions between the QD-OH control and Atto594 labeled DC-SIGN or DC-SIGN-C..... 121

Figure 3.9 (A) Van 't Hoff analyses of the ln$K$d *vs* $1/T$ relationships for QD-DiMan binding with DC-SIGN (blue), DC-SIGN-C (green), and DCSIGNR (red). (B) Comparison of the standard (T = 298 K) enthalpy (blue), entropy (red), and Gibbs free energy (pink) changes of QD-DiMan binding with DC-SIGN, DC-SIGN-C and DC-SIGNR. 123

Figure 3.10 Representative cryo-preserved TEM (contrast inverted HAADF STEM) images of QD-DiMan after binding to (A) DC-SIGN (PQR, = 1:1), (B) DC-SIGN-C (PQR = 1:1), or (C) DC-SIGNR (PQR = 10:1). (D) Quantitative analysis of the number of QDs per cluster for assemblies of QD-DiMan bound with DC-SIGN, DC-SIGN-C, or DC-SIGNR. 126

Figure 3.11 Raw kinetic profile of the fluorescence intensity at 626 nm (red) and 550 nm (green) for the association of (A) QD-DiMan with labelled DC-SIGN, (B) the dissociation of QD-DiMan and labelled DC-SIGN in the presence of excess mannose (Man), and (C) a control containing only labelled DC-SIGN. Kinetic profiles of the FRET ratio ($F = I\mathrm{A}/I\mathrm{D}$) measured for the association of a 1:1 ratio of QD-DiMan with labelled protein (dark colour) and dissociation of bound 1:1 QD-DiMan·lectin complex in excess Man (light colour) for (D) DC-SIGN, (E) DC-SIGN-C, and (F) DC-SIGNR. 128

Figure 4.1 FRET signal with increasing protein to QD molar ratio (PQR) for the titration of DC-SIGN (black circle), DC-SIGNR (red circle), monomeric DC-SIGN carbohydrate-recognition domain (CRD) (blue upward triangle) or monomeric DC-SIGNR CRD (green downward triangle), against a fixed concentration of QDs capped with DHLA-EG$_{11}$-DiMan (QD-DiMan). (Taken from reference [20]). 138

Figure 4.2 (A) Volume population hydrodynamic size distribution histogram, fitted with the Gaussian function, of QR$_{560}$ capped with LA-EG$_{11}$-DiMan (QR-DiMan) measured by DLS. (B) Histograms of the QR core diameter and length distributions obtained from STEM images of QR-DiMan. Mean hydrodynamic diameter and core diameter and length are denoted as $D\mathrm{h}$, $D\mathrm{core}$ and $L\mathrm{core}$, respectively (recorded as mean ± ½ FWHM; $R^2 \geq 0.98$). 140

Figure 4.3 Schematic of DHLA-EG$_{11}$-DiMan capped CdSe/CdS core/shell quantum rods (QR-DiMan) with the chemical structure of deprotonated DHLA-EG$_{11}$-DiMan (LA-EG$_{11}$-DiMan; below)..... 142

Figure 4.4 Schematic depicting the FRET process between an excited glycan-QR (donor) and lectin labelled with Atto594 (acceptor), where $h\nu\mathrm{ex}$ is the excitation energy, $h\nu\mathrm{D}$ is the donor emission energy and $h\nu\mathrm{A}$ is the acceptor emission energy. 144

Figure 4.5 Direct excitation background corrected fluorescence spectra over different concentrations of QR-DiMan binding with Atto594 labelled (A) DC-SIGN and (B) DC-SIGNR. The corresponding FRET ratio–concentration relationships, fitted with the Hill equation (Equation 4.1 for QR-DiMan binding with (C) DC-SIGN or (D) DC-SIGNR. (PQR = 1:1; Fitting parameters are summarised in Table 4.1). 145

Figure 4.6 (A) Direct excitation background corrected fluorescence spectra over different concentrations of QR-DiMan binding with Atto594 labelled DC-SIGNR at a PQR of 10:1. (B) The corresponding FRET ratio–concentration relationships, fitted with the Hill equation (Equation 4.1). (Fitting parameters are summarised in Table 4.1). 147

Figure 4.7 Direct excitation background corrected fluorescence spectra corresponding to titration of DC-SIGN (A) or DC-SIGNR (B) against a fixed concentration of QR-DiMan (10 nM). (C) Plots of the corresponding FRET ratio-PQR relationships fitted with linear fits of the initial PQR data points (PQR ≤ 4 or 12 for DC-SIGN and DC-SIGNR, respectively). The fits give intercept = 0.018 ± 0.002 and 0.0194 ± 0.0006, slope = 0.24 ± 0.01 and 0.0234 ± 0.0002 for DC-SIGN and DC-SIGNR, respectively. (R^2 ≥ 0.99 for all fits) 149

Figure 4.8 Volume population hydrodynamic size distribution histograms fitted with uni- or bi-modal Gaussian functions for QR-DiMan (10 nM) after binding with DC-SIGN at PQRs of (A) 2:1, (B) 4:1, (C) 10:1 and (D) 20:1 or binding with DC-SIGNR at PQRs of (E) 2:1, (F) 4:1, (G) 10:1, and (H) 20:1. (I) The corresponding average $D\mathrm{h}$-PQR relationship (average $D\mathrm{h} = D\mathrm{h},1 \times A1\% + D\mathrm{h},2 \times A2\%,$ where $A1\%$ and $A2\%$ are the percentage area of the Gaussian fits; filled circles: single distribution; half-filled circles: two distributions) fitted with an offset Hill function ($D\mathrm{h},\mathrm{PQR} = D\mathrm{h},0 + D\mathrm{h},\infty - D\mathrm{h},0/(1 + (\mathrm{PQR50}/\mathrm{PQR})n)$ where $D\mathrm{h},0$ was fixed to the $D\mathrm{h}$ of QR-DiMan, $D\mathrm{h},\infty$ = 62 ± 6 and 150 ± 10 nm, $\mathrm{PQR50}$ (PQR at 50% of Dh,∞) = 2.7 ± 0.8 and 1.3 ± 1, and n = 1.3 ± 0.4 and 1 ± 1, for DC-SIGN and DC-SIGNR, respectively; R^2 ≥ 0.99). (DC-SIGN: blue; DC-SIGNR: red). 150

Figure 4.9 Volume population hydrodynamic size distribution histograms fitted with uni- or bi-modal Gaussian functions for a pre-incubated mixture of QR-DiMan (10 nM) with DC-SIGN (40 nM) after addition of free mannose (Man) at concentrations of (I) 0.1, (J) 2, (K) 4 or (L) 100 mM; and a pre-incubated mixture of QR-DiMan (40 nM) with DC-SIGNR (40 nM) after addition of Man at concentrations of (M) 0.1, (N) 2, (O) 4 and (P) 100 mM. (S) Plot of the corresponding average $D\mathrm{h}$ against [Man] (average $D\mathrm{h} = D\mathrm{h},1 \times A1\% + D\mathrm{h},2 \times A2\%,$ where $A1\%$ and $A2\%$ are the percentage area of the Gaussian fits; filled circles: single distribution; half-filled circles: two distributions). (DC-SIGN: blue; DC-SIGNR: red). 151

Figure 4.10 Direct excitation background corrected fluorescence spectra corresponding to increasing concentrations of free mannose to a pre-incubated 4:1 PQR mixture of 10 nM QR-DiMan with (A) DC-SIGN or (B) DC-SIGNR. (C) A plot of the relationship between FRET ratio against mannose concentration for the fluorescence spectra recorded in A and B, fitted with an offset Hill function (Equation 4.2; where $F0$ = 1.049 ± 0.007 and 0.253 ± 0.001, $F\infty$ = 0.071 ± 0.003 and 0.072 ± 0.003, Ki ([Man] at 50% inhibition) = 8.0 ± 0.1 and 4.2 ± 0.2 mM, and n = 1.78 ± 0.05 and 1.21 ± 0.08, for DC-SIGN (blue) and DC-SIGNR (red), respectively; R^2 > 0.999 for both fits). 152

Figure 4.11 S/TEM images of cryo-prepared QR-DiMan (10 nM) after complexation with 4 eq of (A) DC-SIGN, (B) DC-SIGNR, and (C) DC-SIGNR in the presence of Man (2 mM). 153

Figure 4.12 Histograms of the perimeter-perimeter nearest neighbour distance (ppNND), fitted with uni- or multi- modal Gaussian fits, for the S/TEM images of cryo-prepared QR-DiMan incubated with (A) DC-SIGN, (B) DC-SIGNR or (C) pre-incubated sample of QR-DiMan·DC-SIGNR with 2 mM Man. Histograms of the centre-centre nearest neighbour distances (ccNNDs, for 2.9 < ppNND < 7.7 nm), fitted with trimodal Gaussian fits (where applicable), for QR-DiMan incubated with (D) DC-SIGN, (E) DC-SIGNR or (F) pre-incubated QR-DiMan·DC-SIGNR with 2 mM Man. Statistical analysis of the (G) numbers of QRs per cluster (i.e. the number of interconnected QRs with a ppNND < 7.7 nm) and (H) QR nearest neighbour orientations (individual: ppNND > 7.7 nm; parallel: the area of the first ccNND peak (or ccNND <10 nm for DC-SIGN); non-parallel: the sum of the second and third ccNND peak areas (or ccNND >10 nm for DC-SIGN); non-specific: ppNND < 2.9 nm). 155

Figure 4.13 Schematics depicting (1) the simultaneous binding of DC-SIGN to all sections of both QR-DiMan and QD-DiMan; (2) DC-SIGNR crosslinking with spherical QD-DiMan or with the spherical end sections of QR-DiMan into parallel stacks; and (3) the simultaneous binding of DC-SIGNR to the cylindrical section of QR-DiMan, only..... 158

Figure 4.14 Schematic depicting the difference in separation distance between the binding contact area of DC-SIGNR (red) with either the spherical end (left) or cylindrical section (right) of the QR surface. Where d is the separation distance between the protein contact area and the QR-DiMan surface, l is the maximum compression length of the surface glycan ligands, p is the Ca-Ca distance between two CRDs at the furthest separation distance from the QR surface, Dh is the estimated QR-DiMan hydrodynamic diameter, and H is the height of the QR cylindrical section. (Not to scale). 159

Figure 5.1 Hydrodynamic size distribution of G5-Ct, obtained by DLS. Fitted with a lognormal distribution curve (Dh = mean ± ½FWHM; R^2 = 0.998). 169

Figure 5.2 (A) Core diameter distribution of G5-Man, obtained by cryoTEM; (B) hydrodynamic size distribution of G5-Man, obtained by DLS. Both distributions fitted with lognormal distribution curves (where D_{core} or D_h = mean ± ½FWHM; R^2 ≥ 0.99). 170

Figure 5.3 Deconvoluted HRMS of (A) His_6-DC-SIGN monomer; (B) His_6-DC-SIGNR monomer; (C) NSL-DC-SIGNR monomer. 171

Figure 5.4 Hydrodynamic size distributions obtained by DLS for (A) His_{24}-DC-SIGN and (B) His_{24}-DC-SIGNR. 171

Figure 5.5 Schematic depicting LA-EG_2-Man capped G5 (G5-Man) bound to His_6-tagged DC-SIGN/R (His_{24}-DC-SIGN/R)-tethered *via* Ni^{2+} coordination to a tris-NTA DODA-containing DOPC supported lipid bilayer (SLB) on a quartz sensor support. 175

Figure 5.6 QCM-D measurements of the Δf and ΔD time profiles for the formation of SLBs from DOPC SUVs containing tris-NTA DODA in proportions of (A) 0.2%, (B) 0.4%, (C) 1% or (D) 2%. 1. Demonstrates the equilibration of the sensor under buffer; 2. demonstrates a typical frequency and dissipation shift with absorption of the SUVs to the quartz surface; and 3. Demonstrates a slight recovery of frequency and dissipation as SUVs rupture to form the SLB surface. Schematics of these stages depicted below. (Buffer: SUV buffer). 177

Figure 5.7 QCM-D measurements of the Δf and ΔD time profiles for the association and imidazole-driven dissociation of (A) His_{24}-DC-SIGN and (B) His_{24}-DC-SIGNR, onto an SLB bearing 2% tris-NTA; and (C) the interaction of His_{24}-DC-SIGN and His_{24}-DC-SIGNR with SLBs bearing no tris-NTA. 1. Demonstrates the equilibration of the SLB in binding buffer; 2. demonstrates the binding of protein over time; and 3. demonstrates the release of protein *via* displacement of Ni^{2+} by imidazole. (Buffer A: SUV buffer; Buffer B: binding buffer). 179

Figure 5.8 Plot of the dependency of Δf for the association of protein to SLB surface on the percentage of tris-NTA containing lipid for DC-SIGN (blue) and DC-SIGNR (red). Where Δfeq is taken as the difference in average equilibrium frequency 2 mins before addition of protein and that after washing with buffer after addition of protein. Linear fits taken between 0-1% tris-NTA; DC-SIGN: Y = (85 ± 4)X; DC-SIGNR: Y = (84 ± 7)X; R^2 ≥ 0.99. 180

Figure 5.9 QCM-D measurements of the Δf and ΔD time profiles for the association of G5-Man followed by its partial dissociation in binding buffer and complete dissociation in 50 mg mL^{-1} free mannose for SLBs bearing (A) 2% DC-SIGN and (B) 2% DC-SIGNR; and the interaction of G5-Man and Man with SLBs bearing (C) 2% GFP and (D) DOPC only. 1. Demonstrates the equilibration of the SLB-lectin surface in binding buffer; 2. demonstrates the binding of G5-Man with time; and 3. demonstrates the partial dissociation of G5-Man after washing with binding buffer. (Buffer: binding buffer; Man: mannose)..... 182

Figure 5.10 Plots demonstrating the effect of flowrate on the rate of change of frequency for G5-Man binding to SLBs bearing 0.4% DC-SIGN. Plots demonstrate (A) the time profiles of the Δf for flowrates of 5 and 20 µL min^{-1} (dark and light blue, respectively), fitted with linear fits at the turning point of the decay (providing $\mathrm{d}\Delta f/\mathrm{d}t$ of -2.24 ± 0.02 and -5.63 ± 0.08 Hz min^{-1} for 5 and 20 µL min^{-1}, respectively; $R^2 \geq 0.995$); and (C) the turning point $\mathrm{d}\Delta f/\mathrm{d}t$ values against the cubic root of the flowrate, $Q^{1/3}$, fitted with a linear fit, where $Y = (1.4 \pm 0.2)X$; $R^2 = 0.963$. 183

Figure 5.11 Plots of the Δf and ΔD for the titration of G5-Man of concentrations 0.5, 1, 2, 4, 10, 20, 60 nM and subsequent dissociation in buffer with SLBs bearing (A) 0.2% DC-SIGN, (B) 0.4% DC-SIGN, (C) 1% DC-SIGN and (D) 2% DC-SIGN. (Flowrates: 40 µL min^{-1} for 0.5-4 nM; 20 µL min^{-1} for 10-60 nM). 184

Figure 5.12 Plots of the Δf and ΔD for the titration of G5-Man of concentrations 1, 3, 10, 30, 100, 300, 600 nM and subsequent dissociation in buffer with SLBs bearing (A) 0.2% DC-SIGNR, (B) 0.4% DC-SIGN, (C) 1% DC-SIGN and (D) 2% DC-SIGN. (Flowrates: 40 µL min^{-1} for 1-3 nM; 20 µL min^{-1} for 10-30 nM; 10 µL min^{-1} for 100-300 nM; 5 µL min^{-1} for 600 nM). 185

Figure 5.13 Plots of the average $-\Delta f$ signal at 80% ± 40 s of the incubation time for each concentration of G5-Man for the titration of G5-Man against SLBs bearing (A) DC-SIGN and (B) DC-SIGNR; data fitted with Hill equations (Equation 5.2; where $n = 1$ for DC-SIGN and n is floating for DC-SIGNR; $R^2 \geq 0.98$..... 185

Figure 5.14 Schematic depicting the predicted simultaneous binding mode of DC-SIGN (left) and crosslinking binding mode of DC-SGINR (right) with G5-Man upon SLB surfaces. 188

Figure 5.15 Plots of the $-\Delta f$ time profiles for the dissociation of G5-Man from SLB surfaces with varying densities of (A) DC-SIGN and (B) DC-SIGNR. Decays fitted with Equation 5.3; $R^2 \geq 0.98$. .. 190

Figure 5.16 Fluorescence confocal microscopy images of SLBs bearing 2% NSL-DC-SIGNR (A) before and (B) 30 min after addition of G5-Man; and 0.4% NSL-DC-SIGNR (C) before and (D) 30 min after addition of G5-Man. (Scale bar: 50 µm). 193

Figure 5.17 Cryogenic transition electron microscopy (cryoTEM) images of G5-Man on SLBs bearing 2% DC-SIGNR. (A) demonstrates the SLB quality at the same magnification as those in Figure 5.16; (B) demonstrates a typical G5-Man distribution at the edge of a well-formed SLB section; (C) demonstrates a typical G5-Man distribution in the middle of a well-formed SLB; (D) demonstrates the typical secondary cluster formation of G5-Man. 195

Figure 5.18 Cryogenic energy dispersive x-ray (EDX) spectroscopy images of the three GNP coverages observed by cryoTEM of G5-Man bound to 2% DC-SIGNR, namely, regions of no GNP coverage, typical GNP coverage and clustered GNP regions. Images demonstrate, from top-left to bottom-right, the intensities of density (by high-angle annular dark field (HAADF) imaging), and EDX intensities of carbon (C), nitrogen (N), oxygen (O), silicon (Si), nickle (Ni), copper (Cu) and gold (Au). 196

Figure 5.19 Centre to centre inter-GNP nearest neighbour distance (NND) distributions for the cryoTEM images of G5-Man bound to SLBs bearing 2% DC-SIGNR, where (A) is the typical distribution of G5-Man upon well-formed SLB surface (as from Figure 5.17C) and (B) is the G5-Man distribution at regions of secondary clustering (as from Figure 5.17D). (NND = mean ± ½FWHM; $R^2 \geq 0.994$). 197

Figure 6.1. Deconvoluted HRMS spectra of the monomeric extracellular domains of (A) DC-SIGN-Atto488, (B) DC-SIGN-Atto643, (C) DC-SIGNR-Atto488, (D) DC-SIGN-Atto643. 205

Figure 6.2 Deconvoluted HRMS of the monomeric extracellular domains of (A) DC-SIGN with a galactose binding motif (or GBDC-SIGN) and (B) GBDC-SIGNR. 206

Figure 6.3 UV-vis spectra of heterotetrameric (HT)DC-SIGN consisting of a 1:1:2 monomer unit ratio of DC-SIGN-Atto488:DC-SIGN-Atto643:GBDC-SIGN (top) and HTDC-SIGNR consisting of a 1:1:2 monomer unit ratio of DC-SIGNR-Atto488:DC-SIGNR-Atto643:GBDC-SIGNR (bottom). 208

Figure 6.4 Deconvoluted HRMS of the monomer units corresponding to (A) heterotetrameric (HT)DC-SIGN consisting of a 1:1:2 monomer unit ratio of DC-SIGN-Atto488:DC-SIGN-Atto643:GBDC-SIGN and (B) HTDC-SIGNR consisting of a 1:1:2 monomer unit ratio of DC-SIGNR-Atto488:DC-SIGNR-Atto643:GBDC-SIGNR..... 209

Figure 6.5 Deconvoluted HRMS of the monomer units corresponding to (A) DC-SIGN Q274C labelled with spin label MTSSL (DC-SIGN-SL) (B) DC-SIGNR R287C labelled with spin label MTSSL (DC-SIGNR-SL). 210

Figure 6.6 A schematic illustrating the double purification strategy used to label two out of four of the CRDs of DC-SIGN/R. (i) A 1:1:2 ratio of DC-SIGN-Atto488 : DC-SIGN-Atto643 : GBDC-SIGN tetramers were unfolded into monomer units the presence of guanidine·HCl (6 M) and (ii) Renatured by dialysis to remove guanidine, allowing for the formation of a distribution of different heterotetramers. (iii) A galactose-sepharose affinity column was then used to remove tetramers with one or less galactose-binding CRDs due to their poor affinity. (iv) This was repeated with a mannose-sepharose affinity column to remove tetramers containing one or less mannose-binding CRDs. This yielded heterotetramers containing 2 labelled mannose binding CRDs and 2 galactose binding CRDs, labelled on either adjacent or diagonal CRDs, where only those containing both donor (Atto488) and acceptor (Atto643) fluorophores can participate in FRET. 214

Figure 6.7 A schematic representing the two FRET modes that can be determined by smFRET for DC-SIGN with two out of four CRDs labelled with an ATTO488···ATTO643 FRET pair and excited with a 488 nm excitation beam: that between adjacent CRDs (left) and diagonal CRDs (middle). By extracting the resulting FRET efficiencies of the different FRET pair configurations, distances for the 4 side and 2 diagonal inter-CRD arrangements can be obtained *via* Equation 6.2 (right) 215

Figure 6.8 Background corrected fluorescence spectra of (A) HTDC-SIGN (solid line) and a control sample containing a 1:1:2 mixture of DC-SIGN-Atto488:DC-SIGN-Atto643:GBDC-SIGN in their homotetrameric forms (dashed line); and (B) HTDC-SIGNR (solid line) and a control sample containing a 1:1:2 mixture of DC-SIGNR-Atto488:DC-SIGNR-Atto643:GBDC-SIGNR in their homotetrameric forms (dashed line). 216

Figure 6.9 histograms of the population of FRET efficiency measurements detected by smFRET using a acceptor and donor sum threshold, $Tsum$, of 15 a.u. and an acceptor threshold, TA, of 7 a.u. for (A) HTDC-SIGN and (B) HTDC-SIGNR. 218

Figure 6.10 A comparative analysis of mean apparent FRET efficiencies measured by smFRET for HTDC-SIGN (left) and HTDC-SIGNR (right) compared to a donor only control sample (yellow) over increasing acceptor threshold using a constant threshold for the sum of the acceptor and donor intensities of 15 a.u.. The log of the integrated area of each distribution ($\log A$) is represented via the size of the data point. Six sets of data points corresponding to the six possible inter-dye distances were identified (highlighted by different colours). Any data points which overlapped with donor only data points or were likely corresponded to an average of multiple data sets were not assigned to any set (grey). 219

Figure 6.11 Distribution of Cys:Cα…Cys:Cα *vs* dye:O8…dye:O8 distances from MD of DC-SIGN/R. Black crosses/dashed lines denote experimental values recorded in Table 6.1. 220

Figure 6.12 Example MD simulation model reweighted towards the smFRET histogram obtained using $Tsum$ = 15 and TA = 6, for (A) DC-SIGN and (B) DC-SIGNR. 221

Figure 6.13 Comparison of the inter-dye distances obtained by MD sampling, reweighted towards the smFRET histogram obtained using $Tsum$ = 15 and TA = 7, against the inter-dye distances obtained by smFRET compared across a range of TA thresholds. (black diagonal demonstrates $x = y$, *i.e.* perfect correlation). 222

Figure 6.14 Plot of the hyperfine triplet splitting patterns corresponding to the first derivative signals of the interaction of nitroxide spin labelled DC-SIGN (blue) and DC-SIGNR (red) with a fixed wavelength over the magnetic field strength (B) using modulating magnetic field obtained by cw-EPR. $\Delta H0$ denotes the central linewidth, calculated as the difference in B between the signal maxima and minima for each protein. 223

Figure 7.1 Deconvoluted HRMS spectra of a monomeric unit of (A) DC-SIGN's extracellular domain and (B) a DC-SIGN carbohydrate-recognition domain (CRD). 229

Figure 7.2 (A) Number population core diameter distribution histogram and (B) volume population hydrodynamic size distribution histogram, fitted with a lognormal Gaussian fits, for citrate capped 13nm GNPs (G13-Ct; $R^2 \geq 0.96$). 230

Figure 7.3 Volume population hydrodynamic size distributions, fitted with lognormal Gaussian fits, of 13nm GNPs (A) partially PEGylated with HS-EG_7-OH (ppG13-OH) and (B) 13 GNPs fully PEGylated with HS-EG_7-OH (G13-OH). ($R^2 > 0.992$). 231

Figure 7.4 Deconvoluted HRMS spectra of the monomeric units of LA-EG_{11}-Tz-TFP N-terminal labelled (A) ECD and (C) CRD, and (E) LA-EG_{11}-TFP N-terminal labelled CRD, prior to ppG13-OH conjugation; and that of the filtrate collected post-conjugation of ppG13-OH with LA-EG_{11}-Tz-TFP N-terminal labelled (B) ECD and (D) CRD, and (F) LA-EG_{11}-TFP N-terminal labelled CRD. 232

Figure 7.5 Volume population hydrodynamic size distributions, fitted with lognormal Gaussian fits, for 13nm GNPs conjugated with (A) 100 eq of LA-EG_{11}-CRD (G13-CRD_{100}) and (B) 115 eq of LA-EG_{11}-CRD (G13-CRD_{115}). 233

Figure 7.6 Schematic depicting G13-Glycan: 13nm GNPs coated with LA-EG_4-Glycans of either DiMan or Gal. 233

Figure 7.7 Schematic route to prepare an polyvalent lectin-GNPs: (i) Lectin N-terminal α-amine is selectively labelled with an LA-EG_{11} linker by reacting with either LA-EG_{11}-TFP or LA-EG_{11}-Tz-TFP at pH 6.2 to form LA-lectin or LA-Tz-lectin, respectively (where lectin = DC-SIGN CRD or ECD); (ii) Citrate stabilised 13 nm gold nanoparticle (G13) is partially PEGylated with HS-EG_7-OH to form ppG13-OH; (iii) ppG13-OH is then conjugated with LA-lectin *via* self-assembly to form G13-lectin (exemplified with G13-CRD). 236

Figure 7.8 (A) UV-vis spectra of G13 capped with HS-EG_7-OH at varying ligand:GNP molar ratios in binding buffer; (B) the degree of redshift as a function of [ligand]/[GNP] measured by the absorbance at 650 nm (A_{650}), where high absorbance (*i.e.* increased redshift) demonstrates instability in binding buffer; (C) 1.5% agarose gel electrophoresis of G13 capped with HS-EG_7-OH at varying ligand:GNP molar ratios, demonstrating lower HS-EG_7-OH surface coverage (higher citrate content) at lower ratios, signified by reduced gel mobility. 238

Figure 7.9 1.5% agarose gel electrophoresis demonstrating that (A) ppG13-OH, G13-Tz-CRD_{100} and G13-Tz-ECD_{100} have successively slower gel mobility, likely due to an increase in molecular weight with addition of linker-labelled CRD and ECD; and (B) similarly, G13-CRD_{100} and G13-CRD_{115} also display slower gel mobility over ppG13-OH. Additionally, fully PEGylated G13-OH has a slower gel mobility than the partially PEGylated ppG13-OH, indicating the higher ligand surface coverage of the former. (C) Demonstrates the volume population hydrodynamic size distribution histograms, fitted by log-normal Gaussian fits, of ppG13-OH, G13-Tz-CRD_{100} and G13-Tz-ECD_{100}, yielding Dh values of ~17, ~31 and ~144 nm..... 239

Figure 7.10 Schematic depicting the fluorescence quenching of glycan-capped QDs (*e.g.* QD-DiMan) binding with lectin capped GNPs (*e.g.* G13-CRD) with an equilibrium binding dissociation constant, Kd. 240

Figure 7.11 Fluorescence quenching efficiency (QE) assays demonstrating the glycan binding specificity of G13-lectin by: (A) the fluorescence quenching of QD-DiMan (2 nM) by G13-Tz-CRD_{100} (1 nM; QE >80%) and G13-Tz-ECD_{100} (1 nM; QE ~20%), and only marginally by the G13-OH control (QE ~7.5%); (B) marginal fluorescence quenching of QD-Gal (2 nM) negative control by G13-Tz-CRD_{100} (1 nM; QE ~9.5%), and not by G13-OH control. Fluorescence spectra of varying concentrations of QD-DiMan in the absence (solid line) or presence (dash line) of 1:1 mixed (C) G13-Tz-CRD_{100} or (E) G13-CRD_{100}, and the corresponding QD-DiMan QE-concentration relationship, fitted with Equation 7.1, for (D) G13-Tz-CRD_{100} and (F) G13-CRD_{100}..... 242

Figure 7.12 Dose-dependent inhibition of SARS-CoV-2 S-protein-driven entry into Vero76 cells. VSV particles bearing SARS-CoV-2 S protein (Wuhan wildtype, Hu-1) were pre-incubated with (A) G13-Tz-CRD$_{100}$, G13-Tz-ECD$_{100}$ or G13-OH control, or (C) tetrameric DC-SIGN ECD, before addition to target cells. As controls, target cells were pre-incubated with Camostat and Chloroquin before addition of pseudotyped particles. Entry efficiency was determined by quantifying luciferase activity in cell lysates. The orange bars in panel B represent the background luminescence measured in the absence of viral particles. The results of a representative experiment performed in quadruplicates was shown and were confirmed in two separate experiments. Errors bars indicate standard errors of one quadruplicate experiment. (B) The corresponding normalised infection curves, fitted by a modified inhibition model (Equation 7.2), yielding an effective inhibition concentration (EC$_{50}$) of 0.25 ± 0.04 nM and an inhibition coefficient (n) of 0.57 ± 0.06 for G13-Tz-ECD$_{100}$ (red) and EC$_{50}$ = 0.19 ± 0.02 nM and n = 1 for G13-Tz-CRD$_{100}$ (black) (R^2 = 0.97). **244**

Figure 7.13 Normalised dose-dependent luciferase activity data, fitted by the inhibition model (Equation 7.2; n = 1 for all fits), for the inhibition of SARS-CoV-2 S protein-mediated cell entry of Vero76 by G13-CRD$_{100}$ (black) and G13-CRD$_{115}$ (red) against VSV particles bearing the S proteins from SARS-CoV-2 variants: (A) Wuhan Hu-1 (WT), (B) B.1, (C) Delta (B.1.617.2), (D) Omicron (BA.1). Parameters summarised in Table 7.1 (R^2 > 0.92). **246**

Figure 7.14 G13-CRD$_{100}$ and Sotrovimab inhibition of authentic SARS-CoV-2 infections against the (A) Wuhan wild-type (Hu-1) and (B) Omicron BA.1 variants. The absence of bar chart data at high concentrations indicate no observed infections (*i.e.* number of infectious particles surpasses the limit of detection (LOD); where LOD is indicated by the lowest limit of the Y-scale). "No BG" indicates the native infection level in the absence of any inhibitors. **248**

Figure A.1 SARS-CoV-2 S protein glycosylation sites (redrawn based on the data reported from Ref. [9]). **264**

Figure B.1 Normalised intensity spectra of glycan-NP fluorescence absorbance (black) and emission (red) and DC-SIGN-Atto594 absorbance (blue) and emission (green) for (A) QD-EG$_{11}$-DiMan and (C) QR-EG$_{11}$-DiMan; and the spectral overlap as a function of wavelength, $J(\lambda)$, between the glycan-NP emission and DC-SIGN-Atto594 absorbance (Equation B.6) for (B) QD-EG$_{11}$-DiMan and (D) QR-EG$_{11}$-DiMan. **267**

Figure B.2 Plots of the fluorescence intensity *vs.* concentration relationships for (A) QD-DiMan (λex = 450 nm, λem = 559 nm; R^2 = 0.997) and (B) DC-SIGN (Atto594 labelled, λem = 450 nm, λem = 626 nm; R^2 = 0.999) fitted by a linear function..... 269

Figure B.3 (A) Background corrected fluorescence spectra and (B) FRET ratio plot of the dissociation of a 1:1 QD-DiMan·DC-SIGN-C complex with 1×10^6 eq of D-mannose (Man) over time; and (C) background corrected fluorescence spectra and (D) FRET ratio plot of the dissociation and of a 1:1 QD-DiMan·DC-SIGN-C complex with increasing concentrations of Man. 270

Figure C.1 Background corrected fluorescence spectra of the different concentrations of QD-DiMan after mixing with labelled DC-SIGN (PQR = 1:1) at (A) 20 °C, (B) 25 °C and (C) 30 °C; DC-SIGNR (PQR = 10:1) at (D) 20 °C, (E) 25 °C and (F) 30 °C; and DC-SIGN-C (PQR = 1:1) at (G) 20 °C, (H) 25 °C and (I) 30 °C. 271

Figure C.2 Fluorescence intensity time profiles measuring the emission of 550 nm (lighter color) and 628 nm (darker color) at λex = 450 nm for the association of QD-DiMan with (A) DC-SIGN, (B) DC-SIGN-C and (C) DC-SIGNR, and the dissociation of QD-DiMan in the presence of an excess of free mannose with (D) DC-SIGN, (E) DC-SIGN-C and (F) DC-SIGNR..... 272

Figure C.3 Volume population hydrodynamic size distribution histograms fitted with Gaussian fits for the binding of QR-DiMan (10 nM) with DC-SIGN at varying protein:QR molar ratios (PQR) of (A) 0:1, (B) 2:1, (C) 4:1, (D) 6:1, (E) 8:1, (F) 10:1, (G) 15:1 and (H) 20:1. 272

Figure C.4 Volume population hydrodynamic size distribution histograms fitted with uni- or bi- modal Gaussian fits for the binding of QR-DiMan (10 nM) with DC-SIGNR at protein:QR molar ratios (PQR) of (A) 0:1, (B) 2:1, (C) 4:1, (D) 6:1, (E) 8:1, (F) 10:1, (G) 15:1 and (H) 20:1..... 273

Figure C.5 Plots the change in the hydrodynamic size distribution, obtained by DLS, for the change in PQR for the binding of QR-DiMan with DC-SIGN (blue) or DC-SIGNR (red), measured by (A) the mean D_h ± ½FWHM, which demonstrates the similarity in the change in size of the smaller QR-DiMan·DC-SIGN/R complexes with increasing amount of protein, and (B) the percentage population of clustered species ($100\% \times A2/(A1 + A2)$, where Ai is the integrated area of the i^{th} mode of the multimodal Gaussian distribution fit), which demonstrates the increase in clustered species with protein concentration. 273

Figure C.6 Volume population hydrodynamic size distribution histograms, fitted with Gaussian fits for pre-incubated QR-DiMan (10 nM) with DC-SIGN (PQR 4:1) after addition of (A) 0.01 mM, (B) 0.1 mM (C) 1 mM, (D) 2 mM, (E) 4 mM, (F) 6 mM, (G) 10 mM and (H) 100 mM of free mannose (Man). 275

Figure C.7 Volume population hydrodynamic size distribution histograms, fitted with uni- or bi- modal Gaussian fits, for pre incubated QR-DiMan (10 nM) with DC-SIGNR (PQR 4:1) after addition of (A) 0.01 mM, (B) 0.1 mM (C) 1 mM, (D) 2 mM, (E) 4 mM, (F) 6 mM, (G) 10 mM and (H) 100 mM of free mannose (Man). 276

Figure C.8 Plots of the change in the hydrodynamic size distribution, obtained by DLS, for the dissociation of a PQR 4:1 QR-DiMan·DC-SIGN (blue) or QR-DiMan·DC-SIGNR (red) complex with varying mannose concentration ([Man]), measured by (A) the Dh ± ½FWHM and (B) the percentage population of clustered species ($100\% \times A2/(A1 + A2)$, where Ai is the integrated area of the i^{th} mode of the multimodal Gaussian distribution fit). This demonstrates that the decrease in the population of clustered species for DC-SIGNR begins at lower [Man] than the decrease in the size of the smaller QR-DiMan·DC-SIGNR clusters, indicating that the lectin-glycan-QR interactions in the larger complexes are weaker in affinity than those of the smaller complexes. 277

Figure C.9 Typical high angle annular dark field scanning transmission electron microscope (HAADF STEM) images for PQR 4:1 samples of (A-C) QR-DiMan with DC-SIGN; (D-F) QR-DiMan with DC-SIGNR; and (G-I) a pre-incubated sample of QR-DiMan and DC-SIGNR with 2mM of D-mannose (Man). Each incubated for 20 min prior to plunge-freezing and subsequent imaging. 279

Figure C.10 Comparison of inter-QR centre to centre NNDs (ccNND) to inter-QR perimeter to perimeter ppNNDs (NND) for each QR obtained from the STEM images of QR-DiMan incubated with (A) DC-SIGN, (B) DC-SIGNR or (C) a pre-incubated QR-DiMan·DC-SIGNR with D-mannose. Distributions demonstrate the regions associated with crosslinking (ppNND ~2-10 nm, in parallel and non-parallel inter-QR orientations, ccNND <10 nm and ccNND >10 nm, respectively), simultaneous binding (ppNND >10 nm) or non-specific QR adsorption (ppNND <2 nm)..... 280

Figure C.11 Complete pots of the Δf and ΔD against time for the formation of SLBs of different tris-NTA DODA proportions; lectin association; titration of G5-Man; and subsequent dissociation of G5 in buffer, with SLBs bearing (A) 0.2% DC-SIGN, (B) 0.4% DC-SIGN, (C) 1% DC-SIGN and (D) 2% DC-SIGN, (E) 0.2% DC-SIGNR, (F) 0.4% DC-SIGNR, (G) 1% DC-SIGNR, (H) 2% DC-SIGNR. 281

Figure C.12 Cryogenic transition electron microscopy (cryoTEM) images of G5-Man on SLBs bearing 2% DC-SIGNR demonstrating potential tension lines of "islands" of G5-Man-DC-SIGNR-SLB. 282

Figure C.13 Inhibition of pseudo-typed viral particles bearing the vesicular stomatitis virus glycoprotein (VSV-G, as negative control) by various G13-conjugates together with Camostat and chloroquin. No significant or very weak inhibition was observed for G13-Tz-CRD, G13-Tz-ECD and G13-OH, as expected. Camostat showed no apparent inhibition even at 100 μM, while chloroquine showed significant inhibition at 100 μM, as expected. 282

Figure C.14 Summary of dose-dependent inhibition of Vero76 cells using G13-CRD_{100}, G13-CRD_{115} and G13-OH control against: various pseudo-typed virus bearing the S protein from SARS CoV-2 variants (A) WT (Wuhan wild-type), (B) B.1, (C) B.1.617.2 (delta) and (D) BA.1 (Omicron); and (E) a control pseudo-typed virus bearing vesicular stomatitis virus glycoprotein (VSV-G). G13-CRD_{115} gives the highest inhibition while G13-OH control shows no significant inhibition..... 283

Figure D.1 Histograms for the number of "false" FRET events for each FRET efficiency observed by smFRET for DC-SIGN bearing only donor only dye (DC-SIGN-Atto488) over increasing acceptor threshold (TA) for a fixed threshold for the sum of acceptor and donor intensities of 15 a.u.. All histograms were fitted with unimodal Gaussian distribution (parameters summarised in Table D.1). 285

Figure D.2 Plot of the proportion of fluorescence detected as a FRET event by the donor only control sample (DC-SIGN-Atto488) obtained by percentage of FRET events counted with increasing acceptor threshold (TA) compared to that at $TA = 0$. This demonstrates that the percentage of observable donor only events drops below 1% at a TA of 7 a.u.. 287

Figure D.3 Histograms for the number of FRET events for each FRET efficiency observed by smFRET for HTDC-SIGN over increasing acceptor threshold (TA) for a fixed threshold for the sum of acceptor and donor intensities of 15 a.u.. All histograms were fitted with multimodel Gaussian distribution (parameters summarised in Table D.2). 288

Figure D.4 Histograms for the number of FRET events for each FRET efficiency observed by smFRET for HTDC-SIGNR over increasing acceptor threshold (TA) for a fixed threshold for the sum of acceptor and donor intensities of 15 a.u.. All histograms were fitted with multimodel Gaussian distribution (parameters summarised in Table D.2). 289

Figure D.5 Example FRET efficiency distribution derived from the reweighted sampling of the MD simulation of DC-SIGN compared against the FRET efficiency histogram obtained by smFRET at a $T\mathrm{sum}$ of 15 a.u. and a TA of 14 a.u.. 292

Figure E.1 2-dimensional schematic depicting the core size (darker grey) and hydrodynamic size (lighter grey) of a QR. Where $D\mathrm{h}$ and $L\mathrm{h}$ are the glycan-QR hydrodynamic diameter and length, respectively, $D\mathrm{core}$ and $L\mathrm{core}$ are the QR core diameter and length, H is the height of the cylindrical section and $l\mathrm{h}$ is the hydrodynamic length of the ligand. 297

Figure E.2 Histogram of the core diameter distribution, fitted with Gaussian fits, obtained from the STEM images of QD-EG$_{11}$-DiMan reported in Guo, et al., 2017.[8] Where $D\mathrm{core}$ is the mean core diameter. 298

Figure E.3 (A) Dimer of DC-SIGNR CRDs dimers obtained by XRC (PDB: 1xar),[15] and (B) model of tetramer of DC-SIGN CRDs based on superimposition of individual CRDs from PDB: 1k9i.[16] (Red: primary Ca-Ca distance of adjacent CRDs; black: shortest primary Ca-Ca distance of diagonal CRDs; green: longest primary Ca-Ca distance of diagonal CRDs). 298

Figure E.4 Side-on schematic demonstrating separation distance between the lectin binding contact area to the glycan-QR hydrodynamic surface (d) for a rod of hydrodynamic diameter of $D\mathrm{h}$ with respect to the inter-binding site distance furthest from the rod surface (p). Where d can be obtained by Equation E.13. 299

Figure F.1 (A) Skeletal structures of DHLA-EG$_3$-Man and DHLA-EG$_3$-OH. Volume population hydrodynamic diameter distribution histograms, fitted with the Gaussian equation, for QDs capped with (B) 100% DHLA-EG$_3$-Man (QD-Man$_{100\%}$), (C) 75% DHLA-EG$_3$-Man and 25% DHLA-EG$_3$-OH (QD-Man$_{75\%}$/-OH$_{25\%}$), (D) 50% DHLA-EG$_3$-Man and 50% DHLA-EG$_3$-OH (QD-Man$_{50\%}$/-OH$_{50\%}$), and (E) 25% DHLA-EG$_3$-Man and 75% DHLA-EG$_3$-OH (QD-Man$_{25\%}$/-OH$_{75\%}$). 302

Figure F.2 Plots of the relative FRET ratios (*i.e.* $I\mathrm{Ad}\lambda I\mathrm{Dd}\lambda \div p\mathrm{Man}$, where I is the fluorescence intensity under of the acceptor, A, and donor, D, and $p\mathrm{Man}$ is the percentage population of man ligand per QD) against the protein:QD molar ratio (PQR) for QD-Man$_{100\%}$, QD-Man$_{75\%}$/-OH$_{25\%}$, QD-Man$_{50\%}$/-OH$_{50\%}$ and QD-Man$_{25\%}$/-OH$_{75\%}$ (dark to light) incubated with Atto594 labelled (A) DC-SIGN, (B) DC-SIGN-C and (C) DC-SIGN+R. (Data fitted with Hill equations or bimodal Hill equation to highlight the trend in data, not for quantitative evaluation). 303

Figure F.3 (A) Volume population hydrodynamic diameter distribution histogram of QD capped with DHLA-EG_3-DiMan, with a skeletal structure of DHLA-EG_n-DiMan. Binding curves fitted with Equation 3.1, showing the relationship between the FRET ratio at different temperatures for QD-EG_3-DiMan with labelled protein, for QD-EG_3-DiMan with Atto594 labelled (B) DC-SIGN (PQR 1:1) and (C) DC-SIGN-C (PQR 1:1); (D) Van 't Hoff plots fitted with Equation 3.4, showing the relationship between $\ln Kd$ and $1/T$, derived from the FRET binding curves of QD-EG_3-DiMan with DC-SIGN (blue squares), DC-SIGN-C (green circles); and (E) A graphical comparison of the enthalpy and entropy change contributions to the Gibbs free energy change of the binding of DC-SIGN, DC-SIGN-C with QD-EG_n-DiMan with linker lengths n = 3 or 11, at 298 K (see Chapter 3 for n = 11). **305**

List of Abbreviations

A	Acceptor fluorophore
Ab	Antibody
ACE2	Angiotensin-converting enzyme 2
AFM	Atomic force microscopy
AR	Aspect ratio
B_0	Magnetic field strength
BRET	Bioluminescence resonance energy transfer
BSA	Bovine serum albumin
CCL	CC-chemokine ligand
CD	Cluster of differentiation
CF	Correction factor
CNT	Carbon nanotube
Con A	Concanavalin A
COVID-19	Coronavirus disease 2019
CRD	Carbohydrate-recognition domain
Cryo	Cryogenic
Ct	Citrate
CVC	Critical vesicular coverage
Cw-EPR	Continuous wave electron paramagnetic resonance
D	Donor fluorophore
DAMP	Damage associated molecular pattern
DC	Dendritic cell
DC-SIGN	Dendritic cell-specific intercellular adhesion molecule (ICAM)-3 grabbing non-integrin
DC-SIGN-C	DC-SIGN with a C-terminal 16 amino acid deletion
DC-SIGN+R	DC020 with the last 47 amino acids at the C-terminal replaced with the equivalent residues of DC-SIGNR
DC-SIGNR	DC-SIGN related
DEER	Double electron–electron resonance
DENV	Dengue virus
DHLA	Dihydrolipoic acid
DiMan	Mannose-α-1,2-mannose
D_{h}	Mean hydrodynamic diameter
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DODA	Dioctadecylamine
DOPC	1,2-di-(9Z-octadecenoyl)-sn-glycero-3-phosphocholine
EBOV	Ebolavirus
EC_{50}	Effective concentration at 50% inhibition
ECD	Extracellular domain
EDTA	Ethylenediaminetetraacetic acid
EDX	Energy dispersive x-ray spectroscopy
EG	Ethylene glycol
EM	Electron microscopy
Env	Envelope protein

EPR	Electron paramagnetic resonance
F	FRET ratio
Fabs	Fragment antigen-binding region
FRET	Förster resonance energy transfer
FWHM	Full width at half maximum
G	Glycan
GB	Galactose-binding
Gal	Galactose
GFP	Green fluorescent protein
GNP	Gold nanoparticle
GP (or gp)	Glycoprotein
HA	Haemagglutinin
HAADF	High angle annular dark field scanning
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HT	Heterotetrameric
IAV	Influenza A virus
IC_{50}	Concentration at 50% inhibition
ICAM	Intercellular adhesion molecule
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
ILNAR	Interferon-α/β receptor
ITC	Isothermal titration calorimetry
JEV	Japanese encephalitis virus
K_{a}	Equilibrium binding association constant
K_{d}	Equilibrium binding dissociation constant
K_{i}	Equilibrium inhibition constant
k_{off}	Off (binding association) rate coefficient
k_{on}	On (binding dissociation) rate coefficient
L	Lectin
LA	Lipoic acid
LB	lysogeny broth
LC-MS	Liquid chromatography mass spectrometry
LE	Labelling efficiency
LGI	Lectin-glycan interaction
Man	D-mannose
MD	Molecular dynamics
MHC	Major histocompatibility complex
MLGI	Multivalent lectin-glycan interaction
MLV	Marine leukemia virus
MR	Mannose receptor
MS	Mass spectrometry
MTSSL	S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate
MW	Molecular weight

m/z	Mass to charge ratio
NIR	Near-infrared
NMR	Nuclear magnetic resonance
NND	Nearest neighbour distance (pp = perimeter-perimeter; cc = centre-centre)
NP	Nanoparticle
NSET	Nanoscale surface energy transfer
NSL	Non-specifically labelled
OD	Optical density
PAMP	Pathogen associated molecular pattern
PCI	Protein-carbohydrate interaction
PDB	Protein Data Bank
PEG	Poly(ethylene glycol)
PELDOR	Pulse electron-electron double resonance
PMMA	Poly(methyl methacrylate)
PMT	Photomultiplier tube
PNA	Peanut agglutinin
ppGNP	Partially PEGylated GNP
PQR	Protein to quantum dot/rod ratio
PRR	Pattern recognition receptor
QD	Quantum dot
QE	Quenching efficiency
QR	Quantum rod
R_0	Förster radius
R^2	Coefficient of determination
RCA	Ricinus communis agglutinin
RI	Refractive index
RSS	Residual sum of squares
R.t.	Room temperature
QCM(-D)	Quartz crystal microbalance (with dissipation monitoring)
QY	Quantum yield
SARS-CoV	Severe acute respiratory syndrome coronavirus
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SDSL	Site-directed spin label
SLB	Supported lipid bilayer
SmFRET	Single molecule Förster resonance energy transfer
SPR	Surface plasmon resonance
STEM	Scanning transmission electron microscopy
SUV	Small unilamellar vesicle
$t_{½}$	Half-life
TCEP	Tris(2-carboxyethyl)phosphine
TEM	Transmission electron microscopy
T_{FH} cell	Follicular helper T cell
TFP	Tetrafluorophenyl
T_H cell	Helper T cell
TLR	Toll-like receptor

TMPRSS2	Transmembrane protease, serine 2
TOP	Trioctylphosphine
TOPO	Trioctylphosphine oxide
Tris (or tris base)	2-amino-2-(hydroxymethyl)propane-1,3-diol
Tris-NTA	Tris-nitrilotriacetic acid
Tz	Tetrazine
UV-vis	Ultraviolet to visible light
VSV(-G)	Vesicular stomatitis virus (glycoprotein)
WHO	World health organisation
WNV	West Nile virus
XRC	X-ray crystallography
ΔD	Dissipation shift
Δf	Frequency shift
ΔG^{o}	Standard Gibbs free energy change of association
ΔH^{o}	Standard enthalpy change of association
ΔH_0	Central spectral linewidth
ΔS^{o}	Standard entropy change of association
λ_{abs}	Absorption wavelength
λ_{em}	Emission wavelength
λ_{ex}	Excitation wavelength

Chapter 1
Introduction

1.1 Overview

Protein-carbohydrate interactions play an essential role in the communication of cells in the immune system. They are thus also deeply involved in a lot of immune related diseases from viral infection to allergy or cancer promotion, making them widely studied as therapeutic targets. However, due to the multimeric nature of many carbohydrate-binding proteins (*i.e.* lectins), it can be difficult to design potent targeting agents using conventional monovalent molecules. Instead, in order to achieve high affinity, a lectin targeting agent must maximise the number of occupied binding sites using multiple carbohydrate groups and display a good spatial match between the arrangement of its carbohydrates with the specific binding site arrangement of the lectin. Therefore, in order to inform the rational design of glycan-bearing therapeutics which can target such interactions with both high potency and specificity, a comprehensive understanding of the target lectin's multimeric structure and biophysical glycan binding mechanisms are required. However, for a lot of multimeric lectins, these details remain largely unknown.

In recent years, glycosylated nanoparticles have been developed as a way to both target and probe multivalent lectin-glycan interactions (MLGIs), both revealing key information regarding lectin-glyconanoparticle binding affinity and binding modes, as well as displaying high potential as therapeutic agents against diseases such as viral infection. This thesis continues this effort, exploiting the unique multifunctional behaviour of glyconanoparticles, such as glycosylated quantum dots and gold nanoparticles, to investigate the biophysical and structural contributions of MLGIs using a case study pair of lectins, DC-SIGN and DC-SIGNR, which display similar tetravalent and mannose-binding character but different affinities and binding modes towards some of the same glycans.

This chapter provides an introduction into the function and therapeutic relevance of MLGIs in the immune system and an overview of the current advances in nanotechnology-based approaches for probing the binding of lectins, as well as their potential therapeutic applications. Finally, the research aims of this thesis are presented with an overview of how each chapter addresses some of the gaps in the current research alluded to throughout.

1.2 Multivalent Lectin-Glycan Interactions

1.2.1 Protein-Carbohydrate Interactions in the Immune System

Protein-carbohydrate interactions (PCIs) are widespread in nature and play an essential role in the way that cells interact and communicate amongst themselves and with other species within biological systems. In solution, such interactions have a variety of roles, from more fundamental processes such as catalysis and the organisation of the extracellular matrix, to more complex functions such as cell proliferation and opsonisation in the immune system.[1-3] On cell membranes, carbohydrate-binding proteins are used as a way of sensing a cell's external environment: they are often used to detect and inform the rest of the cell (and the cells around it) about any interesting molecules that may be passing by, alternatively they be used to grab onto specific neighbouring cells or draw in immunologically interesting materials from the environment.[1, 4, 5] Due to the degree of sensitivity required by cells for recognition of the wide variety of glycans in nature, carbohydrate-binding proteins have inevitably become indispensable players in the immune system, facilitating processes such as immune cell adhesion and communication (**Figure 1.1C**), as well as recognition of glycans associated with pathogens or damaged cells and, if necessary, internalisation of pathogens or antigens (**Figure 1.1A**) and/or regulation of the rest of the immune system by the modulation of signalling molecules (**Figure 1.1B**).

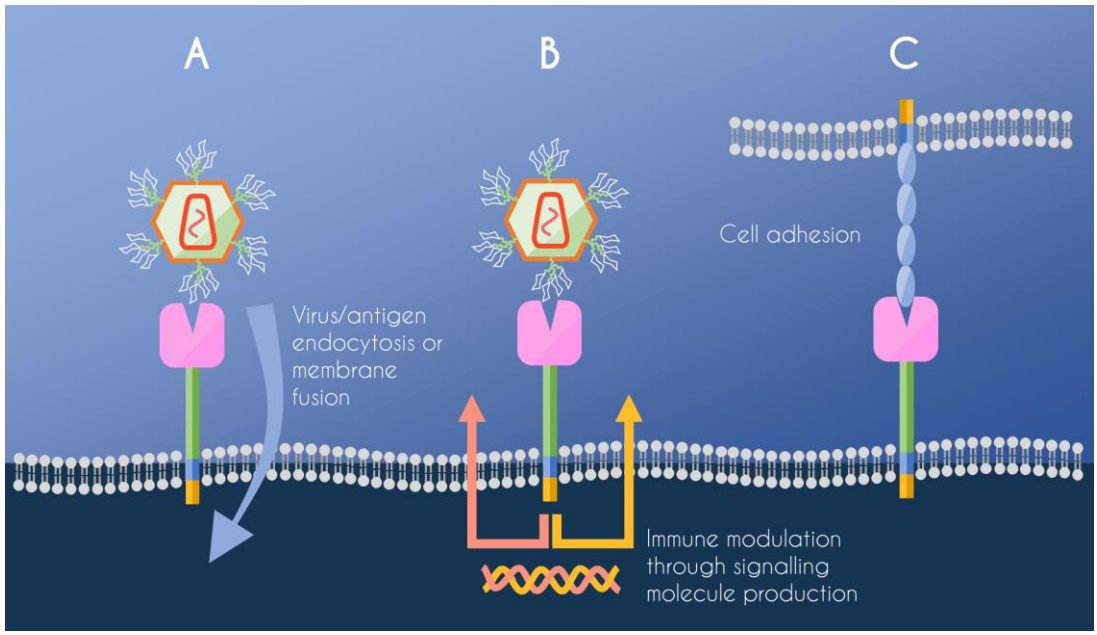

Figure 1.1 Schematic depicting three possible immunological routes involving specific protein-carbohydrate interactions on a cell membrane.

The immunological relevance of such PCIs means that they are also closely associated with many important human diseases. For instance, though such interactions can be used to detect pathogens at the cell surface as part of an immune response, many viruses have evolved to use these interactions to infiltrate host cells. Here, where large pathogens such as bacteria often enter the immune system *via* mechanisms such as phagocytosis, smaller structures such as viruses often rely on receptor-mediated endocytosis or fusion in order to release their genetic material into the host cells.[6] Viruses therefore often make use of specific PCIs to enable initial contact with the host cell. This is facilitated by either the interaction of a virus' specifically arranged surface glycan patterns with the host cell's carbohydrate binding proteins (**Figure 1.1A**), or vice versa.[5, 7] Carbohydrate recognition has also been utilised by the immune system to detect and supress cancer cells and tumour growth.[1] However, some of these interactions can lead to direct promoting effects on cancer growth and metastasis, for example through adhesion of platelets, which act as protection against immune elimination, or endothelial cells, allowing for *trans*-endothelial cell migration of metastatic cells.[1, 8] Additionally, PCIs have also been associated with immune tolerance due to their involvement in the secretion of inflammatory, immune suppressive or adaptive immunity related signalling molecules. These too can lead to harmful effects such as promotion of cancer growth through immune suppression, or a variety of autoimmune and inflammatory diseases such as rheumatoid arthritis, type 1 diabetes and allergy, due to the undesirable activation of inflammatory immune responses.[1, 9]

PCIs are therefore a hot topic in the development of immune-related medicines. For example, endocytosis mechanisms can be exploited for applications such as targeted *in vivo* drug and gene delivery;[10] carbohydrate-induced modulation of immune response can be utilised for new immunotherapeutic development;[11] and targeted inhibitors may be developed to potently block specific viral or bacterial adhesion to cell surfaces.[12]

1.2.2 Lectins, LGIs and MLGIs

PCIs require the accurate detection and discrimination of carbohydrates, which is facilitated by a class of proteins known as *lectins* (aptly named from the Latin verb *lego* (or *lectus*), meaning to choose or pick out).[13] The affinity of a given lectin binding site is dictated by its chemical and spatial complementarity with different mono-/oligo-/poly-saccharides. In animals, lectins have been classified into five main groups: C-type lectins, which bind

via Ca^{2+} coordination; I-type lectins, a glycan binding subset of immunoglobulins; Galectins (or S-type lectins), which bind solely to β-galactosides, often *via* cysteine residues; P-type lectins, which are defined by their cation-dependent binding to mannose-6-phosphate-containing glycoproteins, through non-well-defined binding domains; and pentraxins, which are characteristically pentameric and also Ca^{2+}-dependent.[14] Each of these lectin classes and the individual proteins within them bring different combinations of structures and bonding mechanisms. This is a lectin's primary mode of specificity, where only glycans containing carbohydrate residues that are compatible with the sterics and intermolecular interactions of the lectin's binding sites are able to form strong lectin-glycan interactions (LGIs).

Despite this specificity, LGIs are very weak in order to prevent glycans from becoming kinetically trapped onto the lectin surface.[15] However, this means that monovalent 1:1 lectin-carbohydrate binding is often too dynamic to be sustained long enough to initiate any subsequent biological function. Because of this, many lectins exist in a multimeric form with multiple carbohydrate binding sites, allowing the lectin to bind multivalently to multiple carbohydrate residues (*i.e.* **Figure 1.2)**. These multivalent lectin-glycan interactions (MLGIs) facilitate an additional level of specificity. Here, specificity is not only determined by the *affinity* of each individual carbohydrate binding site towards any particular glycan, but also by the overall binding strength induced by the significant affinity enhancement which results from each additional binding interaction, *i.e.* the *avidity*.

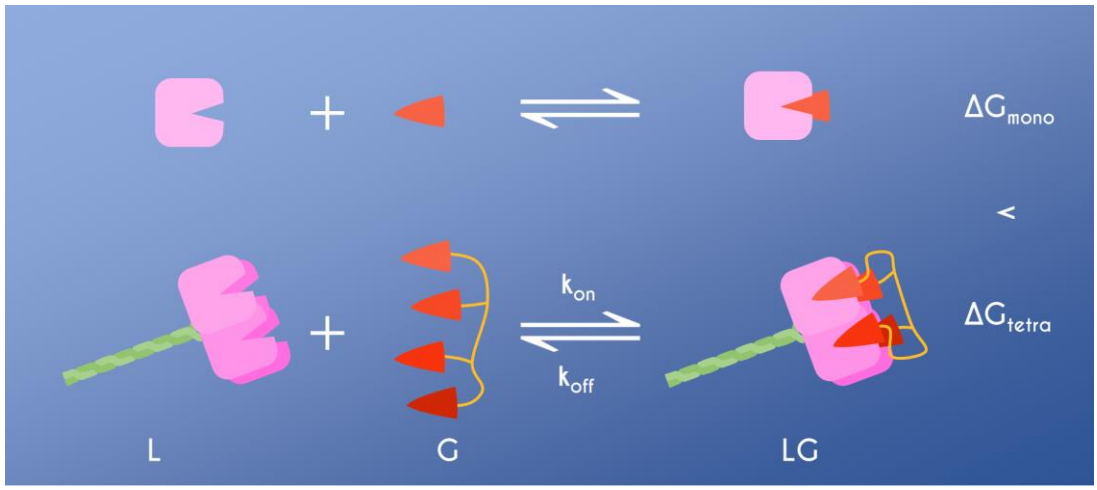

Figure 1.2 Schematic depicting the equilibrium tetravalent binding affinity (ΔG_{tetra}) between a tetravalent lectin L with a tetravalent glycan G.

The affinity or avidity of the association of a lectin-glycan pair (given by **Equation 1.1**; where L is the lectin, G is the glycan and LG is the lectin-glycan complex) can be quantified by the equilibrium binding *dissociation constant* (K_d) which describes the strength of the interaction by the comparative ratio of the concentrations of free unbound species to bound complex present in the system at equilibrium (**Equation 1.2**; where [ ] denotes the concentration).[a] The K_d is related to the overall energy required for LGI to occur (*i.e.* the Gibbs free energy, ΔG) by **Equation 1.3** (where R is the ideal gas constant and T denotes the temperature).[b]

$$\mathrm{L} + \mathrm{G} \rightleftharpoons \mathrm{LG} \qquad 1.1$$

$$K_d = \frac{[L][G]}{[LG]} \qquad 1.2$$

$$\Delta G = RTlnK_d \qquad 1.3$$

The avidity of an MLGI can be compared to their constituent monovalent interactions in two ways: by the degree of cooperativity (α) or the enhancement factor (β).[5] α describes by how much the MLGI enhances the ΔG in comparison to the same number (N) of constituent monovalent interactions (**Equation 1.4**). Naturally, if $\alpha = 1$, then there is no cooperative effect between the formation of one constituent interaction and the formation of the next, they simply act additively. However, if $\alpha > 1$ then the formation of each constituent interaction will *enhance* the formation of the next (*i.e.* positive cooperativity). Conversely, if $\alpha < 1$ then the formation of each constituent interaction will *inhibit* the formation of the next (*i.e.* negative cooperativity).

---

[a] The K_d describes the propensity for the bound species to dissociate, as opposed to the association constant (*i.e.* the equilibrium constant) K_a which describes the propensity for unbound species to associate, where $K_\mathrm{d} = 1/K_\mathrm{a}$. Despite this, K_d is the conventionally used descriptor of affinity in biochemistry, whereby the lower the magnitude, the stronger the interaction.

[b] The Gibbs free energy change describes the overall energy required for a reaction to take place. If $\Delta G > 0$ then the reaction is unfavourable; if $\Delta G < 0$ then the interaction is favourable, whereby, the larger the magnitude, the stronger the interaction.

However, negative cooperation does not mean that the enhancement in affinity of polyvalent interaction in comparison to the monovalent interaction is irrelevant. On the contrary, due to logarithmic relationship between ΔG and K_{d} (**Equation 1.3**), **Equation 1.4** can be rewritten as $K_{\mathrm{d\,poly}} = K_{\mathrm{d\,mono}}{}^{\alpha N}$. This shows that, even if $\alpha \leq 1$, multivalency can still afford very effective affinity enhancement provided $\alpha N > 1$ and $K_{\mathrm{d\,mono}} < 1\mathrm{M}$, as this will still result in $K_{\mathrm{d\,poly}} < K_{\mathrm{d\,mono}}$. Such enhancement can still be highly influential in biological settings, despite displaying negative or no cooperativity. For this reason, the enhancement factor β (**Equation 1.5**) often provides a more useful descriptor for the avidity of MLGIs in comparison to their monovalent constituent interactions.

$$\Delta G_{\mathrm{poly}} = \alpha N \Delta G_{\mathrm{mono}} \qquad 1.4$$

$$\beta = \frac{K_{\mathrm{d\,mono}}}{K_{\mathrm{d\,poly}}} \qquad 1.5$$

The free energy of binding is derived from two constituent energetic terms associated with formation of the lectin-glycan complex: the change in enthalpy[c] (ΔH) and change in entropy[d] (ΔS) (**Equation 1.6**). The enthalpy change for a monovalent LGI is essentially determined by the amount of energy involved in forming physical bonds between the lectin binding site with the carbohydrate binding partner. Thus, in an ideal system, it can be assumed that the enthalpy change associated with each additional constituent monovalent interaction is independent of the number of interactions that have already formed. In this case, the enthalpy change of each interaction simply acts additively, *i.e.* $\Delta H_{\mathrm{poly}} \approx N \Delta H_{\mathrm{mono}}$; thus the higher the multivalency, the more favourable the binding enthalpy. If, however, any established constituent interactions interfere antagonistically with subsequent binding, then this may lead to an enthalpically diminished system, whereby $\Delta H_{\mathrm{poly}} < N \Delta H_{\mathrm{mono}}$. This

---

[c] The enthalpy change of a reaction describes the overall amount of energy required for the conversion of pre-existing bonds to those involved in the formation of the resulting compound. A reaction is enthalpically favourable only if $\Delta H < 0$.

[d] The entropy change of a reaction describes the overall energy change per kelvin associated with the change in the number of degrees of freedom of the system. A reaction is entropically favourable only if $\Delta S > 0$.

can happen if there is too much strain involved in the formation of subsequent interactions which could occur if the arrangement of the polyvalent glycans does not optimally match the specific arrangement of the lectin's carbohydrate binding sites and if the binding partners are too rigid to compensate for these spatial differences.

$$\Delta G_{\mathrm{poly}} = \Delta H_{\mathrm{poly}} - T\Delta S_{\mathrm{poly}} \qquad 1.6$$

Entropy changes associated with binding, which are dependent on the number of degrees of freedom formed and lost, can be broken down into the sum of its components; for MLGIs these can be given as the sum of the translational, rotational, conformational and solvent changes in entropy ($\Delta S_{\mathrm{trans}}$, ΔS_{rot}, ΔS_{conf}, ΔS_{solv}, respectively; **Equation 1.7**). In solution, $\Delta S_{\mathrm{trans}}$ describes the change in the degrees of translational freedom in x, y and z directions and ΔS_{rot} describes the change in the rotational degrees of freedom. These are logarithmically dependent on the mass and size of the molecules involved which, for any 1:1 LGI or MLGI, do not change significantly enough to impose too great a change to the overall entropy.[e] $\Delta S_{\mathrm{trans}}$ and ΔS_{rot} are also dependent on the change in the number of molecules present in the system as each species adds its own contribution to the total number of degrees of freedom. This, therefore, has a much more significant influence on the overall affinity, especially in systems where multiple lectins bind to a single glycan. However, it is worth noting that once the initial binding interaction has formed, $\Delta S_{\mathrm{trans}}$ and ΔS_{rot} are not likely to change upon the formation of additional intramolecular binding events as the total mass, size and number of the complex species will stay the same. Thus, provided the lectin does not bind to multiple glycans at a time, these components are unlikely to influence the β enhancement factor associated with multivalency and therefore impart a negative cooperation element to the overall α.

$$\Delta S = \Delta S_{\mathrm{trans}} + \Delta S_{\mathrm{rot}} + \Delta S_{\mathrm{conf}} + \Delta S_{\mathrm{solv}} \qquad 1.7$$

Instead then, the main entropic contributors to β are likely to be ΔS_{conf} and ΔS_{solv}. ΔS_{solv} is the entropic change associated with the rearrangement of

---

[e] On cell membranes, due to the limited lateral mobility, the number of degrees of freedom are also reduced, thus these effects impose even less significant change to $\Delta S_{\mathrm{trans}}$ and ΔS_{rot}.

solvent molecules upon binding. This can result from the release of solvent molecules from the binding sites or the rearrangement of solvent molecules caused by the change in the hydrophilic surface area of the complex. This means that, unlike $\Delta S_{\mathrm{trans}}$ and ΔS_{rot}, ΔS_{solv} should in theory increase linearly with the number of constituent binding interactions due to the increased liberation of water molecules from the lectin and glycan binding sites. This means that, as with ΔH, $\Delta S_{\mathrm{solv\,poly}} \approx N\Delta S_{\mathrm{solv\,mono}}$, enhancing β. Finally, ΔS_{conf} denotes the entropic influence of internal conformational changes within the protein or glycan upon binding. Given that most binding partners will display some degree of flexibility prior to complexation, LGIs will usually induce a negative ΔS_{conf} contribution as the number of available conformations are reduced upon binding. This effect can then be made substantially larger for MLGIs, depending on the relative flexibility of the protein and the glycan, due to the significant structural reorganisation that may be required to bridge multiple binding sites. Interestingly, this effect can be compensatory, as even if ΔS_{conf} is large enough to induce a negative ΔS_{poly} term, high flexibility in either of the binding partners reduces the effect of steric strain, making it more feasible for optimal ΔH_{poly} to be achieved.

In addition to the thermodynamic stability of LGIs, the actual biological relevance of an interaction is intimately dependent on the binding kinetics. If an interaction is only transient then it will be unlikely to result in any significant biological function, however, if it is too long-lived then the glycan runs the risk of becoming kinetically trapped. The degree of multivalency is therefore key for inducing successful, biologically relevant adhesion. The kinetics can be measured in terms of the on- and off- rates (*i.e.* the association or dissociation of lectin-glycan complex over time). The propensity for which the lectin glycan pair will associate or the complex will dissociate can then be quantified by the on- and off- rate coefficients (k_{on} and k_{off}), respectively. For 1:1 lectin-glycan complexation interactions can be assumed to follow a second order association (**Equation 1.8**) and a first order dissociation (**Equation 1.9**) rate profile; the ratio of which, defines the equilibrium dissociation constant of the reaction (**Equation 1.10**), whereby a reaction is only favourable if $k_{\mathrm{on}} > k_{\mathrm{off}}$. For most systems such as these, the k_{on} is not the determining factor in multivalent enhancement, as it is mainly dictated by the rate of diffusion of the species in solution, whereas, k_{off} is much more dependent on the strength of the interaction.

$$\frac{\mathrm{d[LG]}}{\mathrm{dt}} = k_{\mathrm{on}}[L][G] \qquad 1.8$$

$$\frac{d[LG]}{dt} = -k_{\mathrm{off}}[LG] \qquad 1.9$$

$$K_{\mathrm{d}} = \frac{k_{\mathrm{off}}}{k_{\mathrm{on}}} \qquad 1.10$$

Overall, a lectin's avidity in comparison to its constituent monovalent affinity is essentially dictated by how well the arrangements of the glycan's carbohydrates and the lectin's multiple binding sites match in space. This therefore demonstrates the important role of multivalency in a lectin's *specificity* towards multivalent glycans. Glycans which are able to bridge all of the lectin's binding sites with minimal steric strain on both binding partners will form the strongest binding interactions. Those that impart unfavourable conformational changes or are simply unable to bridge these binding sites will form much weaker interactions and will either dissociate before any biological functionality is achieved or may be forced to form other, less optimal binding configurations.[16] Knowledge regarding the binding mechanisms of MLGIs is therefore key for the informed design of polyvalent glycoconjugates that can target multivalent lectins with both high affinity and specificity. However, despite extensive studies, details regarding the structures of some critically important cell surface lectins remain poorly understood. This is due to limitations of the current methods for probing the structures of such flexible and complex cell membrane lectins. Additionally, little information has been gathered about the kinetics and thermodynamic contributions to MLGIs and how these relate to the lectin structure and overall mechanism of binding. Therefore, new techniques that can improve our structural and mechanistic understanding of multimeric lectins and their MLGIs are urgently needed.

1.2.3 DC-SIGN

Dendritic cells (DCs) are a specific set of immune cells involved in bridging the innate and adaptive immune system by informing other immune cells about the details of an infection. Their roles involve patrolling the body for signs of pathogen infection; assimilating any pathogenic material through receptor-mediated endocytosis; processing and presenting antigens on their

cell surfaces *via* major histocompatibility complex (MHC) class II molecules; and then migrating to the lymph nodes to allow other cells, such as T cells, to detect the presence of the specific pathogen. In addition, they also produce signalling molecules, such as cytokines and chemokines, to promote specific immune response tailored towards the given disease. This is facilitated by the large number of cell membrane pattern recognition receptors (PRR), such as C-type lectin receptors, which can recognise the pathogen- or damage-associated molecular patterns (P/D-AMPs) from any antigenic material (*e.g.* glycoproteins or liposaccharides). Once recognised, a series of intracellular biochemical reactions are initiated, which lead to the expression of specific genes that induce the production of these signalling molecules. The signalling molecules are then released and are either received by the same cell or different cells, to direct cells towards a specific immune response.[17]

DC-SIGN (Dendritic Cell-Specific Intercellular adhesion molecule (ICAM)-3 Grabbing Non-integrin; CD209) is one such PPR, primarily situated on many subsets of dendritic cell.[18, 19] DC-SIGN is a type II membrane C-type lectin consisting of four monomeric units, each comprising of a small cytoplasmic domain at the N-terminal, which contains the key internalisation motifs, followed by a transmembrane and extracellular domain, which bears its glycan-binding functionality (**Figure 1.3**).[20, 21] The extracellular domain consists of a neck of repeating α-helical units which form coiled-coil structures and have been shown to facilitate tetramerisation.[22-25] The neck also acts to project its glycan-binding functionality away from the cell surface by suspending four C-type carbohydrate recognition domains (CRDs) at the C-terminal.[26-28] These CRDs selectively recognise high-mannose and fucose containing glycans, including those found on virus, bacterium, fungus or parasite surfaces, and blood group epitopes such as Lewis^a and Lewis^x.[27, 29, 30]

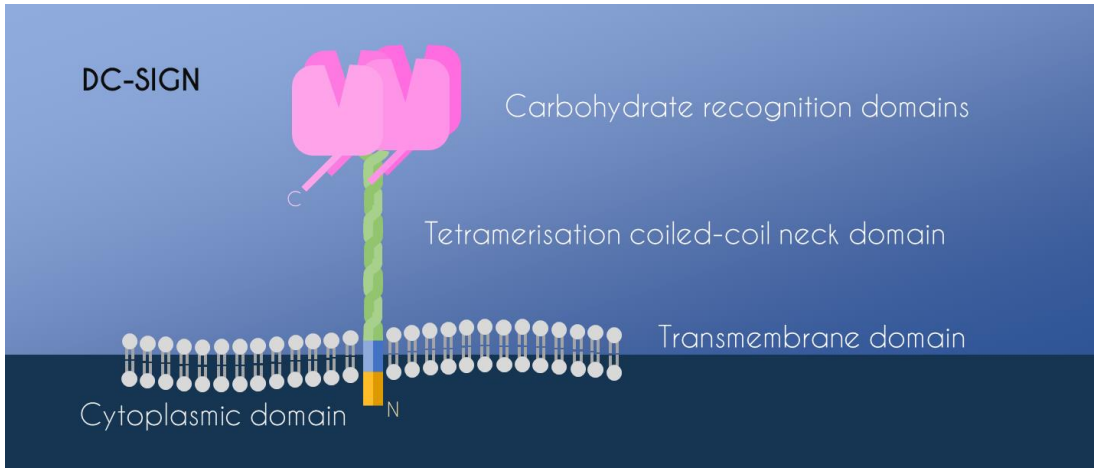

Figure 1.3 Annotated schematic of DC-SIGN.

DC-SIGN has been shown to be deeply involved in a range of immunological processes in the body, such as: cell adhesion, both with ICAM-2 on endothelial cells, which allows the DCs to roll along the endothelial lining and facilitates *trans*-endothelial migration, and ICAM-3 as part of the DC-T cell synapse required to induce T cell activation;[18, 31] pathogen and antigen endocytosis, which is essential for the subsequent presentation of lysed antigenic fragments by MHC class II molecules;[32, 33] and P/DAMP-induced specific cytokine modulation, for immune response regulation.[34-37] Despite these crucial physiological benefits, DC-SIGN was originally identified for its involvement in human immunodeficiency virus (HIV)-1 infection *via* the association of glycoprotein 120 (gp120), a high-mannose containing glycoprotein found on the exterior of HIV-1's viral envelope.[20] This DC-SIGN-associated MLGI allows HIV-1 to enter the dendritic cells, directly infecting the innate immune system, and allowing it to easily transfer to other immune cells, such as the $CD4^+$ T helper cells, because of the natural migration of the DCs to the lymph nodes (**Figure 1.4**).[38] Since then it has also been found to be exploited by a range of other severe viruses such as the Ebolavirus (EBOV),[39, 40] Dengue virus (DENV),[41, 42] Hepatitis C virus (HCV),[43, 44] Japanese encephalitis virus (JEV),[45-47] Measles virus,[48] and more recently, severe acute respiratory syndrome coronavirus (SARS-CoV)-2.[49] The development of highly specific antiviral agents which target DC-SIGN have therefore been of high interest to potently block such viruses from entering the immune cells.[39, 41] Alternatively, targeting agents could be exploited to utilise DC-SIGN's endocytic behaviour to carry specific antigens into the DCs to enhance the immune response against viruses like these which evade the immune system, or from other species which may not ordinarily cause an immune response, such as cancer cells.[10]

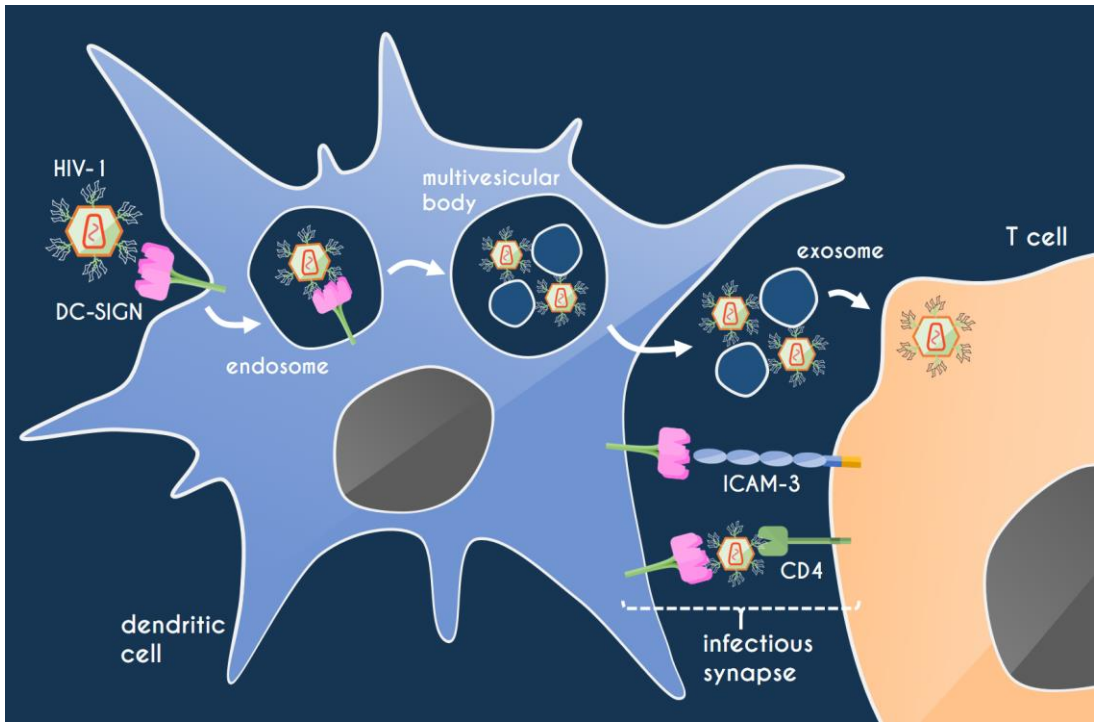

Figure 1.4 DC-SIGN-mediated trans-infection of T cells by HIV-1, either *via* endocytosis and internalisation into multivesicular bodies which then release the virus at the infectious synapse in the presence of exosomes; or *via* DC-SIGN and CD4 binding, both associated with the infectious synapse. Additionally, transient DC-SIGN-ICAM-3 interactions facilitate DC-T cell interaction.[50, 51]

DC-SIGN also displays the remarkable capability to differentiate between different glycan P/DAMPs and modulate the production of different cytokines and chemokines depending on the carbohydrate "fingerprint" of a bound glycoprotein. The specific cytokine signature can stimulate cells to act in different ways, directing the immune system towards either a cellular immune response (*i.e.* by polarising cells which digest pathogens, such as macrophages) or a humoral immune response (*i.e.* by polarising cells which produce antibodies, such as B cells), or towards immune tolerance.[11, 34, 35, 37, 52, 53] This makes DC-SIGN a point of extensive study as it has the potential to be used to tailor immune responses to suit specific needs. For example, towards immune tolerance with respect to autoimmune diseases, allergies and transplant rejections, or towards humoral or cellular immune activation for treating infections or cancers. Though DC-SIGN isn't able to dictate these different immune trajectories on its own, the recognition of different carbohydrate fingerprints have been shown to alter the cytokine expression programme of other PRRs such as toll-like receptor (TLR)2, TLR4 and dectin-1. These differences have been attributed to two general pathways: that

associated with high-mannose containing PAMPs, which direct cytokine production towards a cellular immune response, and that associated with fucose containing PAMPs, which direct cytokine production towards a humoral immune response. The signalling pathways associated with DC-SIGN binding of *high-mannose* containing PAMPs, mainly displayed on many intracellular pathogens such as viruses, works synergistically with other PRRs to promote the transcription of DNA associated with upregulating interleukins (ILs) such as IL-12 and IL-6.[5] IL-12 promotes the differentiation of naïve T helper (T_H) cells into T_H1 cells, which proceed to activate macrophages (involved in cell-mediated immunity).[54, 55] IL-6 promotes the differentiation of T_H cells into T_H17 cells, which are associated with tissue inflammation and many auto-inflammatory diseases.[56] On the other hand, binding of *fucose* containing PAMPs, which are often displayed on parasites,[57] alters this mechanism by inducing the assembly of an alternative DNA transcription complex which competes with that associated with IL-12 and IL-6 gene expression, thus downregulating these cytokines, and instead upregulates chemokines CC-chemokine ligand (CCL)17 and CCL22, and IL-10. CCL17 and CCL22 are associated with attracting T_H2 cells which promote humoral immune responses through the activation of antibody producing B cells.[37, 58] Additionally, this fucose-associated DC-SIGN pathway also facilitates the upregulation of IL-27 which helps generate follicular helper T (T_{FH}) cells which in turn similarly induce the production of antibodies in B cells.[36] IL-10 is an anti-inflammatory cytokine which acts antagonistically with proinflammatory cytokines and is also associated with the suppression of DC and T cell function.[59] Interestingly, the properties IL-10 have made it a point of medicinal interest for applications such as transplant tolerance or induction of the regenerative mechanisms involved in wound healing.[11, 60] Additionally, IL-10 upregulation by DC-SIGN has also been associated with probiotic bacteria to encourage immune tolerance, however, it has also been shown to be exploited by various bacteria and viruses to avoid immune activation, and is also triggered by Lewis antigen epitopes associated with some cancers.[34, 61-63]

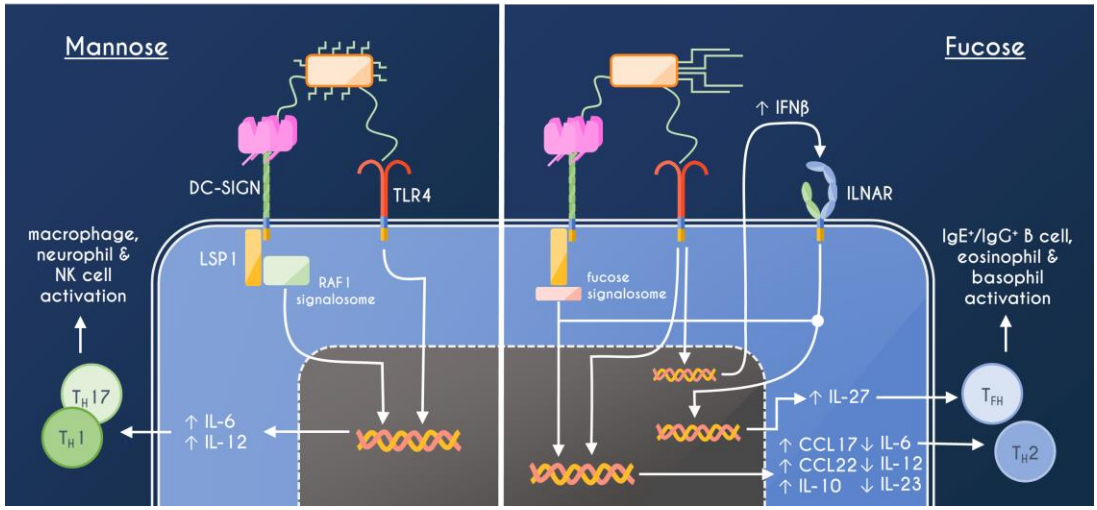

Figure 1.5 DC-SIGN-mediated pathogen recognition and immune regulation. Left: recognition of mannose-containing glycans, in conjunction with activation of other PRRs such as TLR4, enables transcription of genes encoding for IL-12 and IL-6 cytokines, which polarise T_H1 cells and T_H17 that in turn activate cellular immune response (involving macrophages, neurophils and NK cells). Right: recognition of fucose-containing glycans displaces the RAF-1 signalosome (associated with the mannose-associated pathway, thus downregulating IL-12, IL-6 and IL-23) and instead allows the assembly of the fucose signalosome which, in conjugation with other PRRs such as TLR4, enables transcription of anti-inflammatory cytokine IL-10 and chemokines CCL17 and CCL22, which attract T_H2 cells that in turn activate a humoral immune response (involving IgE^+ B cells, eosinophils and basophils), as well as IL-27 (*via* crosstalk with the interferon-α/β receptor (IFNAR)-associated pathway *via* the upregulation interferon-β (IFNβ)) which activates T_{FH} cell differentiation (involved in IgG^+ memory B cell activation).

DC-SIGN also plays a role in the immune regulation of allergens and has shown some association with allergic response. For example, DC-SIGN has been observed to uptake allergens associated with house dust mites (Der p 1 and Der p 2), dogs (Can f 1), peanuts (Ara h 1) and pollen (BG-60).[9] Additionally, over- and under-production of DC-SIGN in relation to other C-type lectins on the same DCs, such as mannose receptor (MR), have been associated with differing T cell response in the presence of allergens. This is currently an interesting point of research as both decreased production of DC-SIGN and increased expression of MR in comparison to DC-SIGN have also been observed in atopic patients.[64-66] Thus, it has been suggested that specific therapeutics which target these lectins may be able to either block allergen uptake or modulate the immune response towards tolerance.[9]

1.2.4 DC-SIGNR

Many closely related variations of DC-SIGN can be found in nature, however, only two have been identified in the human body. The other is known as DC-SIGNR (DC-SIGN-related; Liver (L)-SIGN; CD209L) and is located primarily on the endothelial cells of the liver, lymph nodes and placenta.[67, 68] DC-SIGNR and DC-SIGN share 77% of the same amino acid identity, the same tetrameric architecture and near-identical mannose-binding motifs (**Figure 1.6**).[22, 23, 27, 67, 69] DC-SIGNR also has a high affinity for adhesion molecules such as ICAM-3 and has been shown to participate in the infection of a similar range of viruses.[68, 70] Interestingly, however, despite similarities in the carbohydrate binding domains (CRDs), DC-SIGN and DC-SIGNR show very different selectively towards the same viral glycoproteins.[27, 67, 71, 72] For example, DC-SIGNR is observed bind to HIV-1 much less efficiently than DC-SIGN, however, it has also been found to bind to and effectively transmit the West Nile Virus (WNV), where DC-SIGN shows no such ability.[71, 72] Given the similarities between the two proteins' primary and tertiary structures, these differences in MLGI selectivity are most likely instigated by variation in the quaternary structure of the two proteins, such as the arrangements and orientations of the four binding sites with respect to one another.

Figure 1.6 Comparison of the amino acid sequences of DC-SIGN and DC-SIGNR (asterisks identify identical residues, blue bold shows the transmembrane domain, grey highlights show the 7½ neck repeat units and red bold shows the mannose binding motifs).

These close similarities in amino acid identity yet significant differences in binding activity make DC-SIGN and DC-SIGNR a perfect model lectin pair to investigate how multivalent lectin structure controls properties such as the binding affinity, binding mode, enthalpy, entropy and kinetics of multivalent lectin-glycan binding, distinct from that directly related to chemistry alone. This allows us to gain a deeper understanding the role of spatial selectivity in biology and how these details may be able to aid the design of more specific and potent ligands to target important lectins, such as these.

1.2.5 Limitations and Gaps in Current Research

Given their great biological importance, DC-SIGN and DC-SIGNR have been investigated extensively over the past 20 years, leading to advances in understanding their biological functions, binding behaviours and protein structures. However, despite these steps, there are still large gaps in our understanding of the relationship between the structures and multivalent glycan-binding properties of these lectins, among others. Understanding these MLGIs would allow us to answer questions such as how multivalent lectin binding site arrangements and orientations affect their glycan-binding

mechanisms and affinities; how these binding properties may be influenced by different glycan displays; and what structural aspects of the proteins are responsible for their specificity. Details like these allow us to paint a more comprehensive picture of the mechanisms involved in MLGIs of these proteins, which is crucial for identifying why and how certain viruses differentiate between them, as well as creating rules to guide the design of smarter therapeutics to target lectins, such as DC-SIGN and DC-SIGNR, by exploiting their biological behaviours for medicinal purposes.

The main limitation in the research of DC-SIGN and DC-SIGNR MLGIs comes from the large uncertainty about the two proteins' complete tetrameric structures. This is a classic problem for many complex membrane proteins, as they are often too flexible to be crystallised for structural determination with traditional techniques such as X-ray crystallography (XRC), and too small for state-of-art cryo-electron microscopy techniques. In addition, other conventional biophysical techniques such as isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR), although powerful in revealing quantitative binding thermodynamic and kinetic data, cannot provide important structural information about such interactions. Moreover, ITC has limitations in accurately determining high affinity interactions or interpreting more complex interactions such as crosslinking.[73-75] SPR is also limited by its surface-phase chemistry which can make it difficult to interpret results as affinity can be strongly affected by the kinetics of diffusion to the surface and the surface density and orientation of the surface immobilised binding partner.[76] These techniques therefore miss out on the full range of descriptors for characterising the MLGIs of DC-SIGN and DC-SIGNR, and other lectins, such as: glycan binding site orientation and arrangement, binding mode and valency, and binding affinity, thermodynamics and kinetics, along with how these are dictated by the proteins' amino acid compositions and how these aspects affect one another in relation to different glycans. Therefore, it is crucial to develop alternative techniques that can provide such key structural-mechanistic information to aid our understanding of important MLGIs as well as to inform therapeutic design.

1.3 Glyconanotechnology Strategy and Application in the Study of MLGIs

1.3.1 Introduction to Glyconanotechnology

In general terms, *nanotechnology* is the area of science that deals with the study, manipulation or application of matter at the nanoscale (*i.e.* ~10^{-9}-10^{-7}

metres). Though this encompasses most species on the molecular level, it is generally only applied to new technologies that have been developed at, or for, this scale. Nanoscale technology has captured researchers' imaginations since Richard Feynman famously conceptualised the idea in the late 1950s, leading to breakthroughs in atomic lithography, imaging techniques, self-assembling materials and material property manipulation, among more.[77] It is unsurprising then, that by the 1980s such principles were beginning to be applied to the field of biology and medicine, where such sizes are perfect for the interaction or novel application of proteins and cell components and interfaces. This has led to advances in micro- and molecular biology research; novel therapeutic applications or improved drug potency and targeting; and new diagnostic techniques, exploiting both the complementary size regime and also the unique chemical/physical properties of many materials at this scale.

Glyconanotechnology is therefore the study, manipulation and application of carbohydrate-related biology using nanotechnology. Glyconanotechnology has been used for a range of applications, from therapeutics for inhibiting the adhesion and *trans*-endothelial migration of metastatic cancer cells,[78] novel nanoparticle-based photodynamic or hyperthermia cancer therapies,[79, 80] drug delivery systems,[81] pathogen inhibitors,[82, 83] vaccines,[84, 85] to probes for imaging and diagnostics.[86] For MLGI research, specifically, glyconanotechnology has proved to be incredibly valuable for two main reasons: 1. Nanomaterials, in particular nanoparticles (NPs; broadly defined as materials that are ~10^{-9}-10^{-7} m in three dimensions)[87], provide excellent scaffolds in the design of MLGI-based therapeutics, due to their tuneable size, functionalisation and biocompatibility which can be altered to compliment target lectins/glycans for potential *in vivo* application (see **Section 1.3.2** for more detail); and 2. By exploiting the unique size and physical properties offered by nanotechnology, nano-tools can be developed to probe the molecular scale characteristics (e.g. structures or binding properties) of lectins and MLGIs, which are otherwise inaccessible for more conventional techniques such due to limitations in resolution or practicality (see **Section 1.3.3-1.3.6** for more detail). This thesis therefore focuses on bringing these two aspects together by harnessing the various unique properties of NPs to provide meaningful lectin binding information which can subsequently inform their therapeutic value. Such designs can allow us to both dissect the underlying mechanisms of how lectin structure and glycoconjugate design influence multivalent lectin-glycan binding, and probe their real-world anti-viral efficacy or biological mechanisms.

1.3.2 Glyconanoparticle Design

The utilisation of nanoscale scaffolds as potential lectin targeting therapeutics has been explored since the start of the 1990s, originally by researchers who were exploring linear, dendritic and branched polymer scaffolds, such as Matrosovich, Whitesides and Roy.[88-94] Since the turn of the century, however, the study of glyconanomaterials for targeting MLGIs has expanded significantly to encompass a range of scaffolds with a variety of shapes, sizes, constituent materials and chemical properties, from inorganic scaffolds made from materials such as gold,[95-100] silver,[101, 102] magnetite,[103] and quantum dots;[16, 104] organic structures such as polymer chains,[105-110] dendritic, branched or hyperbranched polymers,[105, 111-114] polymersomes,[115, 116] and nanogels;[117] carbon based particles such as fullerenes,[118-123] nanotubes,[124-130] and graphene sheets;[130-132] and protein based scaffolds.[133] In general glyconanoparticle structure can be characterised by a NP core, of any material, capped polyvalently with glycan containing ligands (**Figure 1.7**).

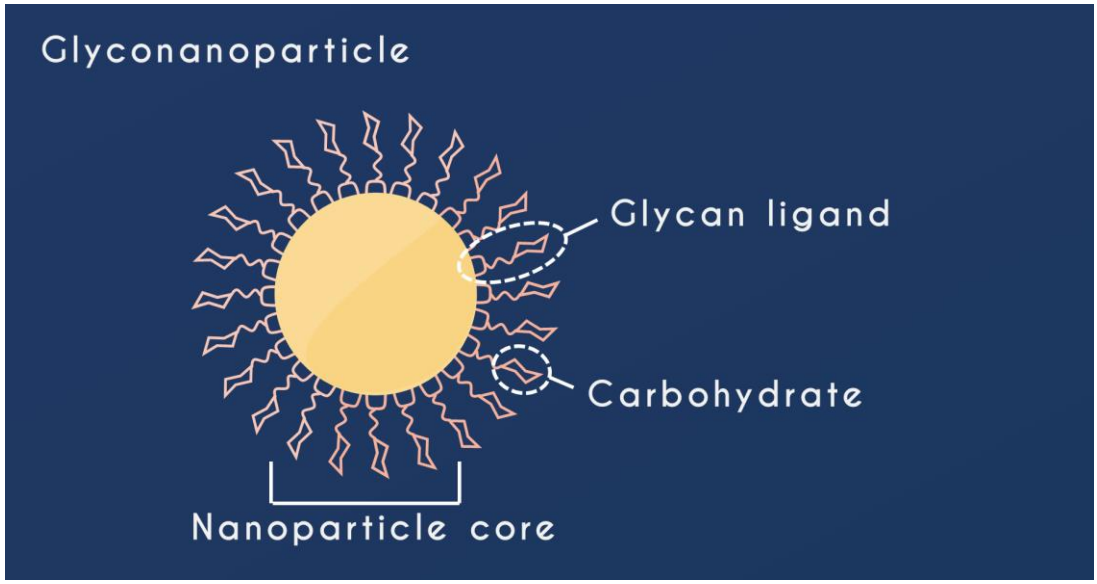

Figure 1.7 Schematic of glyconanoparticle design consisting of carbohydrate bearing glycan ligands polyvalently conjugated onto a nanoparticle core.

NPs have been hailed for their reputation of being relatively inexpensive and straightforward to synthesise which make them desirable for therapeutic application.[134] However, they also provide many other properties which make them particularly advantageous when targeting MLGIs, for example:

1. They can be readily synthesised with sizes tuneable from a few to a few hundred nanometres, matching the size of individual virus surface glycoproteins or even whole virus particles.

2. They are easy to functionalise with specific carbohydrate ligands in a polyvalent manner, allowing for strong multivalent interactions with multimeric target lectins.
3. Their properties can be fine-tuned (e.g. their degree of functionality or scaffold shape and size) in order to control their lectin binding capabilities.
4. Their functional ligands can be designed to provide additional properties to the NP including hydrophilicity, biocompatibility and immune evasion.
5. Many NPs possess unique size-dependent chemical/physical properties (*e.g.* fluorescence, superparamagnetism, photothermal behaviour, high density, non-toxicity, deformability/non-deformability, etc) which can be harnessed to provide multifunctionality by either offering additional therapeutic benefits or providing ways of monitoring their lectin binding behaviour *via* spectroscopic/imagining techniques.

The ability to combine polyvalency and tuneable properties put nanomaterials at an advantage over conventional monovalent drugs when it comes to targeting MLGIs for potential therapeutic applications. This is because many lectins have overlapping monovalent glycan specificity, thus oligo-/poly-valent glycoconjugates with spatial selectivity towards a specific target lectin are able to both increase drug potency and provide specificity against binding to other molecules in the body. In addition, the simple synthesis and potential for multifunctionality make them advantageous over other oligovalent scaffolds produced by more conventional synthetic routes, which are very practical for fine-tuning the positions of a very specific number of carbohydrate moieties but are much harder to adjust in terms of size, shape or functionality.

While a whole plethora of different glyconanoparticles have been investigated for their lectin binding abilities, each bring their own unique set of properties to the table; varying in multifunctionality, size, shape, carbohydrate type, degree of functionalisation, and material properties, e.g. flexibility/rigidity or other chemical/physical properties (**Figure 1.7**). The effects of some of these design elements, such as carbohydrate type, on lectin binding have been closely investigated for a range of lectins, others, such as the influence of shape, remain poorly understood. However, due to their highly varied glycan displays, each glycoconjugate differs in lectin targeting ability. Therefore, the development of tools for identifying the mechanisms behind the relationship between glyconanoparticle design and lectin binding behaviour are of great importance.

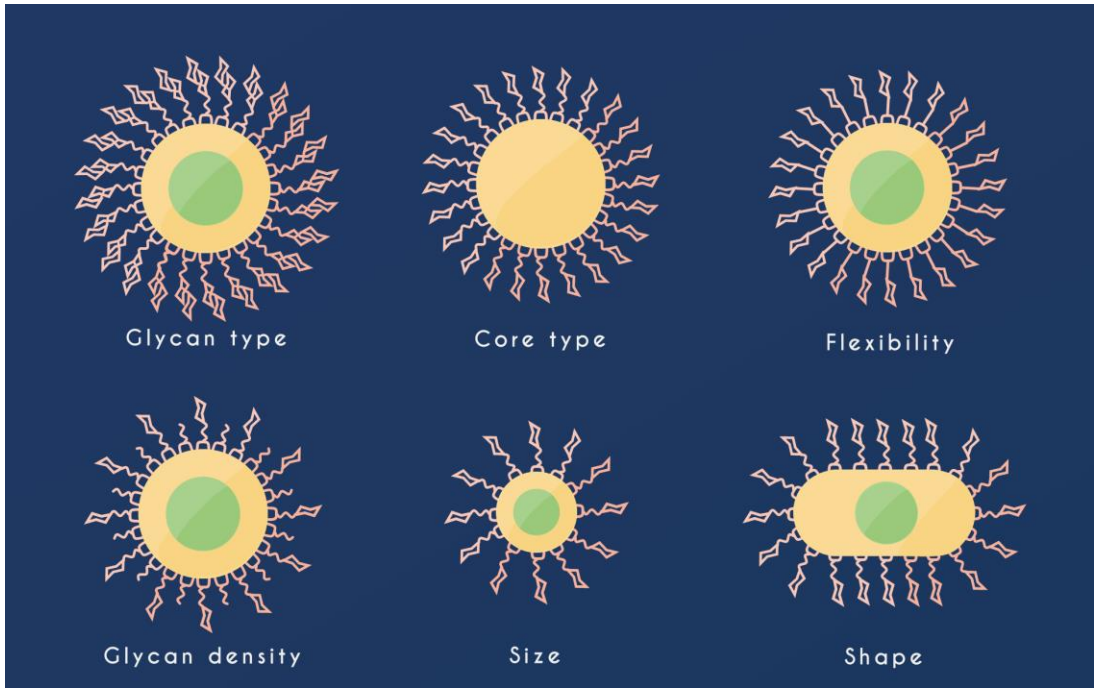

Figure 1.8 Elements of glyconanoparticle design such as the type of functional groups, the core material, the flexibility and density of glycans, and the size and shape of the glyconanoparticle can have a profound effect on its multivalent lectin binding behaviour. However, study of how many of these attributes effect the biophysical mechanisms of lectin binding remains in its infancy.

For monovalent lectin binding, the most important factor in determining the affinity is associated with the NP's surface chemistry, *i.e.* how well the surface glycans bind to individual binding sites. This has been studied in great depth as good monovalent binding affinity will have a high impact on the enhancement observed upon the formation of multiple binding interactions. In general, the most intuitive way to achieve good lectin binding this is to more closely mimic the glycoproteins observed in nature which inherently provide excellent binding capabilities and specificities towards/by target lectins. This is most easily achieved by simply decorating NPs with the key carbohydrates identified in binding. However, this can often be greatly enhanced using higher glycosylated ligands with poly-/oligo-saccharides which more closely mimic glycoproteins found in nature. For DC-SIGN, for example, studies have shown that NPs coated with ligands comprised of oligosaccharides achieve much better antiviral properties over NPs coated with mono-mannose.[104, 135] This is consistent with its preferentiality for high-mannose glycans.[27] Furthermore, binding can also show preferentiality between different oligosaccharide linkage positions; for example, α2,6-sialyllactose coated NPs have been

shown to display preferentiality over α2,3-sialyllactose analogues for the lectin hemagglutinin (HA) from human strains of the influenza A virus (IAV), however, this preferentiality is reversed for HA from avian IAV strains.[136] The main challenge of increasing NP affinity *via* the carbohydrate groups, however, is that synthesis of more complex oligosaccharides can be difficult and time consuming, thus compromises must be made between the added efficacy of enhanced glycosylation and the increased complexity of synthesis. Because of this, some research groups have resorted to deriving more complex glycans from natural glycopeptides (though this has shown mixed benefits)[106, 137] or by synthesising more potent pseudo-saccharide residues, whereby the carbohydrates have been altered to enhance the inhibition of a target virus.[112, 138, 139] In addition to carbohydrate-binding strength, other ligand properties such as ligand flexibility may also influence the binding thermodynamics by inducing additional conformational restrictions.[140] The need for understanding how aspects such as ligand structure and chemistry influence the biophysical contributions to the binding of glycoconjugates with multimeric lectins demonstrates the importance of tools which can probe the thermodynamics and kinetics of such interactions in order to inform the rational design of potent and specific glycoconjugates for targeting such MLGIs.

Apart from the vital dependency between a glyconanoparticle's surface glycans and its capacity to form strong interactions with individual lectin binding sites, its ability to establish effective multivalent enhancement comes down to whether the polyvalent glycan display is able to bridge multiple binding sites at the same time. This is determined by the distances between the surface glycans and the glyconanoparticle's shape, size and flexibility (**Figure 1.9**). For example, one of the key advantages of nanomaterials is their macromolecular scale which provides an ideal size regime compatible with target lectins, allowing the surface glycans to contact more than one of a lectin's binding sites, thus, greatly enhancing the binding affinity. In the case of larger nanoparticles, the scaffold may even be able to bridge multiple cell surface receptors, depending on the lectin density and accessibility, inducing even further multivalency and more closely mimicking the binding of the native glycoproteins on viruses. Such affinity-size relationships have been observed for many scaffolds and are often correlated with either the intra- or inter- lectin binding site distances.[93, 117, 141-145] Thus the glyconanoparticle dimensions are a key factor to consider when designing lectin targeting agents, in order to optimise the degree of multivalency to both increase the binding affinity and the specificity against binding to other molecules with similar binding sites.

However, as discussed earlier, for many lectins, such as DC-SIGN and DC-SIGNR, details such as the intra-lectin binding site distances are unknown, thus, new tools need to be developed to better inform the glyconanoparticle design. In addition, scaffold size has also been associated with increased antiviral capacity simply due to steric shielding effects which provide an additional contribution to the efficacy of glycoconjugate inhibitors as additional lectin binding sites can become sterically inaccessible to viral glycoproteins despite not being directly involved in binding.[146]

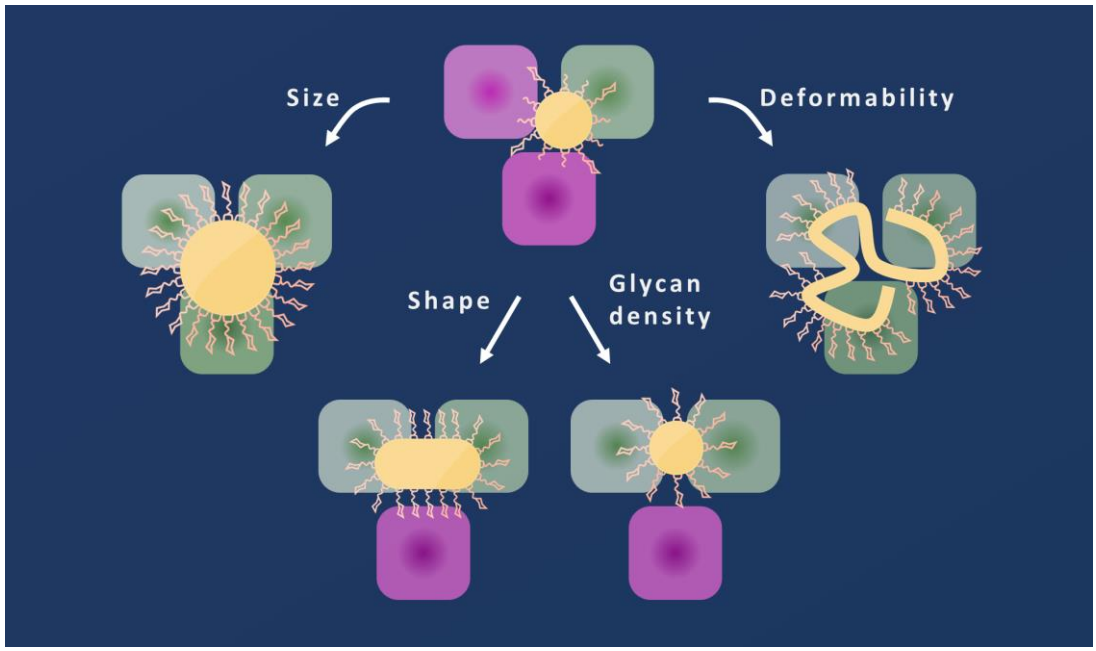

Figure 1.9 Conceptual schematic of different glycan-NPs binding to a trimeric lectin with varying valencies; demonstrating that by changing the size, shape, glycan density and/or deformability, a glycan-NP may be able to bridge a larger number of carbohydrate binding sites in order to improve its overall avidity (green: occupied binding site, pink: unoccupied binding site).

