Université de Lille Faculté des Sciences et Technologies

Ecole Doctorale Biologie-Santé

THÈSE

Pour l'obtention du grade de DOCTEUR

Discipline : Aspects moléculaires et cellulaires de la biologie

Spécialité : Biochimie et Biologie Moléculaire

Présenté et soutenue publiquement par

STEFANIA ROBAKIEWICZ

Le 17 juillet 2020

Minimal Structural Glyco-Epitope For Antibody Recognition

Structure minimale de glyco-épitope pour la récognition des anticorps

Directeur de thèse: Dr. Julie Bouckaert

Composition du jury

Prof. Jesús JIMÉNEZ-BARBERO	Examinateur	CIC bioGUNE
Dr. Françoise JACOB- DUBUISSON	Examinateur	Institut Pasteur de Lille
Prof. Magdalena KRÓL	Examinateur	Warsaw University of Life Sciences
Prof. Francisco CORZANA	Rapporteur	Universidad de La Rioja
Prof. Franck FIESCHI	Rapporteur	Institut de Biologie Structurale
Dr. Julie Bouckaert	Directeur de these	CNRS

GLYCOVAX is a training network for the rational design of the next generation of well-defined glycoconjugate vaccines. The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 675671.

ACKNOWLEDGEMENTS

I would like to thank the GLYCOVAX network and the Computational Molecular Systems Biology team form *Unité de Glycobiologie Structurale et Fonctionnelle* for having trusted me and choosing my candidacy for this position. I need to thank my supervisor, Dr Julie Bouckaert, for providing a side project which in time became the main subject of my studies.

I am deeply grateful to the members of my *Comité de Suivi Individuel*, Dr Françoise Jacob-Dubuisson and Prof. Jesús Jiménez-Barbero, for their guidance, incredible support, and determination to help in any situation.

I would like to express my deepest gratitude to Prof. Jiménez-Barbero for taking me under his wing, and to all his team at CICbioGUNE for welcoming me at their institute and making me feel at home. Their continuous help and support are invaluable and contributed significantly to the outcome of this study. I must thank Prof Nicola Abrescia for his patience, readiness to help, and all the useful things I learned from him about protein expression and cryogenic electron microscopy. I would like to especially thank Dr Ana Arda for being my personal teacher and guide in NMR spectroscopy, Dr Ana Gimeno for her sense of humour and expertise on ITC, Sandra Delgado for her endless energy and immense help around the lab, and Diego Charro for his positive attitude and assistance in protein expression. I must also acknowledge Dr Adriana Rojas who was an incredibly patient and friendly crystallography teacher. Last, but not least, I want to thank all my lab mates, who not only provided support when needed but also made my stay at CICbioGUNE a very pleasant experience. I would not manage to accomplish what I did without the contribution from each and every one of them.

I am grateful to Dr Roberto Adamo and Dr Maria Romano from GSK Siena for hosting me in their laboratory where I have expanded my knowledge on vaccine development. I would also like to thank Dr Barbara Brogioni for her expertise in SPR and her guidance throughout my stay.

I need to thank Dr Niels Reichardt, Dr Sonia Serna, and Begoña Echeverria from CICbiomaGUNE for providing the glycan microarrays and the paucimannose *N*-glycans for the recognition studies.

I feel obliged to thank Prof Savvas Savvides, Dr Kenneth Verstraete, and Ann Dansercoer from VIB for their help in vector development.

I am thankful to Kim van Noort and Ruud Wilbers from Wageningen University and Research Centre for providing the glycoproteins for SPR experiments.

I am grateful to the Doctoral Schol and Prof Philippe Delannoy for their guidance.

I would like to express my enormous gratitude to Dr Robert Schneider for his incredible support and friendship.

I deeply appreciate the constant help and support of my lab friends Aleksandra Łoczechin, Milica Budimir, João Neves, and James Biwi. Their presence made my PhD a little more fun and I thank them for that.

Lastly, I would like to thank my closest friends and family for their endless love and support during this demanding and difficult time. They are all an excellent example of the saying “a friend in need is a friend indeed”. But my deepest gratitude goes to my parents. All I have achieved I owe to them and for this reason, I dedicate this work to my mum and dad.

PUBLICATIONS

Published:

Dumych T, Yamakawa N, Sivignon A, Garenaux E, Robakiewicz S, Coddeville B, Bongiovanni A, Bray F, Barnich N, Szunerits S, Slomianny C, Herrmann M, Gouin SG, Lutsyk AD, Munoz LE, Lafont F, Rolando C, Bilyy R, Bouckaert JMJ. **2018**. Oligomannose-rich membranes of dying intestinal epithelial cells promote host colonization by adherent-invasive *E. coli*. *Front Microbiol*. 9: 742.

Bastin-Héline L, de Fouchier A, Cao S, Koutroumpa F, Caballero-Vidal G, Robakiewicz S, Monsempes C, François MC, Ribeyre T, Maria A, Chertemps T, de Cian A, Walker WB, Wang G, Jacquin-Joly E, Montagné N. **2019**. A novel lineage of candidate pheromone receptors for sex communication in moths. Elife 8: pii: e49826.

Sweeney S, McArdle P, Taciak B, Robakiewicz S, Król M, Murphy PV. **2020**. Migrastatin analogues with an (E)-alkene at the ring C-3: synthesis, conformational analysis and biological evaluation. Arkivoc.

In preparation:

Robakiewicz S, Arda A, Serna S, Gimeno A, Echeverria B, Delgado S, Charro D, Dansercoer A, Verstraete K, van Noort K, Wilbers R, Savvides SN, Abrescia NGA, Reichardt NC, Jiménez-Barbero J, Bouckaert J. The minimal epitope for recognition of paucimannosidic *N*-glycans by Mannitou IgM.

COMMUNICATIONS

Oral presentations:

1. Robakiewicz S, Arda A, Serna S, Gimeno A, Echeverria B, Delgado S, Charro D, Abrescia NGA, Reichardt NC, Jiménez-Barbero J, Bouckaert J. *The Recognition of Paucimannosidic N-Glycans by Mannitou IgM* presented on the Introductory Workshop on Biomedical Glycoscience; 3-5 June 2019; San Sebastian, Spain.

Poster presentation:

1. Robakiewicz S, Arda A, Serna S, Gimeno A, Echeverria B, Delgado S, Charro D, Abrescia NGA, Reichardt NC, Jiménez-Barbero J, Bouckaert J. *Mannitou Antibody as a Potential Diagnostic and Therapeutic Tool* presented on Immunoshape International Symposium on Glycoimmunology; 11-13th July 2018; San Sebastian, Spain.

2. Robakiewicz S, Arda A, Serna S, Gimeno A, Echeverria B, Delgado S, Charro D, Abrescia NGA, Reichardt NC, Jiménez-Barbero J, Bouckaert J. *Mannitou Antibody as a Potential Diagnostic and Therapeutic Tool* presented on the 2nd PSL Chemical Biology Symposium; 17-18th January 2019; Paris, France.

3. Robakiewicz S, Arda A, Serna S, Gimeno A, Echeverria B, Delgado S, Charro D, Abrescia NGA, Reichardt NC, Jiménez-Barbero J, Bouckaert J. *The Recognition of Paucimannosidic N-Glycans by Mannitou IgM* presented on the 25th International Symposium on Glycoconjugates; 25-31st August 2019; Milan, Italy.

ABSTRACT

The biological importance of glycosylation in health and disease is broadly acknowledged. The truncated, mannose-terminating structures consisting of 1–3 mannose residues, two *N*-acetylglucosamines, and a variable number of fucose moieties are termed paucimannose. Paucimannosidic *N*-glycans are abundantly expressed in plants and invertebrates. However, in vertebrates their presence is restricted to some pathophysiological conditions, such as cancer, immune disorders, infections, and inflammation, and in healthy individuals, they are detectable only in trace amounts. Mannitou, a murine monoclonal antibody, has been demonstrated to specifically recognise paucimannose glycoepitopes.

An attempt to characterise Mannitou IgM structure has been made by applying homology modelling, cryo-electron microscopy, and crystallisation techniques. Full-length Mannitou antibody has been generated using hybridoma technology. Recombinant Mannitou Fab has been successfully transiently expressed in HEK293T cells. The binding specificity of Mannitou towards different paucimannose *N*-glycans has been unravelled by a combination of experimental methods. The microarray screening revealed the minimal glyco-epitope to be $Man_2GlcNAc_2$. In turn, $Man_3GlcNAc_2$ manifested one of the strongest interactions with Mannitou antibody. Molecular recognition studies, employing surface plasmon resonance measurements and isothermal titration calorimetry, established a micromolar binding affinity of Manniotu antibody towards $Man_3GlcNAc_2$ glycan. The mapping of the binding epitope by saturation transfer difference nuclear magnetic resonance demonstrated Manα1-3 as the main residue involved in Mannitou antibody recognition. The upregulation of paucimannosidic *N*-glycans in pathophysiological conditions makes Mannitou antibody a promising diagnostic and therapeutic tool.

For determining the minimal carbohydrate structure required for mimicking the antigenic activity of the native MenX polysaccharide, surface plasmon resonance studies have been performed. The experiments involved studying the binding interactions between an anti-MenX antibody and *Neisseria meningitides* serogroup X capsular oligosaccharides of different length. The results suggest that the minimal saccharide

portion capable of ensuring protection against MenX infections may be DP5, making it a promising candidate for vaccine development.

Keywords: paucimannose *N*-glycans; Immunoglobulin M; Mannitou antibody; *Neisseria meningitides* serogroup X; monoclonal antibody; antigen-binding fragment; minimal glyco-epitope; molecular recognition studies; glycan microarrays; Surface Plasmon Resonance; Isothermal Titration Calorimetry; Saturation Transfer Difference Nuclear Magnetic Resonance.

R ÉSUM É

L'importance biologique de la glycosylation pour la santé et la maladie est largement reconnue. Les structures tronquées à terminaison mannose consistant en 1 à 3 résidus de mannose, deux N-acétylglucosamines et un nombre variable de fragments fucose sont appelées paucimannose. Les *N*-glycanes paucimannosidiques sont abondamment exprimés dans les plantes et les invertébrés. Cependant, chez les vertébrés, leur présence est limitée à certaines conditions pathophysiologiques, telles que le cancer, les troubles immunitaires, les infections et l'inflammation, et chez les individus en bonne santé, ils ne sont détectables qu'en quantités infimes. Il a été démontré que le Mannitou, un anticorps monoclonal murin, reconnaît spécifiquement les glycoépitopes de paucimannose.

Une tentative de caractérisation de la structure de Mannitou IgM a été faite en appliquant des techniques de modélisation d'homologie, de microscopie cryoélectronique et de cristallisation. L'anticorps complet de Mannitou a été généré en utilisant la technologie des hybridomes. Le Fab recombinant de Mannitou a été exprimé avec succès de manière transitoire dans des cellules HEK293T. La spécificité de liaison de Mannitou envers différents *N*-glycanes de paucimannose ont été élucidées par une combinaison de méthodes expérimentales. Le criblage par microréseau a révélé que le glyco-épitope minimal était $Man_2GlcNAc_2$. À son tour, $Man_3GlcNAc_2$ a manifesté l'une des interactions les plus fortes avec l'anticorps Mannitou. Des études de reconnaissance moléculaire, utilisant des mesures de résonance plasmonique de surface et une calorimétrie de titrage isotherme, ont établi une affinité de liaison micromolaire de l'anticorps Manniotu envers le glycane $Man_3GlcNAc_2$ (K_d = ~50 µM). La cartographie de l'épitope de liaison par résonance magnétique nucléaire de transfert de saturation a démontré que Manα1-3 est le principal résidu impliqué dans la reconnaissance des anticorps de Mannitou. La régulation positive des *N*-glycanes paucimannosidiques dans des conditions physiopathologiques fait de l'anticorps Mannitou un outil diagnostique et thérapeutique prometteur.

Pour déterminer la structure minimale des glycanes requise pour mimer l'activité antigénique du polysaccharide MenX natif, des études de résonance plasmonique de surface ont été réalisées. Les expériences ont consisté à étudier les interactions de liaison

entre un anticorps anti-MenX et des oligosaccharides capsulaires du sérogroupe X de *Neisseria meningitides* de différentes longueurs. Les résultats suggèrent que la portion minimale de saccharide capable d'assurer une protection contre les infections à MenX pourrait être DP5, ce qui en fait un candidat prometteur pour le développement de vaccins.

Mots-clés: paucimannose *N*-glycanes; Immunoglobuline M; anticorps Mannitou; *Neisseria meningitides* sérogroupe X; anticorps monoclonal; Fab; glyco-épitope minimal; études de reconnaissance moléculaire; microréseaux de glycanes; Résonance Plasmonique de Surface; Calorimétrie de Titrage Isotherme; Résonance Magnétique Nucléaire de Transfert de Saturation.

INDEX

ABBREVIATIONS

α-Glc	α-glucosidase
α-Man	α-mannosidase
AcNPV	*Autographa californica* nuclear polyhedrosis virus
ADCC	Antibody-dependent cellular cytotoxicity
ALG	Asparagine-linked glycosylation
ANN	Artificial neural networks
AOM	Amino octyl mannose
APC	Antigen-presenting cell
Asn	Asparagine
avDP	Average degree of polymerisation
BGH	Bovine growth hormone
BHK 21	Baby hamster kidney cells 21
BSA	Bovine serum albumin
CaMV	Cauliflower mosaic virus
CaPi	Calcium phosphate
CD-MPR	Cation-dependent M-6-P receptor
CDR	Complementarity-determining region
CH	Constant heavy region
CHO	Chinese hamster ovary cells
CL	Constant light region
CMP	Cytidine-5′-monophospho-N-acetylneuraminic acid
CMV	Cytomegalovirus

CNS	Central nervous system
CNX	Calnexin
CRD	Carbohydrate recognition domains
CRM	Cross-reacting material 197
CRT	Calreticulin
Cryo-EM	Cryogenic electron microscopy
DAMPs	Danger-associated molecular patterns
DC	Dendritic cell
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
Dol-P	Dolichyl pyrophosphate
DP	Degree of polymerisation
ds-scFv	Disulphide-bond stabilised scFv
DT	Diphtheria toxin
EBV	Epstein-Barr virus
EBNA1	Epstein-Barr virus nuclear antigen 1
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EF-1α	Elongation factor
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
Endo F	Endo-β-*N*-acetylglucosaminidase F
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
ERGIC-53	ER-Golgi intermediate compartment 53 kDa protein
Fab	Fragment antigen-binding

FBS	Fetal bovine serum
Fc	Fragment crystallisable
Fuc	Fucose
FUT8	α1,6-fucosyltransferase
Fv	Variable fragment
Gal	Galactose
GalNAc	*N*-acetylgalactosamine
GBM	Glioblastoma multiforme
GBP	Glycan-binding protein
Glc	Glucose
GlcNAc	*N*-acetylglucosamine
GnT-I	*N*-acetylglucosaminyltransferase I
GTase	Glycosyltransferase
HAT	Hypoxanthine-aminopterin-thymidine
HBS-EP	HEPES-buffered saline with 0.3 mM EDTA
HC	Heavy chain
HCAb	Heavy chain antibody
HCMV	Human cytomegalovirus
HEK 293	Human embryonic kidney cells
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hex	*N*-acetyl-β-hexosaminidase
HGPRT	Hypoxanthine-guanine phosphoribosyltransferase
HM	Heptyl mannose
HNE	Human neutrophil elastase
HSA	Human serum albumin

IgA	Immunoglobulin A
IgD	Immunoglobulin D
IgE	Immunoglobulin E
IGF-II/CI-MPR	Insulin-like growth factor II/cation-independent M-6-P receptor
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IHC	Immunohistochemistry
IMAC	Immobilised Metal-Ion Affinity Chromatography
IRES	Internal ribosomal entry site
ITC	Isothermal titration calorimetry
LC	Light chain
LLO	Lipid-linked oligosaccharide
mAb	Monoclonal antibody
MAG	Myelin-associated glycoprotein
MALDI-TOF MS	Matrix-assisted laser desorption ionisation – time of flight mass spectrometry
Man	Mannose
MenA	*Neisseria meningitidis* serogroup A
MenW	*Neisseria meningitidis* serogroup W
MenX	*Neisseria meningitidis* serogroup X
MHCII	Class II major histocompatibility complex
miniAb	Minibody
mrCLR	Mannose-recognising C-type lectin receptor
MΘ	Macrophage
MVA	Modified vaccinia virus Ankara

NeuNAc	*N*-acetylneuraminic acid
NHS	*N*-hydroxysuccinimide
NOE	Nuclear Overhauser effect
NPC	Neural progenitor cell
ORF	Open reading frame
OST	Oligosaccharyltransferase
PBS	Phosphate-buffered saline
PBST	Phosphate-buffered saline with Tween 20 detergent
PCR	Polymerase chain reaction
PCT	Pre-crystallisation test
PEG	Polyethylene glycol
PEI	Polyethyleneimine
pI	Isoelectric point
PMP	Paucimannosidic protein
PNGase F	Peptide-*N*-(*N*-acetyl-β-glucosaminyl) asparagine amidase
Pro	Proline
PS	Polysaccharide
RFU	Relative fluorescence unit
RNA	Ribonucleic acid
RNAi	RNA interference
RPMI	Roswell Park Memorial Institute 1640 medium
RT-PCR	Reverse transcription polymerase chain reaction
RU	Resonance unit
S/MAR	Scaffold/matrix attachment region
SA	Sialic acid

SAMPs	Self-associated molecular patterns
scFab	Single-chain Fab
scFv	Single-chain variable fragment
SD	Standard deviation
sdAb	Single-domain antibodies
SDS-PAGE	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
Ser	Serine
SFM	Serum-free medium
SFV	Semliki forest virus
Siglec	Sialic acid-binding immunoglobulin-type lectin
SPR	Surface plasmon resonance
STD - NMR	Saturation transfer difference nuclear magnetic resonance
STDD	Saturation transfer double-difference (spectrum)
SV40	Simian virus 40
T_C	T cytotoxic cell
TEM	Transmission electron microscopy
TGE	Transient gene expression
TGF	Transforming growth factor
TGN	Trans Golgi network
T_H	T helper cell
Thr	Threonine
TLR	Toll-like receptor
TMV	Tobacco mosaic virus
TT	Tetanus toxoid

UDP	Uridine diphosphate
VH	Variable heavy region
VIP36	36 kDa vesicular integral membrane protein
VL	Variable light region

GENERAL INTRODUCTION

1. Glycosylation

Proteins have one or more potential glycosylation sites, which may or may not be glycosylated with different glycan structures. More than 50% of naturally occurring proteins are glycosylated. Protein glycosylation patterns vary depending on the conditions. Heterogeneous glycoforms of the same protein can have different properties or biological activities which is the source of generating heterogeneity of protein forms and functions without the need for genetic alteration. Glycosylation influences solubility, serum half-life, immunogenicity, and biological activity of secreted glycoproteins. Glycosylation mediates cell-cell interactions, such as cell adhesion, cell infection by viruses and microorganisms, fertilisation, and signalling. Glyco-epitopes are recognised by two classes of receptors, glycan-binding proteins (lectins/adhesins) and glycan-binding antibodies.

In humans, the glycan structures attached to glycoproteins are composed of different monosaccharides, including D-α/β-mannose (Man), D-α-glucose (Glc), L-α-fucose (Fuc), D-β-galactose (Gal), D-α/β-*N*-acetylglucosamine (GlcNAc), D-α-*N*-acetylgalactosamine (GalNAc), and sialic acids (SA), e.g. D-α-*N*-acetylneuraminic acids (NeuNAc) (Fig.1). These residues can be linked together in linear or branched chains of varying length and complexity. The monosaccharides in the chain can each be linked together through different carbon atoms, e.g. creating α1-2 or α1-6 linkage. Additionally, each linkage may be either in the α form, where the glycosidic bond lies below the plane of the ring of the left-hand monosaccharide, or in the β form, where it lies above the plane of the ring (Loke et al., 2016).

Compared to proteins, glycan structures and compositions are much more diverse and, unlike proteins, it is not possible to predict their sequences based on a predetermined blueprint as they are not primary gene products. Instead, their "template-free" biosynthesis in the rough endoplasmic reticulum (ER) and Golgi apparatus is mediated by the sequential action of glycosyltransferases and glycosidases. Human cells

have a repertoire of around 250 glycosylation enzymes, each displaying a high substrate specificity, that operate sequentially to yield a varied set of glycosylation patterns existing in cells and tissues.

Despite it is still not fully understood how the glycan repertoire alters in response to changing conditions, it is known that carbohydrates can be either *N*- or *O*-linked to proteins. *N*-linked glycans share a common branched trimannosyl core $Man_3GlcNAc_2$-Asn, which may be elaborated to form a diversity of heterogeneous, often quite large, bi-, tri-, or tetra-antennary glycans. *N*-linked glycans can be classified into three types (Fig.2). Oligomannose or high-mannose *N*-glycans have core mannose residues in addition to those in the trimannosyl core structure (M5-M9). Complex *N*-glycans do not carry mannose residues additional to those of the trimannosyl core but usually both Man α1-3 and Man α1-6 arms are substituted with GlcNAc β1-2. These antennary GlcNAc residues may be extended with various monosaccharide residues to form additional branches, resulting in a high level of complexity. Hybrid *N*-glycans possess structural features of both oligomannose and complex type, with the Man α1-3 arm of the core elongated while the Man α1-6 arm left relatively unprocessed with mannosyl decorations. *O*-linked glycosylation is less comprehended than *N*-linked glycosylation. *O*-linked glycans do not share a single common core and tend to be smaller and less branched than *N*-linked glycans (Brooks 2009; Tjondro et al., 2019).

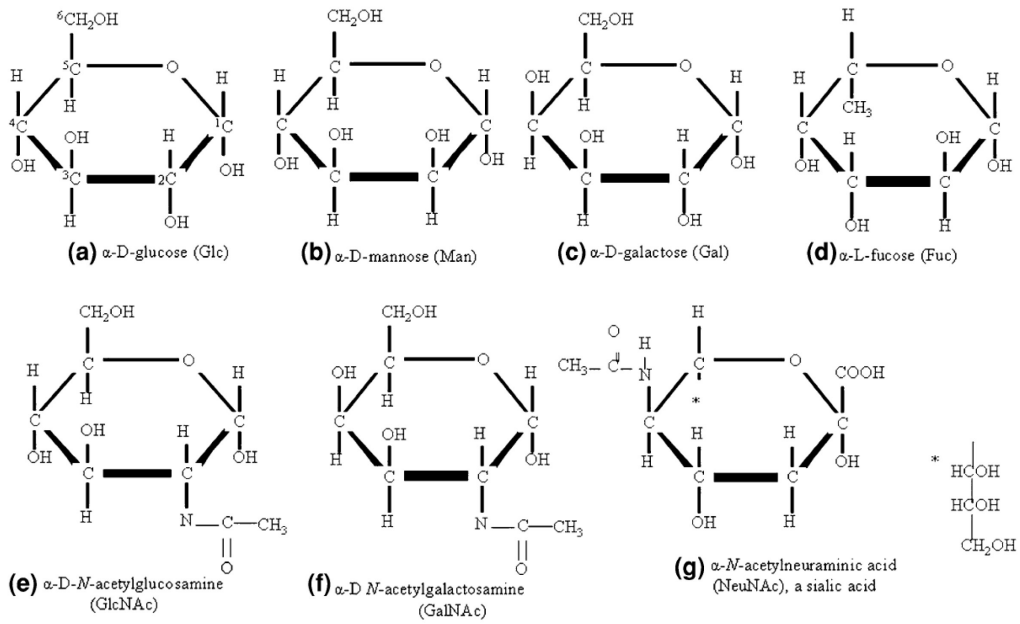


Figure 1. Haworth projection of the monosaccharide structures that make up human glycans. From: Brooks 2009.

1.1. *N*-glycans

Asparagine (*N*)-linked glycosylation is a type of protein glycosylation that uses mannose as a central building block. Mammalian *N*-glycosylation is largely restricted to the Asn-X-Ser/Thr/Cys (X≠Pro) consensus sequences. Protein *N*-glycans share a common trimannosylchitobiose core (Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn). *N*-linked oligosaccharides arise when blocks of 14 sugars are added cotranslationally to newly synthesised polypeptides in the endoplasmic reticulum. These glycans are then subjected to extensive modification as the glycoproteins mature and move through the ER via the Golgi compartments to their final destinations inside and outside the cell. In mature glycoproteins, *N*-linked glycan moieties are structurally diverse. The sugar composition and the number and size of branches in the sugar tree vary among glycoproteins, and cell types, tissues, and species. However, when initially added in the ER to growing nascent polypeptides, the glycans do not display such heterogeneity. The trimming and processing that the glycans undergo when the glycoprotein is still in the ER introduce only limited additional diversity, as the alterations are shared by all glycoproteins. Thus, the spectrum of glycoforms remains rather uniform until the glycoproteins reach the medial stacks of the Golgi apparatus, where structural diversification is introduced through a series of non-uniform modifications. This tremendous diversity in glycoconjugates is especially noticeable in vertebrate and plant cells. The switch from structural uniformity in the ER to diversification in the Golgi complex coincides with a marked change in glycan function. In the early secretory pathway, the glycans play a key role in protein folding, oligomerisation, quality control, sorting, and transport events. Later in the Golgi compartments, the glycans acquire more complex structures and novel functions that they display in the mature proteins (Helenius et al., 2001).

In mammals, *N*-linked protein glycosylation starts with the synthesis of a highly conserved lipid-linked oligosaccharide (LLO) on the ER membrane (Fig.3). Monosaccharide residues are sequentially added to dolichyl pyrophosphate (dol-P) carrier molecule on the cytosolic surface of the ER membrane. When two *N*-acetylglucosamines and five mannoses get connected, the oligosaccharide is flipped to the lumenal side of the membrane, and additional four mannose and three glucose residues

are included by a series of glycosyltransferases (GTase) to form the oligosaccharide core precursor $Glc_3Man_9GlcNAc_2$-P-P-dol (Fig.4). Conservation of this structure for more than a billion years of evolution strongly suggested that it serves a very important purpose. The completed core oligosaccharide is transferred from the dolichyl pyrophosphate carrier to a growing, nascent polypeptide chain, and is coupled through an *N*-glycosidic bond to the side chain of an asparagine residue by the oligosaccharyltransferase (OST) enzyme complex. It recognises a specific conformation of the glycosylation sequon, namely the Asn-X-Ser/Thr consensus sequence, transiently formed when the growing, nascent polypeptide chain emerges from the translocon. Immediately after coupling to the polypeptide chain, the three glucose residues are trimmed by ER α-glucosidase I (cleaves the outermost α1-2-linked glucose) and α-glycosidase II (removes the 2nd and 3rd α1-3-linked glucoses) and the terminal mannoses by α-mannosidase I. When the folded glycoprotein reaches the Golgi complex, the glycan chain undergoes further mannose trimming. A collection of processing mannosidases in the ER and Golgi apparatus remove the four α1-2 mannose residues to yield $Man_5GlcNAc_2$. Following the action of GlcNAc transferase I, which initiates branching of complex oligosaccharides, Golgi α-mannosidase II removes one α1-3- and one αl-6-linked mannose residue to yield $GlcNAcMan_3GlcNAc_2$. During subsequent terminal glycosylation, the attachment of new terminal monosaccharides including *N*-acetylglucosamine, galactose, sialic acid, and fucose may follow. In the end, just five sugars remain of the original core glycan. While the glycoforms in the ER are homogeneous, the Golgi-generated oligosaccharides are highly diverse and differ widely between species (Dean et al., 2014; Helenius et al., 2001; Moremen et al., 1994).

Although the ER and the Golgi apparatus are both components of the secretory pathway, they differ fundamentally in operational principle, overall architecture, and probably evolutionary origin as well. In the ER, most of the biosynthetic machinery faces the cytosol and can directly use precursors, such as sugar nucleotides and amino acids provided by cytosolic enzymes. After synthesis, the products are translocated across the membrane into the lumen. This applies to the polypeptide synthesis that occurs on membrane-bound ribosomes, as well as to the synthesis of phospholipids and, in part, of dolichol precursors for *N*-linked glycans. In those few cases where precursors are transported to the lumenal side, the transfer occurs almost exclusively through lipid-linked intermediates. The ER contains a high concentration of soluble molecular

chaperones and folding enzymes. In the Golgi apparatus, resident proteins (glycosidases, glycosyltransferases, proteases, lectins, permeases, and transport receptors) are membrane-bound. Therefore, the Golgi complex assembles its products on the inside surface of the membrane, with the lumen being occupied by substrate molecules and other cargo. The glycoprotein synthesis in Golgi makes use of soluble precursors, such as sugar nucleotides, that are imported from the cytosol through the membrane by specific transporters without the use of lipid-linked intermediates.

Unlike the ER, the Golgi complex does not have a rigorous system for controlling the fidelity of its biosynthetic processes. Although lacking ER-like quality control and degradation systems, the Golgi apparatus does not export unfinished products thanks to its multicompartmental architecture. As substrate molecules progressively move from the entry side (cis) to the exit side (trans) they are exposed to the full panel of modifying enzymes before being exported.

Like protein translocation, *N*-linked glycosylation evidently belongs to ER functions inherited from the prokaryotic, most likely archaeal, plasma membrane. In archaea, *N*-linked glycans are synthesised in the S-layer of the cell wall. As in eukaryotes, *N*-glycosylation makes use of dolichyl phosphate- and dolichyl pyrophosphate-linked oligosaccharides. In archaea, the oligosaccharides are transferred to asparagine side chains in the same Asn-X-Ser/Thr sequence motif that is used by the ER oligosaccharyltransferase in eukaryotic cells. The evolutionary origin of the Golgi complex together with its glycosylation machinery remain unclear. The sequences of Golgi glycosyltransferases suggest homology with cytosolic enzymes responsible for the use of sugar nucleotides, which suggests a different evolutionary origin to the ER machinery (Helenius et al., 2001).

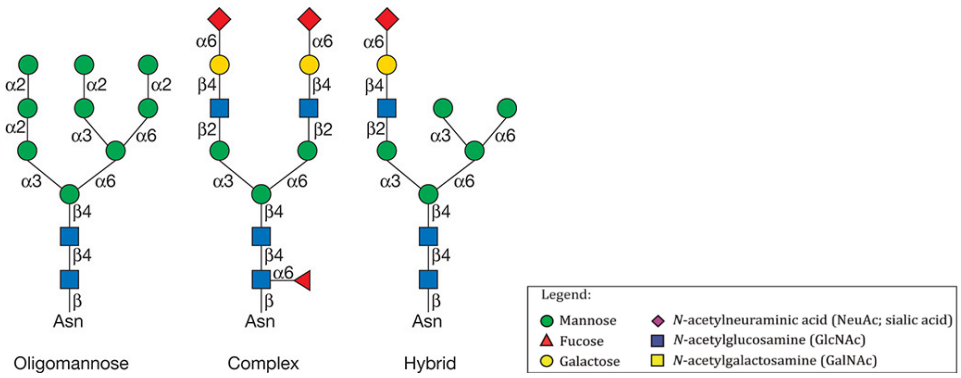


Figure 2. Types of N-glycans at Asn-X-Ser/Thr sequons in eukaryote glycoproteins. Each N-glycan contains the common core $Man_3GlcNAc_2$. From: Stanley et al., 2017.

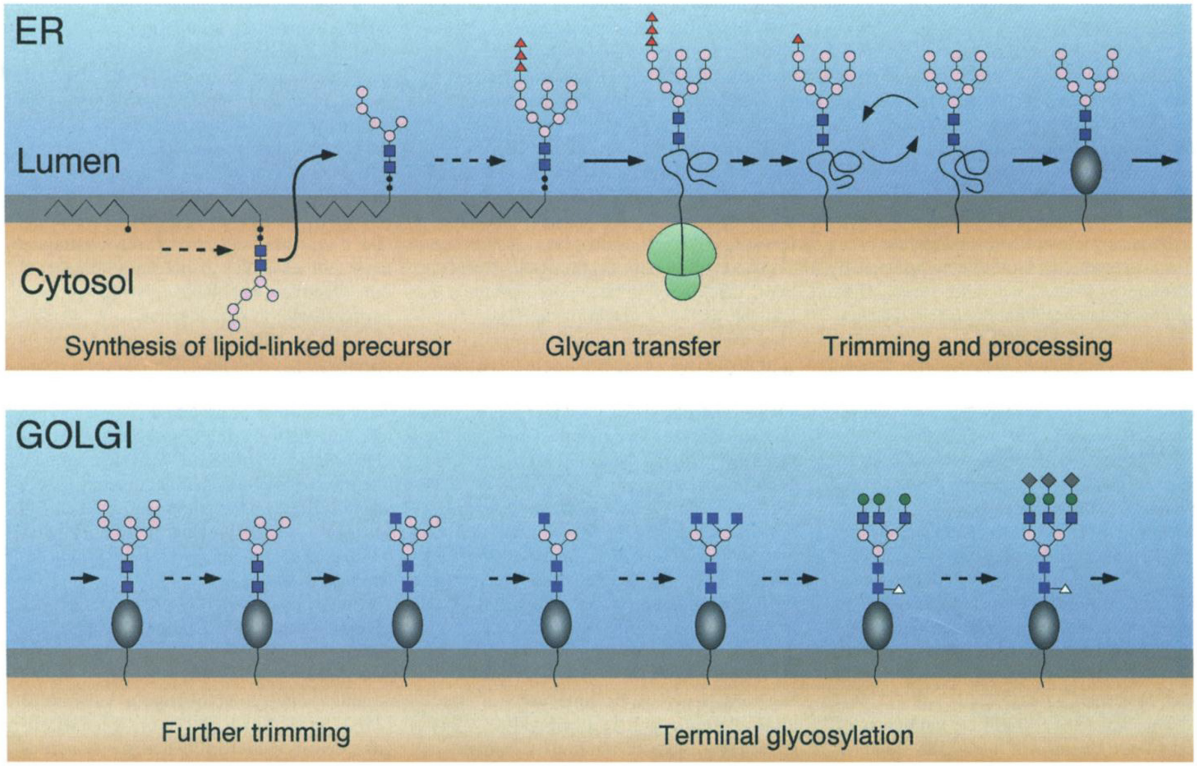


Figure 3. The *N*-linked protein glycosylation pathway. One of many possible terminal glycosylation alternatives is shown. The number of branches generated is variable, as are the number and identity of sugars added. Monosaccharides are added by glycosyltransferases during terminal glycosylation to produce complex *N*-linked glycans. Mature glycoproteins carry a mixture of complex glycans as well as some high-mannose glycans that have avoided terminal glycosylation. From: Helenius et al., 2001.

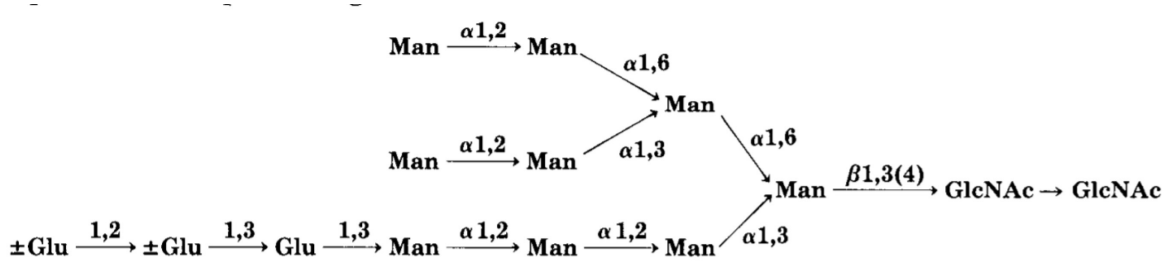


Figure 4. $Glc_3Man_9GlcNAc_2$ structure. From: Li et al., 1978.

1.1.1. Paucimannose *N*-glycans

Unconventional, immuno-centric, α- and β-mannose epitope-rich *N*-glycosylation form, termed paucimannosylation, have been recently discovered in humans. Paucimannosylation is a common modification of proteins expressed by lower organisms. However, recent reports indicate that higher organisms, including humans, also produce paucimannosidic proteins (PMPs) but in a more tissue-specific and physiology-dependent manner (Chatterjee et al., 2019). The truncated, mannose-terminating human paucimannosidic structures usually consist of 1–3 mannose residues, two *N*-acetylglucosamines, and a variable presence of fucose ($Man_{1-3}GlcNAc_2Fuc_{0-1}$) (Schachter 2009).

Although there have been a few reports on protein paucimannosylation in vertebrates over the past decades (Hase et al., 1982; Howard et al., 1982; Mutsaers et al., 1987; Argade et al., 1988), in general mammals have been considered to lack paucimannose *N*-glycans (Sarkar et al., 2006; Schachter 2009; Schachter et al., 2011). Until recently, paucimannose-rich *N*-glycosylation was thought to be limited to plants, e.g. *Arabidopsis thaliana* (Zeng et al., 2018) or *Lotus japonicus* (Dam et al., 2013), and invertebrates, such as *Caenorhabditis elegans* (Cipollo et al., 2002; Schachter 2009; Zhang et al., 2003) and *Drosophila melanogaster* (Altmann et al., 1995; Léonard et al., 2006; Sarkar et al., 2006). While paucimannosylation was found to play an important role in the development (Schachter et al., 2011) and immune response against bacterial pathogens (Shi et al., 2006), its exact biological functions and effector mechanisms in these organisms are yet to be understood. Interestingly, paucimannosidic *N*-glycan expression in vertebrates has been predominantly restricted to some pathophysiological conditions, like cancer (Chatterjee et al., 2019) but also immune disorders, infections, and inflammation (Tjondro et al., 2019). Elevated paucimannosylation has been observed in human colorectal cancer epithelial cells (Balog et al., 2012), human breast cancer epithelial cells (Lee et al., 2014), human pancreatic duct cancer cells (Zipser et al., 2012) ovarian carcinoma (Everest-Dass et al., 2016), lung carcinoma (Wang et al., 2018), skin carcinoma (Möginger et al., 2018), prostate carcinoma (Shah et al., 2015), glioblastoma multiforme (Becker at al., 2019), paraganglioma (Leijon et al., 2017), in pathogen-infected neutrophil-rich sputum derived from individuals with cystic fibrosis and upper

respiratory tract infections (Thaysen-Andersen et al., 2015; Venkatakrishnan et al., 2015), human buccal epithelial cells during *Candida albicans* infection (Everest-Dass et al., 2012), *Mycobacterium tuberculosis*-infected human macrophages (Hare et al., 2017), in exocrine ducts of inflamed mouse pancreas (Zipser et al., 2012), and kidney tissue of mice suffering from systemic lupus erythematosus (Hashii et al. 2009). Mammalian paucimannosidic proteins have also been associated with non-cancer-related cell development, proliferation, and stem cell differentiation. Physiologically, PMPs occur in human adult pancreatic stem cells (Zipser et al., 2012), human embryonic stem cells (Satomaa et al. 2009), human neural stem cells (Zipser et al., 2012; Dahmen et al., 2015), mouse embryonic neural stem cells (Yagi et al., 2012) and mouse early postnatal neural stem cells (Dahmen et al., 2015), murine synaptosomes (Trinidad et al., 2013), and rat brain tissue (Parker et al., 2013).

Paucimannosidic glyco-epitopes were demonstrated to be specifically expressed in mouse early postnatal neural progenitor cells (NPCs), human glioblastoma cells, and human macrophages. It has been established that paucimannosylation is functionally involved in the regulation of cell proliferation and differentiation *in vitro* and *in vivo* in mouse NPCs. Paucimannose *N*-glycans were shown to be predominantly carried by intracellular soluble proteins but were also present on membrane-associated (cell surface) proteins (Dahmen et al., 2015).

The influence on cell proliferation and differentiation of the human glioblastoma cell line A172 provided the first evidence that protein paucimannosylation might influence tumorigenic processes (Dahmen et al., 2015). Further studies by Becker et al. from 2019 have demonstrated the functional involvement of paucimannose *N*-glycans in a variety of tumorigenic processes in two cell lines representing glioblastoma of different aggressiveness, U-87 MG and U-138 MG, and in human patient tissue. Protein paucimannosylation was found at high levels in human glioblastoma multiforme (GBM) tissue, including infiltrated tumour cells, reactive astrocytes, and activated microglia cells, compared to healthy tissue, and reported to play a central role in GBM progression by inhibiting cell adhesion, proliferation, migration, and invasion of cancerous glial cells (Becker et al., 2019). The presence of paucimannosidic structures in reactive astrocytes and microglia in the tumour surrounding tissue is not surprising since protein paucimannosylation has been previously implicated in inflammatory processes (Loke et al., 2016).

PMPs have been functionally associated with roles in innate immunity upon secretion from activated human neutrophils, key granulocytic cells of the innate immune system. In 2015, Thaysen-Andersen et al. were the first to document that human neutrophils produce, store, and upon activation selectively secrete bioactive paucimannosidic glycoproteins into sputum of pathogen-infected lungs. Significant core fucosylation was observed for neutrophil PMPs, including azurocidin, cathepsin G, proteinase 3, neutrophil elastase, and myeloperoxidase. The compartment-specific assembly and storage of inflammation-associated paucimannosidic proteins were shown to be localised in azurophilic granules of neutrophils. The azurophilic granules, also known as primary granules, are key antimicrobial mobile organelles with a lysosome-like character. Their granular content is presented on the cell surface or secreted into the extracellular environment via mechanisms of regulated degranulation upon neutrophil activation (Nauseef et al., 2014). Protein paucimannosylation was found to be abundant in sputum from inflamed lungs irrespective of the infecting pathogen, disease, condition, gender, age, and antibiotic treatment, which suggests that human paucimannose-rich *N*-glycome signature is a general molecular feature of inflamed host microenvironments undergoing pathogenic attacks (Thaysen-Andersen et al., 2015).

Paucimannosidic *N*-glycoproteins are being increasingly associated with important extracellular functions in infection and inflammation. Interestingly, paucimannosidic structures are presented on exposed glycosylation sites of proteins, unlike their spatially more hidden high-mannose counterparts. Therefore, the paucimannose-rich *N*-glycosylation of human neutrophil elastase (HNE) is likely to mediate cell-cell communication with immune cells, including dendritic cells and macrophages that are known to express mannose recognising C-type lectin receptors (mrCLRs) with affinity to mannosylated ligands (Loke et al., 2017). The accessible nature and the spatiotemporal expression of paucimannosidic glyco-epitopes in the extracellular environment under certain physiological conditions are presumably the features that immune cells, and perhaps other cell types, exploit to communicate within the immune system.

In contrast to lower organisms where paucimannosidic structures are almost uniquely generated through the classical ER-Golgi secretory pathway and constitute a ubiquitous class in the *N*-glycome repertoire, mammals appear to use also an alternative ER-Golgi-granule biosynthetic pathway to achieve paucimannosidic *N*-glycosylation

(Fig.5). Similar to the biosynthesis of PMPs in other kingdoms, growing evidence supports a concerted action of *N*-acetyl-β-hexosaminidase (Hex) isoenzymes and linkage-specific α-mannosidases in the production of paucimannosidic glycoepitopes in humans via *N*-acetylglucosaminyltransferase I (GnT-I)-dependent pathway, and possibly also to a lesser extent, the GnT-I-independent truncation pathway. Two mammalian Hex isoenzymes have been associated with PMP formation: heterodimeric Hex A (αβ) and homodimeric Hex B (ββ) comprising combinations of the structurally and catalytically similar α and β subunits encoded in humans by *HEXA* and *HEXB* respectively (Tjondro et al., 2019).

In analogy to the invertebrate system were hexosaminidases are required for the removal of the GlcNAcβ1–2 residue on the Manα1–3 arm to form paucimannosidic structures (Schiller et al., 2012), it has been speculated that vertebrate paucimannosidic *N*-glycans may as well be generated from more elongated structures by β-hexosaminidase A (Hex A) (Moriguchi et al., 2007; Schachter 2009), an enzyme usually located in lysosomes. Some knowledge of the spatiotemporal protein paucimannosylation is available from human neutrophils where paucimannosidic glycoepitopes were suggested to be formed via the action of Hex A by trimming hybrid/complex type *N*-glycan intermediates using biosynthetic machinery assembled during early myeloid maturation of neutrophil precursors in the bone marrow (Loke et al., 2017; Thaysen-Andersen et al., 2015). Surprisingly, paucimannosidic glyco-epitopes have been only marginally detected in lysosomes of neural cells, indicating that they do not originate from lysosomal degradation but may arise from Hex A action outside of lysosomes. Additionally, paucimannose *N*-glycans could rarely be detected in the ER, Golgi compartments, or Golgi-derived endosomes, which excludes the possibility of uptake from the environment and intracellular recycling (Dahmen et al., 2015).

The majority of glycoproteins trafficking through the *N*-glycosylation machinery at the promyelocytic stage of the neutrophil development are directed to the azurophilic granules by vesicles budding from the *cis* Golgi without reaching the late *N*-glycan maturation stage. Paucimannose generation follows the initial synthesis of *cis* Golgi-localised fucosylated hybrid/complex glycan intermediates. Lysosomal glycoprotein degradation, in turn, is facilitated by a suite of exoglycosidases and proteases, leaving the released and partially degraded *N*-glycans without the reducing-end GlcNAcβ and Fucα1-6 residues (Winchester 2005), contrary to the paucimannose species. Moreover, in the

pathogen-infected sputum, the paucimannosidic glycans were conjugated to proteins with preserved bioactivity and structural integrity. This demonstrated that the truncated *N*-glycoproteins in the azurophilic granules of human neutrophils are of non-lysosomal origin and should not be perceived as degradation products for the purpose of salvaging monosaccharides and amino acids but rather as a cellular mechanism to generate a depository of releasable biomolecules displaying unique glyco-epitopes and activities (Thaysen-Andersen et al., 2015). Moreover, non-lysosomal origin and activity of Hex A were shown earlier in plasma membranes of fibroblasts (Mencarelli et al., 2005).

The abundance of Manα1-6-terminating Man_2Fuc structures suggests that a linkage-specific α-mannosidase-driven trimming reaction is often the terminal processing step, and that truncation beyond Man_2Fuc may be rate-limiting in the trimming cascade of the human paucimannosylation machinery. In turn, although still experimentally unverified, the identity of the α-mannosidase displaying Manα1-3 substrate preference to perform the final processing of human Man_2Fuc glycoproteins is likely to be the lysosomal LysMan, given its appropriate substrate preference, pH optimum and subcellular location (Tjondro et al., 2019).

The prevalent core fucosylation of the observed paucimannosidic *N*-glycans indicates that PMPs are formed mainly via the GnT-I-dependent pathway since GlcNAcβ1-2-capped glycoprotein intermediates are substrates required for the Golgi-resident α1,6-fucosyltransferase (FUT8) responsible for fucosylation of the initial GlcNAc residue of the *N*-glycan core (Chatterjee 2019 et al., 2019).

However, the spatiotemporal restricted biosynthetic route of paucimannosidic structures in cancer cells remains elusive. The short, truncated glycans observed in cancer may result from altered expression of various glycoenzymes involved in the glycosylation machinery. It remains to be explored if the paucimannose-generating β-hexosaminidases are aberrantly expressed or if their coding genes HEXA and/or HEXB have a high prevalence for deleterious polymorphisms in GBM (Becker et al., 2019).

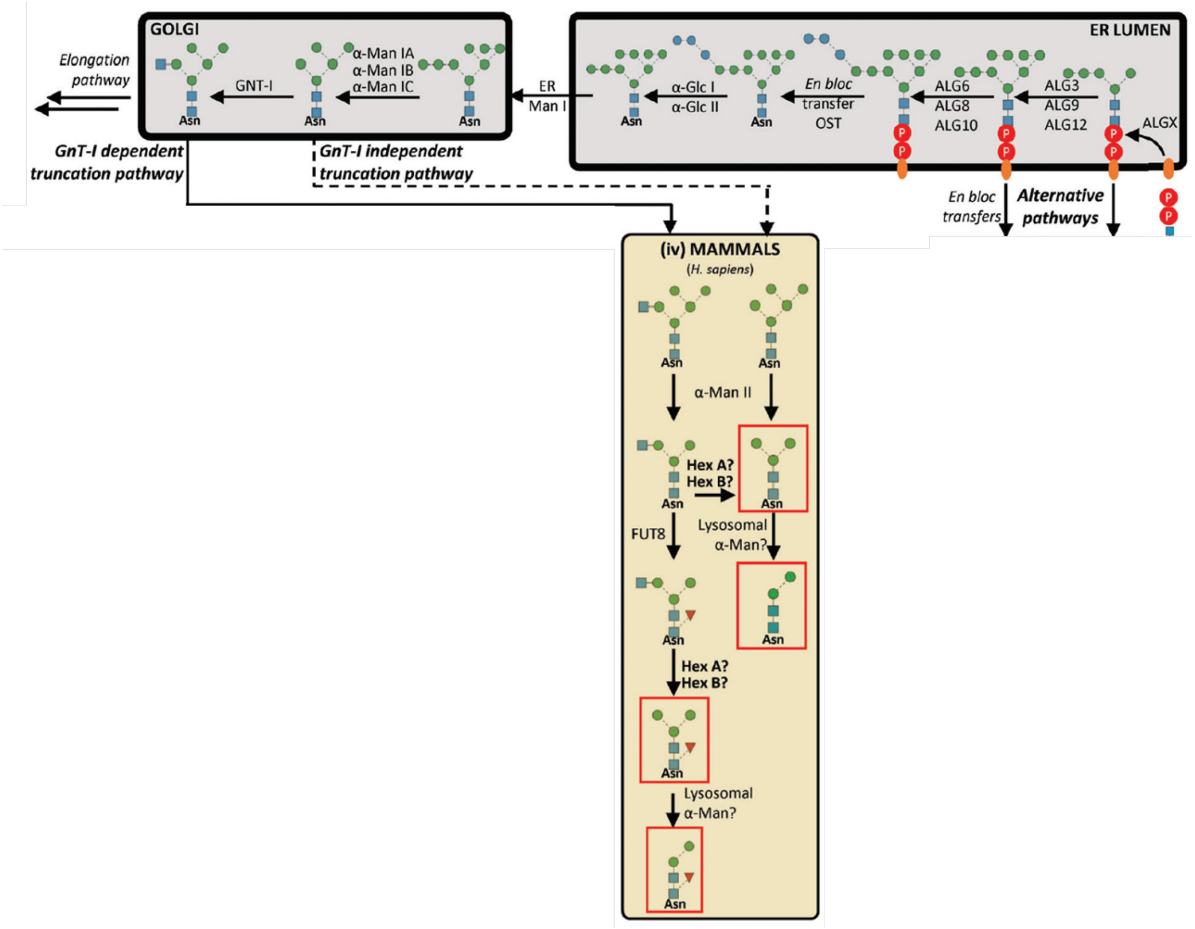


Figure 5. Putative biosynthesis of paucimannose *N*-glycans in mammals. The early stages of the *N*-glycosylation pathway are shared amongst most eukaryotes and involve an *en bloc* transfer of a full-length lipid-linked oligosaccharide precursor ($Glc_3Man_9GlcNAc_2$) to nascent glycoproteins in the endoplasmic reticulum. Following ER-based processing of the oligomannosidic-type *N*-glycoproteins, most eukaryotes are able to produce hybrid-/complex-type glycoproteins in the elongation pathway and/or PMPs via *N*-acetyl-β-hexosaminidase isoenzymes from GlcNAcβ1-2-terminating *N*-glycoprotein intermediates (GnT-I-dependent truncation pathway) and/or via linkage-specific α-mannosidase-based truncation of oligomannosidic-type ($Man_5GlcNAc_2$) *N*-glycoprotein intermediates (GnT-I-independent truncation pathway). ALG, asparagine-linked glycosylation; α-Glc, α-glucosidase; α-Man, α-mannosidase; Asn, asparagine; FUT, fucosyltransferase; GNT, *N*-acetylglucosaminyltransferase; Hex, *N*-acetyl-β-hexosaminidase; OST, oligosaccharyltransferase. Adapted from: Tjondro et al., 2019.

The existence of biologically relevant (non-degraded) human PMPs that may arise from various biosynthetic routes involving the timely expression of Golgi β-hexosaminidases and α-mannosidases has been proposed to explain the generation of paucimannosidic glycoproteins. PMPs have been suggested to occupy multiple environments depending on their tissue origin and function. PMPs mainly reside and function in the peripheral tissue or are rapidly eliminated if secreted into circulation, as their levels are generally low in mammalian blood plasma and other bodily fluids of healthy individuals.

Collectively, these observations prove that paucimannosylation is a recurring feature of mammalian protein *N*-glycosylation and strengthen the association between human PMPs and immune disorders, suggesting a functional involvement of protein paucimannosylation in processes central to infection and inflammation. Furthermore, the occurrence of characteristic paucimannosidic structures on proteins expressed by human cancer and stem cells indicates the involvement of human PMPs in processes related to tumorigenesis and cellular differentiation. The intriguing recent discoveries and the fact that paucimannosidic structures are rarely present in mammals under baseline conditions, make paucimannose *N*-glycans a promising diagnostic marker of pathophysiological states and a potential therapeutic target.

1.2. Biological functions

Considering their ubiquitous presence and abundance in all cellular compartments, extracellular spaces, and body fluids, glycans have numerous biological functions mediated by their own primary structural properties, and/or by modulating functions of proteins and lipids to which they are attached. Biological roles of glycans can be divided into four main categories, namely structural and modulatory functions, intrinsic (intraspecies) recognition, extrinsic (interspecies) recognition, and molecular mimicry of host glycans (Varki 2017).

1.2.1. Structural and modulatory functions

Structural and modulatory roles of glycans include providing physical structure, physical protection and tissue elasticity, water solubility of macromolecules, lubrication, physical expulsion of pathogens, diffusion barriers, glycoprotein folding, protection from proteases, modulation of membrane receptor signalling, membrane organisation, modulation of transmembrane receptor spatial organisation and function, anti-adhesive action, depot functions, nutritional storage, gradient generation, extracellular matrix organisation, protection from immune recognition, effects of glycan branching on glycoprotein function, cell surface glycan:lectin-based lattices, masking or modification of ligands for glycan-binding proteins, tuning a range of functions, molecular functional switching, and epigenetic histone modifications (Varki 2017).

β-Linked homopolymers of glucose, cellulose, or *N*-acetylglucosamine, chitin, are among the most abundant organic molecules, providing strength and rigidity to plant and fungal cell walls and arthropod exoskeletons, since they are very difficult to breakdown by physical, chemical, or enzymatic means (Koch et al., 2015; McFarlane et al., 2014).

The dense layer of mucins that coats many epithelial surfaces such as the inner lining of airways and intestines acts as a barrier and provides protection against the invasion by microorganisms that live within the lumen. Disruption of this layer can have very serious consequences and results in inflammation and carcinogenesis associated with microbial invasion (Bergstrom et al., 2013; Johansson et al., 2013; Lillehoj et al., 2013; Pelaseyed et al., 2014).

Hydrophilic and acidic glycans contribute significantly to the water solubility of the glycosylated proteins in vertebrate internal body fluids such as blood plasma. The exceptionally high concentration of proteins in the blood plasma (~50–70 mg/mL in humans, carrying ~2 mM of bound sialic acids) would probably be impossible without the heavy glycosylation (Varki 2017).

Soluble and membrane-bound mucins on the lining of hollow organs ensures particularly efficient lubrication needed for swallowing food. In turn, in body fluids, such as the synovial fluid in joint cavities and tear fluid in the eyes, it is the hyaluronan that plays a critical lubricating role (Hascall et al., 2009).

Some goblet cells localised in the mouse colonic crypt entrance recognise bacterial products which induce mucin secretion from adjacent goblet cells in the upper crypt. This, in turn, physically expels bacterial intruders that have penetrated the protective inner mucus layer (Birchenough et al., 2016).

Extracellular matrix glycosaminoglycans and/or heavily sialylated glycoproteins can comprise critical diffusion barriers. As an example of this, the heavily sialylated podocalyxin on glomerular podocyte foot processes (Ito et al., 2012) and heparan sulfate glycosaminoglycans within the glomerular basement membrane (van den Born et al., 1993) seem to play important roles in maintaining the integrity of blood plasma filtration by the kidney.

Proteins produced and secreted via the ER-Golgi pathway can be subjected to ER modifications such as *O*-fucosylation (Vasudevan et al., 2015) and *O*-mannosylation (Xu et al., 2015) that facilitates their proper folding in the ER lumen. A major fraction of such ER-synthesised proteins may also be modified by *N*-linked glycosylation at Asn-X-Ser/Thr sequons (Weerapana et al., 2006), and the large, generally hydrophilic carbohydrate chains contribute to correct folding of nascent polypeptides emerging into the lumen of the ER.

Highly glycosylated proteins are protected from protease cleavage probably by steric hindrance or negative charge of the attached glycans. This effect is especially notable for mucins carrying densely packed *O*-glycans (Loomes et al., 1999).

Glycosylation affects the signalling properties of the proteins to which the carbohydrate is attached. As an example, α1-6 core fucosylation of *N*-glycans affects transforming growth factor (TGF) signalling. Dysregulation of TGF-β1 receptor activation leads to abnormal lung development (Wang et al., 2005). Another example may be the

classification of the Fringe molecule as a glycosyltransferase that modifies the important signalling protein Notch and thus modulates Notch–Delta interactions (Bruckner et al., 2000).

Cell surface glycoprotein carbohydrates can modulate membrane domain organisation by their bulk and charge (Wier et al., 1988). Apparently, saccharides on one class of glycoproteins can even modulate the organisation of other classes of glycans on other proteins present on the same cell surface (Cohen et al., 2014).

Bulky glycoproteins in the cell surface glycocalyx is a feature of malignant cells that can indirectly aid cell adhesion and signalling. This facilitates integrin clustering by channelling active integrins into adhesions and alters their state by applying tension to these matrix-bound molecules. This, in turn, promotes focal adhesion assembly and assists integrin-dependent growth factor signalling to support cell growth and survival. Therefore, the bulky glycocalyx is thought to favour metastasis by mechanically enhancing cell surface receptor function (Paszek et al., 2014).

Hyaluronan and polysialic acid are large acidic polymers that can inhibit cell-cell and cell-matrix interactions by both bulk and negative charge. These antiadhesive functions are particularly prominent during phases of development when cell migration is very active (Varki 2017).

Hydrophilic glycans on cell surfaces and extracellular matrices are able to attract and organise water molecules (Espinosa-Marzal et al., 2013). Beyond retaining water and cations, extracellular matrix glycosaminoglycans and polysialic acid can act as depots for growth factors and other biomolecules, which can be stored locally and released when needed, e.g., during injury and wound healing (Varki 2017).

Polymeric glycans like glycogen in animal cells and starch in plant cells serve as long-term storage of glucose as an energy source.

By binding to extracellular matrix glycosaminoglycans, such as heparan sulfate, gradients of growth factors can be generated. This organisation of growth factors by glycosaminoglycans support the morphogen gradients that are crucial during development (Inatani et al., 2003; Schwartz et al., 2014).

In vertebrates, large glycan polymers such as sulfated glycosaminoglycans and hyaluronan, are the main components of the extracellular matrix. These polymers, together with specific proteins, are capable of self-arrangement into larger aggregates to produce structures like basement membranes and cartilage (Aspberg 2012; Iozzo 2005).

The adaptive immune system of vertebrates functions mostly by direct recognition of foreign peptide sequences by the B cell surface Ig receptors and their incorporation into the major histocompatibility receptors for T-cell receptor presentation (Elgert 2009). An attached saccharide can attribute novel specificity to recognition (Galli-Stampino et al., 1997) with larger glycans generally obstructing peptide loading and/or T-cell receptor recognition. This explains a common immune escape strategy of enveloped viruses whose surface glycoproteins tend to be very heavily glycosylated. Sometimes, such protective glycosylation can become so dense that it generates unique clustered epitopes recognised by specific antibodies, such as that seen on the surface of the HIV virion (Garces et al., 2014).

The *N*-linked glycans on cell surface glycoproteins can have varying degrees of branching. The number of *N*-glycans and their degree of branching can collaborate to regulate cell proliferation and differentiation (Lau et al., 2007), as well as thymocyte positive selection (Zhou et al., 2014). Glycan branching is specifically upregulated in T-cell activation and malignant transformation (Lau et al., 2008).

The glycocalyx on the surface of vertebrate cells contains self-organising ordered lattices of glycans and lectins (Bhattacharyya et al., 1989). The formation of galectin-mediated lattices within the glycocalyx alter interactions between cell surface molecules, also affecting their rates of clearance from the cell surface by endocytosis (Dennis et al., 2009; Garner et al., 2008).

Modifications of monosaccharides and/or specific monosaccharides themselves can act as biological masks that prevent the recognition of the underlying glycan by specific glycan-binding proteins. A classic example is the *O*-acetyl modifications of terminal sialic acid which can block the binding of some influenza viruses (Muchmore et al., 1987; Schauer 1985).

The nature, size, number, extent of branching, and degree of sialylation of *N*-glycans may generate numerous glycoforms of a single polypeptide, such as erythropoietin (Yuen et al., 2003), and can modulate the activity of such protein over a range of function, by affecting its interaction with its cognate receptor, and also by altering the rate of clearance from the circulation.

O-GlcNAc modification of nuclear and cytoplasmic proteins has been shown to be a multifunctional molecular switch, which can work with, or in competition against, Ser/Thr phosphorylation, altering the functions of a wide variety of modified proteins

and affecting numerous physiological and pathological processes (Hart 1997; Hart et al., 2011). Another classic example of molecular functional switching is the modulation of IgG effector functions by subtle changes in the structural features of *N*-glycans present in the IgG-Fc region (Chung et al., 2014). Incomplete galactosylation of *N*-glycans has been associated with chronic inflammatory diseases (Rademacher et al., 1994). Furthermore, there are well-defined effects of IgG-Fc glycan core fucosylation on antibody-dependent cellular cytotoxicity (Peipp et al., 2008). Sialylation of a minor fraction of the IgG-Fc *N*-glycans, in turn, seems to convert the IgG molecule into an inflammation inhibitor (Quast et al., 2015).

The addition of *O*-GlcNAc residues to histone proteins surrounding chromosomal DNA is a key component of the histone code that regulates gene expression. *O*-GlcNAcylation targets key transcriptional and epigenetic regulators including RNA polymerase II, histones, and histone deacetylase complexes. *O*-GlcNAc cycling is thought to serve as a homeostatic mechanism linking nutrient availability to higher-order chromatin organisation. Evidence suggests that *O*-GlcNAcylation can also influence X chromosome inactivation and genetic imprinting (Hardivillé et al., 2016; Lewis et al., 2014).

1.2.2. Intrinsic recognition of glycans

Intrinsic (intraspecies) recognition of glycans includes intracellular glycoprotein folding and degradation, intracellular glycoprotein trafficking, triggering of endocytosis and phagocytosis, intercellular signalling, intercellular adhesion, cell-matrix interactions, fertilisation and reproduction, clearance of damaged glycoconjugates and cells, glycans as clearance receptors, danger-associated molecular patterns, self-associated molecular patterns, antigenic epitopes, and xeno-autoantigens (Varki 2017).

Eukaryotic *N*-linked glycans promote the proper folding of newly synthesised polypeptides in the ER. This adaptation allows cells to produce and secrete larger and more complex proteins at higher levels. It also explains why the addition of *N*-linked glycans must occur cotranslationally in the ER before the folding process has begun. When glycosylation is inhibited, the most commonly observed effect is the generation of misfolded, aggregated proteins that fail to reach a functional state. The importance of the

added glycans varies between proteins and depends on the physiological context. Some proteins are completely glycan-dependent, some are partially dependent, whereas many display no dependence at all. Studies indicate that although an *N*-linked glycan does not induce permanent secondary structure of the peptide, it alters the conformational preferences close to the glycosylation site, resulting in more compact conformations. The presence of oligosaccharides has a significant stabilising effect on the folded glycoprotein (Imperiali et al., 1999; Wormald et al., 1999).

The most important indirect effect of glycans on folding involves a unique chaperone system named the calnexin-calreticulin cycle prevalent in the ER of nearly all eukaryotic cells. The membrane-bound calnexin (CNX) and soluble calreticulin (CRT) are homologous ER lectins that transiently bind to all newly synthesised glycoproteins (Fig.6). Calnexin and calreticulin interact with the glycan moieties of substrate glycoproteins after they have been trimmed by glucosidases I and II to the monoglucosylated form $Glc_1Man_{9-6}GlcNAc_2$. CNX and CRT form a complex with ERp57, a thiol oxidoreductase. When the remaining third glucose residue is trimmed by glucosidase II, the complex dissociates. If the glycoprotein is not properly folded at this point, the oligosaccharides are reglucosylated by the ER uridine diphosphate (UDP)-glucose:glycoprotein glucosyltransferase (GT), and the protein reassociates with the CNX and CRT. The cycle is repeated until the protein is either folded or degraded. Substrate recognition by GT is triggered not by a specific signal sequence but by general biophysical properties common for incompletely folded glycoproteins, such as exposed hydrophobic patches or excessive dynamic mobility. By recognising the folding status, GT forces misfolded conformers to remain in the calnexin-calreticulin cycle. The correctly folded glycoproteins are no longer recognised by GT and are allowed to leave the ER to proceed further in the secretory pathway. For many glycoproteins, the interaction with calnexin, calreticulin, and ERp57 slows down the rate of folding but increases its efficiency. The calnexin-calreticulin cycle promotes correct folding, inhibits aggregation of folding intermediates, blocks premature oligomerisation, regulates ER degradation, and provides quality control by preventing incompletely folded glycoproteins to exit the ER and enter the Golgi complex (Hammond et al., 1994; Hebert et al., 1995; Helenius et al., 1997; Parodi et al., 2000; Zapun et al., 1999).

Integral membrane, lumenal, and secretory proteins are synthesised on ribosomes attached to the cytosolic side of the ER. The folding and assembly of these

proteins, as well as a variety of post-translational modifications, occur within the ER. Furthermore, the ER functions as a sorting department in which the fate of newly synthesised proteins is determined. Trimming of the *N*-linked glycans by α1-2 mannosidase I plays a key role in the sorting process leading to selective glycoprotein degradation in the ER. ER-α1-2 mannosidase I cleaves a single α1-2-linked mannose residue from the α1-3 branch of the core oligosaccharide delivering a $Glc_{0-3}Man_8GlcNAc_2$ residue. If a glycoprotein fails to completely fold or assemble to reach its native conformation in the ER, the resulting Man_8 structure serves as a part of the signal triggering the elimination of the glycoprotein by ER-associated degradation (ERAD). However, removal of the mannose is apparently not sufficient as most well-folded proteins are mannose-trimmed before leaving the ER. Considering that mannosidase is relatively slow-acting, it has been suggested that it acts as a molecular clock that regulates the disposal of misfolded glycoproteins and grants protection against premature degradation to the most recently synthesised glycoproteins. ERAD is a highly efficient quality-control system which guarantees that only properly folded proteins and correctly assembled subunits of protein complexes are generated. The existence of an aggressive proteolytic apparatus in the ER causes cells to transport proteins destined for degradation to the cytosol, where they are ubiquitinated. The polyubiquitinated substrates are then recognised and eliminated by the 26S proteasome complex, instead of using a lumenal degradation system that might endanger protease-sensitive folding intermediates. The available data suggest multiple, parallel pathways leading to glycoprotein degradation, with mannosidase I-dependent ERAD system as one of the most distinct sorting criteria (Avezov et al., 2008; Jakob et al., 1998; Plemper et al., 1999; Tokunaga et al., 2000; Wang et al., 2000). Notably, *O*-GlcNAcylation of nucleocytoplasmic proteins can also occur cotranslationally, protecting nascent polypeptide chains from premature degradation by decreasing cotranslational ubiquitinylation (Zhu et al., 2015).

ERGIC-53 (ER-Golgi intermediate compartment 53 kDa protein) and VIP36 (36 kDa vesicular integral membrane protein) are mannose-specific transmembrane proteins with a luminal domain that operate in the early secretory pathway and the Golgi complex. These L-type lectins sharing homologous carbohydrate recognition domains (CRDs) with structural resemblance to leguminous lectins, serve as cargo (correctly folded protein) receptors for transporting certain *N*-linked glycoproteins in the secretory

pathway in animal cells (Hauri et al., 2000; Satoh et al., 2014; Tannous et al., 2015; Yamamoto 2014).

A variety of cell surface receptors recognising terminal glycans can trigger the uptake of molecules (endocytosis), particles (phagocytosis), or even intact cells. The classic example is the asialoglycoprotein receptor of hepatocytes and the mannose receptor of macrophages. The asialoglycoprotein receptor recognises and binds to exposed β-linked galactose residues on desialylated glycoproteins to rapidly clear them away in the liver (Ashwell et al., 1982). In the cis-Golgi compartment, a resident membrane UDP-GlcNAc:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase specifically modifies *N*-glycan moieties in lysosomal hydrolases by adding GlcNAc-phosphate to the C6 hydroxyl of terminal or subterminal mannose residues. Among all the cargo coming from the ER, the phosphotransferase recognises the lysosomal enzymes in a conformation-dependent manner, modifying only the folded forms of the substrate glycoproteins. This recognition requires critically spaced lysine residues at the surface of the enzyme. After the removal of the GlcNAc, the remaining mannose-6-phosphate group is recognised by mannose-6-phosphate receptors (M-6-P receptors) in the trans-Golgi network (TGN), namely the cation-dependent M-6-P receptor (CD-MPR) and insulin-like growth factor II/cation-independent M-6-P receptor (IGF-II/CI-MPR). Their role is to isolate lysosomal enzymes in the TGN by associating with one or more of the oligosaccharides and escort them via clathrin-coated vesicles to endosomes. The acidic medium in the lysosomes promotes dissociation of the Man-6-P–hydrolase complex, whereupon the receptor returns to the TGN for recycling (Hauri et al., 2000; Helenius et al., 2001; Wiederschain 2013; Yamamoto 2014). Such recognition processes may be critical not only for providing antigens to process and present to T cells but also for clearing away damaged glycoconjugates and cells (Varki 2017).

A well-established example of intercellular signalling in vertebrate systems is represented by hyaluronan fragments released during injury that can be detected by Toll-like receptors (TLRs), and therefore triggering host immune responses (Jiang et al., 2011; Taylor et al., 2004).

A family of cell adhesion molecules, selectins, are critical for leukocytes rolling on endothelium prior to their exit from the circulation (Spertini et al., 1991). Selectins recognise a common motif consisting of sialylated fucosylated glycans (Cummings et al., 1992; Kansas et al., 1996; Varki 1994). The selectin-mediated cell adhesion is vital for

cell-cell interactions of leukocytes, platelets, and endothelial cells (McEver 1991; Tedder et al., 1995). Another classic example of Intercellular adhesion is the role of myelin-associated glycoprotein (MAG, Siglec-4) in mediating key interactions between neurons and glia, a process critical for maintaining the stability of the myelin sheath that insulates axons (Sun et al., 2004).

Evidence for critical matrix interactions with cell surface glycans is present in the variety of muscular dystrophies resulting from altered glycosylation of the α-dystroglycan ligand for major matrix proteins such as laminin (Haliloğlu et al., 2004). Another example is the hyaluronan matrices synthesised by stressed cells recruiting inflammatory cells that are first steps in many pathological processes (de La Motte et al., 1993).

Glycan-recognition processes are a critical part of many sperm-egg interactions (Clark et al., 2014), including the migration of sperm to the site of fertilisation (Tecle et al., 2015). Sperm carrying non-species-specific glycan antigens are destroyed by circulating antibodies that can enter the uterine fluid (Ghaderi et al., 2011). In mammals, glycans and glycan-binding proteins are involved in the processes of implantation (Genbacev et al., 2003) and placental functions (Brinkman-Van et al., 2007).

Innate immune cells recognise glycans released from damaged tissue, such as hyaluronan fragments (Jiang et al., 2011) and some matrix proteoglycans (Moreth et al., 2010), as danger-associated molecular patterns (DAMPs) and activate innate immune inflammatory responses.

It has been proposed that glycans could also act as self-associated molecular patterns (SAMPs) (Varki 2011), being recognised by intrinsic inhibitory receptors to maintain the baseline non-activated state of innate immune cells, and to dampen their reactivity following an immune response. A clear example of glycan-based SAMPs is the inhibitory Siglec recognition of cell surface sialoglycans, which may also provide a mechanism for the host to discriminate between non-infectious self and infectious non-self (Chen et al., 2014). Sialoglycan recognition by factor H can blunt immune responses by inhibiting the alternate pathway of complement activation (Blaum et al., 2015).

Intra- and interspecies variations in glycosylation can result in strongly antigenic epitopes, as in the case of ABO blood group polymorphisms (Storry et al., 2009). A significant fraction of circulating immunoglobulins found in humans is directed against foreign glycan antigens. Changes in the levels of anti-glycan antibodies can occur with the

onset of disease, exposure to pathogens, or vaccination (Muthana et al., 2015). Glycan motifs like α-Gal and Neu5Gc represent the major xenoantigens that must be bypassed when attempting xenotransplantation in humans (Lutz et al., 2013).

The non-human sialic acid Neu5Gc can become metabolically incorporated from dietary sources, such as red meat, and appear on the surface of certain human cells as if it was synthesised by the individual. These "xeno-autoantigens" are recognised by pre-existing circulating "xeno-autoantibodies". The resulting "xenosialitis" is suggested as one mechanism for the epidemiological association between red meat consumption and the exacerbation of some common disease states, such as atherosclerosis and carcinoma (Alisson-Silva et al., 2016; Samraj et al., 2014).

The limiting membrane of lysosomes contains a collection of highly glycosylated membrane proteins, the most abundant of which are the homologs LAMP-1 and LAMP-2. Each of them contains up to twenty *N*-linked glycans, that occur in sufficient concentrations to form a continuous carbohydrate coat over the luminal surface of the lysosomal membrane. These glycans serve as protection against degradation by lysosomal proteases (Helenius et al., 2001).

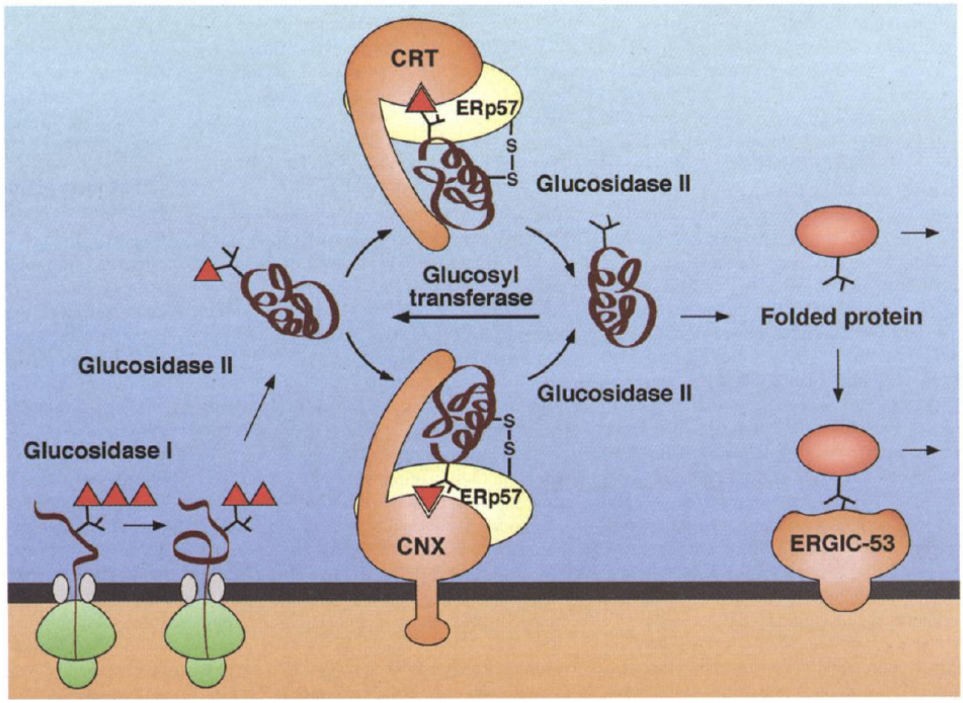


Figure 6. The calnexin-calreticulin cycle. From: Helenius et al., 2001.

1.2.3. Molecular mimicry of host glycans

Microorganisms have evolved to achieve molecular mimicry of host glycans using numerous mechanisms among which the most prominent are the convergent evolution of host-like glycans and appropriation of host glycans.

Most cases of glycan molecular mimicry by pathogens seem to involve convergent evolution of pre-existing pathogen biosynthetic pathways or *de novo* generation of functional genes. Group B Streptococcus polysaccharides display identity to specifics of host glycan structure, such as the Neu5Acα2-3Galβ1-4GlcNAcβ1-, which perfectly matches the structure of *N*-glycan antennae on many human glycoproteins (Michon et al., 1988). Another example is the *Campylobacter* species that carry out near-perfect mimicking of complex brain ganglioside glycans allowing them to imitate endogenous SAMPs and down-regulate innate immune responses by engaging the inhibitory Siglecs (Carlin et al., 2009).

Following the example of sialic acids, the mechanisms of appropriation of host glycans range from simple acquisition of host sialoglycans (Khatua et al., 2012) to the direct transfer of host sialic acids by trans-sialidases (Freire-de-Lima et al., 2015), to the highly efficient uptake of the small amounts of environmental free sialic acids (Johnston et al., 2007), or even the direct utilisation of trace amounts of CMP sialic acid present in host body fluids (Parsons et al., 1988). In addition to acting as SAMPs recognised by Siglecs or limiting complement activation via factor H recruitment, such terminal sialic acids also serve to mask antibody recognition of underlying structures.

2. Antibodies

Vertebrates possess two types of adaptive immunity based on the components the immune system uses to mediate immunity, namely humoral immunity and cell-mediated immunity (Fig.7). Humoral immunity is mediated by antigen-specific blood proteins named antibodies. Antibodies are secreted only by plasma cells, the daughter cells of bone marrow-derived B lymphocytes. Humoral immunity protects against circulating extracellular antigens such as bacteria, microbial exotoxins, and viruses in their extracellular phase since antibodies interact with circulating antigens but are unable to penetrate living cells. Cell-mediated immunity is mediated by antigen-specific cells called thymus-derived lymphocytes. There are at least two main subpopulations of T cells, namely T helper cells (T_H) and T cytotoxic cells (T_C). Cell-mediated immunity protects against intracellular intruders, such as viruses, participates in the defence mechanism against cancer, and plays an important role in transplant rejection. Activated, antigen-specific B and T lymphocytes are referred to as effector cells. B cell-derived plasma cells act as effector cells by producing antibodies, while T_H cells release communication molecules (cytokines) and T_C cells kill target cells. Both humoral and cellular immune responses are evoked and cooperate during an antigenic attack, although one of them predominates depending on the challenge.

Antibodies belong to a family of globular proteins named immunoglobulins (Ig). They are recognition proteins found in serum and other body fluids of vertebrates that bind specifically to the antigens that induce their formation. In the late 1950s and early 1960s, Rodney Porter of Great Britain and Gerald Edelman of the United States elucidated the chemical structure of antibodies. Edelman's and Porter's approaches to the characterisation of the immunoglobulin G (IgG) molecule were different. Edelman defined IgG molecules using chemical solvents, whereas Porter employed protein-degrading enzymes. Our understanding of antibody structure draws from both of the methods and results. Porter and Edelman shared the 1972 Nobel Prize in Medicine for their structural studies of antibodies.

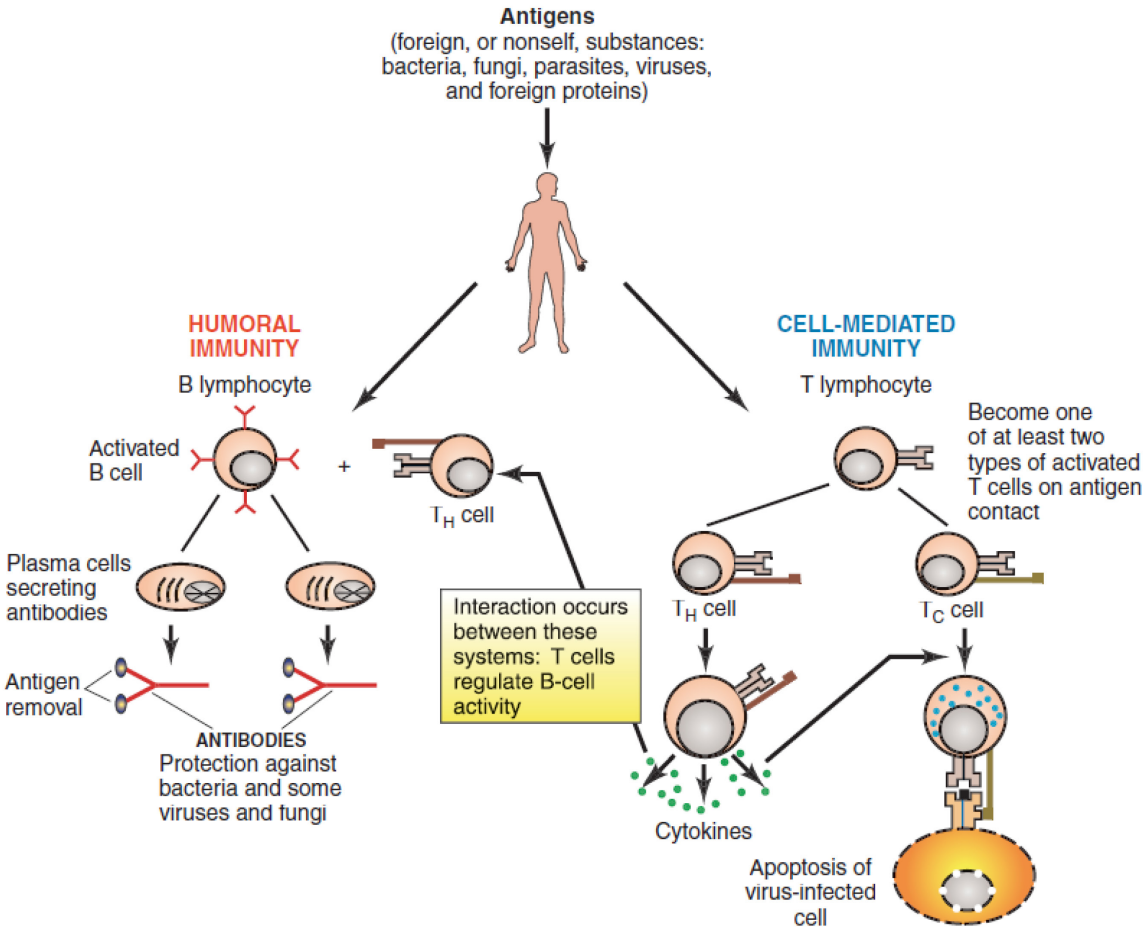


Figure 7. The two types of adaptive immunity: humoral and cell-mediated. From: Elgert 2009.

IgG is used as the prototype antibody to explain the basic structure of all antibodies (Fig.8). The molecular weight of an IgG molecule is ~150 kDa and consists of two Fab fragments (antigen-binding) and one Fc fragment (crystallisable). In an intact molecule, the Fab region binds the antigenic determinant (epitope), whereas the Fc region determines the biological properties of the immunoglobulin, such as serum half-life, interaction with cellular Fc receptors, and the ability to activate complement. The larger antibody subunit (50 kD) is called heavy chain (HC) and the smaller subunit (23 kD) is referred to as the light chain (LC). Each chain consists of a large region that is constant in different types of antibodies, even from unrelated species, and a similar-sized region that is highly variable. The constant region (C) in the carboxyl-terminal end of the chain consists of an amino acid sequence that is almost identical among the majority of immunoglobulins. The opposite amino-terminal end of the chain shows great variability in amino acid sequences and is called the variable (V) region. This demonstrates that antibodies binding different ligands must have different V light (VL) and V heavy (VH) regions. Because the amino acid sequence dictates the three-dimensional structure of a

protein, the unique sequence of amino acid residues for each V region leads to the enormous structural diversity, which accounts for the antibody specificity. The variability of the aminoacid sequences in the variable regions is not random but precisely organised. It is localised within certain sections of the variable region of a chain, with the sections having substantial sequence variation from immunoglobulin to immunoglobulin. The greatest variability in the light chain is around residues 28-34, 50-56, 89-97, and around 31-35, 50-65, 95-102 for the heavy chain. These areas are called hypervariable regions or complementarity-determining regions (CDRs), and moving from the amino-terminal end they are referred to as CDR1, CDR2, and CDR3. Each CDR is about 10 amino acid residues long. Intervening sequences between the CDRs have restricted variability and show little difference in amino acid sequence between chains. These invariant segments make up the framework residues which equal approx. 85% of the variable region. The variable region folds so that the CDRs are exposed on the surface of the immunoglobulin chain. When the light and heavy chains are joined, the CDRs form a cleft that serves as the antigen-binding site of the antibody. Since the amino acid sequences of the CDRs determine the shape and ionic properties of the antigen-binding site, the CDRs are the ones that define the specificity of an antibody (Elgert 2009).

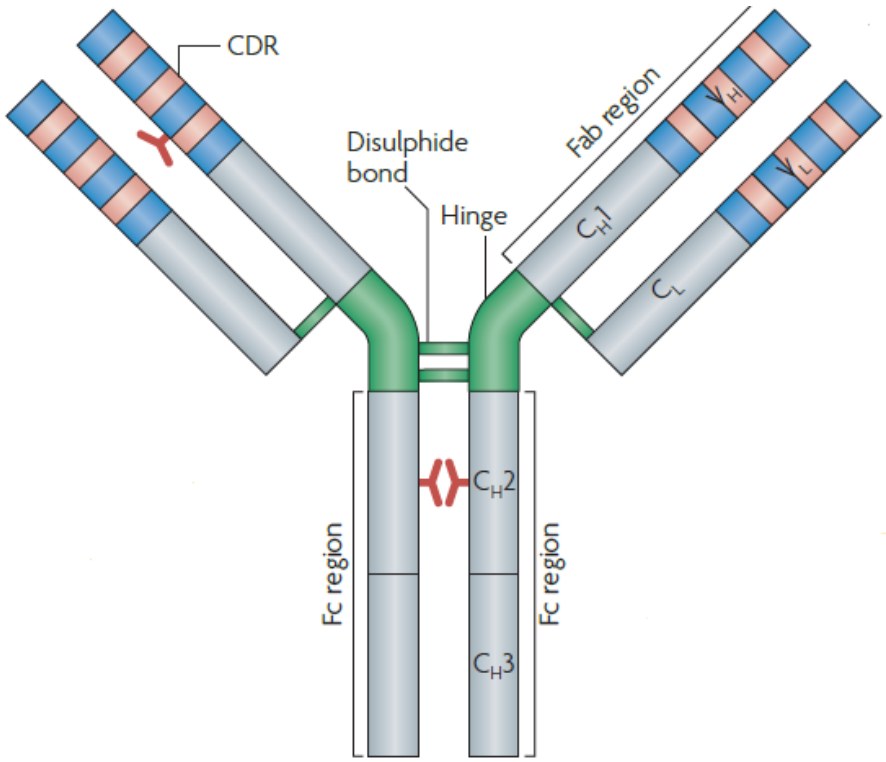


Figure 8. Immunoglobulin G1 structure. Adapted from: Beck et al., 2010.

