Deciphering protective immunity by multimodal mass spectrometry: Towards epitope-focused streptococcal vaccines

Tang, Di

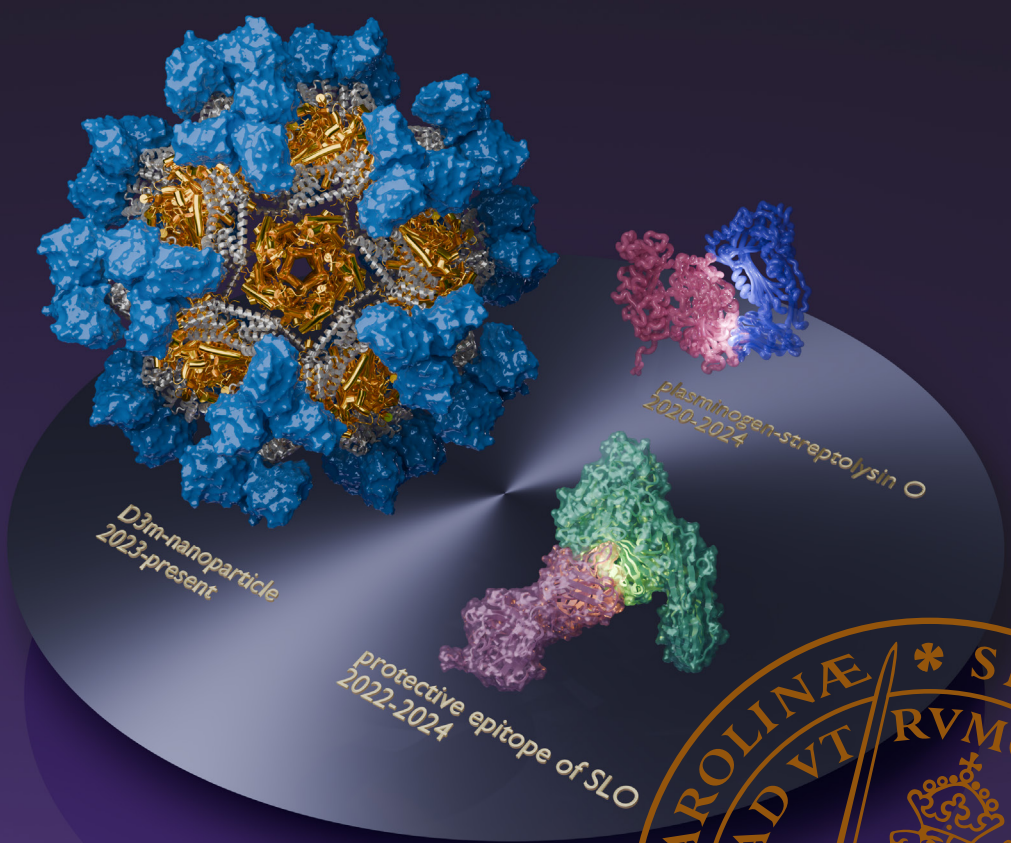
2025

Document Version:
Publisher's PDF, also known as Version of record

Deciphering protective immunity by multimodal mass spectrometry

Towards epitope-focused streptococcal vaccines

DI TANG
DEPARTMENT OF CLINICAL SCIENCES, LUND | FACULTY OF MEDICINE | LUND UNIVERSITY



Deciphering protective immunity by multimodal mass spectrometry
Towards epitope-focused streptococcal vaccines

Deciphering protective immunity by multimodal mass spectrometry

Towards epitope-focused streptococcal vaccines

Di Tang

LUND
UNIVERSITY

DOCTORAL DISSERTATION

Doctoral dissertation for the degree of Doctor of Philosophy (PhD) at the Faculty of Medicine at Lund University to be publicly defended on 14th of March 2025, Friday at 09:00 in Belfrage lecture hall, Biomedical Centre, Lund, Sweden

Faculty opponent
Dr. Martin Beck
Director, Max Planck Institute for Biophysics
Frankfurt, Germany