Despite the correlation between a glyconanoparticle's size and its potential to bridge multiple lectin binding sites, if the glycans are not displayed in an effective manner upon the surface of the nanoparticle, any increase in size will impart little enhancing effect other than that of steric shielding. In this respect, a scaffold's ability to bridge binding sites is very much dictated by how the glycans are decorated on the NP surface. On a fully glycosylated NP, the density of the glycan ligands is determined either by the physical spacing between terminal units (*e.g.* for polymeric or dendritic scaffolds) or else by the surface area occupied by the ligands' anchor groups (*e.g.* for inorganic scaffolds). If the glycan density is too sparse then the distance between

glycans may be too large to bridge multiple binding sites. Conversely, if the glycan density provides a good match with the lectin's binding site arrangement then this may result in a better affinity in comparison to other glycan densities.[111] The glycosylation density of scaffolds can quite easily be altered in order to optimise the binding potency, either by changing the degree of functionalisation (*i.e.* by introducing spacer ligands);[97, 104, 135, 147] the number of glycosylation sites per ligand linker;[99, 135] the curvature of the scaffold (which alters the angle between the ligands);[145] and, provided that the scaffold is curved, the length of the ligand.[16, 104, 144, 145] It is worth noting, however, that the effect of some of these features on the inter-glycan distances may be greatly dependent on the flexibility of the linkers which anchor them to the scaffolds or the flexibility of the NP scaffold itself, as flexible glycans can be more easily rearranged to adjust to the arrangement of the binding sites.

The influence of shape on the glycan presentation is twofold: 1. Changing the shape affects the curvature of the scaffold surface from which the ligands protrude, which in turn changes the angle between the glycan residues and thus the effective surface glycan density; 2. Curvature also affects the effective contact area to which the lectin can bind (for example, an increase in scaffold surface curvature gives rise to larger deviation of the glycan display away from the away from the contact plane of lectin's binding sites. This can induce a steric barrier which may prevent complete binding of all binding sites). Differences in shape have been shown to have substantial influences on the binding or bio-functionality of glyconanoparticles, (e.g. between spherical nanoparticles or nano-rods, or complex structures such as star-like scaffolds or linear polymers).[105, 148-151] However, the mechanisms behind why certain shapes impart specific avidities or biological consequences remains unclear. Moreover, often these effects can often be closely linked to additional influences, such as scaffold size or glycan density. Thus further work is needed to elucidate the direct effects of the influence of shape on MLGIs to improve the rational design of lectin targeting therapeutics.

Finally, due to the simplicity of glycosylation for many nanoscale materials, glycan-mediated lectin targeting has been at the forefront of MLGI therapeutic development. However, lectin-based nanoparticles (lectin-NPs) for targeting glycans may also act as effective therapeutics in processes such as inhibition of viral glycoprotein attachment. Such methods have been explored for the inhibition of IAV, for example, through the design of recombinant hemagglutinin nanoparticles which have shown strong low-nM affinities for sialylated SLB surfaces.[15] Similarly, then, lectin-NPs may also

prove useful in targeting other glycosylation sites associated with viral binding. For example, DC-SIGN and DC-SIGNR have both been identified as being involved in the *trans*-infection of SARS-CoV-2 into $ACE2^+$ cells *via* binding to the glycosylation sites of the virus' spike proteins.[49] Though some inhibitors against this interaction have been developed,[152] a more effective way of inhibiting the virus may be to directly target the glycosylation sites of the spike proteins using lectin-NPs. This may not only inhibit DC-SIGN and DC-SIGNR binding but could also stop the direct association between the spike protein with ACE2, thus greatly improving its *in vivo* application.

Despite excellent advances in glyconanoparticle design over the past few decades, more systematic approaches for informing how differences in scaffold structure and composition affect the mechanisms of lectin binding still need to be developed in order to provide a better informed approach to designing glycoconjugates for therapeutic applications. Due to the multifunctional component of many NPs, these questions can begin to be answered by using their unique physical properties to paint a better picture of how such glyconanoparticles interact with target lectins.

1.3.3 Spectroscopic Methods for Affinity Determination

In order to understand what aspects of glycoconjugate design determine its binding affinity, the unique physical properties of many NPs have been harnessed to provide quantitative data regarding their MLGIs in the presence of target lectins. For example, our group previously employed glycan-conjugated quantum dots (glycan-QDs) and gold NPs (glycan-GNPs) to the study of the MLGIs of DC-SIGN and DC-SIGNR.[16, 99, 104] Both of these approaches take advantage of the characteristic optical properties of these materials at the nanoscale to obtain accurate binding affinity and binding valency information *via* fluorescence spectroscopy techniques.

Quantum dots (QDs) are fluorescent nanocrystals of semiconducting material, often with one or more shell layers of different semiconducting material surrounding the fluorescent core. Unlike many other passive scaffolds employed in glycoconjugate construction, the bright and stable fluorescence of QDs enables the development of multifunctional glycan nanoscale probes which can not only display polyvalent glycan ligands controllably but also offers fluorescent applications. As a result, the glycan-QDs can be harnessed to analyse binding interactions with other fluorescent species *via* a highly robust Förster resonance energy transfer (FRET)-based mechanism. FRET

is the phenomenon whereby the energy absorbed by exciting a “donor” fluorophore is transferred to an “acceptor” fluorophore *via* a non-radiative resonant energy transfer process. This only occurs if the two fluorophores are within close proximity to each other (*i.e.* typically <10 nm apart due to the inverse 6th power relationship between the FRET efficiency to the distance between the two fluorophores) and the emission spectrum of the donor overlaps with the absorption spectrum of the acceptor. If the QD is functionalised with specific carbohydrate ligands, and the corresponding lectin is labelled with a fluorescent dye acceptor, then binding can be observed by measuring the intensity of acceptor fluorescence resulting from QD-dye FRET (**Figure 1.10**). Moreover, because the FRET signal is directly proportional to the degree of lectin·glycan-QD complexation, it is possible to derive quantitative parameters such as the K_d or binding valency (see **Section 2.2.2.1** for more detailed theory).

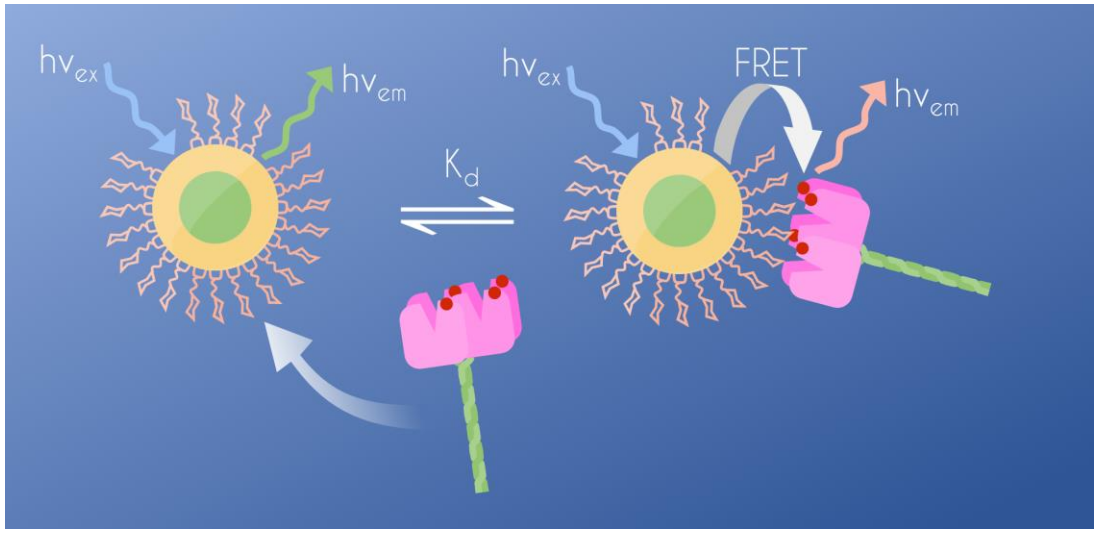

Figure 1.10 Schematic depicting the FRET process between an excited glycan-QD and dye-labelled DC-SIGN.

Compared with other methods, the use of glycan-QD FRET assays to measure LGIs has many advantages, for example: 1. FRET provides quicker data collection in comparison to more technical approaches for affinity measurement, such as ITC; 2. Measuring the dual acceptor and donor fluorescence signal allows for a ratiometric readout which is effectively insensitive to instrument noise and signal fluctuation due to self-calibration, allowing for highly sensitive and robust measurements; 3. Because any non-bound species are likely to be too far apart to participate in FRET, no background signal is produced, allowing for binding interactions to be performed in homogenous solutions without the need of separation; 4. Unlike

surface-phase assays, such as QCM or SPR, QD-FRET does not require the need of surface immobilisation, allowing for more authentic analysis of their independent binding properties. In addition to these benefits, QDs are particularly advantageous for this approach as their broad absorption and narrow symmetric emission allow for a wide choice of excitation wavelengths to minimise the direct excitation of the acceptor dye. Furthermore, due to the high fluorescence signal of QDs, QD-FRET allows for very sensitive detection of binding, requiring only small amounts of material.

As well as QDs, gold NPs (GNPs) offer a similar, highly sensitive, solution based assay for measuring MLGIs. GNPs of >2 nm do not display fluorescence, thus cannot be applied in the same way as QDs. However, gold can still absorb energy from a donor fluorophore in close proximity (the efficiency of which is dependent on the inverse 4^{th} power of the donor-acceptor distance) though a similar energy transfer process *via* the resonance of its surface conduction electrons (*i.e.* SPR). Therefore, by labelling a target lectin with a fluorescent dye donor, the proportion of bound protein can be directly measured by the reduction of dye fluorescence caused by quenching from any bound glycan-GNP. Thus, glycan-GNPs also provide a ratiometric readout strategy for measuring binding (assuming that the measurements before and after incubation of the lectin with glycan-GNP are measured under identical conditions), where the same details can be extracted as with glycan-QDs. Glycan-GNPs also offer other advantages such as gold's excellent biocompatibility compared to the toxicity of QDs, which provides a multifunctional element to glycan-GNPs, potentially allowing for *in vivo* application. Additionally, GNPs provide a much simpler production method compared QDs with additional shell layers.

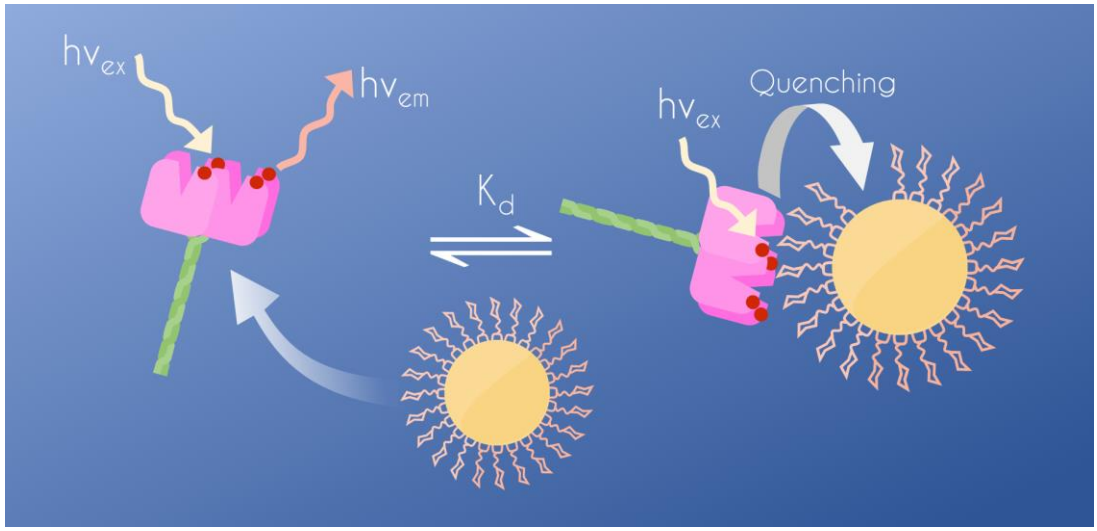

Figure 1.11 Schematic depicting the quenching process between excited dye-labelled DC-SIGN and glycan-GNP.

These glycan-NP based binding affinity readout strategies were pioneered by our group to study the MLGIs of DC-SIGN and DC-SIGNR. Here, the extracellular domains of the two proteins were site-specifically labelled *via* a cysteine site mutation which was positioned close enough to the binding site to permit reasonable energy transfer to a bound glycan-NP without being structurally or chemically essential for binding to occur.[16, 99] Cysteine was chosen due to the fact that all other cysteine residues in the protein are occupied in disulphide bonding, thus this mutation provides a free thiol which can undergo efficient thiol-ene coupling, thus allowing for labelling with maleimide-bearing dyes, such as maleimide-Atto594 which provides a good spectral overlap with CdSe/ZnS QDs with fluorescence spectra at 550 nm (QD_{550nm}).

In order to functionalise QDs or GNPs for this application, thiolate-oligo(ethylene glycol)-glycan ligands were used due to their suitable anchor-linker-glycan design, which provides stable sulphur-coordination to the metal surface; excellent water solubility and resistance against non-specific adsorption from the ethylene glycol linker; and lectin-binding functionality from the terminal glycan.[16, 99, 104] In addition, yne-azide “click” chemistry was employed between the glycan and linker groups, enabling a simple way of creating a library of ligands with different linker lengths or carbohydrate groups (**Figure 1.12**).

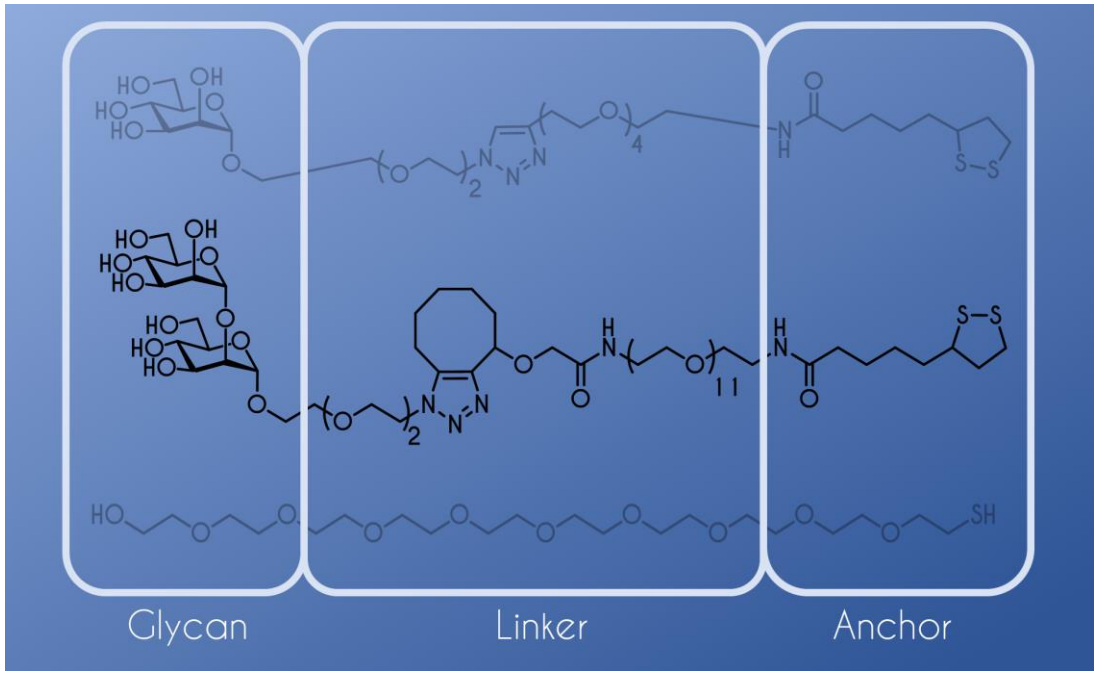

Figure 1.12 Multifunctional anchor-linker-glycan ligand design (top: lipoic acid-tetra(ethylene glycol)-mannose, or LA-EG$_4$-Man, with an yne-azide linkage; middle: lipoic acid-undeca(ethylene glycol)-mannose-α-1,2-mannose, or LA-EG$_{11}$-DiMan, with a cycoloctyne-azide linkage; bottom: thiol-hepta(ethylene glycol), or HS-EG$_7$-OH).

Glycosylated QD$_{550nm}$ were synthesised using ligands containing a lipoic acid anchor, an undeca(ethylene glycol) linker and a mannose-α1,2-mannose (DiMan) glycan (LA-EG$_{11}$-DiMan), yielding QD-EG$_{11}$-DiMan glyconanoparticle probes with ~300 ligands per NP and a hydrodynamic diameter (D_h) of ~9 nm.[104] These probes were able to successfully obtain key quantitative details regarding their MGLI binding affinities in solution *via* ratiometric QD-FRET readout. QD-EG$_{11}$-DiMan glycoconjugates were observed to provide strong binding affinities, with K_d values as low as 2.1 nM for DC-SIGN and 633 nM for DC-SIGNR. QD-EG$_{11}$-DiMan·DC-SIGN complexation displayed an enhancement factor, β, of ~430,000 compared to the binding of monovalent DiMan with an individual DC-SIGN CRD (K_d = 0.9 mM).[153] This emphasises the power of multivalency in biology to generate significantly stronger binding interactions from the combined effect of multiple, much weaker binding partners. Moreover, DC-SIGN binding results in a 300-fold stronger K_d in comparison DC-SIGNR, suggesting that, though their constituent monovalent mannose binding sites are near-identical, DC-SIGN may display a more optimal tetrameric CRD arrangement for the specific glycan display of these particular glycan-QD scaffolds; additionally emphasising the power of multivalency in enhancing the specificity of MLGIs for more specific biological response. Considering that the HIV surface gp120

glycan spike is densely coated with high-mannose glycans and is of comparable size to these scaffolds (~10 nm) this may provide a plausible explanation as to why DC-SIGN is more effective in *trans*-infecting HIV-1 compared to DC-SIGNR.[72, 154]

Mannosylated GNPs were also used by the group to measure DC-SIGN/R binding by capping GNPs with LA-EG_2-DiMan, yielding GNP-EG_2-DiMan with ~700 ligands per GNP and a D_h of ~11 nm.[99] Quenching assays revealed similar trends in the binding affinities to that of QD-EG_{11}-DiMan, with K_d values of ~3.9 nM and ~152 nM for GNP-EG_2-DiMan binding with DC-SIGN and DC-SIGNR, respectively.[99] The weaker affinity of DC-SIGNR compared to DC-SIGN is apparent here, however, the difference is less pronounced using this glycan-GNP method compared to the glycan-QDs. Though this may be due to the various subtle differences in ligand structure, glycan-density or scaffold size, which can lead to more or less optimal binding depending on how well the glycan display matches the CRD arrangement of the proteins, it may also simply be a result of the inner filter effect at high concentrations (*i.e.* reabsorption of emitted light from the GNPs), thus demonstrating a limitation of this approach.

These glycan-NP probes have been employed to investigate the effects of varying other scaffold properties, such as glycan type, linker length, total ligand density or individual ligand glycan valency, which all exhibit different effects on MLGI binding for DC-SIGN and DC-SIGNR. Using Man instead of DiMan, for example, has been observed to greatly reduce the DC-SIGN binding affinity for both QD-EG_{11}-Man (~714 nM) and GNP-EG_2-Man (~33.1 nM).[104] This is still a large enhancement in comparison to monovalent binding between Man with an individual DC-SIGN CRD (3.5 mM), *i.e.* ~5000-100000-fold, however this is between 2-90-fold less than the enhancement observed in DiMan. This affinity amplification has been observed for other lectins, whereby stronger monovalent affinity interactions tend to also lead to a higher enhancement factors with multivalency compared to weaker monovalent interactions.[155] Altering the length of the linker has also been explored. For example, using QD-EG_3-DiMan is found to greatly enhance the binding affinities of both DC-SIGN (~0.61 nM) and DC-SIGNR (~62 nM) compared to the longer QD-EG_{11}-DiMan.[104] This is potentially due to the increased in inter-glycan spacing afforded by the longer ligand length of EG_{11}, which may be less optimal for the CRD arrangements of the two proteins; alternatively, the additional flexibility of EG_{11} may impose unfavourable variations in the thermodynamics (entropy penalty) of the binding interaction.

Apart from varying the ligand length, the inter-glycan distance can also be altered by introducing inert “spacer” ligands to dilute the concentration of the glycan-ligands on the NP surface, thereby extending the separation distances between the terminal glycans. This technique was employed to investigate the effect of diluting DHLA-EG_3-Man coated QDs with inert zwitterionic LA-spacer ligands on the binding of DC-SIGN. Using glycan surface coverages of 100%, 50%, and 25%, it was observed that the relative protein binding efficiency per glycan decreased significantly with decreasing glycan density.[104] This demonstrated that the smaller inter-glycan distances of these scaffolds were more optimal for the tetrameric CRD arrangement of DC-SIGN, though the effect of higher order mannosylated ligands such as those containing DiMan have yet to be explored. Because of this result, higher mannose densities were explored by further mannosylating the NP surface using dendritic ligand linkers which support three glycans per ligand. Here, LA-(EG_2-DiMan)$_3$ ligands were synthesised by “clicking” azide-functionalised DiMan with LA-linkers displaying three alkyne groups.[99] Interestingly, GNPs capped with these LA-(EG-DiMan)$_3$ ligands displayed 3.6-fold stronger binding with DC-SIGNR (K_d = 42 nM) in comparison to the corresponding GNP-EG_2-DiMan scaffolds (K_d = 152 nM), however, GNP-(EG-DiMan)$_3$ displayed a negligible increase in affinity upon binding to DC-SIGN (K_d = 3.9 nM; *vs* 3.6 nM for GNP-EG_2-DiMan).

In addition to these methods, other spectroscopic approaches for lectin·glycan-NP characterisation have been developed. For example, another benefit of these NPs (such as those made from gold or silver) is their absorption spectra and observable broadening/red shift upon agglomeration. This property has been exploited by various groups to measure the crosslinking of lectins such as Con A, RCA_{120} or PNA in the presence of glycan-gold or silver NPs by simply measuring the shift of absorbance with varying lectin concentration or recovery of the native NP absorbance in the presence of a carbohydrate competitor.[147, 156-168] Alternatively, if the protein is fluorescently labelled then the binding of glyconanoparticles can simply be measured by centrifuging lectin·glycan-NP agglomerates out of solution and measuring the decay of fluorescence as the concentration of soluble protein decreases, or recovery of fluorescence in the presence of a competitor. This technique was employed by the Yan group to quantify the binding of labelled Con A with mannosylated GNPs (GNP-Man) of various sizes, ligand lengths and glycan densities by competing against the binding of free mannose.[100] After incubation, the mixture was then centrifuged (thus removing GNP-bound lectin) and the concentration of remaining lectin was measured *via* its resulting

fluorescence intensity. By measuring the proportion of unbound protein over varying lectin concentrations, a value for the K_d of GNP-Man could be obtained relative to that of monovalent Man. This technique was able to reveal that crosslinking was less favourable at larger GNP sizes (*i.e.* K_d of ~25 nM for 30 nm GNPs *vs* 3-4 nM for 7-22 nm GNPs), possibly due to steric hindrance of lectins binding to the glycans of larger particles. Additionally, the binding became stronger with increased linker length which was also attributed to a decreased hindrance for binding as the glycans were protruded away from the GNP surface with larger inter-glycan separation distances. In addition to this approach, alternative techniques have also been investigated where the fluorescent functionality of the NP can be directly harnessed to monitor agglomeration in a similar way. For example, glycan-conjugated materials such as gold nanoclusters (*i.e.* gold NPs of $\lesssim$ 3 nm) and carbon nanotubes (CNTs) also display fluorescence, the latter in the near infra-red (NIR). Thus these materials can be directly used to analyse the crosslinking of lectins by measuring the fluorescence intensity of the supernatant before and after addition of lectin with or without a competitor,[124, 169] or even to directly measure their abundance on immobilised lectin surfaces.[170] Alternatively, materials such as graphene have been exploited for their excellent quenching behaviour. For example, fluorescent aminopyrene molecules grafted with maltose adsorb to the graphene surface and thus display quenched fluorescence, however, upon incubation with Con A, the lectin is able to bind to the maltose and release the fluorophore from the surface, thus resulting in fluorescence recovery which can be used to monitor the amount of binding.[171]

1.3.4 Electron Microscopy and Dynamic Light Scattering

In addition to the spectroscopic functionality, these nanoparticles also allow for simple complexation determination *via* dynamic light scattering (DLS). DLS determines the hydrodynamic sizes of species by measuring the size-dependent intensity fluctuations of monochromic light caused by the Brownian motion of particles in solution, whereby larger particles move with lower velocity and thus scatter light over longer periods of time in comparison to smaller particles. Additionally, for high density NPs, such as metal NPs, electron microscopy (EM) provides a powerful way to image glyconanoparticle complexes, due to the strong absorption of electrons which provides much higher contrast in comparison to organic species. These techniques lead to an additional level of readout, allowing us to see how glycan-NPs organise

themselves in solution in the presence of lectins, which provides useful information about the *binding modes* of such MLGIs and consequently the lectin binding site orientations. Inorganic NP scaffolds are also particularly convenient for this application due to their solid and non-deformable structures.

Scanning or transition EM and DLS have been used widely to monitor proteins such as Con A or PNA for a range of inorganic NPs to provide a means of confirming agglomeration.[147, 156, 157, 159, 160, 164, 165, 167, 168, 172, 173] DLS has even been exploited to obtain K_{d} values by monitoring the dose dependent increase in size with lectin concentration.[174] In addition, DLS and scanning transition EM (STEM) were employed by our group to further analyse the complexation of DC-SIGN and DC-SIGNR with glycan-QDs and glycan-GNPs in order to probe a structural rationale behind their differences in binding affinity.[99, 104] Both glycan-NPs displayed mean hydrodynamic diameters (D_{h}) comparable to the size of the two proteins (D_{h} ~14, 12, 9 and 11 nm for DC-SIGN, DC-SIGNR, QD-EG$_{11}$-DiMan and GNP-EG$_2$-DiMan, respectively). These sizes are easily detectable by DLS, allowing straightforward measurement of the variation of D_{h} with lectin·glycan-QD assembly formation. For example, after incubation of QD-EG$_{11}$-DiMan with DC-SIGN at a protein to QD molar ratio (PQR) of 12.5, a single dominant species with a mean D_h of ~42 nm was observed (**Figure 1.13B**), which is consistent with that expected for a single glycan-QD coated with a monolayer of DC-SIGN protein molecules.[104] In contrast, incubation of QD-EG$_{11}$-DiMan with DC-SIGNR at the same PQR showed much larger and broader hydrodynamic size distributions, with D_{h} values of ~124 and ~205 nm (**Figure 1.13C)**. Such assemblies are significantly larger than those expected for a single glycan-QD coated by a single layer of DC-SIGNR proteins, thus indicating the formation of larger assemblies containing a larger number of binding partners, suggesting that DC-SIGNR was able to crosslink between multiple glycan-QDs at a time. This observation can be backed up by STEM imaging of binding induced NP assemblies: by plunge-freezing and subsequent drying prior to measurement, which allows the NPs to be directly imaged in their native dispersion states without the danger of inducing possible artefacts during the sample drying process.[175] For QD-EG$_{11}$-DiMan in the absence of any protein, the corresponding STEM images show that, in addition to isolated particles, the QDs also shows some degree of adhesion in the binding buffer (**Figure 1.13D**), also observed in the D_h size distributions, likely due to weak interactions between the QD surface DiMan ligands and the Ca^{2+} ions in the buffer. However, upon incubation of DC-SIGN, any

clustered QDs are completely disassembled and only isolated individual QD particles are formed (**Figure 1.13E**).[104] In sharp contrast, the addition of DC-SIGNR leads to the formation of large and extensively aggregated QD assemblies. These organisations of QDs are fully consistent with the hydrodynamic size observations for the two proteins, demonstrating that DC-SIGN bind to single glycan-QDs at a time, whereas DC-SIGNR bridges multiple glycan-QDs resulting in an extensive crosslinked assembly (**Figure 1.13F**). These observations were complimented by similar results for GNP-EG_2-DiMan complexation with DC-SIGN/R,[99] demonstrating the reproducibility of these two approaches and establishing them both as valid methods for determining the binding nature of multivalent lectins to glycan-NPs.

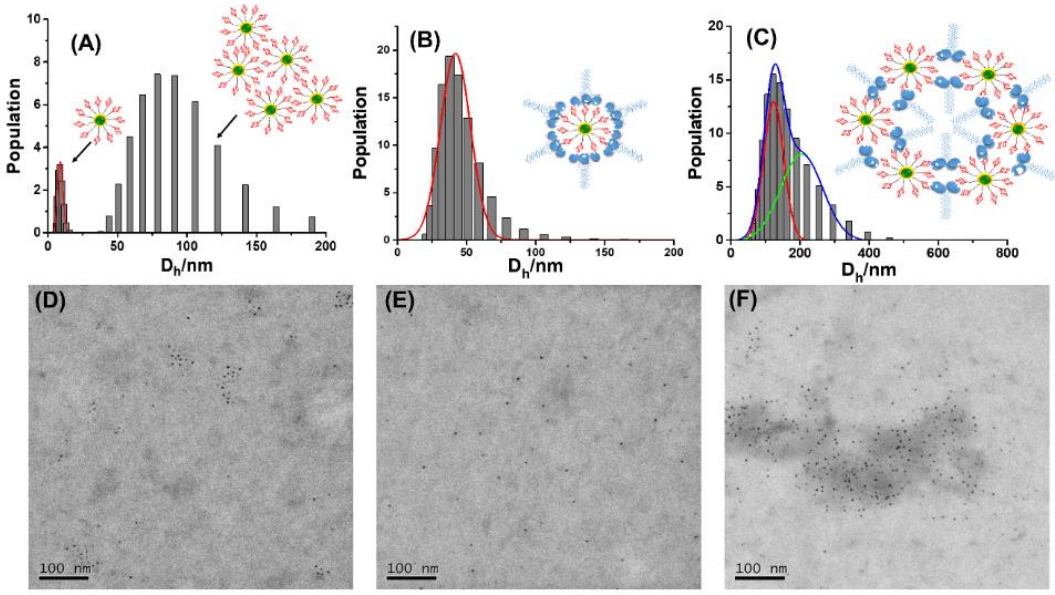

Figure 1.13 Hydrodynamic diameter (D_h) size distribution histograms of (A) QD-EG_{11}-DiMan in HEPES binding buffer; (B) QD-EG_{11}-DiMan incubated with DC-SIGN at a PQR of 12.5; (C) QD-EG_{11}-DiMan incubated with DC-SIGNR at a PQR of 12.5. (D, E, F) show the corresponding STEM images of the cryogenic prepared QD and QD-protein assembly samples. (Figure copied from Reference [104]).

These glycan-NP based multimodal readout approaches therefore distinguish two different binding behaviours for DC-SIGN and DC-SIGNR with NPs of this size: DC-SIGN displays strong binding affinity interactions with individual glycan-NPs, whereas DC-SIGNR forms crosslinking interactions which result in a weaker overall binding affinity. These observations have led to the conclusion that DC-SIGN's four binding sites are oriented in the same direction, allowing it to bind tetravalently to the glycans of individual

glycan-NPs, whereas DC-SIGNR's four binding sites are positioned in an outward-facing orientation, thus the glycan display of a single glycan-NP is unable to bridge all four CRDs which, therefore, forces it to bis-divalently bind with two glycan-NPs to maximise binding multivalency (**Figure 1.14**).

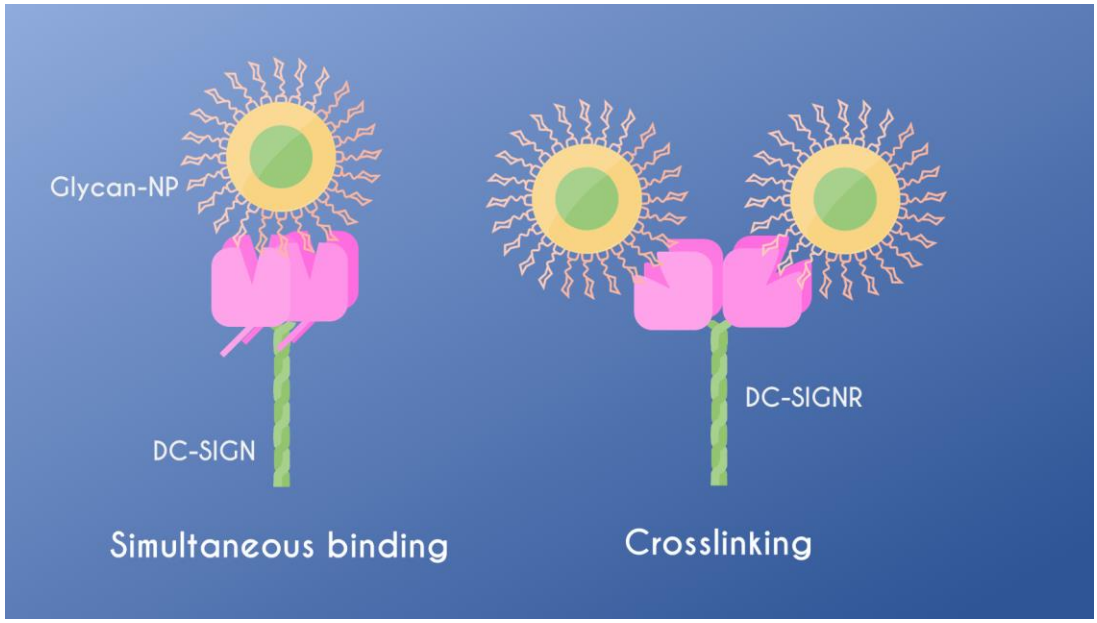

Figure 1.14 Schematic of the binding modes of DC-SIGN (*i.e.* simultaneous binding of all CRDs to an individual NP) and DC-SIGNR (*i.e.* crosslinking with multiple NPs) upon binding to the polyvalent glycan-NPs highlighted in References [99, 104].

These examples demonstrate the power of glycan-NP in establishing crucial details of MGLIs such as binding affinity and binding mode, and show how they can be used to probe differences in NP design for informing the design of more potent and specific polyvalent targeting agents for MGLIs. This has progressed the understanding of the MLGIs of DC-SIGN and DC-SIGNR, however, many questions remain, for example: what structural bases dictate their differences in tetrameric CRD arrangement; what thermodynamic and kinetic contributions dictate their glycan binding affinities; and can fundamental design rules be established to inform the design of glyconanoparticles for specifically targeting lectins with subtle differences in multivalent specificity, such as these? Thus, there is still yet to be explored potential for these glycan-NPs for analysing the biophysical and structural bases behind MLGIs in a way that other techniques cannot.

As with all techniques, however, these glycan-NP probes also have their limitations. Though they can infer key structural information regarding lectin·glycan-NP binding, they aren't able to resolve more specific structural

details regarding multimeric lectins, such as the inter-binding-site distances or the structural rearrangements under binding with glycans. Moreover, though these glycan-NP probes provide a powerful approach for analysing MLGIs in their native states in solution, such assays will hold discrepancies with the true surface chemistry experienced on cell surfaces. Thus, the use of other state-of-the-art techniques which allow for more detailed description of the lectin molecules (such as smFRET or EPR) or permit surface phase analysis (such as QCM or SPR), need to be used in conjunction with the methods highlighted above in order to provide a more comprehensive picture of NP-based MLGIs (see **Section 1.3.5-1.3.6**).

1.3.5 MLGIs on Surfaces

In order to fully understand the binding of membrane proteins with glycans at the cellular level, their MLGIs must be investigated on membrane surfaces. For the measurement of lectin-glycan affinity at cell surfaces, quantitative parameters such as efficacy of viral infection can be obtained by monitoring lectin$^+$ cell infection, or more qualitative information may be obtained by techniques such as confocal microscopy. For example, glycan-NP-mediated inhibition of the infection of DC-SIGN$^+$ and DC-SIGNR$^+$ 293T cells by luciferase gene-encoding murine leukemia virus vector particles bearing EBOV surface glycoproteins (MLV-EBOV-GP) was studied previously by measuring the luciferase activity of infected cells under increasing concentrations of glycan-NP inhibitor.[99, 104] In the absence of inhibitor, the MLV-EBOV-GP pseudo-virus is able to transmit the luciferase gene into the cells *via* the specific binding of DC-SIGN or DC-SIGNR, which results in the subsequent expression of the luciferase protein, thus allowing for the quantitative analysis of luciferase activity by measuring the oxidation of luciferin *via* bioluminescence spectroscopy. In the presence of glycan-NPs, this luciferase activity has been shown to be reduced by 50% at concentrations (IC_{50}) as low as 0.7 nM for QD-EG_3-DiMan and even 95 pM for GNP-EG_2-DiMan.[99, 104] Thus showing the power of glycan-NPs as very potent inhibitors against virus infection and infers the general correlation between the binding affinities obtained by FRET/quenching assays with real-world viral inhibition applications. Despite this, however, there are some discrepancies between trends in solution based affinity and the viral inhibition potency between glycan-NP scaffolds. For example GNP-$(EG\text{-}DiMan)_3$ was found to have a 3.6-fold stronger affinity (lower K_d) than GNP-EG_2-DiMan in binding to DC-SIGNR in solution, however, it gave a 4-fold weaker inhibition (higher IC_{50}

value) against DC-SIGNR mediated MLV-EBOV-GP infection. Thus, some obvious differences between solution phase and cell surface-phase binding are found to occur.

Unfortunately, however, studies such as these only demonstrate the inhibition potency of glyconanoparticles but cannot provide the high level of detail available in solution based assays, such as the binding mode or K_{d}. Thus, it is difficult to obtain specific information about lectin binding mechanisms within the cellular context. Because of this, alternative techniques have been employed to mimic the surface-phase binding systems which are present on cell surfaces. This has included techniques such as SPR and quartz crystal microbalance (QCM) which provide a platform for measuring the binding of glycans onto artificial lectin coated surfaces or vice versa. SPR techniques measure absorption of light from a beam interacting with a conducting metal (often gold) due to the resonance of the conduction electrons, the angle of which is dependent on changes in the refractive index at the solution interface and any changes in mass.[176] QCM directly measures the effect of mass on the resonance frequency of a piezoelectric quartz crystal (which often also uses a layer of gold as a support for functionalisation). Both systems can therefore measure the real-time binding of molecules by flowing material over the sensor surface and monitoring the rate of change of adsorbed analyte as it either associates or dissociates over time. SPR and QCM have therefore been used to measure the binding of many lectins and glycan-NPs either by functionalising sensor surfaces with immobilised lectins and monitoring the association of glycan-NP, or by monitoring the dissociation of freely diffusing lectins from immobilised glycan surfaces in the presence of a glycan-NP competitor. This has allowed for the acquisition of key quantitative binding parameters such as association/dissociation rates and binding strengths, as well as quantitative descriptors of selectivity in the presence of competitors for a range of glyconanoparticles and lectins, including DC-SIGN and DC-SIGNR.[114, 122, 133, 159, 170, 177-186] However, such techniques are limited in the amount of information they can convey, for example, though both techniques can provide some of the key biophysical parameters required to quantify binding, it can be difficult to extrapolate structural details such as binding modes, which are much more easily obtainable by solution-based techniques. Such information is essential to provide understanding regarding the behaviour of multivalent lectin-glycan binding upon surfaces. In addition, using immobilised surfaces can often result in inaccurate binding descriptors, as, by fixing glycans or lectins in place, the system can neither accurately depict the kinetics of fluid cell membranes nor

the collision frequencies or lectin domain formations observed in nature.[33, 187] One approach that could overcome this, would be to support lectins upon artificial membrane surfaces using lipid bilayers which, though do not fully represent the complex nature of cell membranes, do allow for the rearrangement of surface-phase proteins in ways that better mimic the lateral mobility of their native states. Such supported lipid bilayers (SLBs) can be formed by the adsorption and subsequent rupture of nanoscale bilayer liposomes (known as small unilameller vesicles; SUVs) onto smooth, hydrophilic Si-based surfaces such as silica, SiN_3, or mica (though other metal-based hydrophilic surfaces, such as TiO_2, have shown varying degrees success)[188-190]. Conveniently, due to the use of silica-based sensors, both the formation of SLBs and monitoring of subsequent lectin-glycan interactions could be facilitated by QCM-based techniques (see **Section 2.2.1.2** for more detailed theory).

This technique has been harnessed to analyse the kinetics of many ligand-receptor interactions on artificial membrane surfaces, including LGIs such as influenza A hemagglutinin (HA) with sialic acid-coated SLBs.[15] However, apart from this, limited attempt has been made to develop SLBs for the surface-phase analysis of other key membrane lectins such as DC-SIGN and DC-SIGNR. Thus, such analysis could be used in conjugation with the glycan-NP techniques mentioned previously to provide a more detailed story of the binding of lectins to glycan-NPs both in solution and on lipid bilayer surfaces. This would allow us to assess whether there are any mechanistic similarities or discrepancies between translating from solution-phase glycan-NP binding to that on cell membranes. Such information will provide a better picture of how to design more powerful tools for monitoring MLGIs that can inform therapeutic glycoconjugate design at the cellular level.

1.3.6 Spectroscopic Rulers

Although glycan-NPs have been able to obtain key information behind the structural characteristics of lectin-glycan interaction,s obtaining detailed measurement of the glycan binding site arrangements of lectins with unresolved multimeric structures still poses a challenge. For example, though glycan-NP probes been able to resolve the binding site orientation of DC-SIGN and DC-SIGNR, the actual relative positions of these regions in space still remain unsolved. Such details would allow for a more rational design of multivalent glycan-based targeting agents with specific glycan arrangements which perfectly match the inter-binding site distances of these

lectins, allowing a much higher degree of target specificity. Tetrameric models for DC-SIGN and DC-SIGNR have been suggested in the past, based on the crystal structures of individual monomers or dimers, which agree well with the observed MLGI binding activities observed for DC-SIGN and DC-SIGNR MLGIs.[23, 191] However, these are still only theoretical, thus more empirical techniques need to be developed to identify these CRD arrangements in order to better inform more specific glyconanoparticle design. One method for achieving this is to use so-called “spectroscopic rulers”, which are able to empirically measure low-nanometre distances by analysing the relative positions between two or more molecular probes. This approach include techniques such as FRET and electron paramagnetic resonance (EPR) spectroscopy.

In the case of FRET spectroscopy, the distances between a FRET pair of donor and acceptor fluorophores can be measured using the inverse 6th power relationship between distance and the FRET efficiency (see **Equation 2.3**). This therefore provides a very useful way of measuring the distances between two discrete positions on a protein, or any other material, and has been applied in a range of biological scenarios.[192, 193] However, when it comes to systems which may return multiple distances, conventional FRET spectroscopy techniques are greatly limited by the fact that exciting a large volume of sample will always return a fluorescence spectrum corresponding to the average conditions of all the FRET pairs measured. For example, if, the CRDs of DC-SIGN were labelled with a FRET pair of dye molecules, conventional FRET spectroscopy would simply return the average distance between all of the 6 inter-CRD distances (**Figure 1.15 right**). To overcome this problem, a technique known as single-molecule FRET (smFRET) spectroscopy has been developed to measure the fluorescence signal of individual molecules by focusing a beam of light at the excitation wavelength into a small volume of dilute sample, ensuring that no more than one molecule is present in the confocal volume at any time (**Figure 1.15 left**). By measuring the donor and acceptor emissions of individual molecules over an extended period of time, a distribution of FRET events with each FRET efficiency observed can be extracted. Thus, if the FRET pair occupies more than one distance, then multiple distributions of FRET efficiencies will be observed, containing a corresponding number of peaks. This can then be used to obtain the corresponding distances for the different configurations of the system. This approach has been exploited for various biological applications, including to probe membrane proteins to discern their conformational changes or particle interactions,[194] but it is yet to be applied to measure the inter-binding

site arrangement of multimeric lectins. This is because, due to the multimeric nature of many lectins, labelling CRDs with acceptor and donor dye molecules, even *via* specific site mutations, inevitably leads to a distribution of different numbers of dyes in different configurations on the lectin. This is detrimental to FRET readouts as a disproportionate number of acceptor molecules would lead to artificially high FRET efficiencies (see **Equation 2.6**), whereas a disproportionate number of donor molecules may lead to quenching effects, causing artificially low FRET efficiencies. For such techniques to work, only ≤2 dye molecules per lectin can be used, however, such configurations are almost impossible to purify. Therefore, new techniques for labelling 2 out of 4 CRDs must be developed in order to probe multiple distances within the same multimeric lectin.

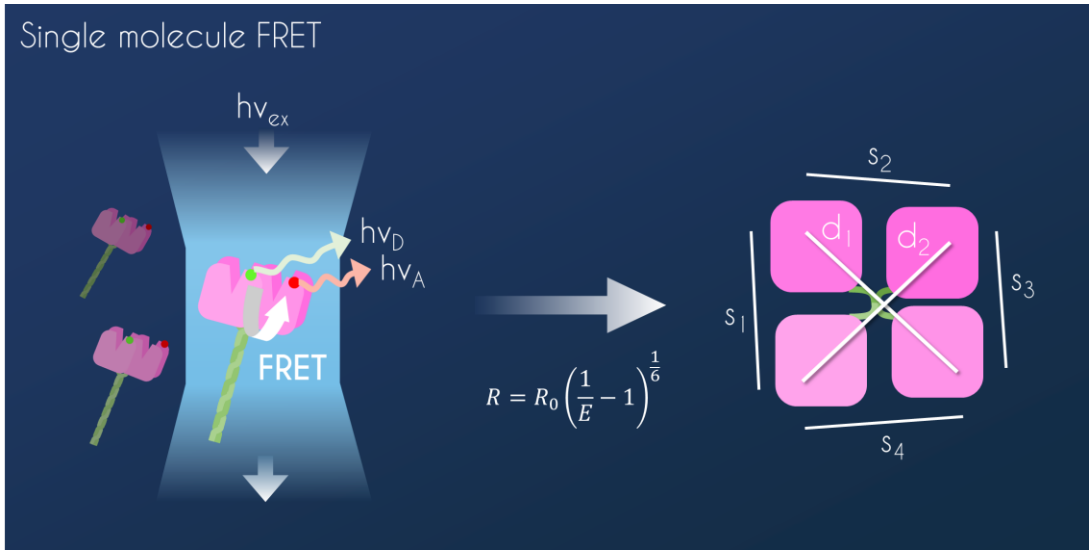

Figure 1.15 Schematic depiction of a single molecule FRET approach to measure inter-binding site distances, demonstrating the passing of single protein molecules through the confocal volume of a focus excitation beam $h\nu_{ex}$ which excites the donor fluorophore (emission $h\nu_D$) which transfers energy to the acceptor fluorophore (emission $h\nu_A$) with a FRET efficiency of E (left). By then obtaining E for all inter-binding site distances, for a tetrameric lectin, values for the 4 side and 2 diagonal distances may be obtained.

Unlike FRET, EPR is not limited by the use of multiple different labelling molecules, and thus does not have the same purification challenges. EPR measures the effect of local environment on the excitation frequency of unpaired electron spins under an external magnetic field (see **Section 2.2.1.3** for more detailed theory). This technique is commonly used to monitor the structural dynamics of proteins over time or conformational changes upon binding with other proteins or ligands, by labelling the protein under

investigation with site directed spin labels (SDSLs).[195, 196] In addition to this, under strict cryogenic environments, EPR techniques such as pulse electron-electron double resonance (PELDOR; or DEER) spectroscopy can be used to measure the dipole coupling between multiple SDSLs (which is inversely proportional to the separation distance between the SDSLs to the 3rd power, and thus can extend much further than FRET).[197] This allows EPR to be used to measure distances between two or more points on a protein. However, unlike smFRET, PELDOR requires much larger sample sizes and measurements must be taken in a frozen state, where temperature effects on lectin conformation could affect the data accuracy compared to normal physiological conditions.

Techniques such as smFRET or EPR have not been applied to the measurement of multimeric lectin binding domains. Therefore, harnessing these methods to resolve currently unknown lectin multimeric arrangements, such as those of DC-SIGN and DC-SIGNR, is of great value. Such information would be able to provide clear structural markers for developing glycoconjugates with matching inter-glycan distances for potent, specific targeting of specific MLGIs for potential therapeutic development, addressing a key challenge currently facing this important research area by avoiding unwanted interactions with other similar proteins in the body.

1.4 Summary and Research Aims

Lectin research has come far within the past few decades, with new insights into their glycan specificities, affinities, structural characteristics and biological pathways. The rise of nanotechnology has provided an ideal platform to facilitate this research, offering vital new methods for probing lectin-glycan interactions and effective lectin binding partners which can be employed to treat disease. Despite these advances, many of the structural characteristics and binding mechanisms for many multimeric cell membrane lectins remain unresolvable by conventional means and therefore require alternative methods of elucidation. Glycan-QDs have proven to be a powerful technique for probing MLGIs through a multimodal readout strategy comprising of FRET, STEM and hydrodynamic size analyses. This technique has been able to reveal the binding mechanisms of a pair of closely related tetrameric lectins, DC-SIGN and DC-SIGNR, demonstrating that subtle differences in their CRD orientations are responsible for their distinct binding affinities and modes with glycan-QDs and in turn affects the glycan-QD inhibition of DC-SIGN/R

mediated viral infections. This thesis therefore builds upon the previous work, and focuses on developing new techniques using glyco-nanotools to evaluate:

1. The thermodynamic and kinetic contributions of lectin·glycan-QD binding, through temperature and time dependent FRET studies, which reveal the energetic and dynamic rationale behind the different specificities of DC-SIGN and DC-SIGNR with glycan-QDs; **Chapter 3**.
2. Structural contributions to the specificity of these proteins, *via* the combination of mutagenesis and glycan-QD-FRET techniques; **Chapter 3**.
3. Scaffold design elements, such as scaffold shape, which contribute to the observed differences in multivalent lectin-glycan-QD binding specificity and affinity; **Chapter 4**.
4. The differences between solution- and surface-phase lectin·glycan-NP binding using SLBs as an intermediate lens for understanding the mechanistic differences between glycan-NP FRET/quenching analysis and cell-based viral inhibition studies; **Chapter 5**.
5. smFRET and EPR as innovative tools to investigate the tetrameric inter-CRD arrangements and dynamics of DC-SIGN and DC-SIGNR, to better understand their relative positions and how these may contribute to their observed specificity differences; **Chapter 6**.
6. Finally, this thesis also explores the use of DC-SIGN for the development of novel DC-SIGN and DC-SIGN CRD-based lectin-NP conjugates as potent inhibitors against SARS-CoV-2 infection by directly targeting viral surface conserved glycans; **Chapter 7**.

This research provides an important advancement to the current study of protein-carbohydrate interactions, utilising and developing state-of-the-art nanotechnologies to provide previously unattained structural and biophysical details regarding the lectins DC-SIGN and DC-SIGNR and their multivalent lectin-glycan interactions. Here, glycan-QDs were able to distinguish that where DC-SIGN and DC-SIGNR both interact with mannosylated species with similar enthalpies, the crosslinking capability of DC-SIGNR results in a large entropic and kinetic penalty, which correlates with the reduced binding efficiency of DC-SIGNR often observed on cells or with other glycoconjugates. However, by lengthening the QD in one dimension, *i.e. via* the use of quantum rods, this crosslinking can be partially overcome, resulting in the addition of stronger simultaneous binding of all four binding sites to individual QRs with a comparable binding strength to that of DC-SIGN. Thus demonstrating the key contribution of shape influences, such as curvature, on the overall avidity

of glycoconjugates to lectin. Furthermore, it was revealed by single molecule FRET that DC-SIGNR has a much broader inter-binding site area than DC-SIGN, which explains why DC-SIGNR is only able to form these stronger tetravalent interactions with larger effective glycan display areas.

In addition to this, preliminary results measuring the binding of DC-SIGN and DC-SIGNR with glycan-GNPs on lipid surfaces, *via* QCM-D, provided strong evidence that the different binding modes of the two lectins (*i.e.* weaker crosslinking of DC-SIGNR and stronger simultaneous binding of DC-SIGN) can also be observed on membrane surfaces. Thus, agreeing with previous cell study results which demonstrate correlations with solution phase experiments.

Finally, by exploiting the strong multivalent mannose-binding capability of DC-SIGN, lectin-conjugated GNP viral neutralising agents have been successfully developed and shown to inhibit SARS-CoV-2 virions *in vitro* with high potency, irrespective of variant type, by targeting the highly mannosylated spike protein.

These data provide useful advancement in answering the questions of “how lectin structure influences glycan-binding behaviour” and “how glycoconjugate design influences lectin-binding behaviour”, using DC-SIGN and DC-SIGNR as a case study pair of lectins. The conclusions obtained herein can therefore be used to inform a more rational design of highly potent and specific lectin targeting therapeutics in the future as well as innovative glycan-targeting therapeutic treatments against diseases such as viral infection.

References

1. G. D. Brown, J. A. Willment and L. Whitehead, *Nat. Rev. Immunol.*, 2018, **18**, 374-389.
2. S. H. Barondes, *Science*, 1984, **223**, 1259-1264.
3. A. E. Kaoutari, F. Armougom, J. I. Gordon, D. Raoult and B. Henrissat, *Nat. Rev. Microbiol.*, 2013, **11**, 497-504.
4. K. Drickamer and M. E. Taylor, *Annu Rev Cell Biol*, 1993, **9**, 237-264.
5. M. Mammen, S. K. Choi and G. M. Whitesides, *Angew. Chem. Int. Ed. Engl.*, 1998, **37**, 2754-2794.
6. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, in *Molecular Biology of the Cell. 4th edition*, Garland Science, New York, 2002.
7. S. Bhatia, L. C. Camacho and R. Haag, *J. Am. Chem. Soc.*, 2016, **138**, 8654-8666.
8. D. Ding, Y. Yao, S. Zhang, C. Su and Y. Zhang, *Oncol Lett*, 2017, **13**, 13-21.
9. F. Salazar, H. F. Sewell, F. Shakib and A. M. Ghaemmaghami, *J. Allergy Clin. Immunol.*, 2013, **132**, 27-36.

10. M. D. Joshi, W. J. Unger, G. Storm, Y. van Kooyk and E. Mastrobattista, *J. Control Release*, 2012, **161**, 25-37.
11. P. Conde, M. Rodriguez, W. van der Touw, A. Jimenez, M. Burns, J. Miller, M. Brahmachary, H. M. Chen, P. Boros, F. Rausell-Palamos, T. J. Yun, P. Riquelme, A. Rastrojo, B. Aguado, J. Stein-Streilein, M. Tanaka, L. Zhou, J. Zhang, T. L. Lowary, F. Ginhoux, C. G. Park, C. Cheong, J. Brody, S. J. Turley, S. A. Lira, V. Bronte, S. Gordon, P. S. Heeger, M. Merad, J. Hutchinson, S. H. Chen and J. Ochando, *Immunity*, 2015, **42**, 1143-1158.
12. S. Pustylnikov, R. S. Dave, Z. K. Khan, V. Porkolab, A. A. Rashad, M. Hutchinson, F. Fieschi, I. Chaiken and P. Jain, *AIDS Res. Hum. Retrovir.*, 2016, **32**, 93-100.
13. W. C. Boyd and E. Shapleigh, *Science*, 1954, **119**, 419-419.
14. H.-J. Gabius, *Eur. J. Biochem.*, 1997, **243**, 543-576.
15. D. Di Iorio, M. L. Verheijden, E. van der Vries, P. Jonkheijm and J. Huskens, *ACS Nano*, 2019, **13**, 3413-3423.
16. Y. Guo, C. Sakonsinsiri, I. Nehlmeier, M. A. Fascione, H. Zhang, W. Wang, S. Pohlmann, W. B. Turnbull and D. Zhou, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 4738-4742.
17. R. A. Goldsby, T. J. Kindt, J. Kuby and B. A. Osborne, *Immunology, Fifth Edition*, W. H. Freeman, New York, 2002.
18. T. B. Geijtenbeek, R. Torensma, S. J. van Vliet, G. C. van Duijnhoven, G. J. Adema, Y. van Kooyk and C. G. Figdor, *Cell*, 2000, **100**, 575-585.
19. B. Jameson, F. Baribaud, S. Pöhlmann, D. Ghavimi, F. Mortari, R. W. Doms and A. Iwasaki, *J Virol*, 2002, **76**, 1866-1875.
20. B. M. Curtis, S. Scharnowske and A. J. Watson, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 8356-8360.
21. D. S. Kwon, G. Gregorio, N. Bitton, W. A. Hendrickson and D. R. Littman, *Immunity*, 2002, **16**, 135-144.
22. D. A. Mitchell, A. J. Fadden and K. Drickamer, *J. Biol. Chem.*, 2001, **276**, 28939-28945.
23. H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327-1335.
24. H. Feinberg, C. K. Tso, M. E. Taylor, K. Drickamer and W. I. Weis, *J. Mol. Biol.*, 2009, **394**, 613-620.
25. Q. D. Yu, A. P. Oldring, A. S. Powlesland, C. K. Tso, C. Yang, K. Drickamer and M. E. Taylor, *J. Mol. Biol.*, 2009, **387**, 1075-1080.
26. D. E. Leckband, S. Menon, K. Rosenberg, S. A. Graham, M. E. Taylor and K. Drickamer, *Biochemistry*, 2011, **50**, 6125-6132.
27. Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis and K. Drickamer, *Nat. Struct. Mol. Biol.*, 2004, **11**, 591-598.
28. H. Feinberg, D. A. Mitchell, K. Drickamer and W. I. Weis, *Science*, 2001, **294**, 2163-2166.
29. P. Valverde, S. Delgado, J. D. Martinez, J. B. Vendeville, J. Malassis, B. Linclau, N. C. Reichardt, F. J. Canada, J. Jimenez-Barbero and A. Arda, *ACS Chem. Biol.*, 2019, **14**, 1660-1671.
30. U. Švajger, M. Anderluh, M. Jeras and N. Obermajer, *Cell. Signal.*, 2010, **22**, 1397-1405.
31. T. B. H. Geijtenbeek, D. J. E. B. Krooshoop, D. A. Bleijs, S. J. van Vliet, G. C. F. van Duijnhoven, V. Grabovsky, R. Alon, C. G. Figdor and Y. van Kooyk, *Nat. Immunol.*, 2000, **1**, 353-357.
32. A. Engering, T. B. Geijtenbeek, S. J. van Vliet, M. Wijers, E. van Liempt, N. Demaurex, A. Lanzavecchia, J. Fransen, C. G. Figdor, V. Piguet and Y. van Kooyk, *J. Immunol.*, 2002, **168**, 2118-2126.
33. A. Cambi, F. de Lange, N. M. van Maarseveen, M. Nijhuis, B. Joosten, E. M. van Dijk, B. I. de Bakker, J. A. Fransen, P. H. Bovee-Geurts, F. N. van Leeuwen, N. F. Van Hulst and C. G. Figdor, *J. Cell. Biol.*, 2004, **164**, 145-155.
34. H. H. Smits, A. Engering, D. van der Kleij, E. C. de Jong, K. Schipper, T. M. van Capel, B. A. Zaat, M. Yazdanbakhsh, E. A. Wierenga, Y. van Kooyk and M. L. Kapsenberg, *J. Allergy Clin. Immunol.*, 2005, **115**, 1260-1267.

35. S. I. Gringhuis, J. den Dunnen, M. Litjens, M. van der Vlist and T. B. Geijtenbeek, *Nat. Immunol.*, 2009, **10**, 1081-1088.
36. S. I. Gringhuis, T. M. Kaptein, B. A. Wevers, M. van der Vlist, E. J. Klaver, I. van Die, L. E. Vriend, M. A. de Jong and T. B. Geijtenbeek, *Nat. Commun.*, 2014, **5**, 5074.
37. S. I. Gringhuis, T. M. Kaptein, B. A. Wevers, A. W. Mesman and T. B. Geijtenbeek, *Nat. Commun.*, 2014, **5**, 3898.
38. T. B. Geijtenbeek, D. S. Kwon, R. Torensma, S. J. van Vliet, G. C. van Duijnhoven, J. Middel, I. L. Cornelissen, H. S. Nottet, V. N. KewalRamani, D. R. Littman, C. G. Figdor and Y. van Kooyk, *Cell*, 2000, **100**, 587-597.
39. C. P. Alvarez, F. Lasala, J. Carrillo, O. Muniz, A. L. Corbi and R. Delgado, *J. Virol.*, 2002, **76**, 6841-6844.
40. G. Simmons, J. D. Reeves, C. C. Grogan, L. H. Vandenberghe, F. Baribaud, J. C. Whitbeck, E. Burke, M. J. Buchmeier, E. J. Soilleux, J. L. Riley, R. W. Doms, P. Bates and S. Pöhlmann, *Virology*, 2003, **305**, 115-123.
41. P. Liu, M. Ridilla, P. Patel, L. Betts, E. Gallichotte, L. Shahidi, N. L. Thompson and K. Jacobson, *Traffic*, 2017, **18**, 218-231.
42. B. Tassaneetrithep, T. H. Burgess, A. Granelli-Piperno, C. Trumpfheller, J. Finke, W. Sun, M. A. Eller, K. Pattanapanyasat, S. Sarasombath, D. L. Birx, R. M. Steinman, S. Schlesinger and M. A. Marovich, *J Exp Med*, 2003, **197**, 823-829.
43. E. G. Cormier, R. J. Durso, F. Tsamis, L. Boussemart, C. Manix, W. C. Olson, J. P. Gardner and T. Dragic, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 14067-14072.
44. P. Y. Lozach, H. Lortat-Jacob, A. de Lacroix de Lavalette, I. Staropoli, S. Foung, A. Amara, C. Houles, F. Fieschi, O. Schwartz, J. L. Virelizier, F. Arenzana-Seisdedos and R. Altmeyer, *J Biol Chem*, 2003, **278**, 20358-20366.
45. M. Shimojima, A. Takenouchi, H. Shimoda, N. Kimura and K. Maeda, *Arch Virol*, 2014, **159**, 2023-2031.
46. P. Wang, K. Hu, S. Luo, M. Zhang, X. Deng, C. Li, W. Jin, B. Hu, S. He, M. Li, T. Du, G. Xiao, B. Zhang, Y. Liu and Q. Hu, *Virology*, 2016, **488**, 108-119.
47. P. Wang, M. Li, W. Lu, D. Zhang, Q. Hu and Y. Liu, *Virol. Sin.*, 2017, **32**, 495-502.
48. L. de Witte, M. Abt, S. Schneider-Schaulies, Y. van Kooyk and T. B. Geijtenbeek, *J Virol*, 2006, **80**, 3477-3486.
49. R. Amraei, W. Yin, M. A. Napoleon, E. L. Suder, J. Berrigan, Q. Zhao, J. Olejnik, K. B. Chandler, C. Xia, J. Feldman, B. M. Hauser, T. M. Caradonna, A. G. Schmidt, S. Gummuluru, E. Mühlberger, V. Chitalia, C. E. Costello and N. Rahimi, *ACS Cent. Sci.*, 2021, **7**, 1156-1165.
50. L. Wu and V. N. KewalRamani, *Nat. Rev. Immunol.*, 2006, **6**, 859-868.
51. C. P. Mason and A. W. Tarr, *Molecules*, 2015, **20**, 2229-2271.
52. D. A. Mitchell, Q. Zhang, L. Voorhaar, D. M. Haddleton, S. Herath, A. S. Gleinich, H. S. Randeva, M. Crispin, H. Lehnert, R. Wallis, S. Patterson and C. R. Becer, *Chem. Sci.*, 2017, **8**, 6974-6980.
53. S. I. Gringhuis, J. den Dunnen, M. Litjens, B. van Het Hof, Y. van Kooyk and T. B. Geijtenbeek, *Immunity*, 2007, **26**, 605-616.
54. C. S. Hsieh, S. E. Macatonia, C. S. Tripp, S. F. Wolf, A. O'Garra and K. M. Murphy, *Science*, 1993, **260**, 547-549.
55. T. R. Mosmann, H. Cherwinski, M. W. Bond, M. A. Giedlin and R. L. Coffman, *J Immunol*, 1986, **136**, 2348-2357.
56. S. Shuttleworth, P. Townsend, F. Silva, A. Cecil, T. Hill, C. Tomassi, H. Rogers and R. Harrison, *Prog Med Chem*, 2011, **50**, 109-133.
57. T. B. Geijtenbeek, S. J. Van Vliet, E. A. Koppel, M. Sanchez-Hernandez, C. M. Vandenbroucke-Grauls, B. Appelmelk and Y. Van Kooyk, *J. Exp. Med.*, 2003, **197**, 7-17.
58. Z. Dembic, *The cytokines of the immune system : the role of cytokines in disease related to immune response*, Academic Press, 2015.
59. D. S. Shouval, J. Ouahed, A. Biswas, J. A. Goettel, B. H. Horwitz, C. Klein, A. M. Muise and S. B. Snapper, *Adv Immunol*, 2014, **122**, 177-210.
60. A. King, S. Balaji, L. D. Le, T. M. Crombleholme and S. G. Keswani, *Adv. Wound Care (New Rochelle)*, 2014, **3**, 315-323.

61. T. B. H. Geijtenbeek , S. J. van Vliet , E. A. Koppel , M. Sanchez-Hernandez , C. M. J. E. Vandenbroucke-Grauls , B. Appelmelk and Y. van Kooyk *Journal of Experimental Medicine*, 2002, **197**, 7-17.
62. R. Mittal, S. Bulgheresi, C. Emami and N. V. Prasadarao, *J Immunol*, 2009, **183**, 6588-6599.
63. M. Nonaka, B. Y. Ma, R. Murai, N. Nakamura, M. Baba, N. Kawasaki, K. Hodohara, S. Asano and T. Kawasaki, *J Immunol*, 2008, **180**, 3347-3356.
64. G. t. Deslée, A.-S. Charbonnier, H. Hammad, G. Angyalosi, I. Tillie-Leblond, A. Mantovani, A. é.-B. Tonnel and J. Pestel, *J. Allergy Clin. Immunol.*, 2002, **110**, 763-770.
65. J. Kayserova, I. Zentsova-Jaresova, V. Budinsky, D. Rozkova, J. Kopecka, E. Vernerova, P. Pohunek, V. Skalicka, R. Spisek and A. Sediva, *Scand J Immunol*, 2012, **75**, 305-313.
66. H. J. Huang, Y. L. Lin, C. F. Liu, H. F. Kao and J. Y. Wang, *Mucosal Immunol.*, 2011, **4**, 519-527.
67. S. Pöhlmann, E. J. Soilleux, F. Baribaud, G. J. Leslie, L. S. Morris, J. Trowsdale, B. Lee, N. Coleman and R. W. Doms, *Proc. Natl. Acad. Sci U. S. A.*, 2001, **98**, 2670-2675.
68. A. A. Bashirova, T. B. H. Geijtenbeek, G. C. F. van Duijnhoven, S. J. van Vliet, J. B. G. Eilering, M. P. Martin, L. Wu, T. D. Martin, N. Viebig, P. A. Knolle, V. N. KewalRamani, Y. van Kooyk and M. Carrington, *J. Exp. Med.*, 2001, **193**, 671-678.
69. E. J. Soilleux, R. Barten and J. Trowsdale, *J. Immunol.*, 2000, **165**, 2937-2942.
70. N. Rahimi, *Biology (Basel)*, 2021, **10**.
71. C. W. Davis, H. Y. Nguyen, S. L. Hanna, M. D. Sanchez, R. W. Doms and T. C. Pierson, *J. Virol.*, 2006, **80**, 1290-1301.
72. N. P. Chung, S. K. Breun, A. Bashirova, J. G. Baumann, T. D. Martin, J. M. Karamchandani, J. W. Rausch, S. F. Le Grice, L. Wu, M. Carrington and V. N. Kewalramani, *J. Biol. Chem.*, 2010, **285**, 2100-2112.
73. A. Velazquez-Campoy and E. Freire, *Nat Protoc*, 2006, **1**, 186-191.
74. J. Rao, J. Lahiri, L. Isaacs, R. M. Weis and G. M. Whitesides, *Science*, 1998, **280**, 708-711.
75. J. D. Chodera and D. L. Mobley, *Annu. Rev. Biophys.*, 2013, **42**, 121-142.
76. V. Porkolab, C. Pifferi, I. Sutkeviciute, S. Ordanini, M. Taouai, M. Thepaut, C. Vives, M. Benazza, A. Bernardi, O. Renaudet and F. Fieschi, *Org. Biomol. Chem.*, 2020, **18**, 4763-4772.
77. S. Bayda, M. Adeel, T. Tuccinardi, M. Cordani and F. Rizzolio, *Molecules*, 2019, **25**, 112.
78. J. Rojo, V. Díaz, J. M. de la Fuente, I. Segura, A. G. Barrientos, H. H. Riese, A. Bernad and S. Penadés, *ChemBioChem*, 2004, **5**, 291-297.
79. L. Lartigue, C. Innocenti, T. Kalaivani, A. Awwad, M. Sanchez Duque Mdel, Y. Guari, J. Larionova, C. Guérin, J. L. Montero, V. Barragan-Montero, P. Arosio, A. Lascialfari, D. Gatteschi and C. Sangregorio, *J Am Chem Soc*, 2011, **133**, 10459-10472.
80. M. Perrier, M. Gary-Bobo, L. Lartigue, D. Brevet, A. Morère, M. Garcia, P. Maillard, L. Raehm, Y. Guari, J. Larionova, J.-O. Durand, O. Mongin and M. Blanchard-Desce, *Journal of Nanoparticle Research*, 2013, **15**, 1602.
81. H. Khan, H. R. Mirzaei, A. Amiri, E. Kupeli Akkol, S. M. Ashhad Halimi and H. Mirzaei, *Semin. Cancer Biol.*, 2021, **69**, 24-42.
82. N. Losada-Garcia, C. Garcia-Sanz, A. Andreu, T. Velasco-Torrijos and J. M. Palomo, *Nanomaterials (Basel)*, 2021, **11**.
83. A. Bernardi, J. Jimenez-Barbero, A. Casnati, C. De Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K. E. Jaeger, M. Lahmann, T. K. Lindhorst, M. Marradi, P. Messner, A. Molinaro, P. V. Murphy, C. Nativi, S. Oscarson, S. Penades, F. Peri, R. J. Pieters, O. Renaudet, J. L. Reymond, B. Richichi, J. Rojo, F. Sansone, C. Schaffer, W. B. Turnbull, T. Velasco-Torrijos, S. Vidal, S. Vincent, T. Wennekes, H. Zuilhof and A. Imberty, *Chem. Soc. Rev.*, 2013, **42**, 4709-4727.
84. K. Niikura, T. Matsunaga, T. Suzuki, S. Kobayashi, H. Yamaguchi, Y. Orba, A. Kawaguchi, H. Hasegawa, K. Kajino, T. Ninomiya, K. Ijiro and H. Sawa, *ACS Nano*, 2013, **7**, 3926-3938.

85. R. Ojeda, J. L. de Paz, A. G. Barrientos, M. Martín-Lomas and S. Penadés, *Carbohydr. Res.*, 2007, **342**, 448-459.
86. N. C. Reichardt, M. Martín-Lomas and S. Penadés, *Chem. Soc. Rev.*, 2013, **42**, 4358-4376.
87. M. Boholm and R. Arvidsson, *NanoEthics*, 2016, **10**, 25-40.
88. W. J. Lees, A. Spaltenstein, J. E. Kingery-Wood and G. M. Whitesides, *J Med Chem*, 1994, **37**, 3419-3433.
89. M. Mammen, G. Dahmann and G. M. Whitesides, *J Med Chem*, 1995, **38**, 4179-4190.
90. A. Spaltenstein and G. M. Whitesides, *J Am Chem Soc*, 1991, **113**, 686-687.
91. M. N. Matrosovich, L. V. Mochaloval, V. P. Marininal, N. E. Byramova and N. V. Bovin, *FEBS*, 1990, **272**, 209-212.
92. R. Roy, F. Andersson, O. and G. Harms, *Angew Chem Int Ed Engl* 1992, **31**, 1478-1481.
93. R. Roy, D. Zanini, S. J. Meunier and D. Romanowska, *J Chem Soc Chem Commun*, 1993, **24**, 1869-1872.
94. L. V. Mochalova, A. B. Tuzikov, V. P. Marinina, A. S. Gambaryan, N. E. Byramova, N. V. Bovin and M. N. Matrosovich, *Antivir Res*, 1994, **23**, 179-190.
95. B. Arnáiz, O. Martínez-Ávila, J. M. Falcon-Perez and S. Penadés, *Bioconjugate Chem.*, 2012, **23**, 814-825.
96. O. Martínez-Ávila, K. Hijazi, M. Marradi, C. Clavel, C. Campion, C. Kelly and S. Penadés, *Chem. Eur. J.*, 2009, **15**, 9874-9888.
97. M. Reynolds, M. Marradi, A. Imberty, S. Penadés and S. Pérez, *Glycoconj. J.*, 2013, **30**, 747-757.
98. X. Jiang, A. Housni, G. Gody, P. Boullanger, M.-T. Charreyre, T. Delair and R. Narain, *Bioconjugate Chem.*, 2010, **21**, 521-530.
99. D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo and D. Zhou, *J. Am. Chem. Soc.*, 2020, **142**, 18022-18034.
100. X. Wang, O. Ramström and M. Yan, *Anal. Chem.*, 2010, **82**, 9082-9089.
101. N. A. Samoilova, M. A. Krayukhina, T. A. Babushkina, I. A. Yamskov, L. M. Likhosherstov and V. E. Piskarev, *J. Appl. Polym. Sci.*, 2017, **134**.
102. V. Ramtenki, D. Raju, U. J. Mehta, C. V. Ramana and B. L. V. Prasad, *New J. Chem.*, 2013, **37**, 3716-3720.
103. M. Álvarez-Paino, G. Marcelo, A. Muñoz-Bonilla, J. Rodríguez-Hernández and M. Fernández-García, *Polym. Chem.*, 2013, **4**, 986-995.
104. Y. Guo, I. Nehlmeier, E. Poole, C. Sakonsinsiri, N. Hondow, A. Brown, Q. Li, S. Li, J. Whitworth, Z. Li, A. Yu, R. Brydson, W. B. Turnbull, S. Pohlmann and D. Zhou, *J. Am. Chem. Soc.*, 2017, **139**, 11833-11844.
105. S. Bhatia, D. Lauster, M. Bardua, K. Ludwig, S. Angioletti-Uberti, N. Popp, U. Hoffmann, F. Paulus, M. Budt, M. Stadtmüller, T. Wolff, A. Hamann, C. Böttcher, A. Herrmann and R. Haag, *Biomaterials*, 2017, **138**, 22-34.
106. T. Tanaka, H. Ishitani, Y. Miura, K. Oishi, T. Takahashi, T. Suzuki, S.-i. Shoda and Y. Kimura, *ACS Macro Lett.*, 2014, **3**, 1074-1078.
107. K. Suzuki, T. Koyama, S. Yingsakmongkon, Y. Suzuki, K. Hatano and K. Matsuoka, *Bioorg. Med. Chem. Lett.*, 2012, **20**, 446-454.
108. M. Nagao, Y. Kurebayashi, H. Seto, T. Tanaka, T. Takahashi, T. Suzuki, Y. Hoshino and Y. Miura, *Polym. J.*, 2016, **48**, 745-749.
109. J. Haldar, L. Alvarez de Cienfuegos, T. M. Tumpey, L. V. Gubareva, J. Chen and A. M. Klibanov, *Pharm. Res.*, 2010, **27**, 259-263.
110. C. R. Becer, M. I. Gibson, J. Geng, R. Ilyas, R. Wallis, D. A. Mitchell and D. M. Haddleton, *J. Am. Chem. Soc.*, 2010, **132**, 15130-15132.
111. S. J. Kwon, D. H. Na, J. H. Kwak, M. Douaisi, F. Zhang, E. J. Park, J. H. Park, H. Youn, C. S. Song, R. S. Kane, J. S. Dordick, K. B. Lee and R. J. Linhardt, *Nat. Nanotechnol.*, 2017, **12**, 48-54.
112. J. Luczkowiak, S. Sattin, I. Sutkeviciute, J. J. Reina, M. Sanchez-Navarro, M. Thepaut, L. Martinez-Prats, A. Daghetti, F. Fieschi, R. Delgado, A. Bernardi and J. Rojo, *Bioconjug Chem.*, 2011, **22**, 1354-1365.

113. M. Ogata, S. Umemura, N. Sugiyama, N. Kuwano, A. Koizumi, T. Sawada, M. Yanase, T. Takaha, J. I. Kadokawa and T. Usui, *Carbohydr. Polym.*, 2016, **153**, 96-104.
114. A. Monaco, V. P. Beyer, R. Napier and C. R. Becer, *Biomacromolecules*, 2020, **21**, 3736-3744.
115. A. Nazemi, S. M. Haeryfar and E. R. Gillies, *Langmuir*, 2013, **29**, 6420-6428.
116. L. Bes, S. Angot, A. Limer and D. M. Haddleton, *Macromolecules*, 2003, **36**, 2493-2499.
117. I. Papp, C. Sieben, A. L. Sisson, J. Kostka, C. Bottcher, K. Ludwig, A. Herrmann and R. Haag, *Chembiochem*, 2011, **12**, 887-895.
118. J. Luczkowiak, A. Muñoz, M. Sánchez-Navarro, R. Ribeiro-Viana, A. Ginieis, B. M. Illescas, N. Martín, R. Delgado and J. Rojo, *Biomacromolecules*, 2013, **14**, 431-437.
119. K. Buffet, E. Gillon, M. Holler, J.-F. Nierengarten, A. Imberty and S. P. Vincent, *Org. Biomol. Chem.*, 2015, **13**, 6482-6492.
120. A. Muñoz, D. Sigwalt, B. M. Illescas, J. Luczkowiak, L. Rodríguez-Pérez, I. Nierengarten, M. Holler, J.-S. Remy, K. Buffet, S. P. Vincent, J. Rojo, R. Delgado, J.-F. Nierengarten and N. Martín, *Nat. Chem.*, 2016, **8**, 50-57.
121. R. Rísquez-Cuadro, J. M. García Fernández, J.-F. Nierengarten and C. Ortiz Mellet, *Chem. Eur. J.*, 2013, **19**, 16791-16803.
122. M. Abellán-Flos, B. J. J. Timmer, S. Altun, T. Aastrup, S. P. Vincent and O. Ramstrom, *Biosens. Bioelectron.*, 2019, **139**, 111328.
123. M. Sánchez-Navarro, A. Muñoz, B. M. Illescas, J. Rojo and N. Martín, *Chemistry*, 2011, **17**, 766-769.
124. A. M. DiLillo, K. K. Chan, X.-L. Sun and G. Ao, *Front. Chem.*, 2022, **10**.
125. M. Pernía Leal, M. Assali, J. J. Cid, V. Valdivia, J. M. Franco, I. Fernández, D. Pozo and N. Khiar, *Nanoscale*, 2015, **7**, 19259-19272.
126. J. Ramos-Soriano and J. Rojo, *Chem. Commun.*, 2021, **57**, 5111-5126.
127. L. Rodríguez-Pérez, J. Ramos-Soriano, A. Pérez-Sánchez, B. M. Illescas, A. Muñoz, J. Luczkowiak, F. Lasala, J. Rojo, R. Delgado and N. Martín, *J Am Chem Soc*, 2018, **140**, 9891-9898.
128. X. Chen, G. S. Lee, A. Zettl and C. R. Bertozzi, *Angew. Chem. Int. Ed.*, 2004, **43**, 6111-6116.
129. P. Wu, X. Chen, N. Hu, U. C. Tam, O. Blixt, A. Zettl and C. R. Bertozzi, *Angew. Chem. Int. Ed. Engl.*, 2008, **47**, 5022-5025.
130. M.-E. Ragoussi, S. Casado, R. Ribeiro-Viana, G. d. l. Torre, J. Rojo and T. Torres, *Chem. Sci.*, 2013, **4**, 4035-4041.
131. T. Koukalová, P. Kovaříček, P. Bojarová, V. L. P. Guerra, V. Vrkoslav, L. Navara, I. Jirka, M. Cebecauer, V. Křen and M. Kalbáč, *Int. J. Mol. Sci.*, 2021, **22**, 6661.
132. Y. Chen, H. Vedala, G. P. Kotchey, A. Audfray, S. Cecioni, A. Imberty, S. Vidal and A. Star, *ACS Nano*, 2012, **6**, 760-770.
133. T. Terada, M. Nishikawa, F. Yamashita and M. Hashida, *Int J Pharm*, 2006, **316**, 117-123.
134. R. Das and B. Mukhopadhyay, *Carbohydr. Res.*, 2021, **507**, 108394.
135. O. Martínez-Ávila, L. M. Bedoya, M. Marradi, C. Clavel, J. Alcamí and S. Penadés, *ChemBioChem*, 2009, **10**, 1806-1809.
136. S. C. Günther, J. D. Maier, J. Vetter, N. Podvalnyy, N. Khanzhin, T. Hennet and S. Stertz, *Sci. Rep.*, 2020, **10**, 768.
137. A. S. Gambaryan, A. B. Tuzikov, N. V. Bovin and M. N. Matrosovich, *Antivir Res*, 2002, **55**, 201–205.
138. S. Mari, H. Posteri, G. Marcou, D. Potenza, F. Micheli, F. J. Cañada, J. Jimenez-Barbero and A. Bernardi, *European Journal of Organic Chemistry*, 2004, **2004**, 5119-5225.
139. J. J. Reina, S. Sattin, D. Invernizzi, S. Mari, L. Martinez-Prats, G. Tabarani, F. Fieschi, R. Delgado, P. M. Nieto, J. Rojo and A. Bernardi, *ChemMedChem*, 2007, **2**, 1030-1036.
140. A. Gimeno, S. Delgado, P. Valverde, S. Bertuzzi, M. A. Berbís, J. Echavarren, A. Lacetera, S. Martín-Santamaría, A. Surolia, F. J. Cañada, J. Jiménez-Barbero and A. Ardá, *Angew Chem Int Ed Engl*, 2019, **58**, 7268-7272.

141. S. Bhatia, D. Lauster, M. Bardua, K. Ludwig, S. Angioletti-Uberti, N. Popp, U. Hoffmann, F. Paulus, M. Budt, M. Stadtmuller, T. Wolff, A. Hamann, C. Bottcher, A. Herrmann and R. Haag, *Biomaterials*, 2017, **138**, 22-34.
142. N. Varga, I. Sutkeviciute, R. Ribeiro-Viana, A. Berzi, R. Ramdasi, A. Daghetti, G. Vettoretti, A. Amara, M. Clerici, J. Rojo, F. Fieschi and A. Bernardi, *Biomaterials*, 2014, **35**, 4175-4184.
143. J. J. Garcia-Vallejo, N. Koning, M. Ambrosini, H. Kalay, I. Vuist, R. Sarrami-Forooshani, T. B. Geijtenbeek and Y. van Kooyk, *Int Immunol*, 2013, **25**, 221-233.
144. C. C. Lin, Y. C. Yeh, C. Y. Yang, G. F. Chen, Y. C. Chen, Y. C. Wu and C. C. Chen, *Chem Commun (Camb)*, 2003, 2920-2921.
145. Y. Y. Chien, M. D. Jan, A. K. Adak, H. C. Tzeng, Y. P. Lin, Y. J. Chen, K. T. Wang, C. T. Chen, C. C. Chen and C. C. Lin, *Chembiochem*, 2008, **9**, 1100-1109.
146. J. Vonnemann, S. Liese, C. Kuehne, K. Ludwig, J. Dernedde, C. Bottcher, R. R. Netz and R. Haag, *J Am Chem Soc*, 2015, **137**, 2572-2579.
147. S. Takae, Y. Akiyama, H. Otsuka, T. Nakamura, Y. Nagasaki and K. Kataoka, *Biomacromolecules*, 2005, **6**, 818-824.
148. J. J. Cid Martín, M. Assali, E. Fernández-García, V. Valdivia, E. M. Sánchez-Fernández, J. M. Garcia Fernández, R. E. Wellinger, I. Fernández and N. Khiar, *J Mater Chem B*, 2016, **4**, 2028-2037.
149. P. M. Chaudhary, S. Sangabathuni, R. V. Murthy, A. Paul, H. V. Thulasiram and R. Kikkeri, *Chem. Commun. (Camb)*, 2015, **51**, 15669-15672.
150. S. Sangabathuni, R. Vasudeva Murthy, P. M. Chaudhary, M. Surve, A. Banerjee and R. Kikkeri, *Nanoscale*, 2016, **8**, 12729-12735.
151. S. Toraskar, M. Gade, S. Sangabathuni, H. V. Thulasiram and R. Kikkeri, *ChemMedChem*, 2017, **12**, 1116-1124.
152. S. Pollastri, C. Delaunay, M. Thépaut, F. Fieschi and A. Bernardi, *Chem Commun (Camb)*, 2022, **58**, 5136-5139.
153. A. Holla and A. Skerra, *Protein Eng. Des. Sel.*, 2011, **24**, 659-669.
154. L. Kong, I. A. Wilson and P. D. Kwong, *Proteins*, 2015, **83**, 590-596.
155. S. H. Liyanage and M. Yan, *Chem. Commun.*, 2020, **56**, 13491-13505.
156. H. Otsuka, Y. Akiyama, Y. Nagasaki and K. Kataoka, *J. Am. Chem. Soc.*, 2001, **123**, 8226-8230.
157. D. C. Hone, A. H. Haines and D. A. Russell, *Langmuir*, 2003, **19**, 7141-7144.
158. C.-S. Tsai, T.-B. Yu and C.-T. Chen, *Chem. Commun.*, 2005, 4273-4275.
159. K. M. Halkes, A. Carvalho de Souza, C. E. P. Maljaars, G. J. Gerwig and J. P. Kamerling, *Eur. J. Org. Chem.*, 2005, **2005**, 3650-3659.
160. C. L. Schofield, A. H. Haines, R. A. Field and D. A. Russell, *Langmuir*, 2006, **22**, 6707-6711.
161. J. M. Bergen, H. A. von Recum, T. T. Goodman, A. P. Massey and S. H. Pun, *Macromol. Biosci.*, 2006, **6**, 506-516.
162. C. L. Schofield, B. Mukhopadhyay, S. M. Hardy, M. B. McDonnell, R. A. Field and D. A. Russell, *Analyst*, 2008, **133**, 626-634.
163. A. Aykaç, M. C. Martos-Maldonado, J. M. Casas-Solvas, I. Quesada-Soriano, F. García-Maroto, L. García-Fuentes and A. Vargas-Berenguel, *Langmuir*, 2014, **30**, 234-242.
164. Y.-J. Chuang, X. Zhou, Z. Pan and C. Turchi, *Biochem. Biophys. Res. Commun.*, 2009, **389**, 22-27.
165. X. L. Hu, H. Y. Jin, X. P. He, T. D. James, G. R. Chen and Y. T. Long, *ACS Appl. Mater. Interfaces*, 2015, **7**, 1874-1878.
166. L.-D. Huang, A. K. Adak, C.-C. Yu, W.-C. Hsiao, H.-J. Lin, M.-L. Chen and C.-C. Lin, *Chem. Eur. J.*, 2015, **21**, 3956-3967.
167. X. Wang, E. Matei, L. Deng, O. Ramström, A. M. Gronenborn and M. Yan, *Chem. Commun. (Camb)*, 2011, **47**, 8620-8622.
168. X. Wang, O. Ramström and M. Yan, *J. Mater. Chem.*, 2009, **19**, 8944-8949.
169. C.-C. Huang, C.-T. Chen, Y.-C. Shiang, Z.-H. Lin and H.-T. Chang, *Anal. Chem.*, 2009, **81**, 875-882.
170. B. N. Murthy, S. Zeile, M. Nambiar, M. R. Nussio, C. T. Gibson, J. G. Shapter, N. Jayaraman and N. H. Voelcker, *RSC Adv.*, 2012, **2**, 1329-1333.
171. Q. Chen, W. Wei and J.-M. Lin, *Biosens. Bioelectron.*, 2011, **26**, 4497-4502.

172. L.-H. Liu, H. Dietsch, P. Schurtenberger and M. Yan, *Bioconjugate Chem.*, 2009, **20**, 1349-1355.
173. G. Sánchez-Pomales, T. A. Morris, J. B. Falabella, M. J. Tarlov and R. A. Zangmeister, *Biotechnol. Bioeng.*, 2012, **109**, 2240-2249.
174. X. Wang, O. Ramström and M. Yan, *Analyst*, 2011, **136**, 4174-4178.
175. N. Hondow, R. Brydson, P. Wang, M. D. Holton, M. R. Brown, P. Rees, H. D. Summers and A. Brown, *J. Nanoparticle Res.*, 2012, **14**.
176. X. Zhu and T. Gao, in *Nano-Inspired Biosensors for Protein Assay with Clinical Applications*, ed. G. Li, Elsevier, 2019, pp. 237-264.
177. M. E. Yakovleva, G. R. Safina and B. Danielsson, *Anal Chim Acta*, 2010, **668**, 80-85.
178. Z. Pei, H. Anderson, T. Aastrup and O. Ramström, *Biosens Bioelectron*, 2005, **21**, 60-66.
179. F. Teillet, B. Dublet, J. P. Andrieu, C. Gaboriaud, G. J. Arlaud and N. M. Thielens, *J Immunol*, 2005, **174**, 2870-2877.
180. N. Frison, M. E. Taylor, E. Soilleux, M. T. Bousser, R. Mayer, M. Monsigny, K. Drickamer and A. C. Roche, *J Biol Chem*, 2003, **278**, 23922-23929.
181. Y. Shinohara, Y. Hasegawa, H. Kaku and N. Shibuya, *Glycobiology*, 1997, **7**, 1201-1208.
182. H. Kato, A. Yashiro, A. Mizuno, Y. Nishida, K. Kobayashi and H. Shinohara, *Bioorg. Med. Chem. Lett.*, 2001, **11**, 2935-2939.
183. C.-C. Lin, Y.-C. Yeh, C.-Y. Yang, G.-F. Chen, Y.-C. Chen, Y.-C. Wu and C.-C. Chen, *Chem. Commun.*, 2003, 2920-2921.
184. T. Hasegawa, T. Fujisawa, M. Numata, M. Umeda, T. Matsumoto, T. Kimura, S. Okumura, K. Sakurai and S. Shinkai, *Chem. Commun.*, 2004, 2150-2151.
185. Y.-K. Lyu, K.-R. Lim, B. Y. Lee, K. S. Kim and W.-Y. Lee, *Chem. Commun.*, 2008, 4771-4773.
186. E. Mahon, Z. Mouline, M. Silion, A. Gilles, M. Pinteala and M. Barboiu, *Chem. Commun.*, 2013, **49**, 3004-3006.
187. M. S. Itano, M. S. Graus, C. Pehlke, M. l. J. Wester, P. Liu, K. A. Lidke, N. L. Thompson, K. Jacobson and A. K. Neumann, *Front. Phys.*, 2014, **2**.
188. F. F. Rossetti, M. Bally, R. Michel, M. Textor and I. Reviakine, *Langmuir*, 2005, **21**, 6443-6450.
189. E. Reimhult, F. Höök and B. Kasemo, *J. Chem. Phys.*, 2002, **117**, 7401-7404.
190. R. P. Richter, R. Bérat and A. R. Brisson, *Langmuir*, 2006, **22**, 3497-3505.
191. G. Tabarani, M. Thepaut, D. Stroebel, C. Ebel, C. Vives, P. Vachette, D. Durand and F. Fieschi, *J. Biol. Chem.*, 2009, **284**, 21229-21240.
192. L. Stryer, *Annu Rev Biochem*, 1978, **47**, 819-846.
193. P. Wu and L. Brand, *Anal Biochem*, 1994, **218**, 1-13.
194. Y. Qiao, Y. Luo, N. Long, Y. Xing and J. Tu, *Micromachines (Basel)*, 2021, **12**, 492.
195. I. D. Sahu and G. A. Lorigan, *Biomolecules*, 2020, **10**, 763.
196. F. Torricella, A. Pierro, E. Mileo, V. Belle and A. Bonucci, *Biochim. Biophys. Acta Proteins Proteom.*, 2021, **1869**, 140653.
197. G. W. Reginsson and O. Schiemann, *Biochem*, 2011, **434**, 353-363.

Chapter 2
General Materials and Methodology

2.1 Materials and Methods

2.1.1 Materials

Phenol (>99.5%) was purchased from Alfa Aesar. D-mannose was purchased from Biosynth Carbosynth. Maleimide-Atto488, Maleimide-Atto643 Maleimide-Atto643 were purchased from ATTO-TEC GmbH. 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-phosphocholine (DOPC) was purchased from Avanti Polar Lipids. Bovine serum albumin (BSA, >99%) was purchased from BioServ UK. commercial CdSe/ZnSe/ZnS Core/Shell/Shell quantum dots (core diameter: 3.9 ± 0.5 nm, λ_{em} = 550 ± 8 nm, quantum yield (QY) = 62%, denoted as QD_{550}) bearing mixed ligands of hexadecylamine, trioctylphosphine and trioctylphosphine oxide in toluene, and CdSe/CdS elongated core/shell quantum rods (QR_{560}; core diameter = 3.1 ± 0.7 nm, core length = 15 ± 6 nm based on our TEM images; 1st excitonic λ_{abs} = 541 nm, λ_{em} = 558 nm, nominal QY = 68 %) capped with mixed ligands of trioctylphosphine oxide, trioctylphosphine and hexadecylamine in hexane, were purchased from Center for Applied Nanotechnology (CAN) GmbH. His_6-Cys peptide was purchased from China peptides. 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES, >99%); 2-amino-2-(hydroxymethyl)propane-1,3-diol (tris base >99.8%); calcium chloride ($CaCl_2$, fused granular); chloroform ($CHCl_3$, >99.8%); ethylenediamine tetraacetic acid, disodium salt dehydrate (ETDA, >99%); hydrochloric acid (HCl, ~37%); sodium chloride (NaCl, >99.5%); sodium hydroxide pellets (NaOH, >99%), and sulphuric acid (H_2SO_4, >95%) were purchased from Fischer Scientific. Hexane (>97%) and methanol (MeOH, >99.9%) and were purchased from Sigma-Aldrich. S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSSL) was purchased from Toronto Research Chemicals. Ethanol absolute (EtOH, >99.97%) was purchased from VWR Chemicals BDH®. Ultra-pure H_2O (resistance >18.2 MΩ.cm) was obtained through an ELGA Purelab classic UVF system.

DHLA-EG_{11}-DiMan, DHLA-EG_3-DiMan, DHLA-EG_3-OH, DHLA-EG_3-Man and LA-EG_2-Man were synthesised in-house using our previously established protocols (ligands and precursors were also provided by Darshita Budhadev, Yuanyuan Liu, Emma Poole and Chadamas Sakonsinsiri in our group).[1-3] LCMS: calculated *m/z* for $C_{60}H_{111}N_5O_{27}S_2$ (DHLA-EG_{11}-DiMan) $[M+H]^{2+}$ 699.84, found 699.92; calculated *m/z* for

$C_{44}H_{78}N_5NaO_{19}S_2$ (DHLA-EG$_3$-DiMan) $[M+Na]^+$ 1068.2, found 1068.6; calculated *m/z* for $C_{32}H_{59}N_5O_9S_2$ (DHLA-EG$_3$-OH) $[M+H]^+$ 722.38, found 722.41; calculated *m/z* for $C_{38}H_{68}N_5O_{14}S_2$ (DHLA-EG$_3$-Man) $[M+H]^+$ 884.11, found 884.46; calculated *m/z* for $C_{27}H_{49}N_4O_{11}S_2$ (LA-EG$_2$-Man) $[M+H]^+$ 669.2834, found 669.2838; calculated *m/z* for $C_{33}H_{58}N_4O_{16}S_2$ (LA-EG$_2$-Man) $[M+H]^+$ 831.3233, found 831.3242. LA-EG$_{11}$-Tz-TFP and LA-EG$_{11}$-TFP linkers and LA-EG$_4$-Gal and LA-EG$_4$-DiMan glycan ligands were synthesised in-house (provided by Darshita Budhadev in our group). HRMS: calculated *m/z* for $C_{69}H_{107}F_4N_5NaO_{21}S_2$ (LA-EG$_{11}$-Tz-TFP) $[M+Na]^+$ 1504.6728, found 1504.6725; calculated *m/z* for $C_{42}H_{69}F_4N_2O_{15}S_2$ (LA-EG$_{11}$-TFP) $[M+H]^+$ 981.4075, found 981.4098; calculated *m/z* for $C_{31}H_{59}N_4O_{13}S_2$ (DHLA-EG$_4$-Gal) $[M+H]^+$ 759.35, found 759.30; calculated *m/z* for $C_{37}H_{69}N_4O_{18}S_2$ (LA-EG$_4$-DiMan) $[M+H]^+$ 921.40; found 921.33. Tris-NTA(-O^{t}Bu)-succ-dioctadecylamine (tris-NTA DODA) was synthesised in-house, as described previously (provided by Changjiang You in the Piehler group).[4]

Common buffers included: loading buffer (25 mM Tris, 1.25 M NaCl, 25 mM $CaCl_2$, pH 7.8); binding buffer (20 mM HEPES, 100 mM NaCl, 10 mM $CaCl_2$, pH 7.8); elution buffer (20 mM HEPES, 100 mM NaCl, 2.5 mM EDTA, pH 7.8); binding buffer for dye labelling (20 mM HEPES, 100 mM NaCl, 10 mM $CaCl_2$, pH 7.2); and elution buffer for labelling (20 mM HEPES, 100 mM NaCl, 2.5 mM EDTA, pH 7.2); tris buffer (10 mM tris base, pH 7.8); dialysis buffer (1.25 M NaCl, 25 mM tris pH 7.8, 25 mM $CaCl_2$); SUV buffer (10 mM HEPES, 150 mM NaCl, pH 7.4). All buffers were made using ultra-pure water.

2.1.2 General instrumentation and computation

Centrifugation was performed using either a Thermo Scientific Haraeus Fresco 17, Heraeus Multifuge 3SR or a Beckman Coulter Avanti JXN-30 centrifuge, depending on the speed and volume, at room temperature (r.t.) unless otherwise stated. Concentration or washing by centrifugation was carried out using Sartorius Stedim Lab 30 kDa molecular weight cut-off (MWCO) Vivaspin 500 and Merck Millipore 10 kDa MWCO Amicon Ultra centrifugal filters for QRs and protein, respectively. Dialysis was performed using Thermofischer Scientific 14000 MWCO BioDesign Dialysis Tubing. Evaporation was performed using at reduced pressure using Genevac Concentrator EZ-2 or a Virtis Benchtop K freeze dryer.

High resolution mass spectrometry (HRMS) was used to analyse all protein samples and was performed using a Bruker Daltonics MicroTOF mass spectrometer. Deconvoluted mass values reported are in Da and protein labelling efficiency (LE) was obtained from the ratio of the integral of the labelled protein HRMS peak to the sum of that of the labelled and unlabelled protein peaks. Liquid chromatography mass spectroscopy (LC-MS) was used to analyse all other samples and was performed using a Bruker AmaZon speed mass spectrometer. Ultraviolet-visible light spectroscopy (UV-vis) was performed on either a Cary 60 UV-vis spectrophotometer using an Agilent Technologies sub-micro 10 mm quartz cell or a Thermo Scientific Nanodrop 2000 spectrophotometer with optical path length of 1 mm using a droplet of sample.

All images were analysed using ImageJ version 1.4.3.67 or 1.53q; all numerical data were analysed using Microsoft Excel 2016; and graphs were plotted using Origin 2019b software.

2.1.3 DHLA-EG$_n$-Glycan

DHLA-EG$_{11}$-DiMan and DHLA-EG$_3$-OH ligands were synthesised by copper-free "clicking" interactions between LA-EG$_{11}$-Cyclooctyne and N$_3$-EG$_2$-DiMan or LA-EG$_3$-Cyclooctyne and N$_3$-EG$_2$-OH followed by reduction using TCEP·HCl as reported previously.[1] All the glycans and linkers were synthesised in-house and purified using our previously established protocols (precursors were also provided by Darshita Budhadev, Yuanyuan Liu, Emma Poole and Chadamas Sakonsinsiri in our group).[1,2] MS: calculated *m/z* for $C_{60}H_{111}N_5O_{27}S_2$ (DHLA-EG$_{11}$-DiMan) $[M+2H]^{2+}$ 699.84, found 699.92; calculated *m/z* for $C_{32}H_{59}N_5O_9S_2$ (DHLA-EG$_3$-OH) $[M+H]^+$ 722.38, found 722.41 (**Figure 2.1**).[f]

---

[f] It is worth noting that a DL-lipoic acid precursor is used for the synthesis of these molecules. Thus, a mixture of enantiomers is present in each of the final products, due to the chiral centre of the lipoic acid substituent. However, this is not a concern for the purpose of nanoparticle conjugation. In addition, although all products are purified *via* column chromatography, the yielding ligand will inevitably contain a mixture of triazole regioisomers due to the random relative orientations of the reacting azide and alkyne groups. (Note: this could be minimised by instead employing copper-catalysed alkyne-azide cycloaddition, which favours the trans regioisomer (depicted in **Figure 2.1**) due to steric hindrance of the other substituents during the reaction).

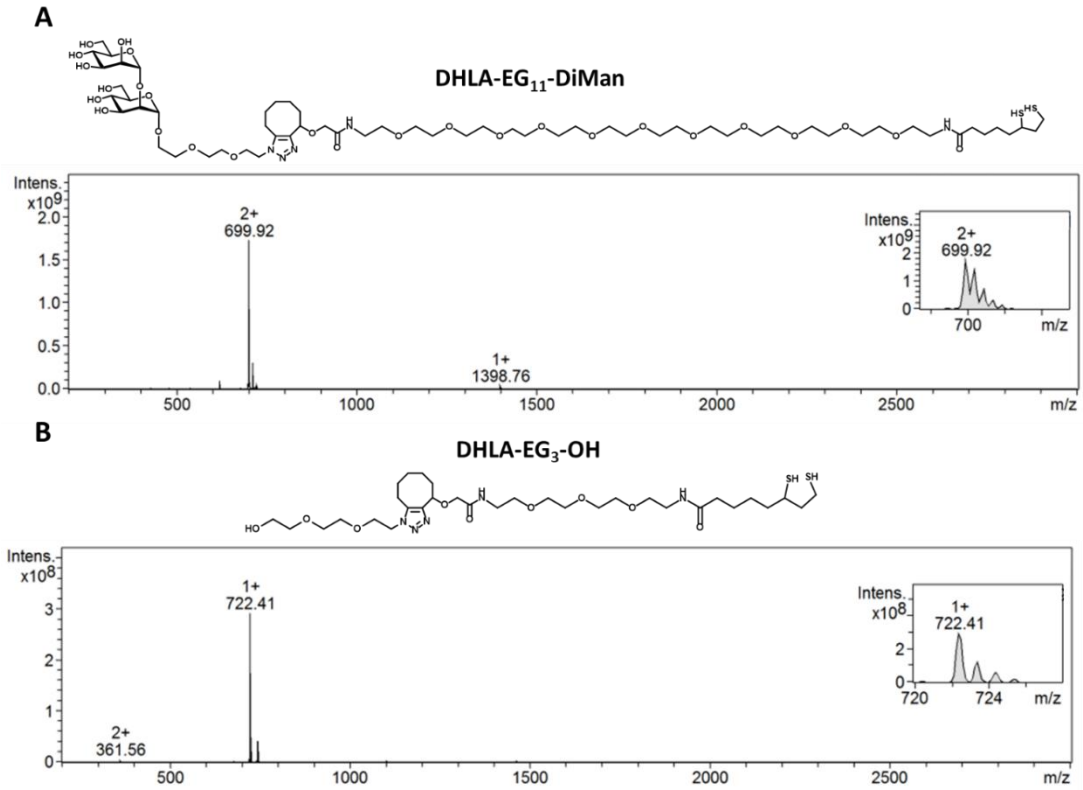

Figure 2.1. Mass spectrometry data and chemical structures of in-house synthesized ligands for functionalize the QD, (A) DHLA-EG$_{11}$-DiMan and (B) DHLA-EG$_3$-OH.[a]

2.1.4 Production and characterisation of proteins

2.1.4.1 Wild-type protein

All proteins used were based off the extracellular domains (ECDs) of DC-SIGN and DC-SIGNR. Plasmids encoding for the wild type amino acid identities DC-SIGN and DC-SIGNR were made in-house (provided by Yuan Guo).

Flakes of a pre-prepared ampicillin-resistant *E. coli* (BL21/DE3) glycerol stock containing the plasmid coding for the desired protein were added to sterilised solutions of LB (2.5 g) in H_2O (100 mL) with ampicillin (100 µL, 50 mg mL^{-1} in H_2O) and were incubated for 15 hr at 30 ^oC and 200rpm. The resulting start culture (30 mL) was then added to sterilised solutions of LB (25 g) in H_2O (1 L) with ampicillin (1 mL, 50 mg mL^{-1} in H_2O) and incubated at 37 ^oC and 200 rpm until the optical density at 550 nm (OD_{550}) was around 0.7-0.8. IPTG (1 mL, 100 mg mL^{-1} in H_2O) was then added and the cell culture was incubated at the same conditions for a further 2.5 hrs. The resulting growth culture was then centrifuged at 5000 rpm at 4 ^oC for 10 min. The

supernatant was discarded, and the pellet was redispersed in tris buffer (10 mM tris base, pH 7.8) and centrifuged at 10000 rpm at 4 ^{o}C for 10 min. The pellet was then collected and stored at -80 ^{o}C.