The numerous possible combinations of available monosaccharides, variable linkages, branching, and lengths, together with a high degree of flexibility of many glycans, require antibodies to use a variety of strategies for carbohydrate recognition. The majority of the observed immunoglobulin diversity stems from the formation of nascent B lymphocytes, where the genes encoding the VH and VL domains are constructed from a limited repertoire of inherited gene segments. These consist of *V* (variable), *D* (diversity) and *J* (joining) gene segments for the heavy chain, and *VJ* segments for the light chain that are located on a separate chromosome and can be either κ or λ type (Fig.9). The recombination events can be quite variable, and substitutions are often incorporated between the gene segments that result in either productive or non-productive immunoglobulins. *V(D)J* recombined genes encoding VH and VL domains are further paired with constant gene segments. Following translation, the antibody polypeptides are modified at glycosylation sites, particularly in the constant regions where these carbohydrate moieties are involved in modulation of effector functions (Haji-Ghassemi et al., 2015).

Edelman showed that each antibody chain has a tandem series of repeating homology units of approx.110 amino acid residues in length called immunoglobulin domains, which fold independently into a compact globular structure. The light chain of IgG has two domains, VL and CL. The heavy chain of IgG has four domains, one VH and three in the CH region; CH1, CH2, and CH3. The heavy-chain V unit shows similarity to the V part of the light chain, while the three C-region units show strong homology to each other and the C region of the light chain. When heavy and light chains are brought together to form an IgG molecule, extensive noncovalent interactions occur between VL and VH, CL and CH1, CH2 and CH2, and CH3 and CH3 domains. Antibodies demonstrate segmental flexibility, which means that the two Fab portions can move relative to one another upon antigen binding. The angle varies from 60 to 180 degrees. This flexible region where the arms meet the stem is called the hinge region and is located between the CH1 and CH2 domains. The proline-rich hinge region is responsible for the rotational flexibility of the Fab domains.

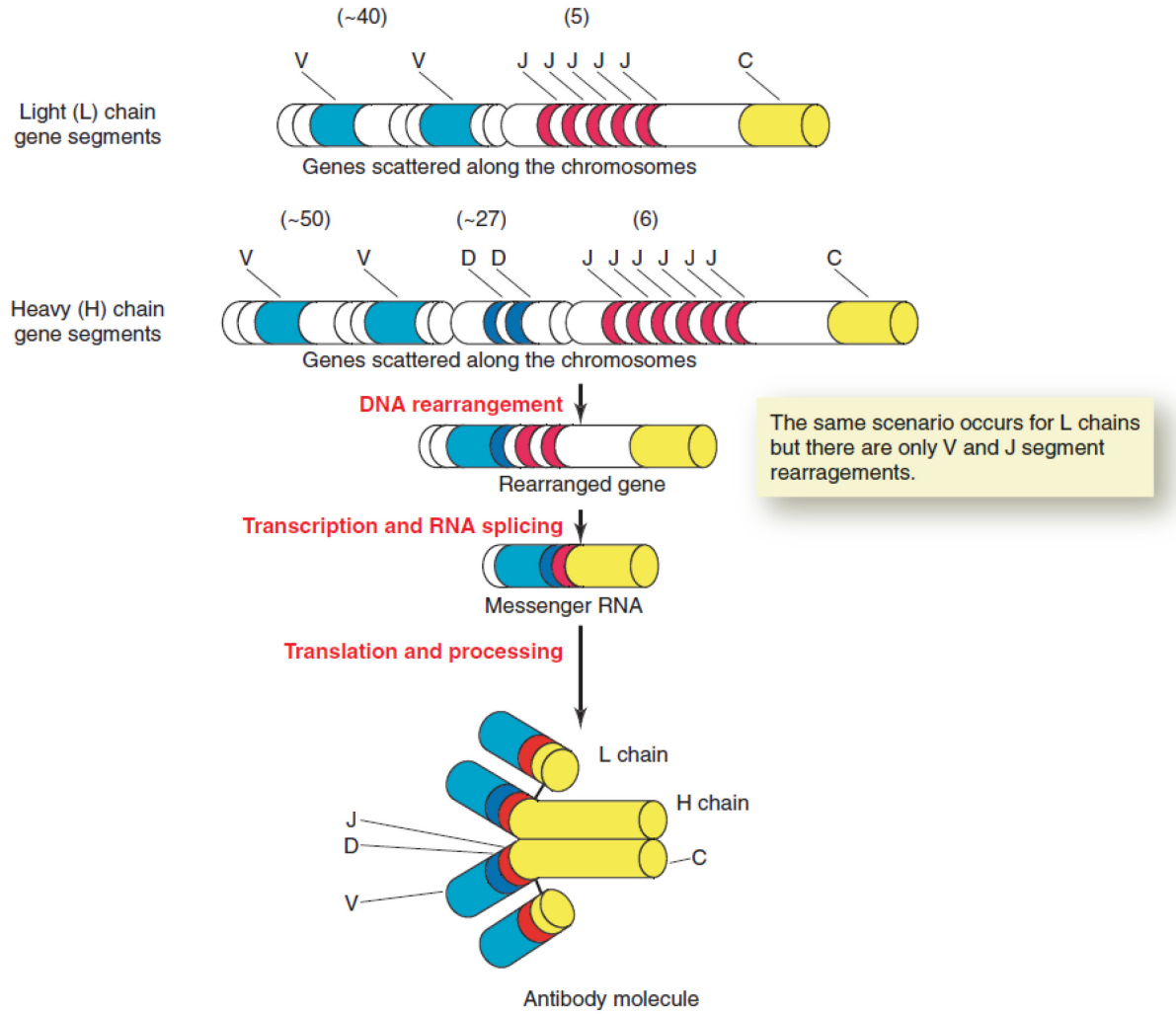


Figure 9. The origin of antibody diversity. From: Elgert 2009.

Humans express five antibody classes, namely IgG, IgA, IgM, IgD, and IgE (Fig.10), which differ in their physicochemical (charge, size, and solubility) and serologic properties (in vitro reactions with antigens). Human IgG has four subclasses, named IgG1, IgG2, IgG3, and IgG4, numbered according to their serum concentrations. Two subclasses have also been identified for human IgA: IgA1 and IgA2

IgG (150 kDa), primarily induced by protein antigens, constitutes about 80% (12.5 mg/mL) of serum antibodies. The four polypeptide chains, two heavy and two light, are covalently held together by disulphide bonds. The four IgG subclasses have 90-95% identity with each other in the C-region domains. The main distinguishing characteristic among the four IgG subclasses is the pattern of interchain linkages in the hinge region. The heavy chain is made up of four domains, one in the variable region and three in the constant region. On the CH2 domain, at position Asn-297, of all IgG heavy chains, a carbohydrate that controls the quaternary structure of this domain is attached. IgG is the

main antibody found in the blood. Nevertheless, it is able to penetrate tissue spaces and coat antigens (opsonisation), which accelerates antigen phagocytic uptake by macrophages and neutrophils and promotes killing by immune cells. IgG is the only antibody class that can cross the placental barrier since only the CH1 and CH3 domains of the IgG heavy chain can bind to placental cells. The human foetus and newborns have limited immune responses. Mechanisms of acquired immunity are not at full strength until some time after birth. The maternal IgG consequently provides some degree of immune protection to the developing foetus. These molecules also are secreted into the mother's milk and, once they have been ingested by an infant, can be transported into the blood, where they confer immunity.

IgA (160 kDa subunit) constitutes 13% (2.1 mg/mL) of human serum antibodies but is the predominant class found in extravascular secretions, where it acts as the first line of defence against pathogens invading the entrances of the body. The IgA present in secretions, such as tears, saliva, nasal secretions, bronchial and digestive tract mucus, and mammary gland secretions, is referred to as secretory IgA. Maternal IgA present in breast milk neutralises pathogens in the infant's gut. IgA1 is the most widespread form in serum, but IgA2 is slightly more prevalent in secretions. The heavy chain is made up of one variable domain and three constant domains. IgA occurs as a dimer joint by a joining chain, a 15 kDa polypeptide consisting of 129 amino acid residues. It is synthesised by plasma cells and attached to IgA, or IgM, either before or at the time of secretion. The J chain attaches to the carboxyl-terminal penultimate cysteine via disulphide bridges to stabilise the multimeric structure of the heavy chain. Dimeric IgA binds to the blood side of the epithelial cells through Fc receptors. Bound IgA is internalised and moves through the cytoplasm of the epithelial cells.

IgD (175 kDa) constitutes less than 1% (40 mg/mL) of human serum antibodies. The constant region of the heavy chain is divided into three domains and consists of 383 amino acid residues. The hinge region contains 64 amino acid residues and is longer than in any other antibody class. IgD is an antibody whose function remains unknown, even though it is one of the main receptors on antigen-naïve mature B cells. As B cells develop, IgD is replaced by other immunoglobulins. IgD is thought to be a regulator of immune responses through its role in antigen internalisation and antigen-triggered B cell growth.

IgE (190 kDa) makes up less than 0.003% (0.4 mg/mL) of human serum antibodies. The heavy chain contains four C-region domains. IgE binds through its Fc part to mast cells or basophils. On subsequent exposure to the same antigen, mast cells and basophils bind the antigen with membrane-bound IgE and trigger allergic reactions. IgE protects against parasites by releasing mediators that attract eosinophils (Elgert 2009).

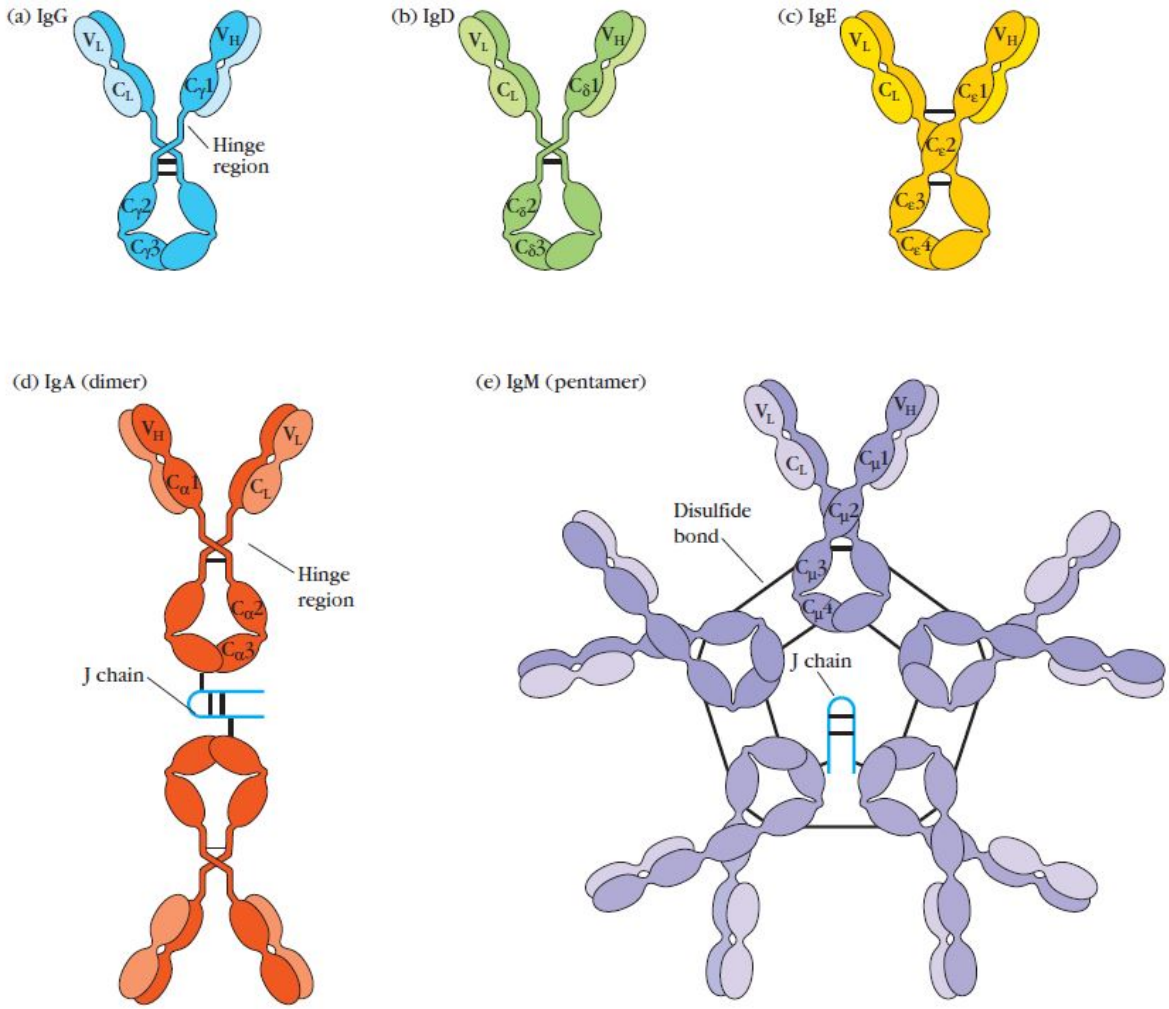


Figure 10. Immunoglobulin classes. From: https://microbenotes.com/antibody-introduction-structure-and-classes/

2.1. Immunoglobulin M

IgM (190 kDa subunit), primarily induced by polysaccharide antigens, is the largest antibody. Secreted IgM is a 950 kDa pentamer that makes up approx. 8% (1.25 mg/mL) of human serum antibodies. The five monomeric molecules are arranged radially, with the Fab regions pointing outward and the Fc portions pointing to the centre of the circle (Fig.11). The IgM heavy chain consists of 576 amino acid residues, with 452 making up the C region. The heavy chain has four C-region domains with interchain disulphide bridges occurring in the CH2, CH3, and CH4 domains at Cys-337, Cys-414, and Cys-575, respectively (Fig.12). Light chains are attached to Cys-136 in the CH1 domain. The CH2 domain of the heavy chain is equivalent to the hinge regions as only IgG, IgA, and IgD molecules have hinge regions. The lack of a hinge region confers a more rigid structure on IgM compared to other immunoglobulin classes. The heavy chains of IgM and IgA possess additional ~18 amino acid residues on the carboxyl-terminal end of the CH4 domain. These areas, called tail pieces, permit immunoglobulin interactions and multimeric molecules formation. The polymerisation of IgM requires Cys-575 located on the tail piece. In contrast to the Ig domains, the tail piece has no defined secondary or tertiary structure. Multimeric IgM also bears a polypeptide joining chain, which is disulphide-linked to the tail pieces and stabilises the multimeric structure. The J chain is attached via Cys-14 and Cys-68 to the penultimate amino acid of two heavy chains in pentameric IgM (Cys-575). In the absence of a J chain during polymerisation, cells secrete solely hexameric IgM (Fig.12). The J chain also binds to the poly-Ig receptor, which mediates epithelial transcytosis (Akhouri et al., 2016; Arnold et al., 2005; Casali 1998; Hiramoto et al., 2018; Li et al., 2020; Müller et al., 2013).

There are five *N*-linked glycosylation sites on the heavy chain. Asn-171, Asn-332, and Asn-395 are occupied by complex glycans, whereas Asn-402 and Asn-563 are occupied by oligomannose glycans. The J chain contains a single *N*-linked glycosylation site at Asn-48. The glycans in these sites terminate predominantly in sialic acid and galactose, but a small population terminates in GlcNAc or mannose. The light chains contain no conserved *N*-linked glycosylation sites. The carbohydrates enhance the solubility of IgM in aqueous solutions and stabilise their tertiary structure by preventing

their aggregation. Glycosylation is crucial for IgM secretion and B cell surface presentation (Arnold et al., 2005; Li et al., 2020).

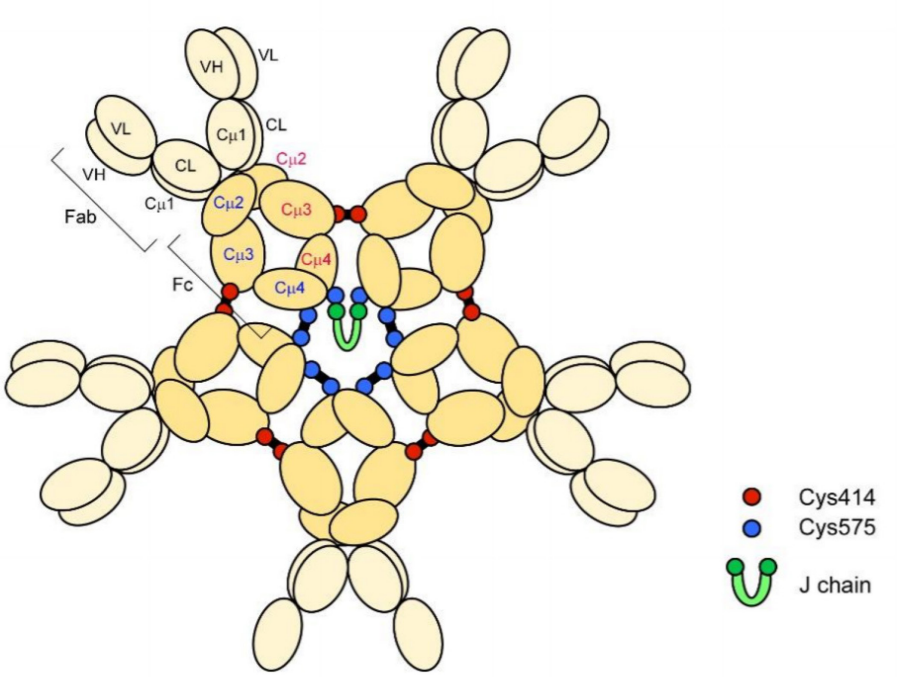


Figure 11. Schematic representation of the conventional pentameric IgM model, a star-shaped symmetric pentagon. The disulphide bonds are depicted in black connecting Cys414 on the CH3 domains and Cys575 on the tail pieces. Adapted from: Hiramoto et al., 2018.

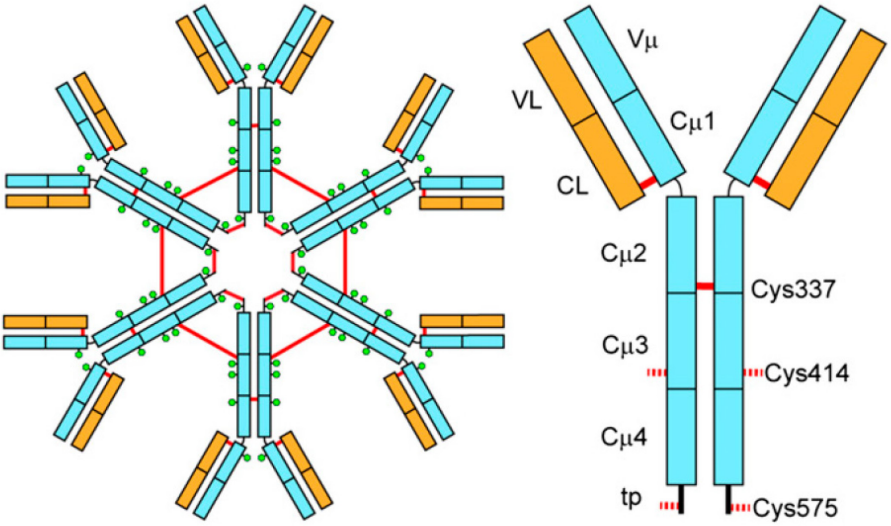


Figure 12. Schematic representation of hexameric (left) and monomeric (right) IgM structures. Heavy chains are depicted in blue and light chains in orange. Green circles indicate glycosylation sites. Red lines represent intersubunit disulphide bridges. The cysteine residues form interdomain disulphide bridges between the CH2 domains (C337) and covalently link the IgM hexamer in CH3 (Cys414) and the C-terminal tail piece (tp) (C575). From: Müller et al., 2013.

IgM is the first line of defence during an immune response to infection (from bacteria, virus, fungi, and parasites) and altered/malignant self-components (e.g. cancer cells) by mediating agglutination reactions, which highlights its role in maintaining immune homeostasis (Fig.13). IgM is the first antibody to be formed by a developing foetus. Early IgM antibodies are secreted before B cells have undergone somatic hypermutations and therefore tend to be of low affinity. To compensate for the reduced binding efficiency of the monomers, IgM forms multimers whose multiple antigen-binding sites confer high overall avidity. Because of its many antigen-binding sites, IgM can quickly clump antigens and directly neutralise them or induce the complement opsonisation of pathogens or apoptotic host cells leading to clearance by phagocytosis without causing an overexuberant inflammatory response. However, hexameric IgM is twenty times more efficient at activating the complement cascade compared with the pentameric form. IgM complement activation is attributed to complement protein C1q binding. The charged residues Asp-417, Glu-418, and His-420 in the CH3 domain of IgM have been proposed as the binding site for C1q. Deficiency in either IgM or the early complement components C1, C4, or C2 causes failure in silent clearance and is strongly linked to autoimmune diseases, such as systemic lupus erythematosus (Boes 2000; Casali 1998; Kaveri et al., 2012; Lintner et al., 2016).

The initial result in nascent B lymphocytes is IgM glycoprotein of a specific sequence. Each B lymphocyte display multiple copies of its particular membrane-bound antibody and IgM in its monomeric form acts as one of the main receptors on the surface of mature B cells, along with IgD. The membrane form of IgM has a different carboxyl-terminal end and is made up of 41 additional amino acid residues, 25 of which form a highly hydrophobic transmembrane segment followed by hydrophilic amino acid residues (Casali 1998; Elgert 2009).

Electron microscopy studies have revealed that considerable conformational changes occur in the IgM pentamer upon antigen binding. In solution, IgM complexes may adopt large stellate structures but upon binding to surface-exposed antigens take the form of staple-like arrangement (Sharp et al., 2019). The widely accepted conventional structural model of pentameric IgM is primarily based on negative-stain electron microscopy (EM) images (Feinstein et al., 1969; Davis et al., 1988), which suggested a star-shaped, symmetric pentagonal structure. More recently, in 2009, Czajkowsky et al. have proposed an improved three-dimensional model based on analysis from cryo–

atomic force microscopy demonstrating that the IgM pentamer is a non-planar, mushroom-shaped molecule with a flexural bias.

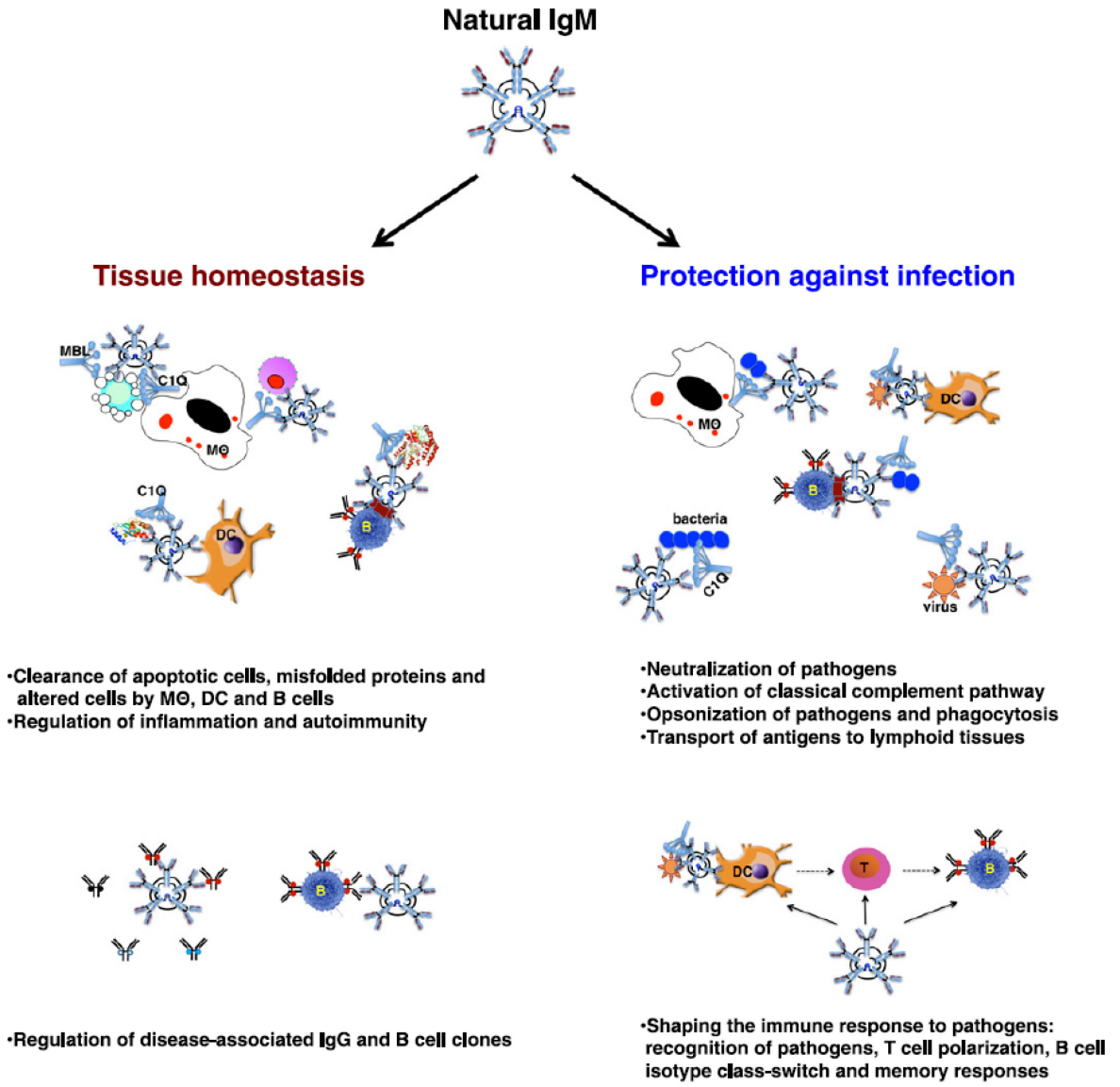


Figure 13. IgM regulates the tissue homeostasis and acts as the first line of defence against invading pathogens by shaping the subsequent immune response. The role of IgM in tissue homeostasis is the activation of complement pathways and prevention of inflammation, autoimmunity, and malignancy. These functions include recognition of apoptotic cells and promoting their clearance by phagocytic cells such as dendritic cells (DCs) and macrophages (MΘ), clearance of altered or malignant cells or misfolded proteins, and regulation of the disease-associated IgG and B cell clones. IgM can confer protection against invading pathogens through their direct neutralisation, activation of classical complement pathways, opsonisation of pathogens and phagocytosis by DCs and MΘ, and transportation of antigens to lymphoid organs for initiating the immune responses by innate immune cells. In addition, nIgM can also regulate the immune response to pathogens by influencing the T cell polarisation and B cell class-switch. From: Kaveri et al., 2012.

2.2. Mannitou antibody

Mannitou is a monoclonal IgM of murine origin, previously termed Laz6-189, that was raised against a specific 130 kDa glycoprotein from leech central nervous system (CNS) (Bajt et al., 1990; Flaster et al., 1983; Zipser et al., 1981). Mannitou antibody has been characterised as being able to specifically recognise the trimannosidic paucimannose glycoepitopes (Zipser et al., 2012).

Using indirect immunohistochemistry anti-paucimannosidic Mannitou antibody was employed to inspect the paucimannosidic *N*-glycosylation in mouse and human tissues. Compared to controls, Mannitou antibody stained all tumour tissues available for analysis, namely human metastatic brain adenocarcinomas, pancreatic adenocarcinoma, lymphomas, and carcinomas of the breast, colon, kidney, liver, ovary, and prostate, as well as human adult pancreatic stem cells and the exocrine ducts of the inflamed mouse pancreas. The cytological localisation of paucimannose was studied in colon carcinoma where Mannitou stained both, intracellular and extracellular granular structures. Intracellularly, paucimannose was associated with vesicles congregated at the apical surface of epithelial cells that represent exocytotic vesicles with similar structures found in the extracellular space, specifically within the colonic ducts. The human pancreatic duct cancer cell line Panc-1 also exhibited strong staining for paucimannose both intracellularly and on the cell surface (Zipser et al., 2012).

The paucimannose-specific Mannitou antibody can modulate cell functions as it was demonstrated to inhibit the proliferation of NPCs *in vivo*. Detection with specific secondary antibodies revealed that Mannitou was incorporated into live human glioblastoma A172 cells and its staining was comparable with that of paucimannose structures. The inhibitory effect of Mannitou might, therefore, be attributed to downregulation or blocking of paucimannosidic glycoepitopes at the cell surface. However, an intracellular action of Mannitou after its internalisation cannot be excluded. Nevertheless, cell viability was not affected by the Mannitou application *in vitro* (Dahmen et al., 2015).

In addition to altered cell proliferation, changes in cell migration and invasion are hallmarks of tumour aggressiveness. Mannitou antibody significantly reduces cell proliferation, cell migration, and invasion. Cell migration and invasion are highly dynamic

processes that rely on an active actin reorganisation. Consequently, in cells treated with Mannitou antibody actin cytoskeletal rearrangements take place. Another hallmark of cancer and a regulatory module of cell migration and invasion into surrounding tissues is cell adhesion. Cell adhesion is significantly reduced in the presence of Mannitou antibody compared to the corresponding controls. Despite all these findings, the exact mode of action of Mannitou antibody remains unknown. It can be speculated that its introduction may simply block paucimannosidic glycoepitopes thereby preventing interaction with still unidentified cis- or trans-ligands, or it may activate signal transduction pathways within the cells downstream of the paucimannose carrying proteins (Becker et al., 2019).

OBJECTIVES

The aim of this study was to produce Mannitou IgM using hybridoma technology to obtain a full-length monoclonal antibody, as well as to generate Mannitou antigen-binding fragments by means of cloning and expression in mammalian cells. These different length antibody structures were required to perform various recognition studies with Mannitou antibody and paucimannose *N*-glycans in order to extract detailed information about the binding interaction. Microarrays were employed to screen among 135 potential oligosaccharide ligands. Surface plasmon resonance enabled kinetic analysis and established the binding affinity of Manniotu antibody towards paucimannosidic *N*-glycans. The thermodynamic parameters of these intermolecular interactions determined by isothermal titration calorimetry provided a detailed insight into the binding mechanism. The mapping of the ligand binding epitope was made possible by saturation transfer difference nuclear magnetic resonance spectroscopy. The next goal was to characterise the structure of Mannitou IgM applying diverse approaches, including homology modelling studies, cryo-electron microscopy, and crystallisation.

The last objective was to investigate the recognition of *Neisseria meningitides* serogroup X capsular oligosaccharides by anti-MenX mAb and Fab. The performed surface plasmon resonance studies contributed to determining the minimal oligosaccharide structure required to mimic the antigenic activity of the native MenX polysaccharide.

Chapter 1.

Mannitou antibody production

I. Generating Mannitou mAb using hybridoma technology

INTRODUCTION

Upon exposure to an antigen, stimulated B-cells give rise to plasma cells which produce highly specific and homogeneous species of immunoglobulins called monoclonal antibodies (mAbs). Back in 1975 George Köhler and Cesar Milstein, working at the Laboratory of Molecular Biology in Cambridge, UK, found a way of mimicking this process to generate large amounts of monoclonal antibodies of predefined specificity. They fused mouse myeloma cells with spleen cells from a mouse immunised with sheep red blood cells and showed that some of the resulting hybrids secreted homogeneous antibodies against the immunising antigen. These hybridoma cells grew continuously in culture, thus "immortalising" a particular antibody, and formed tumours when injected into syngeneic mice. A high titer of the homogeneous antibody was present in the serum and ascites fluid of the tumour-bearing mice (Köhler et al., 1975). The production of monoclonal antibodies by hybridoma technology was a significant milestone in immunology and biomedicine allowing the generation of therapeutic antibodies. The contribution made by Köhler and Milstein was formally recognised in 1984 when they were presented with the Nobel Prize in physiology or medicine.

Hybridoma cells development (Fig.14) begins with retrieving spleen cells from mice that have been immunised with a specific antigen when reaching a sufficient antibody titer in serum. The antibody-producing splenocytes are then fused with cells derived from an immortal tumour of lymphocytes (myeloma) that cannot produce immunoglobulins of their own, so they do not interfere with the production of antibodies by B cells. Mouse spleen cells are mixed with myeloma cells in a ratio of roughly 10:1 in the presence of polyethylene glycol (PEG) to change membrane permeability and allow cell fusion. Since cell fusion is a random process, the cell culture contains a mixture of myeloma–spleen, myeloma–myeloma, and spleen–spleen cell fusions, as well as single myeloma and single spleen cells (Fig.15). The selection of only myeloma–spleen cell fusions is accomplished by culturing the cell mix in hypoxanthine-aminopterin-thymidine (HAT) medium. Aminopterin is a folic acid analogue that blocks the *de novo* biosynthesis of purines and pyrimidines that are vital for DNA synthesis. Myeloma cells used in hybridoma production lack the hypoxanthine-guanine phosphoribosyltransferase (HGPRT) and therefore cannot incorporate exogenous

hypoxanthine to synthesise purines by the salvage pathway. In HAT medium, single myeloma cells and myeloma–myeloma cell fusions die during the first week. Although single spleen cells and splenocyte–splenocyte fusions express HGPRT and therefore are not selected against by HAT medium, these cells naturally die in culture within two weeks because of their lack of immortal growth potential. Only hybridomas resulting from the fusion of HGPRT+ splenocytes with myeloma cells can grow in HAT medium, as they inherit both, the splenocytes ability to synthesise the enzyme, and the immortal growth properties of myeloma cells. It is important to note that the myeloma–spleen cell fusions may emerge as specific antibody producers, nonspecific antibody producers, or nonproducers. The development of appropriate antibody screening assays is thus mandatory for the efficient identification and selection of the specific hybridoma subpopulation in the fusion plate. Culture supernatants from the fusion plates are initially screened for positive hybridoma clones by a number of different immunoassays. While immunofluorescence flow cytometry is often applied to particulate antigens such as whole cells, an enzyme-linked immunosorbent assay (ELISA) is used for soluble antigens such as proteins or polypeptides, and immunohistochemistry (IHC) is developed for tissue antigens. The antibody-secreting hybridoma cells are commonly identified with ELISA. The antigen that was used to immunise the mice with is coated onto the wells and supernatants from each of the clones are poured into the plates. After incubation, an enzyme-coupled anti-mouse antibody is added to the wells. After further incubation, the wells are washed to remove unbound antibodies, and a substrate for the enzyme is added. If the wells contain antibodies bound to the test antigen, the substrate is hydrolysed and a brightly coloured reaction product develops. The degree of colour development indicates the quantity of antibody bound by the antigen. Hybridoma cells from fusion plates further require subcloning to achieve a truly monoclonal population that produces a monospecific antibody. Under the initial conditions in the fusion plates, a well most probably contains no hybridoma clone or more than one hybridoma clone, with or without the ability to produce antibodies. While some hybridoma cells may be genetically unstable at an early stage, the stable clones must be identified and selected as soon as possible. Hybridoma cloning is a time-consuming step in the generation of mAbs. However, it can be accelerated by limiting dilution and microscopic selection of single-cell colonies (Elgert 2009; National Research Council 1999; Price 1985; Zhang 2012).

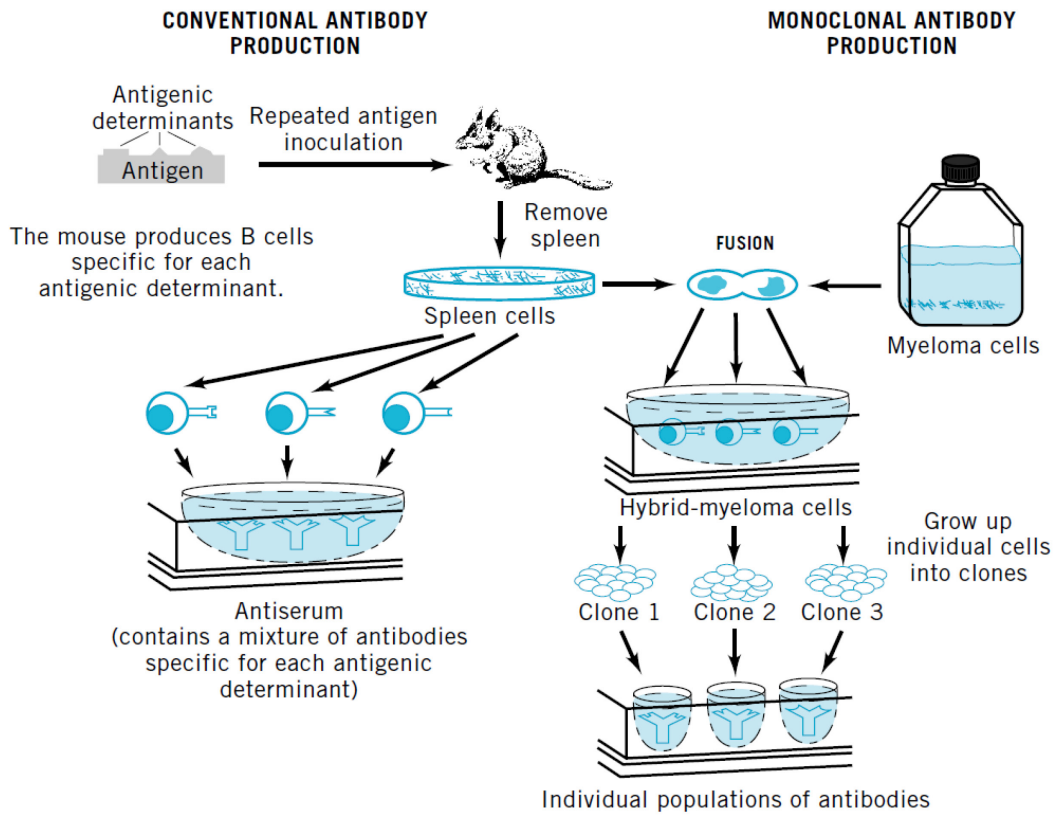


Figure 14. Conventional antibody production compared to monoclonal antibody production. The main difference between the two strategies is the resulting preparations, one is heterogeneous, while the other is homogeneous. B-cell-derived hybridomas can be separated into individual clones and grown indefinitely. One clone gives rise to identical daughter cells, all producing one antibody idiotype directed against a specific antigenic determinant (epitope), a monoclonal antibody. From: Elgert 2009.

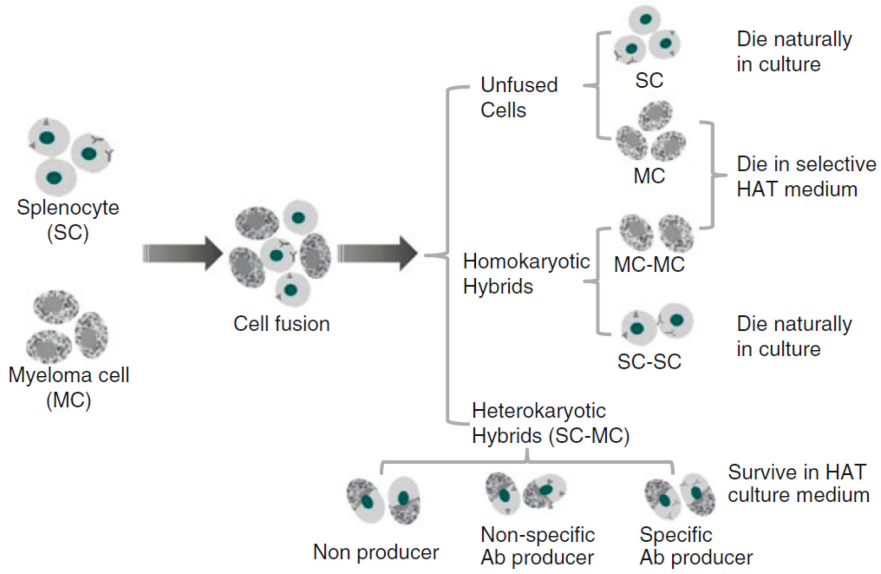


Figure 15. Multiple cell types generated by the fusion of splenocytes with myeloma cells. PEG-mediated cell fusion is likely to result in a mixed population of cells consisting of antibody-producing hybridomas, nonproducing hybridomas, and unfused cells. By hybridoma screening and subcloning, specific hybridoma clones will be identified and isolated from nonspecific antibody producers or nonproducers of myeloma-splenocyte hybridomas. From: Zhang 2012.

MATERIALS & METHODS

1. Hybridoma culture

For the production of the monoclonal Mannitou IgM different hybridoma culturing techniques have been applied. The cells, deposited by Birgit Zipser, were obtained from the Developmental Studies Hybridoma Bank created by the NICHD of the NIH and maintained at The University of Iowa, Department of Biology, Iowa City, IA 52242, USA.

The first method consists of culturing the cells at 37°C, 95% humidity, and 5% CO_2 in Roswell Park Memorial Institute RPMI 1640 medium (Gibco by Life Technologies, USA) supplemented with L-glutamine, 1% penicillin-streptomycin mixture and descending concentrations of fetal bovine serum (FBS). Every second day, the medium is replaced with one having less FBS - 10% FBS, 8% FBS, 6% FBS, 4% FBS, and 2% FBS respectively (Fig.16A). After ten days, the added medium contains no FBS and the cells are left in the incubator for two more weeks. Subsequently, the whole content of the culturing flasks is collected and centrifuged at 200 x *g* for 5 min. The supernatant containing the Mannitou IgM is skimmed off from the pellet and stored at 4°C for further purification.

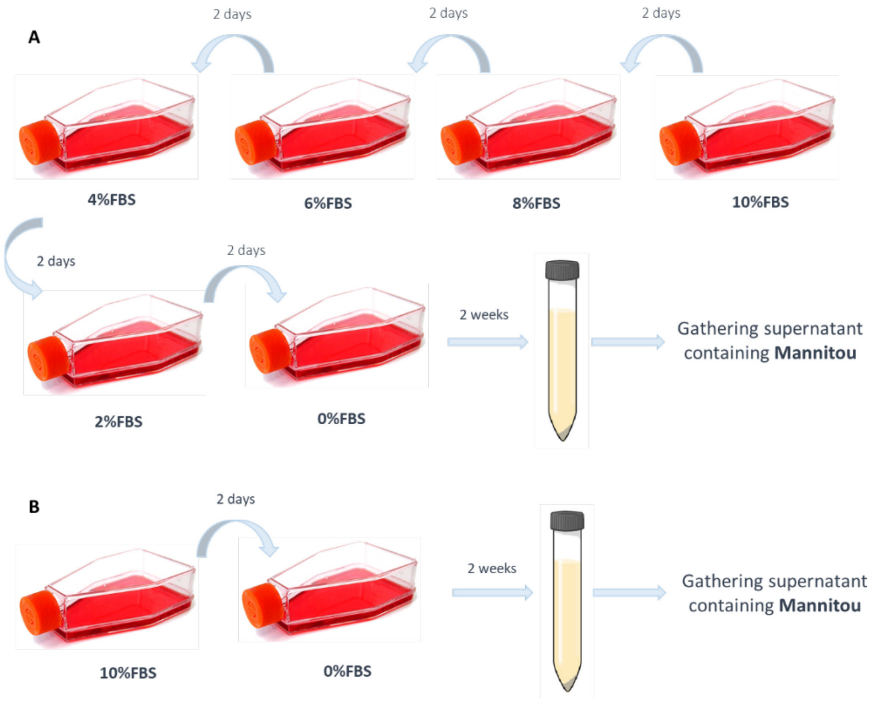


Figure 16. Production of Mannitou IgM by hybridoma cells using RPMI 1640 medium containing descending concentrations of FBS (A) and no FBS (B).

The second culturing technique involves growing the hybridomas at 37°C, 95% humidity, and 5% CO_2 for 2 days in RPMI 1640 medium containing 10% FBS and 1% penicillin-streptomycin. Subsequently, the medium is replaced with 0% FBS RPMI 1640 and the cells are cultured for two more weeks (Fig.16B). Next, the whole content of the culturing flasks is collected and centrifuged at 200 x *g* for 5 min. The supernatant containing the Mannitou IgM is skimmed off from the pellet and stored at 4°C for further purification.

For a higher yield production, a modified approach has been introduced (Fig.17). The method is based on initially culturing the cells at 37°C, 95% humidity, and 5% CO_2 in RPMI 1640 medium (Gibco by Life Technologies, USA) supplemented with 10% FBS and 1% penicillin-streptomycin. The medium is changed each week until the cells reach the recommended density. Once this is achieved, they are transferred to one-litre Wheaton spinner flasks and the medium is replaced with Hybridoma-SFM (Serum-Free Medium) containing L-glutamine (Gibco by Life Technologies, USA). The hybridoma cells are cultured for three weeks for the monoclonal antibody production. Subsequently, the content of the spinner flasks is collected and centrifuged at 200 x *g* for 5 minutes. The harvested supernatant containing the Mannitou IgM is stored at 4°C for further purification.

Figure 17. Production of Mannitou IgM by hybridoma cells in Wheaton spinner flask using SFM. The flask features an adjustable paddle blade impeller for better mixing and aeration. The addition of the bottom dimple improves circulation and reduces the accumulation of cells in the centre of the flask.

2. Mannitou mAb purification

Many purification approaches have been adopted, including affinity chromatography, anion-exchange chromatography, size-exclusion chromatography, and $(NH_4)_2SO_4$ precipitation.

Affinity chromatography separates biomolecules on the basis of a reversible interaction between a protein in the mobile phase and a specific ligand attached to the chromatographic matrix. The sample is applied under conditions that favour the specific binding of the target protein to the complementary binding substance (ligand). Elution is performed specifically, using a competitive ligand, or non specifically, by changing the pH, ionic strength or polarity. Several affinity columns have been used including HiTrap Protein A, HiTrap Protein L, and HiTrap IgM.

HiTrap Protein A is designed for purification and isolation of monoclonal and polyclonal IgG from ascites, serum and cell culture supernatants. Protein A consists of six different regions, five of which show strong, specific binding for the Fc part of IgG. Purified protein A is coupled to highly cross-linked agarose beads by the N-hydroxysuccinimide activation method. Immobilised protein A can bind at least two molecules of IgG per molecule.

HiTrap Protein L allows capturing a broad selection of mouse and rat antibodies and antibody fragments, such as Fab, scFv, and Dab. It exploits the ability of the protein L ligand to bind the variable region of an antibody's kappa light chain. Recombinant protein L has four binding domains and has an affinity for three of four kappa light chain subtypes in humans (1, 3, and 4) and kappa 1 in mice. Since no part of the heavy chain is involved in the binding interaction, protein L binds a wider range of antibody classes than protein A or G. Protein L binds to representatives of all antibody classes, including IgG, IgM, IgA, IgE, and IgD.

HiTrap IgM is designed for purification of monoclonal IgM from hybridoma cell cultures. The columns are packed with a thiophilic adsorption medium with 2-mercaptopyridine coupled to Sepharose High Performance (cross-linked agarose beads). Thiophilic adsorption is based on the ability of some proteins, particularly immunoglobulins, to bind to an immobilised ligand that contains a sulfone group near a thioether. Salts that interact with water molecules, such as potassium and ammonium sulfate, promote this binding by driving the protein into close proximity with the ligand.

The interaction between protein and ligand has been suggested to result from a combined electron donating and accepting action of the ligand, or alternatively as a mixed-mode hydrophilic-hydrophobic interaction.

Ion exchange chromatography is based on the reversible binding of charged sample molecules to oppositely charged groups attached to an insoluble matrix. Substances are bound to ion exchangers when they carry a net charge opposite to that of the ion exchanger. Conditions are then altered so that bound substances are eluted differentially. The elution is usually performed by increases in salt concentration or changes in pH. The pH value at which a biomolecule carries no net charge is called the isoelectric point (pI). When exposed to a pH below its pI, the biomolecule will carry a positive net charge and will bind to a cation exchanger (SP). At pH's above its pI the biomolecule will carry a negative net charge and will bind to an anion exchanger (Q). HiTrap Q is an anion exchange column. The resin has long chains of dextran coupled to a robust, 6% highly cross-linked agarose matrix.

Size-exclusion chromatography is a method in which molecules in solution are separated by their molecular size as they pass through a resin packed in a column. Samples are eluted isocratically. Buffer composition does not directly affect resolution.

Ammonium sulfate precipitation is a method used to separate proteins by altering their solubility in the presence of a high salt concentration. In solution, proteins form hydrogen bonds with water molecules through their exposed polar and ionic groups. When high concentrations of small, highly charged ions such as NH_4^+ and SO_4^{2-} are added, these groups compete with the proteins to bind to the water molecules. This removes the water molecules from the protein and decreases its solubility, resulting in precipitation.

The purification procedures have been executed on ÄKTA Pure chromatography system (GE Healthcare, USA). The successful purification has been performed under 0.5 mL/min flow on a 1 ml HiLoad 16/600 (GE Healthcare, USA) and 1 mL HiTrap NHS-activated HP (GE Healthcare, USA) affinity column, on which a 5 mM amino octyl mannose ligand (AOM) (Fig.18A) has been immobilised. A buffer containing 200 mM $NaHCO_3$ and 500 mM NaCl pH=7.8, serves as a coupling buffer. The elution has been performed with a solution of 100 mM heptyl mannose (HM) (Fig.18B).

The cell culture has been maintained by S. Robakiewicz. During the secondements and other absences, the culture has been maintained by J. Bouckaert and C. Bridot. HiTrap Protein A, HiTrap Protein L, HiTrap IgM, HiTrap Q, and $(NH_4)_2SO_4$ precipitation have been used to purify Mannitou antibody by S. Robakiewicz. During the secondements, J. Bouckaert and C. Bridot have purified the mAb using HiLoad 16/600 and HiTrap NHS-activated columns.

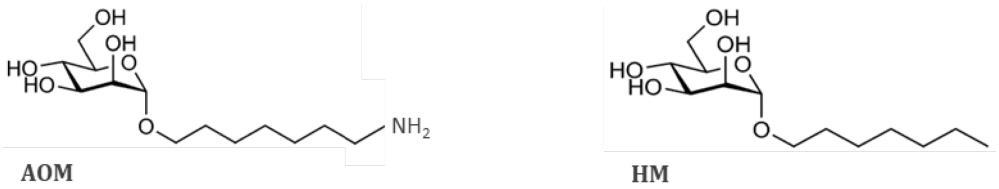


Figure 18. A: Amino octyl mannose structure. B: Heptyl mannose structure.

RESULTS

The Mannitou secreting hybridoma cells have been developed by intraperitoneal immunisation of female Balb/c mice with homogenised material from green horse leech *Haemopis marmorata* nerve cords, containing primary neurones and glia. The antigen has been emulsified with Freund's adjuvant. Spleen cells of the immunised mice have been fused with SP-2 myeloma cells to create the Mannitou mAb producing hybridomas (Bajt et al., 1990; Flaster et al., 1983; Zipser et al., 1981). The line used in this study was obtained from the Developmental Studies Hybridoma Bank.

Many purification approaches have been adopted, including affinity chromatography, anion-exchange chromatography, size-exclusion chromatography, and $(NH_4)_2SO_4$ precipitation. Several columns have been used, namely HiTrap Protein A (Fig.19), HiTrap Protein L (Fig.20), HiTrap IgM (Fig.21), and HiTrap Q (Fig.22). Unfortunately, none of these approaches has proved successful.

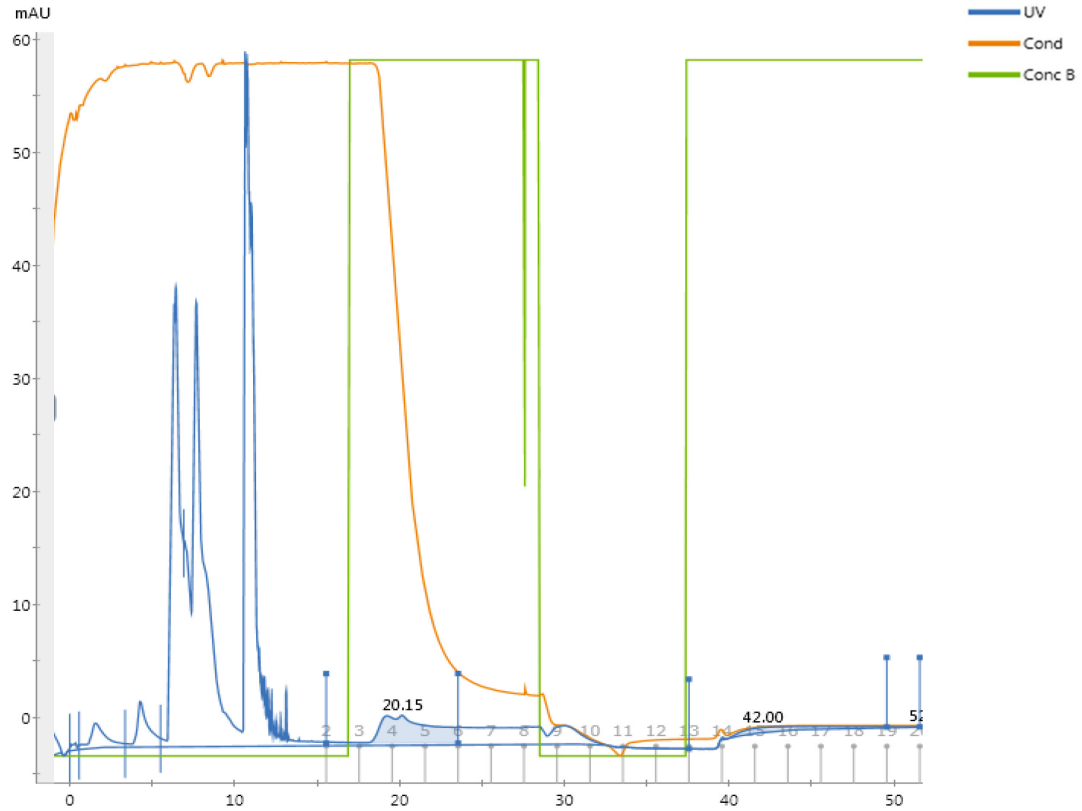


Figure 19. Mannitou IgM purification attempt using HiTrap Protein A.

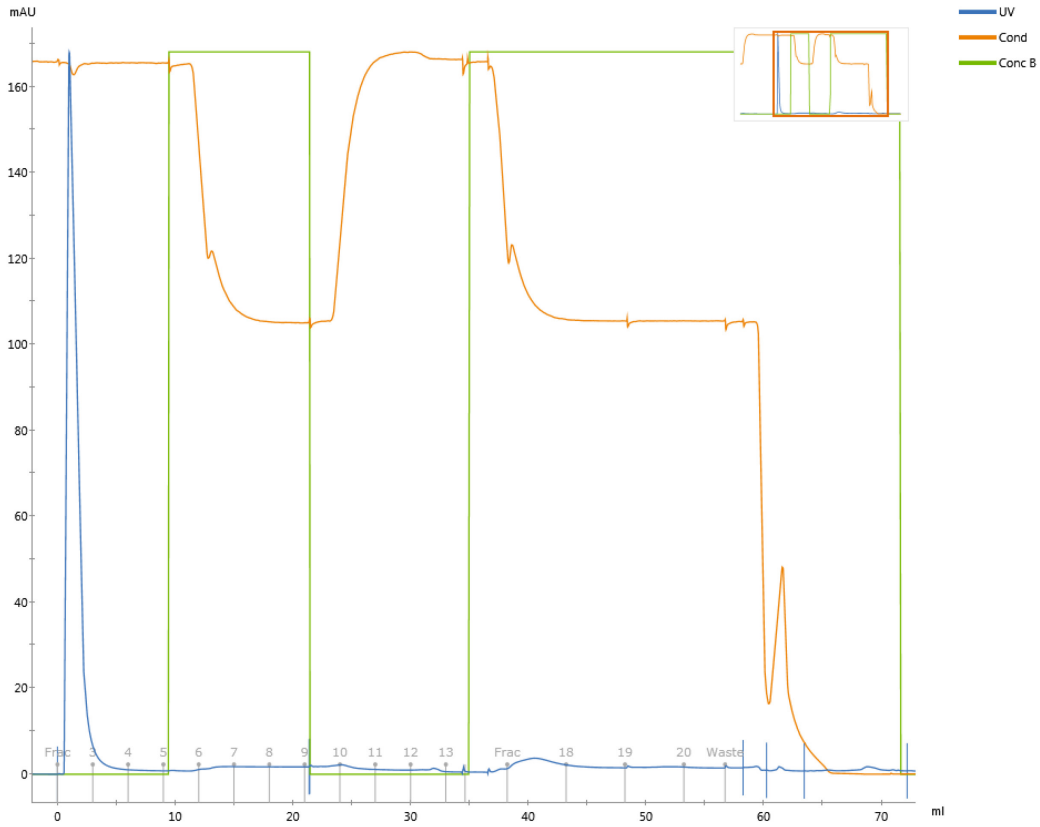


Figure 20. Mannitou IgM purification attempt using HiTrap Protein L.

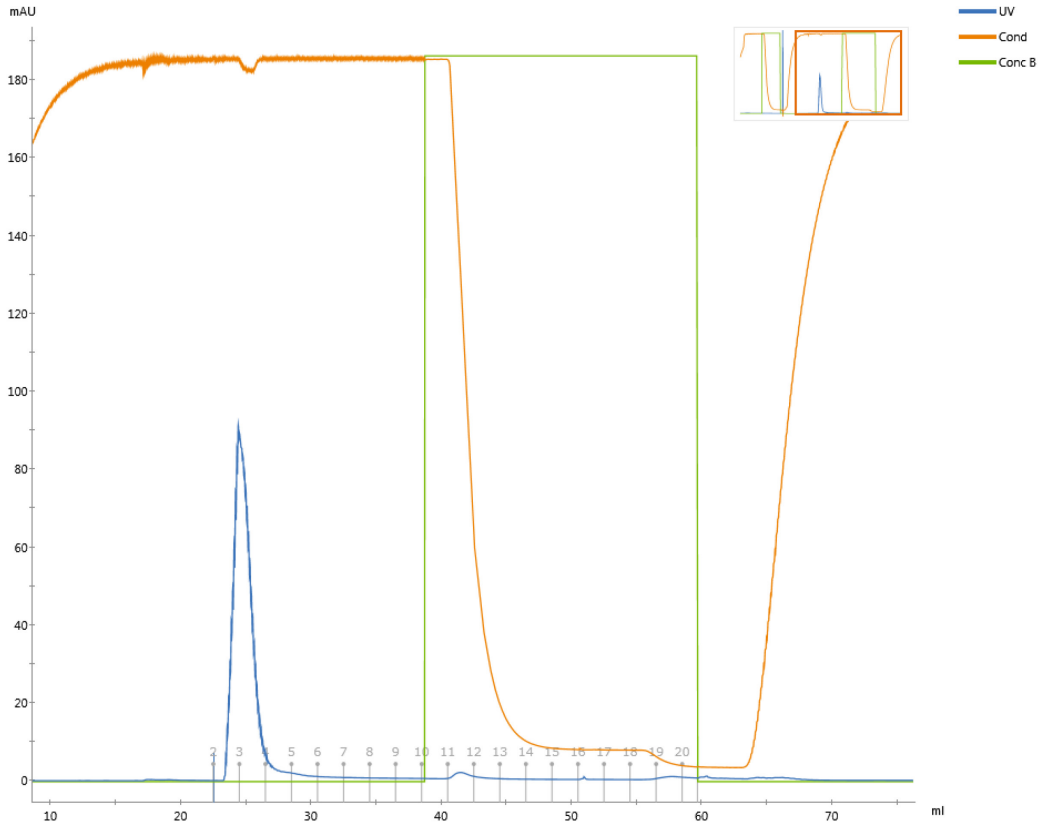


Figure 21. Mannitou IgM purification attempt using HiTrap IgM.

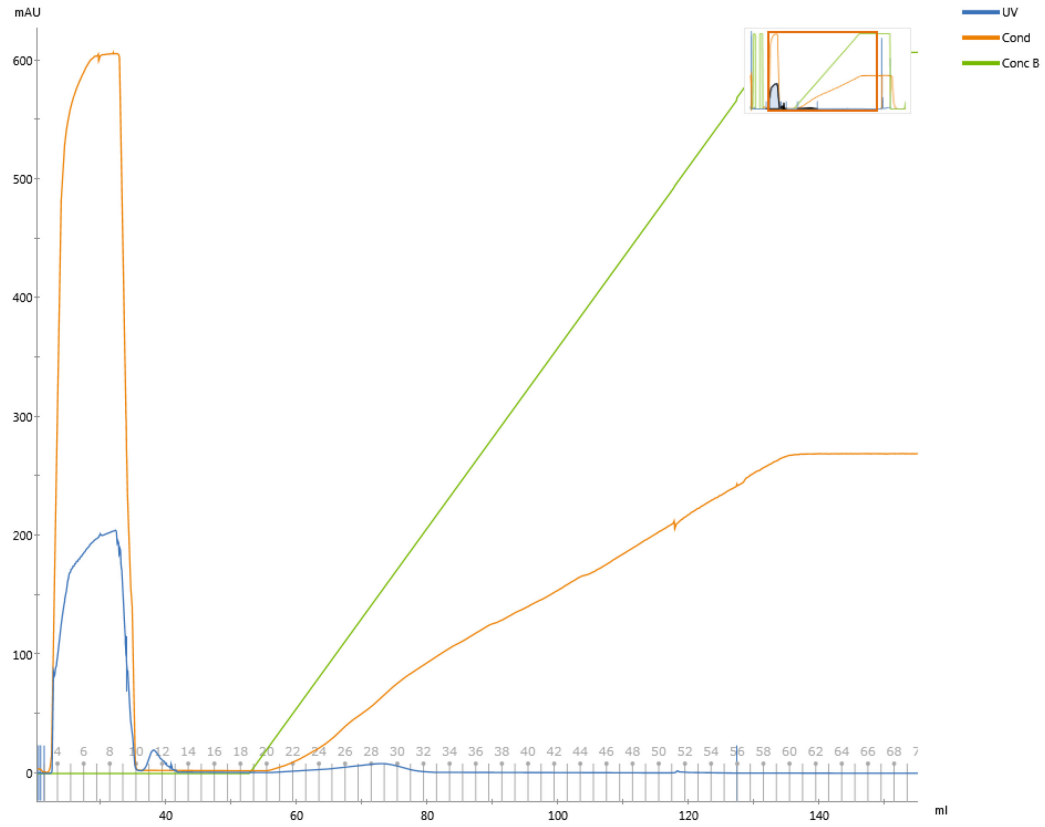


Figure 22. Mannitou IgM purification attempt using HiTrap Q.

Mannitou IgM has been successfully purified from the hybridoma cell culture supernatant employing HiLoad 16/600 (Fig.23) and HiTrap NHS-activated HP affinity column on which amino octyl mannose has been immobilised (Fig.24). To identify the elution peak of Mannitou IgM, the standard ultraviolet UV detection wavelengths have been changed to 218 nm (maximum absorbance for Mannitou) and 246 nm (higher absorbance for heptyl mannose, minimal absorbance for the antibody). Indeed, in what seems to be the Mannitou IgM elution peak, the UV value at 218 nm (violet) is the highest compared to the UV values at 280 nm (blue) and 246 nm (red) (Fig.24 left). For assessing the nature of the eluted product, SDS-PAGE analysis under reducing conditions has been performed. Two bands representing the heavy (75 kDa) and light (25 kDa) chains are visualised on a polyacrylamide gel (Fig.24 right). The excised bands have been subjected to a final validation by sequencing using MALDI-TOF MS Fingerprinting. The results reveal the presence of Mannitou heavy and light chains in the purified sample (Fig.25), confirming both the authenticity and purity of the antibody.

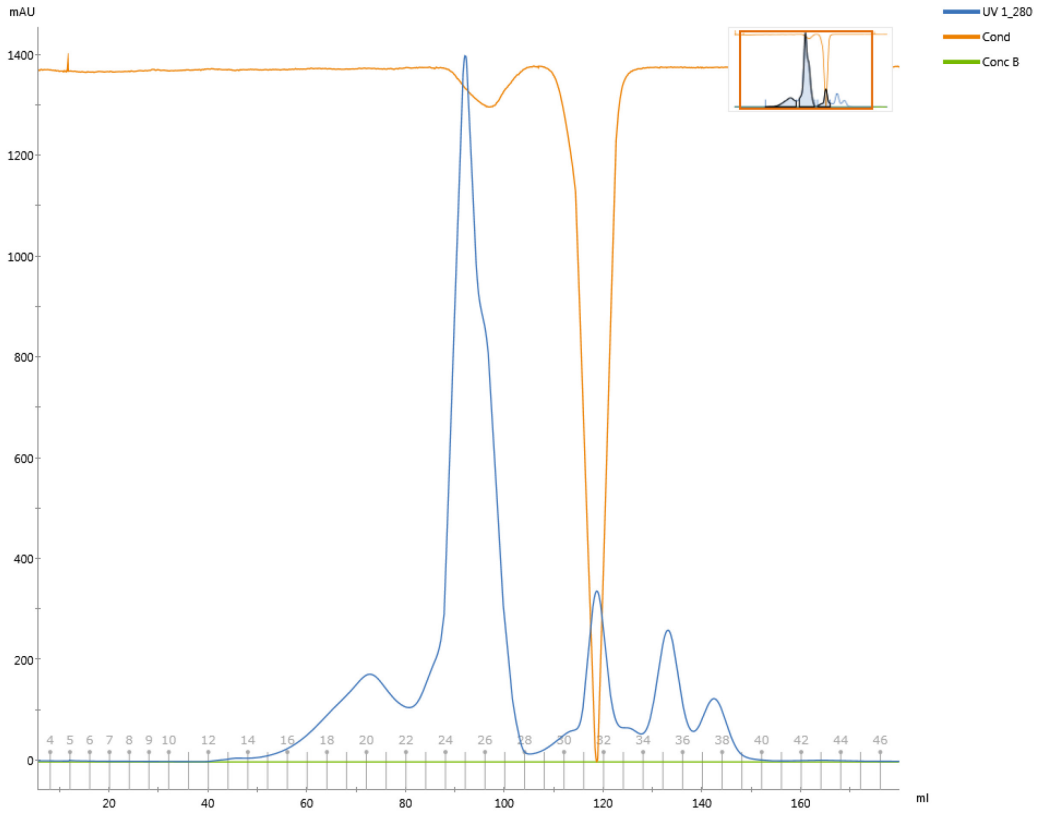


Figure 23. Mannitou IgM purification employing size-exclusion chromatography. The antibody has been purified on a HiLoad 16/600 column.

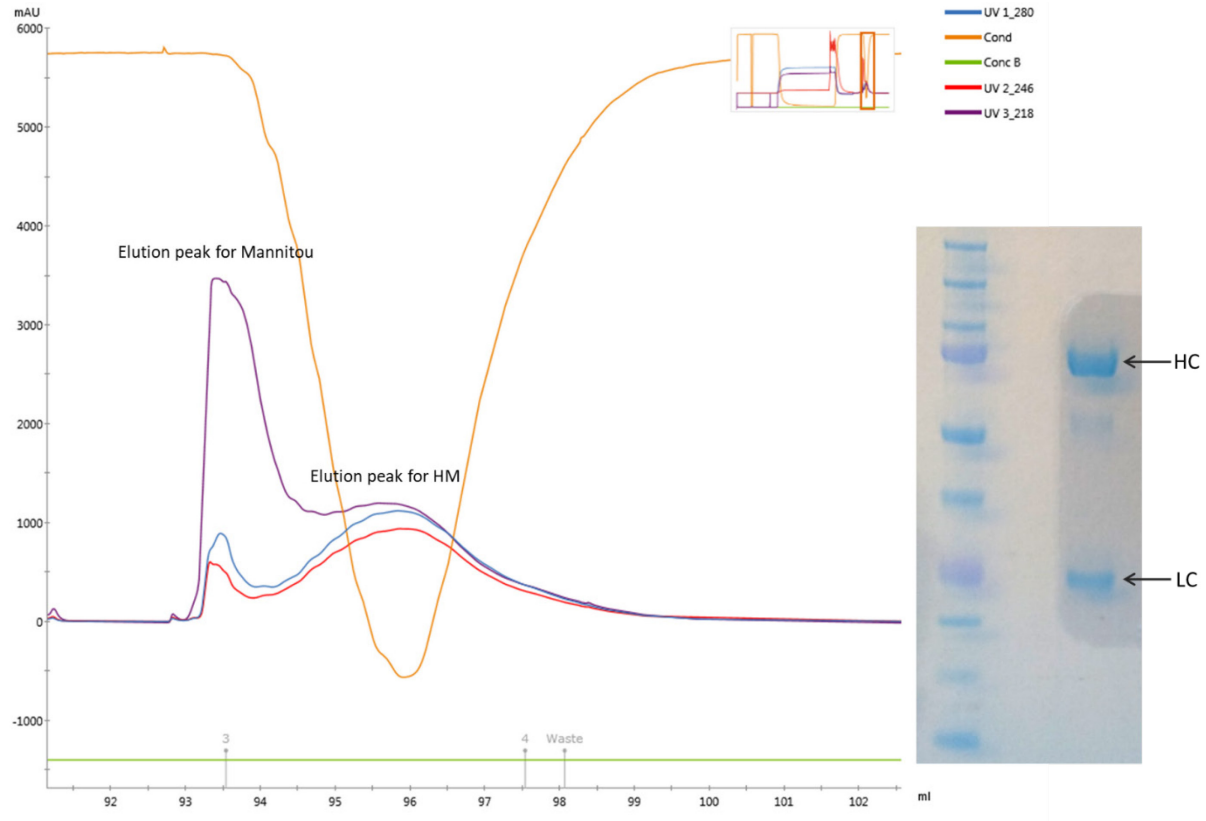


Figure 24. Mannitou IgM purification employing affinity chromatography. Left: Chromatography elution profile of Mannitou IgM. The antibody has been purified on a HiTrap NHS-activated HP column using 100 mM solution of heptyl mannose (HM) for elution. Right: Reducing SDS-PAGE confirmed the presence of both heavy (75 kDa) and light (25 kDa) chains.

Accession	Description	ΣCoverage	Σ# Proteins	Σ# Unique Peptides	Σ# Peptides	Σ# PSMs	Score A(2,4)	Coverage A(2,4)	# Peptides A(2,4)	# PSM A(2,4)	# AAs	MW [kDa]	calc. pI
ManHC	Mannitou IgM Heavy OS=XXxxx GN=MannLC PE=2 SV=	46.75	1	8	8	60	1891.30	46.75	8	60	231	25.7	7.81
P0AA27	Thioredoxin-1 OS=Escherichia coli O157:H7 GN=trxA F	33.94	1	3	3	3	143.68	33.94	3	3	109	11.8	4.88
F2QZP3	Thioredoxin OS=Komagataella phaffii (strain ATCC 762	11.76	1	1	1	1	56.24	11.76	1	1	102	11.3	4.93
F2QY66	Superoxide dismutase [Cu-Zn] OS=Komagataella phaffii	8.44	1	1	1	1	58.31	8.44	1	1	154	15.7	6.40
X2GM95	Serum albumin (Fragment) OS=Cervus nippon PE=2 SV	3.60	1	2	2	2	78.01	3.60	2	2	583	66.1	5.67
F2QNU8	Cerevisin OS=Komagataella phaffii (strain ATCC 76273	2.50	1	1	1	1	56.38	2.50	1	1	559	59.1	6.39
Tse8GST	ToxinaGST OS=xxxx GN=TOXGST PE=1 SV=2 - [TSE8	1.18	1	1	1	1	32.36	1.18	1	1	844	91.9	5.63

Accession	Description	ΣCoverage	Σ# Proteins	Σ# Unique Peptides	Σ# Peptides	Σ# PSMs	Score A(2,4)	Coverage A(2,4)	# Peptides A(2,4)	# PSM A(2,4)	# AAs	MW [kDa]	calc. pI
ManLC	Mannitou IgM Light OS=XXxxx GN=MannLC PE=2 SV=	58.64	1	8	8	54	2032.26	58.64	8	54	220	24.3	7.74
ManHC	Mannitou IgM Heavy OS=XXxxx GN=MannLC PE=2 SV=	28.57	1	5	5	6	298.60	28.57	5	6	231	25.7	7.81
P69777	Major outer membrane lipoprotein OS=Escherichia coli (	15.38	1	1	1	1	48.63	15.38	1	1	78	8.3	9.25
F2QTN6	Uncharacterized protein OS=Komagataella phaffii (strain	14.93	1	1	1	1	73.53	14.93	1	1	134	14.7	4.92
F2QZP3	Thioredoxin OS=Komagataella phaffii (strain ATCC 762	11.76	1	1	1	1	59.74	11.76	1	1	102	11.3	4.93
P0AA27	Thioredoxin-1 OS=Escherichia coli O157:H7 GN=trxA F	11.01	1	1	1	1	53.24	11.01	1	1	109	11.8	4.88
F2QY66	Superoxide dismutase [Cu-Zn] OS=Komagataella phaffii	8.44	1	1	1	1	50.63	8.44	1	1	154	15.7	6.40
B7NRM8	tRNA1(Val) (adenine(37)-N6)-methyltransferase OS=Esc	3.67	1	1	1	1	21.17	3.67	1	1	245	27.3	5.05
F2QNU8	Cerevisin OS=Komagataella phaffii (strain ATCC 76273	2.50	1	1	1	1	46.61	2.50	1	1	559	59.1	6.39

Figure 25. MALDI-TOF MS Fingerprinting results. Top: Mannitou IgM heavy chain sample. Bottom: Mannitou IgM light chain sample.

DISCUSSION

Before the establishment of the hybridoma technology, investigators could generate only polyclonal serum antibodies. This required large numbers of immunised animals and did not provide continuous antibody expression by immortalised cells, which in turn required repeated animal use. The development of the hybridoma technique has significantly reduced the number of lab animals needed to obtain a given antibody. The major advantage of using the continuous hybridoma cultures for mAb production, rather than the conventional route, is almost the unlimited supply of identical antibodies directed toward a single, specific epitope. In general, mAbs are present either in the medium supporting the *in vitro* hybridomas growth or the ascitic fluid in the peritoneal cavity of a mouse inoculated with the hybridoma cells into the abdomen. The huge interest in the mAb production technique derives from two fundamental factors. Firstly, the mAbs generated by an isolated clone are well-defined, unlike the heterogeneous antibody mixture in antiserum. Secondly, the method is ideally suited for the preparation of pure antibodies using non-purified antigens (Milstein et al., 1979).

The *in vitro* growth of hybridoma cultures in batches is technically relatively easy to perform. This technology has low start-up costs, with a similar start-to-finish time to the mouse ascites method, and enables the generation of mAb quantities comparable to those produced by the ascites method. Traditional cell culture media supplemented with serum contain a considerable level of immunoglobulins of animal origin which are difficult to separate from the murine mAbs during purification. Hence, media containing low serum levels (descending concentrations of FBS) and serum-free media are used to cultivate the hybridomas for *in vitro* generation of mAbs. In most cases, hybridoma cells growing in 10% FBS can be adapted within four passages (8-12 days) to grow in less than 1% FBS or FBS-free media. Still, this adaptation can take longer and in 3-5% of the cases, the cells will never adapt to the low FBS levels. Luckily, in the case of Mannitou hybridoma line, low-serum media, as well as serum-free media, have been well-tolerated in terms of viability. However, it cannot be assessed whether this interference does not affect the mAb generation ability of the cells. It has been noted that some cell lines that appear to be maintained adequately in low-serum or serum-free media, as assessed by viability, reach up to 10% lower antibody titers compared to when being maintained in higher-

serum media. Additionally, the quality of serum can vary between manufacturers and even batches. After the adaptation procedure, in order to maximise the yield of mAbs in the culture supernatant, hybridoma cells are allowed to incubate under standard growth conditions almost until the medium is depleted of nutrients. The culture conditions in the absence of animal immunoglobulins and at low levels of other proteins, make the purification of mAbs from the hybridoma supernatant much more effective and expedient. When the cells reach saturated density and the medium turns yellow, the culture supernatant is collected for antibody purification. This approach yields mAbs at concentrations that are typically below 20 μg/mL. Methods that raise the concentration of dissolved oxygen in the medium may increase cell viability and density at which the cells grow, and therefore boost mAb production. For that reason, throughout the course of the study, conventional cell-culture flasks were replaced with spinner flasks to keep the culture medium in constant circulation, and thus permit a more even distribution of nutrients and gases. All these methods can increase productivity substantially. However, antibody concentrations remain in the range of a few micrograms per millilitre. Even though the *in vitro* approach generates mAbs from over 90% of hybridomas, there is a considerable failure rate. If one faces the inability of a hybridoma line to grow and express mAbs *in vitro*, the only way to obtain adequate amounts of the antibody for experimental studies from such cells is to use the ascites method (National Research Council 1999).

One of the main disadvantages of the batch cell-culture method is the large volumes of culture media that need to be processed. This, in turn, leads to low mAb concentration levels and some mAbs are denatured during concentration or purification steps (Lüllau et al., 1996), which significantly decreases the antibody's activity and changes its affinity towards the antigen. During the course of this study, several Mannitou purification approaches were applied, including affinity and size-exclusion chromatography, as to minimise the possibility of its denaturation. Studies can be seriously confounded by purification procedures that alter the native structure of a mAb and result in a loss of activity. Immunoglobulin M often undergoes denaturation during *in vitro* purification techniques, resulting in the loss of complement-binding activity (Roggenbuck et al., 1994). Random antibodies of other isotypes exhibit similar quirks. OKT3 is an excellent example of a mAb with a prominent therapeutic application that cannot be adequately purified from culture supernatant and retain its full function at the same time, so it is produced by the ascites method (Goldstein 1987). Some of these *in*

vitro culture limitations have been overcome by using either semipermeable-membrane-based systems (Marx 1998) or hollow-fibre reactors (Jackson et al., 1996) to produce mAbs in concentrations often as high as those found in ascitic fluid but free of its contaminants, such as mouse immunoglobulins and other proteins. The use of a barrier, either a hollow-fibre or a membrane, with a low-molecular-weight cut-off (10,000-30,000 kD), has been implemented in several devices to permit cells to grow at high densities. Growth of hybridoma cells to higher densities results in greater amounts of mAbs that can be harvested from the culture medium. The objective of the semipermeable-membrane approach is to isolate the cells and the expressed mAbs from the culture medium by creating compartmentalisation. Nutrients and cell waste products can easily diffuse across the barrier allowing the culture to be supplemented with factors that help to optimise the growth of hybridomas. Depleted medium in the larger reservoir can be replaced without losing cells or mAbs. Similarly, cells and mAbs can be harvested independently of the growth medium. In the hollow-fibre bioreactor, the medium is continuously pumped through a circuit that consists of a hollow-fibre cartridge, gas-permeable tubing that oxygenates the media, and a medium reservoir. The hollow-fibre cartridge is composed of multiple fibres that run through a chamber containing hybridoma cells growing at high density. These fibres are semipermeable and serve a purpose similar to that of membrane-based systems. The hollow-fibre bioreactor is designed to provide total yields of 500 mg mAb or more. However, it is technically the most difficult *in vitro* system, as well as the most expensive one. For those reasons, hollow-fibre reactors are used only if large quantities of mAbs are crucially needed.

When a pure product is not necessary for research purposes but the maintenance of high affinity and biological activity is vital, the mouse ascites method is often the preferred alternative to obtain mAbs. The sophisticated eukaryotic machinery ensures the complement-fixing activity of an antibody and natural patterns of post-translational modifications such as glycosylation. *In vitro* hybridoma culture technology may sometimes yield populations of mAbs that are glycosylated at positions different from those expressed in mouse ascites fluid, thereby influencing their antigen-binding capacity and other important biological functions (Spearman et al., 2011). The loss of proper glycosylation might make the antibody product unsuitable for *in vivo* experiments because of increased immunogenicity, reduced binding affinity, or accelerated clearance. In 1995, Leibiger et al. described *in vitro* production of an IgG mAb that contained

terminal mannose moieties at all glycosylation sites. In some cases, such glycosylation substantially affected the antibody, whereas, in other, it was irrelevant. The authors attribute this unusual property to the *in vitro* culture conditions and speculate that the increased *in vivo* clearance of such antibodies was due to binding to mannose receptors (Leibiger et al., 1995). Since IgM is a potent complement-fixing antibody generated early in the human immune response in many infectious diseases, it is advisable to produce this mAb so that its biological activity and high affinity toward the antigen determinant are preserved. However, despite the low concentration levels of Mannitou reached by the presented hybridoma *in vitro* technique, and all previously raised concerns and limitations, the purity of the antibody has always been the crucial determinant. Therefore, for the purpose of this study the *in vitro* hybridoma culture approach was chosen. It ensured the synthesis of a pure Mannitou IgM, free of mouse ascetic fluid contaminants such as immunoglobulins and other proteins. This antibody was later used for numerous assays including microarray screening, surface plasmon resonance, cryo-EM characterisation, and crystallisation. Last but not least, the *in vitro* method was favoured as it does not use lab animals, unlike the *in vivo* approach that causes significant distress and pain to mice.

The identification of specific, high-quality hybridoma clones and subsequent mAb generation is a labour-intensive process. It requires a profound knowledge of multiple disciplines and practice of versatile technical skills, starting from animal handling (for immunisation and sample collection), through immunology (for immunoassays and antibody characterisation), to cellular and molecular biology (cell fusion for hybridoma generation, protein sequencing analysis for antigen preparation, and cell-based assays for hybridoma screening). Hybridoma technology has long been a powerful and indispensable platform for the generation of high-quality monoclonal antibodies. Hybridoma-derived mAbs have not only served as potent tool reagents but have also emerged as the most rapidly expanding class of therapeutics with the establishment of humanised and fully human mAbs (Fig.26) with reduced immunogenicity, destined for clinical application in humans. With the breakthrough in molecular engineering, mAb biopharmaceuticals have been used to aid in successful organ transplantation (Smith 1996), as well as to treat infectious diseases (Meissner et al., 1999), inflammatory diseases (Chan et al., 2010), and cancer (Blattman et al., 2004).

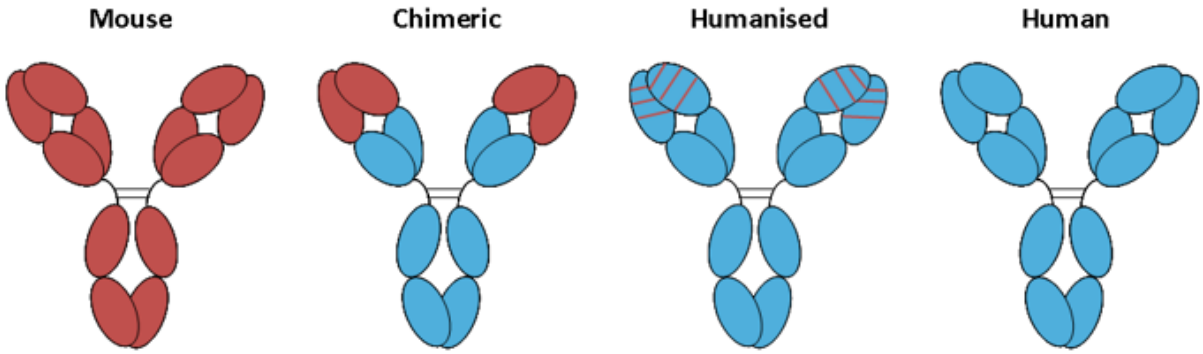


Figure 26. Schematic representation of the progressive humanisation of mAbs. To reduce the immunogenicity of murine antibodies a chimeric mAb has been generated (Boulianne et al., 1984). It consists of human constant regions (blue) and mouse variable regions (red) to retain its specificity. Grafting the complementarity-determining region from a mouse antibody onto a human variable region framework has led to the creation of a humanised mAb (Jones et al., 1986). The drive to reduce immunogenicity by decreasing the mouse content in a mAb culminated in the production of a fully human antibody (Green et al., 1994). The chimeric, humanised and fully human mAbs are much less immunogenic than the original murine antibodies but human anti-antibody responses have still been observed in patients. It should be noted that immunoglobulin sequences are highly homologous across species, thus a fully murine antibody is still relatively close in sequence to a fully-human one.

Chapter 1.

Mannitou antibody production

II. Generating Mannitou Fab by cloning and transient expression in HEK293T cells

INTRODUCTION

Recombinant antibodies are highly specific detection tools used in research (i.a. immunoblot assay, flow cytometry, immunohistochemistry), diagnosis of pathogens and toxins, and the fastest-growing class of therapeutic agents, primarily targeting inflammation sources or tumours (Eisenberg 2012). In order to meet these highly demanding needs, a variety of recombinant protein production systems have been developed over the years.

Immunoglobulin G is a heterotetrameric molecule consisting of two heavy and two light chains connected via disulphide bonds. Moreover, the heavy and light chains are stabilised by intramolecular disulphide bonds as well. These structural properties require a sophisticated folding apparatus as well as an oxidising environment for the generation of disulphide bonds. As a result, many traditionally established expression hosts do not provide these mechanisms for the efficient formulation of IgGs. Consequently, smaller antibody fragments have been developed, which combines both easier and cheaper production, and antigen-binding capacity of a full-length immunoglobulin (Frenzel et al., 2013; Saeed et al., 2017).