Organization: LUND UNIVERSITY

Document name: Doctoral dissertation **Date of issue:** 2025-03-14

Author(s): Di Tang **Sponsoring organization:**

Title and subtitle: Deciphering protective immunity by multimodal mass spectrometry: Towards epitope-focused streptococcal vaccines

Abstract:

Group A *Streptococcus* (GAS) causes infections ranging from mild sore throats to severe invasive disease; however, no licensed vaccine is currently available. In this thesis, the author leverages streptolysin O (SLO), an important GAS protein and therapeutic/vaccine target, as a model system to develop a multimodal protein mass spectrometry (MS) strategy. Through an integrated approach of MS-based proteomics, biochemical assays, and computational modelling, the author reveals that SLO not only forms pores in cell membranes but also modulates human fibrinolytic pathways, potentially enhancing bacterial dissemination. Mapping the immunologically relevant regions (namely epitopes) of SLO reveals a site that can elicit potent protective immune responses. Building on these insights, the author constructs a prototype vaccine by multivalently displaying redesigned SLO-derived epitopes on self-assembling protein nanoparticles, resulting in enhanced immunity and improved protection in a pilot mouse model. To demonstrate the broader applicability of this antibody-guided MS workflow, the author extends it to pneumolysin, a homologous cytolysin from *Streptococcus pneumoniae*, illustrating how the same methodologies can decipher epitope-specific correlates of antibody functionality against other complex bacterial antigens. In summary, this thesis highlights the potential of an integrated MS-based approach to uncover bacterial virulence mechanisms, define epitopes of biological relevance, and facilitate data-driven immunogen design. By combining complementary structural and functional analyses with the promising epitope-focused vaccine modality, the thesis offers a practical roadmap for tackling GAS infections and potentially other challenging microbial pathogens.

Key words: *Streptococcus pyogenes*, streptolysin O, host-pathogen protein-protein interaction, mass spectrometry-based proteomics, integrated structural biology, cholesterol-dependent cytolysin, antibody-guided epitope mapping, structural vaccinology, protein-based nanoparticle

Classification system and/or index terms (if any) Supplementary bibliographical information

Language: English **ISSN and key title**: 1652-8220

Lund University, Faculty of Medicine Doctoral Dissertation Series 2025:24

ISBN: 978-91-8021-677-7

Recipient's notes **Number of pages:** 220

Price Security classification

Signature Date 2025-03-14

Deciphering protective immunity by multimodal mass spectrometry

Towards epitope-focused streptococcal vaccines

Di Tang (唐笛)

LUND
UNIVERSITY

Faculty of Medicine

Department of Clinical Sciences, Lund

Lund University, Faculty of Medicine Doctoral Dissertation Series 2025:24

ISBN 978-91-8021-677-7

ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University

Lund 2025

Dedicated to my mom and dad 献给我的父母

"*You know, the proper method for enquiring after the properties of things, is to deduce them from experiments*."

- Sir Isaac Newton, 1672

Table of Contents

Preface

A PhD journey is fundamentally one of learning, self-reflection, and exploration.

Six years have passed since I arrived in Sweden. From an early age, I have been driven by the desire to challenge my limits and refine my thinking around complex questions. Coming from a humble background, I have always believed that continuous learning is key to charting a path towards broader horizons.

Finishing this thesis evokes a mixture of emotions, marking both the end of one chapter and the start of another. Throughout this journey, I have pushed myself out of my comfort zone on numerous occasions. These efforts result in this thesis, which I am proud to present as a testament to years of hard work and personal growth.

The thesis is structured into three sections: (i) an introduction to the biological background and the methods applied, (ii) four core papers detailing key findings and their broader implications, and (iii) an extended discussion offering perspectives for future directions.

I hope that this work contributes to progress in related fields and provides encouragement to those defining their own paths. May the insights shared here become a meaningful addition to collective scientific endeavours and offer inspiration to those striving to reach their goals.