For 3 L of growth culture, the resulting pellet was then dispersed in cold tris buffer (100 mL) using sonication (50 % cycle, 60 % power, 8 cycles of alternating between 2 min sonication and 1 min at 0 ^{o}C). The cells were then disrupted *via* two cycles in a cell disruptor at 23 kPsi, followed by one cycle of tris buffer (50 mL). The resulting solution was then centrifuged at 20000 rpm at 4 ^{o}C for 15 min. The pellet was then suspended in a buffer of guanidine·HCl (6 M), tris (100 mM, pH 7.0) and β-mercaptoethanol (20 µL) (total volume 60 mL) and further sonicated (50 % cycle, 60 % power, 2 cycles of 2 min sonication and 1 min at 0 ^{o}C followed by 2 min sonication and 30 min at 0 ^{o}C). The solution was then centrifuged at 20000 rpm and 4 ^{o}C for 30 min. The supernatant was added to dialysis buffer (1.25 M NaCl, 25 mM tris pH 7.8, 25 mM $CaCl_2$; total volume 150 mL) and dialysed against the same buffer (2 L) for around 42 hr, where the buffer was replaced 3 times. The dialysed material was then collected and centrifuged at 23000 rpm and 4 ^{o}C for 30 min.

The supernatant was then added to a mannose-sepharose affinity column which had been pre-equilibrated with dialysis buffer. For DC020 and DC020-C, the column was washed with binding buffer (20 mM HEPES, 100 mM NaCl, 10 mM $CaCl_2$, pH 7.8) before being eluted using with elution buffer (20 mM HEPES, 100 mM NaCl, 2.5 mM EDTA, pH 7.8). $CaCl_2$ (pH 7.4) was then added to the eluted fractions to a final concentration of 10 mM. Protein concentrations were obtained by UV-vis spectroscopy using the Beer-Lambert law (**Equation 2.1**; where A_λ is the absorbance a specific wavelength λ, c is the concentration, ε is the extinction coefficient at λ, and l is the sample path length). Here, the protein monomeric concentration is determined by the A_{280} using $\varepsilon_{\mathrm{DC-SIGN}}$ = 70400 $M^{-1}{\cdot}cm^{-1}$ and $\varepsilon_{\mathrm{DC-SIGNR}}$ = 60890 $M^{-1}{\cdot}cm^{-1}$.

$$A_\lambda = c\varepsilon l \qquad 2.1$$

Protein molecular weights (MW) were confirmed from deconvoluted HRMS data, where calculated protein MW values were obtained from the amino acid sequences of one monomeric unit of the corresponding protein. Protein tetrameric hydrodynamic diameters (D_h) were confirmed by DLS analysis.

DC-SIGN: HRMS: calculated MW [DC-SIGN] 39197.22; found 39202.44. DLS: D_h 14.0 ± 4.3 nm.

DC-SIGNR: calculated MW [DC-SIGNR] 37478.99; found 37470.17. DLS: D_h 14.5 ± 3.6 nm.

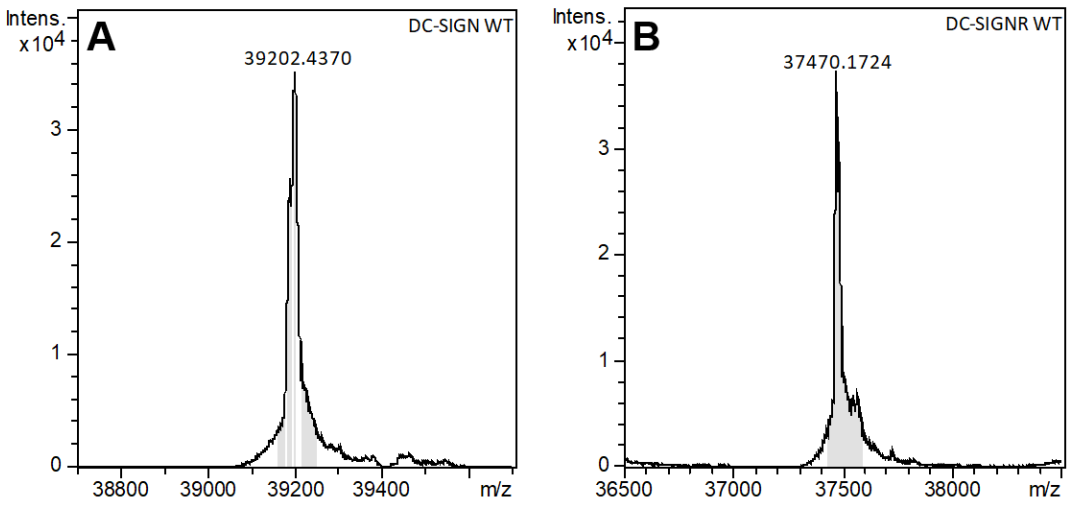

Figure 2.2. Deconvoluted HRMS spectra of the monomeric extracellular domains of (A) DC-SIGN (wild type; WT) and (B) DC-SIGNR (WT).

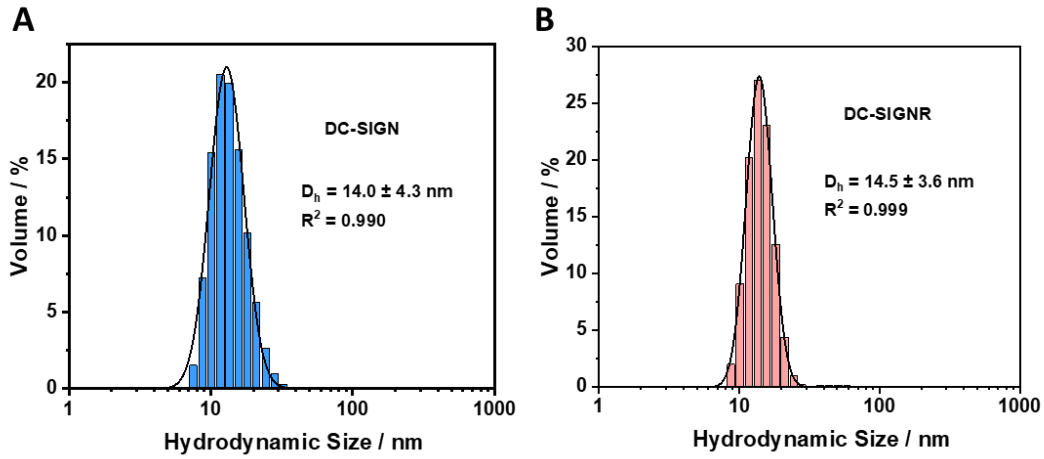

Figure 2.3. Volume population hydrodynamic size distribution histograms fitted with lognormal Gaussian distribution for extracellular segments of wide-type lectins: (A) DC-SIGN (WT) (B) DC-SIGNR (WT), at 40 nM in binding buffer (D_h values given as mean ± ½ FWHM).

2.1.4.2 Labelled protein

Plasmids encoding for DC-SIGN Q274C and DC-SIGNR R287C plasmids were made in-house by site-directed mutagenesis, as described previously.[2]

Proteins were produced from pre-prepared ampicillin-resistant *E. coli* glycerol stocks containing the plasmid coding for the desired protein, as described in **Section 2.1.4.1**. However, during the mannose-sepharose

column chromatography purification, loaded protein was instead washed with binding buffer for dye labelling (20 mM HEPES, 100 mM NaCl, 10 mM $CaCl_2$, pH 7.2) and eluted with elution buffer for labelling (20 mM HEPES, 100 mM NaCl, 2.5 mM EDTA, pH 7.2), before adding $CaCl_2$ (pH 7.4) to the eluted fractions to a final concentration of 25 mM.

The fractions containing protein were combined and concentrated to a final absorbance at 280 nm (A_{280}) of >0.6. Either maleimide-Atto594 (7.36 mM in DMSO, monomer protein:dye molar ratio = 3:1) was then added and immediately mixed by vortex. The solution was then covered and slowly rotated at r.t. for 1 hr, before being stored at 4 ^{o}C overnight. The resulting mixture was then added to a mannose-sepharose affinity column, which had been pre-equilibrated in binding buffer, and washed with binding buffer until the excess dye was removed. The protein was then eluted using elution buffer and $CaCl_2$ (pH 7.4) was then added to each fraction to a final concentration of 10 mM. Concentration was determined using UV-vis adsorption using a modified Beer-Lambert equation (**Equation 2.2**; where [P] is the monomer concentration of protein (P); A_λ is the maximum absorbance peak at wavelength λ (*i.e.* A_{280} for P or A_{603} for Atto594); CF_{280} is the correction factor for the dye at 280 nm (CF_{280} = 0.5 for Atto594), and ε_{P} is the extinction coefficient of monomeric P at 280 nm (ε_{280} = and 60890 $M^{-1}{\cdot}cm^{-1}$ for DC-SIGN Q274C and DC-SIGNR R287C, respectively)).

$$[\mathrm{P}] = \frac{A_{280} - \mathrm{CF}_{\mathrm{dye},280} A_{\mathrm{dye},\lambda}}{\varepsilon_{\mathrm{P},280}\, l} \qquad 2.2$$

Protein molecular weights (MW) were confirmed from deconvoluted HRMS data, where calculated protein MW values were obtained from the amino acid sequences of one monomeric unit of the corresponding protein and LE was obtained from the ratio of the labelled protein to unlabelled protein peak areas.

DC-SIGN-Atto594: HRMS: calculated MW [DC-SIGN Q274C] 39172.22, [DC-SIGN Q274C + 2Ca] 39252.40, [DC-SIGN Q274C-Atto594] 40100.22; found 39253.60 and 40105.75. LE: 87%.

DC-SIGNR-Atto594: HRMS: calculated MW [DC-SIGNR R287C] 37425.95, [DC-SIGNR R287C + 2Ca]: 37506.10, [DC-SIGNR R287C-Atto594] 38353.95; found: 37507.48 and 38358.26. LE: 85%.

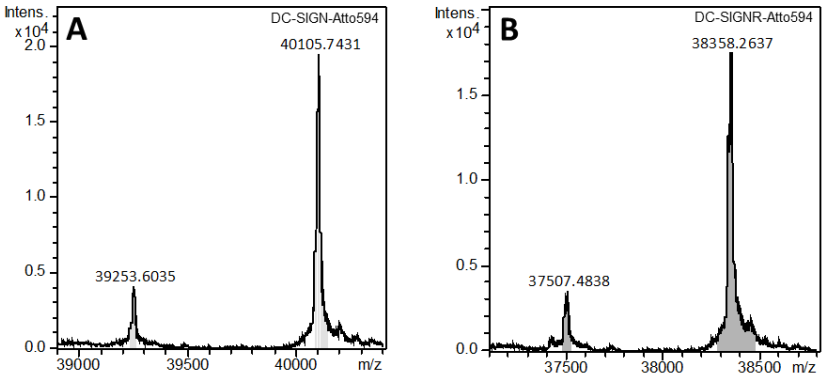

Figure 2.4. Deconvoluted HRMS spectra of the monomeric extracellular domains of (A) DC-SIGN Q274C labelled with maleimide-Atto594 (DC-SIGN-Atto594), (B) DC-SIGNR R287C labelled with maleimide-Atto594 (DC-SIGNR-Atto594).

2.2 Experimental methodology and instrumentation

2.2.1 Experimental Theory

2.2.1.1 QD-FRET Theory

Förster resonance energy transfer (FRET) between donor and acceptor fluorophores can be measured quantitatively *via* the *FRET efficiency* (E), which describes the proportion of energy absorbed by the donor which is transferred to the acceptor. The efficiencies of such dipole-dipole energy transfer processes are well-known to vary with the inverse 6th power of the donor-acceptor distance (R). For 1:1 donor:acceptor FRET systems this can be described using in **Equation 2.3**:

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6} \qquad 2.3$$

R_0, also known as the *Förster radius*, describes the donor-acceptor distance of a specific FRET pair, at which E is 50%. This is characterised by the conditions of the FRET pair system, such as: the dipole orientation factor (κ^2); the refractive index of the solution (n_r); the quantum yield (QY) of the donor (Φ_D; that is, the efficiency by which a fluorophore converts absorbed light into fluorescence); and the spectral overlap (J) between the donor fluorescence and the acceptor absorbance spectra (**Equation 2.5**; where I_D is the fluorescence intensity of the donor, ε_A is the extinction coefficient of the acceptor, and λ is the wavelength of light).

$$R_0 = \left(8.79 \times 10^{-5} \cdot \frac{\kappa^2 \cdot \Phi_{\mathrm{D}} \cdot J}{{n_r}^4}\right)^{\frac{1}{6}} \qquad 2.4$$

$$J = \int J(\lambda)\,\mathrm{d}\lambda = \int \frac{I_{\mathrm{D}}(\lambda)\varepsilon_{\mathrm{A}}(\lambda)\lambda^4}{\int I_{\mathrm{D}}(\lambda)\,\mathrm{d}\lambda}\,\mathrm{d}\lambda \qquad 2.5$$

For polyvalent glycan-QDs, however, their surface areas are often much larger than the binding contact area of many lectins, which means that multiple acceptors can often be assembled onto a single glycan-QD surface. In such cases, E is not only dependent on R and R_0, but also the effective number of acceptors within proximity to each QD (N), in accordance with **Equation 2.6** (where, in order for N to represent the number of bound acceptor, the donor-acceptor distance for each acceptor pair must be the same).[5]

$$E = \frac{1}{1 + \frac{1}{N}\left(\frac{R}{R_0}\right)^6} \qquad 2.6$$

The FRET efficiency can be related to measured fluorescence spectrum of the system *via* the integrated fluorescence intensities corresponding to the donor and acceptor fluorophores, $I_{\mathrm{D}}(\lambda)$ and $I_{\mathrm{A}}(\lambda)$, respectively (**Equation 2.7**; where γ is a correction factor for the quantum yield difference between the donor and acceptor).

$$E = \frac{1}{1 + \gamma \int \frac{I_{\mathrm{D}}(\lambda)}{I_{\mathrm{A}}(\lambda)}\mathrm{d}\lambda} \qquad 2.7$$

As the ratio of the integrated fluorescence intensities should be directly proportional to the ratio of the intensities of each peak, a relationship between the ratio of the acceptor to donor fluorescence intensities can be derived which is shown to be proportional to N (**Equation 2.8**; where α is the factor of proportionality between the donor to acceptor integrated fluorescence ratio to the donor to acceptor peak intensity ratio). This ratio is hereon referred to as the *FRET ratio* (F).

$$F = \frac{I_{\mathrm{A}}}{I_{\mathrm{D}}} = N\left[\gamma \cdot \alpha \cdot \left(\frac{R_0}{R}\right)^6\right] \qquad 2.8$$

Given a constant mean donor-acceptor separation distance, this linear proportionality between the FRET ratio and the number of bound protein therefore allows us to directly measure the proportion of bound protein (N_{bound}/N_{total}) under constant lectin:QD molar ratios, by simply measuring F at different lectin concentrations ($[L]$) (where $F/F_{\max} = N_{bound}/N_{total}$; **Figure 2.5 right**). By monitoring how the proportion of bound protein changes with concentration *via* the FRET ratio, the K_d can be obtained using the Hill equation (**Equation 2.9**; where n denotes the Hill coefficient, a measure of the propensity of a bound protein to assist or hinder subsequent binding, *i.e.* $n > 1,\ n < 1,\ n = 1$ describes positive, negative or no cooperativity, respectively).

$$F = F_{\max} \cdot \frac{N_{bound}}{N_{total}} = F_{\max} \cdot \frac{[\mathrm{L}]^n}{{K_\mathrm{d}}^n + [\mathrm{L}]^n} \qquad 2.9$$

In addition to this, because of the linear dependency of F with N (**Equation 2.8**), the FRET ratio can also be used to analyse the number of proteins able to bind to each glycan-QD (*i.e.* the valency). This can be done by measuring the change in F with the titration of protein against a fixed QD concentration. Assuming ideal N:1 binding of protein to a single QD and identical donor-acceptor separation distances, this should result in a linear F-N relationship which plateaus upon saturation of lectins on the glycan-QD surface (provided $[QD] \gg K_\mathrm{d}$ and assuming no other interactions affect the FRET process, e.g. there is no change of acceptor fluorescence with increasing QD surface coverage). Thus the glycan-QD valency is given by the point at which this linear increase reaches the maximum FRET ratio.

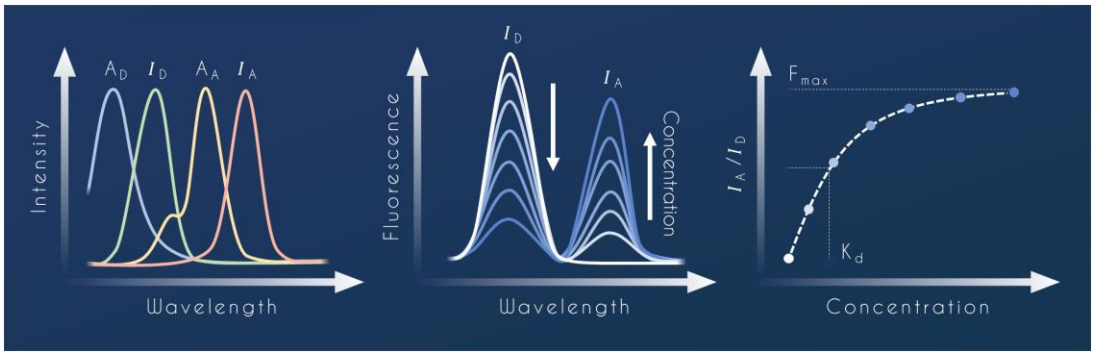

Figure 2.5 Diagram depicting example plots related to Figure 1.10, for: the overlap between the adsorption spectra for donor (QD) and acceptor (dye) fluorophores (A_{D} and A_{A}, respectively) and their emission spectra (with fluorescence intensities I_{D} and I_{A}, respectively) (left); the fluorescence spectra of a FRET pair which increases in binding with increasing concentration (middle); and the change of FRET ratio ($I_{\mathrm{A}}/I_{\mathrm{D}}$) with concentration, showing the maximum FRET ratio (F_{max}) and the dissociation constant (K_{d}) (right).

2.2.1.2 Theory Behind QCM-D Analysis of SLB

Quartz crystal microbalance with dissipation monitoring (QCM-D) is a convenient technique for the setup and monitoring of SLB formation and any subsequent surface interactions. Because quartz displays piezoelectric character (whereby electrical voltage applied to the crystal is converted into mechanical deformation, and *vice versa*), when an alternating current is applied then the crystal will mechanically vibrate in sync. If this vibration frequency matches the crystal's acoustic *resonance frequency* then the crystal is able to freely oscillate which can be monitored *via* its alternating current. The resonance frequency is negatively correlated to the areal mass density (m) of material on the sensor. This means any mass added to the surface (*e.g.* during the formation of SLBs or the subsequent interaction of material) will lead to negative shifts in frequency (f) (**Equation 2.10**; where n is the overtone number, C is the sensitivity constant).

$$\Delta f = -\frac{n\Delta m}{C} \qquad 2.10$$

Additionally, QCM-D monitors the *dissipation of energy* (*i.e.* the loss of energy from the system through heat, per oscillation). This is determined by the mechanical properties of the materials adhered to the surface. Softening of the sensor surface results in an increase in dissipation, whereas more rigid surfaces decrease the dissipation.

QCM-D therefore allows for the quantitative measurement of binding over time under a constant flow of material in buffer. This can be observed during the adsorption of SUVs onto the surface which initially results in a large negative frequency shift (Δf) from the buffer baseline (correlating to an increase in mass) as intact SUVs adsorb to the silica surface (**Figure 2.6**).[6] However, once a critical vesicular coverage (CVC) is obtained, crowding stresses cause rupturing of the adsorbed SUVs, resulting in the formation of the SLB. This is demonstrated by the partial return of the Δf, which forms a new baseline which shifts only slightly upon washing with buffer as any remaining material adsorbed is displaced from the SLB surface. By also monitoring the dissipation, a signature transient shift can be observed which confirms these two phases (*i.e.* the soft intact SUV layer which is much higher in dissipation and the final SLB surface which is much more similar silica baseline).

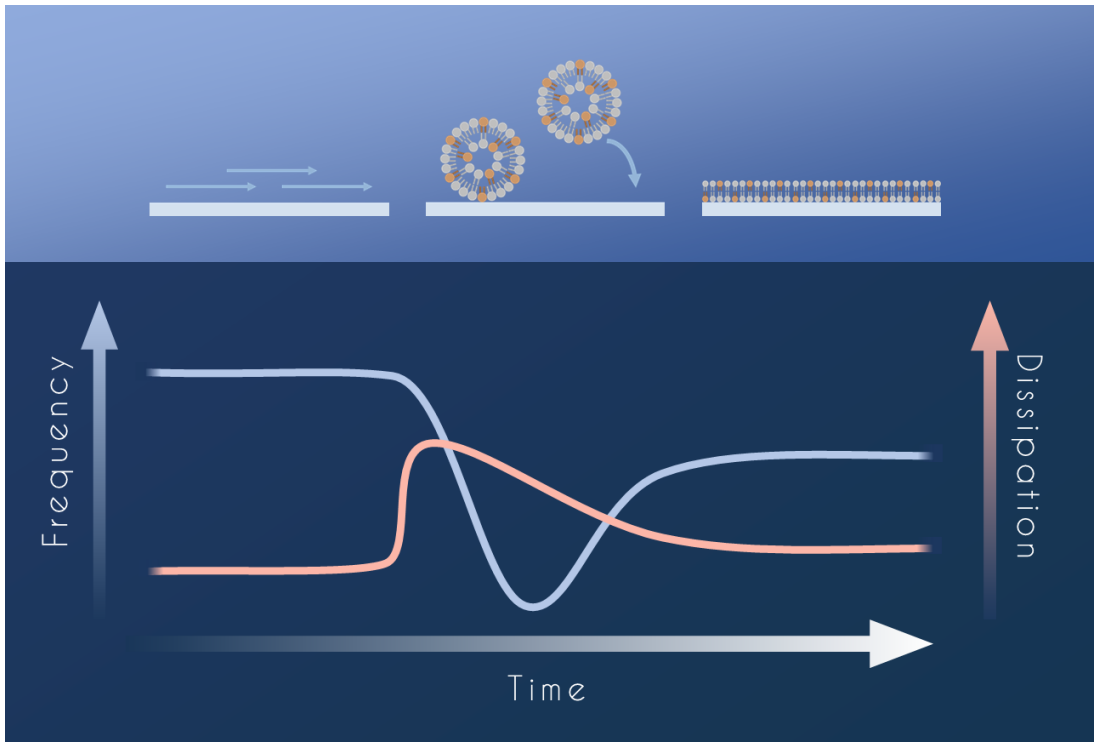

Figure 2.6 QCM-D monitoring of the formation of an SLB *via* the adsorption and subsequent rupturing of SUVs upon a hydrophilic support (reconstructed from Reference [6]).

If the SUV is also comprised of lipids containing a functional tethering group, other materials such as proteins or glycans may be bound to the surface. These functionalised SLBs allow for the concentration- or time-dependent measurement of protein-ligand association by flowing the corresponding analyte (*e.g.* free lectin or glycan) over the semi-immobilised surface and measuring the decrease in Δf. Additionally, the time-dependent

measurement of dissociation can also be obtained by washing the bound surface with pure buffer and measuring the increase in Δf with time as analyte dissociates from the surface. As Δf is proportional to the amount of bound protein, the kinetic parameters such as the apparent-K_{d} can be derived by the Δf-concentration relationship *via* the Hill equation (**Equation 2.9**), and the apparent-k_{off} *via* the first order rate of dissociation (**Equation 1.9**).

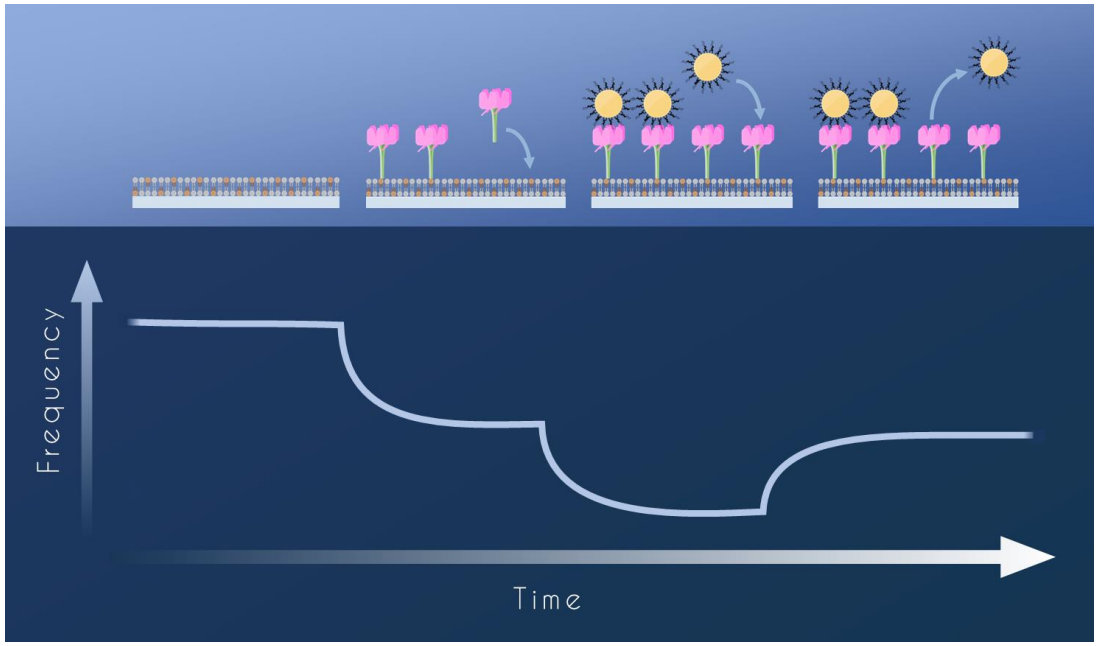

Figure 2.7 QCM-D monitoring of the functionalisation of an SLB surface with protein, and subsequent association and dissociation of analyte.

2.2.1.3 EPR Theory

Electron paramagnetic resonance (EPR analyses the splitting of unpaired electron spin energy levels as they align themselves either parallel ($m_s = -1/2$) or antiparallel ($m_s = 1/2$) to an external magnetic field. In the case of continuous wave (cw)-EPR, the extent to which energy gap (ΔE) between these energy levels is affected by an external magnetic field can then be monitored by scanning the magnetic field strength (B_0) and recording the point at which a fixed microwave frequency (ν) is able to excite the electrons into an antiparallel spin state (**Equation 2.11**; where h is the Planck constant, and μ_B is the Bohr magneton). The extent of the splitting of the spin energy levels is determined by the electron's *g*-factor (g) which is known value for free electrons (*i.e.* g_e) but can be altered by any local magnetic fields.

$$h\nu = \Delta E = g\mu_B B_0 \qquad 2.11$$

Nitroxide SDSLs are particularly suitable for EPR studies as the spin of the N^{14} nuclei (I = 1) couples with the spin of the unpaired electron on the $O^{\bullet}$ which results in a hyperfine triplet splitting pattern. By measuring the change in absorption over modulating ΔB_0, a first derivative spectrum is obtained which can be used to derive key details about the system. For example, as well as the local magnetic environment, the dynamics of the system also influence the signal. Here, the correlation time (τ_c) of the spin label (*i.e.* the average time it takes for a molecule to rotate one radian) is determined by its mobility and effects the signal amplitude and broadening.

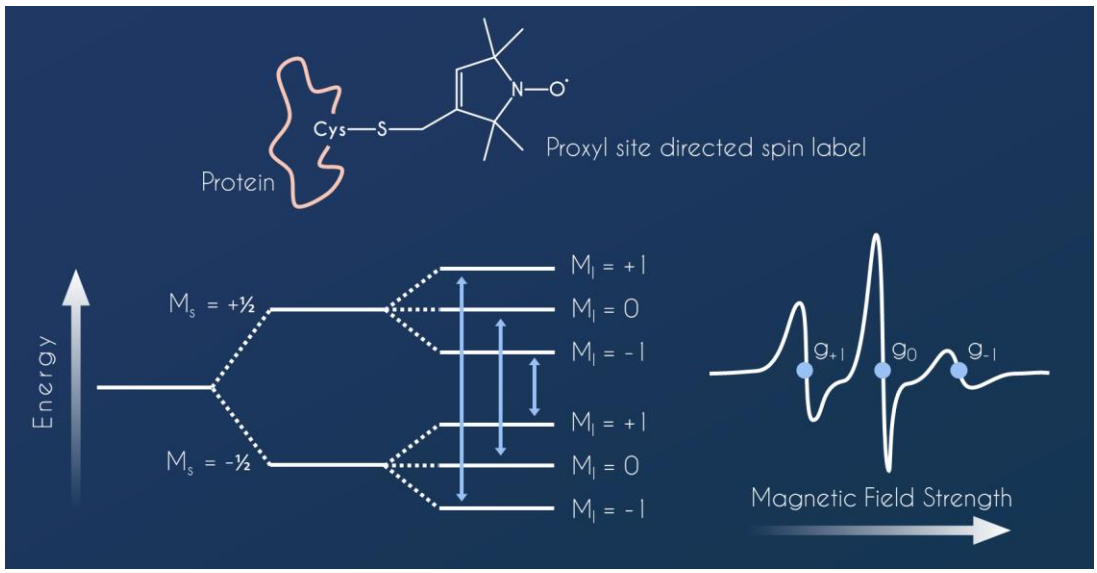

Figure 2.8 (A) schematic of proxyl SDSL at a cysteine residue; (B) energy level diagram demonstrating the hyperfine splitting of electron spin upon coupling with a N^{14} nucleus (C) cw-EPR spectrum corresponding to the hyperfine splitting of nitroxide unpaired electrons.

2.2.2 General Experimental Method

2.2.2.1 FRET assays

All FRET studies were performed using a Cary Eclipse Fluorescence Spectrophotometer and a 0.7 mL SUPRASIL® quartz cuvette (optical path length: 10 mm). Samples were excited with a fixed λ_{ex} = 450 nm (corresponding to λ_{abs} minimum of the Atto594 acceptor to minimise direct excitation background) and the fluorescence spectra were collected from 500 to 750 nm, with intervals ($\Delta\lambda$) of 1 nm. Excitation and emission slit widths and PMT were adjusted to ensure that fluorescence signals were always within

the linear response range of the instrument to ensure high measurement accuracy. While this would affect the absolute fluorescence signals for both the donor (quantum dot/rod QD/R) and acceptor (Atto594), the FRET ratio used in affinity evaluation would be unaffected, due to its ratiometric character. Binding assay samples were prepared by adding labelled protein to glycan-QD/R in a solution of binding buffer with BSA (1 mg/mL), and were incubating for 20 min before measurement. The large excess of BSA was used not only to greatly minimise any nonspecific interactions, but also to drastically reduce the adsorption of the QR and proteins on surfaces, which is especially important for measuring samples at low concentrations (e.g. 20 nM or below) where non-specific surface adsorption can significantly reduce the true solution concentration. All samples were performed in duplicates and each fluorescence spectrum was corrected by subtracting the background spectrum of the labelled lectin, without QR-DiMan, under identical conditions. FRET ratios (F) were obtained as the ratio between the maximum intensity observed from the fluorescence of the acceptor (Atto594; I_{A} at ~627 nm) to that observed from the donor (QD/R; I_{D} at ~550/560 nm, respectively) of the background corrected fluorescence spectra ($F = I_{\mathrm{A}}/I_{\mathrm{D}}$).

2.2.2.2 Hydrodynamic size analysis

All hydrodynamic size measurements were performed by dynamic light scattering (DLS) using a Malvern Zetasizer Nano and 10 mm PMMA cuvettes. Distributions were obtained by averaging a minimum of 3 measurements of 10 runs of 10 s. Mean hydrodynamic diameter, D_{h}, values were obtained by fitting these averaged volume percentage hydrodynamic size distributions with Gaussian (**Equation 2.12**) or lognormal (**Equation 2.14**) distribution curves, as specified. Where, y denotes the volume percentage intensity, x denotes the hydrodynamic size, x_c is the hydrodynamic size distribution centre, w is the standard deviation, A is the integrated area, and FWHM is the full width at half of the maximum amplitude (Gaussian FWHM: **Equation 2.13**; lognormal FWHM: **Equation 2.15**). All D_{h} values are reported as x_c ± ½ FWHM. All protein containing samples were performed using wild type lectin.

$$y = \frac{A}{w\sqrt{\frac{\pi}{2}}} \cdot e^{\frac{-2(x-x_c)^2}{w^2}} \qquad 2.12$$

$$\mathrm{FWHM} = w\sqrt{\ln(4)} \tag{2.13}$$

$$y = \frac{A}{w\sqrt{2\pi}} \cdot \frac{e^{\frac{-\left(\ln\left(\frac{x}{x_c}\right)\right)^2}{2w^2}}}{x} \tag{2.14}$$

$$\mathrm{FWHM} = e^{\ln(x_c)-w^2+w\sqrt{\ln(4)}} - e^{\ln(x_c)-w^2-w\sqrt{\ln(4)}} \tag{2.15}$$

2.2.2.3 STEM analysis

All STEM images were taken using an FEI Titan^3 Themis 300 G2 S/TEM and analysed using analysed using high angle annular dark field scanning transmission electron microscopy (HAADF STEM) mode (performed by Dario Fernandez Ainaga and Nicole Hondow at the University of Leeds).[1, 7] This provides atomic number contrast, permitting imaging of the high atomic number metal nanoparticles (brighter) on the low atomic number background (darker). 3.5 µL of each sample was loaded onto a plasma-cleaned TEM grid with a continuous carbon support film, before blotting and plunge-freezing into liquid ethane. The TEM grids were then warmed to room temperature over several minutes by placing the specimens in a liquid nitrogen cooled storage container in a rotary pumped vacuum desiccator. The samples were then plasma cleaned for 15 s before measurement.

2.2.2.4 Phenol-sulphuric acid carbohydrate quantification

The average number of ligands conjugated to the QR surface was determined using the phenol-sulphuric acid method of carbohydrate quantification.[8] Calibration samples were prepared in duplicates by adding phenol (80 µL, 5 % w/w in H_2O) and sulphuric acid (400 µL) simultaneously to known concentrations of N_3-EG_2-DiMan ligand precursor (80 µL in H_2O). After vortexing briefly, the samples were left at r.t. for 30 min, and their absorbance spectra were recorded by UV-vis spectroscopy using a SUPRASIL® quartz cuvette (path length: 10×2 mm). A calibration curve of absorbance at 490 nm *vs* concentration was then constructed which was fitted with a linear fit, obtaining an extinction coefficient at 490 nm of 18900 ± 300 M^{-1} cm^{-1}.

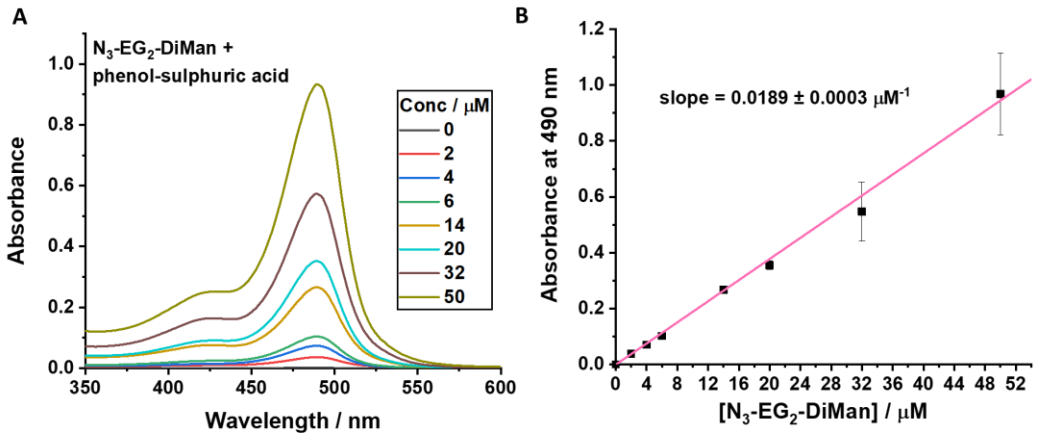

Figure 2.9 (A) The background corrected absorption spectra of varying concentrations of N3-EG2-DiMan samples after reaction with 5 % phenol and sulphuric acid (1:5 v/v ratio) and (B) a plot of sample absorbance at 490 nm against N_3-EG_2-DiMan concentration, fitted by a linear relationship, giving $y = (0.0189 \pm 0.0003) \cdot x$; $R^2 = 0.998$.

The supernatants and washing filtrates obtained after each glycan-NP preparation, were combined, freeze-dried and dissolved in H_2O. Samples were then treated with phenol-sulphuric acid in the same manner as the calibration samples. The calibration extinction coefficient was then used to determine the amount of unbound free ligands for each sample, from the absorbance at 490 nm. The amount of ligand conjugated to each NP was then obtained by subtracting the number of unbound free ligands from the initial amount used for QD conjugation. The number of ligands per NP (N) is therefore the ratio of the amount of conjugated ligand to the amount of NP used (**Equation 2.16**; where n_{total} is the total number of moles used for cap exchange, n_{NP} is the final number of moles of NP, A_{490} is the absorbance of the unconjugated ligand stock at 490 nm, V is the volume of unconjugated ligand stock, ε_{490} is the extinction coefficient (18900 ± 300 $M^{-1}\,cm^{-1}$)

$$N = \frac{n_{\text{total}} - \frac{A_{490}V}{\varepsilon_{490}}}{n_{\text{NP}}} \qquad 2.16$$

References

1. Y. Guo, I. Nehlmeier, E. Poole, C. Sakonsinsiri, N. Hondow, A. Brown, Q. Li, S. Li, J. Whitworth, Z. Li, A. Yu, R. Brydson, W. B. Turnbull, S. Pohlmann and D. Zhou, *J. Am. Chem. Soc.*, 2017, **139**, 11833-11844.
2. Y. Guo, C. Sakonsinsiri, I. Nehlmeier, M. A. Fascione, H. Zhang, W. Wang, S. Pohlmann, W. B. Turnbull and D. Zhou, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 4738-4742.
3. D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo and D. Zhou, *J. Am. Chem. Soc.*, 2020, **142**, 18022-18034.
4. O. Beutel, J. Nikolaus, O. Birkholz, C. You, T. Schmidt, A. Herrmann and J. Piehler, *Angew. Chem. Int. Ed.*, 2014, **53**, 1311-1315.
5. H. Zhang, G. Feng, Y. Guo and D. Zhou, *Nanoscale*, 2013, **5**, 10307-10315.
6. R. P. Richter, R. Bérat and A. R. Brisson, *Langmuir*, 2006, **22**, 3497-3505.
7. N. Hondow, R. Brydson, P. Wang, M. D. Holton, M. R. Brown, P. Rees, H. D. Summers and A. Brown, *J. Nanoparticle Res.*, 2012, **14**.
8. S. K. Saha and C. F. Brewer, *Carbohydr. Res.*, 1994, **254**, 157-167.

Chapter 3
Polyvalent Glycan-Quantum Dots as a Multifunctional Tool for Revealing Thermodynamic, Kinetic and Structural Details of Multivalent Lectin-Glycan Interactions

The following chapter is based on the work from: Hooper, J. et al. *ACS Appl. Mater. Interfaces.* **2022**. 14 (42), 47385–47396.

The author of this thesis wrote the manuscript and conducted all material preparation, FRET experiments and FRET data analysis. STEM was performed and analysed by Dario Fernandez Ainaga and Dr Nicole Hondow. Ligand precursors were synthesised by Dr Yuanyuan Liu and Dr Darshita Budhadev. All plasmids were prepared in-house by Dr Yuan Guo.

3.1 Introduction

Lectin-glycan interactions (LGIs) are widespread and play a pivotal role in biology. As individual LGIs are intrinsically weak, and hence mostly biologically inactive, most lectins form multimeric structures, allowing them to bind multivalently with multivalent glycans to enhance affinity and form biologically relevant interactions.[1] In the immune system, multivalent LGIs (MLGIs) are employed to recognise pathogen associated glycan patterns as a means of activating the host immune defences against infection.[2, 3] However, undesirable nonspecific activation can also lead to persistent inflammation and tissue death.[4, 5] MLGIs are also exploited by pathogens (*e.g.* viruses, bacteria and fungi) to establish attachment on host cells to initiate infection or by cancer cells to suppress host immunity to assist cancer development.[6, 7] Therefore, understanding the mechanisms behind MLGIs is of great importance.

Multivalent glycans are widely employed as research probes for MLGI mechanisms as well as potential therapeutics against specific MLGIs.[8-16] Here, the binding mode between MLGI binding partners is critical. When MLGI binding partners have perfect spatial and/or orientation matches, they will form simultaneous multivalent binding and give a great affinity enhancement, and hence effective therapeutic intervention.[17-22] Whereas, those that do not have such spatial/orientation matches may inter-cross-link with each other, giving rise to a relatively low affinity enhancement and a less effective therapeutic result.[19-21] However, information regarding the majority of MLGI binding

modes and how different binding modes affect the affinity and underlying binding thermodynamics and kinetics remains largely unexplored. This is presumably due to limitations of current biophysical techniques in probing such complex and flexible interactions. For example, isothermal titration calorimetry (ITC)[23, 24] and surface plasmon resonance (SPR)[25] are two of the most widely employed techniques to study the thermodynamics and kinetics of binding interactions, including MLGIs. However, ITC cannot accurately determine the affinity of very strong binding interactions (*e.g.* with an equilibrium dissociation constant, K_{d}, at the low nanomolar level or below).[26, 27] Moreover, cross-linking interactions can make it difficult to interpret the ITC data.[28] It is also difficult to dissect how individual LGIs contribute and control the overall MLGI affinity and specificity by SPR, because these are strongly affected by the density and orientation of the surface immobilised binding partner.[29] In addition, SPR measures the binding interactions occurring at the surface−solution interface, a very different environment from that in solution. Hence, the kinetic data measured by SPR may not be directly transferrable to that in solution. Thus, these conventional biophysical techniques can only provide some, but not a whole set, of key biophysical parameters such as the individual lectin binding thermodynamics, kinetics, binding modes, and binding site orientations, which are important for both the fundamental understanding and therapeutic development against specific MLGIs.

Meanwhile, over the past two decades, the strongly fluorescent quantum-dots (QDs) have emerged as a powerful probe for biological and biomedical research. In particular, the QD's strong and robust fluorescence has been widely exploited as sensitive QD-FRET (Förster resonance energy transfer) readouts in broad biosensing, bioanalytical and diagnostic assays, as well as bioimaging applications.[30-36] Compared to other readout strategies, QD-FRET has the advantages of high sensitivity, simple, separation free detection, and excellent assay robustness because of its ratiometric character. In this regard, our group has recently demonstrated that densely glycosylated QDs (glycan-QDs) can be harnessed as powerful mechanistic probes for MLGIs. We have shown that glycan-QDs can not only provide quantitative MLGI binding affinities *via* QD-FRET readout, but can also dissect their exact binding modes by S/TEM imaging of binding-induced QD assemblies.[19, 20, 37] Using the tetrameric lectins, DC-SIGN and DC-SIGNR (collectively denoted as DC-SIGN/R, hereafter) as model lectins, it was found that despite sharing ~80% amino acid identity, an overall tetrameric architecture and identical monovalent mannose binding motifs,[38, 39] their binding properties with mannose-α-1,2-mannose (DiMan) capped QDs

(QD-DiMan) are very different: DC-SIGN binds strongly *via* simultaneous tetravalent binding, while DC-SIGNR binds much more weakly (>100-fold lower affinity) *via* inter-cross-linking (**Figure 3.1**).[20] It was also revealed that QD-DiMan only potently blocks DC-SIGN-mediated, but not DC-SIGNR-mediated, virus infections, and their potencies are positively linked to their lectin binding affinities.[20] These differences were attributed to subtle orientation differences in their four carbohydrate-recognition domains (CRDs), where, all four CRDs point uprightly, away from the coiled-coil tetramerisation neck, in DC-SIGN, but those in DC-SIGNR are split into two pairs and point sideways **(Figure 3.1**).[19, 20] The CRD orientation difference may account for their distinct glycan binding and virus transmitting properties. For instance, DC-SIGN was found to be more effective in transmitting some HIV strains than DC-SIGNR,[40] while only DC-SIGNR, but not DC-SIGN, could effectively transmit West Neil Virus for infection.[41]

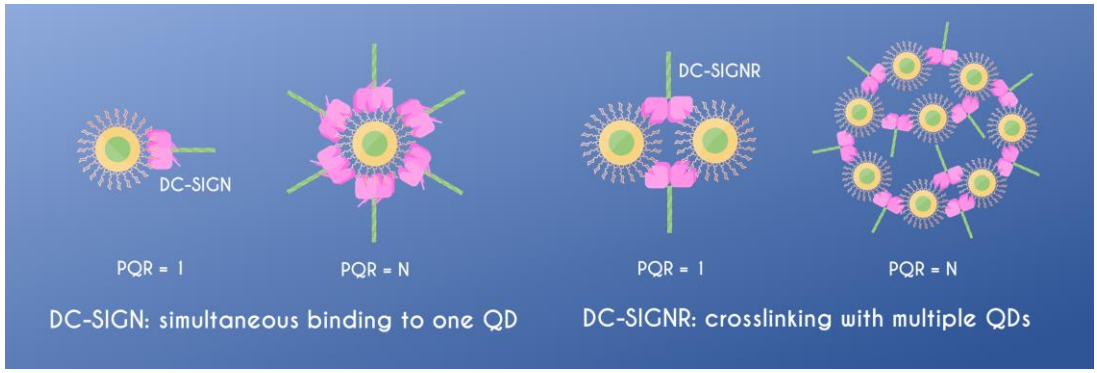

Figure 3.1 Schematic showing the different binding modes for DC-SIGN/R leading to different QD assemblies. The simultaneous tetravalent DC-SIGN·glycan-QD binding leads to individual QD particles, even at high protein:QD ratios (PQRs), whereas the crosslinking between DC-SIGNR and glycan-QDs results in a multiple QDs being assembled together in large-scale assemblies.

The close structural similarity and monovalent mannose specificity yet distinct multivalent QD-DiMan binding modes makes DC-SIGN/R a perfect pair of model lectins to study how binding modes affect MLGI binding thermodynamics and kinetics, as well as the structural bases behind these differences. Moreover, DC-SIGN/R play a key role in facilitating the infection of a wide range of viruses, *e.g.* HIV, HCV, Ebola, Zika and more recently SARS-CoV-2.[40-44] Thus their MLGI biophysical parameters are of great importance and biological significance, not only for understanding their basic structural and biophysical mechanisms but also for guiding the design of multivalent glycan entry inhibitors for blocking DC-SIGN/R-mediated viral

infections. This antiviral mode can avoid virus mutation and the development of resistance, and thus can be advantageous over other antiviral strategies.[8, 9, 13-15] In addition, DC-SIGN-targeting multivalent glycans can be harnessed as potential therapeutics against cancer, allergy and other immune dysregulation diseases, by exploiting DC-SIGN's powerful immune regulation functions.[2, 3, 6, 7]

In this chapter the capability of QD-DiMan MLGI probes is significantly extended by studying their binding thermodynamics *via* measuring their temperature-dependent binding affinities with DC-SIGN and DC-SIGNR in combination with Van 't Hoff analysis. Furthermore, stopped flow fluorescence spectroscopy has been employed to study their binding kinetics. Additionally, these techniques have been applied to identify that a 16 amino acid segment located at the C-terminus of DC-SIGN, which is absent in DC-SIGNR and plays an important role in DC-SIGN's ability in HIV transmission,[40] is critical in defining DC-SIGN's binding thermodynamics and binding mode. This work thus provides a significant development in establishing glycan-QDs as a powerful new platform for studying a MLGIs, extending their capability to probe a range of biophysical parameters, mechanisms, and protein structure-function relationships.

3.2 Experimental

3.2.1 QD-EG_n-Glycan

Cap exchange procedure was adapted from previously described methods of lipoic acid ligand conjugation.[19-20] DHLA-EG_{11}-DiMan was prepared as described in **Section 2.1.3**.

QD-EG_{11}-DiMan. QD_{550} (53 µM in toluene, 1.2 nmol) was precipitated by adding EtOH (1.2 ml) followed by centrifugation at 15000×g for 10 min. The clear supernatant was discarded and the pellet was dissolved in $CHCl_3$. DHLA-EG_{11}-DiMan (2.5 mg, 1.8 µmol) in $CHCl_3$, NaOH (0.1 M in EtOH, 2.2 µmol) and MeOH were then added quickly to the QD solution and the reaction was left to stir, covered, at r.t. for 30 min. Hexane was then added until precipitation was observed and the suspension was then centrifuged at 15000×g for 3 min. The clear supernatant was carefully removed and stored for glycan valency evaluation. The QD pellet was dissolved in H_2O, transferred to a 30 kDa MWCO spin filter and then washed 3× with H_2O, each using a 2 min centrifugation at 15000×g, to remove any unbound free ligands. The supernatant above and the washing solvents were collected and stored for

sulfur-phenol quantification of the surface carbohydrate coverage (see **Section 2.2.2.4**). Mean hydrodynamic diameters (D_h) were obtained by DLS. This yielded a stable QD-DiMan in aqueous solution with a D_h of 12.4 ± 3.0 nm and 210 ± 70 glycans ligands per particle.

QD-EG$_3$-OH. QD$_{550}$ (80 µM in toluene, 2 nmol) was precipitated in the presence of EtOH (1.2 ml) and centrifuged at 15000×g for 10 min. The supernatant was discarded and the pellet was dissolved in $CHCl_3$. DHLA-EG$_3$-OH (1.0 mg, 1.4 µmol) in $CHCl_3$, NaOH (0.1 M in EtOH, 1.68 µmol) and MeOH were then added quickly and the reaction was left to stir, covered, at r.t. for 30 min. Hexane was then added until precipitation was observed, and the suspension was then centrifuged at 15000×g for 3 min. The QD pellet was then dissolved in H_2O followed by washing 3× with H_2O using a 30 kDa MWCO spin filter at 15000×g for 2 min as above. This yielded a stable QD-OH in aqueous solution with a D_h of 9.0 ± 2.9 nm.

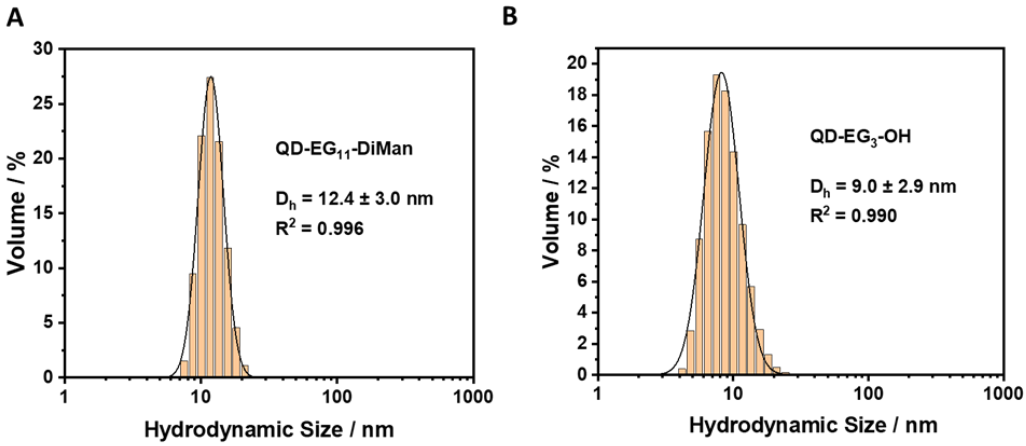

Figure 3.2 Volume populations hydrodynamic size distributions fitted with lognormal Gaussian distribution curves for (A) QD capped with DHLA-EG$_{11}$-DiMan (QD-DiMan) and (B) QD capped with DHLA-EG$_3$-OH (QD-OH).

3.2.2 DC-SIGN, DC-SIGNR and DC-SIGN-C

The DC-SIGN and DC-SIGNR were produced and purified as described in **Section 2.1.4.1**, and their cysteine point mutation variants, DC-SIGN Q274C and DC-SIGNR R287C, were produced, purified and labelled with maleimide-Atto594 as described in **Section 2.1.4.2**. Plasmids encoding for DC-SIGN-C and DC-SIGN-C Q274C variants were prepared by deleting 16 amino acids at the C-terminal of DC-SIGN and DC-SIGN Q274C, respectively, confirmed by DNA sequencing. These proteins were produced and purified as

described in **Section 2.1.4.1** and labelled as described **Section 2.1.4.2**, respectively. The amino acid identities are shown below, whereby asterisks denote residues of similarity between DC-SIGN/-C and DC-SIGNR, and the location of the cysteine mutation is shown in bold.

```
  DC-SIGN: KVPSSISQEQSRQDAIYQNLTQLKAAVGELSEKSKLQEIYQELTQLKAAVGELPEKSKLQEIYQELTRLKAA
DC-SIGN-C: KVPSSISQEQSRQDAIYQNLTQLKAAVGELSEKSKLQEIYQELTQLKAAVGELPEKSKLQEIYQELTRLKAA
 DC-SIGNR: KVPSSLSQEQSEQDAIYQNLTQLKAAVGELSEKSKLQEIYQELTQLKAAVGELPEKSKLQEIYQELTRLKAA
           ***** ***** ************************************************************

  DC-SIGN: VGELPEKSKLQEIYQELTWLKAAVGELPEKSKMQEIYQELTRLKAAVGELPEKSKQQEIYQELTRLKAAVGEL
DC-SIGN-C: VGELPEKSKLQEIYQELTWLKAAVGELPEKSKMQEIYQELTRLKAAVGELPEKSKQQEIYQELTRLKAAVGEL
 DC-SIGNR: VGELPEKSKLQEIYQELTRLKAAVGELPEKSKLQEIYQELTRLKAAVGELPEKSKLQEIYQELTELKAAVGEL
           ****************** ************* ********************** ******** ********

  DC-SIGN: PEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQELTQLKAAVERLCHPCPWEWTFFQGNCYFMSNSQRNWHD
DC-SIGN-C: PEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQELTQLKAAVERLCHPCPWEWTFFQGNCYFMSNSQRNWHD
 DC-SIGNR: PEKSKLQEIYQELTQLKAAVGELPDQSKQQQIYQELTDLKTAFERLCRHCPKDWTFFQGNCYFMSNSQRNWHD
           ***** ******** *********  **** ****** ** * ***** **  ********************

  DC-SIGN: SITACKEVGAQLVVIKSAEEQNFLQLQSSRSNRFTWMGLSDLNQEGTWQWVDGSPLLPSFKQYWNRGEPNNVG
DC-SIGN-C: SITACKEVGAQLVVIKSAEEQNFLQLQSSRSNRFTWMGLSDLNQEGTWQWVDGSPLLPSFKQYWNRGEPNNVG
 DC-SIGNR: SVTACQEVRAQLVVIKTAEEQNFLQLQTSRSNRFSWMGLSDLNQEGTWQWVDGSPLSPSFQRYWNSGEPNNSG
           * *** ** ******* ********** ****** ********************** ***  *** ***** *

  DC-SIGN: EEDCAEFSGNGWNDDKCNLAKFWICKKSAASCSRDEEQFLSPAPATPNPPPA
DC-SIGN-C: EEDCAEFSGNGWNDDKCNLAKFWICKKSAASCSRDE
 DC-SIGNR: NEDCAEFSGSGWNDNRCDVDNYWICKKPAA-CFRDE
            ******** ****  *     ***** ** *  **
```

Proteins were characterised by UV-vis spectroscopy and protein molecular weights (MW) were confirmed from deconvoluted HRMS data, where calculated protein MW values were obtained from the amino acid sequences of one monomeric unit of the corresponding protein and labelling efficiency (LE) was obtained from the ratio of the labelled protein to unlabelled protein peak areas. The wild type tetrameric protein D_h was confirmed by DLS analysis.

DC-SIGN-C: HRMS: calculated MW [DC-SIGN-C]: 37581.41, found 37585.66. DLS: D_h 14.5 ± 3.5 nm.

DC-SIGN-C-Atto594: HRMS: calculated MW [DC-SIGN-C Q274C + 2Ca] 37636.60 and [DC-SIGN-C Q274C-Atto594] 38484.42; found 37636.62 and 38490.71. LE: 75 %.

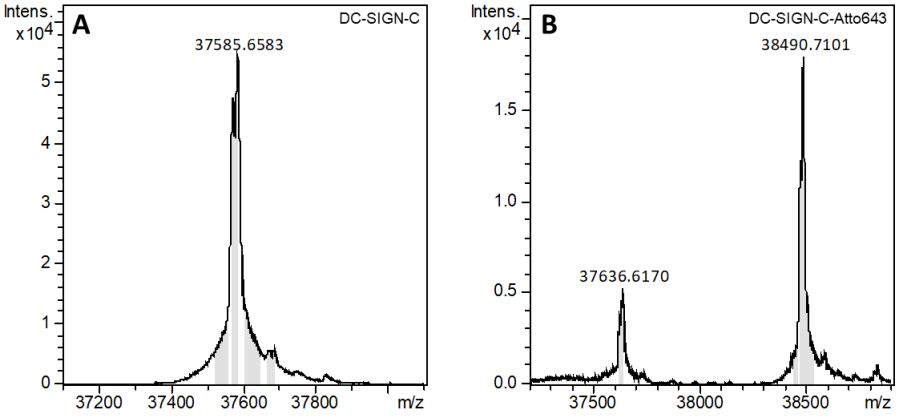

Figure 3.3. Deconvoluted HRMS spectra of the monomeric extracellular domains of (A) DC-SIGN-C, (B) DC-SIGN-C-Atto594.

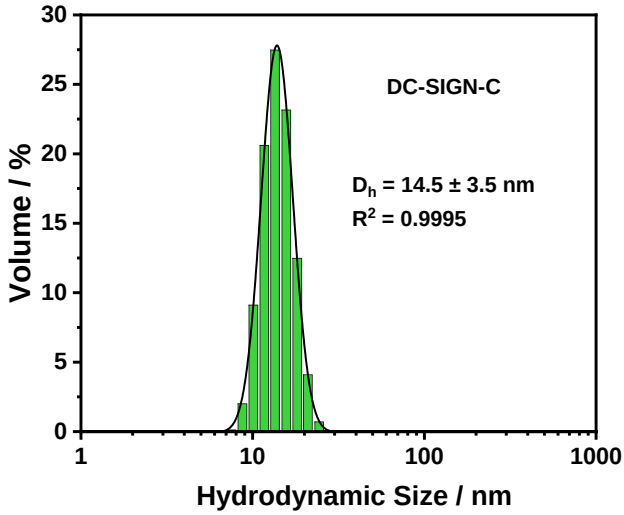

Figure 3.4. Volume population hydrodynamic size distribution histograms fitted with a lognormal Gaussian distribution for (A) DC-SIGN-C at 40 nM in binding buffer (D_h values given as mean ± ½FWHM).

3.2.3 QD-FRET thermodynamic assays

FRET assays were performed as described in **Section 2.2.2.1**. Affinity measurements were obtained by measuring the direct excitation background corrected FRET spectra of lectin and glycan-QD at concentrations of 0.5-60 nM at a protein to QD ratio (PQR) of 1:1 for DC-SIGN and DC-SIGN-C, and concentrations of 0-800 nM at a PQR of 10:1 for DC-SIGNR. Thermodynamic measurements were obtained by repeating these measurements at temperatures of 20, 25 and 30 ^{o}C. Buffer and sample temperatures were controlled by incubation in a water bath or dry bath, respectvely. The cuvette

temperature was maintained by a Cary single cell Peltier accessory temperature control unit, where temperature was sustained using a water pump cooling system.

3.2.4 QD-FRET kinetic assays

Kinetic FRET assays were performed using a TgK scientific SFA-20 Rapid Kinetics stopped-flow accessory in conjunction with a Cary Eclipse Fluorescence Spectrophotometer. Measurements were taken by exciting the sample at λ_{ex} = 450 nm and taking subsequent measurements of the fluorescence intensities at λ_{em} = 550 and 628 nm over time, with a time resolution (Δt) of 0.0125 s. The apparatus consists of two syringes (A and B) fed to a 80 µL high grade Spectrasil B cuvette *via* capillary tubes, which then continues to a switch which is triggered upon injection of 0.3 mL of sample. Before measurement, the system was pre-flushed with H_2O (40 mL), followed by syringe A with BSA (1 mg/ml in binding buffer, 2 mL) and syringe B with binding buffer (2 mL), and finally both syringes with binding buffer (10 mL). All solutions used for measurements were made in binding buffer containing His_6-Cys (5 µg/mL). His_6-Cys has been previously found to enhance fluorescence and reduce nonspecific interactions for the QDs, [19, 20, 45] thus was used instead of BSA which may inhibit collision events under high concentrations. Association measurements were obtained by loading syringe A with QD-DiMan (2.5 mL, 40 nM in buffer) and syringe B with protein (2.5 mL, 40 nM in buffer). Dissociation measurements were obtained by loading syringe A with solution of QD-DiMan and protein (2.5 mL, 40 nM in buffer, PQR 1:1), which had been pre-incubated for 20 min, and syringe B with D-mannose (2.5 mL, 40 mM). Background measurements were obtained by loading syringe A with buffer (2.5 mL) and syringe B with protein (2.5 mL, 40 nM in buffer). For each run, the system was flushed with sample (1.5 mL per syringe) before starting measurements. Each measurement was ran for 60 s before the next injection, where buffer was used to displace the sample once the sample had been completely injected, until the fluorescence signal was observed to drop. Corrected fluorescence intensity profiles were obtained for both association and dissociation experiments by subtracting the background time profiles taken at corresponding injection volumes and λ_{em}, and averaging 3 measurements at each λ_{em} possessing consistent fluorescence intensity plateau values. FRET ratio time profiles were then obtained by the ratio between the averaged corrected fluorescence signals at λ_{em} = 628 nm (I_{A}) and 550 nm (I_{D}), over time. The kinetic profiles for DC-

SIGNR at a 1:1 PQR showed low signal-to-noise due to its weaker binding interaction, thus the data were smoothed by averaging every five time points, giving a Δt of 0.0625 s.

3.2.5 QD STEM imaging

Samples were prepared by adding wild type lectin to a solution of QD-DiMan (40 nM in binding buffer) at PQRs of 1:1 for DC-SIGN and DC-SIGN-C, or 10:1 for DC-SIGNR, and incubating for 20 min. STEM images were obtained as described in **Section 2.2.2.3**.

Software based approaches to analyse the particles was complicated by the significant overlap of particles in the samples with agglomeration (in particular for DC-SIGNR). Consequently, a manual approach to identify individual or groups of QDs was undertaken using the following criteria: (1) Particles were considered to be 'clustered' if there was no clear boundary between their edges at the magnification and focus of the relevant image (*i.e.* $\lesssim$1 nm apart); (2) Once a particle had been counted it was marked to avoid the possibility of double counting. Due to the high contrast difference between the Cd-containing QDs and the carbon support film of the TEM grid images of an appropriate magnification permitted the analysis of between 170 and 400 particles per image.

3.3 Results and Discussion

3.3.1 Materials synthesis and characterisation.

A dihydro-lipoic acid-undeca(ethylene glycol)-mannose-α-1,2-mannose (DHLA-EG_{11}-DiMan) based multifunctional glycan ligand (**Figure 3.5**) was synthesised using our previous procedures.[20] Additionally, a DHLA-tri(ethylene glycol) based ligand with a terminal di(ethylene glycol) group (denoted as DHLA-EG_3-OH) was also synthesised and used as a negative control (**Section 2.1.3** for chemical structure). Each glycan ligand contains three functional domains: a DHLA group for strong QD anchoring *via* chelative zinc-thiolate coordination;[31] a flexible EG_{11} linker for imposing high water solubility, excellent stability and resistance against nonspecific adsorption;[46, 47] and a terminal DiMan group for specific DC-SIGN/R binding.[20]

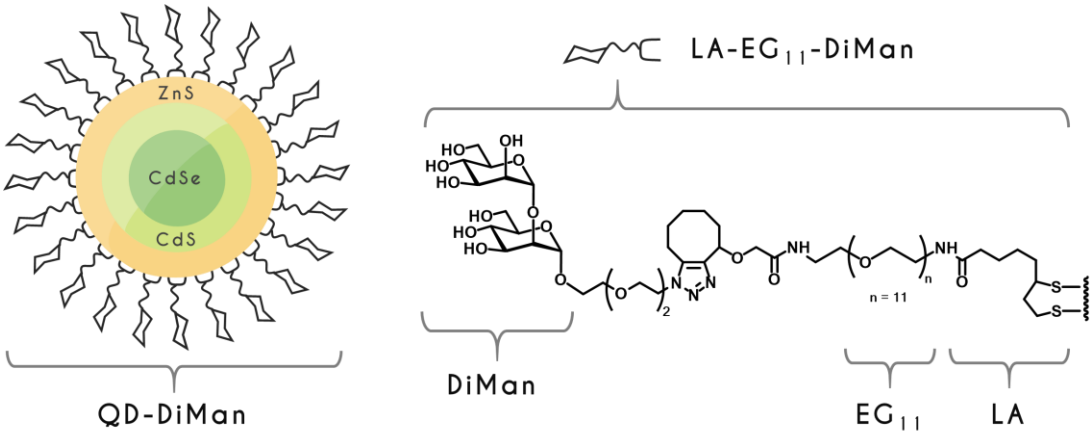

Figure 3.5 Schematic of DHLA-EG$_{11}$-DiMan coated CdSe/CdS/ZnS core/shell/shell QDs (QD-DiMan; left) with the chemical structure of deprotonated DHLA-EG$_{11}$-DiMan (LA-EG$_{11}$-DiMan; right).

A CdSe/CdS/ZnS core/shell/shell QD (λ_{em} ~550 nm, core diameter ~3.9 nm, quantum yield = 62%) was employed to construct the glycan-QDs. It also acted as the donor for developing the QD-FRET-based binding assays. Cap-exchange using deprotonated DHLA-EG$_{11}$-DiMan in a homogeneous solution was employed to make DHLA-EG$_{11}$-DiMan capped QD (denoted as QD-DiMan, hereafter) as reported previously.[19, 20] A QD capped with the DHLA-EG$_3$-OH control ligand (denoted as QD-OH hereafter) was also prepared as a negative control for lectin binding. Both QDs were found to be monodisperse, relatively compact (with hydrodynamic diameters, D_h, of ~12 and ~9 nm for QD-DiMan and QD-OH, respectively; **Figure 3.2**) and highly stable. No changes of physical appearance or precipitation were observed after storage for 1 month in a fridge. The average glycan valency per QD was estimated as 210 ± 70 by measuring the difference between the amount of ligand added and that remaining unbound post cap-exchange (**Section 2.2.2.4**).[20] By calculating the average deflection angle of the glycan ligands, and using the D_{h} value above, the average inter-glycan distance was estimated to be 1.7 ± 0.3 nm (**Appendix B.1**).[36]

The soluble extracellular segments of DC-SIGN/R, which have been shown to faithfully retain the tetrameric structure and glycan binding properties of full length proteins, were used in all glycan-QD binding studies.[19, 20] DC-SIGN with its C-terminal 16 amino acids truncated (denoted as DC-SIGN-C hereafter) was constructed using standard molecular biology techniques and the construct was confirmed by DNA sequencing. Proteins used in FRET assays were modified with a site-specific cysteine mutation at Q274 in DC-SIGN and DC-SIGN-C or R287 for DC-SIGNR, enabling them to be labelled with maleimide-modified Atto594 dye, as described previously, which acted

as a FRET acceptor.[19, 20] The labelling positions lie outside of the lectins' glycan binding pockets; hence, Atto594 labelling does not affect their glycan binding properties as confirmed previously.[20] The QD-Atto594 FRET pair have good spectral overlap with a respectable Förster radius (R_0 ~5.7 nm, **Appendix B.2**), ensuring that efficient FRET can occur provided that the glycan-QD and labelled lectin are bound to one another. Moreover, there is little overlap between the QD and dye emission spectra, making it straightforward to differentiate donor and acceptor fluorescence without the need of spectral deconvolution.[20] The proteins were characterised by high resolution mass spectrometry (HRMS), UV-vis spectroscopy and dynamic light scattering (DLS). All three lectins were found to form stable tetramer with comparable hydrodynamic diameters (D_h) of ~14 nm (**Figure 2.3; Figure 3.4**) The average dye labelling efficiency was ~85% for all DC-SIGN/R and ~75% for DC-SIGN-C (**Figure 2.4; Figure 3.3B**).

3.3.2 Quantifying binding affinity and thermodynamics *via* QD-FRET.

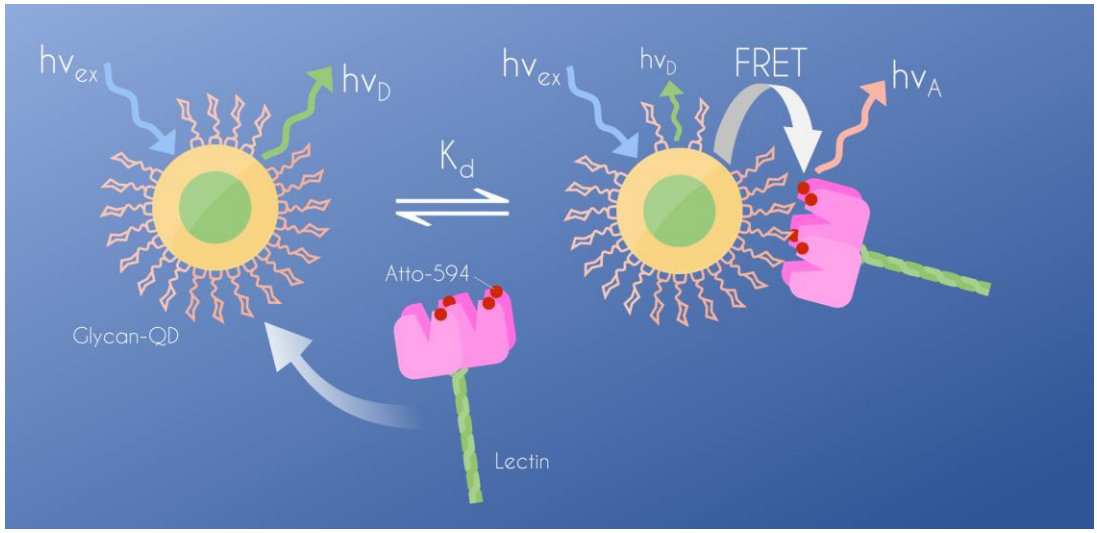

Figure 3.6 Schematic depicting the FRET process between an excited glycan-QD (donor) and lectin labelled with Atto594 (acceptor), where $h\nu_{ex}$ is the excitation energy, $h\nu_D$ is the donor emission energy and $h\nu_A$ is the acceptor emission energy.

The principle of the QD-FRET readout for quantifying the DC-SIGN/R (Atto594 labelled) binding with QD-DiMan is shown schematically in **Figure 3.6**. Since FRET can only happen over a short distance (*e.g.* < 10 nm), any unbound lectins (acceptors) would be too far away to participate in FRET interactions with the QD donor and hence cannot contribute to the FRET signal. Thus, the observed FRET signal is directly linked to the equilibrium of

QD-DiMan-lectin binding, and more specifically, the amount of lectins bound to the QD. This is a distinct advantage of QD-FRET readout, allowing for binding assays to be performed in homogenous solutions without the need of purification or separation.[34, 35, 48] The apparent binding equilibrium dissociation constants, K_{d} (the inverse of the equilibrium association constant, K_{a}, *i.e.* $K_{\mathrm{d}} = 1/K_{\mathrm{a}}$), between QD-DiMan and the lectins were measured *via* our recently established method.[20] Briefly, the fluorescence spectra of pre-mixed QD-DiMan-lectin samples with varying protein concentration, but under a fixed protein:QD molar ratio (PQR) of 1:1 for DC-SIGN or 10:1 for DC-SIGNR, were recorded at a fixed excitation wavelength (λ_{ex}) of 450 nm. This λ_{ex} corresponds to the absorption minimum of the Atto594 receptor, thereby minimising the dye direct excitation background. A higher PQR for DC-SIGNR was used to compensate for its relatively low FRET signal, due to weak binding.

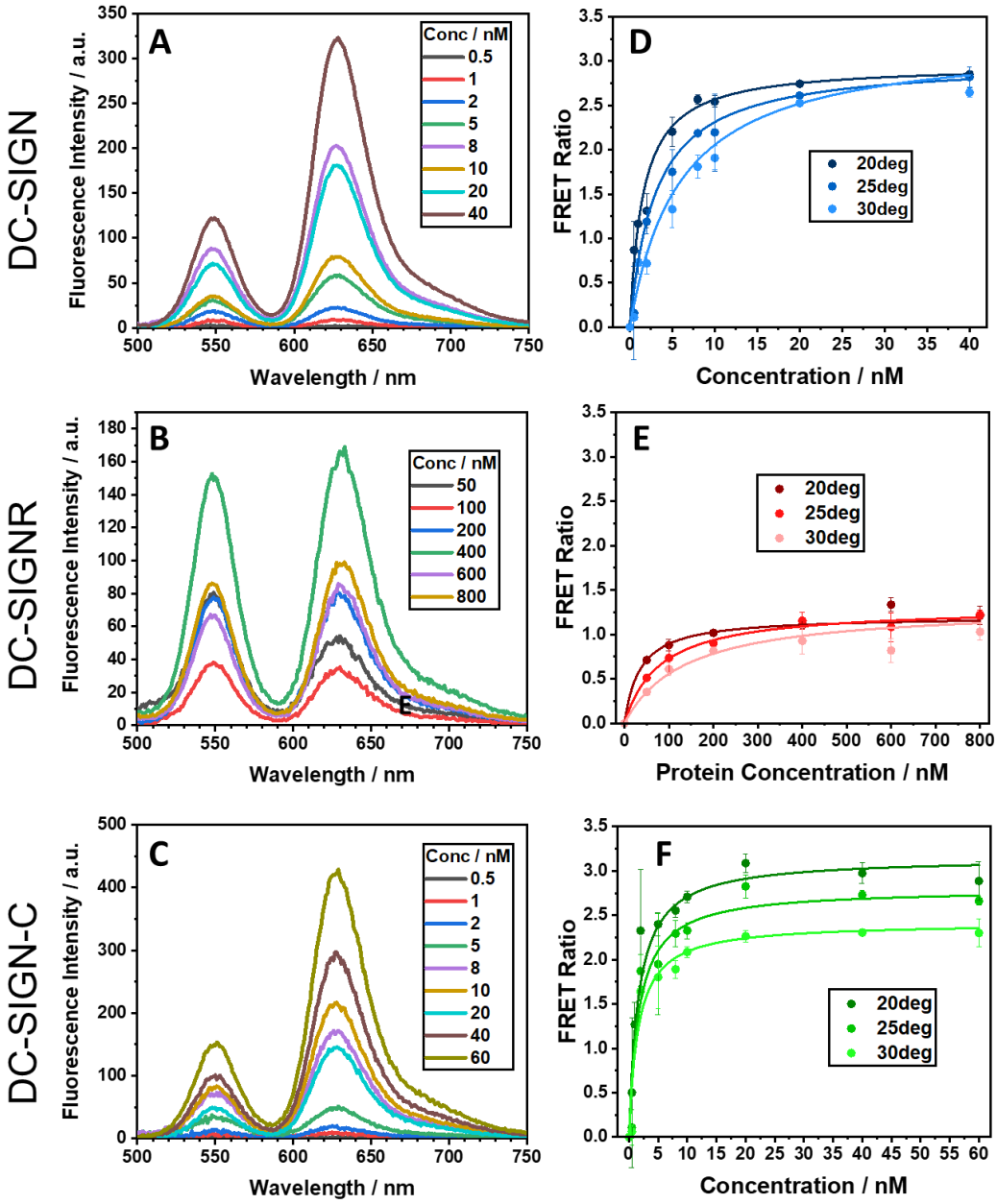

Figure 3.7 Background-corrected fluorescence spectra of different concentrations of a mixture of QD-DiMan with Atto594 labelled lectins at 20 ºC for (A) DC-SIGN (PQR 1:1), (B) DC-SIGNR (PQR 10:1), and (C) DC-SIGN-C at a (PQR 1:1); and the corresponding apparent FRET ratio−protein concentration relationships at three different temperatures, fitted by Equation 3.1, for (D) 1:1 mixed QD and DC-SIGN, (E) 1:10 mixed QD and DC-SIGNR, and (F) 1:1 mixed QD and DC-SIGN-C. Error bars represent the standard deviations (SDs) of triplicate experiments at each concentration.

Table 3.1 Fitting parameters for the FRET curves of QD-DiMan binding with labelled DC-SIGN, DC-SIGN-C and DC-SIGNR at varying temperatures (R^2 >0.99 for all fits, SDs represent fitting errors).

Protein	T / ºC	K_d / nM	F_{max}
DC-SIGN	20	1.54 ± 0.07	3.0 ± 0.1
	25	3.00 ± 0.04	3.00 ± 0.01
	30	5.9 ± 1.7	3.3 ± 0.2
DC-SIGN-C	20	1.62 ± 0.28	3.15 ± 0.08
	25	1.67 ± 0.48	2.80 ±0.07
	30	1.56 ± 0.50	2.42 ± 0.12
DC-SIGNR	20	35 ± 2	1.20 ± 0.02
	25	80 ± 6	1.31 ± 0.04
	30	130 ± 10	1.30 ± 0.09

The corresponding dye direct excitation corrected fluorescence spectra (**Figure 3.7A,B**) revealed that while both the fluorescence intensities of the QD donor (I_{D}; at ~550 nm) and Atto594 acceptor (I_{A}; at ~628 nm) were increased with increasing concentration, I_{A} increased more quickly than I_{D}, leading to an increasing apparent FRET ratio ($F = I_{\mathrm{A}}/I_{\mathrm{D}}$) at higher concentrations before reaching saturation (**Figure 3.7D,E**). In contrast, incubating the labelled lectins with the QD-OH control without the terminal glycan did not produce any noticeable FRET signals, confirming that the FRET signal observed here was due to specific lectin-glycan interactions (**Figure 3.8**). Furthermore, neither the QDs nor the labelled proteins exhibited any significant absorption at λ_{ex} of 450 nm which could affect the FRET measurement *via* inner filter effect. Their absorbance at 450 nm were <0.01 even at the highest concentrations used (80 nM of QD and 800 nM of protein), and their fluorescence intensity-concentration relationships were both perfectly linear across the range of concentrations studied (**Figure B.2**).

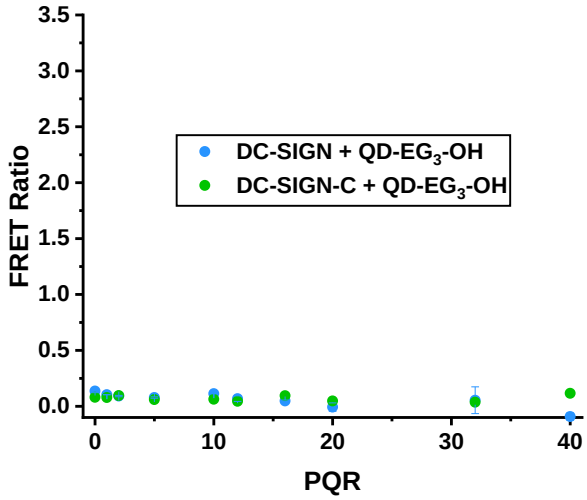

Figure 3.8 Plot of FRET ratio *vs* protein to QD molar ratio (PQR) obtained by titrating varying amounts of DC-SIGN (blue) or DC-SIGN-C (green) to a fixed concentration of QD-OH (20 nM). Neither plots showed significant FRET signals, confirming almost no non-specific interactions between the QD-OH control and Atto594 labeled DC-SIGN or DC-SIGN-C.

The apparent FRET ratio-concentration relationships were then fitted with the Hill equation (**Equation 3.1**) to derive the apparent binding K_{d} values (**Figure 3.7D-E, Table 3.1**). Where [P] is the protein concentration, F is the apparent FRET ratio, F_{max} is the maximal FRET ratio at saturated binding, and *n* is the Hill coefficient, which indicates binding cooperativity (*i.e.* n < 1, = 1, and > 1, indicate binding to be negatively-, none- and positively- cooperative, respectively). Here, n = 1 was assumed as most affinity assays were performed under a PQR of 1, therefore glycan-QDs should be bound with just one lectin; thus no intermolecular lectin-lectin interactions were expected to inhibit or promote the lectin-QD binding.[49]

$$F = \frac{I_{\mathrm{A}}}{I_{\mathrm{D}}} = F_{\mathrm{max}} \cdot \frac{[\mathrm{P}]_{\mathrm{bound}}}{[\mathrm{P}]_{\mathrm{total}}} = F_{\mathrm{max}} \cdot \frac{[\mathrm{P}]^n}{{K_{\mathrm{d}}}^n + [\mathrm{P}]^n} \qquad 3.1$$

Previously, most QD-FRET binding assays were performed by varying the amount of protein (or other binder) whilst maintaining a fixed QD concentration, to obtain the fluorescence-concentration relationships from which apparent K_{d} values were derived.[50] While such a method can provide accurate K_{d} values for weak binders (*e.g.* true K_{d} >> 50% of the QD's saturate binding capacity), it cannot provide accurate measurement for strong binders (*e.g.* true K_{d} < 50% of QD's saturate binding capacity) where the measured

K_d values will simply equal 50% of the QD's saturate binding capacity. In contrast, our above method does not have such limitations and can provide robust K_d measurements for both strong and weak binding partners. This is because the I_A/I_D ratio is linearly proportional to the amount (or fraction, under a fixed PQR) of lectins bound to the QD in the binding system, thus it is independent of the protein concentration or QD's binding capacity, making it a highly robust parameter for K_d quantification.[20, 37]

Consistent with our previous results, the binding affinity of DC-SIGN with QD-DiMan is very strong, with an apparent K_d of ~1.5 nM at 20 ºC, which is >20 fold stronger than that of DC-SIGNR.[20] To obtain the binding thermodynamics, each binding assay was repeated at three different temperatures (20, 25 and 30 ºC, **Figure 3.7D,E**; see **Figure C.1** for the detailed spectra), and their respective apparent K_d values were derived from the Hill fits (**Table 3.1**). Van 't Hoff plots were then constructed to extract the binding enthalpy and entropy changes by combining two Gibbs free energy equations, **Equation 3.2** and **3.3**, and taking a linear fit of the natural log of the K_d against the reciprocal of temperature, $1/T$ (**Equation 3.4**; where, R is the gas constant) as shown in **Figure 3.9A** ($\Delta H = mR$; $\Delta S = -cR$, where $y = mx + c$).

$$\Delta G^o = RT \ln(K_d) \tag{3.2}$$

$$\Delta G^o = \Delta H^o - T \cdot \Delta S^o \tag{3.3}$$

$$\ln(K_d) = \frac{\Delta H^o}{R} \cdot \frac{1}{T} - \frac{\Delta S^o}{R} \tag{3.4}$$

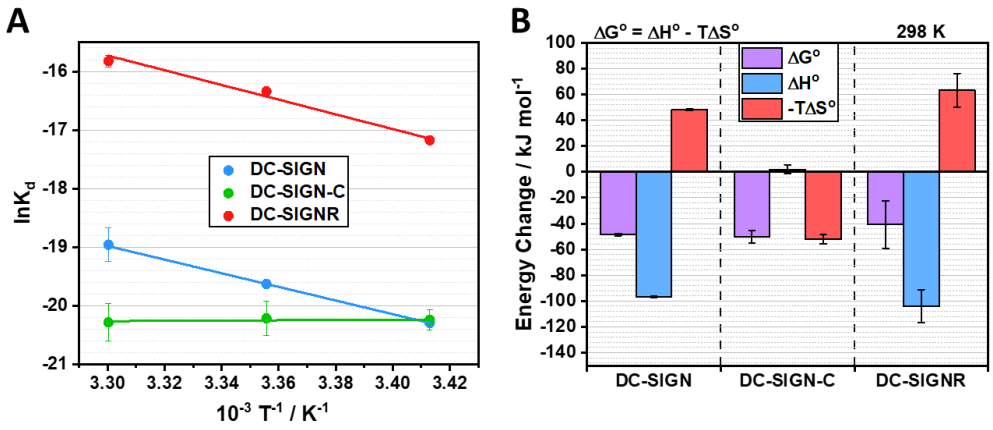

Figure 3.9 (A) Van ' t Hoff analyses of the $\ln(K_d)$ *vs* $1/T$ relationships for QD-DiMan binding with DC-SIGN (blue), DC-SIGN-C (green), and DCSIGNR (red). (B) Comparison of the standard (T = 298 K) enthalpy (blue), entropy (red), and Gibbs free energy (pink) changes of QD-DiMan binding with DC-SIGN, DC-SIGN-C and DC-SIGNR.

Table 3.2 Summary of the binding thermodynamic parameters for QD-DiMan binding with DC-SIGN, DC-SIGN-C, and DC-SIGNR (SDs are propagated from the fitting errors obtained in Figure 3.9, according to Equation 3.4).

Lectin [a]	ΔH^o / kJ mol^{-1}	ΔS^o / J mol^{-1} K^{-1}	ΔG^o / kJ mol^{-1} [b]
DC-SIGN	-96.8 ± 0.6	-161 ± 2	-48.6 ± 0.9
DC-SIGN-C	2 ± 4	174 ± 10	-50 ± 5
DC-SIGNR	-100 ± 10	-201 ± 34	-40 ± 20

[a] ITC measured ΔH^o, ΔS^o and ΔG^o values for CRD·DiMan monovalent binding are -25.8 kJ mol^{-1}, 28.5 J K^{-1} mol^{-1}, and -17.3 kJ mol^{-1}, respectively.[51]

[b] at 298 K.

The binding thermodynamic parameters for QD-DiMan binding with DC-SIGN/R are summarised in Table 3.2. Based on these data, two conclusions can be drawn. (1) Both DC-SIGN/R interactions with QD-DiMan are enthalpy driven with negative binding enthalpy change (ΔH^o) and negative binding entropy change (ΔS^o) terms. Both ΔH^o values for DC-SIGN/R are similar, at ~-100 kJ mol^{-1}, which is approximately 4-fold that of monovalent DC-SIGN CRD·DiMan binding measured by ITC (*e.g.* -25.8 kJ mol^{-1}).[51] Given that

glycan binding does not lead to conformational changes in CRD,[39, 52] and the QD-OH control without the terminal DiMan shows no measurable binding with DC-SIGN (**Figure 3.8**), these results suggest that all four CRDs in both DC-SIGN/R are engaged in glycan binding. This is to be expected for enthalpy driven MLGIs. The excellent consistency between the ΔH^{o} values measured here and that measured from ITC thus confirms our QD-FRET technique is a valid, sensitive new method for investigating the binding thermodynamics of MLGIs. (2) The multivalent binding ΔS^{o} values for QD-DiMan·DC-SIGN/R are ~5.7- and ~7-fold that of the monovalent DC-SIGN CRD·DiMan binding measured by ITC (-28.5 J mol^{-1} K^{-1}), respectively. Thus a larger entropic penalty for DC-SIGNR binding with QD-DiMan is responsible for its lower binding affinity compared to DC-SIGN. The total ΔS^{o} of multivalent interactions consists of changes in translational and rotational entropies with complexation; any changes associated with the conformational changes imposed on each species upon binding; and the entropy changes associated with the rearrangement of solvent molecules.[1] Given that all four CRDs in DC-SIGN/R are engaged in glycan binding and each CRD is most likely to accommodate a single DiMan molecule,[53, 54] entropic contributions from conformation change of monovalent CRD-DiMan binding should be very similar for both lectins. Therefore, the higher entropic penalty observed for DC-SIGNR over DC-SIGN upon binding with QD-DiMan is most likely due to a greater reduction of translational and rotational degrees of freedom by forming a smaller number of larger inter-cross-linked protein-QD complexes. The thermodynamic data obtained here are fully consistent with that expected for MLGIs with different binding modes (*i.e*. simultaneous binding *vs* crosslinking).

3.3.3 Role of C-terminal segment in DC-SIGN multivalent binding.

The CRDs in DC-SIGN/R are linked to the coiled-coil neck domain with some degree of flexibility.[55] A short C-terminal segment of 16-amino acids in length is found at the CRD-neck junction region in DC-SIGN, but it is absent in DC-SIGNR.[38] Thus, this segment may act as a steric wedge to maintain the upright CRD orientation of DC-SIGN, thus determining its multivalent binding properties. To probe this, DC-SIGN-C was constructed, labelled with Atto594 and applied in binding studies with QD-DiMan using the same methods described above.

Interestingly, DC-SIGN-C's overall QD-DiMan binding profile at 20 ^{o}C more closely resembles that of DC-SIGN over DC-SIGNR. (1) Its binding K_{d} is roughly the same as that of DC-SIGN (*e.g.* 1.6 ± 0.3 *vs* 1.54 ± 0.07 nM),

which is >20 fold lower (stronger) than that of DC-SIGNR (35 ± 2 nM; **Figure 3.7C,F; Table 3.1**). (2) Its maximum FRET ratio ($F_{\max}$) is also comparable to that of DC-SIGN, which is >2-fold that of DC-SIGNR despite the PQR used in the latter being 10-fold that of the former (**Figure 3.7F; Table 3.1**). Despite such similarities between DC-SIGN-C and DC-SIGN in QD-DiMan binding at 20 °C, their affinity-temperature dependencies are drastically different. While the K_{d} for QD-DiMan·DC-SIGN binding is increased ~4 fold (*i.e.* affinity is weakened ~4-fold) as temperature is increased from 20 to 30 °C, the K_{d} of DC-SIGN-C remains essentially unchanged. Moreover, the maximum FRET ratio for DC-SIGN binding remains almost constant, while that of DC-SIGN-C is decreased considerably with increasing temperature (**Figure 3.7D,F**).

The Van 't Hoff plot of QD-DiMan·DC-SIGN-C binding therefore shows little change in the $\ln(K_{\mathrm{d}})$ with changing $1/T$, and the binding ΔH° and $-T\Delta S^{\circ}$ (at 298 K) contributions are obtained to be 2 ± 4 and -52 ± 3 kJ mol^{-1}, respectively. This thermodynamic profile contrasts greatly with that of DC-SIGN (ΔH° = -96.8 ± 0.6 kJ mol^{-1} and $-T\Delta S^{\circ}$ = 48.1 ± 0.6 kJ mol^{-1}) or DC-SIGNR ΔH° = -100 ± 10 kJ mol^{-1} and $-T\Delta S^{\circ}$ = 60 ± 10 kJ mol^{-1}). Therefore, the removal of the C-terminal segment in DC-SIGN has shifted its MLGI from being enthalpically driven to entropically driven. Here, the highly favourable binding ΔH° observed in DC-SIGN (-97 kJ mol^{-1}) is diminished completely in DC-SIGN-C (~2 kJ mol^{-1}). However, the binding is compensated with a strongly favorable entropic term ($-T\Delta S$ = -52 ± 3 kJ mol^{-1}), giving rise to almost the same overall binding ΔG° (*e.g.* -49 ± 1 *vs.* -50 ± 5 kJ mol^{-1} for DC-SIGN *vs.* DC-SIGN-C; **Table 3.2; Figure 3.9**).

"Cryo-snapshot" S/TEM imaging was performed to capture the native dispersion states of the QD-DiMan·lectin complexes in solution to probe their binding modes. This was accomplished by rapid plunge freezing of the samples followed by vacuum drying, and finally loading for S/TEM imaging.[20, 56] The corresponding S/TEM images show that, binding of DC-SIGN gives almost exclusively isolated single QD particles (~99%), whereas binding of DC-SIGNR results in most of the QDs (75%) being clustered, with ~20% being in a group of >4 particles (**Figure 3.10**). This result is fully consistent with our previous observations,[20] which also reaffirms the different binding modes between DC-SIGN (simultaneous tetravalent binding to individual QDs) and DC-SIGNR (crosslinking with multiple QDs). Interestingly, binding of DC-SIGN-C gives a particle dispersion that is in between those of DC-SIGN and DC-SIGNR: ~40% of the QDs are isolated, ~55% of particles are in groupings of 2 or 3, and only 5% are in groups of >4 particles (**Figure 3.10D**). This result

shows that the C-terminal segment in DC-SIGN has indeed made a valid, but not the sole, contribution in maintaining its characteristic tetrameric structure and multivalent glycan binding properties. Removal of the C-terminal segment results in DC-SIGN-C losing some of the binding character of DC-SIGN, but gaining some of DC-SIGNR. This is also consistent with previous studies showing that only DC-SIGN$^+$, but not DC-SIGNR$^+$, cells can bind to the HIV-1 virus for efficient viral transmission, and removal of the C-terminal segment in DC-SIGN reduces its virus binding and transmission ability but without completely demolishing it.[40]

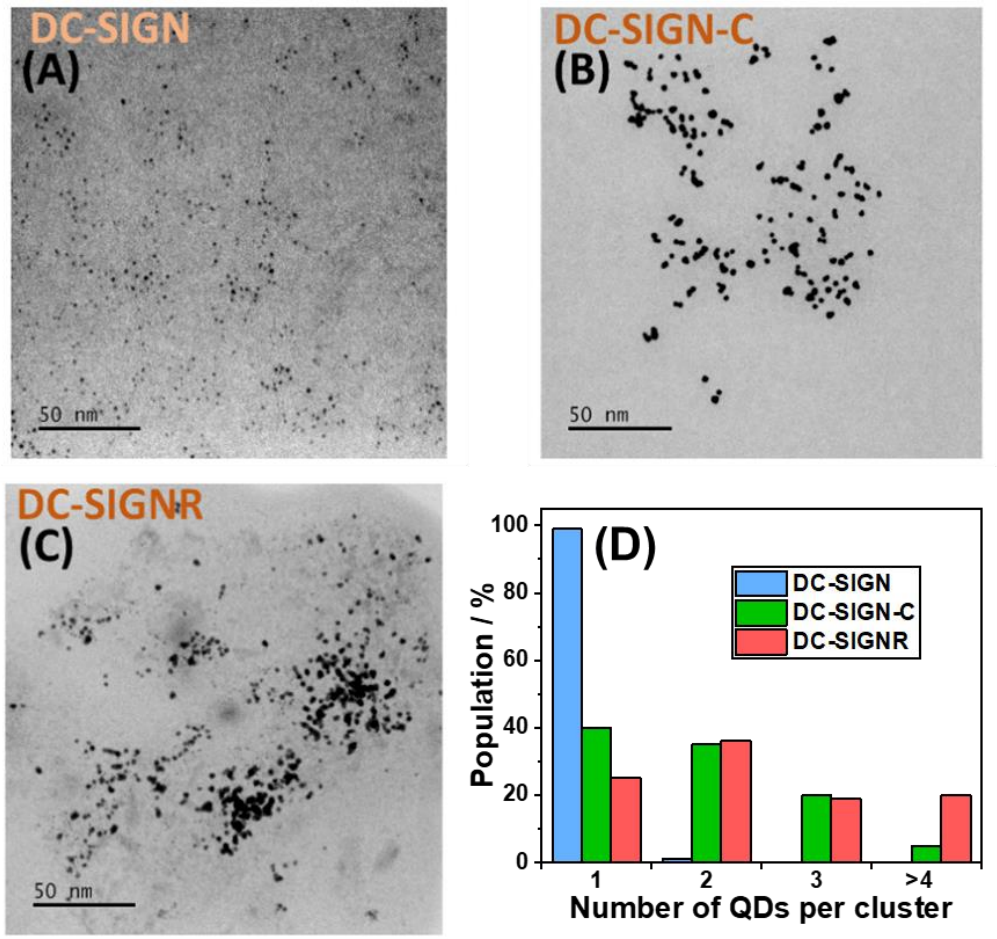

Figure 3.10 Representative cryo-preserved TEM (contrast inverted HAADF STEM) images of QD-DiMan after binding to (A) DC-SIGN (PQR, = 1:1), (B) DC-SIGN-C (PQR = 1:1), or (C) DC-SIGNR (PQR = 10:1). (D) Quantitative analysis of the number of QDs per cluster for assemblies of QD-DiMan bound with DC-SIGN, DC-SIGN-C, or DC-SIGNR.

As the C-terminal segment is located at the flexible CRD-neck junction, it may act as a steric barrier to control the CRD flexibility, forcing each CRD to function as an independent unit to retain MLGI specificity. If this is true, then

the CRDs in DC-SIGN-C would be less restricted and able to change position/orientation relative to one another more freely than that in DC-SIGN. Though this still allows DC-SIGN-C to form stable simultaneous binding to QD-DiMan to provide strong affinity, it would also enhance the probability of the CRDs from one DC-SIGN-C molecule to bind to DiMan ligands from different QDs, leading to lectin-QD clustering, which is unlikely to occur in DC-SIGN. This would likely also lead to CRD-CRD and/or CRD-neck interactions upon QD-DiMan binding, which may account for the observed ΔH^{o} penalty. This steric influence of the C-terminal segment can also rationalise the enhancement of ΔS^{o} in QD-DiMan·DC-SIGN-C binding, whereby the newly found flexibility of the CRDs would allow for the preservation of the flexibility of both the CRDs and flexible EG_{11} chains of the QD-DiMan scaffold upon binding. Moreover, their binding interactions may even relieve some of the steric strains on the CRDs, leading to the positive entropy term. This rationale would also agree with the entropic penalty observed in DC-SIGN, where a more rigid CRD arrangement would cause a loss of the degrees of freedom in the EG_{11} chains, upon binding. These results demonstrate that the combination of mutagenesis, TEM imaging and QD-FRET based biophysical characterisation is a powerful tool to probe structure-function relationships in MLGIs. Here, we reveal that the C-terminal segment may act as a steric wedge between individual CRDs in DC-SIGN to impose good specificity for multivalent binding towards multivalent glycans.

3.3.4 Investigating MLGI kinetics by QD-FRET.

The QD-DiMan·lectin binding kinetics were measured by stopped flow fluorescence spectroscopy. The association rate was obtained by rapid mixing QD-DiMan and labelled lectin into an 80 µL cuvette at a 1:1 molar ratio *via* stopped flow apparatus. Measurements of the QD and dye fluorescence were obtained over time at the wavelengths corresponding to their excitation peak maxima (**Figure 3.11A**) and were corrected by the dye direct excitation signal (**Figure 3.11C**) to provide I_{D} and I_{A} time profiles, respectively (**Figure C.2**). The FRET ratio was obtained as I_A/I_{D} and the averaged FRET ratio-time profiles were fitted by the second order rate equation to derive the apparent on-rate coefficient, k_{on} (**Equation 3.5**; derivation shown in **Appendix E.1.1**), where, $[P]_0$ is the initial protein concentration and a is a constant which accounts for fluorescence degradation.[31, 36]

$$F = F_{\max} \cdot \frac{k_{\mathrm{on}}[\mathrm{P}]_0 t}{1 + k_{\mathrm{on}}[\mathrm{P}]_0 t} - at \qquad 3.5$$

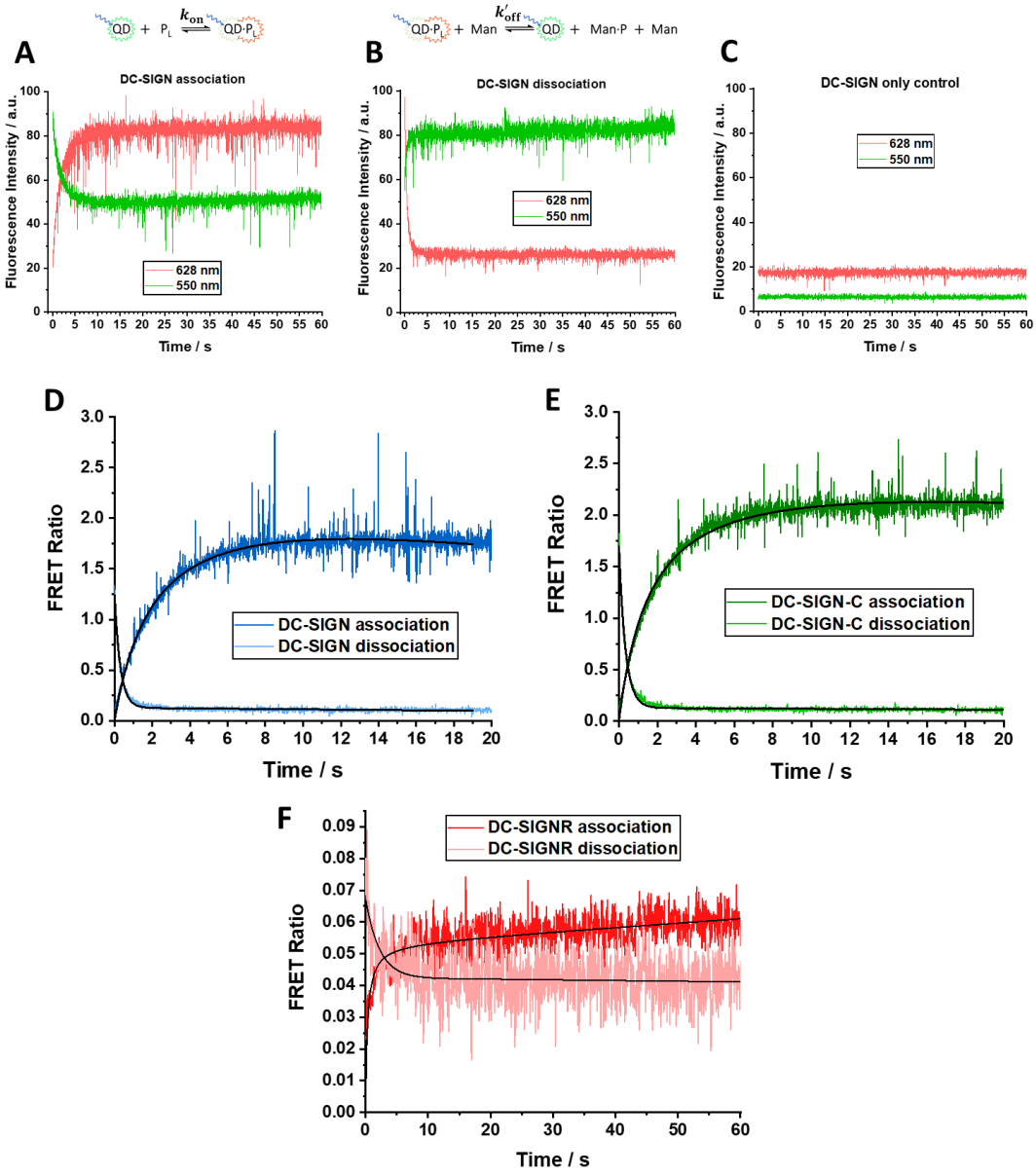

Figure 3.11 Raw kinetic profile of the fluorescence intensity at 626 nm (red) and 550 nm (green) for the association of (A) QD-DiMan with labelled DC-SIGN, (B) the dissociation of QD-DiMan and labelled DC-SIGN in the presence of excess mannose (Man), and (C) a control containing only labelled DC-SIGN. Kinetic profiles of the FRET ratio ($F = I_A/I_D$) measured for the association of a 1:1 ratio of QD-DiMan with labelled protein (dark colour) and dissociation of bound 1:1 QD-DiMan·lectin complex in excess Man (light colour) for (D) DC-SIGN, (E) DC-SIGN-C, and (F) DC-SIGNR.

Table 3.3 Summary Kinetic Parameters for QD-DiMan Binding with DC-SIGN, DC-SIGN-C, and DC-SIGNR (SDs Represent Fitting Errors).

	Association rate			
Lectin	y_{max}	k_{on} **/** $\times 10^7$ **M^{-1} s^{-1}**	$t_{½}$ **/ s**	a **/** $\times 10^{-3}$ **s^{-1}**
DC-SIGN	2.49 ± 0.03	2.24 ± 0.06	1.55 ± 0.04	10.3 ± 0.5
DC-SIGN-C	2.59 ± 0.01	2.92 ± 0.04	1.19 ± 0.02	5.1 ± 0.2
DC-SIGNR	0.0532 ± 0.0004	16 ± 1	0.22 ± 0.01	-2.5 ± 0.2

	Dissociation rate				
Lectin	y_0	y_{eq}	k_{off} **/ s^{-1}**	$t_{½}$ **/ s**	a **/** $\times 10^{-3}$ **s^{-1}**
DC-SIGN	1.261 ± 0.007	0.129 ± 0.001	3.23 ± 0.03	0.215 ± 0.002	1.261 ± 0.007
DC-SIGN-C	1.722 ± 0.006	0.131 ± 0.001	2.95 ± 0.02	0.235 ± 0.002	1.722 ± 0.006
DC-SIGNR	0.067 ± 0.003	0.042 ± 0.001	0.45 ± 0.08	1.5 ± 0.3	0.067 ± 0.003

Both DC-SIGN and DC-SIGN-C showed very similar association FRET ratio-time profiles for $[P]_0$ = 20 nM, which gave maximal FRET ratios similar to those obtained in **Figure 3.7D,F** within 10 s, indicating that saturate binding was achieved (**Figure 3.11D,E**). The second order rate equation fitted nicely for DC-SIGN and DC-SIGN-C, yielding k_{on} values of (2.24 ± 0.06) $\times 10^7$ and (2.92 ± 0.04) $\times 10^7$ M^{-1} s^{-1} and half-lives ($t_{½}$) of 1.55 ± 0.04 and 1.19 ± 0.02 s, respectively (where, $t_{½} = \ln(2)/([P]_0 k_{on})$; **Table 3.3**). DC-SIGN-C association is slightly faster, consistent with its slightly stronger affinity with QD-DiMan found in the thermodynamic studies (**Table 3.1**). A much lower FRET ratio was observed for DC-SIGNR due to its low binding affinity at a 1:1 PQR, which resulted in relatively poor fits due to the low signal-to-noise ratio. This was only slightly improved by taking an average of every 5 measurements, resulting in a time resolution of 0.0625 s. Results showed that

despite a relatively rapid initial association (increase in FRET ratio), DC-SIGNR was not able to reach saturation, even after 60 s, resulting in a negative a-term. Here, the positive a-terms observed for DC-SIGN and DC-SIGN-C (which signify a decrease of FRET over longer timescales) are likely due to the gradual decay in QD fluorescence in binding buffer over time. This has been reported previously for other small molecule ligand-capped QDs.[31, 36] However, the negative a-term for DC-SIGNR thus must be the result of another form of association occurring over a much longer timespan, giving rise to an increasing FRET ratio with time.

