The University of Nottingham

UNITED KINGDOM • CHINA • MALAYSIA

A review of glycoconjugate vaccines: The current state of the technology, considerations for the future, and the gold standard carrier protein, CRM(197)

Abstract

Conjugation of bacterial capsular polysaccharides to immunogenic carrier proteins has been employed for almost five decades to protect infants and adults against bacterial infections. This review investigates the immune mechanisms, conjugation chemistry, and carrier protein interactions that define glycoconjugate technology, and reports on the recent developments that have facilitated a renewed interest into the field of conjugated vaccines. With novel methodologies such as bioconjugation and Generalized Modules for Membrane Antigens (GMMA) reaching maturity, and the carrier protein CRM(197) seeing a research-renaissance, molecular characterization has never been more important. Elucidation of the physiochemical properties of prospective therapeutics is a crucial quality control measure that allows their structural integrity, stability, and immunogenicity to be verified. Thus, a special mention is made to the hydrodynamic methods that ensure current and novel therapeutics remain safe and efficacious (6899 words).

Oliver Meeds
styom@nottingham.ac.uk
University of Nottingham, School of Biosciences,
Sutton Bonington, LE12 5RD

Table of Contents

1. Introduction

Vaccine technology has been propelled into the spotlight in recent years. Armed with almost unlimited public resources, biotechnology companies were able to significantly advance the development and production of vaccines against the SARS-CoV-2 virus, free from the usual economic risks posed by pharmaceutical development (Lalani et al., 2023). Messenger RNA (mRNA) and viral vector vaccines saw the most rapid development and spearheaded the effort to immunise the world against the deadly respiratory disease and return us to some level of normality (Francis et al., 2022). While the pandemic has ebbed away, a shift in the attitude towards health and vaccination remains. For example, a greater acceptance and uptake of influenza vaccination has been observed post pandemic (Kong et al., 2022). During this period of greater intrigue into public health, it would perhaps be wise to shed light on other vaccine technologies that have the potential to prevent lesser-researched and underfunded diseases; technologies such as glycoconjugate vaccines.

One of the phrases frequently broadcasted during the pandemic was ‘spike protein’. Like many viruses, the term relates to the proteinogenic antigen that covers the surface of the SARS-CoV-2 virus, mediating entry into the cell and reproduction (Magazine et al., 2022). This element also elicits an immune response once recognised by the immune system, and antibodies are raised against it (Gao et al., 2020). The spike protein was therefore a primary vaccine candidate against SARS-CoV-2, and it was crucial that the structural domains and corresponding functions were well understood (Xia, 2021). Identification and characterisation of coat protein antigens is a key stage in the development of vaccines against viruses. Structural elucidation can facilitate the elimination of undesirable domains, stabilization of

the antigen such that the most relevant antigenic epitopes are presented, and identification of molecular conjugation sites (Malito et al., 2015). Proteinogenic components of pathogens are not the only vaccine candidates, however. Many different cellular components can act as antigens, can bring about an immune response, and can grant immunity (Weintraub, 2003). This is no better exemplified than in the case of bacterial pathogens whose capsular polysaccharides (CPS) present a highly diverse, highly abundant source of antigenic vaccine candidates (Geno et al., 2015).

CPS are vital to many bacteria. As the component exposed to the extracellular environment, they undergo binding interactions, help the organism evade immune defences, and facilitate the continuation of the life cycle of the cell (Keinhorster et al., 2019, Kenney and Fein, 2011). They are also key virulence factors in immune recognition by their animal hosts (Singh et al., 2018). As a class of high molecular weight extracellular polysaccharides, CPS are carbohydrate containing polymers synthesised via complex biosynthetic pathways and excreted from the bacteria, providing contact with the external environment (De et al., 2010). The rich structural diversity of CPS means that many serotypes may exist for any given bacteria, where each can be identified by unique structures and sero-typical interactions (Jauneikaite et al., 2015). CPS are most common to gram negative bacteria but can be found in gram-positives including pathogens which interact with human epithelial surfaces such as in the gut and respiratory tract (Donnio et al., 2004). These are a particular nuisance as they mimic human tissue to conceal other antigens that would normally illicit an immune response. These bacterial pathogens therefore remain undetected for longer, leading to more chronic infections and disease (Guan et al., 2020). While tricky to detect, healthy individuals will typically elicit a

humoral response to fight off such infections, calling upon a vast repertoire of anti-carbohydrate antibodies (Luetscher et al., 2020).

As native extracellular polysaccharides are non-virulent antigens, a theorised strategy for immunisation against bacterial pathogens was direct injection of purified CPS (Artenstein et al., 1970). The problem with this solution is that polysaccharides are not inherently good at eliciting an immune response. More specifically, native polysaccharides are T-cell independent antigens that can only stimulate antibody production by direct interaction with B-cells, generating an immunoglobulin M (IgM) biased response (Jones et al., 2009). While this offers some protection to the individual, the immune response is significantly lower than it would be in the event of B-T cooperation and antibody class switching to higher affinity immunoglobulin G (IgG) (Kelly et al., 2005). Moreover, during the early stages of human development, the mechanism that stimulates T-cell independent antibody and memory cell production is not yet mature (Zandvoort et al., 2001). Coupled with the heightened disease burden at this age, bacterial infections have historically been a major cause of infant mortality, with pneumonia and meningitis responsible for 23.6% of deaths in children under the age of 5 (Bhutta and Black, 2013). It is therefore imperative to immunise children against bacterial CPS in a way that stimulates B-T cooperation. One solution is to introduce T-cell epitopes by way of a protein carrier molecule - this is the core principle of glycoconjugate vaccines. (van Putten and Tønjum, 2010).

2. Polysaccharide Vaccines

Polysaccharide vaccines are by no means new. The use of native CPS in prevention of pneumonia is reported as early as 1945 with subcutaneous injections of tetravalent formulations of pneumococcus types I, II, V, and VII CPS (MacLeod et al., 1945). However, it wasn't until the 1970s that their importance in preventing bacterial disease was properly recognised with the development and licencing of a 14-valent pneumococcal vaccine. This later developed into a 23-valent formulation that is still used to immunise the elderly against *Streptococcus pneumoniae* (Tomczyk et al., 2014, Gotschlich et al., 1978). A meningococcal vaccine containing four serogroups of bacterial polysaccharide (A, C, Y and 1-135) has also been in use in the USA since 1981 and confers 80-95% bactericidal immunity. Therefore, the efficacy and safety of polysaccharide vaccines is well established (Rouphael et al., 2009). These examples highlight an additional challenge associated with polysaccharides; they are significantly more diverse than proteins (Holt et al., 2020). Where a single proteinogenic antigen may be sufficient in conferring immunity to a virus, a bacterial pathogen can have many different serotypes of capsular polysaccharide, each with the potential to cause disease (Lipsitch, 1997, Cox et al., 2008). *Neisseria meningitidis* has 13 serotypes, with 6 responsible for most incidences of disease. Vaccine formulations against *Neisseria* must therefore protect against each of these 6 serotypes (Lipsitch, 1997, Attarpour-Yazdi et al., 2014).

3. Conjugate Vaccines

The conjugation of polysaccharide antigens to a carrier protein was the most significant development in polysaccharide vaccine technology. The first conjugated vaccine was licenced in 1987 in the form a *Haemophilus influenzae* type b (Hib) conjugate vaccine (Alexander et al.,

1987). This vaccine aimed to reduce the occurrence of Hib associated diseases of which, 90% of cases occur in children under the age of 5 and a further 59% of these cases occur in children younger than 12 months old. Routine vaccinations with Hib vaccines have led to a decline in the incidence of severe Hib infection by more than 90%, a credit to the efficacy and importance of conjugate polysaccharide vaccines (World Health Organisation, 2013). Currently, there are 11 licenced conjugate vaccines in the US offering protection against Pneumococcal, Meningococcal and Hib infections caused by *Streptococcus pneumoniae*, *Neisseria meningitidis* and *Haemophilus influenzae* type b respectively (Center for Biologics Evaluation and Research, 2022).

3.1. Immune Mechanism

Understanding the mechanism of immune response to glycoconjugates is key to the development of effective vaccines. In naïve infants', dendritic cells take up the conjugate vaccine and within days it is transported to lymph nodes where the T-cell dependent antigens of the carrier protein induce germinal centre (GC) formation (Mesin et al., 2016). This requires three crucial cell types: follicular dendritic cells (FDC) which contain and present the antigen, polysaccharide-specific B-cells expressing surface antibodies acting as receptors (BCR), and follicular helper T-cells (T_{FH}) which recognise the protein carrier antigen presented on the surface of B-cells (Rappuoli et al., 2019).

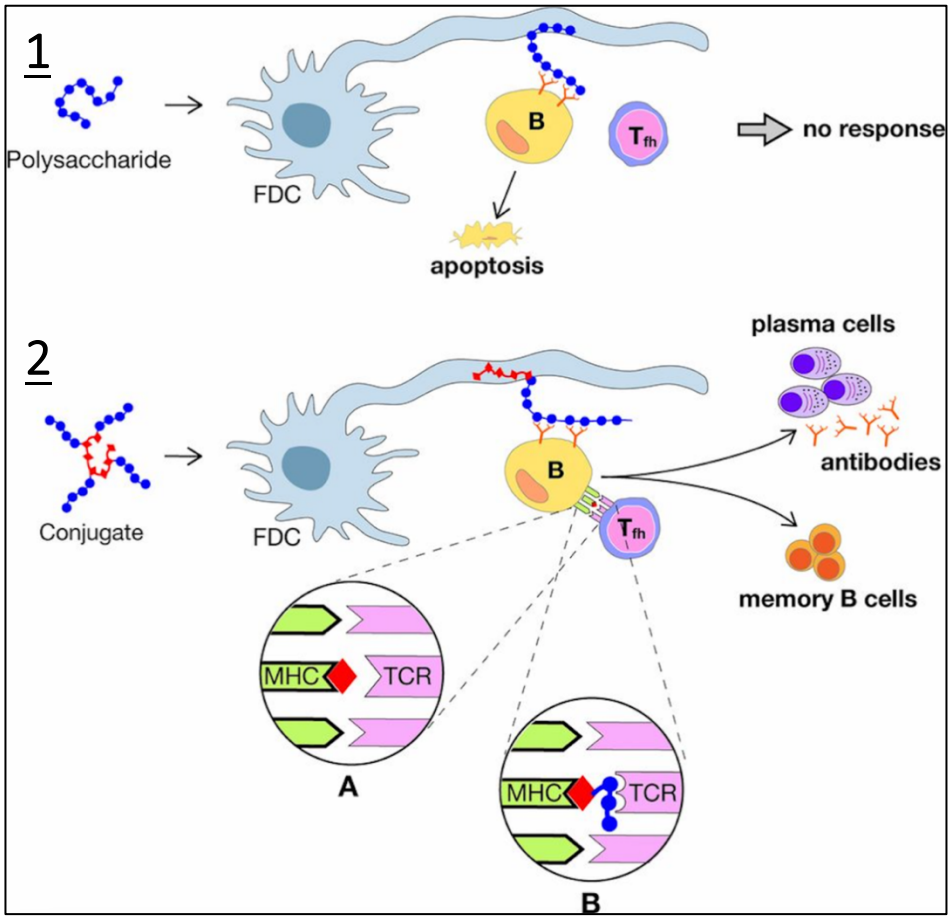

Figure 1. Interaction of polysaccharides-based vaccines with the FDC, B and T_{FH} cells in the GC. 1, Lack of response to native polysaccharides; 2, B-T cooperation elicited by polysaccharide-protein conjugate; A, peptide loading onto MHC; B, TCR recognition of polysaccharide. Adapted from (Rappuoli et al., 2019).

The GC also has distinct light and dark regions. B-cells undergo selection and activation in the light region where they bind and extract conjugate antigens from the FDCs. The peptides are loaded into the cavity of the major histocompatibility complex (MHC) exposing them to T_{FH} cells, which activate the B-cells by direct cell-cell interaction and cytokine secretion, kickstarting the GC reaction. Activated B-cells then enter the dark region of the GC where they multiply rapidly and express cytidine deaminase which introduces random mutations in the variable region of the BCR. B-cells bearing functional receptors re-enter the light zone and once again receive antigens from the FDC to be presented to the T_{FH} cells. Limited numbers of T_{FH} cells ensures that B-cells with improved affinity are preferentially selected and perpetually

enhanced via affinity maturation. Eventually, a small fraction of high affinity B-cells leave the GC to become memory B-cells and polysaccharide-specific antibody secreting plasma cells (Figure 1.2) (Rappuoli et al., 2019, Rappuoli, 2018).

In infants, polysaccharide antigens initially follow the same mechanism. They can bind specific antibodies on the B-cell surface, are extracted from the FDCs, and become internalised. However, they cannot enter the MHC cavity, so the GC reaction can go no further. Interaction of the B-cell with CD4 helper T-cells is needed for B-T interactions. CD4 cells are 'immunodeviant' in neonates and exhibit a shorter-lived effector lineage compared to their adult counterparts. This could explain the absence of this mechanism in children, which matures as the lineage of CD4 cells changes (Cyster and Allen, 2019, Rudd, 2020). Consequently, the necessary steps to produce polysaccharide-specific antibody secreting cells do not take place. The lack of a mature T-cell independent mechanism for antibody proliferation in infants means that they are not afforded immunisation when exposed to polysaccharides alone (Figure 1.1) (Rappuoli et al., 2019).

It is thought that by adolescence we have acquired memory B-cells via direct exposure to polysaccharides from pathogens or via cross reacting polysaccharides from commensal flora or food. These low affinity B-cells can induce T-cell independent extrafollicular proliferation and plasma cell formation in response to polysaccharide vaccines. While this is a preferential response compared to that seen in infants, hypo-responsiveness is often observed after subsequent immunisation resulting from the lack of GC activation and depletion of the memory B-cells pool (Richmond et al., 2000). This leads to a reduced immune response compared unvaccinated people, even to conjugated vaccines, raising concerns about the use

of polysaccharide vaccines. Initial immunisation with conjugates offers better protection and should be used preferentially (Rappuoli, 2018).

3.2. Carrier Proteins

Effective carrier molecules must fit certain criteria. They require approved safety in humans, suitable linkage sites, and many T-cell epitopes (Pichichero, 2013). It is unsurprising then that many existing licenced therapeutic proteins have been reappropriated as carriers for polysaccharide antigens. Of the currently licenced glycoconjugate vaccines, all are conjugated to one of five carrier proteins. Two are bacterial cell membrane proteins: protein D (PD) a highly conserved ~42 kDa surface lipoprotein of *Haemophilus influenzae*, and the outer membrane protein complex of *Neisseria meningitidis* serogroup b (OMPC). Three are secreted toxoid proteins in the form of chemically detoxified tetanus and diphtheria toxoid proteins and a genetically detoxified mutant of the diphtheria toxoid protein CRM_{197} (Micoli et al., 2018). A summary of licenced conjugated vaccines and their carriers is presented in Table (1).

Table 1. Licenced vaccines and their carriers (Bröker et al., 2017, Kallet and Aaron, 2021).

Carrier protein(s)	Polysaccharide(s)	Brand name(s)
TT	Hib	Hiberix™, ActHIB™
	MenA	MenAfriVac™
	MenC	NeisVac-C™
	MenACWY	Nimenrix™
	MenC-Hib	Menitorix™
	MenCY-Hib	MenHibrix™
DT	Hib	ProHIBIT™
	MenACWY	Menactra™

Carrier protein(s)	Polysaccharide(s)	Brand name(s)
CRM	Hib	VaxemHib™
	MenC	Meningitec™, Menjugate™
	MenACWY	Menveo™
	7 pneumococcal serotypes	Prevnar™
	13 pneumococcal serotypes	Prevnar 13™
	20 pneumococcal serotypes	Prevnar 20™
	15 pneumococcal serotypes	Vaxneuvance™
Protein D, TT, DT	10 pneumococcal serotypes	Synflorix™
OMPC	Hib	PedvaxHIB™

Abbreviations: CRM, cross-reacting material 197; DT, diphtheria toxoid; Hib, *Haemophilus influenzae* type b; MenA, meningococcal capsular group A; MenACWY, meningococcal capsular groups A, C, W, Y; MenC, meningococcal capsular group C; MenCY, meningococcal capsular groups C, Y; OMPC, outer membrane protein complex; TT, tetanus toxoid.

3.3. Cross Protection

One of the more novel benefits reported in glycoconjugate vaccine usage is the cross protection offered by the carrier protein (Bröker et al., 2017). As the proteinogenic component of the vaccine is also antigenic, an immune response will be triggered against both the polysaccharide and the protein. This means that the carrier used for conjugation can be tailored to give immunity where it is most needed. Theoretically, if an individual is already immunised against tetanus but not diphtheria, then the most suitable protein carrier for this individual might be diphtheria toxoid or CRM_{197}. In reality, tetanus and diphtheria vaccines are frequently administered together. However, this theory may become more relevant with the development of novel carriers that could offer protection against a broader spectrum of pathogens (Gasparini et al., 2014). All toxoid-based carriers have been reported to significantly

increase antibody levels against their respective pathogens. All evoke antibody titres similar to that of a booster dose of tetanus-diphtheria vaccine (Bröker et al., 2017).

PD is the first non-typal *Haemophilus influenzae* (NTHi) antigen that has been demonstrated as eliciting a protective immune response in humans. Its development as a carrier protein is significant as NTHi has been implicated in middle ear infections following pneumococcal infections of the respiratory tract (Holz et al., 2023). When PD is used to conjugate pneumococcal polysaccharide antigens, it has been shown to reduce acute otitis media (AOM) caused by both the pneumococci and non-typal *H. influenzae* (Forsgren and Riesbeck, 2008).

4. Structures

Considering that the majority of licenced glycoconjugates are conjugated to one of three proteins (Table 1), comparing their structures to elucidate common features and differences may be informative for the future development of novel carriers. It may also allow us to better understand their hydrodynamics and physical characteristics.

4.1. Tetanus Toxoid

The tetanus neurotoxin (TeNT) of *Clostridium tetani* is a ~150 kDa protein which enters the nervous system and inhibits neurotransmitter release by hydrolysis of the 76-Gln-|-Phe-77' bond of synaptobrevin-2 (Schiavo et al., 1992). TeNT is 1315 amino acids in length, split between a light and heavy chain, and bears a closed domain structure stabilised by two disulphide bonds. The light chain is a comprised of a single ~50 kDa N-terminal catalytic domain responsible for the enzymatic activity of the toxin. The heavy chain consists of two

~50 kDa domains: an N-terminal translocation domain responsible for transport across the synaptic vesicle, and a C-terminal domain responsible for neuronal binding (Masuyer et al., 2017). Regarded as one of the most potent neurotoxins, it is essential that TeNT is detoxified to tetanus toxoid (dTT) before conjugation or direct use (Choi et al., 2018). Formaldehyde treatment induces methyl adducts on amino acid residues which are further dehydrated leading to Schiff base modification (Bayart et al., 2017). Schiff bases are then converted to cross-links by reaction with tyrosine residues (Aggerbeck and Heron, 1992). Analysis of dTT revealed most modifications occur in the heavy chain of the protein, and toxicity is negated by preventing the toxin from entering synaptic vesicles (Long et al., 2022). Certain residues of dTT are consistently modified during detoxification. Long et al. (2022) found 15 repeatable modifications, 14 on the heavy chain and 1 on the light chain. However, they also found lots of variation between batches, so the modifications and resulting structure cannot be entirely predicted. This could be problematic considering structural changes could alter the immunogenicity of the protein, although much of the antigenicity of dTT is attributed to the light chain which avoids the brunt of the chemical modification. (Long et al., 2022).

High throughput mapping of TeNT revealed 43 IgG epitopes. Three epitopes, TT-8/G, TT-40/G, and TT-42/G were found to be key epitopes in dictating the efficacy of the vaccine. TT-8/G interferes with the enzymatic toxicity of the virus, while the other two interfere with neural cell receptor binding by the toxin (De-Simone et al., 2023). These three epitopes are visualised in Figure (2).

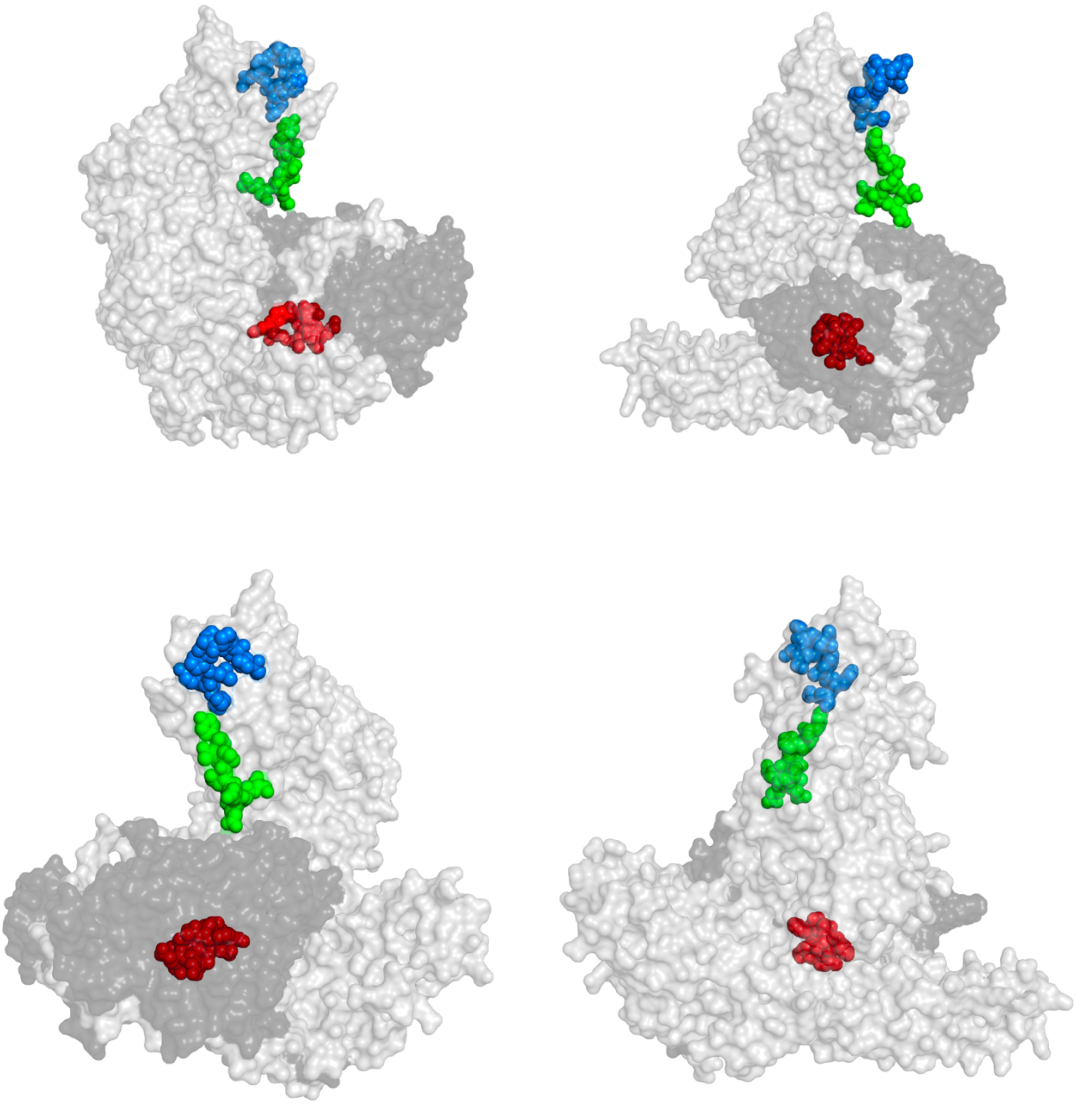

Figure 2. PyMOL representations of tetanus toxin (PDB: 5N0B) rotated through 90° highlighting the light chain (dark grey), heavy chain (white) and significant epitopes: TT-8/G, blue; TT-40/G, green; TT-42/G, blue.

It is interesting to note that the significant epitope TT-8/G, found on the light chain, is buried within the closed structure of the toxin. This could indicate the occurrence of a structural change of the protein *in vivo*, allowing the epitope to become available for binding, although the literature indicates that this epitope is on the surface of the protein and accessible to the solvent. The other significant epitopes are found on the heavy chain, meaning that they could

be subject to structural changes induced by formaldehyde treatment, reducing the overall vaccine efficacy (De-Simone et al., 2023).