Di Tang

26th December 2024

Lund, Sweden

Abstract

Group A *Streptococcus* (GAS) causes infections ranging from mild sore throats to severe invasive disease; however, no licensed vaccine is currently available. In this thesis, the author leverages streptolysin O (SLO), an important GAS protein and therapeutic/vaccine target, as a model system to develop a multimodal protein mass spectrometry (MS) strategy. Through an integrated approach of MS-based proteomics, biochemical assays, and computational modelling, the author reveals that SLO not only forms pores in cell membranes but also modulates human fibrinolytic pathways, potentially enhancing bacterial dissemination. Mapping the immunologically relevant regions (namely epitopes) of SLO reveals a site that can elicit potent protective immune responses. Building on these insights, the author constructs a prototype vaccine by multivalently displaying redesigned SLO-derived epitopes on self-assembling protein nanoparticles, resulting in enhanced immunity and improved protection in a pilot mouse model. To demonstrate the broader applicability of this antibody-guided MS workflow, the author extends it to pneumolysin, a homologous cytolysin from *Streptococcus pneumoniae*, illustrating how the same methodologies can decipher epitope-specific correlates of antibody functionality against other complex bacterial antigens. In summary, this thesis highlights the potential of an integrated MS-based approach to uncover bacterial virulence mechanisms, define epitopes of biological relevance, and facilitate data-driven immunogen design. By combining complementary structural and functional analyses with the promising epitope-focused vaccine modality, the thesis offers a practical roadmap for tackling GAS infections and potentially other challenging microbial pathogens.

Popular science summary

Every year, Group A *Streptococcus* (GAS) causes millions of infections, ranging from mild sore throats to severe, life-threatening conditions. Unfortunately, there is still no licensed vaccine available for GAS.

In this thesis, we employ a series of methods based on mass spectrometry (MS), a powerful tool for analysing biological molecules such as proteins, to gain deeper insights into how these bacteria cause disease and to facilitate the development of better vaccines. A key discovery was that a GAS-produced toxin, streptolysin O (SLO), not only damages human cells by creating pores on the surface but also binds to an important human protein called plasminogen, thereby likely aiding the bacterium's spread.

Further, by integrating multiple MS-based techniques, we were able to pinpoint the precise regions of SLO that the immune system can target most effectively. We then harnessed this crucial component to create a "targeted vaccine" by attaching it to tiny (nanometre-scale) protein particles. In our pilot study, this strategy enabled mice to mount an improve immune protection against invasive bacterial infections.

Finally, we applied the same workflow to investigate pneumolysin, a key toxin produced in *S. pneumoniae*. By mapping how different antibodies recognise and modulate PLY function, we discovered shared "weak spots" that could inform broader vaccine development against other bacteria with similar toxins.

Overall, this thesis demonstrates how powerful, MS-based methodologies can be used to identify the most critical segments of bacterial toxins and direct the body's defences towards them. The overarching goal is to develop safer and more effective vaccines that protect individuals from serious infections and address the growing threat of antibiotic resistance. By uniting fundamental discoveries with emerging vaccine technologies, this work outlines a plan not only for tackling streptococcal diseases but potentially other dangerous pathogens as well. Through these integrated approaches, we may move one step closer to a world less burdened by severe bacterial infections.

Populärvetenskaplig sammanfattning

Varje år orsakar grupp A *streptokocker* (GAS) miljontals infektioner världen över, från enklare halsfluss till allvarliga och ibland livshotande sjukdomar. Trots detta saknas fortfarande ett godkänt vaccin mot GAS.

I denna avhandling har vi använt en serie avancerade metoder baserade på masspektrometri (MS), ett verktyg för att undersöka och analysera proteiner i detalj. Genom dessa tekniker får vi en tydligare bild av hur bakterierna orsakar sjukdom och hur vi på sikt kan utveckla bättre vacciner. En viktig upptäckt var att toxinet streptolysin O (SLO), som produceras av GAS, inte bara orsakar skador på celler genom att göra hål i dem utan också kan binda till ett viktigt protein i kroppen, plasminogen, vilket sannolikt hjälper bakterien att sprida sig.