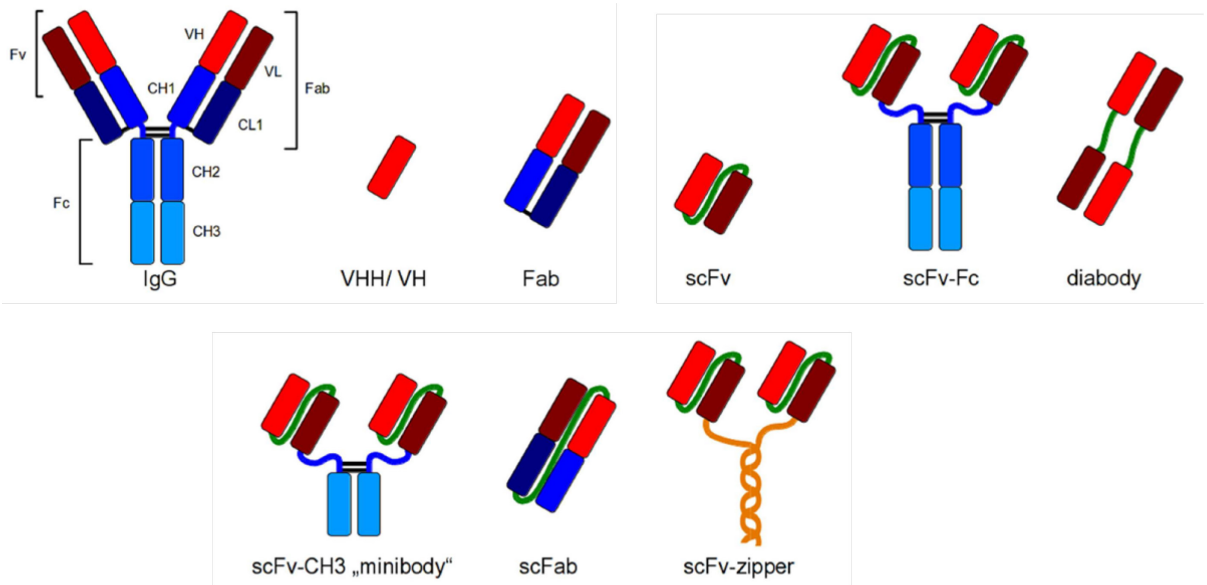


Figure 27. Various recombinant antibodies configurations for different applications. VH: variable region of the heavy chain; VL: variable region of the light chain; CHx: x domain of the heavy-chain constant region; CL1: first domain of the light-chain constant region; in red: variable regions; in blue: constant regions; green: artificial peptide linkers; yellow: dHLX represents amphiphatic helices used for dimerisation of scFv fragments. Adapted from: Frenzel et al., 2013; Nelson 2010.

The smallest antibody fragments are heavy chain antibodies (HCAbs) which are single-domain antibodies (sdAb) (Tang et al., 2013). These antibodies have a single variable domain, called VHH for camelids (Pant et al., 2006) and V-NAR for sharks (Flajnik et al., 2011). These fragments combine the high affinity and specificity of mAbs with the stability and ease of production of small molecules and also can be administered by means other than injection. In turn, the smallest antigen-binding unit with its complete antigen-binding site is the Fv (variable fragment) which consists only of variable regions. A short, flexible, and soluble peptide linker is used to connect the variable regions to form an scFv (single-chain variable fragment) (Bird et al., 1988). scFv are highly soluble, with strong affinity and binding specificity to their target antigen. By combining scFv with the first constant domains of the heavy and light chains, a Fab (antigen-binding fragment) is formed. All these smaller antibody fragments have several advantages over full-length ones, namely better tissue and tumour penetration, rapid blood clearance, short retention time, and reduced immunogenicity (Saeed et al., 2017). Other antibody formats produced in prokaryotic and eukaryotic cells include disulphide-bond stabilised scFv (ds-scFv) (Schmiedl et al., 2000), single-chain Fab (scFab) combining scFv and Fab properties (Hust et al., 2007), as well as multimeric antibody formats like dia-, tria-, or tetra-bodies (Hudson et al., 1999), minibodies (miniAbs) comprising different formats consisting of scFvs linked to oligomerisation domains like immunoglobulin CH3 domain (Hu et al., 1996), leucine zipper, helix turn, helix motif streptavidin, or scFv-scFv tandems (Plückthun et al., 1997) (Fig.27). Depending on their structure and application, recombinant antibodies can be produced in prokaryotic or eukaryotic hosts.

1. Prokaryotic hosts for rAb production

1.1. Gram-negative bacteria

The key step in generating a functional Fv is the secretion of both variable regions to the periplasmic space of *Escherichia coli* where the oxidising environment allows for the correct formation of disulphide bonds (Skerra et al., 1988). The production of recombinant antibodies in the reducing environment of cytoplasm results mainly in non-functional aggregates. Recovering functional antibody fragments from cytoplasmic inclusion bodies by complete denaturation and refolding is usually not efficient (Martineau et al., 1998). For Fab expression in *E. coli,* bicistronic vectors with the first cistron encoding the light chain and the second cistron encoding the Fd fragment are optimal. *Proteus mirabilis* is used for the formation of miniAbs (Kujau et al., 1998) and scFv (Rippmann et al., 1998). Recently, scFv has been produced in *Pseudomonas putidas* (Dammeyer et al., 2011).

1.2. Gram-positive bacteria

Due to the absence of an outer membrane, gram-positive bacteria directly secrete proteins into the medium, which facilitates the production of antibody fragments. *Bacillus brevis* has been applied in the production of scFv (Inoue et al., 1997) and Fab (Shiroza et al., 2003), *Bacillus subtilis* for obtaining scFv (Wu et al., 1998), and *Bacillus megaterium* for both, scFv and Fab (Jordan et al, 2009). So far, *Lactobacillus zeae/casei* (Krüger et al., 2002) and *Lactobacillus paracasei* (Marcotte et al., 2006) have been used for the production of scFv.

2. Eukaryotic hosts for rAb production

Eukaryotic cells possess an advanced folding, post-translational, and secretion apparatus, which enhances the secretory production of antibodies, including full-length immunoglobulins.

2.1. Yeasts

Yeasts combine the properties of eukaryotic cells, short generation time and ease of genetic manipulation, with the robustness and simple medium requirements of unicellular microbial hosts. Furthermore, yeasts do not secrete toxins as bacteria do. *Pichia pastoris* is the main yeast strain used for recombinant antibody production (Buckholz et al., 1991), including scFv (Eldin et al., 1997), scFv-Fc (Liu et al., 2003), and IgG (Potgieter et al., 2010). Other strains like *Saccharomyces cerevisiae* have also been employed in the expression of llama VHH (Frenken et al., 2000), Fab, and IgG (Horwitz et al., 1988). *Hansenula polymorpha*, *Schizosaccharomyces pombe*, *Schwanniomyces occidentalis, Kluyveromyces lactis*, and *Yarrowia lipolytica* are also sometimes applied but they do not play such a significant role (Frenzel et al., 2013). Engineered yeast cells are increasingly being used since yeasts cannot efficiently secrete larger heterologous proteins (>30 kDa), perform proteolysis during high-cell density fermentation, and appropriately glycosylate human glycoproteins. Yeasts tend to hyper glycosylate heterologous proteins even at positions not glycosylated in the native mammalian host. This is very likely to affect the activity of antibodies and is a potential source of immunogenicity and adverse reactions in human patients. The IgG obtained in glycoengineered *P. pastoris* strains, that have been designed to produce humanised glycosylation patterns, are able to mediate antibody effector functions (Hamilton et al., 2003).

2.2. Filamentous fungi

Aspergillus niger subgenus *A. awamori* has been employed in the production of llama VHH (Joosten et al., 2005), scFvs and antibody fusion proteins (Joosten et al., 2003). Efficient heterologous expression and secretion of llama VHH have been achieved in *Aspergillus oryzae* (Okazaki et al., 2012). Fab can be obtained in *Trichoderma reesei* (Nyyssönen et al., 1993).

2.3. Protozoa

Recently, *Leishmania tarentolae* has been explored as an expression system for different recombinant proteins (Niimi et al., 2012). The main advantage of this protozoa is its ability to perform *O*-glycosylation as well as *N*-glycosylation (Klatt et al., 2013).

2.4. Insect cells

Insect cells can be efficiently transfected with insect-specific viruses from the *Baculoviridae* family, particularly the *Autographa californica* nuclear polyhedrosis virus (AcNPV). The flexible envelope of Baculoviruses enables the packing of large heterologous gene sequences (>20 kb). Baculoviral protein expression is usually performed in insect cell lines Sf-9 and Sf-21 of *Spodoptera frugiperda*, DS2 cells of *Drosophila melanogaster*, or High Five cells of *Trichopulsia ni*. In comparison to Sf-9, High Five cells secrete up to 25-fold higher protein levels, have a shorter doubling time, adapt faster to serum-free medium, and grow in suspension culture. In contrast, Sf-9 and Sf-21 cell lines are recommended for high-titer viral stock production, due to higher transfection efficiency (Frenzel et al., 2013). scFv (Kretzschmar et al., 1996), as well as IgG capable of mediating effector functions such as complement binding (Zu Putlitz et al., 1990) and ADCC (Edelman et al., 1997), have been produced in insect cells. Alternatively to baculoviruses, Schneider 2 (S2) cells of *D. melanogaster* can be transfected with expression vectors in a transient or stable manner (Frenzel et al., 2013).

2.5. Mammalian cells

Currently, almost 70% of all recombinant protein pharmaceuticals and 95% of the therapeutic antibodies are obtained from mammalian cells despite relatively high production costs. Nevertheless, the advanced mammalian folding, assembly and post-translational apparatus are capable of generating antibodies that are not only indistinguishable from those produced in the human body but are also non-immunogenic. Moreover, mammalian cell expression system is highly efficient for the secretion of large and complex IgG. All these features allow the generation of high-quality proteins. Nowadays, industrial IgG production levels often exceed 12 g/L, as a result of improved high-producer cell lines, optimised expression media, and prolonged incubation time at high-cell densities. Producer cell lines have been genetically engineered regarding product homogeneity, improved metabolism, reduced apoptosis, and inducible cell cycle arrest, which allows prolonged production time for almost three weeks at high-cell density and viability. The most commonly used mammalian lines are the Chinese hamster ovary cells CHO, human embryonic kidney cells HEK293, baby hamster kidney cells BHK21, mouse myeloma NS0 cells, and embryonic retinal cell line Per.C6 (Frenzel et al., 2013).

2.5.1. Stable antibody production

The generation of stable master cell lines is a prerequisite for the manufacture of biopharmaceutical immunoglobulins in order to guarantee long-term production stability. The main advantages compared to transient expression systems are the high degree of batch-to-batch consistency and the potential of large-scale production. Chinese hamster ovary cells (CHO) are the most widely used host cells for stable antibody expression. The antibody gene expression cassettes have to be stably integrated into the host cell genome. Strong promoters, like the cytomegalovirus (CMV) immediate-early promoter, or the translation elongation factor EF-1α promoter, and polyadenylation sites from the simian virus 40 (SV40), or the bovine growth hormone (BGH) for improved mRNA stability and translation efficiency, are usually implemented into the expression vector. For proper IgG expression, two different genes must be stably transfected into one cell clone, either by co-transfection, or using "double gene" vectors (combining complete expression cassettes for the heavy and light chains on a single plasmid) or by bicistronic expression vectors, since the ratio between the light and heavy chains has a great impact on the secretion level of a functional immunoglobulin. The bicistronic set-up combines the heavy and light chains in a single expression cassette and contains an internal ribosomal entry site (IRES). IRES allows the translation of two or more cistrons from the same transcript. With bicistronic vectors, both chains are encoded on a single mRNA, which leads to a balanced supply of both polypeptides. A different technology requiring both chains to be encoded as a single polypeptide allows the co-expression of the heavy and light chains from a single open reading frame (ORF) (Jostock 2011). In order to overcome the negative effects of the integration site, protective cis-regulatory elements including insulators, boundary elements, scaffold/matrix attachment regions (S/MARs), chromatin opening elements, and antirepressor elements are introduced into the vector. This reduces the influence of heterochromatin and stabilises the transgene expression (Frenzel et al., 2013).

2.5.2. Transient antibody production

Establishing a stable producer cell line is still very expensive, time-consuming and laborious for research applications. In the case of small-scale production, transient antibody expression is much more suitable since it enables fast manufacturing without the need to generate producer cell lines, and therefore is also cost-effective. The human

embryonic kidney cell line HEK293 has been widely used for transient antibody expression as it is highly transfectable and capable of episomal replication. Moreover, HEK293 express the adenoviral E1A transactivator which is a general transcription enhancer. Chinese hamster ovary cells CHO are also gaining popularity, although typically lower yields are achieved. Apart from these two cell lines, African green monkey kidney (COS) cells and baby hamster kidney (BHK) cells may be occasionally employed. Two variants of episomal replicating derivatives are well established: HEK293-T and HEK293-EBNA. In HEK293-T cells the large T antigen of SV40 is constitutively expressed which supports extrachromosomal maintenance of plasmids containing an SV40 origin of replication (SV40ori) and facilitates high amplification rates of these vectors. In HEK293-EBNA cells the copy number of vectors after episomal replication is considerably lower and driven by the EBNA1 transactivator and oriP from Epstein-Barr virus (EBV). Nevertheless, due to high plasmid retention, long transgene expression periods are achieved. Besides those two episomal replication-competent HEK293 derivatives, HKB-11 (HEK293 fusion with a lymphoma-derived cell clone) and 293-F (293 Freestyle) cells are commonly used for transient expression of antibodies and other proteins (Geisse et al., 2009; Jostock 2011).

Whether a transient expression system yields high production values depends on the transfection efficiency. Transient transfection of plasmid DNA in HEK293 cells can be performed with calcium phosphate, cationic liposomes, or polymers like polyethyleneimine (PEI). With many transfection methods, small vectors reach higher gene transfer rates than large ones. Moreover, large vectors tend to result in instabilities and low yields. Thus, for transient expression, minimal vector backbones with only the most essential elements are recommended. In order to introduce into the host cell both heavy and light chain of an antibody, different strategies can be applied (Fig.28). Co-transfection with individual plasmids encoding each chain has the advantage of comparably small vector size and the possibility to change the ratio of heavy to the light chain. Alternatively, both chains can be provided from one plasmid with a monocistronic “tandem” (both cassettes have the same orientation and are positioned sequentially) or “sandwich” set-up (the cassettes are in opposed orientation). Both heavy and light chains can be also produced by bicistronic expression vectors with an internal ribosomal entry site or by a monocistronic single open reading frame set-up (Jostock 2011).

2.6. Transgenic organisms

2.6.1. Plants

Not only can up-scaling be achieved much more easily but also the manufacturing costs are more than ten times lower compared to other expression systems. The generation of a stable genetically modified dicotyledonous plant requires several months of transformation and special regeneration protocols. In principle, it is done by transferring the expression cassette of the transgene with the help of *Agrobacterium tumefaciens*. The gene of interest is cloned into the T-DNA of a binary plasmid (Hoekema et al., 1983). In most cases, the expression of the transgene is under the control of cauliflower mosaic virus (CaMV), a constitutive promoter widely used in plants for recombinant protein synthesis. Alternatively, transient expression systems are implemented that allow achieving time-saving production. *Nicotiana benthamiana* can be transducted with a tobacco mosaic virus TMV-based vector for the secretion of different scFv for the treatment of non-Hodgkin's lymphoma (McCormick et al., 2003). Regardless of the expression system, differences in the glycosylation pattern between plants and mammals are one of the main obstacles to overcome when developing therapeutic antibodies in plant cells. Although plants are able to perform complex glycosylation, these differences, in particular β1,2-xylose and α1,3-fucose, can lead to immunogenicity of the synthesised proteins. For this purpose, different strategies have been developed for recombinant proteins expression with mammalian-like glycosylation patterns. The first approach is the retention of the proteins in the endoplasmic reticulum as ER-associated *N*-glycosylation leads to the generation of oligomannose-type *N*-glycans, which are identical in plant and mammalian cells (Gomord et al., 2005). Another method involves using glycoengineered plants. In most cases, RNA interference (RNAi) is used for down-regulation of endogenous β1,2-xylosyl transferase and α1,3-fucosyl transferase leading to the reduction of the xylosylated and core-fucosylated *N*-glycans (Schähs et al., 2007). The second type of glycoengineering in plants is the co-expression of genes that facilitate the generation of human-like *N*-glycans (Vézina et al, 2009) or even protein sialylation by the co-expression of six mammalian genes (Jez et al., 2013). Generally, plantibodies are produced in tobacco (*N. tabacum* or *N. benthamiana*), but there are also systems using *Arabidopsis thaliana, Medicago sativa, Lemnaminor* (duckweed), rice, lettuce, maize, wheat, tomato, potato and banana. Moreover, monoclonal BY-2 tobacco cells growing in suspension have been developed (Frenzel et al., 2013; Kodati et al., 2016).

2.6.2. Animals

As the humanisation of hybridoma-derived antibody therapeutics is still a laborious and time-consuming procedure, the popularity of human antibody expression in transgenic animals is increasing. Furthermore, antibodies generated in mice or rats may elicit an immune response in human patients. One method for the generation of fully human binders is the antibody phage display technology. Moreover, several transgenic animals have been developed for the production of human monoclonal and polyclonal antibodies. Transgenic cattle carrying human immunoglobulin heavy and kappa light chain loci have been used for the immunisation with an anthrax protective antigen. The resulting polyclonal antibody mixture consisted of entirely human and chimeric immunoglobulins that showed high activity and were protective *in vivo* mouse models (Kuroiwa et al., 2009). Rabbits and cattle have been used for the expression of a bispecific scFv targeting the melanoma-associated proteoglycan and the human CD28 molecule on T cells (Grosse-Hovest et al., 2004).

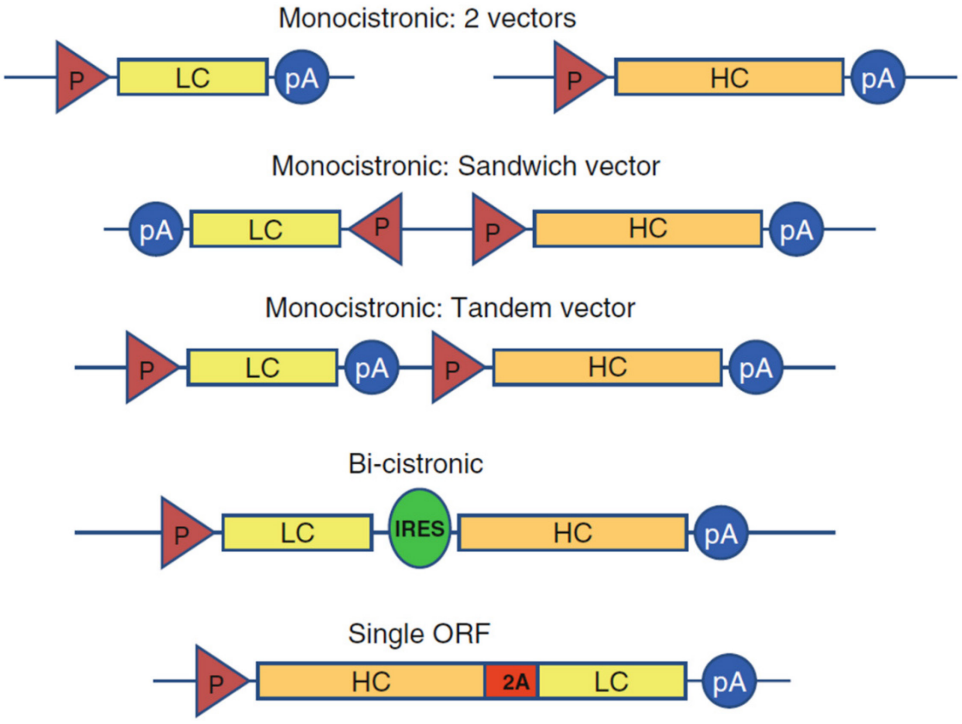


Figure 28. A schematic representation of possible vector set-ups for the transient expression of both antibody chains in mammalian cells. The HC and LC positions can be swapped. The heavy and light chains can be encoded on separate plasmids or combined in single vector with a monocistronic, bicistronic or single ORF set-up. P: promoter, pA: poli(A) tail, IRES: internal ribosomal entry site, 2A: self-processing 2A sequence motive of viral origin. Adapted from: Jostock 2011.

MATERIALS AND METHODS

1. Sequence determination

The genomic sequences of the variable regions (VL and VH domains) of Mannitou antibody have been determined by Fusion Antibodies Ltd by total RNA extraction from hybridoma cells (Laz6-189/Mannitou Ab) and cDNA synthesis by RT-PCR. The positive PCR products are identified by agarose gel electrophoresis. These are then analysed by DNA sequencing on an ABI3130xl Genetic Analyzer (Applied Biosystems, CA, USA). The constant regions have been obtained from GenBank: LC085625.1 (Ojima-Kato et al., 2016) for the light chain, and from UniProtKB: locus A24976 (Schreier et al., 1986) for the heavy chain (Fig.S1).

2. Cloning

The first step in producing the Fab is to obtain the genes encoding its sequence. It has been done by providing the Fab amino acid sequence to the company responsible for the gene synthesis and generation of the construct (GenScript, NJ, USA). For the purpose of this study, two different sequences have been designed - one for the light chain (LC), and one for the heavy-chain Fd (comprising the heavy-chain variable domain VH together with the 1st domain of the heavy-chain constant region CH1; later referred to as HC) carrying a hexahistidine tag (Fig.29). The genes encoding the LC and HC have been codon-optimised, synthesised and subcloned into a pUC57 *E. coli* expression vector.

Mannitou Fab light chain aa sequence:

DVVVTQTPLSLPVSFGDQASISCRSSQSLVNSYGSTYLSWYLHRPGQSPQLLIYGISNRFSGVPD
RFSGSGSGTDFTLTIRTIKPEDLGMYYCLQGTHQPWTFGGGTKLEIKKRADAAPTVSIFPPSSE
QLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDE
YERHNSYTCEATHKTSTSPIVKSFNRKEC

Mannitou Fab heavy chain aa sequence:

EVKLLESGGGLVQPGGSLKLSCAASGFDFSTYWMSWVRQAPGKGLEWIGEINPDSSTINYTPSL
KDKFIISRDNAKNTLYLQMSKVRSEDSVLYYCVRPGTWGYFDYWGQGTTLTVSSASQSFPNVF
PLVSCESPLSDKNLVAMGCLARDFLPSTISFTWNYQNNTEVIQGIRTFPTLRTGGKYLATSQVL
LSPKSILEGSDEYLVCKIHYGGKNRDLHVPIPAVAHHHHHH

Figure 29. Mannitou Fab amino acid sequence. Green: variable region, Pink: constant region. The sequence is composed of the variable regions, determined by sequencing, and the constant regions obtained from GenBank: LC085625.1 (light chain) and UniProtKB: locus A24976 (heavy chain). 7 additional amino acids from the CH2 domain (black) and 6 histidine residues (yellow) have been added to the heavy chain sequence.

2.1. DNA amplification by PCR

Once the constructs are generated, the DNA of interest is amplified by PCR. The primers have been designed using SnapGene 4.1.9 software (GLS Biotech LLC) (Table 1). The PCR is performed under the following conditions:

10 µl Q5 Reaction Buffer	98°C	30 sec	initial denaturation
1 µl dNTPs solution mix	98°C	10 sec	35 cycles
2,5 µl forward primer (10 µM)	55°C	30 sec	
2,5 µl reverse primer (10 µM)	72°C	60 sec	
1 µl DNA template	72°C	2 min	final extension
0,5 µl polymerase			
32,5 µl nuclease-free H_2O			

HC_FP	GAC TAG TAC CGG TGA GGT GAA GCT TCT CGA GTC TGG
HC_RP	GAC TAG TGG TAC CTT AGT GAT GGT GAT GG
LC_FP	GAC TAG TAC CGG TGA TGT TGT GGT GAC TCA AAC TCC ACT C
LC_RP	GAC TAG TGG TAC CTT AAC ACT CTT TCC TGT TG

Table 1. The PCR primers used for the amplification of the HC and LC. FP – forward primer, RP – reverse primer.

2.2. DNA recuperation

Following validation of the presence and size of the inserts by electrophoresis, the DNA is recuperated from the agarose gel using Wizard® SV Gel and PCR Clean-Up System (Promega, USA). The DNA bands are excised from the gel and placed in microcentrifuge tubes. 10 µl Membrane Binding Solution is added per 10 mg of gel slice. The content is vortexed and incubated at 50–65°C until the gel slices completely dissolve. An equal volume of Membrane Binding Solution is added to the PCR product. SV Minicolumn is inserted into the Collection Tube. The dissolved gel mixture and the PCR product are transferred to the Minicolumn assembly, incubated at room temperature for 1 minute and centrifuged at 16 000 × g for 1 minute. The flowthrough is discarded, 700 µl Membrane Wash Solution is added and then centrifuged again at 16 000 × g for 1 minute. The flow-through is discarded, 500 µl Membrane Wash Solution is added and centrifuged at 16 000 × g for 5 minutes. The flowthrough is discarded, and the column assembly is recentrifuged for 1 minute with the microcentrifuge lid open to allow evaporation of any residual ethanol. The Minicolumn is carefully transferred to a clean microcentrifuge tube. 50 µl nuclease-free water is added to the Minicolumn, incubated at room temperature for 1 minute, and then centrifuged at 16 000 × g for 1 minute. The DNA is stored at –20°C.

2.3. DNA ligation into pHL-sec vectors

Before subsequent ligation of the genes into the suitable vector, restriction enzymes have been applied to introduce sticky ends at both the vector and the insert. AgeI and KpnI are endonucleases, therefore the fragments amplified by PCR need to be ~6 bp longer than the actual genes of interest, to allow the restrictases to cleave the phosphodiester bonds within the sequences. The formation of sticky ends is performed at 37°C for 4 hours including the following reagents: 2 µl 10x CutSmart Buffer, 1 µl AgeI, 1 µl KpnI, 2 µl DNA and 14 µl nuclease-free H_2O. After amplifying and purifying the DNA fragments of the LC and HC using Wizard SV Gel and PCR Clean-Up System (Promega, USA), they are inserted into "empty" pHL-sec vectors. The ligation lasts for 1 hour at room temperature using the following reagents: 5 µl 2x Rapid Ligation Buffer, 1 µl T4 DNA Ligase, 2.5 µl pHL-sec vector, 1.5 µl insert. The ligation results in obtaining the expression vectors ready to be transfected into cells.

2.4. Vector amplification

Once the vectors are obtained, they are amplified in competent DH5α *E. coli* cells. The preparation starts with inoculating 2-3 mL culture in LB and shaking it at 180 rpm for 15 hours at 37°C. On the following day, 1:100 dilution is made in LB and left at 37°C until the cells reach the exponential phase - OD_{600} = 0.3-0.4, which usually takes 2-3 hours. The cells are then centrifuged at 10 000 x *g* for 20 minutes at 4°C. After pouring off the supernatant, the pellet is resuspended in 1/10 of the initial volume in ice-cold 100 mM $CaCl_2$. The mix is left on ice up to 30 minutes and afterwards centrifuged at 5 000 x *g* for 5 minutes at 4°C. 10 µl of the ligation product (1-50 ng) is added to 100µl of the n competent DH5α cells and everything is left on ice for 30 minutes. In order to enable the DNA to enter the cells, they are subjected to heat shock: the mix is put in a 42°C water bath for 45 seconds and then immediately on ice for another 2 minutes. Next, 1 mL of LB medium is added and the whole mix is shaken at 250 rpm for 1 hour at 37°C for phenotypic expression. Finally, 200µl of the transformed bacteria are streaked onto already prepared agar plates containing the appropriate selective antibiotic (40 µl ampicillin or carbenicillin/100 µl LB) and left to grow overnight at 37°C (Fig.30).

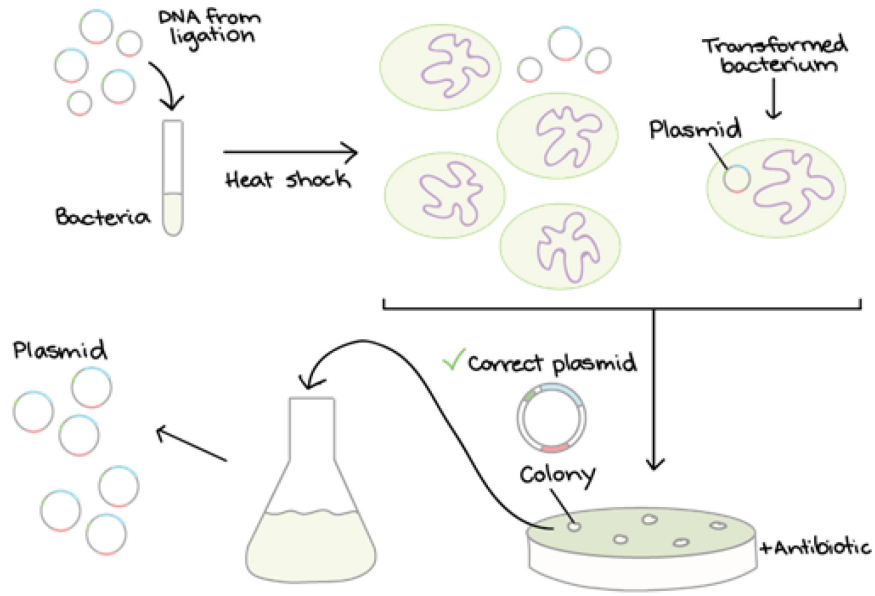


Figure 30. Schematic representation of plasmid amplification using transformed bacteria cells.

On the following day, to verify if the appearing colonies contain the plasmids of interest, colony PCR is performed. Random colonies from each plate are picked up and placed at 95°C for 5 minutes to release their DNA. The colony PCR required the following conditions:

10 µl GoTaq G2 Green Master Mix	95°C	2 min	initial denaturation
2 µl forward primer (10 µM)	95°C	30 sec	30 cycles
2 µl reverse primer (10 µM)	50°C	30 sec	
2 µl DNA template	72°C	1 min	
4 µl nuclease-free H_2O	72°C	5 min	final extension

Once the colonies containing the plasmids of interest are determined, a single colony is picked up to inoculate a starter liquid bacteria culture (in 200 mL LB medium supplemented with 200 µl ampicillin/ carbenicillin). The cultures are left to grow overnight in a thermostatic shaker at 250 rpm at 37°C. On the following day, the cells are harvested and the DNA is purified using EndoFree Plasmid Kit (Qiagen, Germany).

2.5. DNA purification

The bacterial DNA has been purified using EndoFree Plasmid Kit (Qiagen, Germany). The harvested bacterial cultures are centrifuged at 6 000 x *g* for 15 minutes at 4°C. Then, the pellet is resuspended in 10 mL Buffer P1, and another 10 mL of buffer P2 is added and mixed by inverting 5 times. The mix is incubated for 5 minutes at room temperature and 10 mL of pre-chilled Buffer P3 is added and mixed by inverting 5 times. The solution is incubated on ice for 20 minutes and afterwards centrifuged at 20 000 x *g* for 30 minutes at 4°C. In the meantime, a QIAGEN-tip 500 is equilibrated by applying 10 mL Buffer QBT and allowing the column to empty by gravity. Next, the supernatant is poured and allowed to enter the resin by gravity flow. The QIAGEN-tip is washed twice with 30 mL of Buffer QC. The DNA is eluted with 15 mL pre-warmed (65°C) Buffer QF into a clean vessel. The DNA is further precipitated by adding 10,5 mL room-temperature isopropanol and the mixture is centrifuged at 15 000 x *g* for 30 minutes at 4°C. After carefully decanting the supernatant, the DNA pellet is washed with 5 mL room-temperature 70% ethanol and centrifuged at 15 000 x *g* for 10 minutes. After carefully decanting the supernatant, the pellet is left to air-dry for 10 minutes and then redissolved in 1 mL of endotoxin-free Buffer TE, pH=8.

2.6. Control digestion

To verify the sequences of the purified plasmids, control digestion is set up at 37°C for 1 hour using 2 µl 10x Fast Digest Green Buffer, 1 µl Fast Digest Hind III, 1 µl Fast Digest Pvu II, 1 µl DNA and 15 µl nuclease-free H_2O. A simulation of the activity of Hind III and Pvu II restriction enzymes has been done using SnapGene version 4.1.9.

3. Transient expression in human embryonic kidney HEK 293T cells

3.1. Small-scale transfection

Trial transfection has been set up to estimate the best DNA to HEK293T cells ratio. To begin with, a vial of HEK293T cells is rapidly thawed in a water bath at 37°C. The content of the vial is suspended in 4mL Dulbecco's Modified Eagle's Medium DMEM (Thermo Fisher Scientific, USA) supplemented with 10% fetal bovine serum (FBS) and centrifuged at 800 x *g* for 5 minutes. The supernatant is discarded and the pellet is resuspended in 5 mL 10% FBS DMEM. The cells are transferred to a 175cm² flask containing 20 mL 10% FBS and incubated at 37°C in a 5%CO_2 environment. When the cells reach approx. 80% confluence, they are transferred to a 6-well plate. Three days after the passage of the cells, the confluence is approx. 100%.

3.1.1. Transfection

To every well, 2 mL of 2% FBS DMEM is added. To the 1st column, 4 µg of DNA is introduced to each of the two wells – 2 µg of the vector encoding the heavy chain and 2 µg of the vector encoding the light chain. To the 2nd column, 2 µg of DNA is added to each of the two wells – 1 µg of the HC vector and 1 µg of the LC vector. The two wells in the last column are treated as controls and therefore no DNA has been added there. All of the conditions are tested in two repeats. Before the DNA transfection, a mixture of DNA-PEI is made in 1:2 ratio. Thus, for 2 µg of DNA, 4 µg of PEI is used, and for 4 µg of DNA, 8 µg of PEI is used. The transfection cocktail is suspended in 50 µl serum-free DMEM and incubated for 10 minutes before transfecting the cells (Fig.31). The cells are incubated at 37°C for 3 days, and the secreted products are harvested by spinning the media at 6 000

x *g* for 20 minutes. The supernatant is then purified by means of Immobilised Metal-Ion Affinity Chromatography (IMAC).

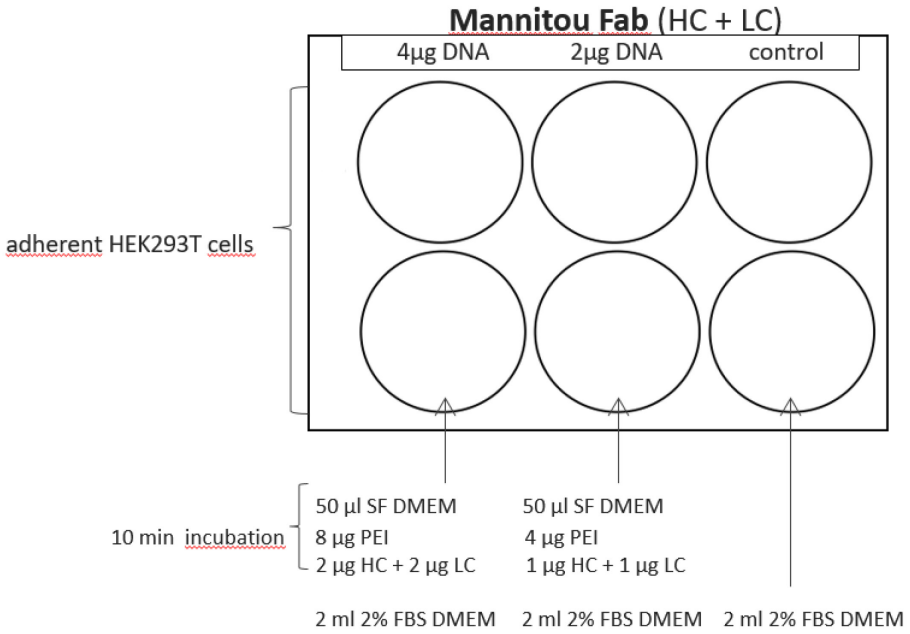


Figure 31. Trail co-transfection of HEK293T cells using two different DNA concentrations to establish the optimal transfection conditions.

3.1.2. SDS-PAGE and Western Blot

After the purification, a verification of the transfection efficacy has been performed by Western Blot analysis. The homogeneity and purity of the Fab have been examined by reducing SDS-PAGE using gradient NuPAGE 4-12% Bis-Tris Protein Gels (Life Technologies, USA). The samples are heated for 10 minutes at 70°C before loading. After running the gels for approx. 1h at 180 V, one of it is stained with GelCode Blue Safe Protein Stain (Thermo Fisher Scientific, USA). Following reducing SDS-PAGE, the other gel is transferred into Towbin buffer (25mM Tris base, 192 mM glycine, 25% methanol) for 15-minute incubation. In the meantime, Immun-Blot PVDF Membrane for Protein Blocking (Bio-Rad, USA) is activated by soaking the membrane in methanol for 2-3 seconds and then incubating in Towbin buffer for 5 minutes. The blotting pads (VWR, USA) also requires soaking in Towbin buffer. The layers are then assembled in the following order: blotting-pad, membrane, gel, blotting-pad. The assembly is run at 15V for 30 minutes. Next, the membrane is washed with 1x PBS and incubated in the blocking buffer (0.1% Tween-20, 2.5% milk, 1xPBS) for 1 hour at room temperature on a rocking platform. The primary monoclonal mouse Penta-His Antibody (QIAGEN, Germany) is

diluted 1:3000 in the blocking buffer and poured on the membrane for overnight incubation at 4°C on a rocking platform. On the following day, the membrane is briefly rinsed with two changes of the washing buffer (0.1% Tween-20, 1xPBS) and then washed twice for 5 minutes with the washing buffer. The secondary goat anti-mouse polyclonal antibody Peroxidase AffiniPure Goat Anti-Mouse IgG, Fcγ fragment specific (Jackson ImmunoResearch Laboratories, Inc., USA) is diluted 1:2000 in the blocking buffer and transferred to the membrane for 1-hour incubation at room temperature on a rocking plate. Afterwards, the membrane is briefly rinsed with two changes of the washing buffer and then washed twice with the washing buffer for 5 minutes. The Amersham ECL Prime Western Blotting Detection Reagent (GE, USA) is left to equilibrate at room temperature for 20 minutes. The detection solutions A (luminol) and B (peroxide) are mixed in 1:1 ratio. The washed membrane is placed protein-side up and the detection reagent is added. The membrane, completely covered with the detection reagent, is incubated for 5 minutes at room temperature.

3.2. Big-scale transfection

3.2.1. Medium composition

To every 500 mL bottle of HyClone DMEM High Glucose (GE Healthcare Life Sciences, UK) medium, 5mL of MEM non-essential amino acids NEAA (1:100) (Gibco Life Technologies, USA) and 5 mL of HyClone L-Glutamine 200mM (1:100) (GE Healthcare Life Sciences, UK) is added. Glutamine is an alternative energy source that supports the rapid division and growth of cells having high energy demands and synthesising large amounts of proteins. Adding supplements of non-essential amino acids, including L-alanine, L-asparagine, L-aspartic acid, L-glycine, L-serine, L-proline, and L-glutamic acid, both stimulates the growth and prolongs the viability of the cells in culture.

3.2.2. Cell progression

Used media are discarded from each 175 cm^2 flask and the cells are washed with 5 mL PBS. 5 mL trypsin solution is added to induce cell detachment from the surface. 15 mL of DMEM supplemented with 10% FBS is added to stop the reaction after 5 minutes. 5 mL of the cell solution is then transferred to a new 175 cm^2 flask and 20 mL of 10% FBS DMEM is added.

3.2.3. Transfection

The content of every 175 cm^2 Nunc EasyFlask (Thermo Fisher Scientific, USA) is transferred to a 2125 cm^2 CELLMASTER (Greiner Bio-One, Austria) expanded-surface polystyrene roller bottle. 200 mL of 10% FBS DMEM is poured into all bottles. The bottles are incubated at 37°C in a roller bottle incubator. After three-day culture, the old media are replaced with 200 mL of 2% FBS DMEM. In one tube, 25 mL of serum-free DMEM and 500 µg of DNA (250 µg of the vector encoding the heavy chain and 250 µg of the vector encoding the light chain) are mixed. In another tube, 25 mL of serum-free DMEM and 1 mg of PEI are mixed. To ensure proper transfection efficiency, the PEI solution is slowly poured into the DNA solution and incubated for 10 minutes at room temperature, to allow the formation of the DNA-PEI complex. The 50 mL transfection cocktail is added to each roller bottle and the cells are placed back in the incubator at 37°C. To serve as a negative control, one bottle is left non-transfected with the DNA-PEI complex. The cells are grown for 7 days until the medium starts turning yellow. The cells are monitored for any unusual behaviour, e.g. cells peeling off, medium changing colour etc. The secreted products are harvested by centrifuging the collected media at 6 000 x *g* for 20 minutes. The supernatant is filtered through a 0.45 µm bottle top membrane, diluted threefold with PBS and stored at 4°C until purification.

3.2.4. Immobilised Metal-Ion Affinity Chromatography (IMAC)

The IMAC has been performed using Chelating Sepharose Fast Flow (GE Healthcare, USA) with Ni^{2+} which selectively retains proteins with exposed histidine. Chelating Sepharose Fast Flow consists of iminodiacetic acid groups coupled to Sepharose 6 Fast Flow by stable ether linkages via a 7-atom spacer. The beads are built of highly cross-linked agarose and their size ranges from 45 to 165 µm. A 0.2 M solution of the desired metal ion (Ni^{2+}/Co^{2+}) is prepared in distilled water. The column is washed with 2 column volumes (CV) of distilled water. Approx. 0.2 CV of the metal ion solution is applied and the column is washed again with at least 5 CV of distilled water to remove the excess of metal ions. The column is equilibrated with 2 CV of the chosen binding buffer (300 mM NaCl, 50 mM NaH_2PO_4, pH=8.0). The prepared resin is then added to the supernatant, previously filtered through a 0.45 µ bottle top filter and diluted x4 in PBS. After incubating at 4°C while gently stirring for 2-3 hours, the mix is poured into a chromatography column and the resin is allowed to sink to the bottom. Once the whole

liquid passed through the column, the resin is washed with 50 mL of the binding buffer. Next, two consecutive washes with ascending concentrations of imidazole (10 mM and then 30 mM) are performed to eliminate as many impurities as possible. Eventually, approx. 5mL of the elution buffer (300 mM NaCl, 50 mM NaH_2PO_4, 250mM imidazole pH=7.0) is added and left on the column. After a 15-minute incubation, the elution fractions are collected and their absorbance is measured.

3.2.5. Size-Exclusion Chromatography (SEC)

Superose 6 Increase 10/300 GL column (GE Healthcare, USA) with 10 µm particle size has been used in this final step of Mannitou Fab purification employing size-exclusion chromatography. The procedure has been executed on the ÄKTA Pure chromatography system (GE Healthcare, USA) under 0.3 mL/min flow. A buffer consisting of 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and 150 mM NaCl, pH=7.4 serves as a running buffer.

The Fab sequence determination has been financed by S. Savvides. The sequence has been designed by S. Robakiewicz and J. Bouckaert. The cloning has been performed by S. Robakiewicz in the lab of S. Savvides with the help of K. Verstraete and A. Dansercoer. The Fab expression and purification have been done by S. Robakiewicz in the labs of J. Jimenez-Barbero and N. Abrescia with the help of S. Delgado and D. Charro.

RESULTS

To determine the approximate end of the CH1 (1st domain of the heavy-chain constant region), and therefore the optimal length of the construct, the Mannitou Fab model has been superimposed onto an already unravelled IgG1 Fab structure (PDB: 5L9D) (Fig.32). The superimposition revealed that the heavy chain should be elongated with approx. 7 amino acids. The **VPIPAVA** motif has been taken from the beginning of the CH2 domain sequence of "*Ig mu chain C region (allele b) – mouse*" (UniProtKB: locus A24976).

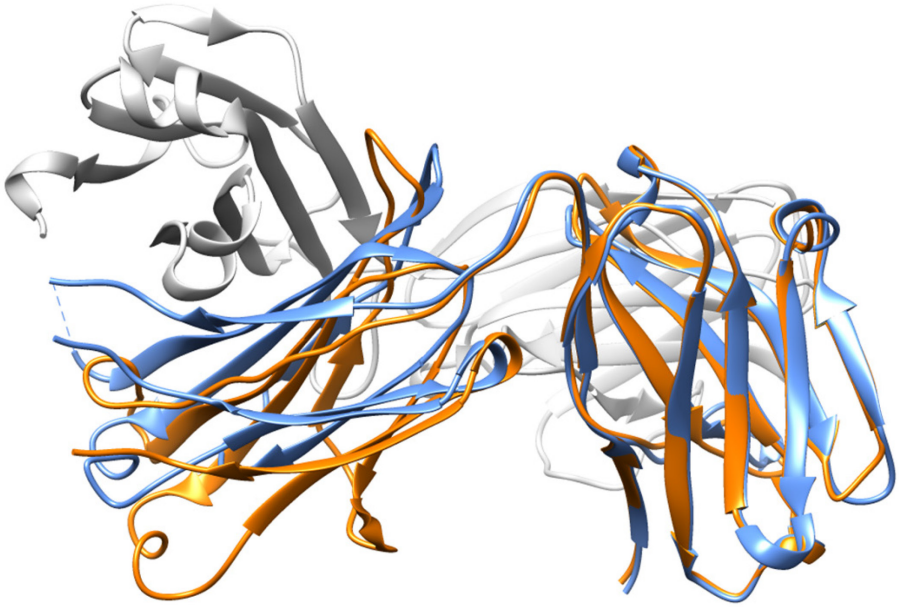


Figure 32. A model of Manniotou Fab HC elongated with 7 additional amino acids (orange), superimposed on mouse IgG1 HC (blue) to determine the expected end of the CH1 domain for Mannitou Fab. The superimposition has been visualised using Chimera 1.12 software.

Following the generation of the constructs, the genes of interest have been amplified by PCR and agarose gel electrophoresis has been performed to confirm their presence and size (Fig.33).

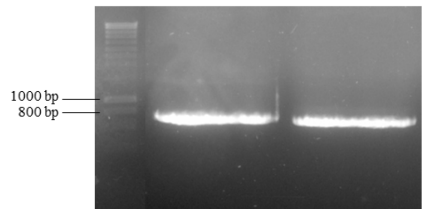


Figure 33. The DNA fragments encoding the HC and LC amplified by PCR and visualised on an agarose gel. The size of the inserts is the following: HC (left) – 702 bp, LC (right) – 663bp.

Subsequent ligation of the gene inserts into “empty” pHL-sec plasmids results in obtaining the desired expression vectors (Fig.34), ready for cell transfection.

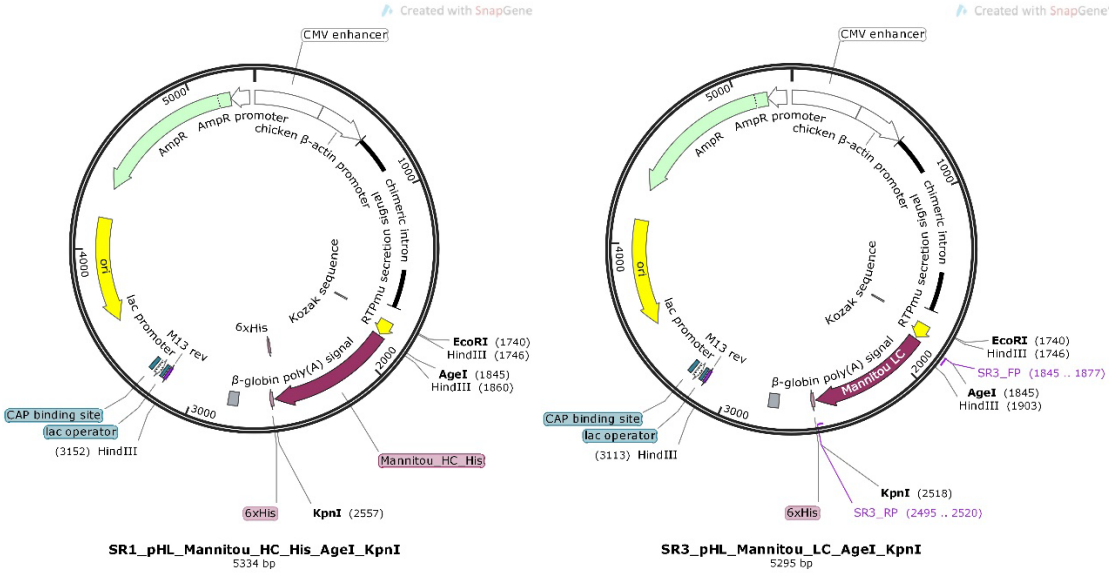


Figure 34. pHL-sec vectors with the gene insert encoding the sequences of the heavy (left) and light chains (right). The vector maps have been generated with SnapGene 2.3.2 software.

To ensure sufficient quantity of the desired vectors, their amplification is effectuated in competent DH5α *E. coli* cells that are capable of taking up exogenous genetic material. The transformed bacteria are allowed to grow overnight on agar plates with ampicillin as the selective antibiotic. As the pHL-sec vectors contain the ampicillin resistance gene, only the cells with the appropriate plasmids should show growth (Fig.35). On the day following the transformation, to verify if the appearing colonies contain the vectors of interest, colony PCR has been performed. The PCR products visualised on agarose gel confirm the presence of the right plasmids in the selected colonies (Fig.36).

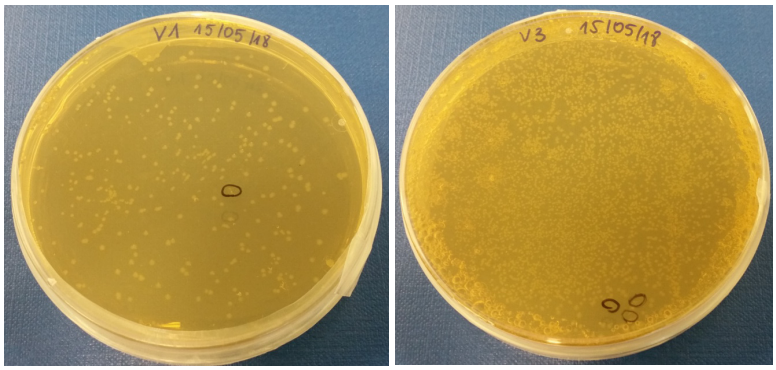


Figure 35. DH5α colonies grown on ampicillin- containing agar plates (40 µl ampicillin /100 µl LB medium) after transformation with vectors encoding the HC (left) and LC (right).

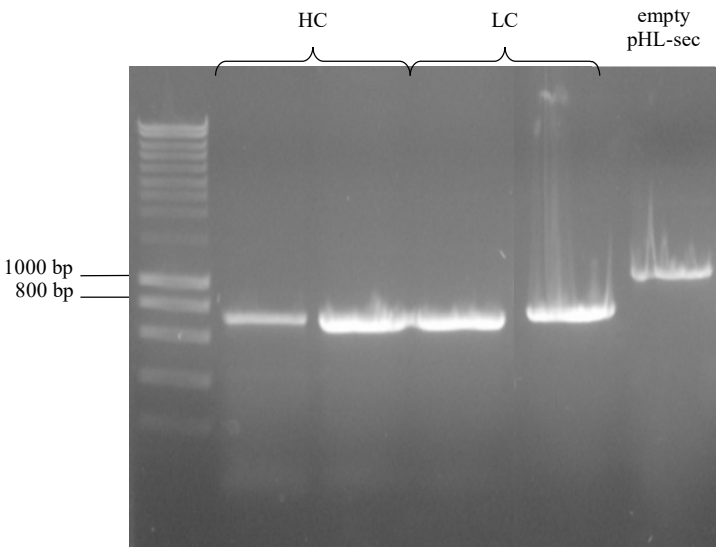


Figure 36. Plasmid inserts amplified by colony PCR and visualised on an agarose gel.

In order to verify the sequences of the purified plasmids, control digestion has been set up. Meanwhile, a simulation of the enzymatic activity of Hind III and Pvu II is performed. The DNA fragments of the digested plasmids visualised on agarose gel correspond to the predicted patterns, which proves that both transfection vectors contain the correctly ligated inserts (Fig.37). Furthermore, to confirm the authenticity of the amplified gene sequences, the resulting clones have been sequenced by ACGT Inc DNA Sequencing Services. The sequencing results confirm the previous findings, as the vectors contain the desired DNA inserts encoding the heavy and light chains of Mannitou Fab.

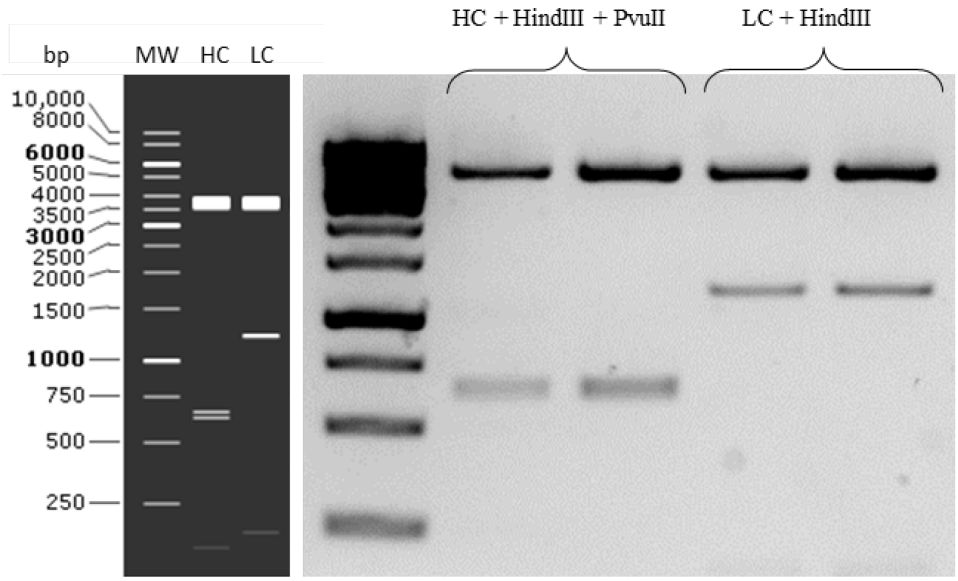


Figure 37. Left: Simulation of the enzymatic activity of Hind III and Pvu II. Lane 1: HC digested with Hind III and Pvu II. Lane 2: LC digested with Hind III; Right: The actual digestion products visualised on an agarose gel. Lane 1&2: HC digested with Hind III and Pvu II. Lane 3&4: LC digested with Hind III.

Trial transfection has been set up to estimate the best DNA to HEK293T cells ratio. The results of the Western Blot visualising the products of the small-scale transfection clearly indicate a presence of a His-tagged protein around the expected size (~25 kDa) for each of the preparations of Mannitou Fab (Fig.38). It is assumed that the appearing bands represent the His_6-tagged heavy chain, since the control (lane 11), containing only the supernatant from non-transfected cells, does not show any signal that would result from the binding to the primary, and subsequently to the secondary antibody.

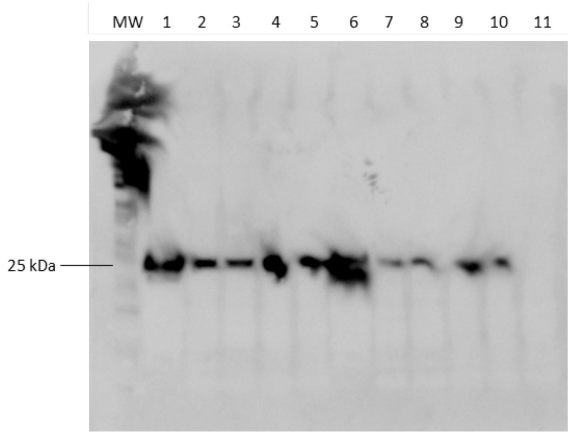


Figure 38. Western Blot analysis after reducing SDS-PAGE showing Mannitou Fab heavy chain. MW: molecular-weight size marker; 1. cell pellet from 2 µg DNA transfection; 2. supernatant from 2 µg DNA transfection (I); 3. supernatant from 2 µg DNA transfection (II); 4. beads supernatant from 2 µg DNA transfection (I); 5. beads supernatant from 2 µg DNA transfection (II); 6. cell pellet from 4 µg DNA transfection; 7. supernatant from 4 µg DNA transfection (I); 8. supernatant from 4 µg DNA transfection (II); 9. beads supernatant from 4 µg DNA transfection (I); 10. beads supernatant from 4 µg DNA transfection (I); 11. control: supernatant from non-transfected cells. The image has been taken by CCD camera Image Quant™ LAS 4000 (GE, USA).

Following the big-scale transfection, Mannitou Fab has been subjected to two-step purification procedure employing both IMAC and SEC techniques. IMAC exploits the molecule's affinity for chelated metal ions. Histidine residues present at the N-terminus of Fab heavy chain form complexes with transition metal ions, such as Co^{2+}, Cu^{2+}, Fe^{3+}, Ni^{2+} and Zn^{2+}. The strength of binding is affected by the buffer's pH and the metal ion selected. Once the purification is completed, SDS-PAGE analysis under reducing conditions is run to check the presence of the anticipated protein. Already after the IMAC, two distinguishable bands appear at the expected size (~25 kDa), suggesting the existence of both heavy and light chains (Fig. 39A). The transformed HEK293Tcells have been cultured for another seven days (14 day-culture in total), to find out whether higher quantities of Mannitou Fab can be generated over time. Once the IMAC finishes and the results are visualised by reducing SDS-PAGE, it becomes clear that the protein production

significantly drops after subsequent 7 day-culture. This is probably due to the fact that over time the cells lose the DNA vector they have been transfected with (Fig. 39B).

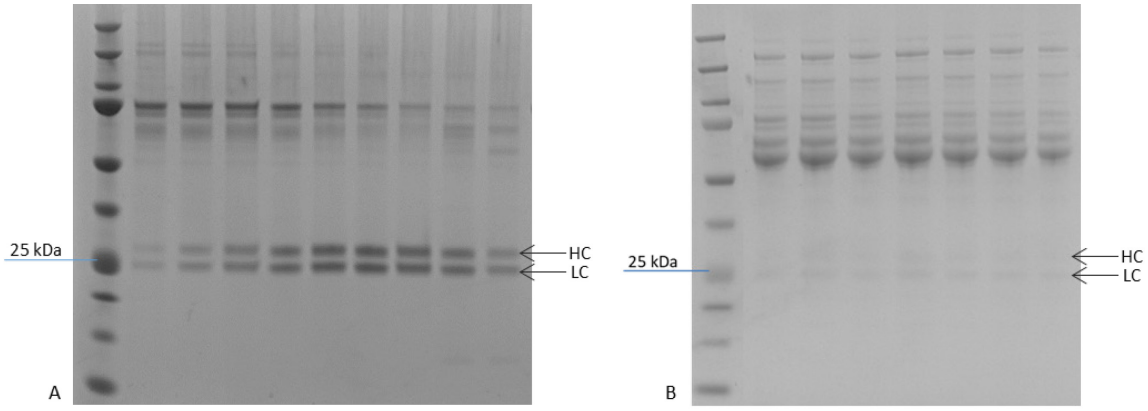


Figure 39. SDS-PAGE results visualising the fractions collected after the IMAC purification. A) Mannitou Fab samples purified using Ni^{2+} after 7-day culture. B) Mannitou Fab samples purified using Ni^{2+} after 14-day culture.

The size-exclusion procedure is the final step of separating the Fab of interest from the other proteins present in the culture medium, i.a. BSA, proteins produced by the cells. Size-exclusion is a chromatographic method in which molecules in solution are separated by their size and molecular weight. When an aqueous solution is used to transport the sample through the column, the technique is known as gel-filtration chromatography. Various columns packed with highly cross-linked, porous beads composed of dextran (Sephadex, Superdex), agarose (Sepharose, Superose), or polyacrylamide (Sephacryl, BioGel P) are available, depending on the nature of the purified protein. In this study, an agarose column has been used, as Mannitou antibody has a high affinity to dextran which results in precipitation of the protein on the column. The chromatographic elution profiles suggest that a well-folded protein has assembled into a Fab, as the peak appears at the expected elution volume, following residual serum albumin (Fig. 40A). After successfully separating Mannitou Fab, SDS-PAGE assay under reducing conditions is run in order to confirm the presence of the anticipated protein. At this stage, mainly the two bands standing for the heavy and light chain distinctly appear on the gel (Fig. 40B). The theoretical molecular weights have been calculated by the ProtParam tool on the ExPASy Server (Gasteiger et al., 2005) and are predicted to be the following: Mannitou Fab MW = 49983, HC MW = 25686, LC MW= 24315. Due to the attached His_6-tag, the heavy chain is slightly larger than the light chain, and therefore two separate bands can be distinguished on the polyacrylamide gel (Fig. 40B).

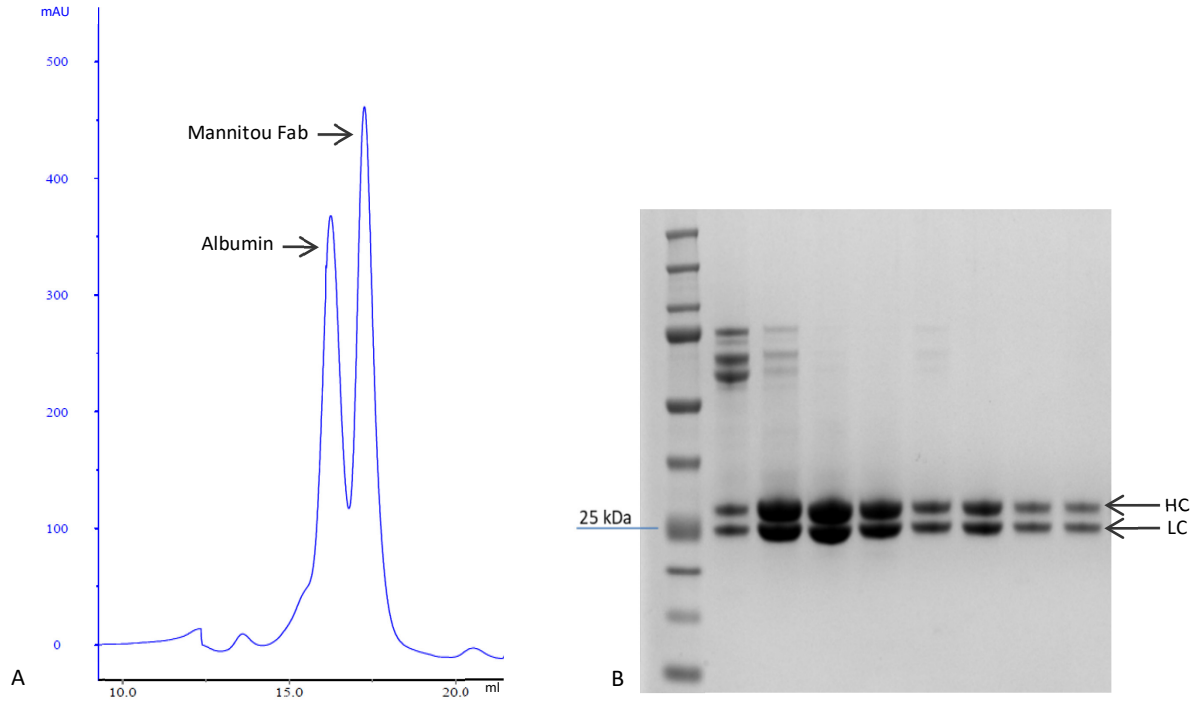


Figure 40. Size-exclusion chromatography of Mannitou Fab. A) Chromatography elution profile of Mannitou Fab. The IMAC-purified antibody (500 µl, 3 mg/mL) has been further purified under 0.3 mL/min flow, using 20 mM HEPES + 150 mM NaCl buffer, pH=7.4, on a Superose 6 Increase 10/300 GL column. B) SDS-PAGE results of the fractions collected from the gel filtration. Both HC and LC are present at the expected height.

The excised bands have been sent for final validation by sequencing using MALDI-TOF MS Fingerprinting. The results reveal the presence of Mannitou Fab heavy chain as well as the light chain in the purified samples (Fig. 41), confirming the proper expression and the authenticity of the antibody.

Accession	Description	ΣCoverage	Σ# Proteins	Σ# Unique Peptides	Σ# Peptides	Σ# PSMs	Score A(2,4)	Coverage A(2,4)	# Peptides A(2,4)	# PSM A(2,4)	# AAs	MW [kDa]	calc. pI
ManHC	Mannitou IgM Fab Heavy OS=XXxxx GN=MannLC PE=2 SV=1 -	46.75	1	8	8	55	1657.95	46.75	8	55	231	25.7	7.81
ManLC	Mannitou IgM Fab Light OS=XXxxx GN=MannLC PE=2 SV=1 - [	43.18	1	5	5	7	251.81	43.18	5	7	220	24.3	7.74
P0AA27	Thioredoxin-1 OS=Escherichia coli O157:H7 GN=trxA PE=1 SV	25.69	1	2	2	2	129.04	25.69	2	2	109	11.8	4.88
B7NGA6	RNA-binding protein Hfq OS=Escherichia coli O17:K52:H18 (str	15.69	2	1	1	7	103.22	15.69	1	7	102	11.2	7.65
F2QTN6	Uncharacterized protein OS=Komagataella phaffii (strain ATCC	15.67	1	2	2	2	109.17	15.67	2	2	134	14.7	4.92
K6ZY70	Uncharacterized protein (Fragment) OS=Phleum pratense OX=1	12.50	6	1	1	1	46.57	12.50	1	1	80	8.9	4.12
F2QZP3	Thioredoxin OS=Komagataella phaffii (strain ATCC 76273 / CB	11.76	1	1	1	1	63.58	11.76	1	1	102	11.3	4.93
X2GM95	Serum albumin (Fragment) OS=Cervus nippon PE=2 SV=1 - [X	11.32	1	7	7	7	238.89	11.32	7	7	583	66.1	5.67
F2QS01	40S ribosomal protein S17-A OS=Komagataella phaffii (strain A	8.82	1	1	1	1	35.58	8.82	1	1	136	15.7	10.30
F2QY66	Superoxide dismutase [Cu-Zn] OS=Komagataella phaffii (strain	8.44	1	1	1	1	61.52	8.44	1	1	154	15.7	6.40
P0A6F5	60 kDa chaperonin OS=Escherichia coli (strain K12) GN=groL P	4.74	2	1	1	1	42.55	4.74	1	1	548	57.3	4.94
XXXX	Pseudomonas prot OS=Sulf PE=1 SV=1 - [DAV_DAV]	4.39	2	1	1	6	124.50	4.39	1	6	569	60.5	5.08
F2R087	Elongation factor EF-1 gamma subunit OS=Komagataella phaffi	4.16	1	1	1	1	41.66	4.16	1	1	409	47.1	6.48
F2QVV6	N-acetyltransferase eco OS=Komagataella phaffii (strain ATCC 7	3.65	1	1	1	1	28.86	3.65	1	1	301	32.2	5.06

Accession	Description	ΣCoverage	Σ# Proteins	Σ# Unique Peptides	Σ# Peptides	Σ# PSMs	Score A(2,4)	Coverage A(2,4)	# Peptides A(2,4)	# PSM A(2,4)	# AAs	MW [kDa]	calc. pI
ManLC	Mannitou IgM Fab Light OS=XXxxx GN=MannLC PE=2 SV=1	52.27	1	6	6	14	572.67	52.27	6	14	220	24.3	7.74
ManHC	Mannitou IgM Fab Heavy OS=XXxxx GN=MannLC PE=2 SV=	30.30	1	5	5	5	264.50	30.30	5	5	231	25.7	7.81
K6Z7K6	GTP-binding nuclear protein (Fragment) OS=Phleum pratens	9.17	10	1	1	1	39.06	9.17	1	1	120	13.6	7.71
C4R707	Histone H2A OS=Komagataella pastoris (strain GS115 / ATC	6.82	2	1	1	1	58.99	6.82	1	1	132	14.3	10.45
F2QR99	Peroxiredoxin (Alkyl hydroperoxide reductase subunit C) OS=	5.64	2	1	1	1	44.53	5.64	1	1	195	21.4	4.94
X2GM95	Serum albumin (Fragment) OS=Cervus nippon PE=2 SV=1 -	5.15	1	3	3	3	85.37	5.15	3	3	583	66.1	5.67
F2QRV5	Uncharacterized leucine-rich repeat-containing protein C926.(	1.24	1	1	1	1	42.77	1.24	1	1	724	80.6	8.79

Figure 41. MALDI-TOF MS Fingerprinting results. Top: Mannitou Fab heavy chain sample. Bottom: Mannitou Fab light chain sample.

DISCUSSION

Antibodies have been selected during evolution for their ability to bind with high affinity and specificity to a wide range of molecules. Because of their stability, they seemed ideal as targeting agents, until it became clear that they might cause serious problems when administered as therapeutics. The first mAbs were of murine origin and upon injection into humans, they were recognised as foreign agents, leading to their elimination by the patient's immune system. Moreover, to be effective, antibodies often need to interact with certain elements of the immune system, such as receptors displayed on effector cells or the complement cascade. Because of their murine nature, these early antibodies did not interact properly with some components of the human immune system and their biological efficacy was severely restricted. In the early 90s, developments in molecular biology made it possible to clone the mAb genes of interest from the parent hybridoma into eukaryotic expression vectors and re-express them i.a. in the format of a Fab. In this way, recombinant versions of any mAb could be obtained from various cell lines in a reproducible fashion. This solved production issues caused by the instability of many hybridoma lines. Cloning antibody genes was the first step towards the modification of antibodies by subcloning, random or directed mutagenesis and molecular evolution procedures, which made it possible to optimise recombinant antibodies generation and introduced the age of antibody engineering (Chames et al., 2009).

The smaller size of antibody fragments enables their penetration into tissues otherwise inaccessible to full-length mAbs. As previously noted, antibody fragments may prove easier and less costly to manufacture, due to relatively small size, which permits the use of both prokaryotic and eukaryotic expression systems. However, they lack the Fc domain that both stabilises full-length mAbs and allows FcR-mediated recycling. As a consequence, antibody fragments are rapidly degraded in humans and have short circulating half-lives. Several effective strategies have been developed to extend their serum half-life, one of them being the chemical addition of polyethylene glycol residues, which considerably increase the size of the fragments. Alternatively, the fusion of recombinant antibody fragments to human serum albumin (HSA) can be used to increase the serum half-life without affecting the binding and activity of the fragments, unlike PEGylation. For most therapeutic applications, a longer half-life is desirable as it decreases the need

for repetitive injections of the molecule to achieve a therapeutically relevant serum concentration (Nelson 2010).

Except in the case of haematological malignancies and diseases, therapeutic antibodies need to penetrate tissues and the extracellular matrix to reach the target cells. Tissue penetration, especially in the case of solid tumours, is a crucial parameter, most of the time severely limiting the overall efficiency of the treatment. Tumours are characterised by heterogeneous and tortuous vasculature, high interstitial fluid pressure and high viscosity of the blood supply. Consequently, mAbs must diffuse against this pressure gradient to penetrate tumours. The major determinant of the speed of diffusion is the molecular size. The rate of diffusion is approximately inversely proportional to the cube root of molecular weight. Consequently, large macromolecules such as mAbs diffuse poorly through larger tumour masses. As a result, solid tumours are more difficult to treat by mAb therapy. mAbs directed against tumour-specific antigens largely remain in the blood and no more than 20% of the administered dose typically interacts with the tumour itself (Beckman et al., 2007; Chames et al., 2009). Strikingly, among the more than 100 mAbs designated as drugs, only fifteen are targeting solid tumours. This clearly reflects the current limitations of mAb treatment considering that over 85% of human cancers are solid tumours (Chiavenna et al., 2017). Because of their reduced size, antibody fragments, such as Fab, usually penetrate tumours more rapidly and efficiently than full-length mAbs. Thanks to their conjugation to either PEG or HAS, the overall tumour uptake of a Fab is not decreasing over a short time.

Expression of the cloned heavy-chain Fd and light-chain genes in *Escherichia coli* offers a simple, rapid, and cheap way of synthesising a recombinant Fab. Nevertheless, successful expression and secretion of correctly folded functional forms of many antibodies from these cells are affected by the variable region sequence. By contrast, mammalian cells contain the appropriate chaperones that enable the proper expression of Fabs, irrespective of their variable region sequence. As most recombinant antibodies are complex molecules that need to be post-translationally modified for full bioactivity, mammalian cells have become the main expression host, as such modifications are only accurately performed by eukaryotic based systems. Pharmaceutical companies routinely create recombinant cell lines that constitutively produce recombinant proteins for extended cultivation times. The stable gene expression is achieved through the insertion of the recombinant gene(s) into the host genome. However, the establishment of a stable

mammalian cell line is a time-consuming and costly process. Faster approaches to recombinant antibody expression are often preferred in research practice. To overcome these difficulties, transient gene expression that can generate Fabs suitable for functional and structural studies serves as a reasonable and convenient alternative. In contrast to stable gene expression, the transient method involves short-term protein production, typically for up to 10 days post-transfection, in the absence of genetic selection of the plasmid DNA. The transient gene expression provides for the successful assembly of the membrane, intracellular, and secreted proteins. One of the few disadvantages of this technology is the fact that new cells have to be transfected for every batch of protein required (Hartley 2012; Wurm 2004).

Different expression vectors, both viral and non-viral, have been designed for transferring the foreign genes into mammalian cells. Fundamental elements of a typical non-viral transient gene expression (TGE) vector include: (1) a constitutive or inducible promoter; (2) a transcription terminator; and (3) a prokaryotic cassette with a replication origin and a selection marker for vector production in bacteria. Most of the commonly used strong constitutive promoters are from viral genomes, including the human cytomegalovirus (HCMV) immediate early promoter, the SV40 early promoter, and Rous sarcoma virus long terminal repeat promoter. Cellular promoters, such as that of the human EF-1α gene, have been shown to be similar in strength to the HCMV promoter. Although most cDNAs can be expressed at high levels from a strong constitutive promoter, insertion of an intron between the promoter and the 5'-end of the cDNA has been shown to increase the recombinant protein expression levels. Mammalian cells expressing the Epstein-Barr virus nuclear antigen 1 (EBNA1) or the SV40 large-T antigen, support the episomal amplification of vectors carrying the EBV or the SV40 origin of replication, respectively. This strategy improves the recombinant protein production levels by increasing the plasmid copy number in the transfected cell population. Semliki Forest virus (SFV) is the most widely used viral expression vector for TGE using mammalian cells. Baculovirus vectors have also been applied in recombinant protein production in both insect and mammalian cells. The disadvantage of using insect cells is that they are unable to produce fully active recombinant mammalian glycoproteins due to the absence of sialylation. Therefore, efforts have been made to engineer the baculovirus expression system that allows the production of humanised recombinant glycoproteins in insect cells. For the expression of recombinant proteins in

mammalian cells, baculovirus vectors have numerous advantages, including the absence of viral replication, lack of cytotoxicity, technical simplicity of vector production, and the large DNA insert capacity (up to 38 kb) allowing the delivery of multiple genes from a single vector. Vaccinia virus vectors have also been used to produce recombinant proteins, like the nerve growth factor, Factor VIII, and human pro-thrombin. The infectious nature of viruses is a major concern, thus to reduce safety risks, less infectious strains such as the modified vaccinia virus Ankara (MVA) have been introduced (Baldi et al., 2007).

Several cell lines have been employed for transient expression of recombinant proteins, but only a few of them present the advantages of high transfectability with common methods, ease of adaptation to suspension cultivation in serum-free conditions, and cost-effective scalability. Created in 1977 by the transformation of primary human embryonic kidney cells with sheared fragments of adenoviral DNA (Ad5), HEK 293 is one of the most commonly used cell lines for transient transfection that constitutively express the adenovirus E1A and E1B genes. In the attempt to specifically improve TGE through promoting episomal replication of the exogenous plasmid DNA, HEK 293 derivatives have been generated. HEK 293-T cells used in this study constitutively express from stably transfected gene the SV40 large T antigen. (Graham et al., 1977).

A multitude of physical and chemical techniques have been developed for the transfer of DNA into cultivated mammalian cells. The physical gene transfer methods including microinjection, electroporation and ballistics, bypass the cellular barriers and deliver naked plasmid DNA directly to the nucleus. However, these methods have not been yet successfully adapted to large-scale TGE. The chemical agents used for transfection can be classified into inorganic compounds (e.g. CaPi-DNA coprecipitation), cationic polymers (e.g. polyethylenimines), and cationic lipids. Calcium phosphate (CaPi) is considered one of the most efficient DNA delivery agents, as it allows the transfer of plasmid DNA to approximately 80–90% HEK 293 cells. Although it is inexpensive and does not interfere with protein recovery, its time-sensitive formation of the coprecipitate makes transfection of larger volumes technically challenging. Our approach involves using linear 25 kDa PEI as it is moderately cytotoxic and currently the most widely used DNA delivery agent. Transfections with linear 25 kDa PEI have been the most productive to date with reported IgG yields up to 80 mg/l in HEK 293 cells (Baldi et al., 2007).

The major improvements in TGE have come via the formulation of serum-free media that support high-density cultivation of cells and efficient transfection methods. Currently, serum-free cultures are the standard for growing mammalian cells in suspension. This stems down from the need to minimise animal-derived components in the production processes, both for biosafety reasons and to facilitate the expression and purification of recombinant proteins. The group of Aricescu found that HEK293T cells surprisingly well tolerate a sudden post-transfection serum depletion. Even the complete absence of fetal bovine serum in the medium added after transfection did not appear to significantly reduce the yields of secreted proteins. This eliminates the main source of protein contamination and enables the possibility of setting up crystallisation screens following a highly simplified purification and concentration protocol. It should be noted that residual bovine serum albumin and other proteins, possibly secreted by cells or resulting from dead cells, are still present but in very small amounts (Aricescu et al., 2006).

The HEK293T transient transfection system chosen to obtain Mannitou Fab is probably best-suited for the generation of secreted proteins intended for structural studies. A protein is best expressed in its native cell type under physiological conditions, where a multitude of molecular systems work together for efficient production and quality control at various stages. This includes synthesis and folding, post-translational modifications, and subcellular targeting. Efficient post-translational processing in the secretory pathway, including the endoplasmic reticulum, Golgi apparatus, vesicles travelling in between them, cell membrane, and lysosomes, is vital for the correct glycosylation and folding, and therefore the full bioactivity of such proteins. Currently, almost all therapeutic antibodies are still produced in mammalian cell lines in order to reduce the risk of immunogenicity due to altered, non-mammalian glycosylation patterns (Aricescu et al., 2006). Another significant advantage of the system is the possibility of co-transfection experiments, which provides a convenient opportunity for high-level expression of protein complexes. Two different vectors encoding the heavy and light chain sequences have been introduced into the cells. Even though the heavy and light chains are expressed separately, the Fab assembles within the endoplasmic reticulum prior to secretion from the cells (Kalwy et al., 2006). The heavy and light chain sequences have been cloned into pHL-sec expression vector which supplies a signal sequence and adds a C-terminal hexahistidine tag for detection and purification. A His_6-tag has been

introduced to the heavy chain fragment to enable the purification by means of IMAC. For polymeric structures composed of heavy and light chains, in the absence of the light chain, the heavy chain is retained in the ER and eventually degraded (Klausner et al., 1990). Since the heavy chain can only be secreted in association with the light chain, exclusively a fully-assembled Fab can be detected. Histidine is strongly involved in the co-ordinate bond with metal ions. When the His-tagged Fab is brought into contact with a carrier on which a metal ion is immobilised, the histidine tag chelates the ion and binds to the carrier. As other proteins do not bind to the carrier, or bind very weakly, the His-tagged Fab can be easily eluted by washing the carrier with an appropriate buffer (Porath 1992). Having validated the homogeneity, purity and the expected dimeric assembly of the generated Fab by MALDI-TOF MS, the next goal was to perform molecular recognition studies using Mannitou antibody.

Chapter 2.

Molecular recognition studies using Mannitou antibody

I. Ligand screening with microarrays of defined *N*-glycans

INTRODUCTION

Information contained in the mammalian glycome is decoded by glycan-binding proteins (GBPs) that mediate diverse functions including host-pathogen interactions, cell trafficking, and transmembrane signalling. Glycan-binding proteins are involved in biological processes that require recognition of selected sets of oligosaccharide ligands. Although the information on the biological role of GBPs is rapidly expanding, challenges remain in identifying the carbohydrate sequences within the glycome that are bound by GBPs, as well as characterising their impact on GBP function. Difficulties arise due to the glycome being generated post-translationally through non-template synthesis directed by over 200 glycosyltransferase genes, whose differential expression and combined specificities determine the unique spectrum of carbohydrate structures produced by a given cell. Lack of tools to interrogate the ability of a protein to bind oligosaccharide ligands or for assessing the detailed specificity of a protein already known to bind carbohydrates has severely limited the pace of biological discovery for GBPs. Recent advances in the construction of glycan microarrays that display diverse glycan libraries directly address these limitations by allowing the specificity of a GBP to be assessed for a wide range of potential oligosaccharide ligands simultaneously. The advances have occurred on two fronts: the generation of glycan libraries to be arrayed on the chip and methods for their attachment to the chip surface (Paulson et al., 2006; Taylor et al., 2009).

The utility of a glycan array depends on an appropriate match between the types of glycan structures it contains and the specificity of the GBP being analysed. The ideal array would contain the entire glycome of an organism on a single chip so that any GBP could be assessed. In reality, however, current arrays are limited to displaying libraries of natural and synthetic glycans that can be practically assembled. Unlike proteins and nucleic acids, it is challenging to chemically synthesise oligosaccharides for microarrays. This is due to the fact that some oligosaccharide chains are linear, others are branched, the monosaccharide building blocks are in α or β anomeric configurations, and adjacent monosaccharides are linked via different carbon atoms in their sugar rings. Multiple strategies are employed to generate oligosaccharides for arraying, namely synthetic approaches (chemical or enzymatic syntheses) and isolation from natural sources. Chemical strategies for glycan elongation include traditional solution-phase chemical

synthesis, automated solid-phase synthesis, and one-pot reactivity-based glycosylation. Chemically obtained compounds are pure and homogeneous. Furthermore, the glycans are usually synthesised to contain a reactive linker that enables site-directed immobilisation. The solid-phase synthesis technique has the advantage of avoiding intermediate isolation and purification steps. An automated solid-phase method that includes selective protection and deprotection steps has been introduced and applied to the synthesis of several glucose- and mannose-containing oligosaccharides (Plante et al., 2001). An alternative synthetic approach is the programmable one-pot method, in which an oligosaccharide of interest is generated by sequential addition of building blocks (thioglycosides) that are either fully protected or have one hydroxyl group exposed (Zhang et al., 1999). Apart from chemical synthesis, oligosaccharides can be prepared by enzyme-catalysed reactions. In nature, there are two classes of enzymes involved in oligosaccharide production: glycosyltransferases catalysing the formation of glycosidic bonds, and glycosidases that hydrolyse these bonds. The biggest advantage of enzymatic synthesis of oligosaccharides is that substrates are used in their natural form. No protecting groups are needed to direct the regio- or stereo-specificity of glycosidic bond formation. Another approach for carbohydrate generation is the isolation from natural sources such as cultured pathogens, plant material, or animal tissues. Despite the samples being readily accessible, isolation procedures can be tedious. Even after extensive purification, sample microheterogeneity owing to carbohydrate diversity may remain and must be taken into account during data analysis. Nevertheless, oligosaccharides with reducing termini are ideal for derivatisation so that they can be immobilised. Free reducing oligosaccharides may be isolated from human or animal milk and urine, or they may be in the form of *N*-linked glycoprotein oligosaccharides released by the peptide-*N*-(*N*-acetyl-b-glucosaminyl) asparagine amidase (PNGase F) and endo-b-*N*-acetylglucosaminidase F (Endo F) (Tarentino et al., 1994), or by hydrazinolysis (Takasaki et al., 1982). *O*-linked glycoprotein oligosaccharides may be released by mild alkaline hydrolysis (Chai et al., 1997) or hydrazinolysis (Patel et al., 1993). Various chemical approaches may be used to obtain oligosaccharide fragments from bacterial and plant polysaccharides, including acid or alkaline hydrolysis, acetolysis, and Smith degradation (Feizi et al., 2003; Paulson et al., 2006).

Multiple different platforms have been described for the display of glycans using a range of immobilisation strategies. These include both non-covalent attachment, as well

as various types of covalent linkages to the surface through the oligosaccharides reducing ends. Glycan array formats based on noncovalent attachment include direct immobilisation of unmodified bacterial polysaccharides and proteoglycans on nitrocellulose or oxidised polystyrene (Wang et al., 2002); binding of lipid-conjugated or linked to alkyl chains glycans to nitrocellulose, polystyrene, polyvinylidene fluoride membranes (Fukui et al., 2002), or coated glass slides; association of glycans linked to fluorous tags with fluoroalkylsilane-coated glass slides (Ko et al., 2005); and affinity adsorption of biotinylated glycans to streptavidin-coated surfaces (Leteux et al., 1999). Array formats using glycans with functionalised spacers that react with a complementary activated surface to form a covalent attachment to plastic, gold, or glass include thiol–maleimide (Park et al., 2002), alkyne–azide (Fazio et al., 2002), NHS–amine (Blixt et al., 2004), benzoquinone–cyclopentadiene coupling (Houseman et al., 2002), and aryl-CF3 diazarine–carbene insertion (Angeloni et al., 2005). With so many alternatives for construction of glycan arrays, it is not possible to conclude that one format is superior to the others. However, many practical issues are ultimately considered during the development of large-scale carbohydrate arrays. From the chemistry perspective, the ease of generating a diverse functionalised glycan library and the efficiency of the coupling reaction are major factors, particularly for small μg quantities of glycans obtained from biological sources. The surface to which the glycans are attached is also critical for the subsequent interrogation with labelled GBP, as a low background is essential for specific binding to be detected. The nature of the spacer between the glycan and the surface can also affect the interaction with GBPs (Feizi et al., 2004; Paulson et al., 2006; Taylor et al., 2009).