4.2. Diphtheria Toxoid

Diphtheria toxin (DT) is an exotoxin of *Corynebacterium diphtheriae* that inhibits protein synthesis. This ~58 kDa protein consist of a single polypeptide chain of 535 amino acids and three domains – receptor (R), translocation (T), and catalytic (C). The R domain is responsible for cell binding and initiates endocytosis. A shift to acidic pH causes conformational change of the protein and subsequent insertion of the T domain into the endosomal membrane. This facilitates the translocation of the C domain into the cytosol which is then cleaved from domains R and T by proteolysis, allowing it to inhibit the EF2 translation factor and terminate protein synthesis (Rodnin et al., 2020). The protein can be proteolyzed with trypsin to two distinct fragments. Fragment A (residues 1 – 190) contains the C domain, and fragment B (residues 191 – 535) contains the remaining R (residues 191 – 378) and T (residues 379 – 535) domains (Bennett et al., 1994). Diphtheria toxin is frequently observed as a dimer consisting of two identical chains, although the dimeric form of DT does not bind to receptors and is therefore nontoxic. The dimer slowly dissociates to form fully toxic monomers (Choe et al., 1992).

Like tetanus toxin, DT is extremely potent. The lethal dose of DT is 0.1 µg of toxin per kg of body weight, so detoxification is essential prior to its use as a therapeutic agent or carrier molecule (Murphy, 1996). Formaldehyde treatment is also the most common method for the conversion of DT to detoxified diphtheria toxin (dDT). Typically, formaldehyde and extra glycine is added to a culture supernatant containing DT, resulting in intramolecular cross-links

and formaldehyde-glycine attachments. These structural changes critically alter both the NADþ-binding cavity and the receptor-binding site of diphtheria toxin (Metz et al., 2020). An investigation into repeatable modifications found more modification of the A chain compared to the B chain, and concluded that A chain modification eliminates toxicity while the B chain retains immunogenicity (Long et al., 2022). Some T-cell epitopes are also affected by these chemical changes, although universally the effects are negligible (De-Simone et al., 2021).

20 IgG epitopes were found on DT after mapping. This is roughly half the number of epitopes that are present on the surface of the TeNT, but DT is almost a third the size of TeNT, so has more epitopes relative to its molecular weight (De-Simone et al., 2021). Regarding linkage sites, lysine is the most frequently modified residue by reductive amination. DT has 39 lysines while TeNT has 107. Glycine is also a commonly modified residue, used for attachments of linker moieties. TeNT has 8 cysteines, 5 of which appear at the surface. DT has 2 cysteine residues with both lying on the surface. The implication here is that there is that TeNT affords more flexibility in glycosylation site selection compared to DT (Figure 3).

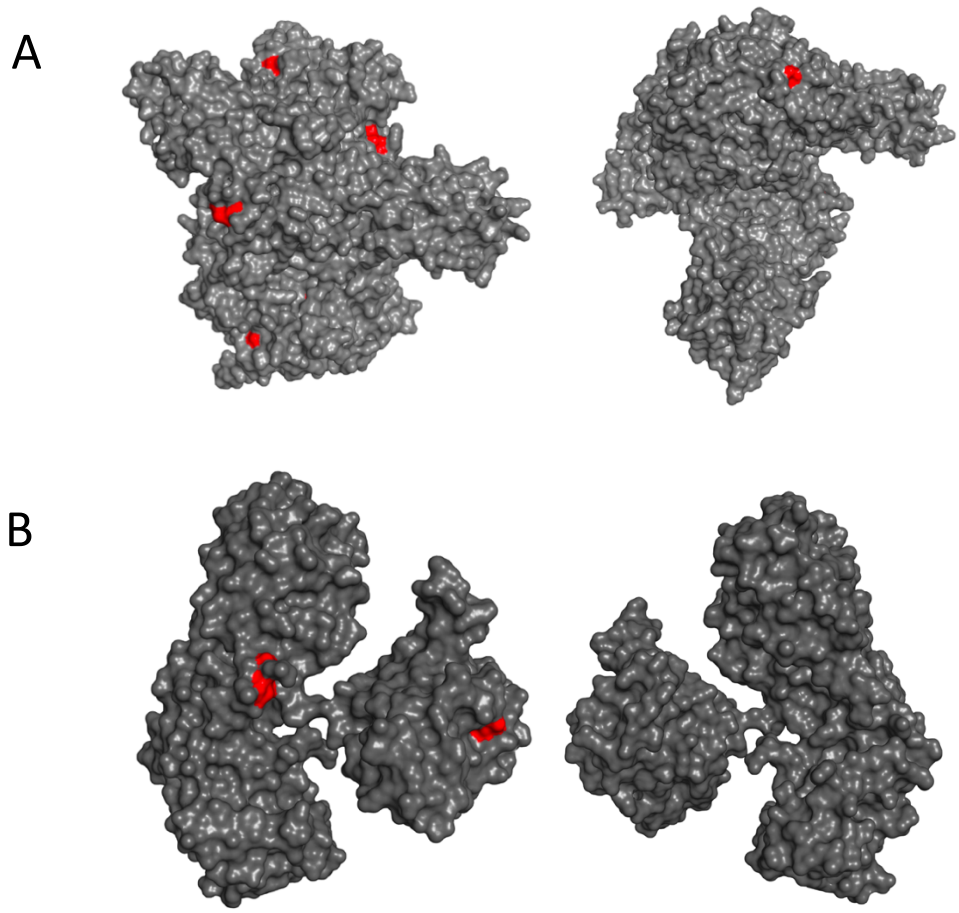

Figure 3. PyMOL representations of both toxoid carriers with surface cysteine residues highlighted in red. A, TeNT (PDB: 5N0B); B, DT (PDB: 1SGK)

4.3. CRM_{197}

CRM_{197} or Cross-Reacting Mutant 197 was first described after discovery of the lysogenic $C7(\beta_{197})^{tox\text{-}crm+}$ strain of *Corynebacterium diphtheria* producing a detoxified form of the diphtheria toxin. This strain was subsequently used to isolate mutant phage producing the same protein. Previous non-toxic forms of the diphtheria toxoid protein had been isolated, but CRM_{197} was special in that it bare such a close resemblance to native DT (Uchida et al., 1972).

The physical and chemical properties of CRM_{197} and DT are almost identical. The two proteins have the same molecular weight of ~58 kDa and are immunologically indistinguishable. The key difference is that CRM_{197} has a toxicity less than one-millionth that of DT, with 0.5 mg failing to produce a reaction, proximal to the site of injection in guinea pigs. The lack of toxicity is a result of a single glycine to glutamate substitution at position 52 (G52E) that renders the protein enzymatically inactive (Uchida et al., 1972, Giannini et al., 1984).

In native DT the mechanism of toxicity is NAD dependent. The C domain of fragment A catalyses the cleavage of the N-glycosidic link between the nicotinamide ring and the N-ribose of NAD. ADP-ribose is then covalently transferred to the modified histidine 715 (diphthamide) of the elongation factor EF-2. This results in the inactivation of EF-2 and protein synthesis ceases, leading to cell death (Malito et al., 2012). The loss of toxicity can be attributed to a significantly reduced ability of CRM_{197} to both bind and hydrolyse NAD in solution. This is due to the change in structure of an active site loop that covers the front of the NAD binding site. In DT, this loop is well ordered in the absence of NAD and disordered in its presence (Bell and Eisenberg, 1997, Bell and Eisenberg, 1996). The loop in the mutant is disordered in both cases and displays an intrinsic flexibility regardless of NAD binding. Due to the active site in itself being structurally unchanged, the loss of actively can be wholly linked to the inability of NAD to bind (Malito et al., 2012). A comparison of the structures reveals the altered active site loop. The increased flexibility of the loop in the unbound state is evidenced by the lack of structure due to low electron density (Figure 4).

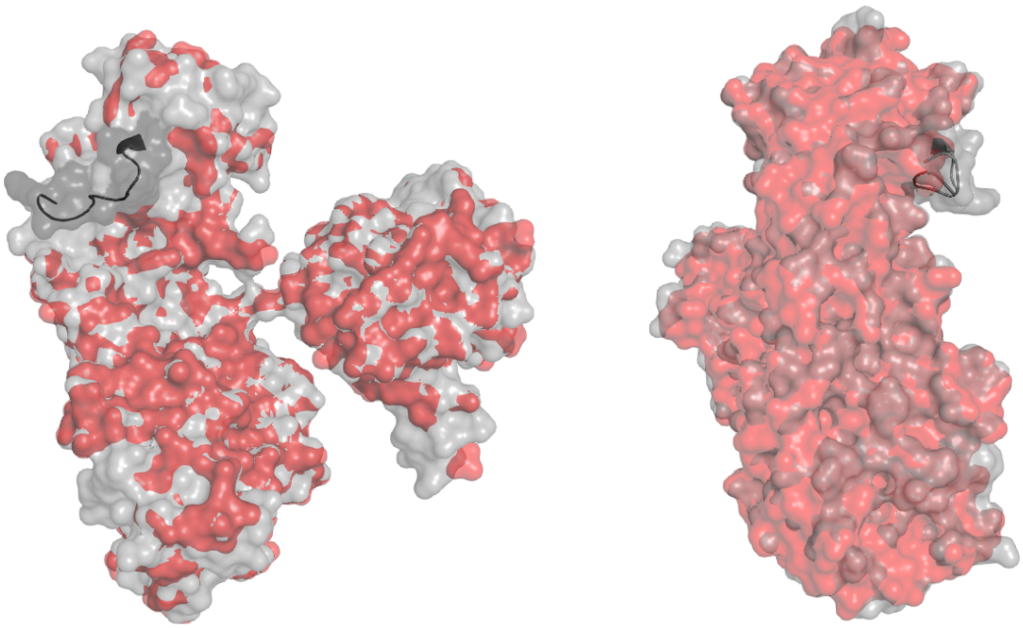

Figure 4. Monomer of nucleotide free native diphtheria toxin (grey, PDB: 1SGK) and CRM_{197} (red, PDB: 4AE0) superimposed in PyMOL. The flexible active site loop (CL2) comprising residues 38–49 (PKSGTQGNYDDD) is displayed in dark grey.

The benefits of a genetically detoxified carrier protein are significant from a structural standpoint. As the mutant always contains the same single amino acid substitution, researchers can be far more confident about the structure they are dealing with. Without the need for chemical processing, the protein will be free from random structural deviations and all the epitopes remain unchanged. The antigenicity of the protein can be more easily predicted and has greater consistency between batches. It has been suggested that this more uniform structure confers improved carrier effect when compared with dDT (Mishra et al., 2018). It should be noted that there is conflicting evidence stating that formaldehyde modification acts to stabilise fragment B, such that it resists proteolysis more readily. Where fragment B of CRM_{197} is prone to degradation in vivo, the structural changes due to chemical modification preserve the fragment and its immunogenicity (Porro et al., 1980).

5. Carrier Production Methods

5.1. Tetanus and Diphtheria Toxins

Methods for production of the TeNT and DT are straightforward. Typically, toxigenic strains of bacteria are grown in liquid medium under conditions that promote toxin formation (Licona-Cassani et al., 2016). *Corynebacterium diphtheriae* and *Clostridium tetani* are cultured for DT and TeNT production respectively. Once a sufficient titre of toxin has been achieved, the supernatant is treated with formalin, an aqueous solution containing formaldehyde, to yield the requisite detoxified toxoid proteins. Toxoid proteins are then purified via ion exchange chromatography, ready for downstream processing (Metz et al., 2013). For diphtheria toxoid production, a hypertoxinogenic strain PW8 is grown in large fermenters for 36 – 45 h or until the concentration of the toxin reaches 150 – 250 limit of flocculation units (Lf/ml), where 1 Lf is equal to roughly 2.5 µg of toxin. This strain is often grown on low-cost media. Formaldehyde treatment of the supernatant then takes 4 – 6 weeks to ensure complete detoxification (Bröker et al., 2011, Chung et al., 2016).

5.2. CRM_{197}

Production of CRM_{197} is more interesting than its native toxin counterpart. As a secreted protein, CRM_{197} can be directly purified from the cell supernatant and requires no further processing. Classically the mutant protein has been expressed in native *Corynebacterium diphtheriae,* although the need for more tightly controlled culture conditions results in higher production costs compared with the native toxin. This has led researchers in search of different production vehicles (Aw et al., 2021). Principally, research has focussed on

production in *E. coli* as CRM_{197} exhibits toxicity towards mammalian cells and spheroplasts of *Saccharomyces cerevisiae.* This has come with several challenges. While *E. coli* produces high yields, the protein frequently forms inclusion bodies which requires intensive downstream refolding steps to yield the native structure, inflating the cost of production (Park et al., 2018). Periplasmic expression promises to overcome this, also reporting high yields of up to 3 $g.L^{-1}$, but significant downstream processing to recover the protein from the periplasmic space also limits the potential implementation of this production method (Goffin et al., 2017). Expression in secretory organisms has been achieved, although B. subtilis yielded only small titres of proteins at 7.1 mg L^{-1} (Jianhua and Petracca, 1999). One of the more promising candidate production organisms is the methylotrophic yeast *Pichia pastoris*. These single eukaryotic cells are tolerant to CRM_{197} and can be grown to high cell density on less expensive media. Crucially, they also secrete CRM_{197} which can be easily purified to >95% purity at a concertation of 114 mg L^{-1} (Aw et al., 2021). This yield is consistent with traditional production in *Corynebacterium diphtheriae,* which ranges from 100 – 150 mg L^{-1}. The most productive recombinant CRM_{197} comes by way of the commercially available Pfenex CRM_{197} produced in an engineered strain of *Pseudomonas fluorescens,* claiming yields of 1 – 2 g L^{-1}. This is also the only commercially available current good manufacturing practice (cGMP) CRM_{197} that is currently used in approved conjugate vaccines; all others are derived from the classic *Corynebacterium diphtheriae* expression system (Hickey et al., 2018).

6. Conjugation Methods

The presence of many -OH groups is usually indicative of good reactivity due to their reactive covalent bonds and the electronegativity of oxygen compared to carbon and hydrogen

(Klecker and Nair, 2017). Despite this, most polysaccharides are not sufficiently reactive for direct conjugation with proteins. Typically, chemical modification is required to introduce reactive functional groups such that the glycan can covalently bind with the carrier protein.

6.1. Reductive Amination

The method employed for conjugation in the majority of licenced glycoconjugate vaccines is reductive amination. This is achieved by targeting neighbouring hydroxyl-containing monosaccharides with sodium periodate. Hydroxyl groups react with sodium periodate to form aldehydes that undergo Schiff base formation with amines on the protein carriers. Generally, polysaccharides activated in this way are readily coupled to lysine residues on carrier proteins by reductive amination (Figure 5) (Morais and Suarez, 2022).

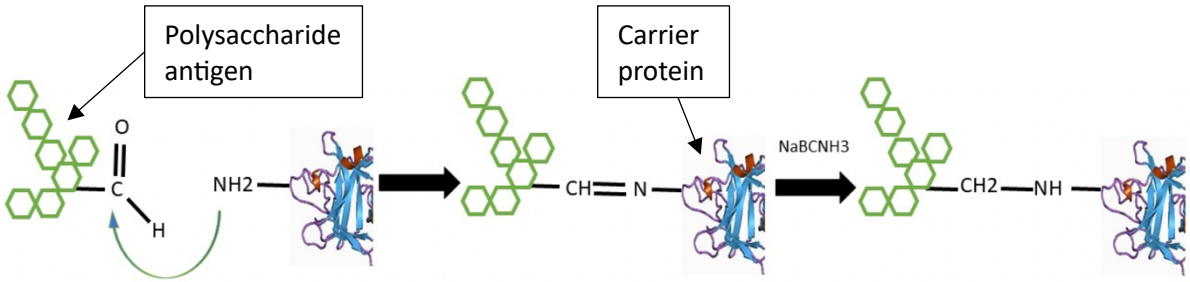

Figure 5. Reductive amination reaction between a carbohydrate oxidised with sodium periodate and the amine group of a carrier protein. Adapted from (Morais and Suarez, 2022).

Chemical modification of the polysaccharide antigen must be carefully performed so to retain the immunogenicity and stability of the vaccine. The oxidation of diols is often localised to saccharide rings where the reaction leads to the irreversible opening of the ring and the formation of a pseudo-sialic acid structure. If many rings along the length of the

polysaccharide are opened the structural change could reduce antibody cross-reactivity and stability (Lees et al., 2008).

6.2. Cyanylation

Another established strategy for polysaccharide activation is to introduce cyanoester groups that will react with amines or hydrazines. This was initially performed with cyanogen bromide (CNBr) and later implemented in the production of the PRP-T Hib conjugate vaccine (Schneerson et al., 1980, Chu et al., 1983). Later iterations of this method use CDAP (1-cyano-4-dimethylaminopyridinium tetrafluoroborate) to circumvent the issue of low efficiency of CNBr which required conjugates to be formulated with linkers such as adipic acid dihydrazine to achieve sufficient levels of conjugation. CDAP reacts with the polysaccharide, exchanging a cyano group for a hydroxyl hydrogen to produce a cyanoester. The cyanoester will then link the epsilon amine of lysines to form a strong covalent bond (Figure 6) (Frasch, 2009).

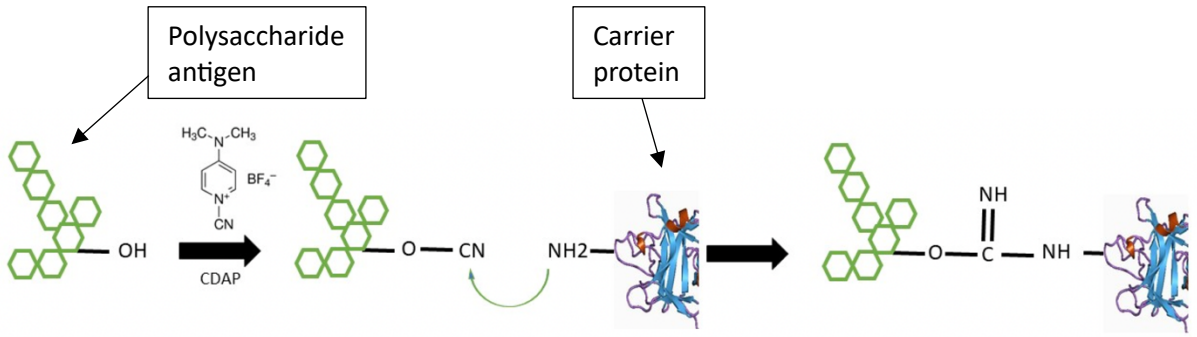

Figure 6. Cyanlyation of a carbohydrate with CDAP and subsequent reaction with an amine of a carrier protein. Adapted from (Morais and Suarez, 2022).

Again, researchers have raised concerns as to the impact on molecular integrity using this method. Optimization of the reaction involves increasing the pH and performing the reaction in the cold due to the instability of CDAP. The increase in the pH means that side reactions are

more likely and may affect the stability of the polysaccharide. Additionally, both cyanylation and reductive amination are considered random methods for conjugation meaning there is little control over the site at which polysaccharides will conjugate (Lees et al., 2020).

6.3. GOase Activation

A novel alternative to chemical modification methods is enzymatic oxidation by the GOase enzyme of the fungus *Fusarium sp.* In certain polysaccharides, GOase generates site specific aldehydes at position C6 of galactose in a reversible reaction (Duke et al., 2022). CPS of *S. pneumoniae* conjugated to CRM_{197} using the GOase method elicited significantly higher IgM and IgG titres compared with chemical oxidation, and equal functional protection measured by opsonophagocytic killing potential assay. This less destructive conjugation method therefore elicits a stronger humoral response, in a way that is non-detrimental to the efficacy of the vaccine (Berti, 2022).

6.4. Site Selective Approaches

Carrier protein mutagenesis can be carried out to introduce a cysteine residue or unnatural amino acid to which functionalized polysaccharide antigens can be selectively coupled. Studies into site selective approaches typically conjugate a synthetic carbohydrate to a bacterial protein (Nilo et al., 2015). In one study, Pillot et al. (2019) made use of the lack of cysteine residues in the structure of pneumococcal surface adhesin A by replacing surface-exposed lysines with single cysteine residues. A synthetic tetrasaccharide representative of a serotype 14 CPS of *Streptococcus pneupmoniae* could then be linked to the single cysteine of the carrier protein by thiol/maleimide coupling chemistry using a sulfhydryl (SH) linker. This demonstrated the specific conjugation of a sugar residue in a fully controlled reaction (Pillot

et al., 2019). Additional control was added to this method by the adoption of amber codon rearrangement for the incorporation of unnatural amino acids (UAA) *in vivo*. Non-canonical amino acids can be coded for by inserting orthogonal aminoacyl-tRNA synthetase/tRNA pairs into normal protein sequences via mutation (Liu and Schultz, 2010). Using this method, the non-canonical amino acid N^{ϵ}-propargyloxycarbonyl-l-lysine (PrK) bearing a bioorthogonal alkyne functional group was incorporated into a pneumococcal protein in place of surface lysines. Click chemistry was then used to link a synthetic azido-functionalised tetrasaccharide by reaction of their alkyne groups. This chemistry refers to efficient yet selective reactions joining small building blocks by formation of a covalent bond between two reactive functional groups performed under mild conditions (Kolb et al., 2001). The resultant conjugates produced similar antigenicity when compared with a randomly conjugated control (Violo et al., 2023).

The benefit of this system is that the only full-length protein synthesis occurs if the UAA is introduced, simplifying the quality control process. It is possible to produce combinatorial mutants harbouring more than one sugar residue, implying that multivalent vaccines could be produced using this method. There are some concerns as to the stability of the protein, however. Compared with random addition of the tetrasacchride, the antibody titres of the mutants were less homogeneous, potentially a result of structural changes due to chemical derivatization and conjugation (Violo et al., 2023).

6.5. Clinical Relevance

In addition to the concerns surrounding conjugation method and molecular stability, the method or chemistry used in conjugation has also been reported to influence the effectiveness of vaccines. *Streptococcus pneumoniae* serotype 19F conjugated to CRM_{197} using

reductive amination raised an additional epitope that isn't present when the same polysaccharide is conjugated to CRM_{197} using cyanylation. It was also discovered that the cyanylated vaccine resulted in greater opsonophagocytic antibodies titres as well as higher levels of cross-opsonophagocytic antibodies against the closely related serotype 19A; clearly suggesting a relationship between conjugation method and antibody functionality (Poolman et al., 2011).

7. Quality Control and Characterization

In vaccine development, one of the most important considerations is quality control, with protocols implemented in each stage of the developments process to validate the structural integrity of the products being produced. Characterisation of the conjugated vaccine itself is arguably the most important step in safeguarding patient health. Immunoassays and physiochemical test ensure that the vaccine is efficacious and meets the strict regulatory guidelines set out by authorities such as the WHO (Jones, 2015, World Health, 2022).

7.1. Analytical Ultracentrifugation

Analytical ultracentrifugation (AUC) is a primary technique used to investigate the hydrodynamics of macromolecules (Lebowitz et al., 2002). In an AUC experiment, high centrifugal forces cause the sedimentation of the macromolecule through its solution. The rate at which the molecule sediments is analysed by absorbance or interference optical systems depending on the whether the macromolecule absorbs UV radiation. This is then used to determine molecular weight and sedimentation coefficient and subsequently to determine the shape and size of the molecule. Two independent methods, sedimentation velocity and

sedimentation equilibrium, are used to determine sedimentation coefficient and molecular weight respectively. Unlike other analytical technique, these experiments can be performed in biological solutions and true solution-state conditions. AUC doesn't rely on interaction with a separation matrix, so data can be analysed directly without the need for comparison with a standard (Edwards et al., 2020).

7.1.1. Sedimentation Velocity

Sedimentation velocity (SV) involves centrifugation at high-speed causing the sample to migrate towards the bottom of the cell independent of back diffusion, creating a boundary between the solvent and solution (Hansen et al., 1994). The solute accelerates initially until a constant speed of sedimentation is reached and then the velocity of movement v can be measured. The rate of the movement of the boundary per unit centrifugal field allows the sedimentation coefficient to be calculated which is dependent on the shape, size, and viscosity of the molecule. Sedimentation coefficient, s, is determined via the svedberg equation:

$$s = \frac{v}{\omega^2 r} = \frac{M(1-\bar{v}\rho_0)}{N_A f} = \frac{M(1-\bar{v}\rho_0)D}{N_A k_B} = \frac{M(1-\bar{v}\rho_0)}{RT} \qquad (1)$$

Where s is the sedimentation coefficient, v is the velocity of boundary movement, $\omega^2 r$ is the angular acceleration, M is the molar mass, $\bar{v}$ is the partial specific volume, ρ_0 is the density of the solvent, N_A is Avogadro's constant, f is the frictional coefficient, D is the diffusion coefficient, k_B is the Boltzmann constant, R is the ideal gas constant, and T is the temperature.

Molecular interactions are identified by SV on a sedimentation distribution. Aggregating molecules sediment slower and are identifiable by multiple peaks corresponding to species with a molecular weight greater than that of the target molecule. Conversely, degradation products or impurities may sediment faster than the target, and therefore appear as peaks with lower corresponding molecular weights (Dam et al., 2005).

One of the pitfalls of SV is the influence nonideality of the analyte caused by co-exclusion and residual polyelectrolyte effects. This can lead to a miscalculation of solution properties. Thus, sedimentation distributions must be produced using the lowest possible concentrations. Once s has been corrected to standard density and viscosity of water at 20°C ($s_{20,w}$), a plot of $s_{20,w}$ against concentration can be extrapolation to zero to find s free from nonideality effects (Harding et al., 2015).