Vidare kunde vi, med hjälp av flera olika MS-baserade metoder, noggrant kartlägga vilka delar av SLO som immunförsvaret känner igen allra bäst. Dessa särskilt avgörande delar av proteinet utnyttjades sedan för att utveckla ett så kallat ”målinriktat vaccin”, där dessa komponenter fästes på mycket små (nanometerstora) proteinpartiklar. I våra inledande studier såg vi att detta förfarande hjälpte möss att få ett bättre immunskydd mot invasiva bakterieinfektioner.

Samma arbetsmetod användes även för att studera pneumolysin, ett toxin från bakterien *S. pneumoniae*. Vi undersökte hur olika antikroppar känner igen och påverkar detta toxin och kunde då identifiera gemensamma svaga punkter som i sin tur kan bana väg för vacciner med bredare skydd mot andra bakterier med liknande toxiner.

Sammanfattningsvis visar denna avhandling hur olika MS-tekniker kan användas för att hitta de allra viktigaste delarna av bakteriella toxiner och rikta kroppens försvar mot just dessa områden. Det övergripande målet är att utveckla säkrare och mer effektiva vacciner som skyddar människor från allvarliga infektioner och samtidigt hjälper oss att hantera det växande problemet med antibiotikaresistens. Genom att kombinera grundläggande forskningsresultat med nya vaccintekniker kan vi förhoppningsvis ta ett steg närmare en framtid där farliga bakterieinfektioner är betydligt mer sällsynta.

List of papers

Papers included in this thesis

1.1.1.1 Paper I

"Streptolysin O Accelerates the Conversion of Plasminogen to Plasmin"

Tang, D., Khakzad, H., Hjortswang, E., Malmström, L., Ekström, S., Happonen, L., & Malmström, J.

Nature Communications 15, p. 1-15 10212, November 2024

Paper II

"Multimodal Mass Spectrometry Identifies a Conserved Protective Epitope in *S. pyogenes* Streptolysin O"

Tang, D., Gueto Tettay, C. A., Hjortswang, E., Ströbaek, J., Ekström, S., Happonen, L., Malmström, L., & Malmström, J.

Analytical Chemistry 96, 22, p. 9060-9068, June 2024

Paper III

"An epitope-focused nanoparticle vaccine elicits a potent protective immune response against *Streptococcus pyogenes*"

Tang, D., Chao, Y., Hjortswang, E., Ströbaek, J., Hultgren, L., Mohanty, T., Karlsson, C., Ekström, S., Happonen, L., Shannon, O., Malmström, L., & Malmström, J. (preliminary author list)

Manuscript

Paper IV

"Mapping the structural epitope landscape of pneumolysin reveals a broadly neutralising trans-phyla B-cell epitope"

Tang, D., Kučinskaitė-Kodzė, I., Ströbaek, J., Gueto-Tettay, C. A., Hultgren, L., Håkansson, A.P., Malmström, L., Ekström, S., Happonen, L., & Malmström, J. (preliminary author list)

Manuscript

Papers not included in this thesis

"The hinge-engineered IgG1-IgG3 hybrid subclass IgGh47 potently enhances Fc-mediated function of anti-streptococcal and SARS-CoV-2 antibodies"

Izadi, A., Karami, Y., Bratanis, E., Wrighton, S., Khakzad, H., Nyblom, M., Olofsson, B., Happonen, L., **Tang, D.**, Sundwall, M., Godzwon, M., Chao, Y., Gomez Toledo, A., Schmidt, T., Ohlin, M., Nilges, M., Malmström, J., Bahnan, W., Shannon, O., Malmström, L., & Nordenfelt, P.

Nature Communications 15, p. 1-22 3600, Apr 2024

"A human monoclonal antibody bivalently binding two different epitopes in streptococcal M protein mediates immune function"

Bahnan, W., Happonen, L., Khakzad, H., Kumra Ahnlide, V., de Neergaard, T., Wrighton, S., Bratanis, E., **Tang, D.**, Hellmark, T., Björck, L., Shannon, O., Malmström, L., Malmström, J., & Nordenfelt, P.

EMBO Molecular Medicine 15, 2, p. 1-21, Feb 2023

"Multienzyme deep learning models improve peptide *de novo* sequencing by mass spectrometry proteomics"

Gueto-Tettay, C. A., **Tang, D.**, Happonen, L., Heusel, M., Khakzad, H., Malmström, J., & Malmström, L.