These kinetic results agree well with that expected for DC-SIGN/R because of their different binding modes. The simultaneous binding of DC-SIGN provides a rapid interaction where, once initial contact between a CRD and QD-DiMan is formed, it becomes kinetically more favourable for the other CRDs in the same lectin to bind due to their close proximity with the ligand. For DC-SIGNR, it is likely that the initial rapid increase in binding is a result of the simultaneous binding of two CRDs with one QD-DiMan to form a QD-DC-SIGNR intermediate unit. The secondary increase in binding, occurring over a much longer timescale, can be attributed to crosslinking interactions. As crosslinking requires multiple QD-DiMan·DC-SIGNR intermediate units to interact with each other to form large assemblies, it would be a much slower process. For DC SIGN-C binding, only minimal amounts of QDs are capable of being extensively crosslinked according to S/TEM imaging (~5%, see **Figure 3.10**). Thus, these may only induce minimal contributions to the overall FRET signal and binding kinetics which would be too small to be resolved by these measurements. Moreover, its similar association rate and maximal FRET ratio to those of DC-SIGN suggest that the small assemblies captured by S/TEM are very dynamic which implies that the CRDs in DC-SIGN-C are more flexible than those of DC-SIGN/R, where crosslinking interactions could be easily reversed in preference of more stable simultaneous binding interactions.

Pseudo-dissociation rates were obtained by injecting a pre-mixed 1:1 solution of QD-DiMan and labelled lectins into a binding buffer containing an excess of free D-mannose. A 10^6 QD molar equivalent of D-mannose was found to effectively compete with lectin-QD-DiMan binding (**Figure B.3**). The presence of free mannose greatly reduces the amount of lectins bound to the QD, leading to a decrease of dye FRET signal together with the simultaneous recovery of the QD fluorescence, and hence a decrease of FRET ratio (**Figure 3.11B**). A rapid decay in FRET ratio is observed by all three lectins, confirming that QD-DiMan·lectin binding is specifically facilitated by carbohydrate-lectin

interactions (**Figure 3.11D-F**). These FRET decay curves were fitted by a pseudo-first order rate equation, **Equation 3.6**, (see **Figure B.3**) as the change in mannose concentration is negligible. Where, k'_{off} is the apparent pseudo-first order dissociation rate coefficient.

$$F = F_{max} \cdot x_0 e^{-k'_{off}t} - at \qquad 3.6$$

As for association, the dissociation rates for DC-SIGN and DC-SIGN-C are similar, with k'_{off} values of 3.23 ± 0.03 and 2.95 ± 0.02 s^{-1} and $t_{½}$ of 0.213 ± 0.002 and 0.235 ± 0.002 s, respectively (where, $t_{½} = \ln(2)/k'_{\mathrm{off}}$). DC-SIGNR appeared to have the slowest rate of dissociation, with k'_{off} and $t_{½}$ values of 0.45 ± 0.08 s^{-1} and 1.5 ± 0.3 s, respectively. This is likely due to the difficulty in dissociating the multiple inter- and intra- QD-DiMan·DC-SIGNR interactions within the extensively crosslinked QD-lectin assemblies.

It is worth noting that the k'_{off} measured in this way is not wholly representative of the true natural dissociation rate, where dissociation and association are in equilibrium and a pair of dissociated binding partners still have chances to re-bind. Here, any dissociated protein binding sites will be rapidly occupied by the competitors, making them unable to re-bind as that which would happen under natural conditions. As a result, the k'_{off} measured in this way should be faster than natural dissociation. This is apparent by using k'_{off} to calculate the apparent binding K'_{d} $(via\ K'_{\mathrm{d}} = k'_{\mathrm{off}}/k_{\mathrm{on}})$ which would yield values of ~140 and ~100 nM for DC-SIGN and DC-SIGN-C, respectively. These values are about two orders of magnitude higher than those measured from the thermodynamic FRET assays above. Therefore, the dissociation rate coefficient derived from competition based kinetic studies must be treated with caution; it may not reflect the true disassociation rate under natural conditions. Such discrepancies are not unexpected and can be quite significant, particularly for multivalent binding systems, where re-association often occurs under natural conditions due to the close proximity of multiple binding pairs within the binding area.

However, by using the K_{d} and k_{on} rate coefficients measured by our QD-FRET thermodynamic (at 20 °C) and kinetic assays, respectively, a more plausible k_{off} of ~0.05 s^{-1} ($k_{\mathrm{off}} = k_{\mathrm{on}}K_{\mathrm{d}}$) is obtained for both DC-SIGN and DC-SIGN-C. As the K_{d} value was measured under equilibrium conditions, this calculated k_{off} should be an accurate reflection of the natural dissociation rate. In fact, this calculated k_{off} value broadly agrees with that measured by SPR (*e.g.* ~0.1 s^{-1}) between surface-immobilised DC-SIGN and DiMan-coated gold nanoparticles (GNPs, ~1.2 nm in diameter), without competitors.[57] Despite

some differences in binding environment (surface immobilised *vs* solution) and core nanoparticle sizes (~4 *vs* ~1.2 nm for QD *vs* GNP), the good agreement between the calculated k_{off} from our QD-FRET assays and that measured by SPR for the same pair of lectin and glycan-nanoparticles demonstrates that our QD-FRET assays are highly credible for probing a variety of important binding thermodynamic and kinetic parameters for MLGIs.

It is worth noting that the k_{on} measured by our QD-FRET assay is almost 1000 fold faster than that measured from SPR using surface-immobilised DC-SIGN (*e.g.* ~10^7 *vs.* ~10^4 $M^{-1}s^{-1}$).[57] We attribute this difference to the different binding environments. As our QD-FRET assays are performed in solution, both binding partners can diffuse freely, greatly increasing the likelihood of collision and thus association. Whereas, in SPR, as one binding partner (*e.g.* DC-SIGN) is immobilised on surface and unable to diffuse, it must rely on the diffusion of the other partner to the surface target sites for any binding to occur. This would result in a significantly slower on-rate than that in solution, which is exactly what has been observed here. This result also implies that the binding kinetics measured by surface assays (*e.g.* SPR, QCM) should not be used to directly predict or explain binding behaviours in solution and *vice versa*, due to the influence of binding environments on kinetics. Instead, all binding assays should be performed under the same conditions as those concerned, or at least as close as possible, in order to obtain meaningful results and explanations. In this regard, the results presented herein establish QD-FRET assays as a powerful new tool for studying solution-phase kinetics and thermodynamics of MLGIs. This method may also be applicable to other types of binding and bio-recognition processes in solution, where other well-established methods such as SPR and QCM are well-suited, but also limited, to studying binding interactions on surfaces with one immobilised binding partner.

3.4 Conclusions

In summary, we have significantly expanded the capability of our glycan-QD method for probing MLGIs. Besides providing quantitative binding affinity and binding mode data,[19, 20] we have developed a sensitive QD-FRET technique for the successful dissection of the thermodynamic and kinetic contributions behind affinity enhancing mechanisms in MLGIs with distinct binding modes, and for identification of lectin structure-function relationships. We have revealed that the lower QD-DiMan binding affinity for the crosslinking DC-SIGNR, over that of the simultaneous tetrameric binding of DC-SIGN, is a

consequence of a larger binding entropy penalty. We have further demonstrated that the removal of a 16 amino acid C-terminal segment in DC-SIGN, absent in DC-SIGNR, greatly affects its QD-DiMan binding thermodynamic profile and completely changes the binding from an enthalpically driven into an entropically driven MLGI. These results have allowed us to hypothesise that the entropic gain in removing the C-terminal segment is the result of an increased freedom of the CRDs, which is not present in DC-SIGN naturally. S/TEM images of the resulting QD-DiMan·DC-SIGN-C complex further supports the idea that the C-terminal segment may play a key role in maintaining the CRD orientation and therefore in controlling the multivalent specificity of DC-SIGN in binding to multivalent glycans. Together, this work has established glycan-QDs as a powerful new platform for probing the biophysical and structural mechanisms of MLGIs in solution. These data are important for guiding the design of multivalent therapeutics against specific MLGIs, particularly those with unknown structures.

References

1. M. Mammen, S. K. Choi and G. M. Whitesides, *Angew. Chem. Int. Ed. Engl.*, 1998, **37**, 2754-2794.
2. G. D. Brown, J. A. Willment and L. Whitehead, *Nat. Rev. Immunol.*, 2018, **18**, 374-389.
3. K. Drickamer and M. E. Taylor, *Curr. Opin. Struct. Biol.*, 2015, **34**, 26-34.
4. W. Wu, C. Liu, C. A. Farrar, L. Ma, X. Dong, S. H. Sacks, K. Li and W. Zhou, *J. Am. Soc. Nephrol.*, 2018, **29**, 168-181.
5. M. Howard, C. A. Farrar and S. H. Sacks, *Semin. Immunopathol.*, 2018, **40**, 75-85.
6. Y. van Kooyk and T. B. Geijtenbeek, *Nat. Rev. Immunol.*, 2003, **3**, 697-709.
7. C. Reily, T. J. Stewart, M. B. Renfrow and J. Novak, *Nat. Rev. Nephrol.*, 2019, **15**, 346-366.
8. A. Bernardi, J. Jimenez-Barbero, A. Casnati, C. De Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K. E. Jaeger, M. Lahmann, T. K. Lindhorst, M. Marradi, P. Messner, A. Molinaro, P. V. Murphy, C. Nativi, S. Oscarson, S. Penades, F. Peri, R. J. Pieters, O. Renaudet, J. L. Reymond, B. Richichi, J. Rojo, F. Sansone, C. Schaffer, W. B. Turnbull, T. Velasco-Torrijos, S. Vidal, S. Vincent, T. Wennekes, H. Zuilhof and A. Imberty, *Chem. Soc. Rev.*, 2013, **42**, 4709-4727.
9. S. Bhatia, L. C. Camacho and R. Haag, *J. Am. Chem. Soc.*, 2016, **138**, 8654-8666.
10. L. L. Kiessling and R. A. Splain, *Annu Rev Biochem*, 2010, **79**, 619-653.
11. E. Pallister, C. J. Gray and S. L. Flitsch, *Curr Opin Struct Biol*, 2020, **65**, 184-192.
12. A. N. Baker, S. J. Richards, C. S. Guy, T. R. Congdon, M. Hasan, A. J. Zwetsloot, A. Gallo, J. R. Lewandowski, P. J. Stansfeld, A. Straube, M. Walker, S. Chessa, G. Pergolizzi, S. Dedola, R. A. Field and M. I. Gibson, *ACS Cent Sci*, 2020, **6**, 2046-2052.
13. J. Ramos-Soriano, M. Ghirardello and M. C. Galan, *Curr Med Chem*, 2022, **29**, 1232-1257.
14. R. Ribeiro-Viana, M. Sanchez-Navarro, J. Luczkowiak, J. R. Koeppe, R. Delgado, J. Rojo and B. G. Davis, *Nat. Commun.*, 2012, **3**, 1303.

15. S. Penadés, B. G. Davis and P. H. Seeberger, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar and P. H. Seeberger, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 2015, pp. 743-753.
16. D. Benito-Alifonso, S. Tremel, B. Hou, H. Lockyear, J. Mantell, D. J. Fermin, P. Verkade, M. Berry and M. C. Galan, *Angew Chem Int Ed Engl*, 2014, **53**, 810-814.
17. P. I. Kitov, J. M. Sadowska, G. Mulvey, G. D. Armstrong, H. Ling, N. S. Pannu, R. J. Read and D. R. Bundle, *Nature*, 2000, **403**, 669-672.
18. T. R. Branson, T. E. McAllister, J. Garcia-Hartjes, M. A. Fascione, J. F. Ross, S. L. Warriner, T. Wennekes, H. Zuilhof and W. B. Turnbull, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 8323-8327.
19. Y. Guo, C. Sakonsinsiri, I. Nehlmeier, M. A. Fascione, H. Zhang, W. Wang, S. Pohlmann, W. B. Turnbull and D. Zhou, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 4738-4742.
20. Y. Guo, I. Nehlmeier, E. Poole, C. Sakonsinsiri, N. Hondow, A. Brown, Q. Li, S. Li, J. Whitworth, Z. Li, A. Yu, R. Brydson, W. B. Turnbull, S. Pohlmann and D. Zhou, *J. Am. Chem. Soc.*, 2017, **139**, 11833-11844.
21. D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo and D. Zhou, *J. Am. Chem. Soc.*, 2020, **142**, 18022-18034.
22. R. P. Galimidi, J. S. Klein, M. S. Politzer, S. Bai, M. S. Seaman, M. C. Nussenzweig, A. P. West, Jr. and P. J. Bjorkman, *Cell*, 2015, **160**, 433-446.
23. T. K. Dam and C. F. Brewer, *Chem Rev*, 2002, **102**, 387-429.
24. W. B. Turnbull and A. H. Daranas, *J. Am. Chem. Soc.*, 2003, **125**, 14859–14866.
25. P. Zihlmann, M. Silbermann, T. Sharpe, X. Jiang, T. Mühlethaler, R. P. Jakob, S. Rabbani, C. P. Sager, P. Frei, L. Pang, T. Maier and B. Ernst, *Chemistry*, 2018, **24**, 13049-13057.
26. A. Velazquez-Campoy and E. Freire, *Nat Protoc*, 2006, **1**, 186-191.
27. J. Rao, J. Lahiri, L. Isaacs, R. M. Weis and G. M. Whitesides, *Science*, 1998, **280**, 708-711.
28. J. D. Chodera and D. L. Mobley, *Annu. Rev. Biophys.*, 2013, **42**, 121-142.
29. V. Porkolab, C. Pifferi, I. Sutkeviciute, S. Ordanini, M. Taouai, M. Thepaut, C. Vives, M. Benazza, A. Bernardi, O. Renaudet and F. Fieschi, *Org. Biomol. Chem.*, 2020, **18**, 4763-4772.
30. K. D. Wegner, Z. Jin, S. Lindén, T. L. Jennings and N. Hildebrandt, *ACS Nano*, 2013, **7**, 7411-7419.
31. W. Wang, Y. Guo, C. Tiede, S. Chen, M. Kopytynski, Y. Kong, A. Kulak, D. Tomlinson, R. Chen, M. McPherson and D. Zhou, *ACS Appl. Mater. Interfaces*, 2017, **9**, 15232-15244.
32. I. L. Medintz, H. T. Uyeda, E. R. Goldman and H. Mattoussi, *Nat. Mater.*, 2005, **4**, 435-446.
33. K. D. Wegner and N. Hildebrandt, *Chem. Soc. Rev.*, 2015, **44**, 4792-4834.
34. M. Cardoso Dos Santos, W. R. Algar, I. L. Medintz and N. Hildebrandt, *Trends Anal. Chem.*, 2020, **125**, 115819.
35. N. Hildebrandt, C. M. Spillmann, W. R. Algar, T. Pons, M. H. Stewart, E. Oh, K. Susumu, S. A. Diaz, J. B. Delehanty and I. L. Medintz, *Chem. Rev.*, 2017, **117**, 536-711.
36. D. Zhou, L. Ying, X. Hong, E. A. Hall, C. Abell and D. Klenerman, *Langmuir*, 2008, **24**, 1659-1664.
37. J. Hooper, Y. Guo and D. Zhou, in *Glycan-Based Cellular Communication: Techniques for Carbohydrate-Protein Interactions*, American Chemical Society, 2020, vol. 1346, ch. 4, pp. 47-66.
38. H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327-1335.
39. H. Feinberg, D. A. Mitchell, K. Drickamer and W. I. Weis, *Science*, 2001, **294**, 2163-2166.
40. N. P. Chung, S. K. Breun, A. Bashirova, J. G. Baumann, T. D. Martin, J. M. Karamchandani, J. W. Rausch, S. F. Le Grice, L. Wu, M. Carrington and V. N. Kewalramani, *J. Biol. Chem.*, 2010, **285**, 2100-2112.

41. C. W. Davis, H. Y. Nguyen, S. L. Hanna, M. D. Sanchez, R. W. Doms and T. C. Pierson, *J. Virol.*, 2006, **80**, 1290-1301.
42. T. B. Geijtenbeek, D. S. Kwon, R. Torensma, S. J. van Vliet, G. C. van Duijnhoven, J. Middel, I. L. Cornelissen, H. S. Nottet, V. N. KewalRamani, D. R. Littman, C. G. Figdor and Y. van Kooyk, *Cell*, 2000, **100**, 587-597.
43. S. Pöhlmann, E. J. Soilleux, F. Baribaud, G. J. Leslie, L. S. Morris, J. Trowsdale, B. Lee, N. Coleman and R. W. Doms, *Proc. Natl. Acad. Sci U. S. A.*, 2001, **98**, 2670-2675.
44. R. Amraei, W. Yin, M. A. Napoleon, E. L. Suder, J. Berrigan, Q. Zhao, J. Olejnik, K. B. Chandler, C. Xia, J. Feldman, B. M. Hauser, T. M. Caradonna, A. G. Schmidt, S. Gummuluru, E. Mühlberger, V. Chitalia, C. E. Costello and N. Rahimi, *ACS Cent. Sci.*, 2021, **7**, 1156-1165.
45. H. Zhang, G. Feng, Y. Guo and D. Zhou, *Nanoscale*, 2013, **5**, 10307-10315.
46. D. Zhou, A. Bruckbauer, L. Ying, C. Abell and D. Klenerman, *Nano Lett.*, 2003, **3**, 1517-1520.
47. D. Zhou, A. Bruckbauer, C. Abell, D. Klenerman and D.-J. Kang, *Adv. Mater.*, 2005, **17**, 1243-1248.
48. D. Zhou, *Biochem. Soc. Trans.*, 2012, **40**, 635-639.
49. D. A. Mitchell, A. J. Fadden and K. Drickamer, *J. Biol. Chem.*, 2001, **276**, 28939-28945.
50. I. L. Medintz, A. R. Clapp, H. Mattoussi, E. R. Goldman, B. Fisher and J. M. Mauro, *Nat. Mater.*, 2003, **2**, 630-638.
51. A. Holla and A. Skerra, *Protein Eng. Des. Sel.*, 2011, **24**, 659-669.
52. M. E. Taylor and K. Drickamer, *FEBS J.*, 2019, **286**, 1800-1814.
53. Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis and K. Drickamer, *Nat. Struct. Mol. Biol.*, 2004, **11**, 591-598.
54. H. Feinberg, R. Castelli, K. Drickamer, P. H. Seeberger and W. I. Weis, *J. Biol. Chem.*, 2007, **282**, 4202-4209.
55. S. Menon, K. Rosenberg, S. A. Graham, E. M. Ward, M. E. Taylor, K. Drickamer and D. E. Leckband, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 11524-11529.
56. N. Hondow, R. Brydson, P. Wang, M. D. Holton, M. R. Brown, P. Rees, H. D. Summers and A. Brown, *J. Nanoparticle Res.*, 2012, **14**.
57. O. Martínez-Ávila, K. Hijazi, M. Marradi, C. Clavel, C. Campion, C. Kelly and S. Penadés, *Chem. Eur. J.*, 2009, **15**, 9874-9888.

Chapter 4
Polyvalent Glycan-Quantum Rods as Multifunctional Mechanistic Probes for Shape-Selective Multivalent Lectin-Glycan Recognition

The following chapter is based on the work from: Hooper, J. et al. *ACS Appl. Nano Mater.* **2023**. 6 (6), 4201-4213.

The author of this thesis wrote the manuscript and conducted all material preparation, FRET experiments and all data analysis. STEM imaging was performed by Dario Fernandez Ainaga and Dr Nicole Hondow. Ligand precursors were synthesised by Dr Darshita Budhadev. All plasmids were prepared in-house by Dr Yuan Guo.

4.1 Introduction

Research into the effects of nanomaterial design in biology has led not only to key advances in the ever-growing field of nanomedicine, but also the discovery of novel tools to answer important biological questions.[1-4] The exploration of fine tuning nanoparticle surface chemistry to achieve biological functionality has been extensively investigated for over 4 decades, however, the study of a nanoparticle's geometric design elements, such as size and shape, and their mechanistic influences in biology, remains in its infancy. Though some studies have shown that varying nanoparticle size and shape can have a strong effect on various biological processes such as cell uptake,[5-7] endocytic pathways,[8-10] cytokine production,[8, 11-13] or pathogen inhibition,[14-16] the relationship between nanoparticle geometry and biological function remains a fairly under-explored topic for a lot of processes. Moreover, often little has been done to elucidate the specific mechanistic differences caused by shape variation at the molecular level, such as binding affinities or binding modes, which are essential for identifying how nanomedicines work and, perhaps more importantly, for guiding the design of more effective therapeutics.

Multivalent lectin-glycan interactions (MLGIs) are widespread and play a pivotal role in pathogen infection, immune regulation and cell-cell communication.[17] It is thus unsurprising that glycan displaying nanomaterials have been widely exploited for potential antiviral and immuno-therapeutic applications, due to simple polyvalent glycan functionalization. For glyconanomaterials, the geometric parameters, such as scaffold size and

shape, inevitably affect their surface glycan display. As natural lectins are often multimeric, variations in glycan display may lead to changes in the number of binding sites that can be occupied at one time, which dictates the strength and mode of lectin-glycan binding and, subsequently, its biological function. A large number of nano-scaffolds with different geometries have been employed to target lectins. These include spherical scaffolds (e.g. inorganic nano-particles, fullerenes, dendrimers, hyper-branched polymers, polymersomes, etc),[18-28] tubular/rod shaped structures (e.g. gold nanorods, carbon nanotubes, cylindrical micelles),[7, 16, 29-32] planar sheets (e.g. graphene nanosheets),[31, 33] as well as structures that are less well-defined (e.g. linear polymers)[15, 34-38] or more complex in shape.[7, 32, 39-41] Since these results are obtained with scaffolds of different glycan composition, softness, size, shape, and glycan density, it is difficult to directly compare results from one another to draw general conclusions. While a few studies have shown that the scaffold size and shape can affect their lectin binding and pathogen inhibition properties, the molecular mechanisms underlying such differences remain unclear.[7, 8, 14-16, 30, 32, 42] Therefore, new tools need to be developed in order to identify the rationales behind these differences allowing us to better understand the molecular mechanisms of shape dependency for such interactions. In this regard, probes that can reveal how scaffold shape affects MLGI binding mode and affinity are highly valuable, allowing us to establish a geometric design rule for glyconanoparticles for potent and specific targeting of particular MLGIs for therapeutic application.

As discussed in previous chapters, our group have recently developed densely glycosylated fluorescent quantum dots (glycan-QDs) as new mechanistic probes for MLGIs.[20, 21] These glycan-QDs were not only able to quantify MLGI affinity *via* a ratiometric QD sensitized dye Förster resonance energy transfer (FRET) readout, but also dissect the exact binding modes and affinity enhancing mechanisms of MLGIs *via* hydrodynamic size analysis and STEM imaging of lectin binding induced glycan-QD assemblies.[20, 21] Using QDs bearing mannose-α-1,2-mannose (DiMan)-glycans (QD-DiMan) as probes, our group was able to extract key structural and mechanistic information for MLGIs of the immunologically important and closely related tetrameric high-mannose-binding viral receptors, DC-SIGN and DC-SIGNR (collectively denoted as DC-SIGN/R hereafter). These studies revealed that, although both lectins bound multivalently with QD-DiMan, these spherical QD-DiMan scaffolds induced different binding modes which resulted in very different affinities. DC-SIGN was found to simultaneously bind to the same QD-DiMan with all 4 carbohydrate-recognition-domains (CRDs) which gave

rise to strong binding (with an apparent equilibrium binding dissociation constant, K_d, of 2.1 nM). Whereas, DC-SIGNR was found to crosslink between QD-DiMan particles (**Figure 3.1**), with a much weaker affinity (K_d = ~633 nM, ~300-fold weaker than that of DC-SIGN).[20] An additional demonstration of this affinity difference was found by monitoring the FRET signal of lectin binding with increasing concentrations of protein (**Figure 4.1**). For DC-SIGN, binding was detected at very low protein to QD molar ratios (PQRs) and increased linearly with PQR until the QD surface was fully saturated with protein. However, for DC-SIGNR, due to low affinity and crosslinking binding nature, saturation occurred at a much higher PQR and significant binding was only observed as PQR was increased above a certain threshold.

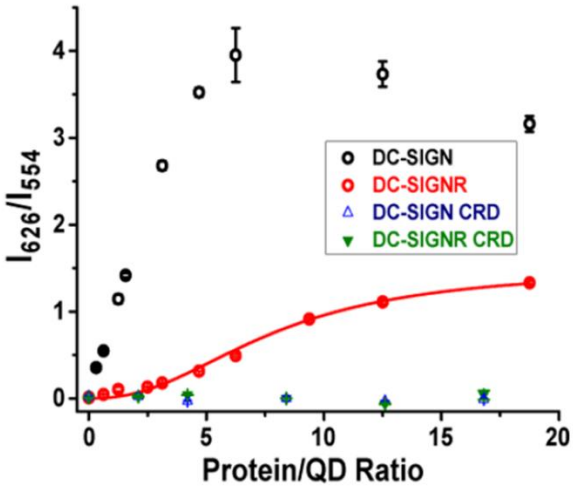

Figure 4.1 FRET signal with increasing protein to QD molar ratio (PQR) for the titration of DC-SIGN (black circle), DC-SIGNR (red circle), monomeric DC-SIGN carbohydrate-recognition domain (CRD) (blue upward triangle) or monomeric DC-SIGNR CRD (green downward triangle), against a fixed concentration of QDs capped with DHLA-EG$_{11}$-DiMan (QD-DiMan). (Taken from reference [20]).

Where QDs comprise of roughly spherical nanoscale crystals of semiconducting material, quantum rods (QRs) are of the same material composition but are elongated in one dimension. QRs have a higher extinction coefficient and single particle brightness than QDs, which is highly beneficial for fluorescence based applications.[43-45] Moreover, their optical properties can be tuned by changing not only the particle size, but also the aspect (length to width) ratio. This feature can be highly beneficial for some applications, e.g. bioluminescence resonance energy transfer (BRET), where an aspect ratio of 3:1 has shown to give the highest BRET efficiency.[46, 47] In addition to these benefits, the distinct curvatures between the central cylindrical section and

spherical ends make QRs an attractive platform for studying the influence of scaffold geometry in controlling MLGI properties. For instance, by conjugating polyvalent glycans on the QR surface (*i.e.* glycan-QR), the glycan displays at the ends will closely resemble that of spherical glycan-QDs, whereas, in the middle, the glycans will be presented more like a curved plane wrapped around the centre of the rod. Therefore, by exploiting both the highly fluorescent properties of QRs with their ability to form two distinct glycan displays, it has been possible to probe the effects of different scaffold geometries on MLGIs using the same glycan-nanoparticle.

This chapter demonstrates (to the best of our knowledge) the first use of glycan-QRs which have been prepared for the application of probing DC-SIGN/R-related MLGIs (**Figure 4.3**). By developing a QR-FRET based ratiometric MLGI affinity readout, together with hydrodynamic size and S/TEM analysis of lectin binding induced QR assemblies, it is shown that QR-DiMan with distinct glycan displaying curvatures can effectively discriminate the MLGI properties between DC-SIGN and DC-SIGNR. DC-SIGN binds strongly and simultaneously to one QR-DiMan regardless of surface curvature with sub-nM K_{d}, comparable to that with QD-DiMan. In contrast, DC-SIGNR binds simultaneously to the glycans displayed on the central cylindrical section of the same QR but cross-links between glycans displayed on the spherical ends of different QRs. This result demonstrates that the QR's unique geometry together with its strong fluorescence and high EM contrast allows us to reveal how glycan presentation, induced by scaffold curvature, affect MLGIs.

4.2 Experimental

4.2.1 Preparation of QR-DiMan.

QR-EG$_{11}$-DiMan (QR-DiMan). Cap exchange procedure was adapted from previously described methods of lipoic acid ligand conjugation.[20, 21, 48] DHLA-EG$_{11}$-DiMan was prepared as described in **Section 2.1.3**. QR$_{560}$ (416 µL, 2.0 nmol) in toluene was dispersed in EtOH (1.2 mL) and span at 15000×g for 5 min. The QR pellet was then dissolved in $CHCl_3$ (300 µL) and then a mixture of DHLA-EG$_{11}$-DiMan (273 µL, 5.4 µmol) in $CHCl_3$ and NaOH (64 µl, 6.48 µmol) in EtOH was added (ligand:QR molar ratio = 2700:1) followed by MeOH (100 µL) to make a homogenous solution. The resulting mixture was immediately covered with aluminum foil and stirred magnetically at room temperature for 30 min. The QRs were pelleted by addition of hexane (600 µL) followed by centrifugation at 15000×g for 3 min. The supernatant was removed and stored for carbohydrate quantification, and the resulting QR

pellet was dissolved in H_2O (300 µL) and washed with H_2O (3 × 200 µL) using a 30 kDa MWCO spin filter. This yielded a stable aqueous QR-DiMan (1.55 nmol, yield = 78 %) with a fluorescence quantum yield (QY) of 6.2% (**Appendix B.2**). All post cap-exchange supernatants and filtrates were combined, dried and used for determining the unbound free ligands using phenol-sulphuric acid carbohydrate quantification against a DiMan calibration standard (**Section 2.2.2.4**). Its surface glycan valency was estimated as 370 ± 30 ligands per QR.

A mean hydrodynamic diameter (D_h) of 20.8 ± 4.8 nm was obtained by dynamic light scattering (DLS; **Figure 4.2A**). Core diameters and lengths distributions were obtained from the area (a) and aspect ratios (AR) of each QR obtained by STEM imaging (where diameter, $d = \sqrt{a/\mathrm{AR}}$; length, $l = \sqrt{a \cdot \mathrm{AR}}$), and were plotted as log histograms and fitted with Gaussian fits to obtain values for the mean core diameter and length of D_{core} = 3.06 ± 0.70 nm and L_{core} = 14.7 ± 5.7 nm, respectively (**Figure 4.2B**). Interestingly, these values are ~64% smaller than the reported sizes prior to ligand conjugation which may suggest some etching of the QR surface upon cap exchange.

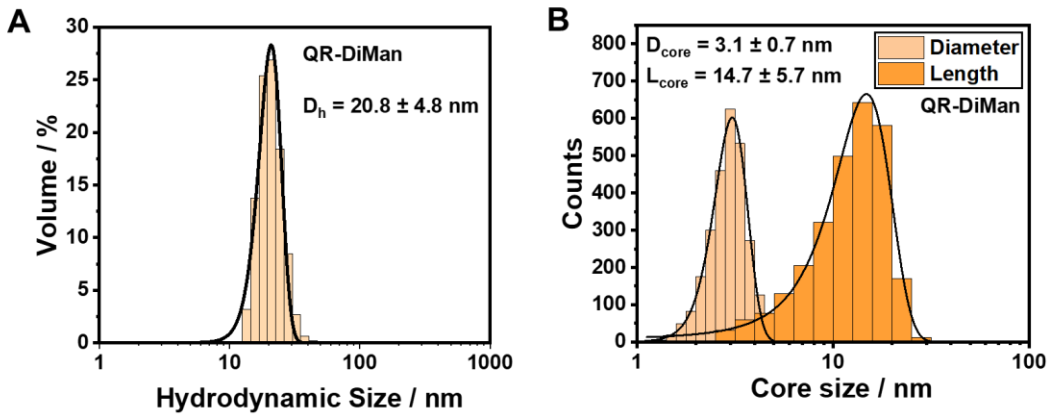

Figure 4.2 (A) Volume population hydrodynamic size distribution histogram, fitted with the Gaussian function, of QR_{560} capped with $LA\text{-}EG_{11}$-DiMan (QR-DiMan) measured by DLS. (B) Histograms of the QR core diameter and length distributions obtained from STEM images of QR-DiMan. Mean hydrodynamic diameter and core diameter and length are denoted as D_h, D_{core} and L_{core}, respectively (recorded as mean ± ½ FWHM; $R^2 \geq 0.98$).

4.2.2 QR-FRET assays

All FRET assays were performed as described in **Section 2.2.2.1**. Samples for affinity determination assays were prepared by adding labelled lectin to a

solution of QR-DiMan in binding buffer (containing BSA, 1 mg mL^{-1}), at final concentrations ranging from 0.1-50 nM for PQRs of 1:1 or 1-800 nM for PQRs of 10:1, and incubating at r.t. for 20 min. FRET spectra were obtained by measuring the direct excitation background corrected fluorescence spectra at each concentration.

Samples for protein titration assays were prepared by adding labelled lectin to QR-DiMan (5 nM, final concentration) in binding buffer (containing BSA, 1 mg mL^{-1}) with lectin concentrations ranging from 0-150 nM for DC-SIGN or 0-400 nM for DC-SIGNR, and incubating at r.t. for 20 min. FRET spectra were obtained by measuring the direct excitation background corrected fluorescence spectra at each concentration.

Samples for mannose dissociation assays were prepared by adding labelled lectin (40 nM, final concentration) to QR-DiMan (10 nM, final concentration) in binding buffer (containing BSA, 1 mg/mL) before incubating at r.t. for 20 min. D-mannose (Man) was then added at concentrations ranging from 0-100 mM, before incubating for a further 20 min and then measuring. Direct excitation background corrected FRET spectra were obtained by subtracting a background of only the labelled lectin (40 nM, final concentration) in binding buffer (containing BSA, 1 mg mL^{-1}) from the fluorescence measurement at each [Man].

4.2.3 Lectin Titration Analysis via DLS

All DLS measurements were performed as described in **Section 2.2.2.2**. Samples for protein titration assays were prepared by adding wild-type lectin to QR-DiMan (10 nM) at lectin concentrations of 0-200 nM. The samples were incubated in binding buffer at r.t. for 20 min before DLS measurement was performed.

Samples for mannose dissociation assays were prepared by adding wild-type lectin (40 nM) to QR-DiMan (10 nM) in binding buffer and incubating at r.t. for 20 min. Man was then added, with concentrations ranging from 0.01-100 mM, before incubating for a further 20 min and measuring by DLS.

Hydrodynamic size distributions were fitted with uni- or bi- modal Gaussian fits. From here, area weighted average D_{h} values were taken (average $D_{\mathrm{h}} = D_{\mathrm{h},1} \times A_1\% + D_{\mathrm{h},2} \times A_2\%$, where $A_1\%$ and $A_2\%$ are the percentage area of the Gaussian fits) and plotted against PQR. Average D_{h} plotted for the titration assays were fitted with an offset Hill equation in order to obtain the D_{h} at plateau.

4.2.4 S/TEM QR images

High angle annular dark field scanning transmission electron microscope (HAADF STEM) images of lectin·glycan-QR complexation were performed as described in **Section 2.2.2.3**. Samples were prepared by adding wild-type protein (160 nM) to QR-DiMan (40 nM) in binding buffer and incubating at r.t. for 20 min before loading onto the TEM grid, blotting and then plunge-freezing. Complex dissociation by D-mannose was prepared in the same way, however, Man (2 mM) was added after the initial incubation step, and incubated for a further 20 min, before loading. Data was collected of 310 QRs from 12 images for QR-DiMan + DC-SIGN, 759 QRs from 15 images for QR-DiMan + DC-SIGNR and 1661 QRs from 15 images for QR-DiMan·DC-SIGNR + Man, taking only particles within a threshold area of 5-500 nm^2.

4.3 Results and Discussion

4.3.1 QR-DiMan Preparation and Characterisation

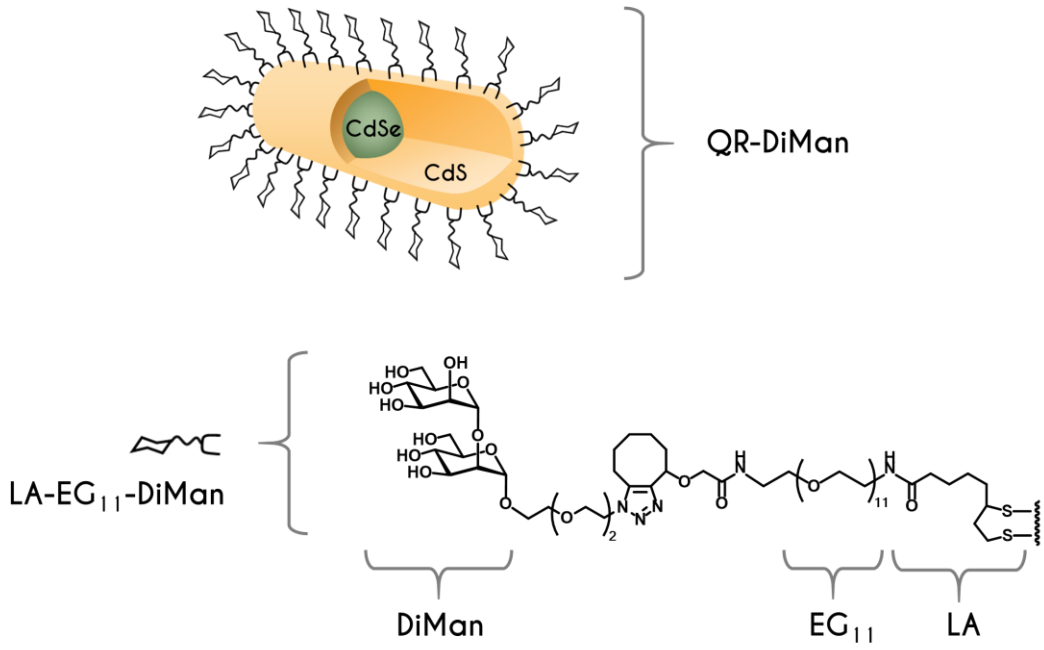

Figure 4.3 Schematic of DHLA-EG_{11}-DiMan capped CdSe/CdS core/shell quantum rods (QR-DiMan) with the chemical structure of deprotonated DHLA-EG_{11}-DiMan (LA-EG_{11}-DiMan; below).

A CdSe/CdS core/shell (dot in a rod) QR with maximal emission at λ_{em} ~560 nm (denoted as QR_{560} hereafter) was chosen to construct the

QR-DiMan *via* our previously described ligand conjugation method.[49] QR_{560} was chosen due to its similar core diameter to the QD scaffolds used in our group's previous studies,[20] allowing for direct comparison of results. A dihydrolipoic acid-undeca(ethylene glycol)-mannose-α-1,2-mannose (DHLA-EG_{11}-DiMan, **Figure 4.3**) based multi-functional ligand was synthesized as described previously.[20] Each ligand contains three unique functional domains: a DHLA group for strong QR surface anchoring;[49] a flexible EG_{11} linker for imposing high water-solubility, stability and resisting non-specific interactions,[50, 51] and a terminal DiMan group for specific lectin binding. QR-DiMan was produced by performing cap-exchange using deprotonated DHLA-EG_{11}-DiMan ligand in a homogeneous $CHCl_3$/MeOH/EtOH solution, giving rise to high cap-exchange efficiency as described previously for QD-DiMan.[20,48] QR-DiMan was found to be relatively compact, with a mean hydrodynamic diameter (D_h) of 20.8 ± 4.8 nm and the first excitonic absorption and emission peaks at 541 nm and 560 nm, respectively. Its fluorescence quantum yield (QY) was determined as 6.2%, this represents a significant reduction of its nominal QY (~68%) prior to cap-exchange (**Appendix B.2**). This result agrees with other literature results, where CdSe/CdS based QDs and QRs have shown to display significantly reduced fluorescence QY after cap-exchange.[46,49] By calculating the difference between the amount of ligand added and that remained in supernatant post cap-exchange *via* the phenol-sulphuric acid carbohydrate quantification, the number of DHLA-EG_{11}-DiMan ligands capped on each QR was estimated as 370 ± 30 (**Section 2.2.2.4**).[19-21]

4.3.2 Binding Affinity Determination *via* FRET

The QR's strong fluorescence was exploited as a ratiometric FRET readout to quantify the binding affinity between QR-DiMan and acceptor fluorophore labelled DC-SIGN/R. As in **Chapter 3**, lectin labelling was achieved by coupling a maleimide-Atto594 dye *via* the site-specifically introduced Q274C and R287C mutations in DC-SIGN and DC-SIGNR, respectively (**Section 2.1.4.2**), which lie close to, but not in, the glycan binding pockets, enabling us to obtain sufficient FRET signals without inhibiting the lectins' glycan binding properties, as confirmed previously.[20, 21, 52] The QR-Atto594 FRET pair has good spectral overlap and a respectable Förster radius R_0 of 4.8 nm (see **Appendix B.2**), ensuring that efficient FRET between QR-DiMan and labelled lectins can happen upon binding. Furthermore, there is little overlap of their

emission spectra, allowing for straightforward separation of the donor and acceptor fluorescence without the need of spectral deconvolution.[52]

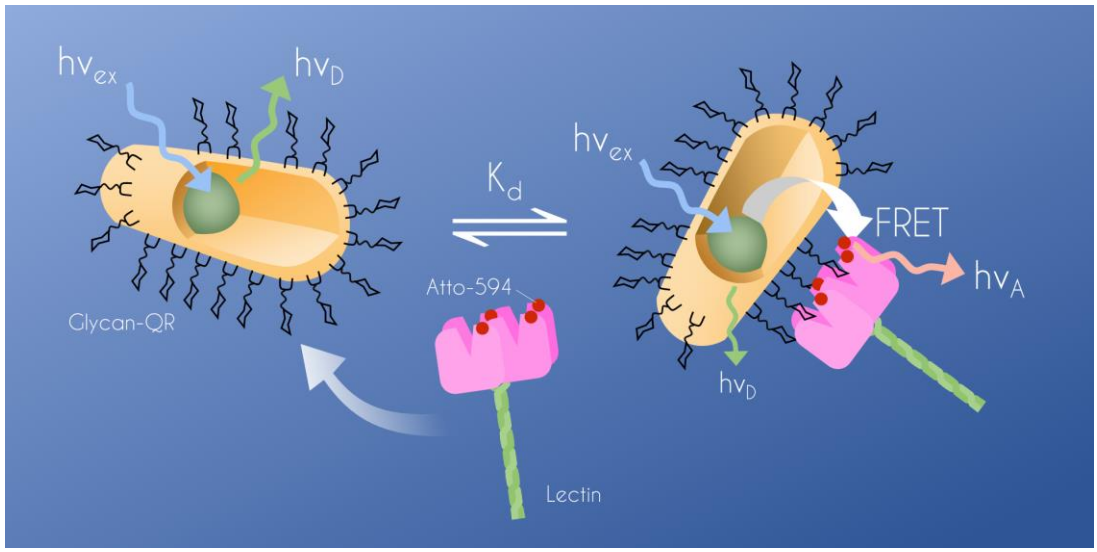

Figure 4.4 Schematic depicting the FRET process between an excited glycan-QR (donor) and lectin labelled with Atto594 (acceptor), where $h\nu_{ex}$ is the excitation energy, $h\nu_D$ is the donor emission energy and $h\nu_A$ is the acceptor emission energy.

The affinity assays were carried out by mixing QR-DiMan with labelled proteins under different concentrations at a fixed protein:QR molar ratio (PQR) of 1:1 in a binding buffer (20 mM HEPES, 100 mM NaCl, 10 mM $CaCl_2$, pH 7.8, containing 1 mg/mL of bovine serum albumin to reduce non-specific adsorption). The resulting fluorescence spectra were recorded using a fixed excitation wavelength (λ_{ex}) of 450 nm, corresponding to the absorption minimum of the Atto594 acceptor to reduce the dye direct excitation background. Exciting an equilibrated mixture of QR-DiMan with labelled lectins resulted in fluorescence of unbound QR (peaking at ~559 nm) or, if binding occurred, energy transfer *via* FRET from the excited QR donor to the Atto594 acceptor, giving rise to simultaneously quenched QR and enhanced Atto594 fluorescence (at ~627 nm; **Figure 4.4**). These fluorescence spectra were corrected by subtracting dye direct excitation background spectra of the labelled lectins, without QR-DiMan, under identical conditions (**Figure 4.5A-C**). The fluorescence spectra were measured over varying concentrations and the resulting dye to QR fluorescence intensity ratios (FRET ratios) were then fitted by the Hill equation to extract the apparent binding K_d values (**Equation 4.1**; where, F is the FRET ratio, I is the peak emission intensity of the donor (D; QR) or acceptor (A; Atto594), [P] is the protein concentration, n is the Hill coefficient and K_d is the apparent

equilibrium binding dissociation constant).[20, 52] Here, n = 1 was assumed because most binding assays were measured under a PQR of 1:1, under which most QRs should be bound by just a single lectin, and hence there should be no positive or negative influence of QR-bound lectins toward further binding of lectins to the same QR-DiMan. As the FRET ratio is proportional to the fraction of protein bound to the QR under a fixed PQR of 1:1,[20] this method is robust and can provide accurate affinity measurement of both strong and weak MLGIs.[52]

$$F = \frac{I_{\mathrm{A}}}{I_{\mathrm{D}}} = F_{\max} \cdot \frac{[\mathrm{P}]_{\mathrm{bound}}}{[\mathrm{P}]_{\mathrm{total}}} = F_{max} \cdot \frac{[\mathrm{P}]^n}{{K_{\mathrm{d}}}^n + [\mathrm{P}]^n} \qquad 4.1$$

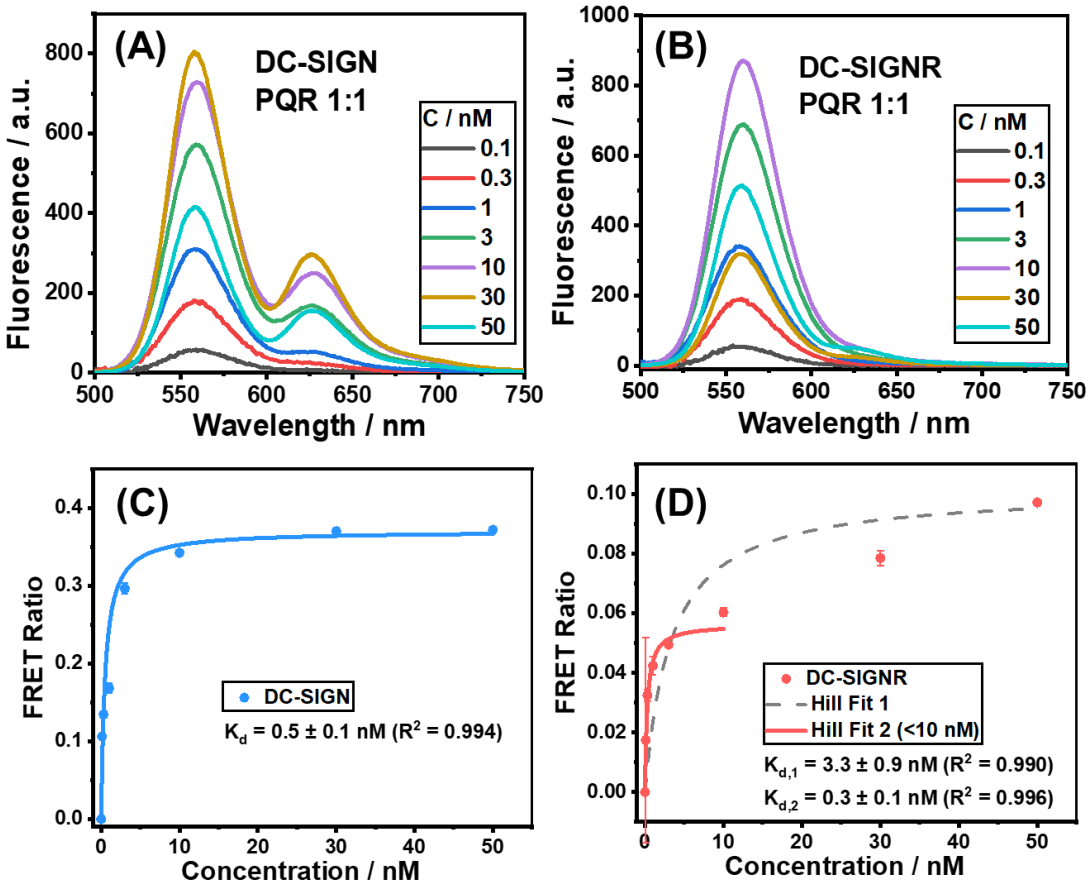

Figure 4.5 Direct excitation background corrected fluorescence spectra over different concentrations of QR-DiMan binding with Atto594 labelled (A) DC-SIGN and (B) DC-SIGNR. The corresponding FRET ratio–concentration relationships, fitted with the Hill equation (Equation 4.1 for QR-DiMan binding with (C) DC-SIGN or (D) DC-SIGNR. (PQR = 1:1; Fitting parameters are summarised in Table 4.1).

The FRET ratio-concentration relationship for DC-SIGN binding with QR-DiMan was fitted very nicely (R^2 > 0.99) by the Hill equation and revealed a K_d of 0.5 ± 0.1 nM at a PQR of 1:1 (**Figure 4.5C; Table 4.1**). This represents a massive 1.8 million-fold enhancement of affinity (β) over the corresponding CRD·DiMan monovalent binding (K_d = 0.9 mM), and a per-glycan normalised affinity enhancement (β/N) of ~4,900 (**Table 4.1**). Interestingly, this affinity is ~4-fold as strong as that of QD-DiMan·DC-SIGN binding (K_d = 2.1 ± 0.5 nM).[20] The sub-nanomolar K_d, here, demonstrates the high suitability of QR-DiMan for potent DC-SIGN targeting. The difference in affinity between QR-DiMan and QD-DiMan in DC-SIGN binding could be due to subtle changes in the inter-glycan distances and/or glycan display curvatures, allowing the former to have a better spatial and/or orientational match with DC-SIGN's four binding sites to form stronger binding than the latter. The FRET ratio for DC-SIGNR binding is considerably lower than that of DC-SIGN at PQR = 1 under equivalent conditions, implying a weaker binding compared to the former. This result is fully consistent with that of DC-SIGN/R binding with QD-DiMan reported previously.[16] The overall FRET ratio-concentration relationship for DC-SIGNR binding with QR-DiMan could be fitted by the Hill equation (R^2 = 0.990), giving an apparent K_d of 3.3 ± 0.9 nM (grey broken line), although several data points were found to deviate considerably from the fitting curve (**Figure 4.5D**). The resulting FRET ratio-concentration relationship appeared to display biphasic binding behaviour, where the FRET ratio exhibits a secondary increase at higher concentrations. By fitting only the first few data points at the low concentration range ([P] ≤ 10 nM), a good fit (R^2 > 0.996) with an apparent K_d of 0.3 ± 0.1 nM was obtained (**Figure 4.5D**). This K_d value is comparable to that of DC-SIGN, suggesting that a similar interaction is taking place, likely involving the same degree of binding multivalency (e.g. binding of all 4 CRDs to the same QR-DiMan). The appearance of the broader (slower increasing signal) secondary binding phase suggests that further binding can occur at higher concentrations, which may indicate the formation of an additional crosslinking interaction, similar to that observed previously with QDs. This is presumably because at higher concentrations, the relatively weak inter-QR binding becomes accessible. In order to obtain a more accurate overall binding affinity between DC-SIGNR and QR-DiMan, the binding assay was performed using a PQR of 10:1 which significantly improved the FRET signals (**Figure 4.6**). Fitting the resulting FRET ratio-[P] relationship with **Equation 4.1**, an apparent binding K_d of 48 ± 9 nM was obtained (**Figure 4.6**), which is ~13-fold stronger than that of the QD-DiMan under equivalent conditions (K_d = ~633 nM, PQR = 10).[20] The binding affinity

enhancement for DC-SIGNR is more substantial than that for DC-SIGN, which is expected due to the presence of the additional high affinity 1:1 binding component as observed with DC-SIGNR under at PQR of 1.

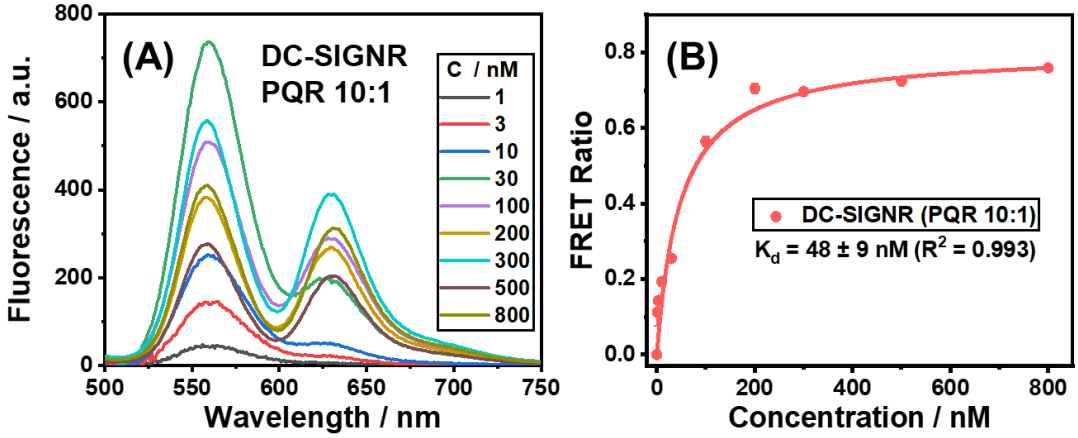

Figure 4.6 (A) Direct excitation background corrected fluorescence spectra over different concentrations of QR-DiMan binding with Atto594 labelled DC-SIGNR at a PQR of 10:1. (B) The corresponding FRET ratio–concentration relationships, fitted with the Hill equation (Equation 4.1). (Fitting parameters are summarised in Table 4.1).

Table 4.1 Summary of fitting parameters obtained from the FRET ratio-concentration relationship for QR-DiMan binding with DC-SIGN and DC-SIGNR fitted with the Hill equation (Equation 4.1). Where, PQR is the protein:QR molar ratio, F_{max} is the FRET ratio at saturation, K_d is the dissociation constant, β is the enhancement factor over the monovalent interaction and N is the number of glycans per QR (R^2 > 0.99 for all fits).

Protein	PQR	F_{max}	K_d / nM	β [a]	β/N	$\frac{K_d^{QD}}{K_d^{QR}}$ [b]
DC-SIGN	1:1	0.37 ± 0.01	0.5 ± 0.1	~1,800,000	~4,900	~4
DC-SIGNR	1:1 [c]	0.056 ± 0.003	0.3 ± 0.1	~3,000,000	~8,100	~2,100
	1:10	0.80 ± 0.03	48 ± 9	~19,000	~51	~13

a $\beta = K_d^{mono}/K_d^{QR}$, K_d^{mono} = 0.9 mM for CRD·DiMan binding.[53]

b K_d^{QD}/K_d^{QR} signifies the enhancement in K_d for QR-DiMan over QD-DiMan. K_d^{QD}= 2.1 ± 0.5 and 633 ± 77 nM for DC-SIGN and DC-SIGNR, respectively.[20]

c Only data points ≤10 nM included in fit.

4.3.3 Binding Mode Determination *via* FRET and Hydrodynamic Size Analysis

In order to more empirically establish the binding modes between the two lectins and QR-DiMan, the effect of titrating protein against a fixed concentration of QR was analysed using FRET and hydrodynamic size analysis. Binding of both lectins with QR-DiMan yielded an initial linear increase of FRET ratio with increasing PQR before reaching saturation (**Figure 4.7**). This behaviour is similar to that observed previously for DC-SIGN binding with QD-DiMan, but is very different from the binding of DC-SIGNR with QD-DiMan, which displays a sigmoidal relationship with very little binding occurring at the low PQRs.[20] This difference agrees with the aforementioned observation that a significant amount of strong affinity, higher order multivalency complexes are established for QR-DiMan·DC-SIGNR complexation, allowing significant binding to occur even at these lower PQRs. By fitting the linear region of the FRET ratio-PQR relationship and taking the intersection with the maximum recorded FRET ratio (**Figure 4.7C**), the "apparent" PQRs required to achieve saturate QR binding (i.e. maximal FRET ratio) are estimated as ~6 for DC-SIGN and ~33 for DC-SIGNR. Please note that these values do not represent the actual number of lectins that are bound to each QR-DiMan but, rather, the number of lectins per QR required to achieve saturate binding because not all added lectins will be able to bind to the QR under the natural association/dissociation equilibrium. As the overall binding affinity of QR-DiMan with DC-SIGNR is significantly weaker than that with DC-SIGN, the proportion of added DC-SIGNR molecules that are bound to the QR would be considerably lower than that for DC-SIGN. In addition, it is also worth noting that these values are likely to be smaller than the "true" PQR required to achieve saturated protein coverage of QR. This is because, according to our previous EDX mapping, the fluorescent core is situated at the centre of the rod,[49] thus only lectins bound within proximity of the central region will be close enough to engage in FRET, due to the inverse sixth power dependency of the FRET efficiency to the distance between the donor and acceptor.[54]

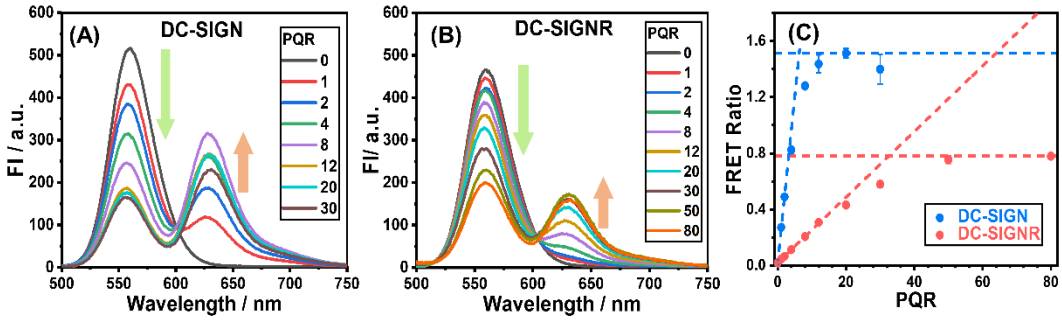

Figure 4.7 Direct excitation background corrected fluorescence spectra corresponding to titration of DC-SIGN (A) or DC-SIGNR (B) against a fixed concentration of QR-DiMan (10 nM). (C) Plots of the corresponding FRET ratio-PQR relationships fitted with linear fits of the initial PQR data points (PQR ≤ 4 or 12 for DC-SIGN and DC-SIGNR, respectively). The fits give intercept = 0.018 ± 0.002 and 0.0194 ± 0.0006, slope = 0.24 ± 0.01 and 0.0234 ± 0.0002 for DC-SIGN and DC-SIGNR, respectively. (R^2 ≥ 0.99 for all fits)

The binding modes of the two proteins with QR-DiMan were further confirmed by analysing the hydrodynamic diameters (D_h) of the resulting QR-lectin complexes. The apparent D_h values were obtained from Gaussian fits of the D_h distribution histograms over a PQR range of 0 to 20. For QR-DiMan·DC-SIGN complexation, only a single size distribution is observed, which plateaus at ~60 nm (**Figure 4.8A-D; Figure C.3Figure** C.5) after a PQR of ~6. This value is similar to the summation of the D_h values of a single QR-DiMan flanked by two proteins (where QR-DiMan and DC-SIGN demonstrate individual D_h values of ~21 and ~14 nm, respectively). This size is therefore likely to be representative of a monolayer of lectin with all CRDs simultaneously specifically bound to a single QR-DiMan particle. QR-DiMan·DC-SIGNR complexation, on the other hand, demonstrates two distinct size distributions (**Figure 4.8E-H**): one plateaus at ~60 nm and matches well with that of DC-SIGN complexation, while the other plateaus at ~140 nm. The similarity in D_h values between the smaller size peak in DC-SIGNR and that observed for DC-SIGN evidently confirms the presence of simultaneous binding in DC-SIGNR. Whereas, the larger size distribution with DC-SIGNR is indicative of the formation of crosslinked clusters, similar to that observed previously with QDs, although the size here is somewhat smaller.[20] Fitting the average D_h (weighted with respect to their integrated areas) against PQR using an offset Hill function (**Figure 4.8I**), it is observed that the average D_h plateaus at 62 ± 6 and 150 ± 10 nm for DC-SIGN and DC-SIGNR, respectively. This demonstrates that the number of DC-SIGNR clustered particles tends to increase with increasing PQR. This result provides further evidence that the high affinity interaction observed at low concentrations for

DC-SIGNR does indeed correlate with the simultaneous binding mode, and the weaker affinity binding correlates with an increase of crosslinked binding at the higher PQRs.

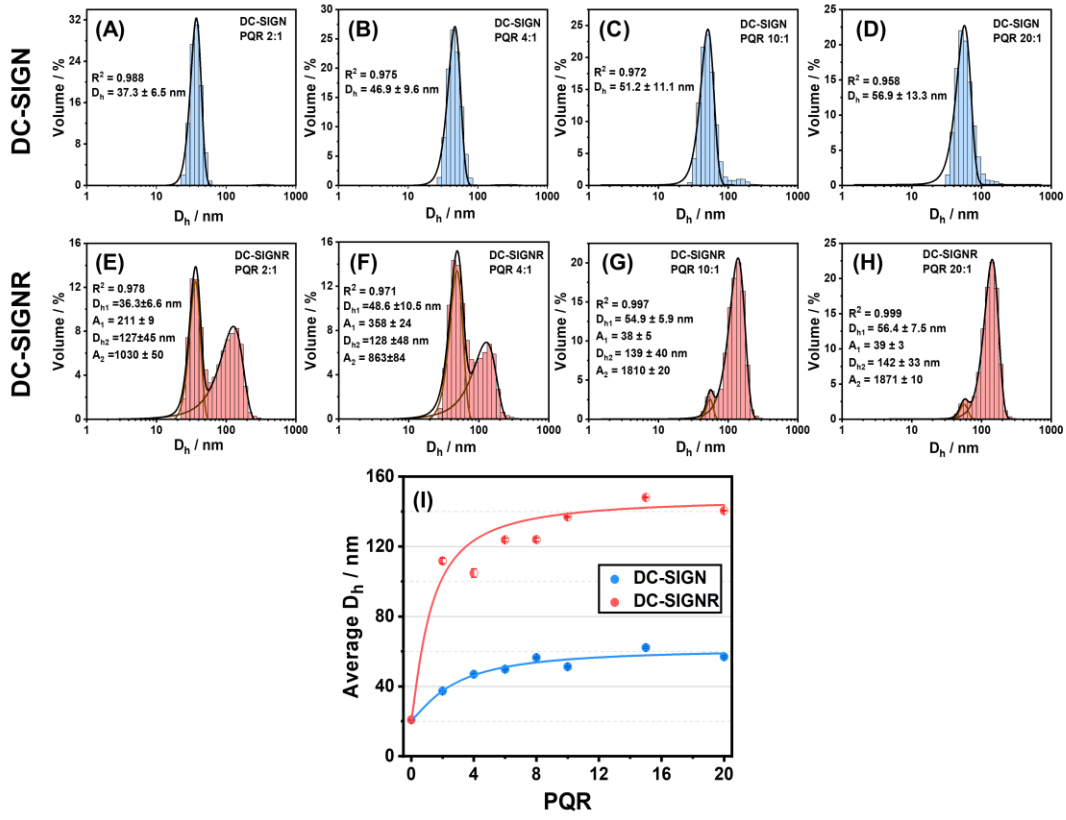

Figure 4.8 Volume population hydrodynamic size distribution histograms fitted with uni- or bi-modal Gaussian functions for QR-DiMan (10 nM) after binding with DC-SIGN at PQRs of (A) 2:1, (B) 4:1, (C) 10:1 and (D) 20:1 or binding with DC-SIGNR at PQRs of (E) 2:1, (F) 4:1, (G) 10:1, and (H) 20:1. (I) The corresponding average D_h-PQR relationship (average $D_h = D_{h,1} \times A_1\% + D_{h,2} \times A_2\%$, where $A_1\%$ and $A_2\%$ are the percentage area of the Gaussian fits; filled circles: single distribution; half-filled circles: two distributions) fitted with an offset Hill function ($D_{h,PQR} = D_{h,0} + (D_{h,\infty} - D_{h,0})/(1 + (PQR_{50}/PQR)^n)$ where $D_{h,0}$ was fixed to the D_h of QR-DiMan, $D_{h,\infty}$ = 62 ± 6 and 150 ± 10 nm, PQR_{50} (PQR at 50% of $D_{h,\infty}$) = 2.7 ± 0.8 and 1.3 ± 1, and n = 1.3 ± 0.4 and 1 ± 1, for DC-SIGN and DC-SIGNR, respectively; $R^2 \geq 0.99$). (DC-SIGN: blue; DC-SIGNR: red).

To confirm the binding affinity-mode relationship, free D-mannose (Man) was added to compete for pre-formed QR-DiMan·DC-SIGN/R complexes prepared under a PQR of 4:1. DLS analysis showed that the amount of clustered QR-DiMan·DC-SIGNR species was decreased even with addition of just 0.1 mM of Man (**Figure 4.9E; Figure C.6Figure** C.8). This is a clear

indication that the crosslinked clusters correspond to the weaker binding mode, which is more easily displaced than the stronger simultaneous binding. The average D_h of the smaller species was also reduced to ~30 nm with ≥ 4 mM Man (**Figure 4.9G**), along with the complete disappearance of the clustered species, indicating the eventual breakdown of both binding modes.

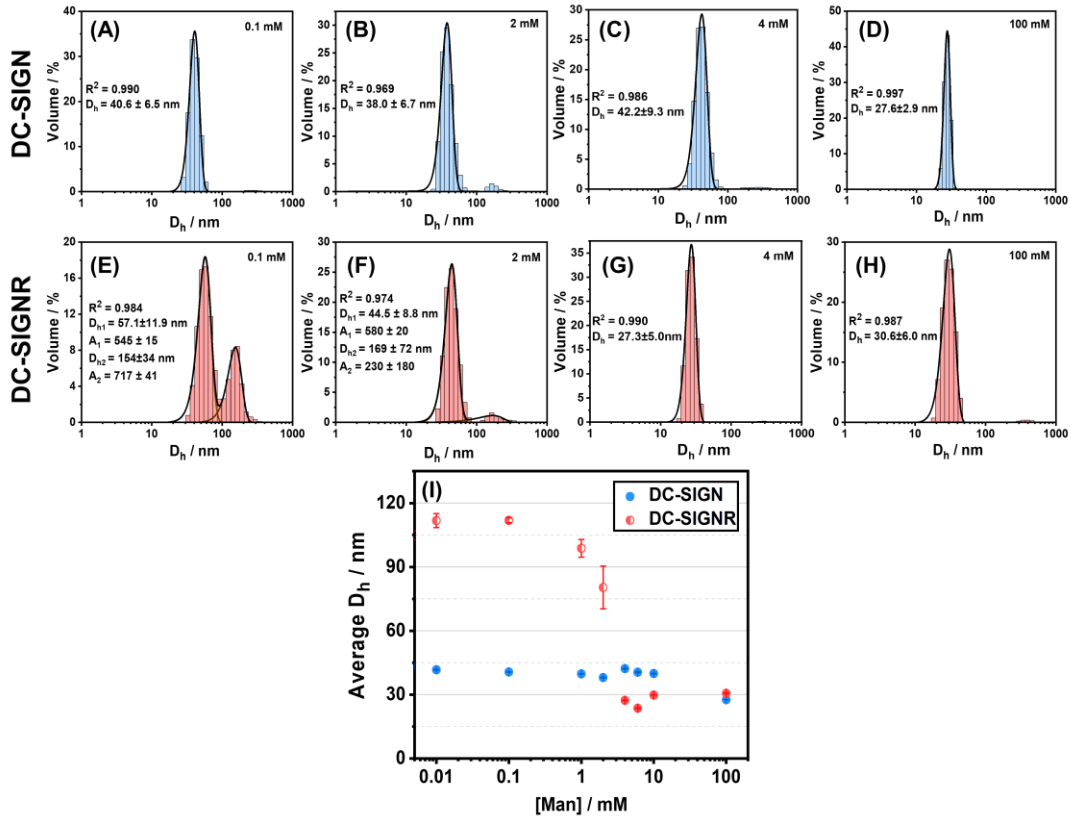

Figure 4.9 Volume population hydrodynamic size distribution histograms fitted with uni- or bi-modal Gaussian functions for a pre-incubated mixture of QR-DiMan (10 nM) with DC-SIGN (40 nM) after addition of free mannose (Man) at concentrations of (I) 0.1, (J) 2, (K) 4 or (L) 100 mM; and a pre-incubated mixture of QR-DiMan (40 nM) with DC-SIGNR (40 nM) after addition of Man at concentrations of (M) 0.1, (N) 2, (O) 4 and (P) 100 mM. (S) Plot of the corresponding average D_h against [Man] (average $D_h = D_{h,1} \times A_1\% + D_{h,2} \times A_2\%$, where $A_1\%$ and $A_2\%$ are the percentage area of the Gaussian fits; filled circles: single distribution; half-filled circles: two distributions). (DC-SIGN: blue; DC-SIGNR: red).

Moreover, addition of Man to QR-DiMan·DC-SIGN/R complexes also led to a significant, dose-dependent reduction in FRET ratio (**Figure 4.10**), indicated by the simultaneous reduction of dye fluorescence and recovery of QR fluorescence. These results are fully consistent with free Man induced lectin·glycan-QR complex dissociation. The FRET ratio change with Man concentration, [Man], was then fitted with an offset Hill equation to obtain the

apparent-inhibition constant (K_i), which represents the "apparent" [Man] required to inhibit binding by 50% (**Equation 4.2**; where F is the FRET ratio at a particular $[\mathrm{Man}]$, and n represents the Hill coefficient; **Figure 4.10C**).

$$F = F_0 + \frac{(F_\infty - F_0)}{1 + \left(\frac{K_i}{[\mathrm{Man}]}\right)^n} \qquad 4.2$$

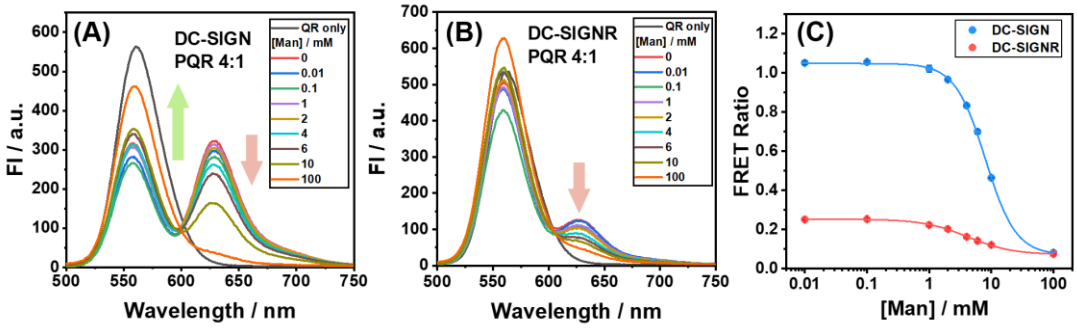

Figure 4.10 Direct excitation background corrected fluorescence spectra corresponding to increasing concentrations of free mannose to a pre-incubated 4:1 PQR mixture of 10 nM QR-DiMan with (A) DC-SIGN or (B) DC-SIGNR. (C) A plot of the relationship between FRET ratio against mannose concentration for the fluorescence spectra recorded in A and B, fitted with an offset Hill function (Equation 4.2; where F_0 = 1.049 ± 0.007 and 0.253 ± 0.001, F_∞ = 0.071 ± 0.003 and 0.072 ± 0.003, K_i ([Man] at 50% inhibition) = 8.0 ± 0.1 and 4.2 ± 0.2 mM, and n = 1.78 ± 0.05 and 1.21 ± 0.08, for DC-SIGN (blue) and DC-SIGNR (red), respectively; R^2 > 0.999 for both fits).

Here, the apparent-K_i values for DC-SIGN and DC-SIGNR were estimated as 8.0 ± 0.1 and 4.2 ± 0.2 mM, respectively. The higher K_i value for DC-SIGN is a reflection of its stronger overall binding affinity with QR-DiMan, and hence requires a higher [Man] in order to displace 50% of binding interactions. However, DC-SIGNR displays a broader decay with increasing [Man] than DC-SIGN does, evidenced by the smaller exponent (n = 1.78 ± 0.05 *vs* 1.21 ± 0.08 for DC-SIGN *vs* DC-SIGNR). This suggests that Man is able to displace a larger amount of DC-SIGNR binding at much lower concentrations than DC-SIGN. It is worth noting that, where the difference in the overall apparent-K_d between DC-SIGN/R binding to QR-DiMan is ~100-fold, the difference in the apparent-K_i with Man is only ~2 fold. This may suggest that, for a PQR of 4:1, the simultaneous binding mode of DC-SIGNR provides a higher contribution to the apparent-K_i than the crosslinking mode, where the latter contributes

mainly to broadening of the decay. Interestingly, because the fluorescent core of the QR has been shown to be located in the centre of the cylindrical middle section of the QR,[49] any binding at the middle section will have a larger influence on the apparent-K_{i}. This result may therefore suggest that the glycan display at the cylindrical section of the rod is better suited to form simultaneous binding with DC-SIGNR than crosslinking interactions. Consistent with this, the decay in D_{h} for the simultaneously bound complexes also occurs at the same [Man] range (**Figure 4.9I**).

Overall, these results collectively demonstrate that our QR-DiMan is able to distinguish between the strong simultaneous- and weaker crosslinking-binding modes by combining both FRET and hydrodynamic size analyses.

4.3.4 Binding Mode Rationale *via* S/TEM Imaging

Electron microscopy was further employed to capture "snapshot" images of the QR-lectin complexes in order to provide a more detailed understanding of their interactions. Here, QR-lectin samples, prepared with a PQR of 4:1, were plunge-frozen and then vacuum dried, before being placed for S/TEM imaging. We have shown previously that this method allows for the successful capture of the native dispersion state of nanoparticle assemblies.[20, 55] Binding of DC-SIGN demonstrated mostly isolated individual QRs (**Figure 4.11A**) which correlated nicely with the single D_{h} distribution for DC-SIGN·QR-DiMan complexes observed by DLS (**Figure 4.8B**). Additionally, binding of DC-SIGNR yielded both clustered QR assemblies and non-clustered individual QRs (**Figure 4.11B**), which again agreed well with the two distinct D_{h} species observed in DLS (**Figure 4.8F**).

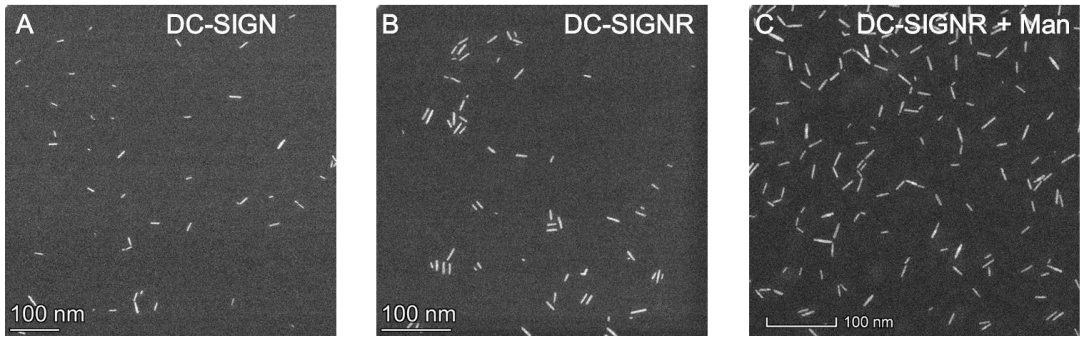

Figure 4.11 S/TEM images of cryo-prepared QR-DiMan (10 nM) after complexation with 4 eq of (A) DC-SIGN, (B) DC-SIGNR, and (C) DC-SIGNR in the presence of Man (2 mM).

To further quantify assembly formation, the inter-QR distances were analysed by measuring the perimeter-perimeter nearest neighbour distances (ppNNDs)

between each QR (i.e. the shortest distance between the perimeter of one QR with that of its nearest neighbouring QR). The distributions of ppNNDs were then plotted as histograms and fitted with Gaussian distribution curves. This analysis revealed three well-defined species for QR-DiMan·DC-SIGNR assemblies with ppNNDs of 1.9 ± 0.8, 4.0 ± 1.1 and 6.7 ± 1.0 nm (mean ± ½FWHM; **Figure 4.12B; Table 4.2**), respectively. DC-SIGN, on the other hand, displayed only one well-defined ppNND distribution at 2.3 ± 1.1 nm (**Figure 4.12; Table 4.2**). These results allowed us to draw three conclusions. (1) The smallest ppNND species (~2 nm) were observed in similar amounts for both DC-SIGN and DC-SIGNR (20-27%). Such distances are comparable to the thickness of the glycan ligand coating, thus these QR assemblies are deemed to result from non-specific interactions. (2) The species with ppNNDs $\gtrsim$ 7.7 nm were found in 71% of QR-DiMan·DC-SIGN and 31% of QR-DiMan·DC-SIGNR complexes and were randomly distributed (**Figure C.10**). These corresponded nicely to single QR-DiMan particles bound with a layer of proteins as observed by DLS, thus confirming that both DC-SIGN and DC-SIGNR were able to bind tetravalently with all CRDs with one QR-DiMan. (3) The well-defined species for DC-SIGNR binding with ppNNDs of 4.0 ± 1.1 and 6.7 ± 1.0 nm were not observed in significant amount with DC-SIGN (**Figure 4.12A,B**), such distances were consistent with the discrete distances expected for DC-SIGNR-bridged QRs. In addition, 49% of QRs with ppNNDs $\lesssim$ 8 nm contained > 2 QRs per cluster for DC-SIGNR, which were only observable in negligible amounts for DC-SIGN (< 3% of QRs, and none with >3 QRs per cluster). These multi-QR assemblies are consistent with the larger D_{h} sizes observed by DLS. Interestingly, these are only limited to a few QRs per cluster, thus do not resemble the extensive intercrosslinked networks observed for QD-DiMan bridged by DC-SIGNR.[20] This is likely due to the dual simultaneous and crosslinking binding mode which imparts a limit to the number of crosslinking interactions possible.

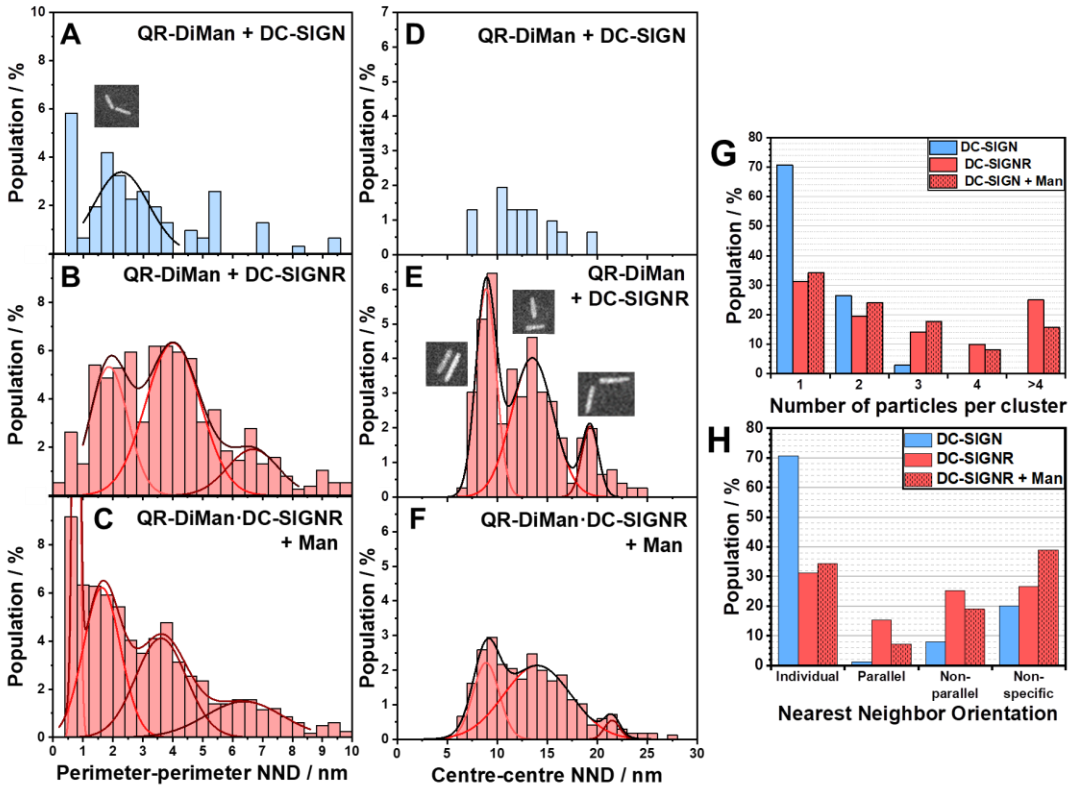

Figure 4.12 Histograms of the perimeter-perimeter nearest neighbour distance (ppNND), fitted with uni- or multi- modal Gaussian fits, for the S/TEM images of cryo-prepared QR-DiMan incubated with (A) DC-SIGN, (B) DC-SIGNR or (C) pre-incubated sample of QR-DiMan·DC-SIGNR with 2 mM Man. Histograms of the centre-centre nearest neighbour distances (ccNNDs, for 2.9 < ppNND < 7.7 nm), fitted with trimodal Gaussian fits (where applicable), for QR-DiMan incubated with (D) DC-SIGN, (E) DC-SIGNR or (F) pre-incubated QR-DiMan·DC-SIGNR with 2 mM Man. Statistical analysis of the (G) numbers of QRs per cluster (i.e. the number of interconnected QRs with a ppNND < 7.7 nm) and (H) QR nearest neighbour orientations (individual: ppNND > 7.7 nm; parallel: the area of the first ccNND peak (or ccNND <10 nm for DC-SIGN); non-parallel: the sum of the second and third ccNND peak areas (or ccNND >10 nm for DC-SIGN); non-specific: ppNND < 2.9 nm).

Table 4.2 Summary of fitting parameters obtained from the Gaussian fits of the nearest neighbour distances (NND) of QR-DiMan particles binding with DC-SIGN, DC-SIGNR or DC-SIGNR with 2 mM Man. Where ppNND is the perimeter-perimeter NND and ccNND is the centre-centre NND (where 2.9 < ppNND < 7.7 nm).

Perimeter-perimeter NND				
Protein	**Mean ppNND / nm**	**FWHM / nm**	**Area / %**	**R^2**
DC-SIGN	2.3 ± 0.1	2.2 ± 0.4	8 ± 1	0.747
DC-SIGNR	1.9 ± 0.2 4.0 ± 0.2 6.7 ± 0.7	1.5 ± 0.5 2.2 ± 0.9 2 ± 1	8 ± 4 15 ± 6 4 ± 3	0.869
DC-SIGNR + Man	1.63 ± 0.09 3.6 ± 0.2 6.4 ± 0.6	1.5 ± 0.4 2.0 ± 0.6 3 ± 1	10 ± 3 9 ± 3 5 ± 2	0.987 [a]
Perimeter-perimeter NND				
Protein	**Mean ccNND / nm**	**FWHM / nm**	**Area / %**	**R^2**
DC-SIGN	-	-	-	-
DC-SIGNR	8.9 ± 0.1 13.5 ± 0.3 19.3 ± 0.3	2.4 ± 0.3 4.9 ± 0.8 2.0 ± 0.6	15 ± 2 21 ± 3 4 ± 1	0.939
DC-SIGNR + Man	8.9 ± 0.1 14.0 ± 0.4 21.5 ± 0.3	3.1 ± 0.4 7.9 ± 0.9 1.8 ± 0.6	7 ± 2 18 ± 2 1.1 ± 0.4	0.980

[a] An additional peak was observed at ~0.8 nm which was an artefact of the image resolution and fit poorly to the data.

To investigate how Man competes with the QR-DiMan·DC-SIGNR binding, S/TEM images were performed on a sample of pre-incubated QR-DiMan·DC-SIGNR complex with 2 mM Man (**Figure 4.11C; Figure 4.12C**). Here, a ~40% reduction of clusters containing > 4 QRs was observed (**Figure 4.12G**), consistent with the significantly reduced mean D_{h} observed by DLS (**Figure 4.9F**). In parallel to the decrease in crosslinking, a ~10% increase in the proportion of individual complexes and a ~50% increase in the

proportion of non-specific interactions was observed (**Figure 4.12H**). This result may explain why the D_h of QR-DiMan did not completely return to its original size after dissociation of protein observed at high Man concentrations.

Given the previously established crosslinking character of spherical QD-DiMan with DC-SIGNR,[20] any deviation in binding mode induced by QRs, with a similar radius and chemical composition, is likely to be imparted by the central cylindrical section, where the glycan display differs most from that of QDs. Therefore, simultaneous binding of DC-SIGNR is most likely to occur in central cylindrical region of the QR while crosslinking is more likely to occur at the spherical ends. This can be investigated by analysing how the QR-DiMan particles orient themselves with respect to each other in the presence of crosslinking DC-SIGNR. This was achieved by measuring the centre-centre nearest neighbour distance (ccNND; the shortest distance between the centre of one QR with that of its nearest neighbour) for all QRs with a ppNND between 4.0 ± 1.1 to 6.7 ± 1.0 nm. As expected, no clear ccNND distribution was observed for QR-DiMan·DC-SIGN complexes (**Figure 4.12D**). In sharp contrast, for QR-DiMan·DC-SIGNR, three discrete distributions were obtained with Gaussian fits with ccNNDs of 8.9 ± 1.2, 13.5 ± 2.4, and 19.3 ± 1.0 nm (mean ± ½FWHM; **Figure 4.12E; Table 4.2**), respectively. These distributions are representative of QRs which are stacked either parallel (i.e. centre to centre), perpendicular (i.e. end to centre) or adjacent (i.e. end to end) to one another, respectively (as depicted in **Figure 4.12E**).

If crosslinking did not discriminate between the different rod regions then, based on surface area alone, the most common inter-QR orientation would be QRs stacked loosely parallel to one another. However, the ccNND distributions show this is not the most prevalent orientation. Instead, only 15% of QRs are stacked parallel to their nearest neighbour, and 25% of QRs display non-parallel nearest neighbour orientations (**Figure 4.12H**), suggesting that crosslinking favours the spherical QR ends over the central cylindrical section. In addition, these parallel interactions were the only nearest neighbour orientation that showed a significant reduction upon addition of the Man competitor. Therefore, it is plausible that these parallel QR stacks are comprised of protein crosslinking at both QR ends, which are then either fully dissociated into isolated particles or partially dissociated into nonparallel interactions in the presence of Man. This is feasible as further crosslinking between adjacent QR-lectin complexes within the same original QR cluster would be much more kinetically favourable than crosslinking multiple QR-lectin complexes that were freely diffusing in solution. Based on

these observations, three interesting conclusions can be deduced: (1) DC-SIGN does not discriminate between the spherical ends and cylindrical section of the QR scaffold; (2) DC-SIGNR crosslinks with spherical ends of the QR; and (3) DC-SIGNR simultaneously binds to cylindrical section of the QR (**Figure 4.13**).

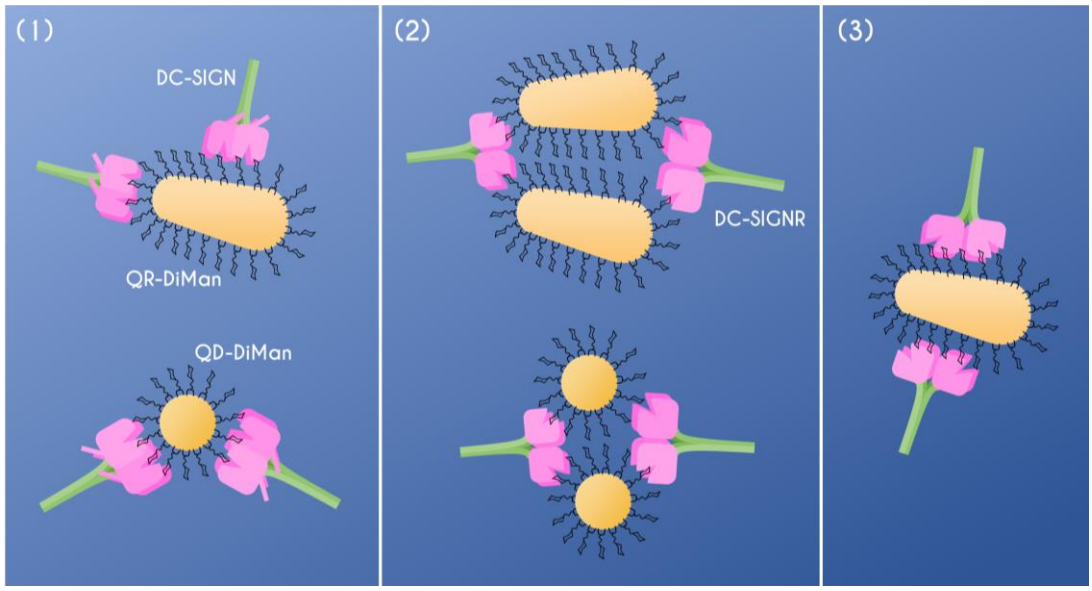

Figure 4.13 Schematics depicting (1) the simultaneous binding of DC-SIGN to all sections of both QR-DiMan and QD-DiMan; (2) DC-SIGNR crosslinking with spherical QD-DiMan or with the spherical end sections of QR-DiMan into parallel stacks; and (3) the simultaneous binding of DC-SIGNR to the cylindrical section of QR-DiMan, only.

These different binding phenomena can be rationalised by considering the relative dimensions of both binding partners. The hydrodynamic dimensions of QR-DiMan can be estimated from the summation of the average QR core dimensions obtained by S/TEM (core length, L_{core} = 14.7 ± 5.7 nm; core diameter, D_{core} = 3.1 ± 0.7 nm; **Figure 4.2B**) and the estimated hydrodynamic surface ligand length (2.9 ± 1.1 nm; derived from the previous QD-DiMan dimensions, D_{h} = 9.5 ± 0.1 nm and D_{core}= 3.7 ± 2.1 nm; **Figure E.1Figure** E.2).[20] This provided a QR-DiMan estimated terminal end D_{h} and cylindrical height (H) of 8.9 ± 2.3 nm and 11.6 ± 5.7 nm, respectively (see **Appendix E.2**). Both of these dimensions are comparable with that of the equivalent QD-DiMan scaffold (D_{h} = 9.5 ± 0.1 nm).[20] Due to the flexible nature of the EG_{11} chains, the surface area and inter-glycan distance of each glycan ligand can be assumed to be roughly the same across the whole QR (1.6 nm^2 and 1.4 nm, respectively; **Appendix B.1**). The crystal structure of DC-SIGNR C-terminal tetrameric fragment (PDB code 1XAR) provides approximate

dimensions of 3.8 × 8.0 nm between the primary Ca^{2+} ions associated with glycan binding (**Figure E.3A**).[56] This means that the binding contact area of each DC-SIGNR is likely to be smaller than both the spherical and cylindrical regions of the QR, and would cover a QR surface area containing ~20 glycans. These simple calculations suggest that neither the QR size nor the inter-glycan distance are likely to be the main factor causing the distinct binding mode for DC-SIGNR between the QR end and middle sections. Instead, the distinct DC-SIGNR binding modes are most likely to be determined by the large differences in the 3-dimensional surface curvatures between these regions. In fact, at the spherical ends of the QR, surface curvature could theoretically impart separation distances (d) as much as ~4 nm from the protein binding sites to the glycan surface (**Equation 4.3**). Here, p is the Ca-Ca distance between two CRDs at the furthest separation distance from the scaffold surface, which is taken as the average diagonal Ca-Ca distance (**Figure 4.14 left**), where p ~ 8.8 nm for DC-SIGNR.

$$d = \frac{D_h - \sqrt{{D_h}^2 - p^2}}{2} \qquad 4.3$$

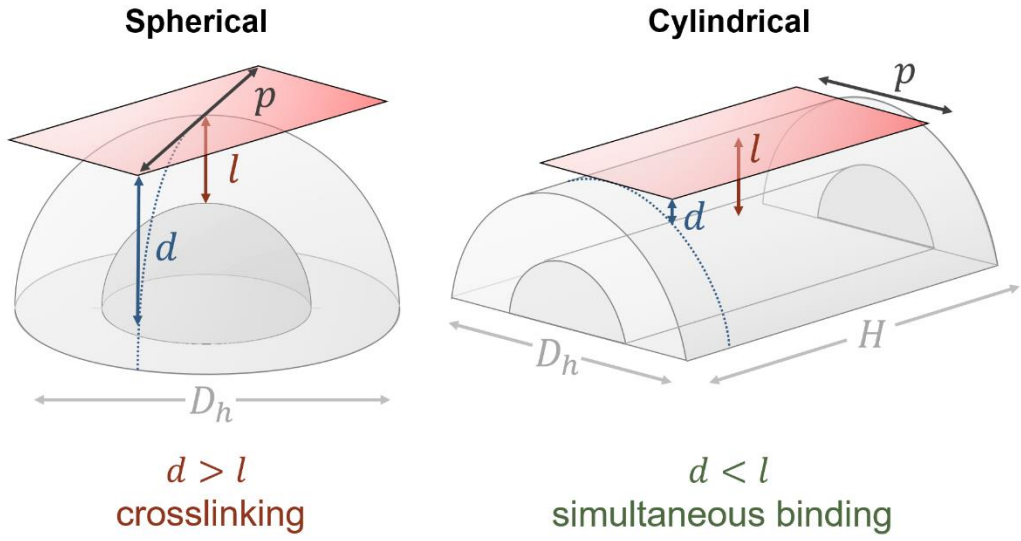

Figure 4.14 Schematic depicting the difference in separation distance between the binding contact area of DC-SIGNR (red) with either the spherical end (left) or cylindrical section (right) of the QR surface. Where d is the separation distance between the protein contact area and the QR-DiMan surface, l is the maximum compression length of the surface glycan ligands, p is the Ca-Ca distance between two CRDs at the furthest separation distance from the QR surface, D_h is the estimated QR-DiMan hydrodynamic diameter, and H is the height of the QR cylindrical section. (Not to scale).

The theoretical maximum length at which the glycan ligands may be able to be compressed (l) can be derived from the estimated length of flexibility in the glycan ligand, i.e. l ~2.1 nm for DHLA-EG$_{11}$-DiMan (estimated by the length of the rigid lipoic acid segment (~0.8 nm; PDB code: LPA) subtracted from the total hydrodynamic ligand length). Therefore, even with compression of the ligands, it would be impossible for all four binding sites to reach the glycan surface simultaneously. This means that the glycan displays at the QR ends are incapable of bridging all four binding sites in DC-SIGNR, thus the protein cannot access its most stable simultaneously bound state. Instead, it is forced to find its next most favourable configuration by crosslinking with other particles to maximize binding enthalpy (see **Chapter 3**). For the cylindrical section of the scaffold, the curvature of the round of the cylinder is the same as the spherical ends, however, along the length of the cylinder, it is roughly flat. This means that the tetrameric lectin has the opportunity to align itself with its longer length parallel to the length of the rod (**Figure 4.14 right**). Therefore, only the short length of the tetramer (i.e. p = 3.8 nm) needs to contribute to the separation distance (*d*) from the rod surface, resulting in a d of only ~0.4 nm. This is well within the maximum compression length of the ligands, therefore allowing glycans to easily bridge all 4 CRDs to give strong simultaneous tetravalent binding. Finally, the tetrameric model of DC-SIGN has been predicted to exhibit a more compact shape with an average diagonal inter-binding site distance of 5.6 nm (**Figure E.3B**).[57] This results in a much smaller separation distance between the protein contact area (d ~1 nm), which is smaller than the maximum compression length of the flexible surface ligands (l = 2.1 nm). This would therefore easily allow the QR surface glycans to bridge all 4 CRDs, regardless of the region of the QR that DC-SIGN binds to, leading to the exclusively simultaneous binding mode as observed here. These calculations therefore show that while the tetrameric structures of both lectins are still unknown, the good agreement between the predicted and observed results suggests that the structural models for DC-SIGN/R are likely to be relatively reliable. Moreover, it also demonstrates that, by taking into account nano-scaffold curvature, estimating the d-l relationship between the protein and nano-scaffold dimensions can provide a useful prediction of the binding mode for multimeric lectins.

4.4 Conclusion

In summary, this chapter presents a new glycan-QR based multifunctional biophysical probe for MLGIs. By combining FRET, hydrodynamic size and S/TEM imaging analysis, geometric influences of glyconano-materials on MLGI properties have been dissected using DC-SIGN and DC-SIGNR, a pair of tetrameric lectin models with almost identical monovalent binding motifs but distinct binding site arrangements. It is demonstrated that, given ample polyvalent glycan density and area, scaffold curvature has a fundamental impact on the binding modes of MLGIs. Here, DC-SIGNR is able to distinguish between the end and middle sections of QR-DiMan, forming strong simultaneous tetravalent binding at the central cylindrical section but bis-bivalent crosslinking at the spherical ends. Whereas, DC-SIGN binds with strong simultaneous tetravalent binding irrespective of the QR section. It has been further predicted that only curvatures affording a separation distance between the protein contact area and the glycan ligand surface (d) smaller than the compression length of the glycan ligands (l) can result in the strong simultaneous binding of all binding sites. However, if d is greater than l, then only weak crosslinking or other lower valency binding interactions can occur. This result thus demonstrates how multimeric lectins like DC-SIGN/R differentiate glycan displays with different geometries, which may help explain some of their differences in virus-binding and transmitting properties. This new information will be useful for guiding the design of specific multivalent glycans for targeting such MLGIs.

References

1. L. Fruk and A. Kerbs, *Bionanotechnology: Concepts and Applications*, Cambridge University Press, Cambridge, 2021.
2. E. N. Zare, X. Zheng, P. Makvandi, H. Gheybi, R. Sartorius, C. K. Y. Yiu, M. Adeli, A. Wu, A. Zarrabi, R. S. Varma and F. R. Tay, *Small*, 2021, **17**, 2007073.
3. A. Albanese, P. S. Tang and W. C. Chan, *Annu. Rev. Biomed. Eng.*, 2012, **14**, 1-16.
4. S. H. Liyanage and M. Yan, *Chem. Commun. (Camb)*, 2020, **56**, 13491-13505.
5. B. D. Chithrani, A. A. Ghazani and W. C. W. Chan, *Nano Lett.*, 2006, **6**, 662-668.
6. A. Albanese and W. C. W. Chan, *ACS Nano*, 2011, **5**, 5478-5489.
7. S. Sangabathuni, R. Vasudeva Murthy, P. M. Chaudhary, M. Surve, A. Banerjee and R. Kikkeri, *Nanoscale*, 2016, **8**, 12729-12735.
8. Z. Li, L. Sun, Y. Zhang, A. P. Dove, R. K. O'Reilly and G. Chen, *ACS Macro Lett.*, 2016, **5**, 1059-1064.
9. X. Xie, J. Liao, X. Shao, Q. Li and Y. Lin, *Sci. Rep.*, 2017, **7**, 3827.
10. L. Ding, C. Yao, X. Yin, C. Li, Y. Huang, M. Wu, B. Wang, X. Guo, Y. Wang and M. Wu, *Small*, 2018, **14**, 1801451.

11. M. Kersting, M. Olejnik, N. Rosenkranz, K. Loza, M. Breisch, A. Rostek, G. Westphal, J. Bünger, N. Ziegler, A. Ludwig, M. Köller, C. Sengstock and M. Epple, *Sci. Rep.*, 2020, **10**, 21591.
12. K. Niikura, T. Matsunaga, T. Suzuki, S. Kobayashi, H. Yamaguchi, Y. Orba, A. Kawaguchi, H. Hasegawa, K. Kajino, T. Ninomiya, K. Ijiro and H. Sawa, *ACS Nano*, 2013, **7**, 3926-3938.
13. S. Guo, H. Li, M. Ma, J. Fu, Y. Dong and P. Guo, *Mol. Ther. Nucleic. Acids*, 2017, **9**, 399-408.
14. J. D. Reuter, A. Myc, M. M. Hayes, Z. Gan, R. Roy, D. Qin, R. Yin, L. T. Piehler, R. Esfand, D. A. Tomalia and J. R. Baker, *Bioconjug. Chem.*, 1999, **10**, 271-278.
15. S. Bhatia, D. Lauster, M. Bardua, K. Ludwig, S. Angioletti-Uberti, N. Popp, U. Hoffmann, F. Paulus, M. Budt, M. Stadtmüller, T. Wolff, A. Hamann, C. Böttcher, A. Herrmann and R. Haag, *Biomaterials*, 2017, **138**, 22-34.
16. S. Toraskar, M. Gade, S. Sangabathuni, H. V. Thulasiram and R. Kikkeri, *ChemMedChem*, 2017, **12**, 1116-1124.
17. G. D. Brown, J. A. Willment and L. Whitehead, *Nat. Rev. Immunol.*, 2018, **18**, 374-389.
18. A. T. Thodikayil, S. Sharma and S. Saha, *ACS Appl. Bio Mater.*, 2021, **4**, 2907-2940.
19. D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo and D. Zhou, *J. Am. Chem. Soc.*, 2020, **142**, 18022-18034.
20. Y. Guo, I. Nehlmeier, E. Poole, C. Sakonsinsiri, N. Hondow, A. Brown, Q. Li, S. Li, J. Whitworth, Z. Li, A. Yu, R. Brydson, W. B. Turnbull, S. Pöhlmann and D. Zhou, *J. Am. Chem. Soc.*, 2017, **139**, 11833-11844.
21. Y. Guo, C. Sakonsinsiri, I. Nehlmeier, M. A. Fascione, H. Zhang, W. Wang, S. Pöhlmann, W. B. Turnbull and D. Zhou, *Angew. Chem. Int. Ed.*, 2016, **55**, 4738-4742.
22. A. Nazemi, S. M. Haeryfar and E. R. Gillies, *Langmuir*, 2013, **29**, 6420-6428.
23. L. Bes, S. Angot, A. Limer and D. M. Haddleton, *Macromolecules*, 2003, **36**, 2493-2499.
24. J. Luczkowiak, A. Muñoz, M. Sánchez-Navarro, R. Ribeiro-Viana, A. Ginieis, B. M. Illescas, N. Martín, R. Delgado and J. Rojo, *Biomacromolecules*, 2013, **14**, 431-437.
25. K. Buffet, E. Gillon, M. Holler, J.-F. Nierengarten, A. Imberty and S. P. Vincent, *Org. Biomol. Chem.*, 2015, **13**, 6482-6492.
26. R. Rísquez-Cuadro, J. M. García Fernández, J.-F. Nierengarten and C. Ortiz Mellet, *Chem. Eur. J.*, 2013, **19**, 16791-16803.
27. D. Arosio, F. Chiodo, J. J. Reina, M. Marelli, S. Penadés, Y. van Kooyk, J. J. Garcia-Vallejo and A. Bernardi, *Bioconjug. Chem.*, 2014, **25**, 2244-2251.
28. A. Bernardi, J. Jimenez-Barbero, A. Casnati, C. De Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K. E. Jaeger, M. Lahmann, T. K. Lindhorst, M. Marradi, P. Messner, A. Molinaro, P. V. Murphy, C. Nativi, S. Oscarson, S. Penades, F. Peri, R. J. Pieters, O. Renaudet, J. L. Reymond, B. Richichi, J. Rojo, F. Sansone, C. Schaffer, W. B. Turnbull, T. Velasco-Torrijos, S. Vidal, S. Vincent, T. Wennekes, H. Zuilhof and A. Imberty, *Chem. Soc. Rev.*, 2013, **42**, 4709-4727.
29. A. M. DiLillo, K. K. Chan, X.-L. Sun and G. Ao, *Front. Chem.*, 2022, **10**, 852988.
30. B.-S. Kim, D.-J. Hong, J. Bae and M. Lee, *J Am Chem Soc*, 2005, **127**, 16333-16337.
31. Y. Chen, H. Vedala, G. P. Kotchey, A. Audfray, S. Cecioni, A. Imberty, S. Vidal and A. Star, *ACS Nano*, 2012, **6**, 760-770.
32. P. M. Chaudhary, S. Sangabathuni, R. V. Murthy, A. Paul, H. V. Thulasiram and R. Kikkeri, *Chem. Commun. (Camb)*, 2015, **51**, 15669-15672.
33. T. Koukalová, P. Kovaříček, P. Bojarová, V. L. P. Guerra, V. Vrkoslav, L. Navara, I. Jirka, M. Cebecauer, V. Křen and M. Kalbáč, *Int. J. Mol. Sci.*, 2021, **22**, 6661.
34. T. Tanaka, H. Ishitani, Y. Miura, K. Oishi, T. Takahashi, T. Suzuki, S.-i. Shoda and Y. Kimura, *ACS Macro Lett.*, 2014, **3**, 1074-1078.
35. K. Suzuki, T. Koyama, S. Yingsakmongkon, Y. Suzuki, K. Hatano and K. Matsuoka, *Bioorg. Med. Chem.*, 2012, **20**, 446-454.