7.1.2. Sedimentation Equilibrium

Sedimentation equilibrium (SE) experiments are performed at lower speeds, where the sedimenting forces are allowed to equilibrate with the back diffusion effects. At equilibrium, an even distribution of concentrations is observed when the flux of sedimenting molecules is equal to the flux of diffusing molecules. Under these conditions there are no friction effects, so the migration of the macromolecule is independent of shape and only a function of the weight average molecular weight (M_w). Centrifugal conditions are related to molecular weight by the equation:

$$\omega^2 r M(1 - \bar{v}\rho_0) = \frac{RT}{c(r)}\frac{dc(r)}{dr} \qquad (2)$$

Where $\omega^2 r$ is centrifugal acceleration, M is the molar mass, $\bar{v}$ is the partial specific volume, ρ_0 is the density of the solvent, R is the ideal gas constant, T is temperature, c is concentration, and r is the radius (Cole et al., 2008).

One limitation of SE experiments is that the sample purity and homogeneity must be more strictly controlled than in SV. Non-homogeneous species are not fractionated during SE, as they are in SV. Thus, contaminants or denaturation products of the analyte will influence the determined molecular weight. Additionally, multiple exponential functions must be solved by software to fit SE concentration gradients; a task that becomes exponentially more difficult with multiple species (Cole et al., 2008).

7.2. SEC-MALS

Size exclusion chromatography coupled with multiple angle light scattering (SEC-MALS) is a technique commonly applied in biomolecular technology to analyse the molecular weight distribution of macromolecular solutions. Solutes are passed through a column filled with a porous matrix at a constant rate. Smaller molecules enter the pores and are retained in the column for longer. Molecules therefore pass through the column in descending order of molecular weight, which can be determined by comparing elution time with known standards. As elution time is also dependent on shape, the standard must have the same conformation as the target molecule. This adds an additional level of complexity to SEC when analysing complex biomolecules capable of multiple conformations such as polysaccharides (Some et al., 2019).

The MALS component works around this problem as an absolute determinate for molecular weight, independent of standards. Spherical particles with sufficient size will cause light to scatter in all directions at equal intensity so it is typically okay to use single angle light in this case. Non-spherical partials will cause light to scatter with variable intensity depending on the angle of incidence. Therefore, multiple angle measurements allow the molecular weight to be determined by the equation:

$$M = \frac{R(0)}{Kc(\frac{dn}{dc})^2} \quad (3)$$

Where M is the molecular weight of the analyte, $R(0)$ is the reduced Rayleigh ratio (intensity of scattered light vs laser intensity) extrapolate to zero angle by the instrument, c is the weight concentration determined by the UV or rDI detector, $\frac{dn}{dc}$ is the refractive index increment of the analyte, and K is an optical constant defined by Equation (4).

$$K = \frac{4\pi^2 n_0^2 (\frac{dn}{dc})^2}{N_A \lambda^4} \quad (4)$$

Where n_0^2 is the solvent viscosity, $\frac{dn}{dc}$ is the refractive increment, N_A is Avogadro's constant and λ is the wavelength (De et al., 2010).

SEC MALS and AUC experiments can be used interchangeably, and both can give an accurate report of aggregation profiles and molecular weight. The former is often preferred due to the significantly quicker analysis times, but AUC acts as a complementary method that can validate the data obtained via SEC-MALS (Ogawa and Hirokawa, 2018).

7.3. Intrinsic Viscosity

Solution viscosity of macromolecules is of significant interest due to the impact of viscosity on the development of new therapeutics. As viscosity depends on molecular shape and size, intermolecular interactions can lead to complications in bioprocessing, formulation development, subcutaneous drug delivery, and manufacturing of biomolecules. Higher concentrations, and even small amino acid changes in proteins can result in increased aggregation and subsequently much higher viscosities (Buck et al., 2015).

Typically, the relative viscosity of a dialysate can be determined by comparing its flow rate against the solution in which it is dissolved, given the densities of each:

$$\eta_r = \frac{t}{t_0} \cdot \frac{\rho}{\rho_0} \qquad (5)$$

Where η_r is relative viscosity, t is time of flow of dialysate, t_0 is time of flow of solution, ρ is density of dialysate, and ρ_0 is density of solution.

Further equations are used to find reduced viscosity by considering the concentration of the dialysate:

$$\eta_{red} = \frac{\eta_{r-1}}{c} = K_H[\eta]^2 c + [\eta] \qquad (6)$$

Where η_{red} is reduced viscosity, c is concentration, K_H is the Huggins constant, and $[\eta]$ is the intrinsic viscosity.

Transformation of the equation allows for $[\eta]$ to be plotted against concentration and extrapolated back to zero to account for nonideality effects. These effects typically occur when intermolecular forces between a solute and solvent are less strong than between solute molecules themselves. An approximation of $[\eta]$ from a single concentration is given by the Solomun-Ciuta Equation (7) but is often less accurate than extrapolation of the physical data (Harding, 1997).

$$[\eta] \sim \frac{\sqrt{2(\eta_r - 1) - 2\ln(\eta_r)}}{c} \quad (7)$$

7.4. Comparison of Hydrodynamic Techniques

While each of the techniques mentioned previously can reveal important molecular information, a comparison of the all the methods highlights the advantages and limitations of each (Table 2). Thus, to obtain a more accurate characterization of the physiochemical properties of a particular biomolecule, all the techniques should be used collectively to validate the results.

Table 2. Comparison of hydrodynamic techniques (Harding et al., 2015).

Consideration	Technique			
	Sedimentation Velocity	**Sedimentation Equilibrium**	**SEC-MALS**	**Capillary Viscosity**
Time	Hours	Day(s)	Minutes	Minutes
Cost	Expensive	Expensive	Less Expensive	Inexpensive
Measurements	Molar mass distribution, S	Absolute M_w and M_z	Molar mass distribution, M_w, M_z, M_n, R_g	$[\eta]$
Analyte volume	~ 400 µl	~ 100 µl	~ 100 µl	1 – 2 ml
Conformation information	Yes	No	Size only	Comprehensive – Axial dimension and flexibility
Interaction information	Yes	No	Yes	Yes
Physiological conditions	Yes	Yes	Yes	Yes
Matrix	No	No	Yes	No

Abbreviations: SEC-MALS, size exclusion chromatography coupled to multiple angle light scattering; S, sedimentation coefficient; M_w, weight average molecular weight; M_z, size average molecular weight; M_n, number average molecular weight; $[\eta]$, intrinsic viscosity.

8. Prospects

8.1. Novel Carriers

Given the application of glycoconjugate vaccines dates back almost half a century, it is surprising that only five carrier proteins are currently utilized, especially considering the benefits that are afforded by cross-protection. Many novel protein carriers are currently being researched as a result. One group recently developed a full length genetically detoxified tetanus toxin, 8MTT. This new mutant has reported similar α-tetanus IgG levels to dTT in mice, yields of up to 500 mg L^{-1} when expressed in *E. coli*, and an antigenic purity well above the

1000 Lf units per mg of protein nitrogen specified by the WHO to ensure tetanus vaccine efficacy (World Health Organization, 2007). This protein marks a significant improvement over current genetically detoxified forms of TT such as the receptor binding domain (THc), where the complete structure induces more protective antibodies and superior protection (Chang et al., 2022).

8.2. Outer Membrane Vesicles

Aside from the current carriers, outer membrane vesicles (OMV) present a unique solution to the issue of glycan antigen delivery. Also referred to as Generalized Modules for Membrane Antigens (GMMA), these genetically engineered gram-negative derived OMV are essentially native components of bacteria that have surface exposed antigens as well as immunostimulatory molecules such as lipoproteins, peptidoglycans, and lipopolysaccharides (LPS). GMMAs are first produced in bacteria engineered for improved OMV production and altered lipid A structure (Rossi et al., 2016). This reduces the reactogenic capacity of the LPS while maintaining its immunopotentiator effect. Antigens are then linked using the chemistry established in glycoconjugate technology, such as SH-Maleimide chemistry and reductive amination. This platform affords the ability to conjugate multiple heterologous glycans offering protection against multiple organisms simultaneously, hosted on carriers that are highly expressed in efficient production systems. This hints towards the potential of GMMAs as a plug and play technology for the rapid and affordable development of much needed, under researched vaccines. Perhaps most significantly Micoli et al. (2020) report improved carrier effects when *Neisseria meningitidis* serogroups C oligosaccharides were conjugated to GMMAs compared with CRM_{197}, especially when the CRM_{197} conjugates required an

Alhydrogel adjuvant to elicit anti-saccharide-specific IgG production, where GMMA conjugates did not (Micoli et al., 2020).

8.3. Bioconjugates

Bioconjugation is a concept that promises to streamline the manufacturing of conjugate vaccines. This comprises complete *in vivo* synthesis of glycoconjugates, typically in *E. coli*, where the protein and glycan construction are performed entirely by the cell. Genetic engineering is employed to introduce one or more N-glycosylation consensus sequence to the carrier protein for site selective conjugation. *E. coli* are engineered to express a bacterial capsule derived o-antigen polysaccharide (O-PS) and an oligosaccharyl transferase which catalases the transfer of the glycan to the asparagine of the N-glycosylation site in the periplasm (Figure 7) (Ihssen et al., 2010).

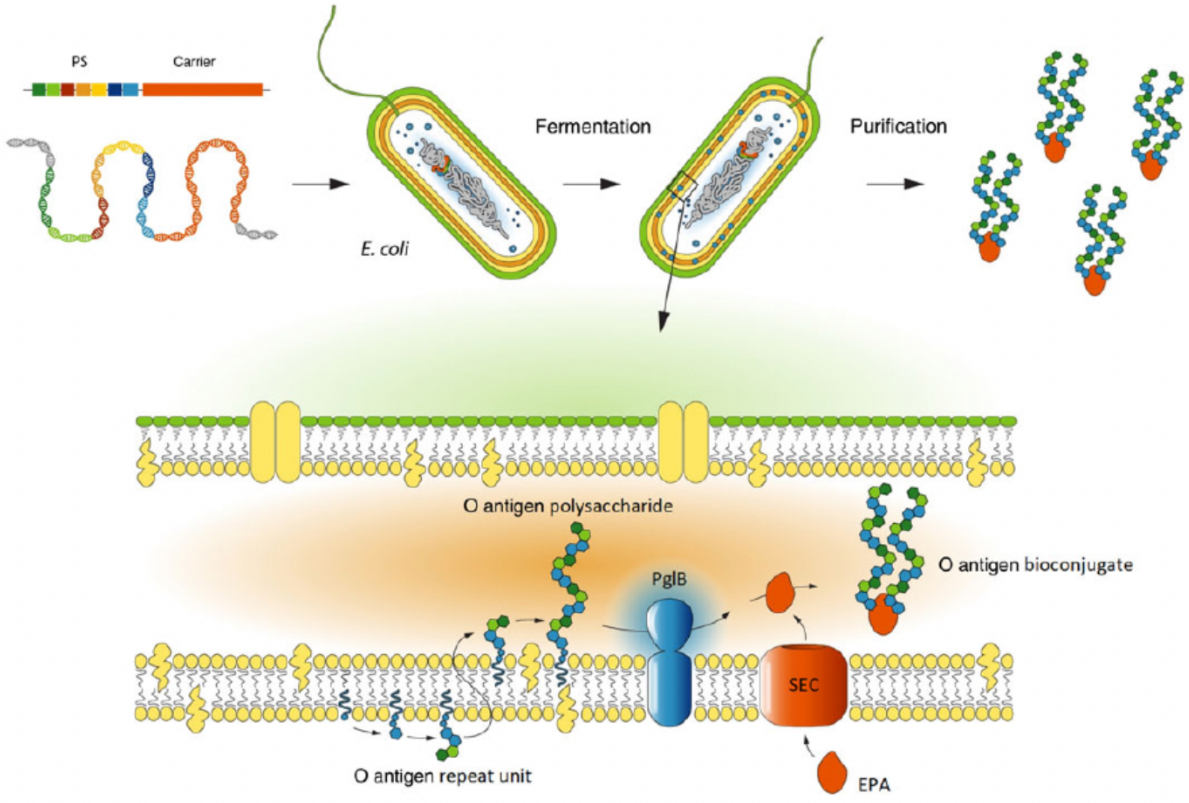

Figure 7. Typical bioconjugation pathway in *E. coli* expressing *Campylobacter* oligosaccharyltransferase (PglB), *Pseudomonas aeruginosa* exoprotein A (EPA), and polysaccharide O antigens. Abbreviation: SEC, secretory pathway (Poolman and Wacker, 2015).

Production strains and conditions can then be optimized to achieve high levels of expression such that the complete glycoconjugate can be purified from cell lysate. A caveat to this method is that the O-PS structure must be known, including the gene cluster required for its expression. The glycosylation of the protein *in vivo* is also difficult to predict and control, and the resulting glycoconjugates are often heterogenous (van der Put et al., 2023). Since the polysaccharides show polydispersity and variable repeating units' number, structural characterization of bioconjugates was considered a developmental barrier. Nicolardi et al. (2022) recently reported an ultrahigh-resolution MALDI-in-Source Decay FT-ICR Mass Spectrometry method permitting the determination of the terminal amino acid sequence and monosaccharide composition of the repeating oligosaccharide units from intact bioconjugated glycoproteins. This method presents a new rapid method for validating the structural integrity of bioconjugates without the need for separation instrumentation (Nicolardi et al., 2022). Bioconjugate technology has been extensively investigated in the search for a *Shigella* vaccine, where chemical methods have failed. The robustness, safety, and efficacy of the bioconjugation platform and *Shigella* O antigens has since been demonstrated. Subsequent development of a quadrivalent bioconjugate vaccine comprising four *Shigella* O-antigens conjugated to recombinant *Pseudomonas aeruginosa* exoprotein A (EPA) is currently undergoing clinical trials and marks the most significant advancement towards a *Shigella* vaccine to date (Martin and Alaimo, 2022).

9. Concluding Remarks

Nevertheless, CRM_{197} remains the gold standard carrier protein for glycoconjugate vaccines, and is almost always used as a comparative control when investigating the efficacy of novel carriers (Micoli et al., 2020). The extensive application, ease of processing, high-yield expression and proven safety and efficacy make CRM_{197} one of the most versatile therapeutic proteins currently available. Recent implications in the use of CRM_{197} as a carrier for cancer vaccines is a testament to this versatility (Xiong et al., 2021). KH-1 is a tumour-associated carbohydrate (TACA) and thymus-independent antigen (TD-Ag) bearing limited immunogenicity, as is the case for most TACAs. Conjugation of KH-1 to CRM_{197} induced higher IgG titres with the potential to recognise and eliminate KH-1-positive cancer cells via complement-dependent cytotoxicity (Liu et al., 2023). Researchers are also cited as investigating CRM_{197} as a carrier to improve the immunogenicity of low molecular weight SARS-CoV-2 immunogenic and antigenic moieties. A protein chimera of CRM_{197} and the Receptor-Binding-Domain (RBD) of SARS-CoV-2 spike protein was expressed in, and purified from, *E. coli*, and was shown to have improved immunogenic properties when compared to RBD alone. The potential of CRM_{197} to confer immunogenicity to low molecular mass antigens reaffirms its value as a research tool (Maltoni et al., 2022). This renewed intrigue has spurred on non-clinical research into the structure of CRM_{197}. The crystal structure of functional monomeric CRM_{197} has now been elucidated and promises to enable more thorough comparisons of the protein in different expression systems and conjugation optimization (Gallagher et al., 2023).

Despite the broad interest in CRM_{197}, the current literature is void of high-quality characterisations of the hydrodynamics of the carrier protein – a collection of physical properties pertinent to the development of new therapeutics. Hydrodynamic characterisation offers a highly accurate report of the structural integrity of a given therapeutic and is a key indicator of safety and stability. This current work aims to produce a thorough characterisation of CRM_{197}, contributing to the current knowledge of the molecule and improving the development of future CRM_{197} based therapies.

Acknowledgements

I would like to thank Stephen Harding, Jennifer Wakefield, and Jacob Pattem at the National Centre for Molecular Hydrodynamics, University of Nottingham, for their invaluable insight and guidance. Further gratitude is given to Ghislane Delpierre and Pierre Duvivier from GSK Vaccines, Rixensart, for their ongoing support throughout this project, and the School of Biosciences, University of Nottingham, for funding this research.

Bibliography

AGGERBECK, H. & HERON, I. 1992. Detoxification of diphtheria and tetanus toxin with formaldehyde. Detection of protein conjugates. *Biologicals,* 20**,** 109-115.

ALEXANDER, W. J., SHAW, J. F. E. & HOLMES, J. R. 1987. Haemophilus-influenzae infections following licensure of the Hib vaccine. *New England Journal of Medicine,* 317**,** 167-168.

ARTENSTEIN, M. S., GOLD, R., ZIMMERLY, J. G., WYLE, F. A., SCHNEIDER, H. & HARKINS, C. 1970. Prevention of Meningococcal disease by group C polysaccharide vaccine. *New England Journal of Medicine,* 282**,** 417-420.

ATTARPOUR-YAZDI, M. M., GHAMARIAN, A., MOUSAVIEHZADEH, M. & DAVOUDI, N. 2014. Identification of the serotypes of bacterial meningitis agents; implication for vaccine usage. *Iran J Microbiol,* 6**,** 211-8.

AW, R., ASHIK, M. R., ISLAM, A. A. Z. M., KHAN, I., MAINUDDIN, M., ISLAM, M. A., AHASAN, M. M. & POLIZZI, K. M. 2021. Production and purification of an active CRM197 in Pichia pastoris and its immunological characterization using a Vi-typhoid antigen vaccine. *Vaccine,* 39**,** 7379-7386.

BAYART, C., PERONIN, S., JEAN, E., PALADINO, J., TALAGA, P. & LE BORGNE, M. 2017. The combined use of analytical tools for exploring tetanus toxin and tetanus toxoid structures. *Journal of Chromatography B,* 1054**,** 80-92.

BELL, C. E. & EISENBERG, D. 1996. Crystal structure of diphtheria toxin bound to nicotinamide adenine dinucleotide. *Biochemistry,* 35**,** 1137-49.

BELL, C. E. & EISENBERG, D. 1997. Crystal structure of nucleotide-free diphtheria toxin. *Biochemistry,* 36**,** 481-8.

BENNETT, M. J., CHOE, S. & EISENBERG, D. 1994. Refined structure of dimeric diphtheria toxin at 2.0 A resolution. *Protein Sci,* 3**,** 1444-63.

BERTI, F. 2022. Expanding polysaccharide-protein coupling of glycoconjugate vaccines. *J Biol Chem,* 298**,** 101755.

BHUTTA, Z. A. & BLACK, R. E. 2013. Global maternal, newborn, and child health — So near and yet so far. *New England Journal of Medicine,* 369**,** 2226-2235.

BRÖKER, M., BERTI, F., SCHNEIDER, J. & VOJTEK, I. 2017. Polysaccharide conjugate vaccine protein carriers as a "neglected valency" – Potential and limitations. *Vaccine,* 35**,** 3286-3294.

BRÖKER, M., COSTANTINO, P., DETORA, L., MCINTOSH, E. D. & RAPPUOLI, R. 2011. Biochemical and biological characteristics of cross-reacting material 197 (CRM197), a non-toxic mutant of diphtheria toxin: Use as a conjugation protein in vaccines and other potential clinical applications. *Biologicals,* 39**,** 195-204.

BUCK, P. M., CHAUDHRI, A., KUMAR, S. & SINGH, S. K. 2015. Highly viscous antibody solutions are a consequence of network formation caused by domain–domain electrostatic complementarities: insights from coarse-grained simulations. *Molecular Pharmaceutics,* 12**,** 127-139.

CENTER FOR BIOLOGICS EVALUATION AND RESEARCH. 2022. *U.S. Food and Drug Administration* [Online]. FDA. Available: https://www.fda.gov/vaccines-blood-biologics/vaccines/vaccines-licensed-use-united-states [Accessed 13/04/2023 2023].

CHANG, M.-J., OLLIVAULT-SHIFLETT, M., SCHUMAN, R., NGOC NGUYEN, S., KALTASHOV, I. A., BOBST, C., RAJAGOPAL, S. P., PRZEDPELSKI, A., BARBIERI, J. T. & LEES, A. 2022. Genetically detoxified tetanus toxin as a vaccine and conjugate carrier protein. *Vaccine,* 40**,** 5103-5113.

CHOE, S., BENNETT, M. J., FUJII, G., CURMI, P. M. G., KANTARDJIEFF, K. A., COLLIER, R. J. & EISENBERG, D. 1992. The crystal structure of diphtheria toxin. *Nature,* 357**,** 216-222.

CHOI, C. W., MOON, J. H., KIM, J. O., YOO, S. H., KIM, H. G., KIM, J. H., PARK, T. J. & KIM, S. S. 2018. Evaluation of potency on diphtheria and tetanus toxoid for adult vaccines by in vivo toxin neutralization assay using national reference standards. *Osong Public Health Res Perspect,* 9**,** 278-282.

CHU, C., SCHNEERSON, R., ROBBINS, J. B. & RASTOGI, S. C. 1983. Further studies on the immunogenicity of Haemophilus influenzae type b and pneumococcal type 6A polysaccharide-protein conjugates. *Infection and Immunity,* 40**,** 245-256.

CHUNG, Y.-J., LEE, J.-A., JUNG, M.-Y., LEE, S.-M., KIM, T.-Y., CHOE, Y.-K. & KIM, I.-H. 2016. Optimization of diphtheria toxin production by Corynebacterium diphtheriae using a casein-based medium in a fermenter. *Biotechnology and Bioprocess Engineering,* 21**,** 537-543.

COLE, J. L., LARY, J. W., T, P. M. & LAUE, T. M. 2008. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol,* 84**,** 143-79.

COX, M. M. J., PATRIARCA, P. A. & TREANOR, J. 2008. FluBlok, a recombinant hemagglutinin influenza vaccine. *Influenza and Other Respiratory Viruses,* 2**,** 211-219.

CYSTER, J. G. & ALLEN, C. D. C. 2019. B cell responses: Cell interaction dynamics and decisions. *Cell,* 177**,** 524-540.

DAM, J., VELIKOVSKY, C. A., MARIUZZA, R. A., URBANKE, C. & SCHUCK, P. 2005. Sedimentation velocity analysis of heterogeneous protein-protein interactions: Lamm equation modeling and sedimentation coefficient distributions c(s). *Biophysical Journal,* 89**,** 619-634.

DE, N., NAVARRO, M. V. A. S., WANG, Q., KRASTEVA, P. V. & SONDERMANN, H. 2010. Chapter 10 - Biophysical assays for protein interactions in the Wsp sensory system and biofilm formation. *Methods in Enzymology.* Academic Press.

DE-SIMONE, S. G., GOMES, L. R., NAPOLEÃO-PÊGO, P., LECHUGA, G. C., DE PINA, J. S. & DA SILVA, F. R. 2021. Epitope mapping of the diphtheria toxin and development of an ELISA-specific diagnostic assay. *Vaccines (Basel),* 9.

DE-SIMONE, S. G., NAPOLEÃO-PÊGO, P., LECHUGA, G. C., CARVALHO, J. P. R. S., GOMES, L. R., CARDOZO, S. V., MOREL, C. M., PROVANCE, D. W. & SILVA, F. R. D. 2023. High-throughput IgG epitope mapping of Tetanus neurotoxin: Implications for immunotherapy and vaccine design. *Toxins,* 15**,** 239.

DONNIO, P. Y., LERESTIF-GAUTIER, A. L. & AVRIL, J. L. 2004. Characterization of Pasteurella spp. strains isolated from human infections. *J Comp Pathol,* 130**,** 137-42.

DUKE, J. A., PASCHALL, A. V., GLUSHKA, J., LEES, A., MOREMEN, K. W. & AVCI, F. Y. 2022. Harnessing galactose oxidase in the development of a chemoenzymatic platform for glycoconjugate vaccine design. *J Biol Chem,* 298**,** 101453.

EDWARDS, G. B., MUTHURAJAN, U. M., BOWERMAN, S. & LUGER, K. 2020. Analytical Ultracentrifugation (AUC): An overview of the application of fluorescence and absorbance AUC to the study of biological macromolecules. *Curr Protoc Mol Biol,* 133**,** e131.

FORSGREN, A. & RIESBECK, K. 2008. Protein D of Haemophilus influenzae: A protective nontypeable H. influenzae antigen and a carrier for Pneumococcal conjugate vaccines. *Clinical Infectious Diseases,* 46**,** 726-731.

FRANCIS, A. I., GHANY, S., GILKES, T. & UMAKANTHAN, S. 2022. Review of COVID-19 vaccine subtypes, efficacy and geographical distributions. *Postgraduate Medical Journal,* 98**,** 389-394.

FRASCH, C. E. 2009. Preparation of bacterial polysaccharide–protein conjugates: Analytical and manufacturing challenges. *Vaccine,* 27**,** 6468-6470.