PLoS Computational Biology 19(1): e1010457, Jan 2023

"Epitope mapping with Sidewinder: An XL-MS and Structural Modeling Approach"

Ströbaek, J., **Tang, D.**, Gueto Tettay, C. A., Gomez Toledo, A., Olofsson, B., Hartman, E., Heusel, M., Malmström, J., & Malmström, L.

Manuscript under revision

"Anti-IL-6R prevents Experimental Type 1 Diabetes"

Luo, Z., **Tang, D.**, Zhao, T., Hamze, N., Berggren, E., Chopra, S., Safi, B., Reynolds, L., Blixt, M., Karlsson, C., Thorvaldson, L., Wang, X., Espes, D., Carlsson, P., Malmström, J., Sandler, S. & Singh, K.

Manuscript

Abbreviations

GAS	Group A *Streptococcus*
SLO	streptolysin O
PPI	protein-protein interaction
PLG	plasminogen
CDC	cholesterol-dependent-cytolysin
tPA	tissue-type plasminogen activator
SKA	streptokinase
PLM	plasmin
PLY	pneumolysin
MS	mass spectrometry
AE-MS	affinity-enrichment MS
XL-MS	cross-linking MS
HDX-MS	hydrogen/deuterium exchange MS
cryo-EM	cryogenic electron microscopy
mAb	monoclonal antibody
CDRs	complementarity-determining regions
Dx	domain x
3-D	three-dimensional
D3m	modified (*de novo* designed) domain 3 of SLO
D3m-NP	D3m-fused icosahedral symmetry protein-based nanoparticle

2 Introduction

2.1 Host-pathogen interaction

2.1.1 Group A *Streptococcus*

GAS biology

The *Streptococcus* genus encompasses numerous bacterial species that significantly influence public health, causing considerable morbidity and mortality. A milestone in this field was Lancefield's study in 1941, which identified the pathogens responsible for most haemolytic infections in humans as Group A *Streptococcus* (GAS, also known as *Streptococcus pyogenes*) [1]. As a strict human pathogen, GAS is associated with a broad range of clinical manifestations, spanning mild illnesses such as pharyngitis to severe invasive diseases, including necrotising fasciitis, streptococcal toxic shock syndrome, and sepsis. Post-infection sequelae can also present as acute rheumatic fever, rheumatic heart disease, and acute post-streptococcal glomerulonephritis [2].

Globally, GAS infections impose a substantial health burden, with over 1.78 million severe incident cases and approximately half a million related deaths annually, in addition to an estimated 727 million new non-invasive infections [3]. Despite the high mortality rate, extensive prevalence, and considerable long-term morbidities associated with GAS, the respective global significance remains under-recognised.

Key virulence factors and pathomechanisms

GAS infections involve highly complex processes driven by both host and bacterial factors that collectively shape pathogenesis. GAS produces a variety of surface-bound and secreted proteins exerting wide-ranging effects on different arms of the immune system. Among these, major virulence factors are integral to bacterial colonisation and the development of invasive disease [4].

Prominent surface-associated virulence determinants include the family of M proteins, which confer immunomodulatory functions [5]; the hyaluronic acid capsule, mediating attachment to epithelial cells [6]; and the recently characterised S protein, which enables GAS to coat with lysed erythrocyte fragments, thereby evading phagocytic killing [7]. Secreted virulence components include chemokine-degrading

proteases, namely the *S. pyogenes* cell envelope proteinase (SpyCEP) and C5α peptidase (ScpA), that cleave essential signals in the innate immune system [8,9], as well as GAS deoxyribonucleases (DNases) that dismantle neutrophil extracellular trap DNA [10]. Streptokinase (SKA) promotes bacterial dissemination by converting plasminogen to plasmin, enhancing proteolysis [11], while immunoglobulin-degrading enzymes such as IdeS and EndoS help GAS evade adaptive immune mechanisms [12,13]. Streptococcal cysteine protease (SpeB) targets a broad range of substrates, including intercellular barrier proteins and gasdermin-A (GSDMA) [14,15]. Secretory potent exotoxins, such as streptolysin O (SLO), form pores in eukaryotic cell membranes, cooperating with NAD^+ glycohydrolase to deplete host intracellular energy stores [16]. In addition, GAS produces multiple superantigens that cross-link T-cell receptors with MHC II molecules, triggering excessive T-cell activation and disproportionate cytokine release [17].