In general, GBPs have a rather low affinity (K_d from mM to μM) for their glycan ligands and would not be expected to bind with sufficient avidity to withstand the extensive washing steps needed to remove unbound molecules. For this reason, binding to arrays is generally enhanced through mimicking naturally occurring multivalent interactions (engagement of multiple binding sites) which leads to improved array sensitivity. Many GBPs are oligomeric, with each subunit typically having a single carbohydrate-recognition domain, and show multivalent interactions with their glycan ligands. Thus, although the CRD within a GBP may have a particular affinity for a ligand, the multivalency enhances binding through increased avidity and allows ligand cross-linking. Typically, higher valency is achieved through the use of labelled secondary

antibodies (Fig.42). After incubation with the fluorescent reagent, the bound fluorophore is detected and fluorescence intensities can be quantified (Cummings et al., 2017; Mende et al., 2019).

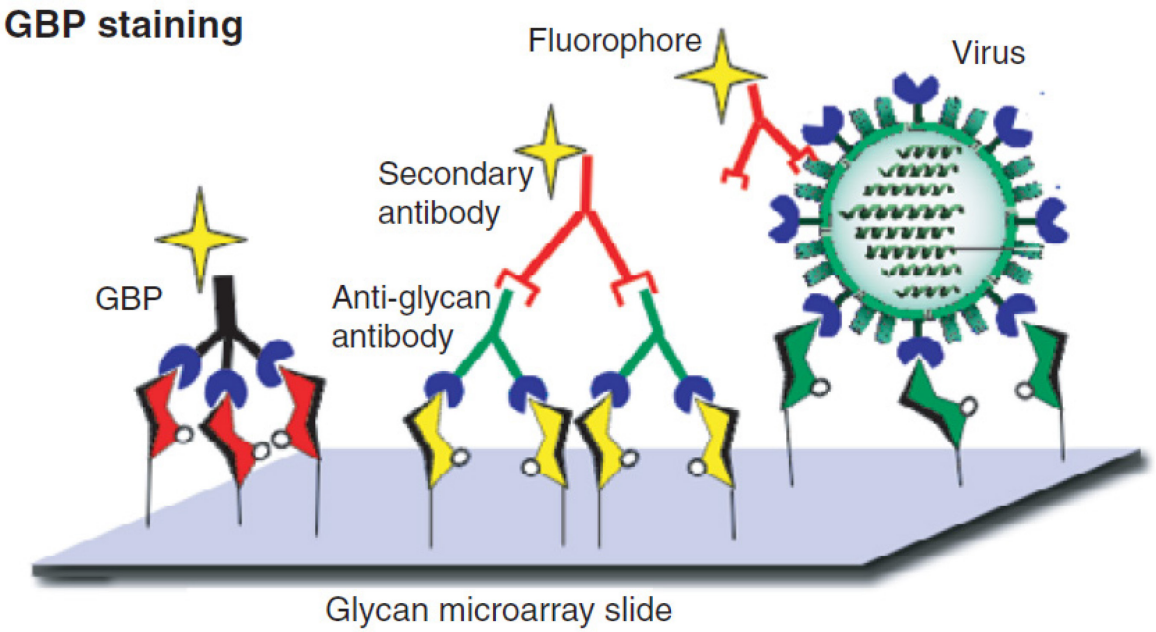


Figure 42. GBP staining. Non-labelled GBPs, antibodies, or pathogens are detected with an overlay of a labelled secondary antibody. Overlaying precomplex GBP with a labelled secondary antibody leads to a higher degree of valency for stable binding. Fluorescently labelled GBPs are detected after staining using a confocal slide scanner or plate reader. Next, the data obtained from the image analysis generate a two-dimensional bar chart to reveal relative signals for each glycan on the array. Adapted from: Paulson et al., 2006.

Since the introduction of the first glycan arrays in 2002, their application for analysis of the specificities of diverse GBPs has rapidly grown. Several array platforms that display diverse glycan libraries have been developed, including The Consortium for Functional Glycomics (CFG) that has produced a mammalian glycan library comprising over 600 synthetic and isolated compounds, with most samples obtained by chemoenzymatic synthesis.

MATERIALS & METHODS

1. Equipment

In this study, NEXTERION Slides H (Schott, Germany) manufactured from high quality, low-fluorescence glass covered by a low-fluorescence coating have been used. The coating consists of a cross-linked, multi-component, thin polymer layer, activated with *N*-hydroxysuccinimide (NHS) esters to provide high covalent immobilisation capacity. The microarray slides have been engineered to exhibit a very low intrinsic non-specific background, without the need for blocking, and high probe loading capacity. The polymer surface has a three-part structure (Fig.43). NHS-ester reactive groups are attached to the cross-linked hydrophilic polymer layer via long, flexible spacers. The terminal amino group of the amino-modified glycans react immediately and irreversibly with the NHS-ester groups to form a covalent amide bond. The flexible spacers tether the immobilised bio-molecules in a quasi-liquid environment that maintains the molecule specificity and chemical conformation. The high accessibility of the tethered biomolecules facilitates interactions with their binding partners in solution. The robust coating matrix is compatible with all common microarray printers and scanners.

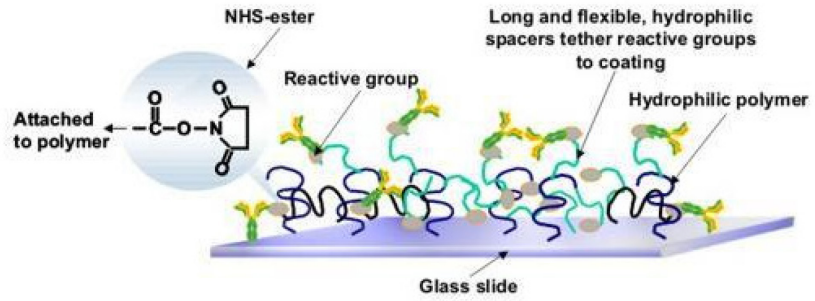


Figure 43. NEXTERION Slide H coating.

2. Microarray preparation and ligand screening

50 µM ligand solutions (1.25 nL, 5 drops, 250 pL drop volume) in 300 mM NaAc, (0.005% Tween-20, pH=8.4) have been spatially arrayed employing a robotic non-contact piezoelectric spotter (SciFLEXARRAYER S11, Scienion, Germany) onto NHS-activated Nexterion Slides H. After printing, the slides are placed in a 75% humidity chamber for 18 hours at 25°C. The remaining NHS groups are quenched with 50 mM solution of ethanolamine in 50 mM sodium borate buffer pH=9.0, for 1 hour. The slides are washed with PBST (PBS with 0.05% Tween-20), PBS, and water, then dried in a slide spinner and stored at -20°C until use. Mannitou mAb has been diluted to 25 µg/mL in PBS (1% BSA, 0.01% Tween-20). The antibody solution (200 µL) is incubated on the microarrays for 1 hour at RT. The slides are then washed with PBST and PBS. Next, they are incubated with Alexa Fluor 555 Goat Anti-Mouse IgM (1:1000) (Thermo Fischer Scientific, USA) in PBS (1% BSA, 0.01% Tween-20) for 1 hour in the dark. The microarrays are washed from the unbound secondary antibody with PBST, PBS, and water, and subsequently dried in a slide spinner. The fluorescence measurements are performed on Agilent G2565BA Microarray Scanner (Agilent Technologies, USA) at 10 µm resolution. The quantification of fluorescence is done using ProScanArray Express software (Perkin Elmer, USA) employing an adaptive circle quantification method from 50 µm (minimum spot diameter) to 300 µm (maximum spot diameter). Average RFU values with local background subtraction of four spots and the standard deviation of the mean are reported using Microsoft Excel and GraphPad Prism software.

The arrays have been prepared and interrogated with the antibody by B. Echeverria, S. Serna, and N. Reichardt.

RESULTS

Microarrays consisting of 135 different glycan structures (Fig.44) have been interrogated with several different preparations of Mannitou mAb at 25 µg/mL (Fig. 45). One of the samples has been tested in two-, four-, and eight-fold dilution (12.5 µg/mL, 6.25 µg/mL, and 3.12 µg/mL) to investigate the effect of concentration on the binding pattern (Fig.46, Fig.S2). The screening reveals that the minimum binding epitope is Manα1-3Manβ1-4GlcNAcβ1-4GlcNAcβ1-N (**G40**). The results clearly show that the binding increases towards the pentasaccharide core of Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ1-N (**G41**). Substitutions in the Manα1-6 branch are well tolerated (**G4, G23, G27, G34, G43, G74, G134**), as well as core Xylβ1-2 (**G2, G7, G9, G10**), and core Fucα1-3 (**G9, G72, G90, G134**), and Fucα1-6 (**G14, G70**) in the reducing end GlcNAc. However, any substitution in the Manα1-3 branch inhibits binding (**G5, G42, G54, G55**), and so does core Fucα1-3 in the distal GlcNAc of the chitobiose core of *N*-glycan (**G83**). This binding pattern is conserved for all the batches and concentrations used (Fig.47).

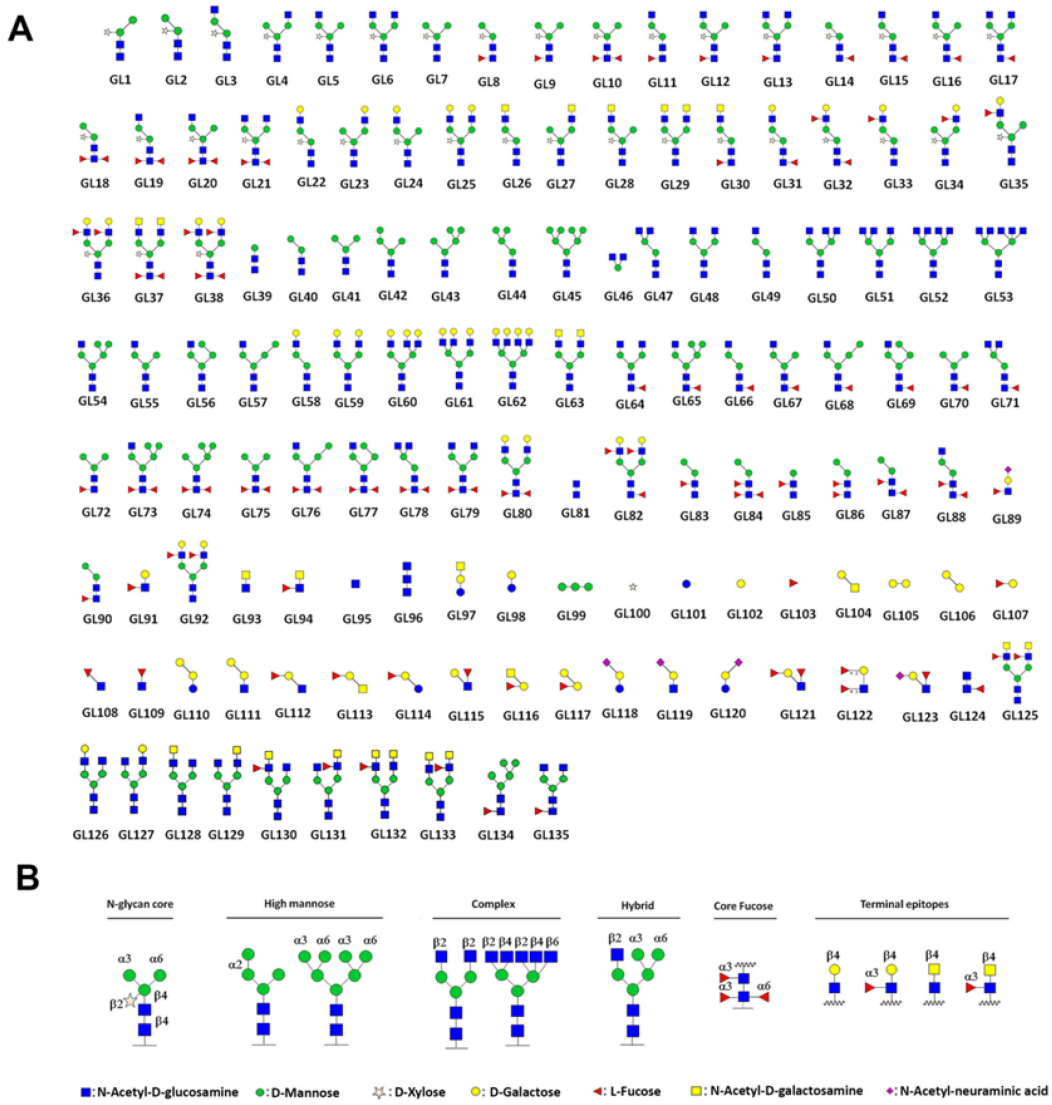


Figure 44. A: *N*-Glycan structures included on the microarrays. B: Glycosidic bond nature of the *N*-glycan structures on the microarrays.

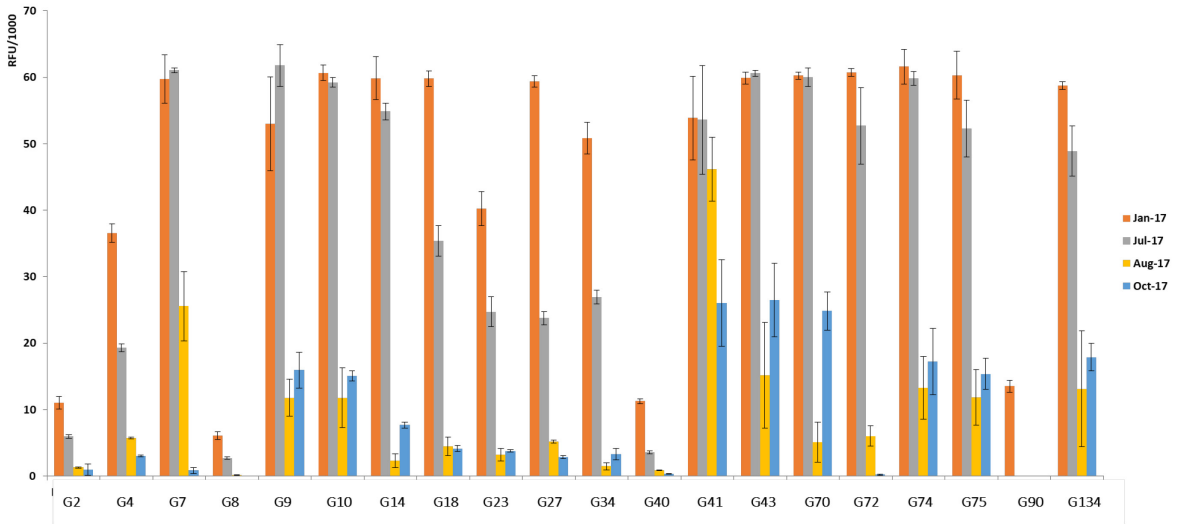


Figure 45. Four different preparations of Mannitou mAb at 25 µg/mL incubated on the microarrays.

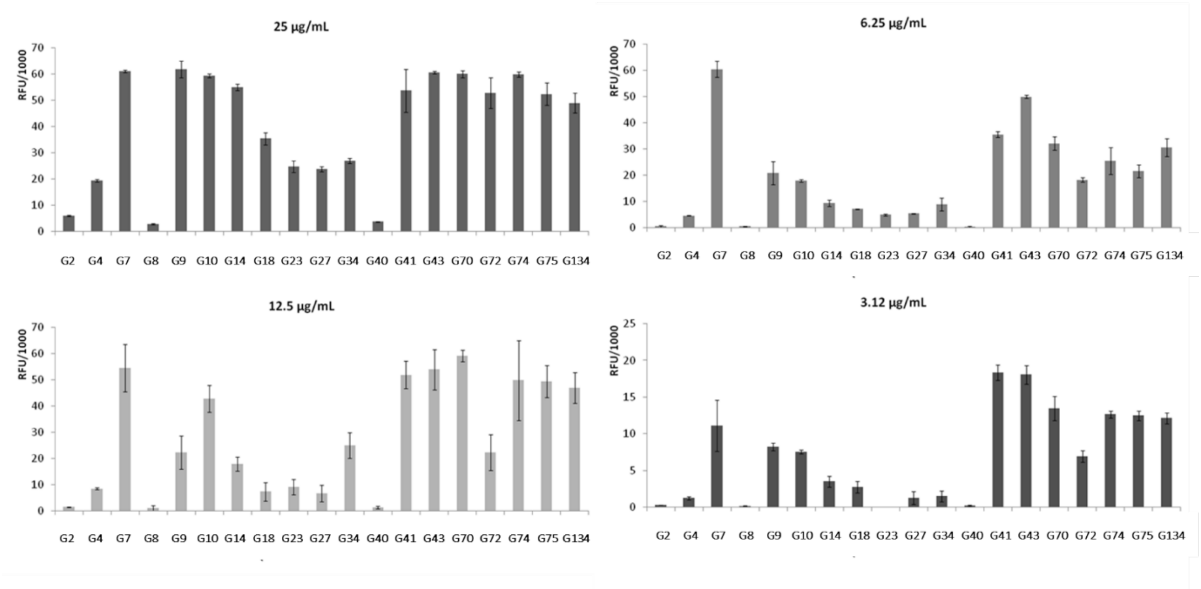


Figure 46. Four different concentrations of one of the Mannitou mAb preparations incubated on the microarrays.

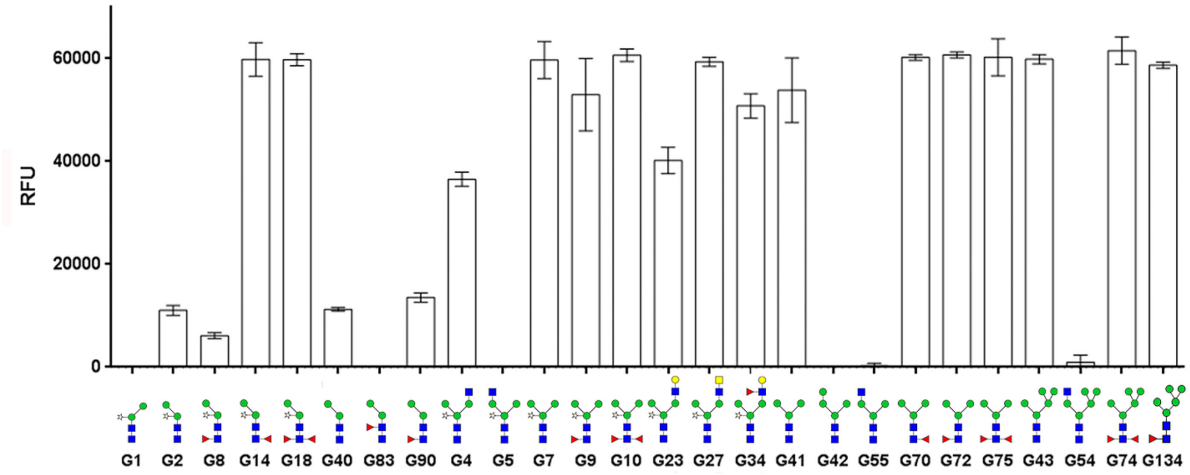


Figure 47. Binding pattern to selected *N*-glycan structures after Mannitou antibody incubation on the microarray (25 µg/mL). Each histogram represents the mean RFU (relative fluorescence unit) values for four spots with the SD (standard deviation) of the mean.

DISCUSSION

Glycan arrays are very well suited for the analysis of antibodies directed against glycans or glycopeptides. Natural antibodies against carbohydrates are found in humans and other mammals. Their origin is a matter of debate, but exposure to microbial carbohydrates is believed to play a major role in their formation. Naturally acquired anti-glycan antibodies vary greatly between individuals. Unusual levels of anti-glycan antibodies in serum are associated with infection, cancer, or autoimmune disease (Muthana et al., 2014). Therefore, anti-glycan antibodies are promising potential biomarkers of pathophysiological condition, and glycan arrays represent a powerful and high-throughput approach to analyse interactions between these antibodies and a broad spectrum of potential carbohydrate antigens. In addition to soluble proteins, whole viruses, bacteria, yeasts, parasites, or mammalian cells can be screened as well if needed (Geissner et al., 2016).

Ideally, any method for characterising protein-carbohydrate interactions would require minimal amounts of material, allow the analysis of many samples, and be compatible with the rather weak binding affinity between the two agents. The microarray format addresses all these issues, as with relatively little antibody quantities parallel screening of many potential ligands can be simultaneously performed in a single experiment. The array allows screening of far more ligands than could be undertaken in any other analysis, including structural studies. Furthermore, features of presentation, such as density and orientation of the oligosaccharides displayed on the surface, can be controlled as to create a high avidity multivalent interaction between naturally oligomeric proteins with multiple binding sites, such as the pentameric Mannitou IgM, and their oligosaccharide antigens (Kiessling et al., 2002). It has to be noted that the affinity for glycans results from a complex balance between entropy costs of immobilising larger saccharides, with resulting favourable interactions, and steric exclusion. Additionally, the chip surfaces are designed to prevent non-specific adsorption of proteins and ensure a homogeneous environment for carbohydrate presentation so that all immobilised glycans have an equal chance of recognition by the interrogating protein (Houseman et al., 2002).

Preferential binding to glycans bearing two GalNAc residues and preferably three Man residues, that are either unsubstituted or substituted only at the α1-6 position has been observed after an interrogation with Mannitou mAb. The fact that the binding gives strong signals with a few structurally related glycans, after which a sharp drop-off in signal is observed, indicates that the association is dependent on the presence of a specific binding epitope. It is very important to point out that glycan array screening at higher antibody concentration reveals binding to the same ligands as at lower concentrations, and this pattern has been reproduced in multiple screens of the array with several different antibody preparations. Glycan array results indicating highly selective binding to a group of glycans can reflect the presence of a pre-formed conformational feature of the ligands, which provides an extended surface for local interactions with the antibody surface at minimal entropy cost associated with glycan immobilisation (Taylor et al., 2009).

These findings are consistent with previously described experiments by Zipser et al. from 2012, where Mannitou antibody was screened against the glycan array 4.0 of The Consortium for Functional Glycomics that comprises 442 mammalian glycans coupled via amino functional groups to amino reactive glass slides. It demonstrated highly restricted specificity for the trimannosyl core of *N*-linked glycans called paucimannose. The affinity of Mannitou was found to be the highest for the $Man_3GlcNAc_2$ structure, which has been subsequently confirmed by the present study. Adding monosaccharides at the non-reducing end of the Manα1-3 branch, led to an almost complete loss of reactivity compared to the initial $Man_3GlcNAc_2$ structure (1% of reactivity), which has been observed again during this screening.

Owing to the complexity and heterogeneity of glycans, the chemical similarity of their monosaccharide subunits and the possibility of many different types of linkages between them, a single analytical method will rarely provide a full reliable compositional, structural, and interaction analysis. Existing glycan array libraries still contain only a small fraction of the carbohydrate diversity found in nature. Therefore, further advances in methods for the synthesis, isolation, and characterisation of glycans are required to expand the currently available glycan libraries. Moreover, glycan presentation on array surfaces can considerably impact the binding interaction, as steric hindrance may prevent binding to some ligands. Lastly, as different formats and standards are used to analyse array information obtained from different laboratories, different array platforms

can yield different results. For these reasons, multiple approaches have been employed to validate glycan array specificity, most commonly involving surface plasmon resonance or isothermal titration calorimetry. Both of these techniques along with STD-NMR have been applied to study the molecular recognition of paucimannose *N*-glycans by Mannitou antibody. SPR defines kinetic parameters associated with the interaction, ITC provides quantitative thermodynamic data, and STD-NMR identifies ligand moieties crucial for the binding. Based on these microarray screening results, $Man_3GlcNAc_2$ (**G41**) and the corresponding α1-6 core fucosylated structure (**G70**) have been selected for further analysis, as $Man_3GlcNAc_2$ manifests one of the strongest interactions with Mannitou, and fucose is an important residue recognised by a multitude of protein receptors. In mammals, core fucose is found exclusively as α1-6-linked to the reducing *N*-acetylglucosamine moiety of the chitobiose core (Kobata 1992; Serna et al., 2011; Tjondro et al., 2019). β1-2-linked xylose and α1-3-linked fucose structures have been excluded from further analyses for being immunogenic glycan motifs from plants and invertebrates but never present in mammals (Brzezicka et al., 2015; Clausen et al., 2017; Paschinger et al., 2009; Tjondro et al., 2019).

Chapter 2.

Molecular recognition studies using Mannitou antibody

II. Kinetic analysis using Surface Plasmon Resonance and Isothermal Titration Calorimetry

INTRODUCTION

Molecular recognition refers to the process in which biological macromolecules interact with each other through non-covalent interactions to form a specific complex. This process has two important defining characteristics, namely specificity and affinity. Virtually all biological processes depend critically on the binding of ligands by specific proteins such as enzymes, antibodies, receptors, lectins, and transport proteins, and the determination of binding affinities for such molecular recognition processes is therefore of central importance.

Protein-ligand binding kinetics describes the process underlying the association between a protein and ligand, particularly focusing on the rate at which these two partners bind to each other. In a simple example, when a protein molecule P and a ligand molecule L with mutual affinity are mixed in a solution, the time-dependent reversible association between them can be formulated as:

$$P + L \rightleftharpoons PL$$

PL – the protein-ligand complex

At equilibrium, the forward binding reaction should be balanced by the reverse unbinding reaction:

$$k_{on}\,[P][L] = k_{off}\,[PL]$$

k_{on}/k_a – association rate constant [1/(M*s)]; concentration-dependent

k_{off}/k_d – dissociation rate constant [1/s]; concentration-independent

The binding affinity is the strength of the interaction between two molecules that bind reversibly. It is translated into physico-chemical terms in the dissociation constant K_d, proportional to the number of occupied binding sites at equilibrium.

$$K_a = k_{on}/k_{off} = [PL]/[P][L] = 1/K_d$$

K_a – equilibrium association/binding constant [1/M]

K_d – equilibrium dissociation constant [M]

Therefore, a fast binding rate accompanied by a slow dissociation rate will give a high association (and respectively low dissociation) constant and, hence, a high binding affinity.

A protein-ligand-solvent system is a thermodynamic system composed of the solute (i.e., the protein and ligand molecules) and the solvent (i.e., liquid water and buffer ions). In such a system, the heat exchange among these substances and how heat transfer is related to various energy changes are dictated by the laws of thermodynamics. As a result, the driving forces that dictate the association between the protein and the ligand are a synthetic result of various interactions and energy exchanges among the protein, ligand, water, and buffer ions. Gibbs free energy is one of the most important thermodynamic quantities for the characterisation of the driving forces. ΔG can be regarded as an approximation of the chemical potential that measures the capacity of a thermodynamic system to do maximum or reversible work at a constant temperature and pressure (isothermal, isobaric). In analogy with any spontaneous process, protein-ligand binding occurs only when the ΔG is negative and the Gibbs free energy of the system is minimised, when the system reaches an equilibrium state at constant pressure and temperature. Because the protein-ligand association extent is determined by the magnitude of the negative ΔG, it can be considered that the change in free energy determines the stability of any given protein-ligand complex, or alternatively, the binding affinity of a ligand to a given protein. It should be noted that the free energy is a state function of the states of a system, and therefore independent of the path leading from one state of the system to another and defined only by the initial and final thermodynamic states.

The standard binding free energy ΔG^o, which refers to the free energy change measured under the conditions of 1 atm pressure, a temperature of 298 K, and the effective reactant (protein and ligand) concentrations of 1 M, is related to the binding affinity constant K_a by the Gibbs relationship. It is apparent that the higher the association constant K_a, the more negative the standard free energy of binding, indicating that the kinetic parameters determine the thermodynamic properties of the complex.

$$\Delta G^o = - RT\ln K_a$$

R – the universal gas constant [8.314 2J/(K*mol)]

T – the absolute temperature [K]

The binding free energy ΔG at any moment in time during an association is given by:

$$\Delta G = \Delta G^o + RT\ln Q$$

Q – the reaction quotient, which is defined as a ratio of the concentration of the protein-ligand complex to the product of the concentrations of the free protein and free ligand at any moment in time; when $Q = K_a$, an association reaction is at thermodynamic equilibrium, and $\Delta G = 0$. ΔG can also be parsed into its enthalpic and entropic contributions.

$$\Delta G = \Delta H - T\Delta S$$

ΔH – change in enthalpy of the system upon ligand binding

ΔS – change in entropy of the system upon ligand binding

T – temperature [K]

Enthalpy is a measure of the total energy of a protein-solvent system. It is the sum of the internal energy (i.e., the energy required to create the system) and the amount of energy required to make room for the system by displacing its environment and establishing its volume and pressure. Enthalpy is a quantifiable state function but the total enthalpy of a protein-solvent system cannot be measured directly. However, the change in enthalpy can be measured when it is expressed as the change in a system's energy. For a binding process, the enthalpy contribution reflects the energy change of the system when the ligand binds to the protein, resulting from the formation and disruption of various non-covalent interactions (hydrogen bonds, ion pairs, van der Waals contacts, electrostatic interactions, dipole-dipole interactions etc.) at the binding interface. However, the change in enthalpy upon binding is a global property of the entire thermodynamic system, including contributions not only from the solute but also from the solvent. Binding enthalpy ΔH is negative in exothermic processes when a formation of energetically favourable non-covalent interactions between atoms occur, and positive in endothermic processes when disruptions of energetically favourable non-covalent interactions happen.

Entropy is a measure of how evenly the heat energy is distributed over the whole thermodynamic system. Such a concept is central to the second law of thermodynamics which states that thermal energy always flows spontaneously from regions of higher temperature to regions of lower temperature. This reduces the state of order of the initial system, and, therefore, entropy could also be viewed as a measure of the disorder among atoms and molecules in a system. When a protein-solvent system is under constant temperature and pressure, the origin of entropy is the heat energy stored in atoms or molecules within the system. Entropy is a spontaneous process of the transfer and

diffusion of thermal energy among various system constituents leading ultimately to an as homogeneous as possible energy distribution over the system. Entropy cannot be measured directly, but entropy change can be measured quantitatively when it is considered as a change in the degrees of freedom of a system. ΔS is a global thermodynamic property of a system. Positive net entropy indicates the overall increase in the degree of freedom of the system, contributing favourably to the binding free energy. Negative net entropy indicates the overall decrease in the degree of freedom of the system, which contributes unfavourably to the binding free energy. The total entropy change associated with binding may be divided into three entropic terms:

$$\Delta S = \Delta S_{solv} + \Delta S_{conf} + \Delta S_{r/t}$$

ΔS_{solv} represents the solvent entropy change arising mainly from surface burial that results in solvent release upon binding, ΔS_{conf} represents the conformational entropy change reflecting the changes in the conformational freedom of both the protein and ligand upon association, $\Delta S_{r/t}$ represents the loss of rotational and translational degrees of freedom of the protein and ligand upon complex formation (Bundle et al., 1994, Du et al., 2016; Li et al., 2014).

The binding enthalpy primarily reflects the strength of non-covalent intermolecular interactions between the ligand and the target protein relative to those existing with the solvent. The change in enthalpy upon binding results principally from 1. the disruption of many individual interactions between the protein and the solvent (loss of hydrogen bonds and van der Waals contacts), and between the ligand and the solvent; 2. the formation of interactions between the protein and the ligand (salt bridges and van der Waals contacts); 3. the solvent reorganisation near the complex surface. These individual components may either contribute favourably or unfavourably, and the net enthalpy change is a result of their combination. The entropy change, on the other hand, mainly reflects two contributions, changes in solvation entropy ΔS_{solv} and changes in conformational entropy ΔS_{conf}. Upon binding, structured water molecules are released from the interacting antibody and carbohydrate surfaces into the bulk solvent. As desolvation (burial of water-accessible surface area) occurs, gain in solvent entropy is observed. This effect is probably the major driving force for the binding, e, and can be optimised, if needed, by the precise and close fit between the epitope and the binding pocket. This gain in solvent entropy is particularly important for hydrophobic groups since water molecules are less mobile near such regions. When the protein-ligand

complex is formed, water molecules are released into the bulk solvent and gain mobility. However, the positive entropy change gets compensated by loss in enthalpy owing to enthalpically weaker hydrogen bonds in bulk water. At the same time, the ligand and certain protein side-chains lose conformational freedom, resulting in a negative change in conformational entropy. The magnitudes of the enthalpy and entropy changes for this process tend to be roughly equal and are the source of the enthalpy-entropy compensation. Such compensation, although frequently observed, is not a requirement. If it was, meaning that changes in ΔH were always compensated by opposing changes in TΔS, optimisation of binding affinities would not be possible, which is not the case (Bronowska 2011; Bundle 1994; Dragan et al., 2017; Leavitt et al., 2001).

1. Surface Plasmon Resonance principle

Surface Plasmon Resonance is an optical-based and label-free method enabling to understand the structure-function relationship by measuring in real time the binding kinetics and affinity of molecular interactions. An SPR instrument consists of a sensor surface made out of a thin gold film, forming the floor of a flow cell through which an aqueous solution flows continuously. The ligand molecules are immobilised on the sensor surface and using a microfluidics system the analyte solution is passed over the modified surface. As the analyte binds the ligand in a time-dependent manner, a proportional increase in the refractive index at the surface is observed. This change, expressed in response units RU and measured in real-time, is proportional to the change in mass concentration on the sensor surface (1 RU = 1 pg/mm^2 for protein). After a certain association time, when the binding phase is complete (all binding sites are occupied), a solution containing no analyte is injected over the sensor surface and the analyte dissociates from the ligand. As this happens, a decrease in RU is observed. The time-dependent RU changes result in binding curves. Various concentrations of the analyte are used to obtain sensorgram data which is later fit to a suitable binding model in order to calculate the kinetic rate constants k_{on} and k_{off}, and the equilibrium constants of the interaction K_a and K_d (Beseničar et al., 2006; Myszka 1997).

Surface Plasmon Resonance is a phenomenon that occurs in thin conducting films at an interface between media of different refractive index (Fig.48 top), as the refractive index of a solution varies depending on the solute content. Under conditions of total

internal reflection, the light leaks an electric field intensity called an evanescent wave field across the interface into the medium of lower refractive index, without actually losing net energy. The amplitude of the evanescent field wave decreases exponentially with distance from the surface, and the effective penetration depth is about half the wavelength of the incident light. This distance is small (~300 nm) in relation to the volume of sample used, so effectively SPR measures the refractive index at the sensor surface. At a certain combination of the angle of incidence and energy (wavelength), free electrons at the metal surface absorb light photons and generate plasmons (electron charge density waves), creating a coherent plasmon oscillation. Under these resonance conditions, a sharp drop in the intensity of the reflected light is observed (Fig.48 bottom). The angle of incidence at which minimum reflection occurs is defined as the SPR angle, which is strongly dependent on the refractive index of the material near the metal surface. Because the evanescent wave field penetrates the solution, conditions for this resonance effect are very sensitive to the refractive index of the solution within the effective penetration depth of the evanescent field. The characteristics of the phenomenon in terms of angle and wavelength of the incident light and the shape of the reflectance minimum vary according to the nature and thickness of the conducting film (Jönsson et al., 1991; Lundström 1994; Zhou et al., 2018).

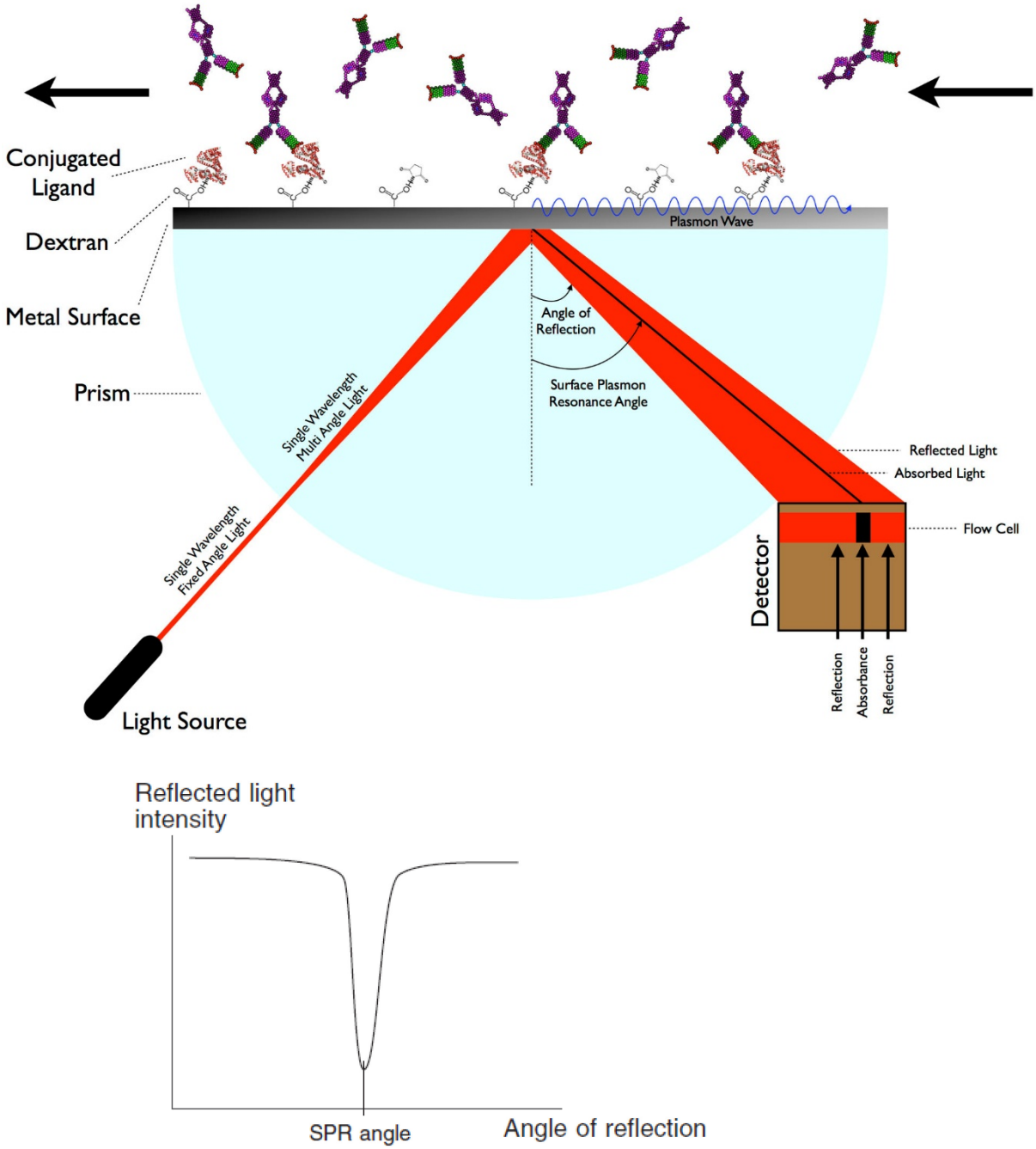


Figure 48. Representation of the Surface Plasmon Resonance principle. From: Biacore Sensor Surface Handbook.

A sensor chip consists of a glass slide coated with a uniform thin layer of gold, approx. 50 nm thick. These components, together with the docking system for mounting the sensor chip in the optical system, provide the conditions required for the generation of an SPR signal. Gold gives a well-defined reflectance minimum at easily handled visible light wavelengths, and is also amenable to covalent bonding of surface matrix layers at the same time, as the metal is largely inert in physiological buffer conditions. To provide

a suitable environment for studying molecular interactions, the gold surface is covered by a linker layer and, on most sensor chip types, a matrix of carboxymethyl dextran. The presence of carboxyl groups in the matrix allows reproducible and generally applicable immobilisation of a wide range of ligands. This dextran matrix not only provides high freedom of rotation to closely mimic solution binding interactions but also extends the surface on which the interactions occur, thus providing better accessibility and increasing instrument sensitivity, It is the surface matrix that determines the properties of the sensor chip with respect to ligand attachment and molecular interaction (Jönsson 1992; Myszka 1997).

In broad terms, there are two main approaches to attaching macromolecular ligands to the sensor chip surface, covalent immobilisation and high-affinity capture. The choice of method depends on the properties of the substance. The immobilisation approaches may exploit the available amine (-NH_2), thiol (-SH), aldehyde (-CHO), hydroxyl (-OH), or carboxyl (-COOH) functional groups on the ligand, to capture nucleic acids, peptides, proteins, antibodies, or carbohydrates. Amine coupling chemistry is the most widely applicable approach for covalently attaching biomolecules to the sensor surface (Fig.49). With this method, a fraction of the carboxy1 groups on the dextran matrix is first activated with a mixture of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and *N*-hydroxysuccinimide (NHS) to give reactive *N*-hydroxysuccinimide esters. The ligand dissolved in a low-ionic-strength buffer, at a pH below its isoelectric point, is flown over the activated surface. The electrostatic attraction between the positively charged ligand and the negatively charged carboxyl groups in the matrix facilitates the ligand concentration on the matrix. During this process, the NHS esters react with amino groups and other nucleophilic groups (the thiol group of cysteine, and to some extent the imidazolium group of histidine and the phenoxy group of tyrosine) to link the ligand covalently to the dextran. Finally, ethanolamine is passed over the sensor chip to inactivate the remaining excess groups by transforming them into amides. Additionally, the solution weakens the electrostatic interaction between the ligand and the matrix and thus removes electrostatically bound material from the surface (Johnsson et al., 1991).

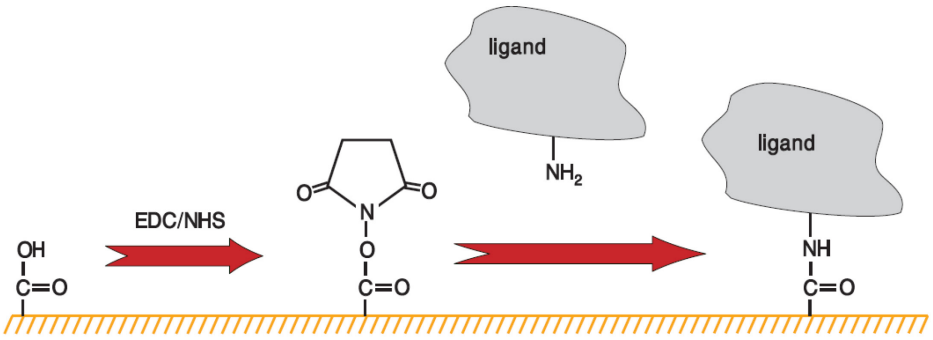


Figure 49. Schematic representation of amine coupling of a ligand to the sensor chip surface. From: Biacore Sensor Surface Handbook.

2. Isothermal Titration Calorimetry principle

Isothermal titration calorimetry is a solution-phase technique directly detecting the absorbed or released heat (binding enthalpy change ΔH) accompanying a binding interaction. It is a titrimetric method in which the titration is performed at constant pressure and temperature. Heat is released or absorbed as a result of the redistribution and formation of non-covalent bonds when the interacting molecules go from free to bound state. An ITC apparatus consists of two identical cells composed of a highly efficient thermal conducting material (Hastelloy or gold) (Fig.50). The reference cell usually contains water, while the sample cell contains one of the binding partners (the sample) and a stirring syringe which holds the other binding partner (the ligand). As modern instruments operate on the heat compensation principle, the response (measured signal) is the time-dependent input of power (µcal/s), applied to the cell heaters, necessary to maintain equal temperatures in the sample and reference cells, as the binding partners are mixed. During the injection of the ligand into the sample cell, heat is absorbed or released depending on whether the macromolecular association reaction is endothermic or exothermic. For an exothermic reaction, the temperature in the sample cell will increase, and the feedback power will be deactivated. For endothermic reactions, the feedback circuit will increase power to the sample cell to maintain the same temperature. The heat absorbed or released during a calorimetric titration is proportional to the fraction of bound ligand. Each ligand injection results in a heat pulse that is integrated with respect to time and normalised for concentration to generate a titration curve of kcal/mol vs molar ratio (ligand/sample). The resulting isotherm is fitted to a binding model to generate the enthalpy (ΔH), affinity (K_d), and

stoichiometry of interaction (n), from which the binding free energy (ΔG) and subsequently entropy (ΔS) can be computed (Leavitt et al., 2001; Pierce 1999).

It is important to highlight that the heat exchange detected by an ITC experiment is the total heat effect in the sample cell upon ligand titration, including not only the heat absorbed or released during the binding interaction but also the heat effects arising from dilution of the ligand and protein, mixing two solutions of different compositions, temperature differences between the sample cell and the syringe etc. As a result, control experiments need to be performed to determine these non-specific heat effects and obtain only the heat exchange during the complex formation (Bronowska 2011; Perozzo et al., 2004).

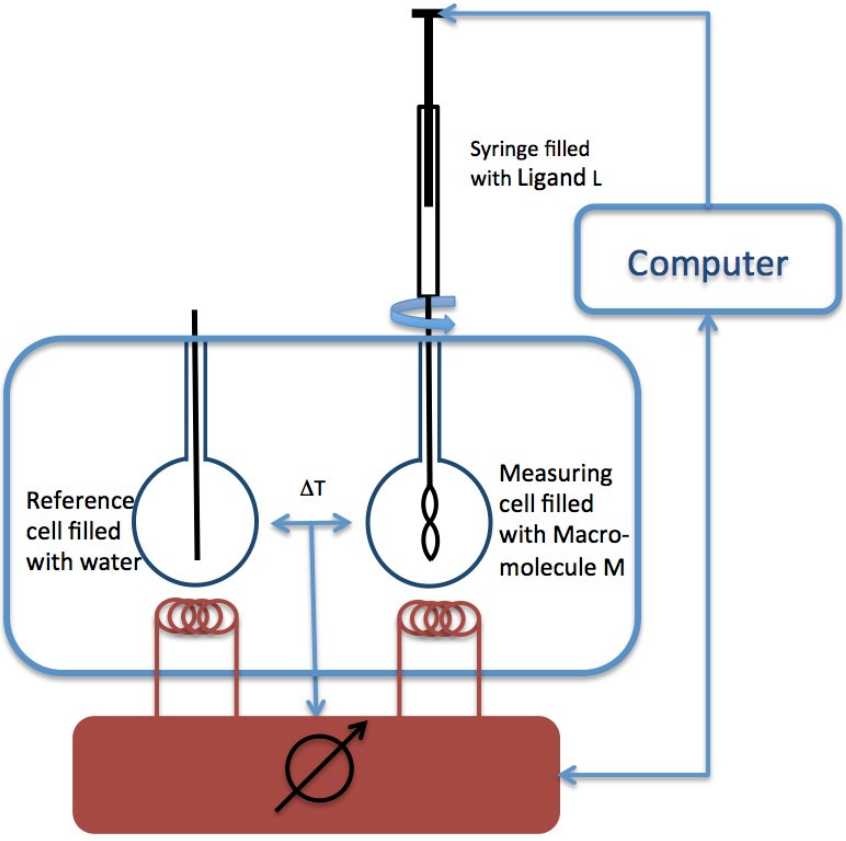


Figure 50. Schematic representation of the working principle of an ITC instrument. The measured signal between the reference and sample cells is the power needed to maintain a zero temperature difference, ΔT ~ 0.

MATERIALS & METHODS

1. SPR measurements

1.1. Equipment

The SPR studies between Mannitou mAb/Fab and selected paucimannose *N*-glycans have been conducted using the Biacore 3000 and Biacore T200 instruments (GE Healthcare Life Sciences, UK). The system allows getting a reliable real-time insight into biological processes in small-scale interaction analysis, providing information on binding specificity, kinetics, affinity, and concentration analysis. It contains all key functionalities needed to define the structure-function relationships and understand the dynamics of molecular pathways. The ligands have been immobilised by amine coupling on a Biacore Sensor Chip CM5 (GE Healthcare Life Sciences, UK) (Fig.51). The CM5 chip carries a matrix of carboxymethylated dextran covalently attached to a gold surface, required for the generation of the surface plasmon resonance response. The flexible dextran matrix extends ~100 nm from the gold surface, allowing relatively free movement of the attached ligand molecules and creates a hydrophilic environment suitable for a wide variety of protein interactions. Its high immobilisation capacity permits a broad range of capture densities and the surface stability provides accuracy and precision, allowing repeated analysis in the same surface conditions.

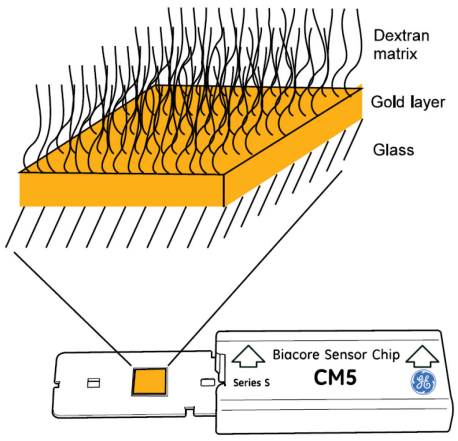


Figure 51. Schematic illustration of the structure of the Biacore Sensor Chip CM5 surface (GE Healthcare Life Sciences, UK). From: Biacore T200 Getting Started.

1.2. Ligand immobilisation

Amine coupling chemistry is one of the most common techniques for covalent immobilisation of a wide range of molecules. This approach results in the formation of a stable bond between the ligand and the sensor surface. Determination of the binding kinetics and affinity between Mannitou mAb/Fab and the studied paucimannose *N*-glycans starts with the CM5 chip surface activation using a mixture of 0.4M 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and 0.1M *N*-hydroxysuccinimide (NHS). The reagent solutions have been equilibrated at room temperature and mixed shortly before use.

The chip for Mannitou Fab experiment has been prepared by immobilising a glycoprotein carrying $Man_3GlcNAc_2$. The 1st flow cell (Fc=1) serves as a reference surface, so it is blocked immediately after activation (blank immobilisation). In the 2nd flow cell (Fc=2) ΔXT/FT *Nicotiana benthamiana* plant-produced helminth glycoprotein omega-1 with $Man_3GlcNAc_2$ *N*-glycan attached (Strasser et al., 2008) is immobilised at 400 RU (400 pg ligand/mm^2 sensor surface) using 10 mM NaAc pH=4.5 as immobilisation buffer. 1M ethanolamine-HCL pH=8.5 is used to inactivate excess reactive groups.

The chip for Mannitou mAb (non-purified and purified) multi-cycle analysis has been prepared by direct immobilisation of $Man_3GlcNAc_2$ and $Man_3GlcNAc_2$(α1-6)Fuc. The 1st and 3rd flow cells (Fc=1, Fc=3) serve as a reference surface, so they are blocked immediately after activation (blank immobilisation). In the 2nd flow cell (Fc=2) $Man_3GlcNAc_2$ is covalently immobilised at 237 RU (237 pg ligand/mm^2 sensor surface) using 10 mM NaAc pH=4.5 as immobilisation buffer. In the 4th flow cell (Fc=4) $Man_3GlcNAc_2$(α1-6)Fuc is covalently immobilised at 163 RU (163pg ligand/mm^2 sensor surface) using 10 mM NaAc pH=4.5 as immobilisation buffer. 1M ethanolamine-HCL pH=8.5 is used to inactivate excess reactive groups.

The chips for Mannitou mAb (non-purified) single-cycle analysis have been prepared by immobilising a glycoprotein carrying the *N*-glycan of choice. The 1st flow cell (Fc=1) serves as a reference surface, so it is blocked immediately after activation (blank immobilisation). In the 2nd, 3rd, and 4th flow cell (Fc=2, Fc=3, Fc=4) ΔXT/FT *Nicotiana benthamiana* plant-produced helminth glycoprotein omega-1 with Man_3(β1-2)$XylGlcNAc_2$(α1-3)Fuc, $Man_3GlcNAc_2$, $Man_3GlcNAc_2$(α1-6)Fuc, or $Man_3GlcNAc_2$(α1-3)Fuc *N*-glycan attached (Strasser et al., 2008) is immobilised using 10 mM NaAc pH=4.5

as immobilisation buffer. 1M ethanolamine-HCL pH=8.5 is used to inactivate excess reactive groups.

The chips for Mannitou mAb (purified) single-cycle analysis have been prepared by immobilising 10 µg/ml of the glycoprotein carrying the *N*-glycan of choice. The 1st flow cell (Fc=1) serves as a reference surface, so it is blocked immediately after activation (blank immobilisation). In the 2nd flow cell (Fc=2) ΔXT/FT *Nicotiana benthamiana* plant-produced helminth glycoprotein omega-1 with $Man_3GlcNAc_2$ *N*-glycan attached (Strasser et al., 2008) is immobilised at 2300 RU (2300 pg ligand/mm^2 sensor surface) using 10 mM NaAc pH=4.5 as immobilisation buffer. In the 3rd flow cell (Fc=3) $Man_3GlcNAc_2$(α1-6)Fuc is covalently immobilised at 2000 RU (2000 pg ligand/mm^2 sensor surface) using 10 mM NaAc pH=4.5 as immobilisation buffer. In the 4th flow cell (Fc=4) $Man_3GlcNAc_2$(α1-3)Fuc is covalently immobilised at 2500 RU (2500 pg ligand/mm^2 sensor surface) using 10 mM NaAc pH=4.5 as immobilisation buffer. 1M ethanolamine-HCL pH=8.5 is used to inactivate excess reactive groups.

1.3. Binding kinetics and affinity analysis

The binding interactions have been studied using HBS-EP, pH=7.4 (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.005% v/v Surfactant P20), and 200 mM $NaHCO_3$, 500 mM NaCl, pH=7.8, as running buffers, at a flow rate of 30 µl/min at 10°C. Both experiments (using Mannitou mAb and Mannitou Fab) begin with three start-up cycles to allow surface stabilisation.

In the multi-cycle analysis using Mannitou Fab, each sample injection (1.5 min contact time) is followed by 10 min dissociation time to allow the bound analyte to dissociate spontaneously from the surface-attached ligand. The Fab has been used injected with increasing concentrations from 0 nM (no analyte), 158 nM, 315 nM, 630 nM, to 1260 nM.

In the multi-cycle analysis using non-purified Mannitou mAb, each sample injection (3 min contact time, 6 min dissociation time) is followed by regeneration with 50mM NaOH to remove the bound analyte from the surface-immobilised ligand. The mAb has been used in 0 nM (no analyte), 2.3 nM, 7.2 nM, 15.9 nM, 31.9 nM, 64.1 nM, 132 nM, 249 nM, 452 nM, 739 nM, and 925 nM concentrations.

In the multi-cycle analysis using purified Mannitou mAb, each sample injection (1 min contact time, 6.5 min dissociation time) is followed by regeneration with 50mM

NaOH to remove the bound analyte from the surface-immobilised ligand. The mAb has been used in 0 nM (no analyte), 0.8 nM, 0.16 nM, 0.33 nM, 0.66 nM, 1.31 nM, 2.63 nM, 5.25 nM, 10.5 nM, 21 nM, 42 nM, 84 nM, 168 nM, and 336 nM concentrations.

In the single-cycle analysis using non-purified Mannitou mAb each sample injection (1.5 min contact time) is followed by 2.5 min dissociation time. The mAb has been used in 13.2 nM, 26.3 nM, 52.6 nM, 105.3 nM, and 210.5 nM concentrations for $Man_3GlcNAc_2$ and $Man_3GlcNAc_2(\alpha1\text{-}3)Fuc$ ligands, and in 19.7 nM, 39.5 nM, 78.9 nM, 157.9 nM, and 315.8 nM concentrations for $Man_3(\beta1\text{-}2)XylGlcNAc_2(\alpha1\text{-}3)Fuc$ and $Man_3GlcNAc_2(\alpha1\text{-}6)Fuc$ ligands.

In the single-cycle analysis using purified Mannitou mAb each sample injection (3 min contact time) is followed by 4 min dissociation time. The mAb has been used in 59.375 nM, 118.75 nM, 237.5 nM, 475 nM, 950 nM concentrations. Data analysis is performed with BIA Evaluation Software.

The chip preparation and SPR measurements have been made by S. Robakiewicz (single-cycle analysis), often with the help of Z. Lens, and by J. Bouckaert and C. Bridot (multi-cycle analysis).

2. ITC measurements

2.1. Equipment

The ITC measurements have been performed using MicroCal PEAQ-ITC (Malvern Panalytical, UK). MicroCal PEAQ-ITC is a highly sensitive and precise, low volume isothermal titration calorimeter for the label-free in solution study of biomolecular interactions. The system directly measures the heat released or absorbed during binding events, from which it calculates binding enthalpy (ΔH), affinity (K_d), stoichiometry (n), free energy (ΔG), and entropy (ΔS). The wide affinity range enables highly reproducible analysis of weak to high-affinity binders (K_d from 10^{-2} to 10^{-9} M) in a single experiment. High signal-to-noise ratio gives more confidence in accessing data quality and relevance of the generated affinity and thermodynamic parameters. MicroCal PEAQ-ITC has a wide range of applications as it enables characterisation of molecular interactions of small molecules, nucleic acids, lipids, proteins, enzymes, and antibodies.

2.2. Binding and kinetic analysis

Prior to the assay, Mannitou Fab has been dialysed against HEPES buffer (20 mM HEPES, 150 mM NaCl) pH=7.4. The same buffer is also used to dissolve the $Man_3GlcNAc_2$ glycan serving as the ligand. 2 µL injections of 3 mM $Man_3GlcNAc_2$ are titrated at 150 s intervals from a computer-controlled 40 µL syringe into the sample cell containing 35 µM Mannitou Fab solution, that is constantly stirred at 750 rpm at 25°C, a condition at which the antibody is considered to be in its native folding state. The heat peaks, generated by the complex association, are integrated to obtain the binding enthalpy change ΔH. The equilibrium dissociation constant K_d is obtained from the slope of the resulting curve fitted to the data points. The data calculation is performed using the Microcal Peaq-ITC Analysis software.

The ITC experiment has been conducted by S. Robakiewicz with the help of A. Gimeno in the lab of J. Jimenez-Barbero.

RESULTS

1. Surface Plasmon Resonance

When studying biological problems at the molecular or cellular level using recombinant antibodies, like Mannitou Fab, it is important to be able to show that the recombinant protein has the same structure and activity as its native counterpart. This is most easily done by confirming that the recombinant antibody binds its natural ligands. Because such interactions involve multiple residues, which are usually far apart in the primary amino acid sequence, they require a correctly folded antibody.

The capacity of SPR to measure the real-time binding data makes it well suited to analyses of kinetics and affinity. Quantitating how macromolecules assemble into complexes and break down over time is required to define the mechanism of binding. This kinetic information is pivotal in relating the structure of biological macromolecules to their function (Myszka 1997). Kinetics and affinity measurements on the sensor surface determine the binding characteristics between the ligand and the analyte. Kinetics deals with time-dependent events and describes how fast molecules bind and dissociate. The kinetics of interaction determines how much complex is formed in a given time. Affinity describes the strength of interaction, independently of time. The affinity of interaction determines how much complex is formed when a mixture of interactants has reached equilibrium. A full kinetic profile can be generated by measuring the interaction as a function of time over a range of analyte concentrations and then fitting the whole data set to a mathematical model describing the interaction mechanism. Kinetic parameters are determined from the association and dissociation phases of the sensorgram. Affinity can be established either from the kinetic rate constants or by measuring the steady-state binding levels of the analyte. The analyte concentrations may be introduced in separate cycles with surface regeneration after each sample injection (multi-cycle analysis) or sequentially in a single cycle with no regeneration between sample injections (single-cycle analysis). In this study, the both multi-cycle (Fig.52, Fig.53, Fig.54) and single-cycle approach (Fig.55, Fig.56) is applied and the affinity is assessed from the kinetic parameters. Different kinetic experiments have been performed using Mannitou Fab (Fig.52) and Mannitou mAb (Fig.53, Fig.54, Fig.55, Fig.56)

respectively. The simplest interaction model that is consistent with the experimental system has been chosen, as the complex that forms follows pseudo-first-order kinetics and it is assumed that the binding is equivalent and independent for all binding sites. 1:1 Langmuir binding model has been fitted to the experimental data from both the association and dissociation phases, after subtracting the reference. A background response will also be generated if there is a difference in the refractive indices of the running and sample buffers. This background response must be subtracted from the sensorgram to obtain the actual binding response. The background response is recorded by injecting the analyte through the reference cell which has no ligand immobilised to the sensor surface.

Mannitou Fab – $Man_3GlcNAc_2$ experiment yields K_d [M] = 4.8 x 10^{-5} ± 1.57 x 10^{-5}. The established kinetic rate constants are k_a [1/Ms] = 7.247 x 10 ± 1.101 x 10, and k_d [1/s] = 3.48 x 10^{-3} ± 1.73 x 10^{-4}.

The kinetic parameters determined from the multi-cycle analysis using Mannitou mAb (non-purified) – $Man_3GlcNAc_2$ are the following: k_a [1/Ms] = 8.074 x 10^3 ± 1.916 x 10, k_d [1/s] = 3.979 x 10^{-4} ± 3.02 x 10^{-6}, K_d [M] = 4.929 x 10^{-8} ± 1.576 x 10^{-7}. The kinetic parameters generated from the Mannitou mAb (non-purified) – $Man_3GlcNAc_2$(α1-6)Fuc multi-cycle analysis are the following: k_a [1/Ms] = 1.368 x 10^4 ± 3.807 x 10, k_d [1/s] = 4.325 x 10^{-4} ± 5.03 x 10^{-6}, K_d [M] = 3.162 x 10^{-8} ± 1.321 x 10^{-7}.

The kinetic parameters determined from the multi-cycle analysis using Mannitou mAb (purified) – $Man_3GlcNAc_2$ are the following: k_a [1/Ms] = 3.008 x 10^5 ± 2.142 x 10^3, k_d [1/s] = 2.066 x 10^{-3} ± 1.72 x 10^{-5}, K_d [M] = 6.868 x 10^{-9} ± 8.03 x 10^{-9}. The kinetic parameters generated from the Mannitou mAb (purified) – $Man_3GlcNAc_2$(α1-6)Fuc multi-cycle analysis are the following: k_a [1/Ms] = 5.045 x 10^5 ± 3.827 x 10^3, k_d [1/s] = 3.052 x 10^{-3} ± 2.15 x 10^{-5}, K_d [M] = 6.05 x 10^{-9} ± 5.62 x 10^{-9}.

The single-cycle analysis for the Mannitou mAb (non-purified) – Man_3(β1-2)$XylGlcNAc_2$(α1-3)Fuc interaction has resulted in K_d [M] = 1.861 x 10^{-8} ± 1.079 x 10^{-8}. The obtained kinetic rate constants are k_a [1/Ms] = 4.19 x 10^4 ± 3.095 x 10^2, and k_d [1/s] = 7.796 x 10^{-4} ± 3.338 x 10^{-6}. The single-cycle analysis for the Mannitou mAb (non-purified) – $Man_3GlcNAc_2$ interaction has resulted in K_d [M] = 5.61 x 10^{-9} ± 3.254 x 10^{-9}. The obtained kinetic rate constants are k_a [1/Ms] = 2.188 x 10^5 ± 3.426 x 10^3, and k_d [1/s] = 1.227 x 10^{-3} ± 1.115 x 10^{-5}. The single-cycle analysis for the Mannitou mAb (non-purified) – $Man_3GlcNAc_2$(α1-6)Fuc has resulted in K_d [M] = 8.721 x 10^{-9} ± 4.171 x

10^{-9}. The obtained kinetic rate constants are k_a [1/Ms] = 9.59 x 10^4 ± 4.998 x 10^3, and k_d [1/s] = 8.364 x 10^{-4} ± 2.085 x 10^{-5}. The single-cycle analysis for the Mannitou mAb (non-purified) – $Man_3GlcNAc_2$(α1-3)Fuc has resulted in K_d [M] = 1.084 x 10^{-8} ± 5.299 x 10^{-9}. The obtained kinetic rate constants are k_a [1/Ms] = 5.62 x 10^4 ± 2.272 x 10^3, and k_d [1/s] = 6.094 x 10^{-4} ± 1.204 x 10^{-5}.

The single-cycle analysis for the Mannitou mAb (purified) – $Man_3GlcNAc_2$ has resulted in K_d [M] = 1.052 x 10^{-7} ± 4.399 x 10^{-8}. The obtained kinetic rate constants are k_a [1/Ms] = 3.628 x 10^3 ± 1.875 x 10^2, and k_d [1/s] = 3.818 x 10^{-4} ± 8.245 x 10^{-6}. The single-cycle analysis for the Mannitou mAb (purified) – $Man_3GlcNAc_2$(α1-6)Fuc has resulted in K_d [M] = 7.654 x 10^{-8} ± 3.716 x 10^{-8}. The obtained kinetic rate constants are k_a [1/Ms] = 4.056 x 10^3 ± 2.983 x 10^2, and k_d [1/s] = 3.104 x 10^{-4} ± 1.109 x 10^{-5}. The single-cycle analysis for the Mannitou mAb (purified) – $Man_3GlcNAc_2$(α1-3)Fuc has resulted in K_d [M] = 1.067 x 10^{-7} ± 4.735 x 10^{-8}. The obtained kinetic rate constants are k_a [1/Ms] = 4.374 x 10^3 ± 3.878 x 10^2, and k_d [1/s] = 4.667 x 10^{-4} ± 1.836 x 10^{-5}.

There are two major tools for assessing the significance of the reported constants: the closeness of fit between the fitted and experimental curves, and the statistical significance of the parameters. Visual inspection of the fitted curves overlaid on the experimental data gives an indication of the closeness of the fits. All assays result in relatively well-fitted curves overlaid onto the experimental data. The chi-square value is a quantitative measure of the closeness of fit, and in an ideal situation will approximate to the square of the short-term noise level. The chi-square value for the Mannitou Fab – $Man_3GlcNAc_2$ experiment is 14.07. The chi-square value for the multi-cycle analysis using non-purified Mannitou mAb – $Man_3GlcNAc_2$ is 32.33, and for Mannitou mAb – $Man_3GlcNAc_2$(α1-6)Fuc it is 3.96. The chi-square value for the multi-cycle analysis using purified Mannitou mAb – $Man_3GlcNAc_2$ is 26.86, and for Mannitou mAb – $Man_3GlcNAc_2$(α1-6)Fuc it is 0.53. The chi-square value for the single-cycle analysis using non-purified Mannitou mAb – Man_3(β1-2)$XylGlcNAc_2$(α1-3)Fuc is 27.66, for Mannitou mAb – $Man_3GlcNAc_2$ it is 8.02, for Mannitou mAb – $Man_3GlcNAc_2$(α1-6)Fuc it is 552.2, and for Mannitou mAb – $Man_3GlcNAc_2$(α1-3)Fuc it is 233.14. The chi-square value for the single-cycle analysis using purified Mannitou mAb – $Man_3GlcNAc_2$ is 4.96, for Mannitou mAb – $Man_3GlcNAc_2$(α1-6)Fuc it is 16.02, for Mannitou mAb – Man3GlcNAc2(α1-3)Fuc it is 11.56. Fitting experimental data to a mathematical interaction model will return values for all parameters in the model, regardless of whether they are significant or not. The

significance of the parameter values returned by the fitting procedure is indicated by the standard error. The higher the standard error, the less significant the parameter is.

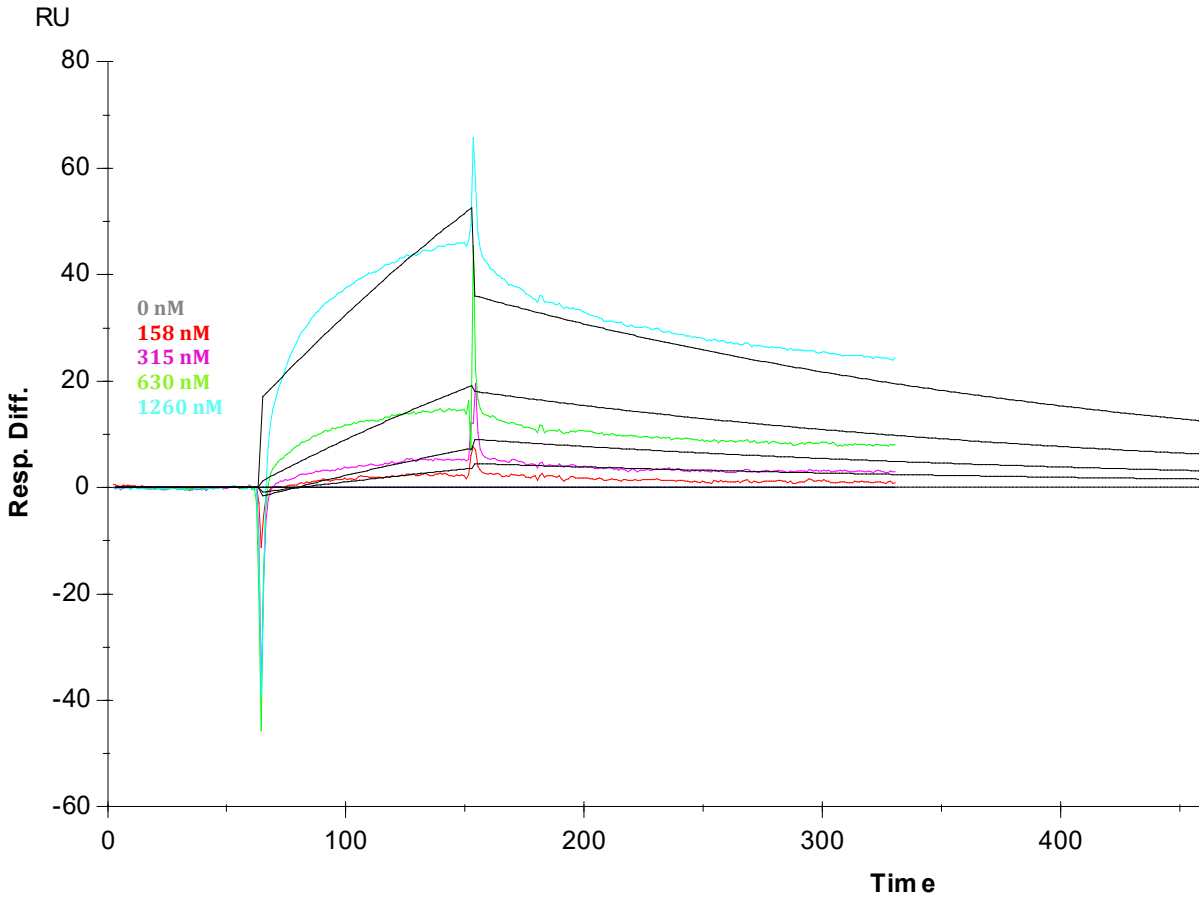


Figure 52. Global fitting of the SPR sensorgram (Fc2-Fc1) representing the kinetic experiments using a concentration series of Mannitou Fab (purified). Global fitting finds the best fit for all of the sensorgrams simultaneously. Selected parameters such as rate and affinity constants are constrained to have a single value for all sensorgrams in the data set. Multi-cycle analysis. The surface-immobilised glycoprotein is a plant-produced omega-1 with $Man_3GlcNAc_2$ *N*-glycan.

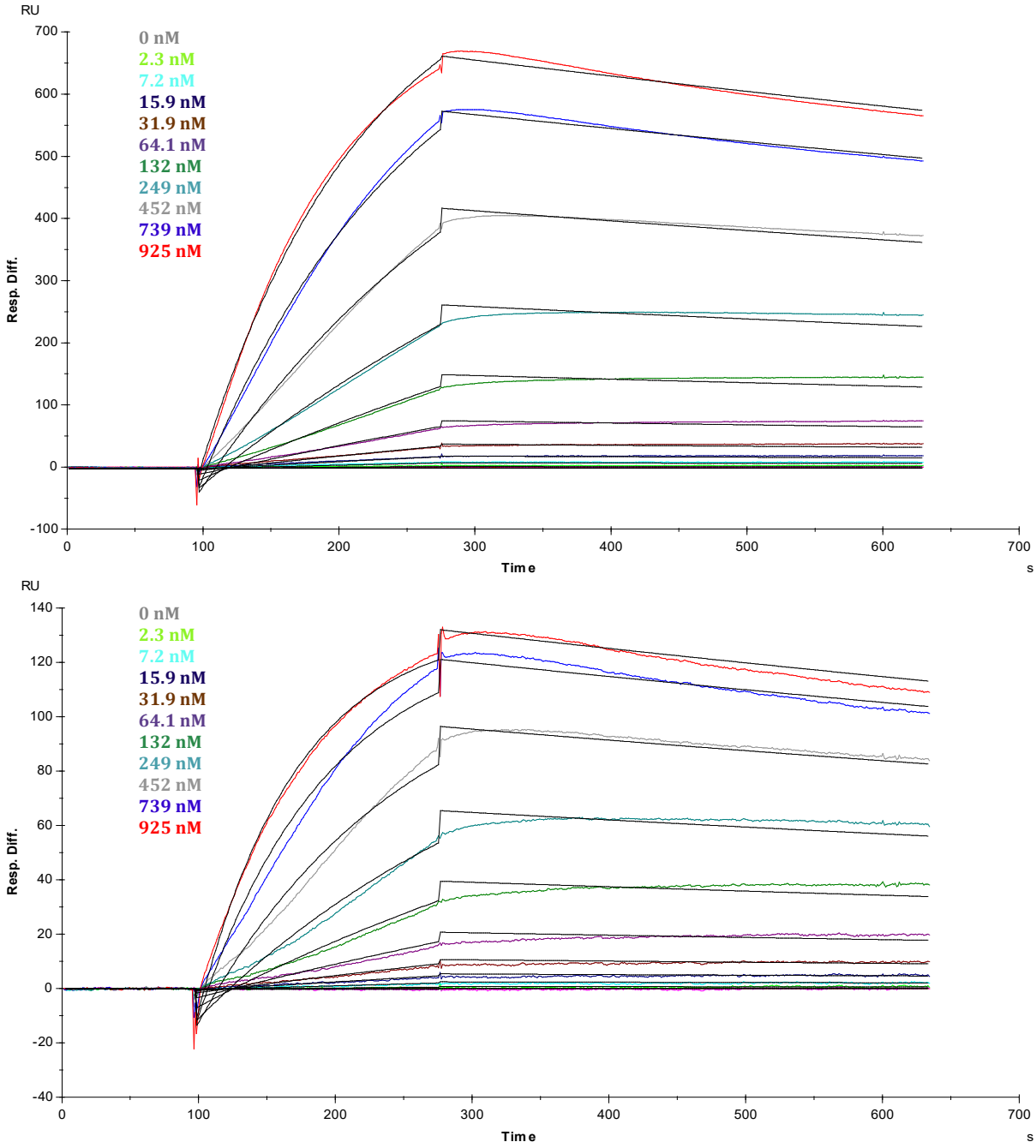


Figure 53. Global fitting of the SPR sensorgrams (Fc2-Fc1) representing the kinetic experiments using a concentration series of Mannitou mAb (non-purified). Multi-cycle analysis. Top: The surface-immobilised *N*-glycan is $Man_3GlcNAc_2$. Bottom: The surface-immobilised *N*-glycan is $Man_3GlcNAc_2(\alpha 1\text{-}6)Fuc$.

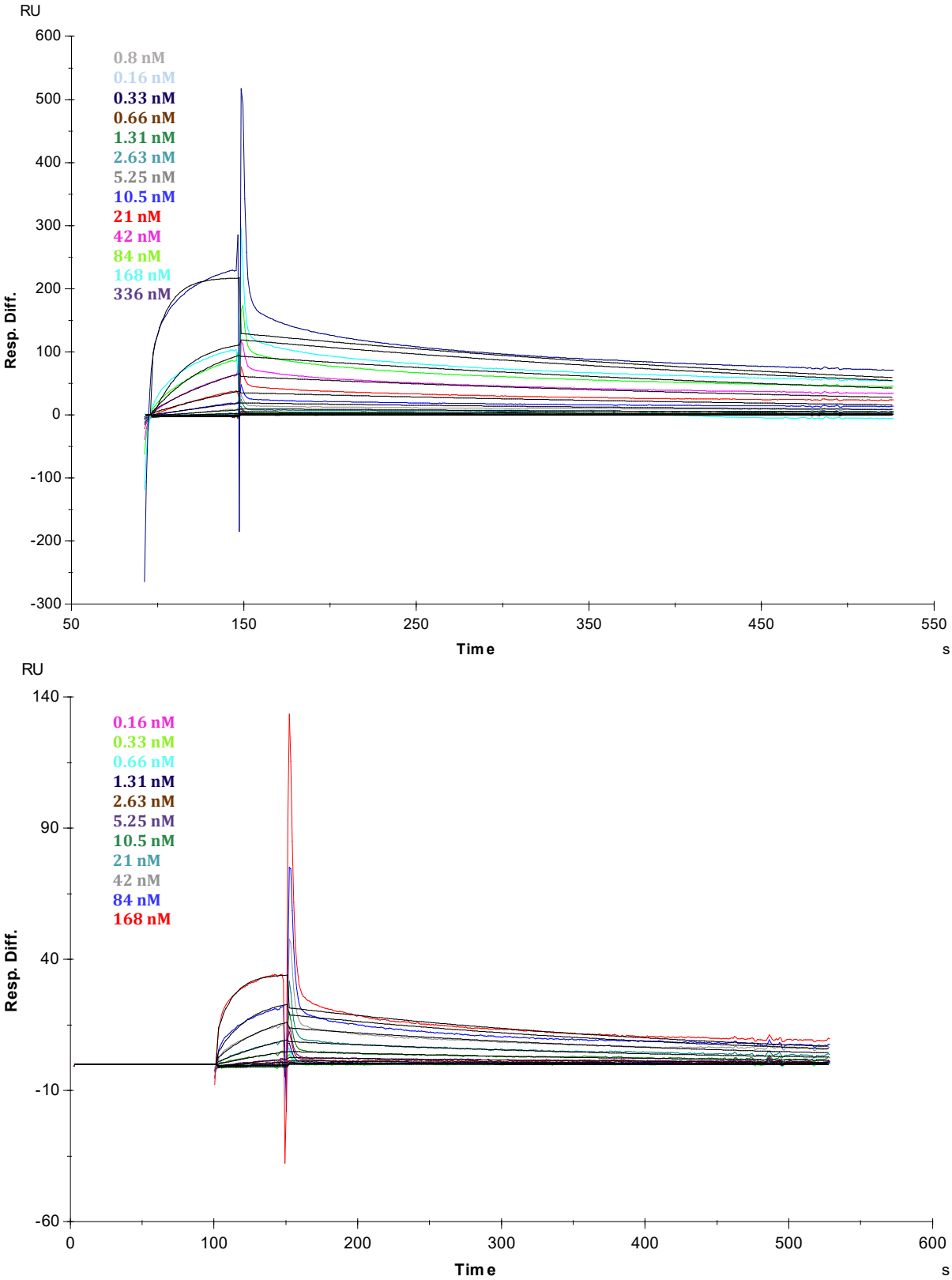
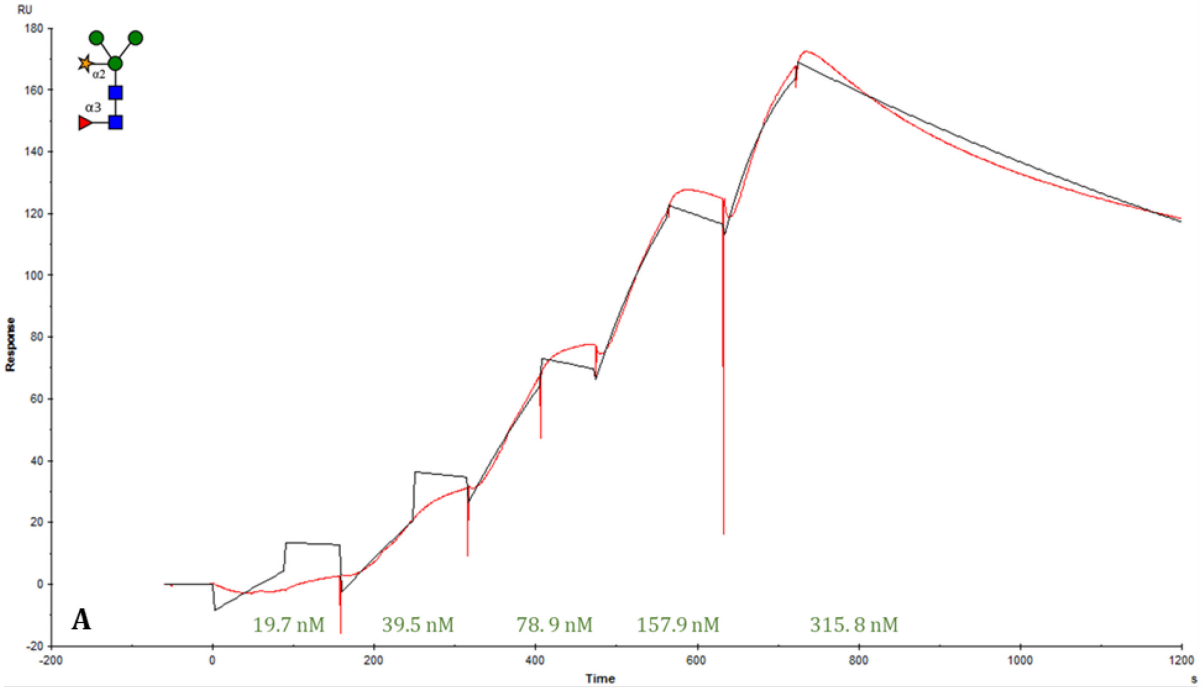
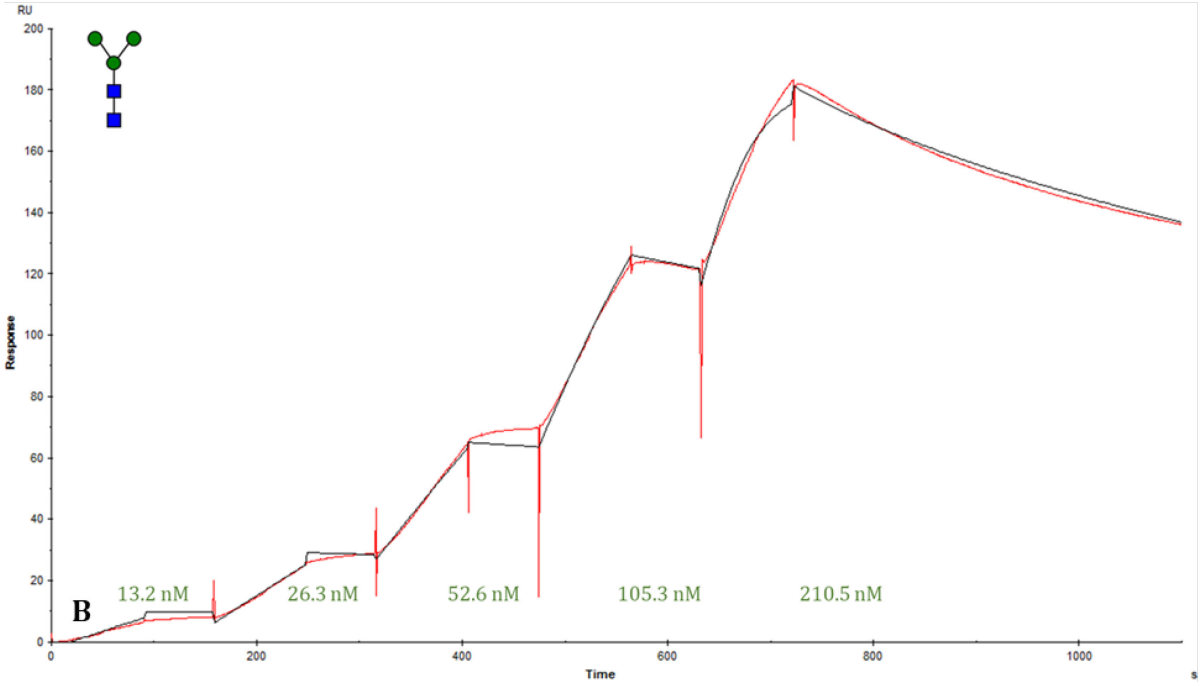


Figure 54. Global fitting of the SPR sensorgrams (Fc2-Fc1) representing the kinetic experiments using a concentration series of Mannitou mAb (purified). Multi-cycle analysis. Top: The surface-immobilised N-glycan is $Man_3GlcNAc_2$. Bottom: The surface-immobilised N-glycan is $Man_3GlcNAc_2(\alpha 1\text{-}6)Fuc$.