GALLAGHER, D. T., OGANESYAN, N. & LEES, A. 2023. Monomeric crystal structure of the vaccine carrier protein CRM(197) and implications for vaccine development. *Acta Crystallogr F Struct Biol Commun,* 79**,** 82-86.

GAO, C., ZENG, J., JIA, N., STAVENHAGEN, K., MATSUMOTO, Y., ZHANG, H., LI, J., HUME, A. J., MÜHLBERGER, E., VAN DIE, I., KWAN, J., TANTISIRA, K., EMILI, A. & CUMMINGS, R. D. 2020. SARS-CoV-2 spike protein interacts with multiple innate immune receptors. *bioRxiv*.

GASPARINI, R., JOHNSTON, W., CONVERSANO, M., GARSCADDEN, A., ALEXANDERIAN, D., GIGLIOLI, N., PERCELL, S., HAN, L. & SMOLENOV, I. 2014. Immunogenicity and safety of combined tetanus, reduced diphtheria, acellular pertussis vaccine when co-administered with quadrivalent meningococcal conjugate and human papillomavirus vaccines in healthy adolescents. *J Vaccines Vaccin,* 5**,** 231.

GENO, K. A., GILBERT, G. L., SONG, J. Y., SKOVSTED, I. C., KLUGMAN, K. P., JONES, C., KONRADSEN, H. B. & NAHM, M. H. 2015. Pneumococcal capsules and their types: Past, present, and future. *Clinical Microbiology Reviews,* 28**,** 871-899.

GIANNINI, G., RAPPUOLI, R. & RATTI, G. 1984. The amino-acid sequence of two non-toxic mutants of diphtheria toxin: CRM45 and CRM197. *Nucleic acids research,* 12**,** 4063-4069.

GOFFIN, P., DEWERCHIN, M., DE ROP, P., BLAIS, N. & DEHOTTAY, P. 2017. High-yield production of recombinant CRM197, a non-toxic mutant of diphtheria toxin, in the periplasm of Escherichia coli. *Biotechnol J,* 12.

GOTSCHLICH, E. C., AUSTRIAN, R., CVJETANOVIC, B. & ROBBINS, J. B. 1978. Prospects for prevention of bacterial-meningitis with polysaccharide vaccines *Bulletin of the World Health Organization,* 56**,** 509-518.

GUAN, L., ZHANG, L., XUE, Y., YANG, J. & ZHAO, Z. 2020. Molecular pathogenesis of the hyaluronic acid capsule of Pasteurella multocida. *Microbial Pathogenesis,* 149**,** 104380.

HANSEN, J. C., LEBOWITZ, J. & DEMELER, B. 1994. Analytical ultracentrifugation of complex macromolecular systems. *Biochemistry,* 33**,** 13155-13163.

HARDING, S. E. 1997. The intrinsic viscosity of biological macromolecules. Progress in measurement, interpretation and application to structure in dilute solution. *Prog Biophys Mol Biol,* 68**,** 207-62.

HARDING, S. E., ADAMS, G. G., ALMUTAIRI, F., ALZAHRANI, Q., ERTEN, T., KÖK, M. S. & GILLIS, R. B. 2015. Ultracentrifuge methods for the analysis of polysaccharides, glycoconjugates, and lignins. *Methods Enzymol,* 562**,** 391-439.

HICKEY, J. M., TOPRANI, V. M., KAUR, K., MISHRA, R. P. N., GOEL, A., OGANESYAN, N., LEES, A., SITRIN, R., JOSHI, S. B. & VOLKIN, D. B. 2018. Analytical comparability assessments of 5 recombinant

CRM197 proteins from different manufacturers and expression systems. *Journal of Pharmaceutical Sciences,* 107**,** 1806-1819.

HOLT, K. E., LASSALLE, F., WYRES, K. L., WICK, R. & MOSTOWY, R. J. 2020. Diversity and evolution of surface polysaccharide synthesis loci in Enterobacteriales. *The ISME Journal,* 14**,** 1713-1730.

HOLZ, E., DARWISH, M., TESAR, D. B. & SHATZ-BINDER, W. 2023. A review of protein- and peptide-based chemical conjugates: Past, present, and future. *Pharmaceutics,* 15**,** 600.

IHSSEN, J., KOWARIK, M., DILETTOSO, S., TANNER, C., WACKER, M. & THÖNY-MEYER, L. 2010. Production of glycoprotein vaccines in Escherichia coli. *Microb Cell Fact,* 9**,** 61.

JAUNEIKAITE, E., TOCHEVA, A. S., JEFFERIES, J. M., GLADSTONE, R. A., FAUST, S. N., CHRISTODOULIDES, M., HIBBERD, M. L. & CLARKE, S. C. 2015. Current methods for capsular typing of Streptococcus pneumoniae. *J Microbiol Methods,* 113**,** 41-9.

JIANHUA, Z. & PETRACCA, R. 1999. Secretory expression of recombinant diphtheria toxin mutants in B. Subtilis. *Journal of Tongji Medical University,* 19**,** 253-256.

JONES, C. 2015. Glycoconjugate vaccines: The regulatory framework. *In:* LEPENIES, B. (ed.) *Carbohydrate-Based Vaccines: Methods and Protocols.* New York, NY: Springer New York.

JONES, H. E., TAYLOR, P. R., MCGREAL, E., ZAMZE, S. & WONG, S. Y. C. 2009. The contribution of naturally occurring IgM antibodies, IgM cross-reactivity and complement dependency in murine humoral responses to pneumococcal capsular polysaccharides. *Vaccine,* 27**,** 5806-5815.

KALLET, A. & AARON, H. 2021. Two New Pneumococcal Vaccines-Prevnar 20 and Vaxneuvance (Reprinted from Medical Letter on Drugs and Therapeutics, vol 63, pg 188-190, 2021). *Jama-Journal of the American Medical Association,* 326**,** 2521-2522.

KEINHORSTER, D., GEORGE, S. E., WEIDENMAIER, C. & WOLZ, C. 2019. Function and regulation of Staphylococcus aureus wall teichoic acids and capsular polysaccharides. *International Journal of Medical Microbiology,* 309.

KELLY, D. F., POLLARD, A. J. & MOXON, E. R. 2005. Immunological memory: the role of B cells in long-term protection against invasive bacterial pathogens. *Jama,* 294**,** 3019-23.

KENNEY, J. P. L. & FEIN, J. B. 2011. Importance of extracellular polysaccharides on proton and Cd binding to bacterial biomass: A comparative study. *Chemical Geology,* 286**,** 109-117.

KLECKER, C. & NAIR, L. S. 2017. Chapter 13 - Matrix chemistry controlling stem cell behavior. *In:* VISHWAKARMA, A. & KARP, J. M. (eds.) *Biology and Engineering of Stem Cell Niches.* Boston: Academic Press.

KOLB, H. C., FINN, M. G. & SHARPLESS, K. B. 2001. Click chemistry: Diverse chemical function from a few good reactions. *Angewandte Chemie-International Edition,* 40**,** 2004-+.

KONG, G., LIM, N. A., CHIN, Y. H., NG, Y. P. M. & AMIN, Z. 2022. Effect of COVID-19 pandemic on influenza vaccination intention: A meta-analysis and systematic review. *Vaccines (Basel),* 10.

LALANI, H. S., NAGAR, S., SARPATWARI, A., BARENIE, R. E., AVORN, J., ROME, B. N. & KESSELHEIM, A. S. 2023. US public investment in development of mRNA covid-19 vaccines: retrospective cohort study. *BMJ,* 380**,** e073747.

LEBOWITZ, J., LEWIS, M. S. & SCHUCK, P. 2002. Modern analytical ultracentrifugation in protein science: A tutorial review. *Protein Science,* 11**,** 2067-2079.

LEES, A., BARR, J. F. & GEBRETNSAE, S. 2020. Activation of soluble polysaccharides with 1-Cyano-4-Dimethylaminopyridine Tetrafluoroborate (CDAP) for use in protein–polysaccharide conjugate vaccines and immunological reagents. III optimization of CDAP activation. *Vaccines,* 8**,** 777.

LEES, A., PUVANESARAJAH, V. & FRASCH, C. E. 2008. Conjugation chemistry. *Pneumococcal Vaccines: The Impact of Conjugate Vaccines***,** 161-174.

LICONA-CASSANI, C., STEEN, J. A., ZARAGOZA, N. E., MOONEN, G., MOUTAFIS, G., HODSON, M. P., POWER, J., NIELSEN, L. K. & MARCELLIN, E. 2016. Tetanus toxin production is triggered by the transition from amino acid consumption to peptides. *Anaerobe,* 41**,** 113-124.

LIPSITCH, M. 1997. Vaccination against colonizing bacteria with multiple serotypes. *Proc Natl Acad Sci U S A,* 94**,** 6571-6.

LIU, C. C. & SCHULTZ, P. G. 2010. Adding new chemistries to the genetic code. *Annu Rev Biochem,* 79**,** 413-44.

LIU, Y., LI, B., ZHENG, X., XIONG, D. & YE, X. 2023. Cancer vaccines based on fluorine-modified KH-1 elicit robust immune response. *Molecules,* 28**,** 1934.

LONG, Z., WEI, C., ROSS, R., LUO, X., MA, X., QI, Y., CHAI, R., CAO, J., HUANG, M. & BO, T. 2022. Effects of detoxification process on toxicity and foreign protein of tetanus toxoid and diphtheria toxoid. *Journal of Chromatography B,* 1207**,** 123377.

LUETSCHER, R. N. D., MCKITRICK, T. R., GAO, C., MEHTA, A. Y., MCQUILLAN, A. M., KARDISH, R., BOLIGAN, K. F., SONG, X., LU, L., HEIMBURG-MOLINARO, J., VON GUNTEN, S., ALTER, G. & CUMMINGS, R. D. 2020. Unique repertoire of anti-carbohydrate antibodies in individual human serum. *Scientific Reports,* 10**,** 15436.

MACLEOD , C. M., HODGES , R. G., HEIDELBERGER , M. & BERNHARD , W. G. 1945. Prevention of Pneumococcal pneumonia by immunization with specific capsular polysaccharides *Journal of Experimental Medicine,* 82**,** 445-465.

MAGAZINE, N., ZHANG, T. Y., WU, Y. Y., MCGEE, M. C., VEGGIANI, G. & HUANG, W. S. 2022. Mutations and evolution of the SARS-CoV-2 spike protein. *Viruses-Basel,* 14.

MALITO, E., BURSULAYA, B., CHEN, C., LO SURDO, P., PICCHIANTI, M., BALDUCCI, E., BIANCUCCI, M., BROCK, A., BERTI, F., BOTTOMLEY, M. J., NISSUM, M., COSTANTINO, P., RAPPUOLI, R. & SPRAGGON, G. 2012. Structural basis for lack of toxicity of the diphtheria toxin mutant CRM197. *Proc Natl Acad Sci U S A,* 109**,** 5229-34.

MALITO, E., CARFI, A. & BOTTOMLEY, M. J. 2015. Protein crystallography in vaccine research and development. *Int J Mol Sci,* 16**,** 13106-40.

MALTONI, G., SCUTTERI, L., MENSITIERI, F., PIAZ, F. D. & HOCHKOEPPLER, A. 2022. High-yield production in Escherichia coli and convenient purification of a candidate vaccine against SARS-CoV-2. *Biotechnology Letters,* 44**,** 1313-1322.

MARTIN, P. & ALAIMO, C. 2022. The ongoing journey of a Shigella bioconjugate vaccine. *Vaccines (Basel),* 10.

MASUYER, G., CONRAD, J. & STENMARK, P. 2017. The structure of the tetanus toxin reveals pH-mediated domain dynamics. *EMBO Rep,* 18**,** 1306-1317.

MESIN, L., ERSCHING, J. & VICTORA, G. D. 2016. Germinal center B cell dynamics. *Immunity,* 45**,** 471-482.

METZ, B., MICHIELS, T., UITTENBOGAARD, J., DANIAL, M., TILSTRA, W., MEIRING, H. D., HENNINK, W. E., CROMMELIN, D. J. A., KERSTEN, G. F. A. & JISKOOT, W. 2020. Identification of formaldehyde-induced modifications in Diphtheria toxin. *J Pharm Sci,* 109**,** 543-557.

METZ, B., TILSTRA, W., VAN DER PUT, R., SPRUIT, N., VAN DEN IJSSEL, J., ROBERT, J., HENDRIKSEN, C. & KERSTEN, G. 2013. Physicochemical and immunochemical assays for monitoring consistent production of tetanus toxoid. *Biologicals,* 41**,** 231-237.

MICOLI, F., ADAMO, R. & COSTANTINO, P. 2018. Protein carriers for glycoconjugate vaccines: History, selection criteria, characterization and new trends. *Molecules,* 23.

MICOLI, F., ALFINI, R., DI BENEDETTO, R., NECCHI, F., SCHIAVO, F., MANCINI, F., CARDUCCI, M., PALMIERI, E., BALOCCHI, C., GASPERINI, G., BRUNELLI, B., COSTANTINO, P., ADAMO, R., PICCIOLI, D. & SAUL, A. 2020. GMMA Is a versatile platform to design effective multivalent vombination vaccines. *Vaccines (Basel),* 8.

MISHRA, R. P. N., YADAV, R. S. P., JONES, C., NOCADELLO, S., MINASOV, G., SHUVALOVA, LUDMILLA A., ANDERSON, WAYNE F. & GOEL, A. 2018. Structural and immunological characterization of E. coli derived recombinant CRM197 protein used as carrier in conjugate vaccines. *Bioscience Reports,* 38.

MORAIS, V. & SUAREZ, N. 2022. Conjugation mechanism for Pneumococcal glycoconjugate vaccines: Classic and emerging methods. *Bioengineering (Basel),* 9.

MURPHY, J. R. 1996. Corynebacterium diphtheriae. *In:* BARON, S. (ed.) *Medical Microbiology.* Galveston (TX): University of Texas Medical Branch at Galveston

NICOLARDI, S., DANUSER, R., DOTZ, V., DOMÍNGUEZ-VEGA, E., AL KAABI, A., BEURRET, M., ANISH, C. & WUHRER, M. 2022. Glycan and protein analysis of glycoengineered bacterial E. coli vaccines by MALDI-in-Source Decay FT-ICR Mass Spectrometry. *Anal Chem,* 94**,** 4979-4987.

NILO, A., PASSALACQUA, I., FABBRINI, M., ALLAN, M., USERA, A., CARBONI, F., BROGIONI, B., PEZZICOLI, A., COBB, J., ROMANO, M. R., MARGARIT, I., HU, Q. Y., BERTI, F. & ADAMO, R. 2015. Exploring the effect of conjugation site and chemistry on the immunogenicity of an anti-group B Streptococcus glycoconjugate vaccine based on GBS67 pilus protein and type V polysaccharide. *Bioconjug Chem,* 26**,** 1839-49.

OGAWA, T. & HIROKAWA, N. 2018. Multiple analyses of protein dynamics in solution. *Biophys Rev,* 10**,** 299-306.

PARK, A. R., JANG, S. W., KIM, J. S., PARK, Y. G., KOO, B. S. & LEE, H. C. 2018. Efficient recovery of recombinant CRM197 expressed as inclusion bodies in E.coli. *PLoS One,* 13**,** e0201060.

PICHICHERO, M. E. 2013. Protein carriers of conjugate vaccines: characteristics, development, and clinical trials. *Hum Vaccin Immunother,* 9**,** 2505-23.

PILLOT, A., DEFONTAINE, A., FATEH, A., LAMBERT, A., PRASANNA, M., FANUEL, M., PIPELIER, M., CSABA, N., VIOLO, T., CAMBERLEIN, E. & GRANDJEAN, C. 2019. Site-specific conjugation for fully controlled glycoconjugate vaccine preparation. *Front Chem,* 7**,** 726.

POOLMAN, J., FRASCH, C., NURKKA, A., KÄYHTY, H., BIEMANS, R. & SCHUERMAN, L. 2011. Impact of the conjugation method on the immunogenicity of Streptococcus pneumoniae serotype 19F polysaccharide in conjugate vaccines. *Clin Vaccine Immunol,* 18**,** 327-36.

POOLMAN, J. & WACKER, M. 2015. Extra-intestinal pathogenic Escherichia coli (ExPEC), a common human pathogen: challenges for vaccine development and progress in the field. *The Journal of infectious diseases,* 213.

PORRO, M., SALETTI, M., NENCIONI, L., TAGLIAFERRI, L. & MARSILI, I. 1980. Immunogenic correlation between cross-reacting material (CRM197) produced by a mutant of Corynebacterium diphtheriae and diphtheria toxoid. *Journal of Infectious Diseases,* 142**,** 716-724.

RAPPUOLI, R. 2018. Glycoconjugate vaccines: Principles and mechanisms. *Sci Transl Med,* 10.

RAPPUOLI, R., DE GREGORIO, E. & COSTANTINO, P. 2019. On the mechanisms of conjugate vaccines. *Proceedings of the National Academy of Sciences,* 116**,** 14-16.

RICHMOND, P., KACZMARSKI, E., BORROW, R., FINDLOW, J., CLARK, S., MCCANN, R., HILL, J., BARKER, M. & MILLER, E. 2000. Meningococcal C polysaccharide vaccine induces immunologic hyporesponsiveness in adults that is overcome by meningococcal C conjugate vaccine. *J Infect Dis,* 181**,** 761-4.

RODNIN, M. V., KASHIPATHY, M. M., KYRYCHENKO, A., BATTAILE, K. P., LOVELL, S. & LADOKHIN, A. S. 2020. Structure of the Diphtheria toxin at acidic pH: Implications for the conformational switching of the translocation domain. *Toxins,* 12**,** 704.

ROSSI, O., CABONI, M., NEGREA, A., NECCHI, F., ALFINI, R., MICOLI, F., SAUL, A., MACLENNAN, C. A., RONDINI, S. & GERKE, C. 2016. Toll-Like receptor activation by generalized modules for membrane antigens from lipid A mutants of Salmonella enterica Serovars typhimurium and Enteritidis. *Clin Vaccine Immunol,* 23**,** 304-14.

ROUPHAEL, N. G., ZIMMER, S. M. & STEPHENS, D. S. 2009. Chapter 53 - Neisseria meningitidis. *In:* BARRETT, A. D. T. & STANBERRY, L. R. (eds.) *Vaccines for Biodefense and Emerging and Neglected Diseases.* London: Academic Press.

RUDD, B. D. 2020. Neonatal T vells: A reinterpretation. *Annu Rev Immunol,* 38**,** 229-247.

SCHIAVO, G., BENFENATI, F., POULAIN, B., ROSSETTO, O., POLVERINO DE LAURETO, P., DASGUPTA, B. R. & MONTECUCCO, C. 1992. Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin. *Nature,* 359**,** 832-5.

SCHNEERSON, R., BARRERA, O., SUTTON, A. & ROBBINS, J. B. 1980. Preparation, characterization, and immunogenicity of Haemophilus influenzae type b polysaccharide-protein conjugates. *The Journal of experimental medicine,* 152**,** 361-376.

SINGH, J. K., ADAMS, F. G. & BROWN, M. H. 2018. Diversity and Function of Capsular Polysaccharide in Acinetobacter baumannii. *Front Microbiol,* 9**,** 3301.

SOME, D., AMARTELY, H., TSADOK, A. & LEBENDIKER, M. 2019. Characterization of proteins by size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). *J Vis Exp*.

TOMCZYK, S., BENNETT, N. M., STOECKER, C., GIERKE, R., MOORE, M. R., WHITNEY, C. G., HADLER, S. & PILISHVILI, T. 2014. Use of 13-valent Pneumococcal conjugate vaccine and 23-valent Pneumococcal polysaccharide vaccine among adults aged >= 65 Years: Recommendations of the advisory committee on immunization practices (ACIP). *Mmwr-Morbidity and Mortality Weekly Report,* 63**,** 822-825.

UCHIDA, T., PAPPENHEIMER, A. M. & HARPER, A. A. 1972. Reconstitution of Diphtheria toxin from two nontoxic cross-reacting mutant proteins. *Science,* 175**,** 901-903.

VAN DER PUT, R. M. F., METZ, B. & PIETERS, R. J. 2023. Carriers and antigens: New developments in glycoconjugate vaccines. *Vaccines (Basel),* 11.

VAN PUTTEN, J. & TØNJUM, T. 2010. Chapter 168 - Neisseria. *In:* COHEN, J., OPAL, S. M. & POWDERLY, W. G. (eds.) *Infectious Diseases (Third Edition).* London: Mosby.

VIOLO, T., LAMBERT, A., PILLOT, A., FANUEL, M., MAC-BÉAR, J., BROUSSARD, C., GRANDJEAN, C. & CAMBERLEIN, E. 2023. Site-selective unnatural amino acid incorporation at single or multiple positions to control sugar-protein connectivity in glycoconjugate vaccine candidates. *Chemistry – A European Journal,* 29**,** e202203497.

WEINTRAUB, A. 2003. Immunology of bacterial polysaccharide antigens. *Carbohydrate Research,* 338**,** 2539-2547.

WORLD HEALTH, O. 2022. *WHO expert committee on biological standardization: seventy-fifth report,* Geneva, World Health Organization.

WORLD HEALTH ORGANISATION 2013. Haemophilus influenzae type b (Hib) vaccination position paper – July 2013. *Wkly Epidemiol Rec,* 88**,** 413-26.

WORLD HEALTH ORGANIZATION 2007. Calibration of replacement international standard of tetanus toxoid for use in flocculation test. *Calibration of replacement international standard of tetanus toxoid for use in flocculation test.*

XIA, X. H. 2021. Domains and functions of spike protein in SARS-Cov-2 in the context of vaccine design. *Viruses-Basel,* 13.

XIONG, A. W., FANG, J. M., REN, S. X., LI, W., WANG, J., ZHAO, Y., CHEN, G. Y., XU, Q. & ZHOU, C. C. 2021. A novel combined conjugate therapeutic cancer vaccine, recombinant EGF-CRM197, in patients with advanced solid tumors: A phase I clinical study. *Frontiers in Oncology,* 11.

ZANDVOORT, A., LODEWIJK, M. E., DE BOER, N. K., DAMMERS, P. M., KROESE, F. G. & TIMENS, W. 2001. CD27 expression in the human splenic marginal zone: the infant marginal zone is populated by naive B cells. *Tissue Antigens,* 58**,** 234-42.

Exploring the hydrodynamic properties of CRM(197) and CRM-conjugated vaccine candidates: implications for innovative therapeutic development

Oliver Meeds

styom@nottingham.ac.uk,
University of Nottingham,
School of Biosciences,
Sutton Bonington, LE12 5RD

(5775 words)

Abstract

Conjugation of thymus independent polysaccharide antigens to immunogenic carrier proteins improves their immunogenicity and confers long lasting immunity. This approach has been employed for decades to protect global populations against bacterial infections. With the omnipresent threat of antibiotic resistance growing ever greater, glycoconjugate technology has seen a resurgence as a strategy to prevent lesser-researched bacterial infections before antibiotic treatment is required. Despite recent advances, the genetically detoxified diphtheria toxin CRM_{197} remains the gold standard carrier protein and has also been investigated as a carrier of cancer antigens. As glycoconjugate technology experiences a research-renaissance an added pressure is placed on manufacturers to produce high quality therapeutic constituents. Improving the knowledge of molecular characteristics of vaccine components is therefore a crucial cog in the developmental machine. In this current work, we report on the high inherent stability of CRM_{197} under appropriate storage conditions but highlight an unusual tendency to aggregate under forces like those exerted on the protein during manufacturing or transportation. Further comment is made regarding the influence of polysaccharide serotype on therapeutic hydrodynamics, emphasizing the importance of structural considerations in vaccine design.

Table of Contents

1. Introduction

Of the many crises currently faced by clinicians, the increased prevalence of antibiotic resistance takes the top spot. Continued misuse of antibiotics, lack of public awareness, and insufficient education has led us to a situation where the number of antibiotic resistance infections is rising year on year (UK Health Security Agency, 2022). Given the drastic consequences that widespread antibiotic resistance would cause to global health and the modest number of newly approved antibiotics in the last decade, researchers are turning towards prevention rather than treatment as a means to subdue bacterial infections within the population (Chahine et al., 2022, Ye and Chen, 2022).

Vaccination is an approach that has been used for decades to prevent bacterial infections. Alongside proteins and other antigenic structures, many bacteria have highly diverse capsular polysaccharides (CPS) coating their exterior, granting the organism binding capability and protection (Singh et al., 2018). These structures can be a nuisance as they help pathogens evade immune detection by mimicking host tissue molecules to conceal antigens and cause chronic disease (Cress et al., 2014). Fortunately, as a structural component with rich serotypic diversity they also present an abundance of bacterial antigens from which vaccine formulations can be developed (Geno et al., 2015).