36. M. Nagao, Y. Kurebayashi, H. Seto, T. Tanaka, T. Takahashi, T. Suzuki, Y. Hoshino and Y. Miura, *Polym. J.*, 2016, **48**, 745-749.
37. J. Haldar, L. Alvarez de Cienfuegos, T. M. Tumpey, L. V. Gubareva, J. Chen and A. M. Klibanov, *Pharm. Res.*, 2010, **27**, 259-263.
38. C. R. Becer, M. I. Gibson, J. Geng, R. Ilyas, R. Wallis, D. A. Mitchell and D. M. Haddleton, *J. Am. Chem. Soc.*, 2010, **132**, 15130-15132.
39. I. Papp, C. Sieben, A. L. Sisson, J. Kostka, C. Bottcher, K. Ludwig, A. Herrmann and R. Haag, *Chembiochem*, 2011, **12**, 887-895.
40. J. Ramos-Soriano and J. Rojo, *Chem. Commun.*, 2021, **57**, 5111-5126.
41. M. Pernía Leal, M. Assali, J. J. Cid, V. Valdivia, J. M. Franco, I. Fernández, D. Pozo and N. Khiar, *Nanoscale*, 2015, **7**, 19259-19272.
42. J. E. Gestwicki, C. W. Cairo, L. E. Strong, K. A. Oetjen and L. L. Kiessling, *J. Am. Chem. Soc.*, 2002, **124**, 14922-14933.
43. A. Fu, W. Gu, B. Boussert, K. Koski, D. Gerion, L. Manna, M. Le Gros, C. A. Larabell and A. P. Alivisatos, *Nano Lett*, 2007, **7**, 179-182.
44. A. E. Albers, E. M. Chan, P. M. McBride, C. M. Ajo-Franklin, B. E. Cohen and B. A. Helms, *J Am Chem Soc*, 2012, **134**, 9565-9568.
45. K. Wu, L. J. Hill, J. Chen, J. R. McBride, N. G. Pavlopolous, N. E. Richey, J. Pyun and T. Lian, *ACS Nano*, 2015, **9**, 4591-4599.
46. R. Alam, D. M. Fontaine, B. R. Branchini and M. M. Maye, *Nano Lett*, 2012, **12**, 3251-3256.
47. R. Alam, L. M. Karam, T. L. Doane, K. Coopersmith, D. M. Fontaine, B. R. Branchini and M. M. Maye, *ACS Nano*, 2016, **10**, 1969-1977.
48. W. Wang, Y. Guo, C. Tiede, S. Chen, M. Kopytynski, Y. Kong, A. Kulak, D. Tomlinson, R. Chen, M. McPherson and D. Zhou, *ACS Appl. Mater. Interfaces*, 2017, **9**, 15232-15244.
49. W. Wang, Y. Kong, J. Jiang, X. Tian, S. Li, U. S. Akshath, C. Tiede, N. Hondow, A. Yu, Y. Guo and D. Zhou, *Nanoscale*, 2020, **12**, 8647-8655.
50. D. Zhou, A. Bruckbauer, C. Abell, D. Klenerman and D.-J. Kang, *Adv. Mater.*, 2005, **17**, 1243-1248.
51. Zhou, A. Bruckbauer, Ying, C. Abell and D. Klenerman, *Nano Lett.*, 2003, **3**, 1517-1520.
52. J. Hooper, Y. Liu, D. Budhadev, D. F. Ainaga, N. Hondow, D. Zhou and Y. Guo, *ACS Appl. Mater. Interfaces*, 2022, **14**, 47385–47396.
53. A. Holla and A. Skerra, *Protein Eng. Des. Sel.*, 2011, **24**, 659-669.
54. T. Förster, *Ann. Phys.*, 1948, **437**, 55-75.
55. N. Hondow, R. Brydson, P. Wang, M. D. Holton, M. R. Brown, P. Rees, H. D. Summers and A. Brown, *J. Nanopart. Res.*, 2012, **14**, 977.
56. H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327-1335.
57. G. Tabarani, M. Thépaut, D. Stroebel, C. Ebel, C. Vivès, P. Vachette, D. Durand and F. Fieschi, *J. Biol. Chem.*, 2009, **284**, 21229-21240.

Chapter 5
Investigating the binding of glycan-gold nanoparticles upon DC-SIGN/R conjugated supported lipid bilayers using QCM-D and cryoTEM

All work in the following chapter was conducted by the author of this thesis. Cryo-EM imaging was facilitated by Dario Fernandez Ainaga and Dr Nicole Hondow. Ligand precursors were synthesised by Dr Darshita Budhadev. Plasmids were co-prepared by Dr Yuan Guo.

5.1 Introduction

So far, this thesis has demonstrated the valuable contribution of glyconanoparticles, such as glycan-coated quantum dots (glycan-QDs), as unique tools for probing the solution-phase affinities, thermodynamics, kinetics and binding modes of multivalent lectin-glycan interactions (MLGIs). Solution-based methods offer a range of advantages, such as the accurate, non-sterically hindered quantification of the binding of individual proteins with easily accessible glycan surfaces, however, such techniques cannot truly represent the surface-phase chemistry observed between glycans and membrane lectins in their native states upon cell or virus surfaces.

Such distinctions were demonstrated by our group's recent study on the use of glycan-coated gold nanoparticles (glycan-GNPs) as potential therapeutics for targeting tetrameric lectins DC-SIGN and DC-SIGNR.[1] In that study, a range of glycans, including mannose (Man) and mannose-α1,2-mannose (DiMan), coupled to monovalent and trivalent ligands, were conjugated onto 5 nm GNPs (G5) and were investigated for their solution-phase binding behaviour and *in vitro* viral inhibition efficacy. Lectin binding was quantified by monitoring the quenching of fluorescence from an excited Atto643 labelled protein to any bound glycan-GNPs *via* a non-radiative energy transfer process (see **Section 1.3.3; Figure 1.11**). This technique achieved comparable K_{d} values with those obtained for QD-EG_{11}-DiMan, where binding between G5-EG_2-DiMan with DC-SIGN/R achieved values of 3.9 and 152 nM, respectively (*vs* 3.0 and 80 nM for QD-EG_{11}-DiMan at 25 ^{o}C; see **Table 3.1**). These studies also revealed that increasing the density of Man or DiMan using trivalent ligands had much more of a strengthening effect on DC-SIGNR than it did DC-SIGN. Interestingly, however, when these glycan-GNPs were applied to the inhibition of ebolavirus glycoprotein (EBOV-GP)-driven infection of DC-SIGN/R^{+} 293T cells, the concentrations required to

achieve 50% inhibition (IC_{50}) were conversely observed to increase (*i.e.* lower in potency) with increased glycan density for both proteins. In addition, though DiMan generally led to stronger binding, the difference in potency of binding between the DiMan-bearing GNPs and the Man-bearing GNPs for DC-SIGNR was much more pronounced in the cell surface-phase inhibition studies compared to the solution-phase quenching studies (*i.e.* $IC_{50}^{Man}/IC_{50}^{DiMan}$ = 12-20-fold *vs* K_d^{Man}/K_d^{DiMan} = 2-3-fold), however, this difference was a lot less pronounced for DC-SIGN (*i.e.* $IC_{50}^{Man}/IC_{50}^{DiMan}$ = 3-10-fold *vs* K_d^{Man}/K_d^{DiMan} = 5-8-fold).

Table 5.1 Summary of parameters for the binding of DC-SIGN/R with different glycan-GNPs, derived from the quenching efficiency, $E = [P]^n/(K_d{}^n + [P]^n)$; and the inhibition of EBOV-GP into DC-SIGN/R$^+$ 293T cells with different glycan-GNPs, derived from the luciferase activity, $A = IC_{50}{}^m/(IC_{50}{}^m + [G]^m)$. Where [L] is the concentration of lectin, K_d is the dissociation constant, n is the Hill coefficient, IC_{50} is the concentration of inhibitor required to achieve 50% luciferase activity, m is the inhibition coefficient and [G] is the concentration of glycan-GNP inhibitor.

	Glycan-GNP	K_d / nM	n	IC_{50} / nM	m
DC-SIGN	G5-Man	33.1 ± 2.1	0.82 ± 0.07	0.26 ± 0.08	1
	G5-$(Man)_3$	18.7 ± 0.3	0.77 ± 0.01	1.57 ± 0.25	1
	G5-DiMan	3.9 ± 0.3	0.59 ± 0.03	0.095 ± 0.017	1
	G5-$(DiMan)_3$	3.6 ± 0.1	0.42 ± 0.01	0.15 ± 0.03	1
DC-SIGNR	G5-Man	214 ± 68	0.68 ± 0.10	7.3 ± 1.2	0.42 ± 0.04
	G5-$(Man)_3$	133 ± 20	0.63 ± 0.07	53 ± 17	0.61 ± 0.14
	G5-DiMan	152 ± 37	0.99 ± 0.14	0.63 ± 0.09	0.56 ± 0.02
	G5-$(DiMan)_3$	42 ± 2	0.48 ± 0.03	2.6 ± 0.1	0.53 ± 0.02

Discrepancies such as these demonstrate the importance of understanding MLGIs not only within the solution-phase but also within their native states upon cell surfaces. Such regimes not only take into account the true kinetic and entropic contributions associated with binding between restricted, laterally mobile protein with freely diffusing glycans, but also incorporate the bigger picture of how proteins might act cooperatively upon cell surfaces to better facilitate glycan binding and internalisation. For example, the formation of clustered proteins on cell membranes, known as microdomains, are understood to promote the binding of extracellular species through the additional multivalent enhancement afforded by multiple receptors interacting with multiple glycans on the same species. Techniques such as electron

microscopy, near-field scanning optical microscopy and confocal fluorescence microscopy have been used in the past to identify DC-SIGN microdomains on cell surfaces, demonstrating clusters of ~80 nm in diameter on naïve dendritic cells which organise to ~200 nm upon binding and internalisation of pathogenic species.[2-5] Such microdomains have been shown to display rapid lateral transportation from the cell's leading edge to zones of potential internalisation upon contact with pathogenic material, demonstrating the importance of such interactions in internalisation.[6] Microdomain stability has been shown to be largely dependent on the presence of DC-SIGN's CRDs,[3] which suggests that they are stabilised by either the extracellular matrix or transmembrane glycans rather than from the cytoplasm, however, microdomain formation may likely play an important role in other cytoplasmic activities such as lectin-associated signalling pathways. These insights demonstrate the importance of understanding the interaction of glycans with lectins such as DC-SIGN an DC-SIGNR in their native membrane-phase environments. However, despite these studies, little has been done to investigate the actual mechanistic modes and affinities of the MLGIs upon membrane surfaces and how differences in binding mechanism may affect clustering and any subsequent biological function. Moreover, other than for inhibition studies and probing microdomains,[2] few studies have been performed to analyse the effect of glyconanoparticles and their design on MLGIs on lipid membranes. Revealing such details would therefore better inform our understanding of MLGIs in their native states and what considerations need to be taken into account in designing glycoconjugates to target such interactions.

So far, surface plasmon resonance (SPR) has been the main technique for measuring the surface-phase kinetics of lectin-glycan interactions,[7-12] however, simply binding glycans or lectins onto a solid support (*i.e.* gold) or immobilising them *via* static interactions, has its limitations. Such techniques can often lead to lectins being positioned in undesirable orientations where the binding sites are not fully accessible.[13] In addition, full immobilisation does not allow for lateral fluidity and thus cannot accurately represent the true dynamic nature of cell membranes. Quartz crystal microbalance (QCM) techniques offer an adequate alternative to SPR, where, instead of providing an optical readout, QCM exploits quartz's piezoelectric readout to measure its mass-dependent acoustic resonance. Here, shifts in the resonance frequency are proportional to the amount additional areal mass bound to the surface (**Equation 5.1**), which is in turn proportional to the amount of bound analyte, thus allowing binding rates and affinities to be derived.[14] Though QCM can

also be used to monitor immobilised binding partners on gold surfaces,[15, 16] because of the natural hydrophilicity of silicon oxide-based surfaces, QCM sensors can also support lipid bilayers,[17] allowing them to monitor interactions upon lipid surfaces which more closely mimic the properties of cell membranes compared to solid surfaces. Additionally, by incorporating lipids that can be functionalised with specific proportions of proteins or ligands, these supported lipid bilayers (SLBs) can be tuned in a simple, controllable manner. This technique thus provides an desirable platform for the detailed study of protein-carbohydrate interactions upon lipid bilayer surfaces without the high complexity of real cell membranes. Additionally, QCM-D (QCM with dissipation monitoring) can also monitor the dissipation of energy from the quartz surface *via* the decay in the amplitude of piezoelectric oscillation, allowing the sensor to observe changes in the rigidity or flexibility of the surface which can provide further information about the changing character of the system upon addition of an analyte.

Glycan-GNPs provide an excellent example glycoconjguate for studying MLGIs upon SLB surfaces. Firstly, the previously performed solution-phase and cell surface studies offer an excellent comparative lens to which SLB assays can add key information to bridge these two environments. Additionally, GNPs provide a much more practical scaffold compared to the QDs discussed in previous chapters. GNPs are not only much simpler to synthesise and functionalise, but they have also been shown to display excellent biocompatibility (in comparison to more cytotoxic metals, such as cadmium), permitting them to be much more readily applied to nanomedicinal applications.[18]

In this study, G5-Man has been employed to investigate the binding of glyconanoparticles with DC-SIGN and DC-SIGNR upon lipid bilayer surfaces. Here, small unilamellar vesicles (SUVs) containing varying concentrations of tris-NTA functionalised lipids were synthesised and converted into SLBs. SLBs were then further functionalised with DC-SIGN and DC-SIGNR by tethering N-terminal oligo-histidine tagged variants of the proteins to tris-NTA *via* Ni^{2+} coordination. QCM-D was then employed to show that DC-SIGN/R functionalised SLBs of various protein densities were successfully formed, stable and were relatively representative of DC-SIGN/R's native surface presentation. Furthermore, DC-SIGN/R functionalised SLBs were able to form strong and specific interactions with G5-Man glycoconjugates which could be completely displaced under high concentrations of mannose. It was additionally observed that, where the affinity of DC-SIGN binding with G5-Man

displayed limited dependency upon the lectin surface density, DC-SIGNR's binding affinity was found to increase significantly (become weaker) with increasing inter-lectin spacing. This indicated G5-Man binding with DC-SIGNR is much more reliant on the interaction of multiple protein molecules. This demonstrates that, for small (4-5 nm) GNP scaffolds, the crosslinking and simultaneous binding mechanisms observed in the solution-phase assays of DC-SIGN/R actually closely resemble those at the membrane surface. Additionally, this difference in density dependency between DC-SIGN and DC-SIGNR affinities may also provide some insight into the observed differences between the solution and cell based assays. In addition, a preliminary effort was made to image such interactions using cryo-TEM and confocal microscopy, which were able to visually characterise the crosslinking of DC-SIGNR by G5-Man on SLB surfaces. Though these techniques are still in their infancy, the development of these strategies for analysing membrane-phase MLGIs shows some promising results for prospective studies in the future.

5.2 Experimental

5.2.1 G5-Ct

GNPs were synthesised *via* a previously established approach.[19] All glassware was washed with aqua regia and oven dried prior to GNP synthesis. Solutions of tannic acid (0.1 mL, 2.5 mM in H_2O, 0.25 mmol) and K_2CO_3 (1 mL, 150 mM in H_2O, 0.15 mmol) were added to a solution of sodium citrate (97 mg, 0.33 mmol) in H_2O (150 mL), and heated at 75 ^oC for 30 min with vigorous stirring. A solution of gold(III) chloride trihydrate (1 mL, 25 mM in H_2O, 0.025 mmol) was then added and the mixture was then heated at 75 ^oC for 1 hr with slow stirring, before cooling to r.t.. Concentrations were obtained by UV-vis spectroscopy (ε_{G5} = 1.1×10^7 M^{-1} cm^{-1}; **Equation 2.1**) and the man hydrodynamic diameter was obtained by DLS.

Citrate capped gold nanoparticles (G5-Ct): Yield: 5.7 nmol. UV-vis: λ_{abs} 509.0 nm. Mean hydrodynamic diameter (D_h) 8.7 ± 2.0 nm.

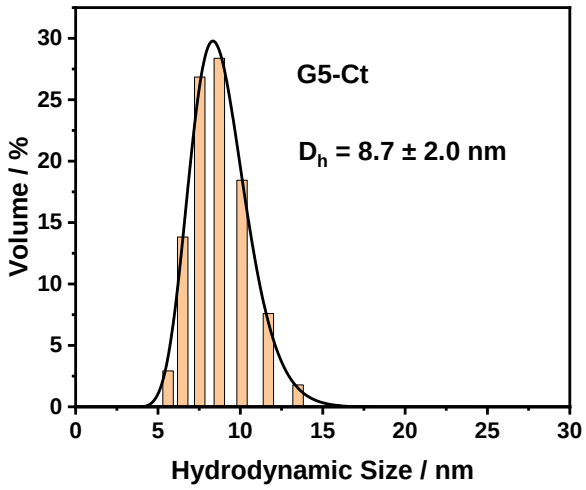

Figure 5.1 Hydrodynamic size distribution of G5-Ct, obtained by DLS. Fitted with a lognormal distribution curve (D_h = mean ± ½FWHM; R^2 = 0.998).

5.2.2 G5-Man

Cap exchange was performed *via* previously established methods. LA-EG$_2$-Man was prepared in-house by our group as described previously.[1]

A solution of LA-EG$_2$-Man (200 µL, 10 mM in H_2O, 2.0 µmol) was added to a solution of G5-Ct (1.33 mL, 1.51 µM, 2.0 nmol, ligand:GNP molar ratio 1000:1) and rotated at r.t. for 36 hr, as described previously.[1] The resulting mixture was then washed with 3 times with 3 mL H_2O using a using a 10 kDa MWCO spin filter at 4000 rpm for 15 mins. Concentrations were obtained by UV-vis spectroscopy, mean hydrodynamic diameters were obtained by DLS and the number of ligands per GNP obtained by sulphur-phenol carbohydrate quantification to be 750 ± 50 (where the extinction coefficient obtained for the sulphur-phenol quantification of LA-EG$_2$-Man was obtained previously to be ε_{490} = 8600 ± 200 M^{-1} cm^{-1}).[1]

LA-EG$_2$-Man capped GNPs (G5-Man): Yield: 1.96 nmol, 98%. UV-vis: λ_{abs} 517 nm. CryoTEM: mean core diameter (D_{core}) 4.9 ± 1.4 nm. DLS: D_{h} 10.4 ± 2.8 nm.

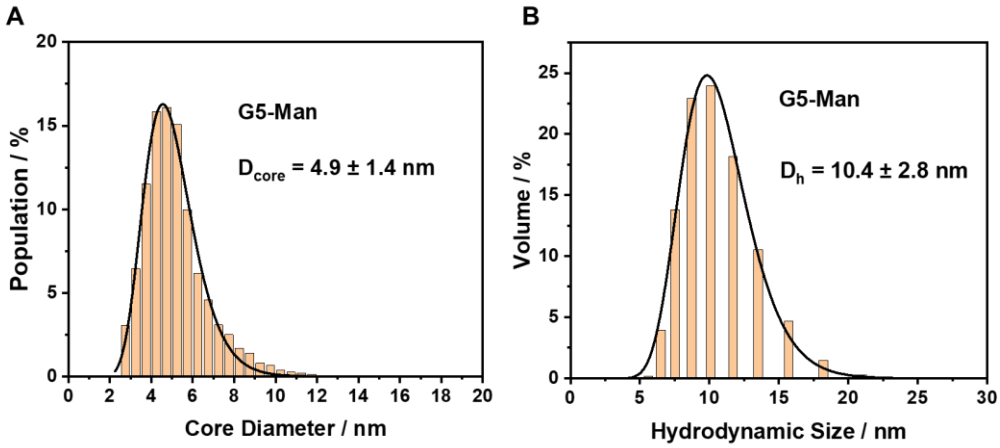

Figure 5.2 (A) Core diameter distribution of G5-Man, obtained by cryoTEM; (B) hydrodynamic size distribution of G5-Man, obtained by DLS. Both distributions fitted with lognormal distribution curves (where D_{core} or D_h = mean ± ½FWHM; $R^2 \geq 0.99$).

5.2.3 Protein Production and Non-Specific Labelling

His_{24}-DC-SIGN and His_{24}-DC-SIGNR were made using the protein production and purification protocol described in **Section 2.1.4.1**, using plasmid encoding for the monomeric His_6-DC-SIGN or His_6-DC-SIGNR, respectively (made in-house by Dr Yuan Guo, as confirmed by DNA sequencing).

To non-specifically label the protein for confocal microscopy, His_{24}-DC-SIGNR was washed 3 times with a buffer of HEPES (20 mM, pH 8.3), NaCl (150 mM) and $CaCl_2$ (100 mM) (herein referred to as non-specific labelling buffer; NSL buffer) using a 10 kDa MWCO spin filter. His_{24}-DC-SIGNR (265 µL, 20.8 µM in NSL buffer, 5.53 nmol) was reacted with NHS-ester-Atto643 (7.4 µL, 10.5 mM in DMSO, 77.3 nmol) by immediate vortexing, followed by rotating at r.t. for 2 hrs. The protein was then purified *via* a mannose-sepharose affinity column which was washed with a buffer of HEPES (20 mM, pH 7.8), NaCl (100 mM) and $CaCl_2$ (10 mM) (herein referred to as binding buffer), to remove any excess dye and denatured protein, before eluting with a buffer of HEPES (20 mM, pH 7.8), NaCl (100 mM) and EDTA (2.5 mM). A solution of $CaCl_2$ (500 mM, pH 8, in H_2O) was then added to a final Ca^{2+} concentration of 10 mM. This yielded non-specifically labelled His_{24}-DC-SIGNR (NSL-DC-SIGNR; 12.2 nmol, 55%).

His_{24}-DC-SIGN: HRMS: calculated monomeric molecular weight (M_w) based on amino acid sequence [protein] 37034.72; found 37027.52. DLS: D_h 12.7 ± 2.0 nm.

His_{24}-DC-SIGNR: HRMS: calculated monomeric M_w based on amino acid sequence [protein] 35343.57; found 35336.14. DLS: D_h 10.2 ± 1.7 nm.

NSL-DC-SIGNR: HRMS: calculated monomeric M_w based on amino acid sequence [protein] 35343.57, [protein+Atto643] 36161.48 and [protein+2Atto643] 36979.36; found 35334.45, 36152.69 and 36970.42; proportion of 1^o labelled protein 69%, proportion of 2^o labelled protein 15%.

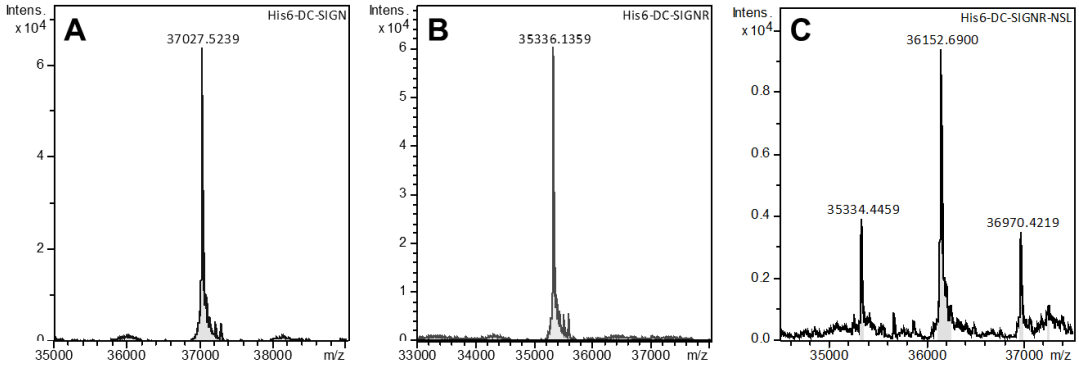

Figure 5.3 Deconvoluted HRMS of (A) His_6-DC-SIGN monomer; (B) His_6-DC-SIGNR monomer; (C) NSL-DC-SIGNR monomer.

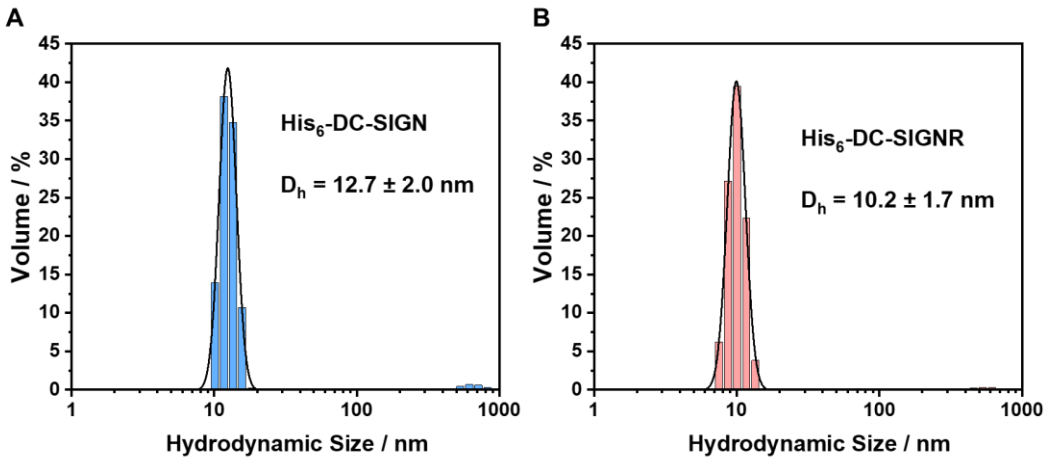

Figure 5.4 Hydrodynamic size distributions obtained by DLS for (A) His_{24}-DC-SIGN and (B) His_{24}-DC-SIGNR.

5.2.4 DOPC and tris-NTA DODA SUVs

SUVs were prepared using a previously established approach.[20] 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-phosphocholine (DOPC; 36 mM in $CHCl_3$) and tris-NTA(-O^{t}Bu)-succ-dioctadecylamine (tris-NTA DODA; 1.56 mM in $CHCl_3$) were combined in molar equivalents of 1:0, 499:1, 249:1, 99:1, 49:1 to a final amount of 4.5 µmol. The lipid mixtures were then dried under N_2 followed by

desiccation for 2 hr, before being dissolved in a buffer of HEPES (10 mM, pH 7.4) and NaCl (150 mM) (herein referred to as SUV buffer) to a final concentration of 2 mg mL^{-1}. The resulting vesicles were reduced in size by 5 cycles of vortexing, freezing in dry ice and thawing in warm water. Each solution was then sonicated (power: 70%; cycle: 1 s on, 1 s off) on ice for 15 min, before being centrifuged for 10 min. The supernatant was then removed and stored under N_2 at 4 ^oC. This yielded small unilaminar vesicles (SUVs) of DOPC, containing 0, 0.2, 0.4, 1 and 2% tris-NTA DODA.

5.2.5 QCM-D

All QCM-D experiments were studied on silica-coated sensors (QSX303; Biolin Scientific, Västra Frölunda, Sweden) using a Q-Sense E4 system equipped with four independent flow modules, connected to a syringe pump (Legato; World Precision Instruments, Stevenage, UK). All measurements were monitored *via* the frequency and dissipation shifts (Δf_i and ΔD_i, respectively) at overtones of i = 3, 5, 7, 9, 11 and 13 (fundamental resonance frequency, f_1 = 5 MHz), which showed comparable data. All buffers used were degassed before each experiment to avoid the introduction of bubbles during measurement. All chambers and tubing were washed with BSA (10 mg mL^{-1}) at a flowrate of 50 µL min^{-1} for 10 min, to passivate the apparatus walls to reduce non-specific adsorption, before being washed with H_2O at a flowrate of 100 µL min^{-1} for 20 min and then dried in air and N_2. The sensor chips were incubated in sodium dodecyl sulphate (SDS) solution (2% w/v in H_2O) for 1 hr before being washed with water, dried under N_2 and then cleaned under UV/ozone for 30 min (immediately before use). Sensor chips were then monitored in air at 24 ^oC for their stability in frequency and dissipation at each overtone, before being equilibrated in SUV buffer for at least 10 minutes, ensuring a 3rd overtone ΔD_3 of 162×10^{-6}, typical of this buffer. All proceeding measurements were monitored using the normalised frequency and dissipation at overtone i = 5 (*i.e.* $\Delta f = \Delta f_5/5$ and $\Delta D = \Delta D_5/5$). Typical measurements were performed at a flowrate of 20 µL min^{-1}, unless otherwise stated. For SLB formation, SUVs (2 mg mL^{-1} in SUV buffer) were diluted to 50 µg mL^{-1} in SUV buffer with 5 mM $NiCl_2$.

5.2.6 Confocal Microscopy

Confocal microscopy samples were prepared using a 4× 200 µL Teflon well plate which had been mounted onto a borosilicate glass slide (Menzel-Glazer;

24 × 24, 1.5 mm) *via* Twinsil® silicone glue. Samples were prepared by adding SUV (25 µL, 0.2 mg mL^{-1} in SUV buffer with 20 mM $NiCl_2$) to SUV buffer (25 µL) and incubating, covered, at r.t. for 15 min, before washing with SUV buffer (10× 100 µL) followed by binding buffer (10× 100 µL). Half of the solution was then replaced with protein (25 µL, 2.24 µM in binding buffer) and was incubated, covered, at r.t. for 60 min, before washing with binding buffer (10× 100 µL).

Confocal microscopy was performed using a Zeiss LSM880 with a Airyscan Inverted Confocal Microscope and samples were excited using a HeNe 633 nm laser. Once images of the SLB-lectin samples were collected, half of solution was replaced with G5-Man (25 µL, 800 nM in binding buffer) and images were taken at various time points during incubation.

5.2.7 CryoTEM

CryoTEM samples were prepared using the confined droplet method. Here a 15 µL droplet of SUV buffer was suspended on a O_2/Ar plasma cleaned silicon nitride coated 3×3 window cryoTEM grid. Half of the droplet was replaced with SUV (7.5 µL, 0.2 mg mL^{-1} in SUV buffer with 20 mM $NiCl_2$) and incubated, covered, at r.t. for 15 min, before removing half of the solution and washing with SUV buffer (10× 7.5 µL) followed by binding buffer (10× 7.5 µL). Protein (7.5 µL, 2.24 µM in binding buffer) was then added and incubated, covered, at r.t. for 60 min, before removing half of the solution and washing with binding buffer (10× 7.5 µL). To this, G5-Man (7.5 µL, 800 nM in binding buffer) was then added and incubated, covered, at r.t. for 60 min, before removing half of the solution and washing with binding buffer (10× 7.5 µL).

CryoTEM was performed by Dario Fernandez Ainaga and Nicole Hondow at the University of Leeds. The grids were then immediately blotted and plunge-frozen into liquid ethane using an FEI Vitrobot mark IV plunge freezer. A Gatan 914 cryo-holder was used for TEM and the temperature was maintained below -160 ^oC to prevent devitrification. Transmission electron microscopy (TEM) was conducted on an FEI Titan^3 Themis G2, operating at 300 kV, fitted with 4 EDX silicon drift detectors, multiple STEM detectors and a Gatan One-View CCD, using probe currents of 20-40 pA. Cryo-EDX spectroscopy was performed using probe currents between 40-60 pA and a dwell time of 23 µs were used.

5.3 Results and Discussion

5.3.1 Materials Synthesis and Characterisation

GNPs of ~5 nm in core diameter (G5) were used to replicate those previously employed in solution phase assays in order to provide a comparative lens to the study.[1] These nanoparticles were synthesised using gold(III) tetrachloride sodium citrate *via* a modified Turkevich method using competing sodium citrate and tannic acid, developed for the production of relatively monodispersed sub-10 nm GNPs.[19] This yielding citrate surface-stabilised G5 (G5-Ct) with a mean core diameter of 4.9 ± 1.4 nm and mean hydrodynamic diameter (D_h) of 8.7 ± 1.9 nm. These were functionalised with Man by substituting the weakly adsorbed citrate ligands with lipoic acid-di(ethylene glycol)-mannose (LA-EG_2-Man). Here, ligands followed a "anchor-linker-glycan" design, where the lipoic acid (LA) anchor unit allowed the formation of highly stable chelating S-Au interactions without the need for reduction of the disulphide bond; di(ethylene glycol) (EG_2) was used as the linker domain to provide additional hydrophilicity to ligand, adequately stabilising the G5 gold core, despite the low number of repeat EG units; and a terminal Man glycan unit was used due to well-known strong specific association with DC-SIGN and DC-SIGNR. In order to closely replicate the scaffolds used in the previous solution-phase studies, LA-EG_2-Man ligands were assembled using alkyne-azide copper-catalysed "click" chemistry between LA-EG_2-CCH and N_3-EG_2-Man precursors. Conjugation of G5-Ct with LA-EG_2-Man yielded G5-Man (**Figure 5.5**) which were characterised by UV-vis spectroscopy, dynamic light scattering (DLS) and sulphur-phenol carbohydrate quantification to provide a maximum adsorption wavelength of 517 nm, a D_h of 10.4 ± 2.8 nm and 750 ± 50 ligands per GNP.

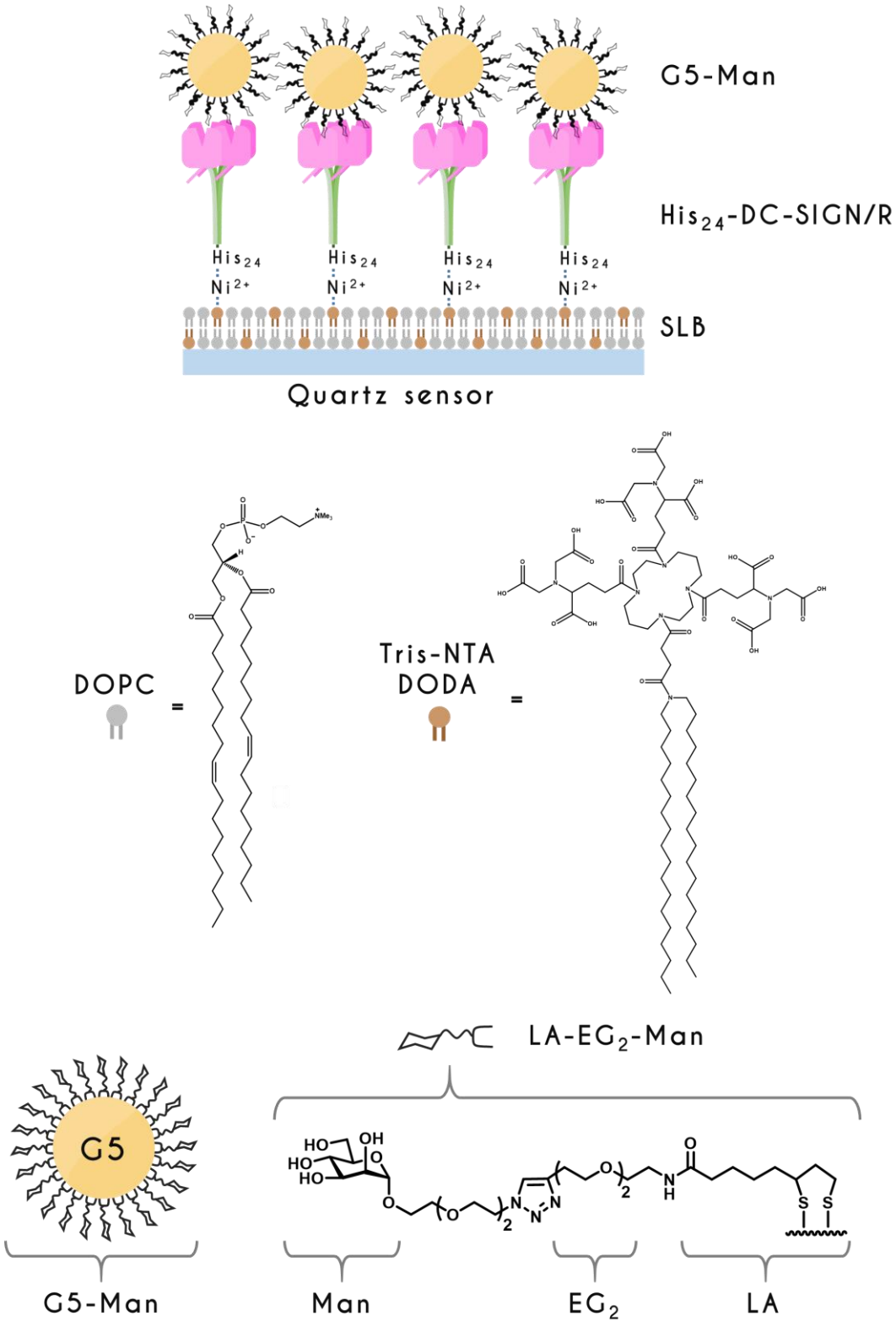

Figure 5.5 Schematic depicting LA-EG$_2$-Man capped G5 (G5-Man) bound to His$_6$-tagged DC-SIGN/R (His$_{24}$-DC-SIGN/R)-tethered *via* Ni^{2+} coordination to a tris-NTA DODA-containing DOPC supported lipid bilayer (SLB) on a quartz sensor support.

Plasmids encoding for N-terminal hexahistidine-tagged variants of DC-SIGN and DC-SIGNR were constructed *via* standard molecular biology techniques

from the sequences corresponding to a slightly N-terminal truncated wild-type extracellular domain of each protein (see **Appendix A.1** for amino acid sequence), and were confirmed by DNA sequencing. Proteins were shown to retain good mannose-specificity during purification with mannose-sepharose affinity column chromatography. The resulting His_{24}-DC-SIGN and His_{24}-DC-SIGNR tetramers were characterised by UV-vis spectroscopy and DLS which provided D_h values of 12.7 ± 2.0 nm and 10.2 ± 1.7 nm, respectively.

5.3.2 SLB-lectin formation

In order to replicate DC-SIGN and DC-SIGNR's native surface-phase presentation, supported lipid bilayers (SLBs) bearing these lectins were employed to mimic the laterally mobile lipid surface associated with their natural cell membrane environment. However, because imbedding full proteins with long cytoplasmic domains into the SLBs could disrupt the stability of the lipid bilayers, a protein immobilisation strategy based on coordination of oligohistidine-tagged protein extracellular domains to functional Ni^{2+}-bearing tris-NTA DODA lipids was employed (**Figure 5.5**). SLBs are prepared by the adsorption and subsequent rupturing of small unilamellar vesicles (SUVs) onto a quartz support.[17] Therefore, DOPC SUVs containing tris-NTA DODA lipids in proportions of 0, 0.2, 0.4, 1 and 2% were prepared with Ni^{2+} containing buffer, to allow for the tethering of histidine-tagged protein *via* Ni^{2+} coordination. The quality of SLB formation was confirmed by monitoring the resonance frequency and energy dissipation *via* QCM-D during the addition of SUVs in SUV buffer with 5 mM $NiCl_2$ onto a quartz sensor chip, at a flowrate of 50 µg mL^{-1} (**Figure 5.6**). Here, a shift in frequency (Δf) and dissipation (ΔD) of ~-40 Hz and 2-3×10^{-6} was observed, respectively, indicating the adsorption of SUVs onto the quartz surface up until the critical vesicular coverage.[17] At this point the Δf then returned to ~-25 Hz and the ΔD returned roughly back to its baseline, signifying the rupturing of adsorbed SUVs and the formation of the SLB. The SLBs were stable under further washing with buffer, displaying minimal change in frequency (*i.e.* <1 Hz).

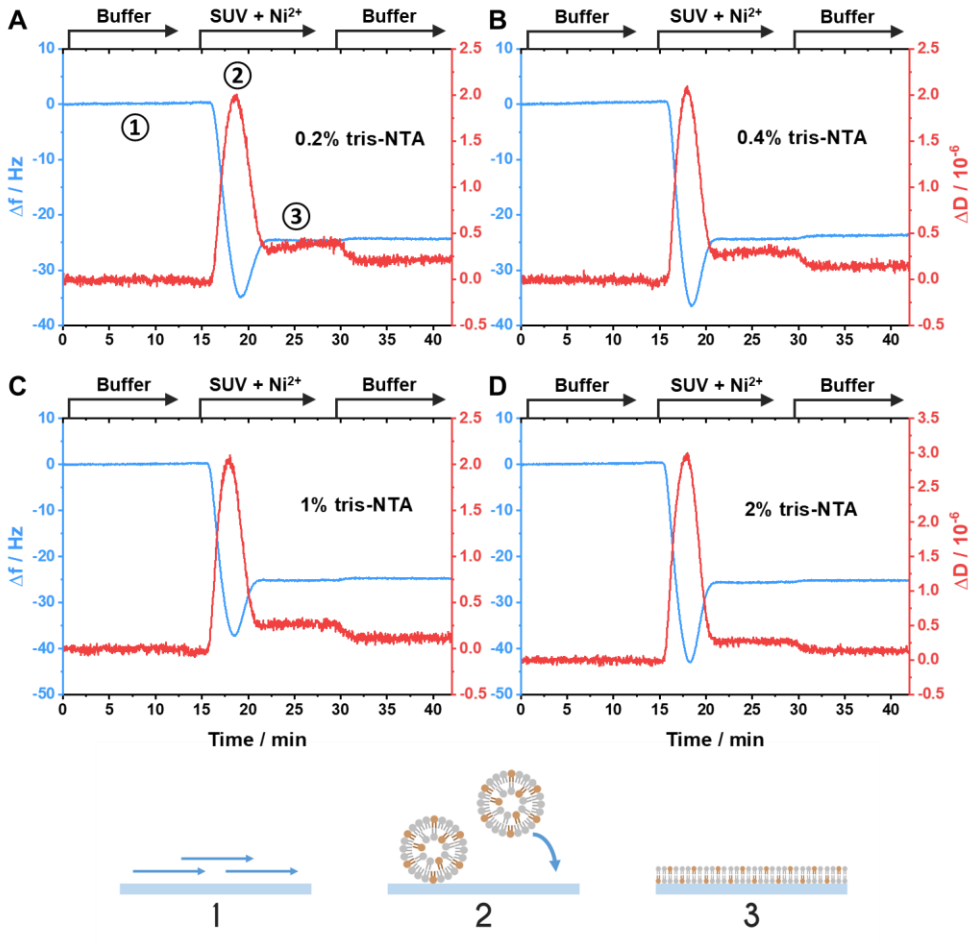

Figure 5.6 QCM-D measurements of the Δf and ΔD time profiles for the formation of SLBs from DOPC SUVs containing tris-NTA DODA in proportions of (A) 0.2%, (B) 0.4%, (C) 1% or (D) 2%. 1. Demonstrates the equilibration of the sensor under buffer; 2. demonstrates a typical frequency and dissipation shift with absorption of the SUVs to the quartz surface; and 3. Demonstrates a slight recovery of frequency and dissipation as SUVs rupture to form the SLB surface. Schematics of these stages depicted below. (Buffer: SUV buffer).

To test the coordination His_{24}-DC-SIGN/R to the tris-NTA-Ni^{2+} bearing SLB surface, SLBs containing 0% and 2% tris-NTA DODA were assembled as described above. In order to ensure the stability of DC-SIGN/R and to replicate the conditions employed in the solution-phase assays, the SLBs were then equilibrated in binding buffer, resulting a small Δf of ~-1 Hz and ΔD of ~0.5×10^{-6}. From here, the SLBs were incubated with 280 nM His_{24}-DC-SIGN/R in binding buffer and the Δf and ΔD was monitored over time (**Figure 5.7**). Both proteins demonstrated no observed binding for 0% tris-NTA (*i.e.* no change of Δf), confirming that protein cannot associate without the presence of tris-NTA. SLBs containing 2% tris-NTA displayed significant binding, saturating at a Δf of ~-100 Hz. Assuming complete surface coverage of protein, the thickness, Δh, of the protein layer can be estimated from the

Sauerbrey equation to be ~13 nm (**Equation 5.1**; where C is the sensors mass sensitivity constant, which is assumed to be 18 ng cm^{-2} Hz^{-1},[1] and ρ is the protein film density, which is assumed to be 1.41 g cm^{-1} [2]).[21, 22]

$$\Delta f = -\frac{\Delta m}{C} = -\frac{\rho}{C}\Delta h \qquad 5.1$$

This correlates nicely with the hydrodynamic sizes measured by DLS (His$_{24}$-DC-SIGN: ~13 nm; His$_{24}$-DC-SIGNR: ~10 nm), indicating that proteins were orientated in an extended structure, protruding outwards from the SLB surface, which matches the predicted native state of these proteins on cell membranes.[23] In addition, an increase in ΔD of ~8-9×10^{-6} is observed with protein binding, indicating an increase in flexibility of the surface. This increase in dissipation is a common property in the formation of protein layers that have been tethered to an SLB surface by flexible linkers such as oligo-histidine, rather than fixed to a solid support.[24] Furthermore, the coiled-coil neck of each tetramer is much smaller than the head group containing the four CRDs, hence packing will be less dense at the neck region (**Figure 5.5**), thus permitting such flexibility at the sensor surface.

These proteins were also able to be completely dissociated under the addition of 500 mM imidazole, indicated by the return of Δf to the baseline. Here, imidazole is able to compete with the His-tag proteins chelating with the Ni^{2+} ions. This result indicates that the binding of protein to the SLB is specifically facilitated by the binding of the histidine-tag at the N-terminal of the proteins to the tris-NTA functionalised lipids *via* Ni^{2+} coordination.

---

[1] C depends largely on the fundamental resonance frequency and material properties of the quartz sensor.

[2] ρ can be estimated from $\rho = 1.41 + 0.145e^{\frac{-M_W}{13000}}$, where the molecular weight M_w of His$_{24}$-DC-SIGN and His$_{24}$-DC-SIGNR is 141 and 148 kDa.

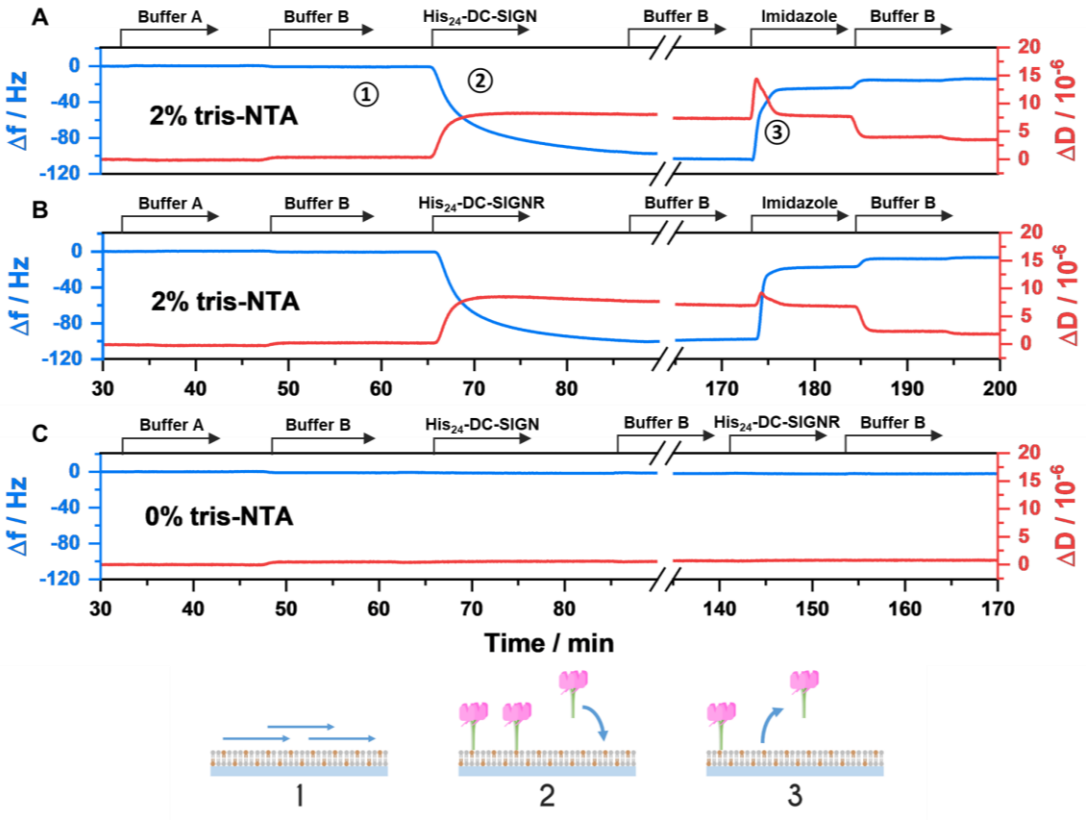

Figure 5.7 QCM-D measurements of the Δf and ΔD time profiles for the association and imidazole-driven dissociation of (A) His_{24}-DC-SIGN and (B) His_{24}-DC-SIGNR, onto an SLB bearing 2% tris-NTA; and (C) the interaction of His_{24}-DC-SIGN and His_{24}-DC-SIGNR with SLBs bearing no tris-NTA. 1. Demonstrates the equilibration of the SLB in binding buffer; 2. demonstrates the binding of protein over time; and 3. demonstrates the release of protein *via* displacement of Ni^{2+} by imidazole. (Buffer A: SUV buffer; Buffer B: binding buffer).

In order to investigate the effect of protein density on binding DC-SIGN/R were then tethered to SLBs with 0, 0.2, 0.4, 1 and 2% tris-NTA DODA. The relationship between the tris-NTA density and the fraction of protein to the SLB was monitored by measuring the Δf. Here, a plot of the equilibrium $|\Delta f|$ against tris-NTA density demonstrated a linear increase with a slope of ~85 Hz per percentage of tris-NTA DODA to DOPC (**Figure 5.8; Figure C.11**). Interestingly, the equilibrium $|\Delta f|$ values at 2% tris-NTA deviated negatively from this linearity. Assuming that the SLB surface has a single DOPC lipid cross-sectional area of ~0.73 nm^2,[25] the distance between tris-NTA molecules at a densities of 0.2, 0.4, 1, 2% can be estimated at ~21, 15, 10 and 7 nm, respectively. For 2% tris-NTA, this distance becomes comparable to the sizes of the DC-SIGN and DC-SIGNR tetramer head groups (where the furthest inter-binding site distances have been estimated at ~6 nm and ~9 nm for DC-SIGN and DC-SIGNR, respectively, from models of the XRC structures of the monomeric or dimeric CRDs of these proteins).[23, 26] Therefore, the $|\Delta f|$ at 2%

tris-NTA may be lower than that expected by the linear trend because the proteins may not be able to completely occupy every tris-NTA molecule at the SLB surface due to steric hindrance once protein has already saturated the surface. This steric hindrance is much less likely to occur at lower tris-NTA densities, thus theoretically allowing the proteins to bind to every tris-NTA molecule. In addition, because species are closer together at higher densities, the amount of the hydrodynamically coupled mass between species also increases. This therefore causes a release of water back into solution, decreasing the amount of mass acting on the surface with increasing protein density, thus further lowering the $|\Delta f|$.[27] (Herein, despite these discrepancies, SLBs bearing DC-SIGN/R with x% tris-NTA to DOPC also contain x% DC-SIGN/R have been abbreviated simply to x% DC-SIGN/R).

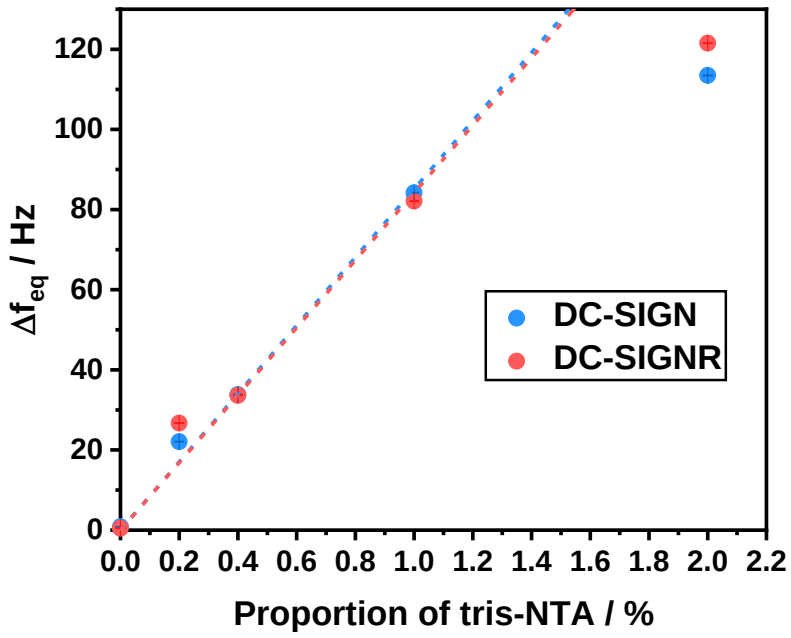

Figure 5.8 Plot of the dependency of Δf for the association of protein to SLB surface on the percentage of tris-NTA containing lipid for DC-SIGN (blue) and DC-SIGNR (red). Where Δf_{eq} is taken as the difference in average equilibrium frequency 2 mins before addition of protein and that after washing with buffer after addition of protein. Linear fits taken between 0-1% tris-NTA; DC-SIGN: Y = (85 ± 4)X; DC-SIGNR: Y = (84 ± 7)X; $R^2 \geq 0.99$.

5.3.3 SLB-lectin·glycan-GNP binding

In order to confirm the binding of G5-Man to these SLB-DC-SIGN/R surfaces, G5-Man was incubated against SLBs of either 2% DC-SIGN, 2% DC-SIGNR, a control SLB bearing 2% His-GFP (His-tagged green fluorescent protein; produced in-house) or an inert SLB containing no tris-NTA. Both the inert and GFP bearing SLBs displayed no significant shift in frequency in the presence of G5-Man (**Figure 5.9A,B**). However, DC-SIGN and DC-SIGNR demonstrated Δf of ~-72 and ~-35 Hz, respectively, demonstrating that binding occurred only through contact with DC-SIGN/R lectins (**Figure 5.9C,D**). Furthermore, DC-SIGN/R retained ~75% of associated glycan-GNPs after washing the surface with binding buffer, indicating strong adsorption between the glycan-GNPs with the protein surface. In order to confirm the mannose-dependent specificity of binding between G5-Man with DC-SIGN/R, the surface was then washed with 50 mg mL^{-1} of D-mannose (Man). The addition of Man induced a significant shift in both frequency and dissipation, observed in all four SLBs tested, which is likely a result of the high viscosity from the high Man concentration. However, upon further washing with pure binding buffer, the frequency was found to return to that prior to the addition of glycan-GNP, indicating the successful dissociation of all glycan-GNP *via* the competition and inhibition of the mannose-binding sites of DC-SIGN/R by free Man.

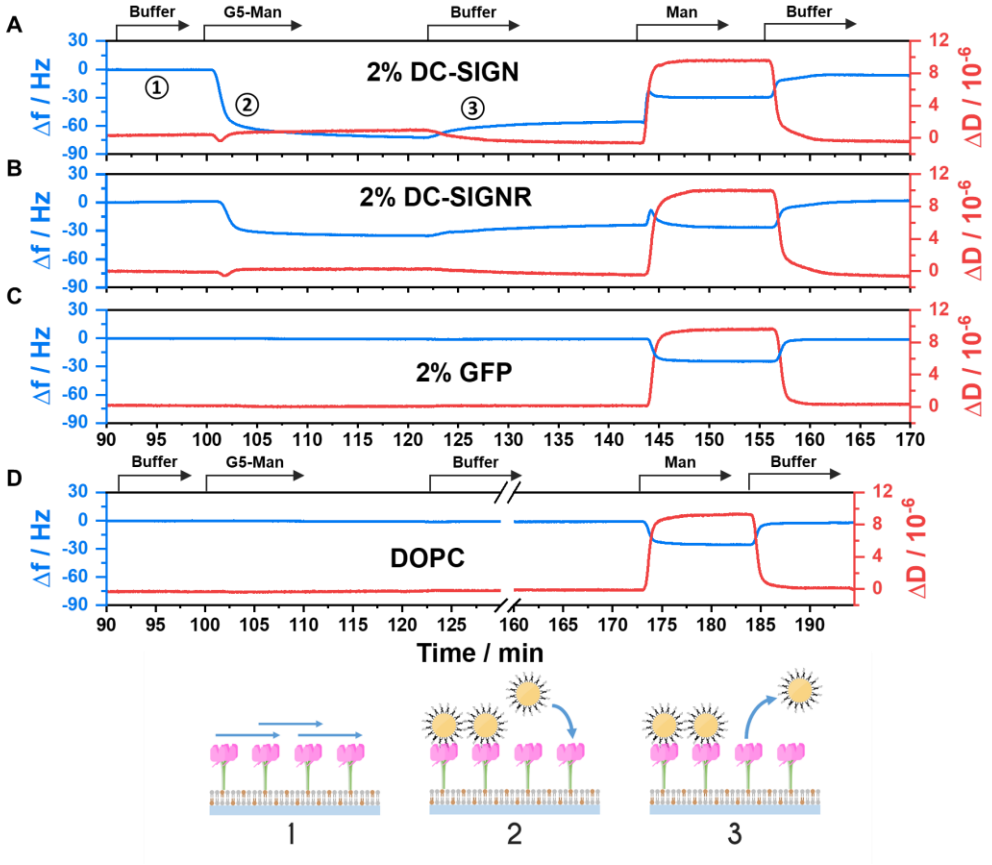

Figure 5.9 QCM-D measurements of the Δf and ΔD time profiles for the association of G5-Man followed by its partial dissociation in binding buffer and complete dissociation in 50 mg mL^{-1} free mannose for SLBs bearing (A) 2% DC-SIGN and (B) 2% DC-SIGNR; and the interaction of G5-Man and Man with SLBs bearing (C) 2% GFP and (D) DOPC only. 1. Demonstrates the equilibration of the SLB-lectin surface in binding buffer; 2. demonstrates the binding of G5-Man with time; and 3. demonstrates the partial dissociation of G5-Man after washing with binding buffer. (Buffer: binding buffer; Man: mannose).

In order to investigate whether the rate at which G5-Man associates to the DC-SIGN/R surface is primarily driven by the binding or by the diffusion of the GNP within the QCM-D chamber, the rate of change of the frequency shift upon addition of GNP was analysed for its dependency on the flowrate. Systems where the rate determining step of binding is driven the diffusion of analyte to the surface are termed “mass transport limited”. For such systems, the rate of uptake of analyte is proportional to the cubic root of the flowrate (Q) and the concentration of analyte.[28] Therefore, the presence of mass transport limitations can be determined by measuring the rate of change of frequency at the turning point of the association curve, against $Q^{1/3}$, for addition of a fixed concentration of analyte.[29] Here, Δf was monitored for the addition of 12.5 nM of G5-Man to 0.4% DC-SIGN at flowrates of 5 and 20 µL min^{-1},

where the steepest region of the decays were fitted with linear fits to obtain the $\mathrm{d}\Delta f/\mathrm{d}t$ (**Figure 5.10A**). These were then plotted against $Q^{1/3}$ (**Figure 5.10B**). Here, though the rate of change of Δf with $Q^{1/3}$ appears not to be strictly linear, an effective gradient of ~1.4 Hz min^{-1} (µL min^{-1})$^{-1/3}$ may be obtained, demonstrating that flowrate does indeed play a significant contribution to the uptake of GNP by the surface. Consequently, in order to obtain the optimal associations rates, high flowrates were employed at low concentrations of GNP in order to minimise mass transport limitation effects.

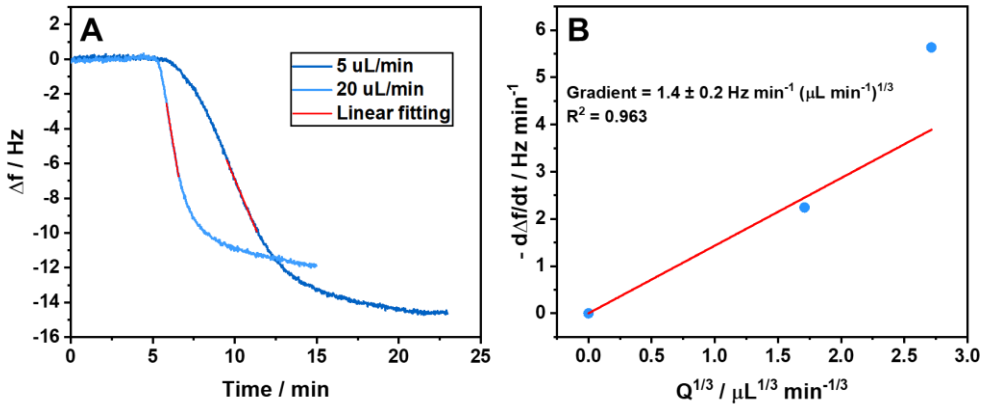

Figure 5.10 Plots demonstrating the effect of flowrate on the rate of change of frequency for G5-Man binding to SLBs bearing 0.4% DC-SIGN. Plots demonstrate (A) the time profiles of the Δf for flowrates of 5 and 20 µL min^{-1} (dark and light blue, respectively), fitted with linear fits at the turning point of the decay (providing $\mathrm{d}\Delta f/\mathrm{d}t$ of -2.24 ± 0.02 and -5.63 ± 0.08 Hz min^{-1} for 5 and 20 µL min^{-1}, respectively; $R^2 \geq 0.995$); and (C) the turning point $\mathrm{d}\Delta f/\mathrm{d}t$ values against the cubic root of the flowrate, $Q^{⅓}$, fitted with a linear fit, where $Y = (1.4 \pm 0.2)X$; $R^2 = 0.963$.

5.3.4 SLB-lectin·glycan-GNP affinity quantification

Due to the proportional relationship between the change in areal mass (Δm) with Δf, highlighted by the Sauerbrey equation (**Equation 5.1**), the rates and affinities of G5-Man binding to the SLB-DC-SIGN/R surfaces can be quantified by measuring the change of Δf with time or concentration. For example, as $-\Delta f$ is proportion to the amount of bound glycan-GNP, by monitoring $-\Delta f$ over increasing concentrations of G5-Man, a Langmuir model can be used to determine the dissociation constant (K_d) (**Equation 5.2**; where $\Delta f_{\max}$ is the Δf at maximum binding; [G] is the glycan-GNP concentration; and n is the Hill coefficient, which is assumed to be $n = 1$ for a Langmuir model). It is important to note here, however, that this approach assumes that Δf is proportional to

the amount of bound GNP at all densities, however, this is not necessarily true, particularly at high protein densities where significant hydrodynamic coupling effects can be observed (**Figure 5.8**). Because of this, the approach described here only conveys an apparent-K_d value. However, this can still provide an informative approximation from which some useful conclusions can be derived.

$$-\Delta f = -\Delta f_{max} \cdot \frac{[G]^n}{{K_d}^n + [G]^n} \qquad 5.2$$

To investigate the influence of SLB-lectin surface density on the binding of G5-Man, the change in Δf was monitored over increasing concentration of G5-Man for both SLB-DC-SIGN (**Figure 5.11**) and SLB-DC-SIGNR (**Figure 5.12**). From here, the Δf corresponding to the point of saturated binding at each concentration (obtained by averaging the Δf signal at 80% ± 40 s of the incubation time for each concentration of G5-Man) was plotted against the glycan-GNP concentration and fitted with the Langmuir equation in order to obtain the corresponding K_d values for each system (**Figure 5.13**).

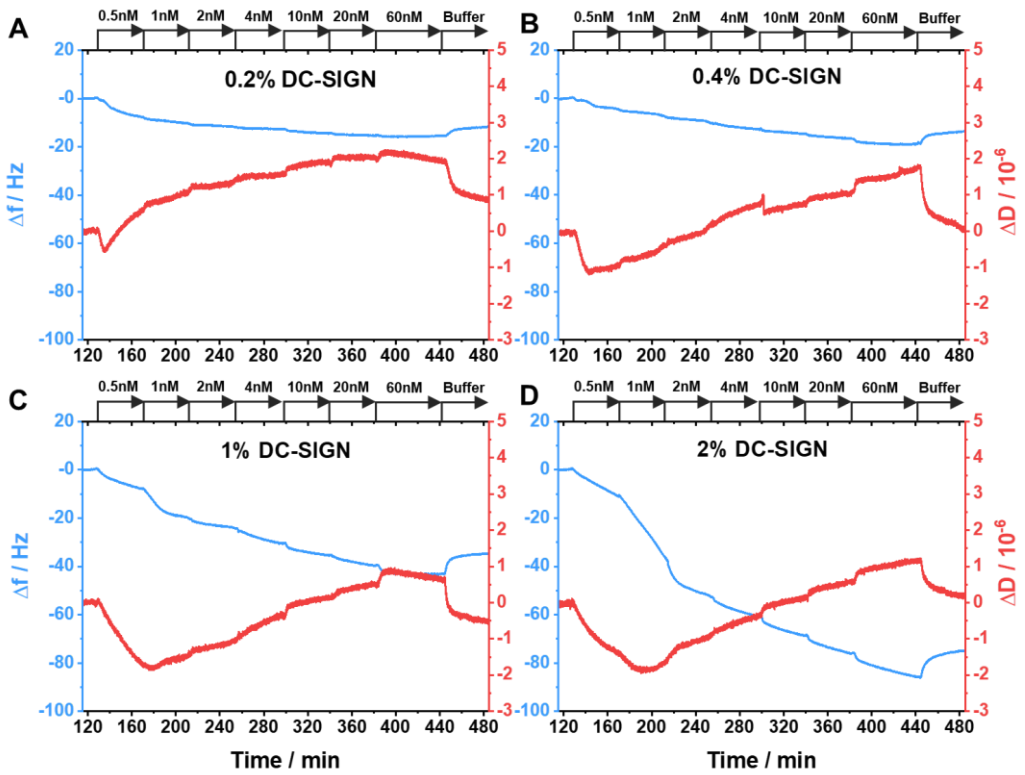

Figure 5.11 Plots of the Δf and ΔD for the titration of G5-Man of concentrations 0.5, 1, 2, 4, 10, 20, 60 nM and subsequent dissociation in buffer with SLBs bearing (A) 0.2% DC-SIGN, (B) 0.4% DC-SIGN, (C) 1% DC-SIGN and (D) 2% DC-SIGN. (Flowrates: 40 µL min^{-1} for 0.5-4 nM; 20 µL min^{-1} for 10-60 nM).

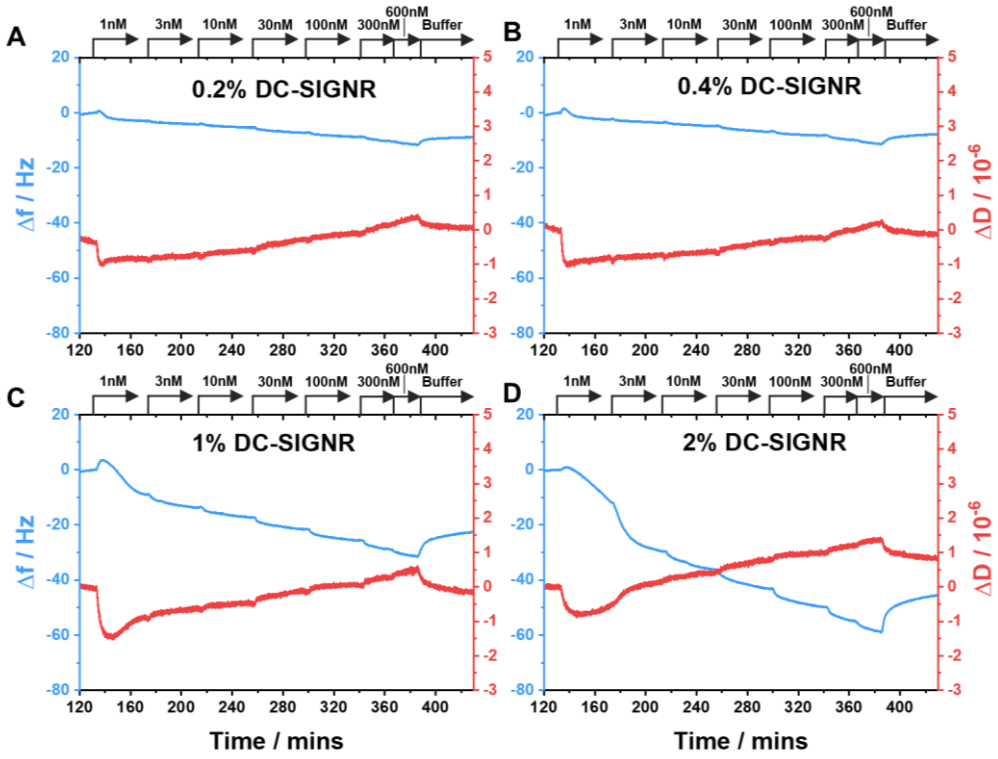

Figure 5.12 Plots of the Δf and ΔD for the titration of G5-Man of concentrations 1, 3, 10, 30, 100, 300, 600 nM and subsequent dissociation in buffer with SLBs bearing (A) 0.2% DC-SIGNR, (B) 0.4% DC-SIGN, (C) 1% DC-SIGN and (D) 2% DC-SIGN. (Flowrates: 40 µL min^{-1} for 1-3 nM; 20 µL min^{-1} for 10-30 nM; 10 µL min^{-1} for 100-300 nM; 5 µL min^{-1} for 600 nM).

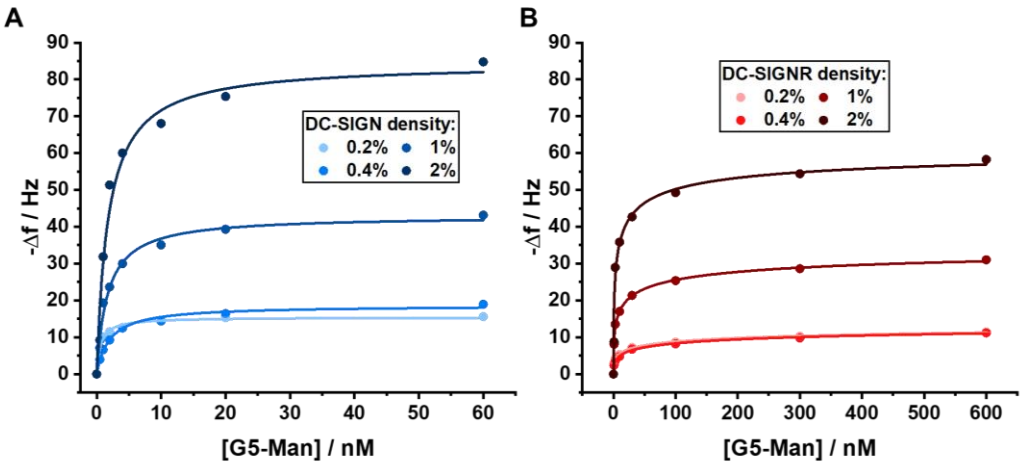

Figure 5.13 Plots of the average $-\Delta f$ signal at 80% ± 40 s of the incubation time for each concentration of G5-Man for the titration of G5-Man against SLBs bearing (A) DC-SIGN and (B) DC-SIGNR; data fitted with Hill equations (Equation 5.2; where n = 1 for DC-SIGN and n is floating for DC-SIGNR; $R^2 \geq 0.98$.

Table 5.2 Summary of parameters for the Langmuir fits of the association of G5-Man to SLBs bearing DC-SIGN or DC-SIGNR of varying densities (Equation 5.2; where $-\Delta f_{\max}$ is the maximum $-\Delta f$ value, K_d is the dissociation constant and n is the cooperativity; n = 1 for DC-SIGN; $R^2 \geq 0.98$).

	DC-SIGN		DC-SIGNR		
Density / %	**$-\Delta f_{\max}$ / Hz**	**Apparent K_d / nM**	**$-\Delta f_{\max}$ / Hz**	**Apparent K_d / nM**	**n**
0.2	15.4 ± 0.3	0.61 ± 0.06	16 ± 1	60 ± 30	0.37 ± 0.03
0.4	18.6 ± 0.5	2.0 ± 0.2	17 ± 2	100 ± 70	0.38 ± 0.04
1	43 ± 1	1.7 ± 0.2	38 ± 3	16 ± 6	0.41 ± 0.04
2	84 ± 4	1.8 ± 0.4	63 ± 6	6 ± 3	0.5 ± 0.1

By comparing the different SLB scaffolds, several interesting observations can be found. Firstly, G5-Man is observed to bind to all four protein densities with strong mid-low nM apparent-K_d values (**Table 5.2**). This demonstrates that binding of surface-bound lectins is similarly as effective as the binding of freely diffusing lectins previously observed using quenching assays performed in solution.[1] Interestingly, binding of G5-Man with 2% DC-SIGN is observed to be 20-fold stronger than the solution-based G5-Man·DC-SIGN binding (1.6 nM *vs* 33.1 nM, respectively). These surface binding K_d values correlate well with the low (sub-nM) IC_{50} values observed for the cell surface-phase inhibition of EBOV-GP by the same glycan-GNP constructs. It is worth noting, however, that for higher protein densities, as well as the limitations of hydrodynamically coupled mass, the accuracy of these results are also limited by the fact that the association of glycan-GNP was unable to reach saturation within the timescales studied, thus the reported Δf values used under-predict the true equilibrium Δf values of binding at each concentration. Despite these limitations, however, these assays show that QCM-D can be used as a powerful tool for providing a more energetic rationale for the efficacy of such inhibitors on membrane surfaces.

Secondly, the dependency of G5-Man binding on the density of SLB-lectin can be compared for SLB-DC-SIGN and SLB-DC-SIGNR. Interestingly, densities between 0.4-2% DC-SIGN provided very little variation in apparent-K_d for G5-Man binding (*i.e.* ~2 nM) demonstrating that binding occurred to a similar extent, regardless of the initial distance between DC-SIGN proteins on

the lipid surface. DC-SIGNR, on the other hand, displayed dramatic variation between lectin densities of 0.2-2%. Here, 2% DC-SIGNR provided the strongest K_d of 6 ± 3 nM, however, this more than doubled to a K_d of 16 ± 6 nM at 1% DC-SIGNR, and increased by another ~5-fold for 0.2-0.4% DC-SIGNR (K_d between ~60-100 nM). This decrease in binding strength with decreasing lectin density infers that the overall apparent strength of binding is very dependent on the initial inter-lectin distance, suggesting that DC-SIGNR is able to form crosslinking interactions between lectins upon the lipid surface. This thus demonstrates an interesting correlation between the binding modes observed in the solution-phase with those on membrane-surfaces for glycan scaffolds of this size. In solution, for example, G5-Man glycans are unable to bridge all of the four binding sites of DC-SIGNR and thus tends to form inter-lectin interactions in order to occupying more CRDs than it would otherwise be sterically permitted with 1:1 lectin:glycan-GNP binding, allowing it to achieve stronger binding.[1] On SLB surfaces, this strong density dependency suggests that the closer the initial inter-DC-SIGNR distance, the easier it is for G5-Man to bridge multiple DC-SIGNRs by binding to the CRDs of adjacent lectins within proximity upon the SLB surface. However, as the protein density decreases and the initial inter-DC-SIGNR distance increases, it may become more difficult for G5-Man to come into contact with a neighbouring lectin, thus a higher proportion of weaker binding interactions may be formed, whereby the glycan-GNP may not be able to occupy all of the binding sites of DC-SIGNR within a reasonable time frame, resulting in a decrease in the overall binding affinity. DC-SIGN, on the other hand, has been observed to form strong stable tetravalent interactions with individual glycan-GNPs in solution-phase experiments due to the complimentary orientations of its binding sites.[1, 30, 31] Thus it can be expected that G5-Man may similarly be able to form strong tetravalent interactions with surface-phase DC-SIGN. This is in agreement with the affinities observed by QCM-D, whereby the K_d values of SLB-DC-SIGN are stronger than those obtained for SLB-DC-SIGNR, which is indicative of stronger simultaneous tetravalent binding mode because of the lower entropic penalty of simultaneous binding in comparison to that of crosslinking (see **Chapter 3**). Confirming this, the lack of density dependency for the glycan-GNP binding strength for SLB-DC-SIGN indicates that the glycan-GNP binding is not dependent on the distance between bound glycan-GNP to the next nearest lectin, indicating that the interaction does not need to form crosslinking in order to obtain a strong multivalent binding interaction.

Though the binding modes appear to be conserved for these glycan-GNPs when translating from solution-phase to surface-phase assays, the

differences in affinity observed by QCM-D present some consequences to consider when investigating the potency of DC-SIGN/R binding upon cell surfaces in comparison to in solution. For example, it is clear that DC-SIGNR is much more dependent on the surface-phase environment than DC-SIGN, as features such as the lectin density play a much bigger role in determining the affinity. This may provide some rationale behind why DC-SIGN demonstrates more consistency between the potencies of cell studies and solution-phase quenching studies in comparison to DC-SIGNR, due to the additional complexity of surface-phase DC-SIGNR·glycan-GNP binding. Interestingly, some studies reveal estimated inter-lectin distances within native DC-SIGN microdomains to be around ~20 nm,[32, 33] similar to that estimated for SLB-lectin densities of 0.2% which interestingly shows the biggest difference in apparent surface-phase K_d between DC-SIGN and DC-SIGNR, thus may provide some mechanistic rationale for the large differences in potency between the inhibition of EBOV-GP on DC-SIGN$^+$ and DC-SIGNR$^+$ 293T cells observed in previous studies.[1]

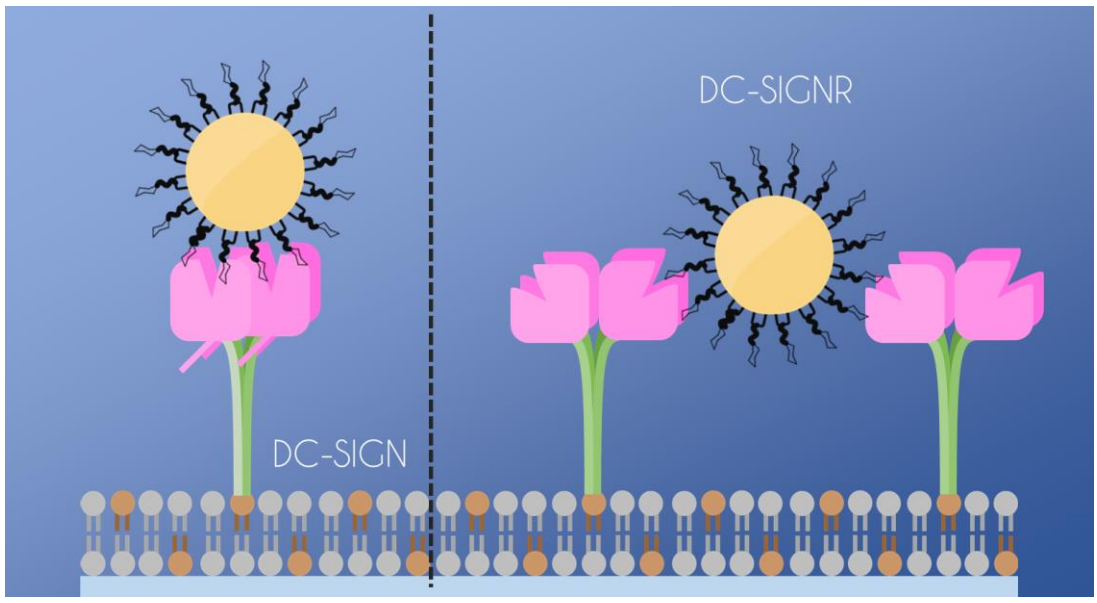

Figure 5.14 Schematic depicting the predicted simultaneous binding mode of DC-SIGN (left) and crosslinking binding mode of DC-SGINR (right) with G5-Man upon SLB surfaces.

It is additionally worth drawing attention to the fact that DC-SIGNR did not to bind well to the same Langmuir model as that of DC-SIGN. Instead K_d values required a modified model, similar to the Hill equation, where $n \neq 1$, in order to provide adequate fits. Such models provided values of $n \approx 0.4$, typical of negative cooperativity whereby bound material inhibits the interaction of any subsequent material. One hypothesis for this may be that as glycan-NPs

begin to bind and crosslink the mobile proteins, the effective density and mobility of free protein available for binding or crosslinking upon the SLB surface begins to decrease, thus making it more difficult for subsequent material to form strong crosslinked interactions involving all four of DC-SIGNR's CRDs. Another interesting observation was that 0.2% DC-SIGN appeared to provide almost 3-fold stronger binding with G5-Man than those of higher densities. However, such results could stem from influences such as the much lower amounts of hydrodynamic coupling or steric hindrance from adjacently bound species. For both of these observations, the failure of glycan-GNP binding reaching true saturation, particularly for DC-SIGNR, may also account for some of these discrepancies.

Another interesting feature of the binding of G5-Man to all SLB-DiMan, is a transient dip in ΔD which appears to suggest an increase in rigidity with low amounts of bound G5-Man. For DC-SIGNR, this ΔD decrease could be caused by an increase in rigidity with crosslinking. This is coupled with a transient increase in Δf which could be related to the release of hydrodynamically coupled solvent as protein clustering occurs. Interestingly, DC-SIGN-SLBs also possess this dip in ΔD which may suggest DC-SIGN is also able to form inter-lectin crosslinking events. The duration of this dip increases with increasing DC-SIGN density; this is likely because crosslinking is much less likely to occur at low densities in comparison to high densities where the probability of crosslinking occurring is much higher, simply because of the close proximity of neighbouring protein. Interestingly, ΔD tends to increase with increased degree of binding indicating a general increase in flexibility with the amount of bound G5-Man regardless of binding mode.

5.3.5 SLB-lectin·glycan-GNP kinetics

In addition to the association profiles, the dissociation of glycan-GNP from the SLB-lectin surface can also be monitored across the different lectins and lectin densities. This was performed following the affinity assays by simply washing the surface with binding buffer. The rate of dissociation of an interaction is usually solely dependent upon the concentration of bound species, however, the rates observed for the systems investigated demonstrated an additional prolonged decay after the initial dissociation phase. In this case, an additional linear term was combined with the dissociation rate model to account for any decay of material from the surface over time (**Equation 5.3**; where $t - t_0$ denotes the time elapsed since the start of the decay, Δf_{eq} is the frequency

shift at equilibrium, k_{off} is the dissociation coefficient and a is the gradient of the linear contribution, or a-term).

$$-\Delta f_t = -\Delta f_{\mathrm{eq}} \cdot \left(1 - e^{-k_{\mathrm{off}} \cdot (t - t_0)}\right) + a \cdot (t - t_0) \qquad 5.3$$

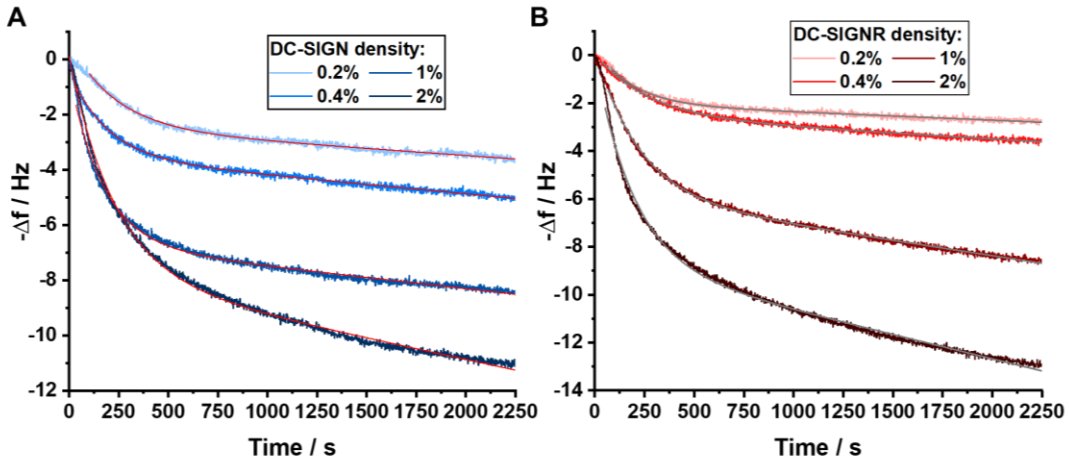

Figure 5.15 Plots of the $-\Delta f$ time profiles for the dissociation of G5-Man from SLB surfaces with varying densities of (A) DC-SIGN and (B) DC-SIGNR. Decays fitted with Equation 5.3; $R^2 \geq 0.98$.

Table 5.3 Summary of the kinetic parameters for the dissociation of G5-Man from SLB-DC-SIGN/R of different densities fitted with Equation 5.3, where k_{off} is the dissociation coefficient, a is the linear decay term and k_{on} is the association coefficient derived by $K_{\mathrm{d}} = k_{\mathrm{off}}/k_{\mathrm{on}}$.

	Density / %	Apparent k_{off} / $\times 10^{-3}$ s^{-1}	a / $\times 10^{-4}$ Hz s^{-1}	Estimated-k_{on} / $\times 10^{5}$ M^{-1} s^{-1}
DC-SIGN	0.2	4.30 ± 0.08	-5.04 ± 0.08	78 ± 9
	0.4	5.28 ± 0.07	-6.80 ± 0.06	30 ± 4
	1	5.87 ± 0.05	-8.11 ± 0.08	35 ± 4
	2	4.35 ± 0.03	-15.5 ± 0.1	26 ± 5
DC-SIGNR	0.2	5.6 ± 0.1	-3.46 ± 0.06	9 ± 5
	0.4	3.81 ± 0.05	-4.86 ± 0.08	4 ± 2
	1	4.57 ± 0.04	-12.70 ± 0.08	30 ± 10
	2	5.15 ± 0.05	-20.3 ± 0.1	70 ± 30

Interestingly all lectins and lectin densities showed similar dissociation rates, with k_{off} values averaging between 4.4 ± 0.7 s^{-1}. This indicates that the off-rate is not the determining factor for the differences in binding affinity. This is likely because the off-rate is solely determined by the strength of binding, rather than other factors such as collision frequency which affect the on-rate. The similarity in k_{off} in indicates that The amount of energy required to displace the glycan-GNPs is roughly the same for both lectins. This suggests tetravalent binding with GNP glycans is achieved for both DC-SIGN and DC-SIGNR, either through one lectin binding to an individual G5-Man (for DC-SIGN) or by two lectins crosslinking a G5-Man (for DC-SIGNR). This means that the observed differences in K_{d} are almost entirely dependent on the association rate, where $K_{\mathrm{d}} = k_{\mathrm{off}}/k_{\mathrm{on}}$. This can be rationalised as DC-SIGNR will take a much longer time forming the initial divalent lectin contact followed by subsequent crosslinking, required to form strong tetravalent binding, in comparison to the much more kinetically simple simultaneous tetravalent binding of DC-SIGN to individual glycan-GNPs. Moreover, the k_{on} is mostly independent of lectin density, which is expected for 1:1 binding, compared to DC-SIGNR, where the k_{on} increases significantly with increasing lectin density due to the closer packing making it easier for G5-Man to crosslink.

Though the k_{off} displays little difference between the initial kinetics each system, the a-term does provide some additional information about how the dissociation evolves over longer time periods. Firstly, it's worth noting that for DC-SIGN, this a-term does not change significantly for densities of 0.2-1%, at a rate of ~0.040 ± 0.009 Hz min^{-1}. This indicates that the a-term likely derives from a systematic deviation with time which is independent of starting frequency (*i.e.* $\sim -\Delta f_{max}$; **Table 5.2**) or the protein density. Interestingly, however, 2% DC-SIGN did show some deviation, demonstrating that high concentrations of DC-SIGN may result in a slightly slower off-kinetics, perhaps as a result of glycan-GNPs finding it easier to rebind or crosslink to neighbouring proteins after dissociation. DC-SIGNR, on the other hand, displays a much more apparent decrease in linearity with density, indicating that crosslinking imparts a biphasic dissociation-phase which occurs over a longer timescale. Here, the magnitude of a is directly proportional to the density of DC-SIGNR and is thus likely directly related to the total amount of bound protein before dissociation.

5.3.6 Imaging glycan-GNP bound SLB-lectin surfaces

Imaging of SLB surfaces has been investigated for a variety of purposes in order to elucidate key structural characteristics of proteins or protein-analyte interactions at membrane surfaces. Techniques such as fluorescence spectroscopy and atomic force microscopy (AFM) can be used to infer a range of information about protein organisation, mobility or topology of lipid bilayers with/without analyte.[34-36] Thus such imaging techniques can be tested for imaging the assembly of lectins in the presence of glycan-GNPs. Because fluorescence microscopy requires the presence of fluorophores, non-fluorescent proteins such as lectins require labelling in order to allow for imaging. As such techniques do not require any site specific labelling (in contrast to FRET) and because His_{24}-DC-SIGN/R do not bear any specific mutations designed for labelling, the lectins can be non-specifically labelled (NSL) and purified in order to remove any denatured protein. Here, His_{24}-DC-SIGN/R were incubated with NHS-ester-Atto463 for 2 hrs and then purified *via* mannose affinity column chromatography in order to yield His_{24}-DC-SIGN-$(Atto643)_x$ and His_{24}-DC-SIGN-$(Atto643)_x$ (herein referred as NSL-DC-SIGN and NSL-DC-SIGNR, respectively).

Though the smooth silicon oxide surfaces used in QCM sensor chips work well for forming SLB surfaces due to their good hydrophilicity, other silicon-based hydrophilic materials, such as mica, borosilicate and silicon nitride, are similarly capable at forming good bilayer coverages. Here, transparent borosilicate glass slides were fixed onto a Teflon sample holder with 4× 200 µL wells. SLBs were assembled by incubating 50 µL of SUVs containing either 0.4% or 2% tris-NTA DODA in SUV buffer with 10 mM $NiCl_2$ for 15 min, followed by washing with SUV buffer and binding buffer. The SLB-lectin was assembled by then incubating with NSL-DC-SIGNR for 60 min, followed by further washing with binding buffer. In order to encourage complete assembly of each material, double concentrations were employed and each phase was incubated for the same amount of time required to achieve complete coverage of the surface as that observed by QCM-D. Fluorescence confocal microscopy showed strong evidence of a thin plane of fluorescence, indicating that SLB-DC-SIGNR had successfully been formed. Once images were taken of the SLB-lectin surface, 800 nM G5-Man was then added and the change in fluorescence signal was measured after 30 min of incubation. Interestingly, however, very little change occurred in the organisation of protein fluorescence with the addition of glycan-GNP for both 0.4% and 2% DC-SIGNR. This suggests one of two things, either: 1. no

crosslinking occurs, despite the evidence presented by the QCM-D assays, or 2. organisation of protein induced by crosslinking is not observable on the micron scale.

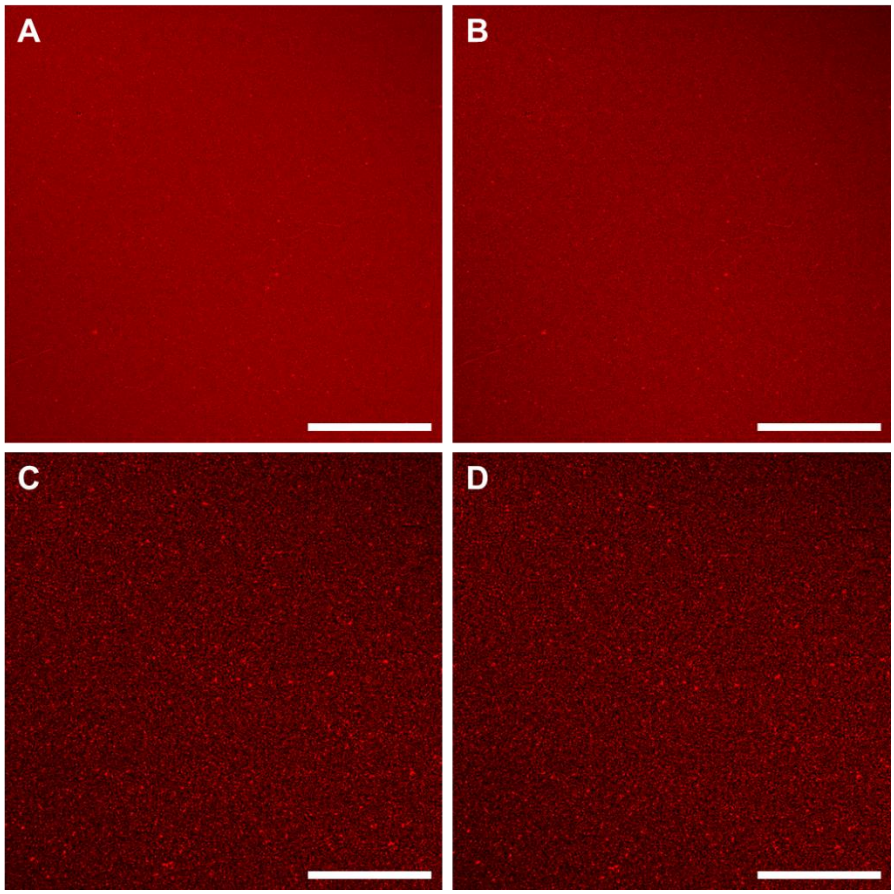

Figure 5.16 Fluorescence confocal microscopy images of SLBs bearing 2% NSL-DC-SIGNR (A) before and (B) 30 min after addition of G5-Man; and 0.4% NSL-DC-SIGNR (C) before and (D) 30 min after addition of G5-Man. (Scale bar: 50 μm).

In order to determine whether the glycan-GNPs did indeed induce organisational change of the SLB-bound protein, a higher resolution method needed to be employed. Electron microscopy provides one such platform with excellent capability for imaging material at resolutions even as low as the sub-nanometre scale. However, though some effort has been devoted to imaging proteins on micelles such as liposomes or polymer nanodiscs,[37] due to the low density of proteins, EM techniques have thus far proved to be unsuitable for inferring the structural characteristics of such systems on SLBs. GNPs, on the other hand, exhibit very high contrast under EM, and can thus be easily imaged to infer structural details about protein interactions, as has been demonstrated in our previous solution-based studies. Moreover, cryogenic transition EM (cryoTEM) provides an excellent platform for probing time

sensitive interactions within their native states in solution by rapidly blotting and freezing samples in liquid ethane before viewing. Here, a new technique has been developed to analyse the binding of glycan-GNPs with SLB-lectin surfaces using cryoTEM imagining.

CryoTEM grids coated with silicon nitride windows provided an excellent platform which can not only support SLBs but also allow the sample to be cryogenically frozen. Here, the confined droplet method was used to assemble the SLB-lectin surfaces, by successive rounds of incubation and washing a 15 µL droplet suspended on a 9 window silicon nitride grid with: i) SUVs containing 2% tris-NTA DODA in Ni^{2+} containing buffer for 15 min (washed firstly with SUV buffer, then binding buffer); ii) DC-SIGNR for 60 min; and finally, iii) G5-Man for 60 min. In order to ensure complete assembly of each material, double concentrations were employed and each phase was incubated for the same amount of time required to achieve complete coverage of the surface, as observed by QCM-D. The grid was then blotted immediately after the final round of washing and rapidly frozen in liquid ethane, before imaging *via* cryoTEM.

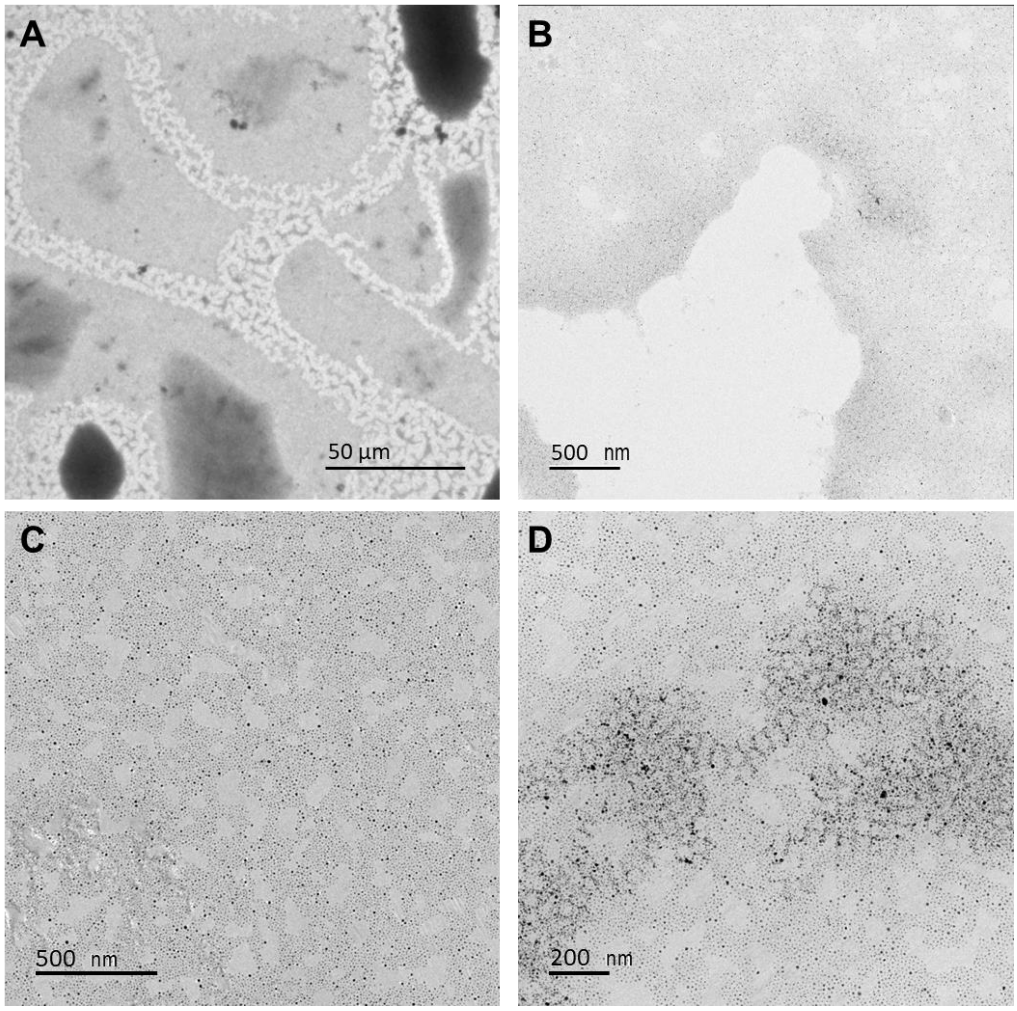

Figure 5.17 Cryogenic transition electron microscopy (cryoTEM) images of G5-Man on SLBs bearing 2% DC-SIGNR. (A) demonstrates the SLB quality at the same magnification as those in Figure 5.16; (B) demonstrates a typical G5-Man distribution at the edge of a well-formed SLB section; (C) demonstrates a typical G5-Man distribution in the middle of a well-formed SLB; (D) demonstrates the typical secondary cluster formation of G5-Man.

Upon initial inspection of the silicon nitride windows from the same scale as that of the fluorescence microscopy images, two main observations can be made which determine the success of SLB formation. Firstly, all intact windows displayed a mixture of large “islands” of around 100 µm, separated by a large number of smaller islands of around 5 µm (**Figure 5.17A**). These islands were found to be composed of extensive distributions of GNPs separated by regions with no nanoparticle coverage, confirmed by elemental analysis of the gold composition in these two regions by cryogenic energy dispersion x-ray (EDX) spectroscopy (**Figure 5.18**). Due to the instability of thin films, these observations suggest that the surface blotting may have resulted in the partial dewetting of the surface, causing the partial rupture of

the SLB. This is evidenced by the apparent tension lines in the small islands (**Figure C.12**), caused by the recoiling of the dewettened film under stress, as well as the accumulation of GNPs at the edges of the islands (**Figure 5.17B**), which may correspond to the angular rims which are shown to be present close to the edge of receding liquid films.[38] Secondly, dark regions of ice were also observed (observed in the bottom left and top right of **Figure 5.17A**), whereby a larger amount of water had retained around some of the regions that had preserved good SLB coverage. However, despite some of these homogeneities in SLB coverage and water retention, there were still large proportions of each grid which displayed both optimal blotting and good SLB coverage and thus are likely to accurately represent the nature of glycan-GNP binding at the SLB-lectin surface.

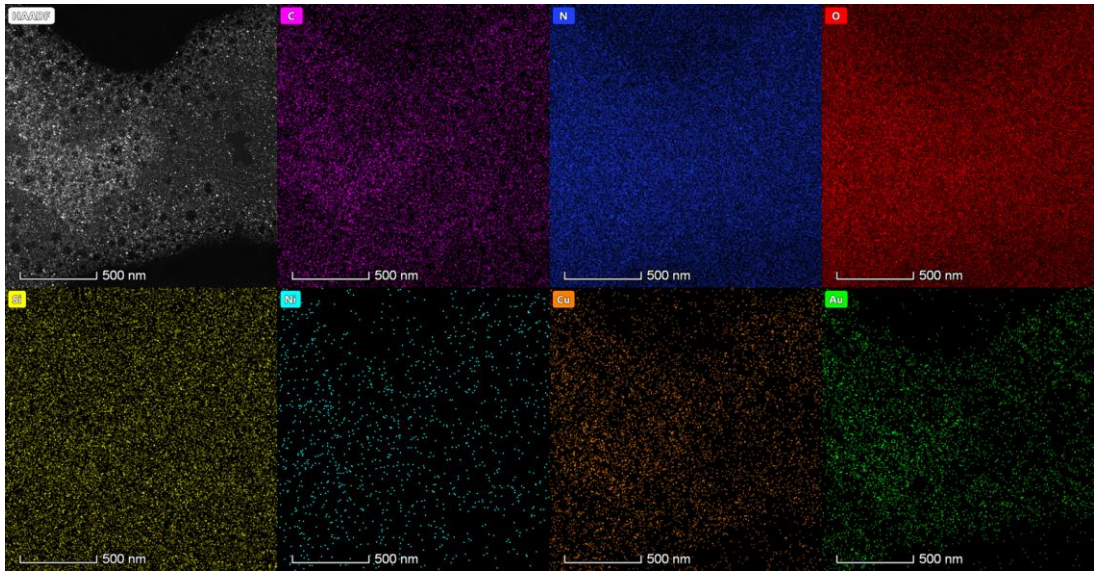

Figure 5.18 Cryogenic energy dispersive x-ray (EDX) spectroscopy images of the three GNP coverages observed by cryoTEM of G5-Man bound to 2% DC-SIGNR, namely, regions of no GNP coverage, typical GNP coverage and clustered GNP regions. Images demonstrate, from top-left to bottom-right, the intensities of density (by high-angle annular dark field (HAADF) imaging), and EDX intensities of carbon (C), nitrogen (N), oxygen (O), silicon (Si), nickle (Ni), copper (Cu) and gold (Au).

The more representative regions were mainly occupied by networks of evenly spaced GNPs, of one layer thick, separated by patches containing no particles (**Figure 5.17C**). Given that the surface was washed a minimum of 10 times between each incubation stage, the presence of GNPs indicated successful formation of SLB-DC-SIGNR. Additionally, the even and non-overlapping organisation of these particles indicated that monolayer coverage of G5-Man

had formed, which is expected for the specific binding of G5-Man to a monolayer of lectins sterically hindering the binding of any subsequent glycan-GNP. Interestingly, the observed networks of GNPs also suggests that crosslinking of DC-SIGNR is likely to have occurred, in agreement with the QCM-D observations, as an absence of crosslinking would likely have resulted in a lack of GNP organisation on the SLB-lectin surface. Furthermore, the nearest neighbour distances of between GNP centres can be measured at this region to provide average centre to centre inter-GNP distance of 9.7 ± 2.8 nm. This agrees well with the length of a single DC-SIGNR binding to two glycan-GNPs, as the inter-binding site distance of the longest length of DC-SIGNR is modelled at around 8.0 nm.[23] Had individual glycan-GNPs bound to individual lectins then the GNP NND would have more closely correspond to the estimated inter-tris-NTA distance for 2% tris-NTA DODA, at ~7 nm, with no observable surface organisation. This therefore indicates that the glycan-GNPs are likely bridged by DC-SIGNR molecules forming extensive crosslinked networks throughout the SLB surface.

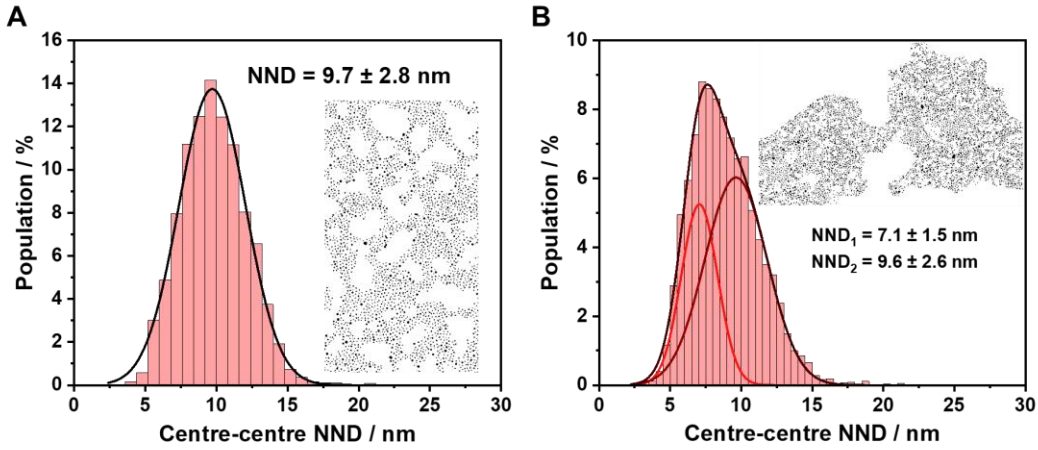

Figure 5.19 Centre to centre inter-GNP nearest neighbour distance (NND) distributions for the cryoTEM images of G5-Man bound to SLBs bearing 2% DC-SIGNR, where (A) is the typical distribution of G5-Man upon well-formed SLB surface (as from Figure 5.17C) and (B) is the G5-Man distribution at regions of secondary clustering (as from Figure 5.17D). (NND = mean ± ½FWHM; $R^2 \geq 0.994$).

Additionally, these 2% SLB-DC-SIGNR·glycan-GNP surfaces also displayed some additional regions, of a few microns in size, which exhibit even more tightly clustered GNPs with NNDs of both 9.6 ± 2.6 nm, similar to those of the typical GNP distribution, but also 7.1 ± 1.5 nm which, interestingly, correspond closely to the estimated inter-tris-NTA distances. This may indicate some simultaneous binding at such high concentrations of protein and G5-Man or

simply some secondary GNP-GNP interactions. These questions could be investigated by analysing the topology of these regions by tilting the sample or *via* other imaging techniques such as AFM.

Here we have demonstrated the application of cryoTEM as a promising tool for inferring the surfaces-phase binding modes of lectins with glycan-GNPs at artificial SLB surfaces. This method can therefore be used to investigate the G5-Man binding of SLB-DC-SIGN, to provide clarification of the differences in binding mode predicted by QCM-D. Additionally, investigation of the evolution of binding over time and with varying protein density would also provide a more visual comparison to clarify the interesting conclusions observed from the affinity measurements obtained here.

5.4 Conclusion

This thesis has thus far demonstrated the power of glyconanoparticles, such as glycan-QDs and glycan-QRs, as solution-based probes for determining the structural and biophysical bases for lectin binding modes and affinities with respect to glyconanoparticle design. However, solution-based methods only provide a partial picture for the true nature of binding for membrane lectins, such as DC-SIGN and DC-SIGNR, where detailed information regarding their surface-phase binding behaviours may be lacking. For this reason we have developed a multimodal supported lipid bilayer approach to determine the binding mode and affinity of membrane lectins by fixing their histidine-tagged extracellular domains onto a Ni^{2+}-bearing supported lipid bilayers and probing their binding characteristics *via* QCM-D and cryoTEM imaging. Here DC-SIGN/R coated SLBs with various lectin densities were successfully synthesised using tris-NTA DODA-containing DOPC small unilamellar vesicles, as confirmed by QCM-D. Mannose functionalised 5nm GNPs (G5-Man) were found to bind more strongly to DC-SIGN/R coated SLBs in comparison to the previously conducted solution-phase assays, despite the slower association kinetics. In addition, it was found that whilst binding with SLB-DC-SIGN is relatively independent of protein density, SLB-DC-SIGNR displayed a strong density dependency, with much higher affinities at higher protein surface concentrations. This result indicates that, as with the solution-based studies, DC-SIGN binds tetravalently to individual G5-Man, whereas DC-SIGNR crosslinks multiple G5-Man particles. To confirm this, SLB-DC-SIGNR was assembled on silicon nitride TEM grids and incubated with G5-Man before being cryogenically frozen for cryoTEM imaging. This method demonstrated the formation of SLB-DC-SIGNR·glycan-GNP surfaces and

gave visual evidence of the formation of glycan-GNP crosslinking *via* the formation of GNP clustered networks with inter-GNP distances similar to that of the predicted inter-binding site length of DC-SIGNR.