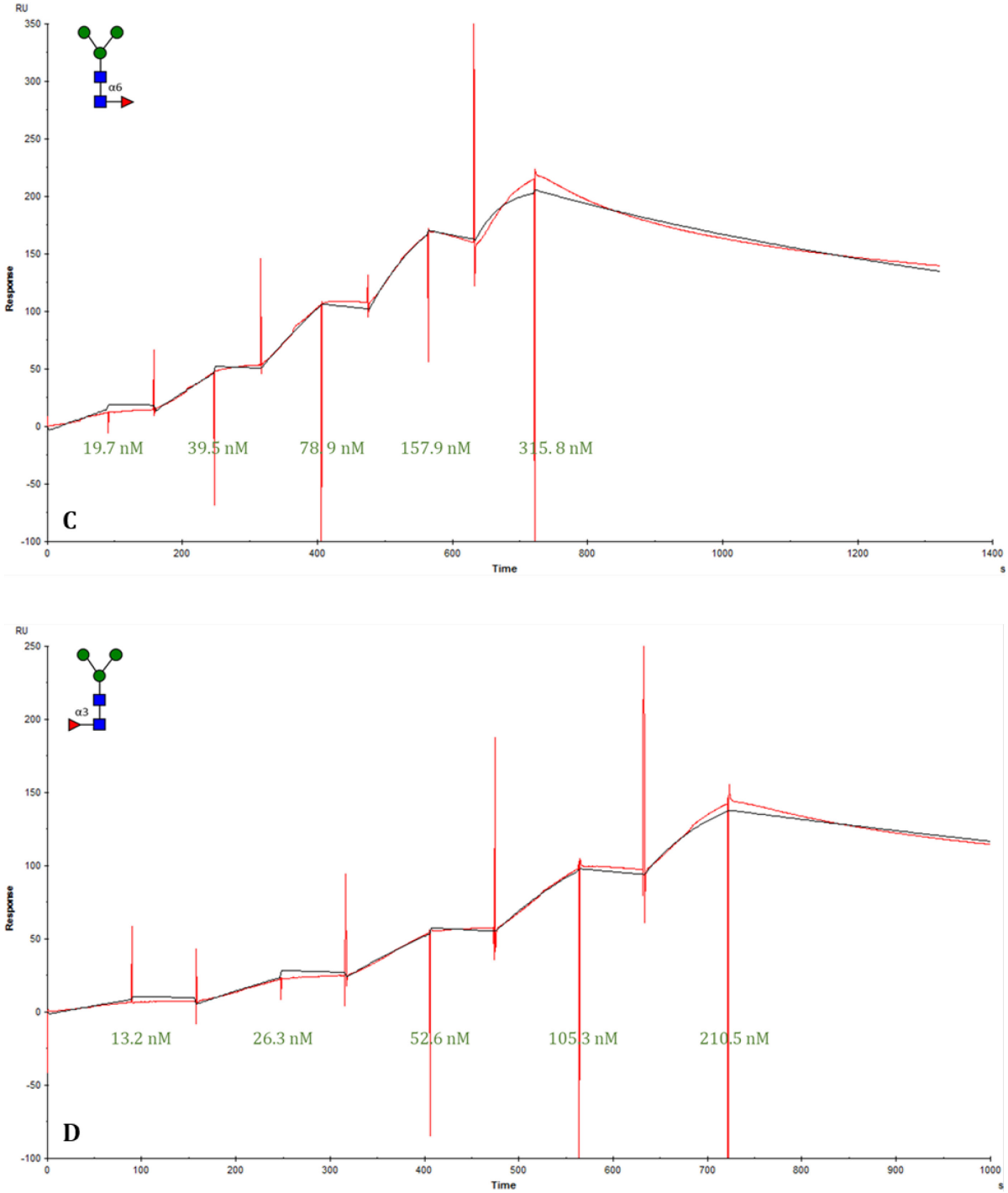
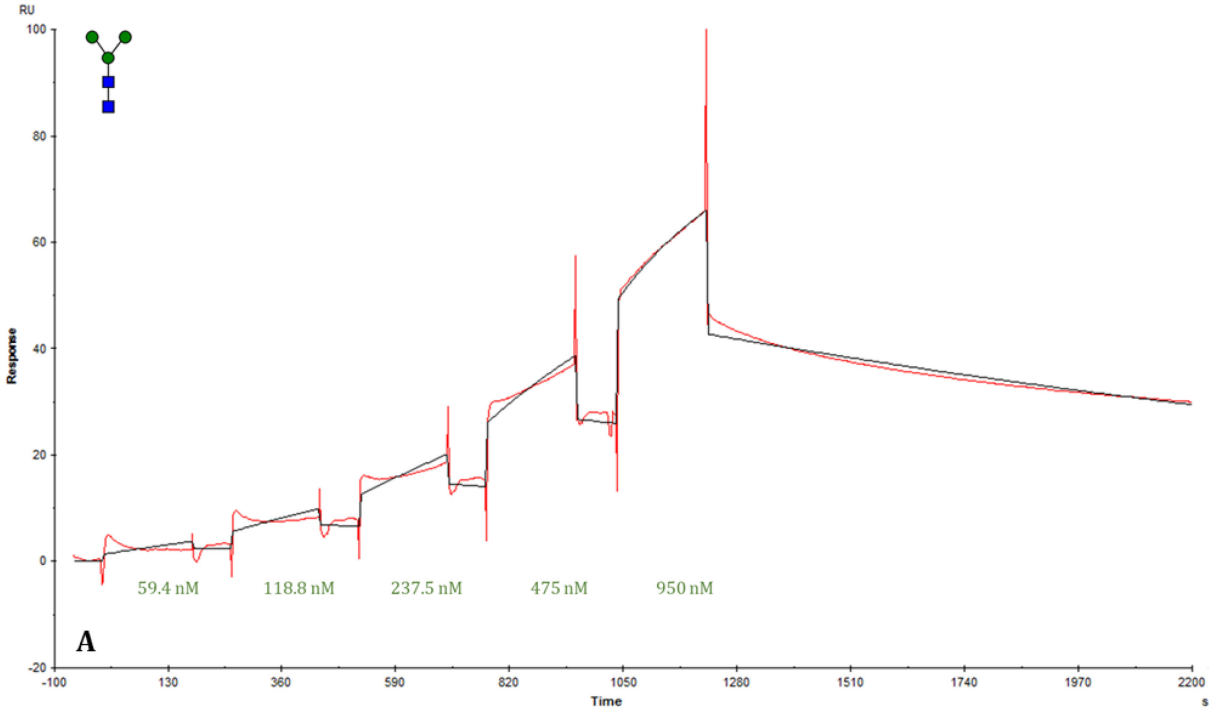
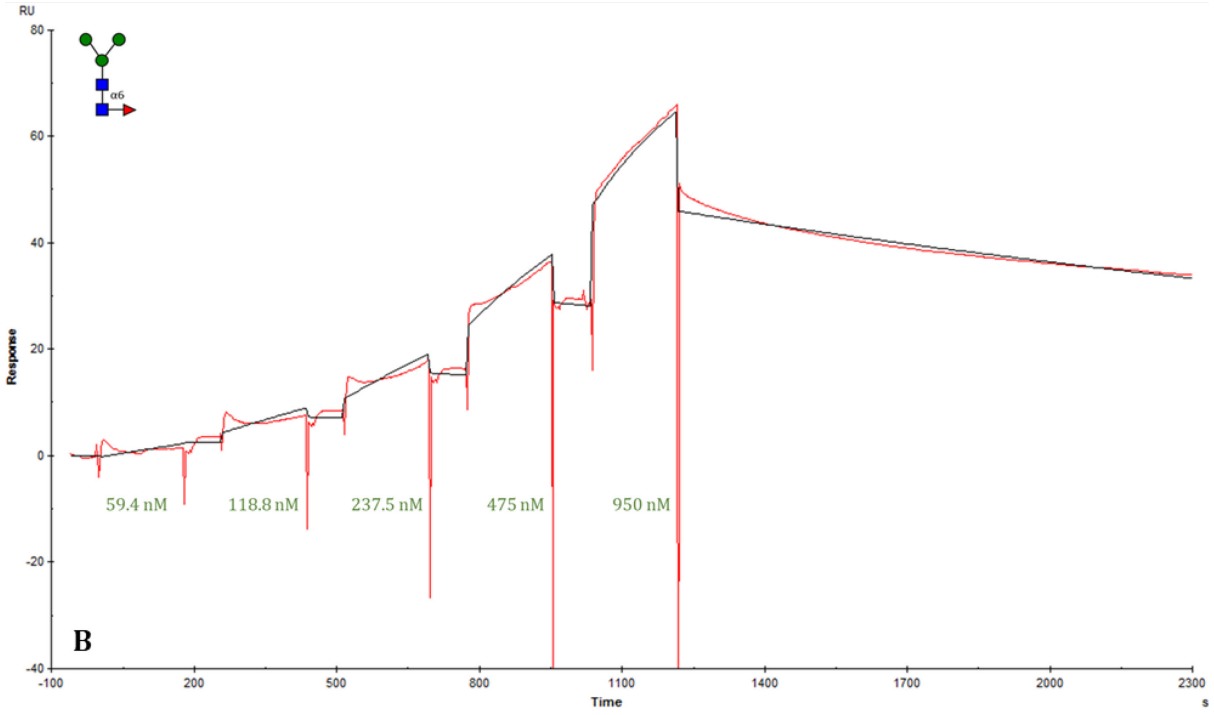


Figure 55. Global fitting of the SPR sensorgrams (Fc2-Fc1) representing the kinetic experiments using a concentration series of Mannitou mAb (non-purified). Single-cycle analysis. A: The surface-immobilised glycoprotein is a plant-produced omega-1 with $Man_3(\beta1\text{-}2)XylGlcNAc_2(\alpha1\text{-}3)$Fuc *N*-glycan. B: The surface-immobilised glycoprotein is a plant-produced omega-1 with $Man_3GlcNAc_2$ *N*-glycan. C: The surface-immobilised glycoprotein is a plant-produced omega-1 with $Man_3GlcNAc_2(\alpha1\text{-}6)$Fuc *N*-glycan. D: The surface-immobilised glycoprotein is a plant-produced omega-1 with $Man_3GlcNAc_2(\alpha1\text{-}3)$Fuc *N*-glycan.





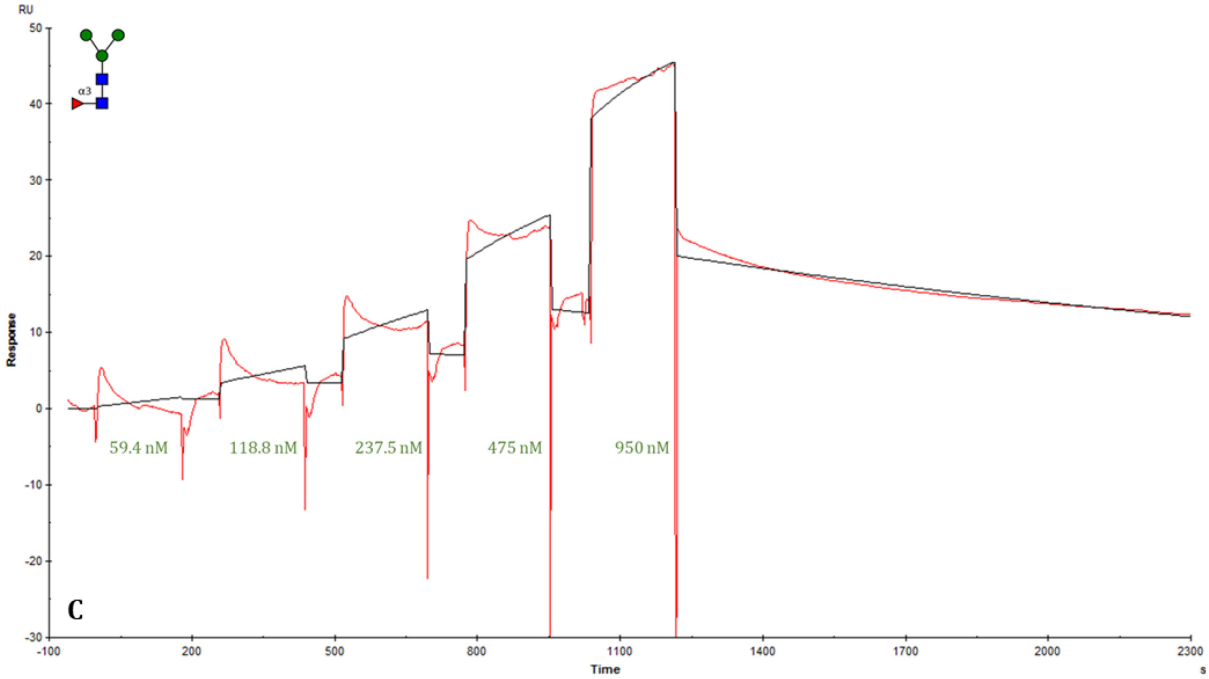


Figure 56. Global fitting of the SPR sensorgrams (Fc2-Fc1) representing the kinetic experiments using a concentration series of Mannitou mAb (purified). Single-cycle analysis. A: The surface-immobilised glycoprotein is a plant-produced omega-1 with $Man_3GlcNAc_2$ *N*-glycan. B: The surface-immobilised glycoprotein is a plant-produced omega-1 with $Man_3GlcNAc_2(\alpha1\text{-}6)Fuc$ *N*-glycan. C: The surface-immobilised glycoprotein is a plant-produced omega-1 with $Man_3GlcNAc_2(\alpha1\text{-}3)Fuc$ *N*-glycan.

2. Isothermal Titration Calorimetry

The binding affinity of Mannitou Fab to $Man_3GlcNAc_2$ has been further verified by means of Isothermal Titration Calorimetry. ITC is considered a gold standard for analysing intermolecular interactions. Apart from the accurate measurement of the equilibrium constants and bringing a highly precise evaluation of the binding stoichiometry of the interaction (molar ratio of reactivity), it provides valuable thermodynamic information represented by binding enthalpy and entropy change. A typical ITC experiment includes ligand titration into a solution containing the bio-macromolecule of interest (e.g., antibody); measurement of the heat released or absorbed upon binding; processing and fitting of the primary ITC data to obtain the dissociation constant (K_d), stoichiometry (n), binding enthalpy (ΔH) and entropy (ΔS), and binding free energy (ΔG) (Fisher et al., 1995).

The primary data is the power applied to the sample cell as a function of time (Fig.57 top) and consists of a series of peaks, with each peak corresponding to the thermal energy released upon binding of $Man_3GlcNAc_2$ to Mannitou Fab. During the first ligand injections, relatively large exothermic heat pulses are observed. After each successive injection, the magnitude of the pulses progressively decreases, until only peaks corresponding to dilution are observed, as ultimately all the binding sites on the antibody become occupied. The transformation of the raw ITC data results in obtaining the differential binding curve representing the heat change per mole injectant as a function of the molar ratio ligand/antibody (Fig.57 bottom). The interpretation of the ITC data involves fitting the binding isotherm to a standard single-site binding model, that assumes the antibody is not undergoing ligand-induced changes in self-association or thermal unfolding, to yield the parameters from a single experiment. Integration of the heat signals generates the binding enthalpy ΔH equal to - 7.12 ± 0.225 kcal/mol. The obtained enthalpy change combined with the binding affinity $K_d = 5.4 \times 10^{-5} \pm 3.24 \times 10^{-6}$ M and stoichiometry of the interaction n = 1, allow the calculation of the free energy of binding ΔG = -5.82 kcal/mol and entropy change $-T\Delta S$ = 1.3 kcal/mol.

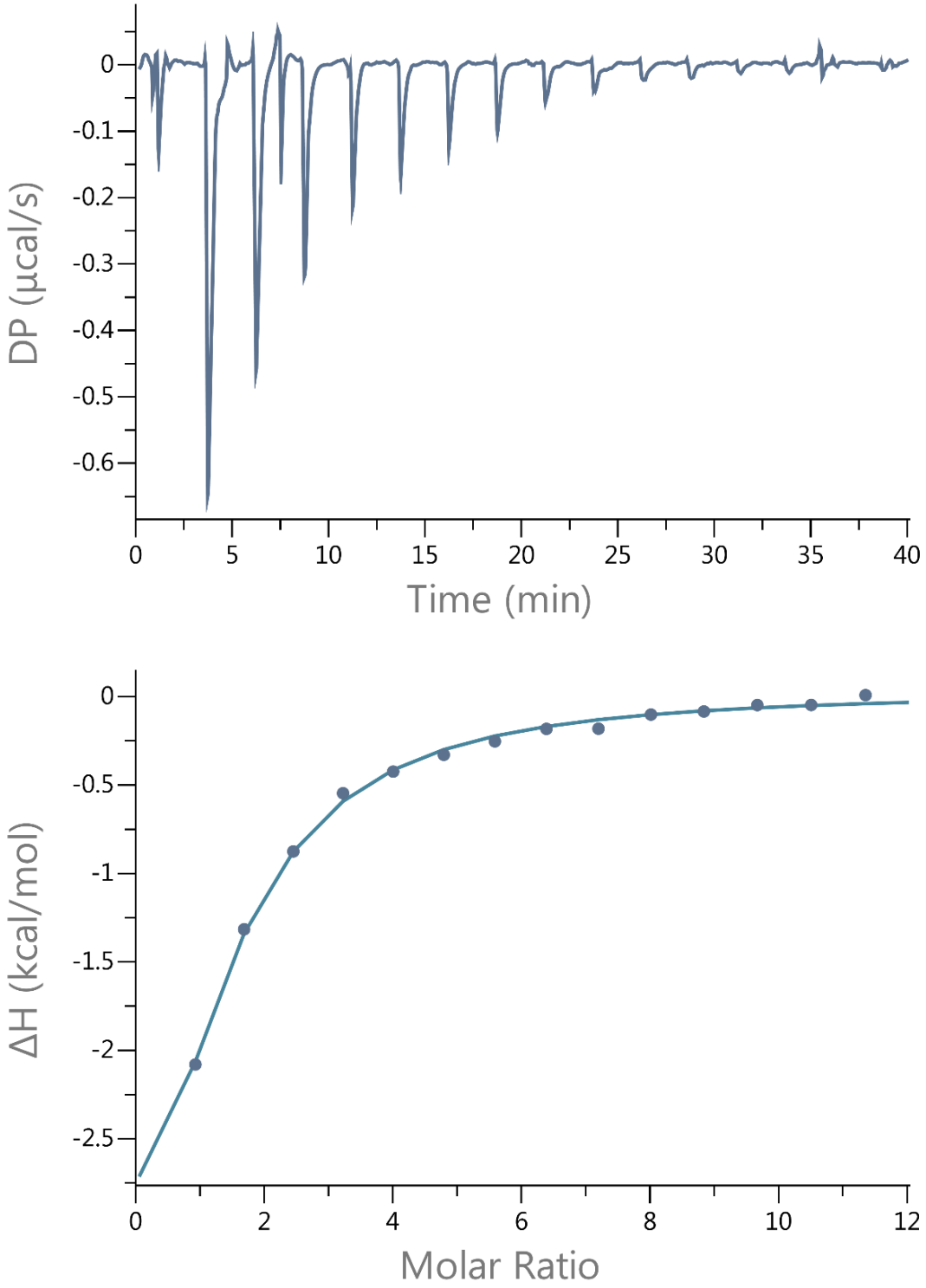


Figure 57. ITC data for titration of $Man_3GlcNAc_2$ into Mannitou Fab. Top: Baseline-substracted raw data thermogram showing power in µcal/s versus time in min. Each downward spikes corresponds to injection of 2µl of the glycan (3 mM) into the antibody solution (35 µM). All available binding sites on Mannitou Fab are saturated by the ligand after approx. 10 injections. Bottom: Integration of the spikes gives kcal/mole of the injected $Man_3GlcNAc_2$, plotted versus the molar ratio of $Man_3GlcNAc_2$ added per mole of Mannitou Fab in the sample cell. The derived binding isotherm represents best-fit of the data to a standard single-site binding model.

DISCUSSION

SPR technology has various advantages including real-time measurements of kinetic parameters of a wide range of interactants (small peptides to viruses), label-free format, small sample consumption, and high sensitivity to detect small molecular weight analytes (<500 Da), even at low-affinity interactions (K_d from 10^{-3} to 10^{-12} M). The regeneratable biosensor surface allows the use of the same sensor chip multiple times (Helmerhorst et al., 2012; McDonnell 2001).

When studying interactions between high molecular weight and low molecular weight substance, it is advantageous to immobilise the low molecular weight component (the glycan antigen) and let the other (the antibody) contribute to the detected mass change, as SPR technology senses changes in the refractive index which correlate with changes in mass concentration on the sensor surface. (Karlsson et al., 1991). A fundamental problem introduced when a ligand is immobilised on a surface is that it is physically separated from the analyte in the bulk solution. In a laminar flow cell, the transport of substance to the sensor surface has to pass the unstirred layer where only diffusion controls the mass transfer. It is crucial to ensure that the supply of analyte to the surface-attached ligand is not limited, otherwise, the analyte concentration near the sensor surface will be different from the bulk concentration, which can hinder the ability to determine rate constants for fast-binding interactions. The mass transport limitation occurs when the binding of the analyte to the ligand is faster than the diffusion of the analyte to the sensor surface. However, surface capacity and flow rate are two experimental parameters that can be adjusted to reduce mass transport effects. High flow rates reduce the height of the unstirred solvent layer, increasing the transport rate of the analyte to and from the sensor surface, and thus limiting the chance of analyte re-binding. Nevertheless, higher flow rates lead to increased sample consumption and for this reason, may not be very practical. The use of low-capacity surfaces, in turn, decreases the demand for the analyte, minimising concentration gradients in the flow cell, the aggregation, avidity, and crowding effects on the sensor surface. Measuring binding interactions on low-capacity surfaces, however, inherently leads to a lower signal-to-noise ratio and may not yield a reliable response. For a detailed kinetic analysis, it is important to improve the quality of the response data by removing as many instrument

artefacts as possible. Matrix effects, refractive index changes, non-specific binding, and instrument drift can all be corrected by using a reference surface, as during the Mannitou antibody – paucimannose *N*-glycans recognition studies. Another potential drawback of the SPR method is that the surface-attached ligand may alter the thermodynamics of binding by constraining the molecule's rotational and diffusional freedom. Additionally, the ligand may not maintain its native configuration upon immobilisation on the sensor chip surface, or alternatively, its non-homogenous orientation may sterically hinder analyte association. To overcome such incidents, the ligand glycan has been also immobilised on the sensor surface in the form of a glycoprotein carrying $Man_3GlcNAc_2$. Furthermore, this approach limits the possibility of non-specific binding, as Mannitou antibody has been found to manifest affinity to dextran. When injected on an empty flow cell, without prior activation and deactivation, a signal is detected, implying an interaction between the antibody and the chip surface. Moreover, the affinity of Mannitou to dextran has also been observed when using the Sephadex column for size-exclusion chromatography. In addition, several attempts have been made to immobilise the studied glycans/glycoproteins on other sensor chips, including Sensor Chip CM4 having a lower degree of carboxymethylation (30% compared to CM5) for decreased immobilisation capacity, and Sensor Chip C1 with a flat carboxylated surface without dextran matrix. The carboxyl groups are directly attached to the surface linker layer which restricts the mobility of the attached ligand and thus reducing the avidity effects observed with multivalent analytes, such as IgM. Unfortunately, all these efforts have been unsuccessful as no adequate immobilisation level has been achieved (Helmerhorst et al., 2012; Malmqvist 1993; Myszka 1997; Rich et al., 2002). Lastly, if a full-length antibody is used as the analyte, just like in the Mannitou mAb – $Man_3GlcNAc_2$ and Mannitou mAb – $Man_3GlcNAc_2$(α1-6)Fuc experiments, the multivalency of the antibody might be an issue. The most common and widely applied binding model is the 1:1 Langmuir model, where the association and dissociation phases of a real-time binding curve are fitted to two separate single-exponential equations. However, with this simple model, the derived interaction parameters may be significantly different from those obtained from solution-based methods such as isothermal titration calorimetry, due to the assumption that there is only one binding site on the antibody (Sun et al., 2017). The adoption of one-to-one binding model, together with possible avidity effects, might explain the significantly lower K_d values for the experiments using the full-length Mannitou antibody compared

to the Fab. All these restraints may raise concerns over the calculated kinetic values and therefore, it has been of crucial importance to compare the obtained SPR parameters against values generated for the same interaction studied with a solution-based technique. For this purpose, Mannitou Fab – $Man_3GlcNAc_2$ analysis has been performed employing isothermal titration calorimetry.

Isothermal Titration Calorimetry has gained a wide range of applications in the field of biophysics, biochemistry, biotechnology, protein engineering, and drug discovery, design and development (Bouchemal 2008; Ghai et al., 2012; Krell 2008; Rajarathnam et al., 2014), as a powerful tool to characterise interactions of biomolecules in a broad spectrum of binding affinities. The structural and dynamic data alone, even when coupled with the most sophisticated computational methods, cannot provide information about the complete thermodynamic profiles consisting of the binding free energy, enthalpy, and entropy, and therefore may not accurately predict the binding affinity (Chaires 2008). ITC is the only technique able to provide quantitative thermodynamic data that is extremely valuable to study the complex stability and elucidate the forces driving the binding process. In addition to its high precision, sensitivity, and reproducibility allowing analysis of weak to high-affinity interaction (K_d from 10^{-2} to 10^{-9} M), it is a non-destructive method that requires no immobilisation, modification, or labelling of the binding partners, allowing the reactants to be studied in their native form and eliminating the risk of any perturbations of the delicately balanced functional chemistry of the biological macromolecules. Additionally, ITC does not set any molecular weight limitations. Moreover, reactants and products do not have to be physically separated after binding equilibrium has been reached as the separation step can perturb the position of equilibrium, especially if off-rates are fast. Furthermore, ITC ensures direct and accurate measurements of equilibrium constants, as opposed to kinetic procedures that access these indirectly, based on interpretation of on- and off-rate processes (Doyle 1997; Doyle et al., 1997).

Nevertheless, there are yet certain disadvantages to the ITC approach. Since heat is a universal signal and each process contributes to the measured global heat effect, the evaluation of the contribution from the binding can be complicated and control experiments need to be performed. Despite the high sensitivity, challenges still exist for extracting heat effects upon complex formation when the binding interaction exhibits a rather small enthalpy change (e.g. non-covalent complexes), resulting in a relatively low

signal to noise ratio. Additionally, when a kinetic process is very slow, it may be overlooked (Doyle 1997; Doyle et al., 1997). The main drawback of ITC when studying Mannitou – paucimannose *N*-glycan interactions, however, is that rather high concentrations of the reactants are needed as the binding affinities are only moderate. This considerably limits the technique's application as it has been extremely time-consuming to prepare both Mannitou Fab and mAb in larger quantities. Moreover, higher concentrations also lead to a probability that self-association of the antibody may occur (Wang et al., 2018). Therefore, it has been very problematic to conduct a reliable ITC experiment with the whole pentameric Mannitou IgM, as no adequate concentration level could be achieved.

If the concentrations of both the protein and the ligand are known, ITC can determine the binding stoichiometry (n) from the molar ratio of the interacting partners at the equivalence point. In data fitting, the parameter n can either be fixed as the number of binding sites per macromolecule or be treated as an additional floating parameter determined from iterative fitting (Perozzo et al., 2004).

The thermodynamic profile indicates the predominant forces that drive the binding interaction. Since the protein-ligand binding event is induced by the decrease in total Gibbs free energy of the system, the change in enthalpy and entropy are the driving forces of this process. The contributions of ΔH and ΔS to ΔG are closely related. The ITC results reveal that the association between Mannitou Fab and $Man_3GlcNAc_2$ (ΔG = -5.82 kcal/mol) is an enthalpy-driven process (ΔH = -7.12 ± 0.225 kcal/mol) counterbalanced by an unfavourable contribution from entropy (-TΔS = 1.3 kcal/mol). However, the favourable negative enthalpy change is large enough to overcompensate for the entropic cost and the total free energy of binding is still a negative value. The enthalpy-dominated induced-fit model (Koshland 1958; Koshland 1959) is relatively realistic and highly probably in the case of Mannitou Fab – $Man_3GlcNAc_2$ interaction since it takes into account the protein conformational flexibility, especially the conformational changes surrounding the binding site. For the binding to take place under the induced fit model, the lack of perfect surface complementary between the binding partners requires multiple tentative collisions to achieve the appropriate shape match between the interacting sites. The initially established contacts between the imperfectly matched surfaces, resulting in negative enthalpy change, should ensure enough strength and longevity of the maturation complex. -, ultimately leading to maturation of the complex

into a fully-bound complex. In addition, this process is accompanied by the release of water molecules from the binding site upon the protein-ligand association. This results in a positive enthalpy change, due to the desolvation energy penalty and the disruption of the original non-covalent interactions surrounding the binding sites, but also in a solvent entropy gain, that in turn contributes favourably to the stability of the newly-formed complex. However, the net entropy change of binding is determined by three entropic parameters, ΔS_{solv}, ΔS_{conf}, and $\Delta S_{r/t}$. In the case of the induced-fit binding, it can be speculated that the ΔS_{conf} value is negative since the formed favourable non-covalent interactions between the binding partners restrict the conformational freedom of the interacting interfaces. It is highly probable that the $\Delta S_{r/t}$ value is also negative due to the loss of rotational and translational degrees of freedom of the binding partners upon association. An unfavourable ΔS_{conf} together with an unfavourable $\Delta S_{r/t}$, tend to compensate for the positive ΔS_{solv} value, ultimately leading to a relatively small net entropy change compared to the net enthalpy change. Indeed, in most induced-fit binding examples (Chang et al., 2004; Corbett et al., 2015; Hariharan et al., 2014), the net entropy changes are all negative values. Nevertheless, it is still possible that the net entropy change could be a positive value contributing favourably to the lowering of the system's free energy. However, the magnitude of the net negative enthalpy change is larger, and thus substantially contributing to the binding by lowering of the system free energy in the induced-fit binding (Du et al., 2016; Li et al., 2014).

In the induced-fit binding process, the formation of new non-covalent interactions between the protein and the ligand are accompanied by the disruption of previous interactions at the binding sites. In order to maintain a stable association of the binding partners in the maturation complex, the favourable enthalpy contribution (negative ΔH), resulting from the newly established interactions, should be large enough to overcompensate for not only the positive enthalpy change, resulting from the disruption of the original interactions, but also the possible unfavourable entropy contribution (negative ΔS).

The thermodynamic parameters portray the molecular-level interactions that drive binding and provide a more detailed insight into the mechanism than is available from the binding affinity alone. One gains a great deal of validation when independent methods, such as the applied SPR and ITC measurements, yield the same value for a specific parameter. In the case of Mannitou Fab – $Man_3GlcNAc_2$ binding event, the

equilibrium dissociation constant calculated from the rate constants (K_d = 48 μM) agrees well with the value determined in solution using titration calorimetry (K_d = 54 μM) providing further support for the kinetic analysis. Additionally, medium micromolar affinity is ideal to further STD-NMR experiments aimed at identifying the exact binding epitope. High-quality data for the kinetic and thermodynamic parameters of intermolecular interactions gives insight into the mechanisms of molecular binding events, which is an essential complement to the structural data in rational drug design and for the optimisation of lead compounds.

Chapter 2.

Molecular recognition studies using Mannitou antibody

III. Epitope mapping by Saturation Transfer Difference Nuclear Magnetic Resonance Spectroscopy

INTRODUCTION

Specific interactions between molecules in solution through non-covalent forces are the fundamental basis of molecular recognition processes. In all cases, the specificity relies on both structural and functional complementarities between the chemical groups of the interacting molecules. Nuclear magnetic resonance (NMR) spectroscopy is an extremely useful technique to investigate transient protein-carbohydrate interactions in solution. It has become an essential tool for characterising events of molecular recognition and obtaining valuable information about the interactions of low-molecular-weight ligands with biologically significant macromolecules, such as antibodies (Angulo et al., 2011).

Saturation transfer difference spectroscopy STD-NMR experiments are based on the nuclear Overhauser effect (NOE) and the observation of the ligand resonance signals. It can be applied as a ligand screening method, to determine in a qualitative manner the binding activity of a compound towards a receptor protein, or as a medium for ligand mapping, identifying moieties critical for binding. The hydrogen atoms of the ligand being in closest contact with the receptor's binding pocket upon attachment show the most intense STD-NMR signals, enabling the mapping of the ligand binding epitope. STD-NMR may be also employed to establish the dissociation constant K_d between the ligand and the receptor. STD-NMR exploits the fact that for relatively weak binding interactions (K_d from 10^{-3} to 10^{-8} M), there is a fast exchange between the free and the bound ligand state leading to a very efficient build-up of saturation of the ligand molecules in solution. STD experiments are very sensitive allowing the detection of very weak interactions, and thus suitable for studying mAb–carbohydrate interactions (Viegas et al., 2011).

Macromolecules, like proteins, consist of a large system of protons tightly coupled by dipole-dipole interactions. The longitudinal relaxation rate of protein protons is dominated by the cross-relaxation rate σ_{intra}. By irradiation of the protein at a resonance where no ligand signals are present, the intramolecular transfer of magnetisation within the protein, called spin diffusion, leads to the fast and uniform saturation of the entire protein (Kalk et al., 1976). When a protein becomes saturated, the intermolecular transfer of magnetisation σ_{inter} from the protein to the binding epitope of the ligand leads to the progressive saturation of the ligand molecules bound to the protein. Upon

dissociation, the saturation of the ligand is transported into the solution where it is detected (Mayer et al., 1999; Mayer et al., 2001).

An STD analysis involves subtracting a spectrum in which the receptor (protein) is selectively saturated (on-resonance spectrum) from a spectrum recorded without protein saturation (off-resonance spectrum) (Fig.58). The on-resonance spectrum is obtained by setting the irradiation frequency at a value where only resonances from the protein nuclei and no resonances from the ligand nuclei are located. This is possible because protein resonances have significant intensity even in the negative ppm region or in the downfield region above 10 ppm outside of the spectral window of low-molecular-weight molecules. The reference off-resonance spectrum is obtained under conditions of thermal equilibrium by applying the irradiation pulse at a frequency where none of the protons resonates, neither of the protein nor of the ligand. Subtraction of the on-resonance from the off-resonance results in a difference spectrum in which only signals of the ligand moieties that received progressive saturation transfer from the protein via spin diffusion, through intermolecular NOE, will remain. Other compounds that may be present in the mixture but do not bind to the receptor will not receive any saturation transfer. Their signals will be of equal intensity in the on-resonance and the off-resonance spectra and, as a consequence, after subtraction, no signals will appear in the difference spectrum (Meyer et al., 2003; Viegas et al., 2011). As saturation is exclusively transferred to molecules bound to the protein, the difference spectrum contains only signals of molecules with a binding affinity. For a molecule that binds to the receptor, only the signals of the hydrogens that are in close contact with the protein and receive magnetisation transfer will appear in the difference spectrum. From those, the ones that are nearest to the protein will display more intense STD signals, owing to a higher degree of saturation. The ligand protons that are further from the target surface will be saturated to a lower degree and their STD intensities will be weaker. Provided that all ligand hydrogens have similar relaxation rates, the degree of saturation received by individual ligand protons reflects the relative proximity of these to the receptor's binding site (Angulo et al., 2011; Yan et al., 2003).

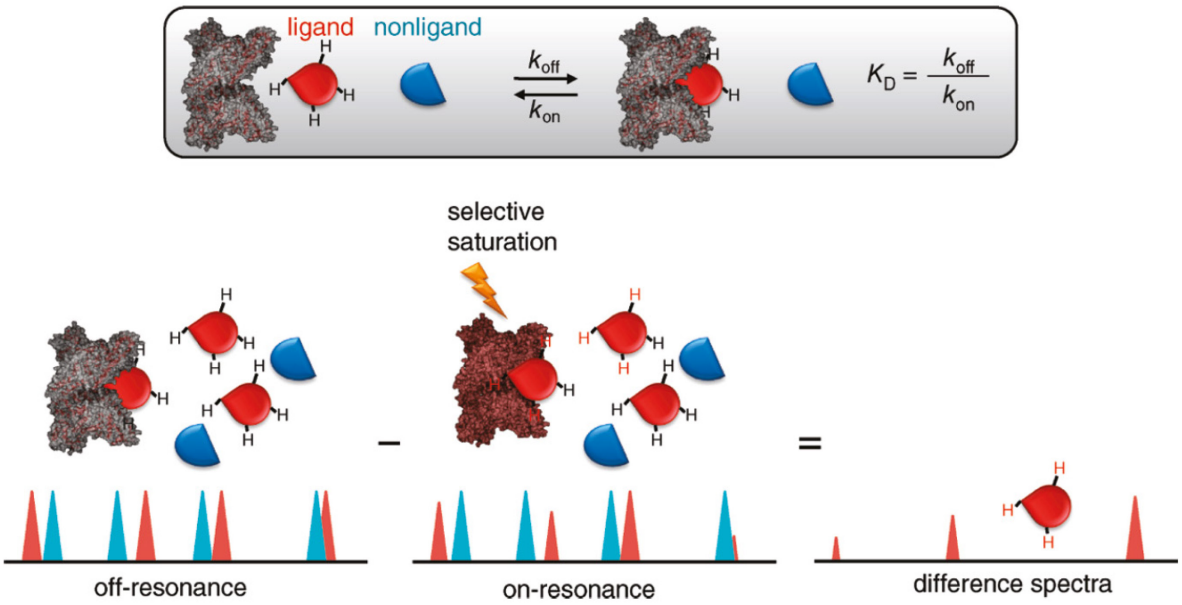


Figure 58. STD-NMR principle. From: Viegas et al., 2011

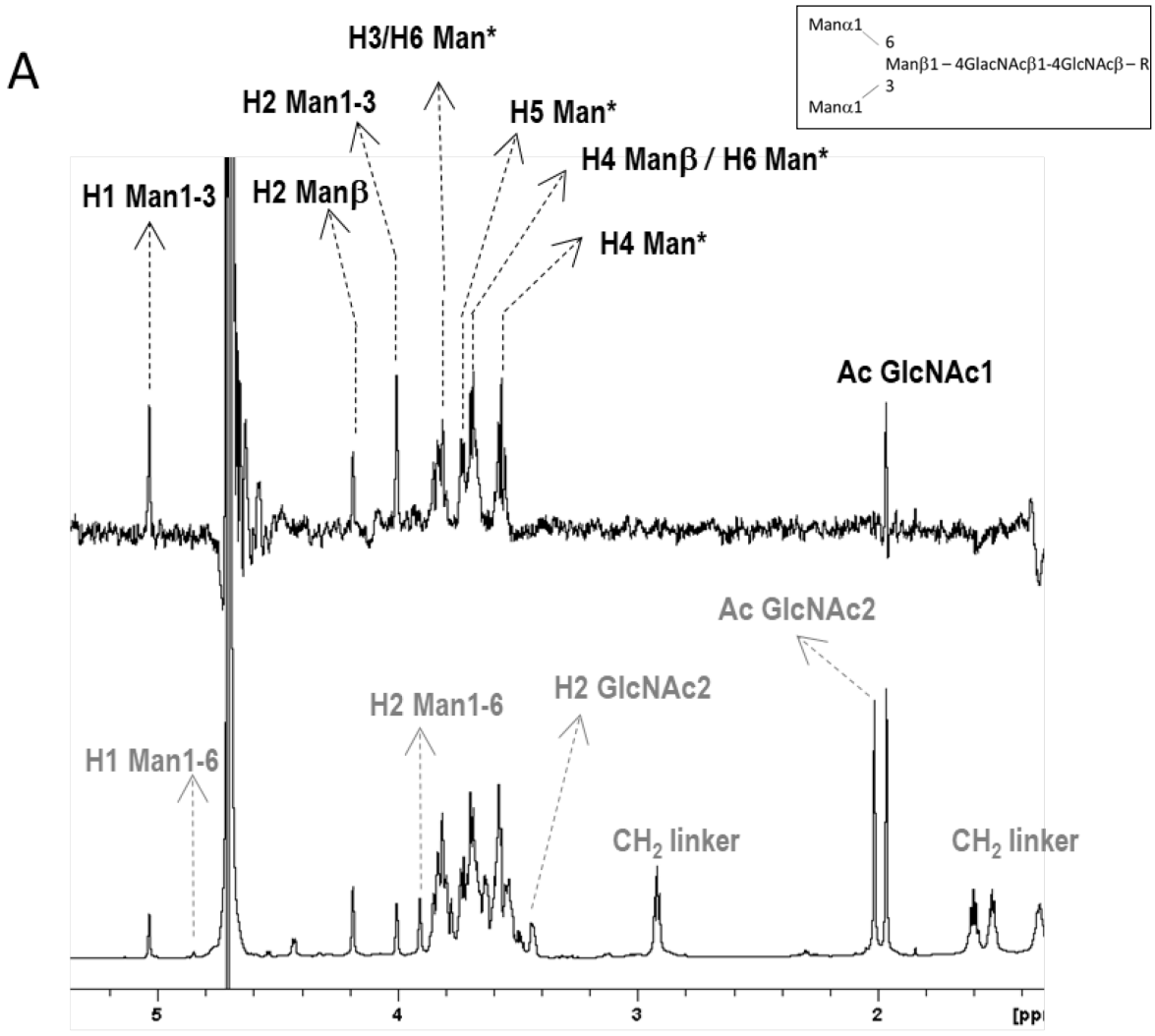
MATERIALS & METHODS

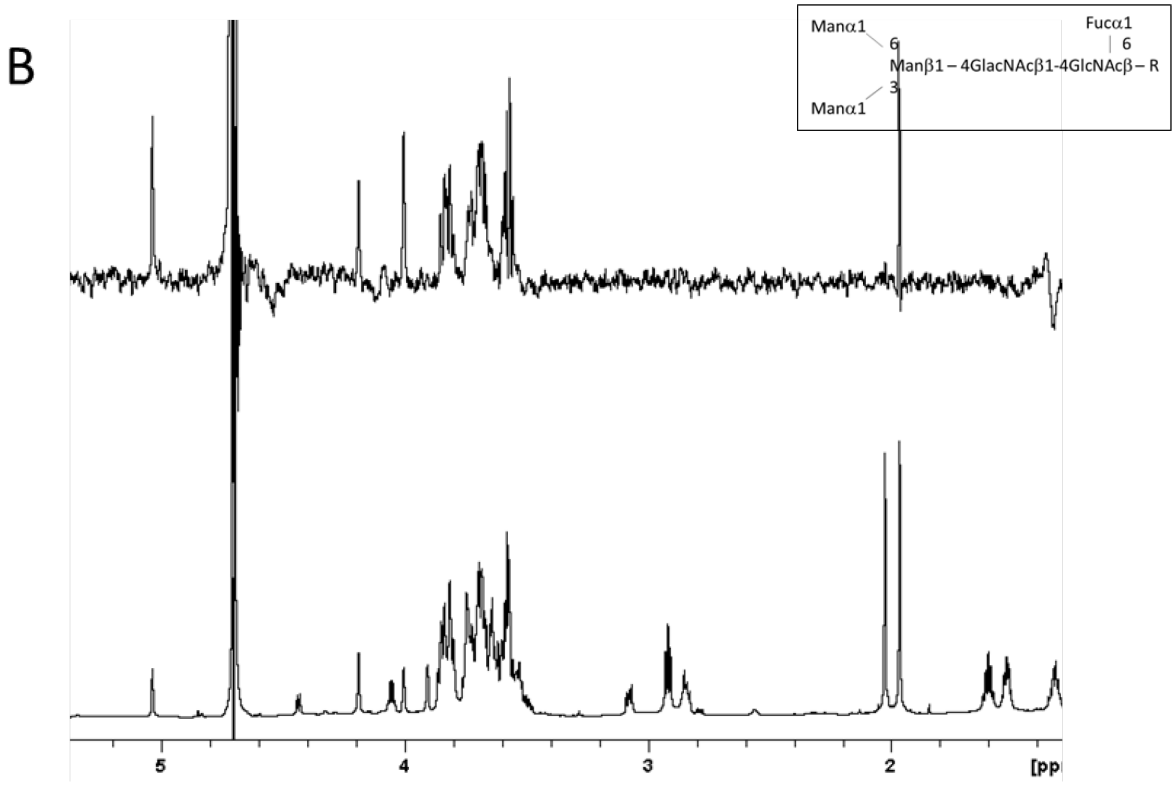
^{1}H-STD NMR experiments have been performed at 25°C on a Bruker AVANCE III spectrometer operating at a proton frequency of 800 MHz and equipped with a cryoprobe. The Mannitou Fab – $Man_3GlcNAc_2$ and Mannitou Fab – $Man_3GlcNAc_2(\alpha 1\text{-}6)Fuc$ samples have been prepared in PBS buffer in D_2O (pD=7.4) in 3 mm NMR tubes. The measurements are taken using a 40:1 ligand to receptor molar ratio with a 1 mM concentration of the respective oligosaccharide and a 25 µM concentration of the antibody. Off-resonance irradiation frequency is set at 100 ppm, where no protein signals are present. Two on-resonance irradiations of the protein are performed at a chemical shift of 0.8 ppm (for the aliphatic region of the protein signals) and 6.9 ppm (for the aromatic region of the protein signal). In order to achieve the desired selectivity and to avoid side-band irradiation, protein saturation has been achieved with a train of Gaussian-shaped pulses of 50 ms duration, with a total irradiation time of 2 s. STD spectra are obtained by subtracting the on-resonance spectra from the off-resonance (reference) spectrum. A separate STD spectrum has been acquired with the antibody alone, in the absence of the glycan ligand, under the same conditions, and subtracted from the STD spectra in the presence of the carbohydrate to obtain saturation transfer double-difference (STDD) spectra. In this way, the resulting STDD spectra are free from the antibody signals in the STD spectra and do not interfere with the analysis of the glycan protons. STD spectra of the free oligosaccharides at 1 mm in D_2O have been acquired as well. These spectra have also been subtracted from the STD spectra of either of the glycans in the presence of the Fab, to ensure that no direct irradiation of the ligand is taking place. Data acquisition and processing are performed with TopSpin 3.0 software (Bruker, Germany) and the figures are built using MestReNova v.8.0.2.

The STD-NMR analysis has been performed by S. Robakiewicz and A. Arda in the lab of J.Jimenez-Barbero.

RESULTS

^{1}H-STD NMR experiments have been performed using two paucimannose *N*-glycans selected from the microarray screening: $Man_3GlcNAc_2$ and $Man_3GlcNAc_2$(α1-6)Fuc. For both the fucosylated and the non-fucosylated glycans, the STD spectra at the aliphatic and the aromatic irradiation frequencies are essentially identical. At the same time, both ligands display very similar STD spectra in which the same protons show analogous STD effects (Fig.59A, Fig.59B). For both of the ligands, the protons showing the strongest STD effect are H1 and H2 of the Manα1-3 residue. On the other hand, the same protons of the Manα1-6 branch demonstrate no STD effect. H2 of the branched βMan residue shows a weaker STD effect. In the more crowded region of 3.6 - 3.9 ppm, an STD effect is also observed for H3, H4 and H5 protons of the αMan residue. Although these protons resonate at the same frequency for Manα1-6 and Manα1-3 residues, the above-mentioned results involving H1 and H2 suggest that they correspond to the Manα1-3 residue. Furthermore, while the Ac protons of the distal GlcNAc residue (GlcNAc2) exhibited no STD effect, the Ac protons of the reducing end GlcNAc residue (GlcNAc1) showed a weak STD effect, which probably results from the presentation of the glycan, where the acetyl group of GlcNAc1 points to the same side as the Manα1-3 residue, while the acetyl group of GlcNAc2 points in the opposite direction (Fig.59C). A lower relative STD value indicates that this Ac is more distant from the antibody that the protons directly involved in the binding. Interestingly, no STD effect was detected for the fucose residue of the core fucosylated glycan, which is in full agreement with the observations that its presence does not affect the binding interaction.





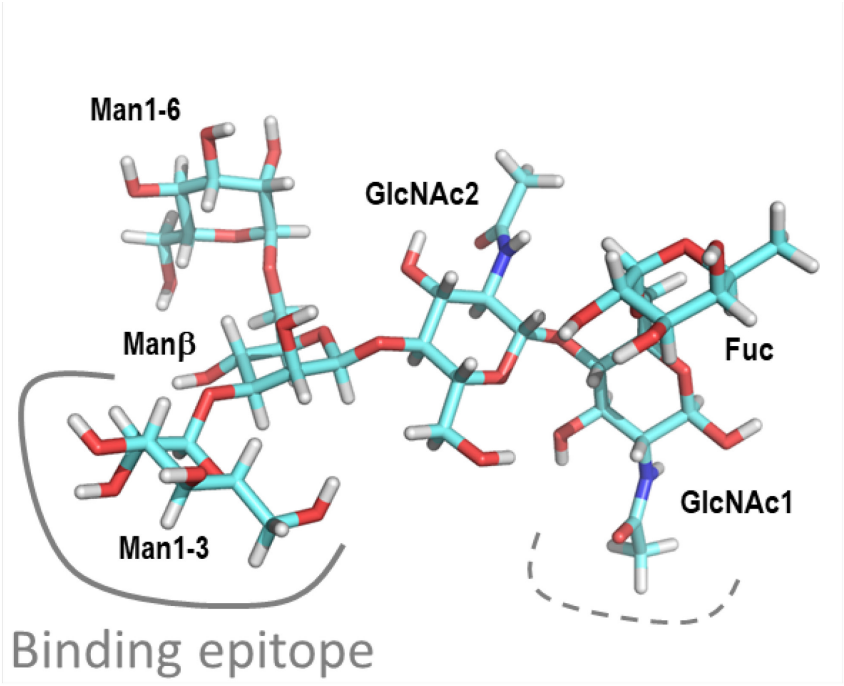


Figure 59. Binding epitope studied by ^{1}H-STD-NMR. (A) ^{1}H-STD experiment for a sample of Mannitou Fab (25µM) with $Man_3GlcNAc_2$ (1mM). Top: STD spectrum, below: reference spectrum (off-resonance). Protons of the ligand showing STD effect (top) are labelled in black, while some of the protons (with non-overlapping signals) showing no STD are labelled in grey (below). * indicates protons corresponding to both Manα1-3 and Manα1-6 residues (overlapping). (B) Same ^{1}H-STD experiment using a sample with the fucosylated *N*-glycan $Man_3GlcNAc_2$(α1-6)Fuc. The STD spectrum is exactly the same as for the non-fucosylated *N*-glycan. (C) Schematic representation of the binding epitope for $Man_3GlcNAc_2$(α1-6)Fuc on a 3D structural model according to STD-NMR. All protons belonging to the Manα1-3 residue show STD, with a very weak contribution of the Ac group on the GlcNAc1 residue.

DISCUSSION

Specific characterisation of which parts of the ligand are in direct contact with the protein is currently left largely to X-ray analyses of co-crystallised ligand-receptor complexes. The advantage of STD-NMR spectroscopy in comparison to other screening techniques used, such as glycan arrays, SPR, or ITC, is the ability to identify the ligand moieties directly involved in the binding to the receptor, the so-called binding epitope. STD-NMR experiments are carried out in solution and can provide detailed and highly relevant information about changes in the conformation of the carbohydrate ligand upon binding to the antibody receptor. The use of STD-NMR allows obtaining information about the binding specificity at the atomic level.

STD-NMR is an extremely sensitive method, therefore really small amounts of the receptor are required. It allows using as little as 1 nmol of protein with a molecular weight >10 kDa. In fact, there is no limit to the size of the protein as an increase of sensitivity is observed along with the increase of the protein size due to a more efficient inter- and intramolecular saturation transfer. STD-NMR technique is insensitive to ligand excess and eliminates the risk of detecting false positive interactions since only saturation transferred to the glycan through the protein can be detected. Additionally, neither the ligand nor the receptor needs to be isotopically labelled or modified for the binding to be detected. The only restrictions set are the long acquisition times and the need for ligand excess. The excess of ligand molecules and longer irradiation/saturation time result in a stronger STD signal, although even at short irradiation times the entire receptor is efficiently saturated. When possible, higher field spectrometers should be used for greater efficiency as sensitivity and spin diffusion increase with field strength (Mayer et al., 1999; Mayer et al., 2001).

STD-NMR effects depend largely on the off-rate. Effective mapping of the ligand binding epitope is possible if the dissociation rate constant k_{off} is greater than, or of the same magnitude as the intermolecular saturation transfer rate within the bound ligand, as the ligand has to leave the binding site before all magnetisation has been equally distributed among all spins in the ligand. Very small k_{off} rates (slow dissociation) of strongly binding molecules cause the ligand to stay in the protein binding pocket for an extended period of time causing the magnetisation to spread over many ligand protons, which in turn results in low signal intensity and reduced sensitivity. If the k_{off} value is

high, the saturation is quickly transferred into the solution. When the ligand is used in large excess, the saturation of free ligands in solution gets amplified because the relaxation of small molecules is slower than the saturation transfer. Large saturation transfer effects are therefore observed for high k_{off} values and long relaxation times, making the experiment more sensitive (Becker et al., 2018; Mayer et al., 2001; Meyer et al., 2003).

The micromolar binding affinity (K_d=48-54 μM) of Mannitou Fab to paucimannose *N*-glycans, Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAcβ1-N and Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4(Fucα1-6)GlcNAcβ1-N, established from the SPR and ITC experiments is in the suitable range for an STD-NMR analysis. The performed spectroscopic studies using solutions of Mannitou Fab with a 40-fold glycan excess ($Man_3GlcNAc_2$ or $Man_3GlcNAc_2$(α1-6)Fuc) has demonstrated the specificity of the antibody for the Manα1-3 residue and, to a lesser extent, the acetyl group of GlcNAc1, revealing that in both cases the binding epitope is identical. Moreover, the STD intensities are exactly the same, indicating that the interactions must be very similar also in terms of affinity. This is in accordance with previous SPR results where Mannitou mAb has been demonstrated to bind both $Man_3GlcNAc_2$ and $Man_3GlcNAc_2$(α1-6)Fuc with nanomolar affinity. Glycan protons that are nearest to the antibody binding site can be easily identified in the difference spectrum as they obtain the highest degree of saturation directly from the receptor. Since the hydrogens showing the strongest signal intensities are H1 and H2 of the Manα1-3 residue, they are the ones being in closest contact with the antibody, and therefore saturated to the highest degree. The H3, H4, and H5 protons of Manα1-3 are also directly involved in the recognition of Mannitou Fab. However, because of signal overlap, their individual contributions to the binding cannot be determined. Finally, the acetyl group of the reducing end GlcNAc residue apparently forms a less intense interaction with the antibody as it produces STD signals of relatively lower intensity. These results presenting the Manα1-3 residue as the main binding epitope, which are in accordance with the previous glycan array analysis pointing to Manα1-3 branch as essential for the binding, explains why any substitution in Manα1-3 arm inhibits the interaction of paucimannose *N*-glycans with Mannitou antibody and confirm that the minimal glyco-epitope recognised by the antibody is the Manα1-3Manβ1-4GlcNAcβ1-4GlcNAcβ1-N structure.

Chapter 3.

Mannitou IgM characterisation

MATERIALS & METHODS

1. Homology modelling of Mannitou Fab

Mannitou Fab was modelled using Phyre^2 server for protein modelling, prediction and analysis (Kelley et al., 2015). HC model has been based on c1clyH template (IgG Fab (human IgG1, kappa). LC model has been based on c1igfM template (IgG1 kappa b13i2 Fab light chain).

Mannitou variable domains have been modelled using PIGSPro server (Lepore et al., 2017). Templates for the heavy chain: for FRH – 1uz6, for H1 – 1uz6, for H2 – 1uz6, for H3 – 1c5d. Templates for the heavy chain: for FRL – 1t66, for L1 – 1t66, for L2 – 1t66, for L3 – 1t66.

2. *N*-glycosylation sites prediction

Potential *N*-glycosylation sites on the full-length monoclonal antibody as well as on the antigen-binding fragment have been predicted using NetNGlyc server (Gupta et al., 2004).

3. Single-particle reconstruction using Cryo-EM

In order to avoid particle overlap, 4 µl of Mannitou mAb at 0.05 mg/mL concentration has been pipetted onto 200-mesh Quantifoil R 2/2 holey-carbon grids and vitrified using Vitrobot Mark III (FEI Company, USA) in liquid ethane. The micrographs are recorded on a JEM-2200FS/CR (JEOL Ltd, Japan) electron microscope operating at 200 kV (4K x 4K CCD camera; magnification 50 000 x; pixel size 2 Å/pixel). The dataset includes 153 733 picks from 610 cryo micrographs.

4. Mannitou Fab crystallisation

Several different crystallisation trials have been set up.

The first crystallisation of Mannitou Fab was set up in July 2018 using five different crystallisation conditions, namely $AmSO_4$ (Fig.S3), Index (Fig.S4), Midas (Fig.S5), Morpheus (Fig.S6), and PACT (Fig.S7). A concentration of 78.8 µM Mannitou Fab (6.8 mg/mL) in 20 mM Tris 150 mM NaCl buffer, pH=7.4, has been reached. The set-ups have been kept at 20°C in Crystal Farm and regularly checked for crystal formation.

The first crystallisation of Mannitou Fab – $Man_3GlcNAc_2$ complex was preceded with the Pre-Crystallisation Test (PCT) used to determine the appropriate antibody concentration for crystallisation screening. The PCT involves testing four different reagent conditions:

- Reagent A1 - 0.1 M TRIS hydrochloride pH 8.5, 2.0 M ammonium sulphate
- Reagent B1 - 0.1 M TRIS hydrochloride pH 8.5, 1.0 M ammonium sulphate
- Reagent A2 - 0.1 M TRIS hydrochloride pH 8.5, 0.2 M magnesium chloride hexahydrate, 30% w/v Polyethylene glycol 4,000
- Reagent B2 - 0.1 M TRIS hydrochloride pH 8.5, 0.2 M magnesium chloride hexahydrate, 15% w/v Polyethylene glycol 4,000

The crystallisation of Mannitou Fab – $Man_3GlcNAc_2$ was set up in November 2018 using Morpheus crystallisation conditions. A concentration of 82.3 µM Mannitou Fab (7.1 mg/mL) in 20 mM HEPES 150 mM NaCl buffer, pH=7.4, has been reached. The set-ups have been kept at 20°C in Crystal Farm and regularly checked for crystal formation.

Another co-crystallisation trial was performed using both $Man_3GlcNAc_2$ and $Man_3GlcNAc_2(\alpha 1\text{-}6)Fuc$ under five different crystallisation conditions, namely JCSG (Fig.S8), MemGold (Fig.S9), Midas, PACT, and PEG. A concentration of 115.9 µM Mannitou Fab (10 mg/mL) has been reached. The set-ups have been kept at 20°C and regularly checked for crystal formation.

The last crystallisation of Mannitou Fab – $Man_3GlcNAc_2$ complex was set up in June 2019 using two different crystallisation conditions, Morpheus and PACT. A concentration of 130.9 µM Mannitou Fab (11.3 mg/mL) in 20 mM HEPES 150 mM NaCl buffer, pH=7.4, has been reached. The set-ups have been kept at 20°C in Crystal Farm and regularly checked for crystal formation.

The homology modelling and *N*-glycosylation sites prediction have been done by S. Robakiewicz. The single-particle reconstruction using Cryo-EM has been attempted by S. Robakiewicz, D. Charro, and N. Abrescia. The Mannitou Fab has been set up for crystallisation by S. Robakiewicz with the help of A. Rojas, and simultaneously by J. Bouckaert and C. Bridot.

RESULTS

1. Homology modelling of Mannitou Fab

The homology modelling of the heavy chain of Mannitou Fab (Fig.60 top) revealed that the predicted structure ends abruptly, in the middle of the β-strand. Therefore, for the purpose of cloning and correct expression in HEK293T cells, the heavy chain sequence has been elongated with seven additional amino acids from the beginning of the CH2 domain sequence (Fig.61 bottom) and superimposed on an IgG1 Fab structure (PDB: 5L9D) (Fig.32) to determine if the elongation is sufficient. The sequence of the light chain (Fig.60) remained intact as its sequence length is identical in mAb and Fab form. Only when having the crystal structure, a validation of the predicted Mannitou Fab model can be made.

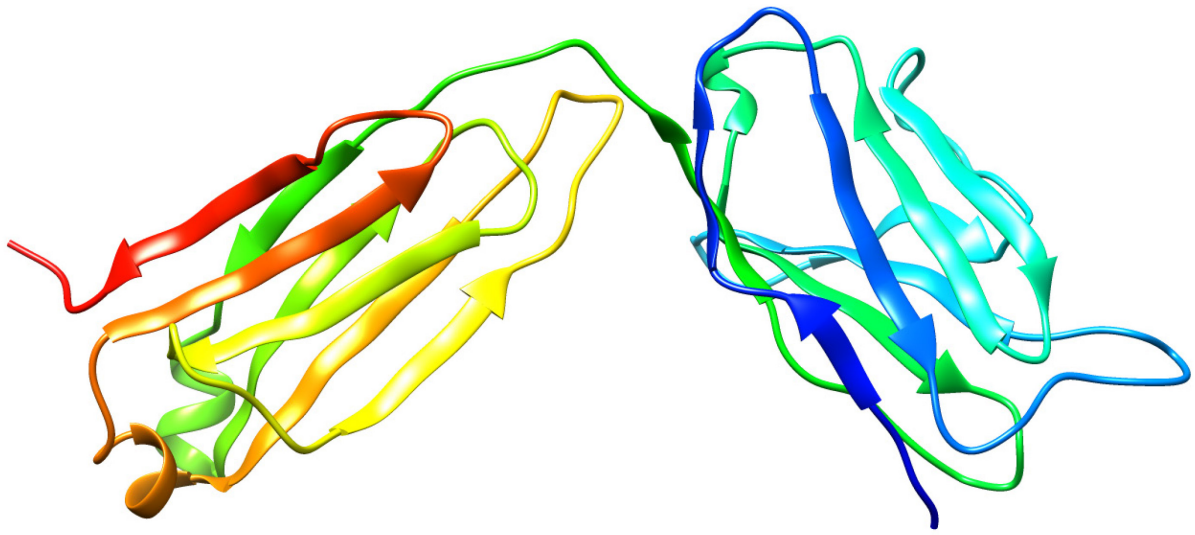


Figure 60. Mannitou Fab light chain modelled with Phyre².

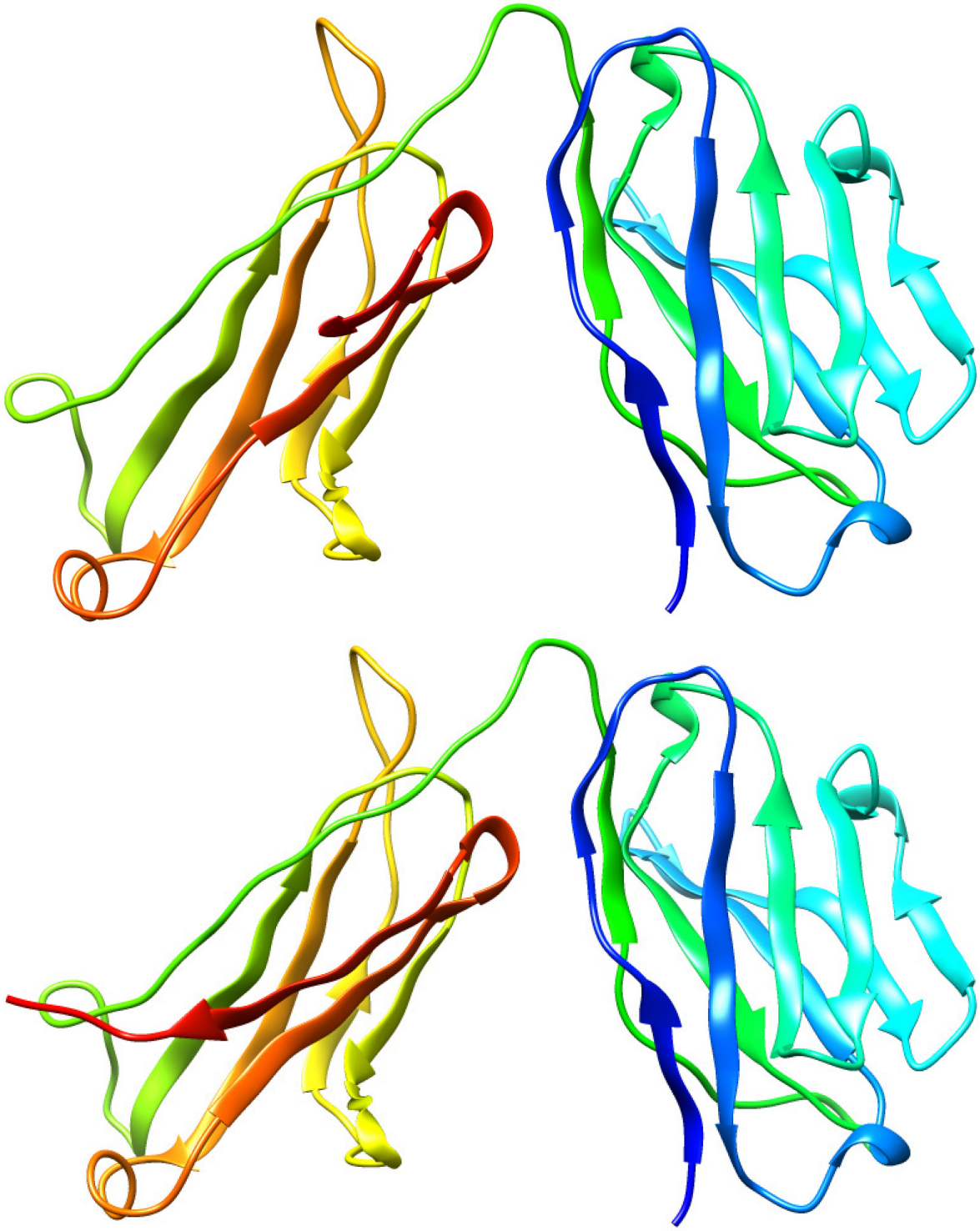


Figure 61. Mannitou Fab heavy chain modelled with Phyre2. Top: Mannitou Fab modelled using the initial heavy chain sequence. Bottom: Mannitou Fab modelled after elongating the initial heavy chain sequence with 7 additional amino acids.

1.2. Variable domains structure prediction

The main chains of the light and heavy chain framework are modelled by homology using, as a template, the corresponding domain from a similar immunoglobulin (Fig.62). The L1-L3 and H1-H2 loops are modelled by inheriting their conformation from an immunoglobulin with the same canonical structure (i.e. bearing in key positions the same amino acids as the target protein). The H3 loop is modelled using loops with highest sequence similarity to the target available. The heavy and the light chain are packed together. The side chains of the molecule are predicted using classical methods for side-chain conformation prediction such as SCWRL4 (Lepore et al., 2017).

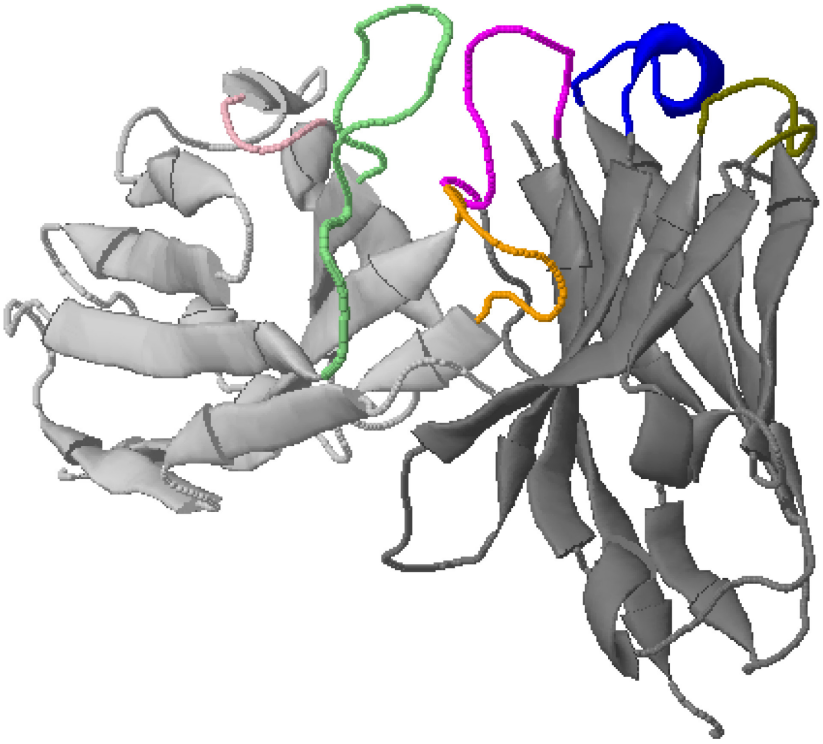


Figure 62. Mannitou Fab variable domains modelled with PIGSPro. CDRs are visualised in colour.

Heavy chain target - template alignment:

target:
EVKLLESGGGLVQPGGSLKLSCAASGFDFST-------YWMSWVRQAPGKGLEWIGEINP-------
DSSTINYTPSLKDKFIISRDNAKNTLYLQMSKVRSEDSVLYYCVRPGTWGYF----------
DYWGQGTTLTVSS

template:
EVKLLESGGGLVQPGGSQKLSCAASGFDFSG-------YWMSWVRQAPGKGLEWIGEINP-------
DSSTINYTPSLKDKFIISRDNAKNTLYLQMSKVRSEDTALYYCAREDGWNYF----------
DYWGQGTTLTVSS

Light chain target - template alignment:

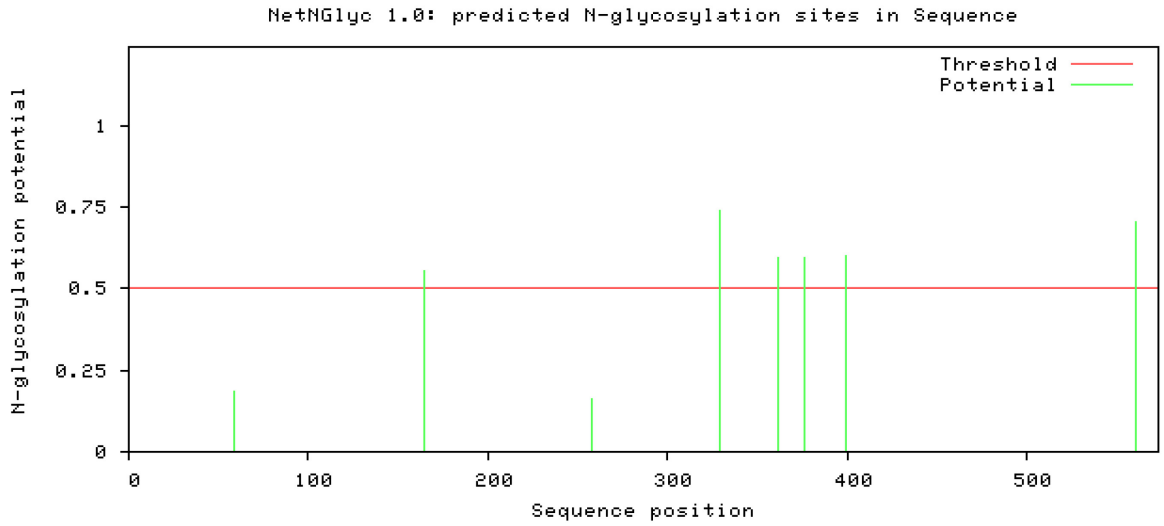
target:
DVVVTQTPLSLPVSFGDQASISCRSSQSLVNSYGS-----TYLSWYLHRPGQSPQLLIYG--------
ISNRFSGVPDRFSGSGSG--------TDFTLTIRTIKPEDLGMYYCLQGTHQP--------
WTFGGGTKLEIK---

template:
DVVMTQTPLSLPVSLGDQASISCRSSQSLVHSNGN-----TYLHWYLQKPGQSPKLLIYK--------
VSNRFSGVPDRFSGSGSG--------TDFTLKISRVEAEDLGVYFCSQSTHVP--------
WTFGGGTKLEIKRAD

2. *N*-glycosylation sites prediction

NetNGlyc predicts *N*-glycosylation sites in human proteins using artificial neural networks (ANN) that examine the sequence context of Asn-X-Ser/Thr sequons (Blom et al., 2004). *N*-glycosylation is known to occur on asparagines within the Asn-X-Ser/Thr sequon, where X is any amino acid except proline. Proline following an asparagine residue is known to preclude *N*-linked glycosylation in most cases by rendering the asparagine inaccessible (conformational constraints). NetNGlyc has been programmed to ignore this proline position. While the of Asn-X-Ser/Thr consensus sequence may be a requirement, it is not always sufficient for the asparagine to be glycosylated. Interestingly, there are a few known examples of *N*-glycosylation occurring within Asn-X-Cys, such as in the leukocyte surface protein (CD69).

NetNGlyc attempts to distinguish glycosylated sequons from non-glycosylated ones (Fig.63). In the sequence output, Asn-X-Ser/Thr sequons are highlighted in blue, and *N*-glycosylated asparagines are red. Any potential crossing of the default threshold of 0.5 represents a predicted glycosylation site, as long as it occurs within the required Asn-X-Ser/Thr sequon. The "potential" score is the averaged output of nine neural networks. For further information, the "jury agreement" column indicates how many of the nine networks support the prediction. When predicting *N*-glycosylation sites with high specificity, only "++" predictions should be considered as asparagines likely to be glycosylated. Taking into account the above-mentioned guidelines, the heavy chain of Mannitou mAb is predicted to be glycosylated at the 329 and 560 positions. No potential glycosylation sites have been found neither for the light chain (Fig.S10) nor for the heavy chain of Mannitou Fab (Fig.S11), although its sequence contains two Asn-X-Ser/Thr sequons.



EVKLLESGGGLVQPGGSLKLSCAASGFDFSTYWMSWVRQAPGKGLEWIGEINPDSSTINYTPSLKDKFIISRDNAKNTLY 80
LQMSKVRSEDSVLYYCVRPGTWGYFDYWGQGTTLTVSSASQSFPNVFPLVSCESPLSDKNLVAMGCLARDFLPSTISFTW 160
NYQNNTEVIQGIRTFPTLRTGGKYLATSQVLLSPKSILEGSDEYLVCKIHYGGKNRDLHVPIPAVAEMNPNVNVFVPPRD 240
GFSGPAPRKSKLICEATNFTPKPITVSWLKDGKLVESGFTTDPVTIENKGSTPQTYKVISTLTISEIDWLNLNVYTCRVD 320
HRGLTFLKNVSSTCAASPSTDILTFTIPPSFADIFLSKSANLTCLVSNLATYETLNISWASQSGEPLETKIKIMESHPNG 400
TFSAKGVASVCVEDWNNRKEFVCTVTHRDLPSPQKKFISKPNEVHKHPPAVYLLPPAREQLNLRESATVTCLVKGFSPAD 480
ISVQWLQRGQLLPQEKYVTSAPMPEPGAPGFYFTHSILTVTEEEWNSGETYTCVVGHEALPHLVTERTVDKSTGKPTLYN 560
VSLIMSDTGGTCY

SeqName	Position	Potential	Jury agreement	N-Glyc results
Mannitou mAb HC	59 NYTP	0.1852	(9/9)	---
Mannitou mAb HC	164 NNTE	0.5563	(6/9)	+
Mannitou mAb HC	258 NFTP	0.1622	(9/9)	---
Mannitou mAb HC	329 NVSS	0.7359	(9/9)	++
Mannitou mAb HC	361 NLTC	0.5932	(8/9)	+
Mannitou mAb HC	376 NISW	0.5957	(8/9)	+
Mannitou mAb HC	399 NGTF	0.5976	(7/9)	+
Mannitou mAb HC	560 NVSL	0.7030	(9/9)	++

Figure 63. Graph illustrating predicted *N*-glycosylation sites across the Mannitou mAb heavy chain. The X-axis represents protein length from *N*- to *C*-terminal. A potential glycosylation position (green vertical lines) crosses the threshold (red horizontal line at 0.5). Asn-X-Ser/Thr sequons are shown in blue. Asparagines predicted to be *N*-glycosylated are shown in red. Note that not all sequons are predicted to be glycosylated. ++ and + indicate predicted *N*-glycosylation sites, while --- stands for a negative site.

3. Single-particle reconstruction using Cryogenic Electron Microscopy

Both negative staining and cryo-electron microscopy are techniques used to provide structural information about macromolecules and their complexes by imaging non-crystalline specimens (single particles) absorbed on electron microscopy grids. More than 153 000 images of Mannitou mAb have been taken at random orientations to be later used to extract structural information (Fig.64). Having a high resolution of a 2D or 3D reconstruction is very difficult, as it depends on many factors, i.a. sample of high purity, having a homogeneous conformation, and forming stable complexes. All images need to be aligned with one another. However, every individual electron micrograph is very noisy and with unknown orientation, so an improvement of the signal-to-noise ratio is performed by averaging a large number of individual projections (Boekema et al. 2009). Due to IgM high flexibility, 3D reconstruction seems impossible. An attempt at 2D reconstruction has been made. The final 2D classification has been performed using 84 176 particles (Fig.65).

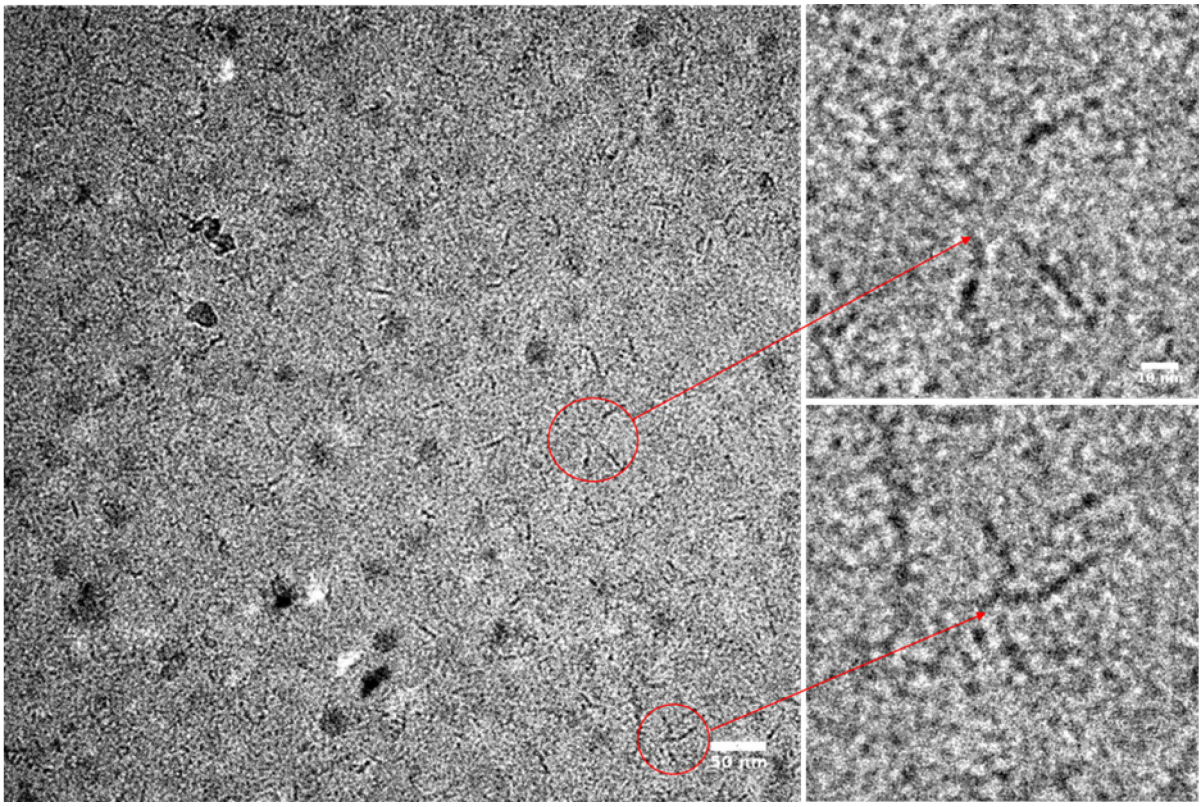


Figure 64. Images taken at 50 000 x magnification using JEM-2200FS/CR transmission electron microscope (JEOL) with an ULTRASCAN 4000 SP (4008 × 4008 pixels) cooled slow-scan CCD camera (GATAN) show that Mannitou IgM appears in a very flexible pentameric form.

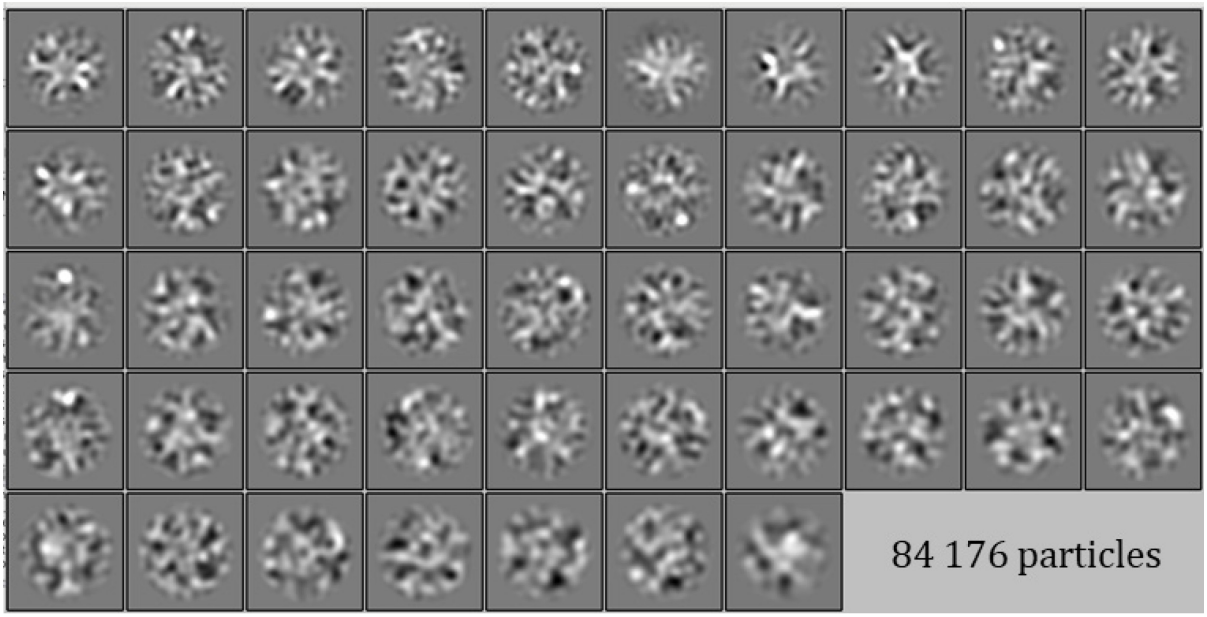


Figure 65. Reference free 2D classifications made with 84 176 particles.

4. Crystallisation of Mannitou Fab and Mannitou Fab – paucimannose *N*-glycans complexes

The crystallisation of a protein requires the creation of a solution that is supersaturated with the macromolecule but exhibits conditions that do not significantly perturb its natural state by reaching the highest possible concentration without causing aggregation or precipitation. The supersaturation is achieved by adding mild precipitating agents, such as neutral salts or polymers, and by manipulating various parameters, including temperature, ionic strength and pH (McPherson et al. 2013). Introducing the sample to a precipitating agent promotes the nucleation of protein crystals in solution, which can result in large three-dimensional crystals growing from the solution. To produce crystals suitable for crystallographic studies, the macromolecule (e.g. antibody, antibody-antigen complex) must be purified to homogeneity, or as close as possible to homogeneity. The homogeneity of the preparation is a key factor in achieving crystals that diffract to high resolution. Among the most widely used techniques to obtain crystals are vapour diffusion, dialysis, batch and liquid-liquid diffusion (Dessau et al. 2011). The crystals provide the X-ray diffraction patterns that in turn serve as the raw data which allow the direct visualisation of the macromolecules or their complexes that the crystals are composed of.

Concentration is a significant crystallisation variable, as too concentrated samples can result in an amorphous precipitate, while too diluted samples yield clear drops. The PCT shows that A2 reagent provides suitable crystallisation conditions as a light granular precipitate is observed (Fig.66). The rest of the drops remain clear after even after 24h. Based on the PCT results, different crystallisation conditions, including $AmSO_4$, Index, Midas, Morpheus, and PACT have been employed to crystallise Mannitou Fab. Moreover, Mannitou Fab – $Man_3GlcNAc_2$ and Mannitou Fab – $Man_3GlcNAc_2(\alpha 1\text{-}6)Fuc$ complexes have been set up for crystallisation using Morpheus, PACT (Fig.67), JCSG, MemGold, and Midas (Fig.68) conditions. Despite several crystallisation and co-crystallisation attempts for over a year, no diffracting crystals have been obtained.

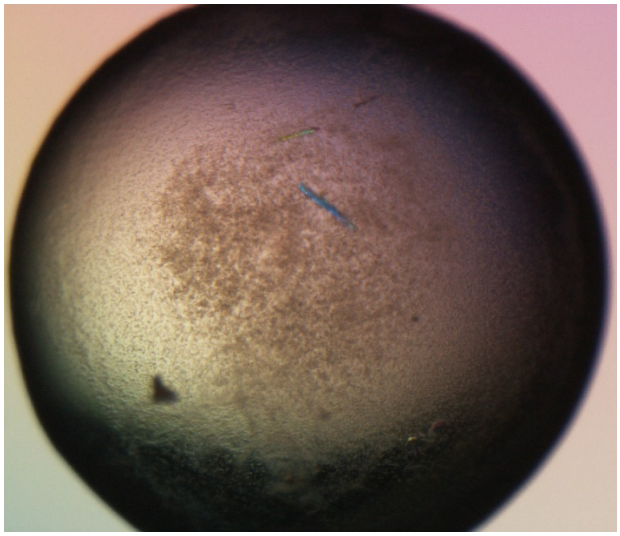


Figure 66. Light granular precipitation observed for A2 reagent during the PCT.

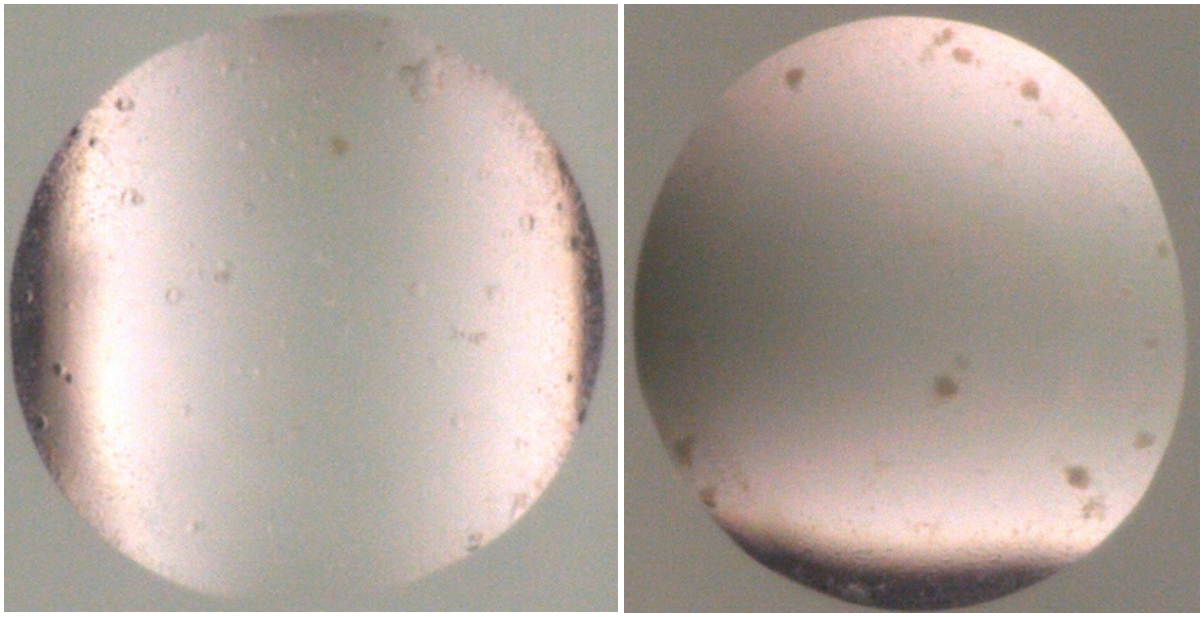


Figure 67. Amorphous precipitation observed under Morpheus (left) and PACT (right) conditions. Mannitou Fab (30 µl) has been used in 130.9 µM concentration and the glycan (1 µl) in 11.6mM concentration.