Initial vaccine formulations looked to protect against pneumococcal infections by taking highly pure extracts of these CPS and administering them intravenously (MacLeod et al., 1945). This type of vaccination proved fairly effective in adults, despite polysaccharide being T-independent antigens (Jones et al., 2009). It would later be discovered that the mechanism leading to the immune response in adults relies heavily on the cooperation of B cells and helper T cells (Rappuoli, 2018). This mechanism is not fully developed in neonates and

children due to their expression of an immunodominant lineage of CD4 helper T cells (Rudd, 2020). This rendered early polysaccharide vaccine formulations ineffective in children. A breakthrough occurred when capsular polysaccharides were conjugated to a carrier protein to produce the first conjugated vaccine. Licenced in 1987, the first *Haemophilus influenzae* type b (Hib) conjugate vaccine comprising purified polysaccharide capsule polyribosylribitol phosphate conjugate to diphtheria toxoid reduced the incidence of sever Hib infections in children under the age of 5 by 90% (Alexander et al., 1987, World Health Organisation, 2013). Conjugation of the polysaccharide antigen to a carrier protein introduces the T cell epitopes absent from the sugar. These facilitate B-T cooperation, allowing high affinity anti-polysaccharide immunoglobulin G's (IgG) to proliferate and generating sufficient memory cells to grant long lasting immunity (Rappuoli et al., 2019).

It is this technology that continues to provide immunity against common bacterial infections for children and adults alike (Center for Biologics Evaluation and Research, 2022). A recent resurgence in glycoconjugation has seen vaccine development considered for historically overlooked bacteria such as *Streptococcus agalactiae* by conjugation of CPS to the carrier protein, CRM_{197} (Carreras-Abad et al., 2020).

CRM_{197} is one of the most frequently used carrier proteins for glycoconjugate vaccine formulations. As a genetically detoxified form of the diphtheria toxin, it requires minimal processing prior to application (Aw et al., 2021). Other carriers such as detoxified tetanus toxin (dTT) and detoxified diphtheria toxin (dDT) require stringent detoxification which can disrupt the protein's structure, introduce random structural changes, and interfere with T cell epitopes (Long et al., 2022). It is thought that the retainment of its native structure leads to an improved carrier effect for CRM_{197} compared to dTT and dDT (Mishra et al., 2018). Despite recent advances in glycoconjugate technology CRM_{197} is still the gold standard carrier protein.

It is often used to compare the efficacy of novel conjugate technologies such as Generalized Modules for Membrane Antigens (GMMA) and has been investigated as a carrier itself to improve the immune response against cancer and SARS-CoV-2 antigens (Micoli et al., 2020, Maltoni et al., 2022, Xiong et al., 2021).

With the development of new therapeutics there is an added emphasis on the physical characterisation of both the individual components and the fully assembled vaccine. Novel vaccine formulations and therapeutics must meet the strict guidelines established by regulatory authorities such as the world health organisation (World Health, 2022). These guidelines are implemented at each stage of development and work to ensure high quality manufacturing of structurally integral components that are both safe and efficacious (Jones, 2015).

This current work aims to contribute to the wealth of knowledge of the carrier protein CMR_{197}, elucidating key information regarding its stability and conformational through a range of hydrodynamic tests. This information is pertinent to its both the manufacture and regulation of the protein. Comparisons will be drawn to the closely related dTT with consideration as to how chemical means of detoxification may affect some of the hydrodynamic properties. Hydrodynamics characterisation of four novel vaccine formulations comprising *Streptococcus agalactiae* CPS conjugated to CRM_{197} will also be performed to investigate how polysaccharide serotype influences therapeutic hydrodynamics.

2. Materials and Methods

2.1. Buffer Preparation

Phosphate buffer saline pH 7.0 adjusted to an ionic strength I = 0.1 mol.L^{-1} (PBS) was produced following the formulation outlined in Table 1 and used as a solvent and reference buffer for all hydrodynamic experiments. The density of PBS at 20.0°C was measured as (1.00334 ± 0.0001) g.mL^{-1} using a densitometer. The viscosity of PBS at 20.0°C was measured as (0.010200 ± 0.00001) Poise using an Anton Paar rolling ball viscometer (Anton Paar GmbH, Graz, Austria).

Table 1. Constituents and their relative masses for the formulation of phosphate buffer saline pH 7.0 adjusted to an ionic strength I = 0.1 mol.L^{-1}.

Component	Mass (g)
Disodium hydrogen orthophosphate dodecahydrate ($Na_2HPO_4 \cdot 12H_2O$)	4.5950 ± 0.0005
Potassium dihydrogen orthophosphate (KH_2PO_4)	1.5610 ± 0.0005
Sodium chloride (NaCl)	2.9230 ± 0.0005
Deionized water (H_2O)	1000

2.2. Samples

Samples of CRM_{197} protein and CRM-Conjugated vaccine candidates were kindly provided by GSK vaccines, Rixensart. The partial specific volume ($\bar{\nu}$) of CRM_{197} was determined to be 0.733 mL.g^{-1} using the routine SEDNTRP (Laue, 1992) from amino acid sequence. $\bar{\nu}$ of each conjugate was determined by applying the protein : polysaccharide concentration ratio and previously determined $\bar{\nu}$'s for CRM_{197} and the unconjugated polysaccharides (MacCalman, 2023)(Table 2).

Table 2. Formulations and supplementary properties for four samples of different serotypes of *Streptococcus agalactiae* conjugated to CRM_{197}.

Conjugate	Serotype	Concentration PS (μg)	Concentration Protein (μg)	$\overline{\nu}$ PS ($mL.g^{-1}$)	$\overline{\nu}$ CRM_{197} ($mL.g^{-1}$)	$\overline{\nu}$ conjugate ($mL.g^{-1}$)	dn/dc ($mL.g^{-1}$)
GBA-CRM-026	Ia	865	783	0.642	0.733	0.683	0.166
GBB-CRM-008B	Ib	1421	1251	0.622	0.733	0.670	0.166
GBD-CRM-005	II	842	669	0.627	0.733	0.666	0.165
GBT-CRM-003	III	1320	861	0.661	0.733	0.661	0.163

N.B. All sample and their information were kindly provided by Ghislain Delpierre of GlaxoSmithKline Pharmaceuticals, Avenue Pascal 2-4-6 1300 Wavre, Belgium. PS = polysaccharide.

2.3. Sample Preparation

Samples of CRM_{197} were thawed for one hour at 37.0°C before being transferred to a Thermofisher Biodesign™ cellulose dialysis tubing 14000 mwco dialysis bag (Thermofisher, Massachusetts, USA) and dialysed in PBS overnight at 4.0°C. Conjugates were stored at 4.0°C and dialysed directly before each experiment. Both CRM_{197} and each of the conjugates were subsequently stored at 4.0°C. Absorbance of stock samples of CRM_{197} were measured using a Jenway 7305 spectrophotometer (Jenway, Staffordshire, UK) absorbing at 280 nm to elucidate concentration by application of the Beer Lambert law:

$$A = \varepsilon c_m l \qquad (1)$$

Where A is absorbance at 280 nm, ε is the extinction coefficient of the sample, c_m is the molar concentration of the sample in $mol.L^{-1}$, and l is the optical pathlength. Molar concentration is then converted to weight concentration by:

$$c = c_m M \qquad (2)$$

Where c is the weight concentration in mg.mL^{-1}, c_m is molar concentration, M is the molecular weight. Stock concentrations of CRM_{197} were (25.0 ± 0.5) mg.mL^{-1}.

Concentrations of each glycoconjugate vaccine candidate were determined by refractometry using an Atago DD-7 digital refractometer (ATAGO CO., Tokyo, Japan) to calculate BRIX % at 20°C. BRIX % is then converted to concentration using the equation:

$$c = BRIX\% \cdot \frac{\frac{dn}{dc_{sample}}}{\frac{dn}{dc_{sucrose}}} \cdot 10 \qquad (3)$$

Where c is weight concentration in mg.mL^{-1}, $\frac{dn}{dc_{sample}}$ is the refractive increment of each sample (Table 2), and $\frac{dn}{dc_{sucrose}}$ is the refractive increment of sucrose at 20.0°C. Stock concentrations of GBA-CRM-026, GBB-CRM-008B, GBD-CRM-005, and GBT-CRM-003 were 2.21 mg.mL^{-1}, 3.18 mg.mL^{-1}, 1.69 mg.mL^{-1}, and 2.46 mg.mL^{-1} respectively.

Samples of CRM_{197} were diluted in PBS to produce concentration series for each of the experiments. A single concentration series was produced for each of the glycoconjugates to perform all the hydrodynamic experiments (Table 3). A full series for the vaccine candidate GBD-CRM-005 could not be produced due to the smaller volume of sample (Table 3).

Table 3. Experimental concentration series for each of hydrodynamic experiments and the samples characterised.

Experimental concentration series (mg.mL^{-1})				
CRM_{197} SV	**CRM_{197} SE**	**CRM_{197} Viscosity**	**GBA-CRM-026, GBB-CRM-008B, GBT-CRM-003**	**GBD-CRM-005**
2.00	1.0	14	2.0	1.5
1.00	0.9	12	1.7	1.3
0.50	0.8	10	1.4	1.1
0.25	0.7	8	1.1	0.9
0.125	0.6	6	0.8	0.7
0.0625	0.5	4	0.5	-
-	0.4	2	0.2	-

2.4. Sedimentation Velocity

Sedimentation velocity analytical ultracentrifugation (SV-AUC) experiments were performed using a Beckman Optima XL-I analytical ultracentrifuge (Beckman Coulter, Brea, CA) equipped with Rayleigh interference and UV optics (280nm). 395 μL of sample and 405 μL of reference buffer were injected into two sector epoxy cells with a 12 mm centre piece and either sapphire or quartz windows. For CRM_{197} alone, cells were loaded into an An-60 Ti Analytical 4-Place Titanium Rotor (Beckman Coulter, Brea, CA). Conjugate vaccines were loaded into an 8 hole An-50 Ti Analytical 8-Place Titanium Rotor (Beckman Coulter, Brea, CA). Experiments were performed at 20.0℃ with an equilibration time of one hour, and a single experiment for CRM_{197} was performed at 30.0℃ with an equilibration/ incubation time of four hours. A temperature correction was applied to the partial specific volume for analysis of data from this experiment (Hinz, 1986):

$$\bar{v}_r = \bar{v}_{25} + 4.25 \times 10^{-4}(T - 25) \qquad (4)$$

Where $\bar{v}_r$ is the partial specific volume at the given temperature, $\bar{v}_{25}$ is the partial specific volume of the sample at 25.0℃, and T is the given temperature of the experiment in ℃. Rotor speed for all CRM_{197} experiments was set at 48000 rpm while the rotor speed for each of the conjugates was set at 40000 rpm.

Data were analysed using SEDFIT (Schuck, 2000, Schuck, 1998). Partial specific volumes, buffer density and buffer viscosity used in the analysis have been previously described (Table 2). Continuous c(*s*) distributions were produced using 200 scans in each case. The 'least squares g(*s*)' method was used to generate distributions of g(*s*) vs *s*, where *s* is the sedimentation coefficient (Svedberg units, S = 10^{-13} s). Sedimentation coefficients were normalised to standard conditions (viscosity and density of water at 20.0℃) to give $s_{20,w}$. A range of 0 – 15 Svedbergs was used for CRM_{197}, with a resolution of 200 points per distribution. A range of 0 – 25 Svedbergs was used for each of the conjugates, with a resolution of 200 points per distribution. A confidence level of 0.95 was used for all analyses. Extrapolation of standard 1 / $s_{20,w}$ vs c plots was performed to eliminate the effects of nonideality and to estimate $s_{20,w}$ at zero concentration ($s^0_{20,w}$). Distributions and plots were exported into graphics package Prism 10 before reporting.

2.5. Sedimentation Equilibrium

Sedimentation equilibrium analytical ultracentrifugation (SE-AUC) experiments were performed using a Beckman Optima XL-I analytical ultracentrifuge (Beckman Coulter, Brea, CA) equipped with Rayleigh interference and UV optics (280nm). 150 μL of sample and reference buffer were injected into two sector titanium cells with a 20 mm centre piece and either sapphire or quartz windows. All samples were loaded into an 8 hole An-50 Ti Analytical 8-Place Titanium Rotor (Beckman Coulter, Brea, CA) and experiments were performed at

20.0℃ after ≥ 1 hour of equilibration time. Samples of CRM_{197} were centrifuged at 13000 rpm until equilibrium was reached. Conjugates were centrifuged at 10000 rpm until equilibrium was reached.

Sedimentation distribution data were analysed using SEDFIT (Schuck, 2000, Schuck, 1998) and weight average molecular weight ($M_{w,app}$) using the SEDFITMSTAR algorithm (Harding et al., 1992, Schuck et al., 2014). The alternative calculation for weight average molecular weight, hinge point (incorporated into SEDFITMSTAR), was also performed and calculates $M_{w,app}$ relative to the initial loading concentration. The z-average molecular weight M_z was also determined using the MFIT algorithm (Ang and Rowe, 2010). A confidence level of 0.95 was used for all analyses. Extrapolation of standard 1 / M_w vs c plots to zero concentration were performed to eliminate the effects of nonideality. Distributions and plots were exported into the graphics package Prism 10 before reporting.

2.6. Rolling ball viscometry

Intrinsic viscosity of each sample was determined using an Anton paar AMVn rolling ball viscometer (Anton Paar GmbH, Graz, Austria). All experiments were performed at 20.0℃. An inversion angle 70.0° for 10 repeats was used to find the average runtime of sample and PBS from which, relative viscosity could be determined using the following equation:

$$\eta_{rel} = \frac{t}{t_0} \qquad (5)$$

where t is the time of flow of the solution, t_0 is the time of flow of the solvent.

The concentration series previously described in Table 2 enabled plots to be produced from the Huggins and Kraemer equations (Huggins, 1942, Kraemer, 1938):

$$\frac{\eta_{sp}}{c} = [\eta'](1 + K_H[\eta']c) \qquad (6)$$

$$\frac{\ln(\eta_{rel})}{c} = [\eta'](1 - K_k[\eta']c) \qquad (7)$$

where η_{sp} is specific viscosity defined as $\eta_{rel} - 1$, η_{rel} is the relative viscosity of the solution, c is concentration, $[\eta']$ is kinematic intrinsic viscosity, and K_H and $K_k$are the Huggins and Kraemer coefficients, respectively. Plots of $\frac{\eta_{sp}}{c}$ vs *c* and $\frac{\ln(\eta_{rel})}{c}$ vs *c* are extrapolated to give $[\eta']$ free from the effects on nonideality. The density correction described by Tanford (1955) was applied to obtain dynamic intrinsic viscosity $[\eta]$ from the previously determined kinematic intrinsic viscosity $[\eta]$ for CRM_{197}. Density correction for the conjugates was negligible, owing to their high viscosity, meaning $[\eta'] \sim [\eta]$:

$$[\eta] = \frac{1 - \bar{\nu}\rho_0}{\rho_0} + [\eta'] \qquad (8)$$

where ρ_0 is the density of the solvent.

2.7. U-tube capillary viscometry

A single measurement for the intrinsic viscosity of CRM_{197} was also taken using a Xylem Analytics U-tube Ostwald capillary viscometer (Xylem Analytics GmbH, Germany) with flow times recorded with a semi-automated Schott Gerate AVS 400 viscometer in a temperature-

controlled water bath at 20.0℃. 2.0 mL of a 10.0 mg.ml^{-1} sample of CRM_{197} was injected into the viscometer and degassed before being submerged in the water bath. The sample was allowed to equilibrate overnight. The flow time of the sample through the viscometer was recorded automatically with 10 repeat runs. Intrinsic viscosity was then determined using the Solomon-Ciuta equation, which allows $[\eta]$ to be estimated from a single concentration point (Solomon and Ciută, 1962):

$$[\eta] \sim \frac{\sqrt{2[(\eta_r - 1) - \ln(\eta_r)]}}{c} \qquad (9)$$

2.8. SEC-MALS

Size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) analysis, as described by Wyatt (1993), was conducted for each of the glycoconjugates. The SEC setup included a Postnova Analysis PN7505 degassing unit (Postnova Analytics GmbH, Landsberg am Lech, Germany), a Shimadzu LC-10 AD HPLC Pump (Shimadzu UK, Milton Keynes, UK), and a Spark-Holland Marathon Basic autosampler (Spark Holland, Emmen, The Netherlands). Samples were separated through a TSK Gel guard column (7.5 × 75 mm) and TSK Gel G5000 and G6000 columns (7.5 × 300 mm) connected in series (Tosoh Biosciences, Tokyo, Japan).

Light scattering intensities were concurrently recorded at 18 angles during elution using a DAWN® HELEOSTM II light scattering photometer, connected sequentially to an Optilab® rEX refractive index detector (Wyatt Technology Corporation, California, USA). Prior to injection into the column, 1.0 mg.mL^{-1} samples of each conjugate were filter sterilized through a 0.45 µm syringe filter. 100 µl aliquots were analysed using the PBS dialysate as eluent, flowing at a rate of 0.8 mL.min^{-1}.

For the estimation of weight average molecular weight, M_w (g.mol^{-1}), and z-average molecular weight, M_z (g.mol^{-1}), we employed ASTRATM (Version 6) software (Wyatt, USA). A 4 mW He-Ne laser with a wavelength of 632.8 nm was utilized, and the refractive increment for each conjugate was referenced from previously reported values (Table 2). It is worth noting that due to the low solute concentrations post-dilution within the columns, nonideality effects were considered negligible.

3. Results and Discussion

3.1. Characterisation of CRM_{197}

3.1.1. Sedimentation Velocity

Three different conditions were employed for the sedimentation velocity characterisation. These were freshly thawed protein that had previously been stored at -80.0℃, protein that had been diluted and stored at 4.0℃ for 3 weeks, and protein that had been incubated at 30.0℃ prior to centrifugation. Each case indicated that the protein has good stability in its natural state.

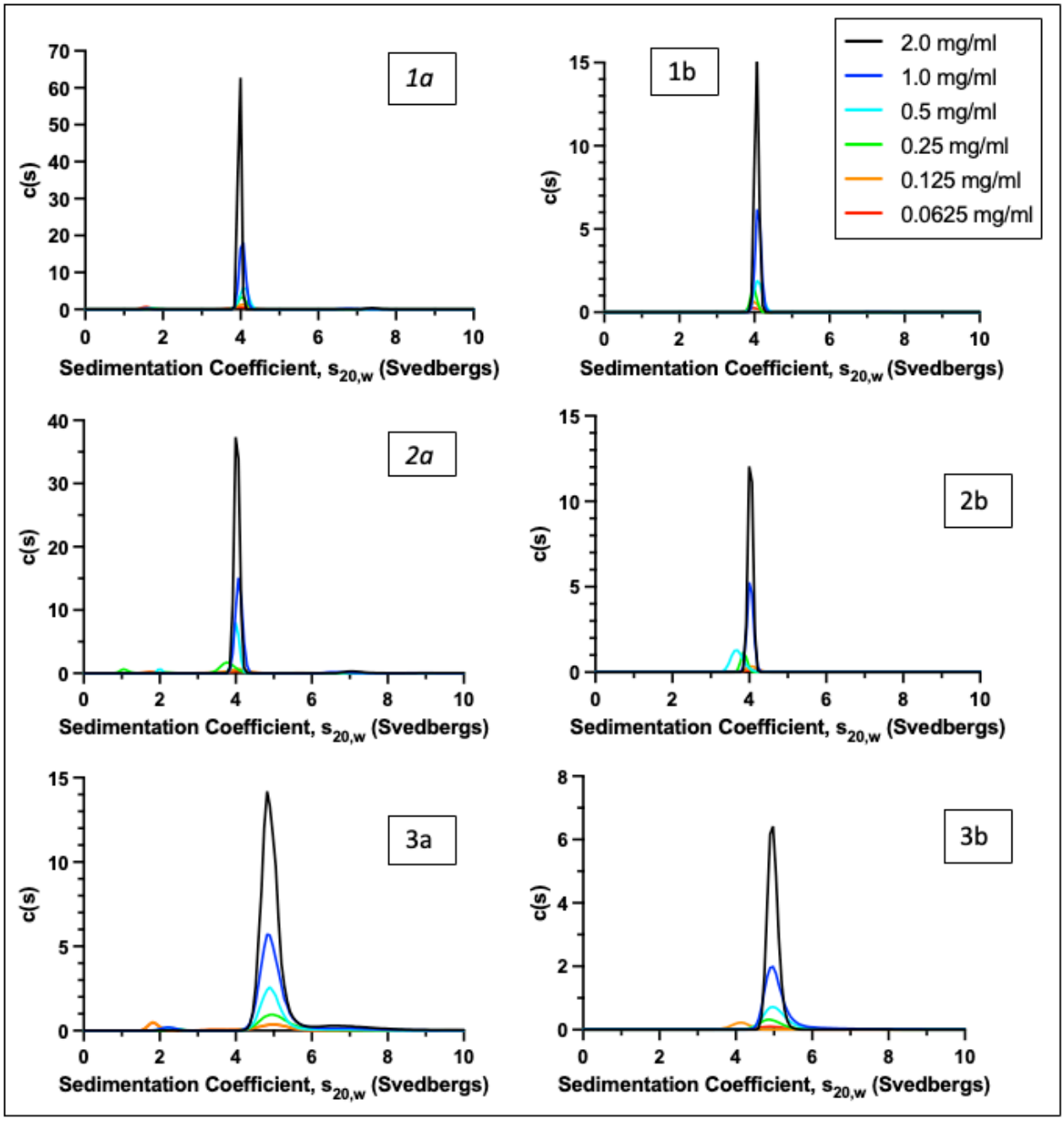

Figure 1. Sedimentation coefficient distributions of CRM_{197} at 6 loading concentrations under different experimental conditions. *1a*, Interference data for freshly thawed protein; *1b* absorbance data for freshly thawed protein; *2a*, Interference data for protein stored at 4.0℃ for three weeks; *2b*, absorbance data for protein stored at 4.0℃ for three weeks; *3a* interference data for protein incubated for 4 hours at 30.0℃; *3b* absorbance data for protein incubated at 30.0℃ for 4 hours.

Each of the distributions reveals that the majority of CRM_{197} remains in its monomeric conformation in solution. For each of the three conditions, the absorbance data show the main peak to be ≥ (95.5 ± 3.3) % of the whole distribution (Figure 1). There is a slight increase in the proportion of the larger dimeric species in both the 3 week and 30.0℃ samples compared to the freshly thawed (Table 4b). The interference distributions show a smaller species as well as an aggregating one. A small peak can be observed around 1 – 2 S on each of the distributions and occupies a larger proportion of the distribution for lower concentrations. For the freshly thawed protein at concentrations of 1.0 and 0.5 $mg.mL^{-1}$, the smaller species makes up 1.1 and 3.6 % of the distribution respectively, while the main peak accounts for 92.1 and 92.7 %. At concentration 0.25 and 0.125 $mg.mL^{-1}$, the smaller species makes up 5.8 and 12.7 % of the distribution respectively, while the main peak accounts for 88.4 and 75.9 % (Figure 1). This trend is observed for all experimental conditions and is responsible for the monomeric peak taking up a lower proportion of the distribution and displaying a larger standard deviation for the interference data than the absorbance (Table 4a). The interference data also shows a larger proportion of the system is aggregating with even the fresh sample displaying (4.4 ± 3.2) % of the distribution as dimer. For both optical systems, incubating the protein and centrifuging at 30.0℃ produced a wider peak for the monomeric species on the sedimentation distribution (Figure 1).

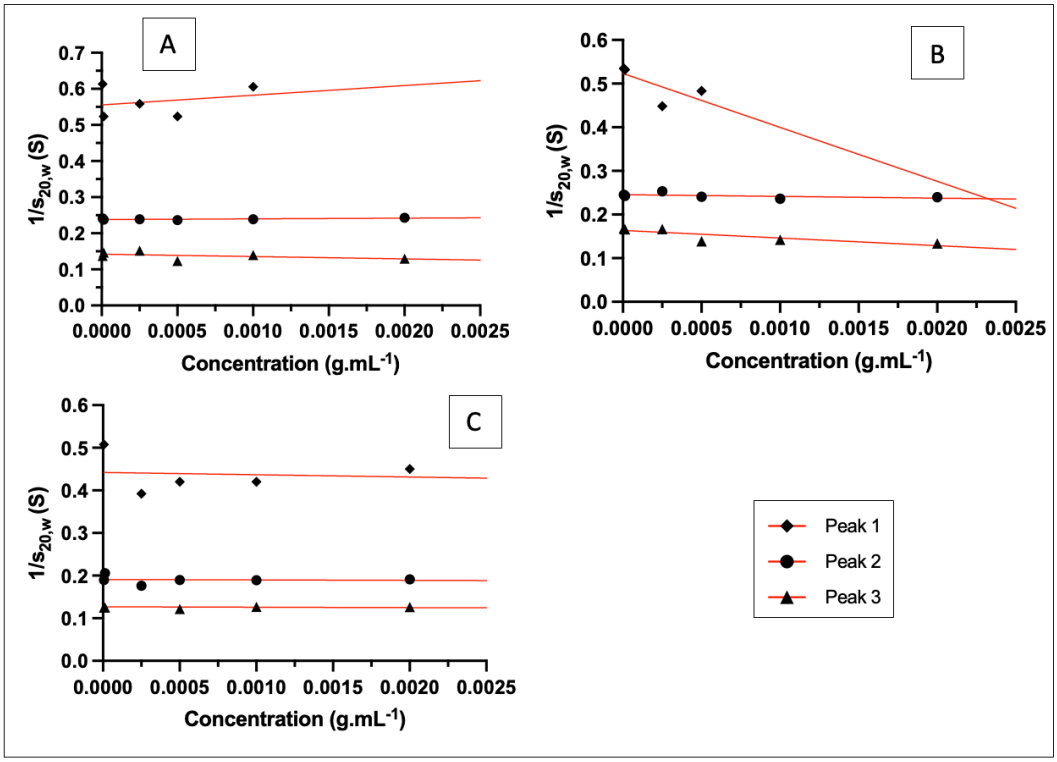

Figure 2. Concentration dependence of the reciprocal of sedimentation coefficient normalised for standard conditions (density and viscosity of water at 20.0°C) for each species observed by interference optics under sedimentation velocity of CRM_{197}. A, directly after thawing; B, after storage at 4.0°C for 3 weeks; C, after incubation at 30.0°C for 4 hours.