The development of this multimodal method for analysing the surface-phase binding of membrane lectins provides an excellent tool for more reliably analysing the binding of glyconanoparticles with lectins by more closely resembling their natural environments. Here, SLBs replicate the fluid nature of cell surfaces and provide a simple platform for controlling protein density, thus providing a good replication of the proteins observed on cell surfaces which cannot be obtained in solution-based assays. This method therefore acts as a bridge between solution-based studies with cell based studies to provide a better understanding of the differences between lectins' affinities and binding modes with their true viral inhibition potencies, thus better informing the design of therapeutics that can target such interactions at cell surface environments.

5.5 Future Direction

While this approach provides some compelling conclusions, the quantitative results obtained by QCM-D can only provide estimated kinetic parameters due to the large influence of hydrodynamically coupled mass. With this in mind, the future direction of this work would be to confirm these predictions using techniques where these factors are less influential, such as spectroscopic ellipsometry which, due to its optical readout, provides a more accurate quantification of adsorbed material. In addition, the cryoTEM experiments here describe only a preliminary methodology for this SLB imaging technique. Thus, further work may be needed to optimise SLB stability and to provide more comparative studies for SLB-DC-SIGN/R using different lectin densities, glycan-GNP concentrations and glycan-GNP incubation time points in order to provide more empirical evidence for the differences between the different SLB systems described in the QCM-D studies.

In addition, investigation of the influence of glyconanoparticle size, shape and glycan composition on lectin binding affinity and mode at lipid bilayer surfaces would also be worth exploring. Such experiments would lead to a better understanding of how different glycoconjugates interact at the membrane surface. Furthermore, such experiments would highlight any discrepancies between solution and surface-phase assays. For example, it is reasonable to hypothesise that with increasing GNP size, interesting steric shielding or inter-lectin multivalency may influence the binding of both

DC-SIGN and DC-SIGNR in ways that cannot be predicted *via* solution-phase assays alone.

The development of these assays therefore provide an important step towards building a more detailed, quantitative and qualitative understanding of how glycans interact with lectins at the cell surface. These key data will help in the effort to design potent and specific glycoconjugates for targeting lectins for specific therapeutic application.

References

1. D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo and D. Zhou, *J. Am. Chem. Soc.*, 2020, **142**, 18022-18034.
2. A. Cambi, F. de Lange, N. M. van Maarseveen, M. Nijhuis, B. Joosten, E. M. van Dijk, B. I. de Bakker, J. A. Fransen, P. H. Bovee-Geurts, F. N. van Leeuwen, N. F. Van Hulst and C. G. Figdor, *J. Cell. Biol.*, 2004, **164**, 145-155.
3. P. Liu, X. Wang, M. S. Itano, A. K. Neumann, K. Jacobson and N. L. Thompson, *Traffic*, 2012, **13**, 715-726.
4. M. S. Itano, M. S. Graus, C. Pehlke, M. l. J. Wester, P. Liu, K. A. Lidke, N. L. Thompson, K. Jacobson and A. K. Neumann, *Front. Phys.*, 2014, **2**.
5. M. Koopman, A. Cambi, B. I. de Bakker, B. Joosten, C. G. Figdor, N. F. van Hulst and M. F. Garcia-Parajo, *FEBS Lett*, 2004, **573**, 6-10.
6. A. K. Neumann, N. L. Thompson and K. Jacobson, *J. Cell Sci.*, 2008, **121**, 634-643.
7. T. Terada, M. Nishikawa, F. Yamashita and M. Hashida, *Int J Pharm*, 2006, **316**, 117-123.
8. F. Teillet, B. Dublet, J. P. Andrieu, C. Gaboriaud, G. J. Arlaud and N. M. Thielens, *J Immunol*, 2005, **174**, 2870-2877.
9. N. Frison, M. E. Taylor, E. Soilleux, M. T. Bousser, R. Mayer, M. Monsigny, K. Drickamer and A. C. Roche, *J. Biol. Chem.*, 2003, **278**, 23922-23929.
10. Y. Shinohara, Y. Hasegawa, H. Kaku and N. Shibuya, *Glycobiology*, 1997, **7**, 1201-1208.
11. A. Monaco, V. P. Beyer, R. Napier and C. R. Becer, *Biomacromolecules*, 2020, **21**, 3736-3744.
12. P. Y. Lozach, H. Lortat-Jacob, A. de Lacroix de Lavalette, I. Staropoli, S. Foung, A. Amara, C. Houles, F. Fieschi, O. Schwartz, J. L. Virelizier, F. Arenzana-Seisdedos and R. Altmeyer, *J Biol Chem*, 2003, **278**, 20358-20366.
13. V. Porkolab, C. Pifferi, I. Sutkeviciute, S. Ordanini, M. Taouai, M. Thepaut, C. Vives, M. Benazza, A. Bernardi, O. Renaudet and F. Fieschi, *Org. Biomol. Chem.*, 2020, **18**, 4763-4772.
14. G. Sauerbrey, *Z. Phys*, 1959, **155**, 206-222.
15. M. E. Yakovleva, G. R. Safina and B. Danielsson, *Anal Chim Acta*, 2010, **668**, 80-85.
16. Z. Pei, H. Anderson, T. Aastrup and O. Ramström, *Biosens Bioelectron*, 2005, **21**, 60-66.
17. R. P. Richter, R. Bérat and A. R. Brisson, *Langmuir*, 2006, **22**, 3497-3505.
18. S. Sangabathuni, R. V. Murthy, P. M. Chaudhary, B. Subramani, S. Toraskar and R. Kikkeri, *Sci Rep*, 2017, **7**, 4239.
19. J. Piella, N. G. Bastús and V. Puntes, *Chem. Mater.*, 2016, **28**, 1066-1075.
20. S. Srimasorn, L. Souter, D. E. Green, L. Djerbal, A. Goodenough, J. A. Duncan, A. R. E. Roberts, X. Zhang, D. Débarre, P. L. DeAngelis, J. C. F. Kwok and R. P. Richter, *Sci. Rep.*, 2022, **12**, 10980.
21. I. Reviakine, D. Johannsmann and R. P. Richter, *Anal. Chem.*, 2011, **83**, 8838-8848.
22. H. Fischer, I. Polikarpov and A. F. Craievich, *Protein Sci.*, 2004, **13**, 2825-2828.

23. H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327-1335.
24. D. Johannsmann, I. Reviakine and R. P. Richter, *Anal. Chem.*, 2009, **81**, 8167-8176.
25. J. F. Nagle and S. Tristram-Nagle, *Biochim. Biophys. Acta*, 2000, **1469**, 159-195.
26. G. Tabarani, M. Thepaut, D. Stroebel, C. Ebel, C. Vives, P. Vachette, D. Durand and F. Fieschi, *J. Biol. Chem.*, 2009, **284**, 21229-21240.
27. I. Carton, A. R. Brisson and R. P. Richter, *Anal. Chem.*, 2010, **82**, 9275-9281.
28. W. T. Hermens, M. Beneš, R. Richter and H. Speijer, *Biotechnol. Appl. Biochem.*, 2004, **39**, 277-284.
29. S. A. van der Meulen, G. V. Dubacheva, M. Dogterom, R. P. Richter and M. E. Leunissen, *Langmuir*, 2014, **30**, 6525-6533.
30. Y. Guo, I. Nehlmeier, E. Poole, C. Sakonsinsiri, N. Hondow, A. Brown, Q. Li, S. Li, J. Whitworth, Z. Li, A. Yu, R. Brydson, W. B. Turnbull, S. Pohlmann and D. Zhou, *J. Am. Chem. Soc.*, 2017, **139**, 11833-11844.
31. Y. Guo, C. Sakonsinsiri, I. Nehlmeier, M. A. Fascione, H. Zhang, W. Wang, S. Pohlmann, W. B. Turnbull and D. Zhou, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 4738-4742.
32. Michelle S. Itano, C. Steinhauer, J. J. Schmied, C. Forthmann, P. Liu, Aaron K. Neumann, Nancy L. Thompson, P. Tinnefeld and K. Jacobson, *Biophys. J.*, 2012, **102**, 1534-1542.
33. B. I. de Bakker, F. de Lange, A. Cambi, J. P. Korterik, E. M. H. P. van Dijk, N. F. van Hulst, C. G. Figdor and M. F. Garcia-Parajo, *ChemPhysChem*, 2007, **8**, 1473-1480.
34. N. F. Bonet, D. G. Cava and M. Vélez, *Front. Mol. Biosci.*, 2022, **9**, 935376-935376.
35. H. Basit, S. G. Lopez and T. E. Keyes, *Methods*, 2014, **68**, 286-299.
36. A. Choucair, M. Chakrapani, B. Chakravarthy, J. Katsaras and L. J. Johnston, *Biochim. Biophys. Acta Biomembr.*, 2007, **1768**, 146-154.
37. D. S. Chorev and C. V. Robinson, *Nat. Chem. Biol.*, 2020, **16**, 1285-1292.
38. A. M. J. Edwards, R. Ledesma-Aguilar, M. I. Newton, C. V. Brown and G. McHale, *Sci. Adv.*, 2016, **2**, e1600183.

Chapter 6
Evaluation of the inter-carbohydrate binding domain distances of DC-SIGN and DC-SIGNR using single molecule FRET

All material production and empirical data collection and analysis in the following chapter was performed by the author of this thesis. Single molecule FRET measurements were performed with the aid of Dr Bogachan Tahirbegi and Dr Liming Ying (Imperial College London). All computation simulations and analysis were performed by Alexander St John. All plasmids were prepared in-house by Dr Yuan Guo.

6.1 Introduction

Within the past few decades, the study of multivalency in medicinal biology has seen an increase in research. This is because, where monovalent protein-ligand interactions which simply involve an individual binding site, such as those of enzymes, can be targeted using conventional monovalent therapeutics, the majority of multivalent targets often rely on their ability to form multiple interactions due to the intrinsically weak nature of their monovalent constituents, thus targeting such interactions for therapeutic applications requires multivalent solutions. This, however, can lead to complications when designing targeting agents as the efficacy of such multivalent interactions can often be greatly dependent on the spatial and orientational arrangements of the protein binding sites with respect to that of the ligand counterpart. Thus, in order to optimise potency and specificity, multivalent therapeutics need to incorporate both a significant degree of high affinity functional groups as well as optimised spatial specificity. However, the structures of many important multimeric targets are completely unknown, thus designing specific ligands which complement the binding site arrangements of such targets can prove to be difficult. This problem is particularly prevalent for membrane proteins, which have a reputation for being challenging to crystallise or image by conventional structural determination techniques such as x-ray crystallography or electron microscopy. Therefore, the development of alternative techniques to determine these key structural parameters of multivalent proteins presents a highly valuable goal in order to better inform the design of multi- or poly- valent therapeutics.

This multivalent spatial specificity has been explored in great depth in the previous chapters for the near-identical tetrameric glycan-binding

proteins, DC-SIGN and DC-SIGNR (collectively referred to as DC-SIGN/R) which have been associated with the infection of a range of serious viruses including HIV-1, ebolavirus and, more recently, SARS-CoV-2.[1-5] It has been shown that, despite their similarities in mannose-binding sites,[6, 7] they demonstrate very different binding affinities and binding modes to the same glycoconjugates (**Chapter 3-5**),[8-10] as well as some of the same pathogens.[11-13] Though these differences have been attributed variations in the specific tetrameric spatial arrangements/originations between their multiple binding sites, there has been limited advances in solving the tetrameric structure of these proteins to compliment these observations. Furthermore, detailed information regarding the specific distances between each binding site would provide key information required to inform the relational design of highly specific and potent therapeutic targeting ligands with glycan arrangements that perfectly complement the tetrameric binding site arrangements of DC-SIGN/R.

Techniques such as Förster resonance energy transfer (FRET) and electron paramagnetic resonance (EPR) have been dubbed the term “spectroscopic rulers” due to their ability to probe the distances between two locations at the sub-10 nm range. FRET is the non-radiative dipole-dipole energy transfer from an excited donor fluorophore to an acceptor fluorophore with overlapping excitation and absorption spectra. The efficiency of this energy transfer process can be obtained by measuring the proportion of fluorescence emitted from the acceptor compared to the total fluorescence from both fluorophores, and is inversely related to the distance between the two fluorophores to the sixth power, thus allowing a measurement of distance to be derived (**Equation 6.2**). A limitation with traditional FRET approaches is that measurements taken from a bulk sample will always return the average of the distances between the fluorophores, which poses a problem if you wish to measure multiple distances or multiple conformational changes within the same sample. Single molecule FRET (smFRET) overcomes this limitation by focusing the excitation beam into a small volume within the sample such that only one molecule may pass the focused light at a time. This means that as each protein passes the focal point, the FRET efficiency can be measured and a distribution of the prevalence of each FRET efficiency can be obtained, from which the distances can be derived. Other techniques such as electron paramagnetic resonance EPR can also be harnessed as “spectroscopic rulers” to measure distances between spin labels and can additionally provide useful information regarding the dynamics of the label environments. However, due to the small sample sizes and room temperature

measurements, smFRET is often a simpler strategy for measuring distances between labels compared to EPR techniques such as pulse electron-electron double resonance (PELDOR) spectroscopy.

SmFRET has been utilised for probing a range of membrane proteins, mainly to characterise protein conformational changes and dynamics.[11, 12] This technique works very well for proteins that are monomeric or heteromultimeric in nature as the distances between two regions can easily be labelled with a FRET pair of dye fluorophores by introducing site mutations at the positions that require measuring. However, for proteins that are homomultimeric in nature, labelling specific monomer units with a single FRET pair of fluorophores poses a challenge, as introducing a site specific mutation will lead to the binding of any number of dyes per protein. Such systems would cause a large amount of interference between the FRET pair and the additional dyes, leading to inaccurate FRET efficiency readouts.

This chapter presents a new technique for site specifically labelling two out of four of the CRDs of DC-SIGN/R with a FRET pair of dyes using a heterotetramerisation and double purification strategy. Because of this, for the first time, estimates for the distances between the CRDs on each monomer unit of the tetrameric proteins DC-SIGN/R have been measured using smFRET (**Figure 6.7**). These distances were then simulated using computational models in order to examine the plausibility of these empirically obtained parameters as reliable descriptors of the inter-CRD distances of the two tetramers.

6.2 Experimental

6.2.1 Donor and Acceptor Labelled DC-SIGN/R

DC-SIGN Q274C, DC-SIGNR R287C were produced and purified as described in **Section 2.1.4.2**. To the concentrated protein (1 eq), either maleimide-Atto488 (9.37 mM in DMSO, 5 eq) or maleimide-Atto643 (9.34 mM in DMSO, 3 eq) with was added with TCEP·HCl (0.2 eq) and immediately mixed by vortex. The solution was then covered and slowly rotated at r.t. for 1 hr, before being stored at 4 ^{o}C overnight. The solution was then added to a mannose-sepharose affinity column, which had been pre-equilibrated in binding buffer, and washed with binding buffer until the excess dye was removed. The protein was then eluted using elution buffer and $CaCl_2$ (pH 7.4) was then added to each fraction to a final concentration of 10 mM. Concentration was determined using UV-vis spectroscopy (**Equation 6.1**).

Protein molecular weights (MW) were confirmed from deconvoluted HRMS data, where calculated protein MW values were obtained from the amino acid sequences of one monomeric unit of the corresponding protein and LE was obtained from the ratio of the labelled protein to unlabelled protein peak areas.

DC-SIGN-Atto488: HRMS: calculated MW [DC-SIGN Q274C] 39172.22, [DC-SIGN Q274C + 2Ca] 39252.40 and [DC-SIGN Q274C-Atto488] 39884.22; found 39253.48 and 39888.90. LE: 72%.

DC-SIGN-Atto643: HRMS: calculated MW [DC-SIGN Q274C] 39172.22, [DC-SIGN Q274C + 2Ca] 39252.40 and [DC-SIGN Q274C-Atto643] 40130.22; found 39251.88 and 40135.25. LE: 74%.

DC-SIGNR-Atto488: HRMS: calculated MW [DC-SIGNR R287C] 37425.95, [DC-SIGNR R287C + 2Ca] 37506.10 and [DC-SIGNR R287C-Atto488] 38137.95; found 37506.81 and 38141.87. LE: 70%.

DC-SIGNR-Atto488: HRMS: calculated MW [DC-SIGNR R287C] 37425.95, [DC-SIGNR R287C + 2Ca] 37506.10 and [DC-SIGNR R287C-Atto488] 38383.95; found 37506.25 and 38389.07. LE: 63%.

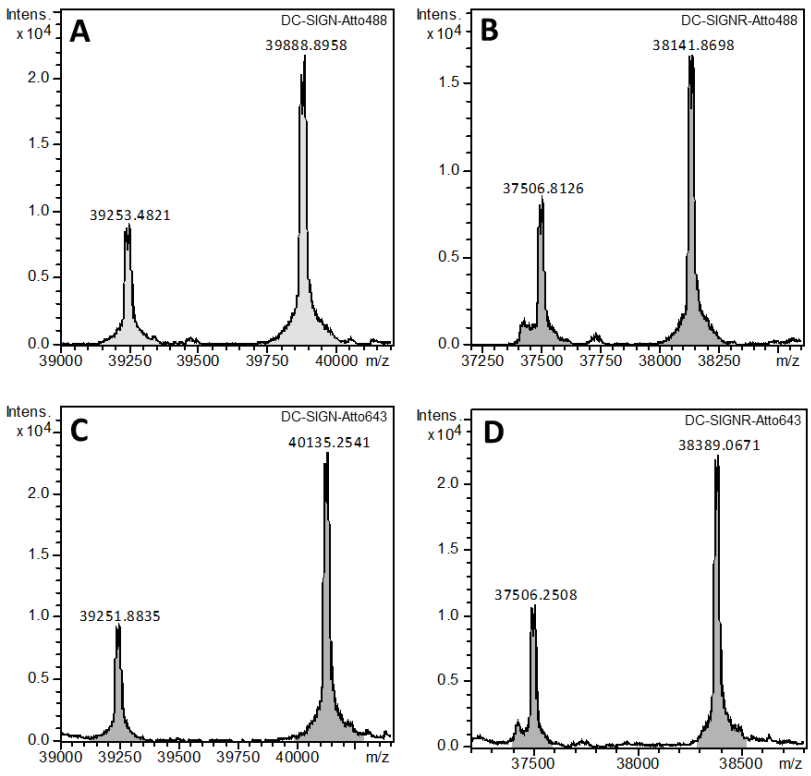

Figure 6.1. Deconvoluted HRMS spectra of the monomeric extracellular domains of (A) DC-SIGN-Atto488, (B) DC-SIGN-Atto643, (C) DC-SIGNR-Atto488, (D) DC-SIGN-Atto643.

6.2.2 GBDC-SIGN

Plasmids encoding for galactose binding (GB)DC-SIGN and GBDC-SIGNR were made in-house *via* conventional mutagenesis techniques, from those encoding for wild type extracellular domains of both proteins, by replacing the "EPNNVGE" amino acid motif associated with mannose binding in DC-SIGN, or "EPNNSGN" in DC-SIGNR, with a "QPDNWYGHGLGGG" motif associated galactose binding, as confirmed by DNA sequencing (see **Appendix A.1** for amino acid identity), as described previously.[14] GBDC-SIGN and GB-DC-SIGNR were produced as described in **Section 2.1.4.1**, however, purification was instead performed using galactose-sepharose affinity column chromatography. Protein concentrations were obtained by UV-vis spectroscopy (**Equation 6.1**). Protein MWs were confirmed from deconvoluted HRMS data, where calculated protein MW values were obtained from the amino acid sequences of one monomeric unit of the corresponding protein.

GBDC-SIGN: HRMS: calculated MW [GBDC-SIGN] 39796.87; found 39787.74.

GBDC-SIGNR: HRMS: calculated MW [GBDC-SIGNR] 38105.71; found 38111.20.

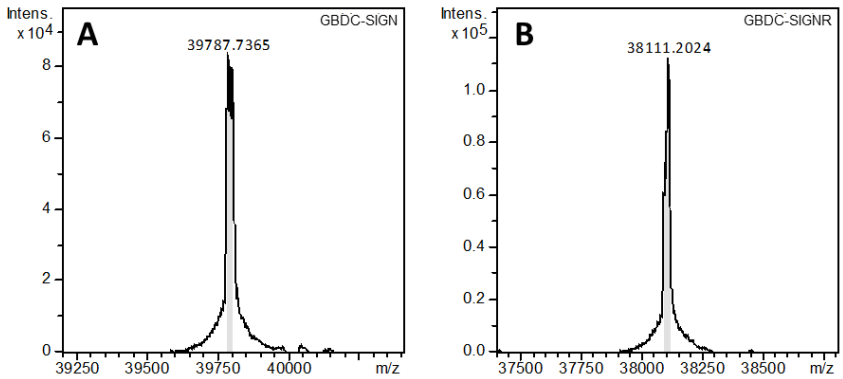

Figure 6.2 Deconvoluted HRMS of the monomeric extracellular domains of (A) DC-SIGN with a galactose binding motif (or GBDC-SIGN) and (B) GBDC-SIGNR.

6.2.3 HTDC-SIGN/R

DC-SIGN-Atto488 (339 µL, 2.66 µM in binding buffer, 0.9 nmol), DC-SIGN-Atto643 (400 µL, 2.25 µM in binding buffer, 0.9 nmol) and GBDC-SIGN (240

µL, 7.5 µM in binding buffer, 1.8 nmol) (or the DC-SIGNR analogues) were then incubated with guanidine·HCl (1.12 g, 11.7 mmol) on ice for 1h with gentle tilting. The mixture was then dialysed against binding buffer (4× 500 mL) before being purified by galactose-sepharose affinity column chromatography by washing with 2.5 column volumes (CV) of binding buffer, to remove any protein bearing ≤1 galactose binding site, followed by elution of the remaining protein with elution buffer. $CaCl_2$ (25 mM, final concentration) was then added to stabilise the protein and the eluted fractions were concentrated and purified by mannose-sepharose affinity column chromatography by washing with 2.5 CV of binding buffer, to remove any protein bearing ≤1 mannose binding site, followed by elution of the remaining protein with elution buffer. $CaCl_2$ (10 mM, final concentration) was then added to stabilise the protein. This yielded heterotetrameric (HT)DC-SIGN and HTDC-SIGN with ≤2 fluorophores per tetramer. Protein concentrations were obtained by UV-vis spectroscopy (**Equation 6.1**; where [P] is the monomeric protein concentration; A_λ is the absorbance maxima, *i.e.* A_{280}, A_{500} and A_{643} for protein, Atto488 and Atto643, respectively; and CF_{280} is the correction factor for each dye at 280 nm, *i.e.* 0.09, 0.04, 0.00 for Atto488, Atto643 and unlabelled protein, respectively; $\varepsilon_{\mathrm{P},280}$ is the protein extinction coefficient of each monomeric unit at 280 nm, *i.e.* 70400, 60890, 77390 and 67880 $M^{-1}\,cm^{-1}$ for DC-SIGN Q274C, DC-SIGNR Q27C, GBDC-SIGN and GBDC-SIGNR, respectively, where the number GB protein is assumed to be 2 per tetramer for HTDC-SIGN/R).

$$[\mathrm{P}] = \frac{A_{280} - \sum CF_{\mathrm{dye},280} \cdot A_{\mathrm{dye},\lambda}}{¼ \sum \varepsilon_{\mathrm{P},280}} \qquad 6.1$$

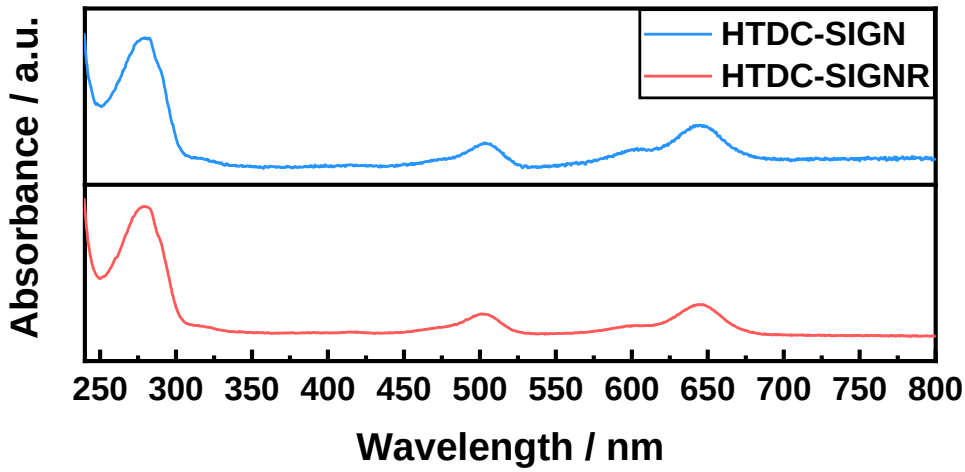

Figure 6.3 UV-vis spectra of heterotetrameric (HT)DC-SIGN consisting of a 1:1:2 monomer unit ratio of DC-SIGN-Atto488:DC-SIGN-Atto643:GBDC-SIGN (top) and HTDC-SIGNR consisting of a 1:1:2 monomer unit ratio of DC-SIGNR-Atto488:DC-SIGNR-Atto643:GBDC-SIGNR (bottom).

Protein MWs were confirmed from deconvoluted HRMS data, where calculated protein MW values were obtained from the amino acid sequences of one monomeric unit of the corresponding protein. The population of each monomer unit was obtained from relative proportion of the areas of each protein peak.

HTDC-SIGN: HRMS: calculated MW [DC-SIGN Q274C] 39172.22, [DC-SIGN Q274C + Ca + Na] 39235.29, found 39239.31 (13%); [DC-SIGN-Atto488] 39884.22, found 39888.28 (13%); [DC-SIGN-Atto643] 40130.22, found 40135.61 (15%); [GBDC-SIGN] 39796.87, found 39788.56 (59%).

HTDC-SIGNR: HRMS: calculated WM: [DC-SIGNR R287C] 37425.95, [DC-SIGNR R287C + 2Ca] 37506.10, found 37507.47 (12%); [DC-SIGNR-Atto488] 38137.95, [GBDC-SIGNR] 38105.71, found 38111.58 (superimposed, 77%); [DC-SIGNR-Atto643] 38383.95, found 38387.49 (12%).

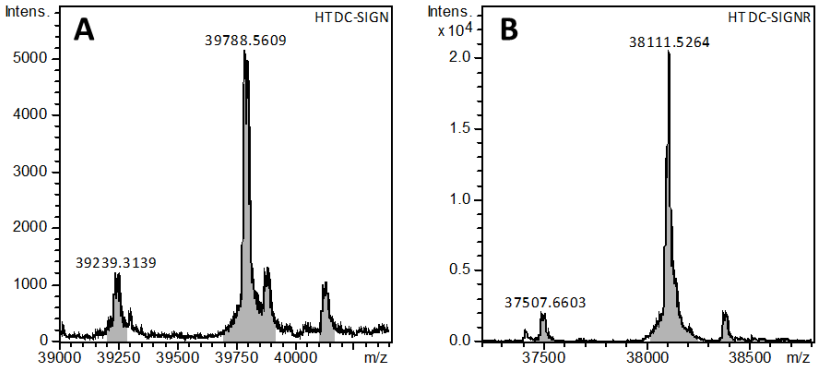

Figure 6.4 Deconvoluted HRMS of the monomer units corresponding to (A) heterotetrameric (HT)DC-SIGN consisting of a 1:1:2 monomer unit ratio of DC-SIGN-Atto488:DC-SIGN-Atto643:GBDC-SIGN and (B) HTDC-SIGNR consisting of a 1:1:2 monomer unit ratio of DC-SIGNR-Atto488:DC-SIGNR-Atto643:GBDC-SIGNR.

6.2.4 Site-Directed Spin-Labelling of DC-SIGN/R

DC-SIGN Q274C and DC-SIGNR R287C were produced and purified as described in **Section 2.1.4.2**. To the concentrated protein (1 eq), MTSSL, thiol-specific spin label (10 eq) was added and vortexed immediately before rotating at r.t. for 1 hr, and leaving overnight at 4 ^{o}C. The solution was then added to a mannose-sepharose affinity column, which had been pre-equilibrated in binding buffer, and washed with binding buffer until the excess dye was removed. The protein was then eluted using elution buffer and $CaCl_2$ (pH 7.4) was then added to each fraction to a final concentration of 10 mM. Protein concentrations were obtained by UV-vis spectroscopy (**Equation 2.1**). Protein MWs were confirmed from deconvoluted HRMS data, where calculated protein MW values were obtained from the amino acid sequences of one monomeric unit of the corresponding protein.

DC-SIGN-SL: HRMS: calculated MW [DC-SIGN Q274C] 39172.22, [DC-SIGN Q274C + 2Ca] 39252.40, [DC-SIGN Q274C-MTSSL] 39356.51; found, 39252.26 and 39361.06.

DC-SIGNR-Atto488: HRMS: calculated MW [DC-SIGNR R287C] 37425.95, [DC-SIGNR R287C + 2Ca] 37506.10, [DC-SIGNR R287C-MTSSL] 37610.24; found, 37507.28 and 37614.91.

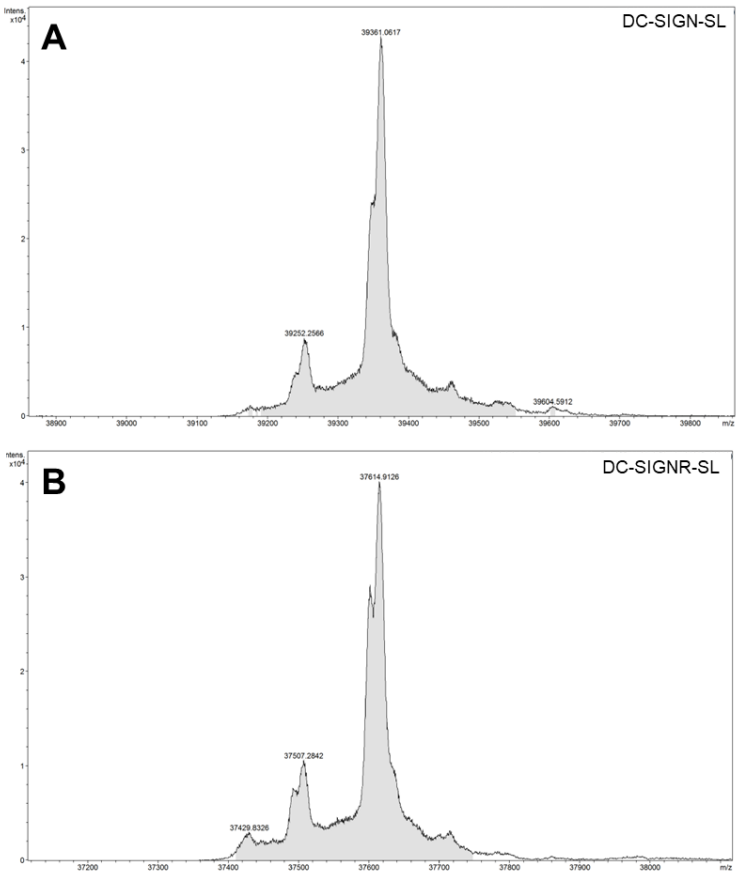

Figure 6.5 Deconvoluted HRMS of the monomer units corresponding to (A) DC-SIGN Q274C labelled with spin label MTSSL (DC-SIGN-SL) (B) DC-SIGNR R287C labelled with spin label MTSSL (DC-SIGNR-SL).

6.2.5 FRET Assays

FRET assays were performed as described in **Section 2.2.2.1**. Samples were prepared using HTDC-SIGN (50 nM) in binding buffer and control samples were prepared using DC-SIGN-Atto488 (12.5 nM), DC-SIGN-Atto643 (12.5 nM) and GBDC-SIGN (25 nM) in binding buffer. All fluorescence spectra were corrected by subtraction of a binding buffer background.

6.2.6 SmFRET

SmFRET samples were prepared by diluting either HTDC-SIGN (<0.1 µL, 330 nM), HTDC-SIGN (<0.1 µL, 1.26 µm) or DC-SIGN-Atto488 (<0.1 µL, 2.17 µM) and mixing it into glass wells with a total volume of 100 µL, containing BSA (final concentration 0.3 mg mL^{-1}), to reduce any non-specific binding of labelled protein to the walls, and GBDC-SIGN/R (final concentration 900 nM),

to avoid dissociation of the tetrameric protein at such low concentrations, in binding buffer. SmFRET was performed using an in-house setup (performed with the help of Bogachan Tahirbegi and Liming Ying, Imperial College London) samples were excited using an excitation wavelength of 488 nm and emission signals were collected at 520 and 643 nm.

6.2.7 Cw-EPR

Continuous wave (cw) electron paramagnetic resonance (cw-EPR) spectroscopy was performed using a Magnettech ESR5000 X-band cwEPR spectrometer (performed with the help of Yue Ma and Christos Pliotas, University of Leeds). Samples were prepared using 25 mM of protein in binding buffer. Measurements were performed simply by scanning the modulated magnetic field strength against a fixed frequency of 100 kHz and monitoring the excitation of the MTSSL spin label. From here the first derivative of the signal was determined. The linewidth ΔH_0 is obtained by the difference in the magnetic field strength obtained between the first derivative excitation peak maxima and minima.

6.2.8 Computational Methods

Molecular dynamics (MD) simulations were performed by Alexander St John in Emanuele Paci's group, University of Leeds. For DC-SIGNR, the tetrameric structure of CRDs with one coiled-coil repeat was taken from the PDB (1XAR).[13] The neck was extended to include a second coiled-coil repeat by overlapping it with a structure of the entire neck (3JQH).[14] For DC-SIGN, the sequence starting from the first residue of the final neck repeat and ending at the C-terminal tail was submitted to AlphaFold 2 Multimer. Q287C mutations for both proteins were introduced using Modeller version 10.1 and dyes were added manually.[15] For the MD, the proteins were described with the AMBER ff14SB force field,[16] except aspartates 114, 149, 160, 161 and glutamates 118, 141, 147, 148 whose carboxylate groups had charges scaled down to 85% in accord with the Electronic Continuum Correction with Rescaling (ECCR) approach.[17] Calcium, potassium and chloride ions were treated with ECCR, its parameters were taken from ref. [18]; the water model used was TIP4P-2005.[18] The dye and linker parameters were taken from AMBER-DYES database in AmberTools20.[19] Atto488 dye used in the FRET experiments as donor was assigned T48 and C3R units with correct sp3 carbon in the maleimide ring linking to the Cys as R-isomer. The acceptor dye Atto643 of

an unknown structure but presumably with two sulphate groups was modelled for simplicity as Atto488. Hydrogens were added to the systems using LEaP module of AMBER20. The systems were immersed in an octahedral box of water molecules,[18] extending up to 12 Å from the protein. Counterions were added to neutralize the charge and match the final concentration of 0.15 M. Stepwise relaxation was carried out according to the published protocols,[20] except that the length of the production MD was 1μs. The 6 Ca-Ca distances (4 Side, 2 Diag) monitored became relatively stable from 500 ns onward and thus the analyses were done on the last 0.5 μs. Five replica simulations were performed for each protein to give a total of 2.5 μs total simulation time for analysis.

The inter-dye and inter-Cys distances (r) were calculated from the distance between the central oxygen atom within each dye during the simulations and converted to FRET efficiencies (E) using **Equation 6.2** where R_0 = 4.8 nm. FRET efficiency histograms were calculated and fitted to the corresponding experimentally determined histograms by reweighting MD snapshots. Reweighting was performed by taking 200 random snapshots from the trajectory and minimising the least squares difference, using the SLSQP algorithm, between the resulting FRET efficiency histogram with the experimentally determined FRET efficiency histogram. Reweighting was performed 10 times independently for both DC-SIGN and DC-SIGNR before calculating the average dye-dye and Cys-Cys distances using the reweighted snapshots.

6.3 Results and Discussion

6.3.1 Double purification strategy for FRET pair labelling DC-SIGN/R

In order to site-specifically label only two out of four of the CRDs of DC-SIGN/R we have designed heterotetrameric analogues of their extracellular domains which comprise of three different monomer units: 2 labelled units and 1 spacer unit. The two labelled units were produced as described previously,[10] by a single site-specific cysteine mutation in proximity of the binding site of the CRD, yielding DC-SIGN Q274C (or DC-SIGNR R287C). This mutation was chosen as there are no other free cysteine residues on the protein and because this site is well documented to retain DC-SIGN's mannose binding specificity.[8-10] These homotetramers were then purified by mannose affinity column chromatography and labelled *via* thiol-maleimide coupling with either a maleimide-Atto488 or a maleimide-Atto643 dye, yielding DC-SIGN-Atto488

and DC-SIGN-Atto643 (λ_{abs} = 500 and 643 nm and λ_{em} = ~520 and ~665 nm, respectively; Atto488···Atto643 FRET pair Förster radius, R_0, = 4.8 nm; labelling efficiency 63-74% by HRMS). The spacer monomer unit was made using a mutated variant of DC-SIGN whereby the mannose binding motif was replaced with that of a galactose binding protein (galactose binding (GB)DC-SIGN), and was left unlabelled.

The three homotetrameric proteins were then mixed in a 1:1:2 ratio of DC-SIGN-Atto488:DC-SIGN-Atto643:GBDC-SIGN, and were unfolded into their monomer constituents with guanidine, before being refolded by dialysis into a distribution of homo and heterotetramers (**Figure 6.6A i,ii**). In order to ensure that each heterotetramer contained only two dyes, alternating mannose and galactose affinity columns were used to purify the protein (**Figure 6.6A iii,iv**). Because the binding of DC-SIGN is highly dependent on multivalent Ca^{2+}-dependent sugar interactions, by passing the protein through a galactose affinity column in the presence of Ca^{2+}, only proteins with >1 galactose-binding CRD will adhere to the column and the rest will elute off. The protein which is retained on the column can then be collected by washing with EDTA which removes the Ca^{2+}. This process can then be repeated using a mannose affinity column in order to remove any tetramers with >3 galactose-binding CRDs. The resulting product is therefore solely comprised of heterotetramers containing two labelled mannose-binding monomer units and two spacer galactose-binding units (herein termed heterotetrameric (HT)DC-SIGN), as demonstrated by HRMS, **Figure 6.4**), with no more than 2 dye molecules present per tetramer, fulfilling the criteria for smFRET measurements. This process was also repeated for DC-SIGNR using DC-SIGNR-Atto488, DC-SIGNR-Atto643 and GBDC-SIGNR, yielding HTDC-SIGNR.

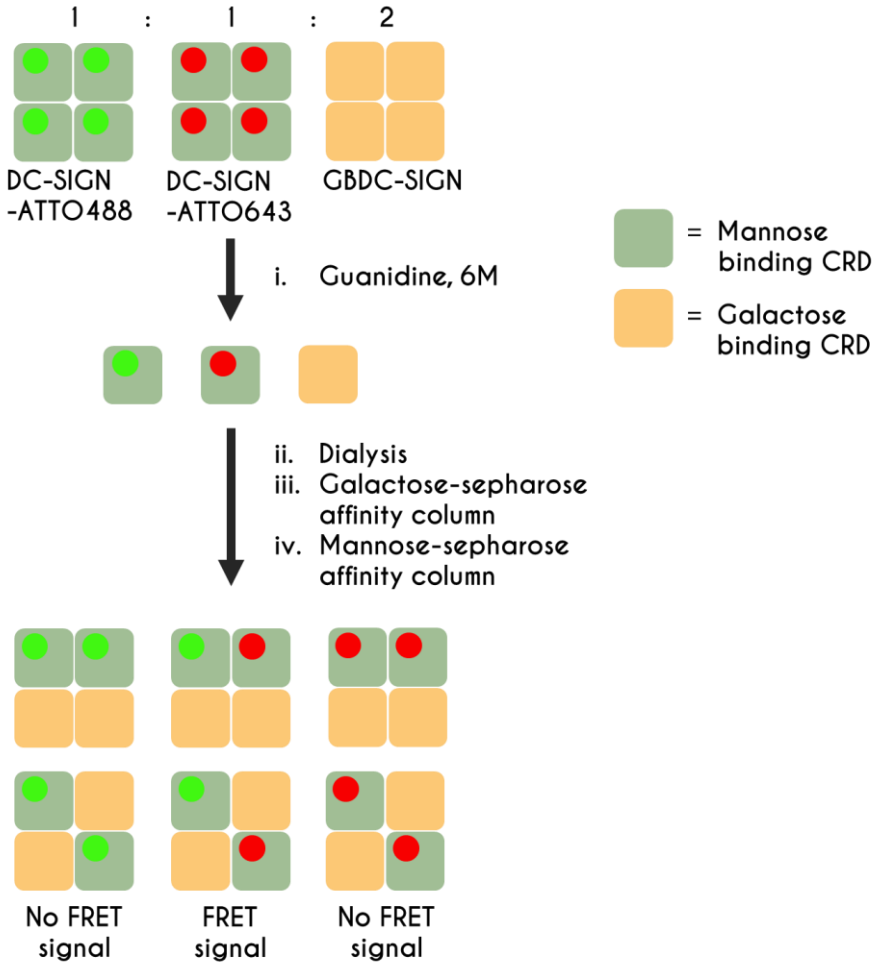

Figure 6.6 A schematic illustrating the double purification strategy used to label two out of four of the CRDs of DC-SIGN/R. (i) A 1:1:2 ratio of DC-SIGN-Atto488 : DC-SIGN-Atto643 : GBDC-SIGN tetramers were unfolded into monomer units the presence of guanidine·HCl (6 M) and (ii) Renatured by dialysis to remove guanidine, allowing for the formation of a distribution of different heterotetramers. (iii) A galactose-sepharose affinity column was then used to remove tetramers with one or less galactose-binding CRDs due to their poor affinity. (iv) This was repeated with a mannose-sepharose affinity column to remove tetramers containing one or less mannose-binding CRDs. This yielded heterotetramers containing 2 labelled mannose binding CRDs and 2 galactose binding CRDs, labelled on either adjacent or diagonal CRDs, where only those containing both donor (Atto488) and acceptor (Atto643) fluorophores can participate in FRET.

Such constructs would inevitably consist of a distribution of 6 arrangements: 4 whereby the labelled CRDs are adjacent to one another (**Figure 6.7 left**) and 2 whereby the labelled CRDs are diagonal to one another (**Figure 6.7 middle**). Thus allowing for the six inter-CRD arrangements to be derived (**Figure 6.7 right**).

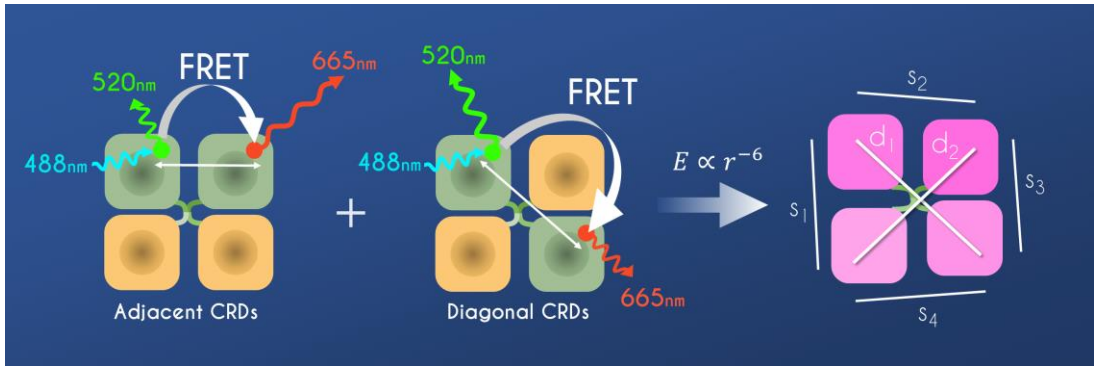

Figure 6.7 A schematic representing the two FRET modes that can be determined by smFRET for DC-SIGN with two out of four CRDs labelled with an ATTO488···ATTO643 FRET pair and excited with a 488 nm excitation beam: that between adjacent CRDs (left) and diagonal CRDs (middle). By extracting the resulting FRET efficiencies of the different FRET pair configurations, distances for the 4 side and 2 diagonal inter-CRD arrangements can be obtained *via* Equation 6.2 (right)

In order to test the FRET response of HTDC-SIGN and HTDC-SIGNR (herein referred to collectively as HTDC-SIGN/R), a simple fluorescence spectroscopic scan was performed on 50 nM of the proteins before and after heterotetramerisation using an excitation wavelength of 480 nm (**Figure 6.8**). Both heterotetramers successfully demonstrated the presence of FRET as indicated by the observed emission at 660 nm corresponding to the fluorescence of the Atto643 acceptor fluorophore, with no observable acceptor emission prior to heterotetramerisation. In addition, HTDC-SIGN appears to demonstrate a much more significant degree of FRET in comparison to HTDC-SIGNR (e.g. the ratio of acceptor to donor fluorescence intensities, I_{660}/I_{521}, for HTDC-SIGNR is ~50% that of HTDC-SIGN). This is an indicator that the average distances between the CRDs of DC-SIGNR may be almost double that of DC-SIGN, which has been predicted by various model systems in the past.[13, 21] Though, it is worth noting that these FRET signals cannot be used to obtain an average inter-dye distance due to the significant proportion of HTDC-SIGN/R that will possess only donor containing fluorophore which contribute significantly to the donor fluorescence intensity signal.

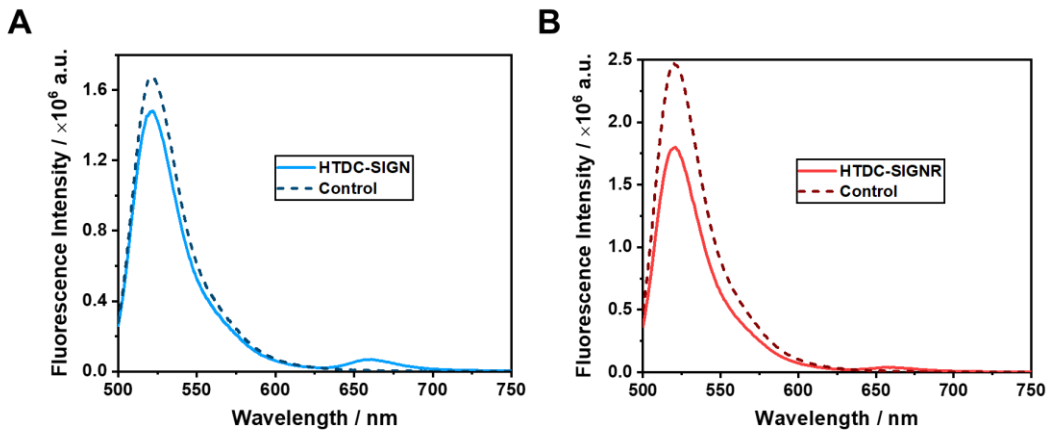

Figure 6.8 Background corrected fluorescence spectra of (A) HTDC-SIGN (solid line) and a control sample containing a 1:1:2 mixture of DC-SIGN-Atto488:DC-SIGN-Atto643:GBDC-SIGN in their homotetrameric forms (dashed line); and (B) HTDC-SIGNR (solid line) and a control sample containing a 1:1:2 mixture of DC-SIGNR-Atto488:DC-SIGNR-Atto643:GBDC-SIGNR in their homotetrameric forms (dashed line).

6.3.2 Inter-CRD Distance Determination *via* smFRET

SmFRET measurements were taken using an excitation beam of 488 nm and fluorescence emission detection channels at 520 nm and 665 nm. In order to minimise the possibility of multiple proteins passing through the confocal volume at once, HTDC-SIGN/R samples were made to pM concentrations in buffer containing by BSA (0.3 mg mL^{-1}), to reduce any non-specific binding of labelled protein to the well-plate walls, as well as an excess of non-fluorescent GBDC-SIGN/R (900 nM), to avoid dissociation of the tetrameric protein at such low concentrations (though we have previously shown these tetramers are stable even at low pM concentrations).[22] During an event whereby a FRET pair traverses the excitation volume, the Atto488 donor is excited and transfers a fraction of its energy to the Atto643 acceptor, the resulting fluorescence emitted by both fluorophores are then detected by the donor and acceptor channels. The FRET efficiency between the two fluorophores can then be obtained as the ratio between the intensity of the acceptor fluorescence signal and the total intensity from both channels. A distribution of all the FRET efficiencies measured is then obtained using a histogram which can be fitted using multiple Gaussian fits to obtain values for the average FRET efficiencies. The average distances between the dyes can then be extracted using **Equation 6.2** (where E is the FRET efficiency; I is the fluorescence signal intensity for donor fluorophore (D) or acceptor fluorophore (A); r is the D-A distance; and R_0 is the Förster radius).[23]

$$E = \frac{I_A}{I_A + I_D} = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6} \quad 6.2$$

Each set of measurements were repeated at least 3 times for HTDC-SIGN, HTDC-SIGNR and a DC-SIGN-Atto488 control, and normalised by the total number of runs performed for each sample. Because of the relatively low signal to noise ratio, a threshold fluorescence intensity is required in order to filter out the "true" FRET events. This was done by ensuring that only events with a sum of acceptor and donor intensities larger than 15 a.u. were counted (*i.e.* sum threshold, $T_{sum} = I_A + I_D = 15$). Though this removed the majority of background noise, such thresholds still detect a significant number of fluorescence events caused by proteins containing only donor fluorophore which fall above T_{sum}. These donor events can be filtered out from the "true" FRET events by simply implementing a second threshold just for the acceptor channel (T_A). By analysing the number of events detected by the DC-SIGN-Atto488 control over increasing T_A (**Figure D.1**), it can be observed that a threshold of T_A = 7 is required to remove at least 99% of donor only fluorescence signal (**Figure D.2**). By then applying thresholds T_{sum} = 15 and T_A = 7 a.u. to the HTDC-SIGN and HTDC-SIGNR signals a representative histogram of "true" FRET efficiency distribution can be obtained (**Figure 6.9**). Before analysing the histograms, it can already be observed that the FRET efficiencies of DC-SIGN are skewed towards the lower inter-dye distances (higher FRET efficiencies) unlike DC-SIGNR which is skewed towards the higher inter-dye distances (low FRET efficiencies). This compliments the bulk FRET, demonstrating further evidence towards the more condensed inter-CRD arrangement of DC-SIGN as proposed by previous models.[13, 21] These histograms can then be fitted with a multimodal Gaussian fit in order to extract the mean FRET efficiencies corresponding to the inter-dye distances of 3.6 ± 0.7 nm, 4.4 ± 0.2 nm, 5.1 ± 0.2 nm and 6 ± 3 nm for DC-SIGN which match well with the proposed inter-binding site distances.[21] For DC-SIGNR, inter-dye distances of 4.7 ± 0.4 nm, 5.01 ± 0.03 nm, 5.6 ± 0.1 nm observed which appear to be in between the longer and shorter adjacent inter-binding site distances observed by previous models.[13] This comparison highlights the limitations of applying both T_{sum} and T_A thresholds. Though using high T_A should filter out any donor only fluorescence events, it also causes any low FRET efficiency events where I_A falls below T_A will be excluded from the distribution. This means that as the inter-dye distances gets larger, the probability of an event being bright enough to render $I_A > T_A$ decreases significantly. This is very evident in for HTDC-SIGN at a T_A of 7 a.u. (**Figure**

6.9A; Table D.1) where the number of events counted for the FRET efficiencies at around 0.2 is much lower than those at 0.4 or above. This means that any higher inter-dye distances for the two proteins cannot be resolved using this method.

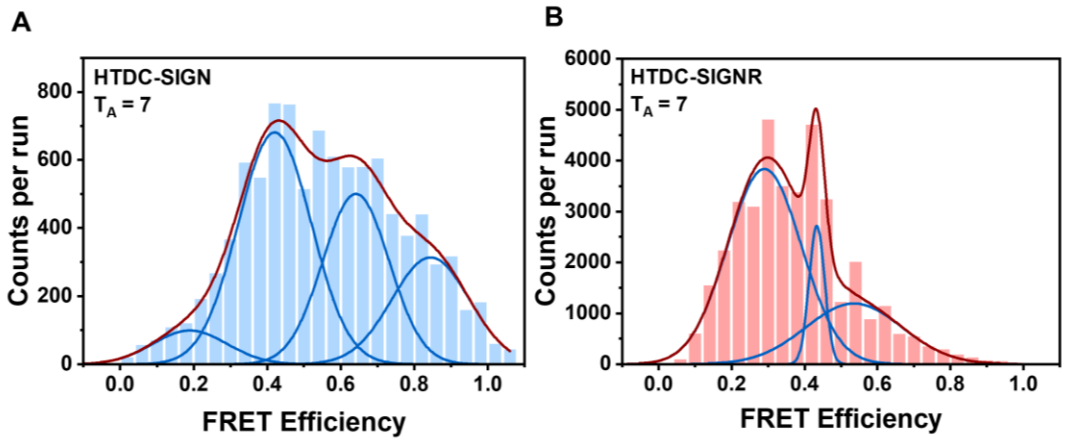

Figure 6.9 histograms of the population of FRET efficiency measurements detected by smFRET using a acceptor and donor sum threshold, T_{sum}, of 15 a.u. and an acceptor threshold, T_A, of 7 a.u. for (A) HTDC-SIGN and (B) HTDC-SIGNR.

In order to scan for these larger inter-dye distances, the FRET efficiency distributions were therefore investigated over a range of T_A thresholds and fitted separately in the same way as above (**Figure D.3Figure** D.4**; Table D.2**). In order to then account for the large number of "false" FRET events that are admitted from those protein containing only the donor fluorophore, the mean FRET efficiencies at each T_A were compared against those observed in the donor only sample. Mean FRET efficiencies that were shown to correlate over multiple T_A but did not correlate with the donor only sample were then grouped into six sets. The mean FRET efficiencies in each set were then averaged in order to determine six approximate inter-dye distances corresponding to the four adjacent and two diagonal distances between the four CRDs.

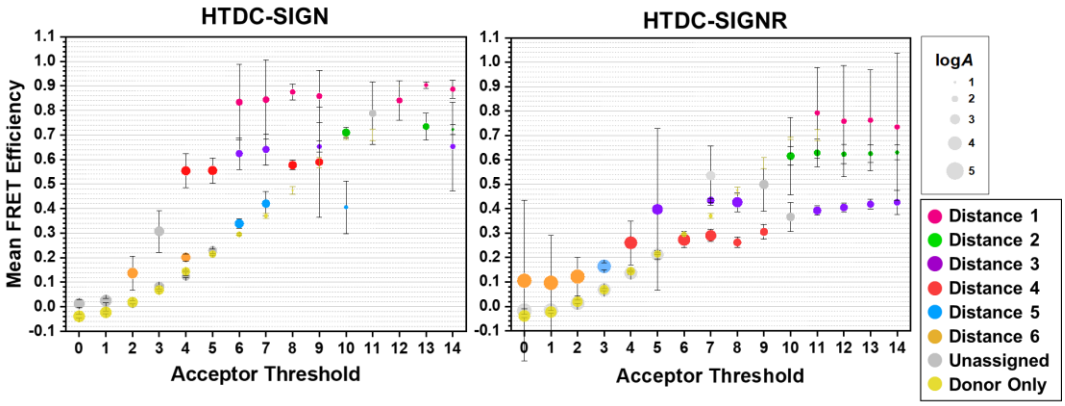

Figure 6.10 A comparative analysis of mean apparent FRET efficiencies measured by smFRET for HTDC-SIGN (left) and HTDC-SIGNR (right) compared to a donor only control sample (yellow) over increasing acceptor threshold using a constant threshold for the sum of the acceptor and donor intensities of 15 a.u.. The log of the integrated area of each distribution ($\log A$) is represented via the size of the data point. Six sets of data points corresponding to the six possible inter-dye distances were identified (highlighted by different colours). Any data points which overlapped with donor only data points or were likely corresponded to an average of multiple data sets were not assigned to any set (grey).

Table 6.1 Summary of average FRET efficiencies and distances obtained by comparing the apparent FRET efficiencies derived from smFRET of HTDC-SIGN and HTDC-SIGNR against a donor only control.

	DC-SIGN		DC-SIGNR	
Inter-dye distance	**FRET efficiency**	**Distance / nm**	**FRET efficiency**	**Distance / nm**
r_{s1}	0.89 ± 0.01	3.36 ± 0.07	0.8 ± 0.1	3.9 ± 0.4
r_{s2}	0.72 ± 0.01	4.11 ± 0.05	0.62 ± 0.02	4.41 ± 0.05
r_{s3}	0.65 ± 0.02	4.33 ± 0.06	0.424 ± 0.006	5.05 ± 0.02
r_{s4}	0.57 ± 0.02	4.56 ± 0.05	0.28 ± 0.01	5.62 ± 0.06
r_{d1}	0.35 ± 0.02	5.32 ± 0.07	0.16 ± 0.02	6.3 ± 0.1
r_{d1}	0.20 ± 0.01	6.04 ± 0.09	0.12 ± 0.07	6.7 ± 0.8

6.3.3 Comparison of Inter-CRD Distances *via* Molecular Dynamics

To gauge the sensitivity of the dyes to confer detailed information on the 3D arrangement of CRDs, the inter-CRD distances were measured between the Cys residues of DC-SIGN/R, where the dyes had been covalently tethered, as well as the dyes themselves. Interestingly, a smaller spread of distances was observed in the former compared to the latter (**Figure 6.11**). This points to a large flexibility of the linkers tethering the dyes as opposed to the relative

stability of the positions of the CRDs. Though this may reflect inherent differences in flexibility between the dye linkers and CRDs, it may also be a limitation of insufficient MD sampling. As the amount of additional sampling may not be a matter of few-fold, but rather several orders of magnitude, the sampling was instead weighted towards those conformations which contribute the most to the observed experimental FRET results. To achieve this, FRET efficiencies based on the six dye-dye distances (4 side (s), 2 diagonal (d)) derived from the MD simulations, were compared against those observed by smFRET efficiency histograms over the different T_A thresholds measured. The contributions of the individual MD snapshots were then reweighted with the condition of maximising the agreement with the experiment. From here, average inter-dye and inter-Cys distances were then recalculated from the reweighted MD snapshots for at each threshold (summarised in **Table D.3Table** D.4).

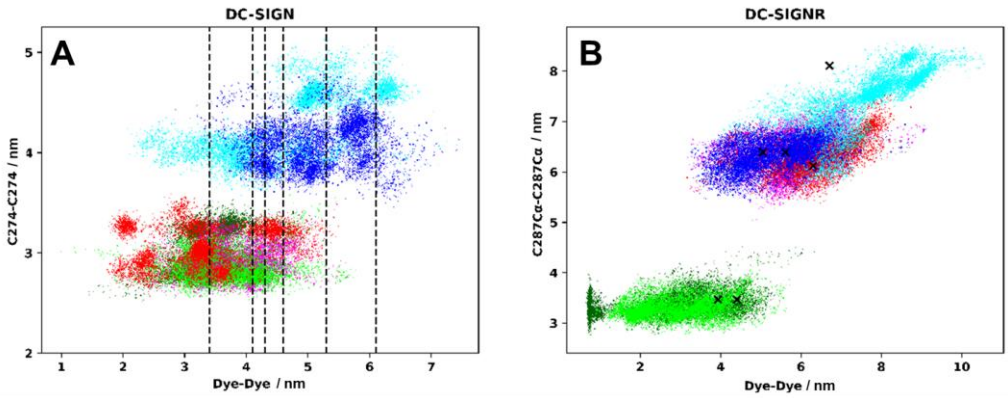

Figure 6.11 Distribution of Cys:Cα…Cys:Cα *vs* dye:O8…dye:O8 distances from MD of DC-SIGN/R. Black crosses/dashed lines denote experimental values recorded in Table 6.1.

The best agreements between the MD FRET efficiencies and the histograms obtained by smFRET for DC-SIGN were observed at thresholds between T_A = 11-14 (providing a residual sum of squares, RSS, values of 0.85-1.02). Moreover, the agreement was found to become worse towards lower T_A (*e.g.* RSS ~4-7 for T_A = 0-4). This may be due to the large uncertainties in accuracy of the larger empirically obtained inter-dye distances, attributed to the limit of detection achievable given the R_0 of the dyes used here, as well as the large contribution of "false" FRET events from protein containing only the donor fluorophore. This will skew the best fit in favour of thresholds with poor resolution of the larger distances (*e.g.* at larger T_A). Interestingly, despite this, DC-SIGNR showed best agreement at a T_A of ~7 (RSS = 1.2). This could be because of the better informed model of DC-SIGNR which already has XRC data relating to a dimer of CRDs,[13] thus the MD sampling fits well even when more of distances are resolved. Saying this, though DC-SIGN favoured higher thresholds, the RSS scores were still reasonable at T_A > 6 for DC-SIGN (*e.g.* RSS = 1.52 at a T_A of 7). As these thresholds resolve more of the FRET

efficiencies, it may be more informative to compare the calculated distances between the MD and the smFRET at these distances.

Table 6.2 Average inter-dye and inter-domain distances calculated from reweighted MD snapshots. Reweighting was performed to fit snapshots to the experimental acceptor threshold of 7 (RSS = 1.52 and 1.2 for DC-SIGN and DC-SIGNR, respectively).

	DC-SIGN		DC-SIGNR	
Distance	**Dye-Dye / nm**	**C274-C274 / nm**	**Dye-Dye / nm**	**C287-C287 / nm**
s1	3.49	3.13	4.19	3.88
s2	3.92	3.1	5.01	5.25
s3	4.17	2.87	5.07	5.37
s4	4.21	2.81	5.33	5.65
d1	5.4	4.43	5.26	6.07
d2	5.53	3.96	5.72	6.11

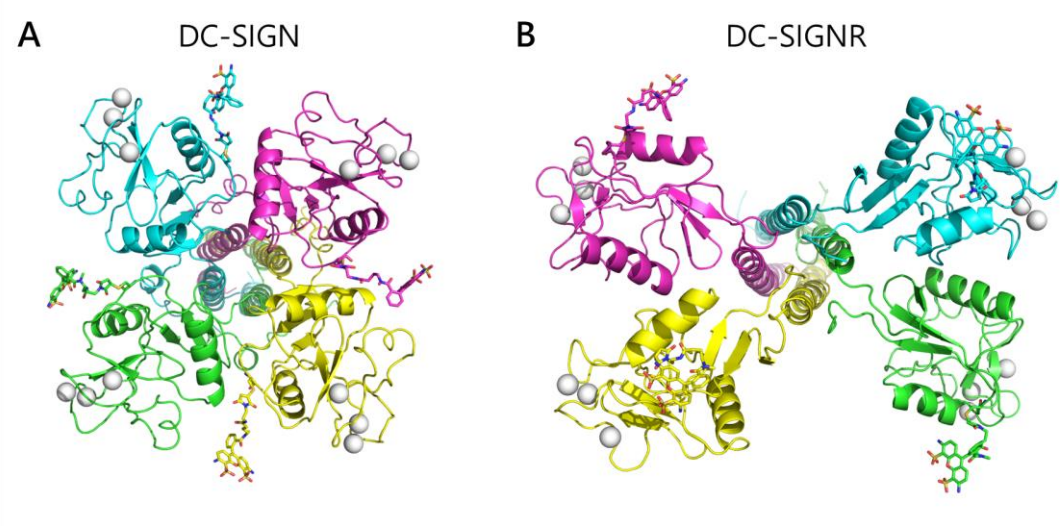

Figure 6.12 Example MD simulation model reweighted towards the smFRET histogram obtained using T_{sum} = 15 and T_A = 6, for (A) DC-SIGN and (B) DC-SIGNR.

By plotting the inter-dye distances derived by MD reweighted sampling at a T_A = 7 (**Table 6.2**) against those derived by smFRET over all thresholds (**Table 6.1**) it can be seen that, for DC-SIGN in particular, a relatively acceptable match can be obtained (**Figure 6.13**). This is incredibly encouraging, particularly because the tetrameric structure of DC-SIGN is completely unresolved outside of MD simulations. DC-SIGNR also shows

reasonable correlation, however, this is skewed towards larger distances in the smFRET. This is either due to the MDs under-predicting these longer distances or to the large uncertainties in the in the smFRET at this distance range. Despite this, these data provide a more empirical model which contributes towards the effort of estimating the inter-binding site distances of DC-SIGN and DC SIGNR, providing to further evidence for a more compact DC-SIGN structure and a more open DC-SIGNR structure (**Figure 6.12**). These data can therefore be used in the design of more specific glycoconjugates which better match these binding site distances to enhance therapeutic specificity and potency towards these lectins.

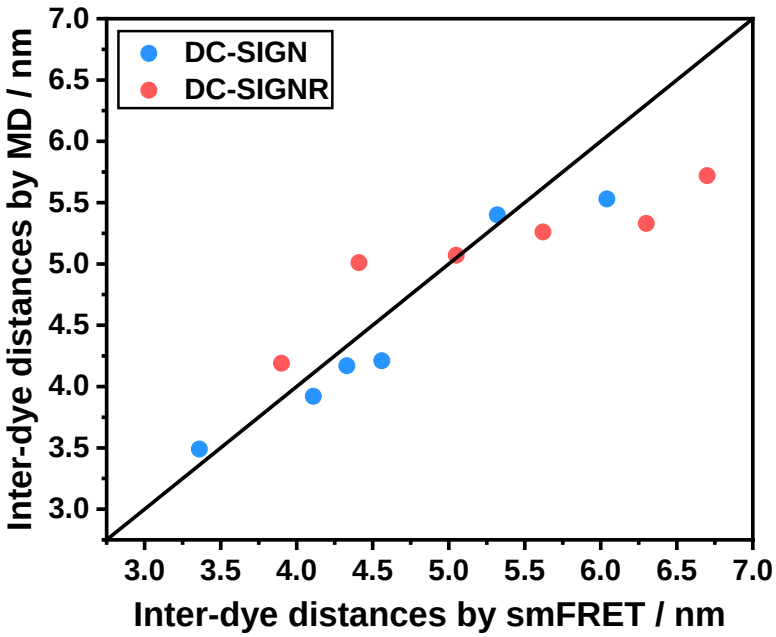

Figure 6.13 Comparison of the inter-dye distances obtained by MD sampling, reweighted towards the smFRET histogram obtained using T_{sum} = 15 and T_A = 7, against the inter-dye distances obtained by smFRET compared across a range of T_A thresholds. (black diagonal demonstrates $x = y$, *i.e.* perfect correlation).

The limited correlation between the MD sampling and the smFRET may also be caused by greater variations in the dynamics of the molecule. This is observable in **Figure 6.11** which demonstrates a wide spread of inter-Cys distances (*i.e.* ±~1 nm) for each of the six distances in DC-SIGNR, compared to only ±~0.5 nm for DC-SIGN. This demonstrates much more flexibility either between the CRDs or the peptide loop bearing the free Cys for DC-SIGNR compared to DC-SIGN. This flexibility was confirmed empirically using continuous wave EPR (cw-EPR) spectroscopy by site-directed labelling of DC-SIGN Q274C and DC-SIGN R287C with a nitroxide spin label. Here the mobility of the spin label effects the broadness of magnetic field strength required to excite its unpaired electrons into their excited states. For

nitroxides, this results in a hyperfine triplet splitting pattern whereby the linewidth of the splitting can be used to qualitatively describe the mobility of the label.

Here the linewidth of the central peak (ΔH_0) can be measured as the ΔB_0 between the maxima and minima of the first derivative absorbance signal. This revealed ΔH_0 of 3.45 mT for DC-SIGN and 4.74 mT for DC-SIGNR (**Figure 6.14**), demonstrating that the labels on DC-SIGNR do indeed possess more flexibility compared to that of DC-SIGN.

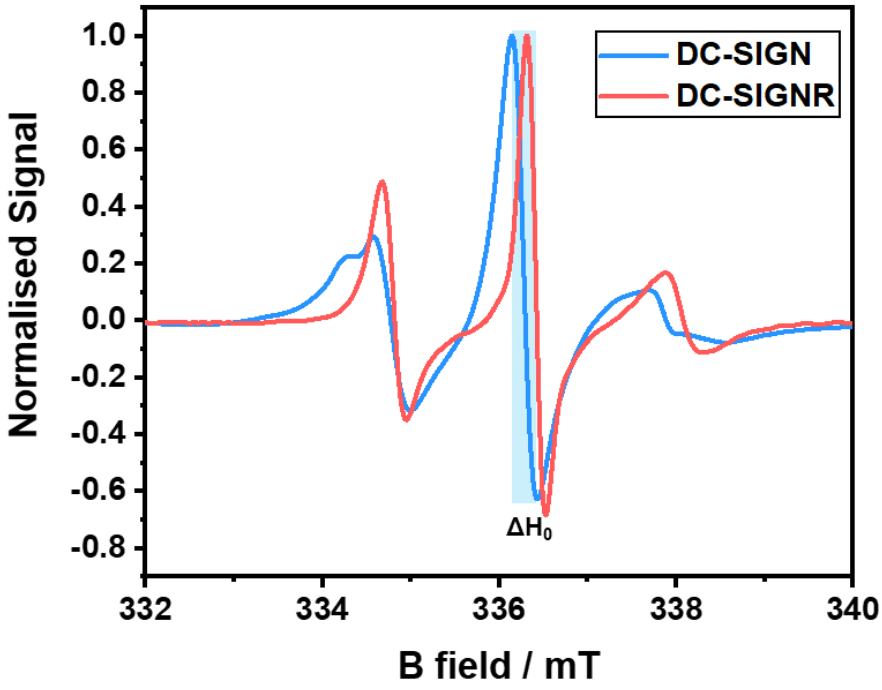

Figure 6.14 Plot of the hyperfine triplet splitting patterns corresponding to the first derivative signals of the interaction of nitroxide spin labelled DC-SIGN (blue) and DC-SIGNR (red) with a fixed wavelength over the magnetic field strength (B) using modulating magnetic field obtained by cw-EPR. ΔH_0 denotes the central linewidth, calculated as the difference in B between the signal maxima and minima for each protein.

6.4 Conclusions

In summary, we have described a new approach to provide reasonable estimations of the inter-carbohydrate recognition domain (CRD) distances of multimeric glycan-binding protein through a combination of empirical and computational techniques using an example pair of closely related tetrameric

proteins, DC-SIGN and DC-SIGNR. We have successfully demonstrated a heterotetramerisation and double purification approach to labelling two out of four of the CRDs of these lectins with a FRET pair of fluorophores, which have allowed us to measure the six inter-CRDs using smFRET. These distances were compared against MD constructs of the tetrameric protein structure predicted XRC and alpha-fold structures. This demonstrated reasonable agreement albeit with limited accuracy at the higher inter-binding site distances due to the limitations of the FRET pair Förster radius. Here we have been able to confirm that DC-SIGN and DC-SIGNR provide different CRD arrangements, whereby DC-SIGN shows a more compact structure compared to DC-SIGNR. We therefore hypothesise that, due to the similarity in the monovalent mannose motifs of DC-SIGN and DC-SIGNR, these differences in tetrameric CRD arrangement may provide a structural rationale for the differences in multivalent glycan-binding mode and viral specificity. Furthermore, the broader tetrameric arrangement of the CRDs of DC-SIGNR agrees well with its ability to crosslink with certain multivalent glycan constructs,[8-10] where the more compact arrangement of DC-SIGN allows it to more easily form strong tetrameric binding to individual multivalent glycans of the same size. In addition, we have also shown by both MD simulations and cw-EPR measurements that DC-SIGNR displays more mobility compared to DC-SIGN at the 287/274 positions, respectively. This suggests that there may be a higher degree of inter-CRD flexibility for DC-SIGNR in comparison to DC-SIGN which may contribute to DC-SIGN's specificity in forming simultaneous binding of all four CRDs.

It is worth noting that the biggest limitation of this smFRET technique comes from significant inverse sixth power relationship between the FRET efficiency and the inter-dye distance. Though this could be overcome by using a FRET pair with a higher R_0, another way of obtaining more accurate measurements of these longer distances may be to employ EPR techniques such as pulse electron-electron double resonance (PELDOR; or DEER) spectroscopy, which, although requires larger sample sizes and cryogenic conditions, demonstrates only an inverse cubic relationship between dipole coupling of spin labels to the inter-spin label distance, thus may be used to scan longer distances.

Despite this however, the distances obtained here are highly valuable and have been able to provide a better informed model of the inter-CRD arrangements of these two proteins. These models therefore provide the important structural details necessary to better inform the rational design of

glycoconjugates with complimentary inter-glycan distances which can more efficiently exploit the benefits of multivalency to provide very specific and potent targeting of these lectins for therapeutic applications.

References

1. N. Rahimi, *Biology (Basel)*, 2020, **10**, 1.
2. T. B. Geijtenbeek, D. S. Kwon, R. Torensma, S. J. van Vliet, G. C. van Duijnhoven, J. Middel, I. L. Cornelissen, H. S. Nottet, V. N. KewalRamani, D. R. Littman, C. G. Figdor and Y. van Kooyk, *Cell*, 2000, **100**, 587-597.
3. C. P. Alvarez, F. Lasala, J. Carrillo, O. Muniz, A. L. Corbi and R. Delgado, *J. Virol.*, 2002, **76**, 6841-6844.
4. S. Pöhlmann, E. J. Soilleux, F. Baribaud, G. J. Leslie, L. S. Morris, J. Trowsdale, B. Lee, N. Coleman and R. W. Doms, *Proc. Natl. Acad. Sci U. S. A.*, 2001, **98**, 2670-2675.
5. R. Amraei, W. Yin, M. A. Napoleon, E. L. Suder, J. Berrigan, Q. Zhao, J. Olejnik, K. B. Chandler, C. Xia, J. Feldman, B. M. Hauser, T. M. Caradonna, A. G. Schmidt, S. Gummuluru, E. Mühlberger, V. Chitalia, C. E. Costello and N. Rahimi, *ACS Cent. Sci.*, 2021, **7**, 1156-1165.
6. H. Feinberg, D. A. Mitchell, K. Drickamer and W. I. Weis, *Science*, 2001, **294**, 2163-2166.
7. Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis and K. Drickamer, *Nat. Struct. Mol. Biol.*, 2004, **11**, 591-598.
8. D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo and D. Zhou, *J. Am. Chem. Soc.*, 2020, **142**, 18022-18034.
9. Y. Guo, I. Nehlmeier, E. Poole, C. Sakonsinsiri, N. Hondow, A. Brown, Q. Li, S. Li, J. Whitworth, Z. Li, A. Yu, R. Brydson, W. B. Turnbull, S. Pohlmann and D. Zhou, *J. Am. Chem. Soc.*, 2017, **139**, 11833-11844.
10. Y. Guo, C. Sakonsinsiri, I. Nehlmeier, M. A. Fascione, H. Zhang, W. Wang, S. Pohlmann, W. B. Turnbull and D. Zhou, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 4738-4742.
11. E. Lerner, T. Cordes, A. Ingargiola, Y. Alhadid, S. Chung, X. Michalet and S. Weiss, *Science*, 2018, **359**, eaan1133.
12. B. Martinac, *Prog. Biophys. Mol. Biol.*, 2017, **130**, 192-197.
13. H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327-1335.
14. H. Feinberg, C. K. Tso, M. E. Taylor, K. Drickamer and W. I. Weis, *J. Mol. Biol.*, 2009, **394**, 613-620.
15. N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M.-y. Shen, U. Pieper and A. Sali, *Curr. Protoc. Bioinformatics*, 2006, **15**, 5.6.1-5.6.30.
16. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696-3713.
17. B. J. Kirby and P. Jungwirth, *J. Phys. Chem. Lett.*, 2019, **10**, 7531-7536.
18. J. L. F. Abascal and C. Vega, *J. Chem. Phys.*, 2005, **123**, 234505.
19. B. Schepers and H. Gohlke, *J. Chem. Phys.*, 2020, **152**, 221103.
20. P. Srb, M. Svoboda, L. Benda, M. Lepšík, J. Tarábek, V. Šícha, B. Grüner, K. Grantz-Šašková, J. Brynda, P. Řezáčová, J. Konvalinka and V. Veverka, *Phys. Chem. Chem. Phys.*, 2019, **21**, 5661-5673.
21. G. Tabarani, M. Thepaut, D. Stroebel, C. Ebel, C. Vives, P. Vachette, D. Durand and F. Fieschi, *J. Biol. Chem.*, 2009, **284**, 21229-21240.
22. B. Tahirbegi, A. J. Magness, M. E. Piersimoni, X. Teng, J. Hooper, Y. Guo, T. Knöpfel, K. R. Willison, D. R. Klug and L. Ying, *Front. Chem.*, 2022, **10**.
23. T. Förster, *Ann. Phys.*, 1948, **437**, 55-75.

Chapter 7
A Polyvalent Nano-Lectin Potently Neutralises SARS-CoV-2 by Targeting Glycans on the Viral Spike Protein

The following chapter demonstrates a collaborative effort between the group of Dr Yuan Guo and Professor Dejian Zhou (University of Leeds) and the group of Professor Dr Stefan Pöhlmann (German Primate Centre, Göttingen, Germany). The author of this thesis prepared and characterised the gold nanoparticles (GNPs), functionalised protein and lectin-GNP glycoconjugates, and performed and analysed fluorescence quenching assays. Dr Darshita Budhadev synthesised and characterised the ligands used for lectin functionalisation and the glycan ligands, and co-led the protein functionalisation and lectin-GNP complexation and characterisation. All lectin plasmids were prepared in-house by Dr Yuan Guo. Pseudotyped virus preparation and inhibition studies and all work with infectious SARS-CoV-2 was conducted by Inga Nehlmeier, Amy Madeleine Kempf, Dr Markus Hoffmann, Nadine Krüger and Professor Dr Stefan Pöhlmann.

7.1 Introduction

The global coronavirus disease 2019 (COVID-19) pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), had a devastating impact on the healthcare systems and economies worldwide. The excess mortality associated with COVID-19 is believed to amount to 18-million from 2020-2021,[1] and tens of millions are suffering from long term physical and mental health problems (*i.e.* long COVID). To combat this, a number of antiviral strategies that target the viral infection process have been developed.[2-7] As binding of the viral surface trimeric spike (S) protein to the angiotensin converting enzyme 2 (ACE2) receptor on the host cell surface is essential for infectious SARS-CoV-2 entry into cells,[8, 9] most antivirals and vaccines (including those under development) target this interaction.[2, 4, 6, 7] For example, neutralising antibodies (Abs) bind to the S protein and block infectious viral entry into cells.[2, 7] However, the emergence of SARS-CoV-2 variants with mutations in the S protein that alter antibody (Ab) epitopes can allow for evasion of neutralising Abs induced upon vaccination and/or infection.[10-14]

Each SARS-CoV-2 S-protein trimer is heavily glycosylated with 66 N-linked glycans, consisting of oligomannose, and other complex- and hybrid-type glycans.[15, 16] Glycosylation plays a critical role in viral pathobiology,

including mediating S-protein folding and stability, camouflaging immunogenic epitopes, facilitating ACE2 binding and viral cell entry.[17] Unlike the frequently changing S protein epitopes targeted by neutralising Abs, no mutations or alterations of S protein glycosylation sites have been reported across the SARS-CoV-2 variants of concern identified by the World Health Organisation (WHO) to date, including the highly mutated Omicron variant (BA.1, BA.4. BA.5; see **Table A.1**). While viral glycans are synthesised by the host cell machinery, they exhibit some unique features that differentiate them from host self-glycans, such as a high content of under-processed oligomannose and high glycan density. These make viral surface glycans an attractive target for developing antivirals. Indeed, a few rare but potent and broadly neutralising Abs target glycans on HIV.[18-20] For example, the Ab 2G12 displays an unusual domain-exchanged structure that brings its two fragment antigen-binding regions (Fabs) into close proximity to create an extended glycan binding surface, allowing 2G12 to form strong multivalent interactions with densely packed glycans on the same glycoprotein trimer spike, gp120, on the HIV surface.[21] This binding is not possible with conventional Abs: their Fabs are too widely (~15 nm) spaced to allow for simultaneous binding to the same gp120 molecule. Unfortunately, due to low natural immunogenicity, glycan-targeting Abs remain rare and, to date, no anti-SARS-CoV-2 neutralising Abs are known to be glycan-targeting. Nevertheless, the success of 2G12 and a few other Abs clearly demonstrates that targeting viral glycans by exploiting multivalency is a viable antiviral strategy.

Multivalent lectin-glycan interactions are widespread and play a key role in pathogen recognition and immune regulation.[22-24] Moreover, lectins have also shown to display useful antiviral activities by binding to viral surface glycans.[25-27] However, some lectins, including the dendritic cell surface lectin, DC-SIGN, have been shown to bind and transmit SARS-CoV-2 to target cells, albeit less effectively than ACE2.[28, 29]

This chapter, therefore describes the design of a new polyvalent lectin-nanoparticle (lectin-NP) antiviral strategy by displaying tetrameric DC-SIGN extracellular domain (ECD) or its monomeric carbohydrate-recognition domain (CRD) polyvalently and flexibly on gold nanoparticle (GNP) scaffolds. It is hypothesised that the flexibly displayed ECDs or CRDs in each lectin-NP will be able to adjust their relative positions, allowing for strong multivalent binding to glycans on viral S proteins to inhibit viral entry. Moreover, lectin-NPs may bind to glycans from different domains on the same trimeric S protein (*i.e.* intra-spike crosslinking) and/or in between neighbouring S proteins on the

same virion particles (inter-spike crosslinking). Such binding may interrupt the S protein conformational changes that are essential for virus entry into cells.[30, 31] A GNP scaffold is chosen here because of excellent biocompatibility, low-/non-cytotoxicity, tuneable size, and robust gold-thiol chemistry for convenient surface modification and bioconjugation.[32, 33] Moreover, glycan- or peptide-GNP conjugates have already been reported for studying SARS-CoV-2 viruses or antiviral IgG Abs, respectively.[34, 35] These advantages therefore make GNPs an excellent, tuneable platform for designing lectin-NPs of specific sizes, lectin valencies and flexibilities, required for potent virus neutralisation which can be readily optimised.

7.2 Experimental

7.2.1 DC-SIGN EDC and CRD

DC-SIGN tetrameric extracellular domain (ECD) and monomeric carbohydrate recognition domain (CRD) were produced and purified as described in **Section 2.1.4.1**. Here, because of the much weaker mannose binding affinity of monovalent CRD, the dialysed cell lysate was first concentrated to around 25 mL for 4 L of growth culture, prior to mannose-sepharose affinity column chromatography. Protein concentrations were obtained by UV-vis spectroscopy using the **Equation 2.1** (where, ε_{280} = 70400 and 52980 $M^{-1} \cdot cm^{-1}$ for ECD and CRD, respectively). Protein molecular weights (MW) were confirmed from deconvoluted HRMS data, where calculated protein MW values were obtained from the amino acid sequences of one monomeric unit of the corresponding protein.

DC-SIGN ECD: HRMS: calculated MW [ECD] 39197.22; found 39201.68.

DC-SIGN CRD: HRMS: calculated MW [CRD] 17794.72, found 17793.32.

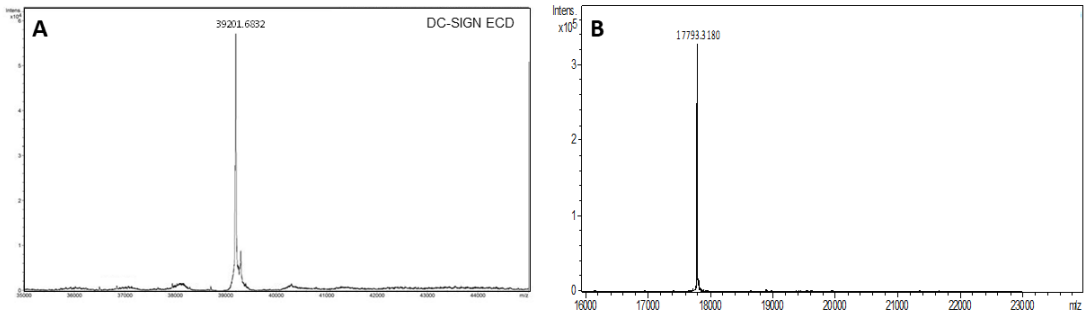

Figure 7.1 Deconvoluted HRMS spectra of a monomeric unit of (A) DC-SIGN's extracellular domain and (B) a DC-SIGN carbohydrate-recognition domain (CRD).

7.2.2 LA N-terminal labelling of ECD and CRD

LA-EG_{11}-TFP or LA-EG_{11}-Tz-TFP linkers were synthesis in-house by Darshita Budhadev in our group. DC-SIGN ECD or CRD N-terminal labelling was performed in a low pH binding buffer (20 mM HEPES, 150 mM NaCl, and 10 mM $CaCl_2$ pH = 6.2; denoted as N-terminal labelling buffer). LA-EG_{11}-TFP or LA-EG_{11}-Tz-TFP (in dry DMSO) was added to CRD in labelling buffer at a linker:CRD molar ratio of 1.5:1. The mixture was mixed on a rotating mixer at room temperature for 40 min and then diluted with binding buffer (20 mM HEPES, 100 mM NaCl, 10 mM $CaCl_2$, pH 7.8). Any unbound free linkers were removed by washing with the binding using 10 kDa MWCO ultra-filtration membrane. MWs were confirmed from deconvoluted HRMS data (see **Figure 7.4A,C,E**), where calculated protein MW values were obtained from the amino acid sequences of one monomeric unit of the corresponding protein.

LA-EG_{11}-Tz-ECD: HRMS: calculated MW [ECD] 39197.22, [LA-EG_{11}-Tz-ECD] 40512.89; found 39202.46 and 40517.99. LE: 18% 1^o labelling.

LA-EG_{11}-Tz-CRD: HRMS: calculated MW [CRD] 17794.72, [LA-EG_{11}-Tz-CRD] 19108.82, found 17793.15 and 19105.76. LE: 22% 1^o labelling.

LA-EG_{11}-CRD: HRMS: calculated MW [CRD] 17794.72, [LA-EG_{11}-CRD], 18610.06 [$(LA\text{-}EG_{11})_2$-CRD] 19424.00; found, 17793.22, 18608.63 and 19424.01. LE: 19% 1^o labelling and 2% 2^o labelling.

7.2.3 G13-Ct

13 nm GNPs (G13) were synthesised *via* previously established procedures.[36, 37] Briefly, $HAuCl_4 \cdot 3H_2O$ (79.2 mg, 0.201 mmol) was dissolved

in H_2O (200 µL) and heated to 130 °C (oil bath temperature). A solution of sodium citrate (Na_3Ct; 228.2 mg, 0.776 mmol) in H_2O (10 mL) was added with rigorous stirring, and the solution turned wine red within a 1 min. The heat was maintained for a further 20 min before the it was allowed to cool for 40 min and then left to stir at r.t. for a further 2 hr. The mean hydrodynamic diameter (D_h) and mean core diameter (D_{core}) of G13 was characterised by dynamic light scattering (DLS) and TEM, respectively. This yielded stable aqueous G13-Ct with a D_h of 15.1 ± 3.2 nm and D_{core} of 12.5 ± 1.0 nm (mean ± ½ FWHM; **Figure 7.2**). The concentration of G13 was determined by the Beer-Lambert law (**Equation 2.1**) using its peak absorbance at 519 nm (A_{519}) and ε_{G13} = 2.32 x 10^8 M^{-1} cm^{-1}.

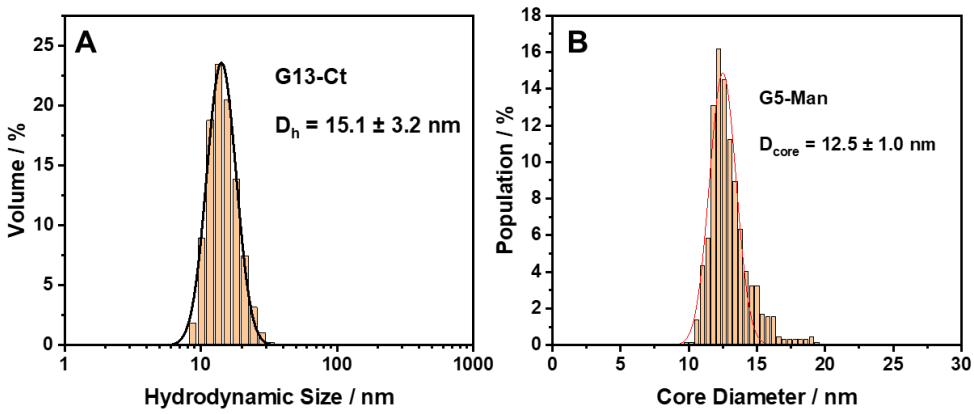

Figure 7.2 (A) Number population core diameter distribution histogram and (B) volume population hydrodynamic size distribution histogram, fitted with a lognormal Gaussian fits, for citrate capped 13nm GNPs (G13-Ct; $R^2 \geq 0.96$).

7.2.4 G13-(EG_7-OH)$_x$

G13-Ct was mixed with HS-EG_7-OH in water at a ligand:GNP molar ratio of 2000:1 or 5000:1 to partially or fully PEGylate the GNP surface, respectively, and were then stirred at r.t. for 48 hr. The resulting mixture was then concentrated using a 100 kDa MWCO filter and was washed with H_2O (3× 30 mL) to remove any unbound free ligands. Concentrations and D_h values were obtained by UV-vis spectroscopy and DLS. This yielded partially PEGylated and fully PEGylated G13 (denoted as ppG13-OH and G13-OH, respectively) which were found to be monodispersed and highly stable in binding buffer, with D_h values of 17.1 ± 4.6 and 17.9 ± 4.8 nm, respectively. ppG13-OH was used for protein conjugation to prevent aggregation in binding buffer, G13-OH was used as a negative control.

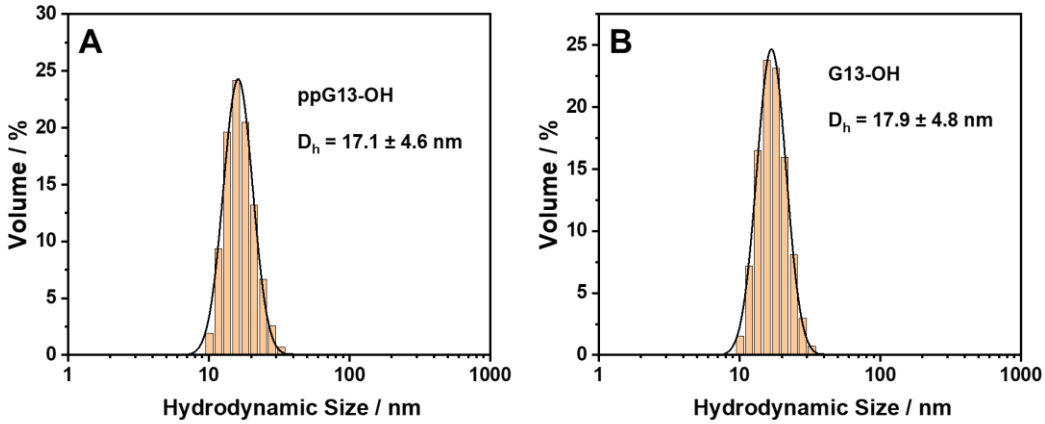

Figure 7.3 Volume population hydrodynamic size distributions, fitted with lognormal Gaussian fits, of 13nm GNPs (A) partially PEGylated with HS-EG$_7$-OH (ppG13-OH) and (B) 13 GNPs fully PEGylated with HS-EG$_7$-OH (G13-OH). (R^2 > 0.992).

7.2.5 G13-ECD and G13-CRD.

ppG13-OH in H_2O was added to ¼ of its volume of a 5x binding buffer (100 mM HEPES, 750 mM NaCl, 50 mM $CaCl_2$, pH 7.8) bring it to the binding buffer concentration. Proceeding this, LA-EG$_{11}$-Tz-ECD, LA-EG$_{11}$-Tz-CRD or LA-EG$_{11}$-CRD was added at a protein:GNP molar ratio of 100:1 or 115:1. The solutions were then stirred at 4 ºC overnight (~16 h). The conjugation mixture was then filtered using a 100 kDa MWCO spin filter (or centrifuged for the larger ECD) and the flow through filtrate (or supernatant) was collected and characterised by HRMS to ensure complete conjugation to the GNP surface (**Figure 7.4B,D,F**). The resulting lectin-GNP conjugates were further washed three times with the binding buffer using the same 100 kDa MWCO spin filter. Concentrations and D_h values were obtained by UV-vis spectroscopy and DLS, respectively (**Figure 7.5; Figure 7.9**). This yielded G13-Tz-ECD$_{100}$ (D_h 144.3 ± 0.6 nm), G13-Tz-CRD$_{100}$ (D_h 30.6 ± 0.2 nm), G13-CRD$_{100}$ (D_h 22.8 ± 5.5 nm) and G13-CRD$_{115}$ (D_h 26.4 ± 5.5 nm).

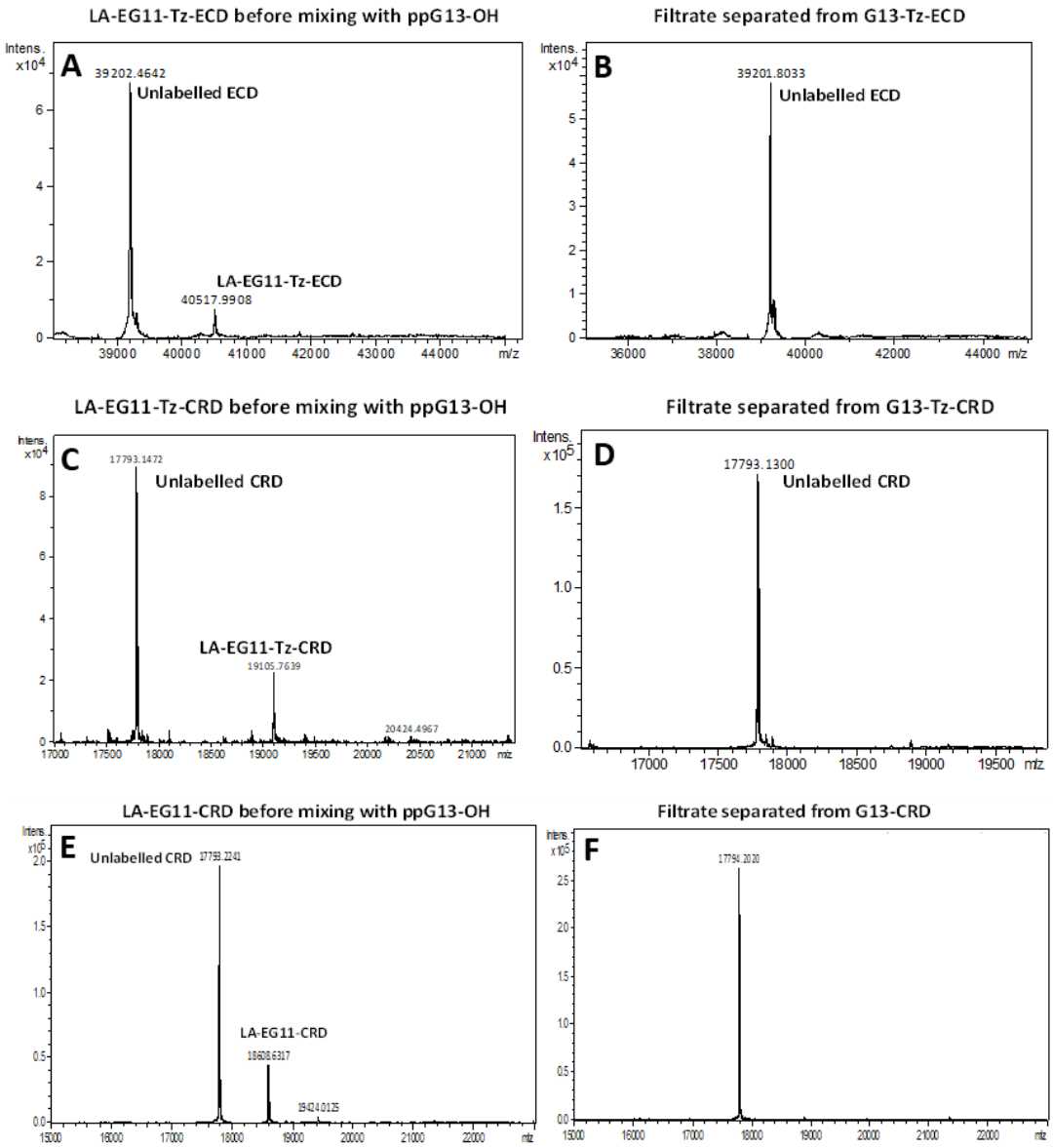

Figure 7.4 Deconvoluted HRMS spectra of the monomeric units of LA-EG$_{11}$-Tz-TFP N-terminal labelled (A) ECD and (C) CRD, and (E) LA-EG$_{11}$-TFP N-terminal labelled CRD, prior to ppG13-OH conjugation; and that of the filtrate collected post-conjugation of ppG13-OH with LA-EG$_{11}$-Tz-TFP N-terminal labelled (B) ECD and (D) CRD, and (F) LA-EG$_{11}$-TFP N-terminal labelled CRD.

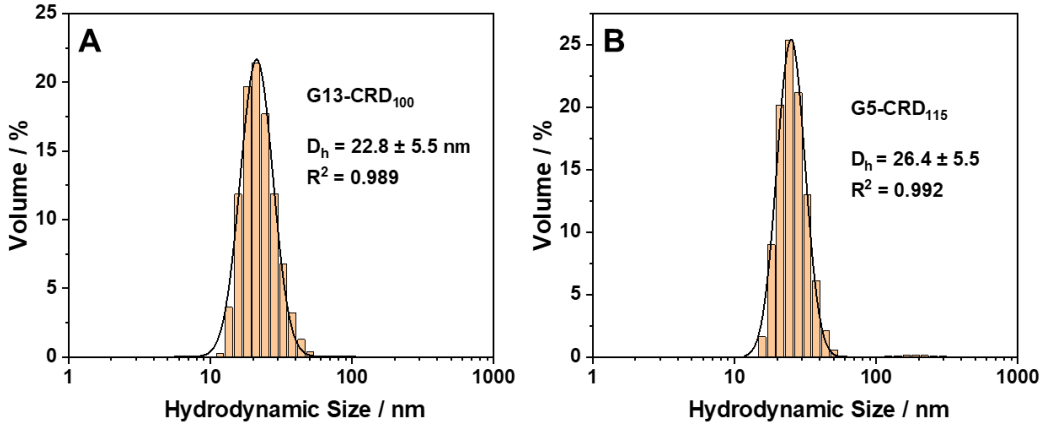

Figure 7.5 Volume population hydrodynamic size distributions, fitted with lognormal Gaussian fits, for 13nm GNPs conjugated with (A) 100 eq of LA-EG$_{11}$-CRD (G13-CRD$_{100}$) and (B) 115 eq of LA-EG$_{11}$-CRD (G13-CRD$_{115}$).

7.2.6 QD-DiMan and QD-Gal

DHLA-EG$_4$-DiMan and DHLA-EG$_4$-Gal ligands were prepared in-house by Darshita Budhadev in our group. Glycan-QDs were prepared as described in **Section 3.2.1**,[38] using CdSe/ZnS core/shell QD$_{600}$ and either DHLA-EG$_4$-DiMan or DHLA-EG$_4$-Gal (ligands prepared in-house) at a ligand:QD molar ratio of 2000:1. The QD concentration was determined by measuring its absorbance at the first exciton peak at 589 nm using an extinction coefficient of 3.3×10^5 $M^{-1}\cdot cm^{-1}$.[38]

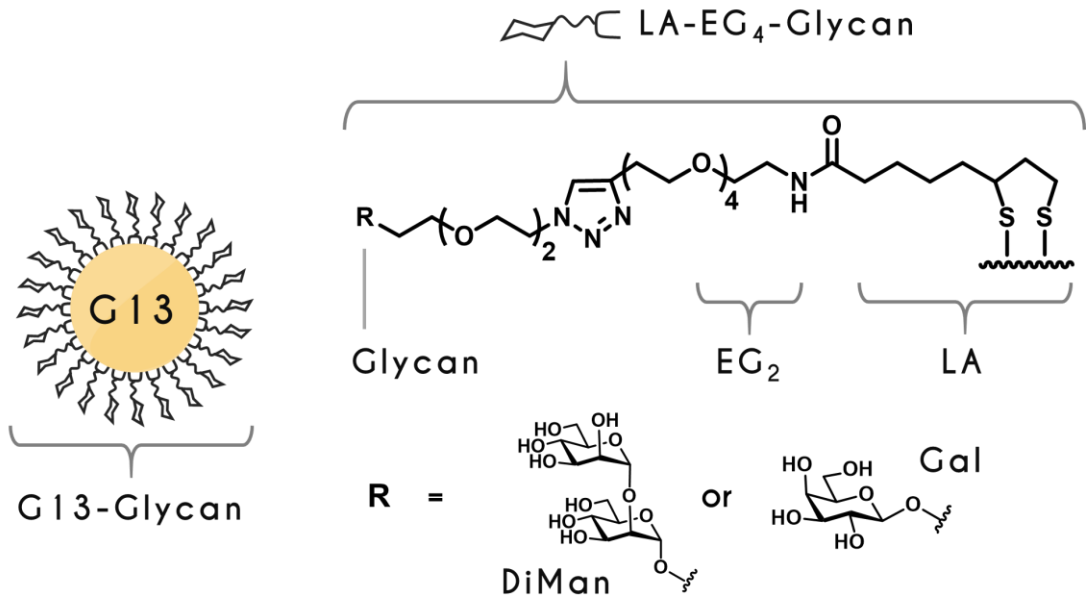

Figure 7.6 Schematic depicting G13-Glycan: 13nm GNPs coated with LA-EG$_4$-Glycans of either DiMan or Gal.

7.2.7 QD Fluorescence Quenching assays

Fluorescence quenching studies were performed using a Cary Eclipse Fluorescence Spectrophotometer and a 0.7 mL SUPRASIL® quartz cuvette (optical path length: 10 mm). Samples were excited with a fixed λ_{ex} = 400 nm and the fluorescence spectra were collected from 570 to 700 nm, with intervals (Δλ) of 1 nm. Excitation and emission slit widths and PMT were adjusted to ensure that fluorescence signals were always within the linear response range of the instrument to ensure high measurement accuracy. All samples were performed in duplicates and compared with and without the presence of the fluorescence acceptor (GNP), whereby the acceptor (QD) emission at λ_{em} of 605 nm was monitored to quantify binding.

Binding viability assays were performed by adding QD-DiMan or QD-Gal control (2 nM, final concentration) to G13-CRD or G13-OH control (1 nM, final concentration) in binding buffer containing BSA (1 mg mL^{-1}) to minimise any nonspecific interactions and absorption of the QD and GNP on surfaces.

Binding affinity assays were performed by adding QD-DiMan to G13-CRD in binding buffer (containing BSA, 1 mg mL^{-1}), at varying concentrations at a fixed protein-conjugate:QD molar ratio (PQR) of 1:1.

7.2.8 Pseudo-SARS CoV-2 virus preparation and inhibition studies.

Pseudotyped virus preparation and inhibition studies were performed by Inga Nehlmeier, Amy Madeleine Kempf, Markus Hoffmann and Nadine Krüger in Stefan Pöhlmann's group at the German Primate Centre, Göttingen, Germany. Vesicular stomatitis virus (VSV) pseudotyped with SARS-CoV-2 S-protein and encoding a luciferase gene was constructed as described previously.[39] To evaluate the inhibitory effect lectin-GNPs on SARS-CoV-2 S protein-driven cell entry, Vero76 cells were seeded in 96-well plates at a density of 2×10^5 cell mL^{-1}. Equal volumes of pseudotype preparations and lectin-GNP were incubated at 37 °C for 120 min. Medium was aspirated from the cells (at 24 hr post seeding) and pseudotype-lectin-GNP mixture (100uL) was added to each well and the cells were incubated at 37°C for 16-18 hr. After that, the cell medium was removed and cells were lysed using PBS supplemented with 0.5% triton X-100 (Carl Roth) for 30 min at r.t.. Then 30 μL of the cell lysates were transferred into white 96-well plates, mixed with

luciferase substrate (Beetle- Juice, PJK) and their luminescence was measured with a Hidex Sense Plate luminometer (Hidex). The luciferase activities of cell lysates from each treatment were measured and normalised by the corresponding control collected in the absence of the lectin-GNP. The normalised infection (NI)-concentration relationship was fitted by the modified inhibition model to derive the apparent viral inhibition potencies (EC_{50} and n values).

7.2.9 Inhibition of authentic SARS-CoV-2 infection by G13-CRD.

All work with infectious SARS-CoV-2 was conducted under BSL-3 conditions at the German Primate Centre, Göttingen, Germany, by the Pöhlmann group. Vero76 cells were seeded in 96-well plates at a density of $2x10^5$ cell mL^{-1}. $G13\text{-}CRD_{100}$ and G13-OH and Sotrovimab were incubated at different concentrations (ranging from 0.1 to 50 nM) at 37 ^oC for 2 hr with SARS-CoV-2 isolate NK, Pango lineage B.1.513 (provided by Stephan Ludwig, Institute of Virology, University of Münster, Münster, Germany) or SARS-CoV-2 isolate Omicron BA.1, Pango lineage B.1. (provided by Christian Drosten, Institute of Virology, Charité-Universitätsmedizin Berlin, Germany), in an inoculation volume of 100 µL. Afterwards, Vero76 cells were infected with the virus-inhibitor mixture at a multiplicity of infection (MOI) of 0.01. After 1 hr, the inoculum was removed, cell cultures were washed with PBS two times, and 100 µL of culture medium was added to the cells. Supernatants were collected at 0 and 48 hr post infection (hpi) and stored at -80 °C until further usage. Viral titres were determined by plaque assay on VeroE6 cells as described previously,[40] given in plaque forming units (PFU / mL).