Midas

MemGold

JCSG

Figure 68. Putative crystal structures of Mannitou Fab – $Man_3GlcNAc_2$ (Midas, MemGold) and Mannitou Fab – $Man_3GlcNAc_2(\alpha 1\text{-}6)Fuc$ (JCSG) complexes. Mannitou Fab has been used in 115.9 µM concentration.

DISCUSSION

A general problem in protein crystallisation is that the crystallisation conditions, which include a combination of a right pH, ionic strength, temperature, protein concentration, the presence of various salts, ligands or additives, the type of precipitant and the actual crystallisation method to use (hanging drop, sitting drop, dialysis, etc.), are practically impossible to predict in advance.

The flexibility of the hinge region of an intact immunoglobulin makes crystallisation very difficult, especially when dealing with a pentameric structure like Mannitou mAb. Structural studies are therefore generally carried out with antigen-binding fragments, not only because they lack the hinge region but also because they are much smaller and do not aggregate during purification or concentration like full-length mAbs. However, crystallising Fabs with bound ligands is made difficult not only by the low affinities of anti-carbohydrate antibodies but also by the packing of Fab molecules in crystals where the constant region of one Fab often lies across the binding site of a neighbouring one, blocking the access of ligands. This mode of crystal contact is unfortunately too common in case of Fab crystallisation outcomes (Davies et al. 1990; Haji-Ghassemi et al., 2015).

The majority of published crystal structures of naturally occurring glycoproteins are derived from proteins that prior to crystallisation either had their glycosylation sites mutated, or had their glycans partially or completely degraded. The practical reason for such a drastic approach is that glycans are often heterogeneous and have a high range of motion hampering the acquisition of an ordered crystal. Even if crystallisation proves possible, the glycans are typically disordered within the resulting image. In instances where glycans are left intact, the glycoproteins are often expressed in heterologous cells, resulting in non-species-specific glycosylation. In turn, when glycosylation sites are mutated or glycosylation is modified, there is a significant possibility that the folded form defined by crystallography may not be the native state. Solving this technical problem remains a major challenge for the future but techniques such as cryo-electron microscopy often help (Varki 2017).

Whereas the structure and assembly of IgG is well-established, for IgM we largely lacked detailed structural information until very recently. Previous studies employing transmission electron microscopy (TEM) have revealed that soluble IgM is an individual pentameric particle of variable shapes, ranging from round or elliptical to oblong. Fab units as well as the core of IgM are flexible and therefore contribute to the multiple conformations of IgM (Akhouri et al., 2016). The different conformations that IgM can assume owing to its flexibility are likely one of the main reason why crystallisation trails of the IgM pentamer have not been successful to date. Given that soluble IgM molecules are dynamic and their conformations vary substantially, we have attempted to use cryo-electron microscopy to obtain 2D and 3D structures of the whole murine Mannitou IgM. Unfortunately, to this day all our efforts have not proven effective. However, based on the negative-stain images, it can be concluded that IgMs are indeed rarely planar star-shaped structures with a central circular region from which a number of arms project out radially. Torsional stress due to the randomly oriented arms carrying Fab units could lead to distorted IgM cores, which results in different conformations, including extended, turtle-shaped, bell-shaped (Akhouri et al., 2016), and mushroom-shaped (Czajkowsky et al., 2009), with the central portion formed by the C-terminal domains protruding out of the plane formed by the Fab domains. Very recently, in 2018, using single-particle negative-stain EM and reference-free 2D class average analysis, Hiramoto et al. have presented IgM as an asymmetric pentagon with a ~ 50° gap (Fig.69). These findings have been further investigated by a cryo-EM tomography study from Sharp et al. in 2019, showing that the IgM pentamer is structurally equivalent to the hexamer with its five arms separated by ~60° and one replaced by a J chain, forming a C6-symmetric dome-shaped structure. These results have been fully confirmed a year later by Li et al., who presented the IgM-Fc pentamer as an asymmetric formation resembling a hexagon with a missing triangle.

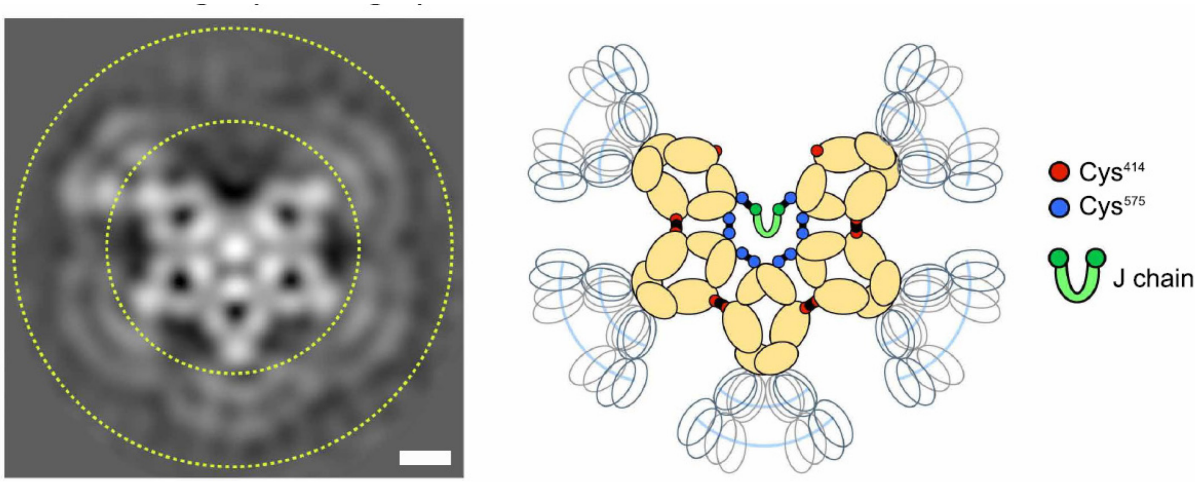


Figure 69. Negative-stain EM image of monoclonal mouse IgM pentamer (left) and a schematic representation of the image. Scale bar: 5 nm. The peripheral region corresponds to the Fab structures (yellow dotted lines), which appears to be highly flexible and thus cannot be observed clearly. Adapted from: Hiramoto et al., 2018.

Since many anti-carbohydrate antibodies cannot be crystallised with (or without) their ligands, homology modelling studies are often necessary, especially when all other alternatives, such as the above-mentioned Cryo-EM, fail. In a series of landmark papers, Chothia and co-workers first demonstrated that the conformation of individual complementarity-determining loops can be grouped into just a few "canonical forms" based on the loop length and position/identity of key residues (Chothia et al., 1987; Chothia et al. 1989, Chothia et al., 1992), and this work has since been expanded (Al-Lazikani et al. 1997; Morea et al. 1997, North et al. 2011). Structural prediction based on this information is limited by the difficulty in modelling the relative orientation of each CDR or even of the heavy and light chain domains, and the highly variable nature of the CDR H3 (Kuroda et al. 2008; Reczko et al. 1995). Further, the contribution of solvent is generally ignored in such studies. The available on-line prediction servers compare the query sequence to antibodies of known structure and select the closest homologs for each CDR.

CDR H3 is the most important of the CDRs, reflecting its unique genetic origin from the D gene segment. Its length and composition help to control the overall shape of the site and contribute disproportionately to the binding of the antigen. There is a considerable spread in lengths of the CDR H3, from 2 to 13 residues. There is no apparent correlation between the size of the carbohydrate epitope and the length of the H3, and while the majority of antibodies show cleft-shaped binding sites, the H3 length and sequence cause wide variations in this feature (Morea et al., 1998; Xu et al., 2000).

Chapter 4.

Neisseria meningitides serogroup X (MenX)

INTRODUCTION

Neisseria meningitidis is one of the main pathogens responsible for bacterial meningitis in infants and adolescent having a significant mortality rate (Harrison et al., 2009). Meningococcus is an aerobic, Gram-negative diplococcus present only in human hosts and is not known to colonise any other animal or environmental niche (Feavers 2001). The bacteria isolated from the bloodstream or the spinal fluid are nearly always coated with a polysaccharide capsule (PS) consisting of repeating sugar units (Yang et al., 2001). Among the thirteen distinct meningococcal serogroups classified based on the structure of their polysaccharide capsule, six of them (A, B, C, W135, Y, and X) are the most common cause of outbreaks of invasive meningococcal disease, including meningitis and septicaemia (Stephens 2009). The highest incidence of meningococcal meningitis occurs in the meningitis belt of sub-Saharan Africa, extending from Senegal to Ethiopia. Since records began, *Neisseria meningitidis* serogroup A (MenA) has been the dominant cause of epidemics of meningococcal meningitis in that region (LaForce et al, 2009), but MenW (Decosas et al., 2002) and MenX (Boisier et al., 2007; Delrieu et al., 2011; Gagneux et al., 2002; Mutonga et al., 2009) have also contributed significantly. Higher case-fatality rates have been reported for meningitis caused by MenX compared with MenA (Boisier et al., 2007; Mutonga et al., 2009), and children between 1–9 years constitute the most affected age group (Boisier et al., 2007; Djibo et al., 2003).

The MenX PS was first characterised and defined as a distinct serogroup in the 1960s (Bories et al., 1966; Evan et al., 1968). It was later revealed the structure of MenX PS to be a homopolymer of 1→4-linked 2-acetamido-2-deoxy-α-D-glucopyranosyl phosphate residues (Bundle et al., 1973; Bundle et al., 1974) (Fig.70). No evidence for *O*-acetylation of the polysaccharide has been found. In turn, it has been previously documented that *O*-acetyl epitope is essential for inducing protective immunity directed against the meningococcal polysaccharide of serogroup A (Xie et al., 2012).

The capsular polysaccharides of *Neisseria meningitidis* are attractive vaccine candidates since they constitute the most highly conserved and most exposed bacterial-surface antigens (Jennings 1990). Currently available meningococcal vaccines have been developed using precisely the capsular polysaccharides of various strains of *Neisseria meningitides*. A monovalent polysaccharide–tetanus toxoid conjugate vaccine against

serogroup A meningococcal disease has recently been licensed under the name of MenAfriVac™ (Frasch et al., 2012; Trotter et al., 2017). Three quadrivalent meningococcal A, C, W135, and Y conjugate vaccines are currently on the market, the first with diphtheria toxoid (Menactra™) (2005) as the carrier protein, the second with CRM_{197}, a non-toxic variant of *Corynebacterium diphtheriae* toxin (Menveo™) (2010), and the third with tetanus toxoid (Nimenrix™) (Miller et al., 2011). Although polysaccharide-protein conjugate vaccines against meningococcal serogroups A, C, W135, and Y are already on the market, a vaccine that could protect against MenX has been developed only at the preclinical level.

Initially, to eradicate infections caused by encapsulated bacteria polysaccharide vaccines were introduced. Although these vaccines were effective in adults and older children, only due to the T-cell independent immune response, they were ineffective in younger children, below two years of age (Mond et al., 1995). To overcome this issue, glycoconjugate vaccines which are highly effective in both adults and children were developed. Glycoconjugates are in general immunogenic from early infancy, induce a longer-lasting immune response, and ensure immunological memory (Pace 2012). Conjugation to a carrier protein provides enhanced immunogenicity by converting the PS into a T-cell–dependent antigen (Costantino et al., 2011). The mechanism of action of glycoconjugate vaccines is based on their uptake, internalisation, and processing by antigen-presenting cells (APC). Subsequent presentation of peptides or glycopeptides on APC surface in association with class II major histocompatibility complex (MHCII) allows the interaction with CD4+ T cells and signalling to B cells for their proliferation and differentiation into antibody-producing and memory cells (Alberts et al., 2002). Additionally, as documented for serogroups A and C, meningococcal conjugate vaccines can reduce the nasopharyngeal carriage of *Neisseria meningitides* which leads to a decrease in transmission (Kristiansen et al., 2013; Ramsay et al., 2003), whereas PS vaccines have not been reported to provide substantial herd immunity (Plotkin 2014).

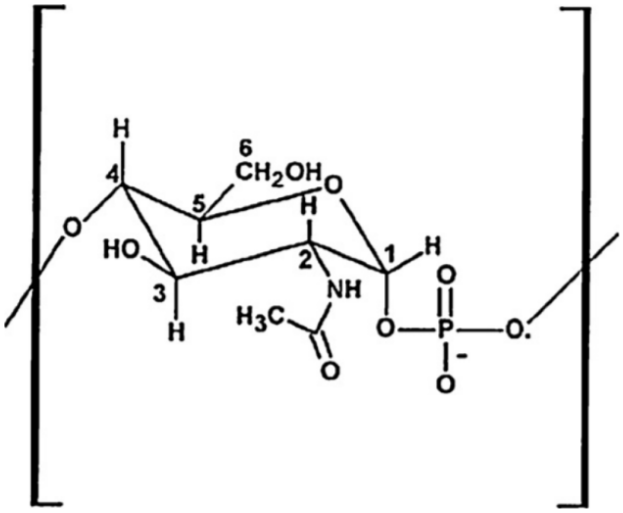


Figure 70. Chair conformation of the serogroup X polysaccharide monomer. From: Xie et al., 2012

MATERIALS & METHODS

1. Equipment

The SPR studies between MenX mAb/Fab and MenX DP15 (degree of polymerisation) have been conducted using the Biacore X100 instrument (GE Healthcare Life Sciences, UK). The system allows getting a reliable real-time insight into biological processes in small-scale interaction analysis, providing information on binding specificity, kinetics, affinity, and concentration analysis. It contains all key functionalities needed to define the structure-function relationships and understand the dynamics of molecular pathways. MenX DP15 has been immobilised by amine coupling on a Biacore Sensor Chip CM5 (GE Healthcare Life Sciences, UK). The CM5 chip carries a matrix of carboxymethylated dextran covalently attached to a gold surface, required for the generation of the surface plasmon resonance response. The flexible dextran matrix extends ~100 nm from the gold surface, allowing relatively free movement of the attached ligand molecules. Its high immobilisation capacity permits a broad range of capture densities and the surface stability provides accuracy and precision, allowing repeated analysis in the same surface conditions.

2. Buffer pH scouting and ligand pre-concentration analysis

2.1. pH scouting using 20 mM NaAc buffer pH=4

HBS-EP pH=7.2 is used as running buffer. Two conjugates, MenX(DP15)–ADH–CRM_{197} (Lot. EB 11Feb13-X) sacc. 108.5 µg/mL, protein 575.3 µg/mL, sacc/prot 0.2 w/w, and $MenX(PS)_{OX}$–CRM_{197} (Lot. EB 19Apr12-PS) sacc. 163.1 µg/mL, protein 507.6 µg/mL, sacc/prot 0.32 w/w, have been prepared in dilutions of 5, 10, 15 µg/mL in 20 mM NaAc buffer pH=4.

2.2. pH scouting using 10 mM NaAc buffer pH=5

HBS-EP pH=7.2 is used as running buffer. Two conjugates, MenX(DP15)–ADH–CRM_{197} (Lot. EB 11Feb13-X) sacc. 108.5 µg/mL, protein 575.3 µg/mL, sacc/prot 0.2 w/w, and $MenX(PS)_{0X}$–CRM_{197} (Lot. EB 19Apr12-PS) sacc. 163.1 µg/mL, protein 507.6 µg/mL, sacc/prot 0.32 w/w, have been prepared in dilutions of 5, 10, 15 µg/mL in 10 mM NaAc buffer pH=5, and subsequently in dilution of 30 µg/mL in 10 mM NaAc buffer pH=5.

3. Ligand immobilisation

The CM5 chip surface is activated with a mixture of 0.4M 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and 0.1M *N*-hydroxysuccinimide (NHS). The reagent solutions have been equilibrated at room temperature and mixed shortly before use. The 1st flow cell (Fc=1) serves as a reference surface, so it is blocked immediately after activation (blank immobilisation). In the 2nd flow cell (Fc=2) MenX(DP15)–ADH–CRM_{197} (30 µg/mL) is covalently immobilised using 10 mM NaAc pH=5 as immobilisation buffer. The target level is set at 500 RU. 1M ethanolamine-HCL pH=8.5 is used to inactivate excess reactive groups. 50 mM NaOH is applied as a wash solution to remove any remaining ligand molecules from the sensor chip surface.

4. Binding analysis

PBS Tween 0.005% pH=7.2 is used as running buffer for the binding analysis at 25°C. The experiment starts with three start-up cycles to allow surface stabilisation. Each sample injection (120 s contact time, 300 s dissociation time) is followed by regeneration with 3.5M $MgCl_2$ (120 s contact time) to remove the bound analyte from the ligand immobilised on the chip surface. The MenX mAb (01 lot 005) has been used in three concentrations, 0.65 µM, 1.3 µM, and 3.3 µM. The MenX Fab (01 lot 001, 01 lot 005) has been used in 10 µM concentration.

5. Kinetic analysis

PBS Tween 0.005% pH=7.2 is used as the running buffer for the kinetic analysis, at a 45 µl/min flow rate at 25°C. The experiment starts with three start-up cycles to allow surface stabilisation. Each sample injection (120 s contact time, 600 s dissociation time) is followed by regeneration with 3.5M $MgCl_2$ (120 s contact time) at a 10 µl/min flow rate to remove the bound analyte from the ligand immobilised on the chip surface. The MenX mAb (01 lot 005) has been used in 0 µM (no analyte), 0.3125 µM, 0.625 µM, 1.25 µM, 2.5 µM, and 5 µM concentrations. The MenX Fab (01 lot 001) has been used in 0 µM (no analyte), 0.625 µM, 1.25 µM, 2.5 µM, 5 µM, and 10 µM concentrations.

6. Inhibition assays

PBS Tween 0.005% pH=7.2 is used as the running buffer for the inhibition assays, at a 45 µl/min flow rate at 25°C. The experiment starts with three start-up cycles to allow surface stabilisation. Each sample injection (120 s contact time, 300 s dissociation time) is followed by regeneration with 3.5M $MgCl_2$ (120 s contact time) to remove the bound analyte from the ligand immobilised on the chip surface. 10 µM MenX mAb (01 lot 005) has been used together with descending concentrations of DP5: 1000 µg/mL, 500 µg/mL, 250 µg/mL, 125 µg/ml, 62.5 µg/mL, 31.3 µg/mL, 15.6 µg/mL, 7.8 µg/mL, 3.9 µg/mL, 2 µg/mL, 1 µg/mL, 0.5 µg/mL, 0 µg/mL (no analyte in the mAb solution). 10 µM MenX mAb (01 lot 005) has been used together with descending concentrations of DP10: 1000 µg/mL, 500 µg/mL, 250 µg/mL, 125 µg/mL, 62.5 µg/mL, 31.3 µg/mL, 15.6 µg/mL, 7.8 µg/mL, 3.9 µg/mL, 2 µg/mL, 1 µg/ml, 0.5 µg/mL, 0 µg/mL (no analyte in the mAb solution). 10 µM MenX mAb (01 lot 005) has been used together with descending concentrations of PS: 1000 µg/mL, 500 µg/mL, 250 µg/mL, 125 µg/mL, 62.5 µg/mL, 31.3 µg/mL, 15.6 µg/mL, 7.8 µg/mL, 3.9 µg/mL, 2 µg/mL, 1 µg/mL, 0.5 µg/mL, 0 µg/mL (no analyte in the mAb solution).

RESULTS

1. Buffer pH scouting and pre-ligand concentration analysis

The procedure has been performed to determine the optimal buffer pH and ligand concentration for amine coupling. The approach involves testing several buffer pH and ligand concentrations prior to direct immobilisation. While the general recommendation to immobilise proteins at pH=5 works adequately in many cases, the coupling conditions frequently need to be optimised for best results. In order to determine the suitable amine coupling conditions without permanently modifying the sensor chip surface, the ligands are injected over a surface that has not been activated with EDC/NHS. Electrostatic pre-concentration of the ligand on the surface is seen as an increase in response and gives an indication of whether the conditions are suitable. Pre-concentration is generally more efficient when the pH of the ligand solution is below the isoelectric point of the ligand so that the ligand carries a net positive charge, but above the pKa (3.5) of carboxyl groups on the surface, so that the surface is negatively charged (Jönsson et al., 1991). However, since amine coupling chemistry requires uncharged amine groups, the immobilisation is more efficient at higher pH. The highest pH conditions that give an adequate response should be therefore applied. In this study, two buffers have been tested: 20 mM NaAc pH=4 (Fig.71) and 10 mM NaAc pH=5 (Fig.72), as low ionic strength favours the electrostatic attraction, and buffers with 10-20 mM total cation concentration are generally recommended. Although 10 mM NaAc pH=5 provides a slightly better binding pattern, when a satisfactory electrostatic pre-concentration cannot be obtained at any pH, the ligand concentration should be increased. Two different ligands, MenX(DP15)–ADH–CRM_{197} and $MenX(PS)_{OX}$–CRM_{197}, have been examined in various concentrations: 5 µg/mL, 10 µg/mL, 15 µg/mL, and 30 µg/mL. Best binding is observed at 30 µg/mL concentration of MenX(DP15)–ADH–CRM_{197} (Fig.72 bottom). No binding is detected for the $MenX(PS)_{OX}$–CRM_{197} ligand, regardless of its concentration or immobilisation buffer pH.

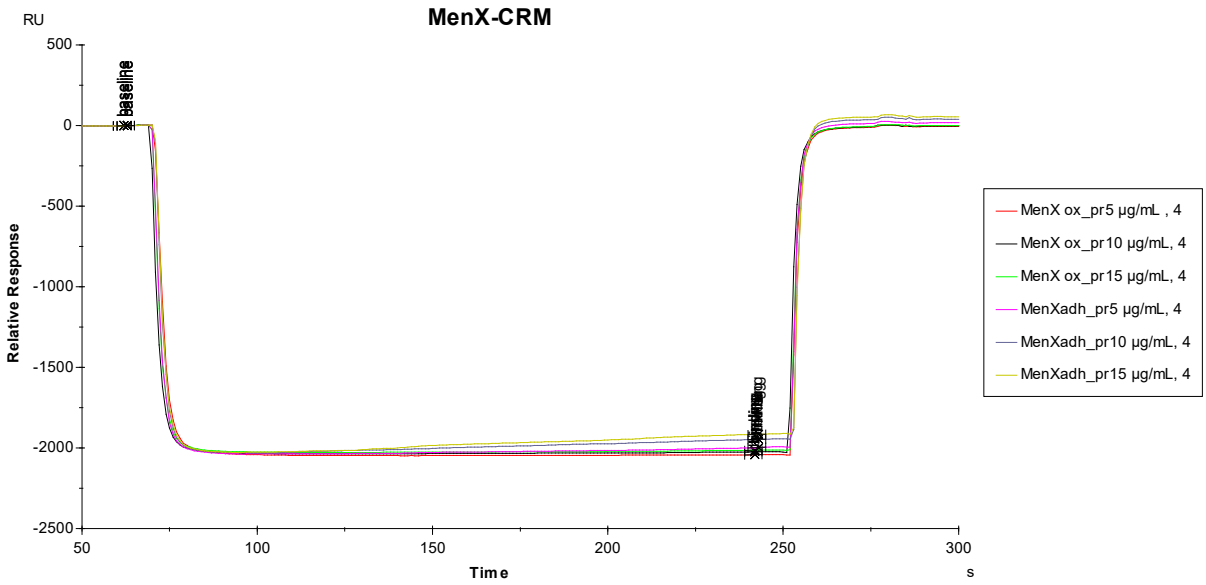


Figure 71. pH scouting results using 20 mM NaAc buffer pH=4.

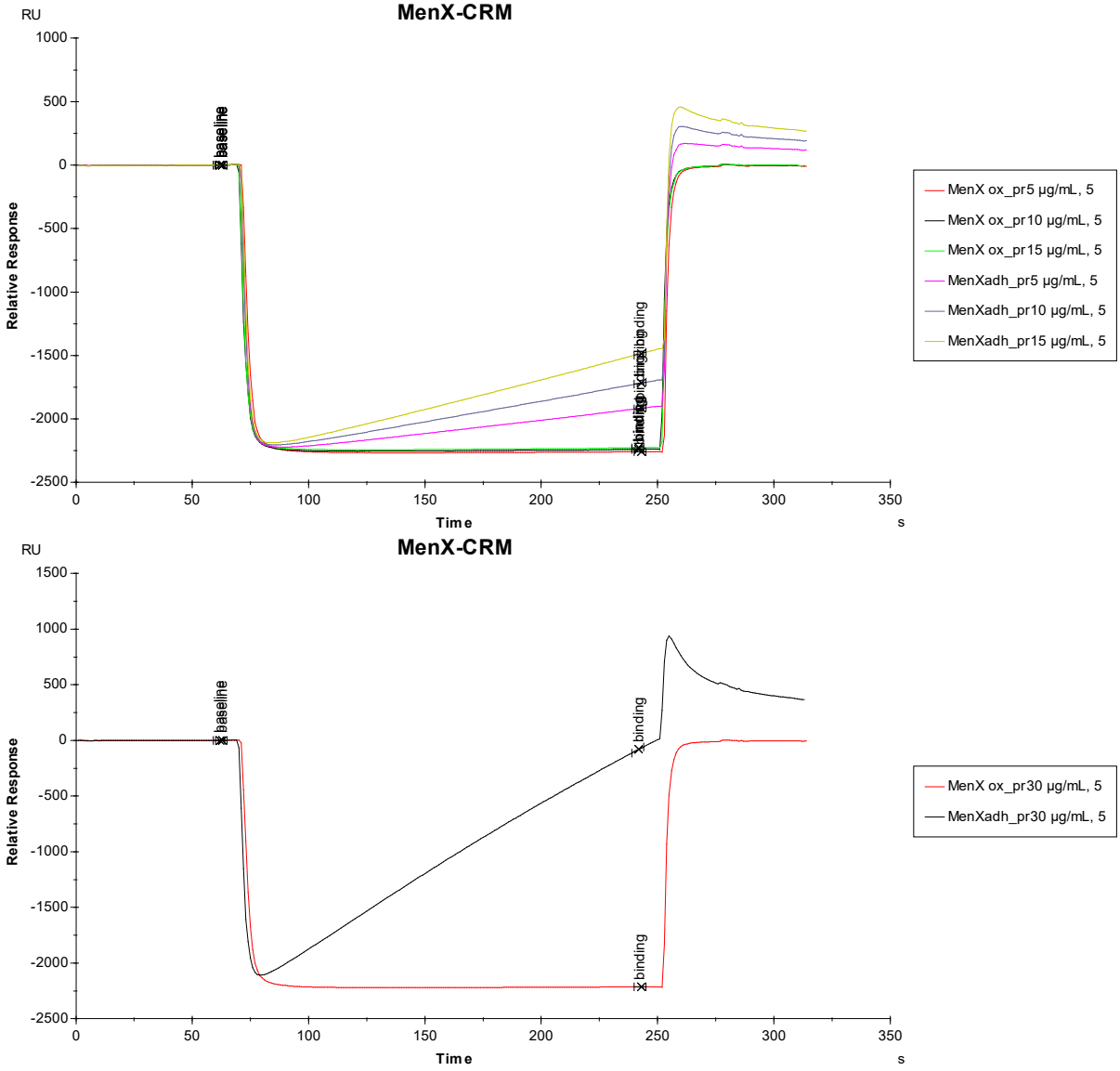
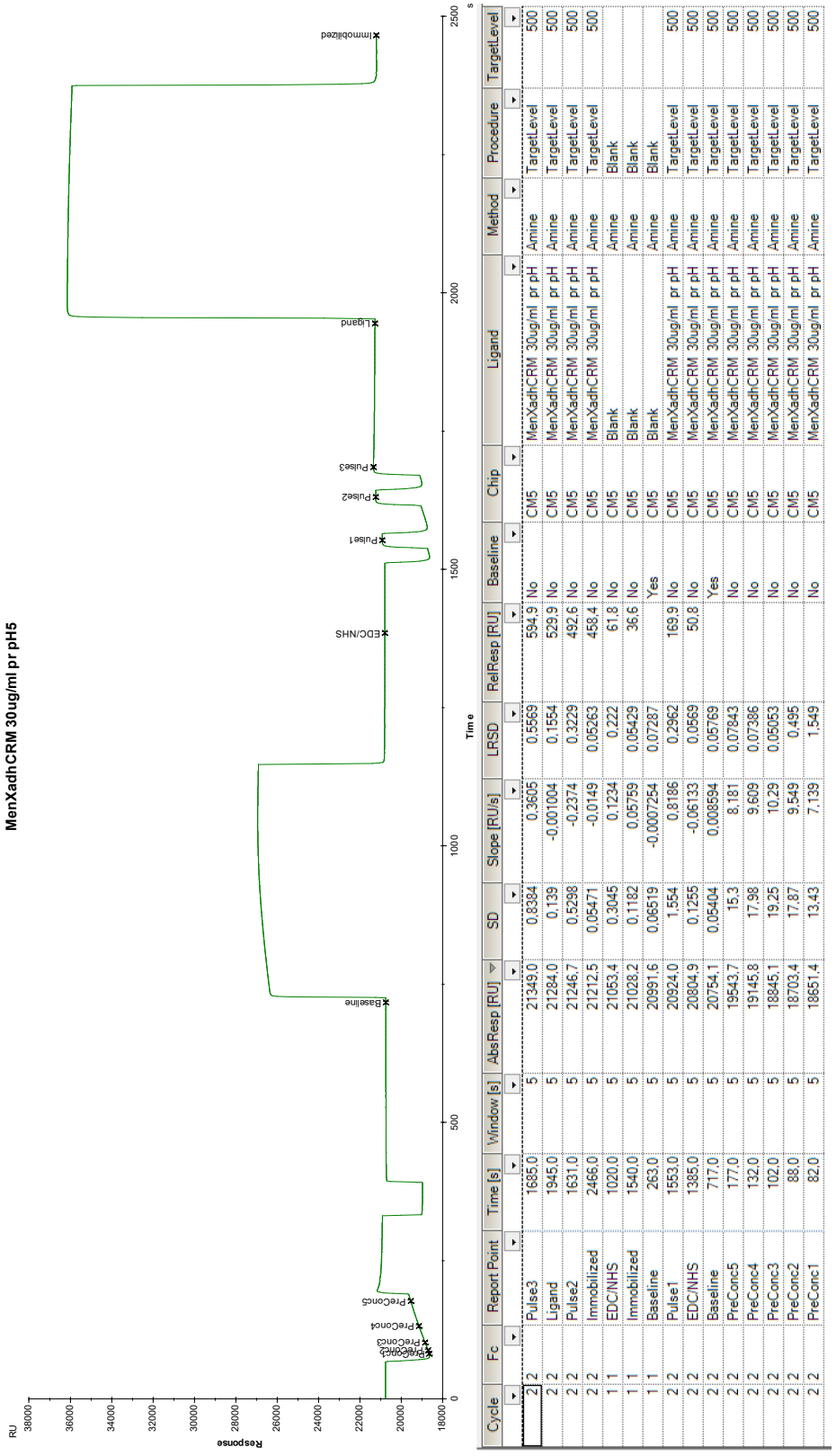


Figure 72. pH scouting results using 10 mM NaAc buffer pH=5.

2. Ligand immobilisation

Amine coupling chemistry is one of the most common techniques for covalent immobilisation of proteins and involves an irreversible chemical attachment of the ligand to the sensor surface. Following this method, carboxyl groups of the dextran matrix are first activated with a mixture of EDC and NHS to create highly reactive *N*-hydroxysuccinimide esters. The ligand is then passed over the surface and the esters react spontaneously with the amine and other nucleophilic groups on the protein to covalently link the ligand to the dextran matrix. This approach results in the formation of a stable bond between the ligand and the sensor surface (Johnsson et al., 1991). MenX(DP15)–ADH–CRM_{197} has been directly immobilised on a CM5 chip at 30 µg/mL concentration in a 10mM NaAc buffer pH=5. As many proteins show limited stability in low ionic strength solutions and at low pH, the ligand solution has been, therefore, prepared directly before use. The target immobilisation level has been set at 500 RU and the final immobilisation achieved is 458 RU (Fig.73). The amount of immobilised ligand only gives an indication of the potential binding capacity. Therefore, a newly prepared surface should always be tested by injection of the analyte to evaluate the functional integrity of the immobilised ligand.



Cycle	Fc	Report Point	Time [s]	Window [s]	AbsResp [RU]	SD	Slope [RU/s]	LRSD	RelResp [RU]	Baseline	Chip	Ligand	Method	Procedure	TargetLevel
2	2	Pulse3	1685,0	5	21349,0	0,8384	0,3605	0,5569	594,9	No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
2	2	Ligand	1945,0	5	21284,0	0,139	-0,001004	0,1554	529,9	No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
2	2	Pulse2	1631,0	5	21246,7	0,5298	-0,2374	0,3223	492,6	No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
2	2	Immobilized	2466,0	5	21212,5	0,05471	-0,0149	0,05263	458,4	No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
1	1	EDC/NHS	1020,0	5	21053,4	0,3045	0,1234	0,222	61,8	No	CM5	Blank	Amine	Blank	
1	1	Immobilized	1540,0	5	21028,2	0,1182	0,05759	0,05423	36,6	No	CM5	Blank	Amine	Blank	
1	1	Baseline	263,0	5	20991,6	0,06519	-0,0007254	0,07287		Yes	CM5	Blank	Amine	Blank	
2	2	Pulse1	1553,0	5	20924,0	1,554	0,8186	0,2962	169,9	No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
2	2	EDC/NHS	1385,0	5	20804,9	0,1255	-0,06133	0,0569	50,8	No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
2	2	Baseline	717,0	5	20754,1	0,05404	0,008594	0,05769		Yes	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
2	2	PreConc5	177,0	5	19543,7	15,3	8,181	0,07843		No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
2	2	PreConc4	132,0	5	19145,8	17,98	9,609	0,07386		No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
2	2	PreConc3	102,0	5	18845,1	19,25	10,29	0,05053		No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
2	2	PreConc2	88,0	5	18703,4	17,87	9,549	0,495		No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500
2	2	PreConc1	82,0	5	18651,4	13,43	7,139	1,549		No	CM5	MenXadhCRM 30ug/ml pr pH	Amine	TargetLevel	500

Flow cell	Procedure	Method	Ligand	Response Bound (RU)	Response Final (RU)	Target reached
1	Blank	Amine		N/A	36,6	N/A
2	Target level	Amine	MenX(DP15)-ADH-CRM$_{197}$ 30 µg/mL	479,1	458,4	Yes

Figure 73. Sensorgram and data showing the immobilisation using 30 µg/mL of MenX(DP15)–ADH–CRM$_{197}$ in 10 mM NaAc buffer pH=5.

3. Binding analysis

When studying biological problems at the molecular or cellular level using recombinant antibodies, like the Fab directed against MenX, it is important to be able to show that the recombinant protein has the same structure and activity as its native counterpart. This is most easily done by confirming that the recombinant antibody binds its natural ligands. Because such interactions involve multiple residues, which are usually far apart in the primary amino acid sequence, they require a correctly folded antibody. The binding analysis has been performed to determine which lots of MenX mAb and MenX Fab are active and can be used for further kinetic analysis. Direct measurement is suitable for macromolecular analytes, with molecular weight >5000 Da, which gives an easily measured response even at low molar concentrations. The duration of the sample injection should ideally be long enough to allow the sensorgrams to flatten out to some extent, although it is not necessary for the binding to reach a steady state. If the injections are too short, the sensitivity of the assay may be impaired. The 120 s of contact time in the conducted experiments have been sufficient to reach the required plateau. Although the reference-subtracted sensorgrams (Fc=2-1) show the actual binding response, it is important to inspect the sensorgrams from both the reference (Fc=1) and active (Fc=2) surfaces, as the sensorgrams from the reference surface may, for example, reveal non-specific binding. MenX mAb (Men X 01 lot 005) shows binding to the surface-attached ligand, even at the lowest concentration of 0.65 µM (Fig.74). In turn, two different lots of MenX Fab (01 lot 001 and 01 lot 005), both applied in concentrations of 10 µM, have demonstrated different interaction patterns (Fig.75). Binding to the immobilised ligand occurs when Men X Fab 01 lot 001 is used as the analyte. No binding interaction is observed for MenX Fab 01 lot 005. For this reason, Men X mAb 01 lot 005 and Men X Fab 01 lot 001 have been employed for further experiments.

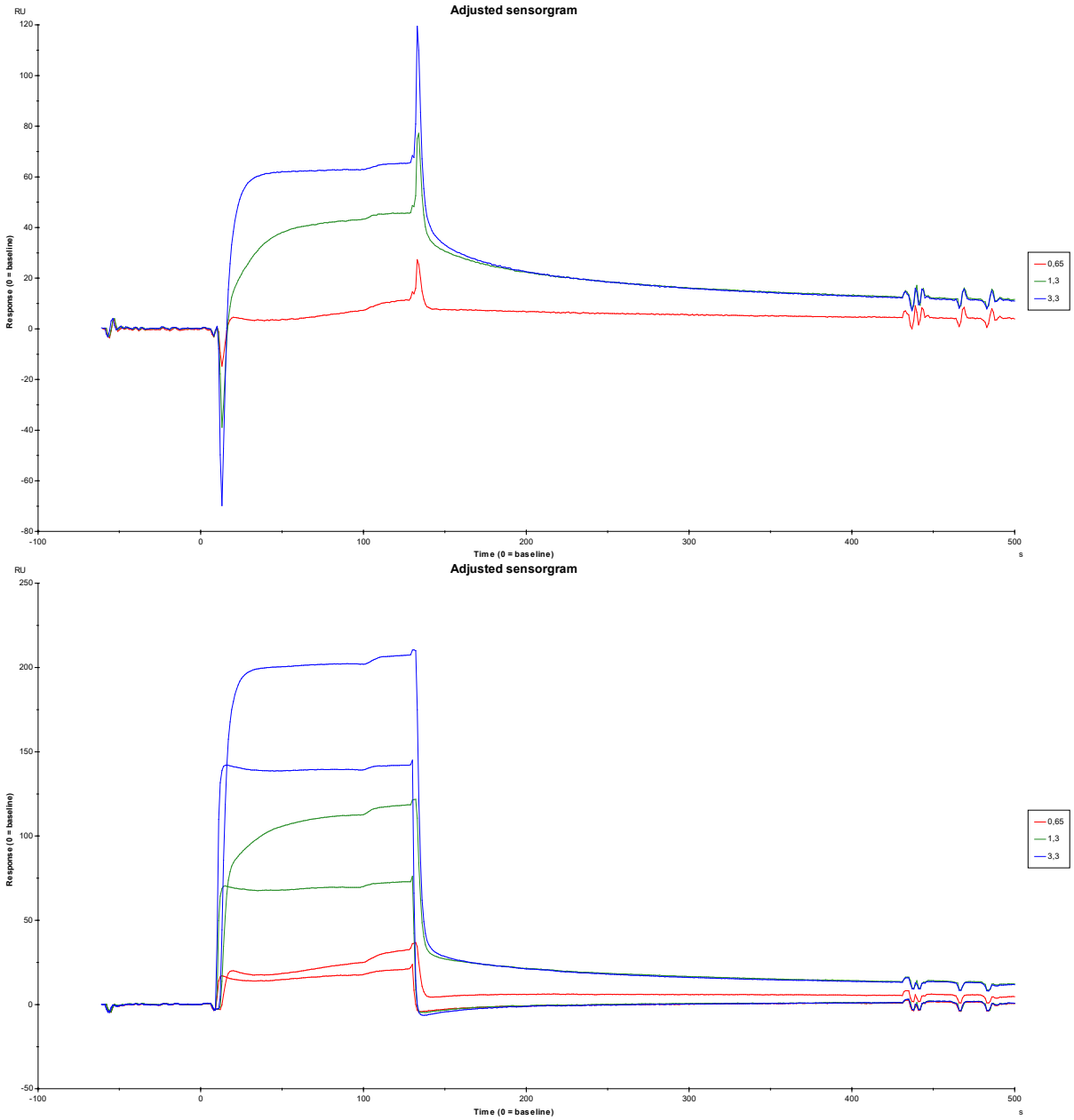


Figure 74. Sensorgrams representing the binding experiments using concentration series (0.65 µM, 1.3 µM, 3.3 µM) of MenX mAb (MenX 01 lot 005). Top: Fc=2-1 Bottom: Fc=1, Fc=2.

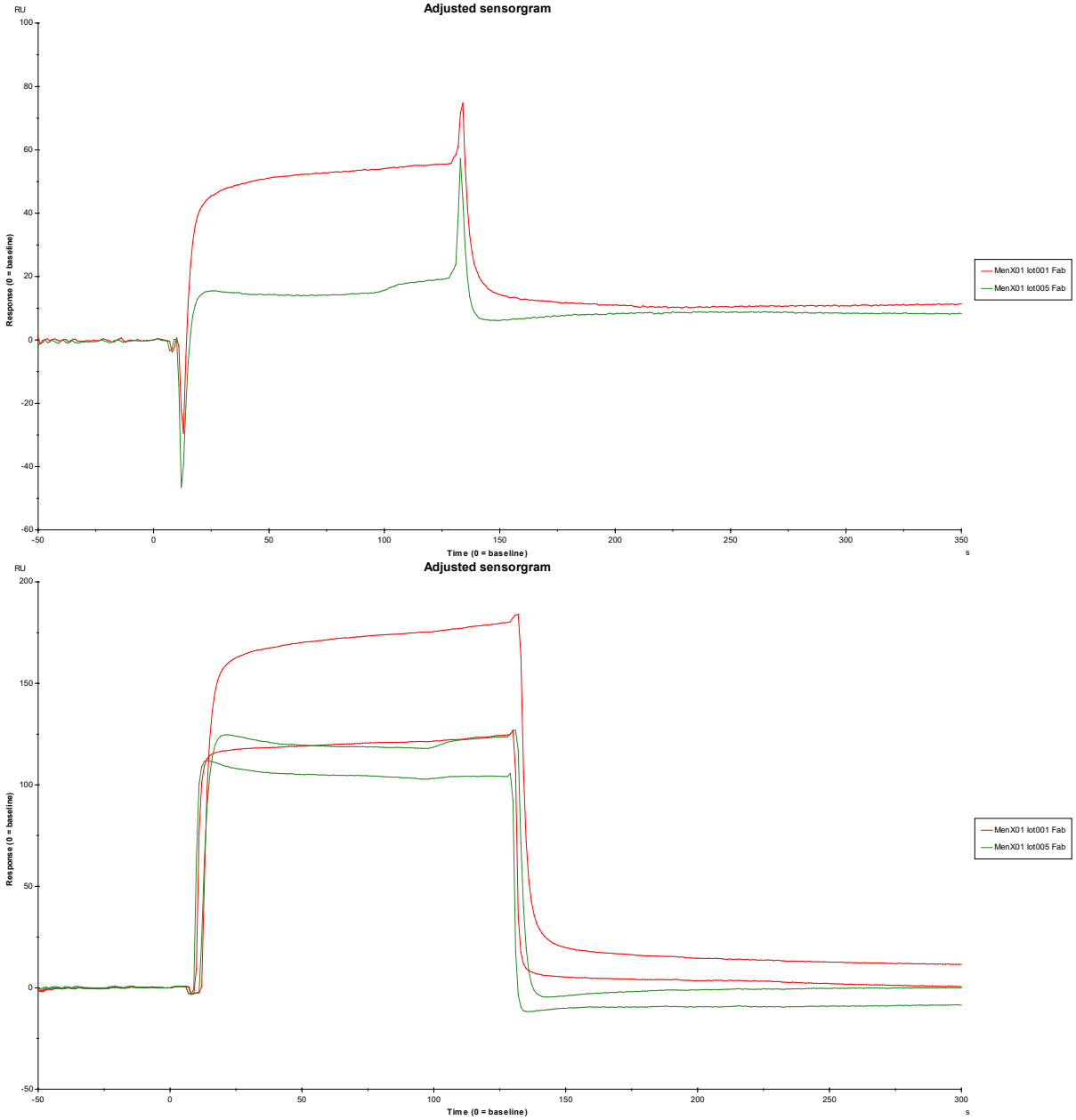


Figure 75. Sensorgrams representing the binding experiments using two different lots of MenX Fab (MenX 01 lot 001 and MenX 01 lot 005) in 10 µM concentration. Top: Fc=2-1 Bottom: Fc=1, Fc=2.

4. Kinetic analysis

Kinetics and affinity measurements on the sensor surface determine the binding characteristics between the ligand and the analyte. Kinetics deals with time-dependent events and describes how fast molecules bind and dissociate. The kinetics of interaction determines how much complex is formed in a given time. Affinity describes the strength of interaction, independently of time. The affinity of interaction determines how much complex is formed when a mixture of interactants has reached equilibrium. Kinetics is analysed by monitoring the interaction as a function of time over a range of analyte concentrations, and then fitting the whole data set to a mathematical model describing the interaction mechanism. Kinetic parameters are determined from the association and dissociation phases of the sensorgram. Affinity is established either from the kinetic rate constants or by measurement of steady-state analyte binding levels. The analyte concentrations may be introduced in separate cycles with surface regeneration after each sample injection (multi-cycle analysis) or sequentially in a single cycle with no regeneration between sample injections (single-cycle analysis). In this study, the multi-cycle approach has been applied and the affinity is assessed from the kinetic parameters. Two different kinetic experiments have been performed using MenX Fab (Fig.76) and MenX mAb (Fig.77) respectively. The simplest interaction model that is consistent with the experimental system has been chosen, namely 1:1 Langmuir binding model. The kinetic parameters determined from the analysis using MenX Fab are the following: k_a [1/Ms] = 3.195 x 10^3 ± 1.8 x 10^2, k_d [1/s] = 3.957 x 10^{-3} ± 2 x 10^{-4}, K_d [M] = 1.238 x 10^{-6} ± 1.11 x 10^{-6}. The kinetic parameters determined from the MenX mAb experiment are the following: k_a [1/Ms] = 4.7 x 10^3 ± 2.3 x 10^2, k_d [1/s] = 3 x 10^{-3} ± 8 x 10^{-5}, K_d [M] = 6.383 x 10^{-7} ± 3.48 x 10^{-7}.

There are two major tools for assessing the significance of the reported constants: the closeness of fit between the fitted and experimental curves, and the statistical significance of the parameters. Visual inspection of the fitted curves overlaid on the experimental data or the residual plots (Fig.78, Fig.79) gives an indication of the closeness of the fits. Ideally, the residuals scatter randomly around zero over a range that corresponds to the short-term noise in the detection system (approximately ±1 to 2 RU). Both assays, using MenX Fab and MenX mAb, result in well-fitted curves overlaid onto the experimental data, as well as plots showing residuals scattered around zero (± 2RU). The

chi-square value is a quantitative measure of the closeness of fit, which describes the deviation between the experimental and fitted curves, and in an ideal situation will approximate to the square of the short-term noise level. The chi-square value for the MenX Fab analysis is 1.95, and for the MenX mAb experiment, it is 3.12. Fitting experimental data to a mathematical interaction model will return values for all parameters in the model, regardless of whether they are significant or not. The significance of the parameter values returned by the fitting procedure is indicated by the standard error or T-value. A high standard error or low T-value indicates that the parameter has little significance. For the 1:1 binding model, an additional indicator of the parameter significance is the U-value. This parameter represents the uniqueness of the calculated rate constants, determined by testing the dependence of fitting on correlated variations between selected variables. Lower values (<25) indicate greater confidence in the reported kinetic constants. The U-value for the MenX mAb experiment is 12, and for the MenX Fab, it is 7.

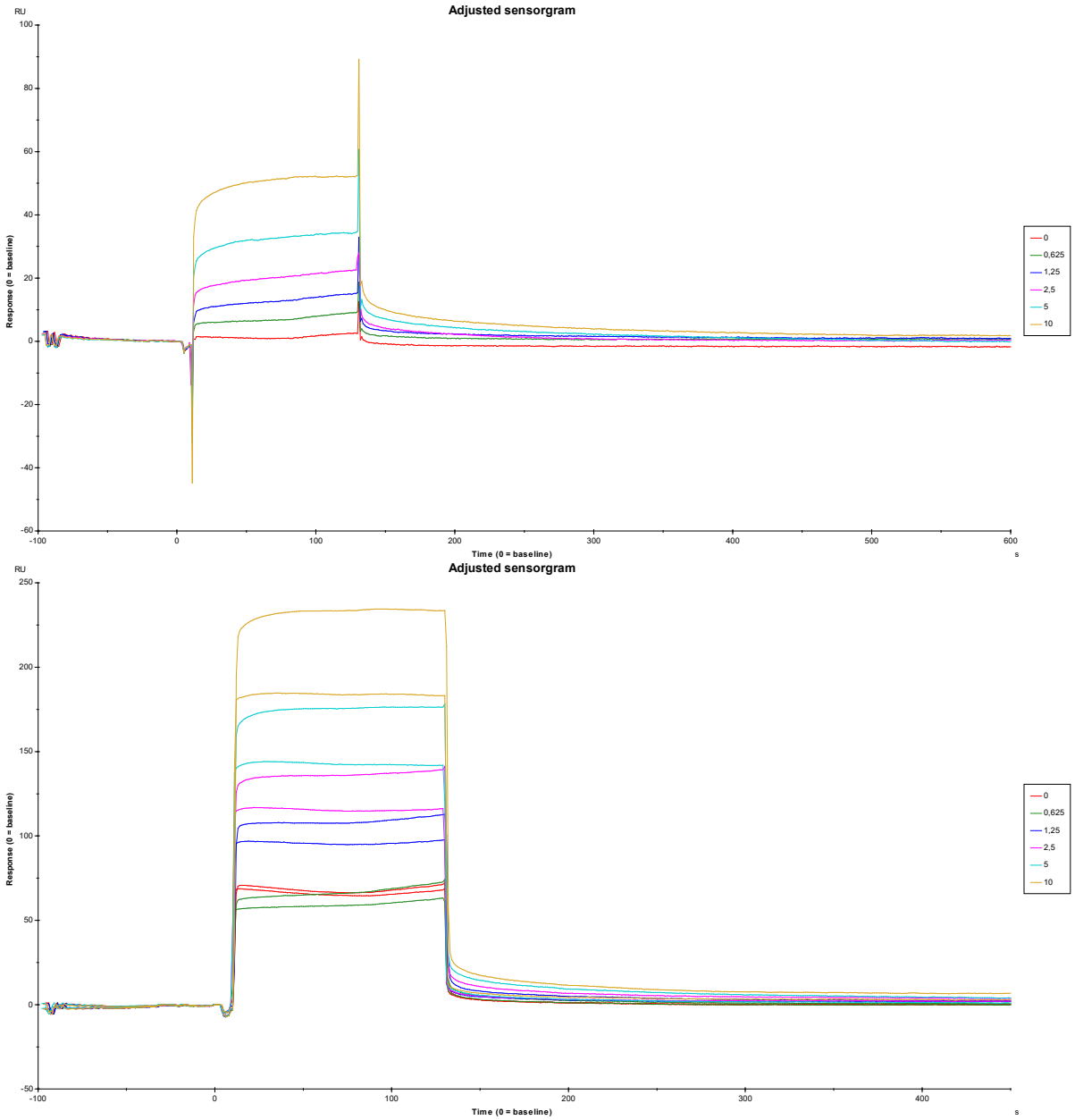


Figure 76. Sensorgrams representing the kinetic experiments using a concentration series of MenX Fab. Top: Fc=2-1 Bottom: Fc=1, Fc=2.

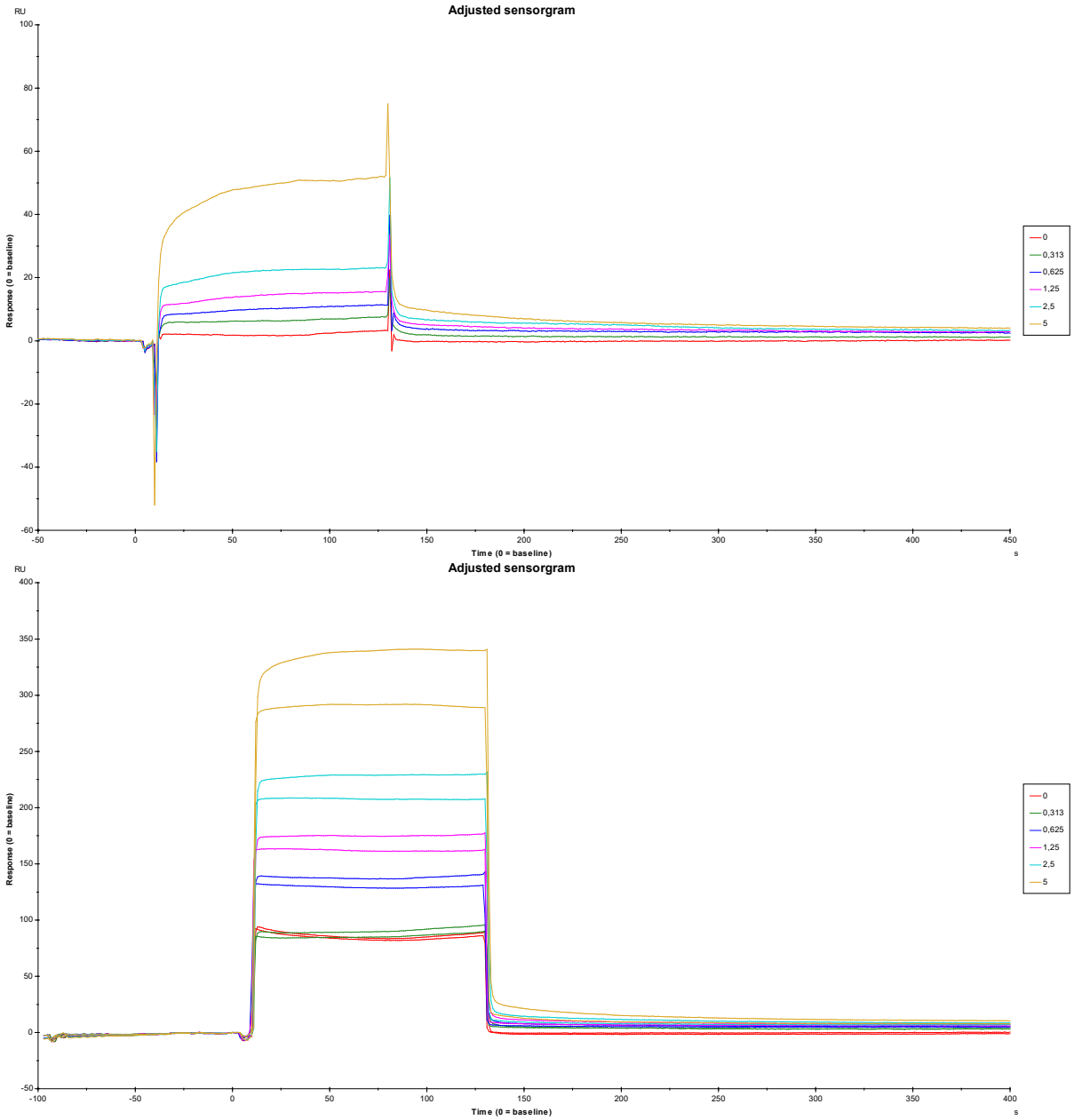


Figure 77. Sensorgrams representing the kinetic exepriments using a concentration series of MenX mAb. Top: Fc=2-1 Bottom: Fc=1, Fc=2.

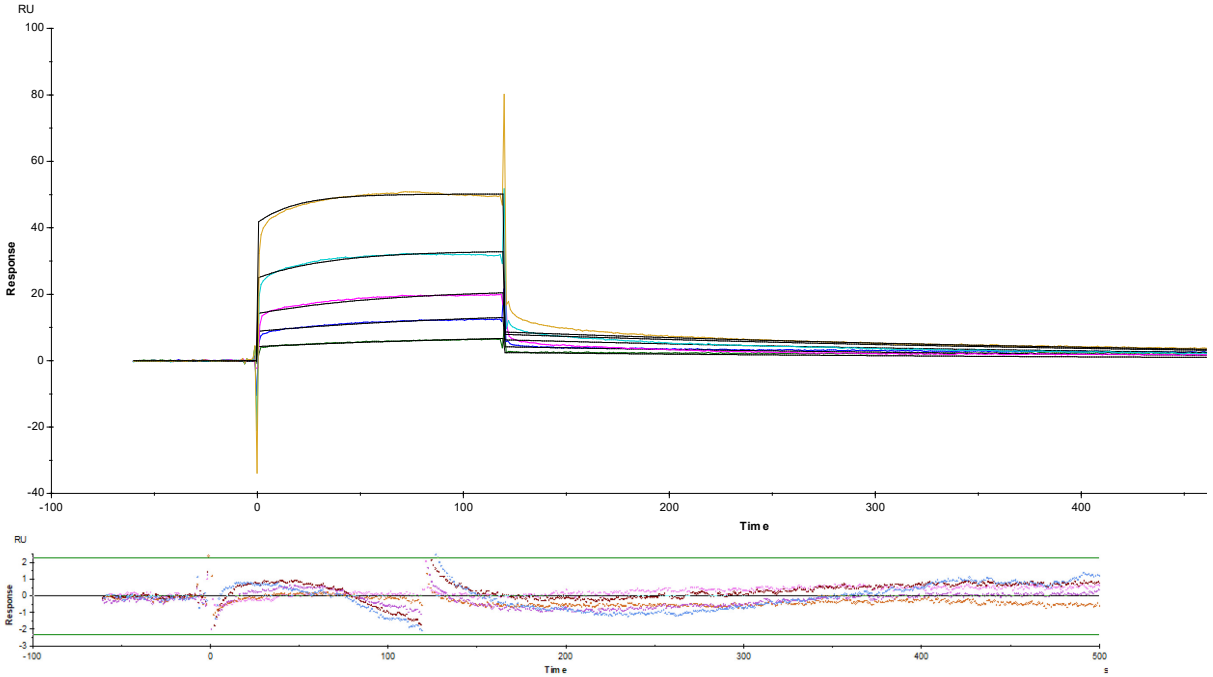


Figure 78. Top: Sensorgram showing a global fitting to the entire set of experimental data (all concentration curves) for MenX Fab. Global fitting finds the best fit for all of the sensorgrams simultaneously. Selected parameters such as rate and affinity constants are constrained to have a single value for all sensorgrams in the data set. Bottom: Residual plot from the global fit.

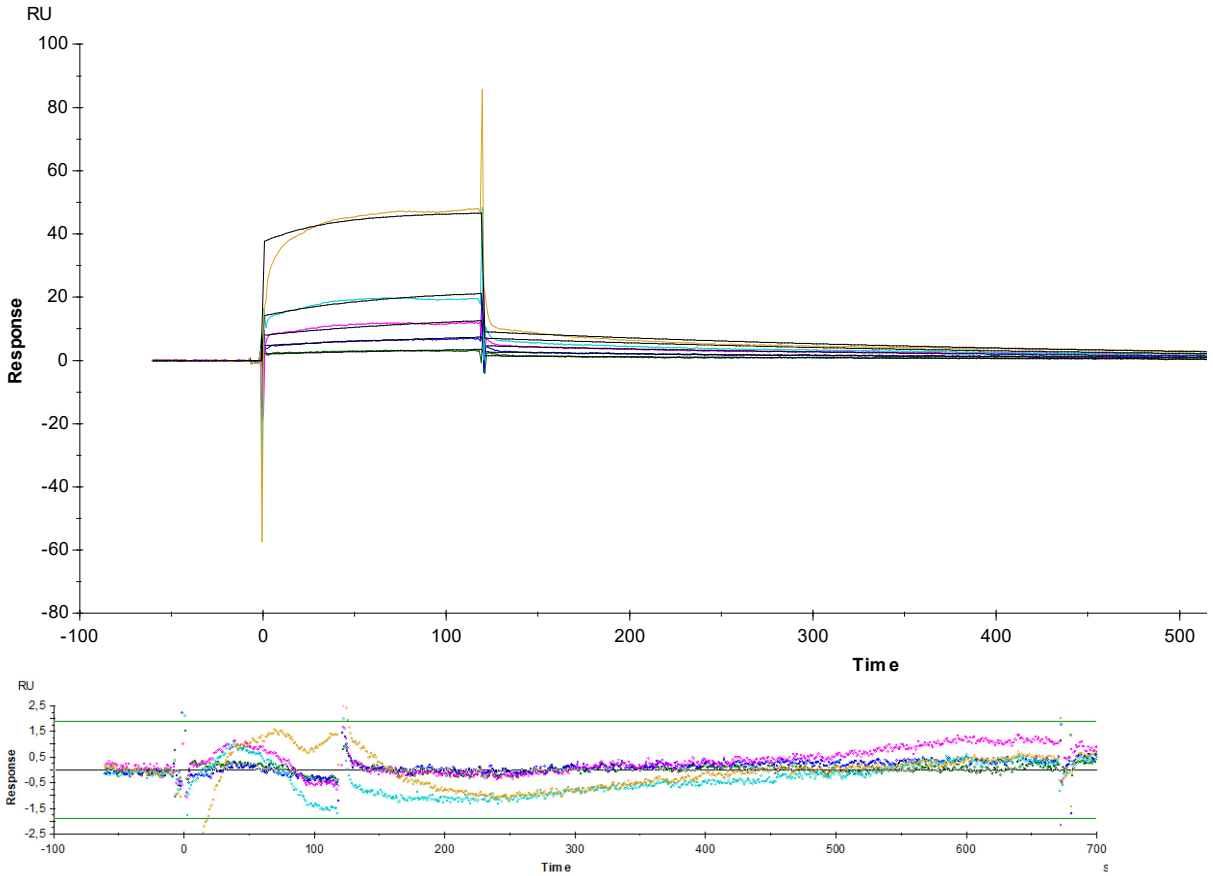


Figure 79. Top: Sensorgram showing a global fitting to the entire set of experimental data (all concentration curves) for MenX mAb. Global fitting finds the best fit for all of the sensorgrams simultaneously. Selected parameters such as rate and affinity constants are constrained to have a single value for all sensorgrams in the data set. Bottom: Residual plot from the global fit.

5. Inhibition assays

Inhibition assays also called solution competition, exploit the ability of the analyte (DPx) to inhibit the binding of the high molecular weight detecting molecule (MenX mAb) to the ligand attached to the chip surface (MenX(DP15)–ADH–CRM_{197}). Obviously, the detecting molecule carrying the bound analyte cannot be able to bind to the surface-attached ligand. Ideally, the detecting molecule should be monovalent for this reason. However, monoclonal antibodies are commonly used as detecting molecules, despite their bivalent binding properties. Even though both antigen-binding sites must be occupied to effectively inhibit the binding of the antibody to the surface, the inhibition assay principle still works reliably. An equal amount of 10 µM MenX mAb has been mixed with various concentrations of DP5, DP10, or PS and incubated to reach equilibrium. Subsequently, the solution mix is injected over the sensor surface. The SPR signal obtained is an indirect measure of the number of free detecting molecules, which is inversely related to the concentration of the analyte in the sample (Lundström 1994). The results of the inhibition assays (Fig.S12, S13, S14) have been analysed and show that the most efficient inhibition is achieved, as expected, with the whole polysaccharide capsule component: 45. 2% inhibition at a 1000 µg/mL PS concentration. The 2nd best choice as a MenX mAb inhibitor is DP5: 32.9% inhibition at a 1000 µg/mL DP5 concentration. The lowest inhibition levels are reached when using DP10: 22.4 % inhibition at a 1000 µg/mL DP10 concentration (Fig.80).

MenX PS		MenX DP10		MenX DP5	
Concentration ug/mL	%Inhibition	Concentration ug/mL	%Inhibition	Concentration ug/mL	%Inhibition
1000	45.2	1000	22.4	1000	32.9
500	37.7	500	19.1	500	29.8
250	29.9	250	6.0	250	23.8
125	19.9	125	8.7	125	19.7
62.5	7.8	62.5	9.0	62.5	12.5
31.25	6.9	31.25	0.0	31.25	8.8
15.63	2.8	15.63	0.0	15.63	5.6
7.81	0.0	7.81	0.0	7.81	0.0
3.91	0.0	3.91	0.0	3.91	4.4
1.95	0.0	1.95	0.0	1.95	4.4
0.98	0.0	0.98	0.0	0.98	5.3
0.49	0.0	0.49	0.0	0.49	7.5

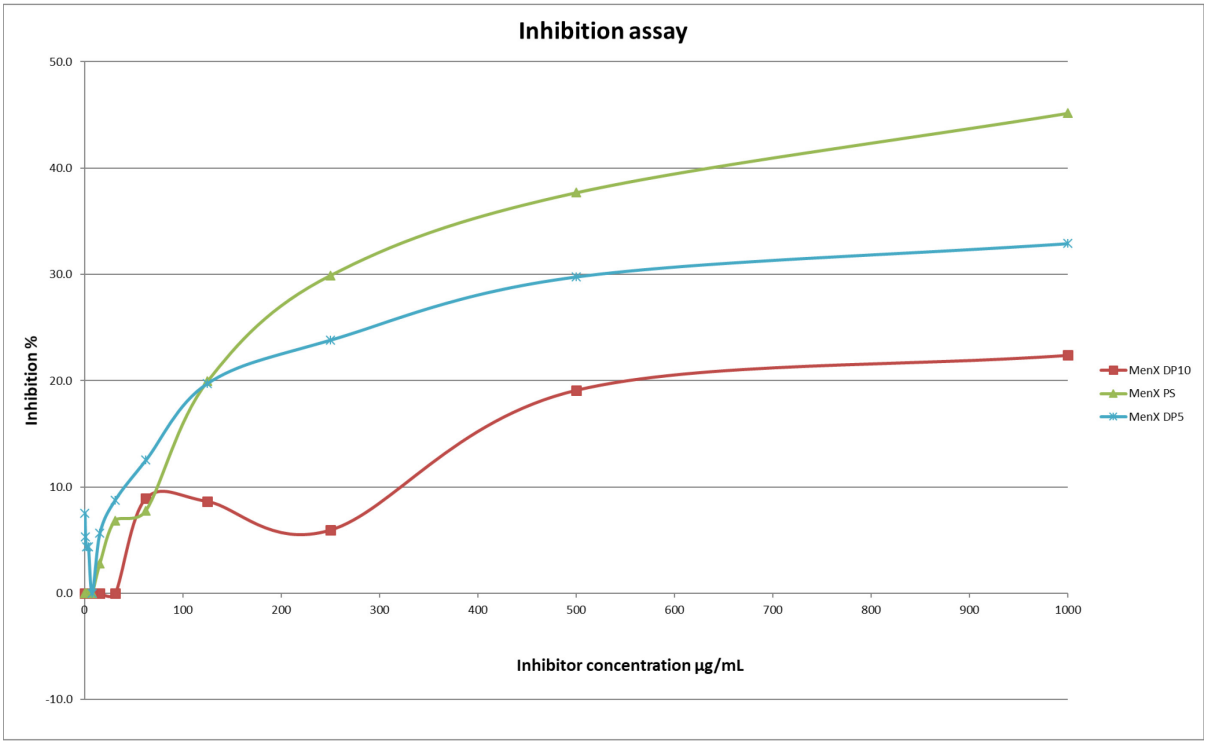


Figure 80. Inhibition assay results using solutions of 5 µM MenX mAb and a wide range of analyte (DP5, DP10, or PS) concentrations. Percentage of inhibition is plotted against the inhibitor (analyte) concentration in µg/mL

DISCUSSION

Systems based on surface plasmon resonance technique can be used to study interactions involving, in principle, any kind of molecule, from organic drug candidates to nucleic acids, proteins, glycoproteins, and even viruses and whole cells. Since the response is a measure of the change in mass concentration, it is proportional to the molecular weight. Smaller molecules give lower molar responses and the practical lower limit for detection is approx. 100 Da. Measurements can be performed on purified interactants, as well as on complex mixtures such as cell culture supernatants or cell extracts. A range of interaction characteristics can be determined from an SPR experiment, including kinetics and affinity assessment (van der Merwe 2001).

Kinetics are analysed by monitoring the interaction as a function of time over a range of analyte concentrations and then fitting the whole data set to a mathematical model describing the interaction. The rate constants reported by the kinetic evaluation have been determined based on a 1:1 Langmuir interaction model and are only valid in that context. If the interaction mechanism is not a simple 1:1 binding, the fitted curves will deviate to some extent from the experimental data and the reported constants will not be a true representation of the interaction kinetics. Moreover, for precise kinetic analysis, the amount of surface-immobilised ligand should not be very hight, so that the maximum response from the analyte binding would be in the region of 100 RU. In this study, the maximum response was recorded at ~ 60 RU during the binding analysis. The main reason for this approach is that low ligand levels help to reduce the rate-limiting effects of mass transfer. Mass transfer is the diffusion-controlled supply of analyte molecules to the surface from the bulk solution. With a high ligand density, the rate at which the analyte binds the ligand may exceed the rate at which the analyte is delivered to the surface. If the transfer is slow in relation to the association rate, the observed binding will be a measure of the diffusion process and not the interaction rates. Consequently, the analysis of association rate under mass transport limited conditions will yield an apparent k_{on} that is slower than the real association rate constant. During the association phase, there is a simultaneous association and dissociation of the analyte, and equilibrium is reached when the association rate equals the dissociation rate. Under ideal experimental conditions, during the dissociation phase, only dissociation of the

analyte should occur. In reality, following the dissociation, before diffusing out of the matrix and being washed from the flow cell, some re-binding of the analyte to the unoccupied ligand may follow. As a consequence, the measured dissociation rate constant would be slower than the real k_{off}. Because mass transport may hinder binding, it is essential to use the lowest density of ligand that gives an adequate level of analyte binding. For this study, the lowest ligand density offering a satisfactory binding level has been established to be ~450 RU. Furthermore, to reduce mass transfer limitation, the characterisation of the antibody-antigen interaction should be run at moderate flow rates and could be slightly increased when needed. All the experiments have been performed at a relatively high flow rate of 45 µl/min. Obtaining accurate kinetic data is a very demanding and time-consuming task, and requires a thorough understanding of binding kinetics and the potential sources of artefacts. Values for affinity constants have been derived from the kinetic measurements. These values, just like the values for rate constants, are only valid in the context of the model used to analyse the binding data. More complex binding mechanisms do not always give straightforward relationships between affinity and kinetics (van der Merwe 2001). Micromolar K_d (M) values for both MenX Fab ($1.238 \times 10^{-6} \pm 1.11 \times 10^{-6}$) and MenX mAb ($6.383 \times 10^{-7} \pm 3.48 \times 10^{-7}$) indicate a medium-strength binding interaction.

The recent removal of meningococcal serogroup A strains from the population in most meningitis belt countries may give an advantage to MenX, previously less able to compete with the more virulent serogroup A. Indeed, a recent study on the MenA conjugate vaccine in Burkina Faso found a significant increase in *Neisseria meningitidis* serogroup X carriage following the introduction of the vaccine (Kristiansen et al., 2013). Although the need for a vaccine against the meningococcal serogroup X has been recognised, no appropriate conjugate has been formulated yet.

The inhibition assay performed with three different saccharides indicate that DP5 is the shortest and most efficient (32.9% inhibition at a 1000 µg/mL concentration) polymer to inhibit binding of MenX mAb to the immobilised MenX(DP15)–ADH–CRM_{197}. This finding might finally provide the long-awaited candidate for a MenX conjugate vaccine since more knowledge on the minimal glyco-epitope required for the immunological activity of MenX PS is still needed. Several attempts to reveal the minimal glyco-epitope have been made in recent years. In a study from 2013 by Micoli et al., three different glycoconjugates were prepared by coupling *Neisseria meningitidis* serogroup X

oligosaccharides of varying chain length to CRM_{197} as the carrier protein, namely $MenX_{OX}$–CRM_{197}, MenX–NH_2–SIDEA–CRM_{197}, and MenX–ADH–SIDEA–CRM_{197}. The average degree of polymerisation (avDP) was of 15-20 and 80-100. All glycoconjugates induced high anti-MenX PS IgG titers in mice following immunisation and generated functional antibodies when tested against African Men X isolates. A step further was a subsequent study in the following year by Morelli et al. in which three fully synthetic MenX PS oligomers (monomer, dimer, and trimer) were conjugated to CRM_{197}. Following immunisation with the obtained glycoconjugates, it was demonstrated that the minimal structure possessing immunogenic activity is the conjugated trimer. While the CRM_{197} conjugates of the monomer and the dimer did not induce polysaccharide specific IgG titers in mice, the conjugated trimer elicited significant anti-MenX PS IgG titers. Nevertheless, its immunogenicity towards the native polysaccharide was shown to be significantly lower to the one achieved with the pentadecamer (avDP15) obtained from the native polymer and conjugated to the same protein carrier, both in terms of IgG levels generated and functionality of the induced antibodies. Importantly, all the conjugates induced very low anti-MenX IgM titers, but the trimer and the pentadecamer antigen conjugates enabled immunoglobulin class switching from IgM to IgG, which is characteristic of the T-cell dependent response. In 2015 Harale et al. reported that a conjugate of a tetrameric repeating unit of MenX PS and the tetanus toxoid (TT) protein was synthesised. It was demonstrated that the tetramer conjugate, as well as the tetrasaccharide alone, were antigenic and able to neutralise the rabbit antiserum against *Neisseria meningitidis* serogroup X, indicating the immunogenic potential of the candidate. Both compounds inhibited the binding of antibodies to the bacterial MenX polysaccharide but the unconjugated tetramer showed lower inhibition (up to 68% inhibition) compared to its conjugate formulation (up to 89% inhibition) at all different concentrations tested (10–1000 mg/mL). Unsurprisingly, the capsular polysaccharide of serogroup X demonstrated the highest inhibition compared to the test candidates at all respective antigen concentrations used. These difference may indicate the need for a slightly bigger molecule to be able to neutralise antibodies.

The use of shorter saccharides has allowed the production of more defined and less cross-linked conjugates, with advantages in relation to reproducibility and vaccine characterisation. Parameters including conjugation chemistry (random against selective carbohydrate modification), linker used for coupling, and saccharide chain length, can

also have an impact on the immunogenicity of the corresponding conjugate vaccine (Micoli et al., 2013).

All the above-mentioned findings suggest that oligomers a little longer than three or four repeating units are probably needed to mimic the antigenic activity of the native MenX polysaccharide. As the minimal PS portion which can confer protection against meningococcal serogroup X infections is still unknown, this study presents DP5 as a promising candidate for the development of a glycoconjugate vaccine. It would be advisable to further confirm the kinetics measurements using e.g. ITC, that could additionally determine thermodynamic binding parameters. Moreover, employing STD-NMR spectroscopy might help identify the binding epitope of the ligand, shining further light on the minimal oligosaccharide structure required for the immunological activity of the MenX PS antigen.

CONCLUSIONS

Pure monoclonal antibodies with preserved immunoreactivity are not only needed for research applications but are also employed to treat various diseases and for diagnostic purposes. The use of pure antibodies is critical since any non-specific protein present in the antibody solution may affect the outcome. Monoclonal antibodies of the IgG class are commonly applied, thus many effective purification schemes have been established for them. In turn, due to the particular structure and oligomerisation of IgM antibodies, they behave differently and it has been much more difficult to efficiently purify them. In addition, IgMs are generally less soluble in buffers as they tend to aggregate, and are also poorly soluble in solutions of high or very low salt. Moreover, their behaviour is strictly individual and varies significantly among different IgMs depending on the species of origin, source, preparation, oligomerisation, and the buffers used. Therefore each IgM is truly unique and purification procedures have to be adjusted to best suit the IgM of interest (Mahassni et al., 2009).

Glycan structures are heterogeneous and diverse because of the possible combinations of available monosaccharides, linkages, branching, and variable lengths of glycan chains. This diversity combined with the difficulty in stabilising the flexible branched structures makes that characterisation of the glycoproteome challenging and time-consuming. Furthermore, the non-template synthesis of carbohydrates considerably limits their generation in large quantities needed for analysis. Consequently, the development of fast and sensitive techniques for glycosylation analysis has been crucial for research (Feizi et al., 2003).

The overall weak binding of carbohydrates is usually accompanied by relatively fast k_{on} and k_{off} rates. Suitable on- and off-rates for the antigen-antibody interaction are critical for the successful implementation of these methods, which otherwise may result in poor signal-to-noise ratios, thus making it difficult to distinguish the free from the bound state. There are different techniques that can be employed to unravel the features of protein-carbohydrate interactions. Generally, the experimental data collected from biophysical techniques, such as surface plasmon resonance or isothermal titration calorimetry, is complemented with information acquired from structural studies,

including NMR spectroscopy, cryo-electron microscopy, and X-ray crystallography. Additionally, molecular modelling methods are frequently used to provide a full picture of the structure under investigation.

The glycan array screening revealed the minimal glyco-epitope to be $Man_2GlcNAc_2$, whereas $Man_3GlcNAc_2$ manifested one of the strongest interactions with Mannitou antibody. Surface plasmon resonance measurements established a micromolar binding affinity of Manniotu Fab towards $Man_3GlcNAc_2$ (K_d = 48 μM), with the value subsequently confirmed by isothermal titration calorimetry (K_d = 54 μM). Additionally, the calorimetric data indicates an enthalpy-driven process (ΔH = -7.12 ± 0.225 kcal/mol) with an unfavourable entropic contribution (-TΔS = 1.3 kcal/mol). The mapping of the binding epitope by saturation transfer difference NMR demonstrated Manα1-3 as the main residue involved in Mannitou antibody recognition, with a notable contribution from βMan and the acetyl group of GlcNAc1. All these findings are consistent with each other and point to Manα1-3 branch as the main binding epitope, explaining why any substitution in this arm inhibits the recognition of paucimannose *N*-glycans by Mannitou antibody.

From the attempts to characterise Manniotu applying approaches, such as homology modelling, cryo-electron microscopy, and crystallisation, it was concluded that this IgM, as any other immunoglobulin M investigated so far, is an incredibly flexible formation. Moreover, it cannot be ruled out that upon binding some part of the oligosaccharide may sustain significant flexibility in the antibody-carbohydrate complex. Elucidation of the molecular basis of the formation of the complexes also requires consideration of the balance between the enthalpic and entropic contribution involved in the binding. The lower affinities observed for carbohydrate-specific antibodies derive from the binding not being driven only by enthalpic factors and emphasise the relative importance of entropic considerations. All this, combined with the difficulties to produce high quantities of pure and active Mannitou IgM makes it incredibly challenging to extract detailed structural information on the antibody and its complexes. Indeed, one of the main reasons for the slow progress in translating the basic research findings on natural IgM to clinical application is the issue inherent to the nature of the molecules themselves. The existence of IgM in either monomeric, pentameric, or hexameric forms presents a dilemma regarding the identification of the structure of the therapeutic molecule that will achieve the highest profiles for efficacy, tolerability, and safety. The pentameric form

poses further potential challenges regarding the stability of such manufactured IgM-based therapeutics. Similar to IgG, the glycosylation pattern of IgM can also play a critical role in its effector functions. These specialised effector functions of IgM provide clinical opportunities but also pose challenges for the purification of well-characterised homogeneous preparations that are required for therapeutic applications.

The upregulation of paucimannosylation in pathophysiological conditions makes it an attractive biomarker for disease states, where Mannitou antibody may serve as a potential diagnostic and therapeutic agent. The evidence that Mannitou antibody inhibits cell proliferation and shows strong staining in metastatic prostate cancer tissue, as compared to weak staining in well-differentiated tissues, makes Mannitou a promising candidate for application in cancer research (Becker et al., 2019; Zipser et al., 2012). The wide repertoire of methods for studying protein-glycan interactions can be used to evaluate the immune response to preventative or therapeutic carbohydrate vaccines.

REFERENCES

1. [No authors listed] **2005**. Menactra: a meningococcal conjugate vaccine. *Med Lett Drugs Ther*. 47: 29–31.

2. [No authors listed] **2010**. A new conjugate meningococcal vaccine (Menveo). *Med Lett Drugs Ther*. 52: 59–60.

3. Akhouri RR, Goel S, Furusho H, Skoglund U, Wahlgren M. **2016**. Architecture of human IgM in complex with *P. falciparum* erythrocyte membrane protein 1. *Cell Rep*. 14: 723–736.

4. Al-Lazikani B, Lesk AM, Chothia C. **1997**. Standard conformations for the canonical structures of immunoglobulins. *J Mol Biol*. 273:927–948.

5. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. **2002**. Molecular Biology of the Cell. Garland Science: New York, NY.

6. Altmann F, Schwihla H, Staudacher E, Glössl J, März L. **1995**. Insect cells contain an unusual, membrane-bound β-*N*-acetylglucosaminidase probably involved in the processing of protein *N*-glycans. *J Biol Chem*. 270: 17344–17349.

7. Angeloni S, Ridet JL, Kusy N, Gao H, Crevoisier F, Guinchard S, Kochhar S, Sigrist H, Sprenger N. **2005**. Glycoprofiling with micro-arrays of glycoconjugates and lectins. *Glycobiology* 15: 31–41.

8. Angulo J, Nieto PM. **2011**. STD-NMR: application to transient interactions between biomolecules-a quantitative approach. *Eur Biophys J*. 40: 1357–1369.

9. Argade S, Hopfer RL, Strang AM, van Halbeek H, Alhadeff JA. **1988**. Structural studies on the carbohydrate moieties of human liver alpha-L-fucosidase. *Arch Biochem Biophys*. 266: 227–247.

10. Aricescu AR, Lu W, Jones EY. **2006**. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr D Biol Crystallogr*. 62: 1243–1250.

11. Arnold JN, Wormald MR, Suter DM, Radcliffe CM, Harvey DJ, Dwek RA, Rudd PM, Sim RB. **2005**. Human serum IgM glycosylation: identification of glycoforms that can bind to mannan-binding lectin. *J Biol Chem*. 280: 29080–29087.

12. Ashwell G, Harford J. **1982**. Carbohydrate-specific receptors of the liver. *Annu Rev Biochem*. 51: 531–554.

13. Aspberg A. **2012**. The different roles of aggrecan interaction domains. *J Histochem Cytochem*. 60: 987–996.

14. Avezov E, Frenkel Z, Ehrlich M, Herscovics A, Lederkremer GZ. **2008**. Endoplasmic reticulum (ER) mannosidase I is compartmentalized and required for N-glycan trimming to Man5-6GlcNAc2 in glycoprotein ER-associated degradation. *Mol Biol Cell*. 19:216–225.

15. Bajt ML, Schmitz B, Schachner M, Zipser B. **1990**. Carbohydrate epitopes involved in neural cell recognition are conserved between vertebrates and leech. *J Neurosci Res*. 27: 276–285.

16. Baldi L, Hacker DL, Adam M, Wurm FM. **2007**. Recombinant protein production by large-scale transient gene expression in mammalian cells: state of the art and future perspectives. *Biotechnol Lett*. 29: 677–684.

17. Balog CI, Stavenhagen K, Fung WL, Koeleman CA, McDonnell LA, Verhoeven A, Mesker WE, Tollenaar RA, Deelder AM, Wuhrer M. **2012**. *N*-glycosylation of colorectal cancer tissues: a liquid chromatography and mass spectrometry-based investigation. *Mol Cell Proteomics* 11: 571–585.

18. Beck A, Wurch T, Bailly C, Corvaia N. **2010**. Strategies and challenges for the next generation of therapeutic antibodies. *Nat Rev Immunol*. 10: 345–352.

19. Becker W, Bhattiprolu KC, Gubensäk N, Zangger K. **2018**. Investigating protein-ligand interactions by solution nuclear magnetic resonance spectroscopy. *Chemphyschem*. 19: 895–906.

20. Becker Y, Förster S, Gielen GH, Loke I, Thaysen-Andersen M, Laurini C, Wehrand K, Pietsch T, Diestel S. **2019**. Paucimannosidic glycoepitopes inhibit tumorigenic processes in glioblastoma multiforme. *Oncotarget* 10: 4449–4465.

21. Beckman RA, Weiner LM, Davis HM. **2007**. Antibody constructs in cancer therapy: protein engineering strategies to improve exposure in solid tumours. *Cancer* 109: 170–179.

22. Bergstrom KS, Xia L. **2013**. Mucin-type O-glycans and their roles in intestinal homeostasis. *Glycobiology* 23: 1026–1037.

23. Beseničar M, Maček P, Lakey JH, Anderluh G. **2006**. Surface plasmon resonance in protein-membrane interactions. *Chem Phys Lipids*. 141: 169–178.

24. Bhattacharyya L, Khan MI, Fant J, Brewer CF. **1989**. Formation of highly ordered cross-linked lattices between asparagine-linked oligosaccharides and lectins observed by electron microscopy. *J Biol Chem*. 264: 11543–11545.

25. Birchenough GM, Nyström EE, Johansson ME, Hansson GC. **2016**. A sentinel goblet cell guards the colonic crypt by triggering Nlrp6-dependent Muc2 secretion. *Science* 352: 1535–1542.

26. Bird RE, Hardman KD, Jacobson JW, Johnson S, Kaufman BM, Lee SM, Lee T, Pope SH, Riordan GS, Whitlow M. **1988**. Single-chain antigen-binding proteins. *Science* 242: 423–426.

27. Blattman JN, Greenberg PD. **2004**. Cancer immunotherapy: a treatment for the masses. *Science* 305: 200–205.

28. Blaum BS, Hannan JP, Herbert AP, Kavanagh D, Uhrin D, Stehle T. **2015**. Structural basis for sialic acid-mediated self-recognition by complement factor H. *Nat Chem Biol*. 11: 77–82.

29. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, Razi N, Stevens DJ, Skehel JJ, van Die I, Burton DR, Wilson IA, Cummings R, Bovin N, Wong CH, Paulson JC. **2004**. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci USA* 101: 17033-17038.

30. Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S. **2004**. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4: 1633–1649.

31. Boekema E. J., Folea M., Kouřil R. **2009**. Single-particle electron microscopy. *Photosynth Res*. 102: 189–196.

32. Boes M. **2000**. Role of natural and immune IgM antibodies in immune responses. *Mol Immunol*. 37: 1141–1149.

33. Boisier P, Nicolas P, Djibo S, Taha MK, Jeanne I, Maïnassara HB, Tenebray B, Kairo KK, Giorgini D, Chanteau S. **2007**. Meningococcal meningitis:

Unprecedented incidence of serogroup X-related cases in 2006 in Niger. *Clin Infect Dis*. 44: 657–663.