Table 4a. Analysis of interference data from sedimentation coefficient distributions and reciprocal sedimentation coefficient plots for each of the experimental conditions of CRM_{197}.

Sample conditions	$s^0_{20,w}$ Peak 1 (S)	Proportion Peak 1(%)	$s^0_{20,w}$ Peak 2 (S)	Proportion Peak 2 (%)	$s^0_{20,w}$ Peak 3 (S)	Proportion (%)
Freshly Thawed	1.80	10.4 ± 11.2	4.20	84.1 ± 14.3	7.04	4.4 ± 3.2
3 Weeks at 4°C	1.91	14.8 ± 6.4	4.07	81.3 ± 13.7	6.11	5.2 ± 2.6
Incubated for 4 hours at 30°C	2.26	8.2 ± 11.8	5.24	84.4 ± 13.9	7.88	7.2 ± 1.2

N.B. Three main peaks of each distribution were analysed, and $s^0_{20,w}$ was determined from the intercept of linear regression for each plot (Figure 2).

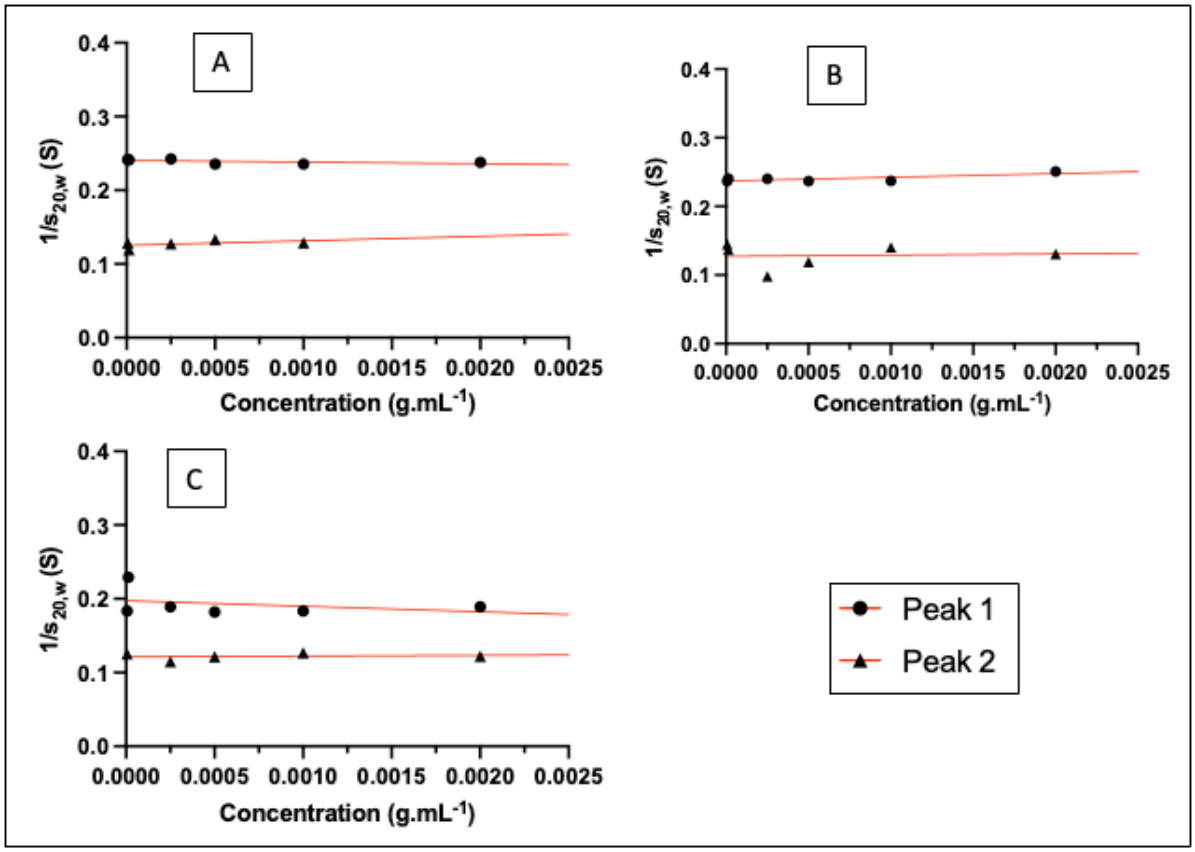

Figure 3. Concentration dependence of the reciprocal of sedimentation coefficient normalised for standard conditions (density and viscosity of water at 20.0°C) for each species observed by absorbance optics (280nm) under sedimentation velocity of CRM_{197}: A, directly after thawing; B, after storage at 4.0°C for 3 weeks; C, after incubation at 30.0°C for 4 hours.

Table 4b. Analysis of absorbance data from sedimentation coefficient distributions and reciprocal sedimentation coefficient plots for each of the experimental conditions of CRM_{197}.

Sample conditions	$s^0_{20,w}$ Peak 1 (S)	Proportion Peak 1(%)	$s^0_{20,w}$ Peak 2 (S)	Proportion Peak 2 (%)
Freshly Thawed	4.16	97.1 ± 1.5	7.97	2.7 ± 1.3
3 Weeks at 4°C	4.22	95.5 ± 3.3	7.85	3.5 ± 3.2
Incubated for 4 hours at 30°C	5.06	97.4 ± 2.5	8.24	3.1 ± 2.5

N.B. Two main peaks of each distribution were analysed, and $s^0_{20,w}$ was determined from the intercept of linear regression for each plot (Figure 3).

For the experiments performed at 20.0℃ both optical systems are in good agreement as to the size of the monomeric peak between ~4.1 S – 4.2 S. There is greater variation between the two optics for the aggregating species. The interference indicates it to be between ~6.1 S – 7.0 S, while absorbance shows the larger species to be between ~7.9 S – 8.0 S: this difference could be a result of the low concentration of the aggregating species. The larger sedimentation coefficient for the samples incubated at 30.0℃ is likely a result of the higher partial specific volume applied during the analysis ($\bar{v} = 0.737$ mL.g^{-1}).

3.1.2. Sedimentation Equilibrium

Like the velocity data, the SE-AUC experiments for CRM_{197} reveal that it is largely stable. Only two of the experimental conditions were analysed using SE-AUC due to time constraints, so only data for the freshly thawed protein and the protein stored at 4.0℃ for 3 weeks are reported here.

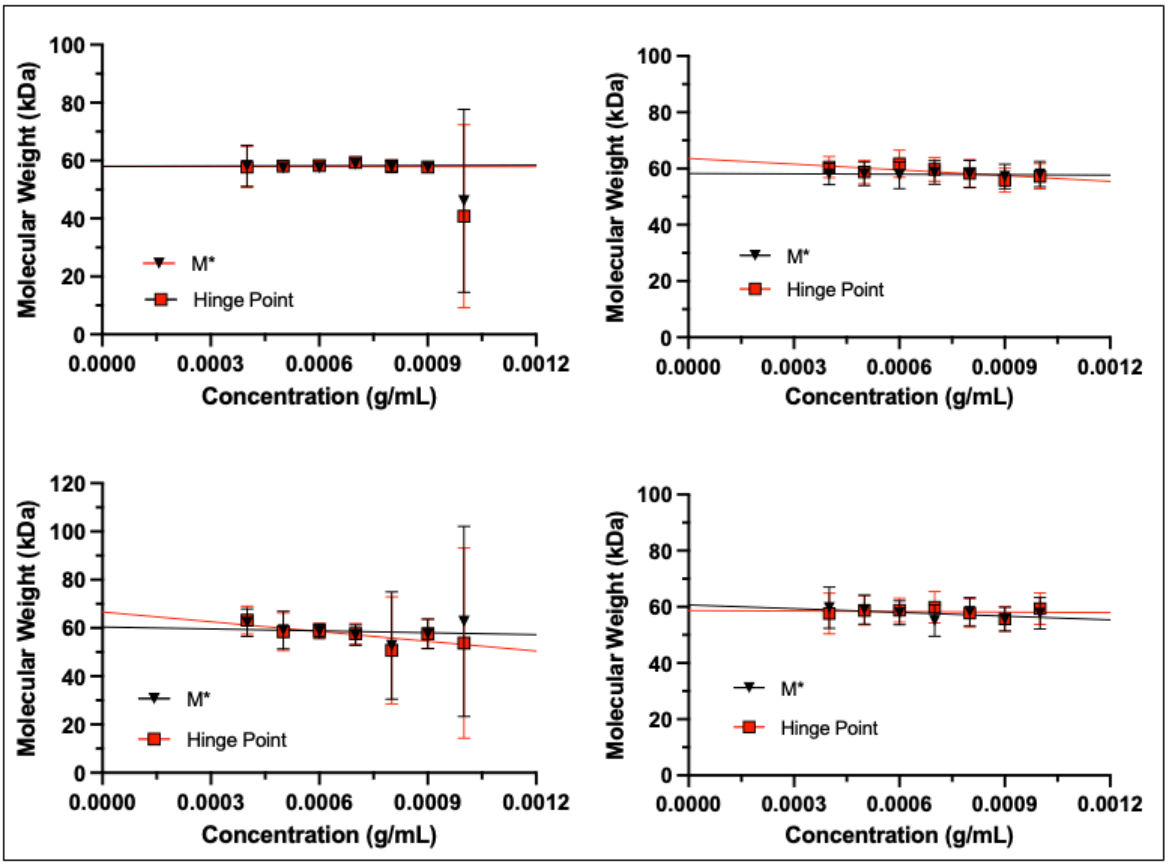

Figure 4. Plots of apparent molecular weight vs concentration for: 1a, interference data for freshly thawed protein; 1b, absorbance data for freshly thawed protein; 2a, interference data for protein sored at 4.0°C for 3 weeks; 2b, absorbance data for protein sored at 4.0°C for 3 weeks. Two different methods for data manipulation are plotted in the form of M* and Hinge point and extrapolated to zero concentration to give an estimate for $M_{w,app}$ free from the effects of nonideality.

Table 5. Sedimentation equilibrium data from both interference and absorbance optics for CRM_{197} under two distinct experimental conditions

Sample Conditions	Interference			Absorbance		
	M* (kDa)	Hinge (kDa)	M_z (kDa)	***M**** (kDa)	Hinge (kDa)	M_z (kDa)
Freshly Thawed	58.04	58.01	58.12	58.28	63.65	58.27
Stored at 4°C for 3 Weeks	60.46	66.66	59.88	60.71	58.67	58.11

N.B. Two different calculations for weight average molecular weight (M_w), *M** and Hinge point (Schuck et al., 2014), are performed. z-average molecular weight (M_z) is also reported. Estimates of molecular weight are made from extrapolations of *M** and Hinge point vs c to eliminate the effects on nonideality (Figure 4), generally with good agreement. Assumed error for monodisperse systems is ± 3%, so estimates to within ~ (± 1.7 kDa).

The *M** data for both the freshly thawed and 3-week-old protein have good agreement between the interference and absorbance data. The *M** estimate for interference is 2.42 kDa larger for the 3-week-old protein than the freshly thawed sample. Similarly, the *M** estimate from absorbance is 2.43 kDa larger for the 3-week-old protein compared to the freshly thawed sample (Table 5). Both estimates for the fresh sample are very similar to the theoretical molecular weight of CRM_{197} taken from the protein sequence as 58.48 kDa (Malito et al., 2012). *M** and Hinge estimations from the interference data of the freshly thawed protein in good agreement, with only a 0.03 kDa discrepancy. The same is not true for any of the absorbance data, or the interference data of the 3-week-old protein (Table 5). This could be due to the greater standard deviation observed for each of the point on these respective plots (Figure 4). The M_z value is 1.76 kDa larger for the stored protein compared to the freshly thawed protein when considering the interference data. This is what would be expected given that a greater proportion of the system is in dimeric form or larger, and that M_z gives more weight to larger molecular species.

3.1.3. Viscosity

Two different methods were employed to determine the intrinsic viscosity of CRM_{197}. The first method used a U-tube capillary viscometer submerged in a temperature-controlled water to measure the flow times of different concentrations of CRM_{197} relative to the PBS buffer in which they were dissolved. It proved difficult to record repeatable flow times for the protein using this method. After two hours of equilibration time, the flow rates between the first repeat and tenth repeat varied by 1.23 seconds. The same was the case for the following 20 repeats, with each subsequent repeat leading to a longer flow time. The standard deviation for all 30 repeats was 0.67 seconds (Table 6).

Table 6. Flow times of a 10.0 mg.mL^{-1} solution of CRM_{197} measured using a U-tube viscometer submerged in a water bath at 20℃.

Repeat	Flow Time (s)		
	Run 1	**Run 2**	**Run 3**
1	98.62	99.94	100.09
2	98.88	99.51	100.28
3	98.97	99.53	101.05
4	99.09	99.77	100.64
5	99.21	99.82	100.55
6	99.23	99.72	100.40
7	99.49	99.98	100.87
8	99.37	99.81	100.74
9	99.40	100.05	101.13
10	99.85	100.27	100.75

N.B. Three independent runs of 10 repeats were performed sequentially, with a single run taking around 40 minutes.

This is despite recording consistent repeatable flowtimes for PBS prior to measuring the flow times of the protein using the same equilibration time. There is the possibility of interaction between the glass and the protein which could be causing localised increases in concentration within the capillary, leading to the increase in flow time. It also seemed to be the case that the protein was aggregating readily under the forces of the flow of solution. There were visible aggregates resembling white strings which formed after multiple runs, particularly in the case of the higher concentrations. The partial blockage of the capillary by these aggregates could have also resulted in the longer flow time and any highly variable flowtimes also. It was possible to take repeatable measurements after 2.0 ml of 10.0 mg.mL^{-1} CRM_{197} was left to equilibrate in the capillary submerged in a temperature-controlled water bath overnight at 20.0℃. The resultant flow time times of (110.09 ± 0.15) s for CRM_{197} and (100.46 ± 0.10) s for PBS gave an unexpectedly high estimate for intrinsic viscosity of 9.30 mL.g^{-1} using the Solomon-Ciuta equation (Equation 9)(Table 7).

Table 7. Flow times for PBS and a 10.0 mg.mL^{-1} solution of CRM_{197} measured using a U-tube viscometer equilibrated overnight in a water bath at 20.0℃.

Repeat	Flow Time PBS (s)	Flow Time CRM_{197} (s)
1	100.25	109.95
2	100.41	110.15
3	100.48	109.95
4	100.43	110.09
5	100.47	109.95
6	100.62	110.41
7	100.52	110.28
8	100.44	110.00
9	100.54	110.04
10	100.44	110.08

After much effort and repeated failures, attempts to measure intrinsic viscosity via this method were abandoned and rolling ball viscosity was instead employed. An explanation for these problems is outlined in the discussion.

By contrast, consistent, repeatable flow times were achieved using rolling ball viscosity which allowed the relative viscosity, η_{rel}, of each concentration of CRM_{197} to be estimated. η_{rel} was then applied to the Kraemer and Huggins equations (Equation 7, 8), to produce a plot of relative viscosity. Extrapolation of these plots gives an estimate for intrinsic viscosity, $[\eta]$, of (4.2 ± 0.1) mL.g^{-1} after correction for density (Equation 8)(Figure 5).

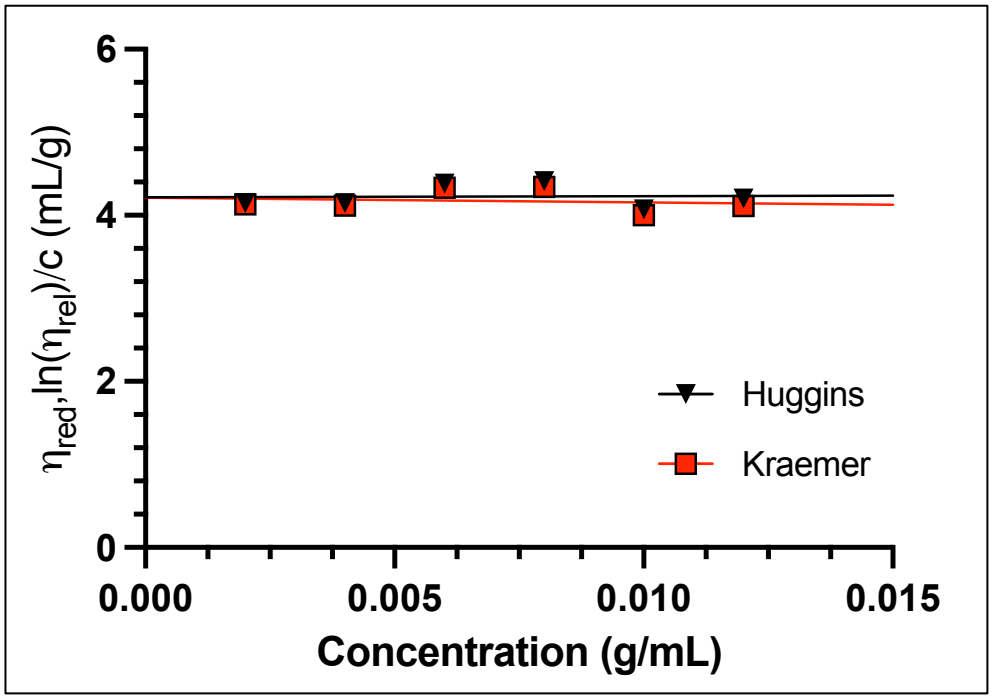

Figure 5. Plots of reduced viscosity (Huggins) and inherent viscosity (Kraemer) from rolling ball measurements of CRM_{197} in PBS buffer at 20.0°C with extrapolation to zero concentration to give an estimate for intrinsic viscosity free from the effects of nonideality.

3.1.4. Conformation

The conformation of monomeric CRM_{197} can be estimated from sedimentation and viscosity data using the ELLIPS suite of macromolecular conformation algorithm (Harding et al., 1997). For an $s_{20,w}$ ~4.2 S, Mw = 58,000 kDa, $\bar{v}$ = 0.733 ml/g: a frictional ratio (f/f_0) = 1.27 is calculated using the using the routine UNIVERSAL PARAM and ELLIPS1. For an (f/f_0) of 1.27, the Perrin (P) universal shape function = 1.13 and 1.10 for a hydration, δ, of 0.30 g/g and 0.40 g/g respectively. A hydration value of 0.30 g/g – 0.40 g/g is inferred from nuclear magnetic resonance of globular proteins, hence the chosen hydration values in place of missing experimental data (Kuntz, 1971). These data correspond to an aspect ratio (a/b) of ~(3.0 ± 0.2) (for a prolate ellipsoid) in ELLIPS1 (Harding et al., 1997).

From $[\eta]$ = 4.20 mL.g^{-1}, $\bar{v}$ = 0.733 mL.g^{-1}, a Simha–Saito (Harding, 1997, Simha, 1940, Saitô, 1951) universal shape function, ν, of 4.06 and 3.70 is calculated for a hydration, δ, of 0.30 g/g and 0.40 g/g respectively. These data correspond to an aspect ratio of 3.0 – 3.4 in ELLIPS1. Estimations for both sedimentation and viscosity data give an aspect ratio (a/b) for CRM_{197} of (3.10 ± 0.30), and both methods have good agreement as to the conformation of the protein as that of a prolate ellipsoid (Figure 6).

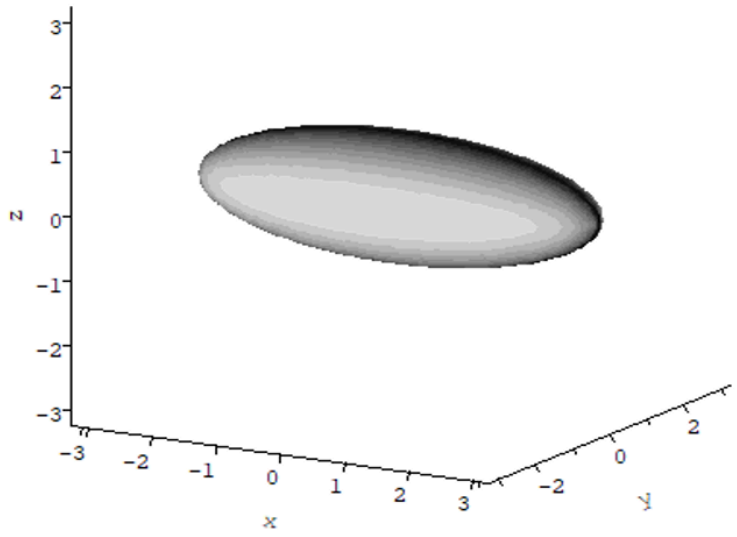

Figure 6. Prolate ellipsoid representation of a monomer of CRM_{197} visualised using the ELLIPS1 programme and displaying an asymmetrical structure of axial ratio 3.1 : 1.

For this protein, the rolling-ball viscometer proved the most appropriate. Attempts to use a standard Ostwald U-tube viscometer proved problematic as indicated above, possibly related to the automatic pumping of the CRM_{197} solutions from the reservoir prior to the measurement of flow times under creeping flow. This is considered further in the discussion.

3.2. Characterisation of glycoconjugate vaccine candidates

3.2.1. Sedimentation Velocity

Sedimentation velocity was performed for each of the glycoconjugate vaccine candidates using Rayleigh interference optics to account for the lack of absorbance of the polysaccharide components of the samples. Analysis of the sedimentation data for each of the prospective conjugate vaccines hints towards a polydisperse system comprising at least two major species (Figure 4). Interestingly, despite the serotype of each conjugate varying, the distribution pattern for each of the four vaccine candidates is very similar. Each sample has a major peak in the region of ~5.0 – 6.0 S, a secondary peak around ~7.5 – 9.5 S, and a tertiary peak around ~11 – 13 S, apart from GBD-CRM-005, for which there are only two distinct species on the distribution (Table 8).

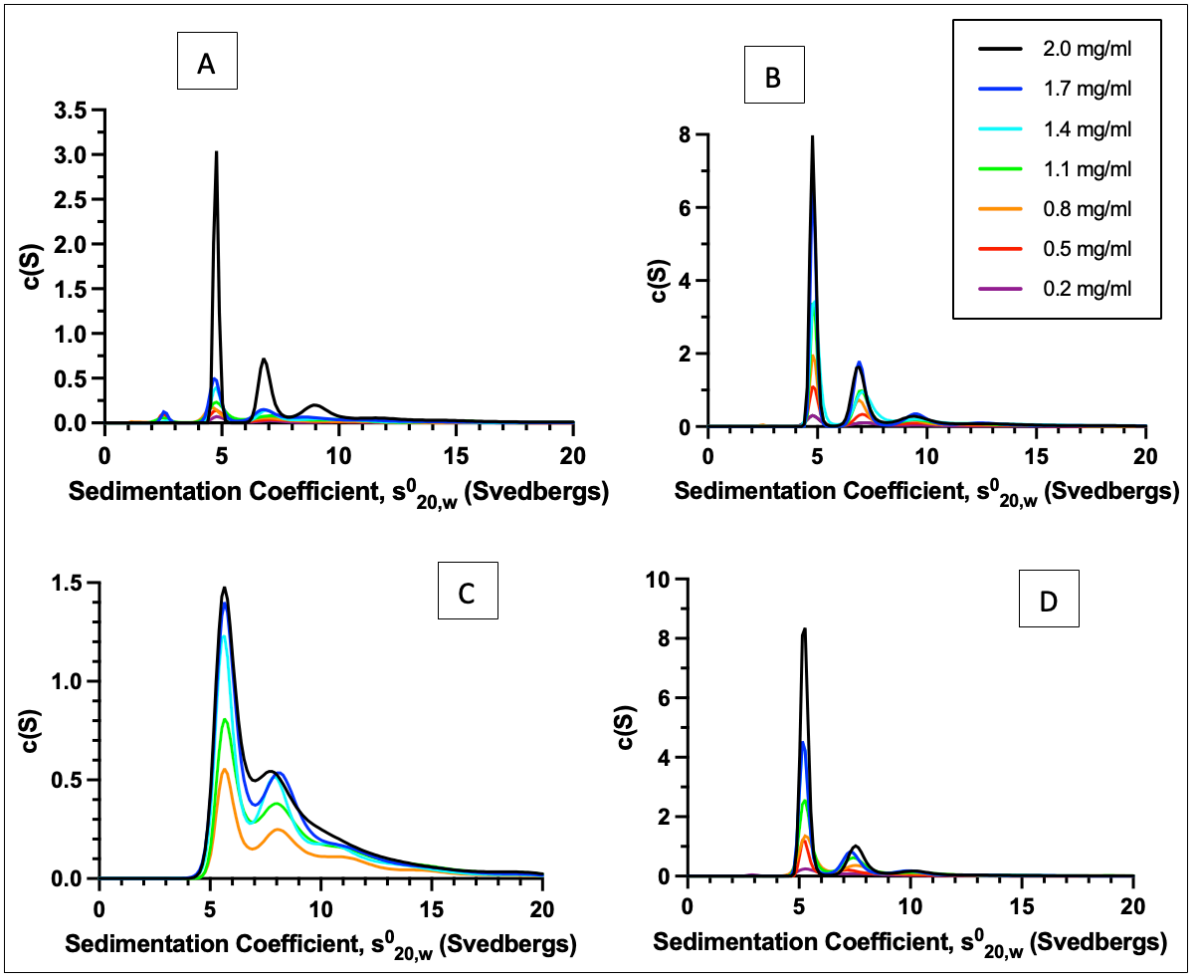

Figure 8. Sedimentation coefficient distributions produced from Rayleigh interference optical data of: A, 7 loading concentrations of GBA-CRM-026; B, 7 loading concentrations of GBB-CRM-008B; C, 5 loading concentrations of GBD-CRM-005; D, 7 loading concentrations of GBT-CRM-003.