7.3 Results

7.3.1 Lectin-GNP Conjugation

The ECD of DC-SIGN is composed of a neck domain, which has been shown to contain the main tetramerisation motifs, as well as the CRDs, which define the protein's glycan binding specificity.[41, 42] Each individual monomeric CRD binds specifically to mannose or fucose containing glycans found on virus surfaces, including SARS-CoV-2, with low to moderate monovalent affinities (K_d of 0.1-3 mM),[43] while the ECD of DC-SIGN contains 4 CRDs and can therefore enhance the binding affinity due to multivalency.[38]

Both DC-SIGN ECD and CRD were recombinantly expressed in *E. coli* and purified by mannose-affinity column chromatography as confirmed by

HRMS (**Figure 7.1**), as reported previously.[44, 45] To facilitate lectin-GNP conjugation, two linker molecules were designed, based upon a general structure of lipoic acid-undecyl-(ethylene glycol)-carboxylic acid tetrafluorobenze ester. Both linkers contain three functional domains: an lipoic acid (LA) group for strong GNP anchoring *via* chelative Au-S bonding; an EG_{11} spacer for good flexibility, water solubility and resistance to non-specific interactions;[46, 47] and a TFP ester for protein labelling *via* reaction with a free surface amine (**Figure 7.7**). The first linker, LA-EG_{11}-Tz-TFP, was prepared by reacting LA-EG_{11}-tetrazine with trans-cyclooctyne-EG_4-TFP ester (TCO-EG_4-TFP) *via* the rapid tetrazine-TCO cycloaddition reaction. While this reaction was rapid, the LA-EG_{11}-Tz-TFP linker was unstable for long-term storage, even at -20 ^oC, and gradually degraded over 4 months. Thus a second linker, LA-EG_{11}-TFP, was prepared *via* the direct esterification of LA-EG_{11}-CO_2H with TFP. LA-EG_{11}-TFP was highly stable, showing minimal degradation after storage for 12 months at -20 ^oC as lyophilised powders.

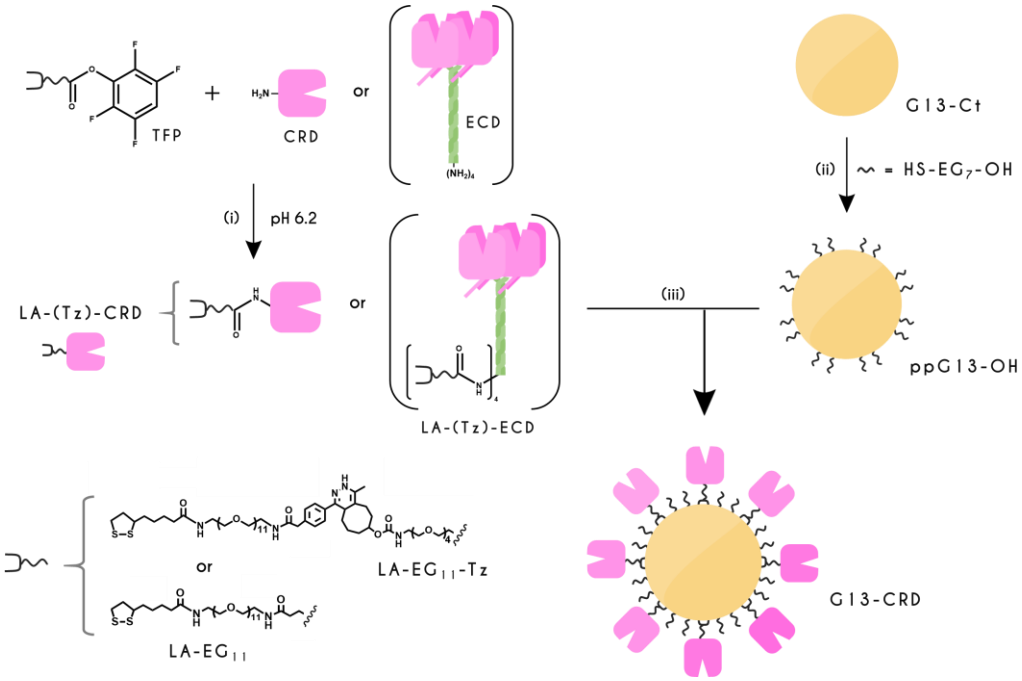

Figure 7.7 Schematic route to prepare an polyvalent lectin-GNPs: (i) Lectin N-terminal α-amine is selectively labelled with an LA-EG_{11} linker by reacting with either LA-EG_{11}-TFP or LA-EG_{11}-Tz-TFP at pH 6.2 to form LA-lectin or LA-Tz-lectin, respectively (where lectin = DC-SIGN CRD or ECD); (ii) Citrate stabilised 13 nm gold nanoparticle (G13) is partially PEGylated with HS-EG_7-OH to form ppG13-OH; (iii) ppG13-OH is then conjugated with LA-lectin *via* self-assembly to form G13-lectin (exemplified with G13-CRD).

To ensure all lectins conjugated on the GNP surface are oriented such that they are available for binding, the lectins were site-specifically labelled with the linkers at the N-terminal α-amine in DC-SIGN ECD or CRD. As the pK_a of N-terminal α-amines are at least 2 units lower than that of other free amines, such as the ε-amines of lysine residues (*e.g.* ~6.0–8.0 *vs* ~10.5),[48] labelling was conducted at a pH 6.2 to increase the probability that only α-amines, but not ε-amines, were unprotonated and available for amidation with the TFP ester (**Figure 7.7i**). Incubating the protein with LA-EG_{11}-Tz-TFP at 1:1.5 molar ratio for ~20 min was sufficient to produce single linker labelled proteins in ~18% and ~22% yields for ECD (denoted as LA-EG_{11}-Tz-ECD) and CRD (denoted as LA-EG_{11}-Tz-CRD), respectively. Extending the incubation time led to the formation of significant amounts of secondary labelling (which could lead to unfavourable lectin orientations or GNP crosslinking). The same condition was used to label CRD with the LA-EG_{11}-TFP linker, giving single linker labelled CRD in yields of ~19% (denoted as LA-EG_{11}-CRD) with a very small amount (~2%) of dual-labelled CRD (denoted as $(LA\text{-}EG_{11})_2$-CRD) (**Figure 7.4B**).

13 nm GNPs (G13), synthesised by citrate reduction of $H[AuCl_4]$ which yielded citrate capped G13 (G13-Ct),[49] were used to construct the antiviral lectin-NPs in two steps. First, G13 was partially PEGylated with a hepta(ethylene glycol) thiol (HS-EG_7-OH; **Figure 7.7ii**) to prevent aggregation during lectin conjugation, which must be performed in binding buffer rather than pure H_2O to ensure stability of the protein. The relative concentration of HS-EG_7-OH required to perform only partial cap exchange of the surface citrate molecules whilst retaining high stability in binding buffer was determined by incubating G13-Ct overnight with varying molar equivalents (eq) of HS-EG_7-OH and measuring the UV-vis spectra. The desired ligand:GNP molar ratio was taken as the lowest amount required to result in GNPs that displayed no red shift in UV-vis spectra in binding buffer (monitored by the change in A_{650}; **Figure 7.8A,B).** The degree of PEGylation was confirmed by monitoring the shift in gel mobility with ligand:GNP molar ratio, where that those with higher citrate content (*i.e.* lower degree of PEGylation) would in theory possess higher negative charge and thus display increased gel mobility (**Figure 7.8C**). Here, 2000 eq was observed to demonstrate both good stability in binding buffer and higher gel mobility (*i.e.* lower HS-EG_7-OH content), thus was used to partially PEGylate G13 (denoted as ppG13-OH). The gel mobility was observed to plateau at 3000 eq, demonstrating complete HS-EG_7-OH coverage, however, 5000 eq was used as the fully PEGylated G13 control GNP (denoted as G13-OH).

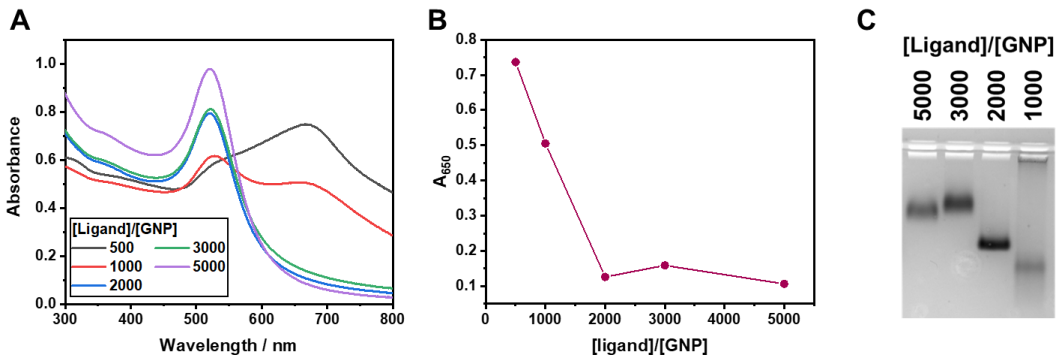

Figure 7.8 (A) UV-vis spectra of G13 capped with HS-EG$_7$-OH at varying ligand:GNP molar ratios in binding buffer; (B) the degree of redshift as a function of [ligand]/[GNP] measured by the absorbance at 650 nm (A_{650}), where high absorbance (*i.e.* increased redshift) demonstrates instability in binding buffer; (C) 1.5% agarose gel electrophoresis of G13 capped with HS-EG$_7$-OH at varying ligand:GNP molar ratios, demonstrating lower HS-EG$_7$-OH surface coverage (higher citrate content) at lower ratios, signified by reduced gel mobility.

For the second step, ppG13-OH was incubated overnight in binding buffer with 100 molar equivalents of linker labelled lectin, yielding G13-lectin *via* self-assembly (**Figure 7.7iii**). Proceeding conjugation, the G13-lectin samples were washed with binding buffer and the resulting wash solutions were analysed by HRMS to ensure that all linker-labelled lectin had been removed. This indicated that complete conjugation of protein to the GNP surface was achieved, demonstrating a lectin valency of 100 lectins per GNP. This was performed for the three linker-labelled lectins (LA-EG$_{11}$-Tz-ECD, LA-EG$_{11}$-Tz-CRD and LA-EG$_{11}$-CRD), yielding G13-Tz-ECD$_{100}$, G13-Tz-CRD$_{100}$ or G13-CRD$_{100}$, respectively. To investigate the effect of CRD valency on antiviral activity, another batch of G13-CRD was prepared at a linker:GNP ratio of 115, yielding G13-CRD$_{115}$. The success of G13-lectin conjugation was supported from the reduced gel mobility as compared against ppG13-OH (**Figure 7.9A,B**)**,** and increased mean hydrodynamic diameters (D_h) following each conjugation step. For example, the D_h of G13-Ct (~15 nm) was increased to ~17 nm for ppG13-OH, and then to ~31 nm for G13-Tz-CRD$_{100}$ or ~140 nm for G13-Tz-ECD$_{100}$ (**Figure 7.9C**). For the smaller linker labelled lectins, D_h values of ~22 nm or 26 nm were demonstrated for G13-CRD$_{100}$ and G13-CRD$_{115}$, respectively.

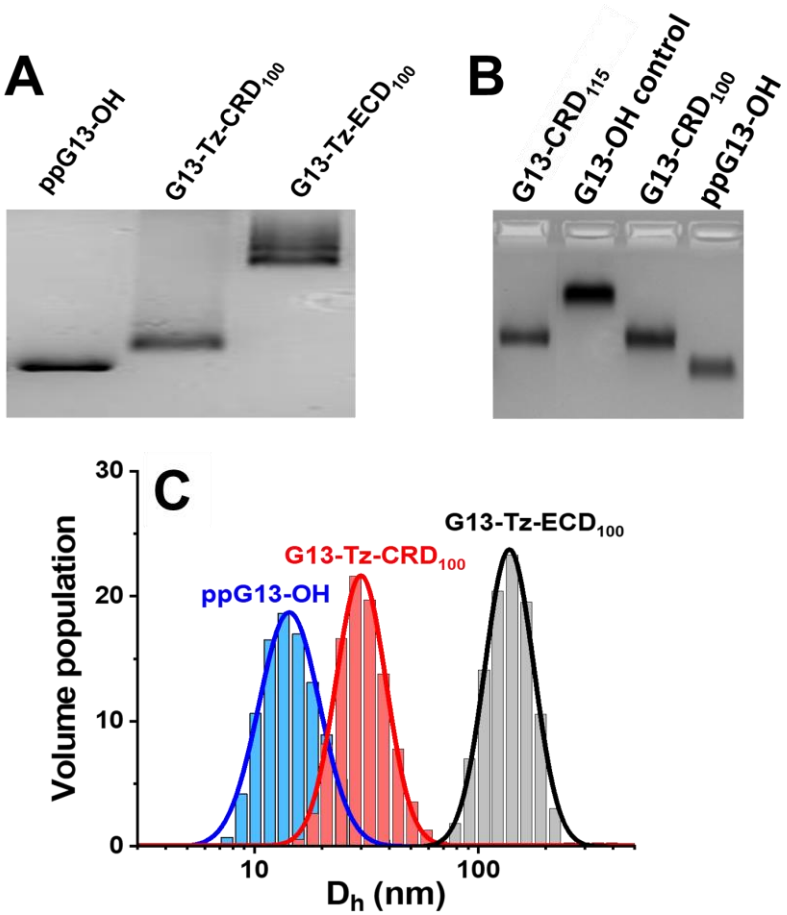

Figure 7.9 1.5% agarose gel electrophoresis demonstrating that (A) ppG13-OH, G13-Tz-CRD$_{100}$ and G13-Tz-ECD$_{100}$ have successively slower gel mobility, likely due to an increase in molecular weight with addition of linker-labelled CRD and ECD; and (B) similarly, G13-CRD$_{100}$ and G13-CRD$_{115}$ also display slower gel mobility over ppG13-OH. Additionally, fully PEGylated G13-OH has a slower gel mobility than the partially PEGylated ppG13-OH, indicating the higher ligand surface coverage of the former. (C) Demonstrates the volume population hydrodynamic size distribution histograms, fitted by log-normal Gaussian fits, of ppG13-OH, G13-Tz-CRD$_{100}$ and G13-Tz-ECD$_{100}$, yielding D_h values of ~17, ~31 and ~144 nm.

7.3.2 Lectin-GNP Glycan Binding

To confirm that G13-Tz-CRD$_{100}$ and G13-CRD$_{100}$ retained CRD's native glycan binding specificity, their binding efficacy was monitored *via* their interaction with glycan coated quantum dots (glycan-QDs). Where quantum dots (QDs) are recognised for their high fluorescence stability, GNP are well-known for their strong fluorescence quenching properties *via* a nanoscale surface energy transfer mechanism (NSET), where the efficiency of energy transfer (E) is inversely proportion to the fluorophore-GNP distance (R) to the fourth power (*i.e.* $E \propto 1/R^4$). NSET is thus able to covers a longer distance

range than the Förster resonance energy transfer (where $E \propto 1/R^6$).[50, 51] Therefore, by incubating lectin-GNP with glycan-QD (**Figure 7.6**), and exciting the QD, binding can be monitored by the quenching of QD fluorescence due to the close proximity of the QD donor fluorophore with the GNP acceptor (**Figure 7.10**). The quenching efficiency (QE; *i.e.* the percentage shift of fluorescence intensity at ~605 nm, I) represents the percentage of added glycan-QDs which have bound to lectin-GNP.

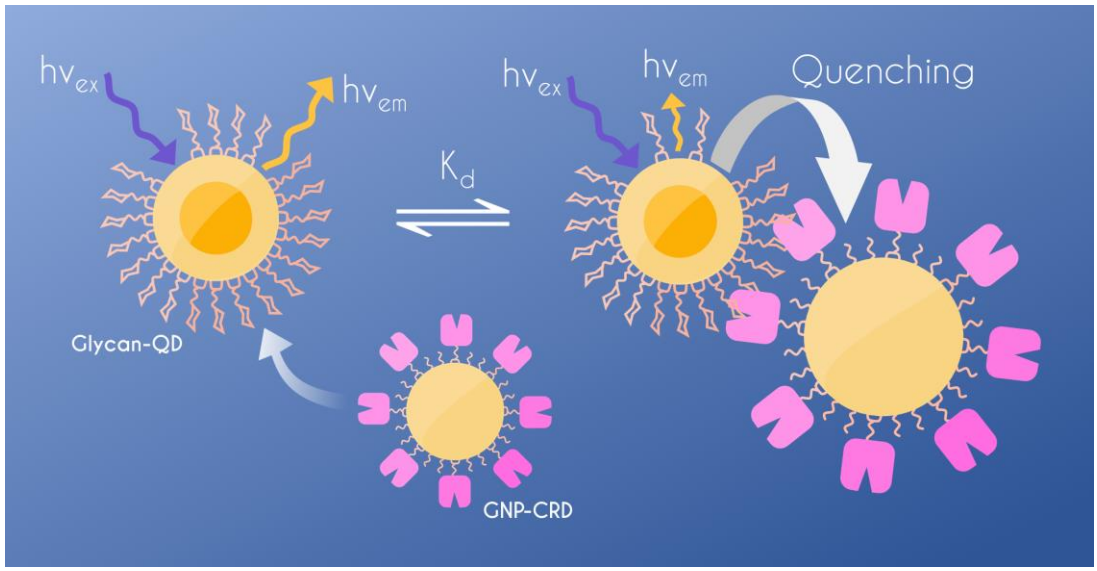

Figure 7.10 Schematic depicting the fluorescence quenching of glycan-capped QDs (*e.g.* QD-DiMan) binding with lectin capped GNPs (*e.g.* G13-CRD) with an equilibrium binding dissociation constant, K_d.

Here, CdSe/ZnS core/shell QD_{600} (λ_{em} = ~600 nm) were incubated with either dihydrolipoic acid-tetra(ethylene glycol)-mannose-α-1,2-mannose (DHLA-EG_4-DiMan) or dihydrolipoic acid-tetra(ethylene glycol)-galactose (DHLA-EG_4-Gal) ligands, yielding QD-DiMan or QD-Gal. Here, QD-DiMan was selected as a positive control due to the specific recognition of DiMan by DC-SIGN and QD-Gal was selected as a negative due to the naturally poor binding affinity of Gal with DC-SIGN.[38, 42]

As expected, the fluorescence of QD-DiMan was quenched efficiently (QE >80%) upon mixing with G13-Tz-CRD_{100} (**Figure 7.11A**). In contrast, the QD-Gal control was quenched much less efficiently (QE ~10%, **Figure 7.11B**), demonstrating that the G13-lectin is specifically binding to the carbohydrates corresponding to their binding sites. In addition to these data, due to the equivalence between the QE and the proportion of bound species, the binding affinity can be obtained using the Hill equation to derive the apparent binding equilibrium dissociation constant, K_d (**Equation 7.1**; where

I is the fluorescence intensity at ~605 nm, I_0 is the I value of the glycan-QD donor prior to addition of the lectin-GNP fluorescence acceptor, [ ] denotes the concentration, and n is the Hill coefficient).[38, 45]

$$\mathrm{QE} = \left(\frac{I - I_0}{I_0}\right) = \frac{[\mathrm{QD}]_{\mathrm{bound}}}{[\mathrm{QD}]_{\mathrm{total}}} = \frac{[\mathrm{QD}]^n}{{K_\mathrm{d}}^n + [\mathrm{QD}]^n} \qquad 7.1$$

This was performed for a G13-Tz-CRD$_{100}$ or G13-CRD$_{100}$ with QD-DiMan at a protein-conjugate:QD molar ratio (PQR) of 1:1, which yielded K_d values of 0.93 ± 0.05 and 0.6 ± 0.1 nM, respectively (**Figure 7.11C-F**). This represents an impressive ~1 million-fold enhancement of affinity over that of monovalent CRD-DiMan binding (K_d ~ 0.9 mM).[52] These results confirm that G13-CRDs have not only retained CRD's native binding specificity with DiMan but also drastically enhanced its affinity by forming multivalent binding. As has been demonstrated in **Chapters 1** and **3**, free tetrameric ECD (before GNP-conjugation) is able to bind strongly with QD-DiMan glycoconjugates with low-sub nM K_d values.[38] However, G13-Tz-ECD$_{100}$ is only able to give a QE about ~¼ of that obtained with G13-Tz-CRD$_{100}$ after mixing with QD-DiMan (**Figure 7.11A**). The ineffective quenching here is attributed to the long rigid coiled-coil neck (~20 nm in length)[41] in ECD, which projects CRDs away from GNP surface, resulting in a large GNP-QD separation distance which therefore results in ineffective quenching after binding.

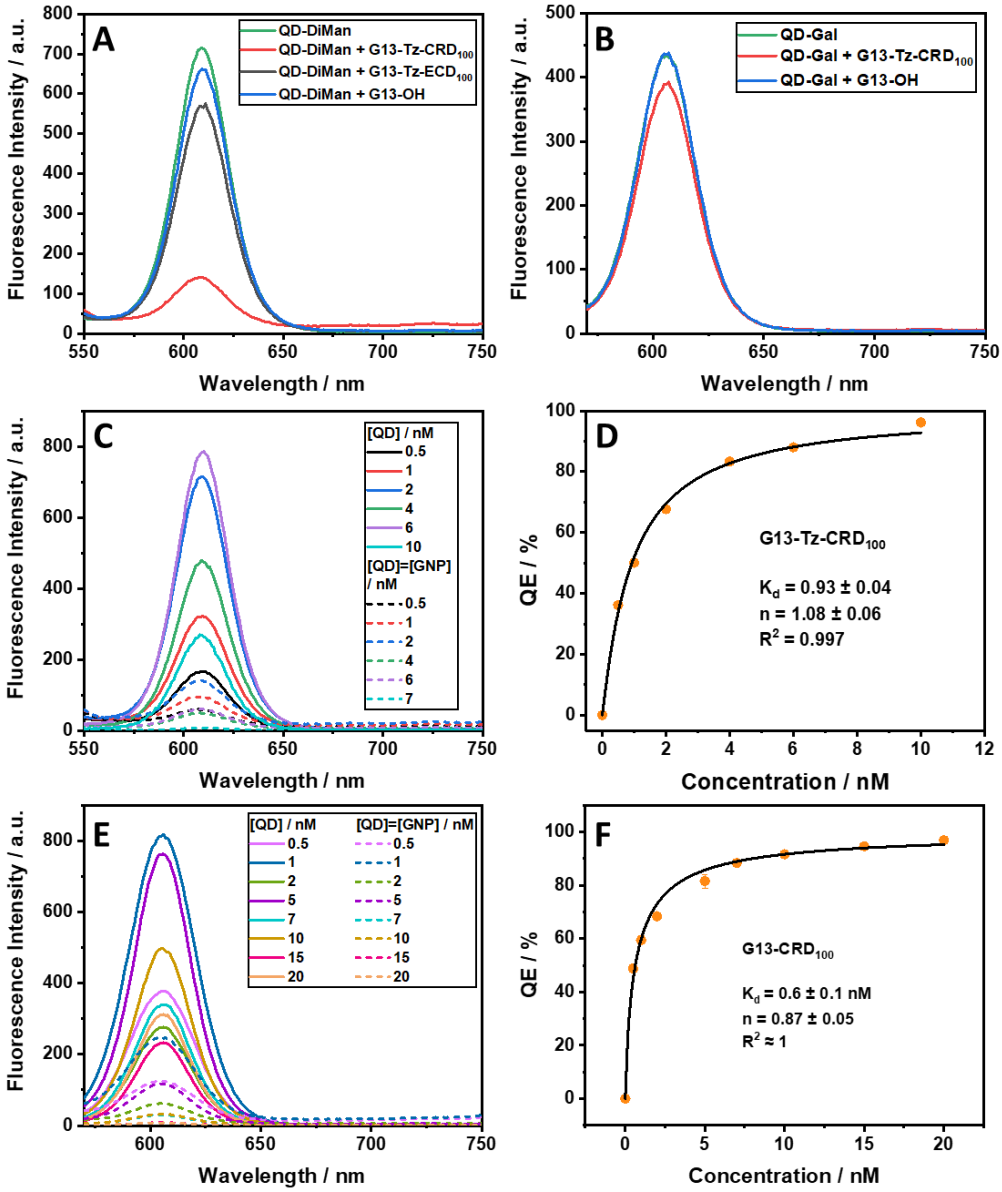

Figure 7.11 Fluorescence quenching efficiency (QE) assays demonstrating the glycan binding specificity of G13-lectin by: (A) the fluorescence quenching of QD-DiMan (2 nM) by G13-Tz-CRD$_{100}$ (1 nM; QE >80%) and G13-Tz-ECD$_{100}$ (1 nM; QE ~20%), and only marginally by the G13-OH control (QE ~7.5%); (B) marginal fluorescence quenching of QD-Gal (2 nM) negative control by G13-Tz-CRD$_{100}$ (1 nM; QE ~9.5%), and not by G13-OH control. Fluorescence spectra of varying concentrations of QD-DiMan in the absence (solid line) or presence (dash line) of 1:1 mixed (C) G13-Tz-CRD$_{100}$ or (E) G13-CRD$_{100}$, and the corresponding QD-DiMan QE-concentration relationship, fitted with Equation 7.1, for (D) G13-Tz-CRD$_{100}$ and (F) G13-CRD$_{100}$.

7.3.3 Inhibition of entry of SARS-CoV-2 S protein bearing pseudotypes into Vero76 cells by lectin-GNPs

Replication-defective single cycle Vesicular Stomatitis Virus (VSV) reporter particles encoding luciferase and bearing the S protein of SARS-CoV-2 were employed to evaluate the inhibitory effect of the lectin-GNPs against SARS-CoV-2 S protein-driven entry into Vero 76 cells. These reporter particles adequately model SARS-CoV-2 entry into cells and thus provide a good model for specific viral inhibition.[10, 11, 53] The natural unconjugated tetrameric ECD did not inhibit entry even at high concentrations of 5 μM (**Figure 7.12C**). In contrast, however, both G13-Tz-ECD$_{100}$ and G13-Tz-CRD$_{100}$ significantly and dose-dependently inhibited SARS-CoV-2 S protein-driven cell entry (**Figure 7.12A**). These inhibition data were then fitted by a modified inhibition model (**Equation 7.2**; where [I] is the inhibitor concentration; NI is the normalised infection, described by the luciferase activity; EC$_{50}$ is the effective inhibitor concentration which affords 50% inhibition; and n is the inhibition coefficient, where n >1, =1 or <1 indicates positive-, none- or negative- inhibition cooperativity, respectively).[45]

$$\mathrm{NI} = \frac{1}{1 + \left(\frac{[\mathrm{I}]}{\mathrm{EC}_{50}}\right)^{n}} \qquad 7.2$$

While the EC$_{50}$ value is clearly important in determining the efficacy of viral inhibition, the inhibiting n value is also of great importance. For example, if three inhibitors exhibited the same EC$_{50}$ but n values of 0.5, 1 and 2, then the [I] required for 99% inhibition will be 9801-, 99-, and 9.9-times the EC$_{50}$ value, respectively. Therefore, an inhibitor with n = 0.5 would be much less effective than that with n = 1 or 2, requiring ~100- or 1000-fold higher concentrations to achieve 99% inhibition. Therefore, viable inhibitors should have n ≥ 1 (with n = 1 being the most widely observed) in order to achieve complete inhibition within a reasonable concentration.

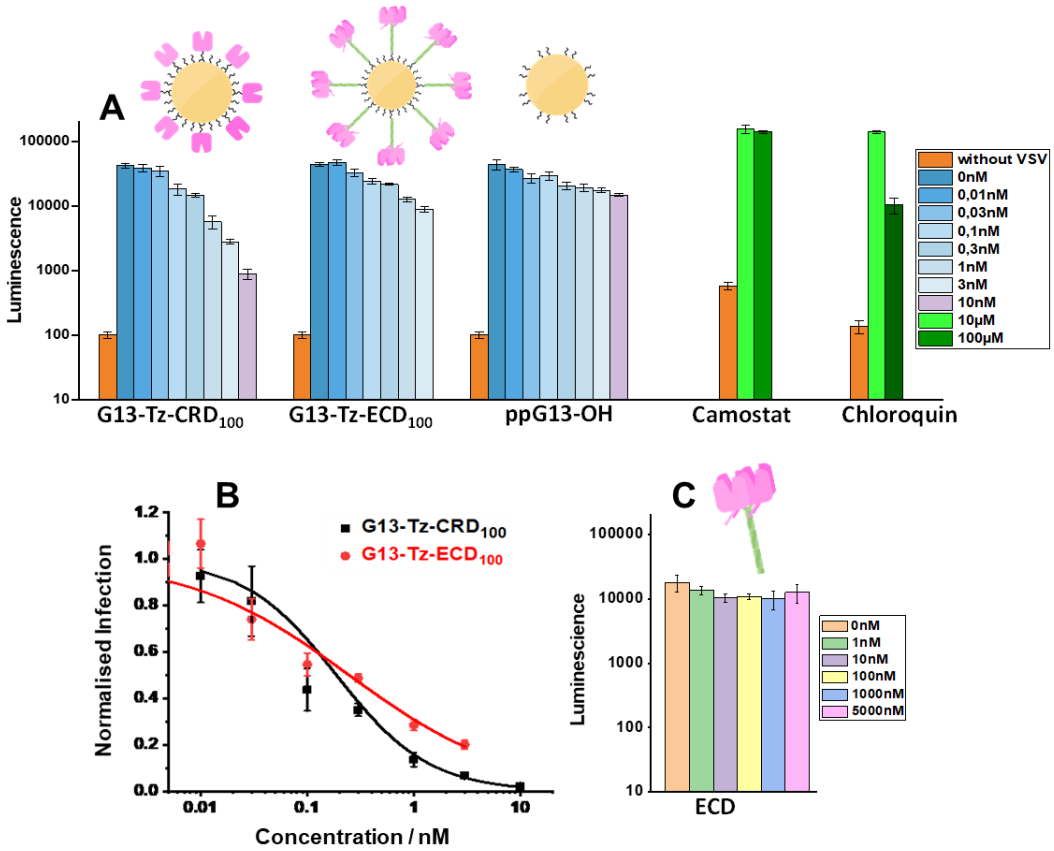

Figure 7.12 Dose-dependent inhibition of SARS-CoV-2 S-protein-driven entry into Vero76 cells. VSV particles bearing SARS-CoV-2 S protein (Wuhan wildtype, Hu-1) were pre-incubated with (A) G13-Tz-CRD$_{100}$, G13-Tz-ECD$_{100}$ or G13-OH control, or (C) tetrameric DC-SIGN ECD, before addition to target cells. As controls, target cells were pre-incubated with Camostat and Chloroquin before addition of pseudotyped particles. Entry efficiency was determined by quantifying luciferase activity in cell lysates. The orange bars in panel B represent the background luminescence measured in the absence of viral particles. The results of a representative experiment performed in quadruplicates was shown and were confirmed in two separate experiments. Errors bars indicate standard errors of one quadruplicate experiment. (B) The corresponding normalised infection curves, fitted by a modified inhibition model (Equation 7.2), yielding an effective inhibition concentration (EC$_{50}$) of 0.25 ± 0.04 nM and an inhibition coefficient (n) of 0.57 ± 0.06 for G13-Tz-ECD$_{100}$ (red) and EC$_{50}$ = 0.19 ± 0.02 nM and n = 1 for G13-Tz-CRD$_{100}$ (black) (R^2 = 0.97).

Fitting **Figure 7.12** with **Equation 7.2** gave comparably low sub-nM EC$_{50}$ values for both G13 -Tz-ECD$_{100}$ and G13-Tz-CRD$_{100}$ (0.25 ± 0.04 and 0.19 ± 0.02 nM, respectively; **Figure 7.12B**), indicating a high antiviral potency. However, the inhibition profile of G13-Tz-ECD$_{100}$ gave n = 0.57 ± 0.06, meaning it is difficult to achieve complete inhibition. In contrast, the inhibition profile of G13-Tz-CRD$_{100}$ yielded n = 1, meaning complete viral inhibition can

be achieved at much more reasonable concentrations. This is evident from the fact that, despite having similar sub-nM EC_{50} values, the normalised infection for G13-Tz-ECD_{100} is >3-fold that of G13-Tz-CRD_{100} at 3 nM (**Figure 7.12**). Therefore, presenting monomeric CRDs flexibly in on a GNP surface, with each CRD serving as an independent binder, is key to viable and potent viral inhibition. Such flexible CRD binding units can readily adjust their relative positions to accommodate viral surface glycans and form strong multivalent binding. In contrast, for ECDs, which each contain four CRDs, the CRD positions are fixed to the coiled coil neck and cannot readily adjust their relative positions to adapt to viral surface glycans, making it difficult for ECD-conjugated GNPs to form strong multivalent binding. As discussed throughout this thesis, this restricted flexibility in CRD presentation is key for facilitating a lectin's selectivity towards spatially matched multivalent glycans. Because of this, natural CRDs often lack the flexibility and adaptability required to achieve complete viral inhibition, making them ineffective as antiviral reagents in comparison to our much more adaptable GNP-CRD conjugates.

Interestingly, neither G13-Tz-CRD_{100} or G13-Tz-ECD_{100} showed significant entry inhibition of pseudo-typed viral particles bearing a control vesicular stomatitis virus glycoprotein (VSV-G), demonstrating the specificity towards targeting SARS-CoV-2 S protein (**Figure C.13**). In addition, the G13-OH control showed no significant inhibition against SARS-CoV-2 S protein-bearing VSV (**Figure 7.12A**), as expected, demonstrating that viral inhibition is due to specific lectin-glycan interactions. Camostat (an inhibitor of the SARS-CoV-2 S protein activating protease TMPRSS2) exhibited no inhibitory effect even at 100 µM, as expected, since Vero cells do not express TMPRSS2. In contrast, chloroquine displayed significant inhibition at ~100 µM, as expected.

The lack of long-term stability for LA-EG_{11}-Tz-TFP means it has to be prepared fresh each time before lectin conjugation, making its use inconvenient. Therefore, lectin-GNP conjugates prepared with the more stable LA-EG_{11}-TFP linker at two different CRD valencies (G13-CRD_{100} and G13-CRD_{115}) were investigated for the inhibition potency of S protein-driven entry of Vero cells. Here, the S proteins of four SARS-CoV-2 variants (Wuhan wild-type Hu-1, B.1, Delta and Omicron BA.1)[10, 11] were investigated. The dose-dependent inhibition data, fitted with **Equation 7.2**, yielded comparable low nM EC_{50} values and n = 1 for G13-CRD_{115} against all four S proteins tested (**Figure 7.13; Table 7.1**), where no inhibition was observed for the

G13-OH control (**Figure C.14**). Interestingly, G13-CRD$_{115}$ showed consistently higher potencies (~3-4 fold) than G13-CRD$_{100}$ (also with n = 1) against each S protein-driven infection (**Figure 7.13; Table 7.1**), suggesting that a higher CRD valency (which also provides in a larger D_h; **Figure 7.5**) is beneficial for improving antiviral potency. Interestingly, clinically approved COVID-19 neutralising Abs or sera from recovered COVID-19 patients or double vaccinated individuals were found to be either ineffective or showed greatly reduced potency against the Omicron variants (both pseudo-typed and authentic viruses).[10, 12] This result demonstrates a great potential for our G13-CRD based lectin-NPs as a new class of anti-SARS-CoV-2 agent. Given that the S protein glycosylation sites are conserved in all SARS-CoV-2 variants of concern, including Omicron BA.4 and BA.5 variants, targeting viral glycans thus represents a viable, robust anti-SARS-CoV-2 strategy.

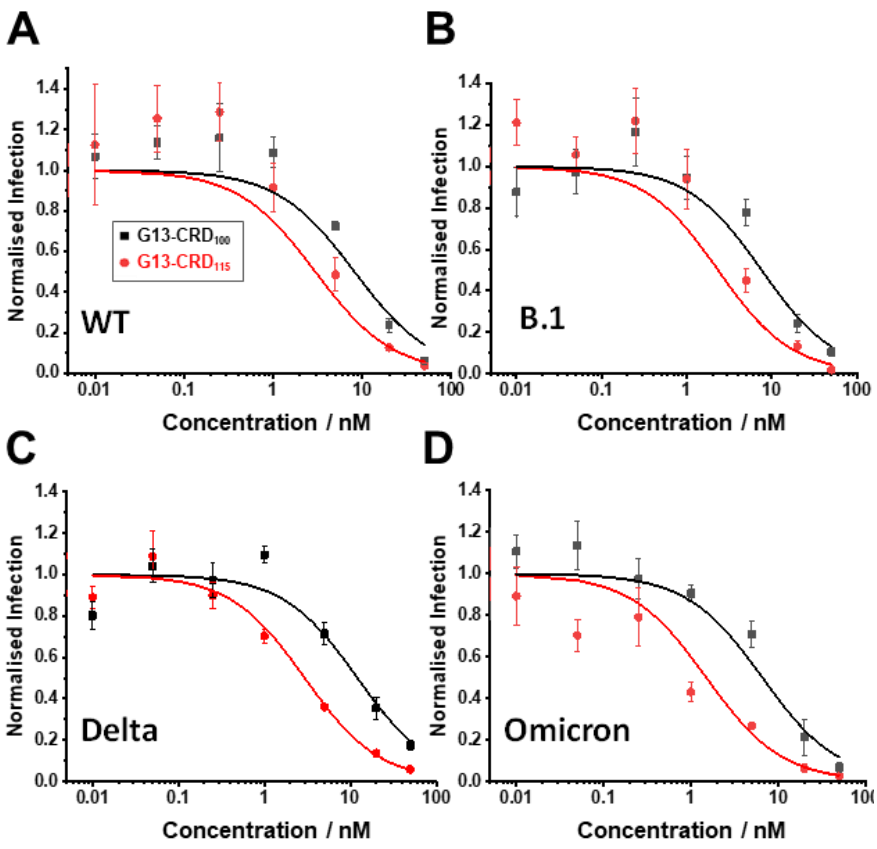

Figure 7.13 Normalised dose-dependent luciferase activity data, fitted by the inhibition model (Equation 7.2; n = 1 for all fits), for the inhibition of SARS-CoV-2 S protein-mediated cell entry of Vero76 by G13-CRD$_{100}$ (black) and G13-CRD$_{115}$ (red) against VSV particles bearing the S proteins from SARS-CoV-2 variants: (A) Wuhan Hu-1 (WT), (B) B.1, (C) Delta (B.1.617.2), (D) Omicron (BA.1). Parameters summarised in Table 7.1 (R^2 > 0.92).

Table 7.1 Summary of inhibition data by G13-CRD$_{100}$ and G13-CRD$_{115}$ against infection of Vero76 cells by pseudo-typed SARS-CoV-2 bearing S proteins from various variants (Figure 7.13; fitted with Equation 7.2; n = 1 for all fits). A higher CRD valency and larger size of G13-CRD appears to be beneficial for boosting antiviral potency.

	G13-CRD$_{100}$ (D_h ~22 nm)		G13-CRD$_{115}$ (D_h ~26 nm)	
SARS-CoV-2 Variant	**EC_{50} / nM**	**R^2**	**EC_{50} / nM**	**R^2**
Wild-type (Hu-1)	8.2 ± 1.7	0.933	3.0 ± 0.5	0.943
D614G wildtype (B.1)	7.6 ± 1.3	0.961	2.3 ± 0.6	0.947
Delta (B.1.617.2)	12.1 ± 2.7	0.952	3.0 ± 0.1	0.992
Omicron (BA.1)	6.6 ± 1.3	0.976	1.5 ± 0.3	0.922

7.3.4 Inhibition of authentic SARS-COV-2 entry into Vero76 cells by lectin-GNPs

The inhibitory effects of G13-CRD against two authentic SARS-CoV-2 viruses, Wuhan Hu-1 (wild type), and Omicron BA.1, were also investigated. Here, Sotrovimab, a clinically approved monoclonal antibody for Covid-19 treatment was analysed as a control. G13-CRD$_{100}$ was highly potent against the wild-type virus, achieving ~92% inhibition at concentrations as low as 0.1 nM (equivalent to ~9 pM EC_{50}, assuming non-cooperative binding, *i.e.* n = 1) and complete inhibition at 10 nM (**Figure 7.14**). In contrast, Sotrovimab was less effective, showing apparently no inhibition at 0.1 nM, although significant inhibition was observed at 1 nM (~85%) and above. The inhibition data of authentic viruses did not follow the classical potency-dose dependence, where the potency increased more rapidly with dose once inhibition was observed, making it difficult to fit the data with inhibition models to derive EC_{50} values. Against Omicron variant BA.1, both Sotrovimab and G13-CRD$_{100}$ showed reduced efficacy and significant inhibition was observed at 5 nM and 10 nM, respectively but both completely inhibited infection at 25 nM. Importantly, the G13-OH control gave no inhibition across the whole concentration range tested, confirming that G13-CRD's inhibitory effect originates from specific CRD-glycan interactions, as expected.

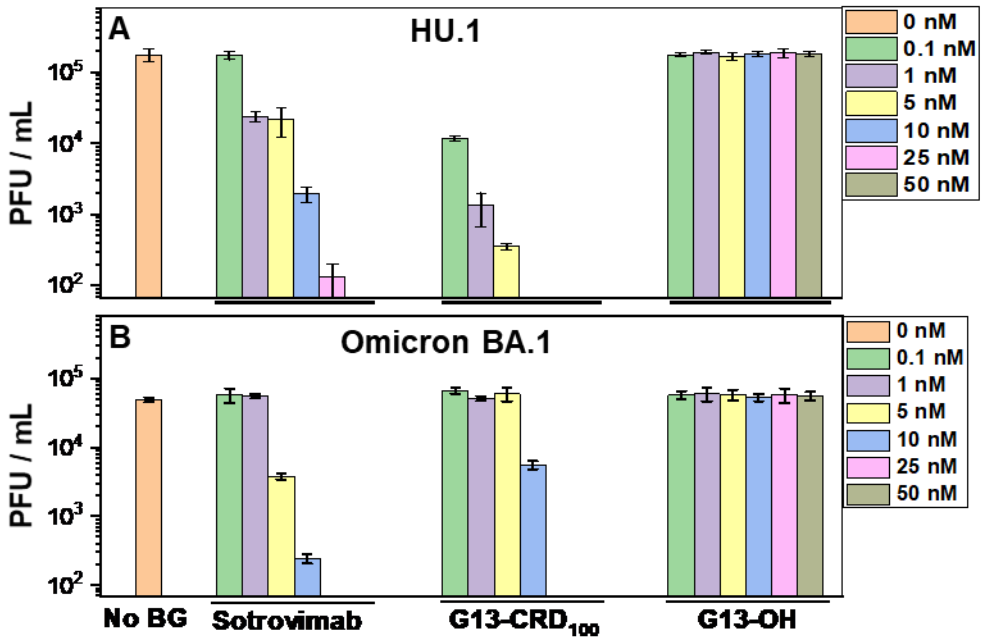

Figure 7.14 G13-CRD$_{100}$ and Sotrovimab inhibition of authentic SARS-CoV-2 infections against the (A) Wuhan wild-type (Hu-1) and (B) Omicron BA.1 variants. The absence of bar chart data at high concentrations indicate no observed infections (*i.e.* number of infectious particles surpasses the limit of detection (LOD); where LOD is indicated by the lowest limit of the Y-scale). "No BG" indicates the native infection level in the absence of any inhibitors.

7.4 Discussion and conclusions

As S protein glycosylation sites are conserved in all four SARS-CoV-2 variants studied (and in all variants of concern identified by WHO to date), individual G13-CRD-S protein interactions are thus expected to be similar across the different variants. As all four viral variants in the pseudotype model were consistently neutralised by G13-CRD with comparable EC_{50} and identical n values (*i.e.* n = 1) matches well with this expectation. This complimentarity implies that these different particles may have exhibited similar numbers of S proteins and inter-S-protein spacing, allowing them to interact with G13-CRD in a similar way. Conversely, a large potency difference for G13-CRD against the authentic Wuhan variant (Hu-1), compared to the analogous pseudo-type virus, could thus be assigned to differences in virion size/shape (~95 nm sphere,[54] *vs.* ~80 nm x 170 nm bullet-shaped,[55] respectively), the number of S proteins or inter-S protein spacing. While individual G13-CRD·S protein interactions may be similar, differences in inter-S protein spacing and surface curvature will likely affect G13-CRD's ability to crosslink neighbouring

S proteins on the virion surface, which is critical for interrupting the S protein's conformational changes and membrane fusion, and thus viral infection. In fact, the G13-CRD was designed to target the inter-S protein spacing of the Wuhan wild-type Hu-1 (~30 nm, see **Equation B.2**; for N of ~40 spikes randomly distributed on a spherical virion of ~95 nm in diameter, D) based on its cryo-EM structure,[54] which correlates well with its excellent inhibition potency against this virus (estimated equivalent EC_{50} of ~9 pM). These results clearly demonstrate the great potential of our design strategy for G13-CRD-based antivirals.

The lower potency of Sotrovimab against the authentic BA.1 over Wuhan Hu-1 variant is assigned to immune evasion mutations of S proteins in BA.1, which weakened Sotrovimab's binding affinity and hence neutralising potency.[12] This result is fully consistent with the significantly reduced potencies observed for most neutralising Abs and sera from past infections and/or vaccinations against the Omicron variant over wild-type Hu-1.[10, 12, 56] As G13-CRD's antiviral action is through blocking viral entry by binding to S protein glycans, it was expected to exhibit similar potencies against both the authentic BA.1 and Wuhan Hu-1 variants as their S-protein glycans are conserved. However, our results show this is not the case, implying that BA.1 variant must have evolved other ways (besides immune evasion mutations in S proteins) which allow it to evade neutralisation by G13-CRD. A likely mechanism here could be altering of the number of S proteins (and hence inter-S protein spacing), rendering these lectin-GNPs no longer spatially matched, thus reducing their ability to crosslink neighbouring S proteins to exert potent neutralisation. This hypothesised binding mechanism is supported by the observation that a potent anti-Ebola virus antibody has been identified by cryo-electron tomography to crosslink neighbouring S proteins *via* its two Fab arms.[57] Unfortunately, the cryo-EM structure of whole authentic Omicron variant with intact S proteins remains to be reported. Such information would help inform the design of specific spatially-matched lectin-NPs to enhance the potent and specific neutralisation of each SARS-CoV-2 variant.

The potential of exploiting multivalency to design potent, broad-spectrum anti-SARS-CoV-2 agents has been recently demonstrated with Abs. Linking IgGs together into an engineered pentameric IgM has been shown to not only greatly enhance IgG's viral neutralising potency (up to 230-fold) but also make it insensitive to a range of known immune evasion mutations. The engineered IgM exhibited high potencies against several SARS-CoV-2

variants (*e.g.* B.1.1.7 (Alpha), P.1 (Gamma) and B.1.351 (Beta)), with *in vivo* rodent models, and thus had great potential to act as a broad spectrum anti-SARS-CoV-2 agent.[58] The IgM's superior antiviral property over IgG is assigned to its larger size and higher binding valency, allowing it to bind/crosslink multiple S-proteins on virion surfaces in a way which is not possible by individual IgGs. The potential of multivalent binding in viral neutralisation has been further demonstrated with HIV. HIV is well-known for its ability of evading IgG neutralisation. Its low number of envelope protein (Env) trimers (~14 per virion) results in large inter-Env distances, making it impossible for individual IgGs to crosslink neighbouring Env proteins. Additionally, Env unique trimer structure with heavy glycan shields also prevent intra-Env crosslinking by IgGs. However, by linking two Fabs together *via* a rigid DNA spacer and systematically tuning the inter-Fab distance, greatly enhanced anti-HIV potency has been achieved for a Fab dimer that has the correct inter-Fab distance for intra-spike crosslinking.[59] Similarly, an engineered tetravalent DVD antibody (containing four variable-domains) has been shown to have 100-fold enhanced potency over its component divalent antibody against Crimean-Congo hemorrhagic fever virus.[60] These examples, as well as our G13-CRD based antivirals, clearly demonstrate that exploiting multivalency is a viable and attractive approach for the development of novel antivirals. While the immune evasion ability of SARS-CoV-2 variants has been almost exclusively considered on the basis of individual Ab-S protein interactions, contributions from multivalency evasion should not be ignored and need to be considered in order to develop more robust antivirals. Compared to antiviral designs, these lectin-GNP antivirals have several advantages. First, the GNP scaffold size, shape, and lectin valency and flexibility can be easily tuned to match the virus of interest. Second, lectins can be cheaply mass produced by recombinant bacterial expression without using animals. Third, viral glycosylation is common and viral glycosylation sites are mostly conserved and thus unlikely to be affected by viral variations. This makes viral glycans a potentially more robust target for developing antivirals over peptide epitopes targeted by most neutralising Abs. Finally, lectin-NP antivirals are especially useful against viral infections that lack effective neutralising Abs, or display Ab-enhanced viral infection (*e.g.* Dengue, Zika).[61, 62] Therefore, the lectin-NP based antivirals reported here represent a highly attractive, robust and more economical alternative to neutralising Abs in the fight against a wide range of viral infections.

References

1. H. Wang, K. R. Paulson, S. A. Pease, S. Watson, H. Comfort, P. Zheng, A. Y. Aravkin, C. Bisignano, R. M. Barber, T. Alam, J. E. Fuller, E. A. May, D. P. Jones, M. E. Frisch, C. Abbafati, C. Adolph, A. Allorant, J. O. Amlag, B. Bang-Jensen, G. J. Bertolacci, S. S. Bloom, A. Carter, E. Castro, S. Chakrabarti, J. Chattopadhyay, R. M. Cogen, J. K. Collins, K. Cooperrider, X. Dai, W. J. Dangel, F. Daoud, C. Dapper, A. Deen, B. B. Duncan, M. Erickson, S. B. Ewald, T. Fedosseeva, A. J. Ferrari, J. J. Frostad, N. Fullman, J. Gallagher, A. Gamkrelidze, G. Guo, J. He, M. Helak, N. J. Henry, E. N. Hulland, B. M. Huntley, M. Kereselidze, A. Lazzar-Atwood, K. E. LeGrand, A. Lindstrom, E. Linebarger, P. A. Lotufo, R. Lozano, B. Magistro, D. C. Malta, J. Månsson, A. M. Mantilla Herrera, F. Marinho, A. H. Mirkuzie, A. T. Misganaw, L. Monasta, P. Naik, S. Nomura, E. G. O'Brien, J. K. O'Halloran, L. T. Olana, S. M. Ostroff, L. Penberthy, R. C. Reiner Jr, G. Reinke, A. L. P. Ribeiro, D. F. Santomauro, M. I. Schmidt, D. H. Shaw, B. S. Sheena, A. Sholokhov, N. Skhvitaridze, R. J. D. Sorensen, E. E. Spurlock, R. Syailendrawati, R. Topor-Madry, C. E. Troeger, R. Walcott, A. Walker, C. S. Wiysonge, N. A. Worku, B. Zigler, D. M. Pigott, M. Naghavi, A. H. Mokdad, S. S. Lim, S. I. Hay, E. Gakidou and C. J. L. Murray, *Lancet*, 2022, **399**, 1513-1536.
2. Y. Cao, B. Su, X. Guo, W. Sun, Y. Deng, L. Bao, Q. Zhu, X. Zhang, Y. Zheng, C. Geng, X. Chai, R. He, X. Li, Q. Lv, H. Zhu, W. Deng, Y. Xu, Y. Wang, L. Qiao, Y. Tan, L. Song, G. Wang, X. Du, N. Gao, J. Liu, J. Xiao, X. D. Su, Z. Du, Y. Feng, C. Qin, C. Qin, R. Jin and X. S. Xie, *Cell*, 2020, **182**, 73-84 e16.
3. L. Enjuanes, I. Sola, S. Zuniga, J. M. Honrubia, M. Bello-Perez, A. Sanz-Bravo, E. Gonzalez-Miranda, J. Hurtado-Tamayo, R. Requena-Platek, L. Wang, D. Munoz-Santos, C. M. Sanchez, A. Esteban and J. Ripoll-Gomez, *Curr. Res. Immunol.*, 2022, **3**, 151-158.
4. F. P. Polack, S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. Pérez Marc, E. D. Moreira, C. Zerbini, R. Bailey, K. A. Swanson, S. Roychoudhury, K. Koury, P. Li, W. V. Kalina, D. Cooper, R. W. Frenck, L. L. Hammitt, Ö. Türeci, H. Nell, A. Schaefer, S. Ünal, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Şahin, K. U. Jansen and W. C. Gruber, *N. Engl. J. Med.*, 2020, **383**, 2603-2615.
5. C. G. Williams, A. S. Jureka, J. A. Silvas, A. M. Nicolini, S. A. Chvatal, J. Carlson-Stevermer, J. Oki, K. Holden and C. F. Basler, *Cell Rep.*, 2021, **36**, 109479.
6. L. R. Baden, H. M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S. A. Spector, N. Rouphael, C. B. Creech, J. McGettigan, S. Khetan, N. Segall, J. Solis, A. Brosz, C. Fierro, H. Schwartz, K. Neuzil, L. Corey, P. Gilbert, H. Janes, D. Follmann, M. Marovich, J. Mascola, L. Polakowski, J. Ledgerwood, B. S. Graham, H. Bennett, R. Pajon, C. Knightly, B. Leav, W. P. Deng, H. H. Zhou, S. Han, M. Ivarsson, J. Miller, T. Zaks and C. S. Grp, *N. Engl. J. Med.*, 2021, **384**, 403-416.
7. A. Baum, B. O. Fulton, E. Wloga, R. Copin, K. E. Pascal, V. Russo, S. Giordano, K. Lanza, N. Negron, M. Ni, Y. Wei, G. S. Atwal, A. J. Murphy, N. Stahl, G. D. Yancopoulos and C. A. Kyratsous, *Science*, 2020, **369**, 1014-+.
8. M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Kruger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N. H. Wu, A. Nitsche, M. A. Muller, C. Drosten and S. Pohlmann, *Cell*, 2020, **181**, 271-+.
9. M. Hoffmann, H. Kleine-Weber and S. Pohlmann, *Mol. Cell*, 2020, **78**, 779-784 e775.
10. M. Hoffmann, N. Kruger, S. Schulz, A. Cossmann, C. Rocha, A. Kempf, I. Nehlmeier, L. Graichen, A. S. Moldenhauer, M. S. Winkler, M. Lier, A. Dopfer-Jablonka, H. M. Jack, G. M. N. Behrens and S. Pohlmann, *Cell*, 2022, **185**, 447-456 e411.
11. M. Hoffmann, P. Arora, R. Gross, A. Seidel, B. F. Hornich, A. S. Hahn, N. Kruger, L. Graichen, H. Hofmann-Winkler, A. Kempf, M. S. Winkler, S. Schulz, H. M. Jack, B. Jahrsdorfer, H. Schrezenmeier, M. Muller, A. Kleger, J. Munch and S. Pohlmann, *Cell*, 2021, **184**, 2384-+.

12. W. Dejnirattisai, J. Huo, D. Zhou, J. Zahradnik, P. Supasa, C. Liu, H. M. E. Duyvesteyn, H. M. Ginn, A. J. Mentzer, A. Tuekprakhon, R. Nutalai, B. Wang, A. Dijokaite, S. Khan, O. Avinoam, M. Bahar, D. Skelly, S. Adele, S. A. Johnson, A. Amini, T. G. Ritter, C. Mason, C. Dold, D. Pan, S. Assadi, A. Bellass, N. Omo-Dare, D. Koeckerling, A. Flaxman, D. Jenkin, P. K. Aley, M. Voysey, S. A. Costa Clemens, F. G. Naveca, V. Nascimento, F. Nascimento, C. Fernandes da Costa, P. C. Resende, A. Pauvolid-Correa, M. M. Siqueira, V. Baillie, N. Serafin, G. Kwatra, K. Da Silva, S. A. Madhi, M. C. Nunes, T. Malik, P. J. M. Openshaw, J. K. Baillie, M. G. Semple, A. R. Townsend, K. A. Huang, T. K. Tan, M. W. Carroll, P. Klenerman, E. Barnes, S. J. Dunachie, B. Constantinides, H. Webster, D. Crook, A. J. Pollard, T. Lambe, O. Consortium, I. C. Consortium, N. G. Paterson, M. A. Williams, D. R. Hall, E. E. Fry, J. Mongkolsapaya, J. Ren, G. Schreiber, D. I. Stuart and G. R. Screaton, *Cell*, 2022, **185**, 467-484 e415.
13. Y. Cao, J. Wang, F. Jian, T. Xiao, W. Song, A. Yisimayi, W. Huang, Q. Li, P. Wang, R. An, J. Wang, Y. Wang, X. Niu, S. Yang, H. Liang, H. Sun, T. Li, Y. Yu, Q. Cui, S. Liu, X. Yang, S. Du, Z. Zhang, X. Hao, F. Shao, R. Jin, X. Wang, J. Xiao, Y. Wang and X. S. Xie, *Nature*, 2022, **602**, 657-663.
14. E. Callaway, *Nature*, 2021, **589**, 500-501.
15. H. Yao, Y. Song, Y. Chen, N. Wu, J. Xu, C. Sun, J. Zhang, T. Weng, Z. Zhang, Z. Wu, L. Cheng, D. Shi, X. Lu, J. Lei, M. Crispin, Y. Shi, L. Li and S. Li, *Cell*, 2020, **183**, 730-738 e713.
16. Y. Watanabe, J. D. Allen, D. Wrapp, J. S. McLellan and M. Crispin, *Science*, 2020, **369**, 330-333.
17. L. Casalino, Z. Gaieb, J. A. Goldsmith, C. K. Hjorth, A. C. Dommer, A. M. Harbison, C. A. Fogarty, E. P. Barros, B. C. Taylor, J. S. McLellan, E. Fadda and R. E. Amaro, *ACS Cent. Sci.*, 2020, **6**, 1722-1734.
18. L. M. Walker, M. Huber, K. J. Doores, E. Falkowska, R. Pejchal, J. P. Julien, S. K. Wang, A. Ramos, P. Y. Chan-Hui, M. Moyle, J. L. Mitcham, P. W. Hammond, O. A. Olsen, P. Phung, S. Fling, C. H. Wong, S. Phogat, T. Wrin, M. D. Simek, G. P. I. Protocol, W. C. Koff, I. A. Wilson, D. R. Burton and P. Poignard, *Nature*, 2011, **477**, 466-470.
19. A. Trkola, M. Purtscher, T. Muster, C. Ballaun, A. Buchacher, N. Sullivan, K. Srinivasan, J. Sodroski, J. P. Moore and H. Katinger, *J. Virol.*, 1996, **70**, 1100-1108.
20. R. Pejchal, K. J. Doores, L. M. Walker, R. Khayat, P. S. Huang, S. K. Wang, R. L. Stanfield, J. P. Julien, A. Ramos, M. Crispin, R. Depetris, U. Katpally, A. Marozsan, A. Cupo, S. Maloveste, Y. Liu, R. McBride, Y. Ito, R. W. Sanders, C. Ogohara, J. C. Paulson, T. Feizi, C. N. Scanlan, C. H. Wong, J. P. Moore, W. C. Olson, A. B. Ward, P. Poignard, W. R. Schief, D. R. Burton and I. A. Wilson, *Science*, 2011, **334**, 1097-1103.
21. D. A. Calarese, C. N. Scanlan, M. B. Zwick, S. Deechongkit, Y. Mimura, R. Kunert, P. Zhu, M. R. Wormald, R. L. Stanfield, K. H. Roux, J. W. Kelly, P. M. Rudd, R. A. Dwek, H. Katinger, D. R. Burton and I. A. Wilson, *Science*, 2003, **300**, 2065-2071.
22. C. R. Bertozzi and L. L. Kiessling, *Science*, 2001, **291**, 2357-2364.
23. K. Drickamer and M. E. Taylor, *Curr. Opin. Struct. Biol.*, 2015, **34**, 26-34.
24. A. Varki, *Essentials of glycobiology*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, Third edition. edn., 2017.
25. E. M. Coves-Datson, S. R. King, M. Legendre, A. Gupta, S. M. Chan, E. Gitlin, V. V. Kulkarni, J. Pantaleon Garcia, D. F. Smee, E. Lipka, S. E. Evans, E. B. Tarbet, A. Ono and D. M. Markovitz, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 2122-2132.
26. K. O. Francois and J. Balzarini, *Med. Res. Rev.*, 2012, **32**, 349-387.
27. C. A. Mitchell, K. Ramessar and B. R. O'Keefe, *Antiviral Res.*, 2017, **142**, 37-54.
28. F. A. Lempp, L. B. Soriaga, M. Montiel-Ruiz, F. Benigni, J. Noack, Y. J. Park, S. Bianchi, A. C. Walls, J. E. Bowen, J. Zhou, H. Kaiser, A. Joshi, M. Agostini, M. Meury, E. Dellota, Jr., S. Jaconi, E. Cameroni, J. Martinez-Picado, J. Vergara-Alert, N. Izquierdo-Useros, H. W. Virgin, A. Lanzavecchia, D. Veesler, L. A. Purcell, A. Telenti and D. Corti, *Nature*, 2021, **598**, 342-347.
29. R. Amraei, W. Yin, M. A. Napoleon, E. L. Suder, J. Berrigan, Q. Zhao, J. Olejnik, K. B. Chandler, C. Xia, J. Feldman, B. M. Hauser, T. M. Caradonna, A. G. Schmidt, S.

Gummuluru, E. Muhlberger, V. Chitalia, C. E. Costello and N. Rahimi, *ACS Cent. Sci.*, 2021, **7**, 1156-1165.
30. A. C. Walls, Y. J. Park, M. A. Tortorici, A. Wall, A. T. McGuire and D. Veesler, *Cell*, 2020, **181**, 281-292 e286.
31. D. J. Benton, A. G. Wrobel, P. Xu, C. Roustan, S. R. Martin, P. B. Rosenthal, J. J. Skehel and S. J. Gamblin, *Nature*, 2020, **588**, 327-330.
32. E. C. Dreaden, A. M. Alkilany, X. Huang, C. J. Murphy and M. A. El-Sayed, *Chem. Soc. Rev.*, 2012, **41**, 2740-2779.
33. M. C. Daniel and D. Astruc, *Chem. Rev.*, 2004, **104**, 293-346.
34. A. N. Baker, S. J. Richards, C. S. Guy, T. R. Congdon, M. Hasan, A. J. Zwetsloot, A. Gallo, J. R. Lewandowski, P. J. Stansfeld, A. Straube, M. Walker, S. Chessa, G. Pergolizzi, S. Dedola, R. A. Field and M. I. Gibson, *ACS Cent. Sci.*, 2020, **6**, 2046-2052.
35. T. T. S. Lew, K. M. M. Aung, S. Y. Ow, S. N. Amrun, L. Sutarlie, L. F. P. Ng and X. Su, *ACS Nano*, 2021.
36. L. Song, V. H. Ho, C. Chen, Z. Yang, D. Liu, R. Chen and D. Zhou, *Adv. Healthc. Mater.*, 2013, **2**, 275-280.
37. J. Liu and Y. Lu, *Nature Protocols*, 2006, **1**, 246-252.
38. Y. Guo, I. Nehlmeier, E. Poole, C. Sakonsinsiri, N. Hondow, A. Brown, Q. Li, S. Li, J. Whitworth, Z. Li, A. Yu, R. Brydson, W. B. Turnbull, S. Pohlmann and D. Zhou, *J. Am. Chem. Soc.*, 2017, **139**, 11833-11844.
39. M. Z. Chaudhry, K. Eschke, M. Hoffmann, M. Grashoff, L. Abassi, Y. Kim, L. Brunotte, S. Ludwig, A. Kröger, F. Klawonn, S. H. Pöhlmann and L. Cicin-Sain, *J. Virol.*, 2022, **96**, e02186-02121.
40. M. Hoffmann, H. Hofmann-Winkler, J. C. Smith, N. Kruger, P. Arora, L. K. Sorensen, O. S. Sogaard, J. B. Hasselstrom, M. Winkler, T. Hempel, L. Raich, S. Olsson, O. Danov, D. Jonigk, T. Yamazoe, K. Yamatsuta, H. Mizuno, S. Ludwig, F. Noe, M. Kjolby, A. Braun, J. M. Sheltzer and S. Pohlmann, *EBioMedicine*, 2021, **65**, 103255.
41. H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327-1335.
42. Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis and K. Drickamer, *Nat. Struct. Mol. Biol.*, 2004, **11**, 591-598.
43. H. Feinberg, R. Castelli, K. Drickamer, P. H. Seeberger and W. I. Weis, *J. Biol. Chem.*, 2007, **282**, 4202-4209.
44. Y. Guo, C. Sakonsinsiri, I. Nehlmeier, M. A. Fascione, H. Zhang, W. Wang, S. Pohlmann, W. B. Turnbull and D. Zhou, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 4738-4742.
45. D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo and D. Zhou, *J. Am. Chem. Soc.*, 2020, **142**, 18022-18034.
46. D. J. Zhou, A. Bruckbauer, C. Abell, D. Klenerman and D. J. Kang, *Adv. Mater.*, 2005, **17**, 1243-+.
47. D. Zhou, A. Bruckbauer, L. M. Ying, C. Abell and D. Klenerman, *Nano Lett.*, 2003, **3**, 1517-1520.
48. C. B. Rosen and M. B. Francis, *Nat. Chem. Biol.*, 2017, **13**, 697-705.
49. L. Song, V. H. Ho, C. Chen, Z. Yang, D. Liu, R. Chen and D. Zhou, *Adv Healthc Mater*, 2013, **2**, 275-280.
50. T. Pons, I. L. Medintz, K. E. Sapsford, S. Higashiya, A. F. Grimes, D. S. English and H. Mattoussi, *Nano Lett.*, 2007, **7**, 3157-3164.
51. T. L. Jennings, M. P. Singh and G. F. Strouse, *J. Am. Chem. Soc.*, 2006, **128**, 5462-5467.
52. A. Holla and A. Skerra, *Protein Eng. Des. Sel.*, 2011, **24**, 659-669.
53. F. Schmidt, Y. Weisblum, F. Muecksch, H. H. Hoffmann, E. Michailidis, J. C. C. Lorenzi, P. Mendoza, M. Rutkowska, E. Bednarski, C. Gaebler, M. Agudelo, A. Cho, Z. Wang, A. Gazumyan, M. Cipolla, M. Caskey, D. F. Robbiani, M. C. Nussenzweig, C. M. Rice, T. Hatziioannou and P. D. Bieniasz, *J Exp Med*, 2020, **217**.

54. Z. Ke, J. Oton, K. Qu, M. Cortese, V. Zila, L. McKeane, T. Nakane, J. Zivanov, C. J. Neufeldt, B. Cerikan, J. M. Lu, J. Peukes, X. Xiong, H. G. Krausslich, S. H. W. Scheres, R. Bartenschlager and J. A. G. Briggs, *Nature*, 2020, **588**, 498-502.
55. P. Ge, J. Tsao, S. Schein, T. J. Green, M. Luo and Z. H. Zhou, *Science*, 2010, **327**, 689-693.
56. M. McCallum, N. Czudnochowski, L. E. Rosen, S. K. Zepeda, J. E. Bowen, A. C. Walls, K. Hauser, A. Joshi, C. Stewart, J. R. Dillen, A. E. Powell, T. I. Croll, J. Nix, H. W. Virgin, D. Corti, G. Snell and D. Veesler, *Science*, 2022, **375**, 864-868.
57. E. E. Tran, E. A. Nelson, P. Bonagiri, J. A. Simmons, C. J. Shoemaker, C. S. Schmaljohn, G. P. Kobinger, L. Zeitlin, S. Subramaniam and J. M. White, *J. Virol.*, 2016, **90**, 7618-7627.
58. Z. Ku, X. Xie, P. R. Hinton, X. Liu, X. Ye, A. E. Muruato, D. C. Ng, S. Biswas, J. Zou, Y. Liu, D. Pandya, V. D. Menachery, S. Rahman, Y. A. Cao, H. Deng, W. Xiong, K. B. Carlin, J. Liu, H. Su, E. J. Haanes, B. A. Keyt, N. Zhang, S. F. Carroll, P. Y. Shi and Z. An, *Nature*, 2021, **595**, 718-723.
59. R. P. Galimidi, J. S. Klein, M. S. Politzer, S. Bai, M. S. Seaman, M. C. Nussenzweig, A. P. West, Jr. and P. J. Bjorkman, *Cell*, 2015, **160**, 433-446.
60. J. M. Fels, D. P. Maurer, A. S. Herbert, A. S. Wirchnianski, O. Vergnolle, R. W. Cross, D. M. Abelson, C. L. Moyer, A. K. Mishra, J. T. Aguilan, A. I. Kuehne, N. T. Pauli, R. R. Bakken, E. K. Nyakatura, J. Hellert, G. Quevedo, L. Lobel, S. Balinandi, J. J. Lutwama, L. Zeitlin, T. W. Geisbert, F. A. Rey, S. Sidoli, J. S. McLellan, J. R. Lai, Z. A. Bornholdt, J. M. Dye, L. M. Walker and K. Chandran, *Cell*, 2021, **184**, 3486-3501 e3421.
61. S. B. Halstead and E. J. O'Rourke, *Nature*, 1977, **265**, 739-741.
62. W. Dejnirattisai, A. Jumnainsong, N. Onsirisakul, P. Fitton, S. Vasanawathana, W. Limpitikul, C. Puttikhunt, C. Edwards, T. Duangchinda, S. Supasa, K. Chawansuntati, P. Malasit, J. Mongkolsapaya and G. Screaton, *Science*, 2010, **328**, 745-748.

Chapter 8
Concluding Remarks and Future Directions

Multimeric glycan binding proteins (lectins) are widespread in biology and play pivotal role in the immune system as pattern recognition receptors, engaging in processes such as cell-cell adhesion, pathogen/antigen internalisation and pathogen-associated cytokine/chemokine modulation. However, due to their close association with pathogens and other immune processes, they have been associated with a range of immunological diseases, such as viral infection, which make them a sought after target for many therapeutic applications. These lectin-glycan interactions (LGIs) are facilitated by the use of multivalency, whereby, because monovalent LGIs often demonstrate relatively weak, bio-inactive affinities, the integration of multiple binding sites per lectin, and/or the engagement of multiple lectins, is required to afford the significant binding affinity enhancement needed to facilitate any subsequent biological functionality. In addition, by utilising multiple binding sites in specific spatial arrangements, multivalent lectins can not only differentiate between the different types of carbohydrates which best compliment their individual binding sites, but are also able to distinguish between the specific spatial arrangements of glycans, whereby glycans which match poorly with the arrangements of the lectin binding sites may form much weaker interactions or resort to less favourable binding modes. These properties mean that conventional monovalent therapeutics are often unsuitable for targeting lectins, thus targeting agents bearing multiple glycans have been developed (see **Section 1.3.2**). However, the multivalent specificity of such lectins thus adds an additional level of complexity which needs to be taken into account when designing such glyco-therapeutics. Aspects such as glycoconjugate size, shape, flexibility and glycan density all influence the efficiency and specificity of binding. However, despite this, limited research has been done to elucidate the structural and biophysical bases behind how these the key design elements influence the targeting of specific lectins. In order to solve this, however, good understanding of the biophysical and structural mechanisms behind multivalent lectin binding, and how glycan design affects these mechanisms, needs to be elucidated in order to better inform the rational design of potential therapeutics which target such lectins.

8.1 Determinations of Biophysical Contributions Influencing MLGIs

In this thesis, two closely related tetrameric mannose-binding membrane lectins, DC-SIGN and DC-SIGNR (referred to as DC-SIGN/R), were chosen as a case study for investigating how differences lectin structure affects glycan binding and how this influences glycoconjugate design. These proteins share ~80% of the same amino acid identity, have near-identical mannose binding motifs and have both been associated with viral infection.[1, 2] However, despite these similarities, they have been demonstrated to show large differences in binding to some of the same viruses.[3-6] Furthermore, our group previously demonstrated that the two proteins display large variability in binding affinities and binding modes to the same glyconanoparticles (glycan-NPs; i.e. glycan-quantum dots and glycan-gold nanoparticles), whereby DC-SIGN has been shown to bind with all four binding sites to individual glycan-NPs with a much higher affinity compared to DC-SIGNR which was observed to crosslink between glycan-NPs with outward facing binding site orientations.[7, 8] Here, a highly sensitive FRET-based glycan-quantum dot (glycan-QD) approach was developed to dissect the thermodynamic and kinetic contributions towards these different binding behaviours (**Chapter 3**). Here it was observed that both DC-SIGN and DC-SIGNR MLGIs resulted in similar enthalpically driven interactions which were ~4-fold that of the monovalent binding observed in the literature, which was thus attributed to the binding of all four binding sites for both proteins. Furthermore, the lower affinity enhancement of DC-SIGNR's crosslinking binding mode was attributed to a larger entropic penalty and a much slower overall rate of association in comparison to DC-SIGN, which was attributed to the slow formation of crosslinking over time.

This FRET-based approach of analysing the solution based thermodynamics and kinetics thus provides an excellent platform for the investigation of MLGIs without the additional complexities found in their native surface-phase environments (*i.e.* neighbouring receptors, lateral fluidity, diffusion kinetics, etc). Furthermore, these techniques were able to be applied to analyse the energetic and kinetic contributions of a key structural component of DC-SIGN (see **Section 8.2**). Therefore, these techniques could easily be applied to analyse the enthalpic, entropic and kinetic influences of different glycoconjugate designs on binding to such lectins. For example, it was hypothesised that the main contribution to the entropic penalty for both DC-SIGN and DC-SIGNR may have been attributed to the long flexible

undeca(ethylene glycol) (EG_{11}) linkers which tether the glycans to the QD surface. Thus, using smaller and/or more rigid linkers may reduce this entropic effect, providing a stronger overall MLGI affinity enhancement. Preliminarily results measuring the enthalpic contributions to DC-SIGN binding with mono-mannose capped QDs (QD-DiMan) with shorter tri(ethylene glycol) (EG_3) linkers demonstrated that indeed a large entropic enhancement can be observed (**Appendix F.2**). However, this system also resulted in a low enthalpic enhancement in comparison to the literature monovalent glycan binding affinity. This may be caused by a lack of complementarity between more rigid glycan arrangements with the binding site arrangement of DC-SIGN. Thus, further investigation will need to be conducted to elucidate how different design rules, such as these, individually contribute to the biophysical parameters of MLGIs, in order to ensure the optimal potency and specificity of lectin targets.

8.2 Determination of Structural Contributions Influencing MLGIs

In addition to using QD-FRET to elucidate key differences between the MLGIs of DC-SIGN and DC-SIGNR, these techniques were also employed to investigate a structural rationale behind the multivalent specificity of DC-SIGN (**Chapter 3**). Here, a small unstructured 16-amino acid motif at the C-terminal, located at the flexible junction between the neck domain and the carbohydrate recognition domain (CRD) of DC-SIGN, which is not present in DC-SIGNR, was investigated for its structural influence on binding. Here, it was observed that the removal of this C-terminal segment (termed DC-SIGN-C) resulted in a complete shift of the enthalpy-entropy profile (*i.e.* rendering binding entropically driven), despite retaining a similar overall binding affinity, and was able to engage in a small degree of crosslinking (as observed by STEM imaging). This C-terminal segment was thus associated with providing a structural restraint to the CRD arrangement of DC-SIGN; where, though DC-SIGN-C still retained good binding, it resulted in a more flexible inter-CRD arrangement which was able to form entropically favourable interactions with flexible glycan-QDs. These results suggested that this region is associated with ensuring the spatial specificity of DC-SIGN towards specific glycans. Interestingly, a preliminary study into investigating the effects DC-SIGN-C with QD-Man of varying glycan densities (achieved using various proportions of an inert OH spacer ligand; **Appendix F.1**), showed that, where DC-SIGN tended to be highly dependent on the inter-glycan spacing, this specificity was

completely lost without the C-terminal segment. This is consistent with the aforementioned hypothesis, where without the C-terminal segment, the CRDs may be more able to rearrange themselves to better match the different arrangements of the glycan display.

Though these studies provided some useful insight into the structural rationale for the specificity of DC-SIGN, more studies may need to be employed to provide a structural rationale behind the differences between DC-SIGN and DC-SIGNR tetrameric binding. A preliminary investigation into the influence of the amino acids before this C-terminal segment was conducted by replacing the C-terminal 47-amino acids of DC-SIGN with the corresponding 30 C-terminal amino acids of DC-SIGNR (termed DC-SIGN+R). Interestingly, by replacing this segment a complete collapse in the binding capacity of DC-SIGN is observed upon binding of DC-SIGN+R with QD-Man (**Appendix F.1**). This may suggest that, with this region being so close to the flexible neck/CRD junction, it may play a key part in the different CRD arrangements between DC-SIGN and DC-SIGNR. However, more comparative studies need to be conducted to clarify the thermodynamic/kinetic changes DC-SIGN+R may afford.

As knowledge of the multimeric arrangements of the four CRDs is so crucial to optimising the multivalency of MLGIs, this thesis therefore demonstrated a new technique which employs single molecule FRET to elucidate the inter-binding site arrangements of DC-SIGN and DC-SIGNR (**Chapter 6**). This was achievable by a double-purification approach to label only 2 out of 4 of the CRDs in DC-SIGN/R with a FRET pair dyes through the heterotetramerisation of labelled mannose binding and unlabelled galactose binding monomeric units. This was compared against MD simulations of the two proteins to derive more empirically informed models for their CRD arrangements. These may be used to inform the specific arrangements of glycoconjugates designed for targeting these proteins with very high specificity.

Moreover, this thesis also describes how the tetrameric surface area of these lectins influences whether or not glycoconjugates are able to bridge all of the CRDs to form strong simultaneous binding. Here, glycan-conjugated quantum rods (QRs) were developed and exploited for their QR-based FRET readout for characterising binding affinity, and hydrodynamic size analysis and TEM imaging for characterisation complexation (**Chapter 4**). This technique was able to reveal that DC-SIGNR was able to crosslink with the spherical end sections of the glycan-QRs (which most represented the glycan

display of glycan-QDs), however, it was able to form much stronger simultaneous binding of all four CRDs with the cylindrical central region of the QRs. This was found to result from the more planar surface curvature of the cylindrical section which was able to bridge the CRDs in a way that the spherical section could not. This demonstrates both the importance of being able to identify the binding site area of multimeric lectins as well as the rational design of glycoconjugate shape and size to be able to match this contact area in order to optimise multivalency.

8.3 Determination of Surface-Phase Influences on MLGIs

This thesis mainly demonstrates the use of solution based studies in order to elucidate raw biophysical and structural parameters related to multivalent lectin·glycan-NP binding without the additional complexities present upon surfaces. However, such systems do not represent the true states of membrane receptors in their native settings. This has been observed by our group when translating solution-phase assays to the cell studies.[7] Here, we have demonstrated a new lectin-functionalised supported lipid bilayer (SLB) model system which has been used to better mimic the native states of lectins on cell membranes (**Chapter 5**). QCM-D was employed to analyse the binding behaviours of mono-mannose capped gold nanoparticles (GNP-Man) upon DC-SIGN- and DC-SIGNR- functionalised SLBs using different lectin densities. Here it was observed that, where DC-SIGN binding appears to be relatively independent of the initial inter-lectin distance, the affinity of GNP-Man·DC-SIGNR binding appears to weaken at larger initial inter-lectin spacings. Here it was hypothesised that DC-SIGNR may be able to form crosslinking interactions with GNP-Man, as observed in solution, whereby closer inter-lectin surface arrangements encourage crosslinking much more easily. These GNP-Man·DC-SIGNR constructs were also trialled under cyro-TEM imaging and demonstrated good evidence for these crosslinking interactions. However, these have not been analysed for DC-SIGN or for other lectin surface densities, which may provide further insight into the physical binding arrangements of GNP-Man at the lectin-SLB surfaces. In addition, QCM only provides estimations of the binding affinities due to the large uncertainties of the true degree of binding caused by hydrodynamically coupled mass. Thus, more robust techniques, such as spectroscopic ellipsometry, may be able to provide more accurate measurements of these surface phase biophysical parameters.

Additionally, investigation of the influence of glyconanoparticle size, shape and glycan composition on lectin binding affinity and mode at lipid bilayer surfaces would also be worth exploring. Such experiments would lead to a better understanding of how different glycoconjugates interact at the membrane surface. Furthermore, such studies would highlight the discrepancies between solution and surface-phase assays. For example, it is reasonable to hypothesise that with increasing GNP size, interesting steric shielding or inter-lectin multivalency may influence the binding of both DC-SIGN and DC-SIGNR in ways that cannot be predicted *via* solution-phase assays alone.

8.4 Therapeutic Application

Though the studies highlighted here have provided some important details for understanding the effect of lectin multimeric structure on lectin-glycan binding and the influence of glycoconjugate design on the specificity and potency of MLGIs, the long-term aim for these findings is to inform the rational design of glyco-therapeutics for medical application. For example, using these findings some key themes can be observed: 1. Glyconanoparticle surface curvature needs to allow for adequate bridging of all four of the CRDs (as demonstrated in **Chapter 4**), which is particularly important for DC-SIGNR which displays a much broader inter-CRD arrangement (as demonstrated by smFRET; **Chapter 6**); 2. Glycan ligand flexibility may need to be kept minimal in order to achieve a lower entropic penalty upon binding (as demonstrated in **Chapter 3**); 3. However, in order for this to still return optimal enthalpy, the glycans need to be arranged to specifically match the inter-binding site arrangements, which can be acquired from the smFRET-informed MD model (**Chapter 6**). Such designs should then be investigated for their surface-phase binding using the lectin-SLB models designed in **Chapter 5**, in order to investigate the more complex influences of inter-lectin multivalency at membrane surfaces. These rationally designed glycoconjugates may be highly potent and highly specific, making them excellent for targeting lectins for therapeutic applications such as antivirals.

Beyond the biophysical and structural mechanisms behind the interactions of such glycoconjugates, however, it is also important to understand their biological consequences with lectins at cell surfaces (*e.g.* internalisation, cytokine modulation, etc). This may in fact vary with differences in affinity/avidity, binding mode, binding kinetics or inter-lectin

multivalency, however, this has not yet been investigated. To probe this, a library of glycan-NPs which incorporate different design elements (*i.e.* carbohydrate type, size, shape, ligand flexibility or glycan density) could be investigated to monitor the effect of these on DC-SIGN-associated processes within the context of their native cell environments. For example, for DC-SIGN$^+$ dendritic cells, it has been shown that the type of carbohydrate can have a profound influence on the intracellular signalling pathways required for different cytokine expression (see **Section 1.2.3**). However, no studies have investigated whether these other glycoconjugate design elements influence lectin-associated cytokine expression. Such influences would have huge medical implications as different ligand designs could potentially be used to activate the immune system or encourage the different immune responses towards humoral or cellular immune responses or immune tolerance, *via* DC-SIGN mediated signalling pathways. This could apply to various immune-related problems such as treating cancer cells or diseases such as HIV, or by helping relax the immune system in cases such as autoimmune diseases, transplant rejections or allergies.

In addition to potential glycan-based nano-therapeutics, this thesis demonstrates the development of potential lectin-based nano-therapeutics. Here, the successful functionalisation and conjugation of lectins onto nanoparticles has demonstrated to yield highly potent neutralisers of different variants of the SARS-CoV-2 virus. This is accomplished by targeting and binding to the glycosylation sites of the SARS-CoV-2 S protein which are conserved within all of the SARS-CoV-2 variants of concern to the World Health Organisation (WHO), to date, and thus provide an excellent alternative antiviral strategy for targeting a broad range of virus variants. Though this antiviral strategy has proven to be very effective for the inhibition of viral entry in a very controlled environment, more research may need to be done to ensure that such lectin-NP designs provide robust specificity towards virus targets, *i.e.* either in more complex systems such as animal models or against other glycosylated species. For example, our proposed mechanism for inhibition is that the lectin-NPs crosslink neighbouring S proteins on the same virus. If this is true then the size and valency of the lectin-NP and the number of lectins on the virus surface all play a key role in the binding potency, and thus specificity, of the viral inhibitor. Such differentiation has been observed in our studies, however, more concrete evidence for this mechanistic route needs to be obtained by imaging the lectin-NPs on the virion surface, *e.g. via* electron microscopy. If this is indeed the case then by identifying the inter-S protein distance of glycosylated proteins on similar viruses, such lectin-NPs

can be rationally designed to match these distances to provide more potent and specific inhibitors against such viruses.

References

1. H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327-1335.
2. H. Feinberg, D. A. Mitchell, K. Drickamer and W. I. Weis, *Science*, 2001, **294**, 2163-2166.
3. M. Shimojima, A. Takenouchi, H. Shimoda, N. Kimura and K. Maeda, *Arch. Virol.*, 2014, **159**, 2023-2031.
4. N. P. Chung, S. K. Breun, A. Bashirova, J. G. Baumann, T. D. Martin, J. M. Karamchandani, J. W. Rausch, S. F. Le Grice, L. Wu, M. Carrington and V. N. Kewalramani, *J. Biol. Chem.*, 2010, **285**, 2100-2112.
5. C. W. Davis, H. Y. Nguyen, S. L. Hanna, M. D. Sanchez, R. W. Doms and T. C. Pierson, *J. Virol.*, 2006, **80**, 1290-1301.
6. A. Marzi, T. Gramberg, G. Simmons, P. Möller, A. J. Rennekamp, M. Krumbiegel, M. Geier, J. Eisemann, N. Turza, B. Saunier, A. Steinkasserer, S. Becker, P. Bates, H. Hofmann and S. Pöhlmann, *J. Virol.*, 2004, **78**, 12090-12095.
7. D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo and D. Zhou, *J. Am. Chem. Soc.*, 2020, **142**, 18022-18034.
8. Y. Guo, I. Nehlmeier, E. Poole, C. Sakonsinsiri, N. Hondow, A. Brown, Q. Li, S. Li, J. Whitworth, Z. Li, A. Yu, R. Brydson, W. B. Turnbull, S. Pohlmann and D. Zhou, *J. Am. Chem. Soc.*, 2017, **139**, 11833-11844.

Appendix A
General Information

A.1 Amino acid sequences

Comparative amino acid sequence of all protein variants with key motifs highlighted (yellow: His_6 residue; grey: position of Cys mutation for labelling; blue: galactose binding motif; green: segment of DC-SIGN replaced with that of DC-SIGNR):

```
CRD               ------------------------------------------------------------
DC-SIGN           AKVPSSISQEQSRQDAIYQNLTQLKAAVGELSEKSKLQEIYQELTQLKAAVGELPEKSKL
DC-SIGN-C         AKVPSSISQEQSRQDAIYQNLTQLKAAVGELSEKSKLQEIYQELTQLKAAVGELPEKSKL
GBDC-SIGN         AKVPSSISQEQSRQDAIYQNLTQLKAAVGELSEKSKLQEIYQELTQLKAAVGELPEKSKL
His6-DC-SIGN      ----------------------AHHHHHHGELSEKSKLQEIYQELTQLKAAVGELPEKSKL
DC-SIGN+R         AKVPSSISQEQSRQDAIYQNLTQLKAAVGELSEKSKLQEIYQELTQLKAAVGELPEKSKL
DC-SIGNR          AKVPSSLSQEQSEQDAIYQNLTQLKAAVGELSEKSKLQEIYQELTQLKAAVGELPEKSKL
GBDC-SIGNR        AKVPSSLSQEQSEQDAIYQNLTQLKAAVGELSEKSKLQEIYQELTQLKAAVGELPEKSKL
His6-DC-SIGNR     ----------------------AHHHHHHGELSEKSKLQEIYQELTQLKAAVGELPEKSKL

CRD               ------------------------------------------------------------
DC-SIGN           QEIYQELTRLKAAVGELPEKSKLQEIYQELTWLKAAVGELPEKSKMQEIYQELTRLKAAV
DC-SIGN-C         QEIYQELTRLKAAVGELPEKSKLQEIYQELTWLKAAVGELPEKSKMQEIYQELTRLKAAV
GBDC-SIGN         QEIYQELTRLKAAVGELPEKSKLQEIYQELTWLKAAVGELPEKSKMQEIYQELTRLKAAV
His6-DC-SIGN      QEIYQELTRLKAAVGELPEKSKLQEIYQELTWLKAAVGELPEKSKMQEIYQELTRLKAAV
DC-SIGN+R         QEIYQELTRLKAAVGELPEKSKLQEIYQELTWLKAAVGELPEKSKMQEIYQELTRLKAAV
DC-SIGNR          QEIYQELTRLKAAVGELPEKSKLQEIYQELTRLKAAVGELPEKSKLQEIYQELTRLKAAV
GBDC-SIGNR        QEIYQELTRLKAAVGELPEKSKLQEIYQELTRLKAAVGELPEKSKLQEIYQELTRLKAAV
His6-DC-SIGNR     QEIYQELTRLKAAVGELPEKSKLQEIYQELTRLKAAVGELPEKSKLQEIYQELTRLKAAV

CRD               ------------------------------------------------------------
DC-SIGN           GELPEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQ
DC-SIGN-C         GELPEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQ
GBDC-SIGN         GELPEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQ
His6-DC-SIGN      GELPEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQ
DC-SIGN+R         GELPEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQELTRLKAAVGELPEKSKQQEIYQ
DC-SIGNR          GELPEKSKLQEIYQELTELKAAVGELPEKSKLQEIYQELTQLKAAVGELPDQSKQQQIYQ
GBDC-SIGNR        GELPEKSKLQEIYQELTELKAAVGELPEKSKLQEIYQELTQLKAAVGELPDQSKQQQIYQ
His6-DC-SIGNR     GELPEKSKLQEIYQELTELKAAVGELPEKSKLQEIYQELTQLKAAVGELPDQSKQQQIYQ

CRD               --------AERLCHPCPWEWTFFQGNCYFMSNSQRNWHDSITACKEVGAQLVVIKSAEEQ
DC-SIGN           ELTQLKAAVERLCHPCPWEWTFFQGNCYFMSNSQRNWHDSITACKEVGAQLVVIKSAEEQ
DC-SIGN-C         ELTQLKAAVERLCHPCPWEWTFFQGNCYFMSNSQRNWHDSITACKEVGAQLVVIKSAEEQ
GBDC-SIGN         ELTQLKAAVERLCHPCPWEWTFFQGNCYFMSNSQRNWHDSITACKEVGAQLVVIKSAEEQ
His6-DC-SIGN      ELTQLKAAVERLCHPCPWEWTFFQGNCYFMSNSQRNWHDSITACKEVGAQLVVIKSAEEQ
DC-SIGN+R         ELTQLKAAVERLCHPCPWEWTFFQGNCYFMSNSQRNWHDSITACKEVGAQLVVIKSAEEQ
DC-SIGNR          ELTDLKTAFERLCRHCPKDWTFFQGNCYFMSNSQRNWHDSVTACQEVRAQLVVIKTAEEQ
GBDC-SIGNR        ELTDLKTAFERLCRHCPKDWTFFQGNCYFMSNSQRNWHDSVTACQEVRAQLVVIKTAEEQ
His6-DC-SIGNR     ELTDLKTAFERLCRHCPKDWTFFQGNCYFMSNSQRNWHDSVTACQEVRAQLVVIKTAEEQ

CRD               NFLQLQSSRSNRFTWMGLSDLNQEGTWQWVDGSPLLPSFKQYWNRGEPNN------VGEE
DC-SIGN           NFLQLQSSRSNRFTWMGLSDLNQEGTWQWVDGSPLLPSFKQYWNRGEPNN------VGEE
DC-SIGN-C         NFLQLQSSRSNRFTWMGLSDLNQEGTWQWVDGSPLLPSFKQYWNRGEPNN------VGEE
GBDC-SIGN         NFLQLQSSRSNRFTWMGLSDLNQEGTWQWVDGSPLLPSFKQYWNRGQPDNWYGHGLGGGE
His6-DC-SIGN      NFLQLQSSRSNRFTWMGLSDLNQEGTWQWVDGSPLLPSFKQYWNRGEPNN------VGEE
DC-SIGN+R         NFLQLQSSRSNRFTWMGLSDLNQEGTWQWVDGSPLLPSFKQYWNRGEPNN------VGEE
DC-SIGNR          NFLQLQTSRSNRFSWMGLSDLNQEGTWQWVDGSPLSPSFQRYWNSGEPNN------SGNE
```

```
GBDC-SIGNR          NFLQLQTSRSNRFSWMGLSDLNQEGTWQWVDGSPLSPSFQRYWNSGQPDNWYGHGLGGGE
His6-DC-SIGNR       NFLQLQTSRSNRFSWMGLSDLNQEGTWQWVDGSPLSPSFQRYWNSGEPNN------SGNE


CRD                 DCAEFSGNGWNDDKCNLAKFWICKKSAASCSRDEEQFLSPAPATPNPPPA
DC-SIGN             DCAEFSGNGWNDDKCNLAKFWICKKSAASCSRDEEQFLSPAPATPNPPPA
DC-SIGN-C           DCAEFSGNGWNDDKCNLAKFWICKKSAASCSRDE----------------
GBDC-SIGN           DCAEFSGNGWNDDKCNLAKFWICKKSAASCSRDEEQFLSPAPATPNPPPA
His6-DC-SIGN        DCAEFSGNGWNDDKCNLAKFWICKKSAASCSRDEEQFLSPAPATPNPPPA
DC-SIGN+R           DCAEFSGSGWNDNRCDVDNYWICKKPAA-CFRDE----------------
DC-SIGNR            DCAEFSGSGWNDNRCDVDNYWICKKPAA-CFRDE----------------
GBDC-SIGNR          DCAEFSGSGWNDNRCDVDNYWICKKPAA-CFRDE----------------
His6-DC-SIGNR       DCAEFSGSGWNDNRCDVDNYWICKKPAA-CFRDE----------------
```

A.2 SARS-CoV-2 S protein mutations

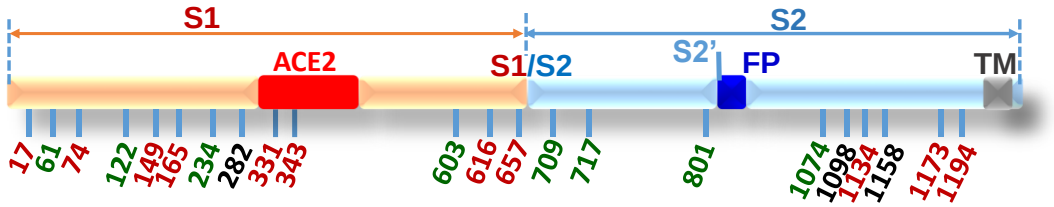

Figure A.1 SARS-CoV-2 S protein glycosylation sites (redrawn based on the data reported from Ref. [9]).

Table A.1 Summary of SARS-CoV-2 protein mutation sites in different variants (data obtained from Ref. [10]).

WHO label	Lineage	Spike mutations or deletions (Δ)
Alpha (α)	B.1.1.7	Δ69-70; Δ144; E484K; 452R; N501Y; A570D; D614G; P681H; T716I; S982A; D1118H
Beta (β)	B.1.351	K417N; E484K; N501Y; D614G; A701V
Gama (γ)	P1	K417T; E484K; N501Y; D614G; H655Y
Delta (δ)	B.1.617.2	L452R; T478K; N501Y; D614G; H655Y
Omicron	B.1.1.529 BA.1	A67V; Δ69-70; T95I; G142D; Δ143-145; Δ211-212; ins214EPE; G339D; S371L; S373P; S375F; K417N; N440K; G446S; S477N; T478K; E484A; Q493K; G496S; Q498R; N501Y; Y505H; T547K; D614G; H655Y; N679K; P681H; N764K; D796Y; N856K; Q954H; N969K; L981F

BA.4	T19I; LPPA24-27S; Δ69-70; G142D; V213G; G339D; S371F; S373P; S375F; T376A; D405N; R408S; K417N; N440K; L452R; S477N; T478K; E484A; F486V; Q498R; N501Y; Y505H; D614G; H655Y; N679K; P681H; N764K; D796Y; Q954H; N969K
BA.5	T19I; LPPA24-27S; Δ69-70; G142D; V213G; G339D; S371F; S373P; S375F; T376A; D405N; R408S; K417N; N440K; L452R; S477N; T478K; E484A; F486V; Q498R; N501Y; Y505H; D614G; H655Y; N679K; P681H; N764K; D796Y; Q954H; N969K

Appendix B
Glycan-NP Characterisation

B.1 NP Glycan Display

The average inter-glycan distance of each ligand (X) on spherical NP surfaces can be calculated from the hydrodynamic diameter (D_h) using **Equation B.1-B.2**.[8, 11] k is the average surface area of the glycan ligand which can be obtained from the NP surface area (A) divided by the number of ligands per NP (N), shown in **Equation B.1**. θ is the average deflection angle between glycans (Eq B.3).

$$k = \frac{A}{N} = \frac{\pi {D_\mathrm{h}}^2}{N} \quad \text{B.1}$$

$$X = 2\sqrt{\frac{k}{\pi}} = \frac{2D_\mathrm{h}}{\sqrt{N}} \quad \text{B.2}$$

$$\theta = 360 \cdot \frac{X}{\pi D_h} \quad \text{B.3}$$

Rods can be assumed to be formed for a cylindrical central section of height H and diameter D_h, sandwiched between two hemispheres of diameter D_h (with a total rod length of L_h). Here, k is instead obtained by **Equation B.5**. θ and X can then be calculated for the spherical end sections in the same way as that of spherical NPs (**Equation B.2, B.3**) from the estimated D_h derived in **Appendix E.2**.

$$A = \pi D_\mathrm{h} H + \pi {D_\mathrm{h}}^2 = \pi D_\mathrm{h} L_\mathrm{h} \quad \text{B.4}$$

$$k = \frac{\pi D_\mathrm{h} L}{N} \quad \text{B.5}$$

Table B.1 Summary of physical parameters derived from the hydrodynamic diameter (D_h) and glycan valency (N) of glycan-NPs. Where, L_h is the hydrodynamic length, k is the average surface area of each ligand, θ is the average deflection angle, and X is the inter-glycan distance.

Sample	D_h (L_h) / nm	N	k / nm^2	X / nm	θ / deg
QD-DiMan	12.4 ± 3.0	210 ± 70	2.3 ± 0.7	1.7 ± 0.3	16 ± 3
QR-DiMan	8.9 ± 2.1 (20.6 ± 6.1)	370 ± 30	1.6 ± 0.6	1.4 ± 0.3	18 ± 6 [a]
G5-Man	10.4 ± 2.8	750 ± 50	0.5 ± 0.2	0.8 ± 0.2	8 ± 3

[a] At the spherical end regions

B.2 Spectral overlap and Förster radius of QD-Atto594 FRET pair

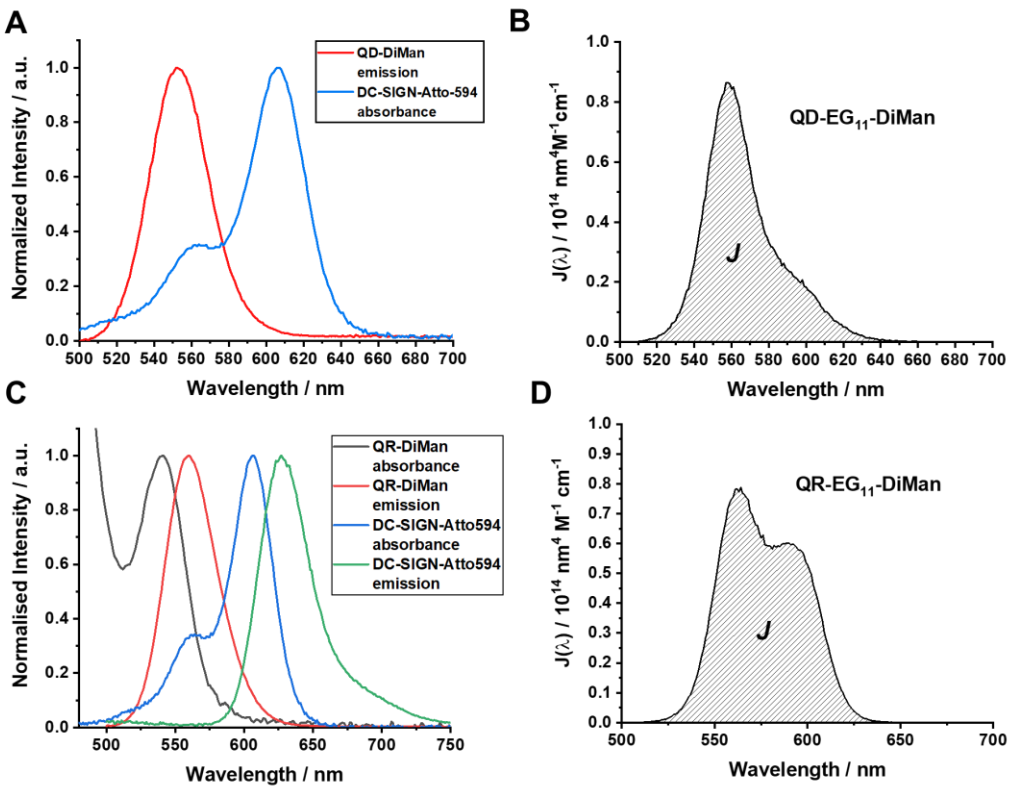

Figure B.1 Normalised intensity spectra of glycan-NP fluorescence absorbance (black) and emission (red) and DC-SIGN-Atto594 absorbance (blue) and emission (green) for (A) QD-EG$_{11}$-DiMan and (C) QR-EG$_{11}$-DiMan; and the spectral overlap as a function of wavelength, $J(\lambda)$, between the glycan-NP emission and DC-SIGN-Atto594 absorbance (Equation B.6) for (B) QD-EG$_{11}$-DiMan and (D) QR-EG$_{11}$-DiMan.

The spectral overlap, $J(\lambda)$, was calculated by **Equation B.6**, where $I_D(\lambda)$ is the normalized donor fluorescence intensity as a function of wavelength and $\varepsilon_A(\lambda)$ is the extinction coefficient of the acceptor as a function of wavelength (see **Figure B.1**.[8, 12, 13] The J value is defined as the integral of the spectral overlap.

$$J = \int J(\lambda)\,\mathrm{d}\lambda = \int \frac{I_{\mathrm{D}}(\lambda)\varepsilon_{\mathrm{A}}(\lambda)\lambda^4}{\int I_{\mathrm{D}}(\lambda)\,\mathrm{d}\lambda}\,\mathrm{d}\lambda \qquad \text{B.6}$$

The Förster radius (R_0) was obtained using **Equation B.7**; where Φ_D is the quantum yield (QY; assumed to be 50% for QD-DiMan), κ^2 is the dipole orientation factor (κ^2 = 2/3, assuming randomly oriented dipoles) and n_r is the refractive index (n_r = 1.33 for the binding buffer).[5,9] This provided an R_0 value of 57 Å.

$$R_0 = \left(8.79 \times 10^{-5} \cdot \frac{\kappa^2 \cdot \Phi_D \cdot J}{{n_r}^4}\right)^{\frac{1}{6}} \qquad \text{B.7}$$

For QRs the QY was obtained by comparison against a Rhodamine 6G in ethanol (QY = 95%, λ_{ex} = 480 nm) reference. Here, the QY before and after conjugation with DHLA-EG_{11}-DiMan was found to be 49 % and 6.2 %, respectively, indicating an 8-fold reduction of QY after cap-exchange. This is likely caused by cap-exchange induced surface defects due to ligand etching of the outer protective shell of the QR surface.[14]

Table B.2 Summary of parameters for obtaining the Förster radius (R_0) of each NP-Atto594 FRET pair, derived from Equation B.6-B.7 from their corresponding quantum yields (QY) and spectral overlap J-values.

NP-glycan	QY / %	J / $\times 10^{15}$ nm^4 M^{-1} cm^{-1}	R_0 / nm
QD-DiMan	50	3.43	5.7
QR-DiMan	6.2	4.22	4.8

B.3 QD and Atto594 labelled lectin fluorescence-concentration relationships

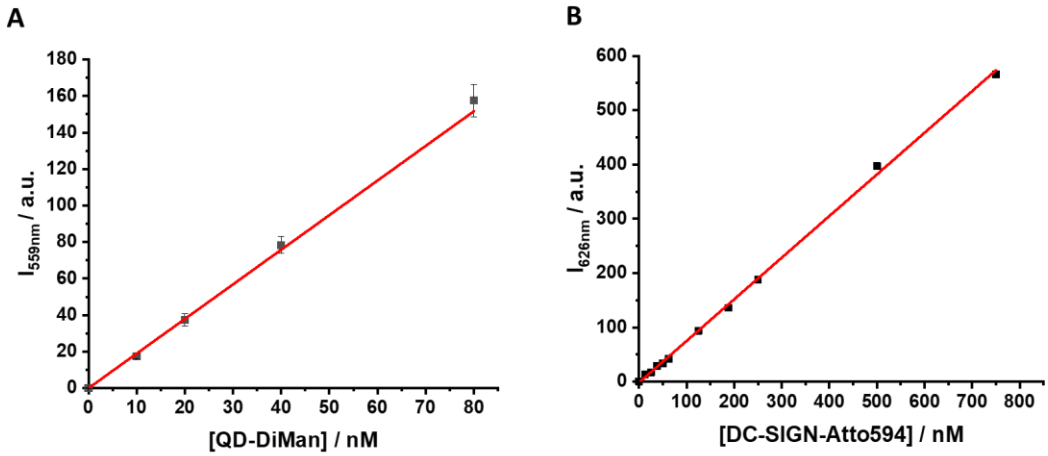

Figure B.2 Plots of the fluorescence intensity *vs.* concentration relationships for (A) QD-DiMan (λ_{ex} = 450 nm, λ_{em} = 559 nm; R^2 = 0.997) and (B) DC-SIGN (Atto594 labelled, λ_{em} = 450 nm, λ_{em} = 626 nm; R^2 = 0.999) fitted by a linear function.

Both the QD and the protein fluorescence increase linearly with concentration across the whole range studied (*e.g.* 0-80 nM for QD-DiMan and 0-800 nM for DC-SIGN-Atto594), confirming no measureable inner filter effect (the presence of significant inner filter effect would cause fluorescence signals to deviate downwardly from linear at high concentrations).

These results also confirm that neither the QD nor the protein (Atto594 labeled) fluorescence quantum yield is affected by concentration within the range studied.