34. Bories S, Slaterus KW, Faucon R, Audiffren P, Vandekerkove M. **1966**. Peut-on individualiser deux nouveaux groupes sérologiques de *Neisseria meningitidis? Med Trop*. 26: 603–616.

35. Bouchemal K. **2008**. New challenges for pharmaceutical formulations and drug delivery systems characterization using isothermal titration calorimetry. *Drug Discov Today* 13: 960–972.

36. Boulianne GL, Hozumi N, Shulman MJ. **1984**. Production of functional chimaeric mouse/human antibody. *Nature* 312: 643–646.

37. Brinkman-Van der Linden EC, Hurtado-Ziola N, Hayakawa T, Wiggleton L, Benirschke K, Varki A, Varki N. **2007**. Human-specific expression of Siglec-6 in the placenta. *Glycobiology* 17: 922–931.

38. Bronowska AK. **2011**. Thermodynamics of Ligand-Protein Interactions: Implications for Molecular Design. In *Thermodynamics - Interaction Studies - Solids, Liquids and Gases*. Moreno Piraján JC, Ed. InTech: Rijeka, Croatia; pp 1–48.

39. Brooks SA. **2009**. Strategies for analysis of the glycosylation of proteins: current status and future perspectives. *Mol Biotechnol*. 43: 76–88.

40. Bruckner K, Perez L, Clausen H, Cohen S. **2000**. Glycosyltransferase activity of Fringe modulates Notch-Delta interactions. *Nature* 406: 411–415.

41. Brzezicka K, Echeverria B, Serna S, van Diepen A, Hokke CH, Reichardt NC. **2015**. Synthesis and microarray-assisted binding studies of core xylose and fucose containing *N*-glycans. *ACS Chem Biol*. 10: 1290–1302.

42. Buckholz RG, Gleeson MA. **1991**. Yeast systems for the commercial production of heterologous proteins. *Biotechnology (NY)* 9: 1067–1072.

43. Bundle DR, Sigurskjold BW. **1994**. Determination of accurate thermodynamics of binding by titration microcalorimetry. *Meth Enzymol*. 247: 288–305.

44. Bundle DR, Jennings HJ, Kenny CP. **1973**. An improved procedure for the isolation of meningococcal, polysaccharide antigens, and the structural determination of the antigen from serogroup X. *Carbohydr Res*. 26: 268–270.

45. Bundle DR, Jennings HJ, Kenny CP. **1974**. Studies on the group-specific polysaccharide of *Neisseria meningitidis* serogroup X and an improved procedure for its isolation. *J Biol Chem*. 249: 4797–4801.

46. Bundle DR, Smith IC, Jennings HJ. **1974**. Determination of the structure and conformation of bacterial polysaccharides by carbon 13 nuclear magnetic resonance. Studies on the group-specific antigens of *Neisseria meningitidis* serogroups A and X. *J Biol Chem.* 249: 2275–2281.

47. Carlin AF, Uchiyama S, Chang YC, Lewis AL, Nizet V, Varki A. **2009**. Molecular mimicry of host sialylated glycans allows a bacterial pathogen to engage neutrophil Siglec-9 and dampen the innate immune response. *Blood* 113: 3333–3336.

48. Casali P. **1998**. IgM. In *Encyclopedia of Immunology*; Delves PJ, Ed. Academic Press: Cambridge, MA; pp 1212–1217.

49. Chai W, Feizi T, Yuen CT, Lawson AM. **1997**. Nonreductive release of *O*-linked oligosaccharides from mucin glycoproteins for structure/function assignments as neoglycolipids: Application in the detection of novel ligands for E-selectin. *Glycobiology* 7: 861–872.

50. Chaires JB. **2008**. Calorimetry and thermodynamics in drug design. *Annu Rev Biophys*. 37: 135–151.

51. Chames P, Van Regenmortel M, Weiss E, Baty D. **2009**. Therapeutic antibodies: successes, limitations and hopes for the future. *Br J Pharmacol.* 157: 220–233.

52. Chan AC, Carter PJ. **2010**. Therapeutic antibodies for autoimmunity and inflammation. *Nat Rev Immunol.* 10: 301–316.

53. Chang CE, Gilson MK. **2004**. Free energy, entropy, and induced fit in host-guest recognition: Calculations with the second-generation mining minima algorithm. *J Am Chem Soc*. 126: 13156–13164.

54. Chatterjee S, Lee LY, Kawahara R, Abrahams JL, Adamczyk B, Anugraham M, Ashwood C, Sumer-Bayraktar Z, Briggs MT, Chik JHL, Everest-Dass A, Förster S, Hinneburg H, Leite KRM, Loke I, Möginger U, Moh ESX, Nakano M, Recuero S, Sethi MK, Srougi M, Stavenhagen K, Venkatakrishnan V, Wongtrakul-Kish K, Diestel S, Hoffmann P, Karlsson NG, Kolarich D, Molloy MP, Muders MH, Oehler MK, Packer NH, Palmisano G, Thaysen-Andersen M. **2019**. Protein paucimannosylation is an enriched *N*-glycosylation signature of human cancers. *Proteomics* 19: e1900010.

55. Chen GY, Brown NK, Zheng P, Liu Y. **2014**. Siglec-G/10 in selfnonself discrimination of innate and adaptive immunity. *Glycobiology* 24: 800–806.

56. Chiavenna SM, Jaworski JP, Vendrell A. **2017**. State of the art in anti-cancer mAbs. *J Biomed Sci*. 24: 15.

57. Chothia C, Lesk AM. **1987**. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol*. 196: 901–917.

58. Chothia C, Lesk AM, Gherardi E, Tomlinson IM, Walter G, Marks JD, Llewelyn MB, Winter G. **1992**. Structural repertoire of the human V(H) segments. *J Mol Biol*. 227: 799–817.

59. Chothia C, Lesk AM, Tramontano A, Levitt M, Smithgill SJ, Air G, Sheriff S, Padlan EA, Davies D, Tulip WR, et al. **1989**. Conformations of immunoglobulin hypervariable regions. *Nature* 342: 877–883.

60. Chung AW, Crispin M, Pritchard L, Robinson H, Gorny MK, Yu X, Bailey-Kellogg C, Ackerman ME, Scanlan C, Zolla-Pazner S, Alter G. **2014**. Identification of antibody glycosylation structures that predict monoclonal antibody Fc-effector function. *AIDS* 28: 2523–2530.

61. Cipollo JF, Costello CE, Hirschberg CB. **2002**. The fine structure of *Caenorhabditis elegans N*-glycans. *J Biol Chem*. 277: 49143–49157.

62. Clark GF. **2014**. A role for carbohydrate recognition in mammalian sperm-egg binding. *Biochem Biophys Res Commun*. 450: 1195–1203.

63. Clausen H, Wandall HW, Steentoft C, Stanley P, Schnaar RL. **2017**. Glycosylation Engineering. In *Essentials of Glycobiology*; Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH, Ed. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.

64. Cohen M, Varki A. **2014**. Modulation of glycan recognition by clustered saccharide patches. *Int Rev Cell Mol Biol*. 308: 75–125.

65. Corbett PT, Tong LH, Sanders JK, Otto S. **2005**. Diastereoselective amplification of an induced-fit receptor from a dynamic combinatorial library. *J Am Chem Soc*. 127: 8902–8903.

66. Costantino P, Rappuoli R, Berti F. **2011**. The design of semi-synthetic and synthetic glycoconjugate vaccines. *Expert Opin Drug Discov*. 6: 1045–1066.

67. Cummings RD, Schnaar RL, Esko JD, Drickamer K, Taylor ME. **2017**. Principles of Glycan Recognition. In *Essentials of Glycobiology*; Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH, Ed. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.

68. Cummings RD, Smith DF. **1992**. The selectin family of carbohydratebinding proteins: Structure and importance of carbohydrate ligands for cell adhesion. *BioEssays* 14: 849–856.

69. Czajkowsky DM, Sha Z. **2009**. The human IgM pentamer is a mushroom-shaped molecule with a flexural bias. *Proc Natl Acad Sci USA* 106: 14960–14965.

70. Dahmen AC, Fergen MT, Laurini C, Schmitz B, Loke I, Thaysen-Andersen M, Diestel S. **2015**. Paucimannosidic glycoepitopes are functionally involved in proliferation of neural progenitor cells in the subventricular zone. *Glycobiology* 25: 869–880.

71. Dam S, Thaysen-Andersen M, Stenkjær E, Lorentzen A, Roepstorff P, Packer NH, Stougaard J. **2013**. Combined *N*-glycome and *N*-glycoproteome analysis of the *Lotus japonicus* seed globulin fraction shows conservation of protein structure and glycosylation in legumes. *J Proteome Res*. 12: 3383–3392.

72. Dammeyer T, Steinwand M, Krüger SC, Dübel S, Hust M, Timmis KN. **2011**. Efficient production of soluble recombinant single chain Fv fragments by a *Pseudomonas putida* strain KT2440 cell factory. *Microb Cell Fact.* 10: 11.

73. Davies DR, Padlan EA, Sheriff S. **1990**. Antibody-antigen complexes. *Annu Rev Biochem*. 59: 439–473.

74. Davis AC, Roux KH, Shulman MJ. **1988**. On the structure of polymeric IgM. *Eur J Immunol*. 18: 1001–1008.

75. de La Motte CA, Hascall VC, Calabro A, Yen-Lieberman B, Strong SA. **1999**. Mononuclear leukocytes preferentially bind via CD44 to hyaluronan on human intestinal mucosal smooth muscle cells after virus infection or treatment with poly(I.C). *J Biol Chem*. 274: 30747–30755.

76. Dean N, Gao XD. **2014**. Dolichyl-Phosphate (UDP-N-Acetylglucosamine) N-Acetylglucosaminephospho Transferase 1 (GlcNAc-1-P Transferase) (DPAGT1). In *Handbook of Glycosyltransferases and Related Genes*; Taniguchi N, Honke K, Fukuda M, Narimatsu H, Yamaguchi Y, Angata T, Ed. Springer: Tokyo, Japan; pp 1223–1230.

77. Decosas J, Koama JB. **2002**. Chronicle of an outbreak foretold: Meningococcal meningitis W135 in Burkina Faso. *Lancet Infect Dis*. 2: 763–765.

78. Delrieu I, Yaro S, Tamekloé TAS, Njanpop-Lafourcade BM, Tall H, Jaillard P, Ouedraogo MS, Badziklou K, Sanou O, Drabo A, Gessner BD, Kambou JL, Mueller JE. **2011**. Emergence of epidemic *Neisseria meningitidis* serogroup X meningitis in Togo and Burkina Faso. *PLoS One* 6: e19513.

79. Dennis JW, Lau KS, Demetriou M, Nabi IR. **2009**. Adaptive regulation at the cell surface by N-glycosylation. *Traffic* 10: 1569–1578.

80. Dessau M. A. Modis Y. **2011**. Protein crystallization for X-ray crystallography. *J Vis Exp*. 47: e2285.

81. Djibo S, Nicolas P, Alonso JM, Djibo A, Couret D, Riou JY, Chippaux JP. **2003**. Outbreaks of serogroup X meningococcal meningitis in Niger 1995–2000. *Trop Med Int Health* 8: 1118–1123.

82. Doyle ML, Hensley P. **1997**. Experimental Dissection of Protein-Protein Interactions in Solution. In *Advances in Molecular and Cell Biology*. Bittar E, Allewell NM, Woodward C, Ed. JAI Press Inc.: Greenwich, CT; pp 279–337.

83. Doyle ML. **1997**. Characterization of binding interactions calorimetry. *Curr Opin Biotechnol*. 8: 31–35.

84. Dragan AI, Read CM, Crane-Robinson C. **2017**. Enthalpy–entropy compensation: the role of solvation. *Eur Biophys J*. 46: 301–308.

85. Du X, Li Y, Xia YL, Ai SM, Liang J, Sang P, Ji XL, Liu SQ. **2016**. Insights into protein–ligand interactions: mechanisms, models, and methods. *Int J Mol Sci*. 17: 144.

86. Edelman L, Margaritte C, Chaabihi H, Monchâtre E, Blanchard D, Cardona A, Morin F, Dumas G, Petres S, Kaczorek M. **1997**. Obtaining a functional recombinant anti-rhesus (D) antibody using the baculovirus-insect cell expression system. *Immunology* 91: 13–19.

87. Eisenberg SA. **2012**. Biologic therapy. *J Infus Nurs*. 35: 301–313.

88. Eldin P, Pauza ME, Hieda Y, Lin G, Murtaugh MP, Pentel PR, Pennell CA. **1997**. High-level secretion of two antibody single-chain Fv fragments by *Pichia pastoris*. *J Immunol Methods* 201: 67–75.

89. Elgert KD. **2009**. Antibody Structure and Function. In *Immunology: Understanding The Immune System*. Wiley-Blackwell: 58–78.

90. Elgert KD. **2009**. Introduction to the Immune System. In *Immunology: Understanding The Immune System*. Wiley-Blackwell: 1–25.

91. Espinosa-Marzal RM, Fontani G, Reusch FB, Roba M, Spencer ND, Crockett R. **2013**. Sugars communicate through water: Oriented glycans induce water structuring. *Biophys J*. 104: 2686–2694.

92. Evans JR, Artenstein MS, Hunter DH. **1968**. Prevalence of meningococcal serogroups and description of three new groups. *Am J Epidemiol*. 87: 643–646.

93. Everest-Dass AV, Briggs MT, Kaur G, Oehler MK, Hoffmann P, Packer NH. **2016**. *N*-glycan MALDI imaging mass spectrometry on formalin-fixed paraffin-embedded tissue enables the delineation of ovarian cancer tissues. *Mol Cell Proteomics* 15: 3003–3016.

94. Everest-Dass AV, Jin D, Thaysen-Andersen M, Nevalainen H, Kolarich D, Packer NH. **2012**. Comparative structural analysis of the glycosylation of salivary and buccal cell proteins: Innate protection against infection by *Candida albicans*. *Glycobiology* 22: 1465–1479.

95. Fazio F, Bryan MC, Blixt O, Paulson JC, Wong CH. **2002**. Synthesis of sugar arrays in microtiter plate. *J Am Chem Soc*. 124: 14397–14402.

96. Feavers IM. **2001**. Meningococcal Vaccines and Vaccine Developments. In *Meningococcal Vaccines*; Pollard AJ, Maiden MC, Ed. Humana Press Inc: Totowa, NJ; pp 1–22.

97. Feinstein A, Munn EA. **1969**. Conformation of the free and antigen-bound IgM antibody molecules. *Nature* 224: 1307–1309.

98. Feizi T, Chai W. **2004**. Oligosaccharide microarrays to decipher the glyco code. *Nat Rev Mol Cell Biol*. 5: 582–588.

99. Feizi T, Fazio F, Chai W, Wong CH. **2003**. Carbohydrate microarrays - a new set of technologies at the frontiers of glycomics. *Curr Opin Struct Biol*. 13: 637–645.

100. Fisher HF, Singh N. **1995**. Calorimetric methods for interpreting protein–ligand interactions. *Methods Enzymol*. 259: 194–221.

101. Flajnik MF, Deschacht N, Muyldermans S. **2011**. A Case of convergence: Why did a simple alternative to canonical antibodies arise in sharks and camels? *PLoS Biol.* 9: e1001120.

102. Flaster MS, Schlcy C, Zipser B. **1983**. Generating monoclonal antibodies against excised gel bands to correlate immunocytochemical and biochemical data. *Brain Res* 277: 196–199.

103. Frasch CE, Preziosi M-P, LaForce FM. **2012**. Development of a group A meningococcal conjugate vaccine, MenAfriVac(TM). *Hum Vaccin Immunother*. 8: 715–724.

104. Freire-de-Lima L, Fonseca LM, Oeltmann T, Mendonça-Previato L, Previato JO. **2015**. The trans-sialidase, the major Trypanosoma cruzi virulence factor: Three decades of studies. *Glycobiology* 25: 1142–1149.

105. Frenken LG, van der Linden RH, Hermans PW, Bos JW, Ruuls RC, de Geus B, Verrips CT. **2000**. Isolation of antigen-specific llama VHH antibody fragments and their high-level secretion by *Saccharomyces cerevisiae*. *J Biotechnol.*78: 11–21.

106. Frenzel A, Hust M, Schirrmann T. **2013**. Expression of recombinant antibodies. *Front Immunol.* 4: 217.

107. Fukui S, Feizi T, Galustian C, Lawson AM, Chai W. **2002**. Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate–protein interactions. *Nat Biotechnol*. 20: 1011–1017.

108. Gagneux SP, Hodgson A, Smith TA, Wirth T, Ehrhard I, Morelli G, Genton B, Binka FN, Achtman M, Pluschke G. **2002**. Prospective study of a serogroup X *Neisseria meningitidis* outbreak in northern Ghana. *J Infect Dis*. 185: 618–626.

109. Galli-Stampino L, Meinjohanns E, Frische K, Meldal M, Jensen T, Werdelin O, Mouritsen S. **1997**. T-cell recognition of tumor-associated carbohydrates: The nature of the glycan moiety plays a decisive role in determining glycopeptide immunogenicity. *Cancer Res*. 57: 3214–3222.

110. Garces F, Sok D, Kong L, McBride R, Kim HJ, Saye-Francisco KF, Julien JP, Hua Y, Cupo A, Moore JP, Paulson JC, Ward AB, Burton DR, Wilson IA. **2014**. Structural evolution of glycan recognition by a family of potent HIV antibodies. *Cell* 159: 69–79.

111. Garner OB, Baum LG. **2008**. Galectin-glycan lattices regulate cell-surface glycoprotein organization and signalling. *Biochem Soc Trans*. 36: 1472–1477.

112. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. **2005**. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Walker JM, Ed. Humana Press: Totowa, NJ; pp 571–607.

113. Geisse, S.; Fux, C. **2009**. Recombinant protein production by transient gene transfer into mammalian cells. *Methods Enzymol.* 463: 223–238.

114. Geissner A, Seeberger PH. **2016**. Glycan arrays: From basic biochemical research to bioanalytical and biomedical applications. *Annu Rev Anal Chem*. 9: 223–47.

115. Genbacev OD, Prakobphol A, Foulk RA, Krtolica AR, Ilic D, Singer MS, Yang ZQ, Kiessling LL, Rosen SD, Fisher SJ. **2003**. Trophoblast L-selectin-mediated adhesion at the maternal-fetal interface. *Science* 299: 405–408.

116. Ghaderi D, Springer SA, Ma F, Cohen M, Secrest P, Taylor RE, Varki A, Gagneux P. **2011**. Sexual selection by female immunity against paternal antigens can fix loss of function alleles. *Proc Natl Acad Sci USA* 108: 17743–17748.

117. Ghai R, Falconer RJ, Collins BM. **2012**. Applications of isothermal titration calorimetry in pure and applied research - survey of the literature from 2010. *J Mol Recognit.* 25: 32–52.

118. Goldstein G. **1987**. Monoclonal antibody specificity: Orthoclone OKT3 T-cell blocker. *Nephron*. 46 suppl 1: 5–11.

119. Gomord V, Chamberlain P, Jefferis R, Faye L. **2005**. Biopharmaceutical production in plants: problems, solutions and opportunities. *Trends Biotechnol.* 23: 559–565.

120. Graham FL, Smiley J, Russell WC, Nairn R.**1977**. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol*. 36: 59–74.

121. Green LL, Hardy MC, Maynard-Currie CE, Tsuda H, Louie DM, Mendez MJ, Abderrahim H, Noguchi M, Smith DH, Zeng Y, David NE, Sasai H, Garza D, Brenner DG, Hales JF, McGuinness RP, Capon DJ, Klapholz S, Jakobovits A. **1994**. Antigen-specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs. *Nat Genet.* 7: 13–21.

122. Grosse-Hovest L, Müller S, Minoia R, Wolf E, Zakhartchenko V, Wenigerkind H, Lassnig C, Besenfelder U, Müller M, Lytton SD, Jung G, Brem G. **2004**. Cloned transgenic farm animals produce a bispecific antibody for T cell-mediated tumor cell killing. *Proc Natl Acad Sci USA* 101: 6858–6863.

123. Gupta R, Jung E, Brunak S. **2004**. Prediction of N-glycosylation sites in human proteins. http://www.cbs.dtu.dk/services/NetNGlyc/

124. Haji-Ghassemi O, Blackler RJ, Martin Young N, Evans SV. **2015**. Antibody recognition of carbohydrate epitopes. *Glycobiology* 25: 920–952.

125. Haliloğlu G, Topaloglu H. **2004**. Glycosylation defects in muscular dystrophies. *Curr Opin Neurol*. 17: 521–527.

126. Hamilton SR, Bobrowicz P, Bobrowicz B, Davidson RC, Li H, Mitchell T, Nett JH, Rausch S, Stadheim TA, Wischnewski H, Wildt S, Gerngross TU. **2003**. Production of complex human glycoproteins in yeast. *Science* 301: 1244–1246.

127. Hammond C, Braakman I, Helenius A. **1994**. Role of *N*-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proc Natl Acad Sci USA* 91: 913–917.

128. Harale KR, Dumare NB, Singh D, Misra AK, Chhikara MK. **2015**. Synthesis of a tetrasaccharide and its glycoconjugate corresponding to the capsular polysaccharide of *Neisseria meningitidis* serogroup X and its immunochemical studies. *RSC Adv*. 5: 41332–41340.

129. Hardivillé S, Hart GW. **2016**. Nutrient regulation of gene expression by *O*-GlcNAcylation of chromatin. *Curr Opin Chem Biol*. 33: 88–94.

130. Hare NJ, Lee LY, Loke I, Britton WJ, Saunders BM, Thaysen-Andersen M. **2017**. *Mycobacterium tuberculosis* infection manipulates the glycosylation machinery and the *N*-glycoproteome of human macrophages and their microparticles. *J Proteome Res*. 16: 247–263.

131. Hariharan P, Guan L. **2014**. Insights into the inhibitory mechanisms of the regulatory protein IIA(Glc) on melibiose permease activity. *J Biol Chem*. 289: 33012–33019.

132. Harrison LH, Trotter CL, Ramsay ME. **2009**. Global epidemiology of meningococcal disease. *Vaccine* 27: B51–B63.

133. Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O. **2011**. Cross talk between *O*-GlcNAcylation and phosphorylation: Roles in signaling, transcription, and chronic disease. *Annu Rev Biochem*. 80: 825–858.

134. Hart GW. **1997**. Dynamic *O*-linked glycosylation of nuclear and cytoskeletal proteins. *Annu Rev Biochem*. 66: 315–335.

135. Hartley JL. **2012**. Why Proteins in Mammalian Cells?. In *Protein Expression in Mammalian Cells*; Hartley J, Ed. Humana Press: Totowa, NJ; pp 1–12.

136. Hascall V, Esko JD. **2009**. Hyaluronan. In *Essentials of Glycobiology*; Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME, Ed. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY; pp 219–228.

137. Hase S, Okawa K, Ikenaka T. **1982**. Identification of the trimannosyl chitobiose structure in sugar moieties of Japanese quail ovomucoid. *J Biochem*. 91: 735–737.

138. Hashii N, Kawasaki N, Itoh S, Nakajima Y, Kawanishi T, Yamaguchi T. **2009**. Alteration of N-glycosylation in the kidney in a mouse model of systemic lupus erythematosus: Relative quantification of N-glycans using an isotopetagging method. *Immunology* 126: 336–345.

139. Hauri HP, Appenzeller C, Kuhn F, Nufer O. **2000**. Lectins and traffic in the secretory pathway. *FEES Lett*. 476: 32–37.

140. Hebert DN, Foellmer B, Helenius A. **1995**. Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. *Cell* 81: 425–433.

141. Helenius A, Aebi M. **2001**. Intracellular functions of N-linked glycans. *Science* 291: 2364–2369.

142. Helenius A, Trombetta ES, Hebert DN, Simons JF. **1997**. Calnexin, calreticulin and the folding of glycoproteins. *Trends Cell Biol*. 7: 193–200.

143. Helmerhorst E, Chandler DJ, Nussio M, Mamotte CD. **2012**. Real-time and label-free bio-sensing of molecular interactions by Surface Plasmon Resonance: A laboratory medicine perspective. *Clin Biochem Rev*. 33: 161–173.

144. Hiramoto E, Tsutsumi A, Suzuki R, Matsuoka S, Arai S, Kikkawa M, Miyazaki T. **2018**. The IgM pentamer is an asymmetric pentagon with an open groove that binds the AIM protein. *Sci Adv*. 4: eaau1199.

145. Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA. **1983**. Abinary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti- plasmid. *Nature* 303: 179–180.

146. Horwitz AH, Chang CP, Better M, Hellstrom KE, Robinson RR. **1988**. Secretion of functional antibody and Fab fragment from yeast cells. *Proc Natl Acad Sci USA* 85: 8678–8682.

147. Houseman BT, Mrksich M. **2002**. Carbohydrate arrays for the evaluation of protein binding and enzymatic modification. *Chem Biol*. 9: 443–454.

148. Howard DR, Natowicz M, Baenziger JU. **1982**. Structural studies of the endoglycosidase H-resistant oligosaccharides present on human beta-glucuronidase. *J Biol Chem*. 257: 10861–10868.

149. Hu S, Shively L, Raubitschek A, Sherman M, Williams LE, Wong JY, Shively JE, Wu AM. **1996**. Minibody: a novel engineered anti-carcinoembryonic antigen antibody fragment (single-chain Fv-CH3) which exhibits rapid, high-level targeting of xenografts. *Cancer Res.* 56: 3055–3061.

150. Hudson PJ, Kortt AA. **1999**. High avidity scFv multimers; diabodies and triabodies. *J Immunol Methods* 231: 177–189.

151. Hust M, Jostock T, Menzel C, Voedisch B, Mohr A, Brenneis M, Kirsch M, Meier D, Dübel S. **2007**. Single chain Fab (scFab) fragment. *BMC Biotechnol.* 7: 14.

152. Inatani M, Irie F, Plump AS, Tessier-Lavigne M, Yamaguchi Y. **2003**. Mammalian brain morphogenesis and midline axon guidance require heparan sulfate. *Science* 302: 1044–1046.

153. Inoue Y, Ohta T, Tada H, Iwasa S, Udaka S, Yamagata H. **1997**. Efficient production of a functional mouse/human chimeric Fab against human urokinase-type plasminogen activator by *Bacillus brevis*. *Appl Microbiol Biotechnol.* 48: 487–492.

154. Iozzo RV. **2005**. Basement membrane proteoglycans: From cellar to ceiling. *Nat Rev Mol Cell Biol*. 6: 646–656.

155. Ito M, Sugihara K, Asaka T, Toyama T, Yoshihara T, Furuichi K, Wada T, Asano M. **2012**. Glycoprotein hyposialylation gives rise to a nephrotic-like syndrome that is prevented by sialic acid administration in GNE V572L point-mutant mice. *PLoS One* 7: e29873.

156. Jackson LR, Trudel LJ, Fox JG, Lipman NS. **1996**. Evaluation of hollow fiber bioreactors as an alternative to murine ascites production for small scale monoclonal antibody production. *J Immunol Methods* 189: 217–231.

157. Jakob CA, Burda P, Roth J, Aebi M. **1998**. Degradation of misfolded endoplasmic reticulum glycoproteins in *Saccharomyces cerevisiae* is determined by a specific oligosaccharide structure. *J Cell Biol*. 142: 1223–1233.

158. Jennings HJ. **1990**. Capsular polysaccharides as vaccine candidates. *Curr Top Microbiol Immunol.* 150: 97–127.

159. Jez J, Castilho A, Grass J, Vorauer-Uhl K, Sterovsky T, Altmann F, Steinkellner H. **2013**. Expression of functionally active sialylated human erythropoietin in plants. *Biotechnol J.* 8: 371–382.

160. Jiang D, Liang J, Noble PW. **2011**. Hyaluronan as an immune regulator in human diseases. *Physiol Rev*. 91: 221–264.

161. Johansson ME, Sjövall H, Hansson GC. **2013**. The gastrointestinal mucus system in health and disease. *Nat Rev Gastroenterol Hepatol*. 10: 352–361.

162. Johnsson B, Löfås S, Lindquist G. **1991**. Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Anal Biochem*. 198: 268–277.

163. Johnston JW, Zaleski A, Allen S, Mootz JM, Armbruster D, Gibson BW, Apicella MA, Munson RSJ. **2007**. Regulation of sialic acid transport and catabolism in Haemophilus influenzae. *Mol Microbiol*. 66: 26–39.

164. Jones PT, Dear PH, Foote J, Neuberger MS, Winter G. **1986**. Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature* 321: 522–525.

165. Joosten V, Gouka RJ, van den Hondel CAMJJ, Verrips CT, Lokman BC. **2005**. Expression and production of llama variable heavy-chain antibody fragments (V(HH)s) by *Aspergillus awamori*. *Appl Microbiol Biotechnol.* 66: 384–392.

166. Joosten V, Lokman C, Van Den Hondel CA, Punt PJ. **2003**. The production of antibody fragments and antibody fusion proteins by yeasts and filamentous fungi. *Microb Cell Fact.* 2: 1.

167. Jordan E, Al-Halabi L, Schirrmann T, Hust M. **2009**. Antibody production by the Gram-positive bacterium *Bacillus megaterium*. *Methods Mol Biol.* 525: 509–516.

168. Jostock T. **2011**. Expression of Antibody in Mammalian Cells. In *Antibody Expression and Production*; Al-Rubeai M, Ed. Springer: Dordrecht, The Netherlands; pp 1–24.

169. Jönsson U, Fägerstam L, Ivarsson B, Johnsson B, Karlsson R, Lundh K, Löfås S, Persson B, Roos H, Rönnberg I, Sjölander S, Stenberg E, Ståhlberg R, Urbaniczky C, östlin H, Malmqvist M. **1991**. Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. *Biotechniques* 11: 620–627.

170. Jönsson U. **1992**. Real-time biospecific interaction analysis. *Biosensors '92 Proceedings*: 260–266.

171. Kalk A, Berendsen HJC. **1976**. Proton magnetic relaxation and spin diffusion in proteins. *J Magn Reson*. 24: 343–366.

172. Kalwy S, Rance J, Young R. **2006**. Towards more efficient protein expression. *Mol Biotechnol*. 34: 151–156.

173. Kansas GS. **1996**. Selectins and their ligands: Current concepts and controversies. *Blood* 88: 3259–3287.

174. Karlsson R, Michaelsson A, Mattsson L. **1991**. Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system. *J Immunol Methods*. 145: 229–40.

175. Kaveri SV, Silverman GJ, Bayry J. **2012**. Natural IgM in immune equilibrium and harnessing their therapeutic potential. *J Immunol*. 188: 939–945.

176. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. **2015**. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 10: 845–858.

177. Khatua B, Bhattacharya K, Mandal C. **2012**. Sialoglycoproteins adsorbed by Pseudomonas aeruginosa facilitate their survival by impeding neutrophil extracellular trap through siglec-9. *J Leukoc Biol*. 91: 641–655.

178. Kiessling LL, Cairo CW. **2002**. Hitting the sweet spot. *Nat Biotechnol*. 20: 234–235.

179. Klatt S, Rohe M, Alagesan K, Kolarich D, Konthur Z, Hartl D. **2013**. Production of glycosylated soluble amyloid precursor protein alpha (sAPPalpha) in *Leishmania tarentolae*. *J Proteome Res.* 12: 396–403.

180. Klausner RD, Sitia R. **1990**. Protein degradation in the endoplasmic reticulum. *Cell* 62: 611–614.

181. Ko KS., Jaipuri FA, Pohl NL. **2005**. Fluorous-based carbohydrate microarrays. *J Am Chem Soc*. 127: 13162–13163.

182. Kobata A. **1992**. Structures and functions of the sugar chains of glycoproteins. *Eur J Biochem*. 209: 483–501.

183. Koch BE, Stougaard J, Spaink HP. **2015**. Keeping track of the growing number of biological functions of chitin and its interaction partners in biomedical research. *Glycobiology* 25: 469–482.

184. Kodati B, Dorbha S, Kunaparaju RK. **2016**. Heterologus protein expression in different host systems. *J Chem Pharm Res.* 8: 1068–1074.

185. Köhler G, Milstein C. **1975**. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256: 495–497.

186. Koshland DE. **1958**. Application of a theory of enzyme specificity to protein synthesis. *Proc Natl Acad Sci USA* 44: 98–104.

187. Koshland DE. **1959**. Enzyme flexibility and enzyme action. *J Cell Comp Physiol* 54: 245–258.

188. Krell T. **2008**. Microcalorimetry: A response to challenges in modern biotechnology. *Microb Biotechnol*. 1: 126–136.

189. Kretzschmar T, Aoustin L, Zingel O, Marangi M, Vonach B, Towbin H, Geiser M. **1996**. High-level expression in insect cells and purification of secreted monomeric single-chain Fv antibodies. *J Immunol Methods* 195: 93–101.

190. Kristiansen PA, Diomandé F, Ba AK, Sanou I, Ouédraogo AS, Ouédraogo R, Sangaré L, Kandolo D, Aké F, Saga IM, Clark TA, Misegades L, Martin SW, Thomas JD, Tiendrebeogo SR, Hassan-King M, Djingarey MH, Messonnier NE, Préziosi MP, Laforce FM, Caugant DA. **2013**. Impact of the serogroup A meningococcal conjugate vaccine, MenAfriVac, on carriage and herd immunity. *Clin Infect Dis*. 56: 354–363.

191. Krüger C, Hu Y, Pan Q, Marcotte H, Hultberg A, Delwar D, van Dalen PJ, Pouwels PH, Leer RJ, Kelly CG, van Dollenweerd C, Ma JK, Hammarström L. **2002**. *In situ* delivery of passive immunity by lactobacilli producing single-chain antibodies. *Nat Biotechnol.* 20: 702–706.

192. Kujau MJ, Hoischen C, Riesenberg D, Gumpert J. **1998**. Expression and secretion of functional miniantibodies McPC603scFvDhlx in cell-wall-less L-form strains of *Proteus mirabilis a*nd *Escherichia coli*: a comparison of the synthesis capacities of L-form strains with an *E. coli* producer strain. *Appl Microbiol Biotechnol.* 49: 51–58.

193. Kuroda D, Shirai H, Kobori M, Nakamura H. **2008**. Structural classification of CDR-H3 revisited: A lesson in antibody modeling. *Proteins* 73: 608–620.

194. Kuroiwa Y, Kasinathan P, Sathiyaseelan T, Jiao J, Matsushita H, Sathiyaseelan J, Wu H, Mellquist J, Hammitt M, Koster J, Kamoda S, Tachibana

K, Ishida I, Robl JM. **2009**. Antigen-specific human polyclonal antibodies from hyperimmunized cattle. *Nat Biotechnol.* 27: 173–181.

195. LaForce FM, Ravenscroft N, Djingarey M, Viviani S. **2009**. Epidemic meningitis due to group A *Neisseria meningitidis* in the African meningitis belt: A persistent problem with an imminent solution. *Vaccine* 27: B13–B19.

196. Lau KS, Dennis JW. **2008**. *N*-Glycans in cancer progression. *Glycobiology* 18: 750–760.

197. Lau KS, Partridge EA, Grigorian A, Silvescu CI, Reinhold VN, Demetriou M, Dennis JW. **2007**. Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell* 129: 123–134.

198. Leavitt S, Freire E. **2001**. Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr Opin Struct Biol*. 11: 560–566.

199. Lee LY, Lin CH, Fanayan S, Packer NH, Thaysen-Andersen M. **2014**. Differential site accessibility mechanistically explains subcellular-specific *N*-glycosylation determinants. *Front Immunol*. 5: 404.

200. Leibiger H, Hansen A, Schoenherr G, Seifert M, Wüstner D, Stigler R, Marx U. **1995**. Glycosylation analysis of a polyreactive human monoclonal IgG antibody derived from a human-mouse heterohybridoma. *Mol Immunol.* 32: 595–602.

201. Leijon H, Kaprio T, Heiskanen A, Satomaa T, Hiltunen JO, Miettinen MM, Arol J, Haglund C. **2017**. *N*-glycomic profiling of pheochromocytomas and paragangliomas separates metastatic and nonmetastatic disease. *J Clin Endocrinol Metab*. 102: 3990–4000.

202. Léonard R, Rendic D, Rabouille C, Wilson IB, Préat T, Altmann F. **2006**. The *Drosophila* fused lobes gene encodes an *N*-acetylglucosaminidase involved in *N*-glycan processing. *J Biol Chem*. 281: 4867–4875.

203. Lepore R, Olimpieri PP, Messih MA, Tramontano A. **2017**. PIGSPro: Prediction of immunoglobulin structures v2. *Nucleic Acids Res*. 45: W17–W23

204. Leteux C, Stoll MS, Childs RA, Chai W, Vorozhaikina M, Feizi T. **1999**. Influence of oligosaccharide presentation on the interactions of carbohydrate sequence-specific antibodies and the selectins. Observations with biotinylated oligosaccharides. *J Immunol Methods*. 227: 109–119.

205. Lewis BA, Hanover JA. **2014**. O-GlcNAc and the epigenetic regulation of gene expression. *J Biol Chem*. 289: 34440–34448.

206. Li E, Tabas I, Kornfeld S. **1978**. The synthesis of complex-type oligosaccharides. I. Structure of the lipid-linked oligosaccharide precursor of the complex-type oligosaccharides of the vesicular stomatitis virus G protein. *J Biol Chem*. 253: 7762–7770.

207. Li H, Xie Y, Liu C, Liu S. **2014**. Physicochemical bases for protein folding, dynamics, and protein-ligand binding. *Sci China Life Sci*. 57: 287–302.

208. Li Y, Wang G, Li N, Wag Y, Zhu Q, Chu H, Wu W, Tan Y, Yu F, Su X-D, Gao N, Xiao J. **2020**. Structural insights into immunoglobulin M. *Science* 367: 1014–1017.

209. Lillehoj EP, Kato K, Lu W, Kim KC. **2013**. Cellular and molecular biology of airway mucins. *Int Rev Cell Mol Biol*. 303: 139–202.

210. Lintner KE, Wu YL, Yang Y, Spencer CH, Hauptmann G, Hebert LA, Atkinson JP, Yu CY. **2016**. Early components of the complement classical activation pathway in human systemic autoimmune diseases. *Front Immunol*. 7: 36.

211. Liu J, Wei D, Qian F, Zhou Y, Wang J, Ma Y, Han Z. **2003**. pPIC9- Fc: a vector system for the production of single-chain Fv-Fc fusions in *Pichia pastoris* as detection reagents *in vitro*. *J Biochem.* 134: 911–917.

212. Loke I, Kolarich D, Packer NH, Thaysen-Andersen M. **2016**. Emerging roles of protein mannosylation in inflammation and infection. *Mol Aspects Med*. 51: 31–55.

213. Loke I, Østergaard O, Heegaard NHH, Packer NH, Thaysen-Andersen M. **2017**. Paucimannose-rich N-glycosylation of spatiotemporally regulated human neutrophil elastase modulates its immune functions. *Mol Cell Proteomics* 16: 1507–1527.

214. Loomes KM, Senior HE, West PM, Roberton AM. **1999**. Functional protective role for mucin glycosylated repetitive domains. *Eur J Biochem*. 266: 105–111.

215. Lüllau E, Heyse S, Vogel H, Marison I, von Stockar U, Kraehenbuhl JP, Corthésy B. **1996**. Antigen binding properties of purified immunoglobulin A and reconstituted secretory immunoglobulin A antibodies. *J Biol Chem.* 271: 16300–16309.

216. Lundström I. **1994**. Real-time biospecific interaction analysis. *Biosens. Bioelectron*. 9: 725–736.

217. Lutz AJ, Li P, Estrada JL, Sidner RA, Chihara RK, Downey SM, Burlak C, Wang ZY, Reyes LM, Ivary B, Yin F, Blankenship RL, Paris LL, Tector AJ. **2013**. Double knockout pigs deficient in N-glycolylneuraminic acid and Galactose alpha-1,3-Galactose reduce the humoral barrier to xenotransplantation. *Xenotransplantation* 20: 27–35.

218. Mahassni SH, Klapper DG, Hiskey RG. **2009**. Purification of a murine IgM monoclonal antibody. *Hybridoma (Larchmt)* 28: 189–197.

219. Malmqvist M. **1993**. Surface plasmon resonance for detection and measurement of antibody-antigen affinity and kinetics. *Curr Opin Immunol*. 5: 282–286.

220. Marcotte H, Kõll-Klais P, Hultberg A, Zhao Y, Gmür R, Mändar R, Mikelsaar M, Hammarström L. **2006**. Expression of single-chain antibody against RgpA protease of *Porphyromonas gingivalis* in *Lactobacillus*. *J Appl Microbiol.* 100: 256–263.

221. Martineau P, Jones P, Winter G. **1998**. Expression of an antibody fragment at high levels in the bacterial cytoplasm. *J Mol Biol.* 280: 117–127.

222. Marx U. **1998**. Membrane-based cell culture technologies: A scientifically and economically satisfactory alternative to malignant ascites production for monoclonal antibodies. *Res Immunol.* 149: 557–559.

223. Mayer M, Meyer B. **1999**. Characterization of ligand binding by Saturation Transfer Difference NMR Spectroscopy. *Angew Chem Int Ed*. 38: 1784–1788.

224. Mayer M, Meyer B. **2001**. Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. *J Am Chem Soc*. 123: 6108–6117.

225. McCormick AA, Reinl SJ, Cameron TI, Vojdani F, Fronefield M, Levy R, Tusé D. **2003**. Individualized human scFv vaccines produced in plants: humoral anti-idiotype responses in vaccinated mice confirm relevance to the tumour Ig. *J Immunol Methods* 278: 95–104.

226. McDonnell JM. **2001**. Surface plasmon resonance: towards an understanding of the mechanisms of biological molecular recognition. *Curr Opin Chem Biol*. 5: 572–577.

227. McEver RP. **1991**. Selectins: Novel receptors that mediate leukocyte adhesion during inflammation. *Thromb Haemost*. 65: 223–228.

228. McFarlane HE, Doring A, Persson S. **2014**. The cell biology of cellulose synthesis. *Annu Rev Plant Biol*. 65: 69–94.

229. McPherson A., Gavira J. A. **2013**. Introduction to protein crystallization. *Acta Crystallogr F Struct Biol Commun*. 70: 2–20.

230. Meissner HC, Welliver RC, Chartrand SA, Law BJ, Weisman LE, Dorkin HL, Rodriguez WJ. **1999**. Immunoprophylaxis with palivizumab, a humanized respiratory syncytial virus monoclonal antibody, for prevention of respiratory syncytial virus infection in high-risk infants: a consensus opinion. *Pediatr Infect Dis J.* 18: 223–231.

231. Mencarelli S, Cavalieri C, Magini A, Tancini B, Basso L, Lemansky P, Hasilik A, Li YT, Chigorno V, Orlacchio A, Emiliani C, Sonnino S. **2005**. Identification of plasma membrane associated mature ß-hexosaminidase A, active towards GM2 ganglioside, in human fibroblasts. *FEBS Lett*. 579: 5501–5506.

232. Mende M, Bordoni V, Tsouka A, Loeffler FF, Delbianco M, Seeberger PH. **2019**. Multivalent glycan arrays. *Faraday Discuss*. 219: 9–32.

233. Meyer B, Peters T. **2003**. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew Chem Int Ed Engl*. 42: 864–890.

234. Michon F, Brisson JR, Dell A, Kasper DL, Jennings HJ. **1988**. Multiantennary group-specific polysaccharide of group B Streptococcus. *Biochemistry* 27: 5341–5351.

235. Micoli F, Romano MR, Tontini M, Cappelletti E, Gavini M, Proietti D, Rondini S, Swennen E, Santini L, Filippini S, Balocchi C, Adamo R, Pluschke G, Norheim

G, Pollard A, Saul A, Rappuoli R, MacLennan CA, Berti F, Costantino P. **2013**. Development of a glycoconjugate vaccine to prevent meningitis in Africa caused by meningococcal serogroup X. *Proc Natl Acad Sci USA.* 110: 19077–19082.

236. Miller JM, Mesaros N, Van Der Wielen M, Baine Y. **2011**. Conjugate meningococcal vaccines development: GSK biologicals experience. *Adv Prev Med*. 846756: 1–17.

237. Milstein C, Galfre G, Secher DS, Springer T. **1979**. Monoclonal antibodies and cell surface antigens. *Cell Biol Int Rep*. 3: 1–16.

238. Möginger U, Grunewald S, Hennig R, Kuo CW, Schirmeister F, Voth H, Rapp E, Khoo KH, Seeberger PH, Simon JC, Kolarich D. **2018**. Alterations of the human skin *N*- and *O*-glycome in basal cell carcinoma and squamous cell carcinoma. *Front Oncol*. 8: 70.

239. Mond JJ, Lees A, Snapper CM. **1995**. T cell-independent antigens type 2. *Annu Rev Immunol*. 13: 655–692.

240. Morea V, Tramontano A, Rustici M, Chothia C, Lesk AM. **1997**. Antibody structure, prediction and redesign. *Biophys Chem*. 68: 9–16.

241. Morea V, Tramontano A, Rustici M, Chothia C, Lesk AM. **1998**. Conformations of the third hypervariable region in the VH domain of immunoglobulins. *J Mol Biol*. 275: 269–294.

242. Morelli L, Cancogni D, Tontini M, Nilo A, Filippini S, Costantino P, Romano MR, Berti F, Adamo R, Lay L. **2014**. Synthesis and immunological evaluation of protein conjugates of *Neisseria meningitidis* X capsular polysaccharide fragments. *Beilstein J Org Chem*. 10: 2367–2376.

243. Moremen KW, Trimble RB, Herscovics A. **1994**. Glycosidases of the asparagine-linked oligosaccharide processing pathway. *Glycobiology* 4: 113–125.

244. Moreth K, Brodbeck R, Babelova A, Gretz N, Spieker T, Zeng-Brouwers J, Pfeilschifter J, Young MF, Schaefer RM, Schaefer L. **2010**. The proteoglycan biglycan regulates expression of the B cell chemoattractant CXCL13 and aggravates murine lupus nephritis. *J Clin Invest*. 120: 4251–4272.

245. Moriguchi K, Takemoto T, Aoki T, Nakakita S, Natsuka S, Hase S. **2007**. Free oligosaccharides with Lewis x structure expressed in the segmentation period of zebrafish embryo. *J Biochem*. 142: 213–227.

246. Muchmore E, Varki A. **1987**. Inactivation of influenza C esterase decreases infectivity without loss of binding; a probe for 9-Oacetylated sialic acids. *Science* 236: 1293–1295.

247. Müller R, Gräwert MA, Kern T, Madl T, Peschek J, Sattler M, Groll M, Buchner J. **2013**. High-resolution structures of the IgM Fc domains reveal principles of its hexamer formation. *Proc Natl Acad Sci USA*110: 10183–10188.

248. Muthana SM, Gildersleeve JC. **2014**. Glycan microarrays: powerful tools for biomarker discovery. *Cancer Biomark*. 14: 29–41.

249. Muthana SM, Xia L, Campbell CT, Zhang Y, Gildersleeve JC. **2015**. Competition between serum IgG, IgM, and IgA anti-glycan antibodies. *PLoS One* 10: e0119298.

250. Mutonga DM, Pimentel G, Muindi J, Nzioka C, Mutiso J, Klena JD, Morcos M, Ogaro T, Materu S, Tetteh C, Messonnier NE, Breiman RF, Feikin DR. **2009**. Epidemiology and risk factors for serogroup X meningococcal meningitis during an outbreak in western Kenya, 2005–2006. *Am J Trop Med Hyg*. 80: 619–624.

251. Mutsaers JH, Van Halbeek H, Vliegenthart JF, Tager JM, Reuser AJ, Kroos M, Galjaard H. **1987**. Determination of the structure of the carbohydrate chains of acid alpha-glucosidase from human placenta. *Biochim Biophys Acta* 911: 244–251.

252. Myszka DG. **1997**. Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors. *Curr Opin Biotechnol*. 8: 50–57.

253. National Research Council **1999**. Monoclonal Antibody Production. The National Academies Press: Washington, DC.

254. Nauseef WM, Borregaard N. **2014**. Neutrophils at work. *Nat Immunol*. 15: 602–611.

255. Nelson AL. **2010**. Antibody fragments: hope and hype. *MAbs* 2: 77–83.

256. Niimi T. **2012**. Recombinant protein production in the eukaryotic protozoan parasite *Leishmania tarentolae*: a review. *Methods Mol Biol.* 824: 307–315.

257. North B, Lehmann A, Dunbrack RL. **2011**. A new clustering of antibody CDR loop conformations. *J Mol Biol*. 406: 228–256.

258. Nyyssönen E, Penttilä M, Harkki A, Saloheimo A, Knowles JK, Keräne S. **1993**. Efficient production of antibody fragments by the filamentous fungus *Trichoderma reesei*. *Biotechnology (NY)* 11: 591–595.

259. Ojima-Kato T, Fukui K, Yamamoto H, Hashimura D, Miyake S, Hirakawa Y, Yamasaki T, Kojima T, Nakano H. **2016**. 'Zipbody' leucine zipper-fused Fab in *E. coli in vitro* and *in vivo* expression systems. *Protein Eng Des Sel*. 29:149-157.

260. Okazaki F, Aoki J, Tabuchi S, Tanaka T, Ogino C, Kondo A. **2012**. Efficient heterologous expression and secretion in *Aspergillus oryzae* of a llama variable heavy-chain antibody fragment V(HH) against EGFR. *Appl Microbiol Biotechnol.* 96: 81–88.

261. Pace D. **2013**. Glycoconjugate vaccines. *Expert Opin Biol Ther*. 13: 11–33.

262. Pant N, Hultberg A, Zhao Y, Svensson L, Pan-Hammarstrom Q, Johansen K, Pouwels PH, Ruggeri FM, Hermans P, Frenken L, Boren T, Marcotte H, Hammarstrom L. **2006**. Lactobacilli expressing variable domain of llama heavy-chain antibody fragments (lactobodies) confer protection against rotavirus-induced diarrhoea. *J Infect Dis.* 194: 1580–1588.

263. Park S, Shin I. **2002**. Fabrication of carbohydrate chips for studying protein-carbohydrate interactions. *Angew Chem Int Ed Engl*. 41: 3180–3182.

264. Parker BL, Thaysen-Andersen M, Solis N, Scott NE, Larsen MR, Graham ME, Packer NH, Cordwell SJ. **2013**. Site-specific glycan-peptide analysis for determination of *N*-glycoproteome heterogeneity. *J Proteome Res*. 12: 5791–5800.

265. Parodi AJ. **2000**. Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation. *Biochem J*. 348:1–13.

266. Parsons NJ, Patel PV, Tan EL, Andrade JR, Nairn CA, Goldner M, Cole JA, Smith H. **1988**. Cytidine 5'-monophospho-N-acetyl neuraminic acid and a low molecular weight factor from human blood cells induce lipopolysaccharide alteration in gonococci when conferring resistance to killing by human serum. *Microb Pathog*. 5: 303–309.

267. Paschinger K, Rendić D, Wilson IB. **2009**. Revealing the anti-HRP epitope in Drosophila and Caenorhabditis. *Glycoconj J*. 26: 385–395.

268. Paszek MJ, DuFort CC, Rossier O, Bainer R, Mouw JK, Godula K, Hudak JE, Lakins JN, Wijekoon AC, Cassereau L, Rubashkin MG, Magbanua MJ, Thorn KS, Davidson MW, Rugo HS, Park JW, Hammer DA, Giannone G, Bertozzi CR, Weaver VM. **2014**. The cancer glycocalyx mechanically primes integrin-mediated growth and survival. *Nature* 511: 319–325.

269. Patel T, Bruce J, Merry A, Bigge C, Wormald M, Jaques A, Parekh R. **1993**. Use of hydrazine to release in intact and unreduced form both *N*- and *O*-linked oligosaccharides from glycoproteins. *Biochemistry* 32: 679–693.

270. Paulson JC, Blixt O, Collins BE. **2006**. Sweet spots in functional glycomics. *Nat Chem Biol*. 2: 238–48.

271. Peipp M, Lammerts van Bueren JJ, Schneider-Merck T, Bleeker WW, Dechant M, Beyer T, Repp R, van Berkel PH, Vink T, van de Winkel JG, Parren PW, Valerius T. **2008**. Antibody fucosylation differentially impacts cytotoxicity mediated by NK and PMN effector cells. *Blood* 112: 2390–2399.

272. Pelaseyed T, Bergstrom JH, Gustafsson JK, Ermund A, Birchenough GM, Schutte A, van der Post S, Svensson F, Rodriguez-Pineiro AM, Nystrom EE, Wising C, Johansson ME, Hansson GC. **2014**. The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunol Rev*. 260: 8–20.

273. Perozzo R, Folkers G, Scapozza L.**2004**. Thermodynamics of protein–ligand interactions: history, presence, and future aspects. *J Recept Signal Transduct Res*. 24: 1–52.

274. Pierce MM. **1999**. Isothermal titration calorimetry of protein-protein interactions. *Methods* 19: 213–221.

275. Plante OJ, Palmacci ER, Seeberger PH. **2001**. Automated solid-phase synthesis of oligosaccharides. *Science* 291: 1523–1527.

276. Plotkin S. **2014**. History of vaccination. *Proc Natl Acad Sci USA* 111: 12283–12287.

277. Plückthun A, Pack P. **1997**. New protein engineering approaches to multivalent and bispecific antibody fragments. *Immunotechnology* 3: 83–105.

278. Porath J. **1992**. Immobilized Metal Ion Affinity Chromatography. *Protein Expr Purif.* 3: 263–281.

279. Potgieter TI, Kersey SD, Mallem MR, Nylen AC, d'Anjou M. **2010**. Antibody expression kinetics in glycoengineered *Pichia pastoris*. *Biotechnol Bioeng.* 106*:* 918–927.

280. Price PJ. **1985**. Hybridoma Technology. In *Advances in Cell Culture*; Maramorosch K, Ed. Academic Press Inc: Orlando, FL; pp 157–177.

281. Rajarathnam K, Rösgen J. **2014**. Isothermal titration calorimetry of membrane proteins - Progress and challenges. *Biochim Biophys Acta* 1838: 69–77.

282. Ramsay ME, Andrews NJ, Trotter CL, Kaczmarski EB, Miller E. **2003**. Herd immunity from meningococcal serogroup C conjugate vaccination in England: Database analysis. *BMJ* 326: 365–366.

283. Reczko M, Martin ACR, Bohr H, Suhai S. **1995**. Prediction of hypervariable CDR-H3 loop structures in antibodies. *Protein Eng*. 8: 389–395.

284. Rich RL, Myszka DG. **2000**. Advances in surface plasmon resonance biosensor analysis. *Curr Opin Biotechnol*. 11: 54–61.

285. Rippmann JF, Klein M, Hoischen C, Brocks B, Rettig WJ, Gumpert J, Pfizenmaier K, Mattes R, Moosmayer D. **1998**. Procaryotic expression of single-chain variable-fragment (scFv) antibodies: secretion in L-form cells of *Proteus mirabilis* leads to active product and overcomes the limitations of periplasmic expression in *Escherichia coli. Appl Environ Microbiol.* 64: 4862–4869.

286. Roggenbuck D, Marx U, Kiessig ST, Schoenherr G, Jahn S, Porstmann T. **1994**. Purification and immunochemical characterization of a natural human polyreactive monoclonal IgM antibody. *J Immunol Methods* 167: 207–218.

287. Saeed AF, Wang R, Ling S, Wang S. **2017**. Antibody engineering for pursuing a healthier future. *Front Microbiol*. 8: 495.

288. Sarkar M, Leventis PA, Silvescu CI, Reinhold VN, Schachter H, Boulianne GL. **2006**. Null mutations in *Drosophila N*-acetylglucosaminyltransferase I produce defects in locomotion and a reduced life span. *J Biol Chem*. 281, 12776–12785.

289. Satoh T, Suzuki K, Yamaguchi T, Kato K. **2014**. Structural Basis for Disparate Sugar-Binding Specificities in the Homologous Cargo Receptors ERGIC-53 and VIP36. *PLoS One* 9: e87963.

290. Satomaa T, Heiskanen A, Mikkola M, Olsson C, Blomqvist M, Tiittanen M, Jaatinen T, Aitio O, Olonen A, Helin J, Hiltunen J, Natunen J, Tuuri T, Otonkoski T, Saarinen J, Laine J. **2009**. The N-glycome of human embryonic stem cells. *BMC Cell Biol*. 10: 42.

291. Schachter H, Boulianne G. **2011**. Life is sweet! A novel role for *N*-glycans in *Drosophila* lifespan. *Fly* 5: 18–24.

292. Schachter H. **2009**. Paucimannose N-glycans in *Caenorhabditis elegans* and *Drosophila melanogaster*. *Carbohydr Res*. 344: 1391–1396.

293. Schähs M, Strasser R, Stadlmann J, Kunert R, Rademacher T, Steinkellner H. **2007**. Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern. *Plant Biotechnol J.* 5: 657–663.

294. Schauer R. **1985**. Sialic acids and their role as biological masks. *Trends Biochem Sci*. 10: 357–360.

295. Schiller B, Hykollari A, Yan S, Paschinger K, Wilson IB. **2012**. Complicated *N*-linked glycans in simple organisms. *Biol Chem*. 393: 661–673.

296. Schmiedl A, Breitling F, Winter CH, Queitsch I, Dübel S. **2000**. Effects of unpaired cysteines on yield, solubility and activity of different recombinant antibody constructs expressed in E. coli. *J Immunol Methods* 242: 101–114.

297. Schreier PH, Quester S, Bothwell A. **1986**. Allotypic differences in murine mu genes. *Nucleic Acids Res*. 14: 2381–2389.

298. Schwartz NB, Domowicz MS. **2014**. Chemistry and function of glycosaminoglycans in the nervous system. *Adv Neurobiol*. 9: 89–115.

299. Serna S, Yan S, Martin-Lomas M, Wilson IBH, Reichardt NC. **2011**. Fucosyltransferases as synthetic tools: glycan array based substrate selection and core fucosylation of synthetic *N*-glycans. *J Am Chem Soc.* 133: 16495–16502.

300. Shah P, Wang X, Yang W, Toghi Eshghi S, Sun S, Hoti N, Chen L, Yang S, Pasay J, Rubin A, Zhang H. **2015**. Integrated proteomic and glycoproteomic analyses of prostate cancer cells reveal glycoprotein alteration in protein abundance and glycosylation. *Mol Cell Proteomics* 14: 2753–2763.

301. Sharp TH, Boyle AL., Diebolder CA, Kros A, Koster AJ, Gros P. **2019**. Insights into IgM-mediated complement activation based on in situ structures of IgM-C1-C4b. *Proc Natl Acad Sci USA*. 116: 11900–11905.

302. Shi H, Tan J, Schachter H. **2006**. *N*-glycans are involved in the response of *Caenorhabditis elegans* to bacterial pathogens. *Methods Enzymol*. 417: 359–389.

303. Shiroza T, Shinozaki-Kuwahara N, Hayakawa M, Shibata Y, Hashizume T, Fukushima K, Udaka S, Abiko Y. **2003**. Production of a single-chain variable fraction capable of inhibiting the *Streptococcus mutans* glucosyltransferase in *Bacillus brevis*: construction of a chimeric shuttle plasmid secreting its gene product. *Biochim Biophys Acta* 1626: 57–64.

304. Skerra A, Plückthun A. **1988**. Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli*. *Science* 240: 1038–1041.

305. Smith SL. **1996**. Ten years of Orthoclone OKT3 (muromonab-CD3): a review. *J Transpl Coord.* 6: 109–119.

306. Spearman M, Dionne B, Butler M. **2011**. The Role of Glycosylation in Therapeutic Antibodies. In *Antibody Expression and Production*; Al-Rubeai M, Ed. Springer: Dordrecht, The Netherlands; pp 251–292.

307. Spertini O, Luscinskas FW, Kansas GS, Munro JM, Griffin JD, Gimbrone MAJ, Tedder TF. **1991**. Leukocyte adhesion molecule-1 (LAM-1, L-selectin) interacts with an inducible endothelial cell ligand to support leukocyte adhesion. *J Immunol*. 147: 2565–2573.

308. Stanley P, Taniguchi N, Aebi M. **2017**. N-Glycans. In Essentials of Glycobiology; Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH, Ed. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.

309. Stephens DS. **2009**. Biology and pathogenesis of the evolutionarily successful, obligate human bacterium *Neisseria meningitides*. *Vaccine* 27: 71–78.

310. Storry JR, Olsson ML. **2009**. The ABO blood group system revisited: a review and update. *Immunohematology* 25: 48–59.

311. Strasser R, Stadlmann J, Schähs M, Stiegler G, Quendler H, Mach L, Glössl J, Weterings K, Pabst M, Steinkellner H. **2008**. Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like *N*-glycan structure. *Plant Biotechnol J*. 6: 392–402.

312. Sun YS, Landry JP, Zhu XD. **2017**. Evaluation of Kinetics Using Label-Free Optical Biosensors. *Instrum Sci Technol*. 45: 486–505.

313. Takasaki S, Mizuochi T, Kobata A. **1982**. Hydrazinolysis of asparagine-linked sugar chains to produce free oligosaccharides. *Methods Enzymol*. 83: 263–268.

314. Tang Z, Feng M, Gao W, Phung Y, Chen W, Chaudhary A, St Croix B, Qian M, Dimitrov DS, Ho M. **2013**. A human single-domain antibody elicits potent antitumor activity by targeting an epitope in mesothelin close to the cancer cell surface. *Mol Cancer Ther*. 12: 416–426.

315. Tannous A, Pisoni GB, Hebert DN, Molinari M. **2015**. N-linked sugar-regulated protein folding and quality control in the ER. *Semin Cell Dev Biol*. 41: 79–89.

316. Tarentino AL, Plummer TH. **1994**. Enzymatic deglycosylation of asparagine-linked glycans: purification, properties, and specificity of oligosaccharide-cleaving enzymes from *Flavobacterium meningosepticum*. *Methods Enzymol*. 230: 44–57.

317. Taylor KR, Trowbridge JM, Rudisill JA, Termeer CC, Simon JC, Gallo RL. **2004**. Hyaluronan fragments stimulate endothelial recognition of injury through TLR4. *J Biol Chem*. 279: 17079–17084.

318. Taylor ME, Drickamer K. **2009**. Structural insights into what glycan arrays tell us about how glycan-binding proteins interact with their ligands. *Glycobiology* 19: 1155–62.

319. Tecle E, Gagneux P. **2015**. Sugar-coated sperm: Unraveling the functions of the mammalian sperm glycocalyx. *Mol Reprod Dev*. 82: 635–650.

320. Tedder TF, Steeber DA, Chen A, Engel P. **1995**. The selectins: Vascular adhesion molecules. *FASEB J*. 9: 866–873.

321. Thaysen-Andersen M, Venkatakrishnan V, Loke I, Laurini C, Diestel S, Parker BL, Packer NH. **2015**. Human neutrophils secrete bioactive paucimannosidic proteins from azurophilic granules into pathogen-infected sputum. *J Biol Chem*. 290: 8789–8802.

322. Tjondro HC, Loke I, Chatterjee S, Thaysen-Andersen M. **2019**. Human protein paucimannosylation: cues from the eukaryotic kingdoms. *Biol Rev*. 94: 2068–2100.

323. Tokunaga F, Brostrom C, Koide T, Arvan P. **2000**. Endoplasmic reticulum (ER) - associated degradation of misfolded N-linked glycoproteins is suppressed upon inhibition of ER mannosidase I. *J Biol Chem*. 275: 40757–40764.

324. Trinidad JC, Schoepfer R, Burlingame AL, Medzihradszky KF. **2013**. *N*- and *O*-glycosylation in the murine synaptosome. *Mol Cell Proteomics* 12: 3474–3488.

325. Trotter CL, Lingani C, Fernandez K, Cooper LV, Bita A, Tevi-Benissan C, Ronveaux O, Préziosi MP, Stuart JM. **2017**. Impact of MenAfriVac in nine countries of the African meningitis belt, 2010-15: an analysis of surveillance data. *Lancet Infect Dis*. 17: 867–872.

326. van den Born J, van den Heuvel LP, Bakker MA, Veerkamp JH, Assmann KJ, Weening JJ, Berden JH. **1993**. Distribution of GBM heparan sulfate proteoglycan core protein and side chains in human glomerular diseases. *Kidney Int*. 43: 454–463.

327. van der Merwe PA. **2001**. Surface Plasmon Resonance. In *Protein-Ligand Interactions: Hydrodynamics and Calorimetry*; Harding SE, Chowdhry BZ, Ed. Oxford University Press: Oxford, UK; pp 137–170.

328. Varki A. **1994**. Selectin ligands. *Proc Natl Acad Sci USA* 91: 7390–7397.

329. Varki A. **2011**. Since there are PAMPs and DAMPs, there must be SAMPs? Glycan "self-associated molecular patterns" dampen innate immunity, but pathogens can mimic them. *Glycobiology* 21: 1121–1124.

330. Varki A. **2017**. Biological roles of glycans. *Glycobiology* 27: 3–49.

331. Vasudevan D, Takeuchi H, Johar SS, Majerus E, Haltiwanger RS. **2015**. Peters plus syndrome mutations disrupt a noncanonical ER quality-control mechanism. *Curr Biol*. 25: 286–295.

332. Venkatakrishnan V, Thaysen-Andersen M, Chen SC, Nevalainen H, Packer NH. **2015**. Cystic fibrosis and bacterial colonization define the sputum *N*-glycosylation phenotype. *Glycobiology* 25: 88–100.

333. Vézina LP, Faye L, Lerouge P, D'Aoust MA, Marquet-Blouin E, Burel C, Lavoie PO, Bardor M, Gomord V. **2009**. Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. *Plant Biotechnol J.* 7: 442– 455.

334. Viegas A, Manso J, Nobrega FL, Cabrita EJ. **2011**. Saturation-Transfer Difference (STD) NMR: A simple and fast method for ligand screening and characterization of protein binding. *J Chem Educ*. 88: 990–994.

335. Wang D, Liu S, Trummer BJ, Deng C,Wang A. **2002**. Carbohydrate microarrays for the recognition of cross-reactive molecular markers ofmicrobes and host cells. *Nat Biotechnol*. 20: 275–281.

336. Wang J, White AL. **2000**. Role of Calnexin, calreticulin, and endoplasmic reticulum mannosidase I in apolipoprotein(a) intracellular targeting. *Biochemistry* 39: 8993–9000.

337. Wang W, Ignatius AA, Ohtake S, Yang TC. **2018**. Introduction to High-Concentration Proteins. In *Challenges in Protein Products*. Warne NW, Mahler HC, Ed. Springer International Publishing: Cham, Switzerland; pp 99–123.

338. Wang X, Deng Z, Huang C, Zhu T, Lou J, Wang L, Li Y. **2018**. Differential *N*-glycan patterns identified in lung adenocarcinoma by *N*-glycan profiling of formalin-fixed paraffin-embedded (FFPE) tissue sections. *J Proteomics* 172: 1–10.

339. Wang X, Inoue S, Gu J, Miyoshi E, Noda K, Li W, Mizuno-Horikawa Y, Nakano M, Asahi M, Takahashi M, Uozumi N, Ihara S, Lee SH, Ikeda Y, Yamaguchi Y, Aze Y, Tomiyama Y, Fujii J, Suzuki K, Kondo A, Shapiro SD, Lopez-Otin C, Kuwaki T, Okabe M, Honke K, Taniguchi N. **2005**. Dysregulation of TGF-beta1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucosedeficient mice. *Proc Natl Acad Sci USA* 102: 15791–15796.

340. Weerapana E, Imperiali B. **2006**. Asparagine-linked protein glycosylation: From eukaryotic to prokaryotic systems. *Glycobiology* 16: 91R–101R.

341. Wiederschain GY. **2013**. Glycobiology: Progress, problems, and perspectives. *Biochemistry* 78: 679–696.

342. Wier M, Edidin M. **1988**. Constraint of the translational diffusion of a membrane glycoprotein by its external domains. *Science* 242: 412–414.

343. Winchester B. **2005**. Lysosomal metabolism of glycoproteins. *Glycobiology* 15: 1R–15R.

344. Wormald MR, Dwek RA. **1999**. Glycoproteins: glycan presentation and protein-fold stability. *Structure* 7: R155-R160.

345. Wu SC, Ye R, Wu XC, Ng SC, Wong SL. **1998**. Enhanced secretory production of a single-chain antibody fragment from *Bacillus subtilis* by co-production of molecular chaperones. *J Bacteriol.* 180: 2830–2835.

346. Wurm FM. **2004**. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol*. 22: 1393–1398.

347. Xie O, Bolgiano B, Gao F, Lockyer K, Swann C, Jones C, Delrieu I, Njanpop-Lafourcade BM, Tamekloe TA, Pollard AJ, Norheim G. **2012**. Characterization of size, structure and purity of serogroup X Neisseria meningitidis polysaccharide, and development of an assay for quantification of human antibodies. *Vaccine* 30: 5812–5823.

348. Xu C, Ng DT. **2015**. O-mannosylation: The other glycan player of ER quality control. *Semin Cell Dev Biol*. 41: 129–134.

349. Yagi H, Saito T, Yanagisawa M, Yu RK, Kato K. **2012**. Lewis X-carrying N-glycans regulate the proliferation of mouse embryonic neural stem cells via the Notch signaling pathway. *J Biol Chem*. 287: 24356–24364.

350. Yamamoto K. **2014**. Intracellular lectins are involved in quality control of glycoproteins. *Proc Jpn Acad Ser B Phys Biol Sci*. 90: 67–82.

351. Yan J, Kline AD, Mo H, Shapiro MJ, Zartler ER. **2003**. The effect of relaxation on the epitope mapping by saturation transfer difference NMR. *J Magn Reson*. 163: 270–276.

352. Yang Q, Jennings H. **2001**. Purification of Capsular Polysaccharide. In *Meningococcal Vaccines*; Pollard AJ, Maiden MC, Ed. Humana Press Inc: Totowa, NJ; pp 41–47.

353. Yuen CT, Storring PL, Tiplady RJ, Izquierdo M, Wait R, Gee CK, Gerson P, Lloyd P, Cremata JA. **2003**. Relationships between the N-glycan structures and biological activities of recombinant human erythropoietins produced using different culture conditions and purification procedures. *Br J Haematol*. 121: 511–526.

354. Zapun A, Jakob CA, Thomas DY, Bergeron JJM. **1999**. Protein folding in a specialized compartment: the endoplasmic reticulum. *Structure* 7: R173–R182.

355. Zeng W, Ford KL, Bacic A, Heazlewood JL. **2018**. *N*-linked glycan micro-heterogeneity in glycoproteins of Arabidopsis. *Mol Cell Proteomics* 17: 413–421.

356. Zhang C. **2012**. Hybridoma Technology for the Generation of Monoclonal Antibodies. In *Antibody Methods and Protocols*; Proetzel G, Ebersbach H, Ed. Humana Press: New York, NY; pp 117–135.

357. Zhang W, Cao P, Chen S, Spence AM, Zhu S, Staudacher E, Schachter H. **2003**. Synthesis of paucimannose *N*-glycans by *Caenorhabditis elegans* requires prior actions of UDP-*N*-acetyl-D-glucosamine:α-3-D-mannoside β1,2-*N*-acetylglucosaminyltransferase I, α 3,6-mannosidase II and a specific membrane-bound β -*N*-acetylglucosaminidase. *Biochem J*. 372: 53–64.

358. Zhang ZY, Ollmann IR, Ye X-S, Wischnat R, Baasov T, Wong CH. **1999**. Programmable one-pot oligosaccharide synthesis. *J Am Chem Soc*. 121: 734–753.

359. Zhou M, Li Q, Kong W, Wang R. **2018**. Experimental Methods Used for Identifying Small-Molecule Inhibitors of Protein-Protein Interaction. In *Targeting Protein-Protein Interactions by Small Molecules*. Sheng C, Georg GI, Ed. Springer: Singapore; pp 95–133.

360. Zhou RW, Mkhikian H, Grigorian A, Hong A, Chen D, Arakelyan A, Demetriou M. **2014**. *N*-glycosylation bidirectionally extends the boundaries of thymocyte positive selection by decoupling Lck from Ca2+ signaling. *Nat Immunol*. 15:1038–1045.

361. Zhu Y, Liu TW, Cecioni S, Eskandari R, Zandberg WF, Vocadlo DJ. **2015**. O-GlcNAc occurs cotranslationally to stabilize nascent polypeptide chains. *Nat Chem Biol*. 11: 319–325.

362. Zipser B, Bello-DeOcampo D, Diestel S, Tai MH, Schmitz B. **2012**. Mannitou monoclonal antibody uniquely recognizes paucimannose, a marker for human cancer, stemness, and inflammation. *J Carbohydr Chem*. 31: 504–518.

363. Zipser B, McKay R. **1981**. Monoclonal antibodies distinguish identifiable neurones in the leech. *Nature* 289: 549–554.

364. Zu Putlitz J, Kubasek WL, Duchêne M, Marget M, von Specht BU, Domdey H. **1990**. Antibody production in baculovirus-infected insect cells. *Biotechnology (NY)* 8: 651–654.

APPENDIX

Mannitou mAb light chain aa sequence:

DVVVTQTPLSLPVSFGDQASISCRSSQSLVNSYGSTYLSWYLHRPGQSPQLLIYGISNRFSGVPD
RFSGSGSGTDFTLTIRTIKPEDLGMYYCLQGTHQPWTFGGGTKLEIKKRADAAPTVSIFPPSSE
QLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDE
YERHNSYTCEATHKTSTSPIVKSFNRKEC

Mannitou mAb heavy chain aa sequence:

EVKLLESGGGLVQPGGSLKLSCAASGFDFSTYWMSWVRQAPGKGLEWIGEINPDSSTINYTPSL
KDKFIISRDNAKNTLYLQMSKVRSEDSVLYYCVRPGTWGYFDYWGQGTTLTVSSASQSFPNVF
PLVSCESPLSDKNLVAMGCLARDFLPSTISFTWNYQNNTEVIQGIRTFPTLRTGGKYLATSQVL
LSPKSILEGSDEYLVCKIHYGGKNRDLHVPIPAVAEMNPNVNVFVPPRDGFSGPAPRKSKLICE
ATNFTPKPITVSWLKDGKLVESGFTTDPVTIENKGSTPQTYKVISTLTISEIDWLNLNVYTCRV
DHRGLTFLKNVSSTCAASPSTDILTFTIPPSFADIFLSKSANLTCLVSNLATYETLNISWASQSGE
PLETKIKIMESHPNGTFSAKGVASVCVEDWNNRKEFVCTVTHRDLPSPQKKFISKPNEVHKHP
PAVYLLPPAREQLNLRESATVTCLVKGFSPADISVQWLQRGQLLPQEKYVTSAPMPEPGAPGF
YFTHSILTVTEEEWNSGETYTCVVGHEALPHLVTERTVDKSTGKPTLYNVSLIMSDTGGTCY

Figure S1. Mannitou mAb amino acid sequence. Green: variable region, Pink: constant region. CH1 domain is underlined. The sequence is composed of the variable regions, determined by sequencing, and the constant regions obtained from GenBank: LC085625.1 (light chain) and UniProtKB: locus A24976 (heavy chain).

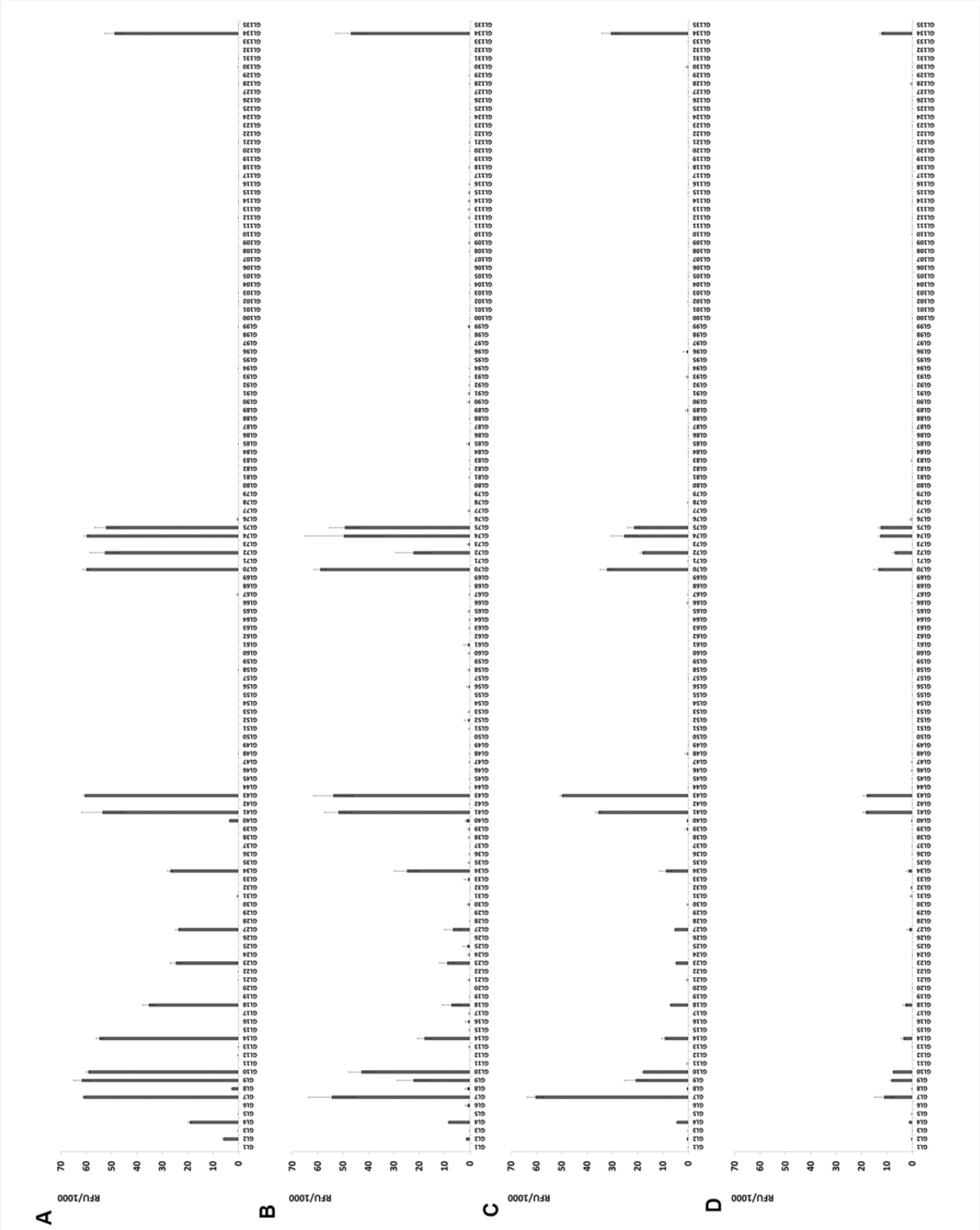


Figure S2. Glycan microarray results using different concentrations of Mannitou antibody. A: 500 nM, B: 250 nM, C: 125 nM, D: 62.5 nM. The pattern of behaviour is maintained among the applied concentrations. Each histogram represents the mean RFU (relative fluorescence unit) values for four spots with the SD (standard deviation) of the mean.

The $AmSO_4$ Suite Composition Table

Number	Salt	Buffer	Precipitant	Cat. no. (Refill-Hit Solution, 4 x 12.5 ml tubes)
1			2.2 M Ammonium sulfate	134401
2	0.2 M Ammonium acetate		2.2 M Ammonium sulfate	134402
3	0.2 M Ammonium chloride		2.2 M Ammonium sulfate	134403
4	0.2 M Ammonium phosphate		2.2 M Ammonium sulfate	134404
5	0.2 M Ammonium fluoride		2.2 M Ammonium sulfate	134405
6	0.2 M Ammonium formate		2.2 M Ammonium sulfate	134406
7	0.18 M tri-Ammonium citrate		2.2 M Ammonium sulfate	134407
8	0.2 M di-Ammonium phosphate		2.2 M Ammonium sulfate	134408
9	0.2 M Ammonium iodide		2.2 M Ammonium sulfate	134409
10	0.2 M Ammonium nitrate		2.2 M Ammonium sulfate	134410
11	0.2 M di-Ammonium tartrate		2.2 M Ammonium sulfate	134411
12	0.2 M Cadmium chloride		2.2 M Ammonium sulfate	134412
13	0.2 M Cadmium sulfate		2.2 M Ammonium sulfate	134413
14	0.2 M Cesium chloride		2.2 M Ammonium sulfate	134414
15	0.2 M Cesium sulfate		2.2 M Ammonium sulfate	134415
16	0.2 M Ammonium bromide		2.2 M Ammonium sulfate	134416
17	0.2 M Lithium acetate		2.2 M Ammonium sulfate	134417
18	0.2 M Lithium chloride		2.2 M Ammonium sulfate	134418
19	0.2 M tri-Lithium citrate		2.2 M Ammonium sulfate	134419
20	0.2 M Lithium nitrate		2.2 M Ammonium sulfate	134420
21	0.2 M Lithium sulfate		2.2 M Ammonium sulfate	134421
22	0.2 M Potassium acetate		2.2 M Ammonium sulfate	134422
23	0.2 M Potassium bromide		2.2 M Ammonium sulfate	134423
24	0.2 M Potassium chloride		2.2 M Ammonium sulfate	134424
25	0.2 M tri-Potassium citrate		2.2 M Ammonium sulfate	134425
26	0.2 M Potassium phosphate		2.2 M Ammonium sulfate	134426
27	0.2 M Potassium fluoride		2.2 M Ammonium sulfate	134427
28	0.2 M Potassium formate		2.2 M Ammonium sulfate	134428
29	0.2 M di-Potassium phosphate		2.2 M Ammonium sulfate	134429
30	0.2 M Potassium iodide		2.2 M Ammonium sulfate	134430
31	0.2 M Potassium nitrate		2.2 M Ammonium sulfate	134431
32	0.2 M K/Na tartrate		2.2 M Ammonium sulfate	134432
33	0.2 M Potassium sulfate		2.2 M Ammonium sulfate	134433
34	0.2 M Potassium thiocyanate		2.2 M Ammonium sulfate	134434
35	0.2 M Sodium acetate		2.2 M Ammonium sulfate	134435
36	0.2 M Sodium bromide		2.2 M Ammonium sulfate	134436
37	0.2 M Sodium chloride		2.2 M Ammonium sulfate	134437
38	0.2 M tri-Sodium citrate		2.2 M Ammonium sulfate	134438
39	0.2 M Sodium phosphate		2.2 M Ammonium sulfate	134439
40	0.2 M Sodium fluoride		2.2 M Ammonium sulfate	134440
41	0.2 M Sodium formate		2.2 M Ammonium sulfate	134441
42	0.2 M di-Sodium phosphate		2.2 M Ammonium sulfate	134442
43	0.2 M Sodium iodide		2.2 M Ammonium sulfate	134443
44	0.2 M Sodium malonate		2.2 M Ammonium sulfate	134444
45	0.2 M Sodium nitrate		2.2 M Ammonium sulfate	134445
46	0.2 M Sodium sulfate		2.2 M Ammonium sulfate	134446
47	0.2 M di-Sodium tartrate		2.2 M Ammonium sulfate	134447
48	0.2 M Sodium thiocyanate		2.2 M Ammonium sulfate	134448

Number	Salt	Buffer	Precipitant	Cat. no. (Refill-Hit Solution, 4 x 12.5 ml tubes)
49		0.1 M Citric acid pH 4.0	0.8 M Ammonium sulfate	134449
50		0.1 M Citric acid pH 5.0	0.8 M Ammonium sulfate	134450
51		0.1 M MES pH 6.0	0.8 M Ammonium sulfate	134451
52		0.1 M HEPES pH 7.0	0.8 M Ammonium sulfate	134452
53		0.1 M Tris pH 8.0	0.8 M Ammonium sulfate	134453
54		0.1 M Bicine pH 9.0	0.8 M Ammonium sulfate	134454
55		0.1 M Citric acid pH 4.0	1.6 M Ammonium sulfate	134455
56		0.1 M Citric acid pH 5.0	1.6 M Ammonium sulfate	134456
57		0.1 M MES pH 6.0	1.6 M Ammonium sulfate	134457
58		0.1 M HEPES pH 7.0	1.6 M Ammonium sulfate	134458
59		0.1 M Tris pH 8.0	1.6 M Ammonium sulfate	134459
60		0.1 M Bicine pH 9.0	1.6 M Ammonium sulfate	134460
61		0.1 M Citric acid pH 4.0	2.4 M Ammonium sulfate	134461
62		0.1 M Citric acid pH 5.0	2.4 M Ammonium sulfate	134462
63		0.1 M MES pH 6.0	2.4 M Ammonium sulfate	134463
64		0.1 M HEPES pH 7.0	2.4 M Ammonium sulfate	134464
65		0.1 M Tris pH 8.0	2.4 M Ammonium sulfate	134465
66		0.1 M Bicine pH 9.0	2.4 M Ammonium sulfate	134466
67		0.1 M Citric acid pH 4.0	3.2 M Ammonium sulfate	134467
68		0.1 M Citric acid pH 5.0	3.2 M Ammonium sulfate	134468
69		0.1 M MES pH 6.0	3.2 M Ammonium sulfate	134469
70		0.1 M HEPES pH 7.0	3.2 M Ammonium sulfate	134470
71		0.1 M Tris pH 8.0	3.2 M Ammonium sulfate	134471
72		0.1 M Bicine pH 9.0	3.2 M Ammonium sulfate	134472
73	0.1 M tri-Sodium citrate		0.5 M Ammonium sulfate; 1.0 M Lithium Sulfate	134473
74			1.0 M Ammonium sulfate	134474
75		0.1 M Sodium acetate pH 4.6	1.0 M Ammonium sulfate	134475
76		0.1 M HEPES sodium salt pH 7.5	1.0 M Ammonium sulfate; 2% (w/v) PEG 400	134476
77		0.1 M Tris·HCl pH 8.5	1.0 M Ammonium sulfate	134477
78	0.05 M tri-Sodium citrate		1.2 M Ammonium sulfate; 3% (w/v) Isopropanol	134478
79		0.1 M Tris·HCl pH 8.5	1.5 M Ammonium sulfate; 15% (w/v) Glycerol	134479
80	0.5 M Lithium chloride		1.6 M Ammonium sulfate	134480
81	1.0 M Lithium sulfate		1.6 M Ammonium sulfate	134481
82	0.2 M Sodium chloride	0.1 M HEPES sodium salt pH 7.5	1.6 M Ammonium sulfate	134482
83		0.1 M HEPES sodium salt pH 7.5	1.6 M Ammonium sulfate; 2% (w/v) PEG 1000	134483
84		0.1 M MES sodium salt pH 6.5	1.8 M Ammonium sulfate	134484
85	2.0 M Sodium chloride		2.0 M Ammonium sulfate	134485
86		0.1 M Sodium acetate pH 4.6	2.0 M Ammonium sulfate	134486
87		0.1 M MES sodium salt pH 6.5	2.0 M Ammonium sulfate; 5% (w/v) PEG 400	134487
88		0.1 M Tris·HCl pH 8.5	2.0 M Ammonium sulfate	134488
89			2.2 M Ammonium sulfate	134489
90			2.2 M Ammonium sulfate; 20% (w/v) Glycerol	134490
91	0.1 M tri-Sodium citrate		2.4 M Ammonium sulfate	134491
92			3.0 M Ammonium sulfate; 1% (w/v) MPD	134492
93			3.0 M Ammonium sulfate; 10% (w/v) Glycerol	134493
94		0.1 M HEPES sodium salt pH 7.5	3.5 M Ammonium sulfate	134494
95		0.1 M MES sodium salt pH 6.5	3.5 M Ammonium sulfate; 1% (w/v) MPD	134495
96			3.5 M Ammonium sulfate	134496

Figure S3. The $AmSO_4$ composition table.