Table 8. Analysis of interference data from sedimentation coefficient distributions and reciprocal sedimentation coefficient plots for each of the conjugate vaccine candidates.

Conjugate	Peak 1		Peak 2		Peak 3	
	$s^0_{20,w}$ (S)	Proportion (%)	$s^0_{20,w}$ (S)	Proportion (%)	$s^0_{20,w}$ (S)	Proportion (%)
GBA-CRM-026	5.06	37.2 ± 5.2	8.12	30.2 ± 11.9	10.83	24.8 ± 14.7
GBB-CRM-008B	5.02	46.3 ± 4.4	7.51	28.4 ± 0.9	11.88	18.6 ± 7.6
GBD-CRM-005	6.01	44.0 ± 4.6	8.56	46.8 ± 13.5	No Peak	No Peak
GBT-CRM-003	5.69	60.1 ± 5.3	9.41	30.3 ± 10.6	12.85	11.6 ± 2.6

N.B. Three main peaks of each distribution were analysed for all samples except GBD-CRM-005. $s^0_{20,w}$ was determined from the intercept of linear regression for each plot (Figure 9).

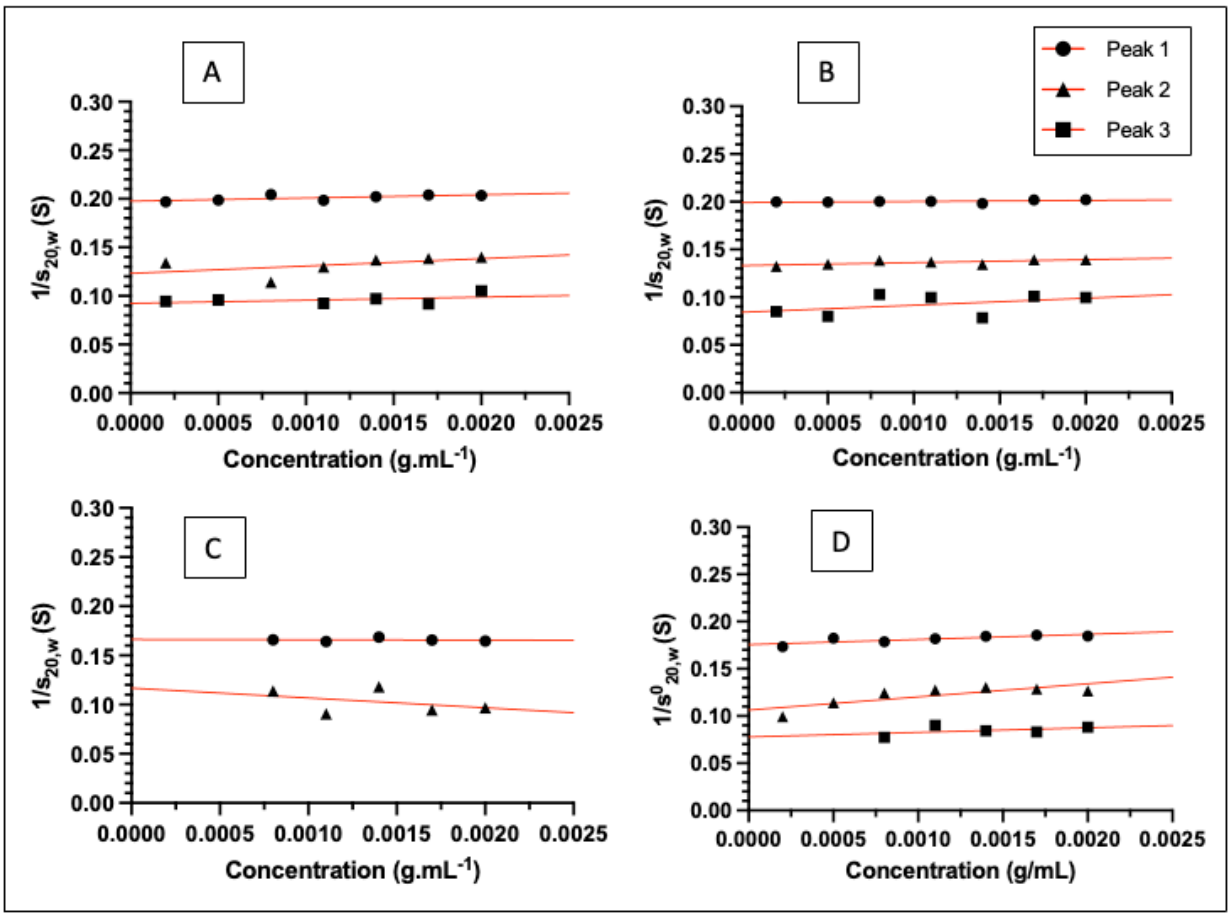

Figure 9. Concentration dependence of the reciprocal of sedimentation coefficient normalised for standard conditions (density and viscosity of water at 20.0°C) for each conjugate observed by interference optics under. A, GBA-CRM-026; B, GBB-CRM-008B; C, GBD-CRM-005; D, GBT-CRM-003.

While the shape of each distribution shows similarity, the relative proportion of each species varies between each of the samples. The primary peak of GBA-CRM-026 occupies the smallest proportion of the distribution relative to the other conjugates at 37.2%, while the primary peak of GBT-CRM-003 makes up almost double that at 60.1% of the distribution (Table 8). It is also generally the case that proportion peak 1 (%) > proportion peak 2 (%) > proportion peak 3 (%), except for GBD-CRM-005 where the second peak occupies a larger percentage of the distribution, although the primary peak has a greater amplitude (Figure 8).

Absorbance data were also collected for a number of cells for three of the conjugates to test which species were proteinogenic.

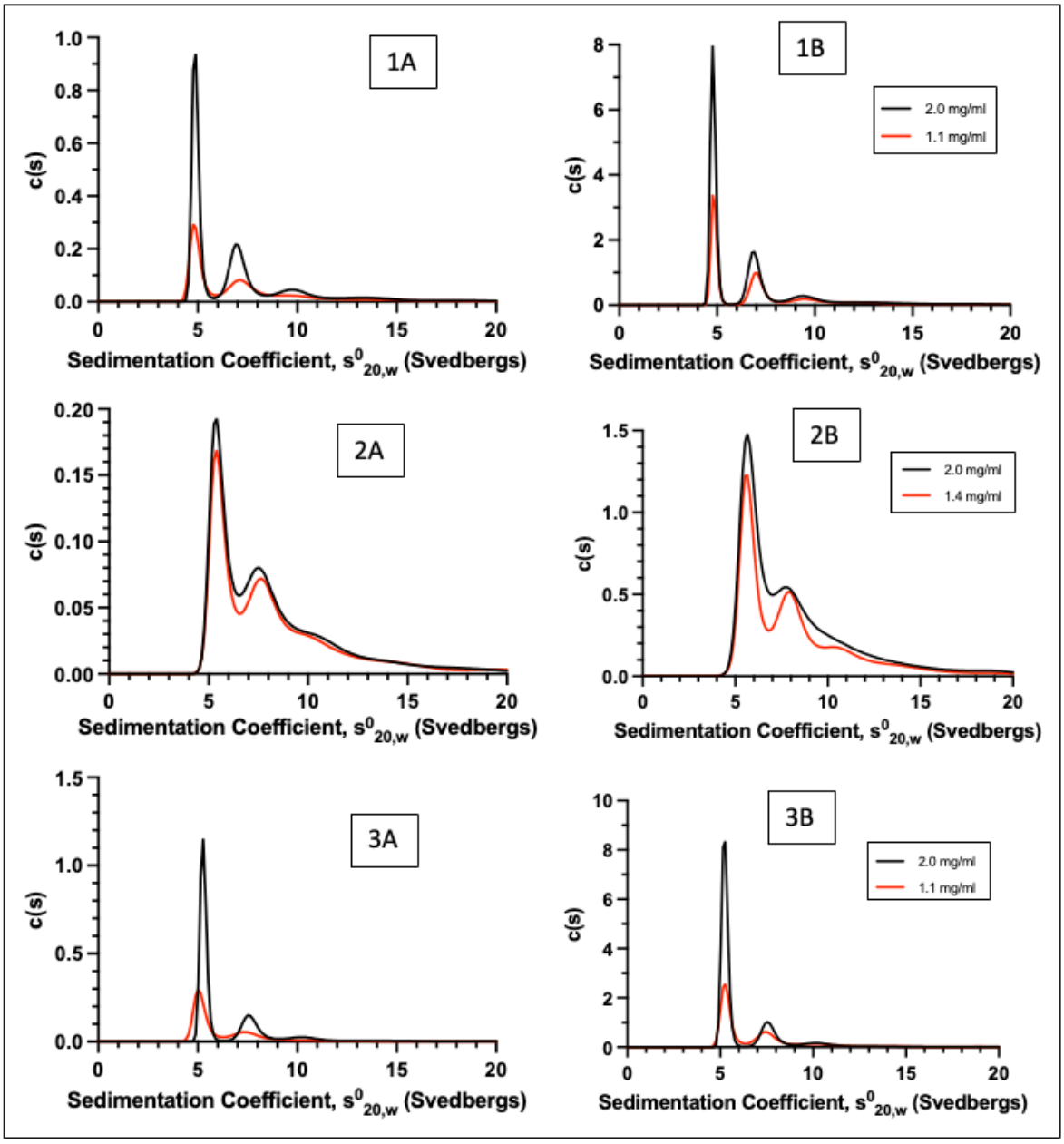

Figure 8. Sedimentation coefficient distributions of two loading concentrations of: 1, GBB-CRM-008B; 2, GBD-CRM-005; 3, GBT-CRM-003 produced from sedimentation velocity at rotor speed 40,000 rpm in PBS buffer. A, corresponds to absorbance data; B, corresponds to interference data

Looking at the distributions produced from both sets of data side by side, it is clear to see that each of the peaks corresponds to a proteinogenic species. That means that the peak could be

either an unbound protein, conjugated glycoconjugate, aggregating protein, or aggregating glycoconjugate. These data also imply that none of the species in the sample are unconjugated polysaccharide.

Sedimentation coefficient is a manifestation of both molecular shape and molecular weight; thus it is possible to estimate the molecular weight by further estimation of diffusion or frictional coefficient as well as sedimentation coefficient (Schuck, 1998). SEDFIT therefore gives an estimate for the molecular weight of each peak. Each of the major peaks for all of the conjugates has a molecular weight of 50,000 – 60,000 Da. It is reasonable to assume that this species is unconjugated protein where the change in molecular weight could be the result of altered structure during chemical conjugation.

3.2.2. Sedimentation Equilibrium

Absolute molecular weight determination proved difficult due to the polydisperse nature of each of the samples. Sedimentation equilibrium gives the most accurate estimation for molecular weight when a highly pure monodisperse system is analysed. Due to the presence of multiple species of highly variable molecular weight, it is near impossible to achieve a state where all of the species in such a sample are in equilibrium. This is in stark contrast for the charachterisation of a globular protein like CRM_{197} which we considered above. We can however give an estimate as to the average molecular weight for all the species in that sample. Extrapolations of concentration dependence plots give estimates for the $M_{w,app}$ and M_z for each of the glycoconjugates (Figure 9, 10). The range of estimates for M^* is 343.3 – 424.3 kDa, while the range of hinge point estimates is slightly smaller at 262.3 – 302.9 kDa (Table 9). The differences in $M_{w,app}$ do not seem to relate to the difference in molecular weight of the polysaccharides each of the conjugates contains, as the molecular weight of the sized serotype

Ia, Ib, and II polysaccharides are reported as 189, 156 and 161 kDa (accurate to ± 5 – 10 %) respectively, while the molecular weight of the native serotype III polysaccharide is reported as 102 kDa (MacCalman, 2023). These molecular weights lead one to believe that GBA-CRM-026 should have the largest molecular weight, assuming the conjugation ratios and methods to be the same, however this is not the case (Table 9). Given the disagreement between the two estimations for M_w, it is hard to give any weight to interpretations from the sedimentation equilibrium data at this stage (Figure 9).

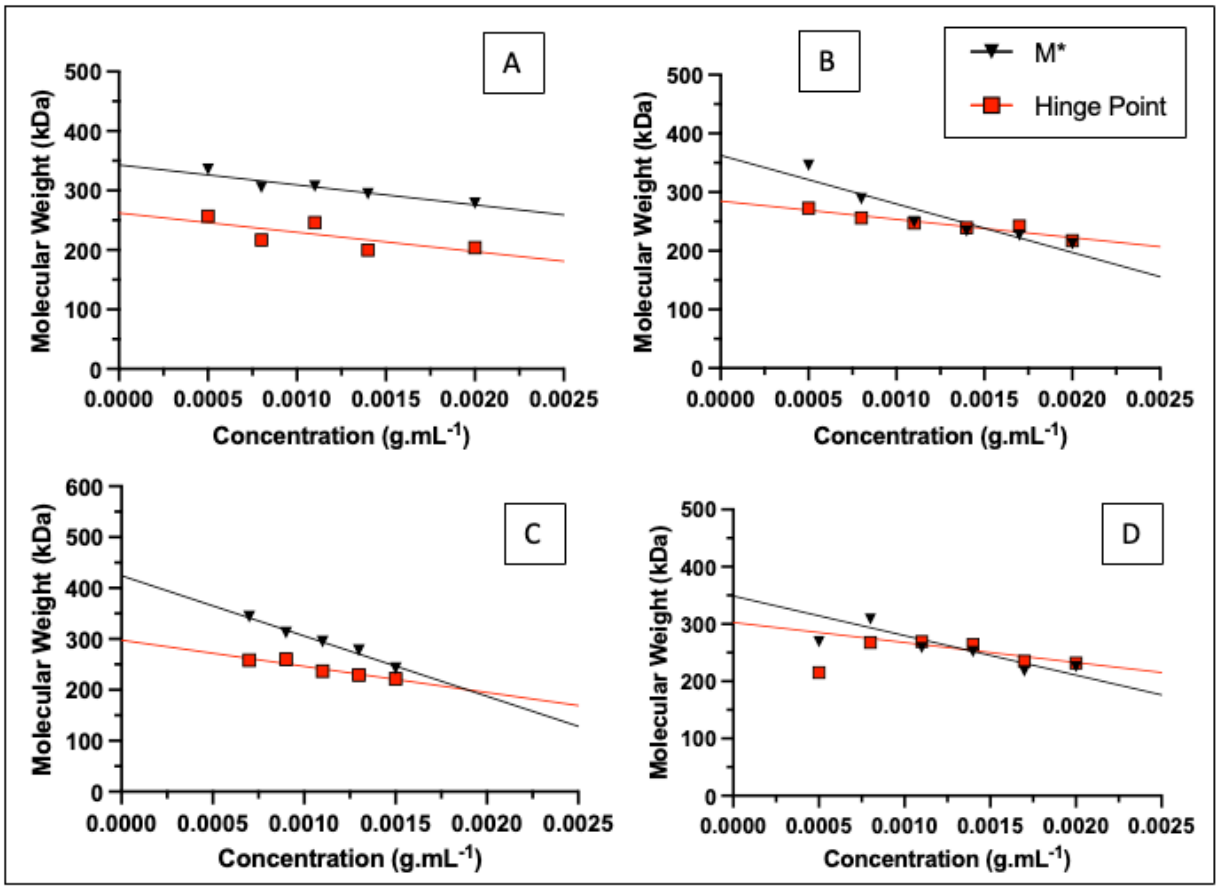

Figure 9. Plots of apparent molecular weight, $M_{w,app}$, vs concentration for: A, GBA-CRM-026; B, GBB-CRM-008B; C, GBD-CRM-005; D, GBT-CRM-003 in PBS buffer, pH = 7.0, I = 0.1 mol.L^{-1} and a rotor speed of 10,000 rpm.

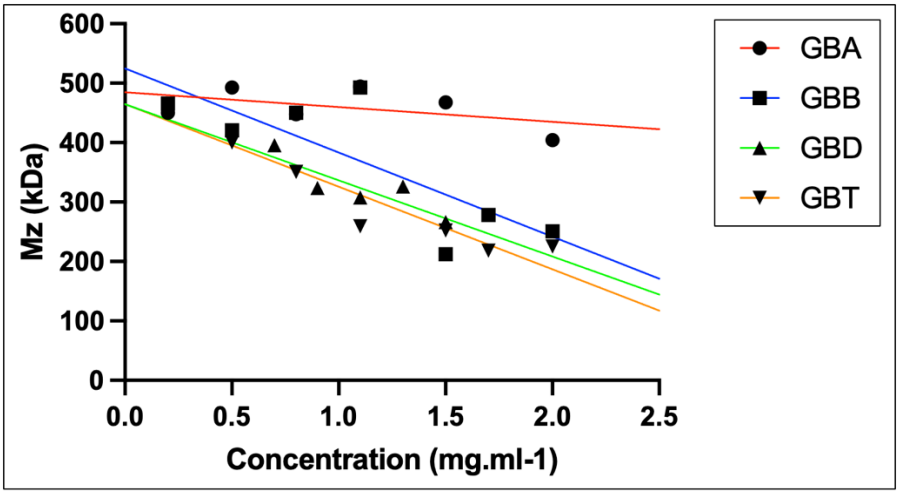

Figure 10. Weight average molecular weight, M_z, vs concentration plots produced by sedimentation equilibrium of each of the 4 glycoconjugates in PBS buffer, pH = 7.0, I = 0.1 mol.L^{-1} and a rotor speed of 10,000 rpm.

Table 9. Sedimentation equilibrium data from for each of the conjugate vaccine candidates.

Sample	$M_{w,app}$ M*(kDa)	$M_{w,app}$ Hinge (kDa)	M_z (kDa)
GBA-CRM-026	343.3	262.3	485.8
GBB-CRM-008B	362.7	284.7	450.2
GBD-CRM-005	424.3	297.7	464.7
GBT-CRM-003	349.1	302.9	464.5

N.B Two different calculations for weight average molecular weight (M_w), M^* and Hinge point, are performed. z-average molecular weight (M_z) is also reported. Estimates of molecular weight are made from extrapolations of M^* and Hinge point vs c to eliminate the effects on nonideality (Figure 9). Estimates of M_z are made from extrapolations of concentration dependence plots extrapolated to zero concentration (Figure 10). The differences between the M^* and Hinge methods are most likely due to the presence of aggregates which have a greater effect on the M^* estimates (Schuck et al., 2014).

3.2.3. SEC-MALS

SEC-MALS analysis was also performed for each of the glycoconjugate vaccine candidates.

Similarly, to the initial sedimentation velocity data, SEC-MALS revealed two distinct species

present in each of the samples. Where the two methods differ is the size estimations for each species. The SEC-MALS data give a much higher estimate for both the weight average and z-average molecular weights compared to sedimentation equilibrium (Table 10). While some larger species were present in the mega Dalton range during SE-AUC, they mostly occupied <1 % of the sedimentation distribution and were therefore not optically registered. The mass fraction calculated during SEC-MALS for each of these large peaks was between 29-40 %. It could be the case that this peak is in fact multiple species whose combined molecular weights are leading to this large M_w. This is evidenced by the slight hump to the right of the main peak (Figure 11). Additionally, the column may have been contaminated by larger molecular weight species. Given that the sedimentation data indicated that the conjugate vaccine candidates are rather similar regarding molecular weight and sedimentation coefficient, the large differences observed in the SEC-MALS data are surprising (Table 8, Table 9).

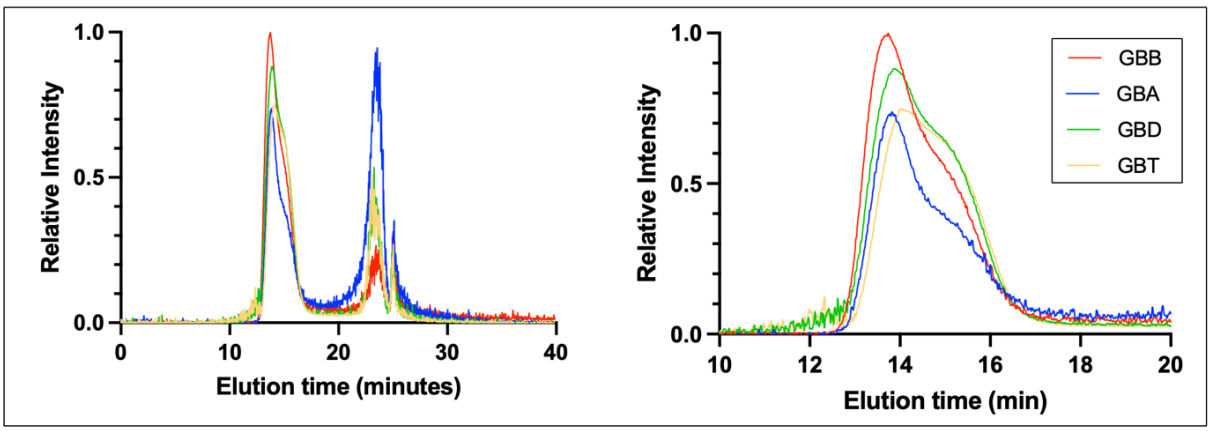

Figure 11. Combined multi angle light scattering data from SEC-MALS analysis of each conjugate, including a highlighted view of the first peak.

Table 10. Average molecular weight estimates from SEC-MALS analysis of each conjugate vaccine candidate.

Sample	Peak 1		Peak 2	
	M_w (kDa)	M_z (kDa)	M_w (kDa)	M_z (kDa)
GBA-CRM-026	2374	3120	759	798
GBB-CRM-008B	1679	2098	459	495
GBD-CRM-005	944	1129	332	346
GBT-CRM-003	725	817	274	286

N.B. Estimates for molecular weight from SEC-MALS are accurate to within ~ (± 5 %).

3.2.4. Viscosity

Estimates for intrinsic viscosity from rolling ball viscometry of each conjugate shows that there is a lot of variation between samples (Table 11). It is possible that this variation is caused by differences in the structures and viscosities of the polysaccharide serotypes conjugated to each sample but again, the previously reported intrinsic viscosities for each serotype disagree with this proposal. The intrinsic viscosities of the sized serotype Ia, Ib, and II polysaccharides are reported as 48.7, 62.1 and 55.0 $mL.g^{-1}$ respectively, while the intrinsic viscosity of the native serotype III polysacharide is reported as 44.7 $mL.g^{-1}$ (MacCalman, 2023). While both the sized serotype II polysaccharide and the serotype II conjugate (GBB-CRM-008B) have the highest estimated viscosity, the serotype III conjugate (GBT-CRM-005) does not have the lowest viscosity, while the serotype III polysaccharide does (Table 11). Perhaps it is the case that the conjugation efficiency of each polysaccharide is different and therefore a different number of polysaccharide chains ends up on the final molecule leading to variation in the viscosity. It should be noted that some of the data for the conjugates is poor. Specifically, the relative viscosity of GBA-CRM-026 should not be decreasing with concentration so some instrumental error may be involved in the intrinsic viscosity estimates.