B.4 QD-DiMan·Lectin Dissociation Tests

A dissociation viability test was performed on the strongest binding protein (DC-SIGN-C) in binding buffer containing 10 µg/mL of His_6-Cys as before to determine the incubation time and mannose concentration required for dissociation of a 1:1 QD-DiMan·lectin complex. Here, DC-SIGN-C and QD-DiMan was mixed at 1:1 molar ratio (both at 44 nM, 360 µL) and incubated at r.t. for 20 min, and then the fluorescence spectrum was recorded. Then 40 µL D-mannose (400 mM, 1×10^6 eq) was added and fluorescence spectra were recorded every 2 min. After 20 min, additional Mannose was added up to 4×10^6 eq. Fluorescence spectra were background corrected by their corresponding protein only controls with the same amount of D-mannose.

FRET ratios were then obtained and plotted against time or amount of D-mannose.

An incubation time of 2 min and 1×10^6 eq of mannose to the QD concentration were found to be sufficient for maximum dissociation of DC-SIGN-C from QD-DiMan.

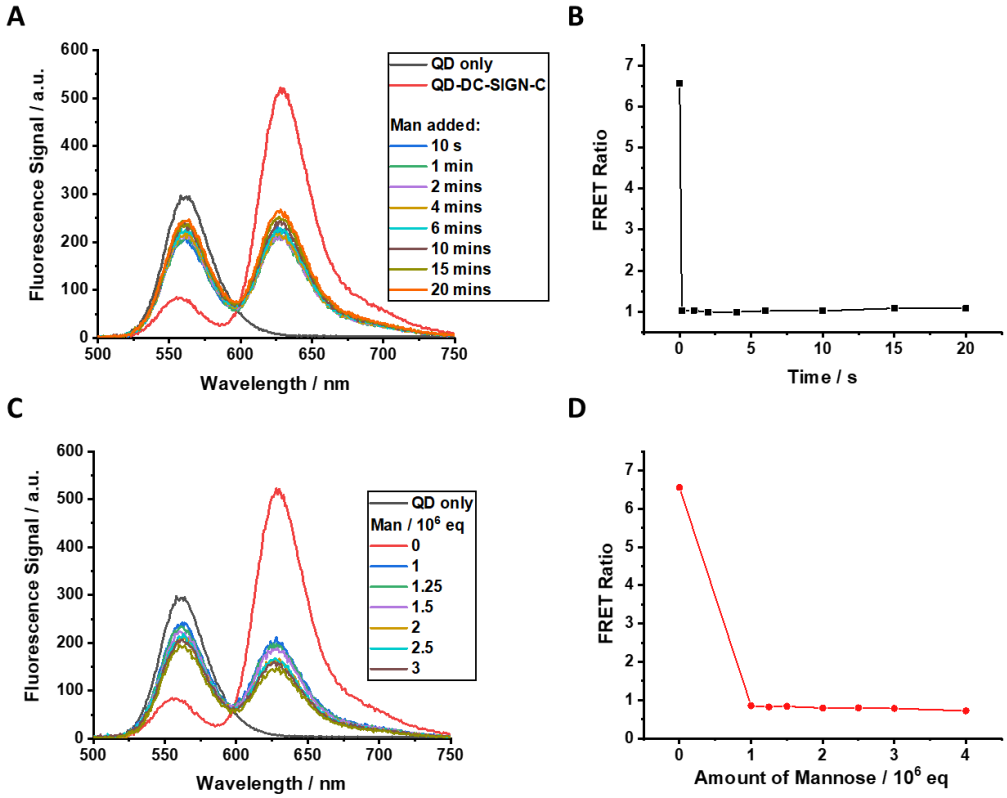

Figure B.3 (A) Background corrected fluorescence spectra and (B) FRET ratio plot of the dissociation of a 1:1 QD-DiMan·DC-SIGN-C complex with 1×10^6 eq of D-mannose (Man) over time; and (C) background corrected fluorescence spectra and (D) FRET ratio plot of the dissociation and of a 1:1 QD-DiMan·DC-SIGN-C complex with increasing concentrations of Man.

Appendix C
Lectin-Glycan Binding Assays

C.1 Thermodynamic QD-FRET studies

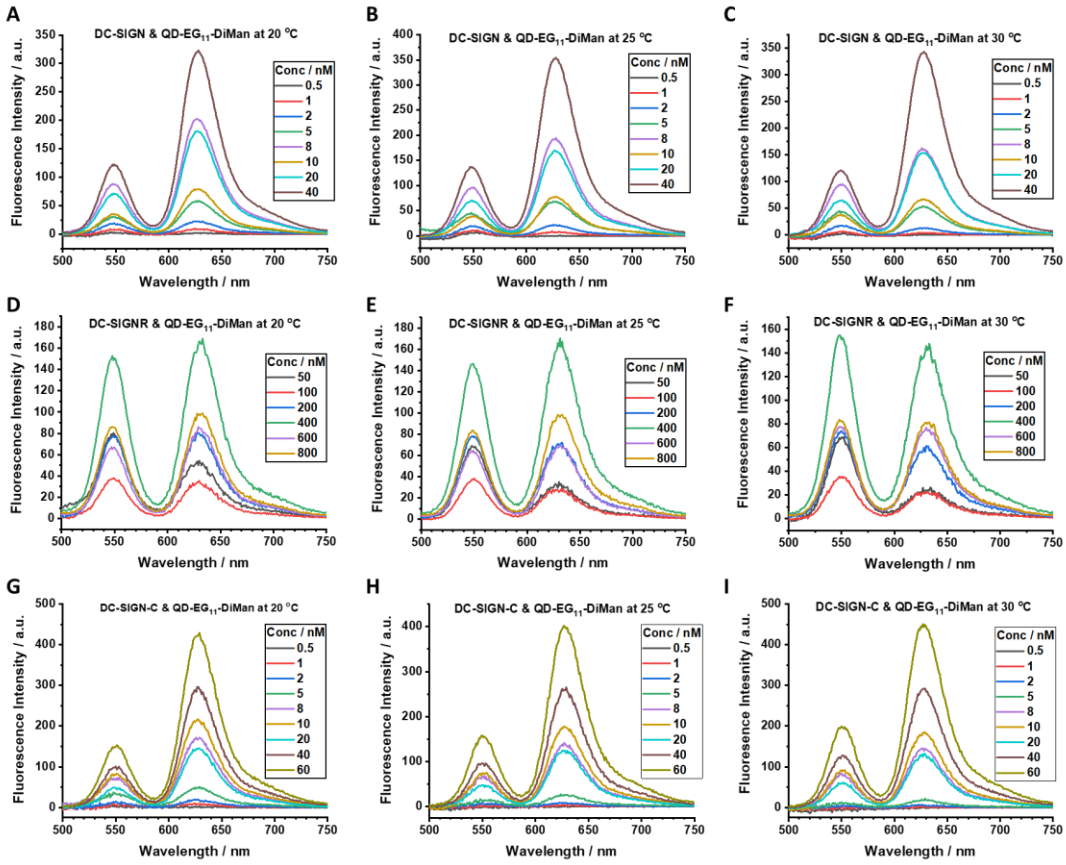

Figure C.1 Background corrected fluorescence spectra of the different concentrations of QD-DiMan after mixing with labelled DC-SIGN (PQR = 1:1) at (A) 20 °C, (B) 25 °C and (C) 30 °C; DC-SIGNR (PQR = 10:1) at (D) 20 °C, (E) 25 °C and (F) 30 °C; and DC-SIGN-C (PQR = 1:1) at (G) 20 °C, (H) 25 °C and (I) 30 °C.

C.2 Kinetic QD-FRET studies

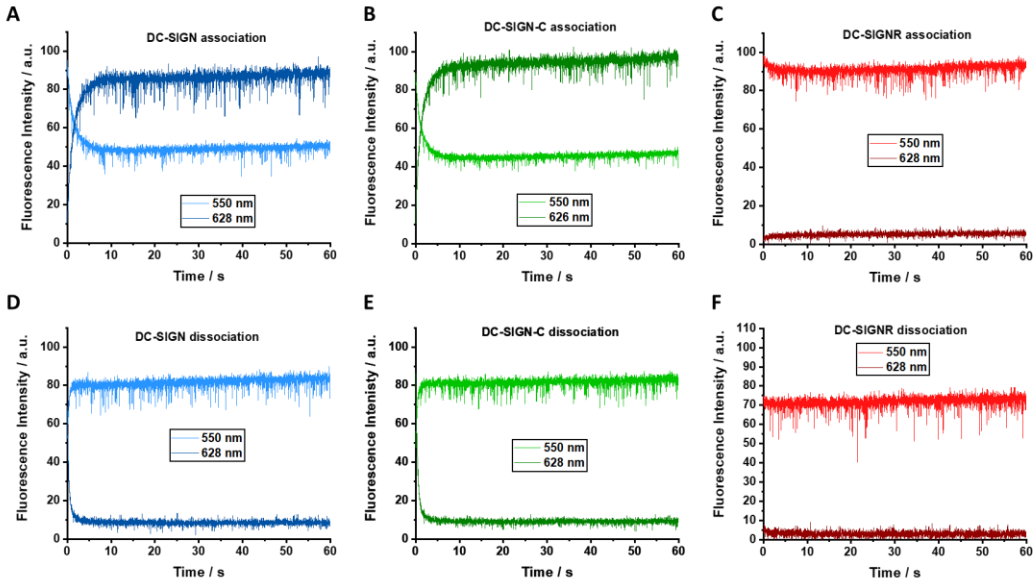

Figure C.2 Fluorescence intensity time profiles measuring the emission of 550 nm (lighter color) and 628 nm (darker color) at λ_{ex} = 450 nm for the association of QD-DiMan with (A) DC-SIGN, (B) DC-SIGN-C and (C) DC-SIGNR, and the dissociation of QD-DiMan in the presence of an excess of free mannose with (D) DC-SIGN, (E) DC-SIGN-C and (F) DC-SIGNR.

C.3 QR Hydrodynamic Size Studies

C.3.1 Lectin Titration Analysis *via* DLS

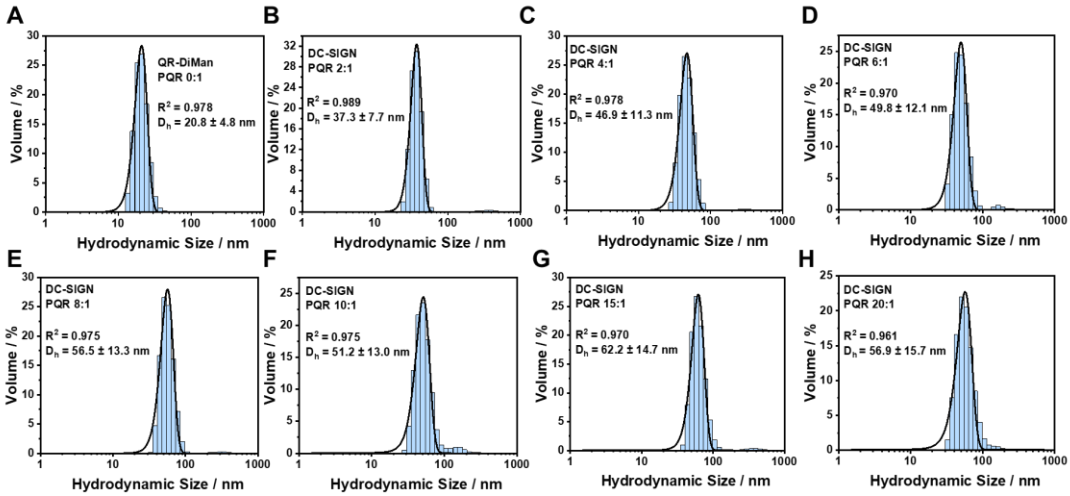

Figure C.3 Volume population hydrodynamic size distribution histograms fitted with Gaussian fits for the binding of QR-DiMan (10 nM) with DC-SIGN at varying protein:QR molar ratios (PQR) of (A) 0:1, (B) 2:1, (C) 4:1, (D) 6:1, (E) 8:1, (F) 10:1, (G) 15:1 and (H) 20:1.

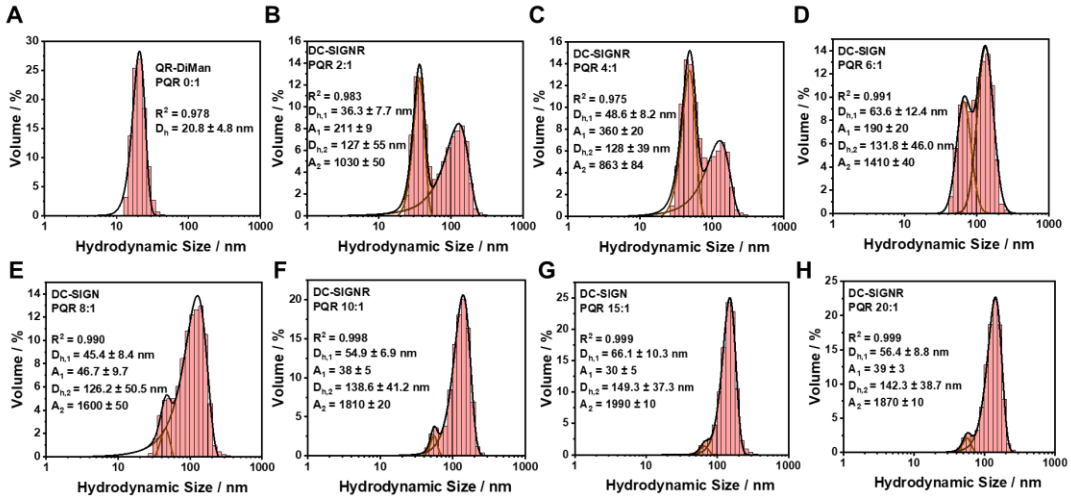

Figure C.4 Volume population hydrodynamic size distribution histograms fitted with uni- or bi- modal Gaussian fits for the binding of QR-DiMan (10 nM) with DC-SIGNR at protein:QR molar ratios (PQR) of (A) 0:1, (B) 2:1, (C) 4:1, (D) 6:1, (E) 8:1, (F) 10:1, (G) 15:1 and (H) 20:1.

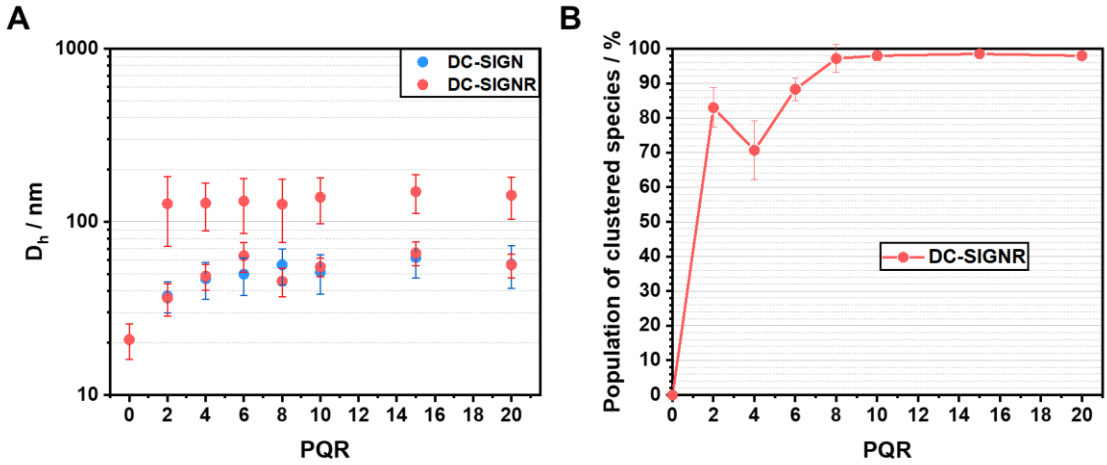

Figure C.5 Plots the change in the hydrodynamic size distribution, obtained by DLS, for the change in PQR for the binding of QR-DiMan with DC-SIGN (blue) or DC-SIGNR (red), measured by (A) the mean D_h ± ½FWHM, which demonstrates the similarity in the change in size of the smaller QR-DiMan·DC-SIGN/R complexes with increasing amount of protein, and (B) the percentage population of clustered species ($100\% \times A_2/(A_1 + A_2)$, where A_i is the integrated area of the i^{th} mode of the multimodal Gaussian distribution fit), which demonstrates the increase in clustered species with protein concentration.

Table C.1 Summary of parameters for the multimodal Gaussian fits of the hydrodynamic size distribution histograms for the binding of QR-DiMan with DC-SIGN/R. Where PQR is the protein to QR molar ratio; D_h is the mean hydrodynamic size; FWHM is the full width at half maximum; and A is the integrated area. (All errors represent the standard deviation as observed by the fitting).

Lectin	PQR	D_h / nm	FWHM / nm	A	R^2
DC-SIGN	0	20.9 ± 0.2	9.7 ± 0.4	290 ± 10	0.978
	2	37.3 ± 0.2	15.3 ± 0.4	530 ± 10	0.988
	4	46.9 ± 0.4	22.7 ± 0.9	660 ± 20	0.975
	6	49.8 ± 0.5	24 ± 1	680 ± 30	0.970
	8	56.5 ± 0.4	27 ± 1	790 ± 30	0.975
	10	51.2 ± 0.4	26 ± 1	680 ± 30	0.975
	15	62.2 ± 0.5	29 ± 1	840 ± 30	0.970
	20	56.9 ± 0.6	31 ± 1	750 ± 30	0.961
DC-SIGNR	2	36.3 ± 0.3 127 ± 2	15.5 ± 0.6 111 ± 6	211 ± 9 1030 ± 50	0.981
	4	48.6 ± 0.5 128 ± 3	16 ± 2 79 ± 2	360 ± 20 860 ± 80	0.991
	6	63.6 ± 0.7 132 ± 1	25 ± 2 92 ± 3	190 ± 20 1410 ± 40	0.991
	8	45 ± 1 126 ± 1	17 ± 3 101 ± 3	47 ± 10 1600 ± 50	0.987
	10	54.9 ± 0.8 138.7 ± 0.4	14 ± 3 75 ± 1	38 ± 5 1810 ± 20	0.998
	15	66 ± 1 149.3 ± 0.2	21 ± 3 75 ± 1	30 ± 5 1990 ± 10	0.999
	20	56.4 ± 0.7 142.3 ± 0.2	18 ± 2 77 ± 1	39 ± 3 1871 ± 10	0.999

C.3.2 QR-DiMan·Lectin Dissociation with Mannose *via* DLS

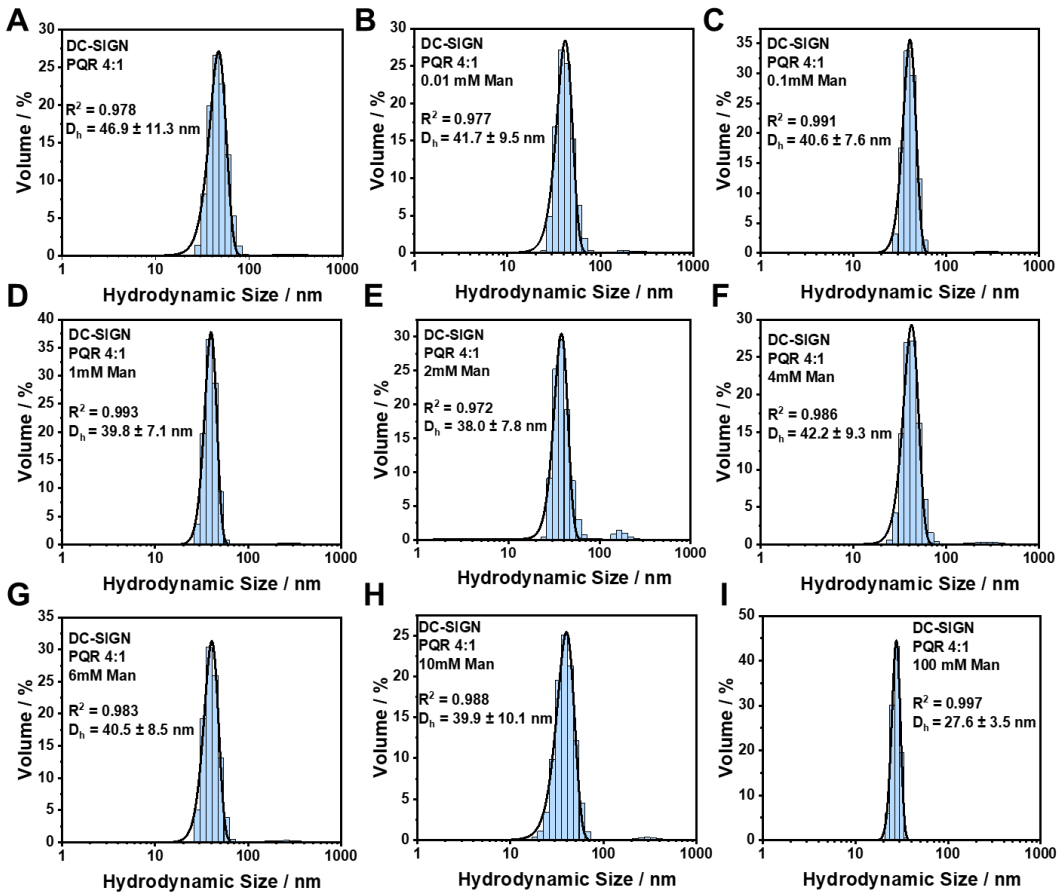

Figure C.6 Volume population hydrodynamic size distribution histograms, fitted with Gaussian fits for pre-incubated QR-DiMan (10 nM) with DC-SIGN (PQR 4:1) after addition of (A) 0.01 mM, (B) 0.1 mM (C) 1 mM, (D) 2 mM, (E) 4 mM, (F) 6 mM, (G) 10 mM and (H) 100 mM of free mannose (Man).

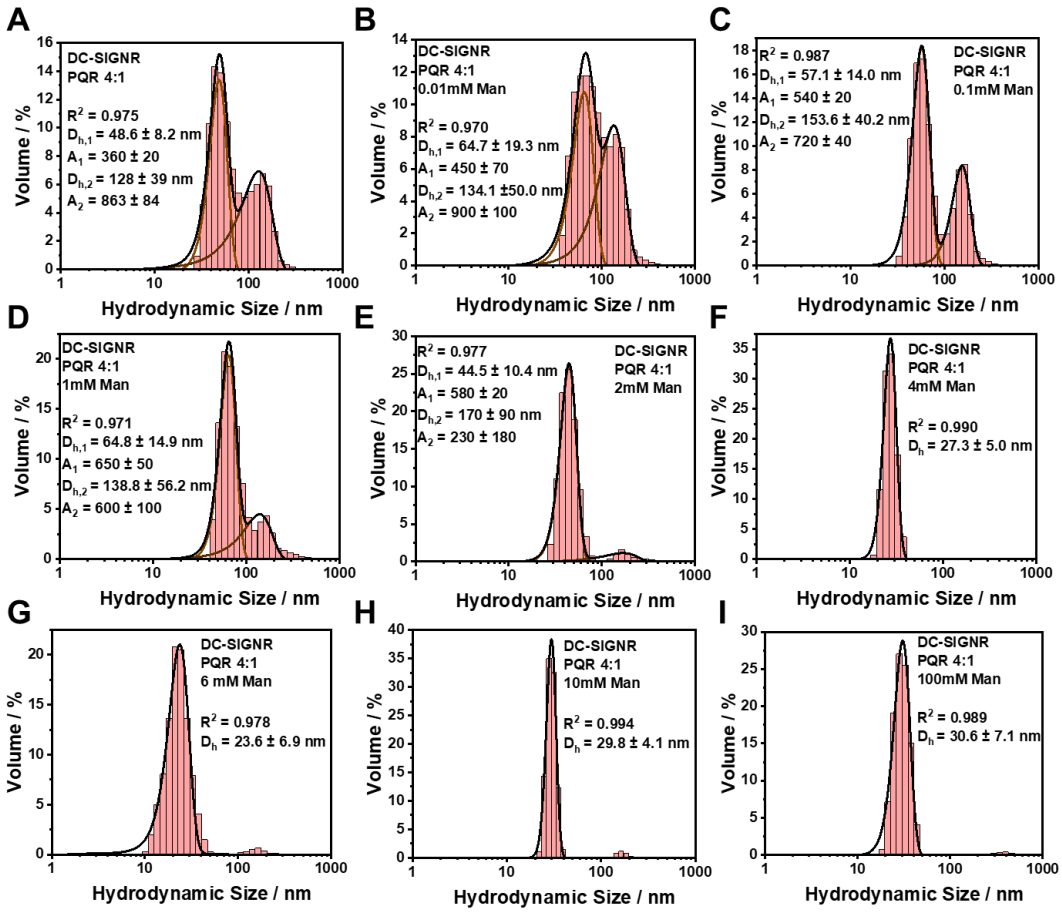

Figure C.7 Volume population hydrodynamic size distribution histograms, fitted with uni- or bi- modal Gaussian fits, for pre incubated QR-DiMan (10 nM) with DC-SIGNR (PQR 4:1) after addition of (A) 0.01 mM, (B) 0.1 mM (C) 1 mM, (D) 2 mM, (E) 4 mM, (F) 6 mM, (G) 10 mM and (H) 100 mM of free mannose (Man).

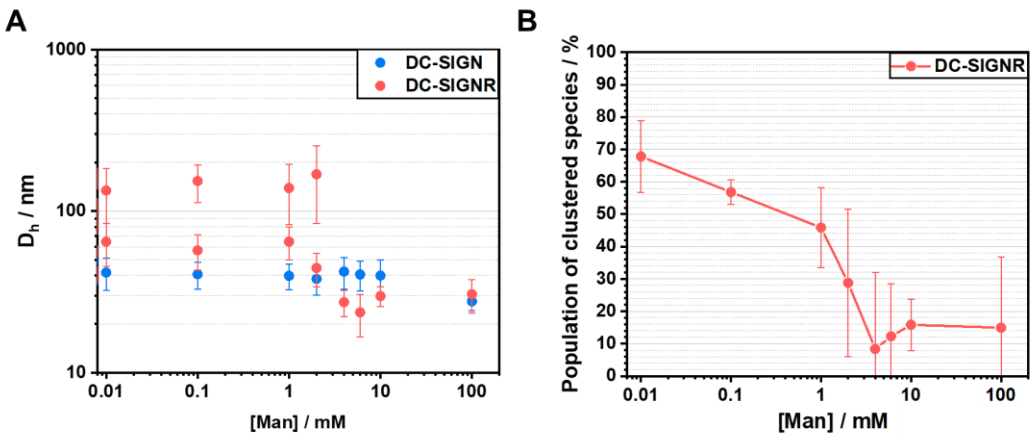

Figure C.8 Plots of the change in the hydrodynamic size distribution, obtained by DLS, for the dissociation of a PQR 4:1 QR-DiMan·DC-SIGN (blue) or QR-DiMan·DC-SIGNR (red) complex with varying mannose concentration ([Man]), measured by (A) the D_h ± ½FWHM and (B) the percentage population of clustered species ($100\% \times A_2/(A_1 + A_2)$, where A_i is the integrated area of the i^{th} mode of the multimodal Gaussian distribution fit). This demonstrates that the decrease in the population of clustered species for DC-SIGNR begins at lower [Man] than the decrease in the size of the smaller QR-DiMan·DC-SIGNR clusters, indicating that the lectin-glycan-QR interactions in the larger complexes are weaker in affinity than those of the smaller complexes.

Table C.2 Summary of parameters for the multimodal Gaussian fits of the D_h distribution histograms for the dissociation of QR-DiMan·DC-SIGN/R with free D-mannose. Where [Man] is the concentration of D-mannose; D_h is the mean hydrodynamic size; FWHM is the full width at half maximum; and A is the integrated area (All errors represent the standard deviation as observed by the fitting).

Lectin	[Man] / mM	D_h / nm	FWHM / nm	A	R^2
DC-SIGN	0	43.4 ± 0.2	16.7 ± 0.5	610 ± 20	0.985
	0.01	41.7 ± 0.3	19 ± 0.7	570 ± 20	0.976
	0.1	40.6 ± 0.1	15.2 ± 0.3	580 ± 10	0.99
	1	39.8 ± 0.1	14.1 ± 0.3	568 ± 10	0.993
	2	38 ± 0.3	15.7 ± 0.6	510 ± 20	0.971
	4	42.2 ± 0.2	18.6 ± 0.5	580 ± 10	0.985
	6	40.5 ± 0.2	17 ± 0.5	570 ± 20	0.982
	10	39.9 ± 0.2	20.3 ± 0.5	550 ± 10	0.988
	100	27.6 ± 0.1	6.9 ± 0.1	328 ± 4	0.997
DC-SIGNR	0	48.6 ± 0.5 128 ± 3	25 ± 1 110 ± 10	360 ± 20 860 ± 80	0.971
	0.01	65 ± 1 134 ± 5	39 ± 3 100 ± 10	450 ± 70 900 ± 100	0.965
	0.1	57.1 ± 0.3 154 ± 2	28 ± 0.8 80 ± 5	540 ± 10 720 ± 40	0.984
	1	64.8 ± 0.6 139 ± 9	30 ± 2 110 ± 30	650 ± 50 600 ± 100	0.966
	2	44.5 ± 0.3 170 ± 30	20.8 ± 0.8 200 ± 100	580 ± 20 200 ± 200	0.974
	4	27.3 ± 0.1	10 ± 0.3	390 ± 10	0.988
	6	23.6 ± 0.2	13.8 ± 0.6	310 ± 10	0.975
	10	29.8 ± 0.1	8.3 ± 0.2	339 ± 7	0.993
	100	30.6 ± 0.2	14.3 ± 0.4	440 ± 10	0.986

C.4 QR STEM Images

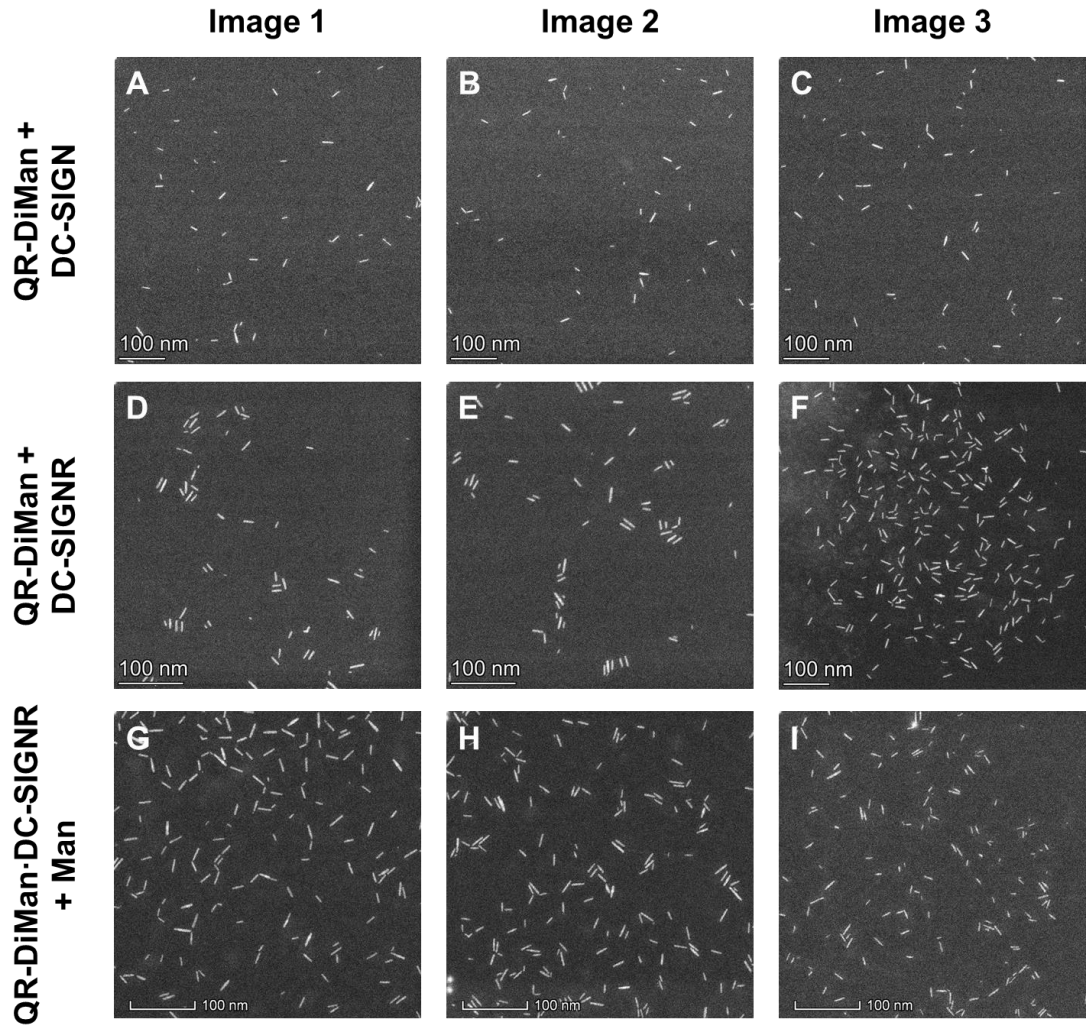

Figure C.9 Typical high angle annular dark field scanning transmission electron microscope (HAADF STEM) images for PQR 4:1 samples of (A-C) QR-DiMan with DC-SIGN; (D-F) QR-DiMan with DC-SIGNR; and (G-I) a pre-incubated sample of QR-DiMan and DC-SIGNR with 2mM of D-mannose (Man). Each incubated for 20 min prior to plunge-freezing and subsequent imaging.

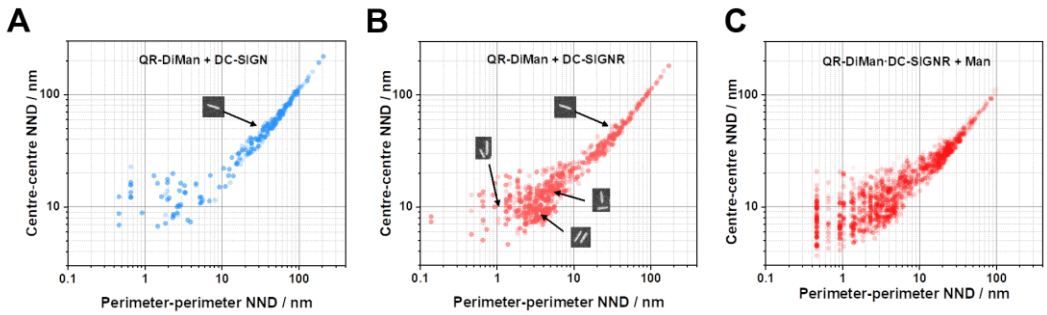

Figure C.10 Comparison of inter-QR centre to centre NNDs (ccNND) to inter-QR perimeter to perimeter ppNNDs (NND) for each QR obtained from the STEM images of QR-DiMan incubated with (A) DC-SIGN, (B) DC-SIGNR or (C) a pre-incubated QR-DiMan·DC-SIGNR with D-mannose. Distributions demonstrate the regions associated with crosslinking (ppNND ~2-10 nm, in parallel and non-parallel inter-QR orientations, ccNND <10 nm and ccNND >10 nm, respectively), simultaneous binding (ppNND >10 nm) or non-specific QR adsorption (ppNND <2 nm).

C.5 G5-Man bound DC-SIGNR-SLB

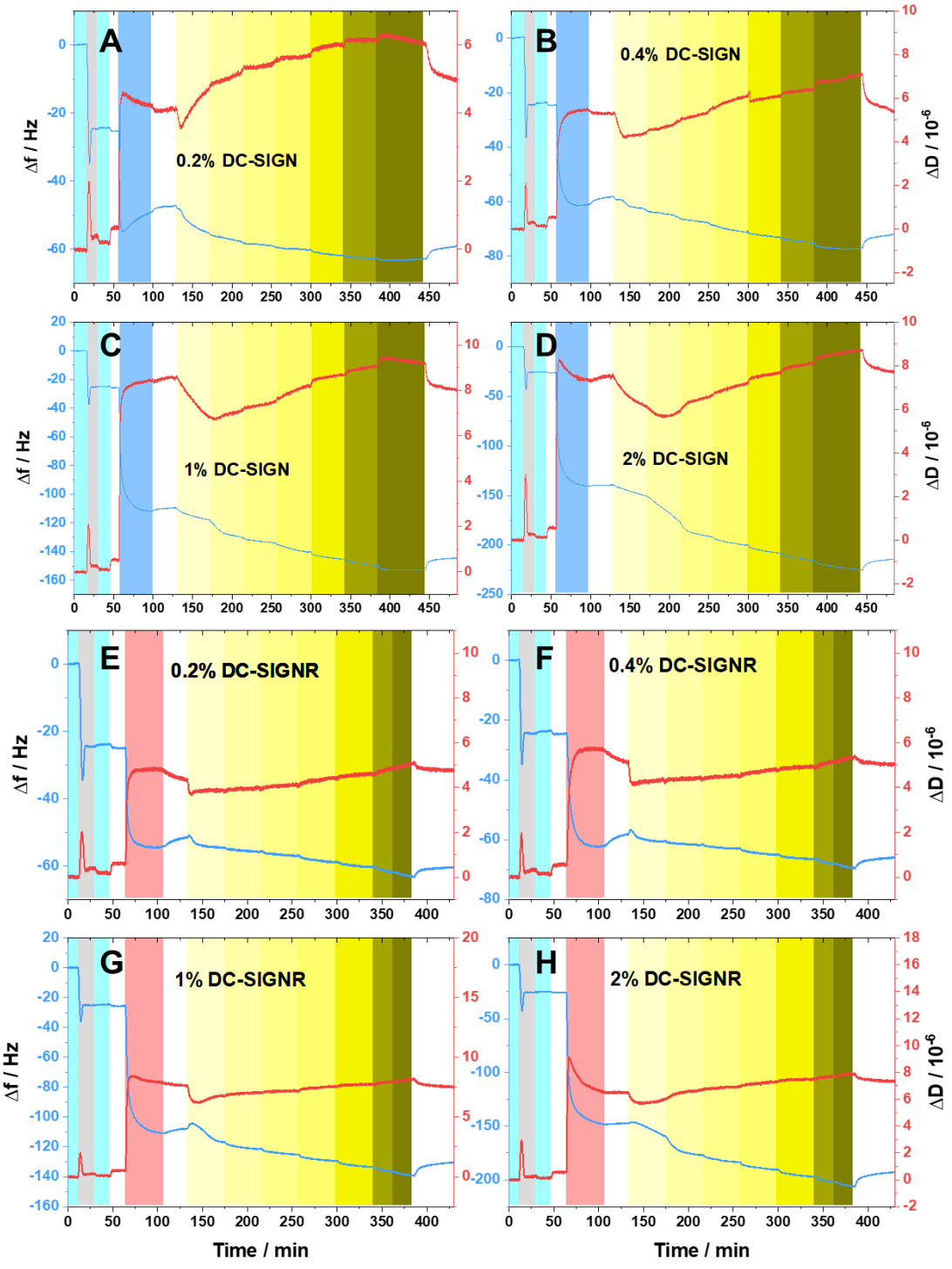

Figure C.11 Complete pots of the Δf and ΔD against time for the formation of SLBs of different tris-NTA DODA proportions; lectin association; titration of G5-Man; and subsequent dissociation of G5 in buffer, with SLBs bearing (A) 0.2% DC-SIGN, (B) 0.4% DC-SIGN, (C) 1% DC-SIGN and (D) 2% DC-SIGN, (E) 0.2% DC-SIGNR, (F) 0.4% DC-SIGNR, (G) 1% DC-SIGNR, (H) 2% DC-SIGNR.

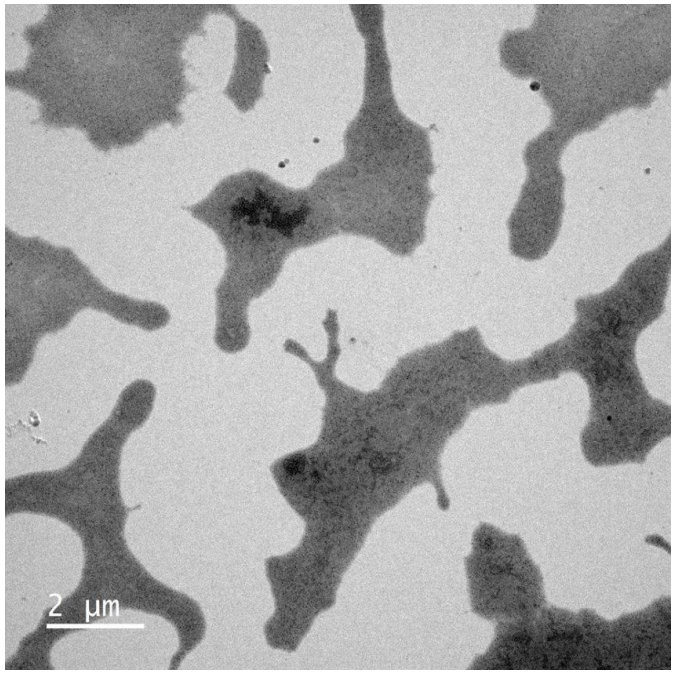

Figure C.12 Cryogenic transition electron microscopy (cryoTEM) images of G5-Man on SLBs bearing 2% DC-SIGNR demonstrating potential tension lines of "islands" of G5-Man-DC-SIGNR-SLB.

C.6 Inhibition of pseudo-typed VSV entry into Vero76 cells by lectin-GNPs

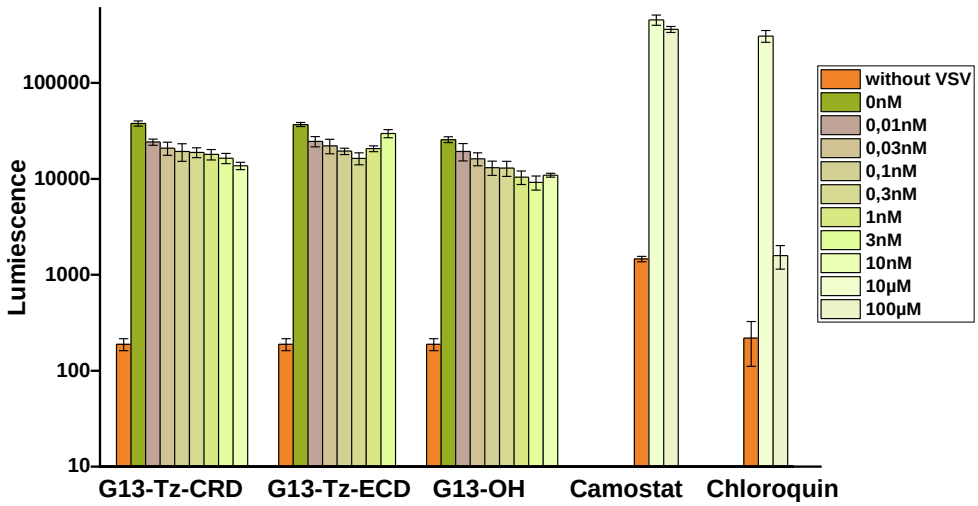

Figure C.13 Inhibition of pseudo-typed viral particles bearing the vesicular stomatitis virus glycoprotein (VSV-G, as negative control) by various G13-conjugates together with Camostat and chloroquin. No significant or very weak inhibition was observed for G13-Tz-CRD, G13-Tz-ECD and G13-OH, as expected. Camostat showed no apparent inhibition even at 100 µM, while chloroquine showed significant inhibition at 100 µM, as expected.

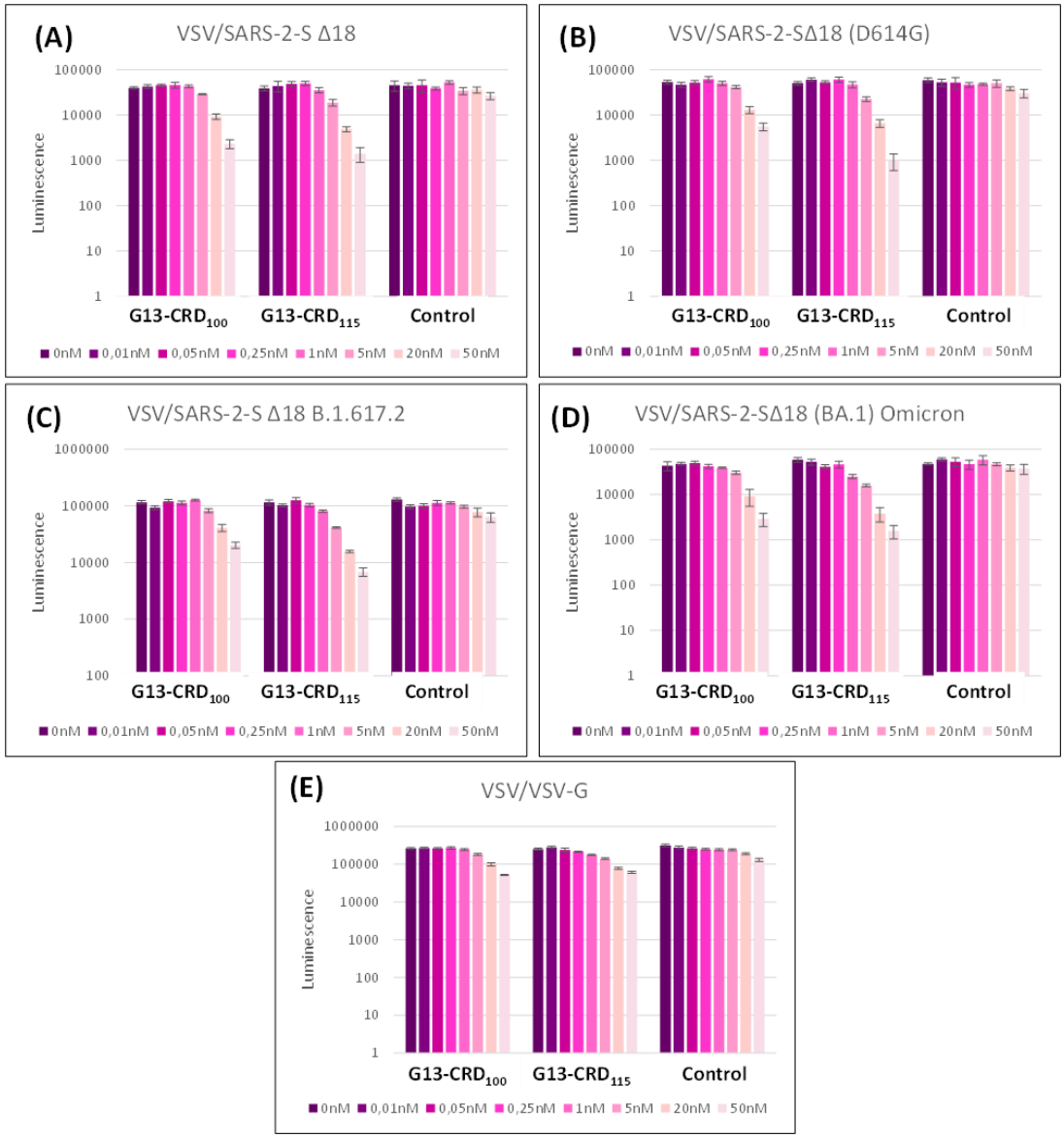

Figure C.14 Summary of dose-dependent inhibition of Vero76 cells using G13-CRD$_{100}$, G13-CRD$_{115}$ and G13-OH control against: various pseudo-typed virus bearing the S protein from SARS CoV-2 variants (A) WT (Wuhan wild-type), (B) B.1, (C) B.1.617.2 (delta) and (D) BA.1 (Omicron); and (E) a control pseudo-typed virus bearing vesicular stomatitis virus glycoprotein (VSV-G). G13-CRD$_{115}$ gives the highest inhibition while G13-OH control shows no significant inhibition.

Appendix D
Lectin Inter-CRD Measurements

D.1 SmFRET FRET Efficiency Distributions

D.1.1 DC-SIGN-Atto488 Control

In order to determine the acceptor intensity threshold which results in <1% of events from protein bearing only donor fluorophore, the fluorescence of a donor only control protein (DC-SIGN-Atto488) was monitored. To avoid the contribution of noise, a threshold of the sum of donor and acceptor intensities of 15 a.u. was applied and the resulting number of donor events was normalised to 100%. The decrease in donor only events over increasing acceptor intensity threshold was then recorded (**Figure D.1**) to find the acceptor threshold required to filter out 99% counted donor only fluorescence events (**Figure D.2**). All histograms were fitted with unimodal Gaussian distribution curves in order to determine the mean FRET efficiency for these observed “false” FRET events at different acceptor thresholds (**Table D.1**).

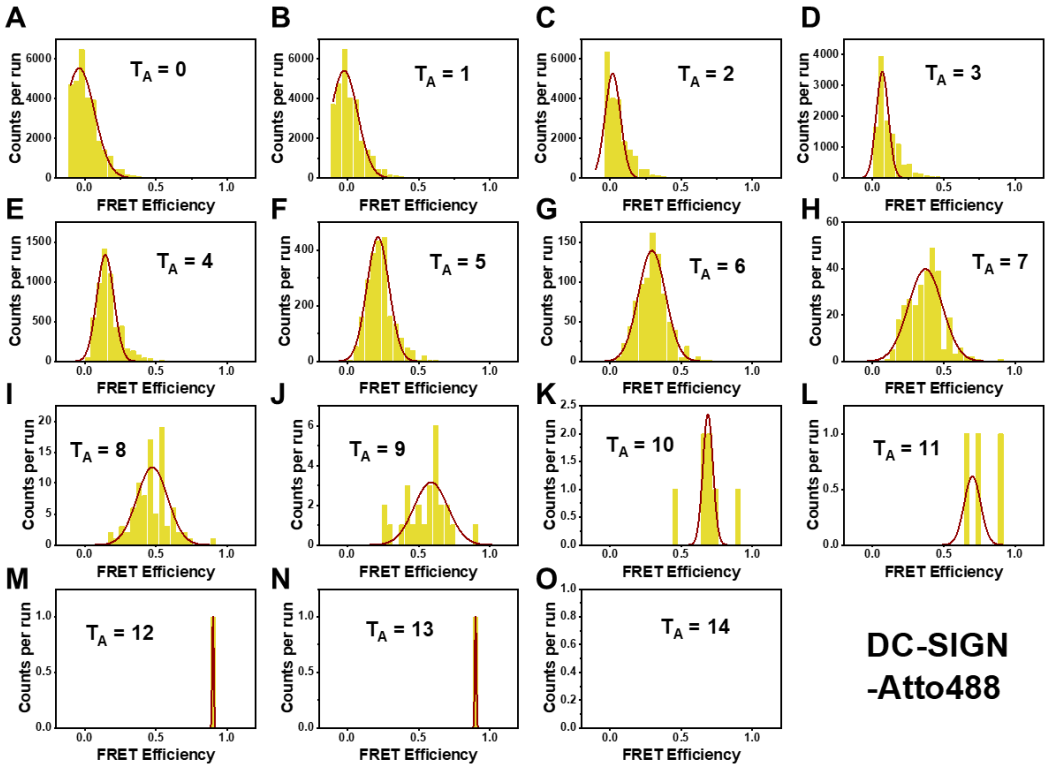

Figure D.1 Histograms for the number of "false" FRET events for each FRET efficiency observed by smFRET for DC-SIGN bearing only donor only dye (DC-SIGN-Atto488) over increasing acceptor threshold (T_A) for a fixed threshold for the sum of acceptor and donor intensities of 15 a.u.. All histograms were fitted with unimodal Gaussian distribution (parameters summarised in Table D.1).

Table D.1 Summary of parameters for the unimodal Gaussian fits of the FRET efficiency distributions recorded for DC-SIGN-Atto488 donor only control sample with a fixed threshold for the sum of acceptor and donor intensities of 15 a.u.. FWHM is the full width at half maximum and A is the integrated area under the fit. *These values do not represent "real" FRET events but demonstrate the how much of the control is detected using these thresholds.

Acceptor Threshold	Mean FRET efficiency*	FWHM	Area, A	R^2
0	-0.039 ± 0.007	0.25 ± 0.02	1480 ± 90	0.97308
1	-0.023 ± 0.005	0.22 ± 0.01	1280 ± 70	0.96784
2	0.019 ± 0.006	0.13 ± 0.02	710 ± 70	0.80543
3	0.068 ± 0.004	0.101 ± 0.01	370 ± 30	0.86492
4	0.144 ± 0.003	0.141 ± 0.007	202 ± 8	0.96172
5	0.215 ± 0.003	0.184 ± 0.008	88 ± 3	0.96746
6	0.295 ± 0.004	0.226 ± 0.01	34 ± 1	0.9667
7	0.37 ± 0.01	0.28 ± 0.02	11.9 ± 0.9	0.87253
8	0.47 ± 0.01	0.26 ± 0.03	3.4 ± 0.4	0.72465
9	0.59 ± 0.02	0.28 ± 0.05	0.9 ± 0.1	0.55969
10	0.688 ± 0.005	0.08 ± 0.01	0.21 ± 0.03	0.75324
11	0.70 ± 0.02	0.14 ± 0.05	0.09 ± 0.03	0.24837
12	0.9	0.01	0.01	1
13	0.9	0.01	0.01	1
14	-	-	-	-

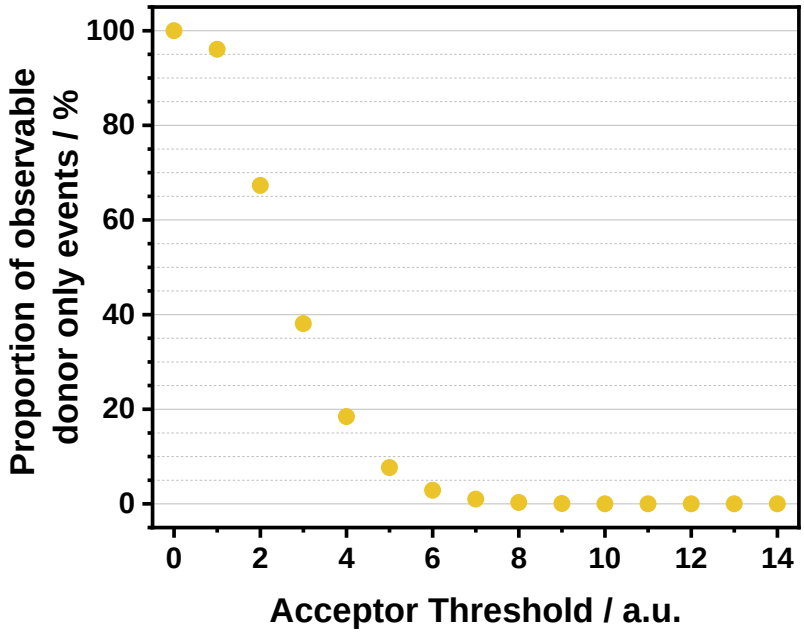

Figure D.2 Plot of the proportion of fluorescence detected as a FRET event by the donor only control sample (DC-SIGN-Atto488) obtained by percentage of FRET events counted with increasing acceptor threshold (T_A) compared to that at T_A = 0. This demonstrates that the percentage of observable donor only events drops below 1% at a T_A of 7 a.u..

D.1.2 HTDC-SIGN/R smFRET signal

The smFRET signals of HTDC-SIGN and HTDC-SIGNR were then counted over varying acceptor intensity threshold with a fixed acceptor and donor intensity sum threshold of 15 a.u., and the resulting histograms were fitted with multimodal Gaussian fits where the number of terms was determined by manually finding the highest achievable R^2 value (**Figure D.3; Figure D.4**). All parameters are recorded in **Table D.2**.

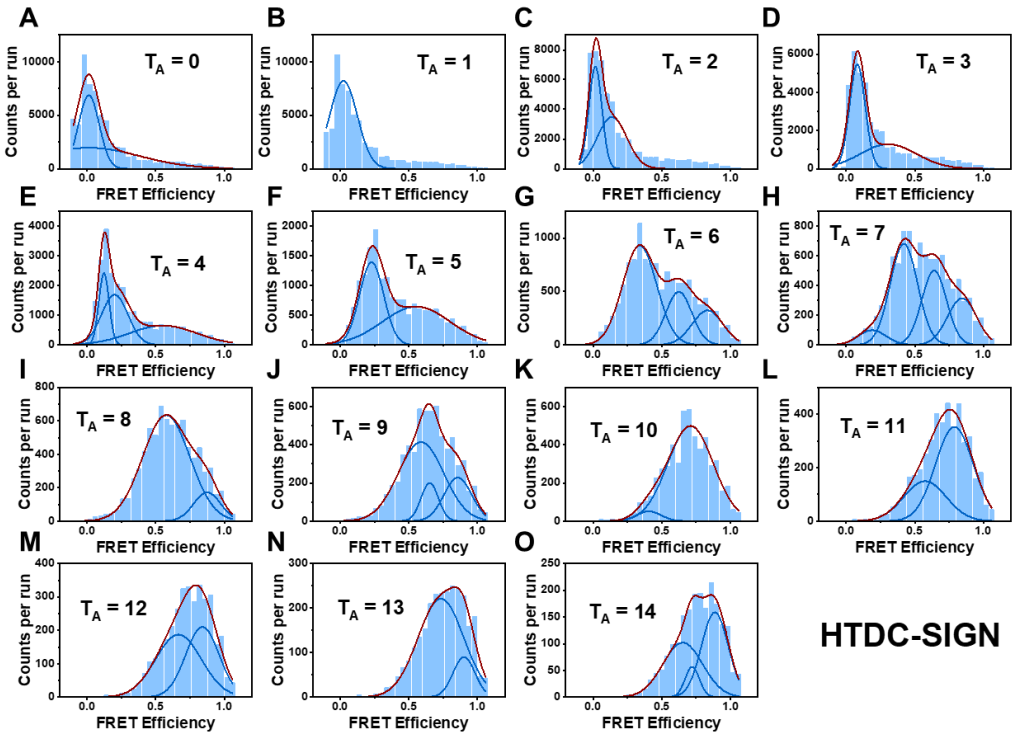

Figure D.3 Histograms for the number of FRET events for each FRET efficiency observed by smFRET for HTDC-SIGN over increasing acceptor threshold (T_A) for a fixed threshold for the sum of acceptor and donor intensities of 15 a.u.. All histograms were fitted with multimodel Gaussian distribution (parameters summarised in Table D.2).

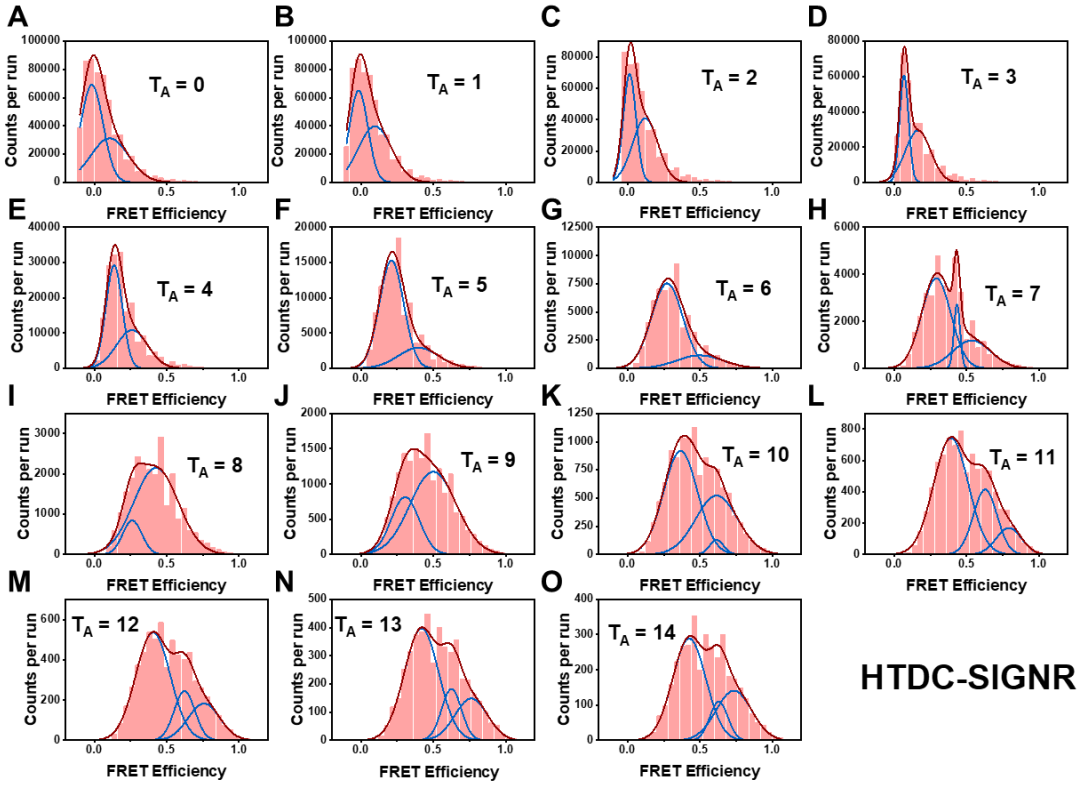

Figure D.4 Histograms for the number of FRET events for each FRET efficiency observed by smFRET for HTDC-SIGNR over increasing acceptor threshold (T_A) for a fixed threshold for the sum of acceptor and donor intensities of 15 a.u.. All histograms were fitted with multimodel Gaussian distribution (parameters summarised in Table D.2).

Table D.2 Summary of parameters for the multimodal Gaussian fits of the FRET efficiency distributions recorded for HTDC-SIGN and HTDC-SIGNR with a fixed threshold for the sum of acceptor and donor intensities of 15 a.u.. FWHM is the full width at half maximum and A is the integrated area under the fit.

HTDC-SIGN				
Acceptor Threshold	**Mean FRET Efficiency**	**FWHM**	**Area**	**R^2**
0	0.01 ± 0.02 0.01 ± 0.6	0.18 ± 0.05 0.9 ± 0.6	1300 ± 700 2000 ± 3000	0.93292
1	0.025 ± 0.009	0.24 ± 0.02	2100 ± 200	0.87421
2	0.016 ± 0.006 0.14 ± 0.07	0.11 ± 0.03 0.25 ± 0.09	800 ± 500 900 ± 500	0.93109

3	0.081 ± 0.005 0.31 ± 0.09	0.13 ± 0.02 0.5 ± 0.2	800 ± 200 700 ± 200	0.92043
4	0.123 ± 0.003 0.2 ± 0.01 0.55 ± 0.07	0.09 ± 0.01 0.23 ± 0.03 0.6 ± 0.1	230 ± 50 400 ± 100 400 ± 100	0.98382
5	0.23 ± 0.006 0.56 ± 0.05	0.21 ± 0.02 0.58 ± 0.1	310 ± 60 400 ± 70	0.95319
6	0.34 ± 0.02 0.62 ± 0.06 0.8 ± 0.2	0.29 ± 0.03 0.2 ± 0.1 0.3 ± 0.2	290 ± 40 100 ± 100 100 ± 100	0.9584
7	0.2 ± 0.4 0.42 ± 0.05 0.64 ± 0.06 0.8 ± 0.2	0.2 ± 0.4 0.2 ± 0.2 0.2 ± 0.2 0.2 ± 0.2	30 ± 100 200 ± 200 100 ± 200 100 ± 100	0.96383
8	0.58 ± 0.02 0.88 ± 0.03	0.41 ± 0.03 0.23 ± 0.07	270 ± 30 40 ± 20	0.97546
9	0.6 ± 0.2 0.65 ± 0.02 0.9 ± 0.1	0.4 ± 0.2 0.1 ± 0.1 0.2 ± 0.1	200 ± 200 30 ± 80 100 ± 200	0.97301
10	0.4 ± 0.1 0.71 ± 0.02	0.2 ± 0.2 0.39 ± 0.05	10 ± 20 210 ± 20	0.9411
11	0.6 ± 0.5 0.8 ± 0.1	0.4 ± 0.4 0.32 ± 0.1	100 ± 200 100 ± 200	0.96965
12	0.7 ± 0.5 0.84 ± 0.08	0.4 ± 0.4 0.3 ± 0.2	100 ± 200 100 ± 200	0.97407
13	0.73 ± 0.05 0.9 ± 0.01	0.39 ± 0.06 0.19 ± 0.09	90 ± 30 20 ± 20	0.97898
14	0.7 ± 0.2	0.3 ± 0.2	40 ± 40	0.98206

	0.72 ± 0.02 0.89 ± 0.04	0.12 ± 0.08 0.23 ± 0.04	10 ± 10 40 ± 30	
HTDC-SIGNR				
Acceptor Threshold	**Mean FRET Efficiency**	**FWHM**	**Area**	**R^2**
0	-0.017 ± 0.008 0.1 ± 0.3	0.09 ± 0.04 0.2 ± 0.1	10000 ± 20000 10000 ± 30000	0.97898
1	-0.017 ± 0.008 0.1 ± 0.2	0.08 ± 0.03 0.14 ± 0.1	10000 ± 20000 10000 ± 20000	0.97114
2	0.014 ± 0.008 0.12 ± 0.08	0.05 ± 0.02 0.1 ± 0.05	8000 ± 7000 9000 ± 7000	0.92197
3	0.068 ± 0.002 0.16 ± 0.02	0.036 ± 0.003 0.1 ± 0.01	4700 ± 700 6500 ± 900	0.9809
4	0.136 ± 0.005 0.26 ± 0.09	0.07 ± 0.01 0.13 ± 0.05	4000 ± 2000 3000 ± 2000	0.96903
5	0.21 ± 0.02 0.4 ± 0.3	0.1 ± 0.02 0.2 ± 0.2	3000 ± 2000 1000 ± 2000	0.94923
6	0.27 ± 0.03 0.5 ± 0.6	0.13 ± 0.03 0.2 ± 0.3	2000 ± 2000 500 ± 2000	0.93764
7	0.29 ± 0.02 0.434 ± 0.008 0.5 ± 0.1	0.12 ± 0.02 0.03 ± 0.02 0.16 ± 0.09	1000 ± 300 150 ± 50 400 ± 300	0.97002
8	0.26 ± 0.02 0.43 ± 0.04	0.08 ± 0.04 0.18 ± 0.03	100 ± 200 800 ± 200	0.92917
9	0.31 ± 0.03 0.5 ± 0.1	0.11 ± 0.05 0.19 ± 0.06	200 ± 300 500 ± 300	0.9636
10	0.37 ± 0.06 0.62 ± 0.04 0.6 ± 0.2	0.14 ± 0.03 0.06 ± 0.07 0.17 ± 0.07	300 ± 200 20 ± 40 200 ± 200	0.97854

	0.39 ± 0.02	0.14 ± 0.01	220 ± 30	
11	0.63 ± 0.06	0.1 ± 0.05	100 ± 100	0.98386
	0.8 ± 0.2	0.1 ± 0.09	40 ± 80	
	0.41 ± 0.02	0.14 ± 0.02	160 ± 20	
12	0.62 ± 0.04	0.08 ± 0.05	40 ± 100	0.97647
	0.8 ± 0.2	0.1 ± 0.1	100 ± 100	
	0.42 ± 0.02	0.14 ± 0.02	120 ± 20	
13	0.63 ± 0.04	0.08 ± 0.05	30 ± 70	0.96933
	0.8 ± 0.2	0.1 ± 0.1	40 ± 70	
	0.43 ± 0.05	0.14 ± 0.03	80 ± 40	
14	0.63 ± 0.03	0.07 ± 0.08	20 ± 70	0.95747
	0.7 ± 0.3	0.1 ± 0.2	40 ± 100	

D.2 Reweighted MD Distances

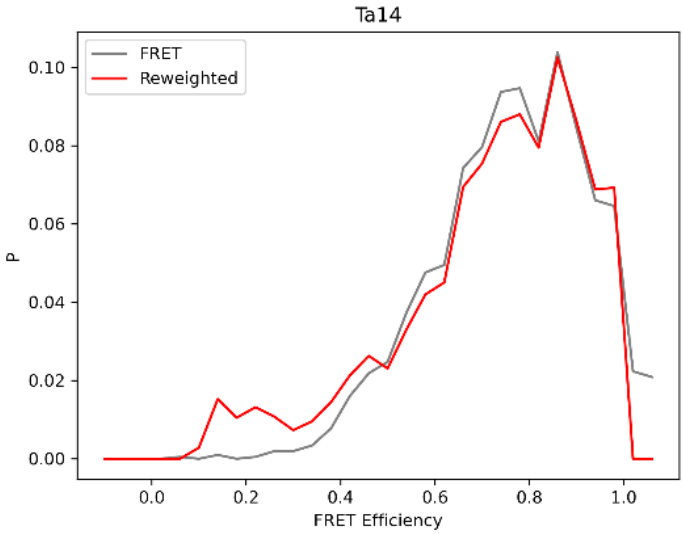

Figure D.5 Example FRET efficiency distribution derived from the reweighted sampling of the MD simulation of DC-SIGN compared against the FRET efficiency histogram obtained by smFRET at a T_{sum} of 15 a.u. and a T_{A} of 14 a.u..

Table D.3 Distances between free C274 (*i.e.* Cα…Cα) and dye molecules (*i.e.* O8…O8) derived from the MD sampling reweighted towards the smFRET histograms obtained at each T_A for each of the 4 side (s) and 2 diagonal (d) inter-CRD distances of DC-SIGN.

		Cα…Cα Distance / nm						O8…O8 Distance / nm					
T_A	RSS	s1	s2	s3	s4	d1	d2	s1	s2	s3	s4	d1	d2
0	6.94	3.2	2.78	2.84	3.03	4.45	3.88	4.19	4.27	4.7	4.57	6.01	6.28
1	6.97	3.17	2.8	2.84	3.02	4.38	3.94	4.24	4.4	4.64	4.5	6.08	6.29
2	6.99	3.2	2.79	2.83	3.04	4.46	3.87	4.12	4.23	4.54	4.55	5.89	6.22
3	5.94	3.17	2.77	2.84	3.03	4.4	3.9	4.2	4.28	4.56	4.55	5.95	6.26
4	3.99	3.16	2.81	2.84	3.05	4.43	3.92	4.06	4.36	4.46	4.63	5.92	6.21
5	2.71	3.18	2.83	2.84	3.08	4.46	3.92	3.94	4.19	4.34	4.54	5.79	6.03
6	1.88	3.13	2.88	2.83	3.09	4.44	3.95	3.64	4.15	4.28	4.2	5.56	5.71
7	1.52	3.13	2.87	2.81	3.1	4.43	3.96	3.49	4.17	4.21	3.92	5.4	5.53
8	1.56	3.09	2.91	2.83	3.12	4.42	3.98	3.42	4.19	4.04	3.87	5.4	5.35
9	1.42	3.1	2.88	2.84	3.11	4.41	3.97	3.51	4.09	3.92	3.68	5.27	5.21
10	1.23	3.12	2.87	2.83	3.13	4.43	3.96	3.53	4.03	3.87	3.37	5.28	4.94
11	1.02	3.09	2.9	2.83	3.09	4.33	4.03	3.44	3.98	3.73	3.33	5.12	4.86
12	0.91	3.06	2.89	2.82	3.06	4.28	4.05	3.44	3.89	3.63	3.15	4.92	4.8
13	0.85	3.06	2.91	2.82	3.05	4.28	4.05	3.41	3.86	3.53	3.03	4.84	4.71
14	1.02	3.06	2.91	2.83	3.06	4.3	4.05	3.41	3.86	3.51	2.99	4.77	4.72

Table D.4 Distances between free C274 (*i.e.* Cα…Cα) and dye molecules (*i.e.* O8…O8) derived from the MD sampling reweighted towards the smFRET histograms obtained at each T_A for each of the 4 side (s) and 2 diagonal (d) inter-CRD distances of DC-SIGN.

		Cα…Cα Distance / nm						O8…O8 Distance / nm					
T_A	RSS	s1	s2	s3	s4	d1	d2	s1	s2	s3	s4	d1	d2
0	6.7	3.42	5.98	4.93	5.97	6.78	6.38	3.59	6.05	5.01	5.88	6.79	6.13
1	6.7	3.41	6.16	4.91	5.96	6.68	6.43	3.48	6.23	5.02	5.82	6.71	6.16
2	6.6	3.41	6.12	4.90	6.05	6.79	6.49	3.71	6.08	5.04	6.14	6.66	6.32

3	3.2	3.42	6.15	4.89	5.97	6.81	6.43	3.57	6.18	5.22	6.09	6.66	6.36
4	1.7	3.53	5.97	5.10	5.84	6.50	6.40	3.78	5.97	5.31	5.87	6.38	6.18
5	1.4	3.72	5.80	5.14	5.78	6.24	6.25	4.07	5.71	5.25	5.67	5.96	5.76
6	1.4	3.83	5.57	5.14	5.78	6.21	6.14	4.16	5.29	5.06	5.52	5.76	5.44
7	1.2	3.88	5.37	5.25	5.65	6.11	6.07	4.19	5.07	5.01	5.33	5.72	5.26
8	1.2	3.83	5.20	5.32	5.59	6.07	5.99	4.02	4.83	5.02	5.05	5.71	5.01
9	1.4	3.83	4.96	5.31	5.55	6.19	5.94	3.92	4.76	4.93	4.96	5.73	4.94
10	1.3	3.88	4.81	5.41	5.45	6.18	5.84	3.90	4.59	4.90	4.93	5.72	4.82
11	1.3	3.88	4.85	5.35	5.44	6.16	5.82	3.82	4.58	4.83	4.83	5.68	4.80
12	1.5	3.86	4.87	5.42	5.46	6.24	5.86	3.64	4.58	4.82	4.72	5.78	4.76
13	1.4	3.82	4.95	5.34	5.51	6.08	5.86	3.68	4.61	4.81	4.82	5.56	4.62
14	1.4	3.85	4.94	5.28	5.47	6.25	5.87	3.64	4.51	4.70	4.69	5.68	4.60

Appendix E
Equation Derivations

E.1.1 Derivation of Second Order FRET Rate Equation for Association

Association kinetics was performed by investigating the interaction between 1 eq of protein (P) with 1 eq of QD (**Equation E.1**) so that a second order rate equation could be used to obtain the association rate coefficient for association, k_{on}, where the rate of change of protein concentration, $\mathrm{d[P]}/\mathrm{d}t$, is directly proportional to the concentration squared, as [QD] = [P] (**Equation E.2**). This can then be integrated to provide the relationship between the concentration and time (**Equation E.3**).

$$\mathrm{QD} + \mathrm{P} \rightarrow \mathrm{QD} \cdot \mathrm{P} \qquad \text{E.1}$$

$$\frac{\mathrm{d[P]}}{\mathrm{d}t} = -k_{\mathrm{on}}[\mathrm{QD}][\mathrm{P}] = -k_{\mathrm{on}}[\mathrm{P}]^2 \qquad \text{E.2}$$

$$[\mathrm{P}]_t = \frac{[\mathrm{P}]_0}{1 + k_{\mathrm{on}}[\mathrm{P}]_0 t} \qquad \text{E.3}$$

It can be assumed that the FRET ratio associated with the QD-glycan-protein binding interaction is linearly proportional to the fraction of bound protein (**Equation E.6**). As the amount of bound protein is simply the amount of protein that is not free at any time, the concentration of bound protein, $[\mathrm{QD} \cdot \mathrm{P}]_t$, is equal to $[\mathrm{P}]_0 - [\mathrm{P}]_t$. uation can then be combined with **Equation E.3** to obtain the relationship between the fraction of bound protein and time (**Equation E.5**).

$$\theta_t = \frac{n_{\mathrm{bound\ P}}}{n_{\mathrm{total\ P}}} = \frac{[\mathrm{QD} \cdot \mathrm{P}]_t}{[\mathrm{P}]_t + [\mathrm{QD} \cdot \mathrm{P}]_t} = \frac{[\mathrm{P}]_0 - [\mathrm{P}]_t}{[\mathrm{P}]_0} \qquad \text{E.4}$$

$$\theta_t = \frac{k_{\mathrm{on}}[\mathrm{P}]_0 t}{1 + k_{\mathrm{on}}[\mathrm{P}]_0 t} \qquad \text{E.5}$$

The FRET ratio can then be obtained as **Equation E.6** using two assumptions: firstly, that the maximum FRET ratio occurs when all protein has bound, i.e. at θ_t = 1, and secondly, that the FRET ratio decays linearly with time due to the decay of fluorescence with prolonged light exposure, where a denotes the linear decay factor.[12] This can then be inserted into **Equation E.5**

to provide the relationship between the FRET ratio and time (**Equation E.7**) from which the association FRET time profiles are fitted, unless otherwise stated.

$$F_t = F_{\max}\theta_t - at \tag{E.6}$$

$$F_t = F_{max}\frac{k_{\mathrm{on}}[\mathrm{P}]_0 t}{1 + k_{\mathrm{on}}[\mathrm{P}]_0 t} - at \tag{E.7}$$

E.1.2 Pseudo-First Order FRET Rate Equation for Dissociation

Dissociation kinetics were performed using an excess of free D-mannose (Man) to readily compete with the 1:1 QD-glycan-protein binding interaction (**Equation E.8**). Due to this, a pseudo-first rate order can be derived to obtain the pseudo-dissociation rate coefficient, k'_{off}, where the rate of change of the concentration of bound protein, $\mathrm{d}[\mathrm{QD}\cdot\mathrm{P}]/\mathrm{d}t$, is linearly proportional to this concentration due to the negligible change in the concentration of mannose (**Equation E.9**). This can be integrated to obtain the relationship between complex concentration and time (**Equation E.10**).

$$\mathrm{QD}\cdot\mathrm{P} + \mathrm{Man} \rightarrow \mathrm{QD} + \mathrm{P} + \mathrm{Man} \tag{E.8}$$

$$\frac{\mathrm{d}[\mathrm{QD}\cdot\mathrm{P}]}{\mathrm{d}t} = -k[\mathrm{QD}\cdot\mathrm{P}][\mathrm{Man}] = -k'_{\mathrm{off}}[\mathrm{QD}\cdot\mathrm{P}] \tag{E.9}$$

$$[\mathrm{QD}\cdot\mathrm{P}]_t = [\mathrm{QD}\cdot\mathrm{P}]_0 e^{-k'_{\mathrm{off}}\,t} \tag{E.10}$$

The change in FRET ratio over time (**Equation E.12**) can then be derived using **Equation E.4** and **Equation E.6** to obtain the relationship between the FRET ratio and the concentration of bound protein (**Equation E.11**), where $F_0 = F_{\max}$ and the protein concentration is equal to $[\mathrm{QD}\cdot\mathrm{P}]_0 - [\mathrm{QD}\cdot\mathrm{P}]_t$, which is inserted into (**Equation E.10**). **Equation E.12** is used to fit all dissociation FRET-time profiles, unless otherwise state.

$$F_t = F_0\,\theta_t - at = F_0 \cdot \frac{[\mathrm{QD}\cdot\mathrm{P}]_t}{[\mathrm{P}]_t + [\mathrm{QD}\cdot\mathrm{P}]_t} - at = F_0\frac{[\mathrm{QD}\cdot\mathrm{P}]_t}{[\mathrm{QD}\cdot\mathrm{P}]_0} - at \tag{E.11}$$

$$F_t = F_0 e^{-k'_{\text{off}}\, t} - at \qquad \text{E.12}$$

E.2 Model of QR Glycan Display and Lectin Binding

Though exact values for hydrodynamic length (L_h) and diameter (D_h) cannot be obtained due to the limitations of DLS. Approximate values for these can be estimated by adding the average hydrodynamic length of the bound LA-EG$_{11}$-DiMan ligands (l_{h}) onto the known core dimensions of the QR. l_{h} can be determine from the QD-EG$_{11}$-DiMan scaffolds described previously in Guo, *et al.*, 2017.[8] Here, the hydrodynamic size and core size were obtained to be 9.5 ± 0.1 nm and 3.7 ± 2.1 nm (**Figure E.2**), obtained by DLS and STEM, respectively, which provides a $l_{h,ligand}$ for the surface LA-EG$_{11}$-DiMan ligands of 2.9 ± 1.1 nm (where, $l_{\text{h}} = ½(D_{\text{h}}^{\text{QD}} - D_{\text{core}}^{\text{QD}})$; **Figure E.1**). The average core length and diameter of QR-DiMan were obtained by STEM to be 14.8 ± 5.7 nm and 3.1 ± 0.7 nm, respectively (**Figure 4.2**). Thus, the average hydrodynamic length and diameters can be estimated to be around L_{h} = 20.6 ± 6.1 nm and D_{h} = 8.9 ± 2.1 nm, respectively.

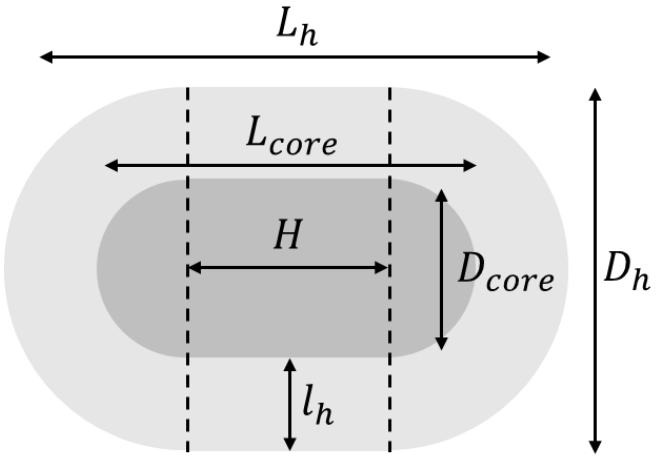

Figure E.1 2-dimensional schematic depicting the core size (darker grey) and hydrodynamic size (lighter grey) of a QR. Where D_{h} and L_{h} are the glycan-QR hydrodynamic diameter and length, respectively, D_{core} and L_{core} are the QR core diameter and length, H is the height of the cylindrical section and l_{h} is the hydrodynamic length of the ligand.

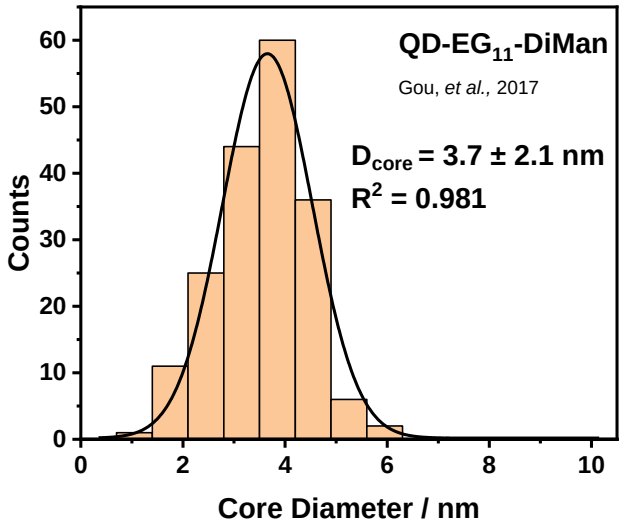

Figure E.2 Histogram of the core diameter distribution, fitted with Gaussian fits, obtained from the STEM images of QD-EG$_{11}$-DiMan reported in Guo, et al., 2017.[8] Where D_{core} is the mean core diameter.

To compare the protein binding contact area with the dimensions of the glycan-QR, the inter-binding site dimensions for DC-SIGN and DC-SIGNR can be obtained based on the distances between the primary Ca^{2+} ions associated with binding from the models based on 1xar and 1k9i crystal structures of the CRDs.[15, 16] This provides DC-SIGNR dimensions of approx. 3.8 × 8.0 nm with an average diagonal primary Ca-Ca distance of ~8.8 nm (**Figure E.3A**), and DC-SIGN dimensions of approx. 3.9 × 4.0 nm with an average diagonal primary Ca-Ca distance of ~5.6 nm (**Figure E.3B**).

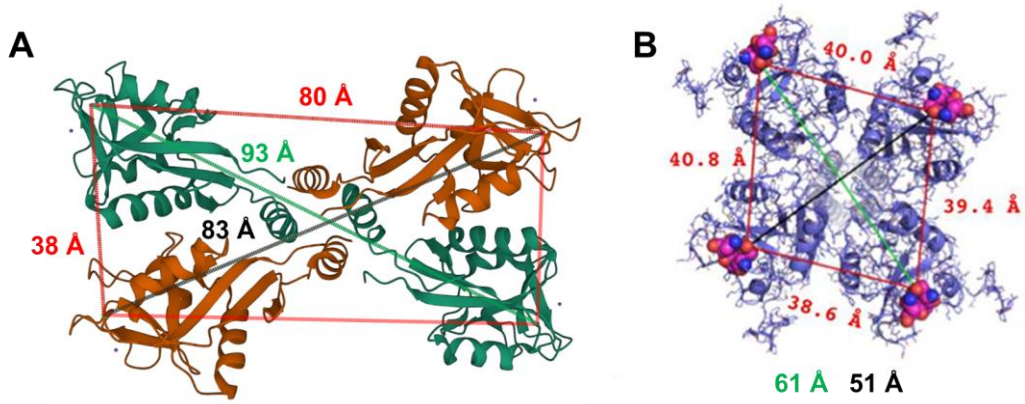

Figure E.3 (A) Dimer of DC-SIGNR CRDs dimers obtained by XRC (PDB: 1xar),[15] and (B) model of tetramer of DC-SIGN CRDs based on superimposition of individual CRDs from PDB: 1k9i.[16] (Red: primary Ca-Ca distance of adjacent CRDs; black: shortest primary Ca-Ca distance of diagonal CRDs; green: longest primary Ca-Ca distance of diagonal CRDs).

To identify the role of curvature, the theoretical separation distance between the protein binding contact area and the QR glycan surface can be estimated by assuming that the rod is a hard body and that the contact surface of the protein is a rigid plane. Here, the minimum separation distance of the extremities of the contact plane with the hydrodynamic surface of the glycan-QR, d, can be obtained by **Equation E.13**, as shown in **Figure E.4** (where p is the primary Ca-Ca distance between the binding sites furthest from the QR surface, d is the separation distance and D_h is the hydrodynamic diameter). This equation can also be represented in terms of surface curvature, κ, in order to provide a more universal description of separation distance, (where κ is the inverse of the radius of curvature).

$$d = \frac{D_h - \sqrt{{D_h}^2 - p^2}}{2} = \frac{2 - \sqrt{4 - p^2\kappa^2}}{2\kappa} \qquad \text{E.13}$$

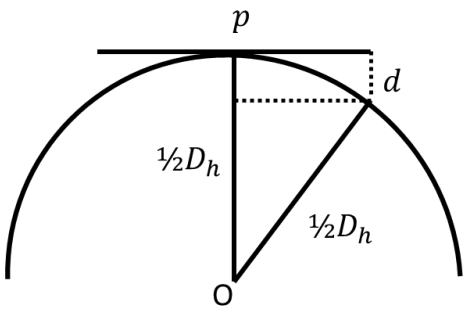

Figure E.4 Side-on schematic demonstrating separation distance between the lectin binding contact area to the glycan-QR hydrodynamic surface (d) for a rod of hydrodynamic diameter of D_h with respect to the inter-binding site distance furthest from the rod surface (p). Where d can be obtained by Equation E.13.

For DC-SIGNR at the spherical ends, p equals to the diagonal of the contact area which is ~8.8 nm, which equates to a d of ~4 nm. At the cylindrical ends, however, the longer length of DC-SIGNR can align itself along the longer length of the QR meaning only the shorter inter-CRD distance (p = 3.8 nm) contributes to separation distance, leading to a 10-fold smaller d of 0.4 nm. For DC-SIGN, a p of ~5.6 nm leads to a d of ~1 nm. In order to estimate the theoretical separation distance required to enable binding of all CRDs to the glycan surface, the compression length of the LA-EG$_{11}$-DiMan ligand can be estimated by subtracting the length of the rigid LA substituent (l_{LA} ~0.8 nm;

PDB ID: LPA) from the average hydrodynamic ligand length ($l = l_{\mathrm{h}} - l_{\mathrm{LA}}$ ~ 2.9 – 0.8 = 2.1 nm). The difference between d and l therefore provide a rationale behind the observed differences in binding modes between the two proteins and the two QR regions, where if $d > l$ then the glycan surface is unable to bridge all four CRDs, thus it is must find glycans on neighbouring QR-DiMan particles in order to maximally occupy its binding sites. However, if $d < l$ then the glycans only need to be compressed slightly to bridge all binding sites, allowing for simultaneous binding.

Appendix F
Preliminary Experiments

F.1 Investigating QD-Man Glycan Density on MLGIs

QD_{550} was capped with dihydrolipoic acid-tri(ethylene glycol)-mannose (denoted as DHLA-EG_3-Man; **Figure F.1**) and LA-EG_3-OH, as described in **Section 3.2.1**, using QD:Man:OH molar ratios of 1:700:0, 1:525:175 and 1:350:350, 1:175:525 to yield QD-$Man_{100\%}$, QD-$Man_{75\%}$/-$OH_{25\%}$, QD-$Man_{50\%}$/-$OH_{50\%}$, QD-$Man_{25\%}$/-$OH_{75\%}$ with D_{h} values of 9 ± 2, 9 ± 3, 8 ± 2 and 8 ± 2 nm, respectively.

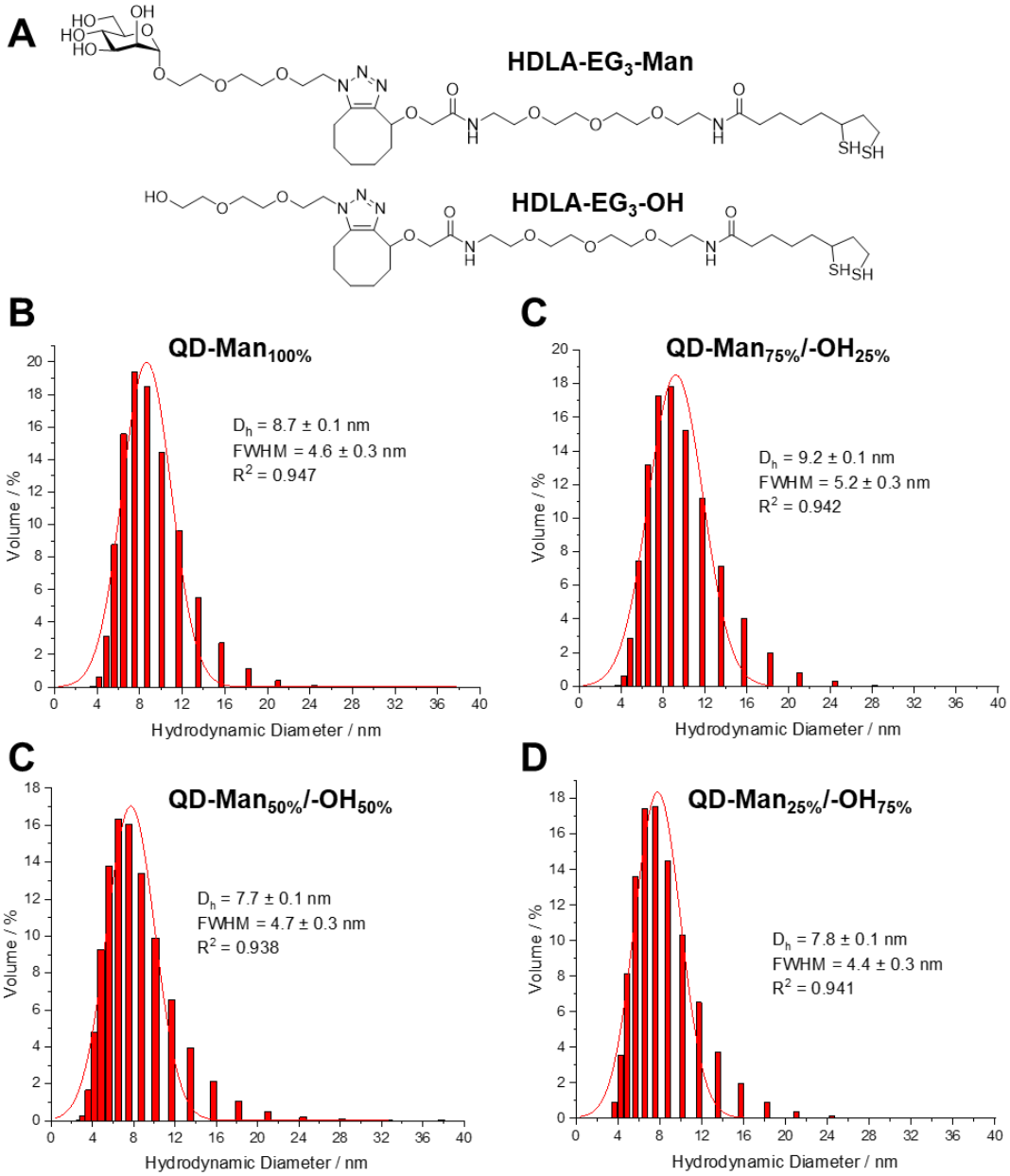

Figure F.1 (A) Skeletal structures of DHLA-EG$_3$-Man and DHLA-EG$_3$-OH. Volume population hydrodynamic diameter distribution histograms, fitted with the Gaussian equation, for QDs capped with (B) 100% DHLA-EG$_3$-Man (QD-Man$_{100\%}$), (C) 75% DHLA-EG$_3$-Man and 25% DHLA-EG$_3$-OH (QD-Man$_{75\%}$/-OH$_{25\%}$), (D) 50% DHLA-EG$_3$-Man and 50% DHLA-EG$_3$-OH (QD-Man$_{50\%}$/-OH$_{50\%}$), and (E) 25% DHLA-EG$_3$-Man and 75% DHLA-EG$_3$-OH (QD-Man$_{25\%}$/-OH$_{75\%}$).

QD-Man of varying glycan densities (40 nM, final concentration) were incubated with DC-SIGN Q274C-Atto594, DC-SIGN-C Q274C-Atto594 and DC-SIGN+R Q274C-Atto594 (produced, purified and labelled as described in **Section 2.1.4.2**; see **Appendix A.1** for amino acid identity) at varying protein:QD molar ratios (PQRs) in binding buffer (100 µL, final volume) in a 96-well plate (Greiner Bio-One µ-Clear®), which had been passivated with

BSA (1 mg mL^{-1}), at r.t. for 20 min. Fluorescence measurements were measured using a BGM Labtech CLARIOstar Plus microplate reader and FRET ratios were obtained as the ratio between the areas under the dye and QD fluorescence peaks, at 626 nm and 551 nm, respectively. FRET ratios were divided by the percentage of glycan surface coverage and plotted against PQR, as displayed in **Figure F.2**. Results demonstrate that 100% Man coverage is the optimal glycan valency for DC-SIGN (resulting in stronger binding or higher binding affinity or a closer donor-acceptor distance), where all Man coverages result in similar relative FRET ratio signals. DC-SIGN-C displays the opposite, where 100% Man coverage is less optimal and all other densities produce similarly stronger relative FRET ratio signal, demonstrating a loss of specificity upon removal of the C-terminal segment. DC-SIGN+R appears to show a near complete loss of binding at high Man coverages which increases with decreasing Man coverage, thus this mutation may demonstrate more DC-SIGNR binding behaviour (*e.g.* crosslinking) at higher Man coverages.

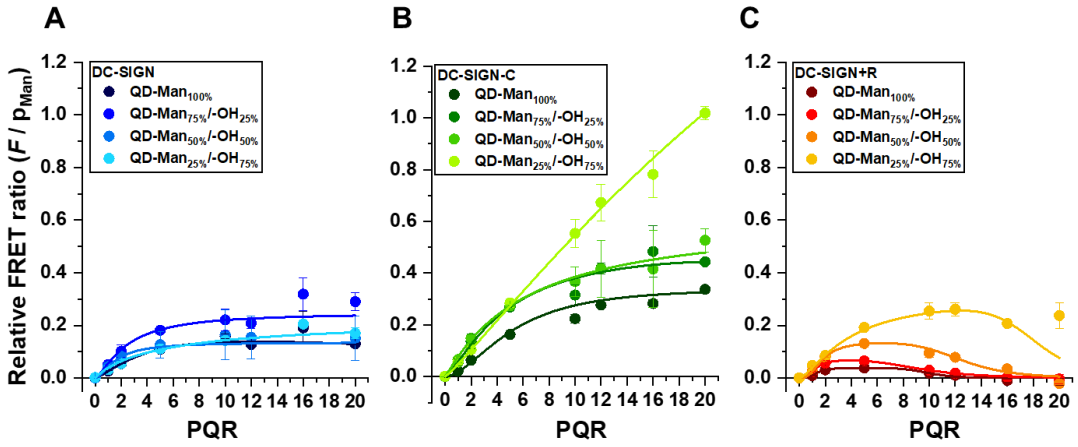

Figure F.2 Plots of the relative FRET ratios (*i.e.* $\frac{\int I_A d\lambda}{\int I_D d\lambda} \div p_{Man}$, where I is the fluorescence intensity under of the acceptor, A, and donor, D, and p_{Man} is the percentage population of man ligand per QD) against the protein:QD molar ratio (PQR) for QD-Man$_{100\%}$, QD-Man$_{75\%}$/-OH$_{25\%}$, QD-Man$_{50\%}$/-OH$_{50\%}$ and QD-Man$_{25\%}$/-OH$_{75\%}$ (dark to light) incubated with Atto594 labelled (A) DC-SIGN, (B) DC-SIGN-C and (C) DC-SIGN+R. (Data fitted with Hill equations or bimodal Hill equation to highlight the trend in data, not for quantitative evaluation).

F.2 QD-EG$_3$-DiMan·lectin Binding Thermodynamics

QD-EG$_3$-DiMan was prepared as described in **Section 2.1.3**, by capping QD$_{550}$ with dihydrolipoic acid-tri(ethylene glycol)-mannose-α-1,2-mannose (DHLA-EG$_3$-DiMan), yielding a D_h of 9.1 ± 2.5 nm. QD-EG$_3$-DiMan was analysed for its binding thermodynamics with Atto594 labelled DC-SIGN and DC-SIGN-C, as described in **Section 3.3.2**, using temperatures of 25, 30 and 35 ^oC. Dissocation constant (K_d) values were obtained by the Hill equation (**Equation 3.1**) and enthalpy and entropy change terms (ΔH^o and ΔS^o, respectively) were obtained by Van 't Hoff analysis (**Equation 3.4**). K_d values for DC-SIGN binding with QD-EG$_3$-DiMan binding are not too dissimilar from that of QD-EG$_{11}$-DiMan at 25 ^oC (6.1 ± 0.6 nM for QD-EG$_3$-DiMan and 3.00 ± 0.04 nM, respectively). However, the thermodynamic profile is shifted, becoming less negative for both enthalpy and entropy (ΔH = -15 ± 3 kJ mol^{-1}; ΔS = 108 ± 11 J K^{-1} mol^{-1}). This entropic enhancement in comparison to the EG$_{11}$ may be because there is less change in flexibility upon binding to the rigid EG$_3$ ligand, however, this also affords a reduced enthalpy term which may be because the rigid linker also forces the CRDs into a strained arrangement in order to achieve binding. DC-SIGN-C shifts to a very similar thermodynamic profile (ΔH = -13 ± 7 kJ mol^{-1}; ΔS = 114 ± 24 J K^{-1} mol^{-1}), whereby its enthalpy and entropy changes are enhanced and reduced, respectively, in comparison to QD-EG$_{11}$-DiMan binding. This could suggest that the flexible CRD arrangement of DC-SIGN-C is restricted by the rigid EG$_3$ linker.

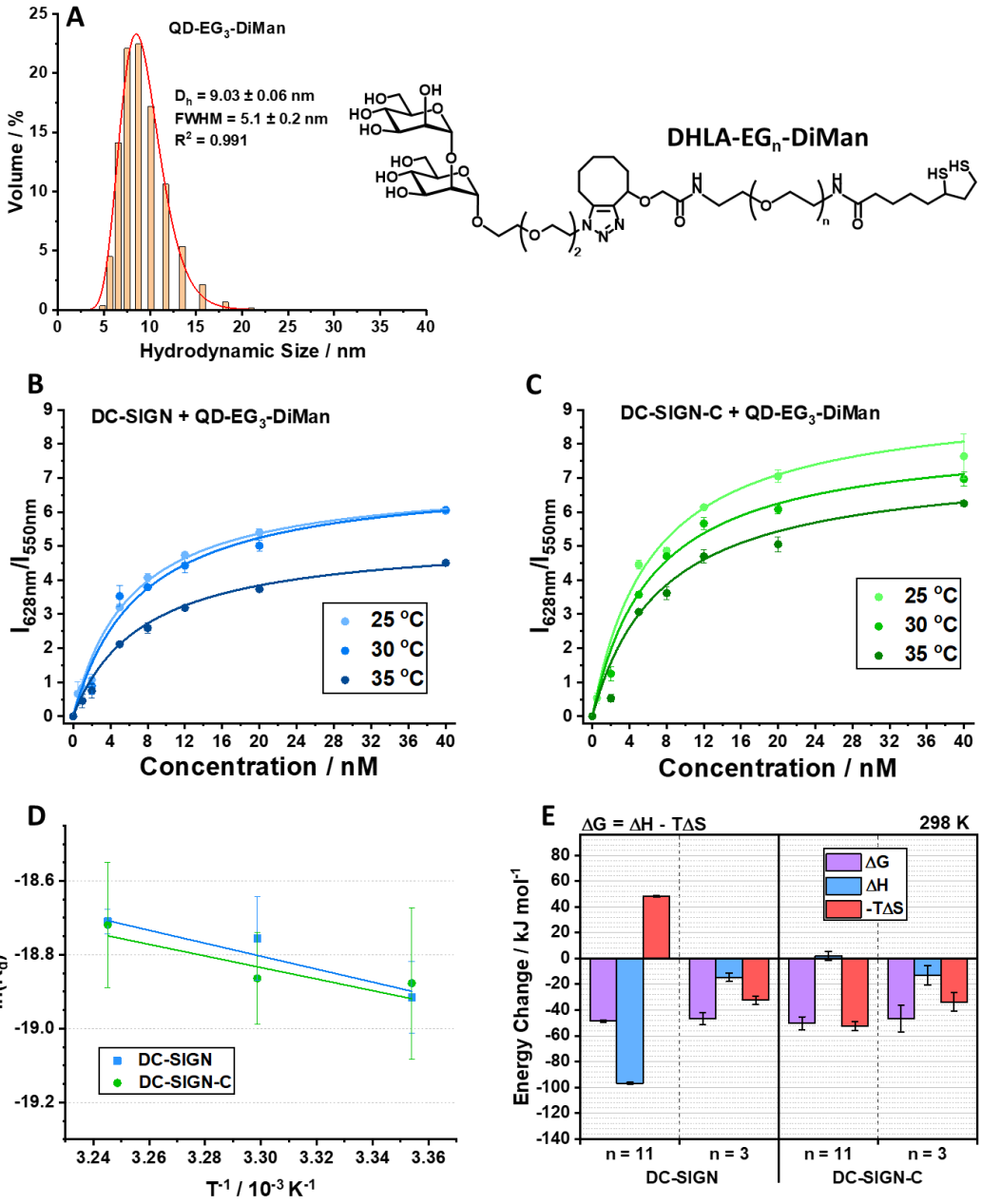

Figure F.3 (A) Volume population hydrodynamic diameter distribution histogram of QD capped with DHLA-EG3-DiMan, with a skeletal structure of DHLA-EGn-DiMan. Binding curves fitted with Equation 3.1, showing the relationship between the FRET ratio at different temperatures for QD-EG3-DiMan with labelled protein, for QD-EG3-DiMan with Atto594 labelled (B) DC-SIGN (PQR 1:1) and (C) DC-SIGN-C (PQR 1:1); (D) Van 't Hoff plots fitted with Equation 3.4, showing the relationship between $\ln K_d$ and $1/T$, derived from the FRET binding curves of QD-EG3-DiMan with DC-SIGN (blue squares), DC-SIGN-C (green circles); and (E) A graphical comparison of the enthalpy and entropy change contributions to the Gibbs free energy change of the binding of DC-SIGN, DC-SIGN-C with QD-EGn-DiMan with linker lengths n = 3 or 11, at 298 K (see Chapter 3 for n = 11).

References

1. H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327-1335.
2. H. Feinberg, D. A. Mitchell, K. Drickamer and W. I. Weis, *Science*, 2001, **294**, 2163-2166.
3. M. Shimojima, A. Takenouchi, H. Shimoda, N. Kimura and K. Maeda, *Arch. Virol.*, 2014, **159**, 2023-2031.
4. N. P. Chung, S. K. Breun, A. Bashirova, J. G. Baumann, T. D. Martin, J. M. Karamchandani, J. W. Rausch, S. F. Le Grice, L. Wu, M. Carrington and V. N. Kewalramani, *J. Biol. Chem.*, 2010, **285**, 2100-2112.
5. C. W. Davis, H. Y. Nguyen, S. L. Hanna, M. D. Sanchez, R. W. Doms and T. C. Pierson, *J. Virol.*, 2006, **80**, 1290-1301.
6. A. Marzi, T. Gramberg, G. Simmons, P. Möller, A. J. Rennekamp, M. Krumbiegel, M. Geier, J. Eisemann, N. Turza, B. Saunier, A. Steinkasserer, S. Becker, P. Bates, H. Hofmann and S. Pöhlmann, *J. Virol.*, 2004, **78**, 12090-12095.
7. D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo and D. Zhou, *J. Am. Chem. Soc.*, 2020, **142**, 18022-18034.
8. Y. Guo, I. Nehlmeier, E. Poole, C. Sakonsinsiri, N. Hondow, A. Brown, Q. Li, S. Li, J. Whitworth, Z. Li, A. Yu, R. Brydson, W. B. Turnbull, S. Pohlmann and D. Zhou, *J. Am. Chem. Soc.*, 2017, **139**, 11833-11844.
9. Y. Watanabe, J. D. Allen, D. Wrapp, J. S. McLellan and M. Crispin, *Science*, 2020, **369**, 330-333.
10. Sars-CoV-2 circulating variants, https://viralzone.expasy.org/9556, (accessed 20/12/2022, 2022).
11. H. D. Hill, J. E. Millstone, M. J. Banholzer and C. A. Mirkin, *ACS Nano*, 2009, **3**, 418-424.
12. D. Zhou, L. Ying, X. Hong, E. A. Hall, C. Abell and D. Klenerman, *Langmuir*, 2008, **24**, 1659-1664.
13. L. J. Charbonnière and N. Hildebrandt, *Eur. J. Inorg. Chem.*, 2008, **2008**, 3241-3251.
14. D. Zhou, Y. Li, E. A. H. Hall, C. Abell and D. Klenerman, *Nanoscale*, 2011, **3**, 201-211.
15. H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327-1335.
16. G. Tabarani, M. Thépaut, D. Stroebel, C. Ebel, C. Vivès, P. Vachette, D. Durand and F. Fieschi, *J. Biol. Chem.*, 2009, **284**, 21229-21240.