Index™ — HR2-144 Reagent Formulation

Tube #	Salt	Buffer ◊	Precipitant
1.	None	0.1 M Citric acid pH 3.5	2.0 M Ammonium sulfate
2.	None	0.1 M Sodium acetate trihydrate pH 4.5	2.0 M Ammonium sulfate
3.	None	0.1 M BIS-TRIS pH 5.5	2.0 M Ammonium sulfate
4.	None	0.1 M BIS-TRIS pH 6.5	2.0 M Ammonium sulfate
5.	None	0.1 M HEPES pH 7.5	2.0 M Ammonium sulfate
6.	None	0.1 M Tris pH 8.5	2.0 M Ammonium sulfate
7.	None	0.1 M Citric acid pH 3.5	3.0 M Sodium chloride
8.	None	0.1 M Sodium acetate trihydrate pH 4.5	3.0 M Sodium chloride
9.	None	0.1 M BIS-TRIS pH 5.5	3.0 M Sodium chloride
10.	None	0.1 M BIS-TRIS pH 6.5	3.0 M Sodium chloride
11.	None	0.1 M HEPES pH 7.5	3.0 M Sodium chloride
12.	None	0.1 M Tris pH 8.5	3.0 M Sodium chloride
13.	None	0.1 M BIS-TRIS pH 5.5	0.3 M Magnesium formate dihydrate
14.	None	0.1 M BIS-TRIS pH 6.5	0.5 M Magnesium formate dihydrate
15.	None	0.1 M HEPES pH 7.5	0.5 M Magnesium formate dihydrate
16.	None	0.1 M Tris pH 8.5	0.3 M Magnesium formate dihydrate
17.	None	None - pH 5.6	1.26 M Sodium phosphate monobasic monohydrate 0.14 M Potassium phosphate dibasic
18.	None	None - pH 6.9	0.49 M Sodium phosphate monobasic monohydrate 0.91 M Potassium phosphate dibasic
19.	None	None - pH 8.2	0.056 M Sodium phosphate monobasic monohydrate 1.344 M Potassium phosphate dibasic
20.	None	0.1 M HEPES pH 7.5	1.4 M Sodium citrate tribasic dihydrate
21.	None	None	1.8 M Ammonium citrate tribasic pH 7.0
22.	None	None	0.8 M Succinic acid pH 7.0
23.	None	None	2.1 M DL-Malic acid pH 7.0
24.	None	None	2.8 M Sodium acetate trihydrate pH 7.0
25.	None	None	3.5 M Sodium formate pH 7.0
26.	None	None	1.1 M Ammonium tartrate dibasic pH 7.0
27.	None	None	2.4 M Sodium malonate pH 7.0
28.	None	None	35% v/v Tacsimate™ pH 7.0
29.	None	None	60% v/v Tacsimate™ pH 7.0
30.	0.1 M Sodium chloride	0.1 M BIS-TRIS pH 6.5	1.5 M Ammonium sulfate
31.	0.8 M Potassium sodium tartrate tetrahydrate	0.1 M Tris pH 8.5	0.5% w/v Polyethylene glycol monomethyl ether 5,000
32.	1.0 M Ammonium sulfate	0.1 M BIS-TRIS pH 5.5	1% w/v Polyethylene glycol 3,350
33.	1.1 M Sodium malonate pH 7.0	0.1 M HEPES pH 7.0	0.5% v/v Jeffamine® ED-2001 pH 7.0
34.	1.0 M Succinic acid pH 7.0	0.1 M HEPES pH 7.0	1% w/v Polyethylene glycol monomethyl ether 2,000
35.	1.0 M Ammonium sulfate	0.1 M HEPES pH 7.0	0.5% w/v Polyethylene glycol 8,000
36.	15% v/v Tacsimate™ pH 7.0	0.1 M HEPES pH 7.0	2% w/v Polyethylene glycol 3,350
37.	None	None	25% w/v Polyethylene glycol 1,500
38.	None	0.1 M HEPES pH 7.0	30% v/v Jeffamine® M-600® pH 7.0
39.	None	0.1 M HEPES pH 7.0	30% v/v Jeffamine® ED-2001 pH 7.0
40.	None	0.1 M Citric acid pH 3.5	25% w/v Polyethylene glycol 3,350
41.	None	0.1 M Sodium acetate trihydrate pH 4.5	25% w/v Polyethylene glycol 3,350
42.	None	0.1 M BIS-TRIS pH 5.5	25% w/v Polyethylene glycol 3,350
43.	None	0.1 M BIS-TRIS pH 6.5	25% w/v Polyethylene glycol 3,350
44.	None	0.1 M HEPES pH 7.5	25% w/v Polyethylene glycol 3,350
45.	None	0.1 M Tris pH 8.5	25% w/v Polyethylene glycol 3,350
46.	None	0.1 M BIS-TRIS pH 6.5	20% w/v Polyethylene glycol monomethyl ether 5,000
47.	None	0.1 M BIS-TRIS pH 6.5	28% w/v Polyethylene glycol monomethyl ether 2,000
48.	0.2 M Calcium chloride dihydrate	0.1 M BIS-TRIS pH 5.5	45% v/v (+/-)-2-Methyl-2,4-pentanediol

Index™ HR2-144 Reagent Formulation

Tube #	Salt	Tube #	Buffer ◊	Tube #	Precipitant
49.	0.2 M Calcium chloride dihydrate	49.	0.1 M BIS-TRIS pH 6.5	49.	45% v/v (+/-)-2-Methyl-2,4-pentanediol
50.	0.2 M Ammonium acetate	50.	0.1 M BIS-TRIS pH 5.5	50.	45% v/v (+/-)-2-Methyl-2,4-pentanediol
51.	0.2 M Ammonium acetate	51.	0.1 M BIS-TRIS pH 6.5	51.	45% v/v (+/-)-2-Methyl-2,4-pentanediol
52.	0.2 M Ammonium acetate	52.	0.1 M HEPES pH 7.5	52.	45% v/v (+/-)-2-Methyl-2,4-pentanediol
53.	0.2 M Ammonium acetate	53.	0.1 M Tris pH 8.5	53.	45% v/v (+/-)-2-Methyl-2,4-pentanediol
54.	0.05 M Calcium chloride dihydrate	54.	0.1 M BIS-TRIS pH 6.5	54.	30% v/v Polyethylene glycol monomethyl ether 550
55.	0.05 M Magnesium chloride hexahydrate	55.	0.1 M HEPES pH 7.5	55.	30% v/v Polyethylene glycol monomethyl ether 550
56.	0.2 M Potassium chloride	56.	0.05 M HEPES pH 7.5	56.	35% v/v Pentaerythritol propoxylate (5/4 PO/OH)
57.	0.05 M Ammonium sulfate	57.	0.05 M BIS-TRIS pH 6.5	57.	30% v/v Pentaerythritol ethoxylate (15/4 EO/OH)
58.	None	58.	0.1 M BIS-TRIS pH 6.5	58.	45% v/v Polypropylene glycol P 400
59.	0.02 M Magnesium chloride hexahydrate	59.	0.1 M HEPES pH 7.5	59.	22% w/v Poly(acrylic acid sodium salt) 5,100
60.	0.01 M Cobalt(II) chloride hexahydrate	60.	0.1 M Tris pH 8.5	60.	20% w/v Polyvinylpyrrolidone K 15
61.	0.2 M L-Proline	61.	0.1 M HEPES pH 7.5	61.	10% w/v Polyethylene glycol 3,350
62.	0.2 M Trimethylamine N-oxide dihydrate	62.	0.1 M Tris pH 8.5	62.	20% w/v Polyethylene glycol monomethyl ether 2,000
63.	5% v/v Tacsimate™ pH 7.0	63.	0.1 M HEPES pH 7.0	63.	10% w/v Polyethylene glycol monomethyl ether 5,000
64.	0.005 M Cobalt(II) chloride hexahydrate 0.005 M Nickel(II) chloride hexahydrate 0.005 M Cadmium chloride hydrate 0.005 M Magnesium chloride hexahydrate	64.	0.1 M HEPES pH 7.5	64.	12% w/v Polyethylene glycol 3,350
65.	0.1 M Ammonium acetate	65.	0.1 M BIS-TRIS pH 5.5	65.	17% w/v Polyethylene glycol 10,000
66.	0.2 M Ammonium sulfate	66.	0.1 M BIS-TRIS pH 5.5	66.	25% w/v Polyethylene glycol 3,350
67.	0.2 M Ammonium sulfate	67.	0.1 M BIS-TRIS pH 6.5	67.	25% w/v Polyethylene glycol 3,350
68.	0.2 M Ammonium sulfate	68.	0.1 M HEPES pH 7.5	68.	25% w/v Polyethylene glycol 3,350
69.	0.2 M Ammonium sulfate	69.	0.1 M Tris pH 8.5	69.	25% w/v Polyethylene glycol 3,350
70.	0.2 M Sodium chloride	70.	0.1 M BIS-TRIS pH 5.5	70.	25% w/v Polyethylene glycol 3,350
71.	0.2 M Sodium chloride	71.	0.1 M BIS-TRIS pH 6.5	71.	25% w/v Polyethylene glycol 3,350
72.	0.2 M Sodium chloride	72.	0.1 M HEPES pH 7.5	72.	25% w/v Polyethylene glycol 3,350
73.	0.2 M Sodium chloride	73.	0.1 M Tris pH 8.5	73.	25% w/v Polyethylene glycol 3,350
74.	0.2 M Lithium sulfate monohydrate	74.	0.1 M BIS-TRIS pH 5.5	74.	25% w/v Polyethylene glycol 3,350
75.	0.2 M Lithium sulfate monohydrate	75.	0.1 M BIS-TRIS pH 6.5	75.	25% w/v Polyethylene glycol 3,350
76.	0.2 M Lithium sulfate monohydrate	76.	0.1 M HEPES pH 7.5	76.	25% w/v Polyethylene glycol 3,350
77.	0.2 M Lithium sulfate monohydrate	77.	0.1 M Tris pH 8.5	77.	25% w/v Polyethylene glycol 3,350
78.	0.2 M Ammonium acetate	78.	0.1 M BIS-TRIS pH 5.5	78.	25% w/v Polyethylene glycol 3,350
79.	0.2 M Ammonium acetate	79.	0.1 M BIS-TRIS pH 6.5	79.	25% w/v Polyethylene glycol 3,350
80.	0.2 M Ammonium acetate	80.	0.1 M HEPES pH 7.5	80.	25% w/v Polyethylene glycol 3,350
81.	0.2 M Ammonium acetate	81.	0.1 M Tris pH 8.5	81.	25% w/v Polyethylene glycol 3,350
82.	0.2 M Magnesium chloride hexahydrate	82.	0.1 M BIS-TRIS pH 5.5	82.	25% w/v Polyethylene glycol 3,350
83.	0.2 M Magnesium chloride hexahydrate	83.	0.1 M BIS-TRIS pH 6.5	83.	25% w/v Polyethylene glycol 3,350
84.	0.2 M Magnesium chloride hexahydrate	84.	0.1 M HEPES pH 7.5	84.	25% w/v Polyethylene glycol 3,350
85.	0.2 M Magnesium chloride hexahydrate	85.	0.1 M Tris pH 8.5	85.	25% w/v Polyethylene glycol 3,350
86.	0.2 M Potassium sodium tartrate tetrahydrate	86.	None	86.	20% w/v Polyethylene glycol 3,350
87.	0.2 M Sodium malonate pH 7.0	87.	None	87.	20% w/v Polyethylene glycol 3,350
88.	0.2 M Ammonium citrate tribasic pH 7.0	88.	None	88.	20% w/v Polyethylene glycol 3,350
89.	0.1 M Succinic acid pH 7.0	89.	None	89.	15% w/v Polyethylene glycol 3,350
90.	0.2 M Sodium formate	90.	None	90.	20% w/v Polyethylene glycol 3,350
91.	0.15 M DL-Malic acid pH 7.0	91.	None	91.	20% w/v Polyethylene glycol 3,350
92.	0.1 M Magnesium formate dihydrate	92.	None	92.	15% w/v Polyethylene glycol 3,350
93.	0.05 M Zinc acetate dihydrate	93.	None	93.	20% w/v Polyethylene glycol 3,350
94.	0.2 M Sodium citrate tribasic dihydrate	94.	None	94.	20% w/v Polyethylene glycol 3,350
95.	0.1 M Potassium thiocyanate	95.	None	95.	30% w/v Polyethylene glycol monomethyl ether 2,000
96.	0.15 M Potassium bromide	96.	None	96.	30% w/v Polyethylene glycol monomethyl ether 2,000

Figure S4. The Index composition table.

MIDAS™ **Box 1 of 2** **MD1-59**

Tube No.	% conc	Precipitant	% conc	Salt/Additive	pH	% conc	Buffer
1-1	50 % v/v	polypropylene glycol 400	5 %	dimethyl sulfoxide	6	0.1 M	HEPES-NaOH
1-2	12 % w/v	polyvinyl pyrrolidone K15	-	-	5.5	0.1 M	MES- NaOH
1-3	45 % w/v	polyacrylate 2100, sodium salt	-	-	6.5	0.1 M	HEPES-NaOH
1-4	14 % v/v	acrylic acid/maleic acid copolymer (50:50), sodium salt	-	-	-	-	-
1-5	12.5 % w/v	polyacrylate 2100, sodium salt	0.5 M	ammonium phosphate	-	-	-
1-6	19 % v/v	acrylic acid/maleic acid copolymer (50:50), sodium salt	-	-	8.5	0.1 M	Tris-HCl
1-7	10 % v/v	polypropylene glycol 400	-	-	-	-	-
1-8	5 % w/v	polyacrylate 2100, sodium salt	-	-	-	-	-
1-9	25 % v/v	pentaerythritol propoxylate (5/4 PO/OH)	-	-	6	0.1 M	MES- NaOH
1-10	24 % w/v	polyvinyl pyrrolidone K15	0.1 M	sodium sulfate	-	-	-
1-11	35 % v/v	pentaerythritol ethoxylate (15/4 EO/OH)	0.2 M	calcium chloride	6.5	0.1 M	HEPES-NaOH
1-12	35 % v/v	polypropylene glycol 400	-	-	7	0.1 M	K/Na Phosphate
1-13	20% v/v	Jeffamine D2000	0.2 M	sodium chloride	5.5	0.1 M	MES- NaOH
	10 % v/v	Jeffamine M2005	-	-	-	-	-
1-14	15 % v/v	pentaerythritol propoxylate (5/4 PO/OH	0.2 M	sodium thiocyanate	7	0.1 M	HEPES-NaOH
1-15	5 % w/v	polyvinyl alcohol type II	0.2 M	potassium acetate	7	0.1 M	HEPES-NaOH
	10 % v/v	Jeffamine T403	-	-	-	-	-
1-16	45 % v/v	pentaerythritol propoxylate (5/4 PO/OH)	0.2 M	sodium chloride	6	0.1 M	MES- NaOH
1-17	8 % w/v	polyvinyl alcohol type II	10 % v/v	1- propanol	7	0.1 M	HEPES-NaOH
1-18	30 % w/v	polyvinyl pyrrolidone K15	0.1 M	lithium sulfate	7	0.1 M	HEPES-NaOH
1-19	40 % v/v	polypropylene glycol 400	0.2 M	imidazole	7	-	-
1-20	8 % w/v	acrylic acid/maleic acid copolymer (50:50), sodium salt	0.06 M	lithium sulfate	7.5	0.1 M	HEPES-NaOH
	3 % v/v	pentaerythritol ethoxylate (3/4 EO/OH)	-	-	-	-	-
1-21	35 % v/v	Jeffamine SD2001	0.1 M	sodium chloride	8	0.1 M	Tris-HCl
1-22	30 % v/v	Jeffamine M600	10 % v/v	dimethyl sulfoxide	-	-	-
1-23	20 % v/v	polypropylene glycol 400	10 % v/v	1-propanol	-	-	-
1-24	28 % w/v	acrylic acid/maleic acid copolymer (50:50), sodium salt	-	-	6.5	0.1 M	HEPES-NaOH
1-25	15 % w/v	Jeffamine ED2003	10 % v/v	ethanol	-	-	-
1-26	30 % w/v	Jeffamine ED2003	0.2 M	sodium chloride	6	0.1 M	MES- NaOH
1-27	25 % v/v	Jeffamine SD2001	0.1 M	sodium malonate	5.5	0.1 M	MES- NaOH
1-28	15 % v/v	pentaerythritol propoxylate (5/4 PO/OH)	0.2 M	sodium chloride	6	0.1 M	MES- NaOH
1-29	35 % v/v	pentaerythritol ethoxylate (3/4 EO/OH)	0.2 M	Magnesium chloride	-	-	
1-30	40 % v/v	pentaerythritol propoxylate (5/4 PO/OH)	15 % v/v	ethanol	-	-	-
1-31	50 % v/v	pentaerythritol propoxylate (5/4 PO/OH)	-	-	8	0.1 M	Tris-HCl
1-32	12.5 % w/v	polyvinyl pyrrolidone K15	0.2 M	sodium chloride	8	0.1 M	Tris-HCl
	10 % w/v	PEG 4000	-	-	-	-	-
1-33	25 % v/v	pentaerythritol propoxylate (5/4 PO/OH)	0.1 M	sodium chloride	-	-	-
	10 % v/v	dimethyl sulfoxide,	-	-	-	-	-
1-34	35 % w/v	polyacrylate 2100, sodium salt	0.2 M	ammonium sulfate	7.5	0.1 M	HEPES-NaOH
1-35	30 % v/v	pentaerythritol ethoxylate (15/4 EO/OH)	0.1 M	magnesium formate	8.5	0.1 M	Tris-HCl
1-36	20 % v/v	Glascol W13	0.2 M	sodium sulfate	7.5	0.1 M	HEPES-NaOH
1-37	60 % v/v	polypropylene glycol 400	-	-	8	0.1 M	Tris-HCl
1-38	30 % v/v	pentaerythritol ethoxylate (15/4 EO/OH)	-	-	7.5	0.1 M	HEPES-NaOH
	6 % w/v	polyvinyl pyrrolidone K15	-	-	-	-	-
1-39	45 % v/v	polypropylene glycol 400	10 % v/v	ethanol	-	-	-
1-40	10 % v/v	pentaerythritol ethoxylate (3/4 EO/OH)	10 % v/v	1-butanol	-	-	-
1-41	12.5 % w/v	polyacrylate 2100, sodium salt	-	-	7	0.1 M	HEPES-NaOH
	6 % v/v	Jeffamine SD2001	-	-	-	-	-
1-42	6 % w/v	polyvinyl pyrrolidone K15	-	-	6.5	0.1 M	HEPES-NaOH
1-43	20 % w/v	Jeffamine ED2003	-	-	6.5	0.1 M	HEPES-NaOH
1-44	20 % v/v	glycerol ethoxylate	10 % v/v	tetrahydrofuran	8.0	0.1 M	Tris-HCl
1-45	25 % v/v	Jeffamine D2000	0.2 M	imidazole	7	-	-
1-46	30 % v/v	Jeffamine SD2001	0.2 M	potassium chloride	6.5	0.1 M	HEPES-NaOH
1-47	30 % v/v	polypropylene glycol 400	0.1 M	sodium chloride	-	-	-
1-48	20 % v/v	Jeffamine SD2001	15 % v/v	1-propanol	-	-	-

MIDAS™ Box 2 of 2 MD1-59

Tube No.	% conc	Precipitant	% conc	Salt/Additive	pH	% conc	Buffer
2-1	25 % v/v	Jeffamine T403	0.2 M	lithium sulfate	8	0.1 M	Tris-HCl
2-2	35 % v/v	pentaerythritol propoxylate (5/4 PO/OH)	0.2 M	potassium acetate	-	-	-
2-3	20 % v/v	pentaerythritol ethoxylate (15/4 EO/OH)	0.2 M	potassium chloride	9.5	0.1 M	Glycine
2-4	40 % v/v	pentaerythritol propoxylate (5/4 PO/OH)	0.2 M	sodium thiocyanate	7	0.1 M	HEPES-NaOH
2-5	15 % v/v	Jeffamine T403	0.2 M	potassium chloride	6.5	0.1 M	HEPES-NaOH
	15 % v/v	Jeffamine ED2003	-	-	-	-	-
2-6	15 % v/v	pentaerythritol ethoxylate (15/4 EO/OH),	0.2 M	potassium acetate	6	0.1 M	MES- NaOH
	3 % v/v	Jeffamine T403	-	-	-	-	-
2-7	30 % w/v	polyacrylate 2100, sodium salt	0.1 M	sodium malonate	7	0.1 M	HEPES-NaOH
2-8	10 % v/v	Jeffamine D2000	10 % v/v	ethanol	-	-	-
	10 % v/v	Jeffamine M2005	-	-	-	-	-
2-9	25 % w/v	Jeffamine ED2003	0.1 M	lithium sulfate	8	0.1 M	Tris-HCl
2-10	10 % v/v	Jeffamine T403	-	-	8	0.1 M	Tris-HCl
	10 % w/v	Jeffamine ED2003	-	-	-	-	-
2-11	25 % w/v	polyacrylate 2100, sodium salt	0.1 M	lithium sulfate	6.5	0.1 M	HEPES-NaOH
2-12	15 % w/v	polyacrylate 2100, sodium salt	0.2 M	magnesium chloride	7.5	0.1 M	HEPES-NaOH
2-13	40 % v/v	Jeffamine D2000	-	-	6.5	0.1 M	HEPES-NaOH
2-14	10 % w/v	polyacrylate 2100, sodium salt	0.5 M	sodium chloride	8	0.1 M	Tris-HCl
2-15	14 % v/v	Jeffamine ED900	-	-	7	0.1 M	K/Na Phosphate
	11 % v/v	Jeffamine SD2001	-	-	-	-	-
2-16	20 % w/v	polyacrylate 2100, sodium salt	0.2 M	sodium chloride	9	0.1 M	Bicine
2-17	20 % v/v	Jeffamine D2000	0.2 M	sodium malonate	5.5	0.1 M	MES- NaOH
2-18	30 % v/v	Jeffamine M2070	0.2 M	potassium chloride	8	0.1 M	Tris-HCl
2-19	20 % v/v	Jeffamine M2070	20 % v/v	dimethyl sulfoxide	-	-	-
2-20	40 % w/v	pentaerythritol propoxylate (17/8 PO/OH)	0.2 M	magnesium chloride	5.5	0.1 M	MES- NaOH
2-21	20 % w/v	polyacrylate 5100, sodium salt	-	-	8	0.1 M	Tris-HCl
2-22	28 % v/v	poly(ethylene imine) branched	-	-	7	0.1 M	HEPES-NaOH
2-23	20 % v/v	Sokalan® CP 7	0.1 M	ammonium formate	7	0.1 M	HEPES-NaOH
2-24	20 % w/v	Sokalan® HP 56	0.2 M	sodium sulfate	8	0.1 M	Tris-HCl
2-25	25 % v/v	Sokalan® CP 7	0.1 M	potassium chloride	7	0.1 M	HEPES-NaOH
2-26	20 % v/v	Sokalan® CP 5	0.3 M	ammonium formate	7	0.1 M	HEPES-NaOH
2-27	40 % v/v	glycerol ethoxylate	-	-	-	-	-
2-28	30 % v/v	glycerol ethoxylate	-	-	8.5	0.1 M	Tris-HCl
2-29	15 % v/v	Sokalan® HP 66 K	-	-	7	0.1 M	HEPES-NaOH
	3 % v/v	poly(ethylene imine)	-	-	-	-	-
2-30	35% v/v	glycerol ethoxylate	0.2 M	lithium citrate	-	-	-
2-31	30 % v/v	glycerol ethoxylate	0.2 M	ammonium acetate	6.5	0.1 M	MES- NaOH
2-32	20 % v/v	Sokalan® CP 42	5% v/v	methanol	8	0.1 M	Tris-HCl
2-33	25 % v/v	Sokalan® CP 42	10 % v/v	tetrahydrofuran	7	0.1 M	Tris-HCl
2-34	20 % v/v	Sokalan® CP 42	0.1 M	lithium acetate	6	0.1 M	Bis-Tris- NaOH
2-25	15 % v/v	Sokalan® CP 12 S	0.1 M	lithium citrate	5.5	0.1 M	Bis-Tris- NaOH
2-36	15 % v/v	Sokalan® CP 5	-	-	6	0.1 M	Bis-Tris- NaOH
2-37	25 % v/v	Sokalan® CP 42	-	-	6	0.1 M	Bis-Tris- NaOH
2-38	25 % v/v	Sokalan® HP 66 K	0.2 M	ammonium acetate	7	0.1 M	HEPES-NaOH
2-39	20 % v/v	glycerol ethoxylate	-	-	8.5	0.1 M	Tris-HCl
	3 % v/v	poly(ethylene imine)	-	-	-	-	-
2-40	25 % v/v	glycerol ethoxylate	0.2 M	ammonium chloride	7.5	0.1 M	HEPES-NaOH
2-41	40 % v/v	Glascol® W13	0.2 M	potassium citrate	-	-	-
2-42	30 % w/v	polyacrylate 5100, sodium salt	10 % v/v	ethanol	6	0.1 M	MES- NaOH
2-43	15 % v/v	Sokalan® CP 42	0.2 M	potassium citrate	-	-	-
2-44	30 % v/v	Sokalan® CP 42	-	-	8.5	0.1 M	Tris-HCl
2-45	25 % w/v	Sokalan® HP 56	0.2 M	ammonium acetate	7	0.1 M	HEPES-NaOH
2-46	25 % v/v	Sokalan® CP 5	-	-	8.5	0.1 M	Tris-HCl
2-47	10 % w/v	poly(vinyl pyrrolidone) K15	0.2 M	ammonium formate	-	-	-
	20 % w/v	PEG 4000	-	-	-	-	-
2-48	15 % w/v	poly(vinyl pyrrolidone) K15	-	-	8	0.1 M	Tris-HCl
	25 % w/v	PEG MME 5000	-	-	-	-	-

Figure S5. The Midas composition table.

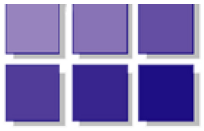
Table 5- MD1-46 (MD1-47 HT) Morpheus™ Conditions 1-48 (A1-D12)

Well	Tube	Ligand stock	Conc.	buffer stock	Conc.	pH @ (20°C)	precipitant stock	Conc.
A1	1-1	Divalents	0.06M	Buffer 1	0.1M	6.5	P550MME_P20K	30.0%
A2	1-2	Divalents	0.06M	Buffer 1	0.1M	6.5	EDO_P8K	30.0%
A3	1-3	Divalents	0.06M	Buffer 1	0.1M	6.5	GOL_P4K	30.0%
A4	1-4	Divalents	0.06M	Buffer 1	0.1M	6.5	MPD_P1K_P3350	37.5%
A5	1-5	Divalents	0.06M	Buffer 2	0.1M	7.5	P550MME_P20K	30.0%
A6	1-6	Divalents	0.06M	Buffer 2	0.1M	7.5	EDO_P8K	30.0%
A7	1-7	Divalents	0.06M	Buffer 2	0.1M	7.5	GOL_P4K	30.0%
A8	1-8	Divalents	0.06M	Buffer 2	0.1M	7.5	MPD_P1K_P3350	37.5%
A9	1-9	Divalents	0.06M	Buffer 3	0.1M	8.5	P550MME_P20K	30.0%
A10	1-10	Divalents	0.06M	Buffer 3	0.1M	8.5	EDO_P8K	30.0%
A11	1-11	Divalents	0.06M	Buffer 3	0.1M	8.5	GOL_P4K	30.0%
A12	1-12	Divalents	0.06M	Buffer 3	0.1M	8.5	MPD_P1K_P3350	37.5%
B1	1-13	Halogens	0.09M	Buffer 1	0.1M	6.5	P550MME_P20K	30.0%
B2	1-14	Halogens	0.09M	Buffer 1	0.1M	6.5	EDO_P8K	30.0%
B3	1-15	Halogens	0.09M	Buffer 1	0.1M	6.5	GOL_P4K	30.0%
B4	1-16	Halogens	0.09M	Buffer 1	0.1M	6.5	MPD_P1K_P3350	37.5%
B5	1-17	Halogens	0.09M	Buffer 2	0.1M	7.5	P550MME_P20K	30.0%
B6	1-18	Halogens	0.09M	Buffer 2	0.1M	7.5	EDO_P8K	30.0%
B7	1-19	Halogens	0.09M	Buffer 2	0.1M	7.5	GOL_P4K	30.0%
B8	1-20	Halogens	0.09M	Buffer 2	0.1M	7.5	MPD_P1K_P3350	37.5%
B9	1-21	Halogens	0.09M	Buffer 3	0.1M	8.5	P550MME_P20K	30.0%
B10	1-22	Halogens	0.09M	Buffer 3	0.1M	8.5	EDO_P8K	30.0%
B11	1-23	Halogens	0.09M	Buffer 3	0.1M	8.5	GOL_P4K	30.0%
B12	1-24	Halogens	0.09M	Buffer 3	0.1M	8.5	MPD_P1K_P3350	37.5%
C1	1-25	NPS	0.09M	Buffer 1	0.1M	6.5	P550MME_P20K	30.0%
C2	1-26	NPS	0.09M	Buffer 1	0.1M	6.5	EDO_P8K	30.0%
C3	1-27	NPS	0.09M	Buffer 1	0.1M	6.5	GOL_P4K	30.0%
C4	1-28	NPS	0.09M	Buffer 1	0.1M	6.5	MPD_P1K_P3350	37.5%
C5	1-29	NPS	0.09M	Buffer 2	0.1M	7.5	P550MME_P20K	30.0%
C6	1-30	NPS	0.09M	Buffer 2	0.1M	7.5	EDO_P8K	30.0%
C7	1-31	NPS	0.09M	Buffer 2	0.1M	7.5	GOL_P4K	30.0%
C8	1-32	NPS	0.09M	Buffer 2	0.1M	7.5	MPD_P1K_P3350	37.5%
C9	1-33	NPS	0.09M	Buffer 3	0.1M	8.5	P550MME_P20K	30.0%
C10	1-34	NPS	0.09M	Buffer 3	0.1M	8.5	EDO_P8K	30.0%
C11	1-35	NPS	0.09M	Buffer 3	0.1M	8.5	GOL_P4K	30.0%
C12	1-36	NPS	0.09M	Buffer 3	0.1M	8.5	MPD_P1K_P3350	37.5%
D1	1-37	Alcohols	0.12M	Buffer 1	0.1M	6.5	P550MME_P20K	30.0%
D2	1-38	Alcohols	0.12M	Buffer 1	0.1M	6.5	EDO_P8K	30.0%
D3	1-39	Alcohols	0.12M	Buffer 1	0.1M	6.5	GOL_P4K	30.0%
D4	1-40	Alcohols	0.12M	Buffer 1	0.1M	6.5	MPD_P1K_P3350	37.5%
D5	1-41	Alcohols	0.12M	Buffer 2	0.1M	7.5	P550MME_P20K	30.0%
D6	1-42	Alcohols	0.12M	Buffer 2	0.1M	7.5	EDO_P8K	30.0%
D7	1-43	Alcohols	0.12M	Buffer 2	0.1M	7.5	GOL_P4K	30.0%
D8	1-44	Alcohols	0.12M	Buffer 2	0.1M	7.5	MPD_P1K_P3350	37.5%
D9	1-45	Alcohols	0.12M	Buffer 3	0.1M	8.5	P550MME_P20K	30.0%
D10	1-46	Alcohols	0.12M	Buffer 3	0.1M	8.5	EDO_P8K	30.0%
D11	1-47	Alcohols	0.12M	Buffer 3	0.1M	8.5	GOL_P4K	30.0%
D12	1-48	Alcohols	0.12M	Buffer 3	0.1M	8.5	MPD_P1K_P3350	37.5%

Table 5 contd. - MD1-46 (MD1-47 HT) Morpheus™ Conditions 49-96 (E1-H12)

Well	Tube	Ligand stock	Conc.	Buffer stock	Conc.	pH @ (20°C)	Precipitant stock	Conc.
E1	2-1	Ethylene Glycols	0.12M	Buffer 1	0.1M	6.5	P550MME_P20K	30.0%
E2	2-2	Ethylene Glycols	0.12M	Buffer 1	0.1M	6.5	EDO_P8K	30.0%
E3	2-3	Ethylene Glycols	0.12M	Buffer 1	0.1M	6.5	GOL_P4K	30.0%
E4	2-4	Ethylene Glycols	0.12M	Buffer 1	0.1M	6.5	MPD_P1K_P3350	37.5%
E5	2-5	Ethylene Glycols	0.12M	Buffer 2	0.1M	7.5	P550MME_P20K	30.0%
E6	2-6	Ethylene Glycols	0.12M	Buffer 2	0.1M	7.5	EDO_P8K	30.0%
E7	2-7	Ethylene Glycols	0.12M	Buffer 2	0.1M	7.5	GOL_P4K	30.0%
E8	2-8	Ethylene Glycols	0.12M	Buffer 2	0.1M	7.5	MPD_P1K_P3350	37.5%
E9	2-9	Ethylene Glycols	0.12M	Buffer 3	0.1M	8.5	P550MME_P20K	30.0%
E10	2-10	Ethylene Glycols	0.12M	Buffer 3	0.1M	8.5	EDO_P8K	30.0%
E11	2-11	Ethylene Glycols	0.12M	Buffer 3	0.1M	8.5	GOL_P4K	30.0%
E12	2-12	Ethylene Glycols	0.12M	Buffer 3	0.1M	8.5	MPD_P1K_P3350	37.5%
F1	2-13	Monosaccharides	0.12M	Buffer 1	0.1M	6.5	P550MME_P20K	30.0%
F2	2-14	Monosaccharides	0.12M	Buffer 1	0.1M	6.5	EDO_P8K	30.0%
F3	2-15	Monosaccharides	0.12M	Buffer 1	0.1M	6.5	GOL_P4K	30.0%
F4	2-16	Monosaccharides	0.12M	Buffer 1	0.1M	6.5	MPD_P1K_P3350	37.5%
F5	2-17	Monosaccharides	0.12M	Buffer 2	0.1M	7.5	P550MME_P20K	30.0%
F6	2-18	Monosaccharides	0.12M	Buffer 2	0.1M	7.5	EDO_P8K	30.0%
F7	2-19	Monosaccharides	0.12M	Buffer 2	0.1M	7.5	GOL_P4K	30.0%
F8	2-20	Monosaccharides	0.12M	Buffer 2	0.1M	7.5	MPD_P1K_P3350	37.5%
F9	2-21	Monosaccharides	0.12M	Buffer 3	0.1M	8.5	P550MME_P20K	30.0%
F10	2-22	Monosaccharides	0.12M	Buffer 3	0.1M	8.5	EDO_P8K	30.0%
F11	2-23	Monosaccharides	0.12M	Buffer 3	0.1M	8.5	GOL_P4K	30.0%
F12	2-24	Monosaccharides	0.12M	Buffer 3	0.1M	8.5	MPD_P1K_P3350	37.5%
G1	2-25	Carboxylic acids	0.10M	Buffer 1	0.1M	6.5	P550MME_P20K	30.0%
G2	2-26	Carboxylic acids	0.10M	Buffer 1	0.1M	6.5	EDO_P8K	30.0%
G3	2-27	Carboxylic acids	0.10M	Buffer 1	0.1M	6.5	GOL_P4K	30.0%
G4	2-28	Carboxylic acids	0.10M	Buffer 1	0.1M	6.5	MPD_P1K_P3350	37.5%
G5	2-29	Carboxylic acids	0.10M	Buffer 2	0.1M	7.5	P550MME_P20K	30.0%
G6	2-30	Carboxylic acids	0.10M	Buffer 2	0.1M	7.5	EDO_P8K	30.0%
G7	2-31	Carboxylic acids	0.10M	Buffer 2	0.1M	7.5	GOL_P4K	30.0%
G8	2-32	Carboxylic acids	0.10M	Buffer 2	0.1M	7.5	MPD_P1K_P3350	37.5%
G9	2-33	Carboxylic acids	0.10M	Buffer 3	0.1M	8.5	P550MME_P20K	30.0%
G10	2-34	Carboxylic acids	0.10M	Buffer 3	0.1M	8.5	EDO_P8K	30.0%
G11	2-35	Carboxylic acids	0.10M	Buffer 3	0.1M	8.5	GOL_P4K	30.0%
G12	2-36	Carboxylic acids	0.10M	Buffer 3	0.1M	8.5	MPD_P1K_P3350	37.5%
H1	2-37	Amino acids	0.10M	Buffer 1	0.1M	6.5	P550MME_P20K	30.0%
H2	2-38	Amino acids	0.10M	Buffer 1	0.1M	6.5	EDO_P8K	30.0%
H3	2-39	Amino acids	0.10M	Buffer 1	0.1M	6.5	GOL_P4K	30.0%
H4	2-40	Amino acids	0.10M	Buffer 1	0.1M	6.5	MPD_P1K_P3350	37.5%
H5	2-41	Amino acids	0.10M	Buffer 2	0.1M	7.5	P550MME_P20K	30.0%
H6	2-42	Amino acids	0.10M	Buffer 2	0.1M	7.5	EDO_P8K	30.0%
H7	2-43	Amino acids	0.10M	Buffer 2	0.1M	7.5	GOL_P4K	30.0%
H8	2-44	Amino acids	0.10M	Buffer 2	0.1M	7.5	MPD_P1K_P3350	37.5%
H9	2-45	Amino acids	0.10M	Buffer 3	0.1M	8.5	P550MME_P20K	30.0%
H10	2-46	Amino acids	0.10M	Buffer 3	0.1M	8.5	EDO_P8K	30.0%
H11	2-47	Amino acids	0.10M	Buffer 3	0.1M	8.5	GOL_P4K	30.0%
H12	2-48	Amino acids	0.10M	Buffer 3	0.1M	8.5	MPD_P1K_P3350	37.5%

Figure S6. The Morpheus composition table.



moleculardimensions.com

PACT *premier*™ **MD1-29**

Tube No.	Buffer/Salt	Buffer/Salt	pH	Precipitant
P1-1	0.1 M SPG buffer	None	4.0	25 % w/v PEG 1500
P1-2	0.1 M SPG buffer	None	5.0	25 % w/v PEG 1500
P1-3	0.1 M SPG buffer	None	6.0	25 % w/v PEG 1500
P1-4	0.1 M SPG buffer	None	7.0	25 % w/v PEG 1500
P1-5	0.1 M SPG buffer	None	8.0	25 % w/v PEG 1500
P1-6	0.1 M SPG buffer	None	9.0	25 % w/v PEG 1500
P1-7	0.2 M sodium chloride	0.1 M sodium acetate	5.0	20 % w/v PEG 6000
P1-8	0.2 M ammonium chloride	0.1 M sodium acetate	5.0	20 % w/v PEG 6000
P1-9	0.2 M lithium chloride	0.1 M sodium acetate	5.0	20 % w/v PEG 6000
P1-10	0.2 M magnesium chloride	0.1 M sodium acetate	5.0	20 % w/v PEG 6000
P1-11	0.2 M calcium chloride	0.1 M sodium acetate	5.0	20 % w/v PEG 6000
P1-12	0.01 M zinc chloride	0.1 M sodium acetate	5.0	20 % w/v PEG 6000
P1-13	0.1 M MIB buffer	None	4.0	25 % w/v PEG 1500
P1-14	0.1 M MIB buffer	None	5.0	25 % w/v PEG 1500
P1-15	0.1 M MIB buffer	None	6.0	25 % w/v PEG 1500
P1-16	0.1 M MIB buffer	None	7.0	25 % w/v PEG 1500
P1-17	0.1 M MIB buffer	None	8.0	25 % w/v PEG 1500
P1-18	0.1 M MIB buffer	None	9.0	25 % w/v PEG 1500
P1-19	0.2 M sodium chloride	0.1 M MES	6.0	20 % w/v PEG 6000
P1-20	0.2 M ammonium chloride	0.1 M MES	6.0	20 % w/v PEG 6000
P1-21	0.2 M lithium chloride	0.1 M MES	6.0	20 % w/v PEG 6000
P1-22	0.2 M magnesium chloride	0.1 M MES	6.0	20 % w/v PEG 6000
P1-23	0.2 M calcium chloride	0.1 M MES	6.0	20 % w/v PEG 6000
P1-24	0.01 M zinc chloride	0.1 M MES	6.0	20 % w/v PEG 6000
P1-25	0.1 M PCTP buffer	None	4.0	25 % w/v PEG 1500
P1-26	0.1 M PCTP buffer	None	5.0	25 % w/v PEG 1500
P1-27	0.1 M PCTP buffer	None	6.0	25 % w/v PEG 1500
P1-28	0.1 M PCTP buffer	None	7.0	25 % w/v PEG 1500
P1-29	0.1 M PCTP buffer	None	8.0	25 % w/v PEG 1500
P1-30	0.1 M PCTP buffer	None	9.0	25 % w/v PEG 1500
P1-31	0.2 M sodium chloride	0.1 M HEPES	7.0	20 % w/v PEG 6000
P1-32	0.2 M ammonium chloride	0.1 M HEPES	7.0	20 % w/v PEG 6000
P1-33	0.2 M lithium chloride	0.1 M HEPES	7.0	20 % w/v PEG 6000
P1-34	0.2 M magnesium chloride	0.1 M HEPES	7.0	20 % w/v PEG 6000
P1-35	0.2 M calcium chloride	0.1 M HEPES	7.0	20 % w/v PEG 6000
P1-36	0.01 M zinc chloride	0.1 M HEPES	7.0	20 % w/v PEG 6000
P1-37	0.1 M MMT buffer	None	4.0	25 % w/v PEG 1500
P1-38	0.1 M MMT buffer	None	5.0	25 % w/v PEG 1500
P1-39	0.1 M MMT buffer	None	6.0	25 % w/v PEG 1500
P1-40	0.1 M MMT buffer	None	7.0	25 % w/v PEG 1500
P1-41	0.1 M MMT buffer	None	8.0	25 % w/v PEG 1500
P1-42	0.1 M MMT buffer	None	9.0	25 % w/v PEG 1500
P1-43	0.2 M sodium chloride	0.1 M Tris	8.0	20 % w/v PEG 6000
P1-44	0.2 M ammonium chloride	0.1 M Tris	8.0	20 % w/v PEG 6000
P1-45	0.2 M lithium chloride	0.1 M Tris	8.0	20 % w/v PEG 6000
P1-46	0.2 M magnesium chloride	0.1 M Tris	8.0	20 % w/v PEG 6000
P1-47	0.2 M calcium chloride	0.1 M Tris	8.0	20 % w/v PEG 6000
P1-48	0.002 M zinc chloride	0.1 M Tris	8.0	20 % w/v PEG 6000

moleculardimensions.com

PACT *premier*™ **MD1-29**

Tube No.	Salt	Buffer System	pH	Precipitant
P2-1	0.2 M sodium fluoride	None		20 % w/v PEG 3350
P2-2	0.2 M sodium bromide	None		20 % w/v PEG 3350
P2-3	0.2 M sodium iodide	None		20 % w/v PEG 3350
P2-4	0.2 M potassium thiocyanate	None		20 % w/v PEG 3350
P2-5	0.2 M sodium nitrate	None		20 % w/v PEG 3350
P2-6	0.2 M sodium formate	None		20 % w/v PEG 3350
P2-7	0.2 M sodium acetate	None		20 % w/v PEG 3350
P2-8	0.2 M sodium sulfate	None		20 % w/v PEG 3350
P2-9	0.2 M potassium/sodium tartrate	None		20 % w/v PEG 3350
P2-10	0.02 M sodium/potassium phosphate	None		20 % w/v PEG 3350
P2-11	0.2 M sodium citrate	None		20 % w/v PEG 3350
P2-12	0.2 M sodium malonate	None		20 % w/v PEG 3350
P2-13	0.2 M sodium fluoride	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-14	0.2 M sodium bromide	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-15	0.2 M sodium iodide	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-16	0.2 M potassium thiocyanate	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-17	0.2 M sodium nitrate	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-18	0.2 M sodium formate	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-19	0.2 M sodium acetate	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-20	0.2 M sodium sulfate	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-21	0.2 M potassium/sodium tartrate	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-22	0.02 M sodium/potassium phosphate	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-23	0.2 M sodium citrate	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-24	0.2 M sodium malonate	0.1 M Bis Tris propane	6.5	20 % w/v PEG 3350
P2-25	0.2 M sodium fluoride	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-26	0.2 M sodium bromide	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-27	0.2 M sodium iodide	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-28	0.2 M potassium thiocyanate	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-29	0.2 M sodium nitrate	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-30	0.2 M sodium formate	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-31	0.2 M sodium acetate	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-32	0.2 M sodium sulfate	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-33	0.2 M potassium/sodium tartrate	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-34	0.02 M sodium/potassium phosphate	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-35	0.2 M sodium citrate	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-36	0.2 M sodium malonate	0.1 M Bis Tris propane	7.5	20 % w/v PEG 3350
P2-37	0.2 M sodium fluoride	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-38	0.2 M sodium bromide	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-39	0.2 M sodium iodide	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-40	0.2 M potassium thiocyanate	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-41	0.2 M sodium nitrate	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-42	0.2 M sodium formate	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-43	0.2 M sodium acetate	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-44	0.2 M sodium sulfate	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-45	0.2 M potassium/sodium tartrate	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-46	0.02 M sodium/potassium phosphate	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-47	0.2 M sodium citrate	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350
P2-48	0.2 M sodium malonate	0.1 M Bis Tris propane	8.5	20 % w/v PEG 3350

Figure S7. The PACT composition table.

JCSG-*plus* HT-96 **Rows A- D** **MD1-40**

Tube #	Salt	Buffer	pH	Precipitant
A1	0.2 M lithium sulfate	0.1 M sodium acetate	4.5	50 % v/v PEG 400
A2	None	0.1 M Citrate	5.5	20 % w/v PEG 3K
A3	0.2 M di-ammonium hydrogen citrate	None	-	20 % w/v PEG 3350
A4	0.02 M calcium chloride	0.1 M Sodium acetate	4.6	30 % v/v MPD
A5	0.2 M magnesium formate	None	-	20 % w/v PEG 3350
A6	0.2 M lithium sulfate	0.1 M phosphate/citrate	4.2	20 % w/v PEG 1K
A7	None	0.1 M CHES	9.5	20 % w/v PEG 8K
A8	0.2 M ammonium formate	None	-	20 % w/v PEG 3350
A9	0.2 M ammonium chloride	None	-	20 % w/v PEG 3350
A10	0.2 M potassium formate	None	-	20 % w/v PEG 3350
A11	0.2 M ammonium dihydrogen phosphate	0.1 M Tris	8.5	50 % v/v MPD
A12	0.2 M potassium nitrate	None	-	20 % w/v PEG 3350
B1	None	0.1 M Na Citrate	4.0	0.8 M ammonium sulfate
B2	0.2 M sodium thiocyanate	None	-	20 % w/v PEG 3350
B3	None	0.1 M Bicine	9.0	20 % w/v PEG 6K
B4	None	0.1 M HEPES	7.5	10 % w/v PEG 8K/ 8 % v/v Ethylene glycol
B5	None	0.1 M sodium cacodylate	6.5	40 % v/v MPD/ 5 % w/v PEG 8000
B6	None	0.1 M phosphate/citrate	4.2	40 % v/v Ethanol/ 5 % w/v PEG 1000
B7	None	0.1 M sodium acetate	4.6	8 % w/v PEG 4K
B8	0.2 M magnesium chloride	0.1 M Tris	7.0	10 % w/v PEG 8K
B9	None	0.1 M Na citrate	5.0	20 % w/v PEG 6K
B10	0.2 M magnesium chloride	0.1 M sodium cacodylate	6.5	50 % v/v PEG 200
B11	None	None	6.5	1.6 M tri-sodium citrate
B12	0.2 M tri-potassium citrate	None	-	20 % w/v PEG 3350
C1	0.2 M sodium chloride	0.1 M phosphate/citrate	4.2	20 % w/v PEG 8K
C2	1.0 M lithium chloride	0.1 M Na citrate	4.0	20 % w/v PEG 6K
C3	0.2 M ammonium nitrate	None	-	20 % w/v PEG 3350
C4	None	0.1 M HEPES	7.0	10 % w/v PEG 6K
C5	None	0.1 M Na HEPES	7.5	0.8 M sodium dihydrogen phosphate 0.8 M potassium dihydrogen phosphate
C6	None	0.1 M phosphate/citrate	4.2	40 % v/v PEG 300
C7	0.2 M zinc acetate	0.1 M sodium acetate	4.5	10 % w/v PEG 3K
C8	None	0.1 M Tris	8.5	20 % v/v Ethanol
C9	None	0.1 M Na/K phosphate	6.2	25 % v/v 1,2-propanediol 10 % v/v Glycerol
C10	None	0.1 M Bicine	9.0	10 % w/v PEG 20,000/ 2% v/v Dioxane
C11	None	0.1 M sodium acetate	4.6	2.0 M ammonium sulfate
C12	None	None	-	10 % w/v PEG 1000/ 10 % w/v PEG 8000
D1	None	None	-	24 % w/v PEG 1500/ 20 % v/v Glycerol
D2	0.2 M magnesium chloride	0.1 M Na HEPES	7.5	30 % v/v PEG 400
D3	0.2 M sodium chloride	0.1 M Na/K phosphate	6.2	50 % v/v PEG 200
D4	0.2 M lithium sulfate	0.1 M sodium acetate	4.5	30 % w/v PEG 8K
D5	None	0.1 M HEPES	7.5	70 % v/v MPD
D6	0.2 M magnesium chloride	0.1 M Tris	8.5	20 % w/v PEG 8K
D7	0.2 M lithium sulfate	0.1 M Tris	8.5	40 % v/v PEG 400
D8	None	0.1 M Tris	8.0	40 % v/v MPD
D9	0.17 M ammonium sulfate	None	-	25.5 % w/v PEG 4K/ 15 % v/v Glycerol
D10	0.2 M calcium acetate	0.1 M Sodium cacodylate	6.5	40 % v/v PEG 300
D11	0.14 M calcium chloride	0.07 M Sodium acetate	4.6	14 % v/v 2-propanol/ 30 % v/v Glycerol
D12	0.04 M potassium dihydrogen phosphate	None	-	16 % w/v PEG 8K/ 20 % v/v Glycerol

JCSG-*plus* HT-96 **Rows E- H** **MD1-40**

Tube	Salt	Buffer	pH	Precipitant
E1	None	0.1 M sodium cacodylate	6.5	1.0 M tri-sodium citrate
E2	0.2 M sodium chloride	0.1 M sodium cacodylate	6.5	2.0 M ammonium sulfate
E3	0.2 M sodium chloride	0.1 M HEPES	7.5	10 % v/v 2-propanol
E4	0.2 M lithium sulfate	0.1 M Tris	8.5	1.26 M ammonium sulfate
E5	None	0.1 M CAPS	10.5	40 % v/v MPD
E6	0.2 M zinc acetate	0.1 M imidazole	8.0	20 % w/v PEG 3K
E7	0.2 M zinc acetate	0.1 M sodium cacodylate	6.5	10 % v/v 2-propanol
E8	None	0.1 M sodium acetate	4.5	1.0 M di-ammonium hydrogen phosphate
E9	None	0.1 M MES	6.5	1.6 M magnesium sulfate
E10	None	0.1 M Bicine	9.0	10 % w/v PEG 6K
E11	0.16 M calcium acetate	0.08 M sodium cacodylate	6.5	14.4 % w/v PEG 8K/ 20 % v/v glycerol
E12	None	0.1 M imidazole	8.0	10 % w/v PEG 8K
F1	0.05 M caesium chloride	0.1 M MES	6.5	30 % v/v Jeffamine M-600
F2	None	0.1 M Na Citrate	5.0	3.2 M ammonium sulfate
F3	None	0.1 M Tris	8.0	20 % v/v MPD
F4	None	0.1 M HEPES	7.5	20 % v/v Jeffamine M-600
F5	0.2 M magnesium chloride	0.1 M Tris	8.5	50 % v/v ethylene glycol
F6	None	0.1 M Bicine	9.0	10 % v/v MPD
F7	None	None	7.0	0.8 M succinic acid
F8	None	None	7.0	2.1 M DL-malic acid
F9	None	None	7.0	2.4 M sodium malonate
F10	1.1 M sodium malonate	0.1 M HEPES	7.0	0.5 % v/v Jeffamine ED-2001
F11	1.0 M succinic acid	0.1 M HEPES	7.0	1 % w/v PEG 2000 MME
F12	None	0.1 M HEPES	7.0	30 % v/v Jeffamine M-600
G1	None	0.1 M HEPES	7.0	30 % v/v Jeffamine ED-2001
G2	0.02 M magnesium chloride	0.1 M HEPES	7.5	22 % w/v polyacrylic acid 5100 sodium salt
G3	0.01 M cobalt chloride	0.1 M Tris	8.5	20 % w/v polyvinylpyrrolidone K15
G4	0.2 M tri-methylamine N-oxide	0.1 M Tris	8.5	20 % w/v PEG 2K MME
G5	0.005 M cobalt chloride 0.005 M cadmium chloride 0.005 M magnesium chloride 0.005 M nickel chloride	0.1 M HEPES	7.5	12 % w/v PEG 3350
G6	0.2 M sodium malonate	None	7.0	20 % w/v PEG 3350
G7	0.1 M succinic acid	None	7.0	15 % w/v PEG 3350
G8	0.15 M DL- malic acid	None	7.0	20 % w/v PEG 3350
G9	0.1 M potassium thiocyanate	None	None	30 % w/v PEG 2K MME
G10	0.15 M potassium bromide	None	None	30 % w/v PEG 2K MME
G11	None	0.1 M Bis Tris	5.5	2.0 M ammonium sulfate
G12	None	0.1 M Bis Tris	5.5	3.0 M sodium chloride
H1	None	0.1 M Bis Tris	5.5	0.3 M magnesium formate
H2	1.0 M ammonium sulfate	0.1 M Bis Tris	5.5	1 % w/v PEG 3350
H3	None	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
H4	0.2 M calcium chloride	0.1 M Bis Tris	5.5	45 % v/v MPD
H5	0.2 M ammonium acetate	0.1 M Bis Tris	5.5	45 % v/v MPD
H6	0.1 M ammonium acetate	0.1 M Bis Tris	5.5	17 % w/v PEG 10K
H7	0.2 M ammonium sulfate	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
H8	0.2 M sodium chloride	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
H9	0.2 M lithium sulfate	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
H10	0.2 M ammonium acetate	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
H11	0.2 M magnesium chloride	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
H12	0.2 M ammonium acetate	0.1 M HEPES	7.5	45 % v/v MPD

Figure S8. The JCSG composition table.

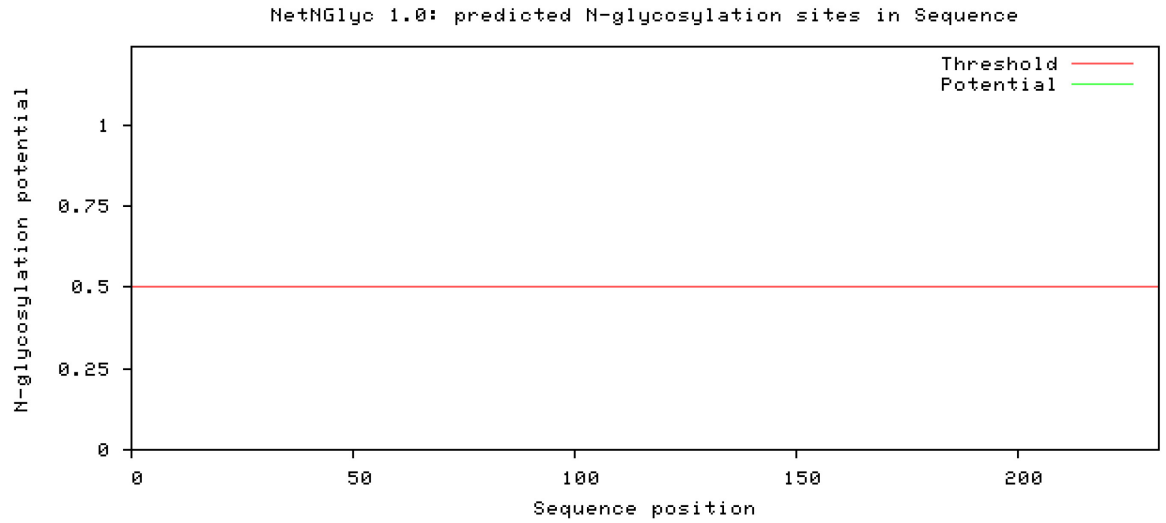
MemGold™ Conditions 1-48 (Box 1) MD1–39

Tube #	Conc.	Salt	Conc.	Buffer	pH	Conc.	Precipitant
1-1	2.2 M	Ammonium sulfate	0.08 M	Sodium citrate	5.2		
1-2	1.2 M	Sodium citrate tribasic dihydrate	0.01 M	Tris	8.0		
1-3			0.015 M	Tricine	8.5	24 % w/v	PEG 4000
1-4	0.36 M	Sodium chloride	0.015 M	Sodium phosphate	7.0	9.9 % w/v	PEG 4000
						0.1 % w/v	Sodium azide
1-5	0.3 M	Sodium chloride	0.01 M	Tris	8.0	27.5 % w/v	PEG 4000
1-6			0.225 M	MES/Bis-Tris	6.6	6.6 % w/v	PEG 6000
1-7	0.1 M	Ammonium sulfate	0.1 M	HEPES	7.5	12 % w/v	PEG 4000
						22 % v/v	Glycerol
1-8	0.01 M	Magnesium sulfate heptahydrate	0.02 M	MES	6.5	7.7 % w/v	PEG 1500
	0.02 M	Sodium chloride					
	0.02 M	Calcium chloride dihydrate					
1-9	2.5 M	Ammonium sulfate	0.05 M	HEPES	7.5		
1-10	1.1 M	Sodium citrate tribasic dihydrate	0.0665 M	HEPES	7.5		
1-11	3.3 M	Ammonium sulfate	0.15 M	Potassium phosphate	6.5		
1-12	0.1 M	Magnesium acetate tetrahydrate	0.1 M	Sodium citrate	5.8	14 % w/v	PEG 5000 MME
1-13	0.1 M	Soduim chloride	0.02 M	Sodium citrate	5.6	11 % w/v	PEG 3350
1-14	0.1 M	Sodium chloride	0.02 M	Sodium citrate	5.6	5.5 % w/v	PEG 3350
1-15	0.05 M	Calcium chloride dihydrate	0.1 M	Tris	8.2	32 % v/v	PEG 400
	0.05 M	Barium chloride dihydrate					
1-16	0.05 M	Sodium chloride	0.1 M	Sodium phosphate	6.2	16 % w/v	PEG 4000
1-17	0.1 M	Magnesium chloride hexahydrate	0.03 M	Tris-HCl	8.2	19 % w/v	PEG 4000
1-18	0.2 M	Sodium chloride	0.025 M	HEPES	7.5	13 % w/v	PEG 4000
1-19			0.1 M	HEPES	7.5	11 % w/v	PEG 3350
1-20	0.1 M	Sodium chloride	0.02 M	KMES	6.7	6.6 % w/v	PEG 4000
1-21	0.1 M	Potassium chloride	0.02 M	Tris	7.0	20 % w/v	PEG 4000
1-22	0.05 M	Magnesium chloride hexahydrate	0.1 M	Sodium cacodylate	6.7	6.6 % w/v	PEG 3350
						0.1 % w/v	Sodium azide
1-23	0.2 M	Potassium chloride	0.1 M	Sodium citrate	5.5	37 % v/v	Pentaerythritol propoxylate (5/4 PO/OH)
1-24			0.1 M	Tris	8.0	5.5 % w/v	PEG 4000
1-25	0.1 M	Sodium chloride	0.02 M	Tris	7.0	7.7 % w/v	PEG 4000
1-26	0.1 M	Magnesium chloride hexahydrate	0.1 M	Tris	7.5	22 % v/v	PEG 400
1-27	0.04 M	Sodium chloride	0.04 M	Tris	8.0	27 % v/v	PEG 350 MME
1-28	0.05 M	Sodium chloride	0.1 M	Sodium citrate	6.0	22 % v/v	PEG 400
	0.02 M	Magnesium chloride hexahydrate					
1-29			0.1 M	Sodium acetate	5.5	8.8 % w/v	PEG 2000 MME
1-30			0.4 M	Ammonium acetate	8.0	13 % w/v	PEG 2000 MME
1-31			0.02 M	Bis-Tris	7.0	15 % w/v	PEG 2000
1-32	0.1 M	Sodium chloride	0.02 M	Tris	7.5	11 % w/v	PEG 1500
	0.1 M	Magnesium chloride hexahydrate					
1-33	0.1 M	Sodium chloride	0.1 M	HEPES	8.0	11 % w/v	PEG 1500
	0.1 M	Magnesium chloride hexahydrate					
1-34	0.2 M	Sodium acetate trihydrate	0.1 M	HEPES	7.0	22 % w/v	PEG 3000
	0.2 M	Potassium chloride					
1-35	0.02 M	Nickel(II) sulfate hexahydrate	0.01 M	HEPES	7.0	33 % v/v	Jeffamine® M-600
1-36	0.15 M	Sodium chloride	0.1 M	Tris	8.0	13 % w/v	PEG 6000
1-37	0.2 M	Calcium chloride dihydrate	0.1 M	HEPES	7.5	53 % v/v	PEG 400
1-38	0.05 M	Magnesium acetate tetrahydrate	0.05 M	Sodium acetate	5.0	28 % v/v	PEG 400
1-39			0.05 M	HEPES	7.5	22 % w/v	PEG 4000
1-40	0.2 M	Calcium chloride dihydrate	0.1 M	Tris-HCl	8.0	44 % v/v	PEG 400
1-41	0.05 M	Magnesium acetate tetrahydrate	0.05 M	Sodium acetate	5.4	24 % v/v	PEG 400
1-42	0.2 M	Calcium chloride dihydrate	0.1 M	MES	6.5	26 % v/v	PEG 350 MME
1-43	0.1 M	Potassium chloride	0.1 M	Tris	8.5	39 % v/v	PEG 400
1-44	0.05 M	Magnesium chloride hexahydrate	0.1 M	Glycine	9.0	22 % v/v	PEG 400
1-45	0.1 M	Ammonium sulfate	0.1 M	Glycine	3.8	28 % v/v	Triethylene glycol (TEG)
1-46	0.15 M	Sodium formate	0.1 M	HEPES	7.2	18 % w/v	PEG 3350
1-47			0.2 M	Sodium acetate	6.8	8.8 % w/v	PEG 6000
1-48	0.2 M	Potassium chloride	0.1 M	MES	6.5	18 % w/v	PEG 6000

MemGold™ Conditions 1-48 (Box 2) MD1–39

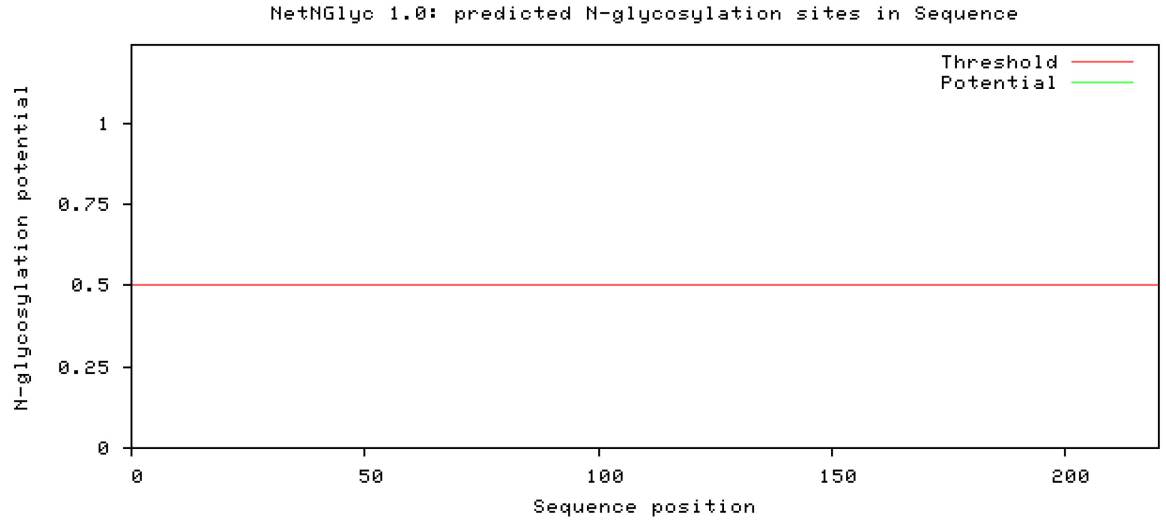
Tube #	Conc.	Salt	Conc.	Buffer	pH	Conc.	Precipitant
2-1	0.22 M	Sodium citrate tribasic dihydrate	0.1 M	Tris	8.0	35 % v/v	PEG 400
2-2			0.1 M	Sodium acetate	4.5	17 % v/v	PEG 400
2-3	1.0 M	Lithium sulfate	0.02 M	Tris	8.5	1.8 % w/v	PEG 8000
2-4			0.02 M	Tris	7.5	22 % v/v	PEG 500 MME
2-5	0.05 M	Sodium chloride	0.02 M	Glycine	10.0	33 % w/v	PEG 1000
2-6	0.2 M	Magnesium chloride hexahydrate	0.1 M	Tris	8.5	25 % w/v	PEG 4000
2-7	0.2 M	Magnesium chloride hexahydrate	0.1 M	Sodium cacodylate	6.5	31 % w/v	PEG 2000
2-8			0.64 M	Sodium acetate	4.6	18 % w/v	PEG 3350
2-9	0.1 M	Sodium chloride	0.1 M	Tris-HCl	8.0	33 % v/v	PEG 400
	0.1 M	Cadmium chloride hemi(pentahydrate)					
2-10			0.1 M	BICINE	8.9	31 % w/v	PEG 2000
2-11	0.05 M	Sodium sulfate	0.05 M	Tris	8.5	35 % v/v	PEG 400
	0.05 M	Lithium sulfate					
2-12	0.1 M	Sodium chloride	0.05 M	Glycine	9.5	33 % v/v	PEG 300
2-13	0.3 M	Magnesium nitrate hexahydrate	0.1 M	Tris	8.0	23 % w/v	PEG 2000
2-14	0.12 M	Lithium sulfate	0.02 M	Tris	7.5	20 % v/v	PEG 300
			0.1 M	Sodium citrate	5.0		
2-15	0.1 M	Sodium chloride	0.12 M	Tris	9.4	20 % v/v	PEG 400
2-16	0.2 M	Sodium chloride	0.1 M	HEPES	7.0	22 % v/v	PEG 500 MME
2-17	0.1 M	Sodium chloride	0.1 M	Tris	8.0	21 % v/v	PEG 400
	0.325 M	Sodium acetate trihydrate					
2-18	0.02 M	Sodium citrate tribasic dihydrate	0.08 M	Sodium phosphate	6.2	18 % w/v	PEG 2000
2-19	0.02 M	Potassium nitrate	0.03 M	Potassium citrate	6.5	7.7 % w/v	PEG 4000
2-20	0.1 M	Sodium chloride	0.1 M	Tris	8.5	30 % w/v	PEG 2000 MME
	0.005 M	Magnesium chloride hexahydrate					
2-21	0.2 M	Calcium chloride dihydrate	0.1 M	HEPES	7.0	33 % v/v	PEG 400
2-22	0.1 M	Calcium chloride dihydrate	0.1 M	Tris	6.5	13 % w/v	PEG 2000 MME
2-23	0.2 M	Ammonium sulfate	0.02 M	Sodium acetate	4.0	33 % v/v	PEG 200
	0.02 M	Sodium chloride					
2-24	0.07 M	Sodium chloride	0.05 M	Sodium citrate	4.5	22 % v/v	PEG 400
2-25	0.2 M	Ammonium sulfate	0.1 M	Sodium acetate	4.6	28 % v/v	PEG 500 MME
2-26			0.05 M	Glycine	9.0	55 % v/v	PEG 400
2-27	0.1 M	Magnesium chloride hexahydrate	0.1 M	Tris	8.5	33 % v/v	PEG 400
	0.1 M	Sodium chloride					
2-28	0.05 M	Citric acid				19 % w/v	PEG 1000
	0.1 M	Lithium sulfate					
	0.05 M	Sodium phosphate dibasic dihydrate					
2-29	0.2 M	Magnesium chloride hexahydrate	0.025 M	Sodium citrate	4.0	33 % v/v	PEG 400
	0.1 M	Potassium chloride					
2-30	0.05 M	Zinc acetate dihydrate	0.05 M	MES	6.1	11 % w/v	PEG 8000
2-31	0.3 M	Magnesium nitrate hexahydrate	0.1 M	Tris	8.0	22 % w/v	PEG 8000
2-32	0.1 M	Sodium chloride	0.1 M	MES	6.5	33 % v/v	PEG 400
						4 % v/v	Ethylene glycol
2-33	0.05 M	Sodium chloride	0.1 M	Sodium citrate	5.5	26 % v/v	PEG 400
2-34	0.1 M	Lithium sulfate	0.1 M	Glycine	9.3	30 % v/v	PEG 400
2-35	0.15 M	Potassium citrate tribasic monohydrate				22 % w/v	PEG 6000
	0.05 M	Lithium citrate tribasic tetrahydrate					
	0.1 M	Sodium phosphate monobasic monohydrate					
2-36	0.001 M	Zinc sulfate heptahydrate	0.05 M	HEPES	7.8	28 % v/v	PEG 600
2-37	0.1 M	Sodium chloride	0.1 M	Sodium phosphate	7.0	33 % v/v	PEG 300
2-38	0.1 M	Sodium chloride	0.05 M	BICINE	9.0	33 % v/v	PEG 300
2-39	0.05 M	Zinc acetate dihydrate	0.1 M	Sodium cacodylate	6.0	6.6 % w/v	PEG 8000
						6 % v/v	Ethylene glycol
2-40	0.2 M	Lithium sulfate	0.1 M	Sodium citrate	3.5	28 % v/v	PEG 400
2-41	0.1 M	Sodium chloride	0.1 M	Tris	7.5	11 % w/v	PEG 4000
2-42	0.05 M	Lithium sulfate	0.1 M	Tricine	7.4	7 % w/v	PEG 3000
2-43	0.2 M	Calcium chloride dihydrate	0.1 M	MES	6.5	33 % v/v	PEG 400
2-44	1.0 M	Sodium chloride	0.1 M	Sodium citrate	6.0	28 % w/v	PEG 4000
2-45			0.1 M	HEPES	7.5	11 % w/v	PEG 4000
2-46	0.002 M	Zinc sulfate heptahydrate	0.08 M	HEPES	7.0	25 % v/v	Jeffamine® ED-2003
2-47	0.001 M	Cadmium chloride hemi(pentahydrate)	0.1 M	MES	6.5	30 % v/v	PEG 400
	0.03 M	Magnesium chloride hexahydrate					
2-48	3.0 M	Sodium chloride	0.1 M	Bis-Tris propane	7.0		

Figure S9. The MemGold composition table.



EVKLLESGGGLVQPGGSLKLSCAASGFDFSTYWMSWVRQAPGKGLEWIGEINPDSSTINYTPSLKDKFIISRDNAKNTLY 80
LQMSKVRSEDSVLYYCVRPGTWGYFDYWGQGTTLTVSSASQSFPNVFPLVSCESPLSDKNLVAMGCLARDFLPSTISFTW 160
NYQNNTEVIQGIRTFPTLRTGGKYLATSQVLLSPKSILEGSDEYLVCKIHYGGKNRDLHVPIPAVAHHHHHH 240

Figure S10. The graph illustrating predicted *N*-glycosylation sites across the Mannitou Fab heavy chain. X-axis represents protein length from *N*- to C-terminal. A potential glycosylation position (green vertical lines) should cross the threshold (red horizontal line at 0.5). Asn-X-Ser/Thr sequons are shown in blue. Asparagines predicted to be *N*-glycosylated would be shown in red.



DVVVTQTPLSLPVSFGDQASISCRSSQSLVNSYGSTYLSWYLHRPGQSPQLLIYGISNRFSGVPDRFSGSGSGTDFTLTI 80
RTIKPEDLGMYYCLQGTHQPWTFGGGTKLEIKKRADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSE 160
RQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRKEC 240

Figure S11. The graph illustrating predicted *N*-glycosylation sites across the Mannitou light chain. X-axis represents protein length from *N*- to C-terminal. A potential glycosylation position (green vertical lines) should cross the threshold (red horizontal line at 0.5). Asn-X-Ser/Thr sequons would be shown in blue. Asparagines predicted to be *N*-glycosylated would be shown in red.

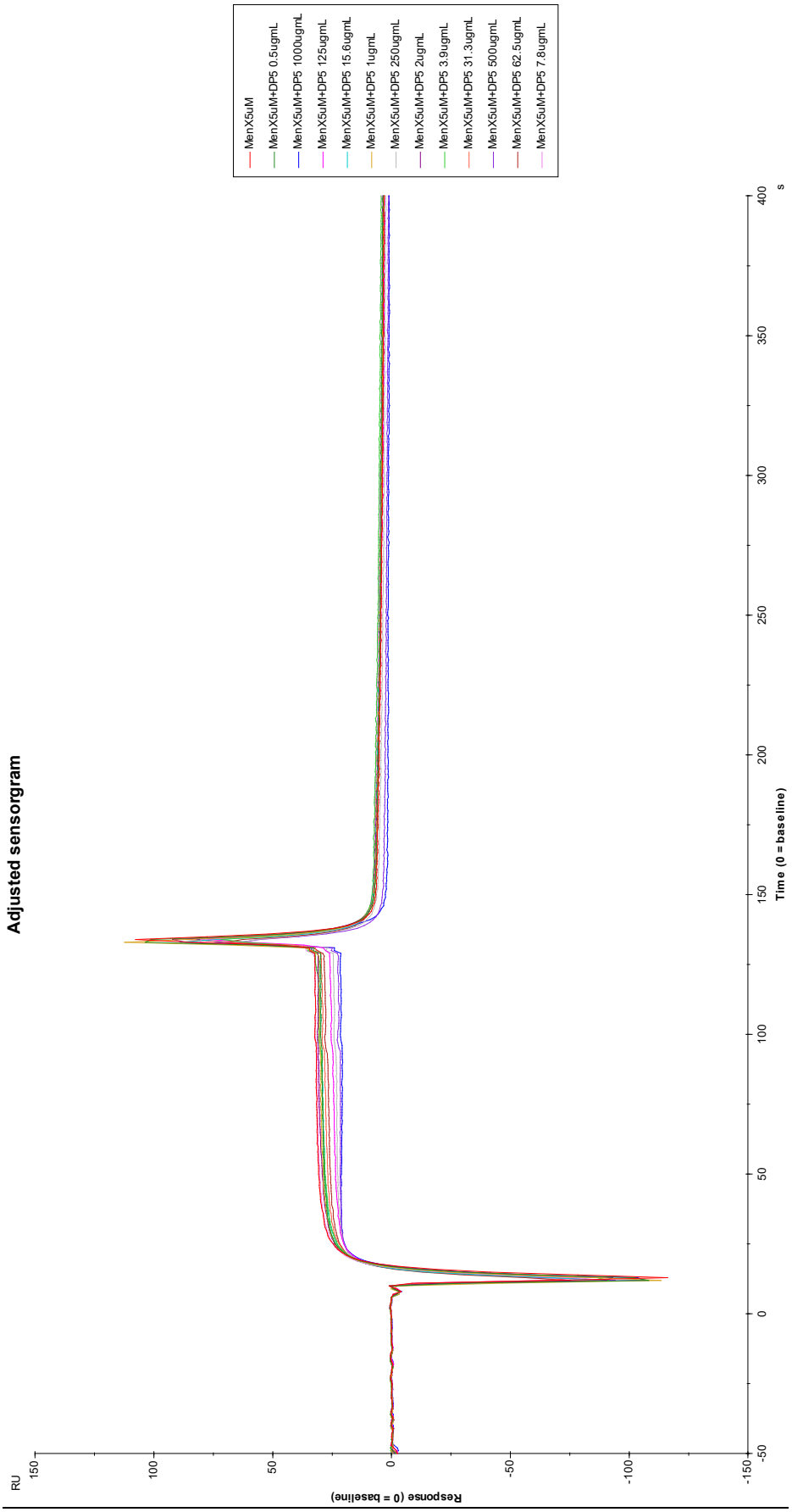


Figure S12. Inhibition assay sensorgram showing competition between DP5 and the surface-attached ligand (MenX(DP15)–ADH–CRM$_{197}$).

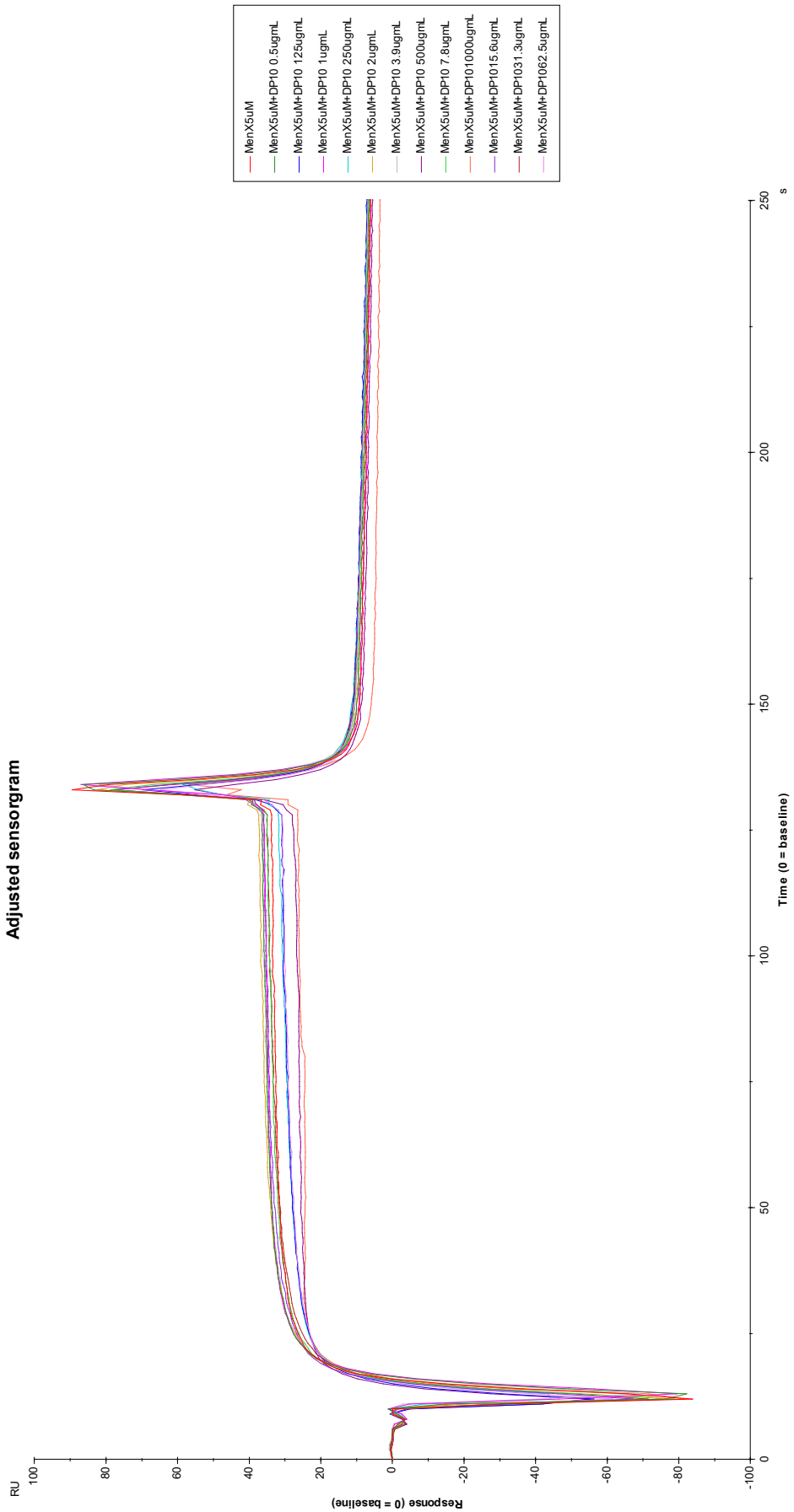


Figure S13. Inhibition assay sensorgram showing competition between DP10 and the surface-attached ligand (MenX(DP15)–ADH–CRM$_{197}$).

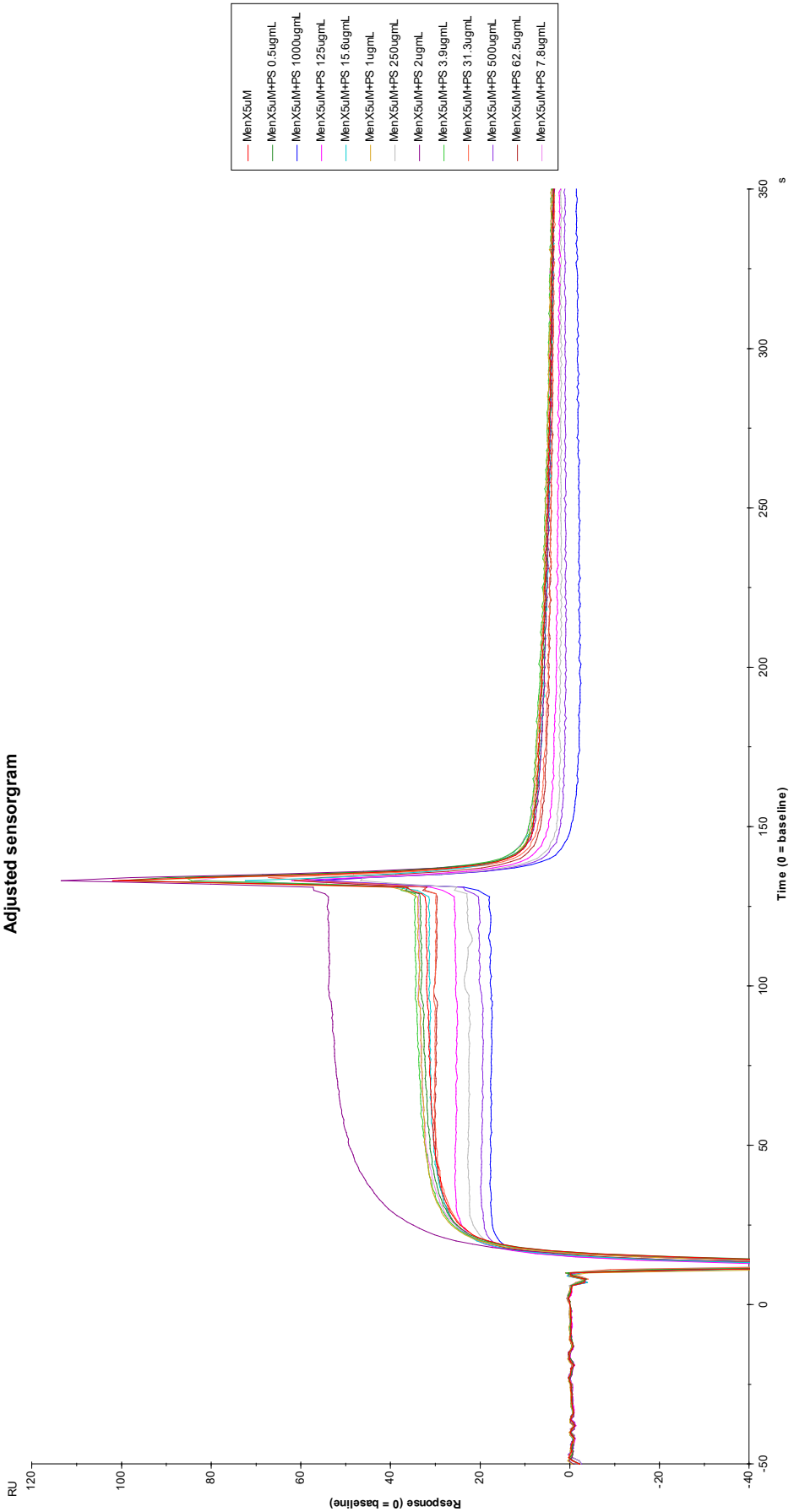


Figure S14. Inhibition assay sensorgram showing competition between PS and the surface-attached ligand (MenX(DP15)–ADH–CRM$_{197}$).