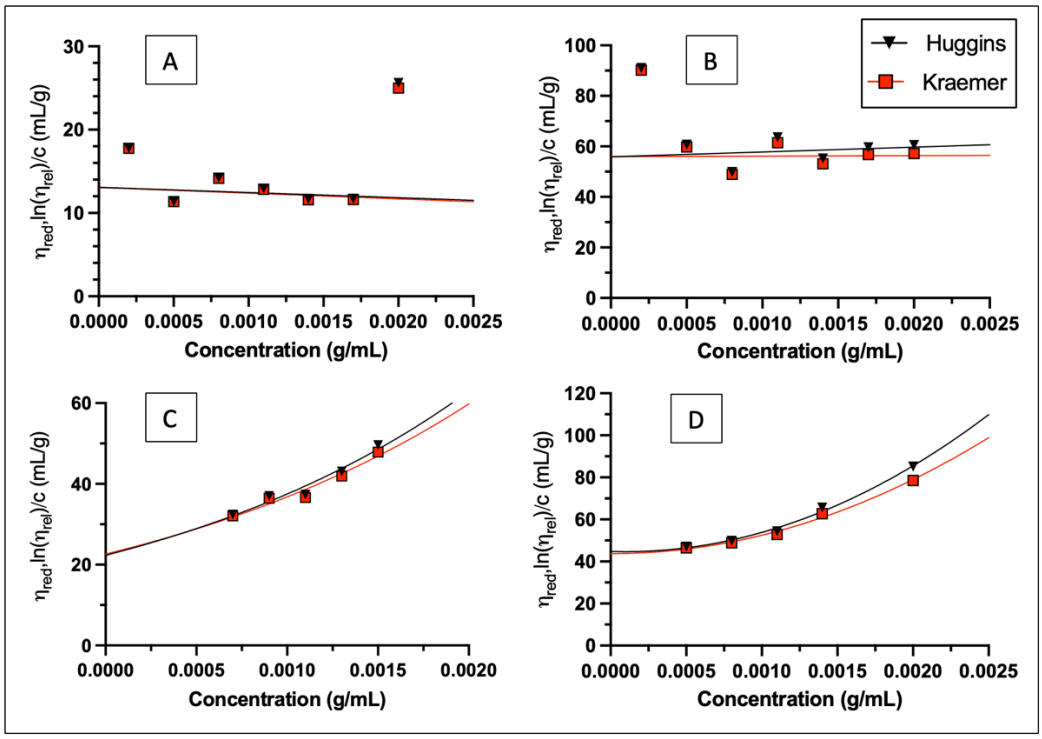

Figure 12. Plots of reduced viscosity (Huggins) and inherent viscosity (Kraemer) with extrapolation to zero concentration to give an estimate for intrinsic viscosity free from the effects of nonideality for: A, GBA-CRM-026; B, GBB-CRM-008B; C, GBD-CRM-005; D, GBT-CRM-003. (Anomalous results are either omitted or excluded from the trend lines).

Table 11. Intrinsic viscosities for each of the conjugate vaccine candidates estimated from extrapolations of concentration dependence plots (Figure 12).

Sample	Intrinsic Viscosity, $[\eta]$ (mL.g^{-1})
GBA-CRM-026	13.1 ± 0.1
GBB-CRM-008B	55.9 ± 0.1
GBD-CRM-005	22.5 ± 0.2
GBT-CRM-003	44.3 ± 0.7

4. Discussion

Despite the sample of CRM_{197} stored at 4.0°C and the sample incubated at 30.0°C having slightly more aggregating species than the freshly thawed sample, both still fall well within the 90% protein purity requirement specified by the world health organisation for the control of CRM_{197} for use in glycoconjugate vaccines (World Health and Standardization, 2013). Under all experimental conditions CRM_{197} has significantly less aggregate than the related carrier protein, dTT. Previous sedimentation velocity experiments analysing samples of dTT found they comprised an average of (86 ± 2) % monomer and (14 ± 2) % dimer across a range of 7 loading concentration from 0.1 – 2.0 $mg.mL^{-1}$ (Abdelhameed et al., 2012). This shows there is significantly less interaction between monomers of CRM_{197} than the related carrier protein that sees frequently and continued used in glycoconjugate vaccine formulations (Kensinger and Arunachalam, 2022). This increase in aggregation could be due to the introduction of random structural changes during the detoxification process. Typically, formaldehyde treatment causes Schiff base modification in tetanus toxin which undergo cross linking reactions with tyrosine residues (Long et al., 2022). Given that charge state and hydrophobicity are known factors in aggregation, the introduction of cross links could alter either of these properties and therefore induce aggregation (Trainor et al., 2017). It appears that the retention of the native structure due to genetic detoxification of CRM_{197} is implicit in the lower degree of aggregation when compared to chemical detoxification (Uchida et al., 1972). Although dTT and CRM_{197} are not the same protein, they are closely related. Much of the modification induced by formaldehyde treatment of dTT is also observed after detoxification of diphtheria toxin (Metz et al., 2020). I would therefore expect to see a similar degree of aggregation if detoxified diphtheria toxin were to be analysed, and this would be

the next logical comparison against CRM_{197}. The superiority of a genetically detoxified toxoid carrier compared to a chemically detoxified one is given more credence by recent efforts to develop, and optimise the production of, a full length genetically detoxified tetanus toxin (Chang et al., 2022).

Where the stability of CRM_{197} does come into question is under agitation. As previously reported, repeated flow of samples of CRM_{197} through a capillary viscometer caused the formation of visible aggregates. Experiments investigating the effects of shear flow on protein unfolding found that even very high shear forces induced by passing a model globular protein, horse cytochrome c, through a ~ 160 μm at a rate of up to 10 $m.s^{-1}$ had little to no effect on the stability of the folded protein (Jaspe and Hagen, 2006). As these flow rates are significantly higher than those applied to CRM_{197} it is unlikely that the protein is aggregating under shear force alone. Studies into large extensional flow under pressure (e.g. in syringes) tell a different story, however. Dobson et al. (2017) revealed that extensional flow conditions induce unfolding and aggregation of several proteins, including bovine serum albumin and various immunoglobulins. Regarding Ostwald viscometry, the movement of the sample from the reservoir to the capillary under constant force applied by the Schott Gerate AVS 400 viscometer prior to the creeping flow measurements under gravity may have caused a rapid local increase in fluid viscosity generating an extensional flow (Willis et al., 2018). It is possible that this flow prior to the start of viscometer timing caused CRM_{197} to unfold with time exposing otherwise concealed sequences who have a higher propensity for aggregation and led to the visible aggregate observed after the experiments (Dobson et al., 2017). Dynamic light scattering of samples before and after U-tube viscosity measurements would give a good indication as to the level of aggregation caused so would be a good suggestion for further study.

An interesting trend was the tendency for lower concentrations of freshly thawed CRM_{197} to have higher aggregation. This is in direct conflict with what is usually expected for protein as aggregation is a second order reaction, where an increase in concentration would lead to an increase in collisions and therefore greater aggregation (Lumry and Eyring, 1954). For absorbance, 2.0 and 1.0 $mg.mL^{-1}$ samples had 0.0 and 1.7 % dimer respectively, while 0.125 and 0.0625 $mg.mL^{-1}$ had 4.0 and 4.3 % respectively (Appendix 1). For interference, 2.0 and 1.0 $mg.mL^{-1}$ samples had 3.0 and 3.6 % dimer respectively, while 0.125 and 0.0625 $mg.mL^{-1}$ had 6.3 and 10.0 % respectively (Appendix 2). Inverse relationships between concentration and aggregation have been reported for proteins that experience agitation. This is due to the increased air/water interface to protein ratio (Treuheit et al., 2002). CRM_{197} may be sensitive to these effects also, so perhaps more care should have been taken when handling the protein. These effects were not previously observed for dTT, which exhibited a normal concentration, aggregation relationship (Abdelhameed et al., 2012). In this instance, while the formaldehyde treatment leads to overall increases in aggregation, dTT may have an increased resistance to unfolding due to cross-linking and therefore is less prone to aggregation under agitation (Thaysen-Andersen et al., 2007). It should be noted that a concentration of 0.0625 $mg.mL^{-1}$ is just below the minimum sample concentration for absorbance analysis of typical proteins (~ 0.08 $mg.mL^{-1}$) – depending on the extinction coefficient and is approaching the minimum sample concentration for interference analysis of typical proteins (~ 0.06 $mg.mL^{-1}$) – see for example Cole et al. (2008). The sensitivity of the two optical systems may therefore be too low to give accurate data at this concentration.

It is difficult to meaningfully comment on the hydrodynamics of each of the conjugate vaccines at this stage given the high degree of polydispersity observed, leading to differences in the

hinge and *M** methods for molecular weight extraction from sedimentation equilibrium and with SEC-MALS. The age of the samples also ranged from 6 – 7 years so the stability and structure could have changed significantly over that period. Assuming each of the conjugates was produced and manufactured identically, it could be inferred that the serotype of polysaccharide significantly influences the hydrodynamic properties of the whole conjugate. If we consider the initial peak for each of the velocity experiments to be unbound CRM_{197} then the serotype also affects conjugation efficiency. For example, ~37.2 % of GBA-CRM-026 and ~60.1 % of GBT-CRM-003 was monomeric CRM_{197}, indicating that serotype 1a had a better binding affinity for the carrier protein than serotype III (Table 2). Before further conclusions can be made, further sedimentation equilibrium measurements will be needed at different equilibrium speeds: lower speeds to focus on the high molecular weight aggregates, and higher speeds focusing on the macromolecular component.

5. Conclusions

The results reported here have implications in the development process of conjugate vaccines or therapeutics. While CRM_{197} appears to aggregate less than the related dTT under regular storage conditions, its tendency to aggregate under agitation or under specific flow conditions could be overlooked during manufacturing. This could be particularly important during transportation or large-scale production of CRM_{197} where these conditions are common. To make more meaningful conclusions about the influence of chemical detoxification on the hydrodynamics and aggregation of carrier proteins, further experiments and comparison should be performed with the chemically detoxified diphtheria toxin. Regarding the conjugate vaccines, repeat experiments would have to be conducted using fresher samples to properly

investigate the long-term stability of the molecules as well as the influence of serotype on the hydrodynamic properties.

6. Acknowledgments

Thank you to Jennifer Wakefield, Jacob Pattem and Vlad Dinu at the National Centre for Molecular Hydrodynamics (NCMH), University of Nottingham, for their help and guidance throughout my research. I would like especially thank Stephen Harding, NCMH, University of Nottingham, for his invaluable guidance, support, and mentorship over the course of the past year. Further gratitude is given to Ghislain Delpierre and Pierre Duvivier from GSK Vaccines, Rixensart, for their assistance throughout the project, and for supplying the proteins and glycoconjugates that made the project possible. Final thanks to the School of Biosciences, University of Nottingham, for funding this research.

7. References

ABDELHAMEED, A. S., MORRIS, G. A., ADAMS, G. G., ROWE, A. J., LALOUX, O., CERNY, L., BONNIER, B., DUVIVIER, P., CONRATH, K., LENFANT, C. & HARDING, S. E. 2012. An asymmetric and slightly dimerized structure for the tetanus toxoid protein used in glycoconjugate vaccines. *Carbohydrate Polymers,* 90**,** 1831-1835.

ALEXANDER, W. J., SHAW, J. F. E. & HOLMES, J. R. 1987. Haemophilus-influenzae infections following licensure of the Hib vaccine. *New England Journal of Medicine,* 317**,** 167-168.

ANG, S. & ROWE, A. J. 2010. Evaluation of the information content of sedimentation equilibrium data in self-interacting systems. *Macromol Biosci,* 10**,** 798-807.

AW, R., ASHIK, M. R., ISLAM, A. A. Z. M., KHAN, I., MAINUDDIN, M., ISLAM, M. A., AHASAN, M. M. & POLIZZI, K. M. 2021. Production and purification of an active CRM197 in Pichia pastoris and its immunological characterization using a Vi-typhoid antigen vaccine. *Vaccine,* 39**,** 7379-7386.

CARRERAS-ABAD, C., RAMKHELAWON, L., HEATH, P. T. & LE DOARE, K. 2020. A Vaccine Against Group B Streptococcus: Recent Advances. *Infect Drug Resist,* 13**,** 1263-1272.

CENTER FOR BIOLOGICS EVALUATION AND RESEARCH. 2022. *U.S. Food and Drug Administration* [Online]. FDA. Available: https://www.fda.gov/vaccines-blood-biologics/vaccines/vaccines-licensed-use-united-states [Accessed 13/04/2023 2023].

CHAHINE, E. B., DOUGHERTY, J. A., THORNBY, K. A. & GUIRGUIS, E. H. 2022. Antibiotic Approvals in the Last Decade: Are We Keeping Up With Resistance? *Ann Pharmacother,* 56**,** 441-462.

CHANG, M.-J., OLLIVAULT-SHIFLETT, M., SCHUMAN, R., NGOC NGUYEN, S., KALTASHOV, I. A., BOBST, C., RAJAGOPAL, S. P., PRZEDPELSKI, A., BARBIERI, J. T. & LEES, A. 2022. Genetically detoxified tetanus toxin as a vaccine and conjugate carrier protein. *Vaccine,* 40**,** 5103-5113.

COLE, J. L., LARY, J. W., T, P. M. & LAUE, T. M. 2008. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol,* 84**,** 143-79.

CRESS, B. F., ENGLAENDER, J. A., HE, W., KASPER, D., LINHARDT, R. J. & KOFFAS, M. A. G. 2014. Masquerading microbial pathogens: capsular polysaccharides mimic host-tissue molecules. *FEMS Microbiology Reviews,* 38**,** 660-697.

DOBSON, J., KUMAR, A., WILLIS, L. F., TUMA, R., HIGAZI, D. R., TURNER, R., LOWE, D. C., ASHCROFT, A. E., RADFORD, S. E., KAPUR, N. & BROCKWELL, D. J. 2017. Inducing protein aggregation by extensional flow. *Proceedings of the National Academy of Sciences,* 114**,** 4673-4678.

GENO, K. A., GILBERT, G. L., SONG, J. Y., SKOVSTED, I. C., KLUGMAN, K. P., JONES, C., KONRADSEN, H. B. & NAHM, M. H. 2015. Pneumococcal capsules and their types: Past, present, and future. *Clinical Microbiology Reviews,* 28**,** 871-899.

HARDING, S., HORTON, J. & MORGAN, P. 1992. MSTAR: a FORTRAN program for the model independent molecular weight analysis of macromolecules using low speed or high speed sedimentation equilibrium. *Analytical ultracentrifugation in biochemistry and polymer science. Royal Society of Chemistry, Cambridge***,** 275-294.

HARDING, S. E. 1997. The intrinsic viscosity of biological macromolecules. Progress in measurement, interpretation and application to structure in dilute solution. *Progress in biophysics and molecular biology,* 68 2-3**,** 207-62.

HARDING, S. E., HORTON, J. C. & CÖLFEN, H. 1997. The ELLIPS suite of macromolecular conformation algorithms. *Eur Biophys J,* 25**,** 347-59.

HINZ, H. J. 1986. *Thermodynamic Data for Biochemistry and Biotechnology,* New York, Springer-Verlag.

HUGGINS, M. L. 1942. The Viscosity of Dilute Solutions of Long-Chain Molecules. IV. Dependence on Concentration. *Journal of the American Chemical Society,* 64**,** 2716-2718.

JASPE, J. & HAGEN, S. J. 2006. Do protein molecules unfold in a simple shear flow? *Biophys J,* 91**,** 3415-24.

JONES, C. 2015. Glycoconjugate vaccines: The regulatory framework. *In:* LEPENIES, B. (ed.) *Carbohydrate-Based Vaccines: Methods and Protocols.* New York, NY: Springer New York.

JONES, H. E., TAYLOR, P. R., MCGREAL, E., ZAMZE, S. & WONG, S. Y. C. 2009. The contribution of naturally occurring IgM antibodies, IgM cross-reactivity and complement dependency in murine humoral responses to pneumococcal capsular polysaccharides. *Vaccine,* 27**,** 5806-5815.

KENSINGER, R. & ARUNACHALAM, A. B. 2022. Preclinical development of the quadrivalent meningococcal (ACYW) tetanus toxoid conjugate vaccine, MenQuadfi®. *Glycoconjugate Journal,* 39**,** 381-392.

KRAEMER, E. O. 1938. Molecular Weights of Celluloses and Cellulose Derivates. *Industrial & Engineering Chemistry,* 30**,** 1200-1203.

KUNTZ, I. D., JR. 1971. Hydration of macromolecules. III. Hydration of polypeptides. *Journal of the American Chemical Society,* 93**,** 514-516.

LAUE, T. M. 1992. Computer-aided interpretation of analytical sedimentation data for proteins. *Analytical ultracentrifugation in biochemistry and polymer science.***,** 90-125.

LONG, Z., WEI, C., ROSS, R., LUO, X., MA, X., QI, Y., CHAI, R., CAO, J., HUANG, M. & BO, T. 2022. Effects of detoxification process on toxicity and foreign protein of tetanus toxoid and diphtheria toxoid. *Journal of Chromatography B,* 1207**,** 123377.

LUMRY, R. & EYRING, H. 1954. Conformation Changes of Proteins. *The Journal of Physical Chemistry,* 58**,** 110-120.

MACCALMAN, T. 2023. Hydrodynamics and stability of Group B Streptococcus polysaccharides and CRM197 protein.

MACLEOD , C. M., HODGES , R. G., HEIDELBERGER , M. & BERNHARD , W. G. 1945. Prevention of Pneumococcal pneumonia by immunization with specific capsular polysaccharides *Journal of Experimental Medicine,* 82**,** 445-465.

MALITO, E., BURSULAYA, B., CHEN, C., SURDO, P. L., PICCHIANTI, M., BALDUCCI, E., BIANCUCCI, M., BROCK, A., BERTI, F., BOTTOMLEY, M. J., NISSUM, M., COSTANTINO, P., RAPPUOLI, R. & SPRAGGON, G. 2012. Structural basis for lack of toxicity of the diphtheria toxin mutant CRM197. *Proceedings of the National Academy of Sciences,* 109**,** 5229-5234.

MALTONI, G., SCUTTERI, L., MENSITIERI, F., PIAZ, F. D. & HOCHKOEPPLER, A. 2022. High-yield production in Escherichia coli and convenient purification of a candidate vaccine against SARS-CoV-2. *Biotechnology Letters,* 44**,** 1313-1322.

METZ, B., MICHIELS, T., UITTENBOGAARD, J., DANIAL, M., TILSTRA, W., MEIRING, H. D., HENNINK, W. E., CROMMELIN, D. J. A., KERSTEN, G. F. A. & JISKOOT, W. 2020. Identification of formaldehyde-induced modifications in Diphtheria toxin. *J Pharm Sci,* 109**,** 543-557.

MICOLI, F., ALFINI, R., DI BENEDETTO, R., NECCHI, F., SCHIAVO, F., MANCINI, F., CARDUCCI, M., PALMIERI, E., BALOCCHI, C., GASPERINI, G., BRUNELLI, B., COSTANTINO, P., ADAMO, R., PICCIOLI, D. & SAUL, A. 2020. GMMA Is a versatile platform to design effective multivalent vombination vaccines. *Vaccines (Basel),* 8.

MISHRA, R. P. N., YADAV, R. S. P., JONES, C., NOCADELLO, S., MINASOV, G., SHUVALOVA, LUDMILLA A., ANDERSON, WAYNE F. & GOEL, A. 2018. Structural and immunological characterization of E. coli derived recombinant CRM197 protein used as carrier in conjugate vaccines. *Bioscience Reports,* 38.

RAPPUOLI, R. 2018. Glycoconjugate vaccines: Principles and mechanisms. *Sci Transl Med,* 10.

RAPPUOLI, R., DE GREGORIO, E. & COSTANTINO, P. 2019. On the mechanisms of conjugate vaccines. *Proceedings of the National Academy of Sciences,* 116**,** 14-16.

RUDD, B. D. 2020. Neonatal T vells: A reinterpretation. *Annu Rev Immunol,* 38**,** 229-247.

SAITÔ, N. 1951. The effect of the Brownian motion on the viscosity of solutions of macromolecules, I. Ellipsoid of revolution. *Journal of the Physical Society of Japan,* 6**,** 297-301.

SCHUCK, P. 1998. Sedimentation analysis of noninteracting and self-associating solutes using numerical solutions to the Lamm equation. *Biophys J,* 75**,** 1503-12.

SCHUCK, P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical journal,* 78**,** 1606-1619.

SCHUCK, P., GILLIS, R. B., BESONG, T. M., ALMUTAIRI, F., ADAMS, G. G., ROWE, A. J. & HARDING, S. E. 2014. SEDFIT-MSTAR: molecular weight and molecular weight distribution analysis of polymers by sedimentation equilibrium in the ultracentrifuge. *Analyst,* 139**,** 79-92.

SIMHA, R. 1940. The influence of Brownian movement on the viscosity of solutions. *The Journal of physical chemistry,* 44**,** 25-34.

SINGH, J. K., ADAMS, F. G. & BROWN, M. H. 2018. Diversity and Function of Capsular Polysaccharide in Acinetobacter baumannii. *Front Microbiol,* 9**,** 3301.

SOLOMON, O. F. & CIUTĂ, I. Z. 1962. Détermination de la viscosité intrinsèque de solutions de polymères par une simple détermination de la viscosité. *Journal of Applied Polymer Science,* 6**,** 683-686.

TANFORD, C. 1955. Intrinsic viscosity and kinematic viscosity. *The Journal of Physical Chemistry,* 59**,** 798-799.

THAYSEN-ANDERSEN, M., JØRGENSEN, S. B., WILHELMSEN, E. S., PETERSEN, J. W. & HØJRUP, P. 2007. Investigation of the detoxification mechanism of formaldehyde-treated tetanus toxin. *Vaccine,* 25**,** 2213-2227.

TRAINOR, K., BROOM, A. & MEIERING, E. M. 2017. Exploring the relationships between protein sequence, structure and solubility. *Curr Opin Struct Biol,* 42**,** 136-146.

TREUHEIT, M. J., KOSKY, A. A. & BREMS, D. N. 2002. Inverse Relationship of Protein Concentration and Aggregation. *Pharmaceutical Research,* 19**,** 511-516.

UCHIDA, T., PAPPENHEIMER, A. M. & HARPER, A. A. 1972. Reconstitution of Diphtheria toxin from two nontoxic cross-reacting mutant proteins. *Science,* 175**,** 901-903.

UK HEALTH SECURITY AGENCY 2022. English Surveillance Programme for Antimicrobial Utilisation and Resistance (ESPAUR) Report 2021 to 2022.

WILLIS, L. F., KUMAR, A., DOBSON, J., BOND, N. J., LOWE, D., TURNER, R., RADFORD, S. E., KAPUR, N. & BROCKWELL, D. J. 2018. Using extensional flow to reveal diverse aggregation landscapes for three IgG1 molecules. *Biotechnol Bioeng,* 115**,** 1216-1225.

WORLD HEALTH, O. 2022. *WHO expert committee on biological standardization: seventy-fifth report,* Geneva, World Health Organization.

WORLD HEALTH, O. & STANDARDIZATION, W. H. O. E. C. O. B. 2013. *WHO Expert Committee on Biological Standardization, sixtieth report,* Geneva, World Health Organization.

WORLD HEALTH ORGANISATION 2013. Haemophilus influenzae type b (Hib) vaccination position paper – July 2013. *Wkly Epidemiol Rec,* 88**,** 413-26.

WYATT, P. J. 1993. Light scattering and the absolute characterization of macromolecules. *Analytica chimica acta,* 272**,** 1-40.

XIONG, A. W., FANG, J. M., REN, S. X., LI, W., WANG, J., ZHAO, Y., CHEN, G. Y., XU, Q. & ZHOU, C. C. 2021. A novel combined conjugate therapeutic cancer vaccine, recombinant EGF-CRM197, in patients with advanced solid tumors: A phase I clinical study. *Frontiers in Oncology,* 11.

YE, J. & CHEN, X. 2022. Current Promising Strategies against Antibiotic-Resistant Bacterial Infections. *Antibiotics (Basel),* 12.

8. Appendix

Appendix 1. Relative proportions dimer for 6 loading concentrations of CRM_{197} under three different experimental conditions determined by sedimentation velocity. Data collected by absorbance optics (280 nm).

Concentration (mg.mL^{-1})	**Proportion of dimeric peak (%)**		
	Freshly Thawed	**3 weeks at 4.0℃**	**4 hours at 30.0℃**
2.00	-	2.2	0.9
1.00	1.7	4.6	2.4
0.50	1.7	9.1	1.9
0.25	1.8	0.3	7.4
0.125	4.0	0.9	-
0.0625	4.3	3.9	2.7

Appendix 2. Relative proportions dimer for 6 loading concentrations of CRM_{197} under three different experimental conditions determined by sedimentation velocity. Data collected by Rayleigh interference optics.

Concentration (mg.mL^{-1})	**Proportion of dimeric peak (%)**		
	Freshly Thawed	**3 weeks at 4.0℃**	**4 hours at 30.0℃**
2.00	3.0	3.0	7.4
1.00	3.6	4.1	5.9
0.50	1.9	1.9	6.4
0.25	2.8	6.1	-
0.125	6.3	7.8	8.9
0.0625	10.0	8.1	7.4