2.1.2 Host immunity and bacterial evasion

Determining the immune processes involved in natural protection against GAS infections is vital for understanding the underlining biology and for guiding the development of vaccines and innovative therapies. Although the precise mechanisms of protective immunity to GAS remain incompletely understood, it is widely accepted that the principal processes include opsonophagocytosis, neutralisation of various toxins and virulence elements (including bacterial proteases), and inhibition of bacterial adhesion. These functions are mediated by B-cell-derived antibodies and phagocytes, with support from helper T cells, pro-inflammatory cytokines, chemokines, and other innate factors [18].

In human challenge trial studies and murine pharyngitis models, early infection evokes a potent acute-phase pro-inflammatory response, marked by systematic rises in IL-1β, IL-6, IL-18, and IFNγ [19,20]. This inflammatory profile, driven in part by IL-1 signalling, can unexpectedly benefit GAS by recruiting neutrophils that help the pathogen outcompete other members of microbiota [19]. Concurrently, GAS resists immune clearance via the hyaluronic acid capsule and DNases, thereby protecting bacteria from neutrophil-mediated killing [21,22].

Superantigens further contribute by connecting MHC-II and T-cell receptors, causing a cytokine storm that includes enhanced IFNγ production, thereby intensifying inflammation [23,24]. Although T cells usually aid in infection control, superantigen-mediated hyperactivation can actually promote GAS replication by undermining effective bactericidal activities [23]. In skin infections, invasive diseases also stimulate a pro-inflammatory cytokine profile [25], though IL-1β appears protective in this setting by driving emergency granulopoiesis and restricting bacterial spread [26].

Of particular note is the pyroptosis pathway, governed by the gasdermin family, which can limit intracellular GAS survival [27]. Specifically, SpeB-mediated cleavage of gasdermin A in skin cells incites beneficial cell lysis, depriving GAS of a favourable intracellular environment [15,28].

During these escalating host-pathogen confrontations, GAS fine-tunes diverse virulence factors through transcriptional and post-translational mechanisms. CovRS (or CsrRS) detects the human antimicrobial peptide LL-37 [29], which accumulates in saliva and infected tissues with inflammation [30], enabling GAS to sense immune activation and adjust expression of the capsule, M protein, streptococcal cysteine protease, and SLO. Additionally, reactive oxygen species produced by neutrophils can inactivate streptococcal cysteine protease or SLO; however, GAS counteracts this by leveraging glutathione released from lysed cells, thus restoring toxin function and potentially upregulating superantigen activity [31,32].

Altogether, these observations highlight the precarious balance between immune-mediated pathogen clearance and GAS exploitation of inflammation. On the other hand, while the major virulence factors of GAS and their immunogenic properties have been well characterised [33,34], our understanding of the detailed molecular aspects of relevant interactions and corresponding antibody responses remains limited. In particular, there are significant gaps in knowledge and understanding regarding which epitopes are naturally recognised and targeted, the *in vivo* relevance of these antibodies, and the specific antibody characteristics (such as isotype, subclass, post-translational modifications[PTMs], and clonality) that influence antibody protective capacity.

2.1.3 Antibody and epitope

Antigen, epitope, and antibody

An antigen is any substance that can trigger an immune response, typically a foreign protein derived from bacteria or viruses. Each antigen contains smaller regions called epitopes, which are recognised by the induced antibodies with high specificity. An antibody (Ab), or immunoglobulin (Ig), is a Y-shaped protein made up of two identical heavy chains and two identical light chains. The tips of its "arms" have variable regions featuring hypervariable (HV) loops that form part of the complementarity-determining regions (CDRs), providing the binding specificity for an epitope. By binding these epitopes, antibodies can neutralise pathogens or mark them for clearance by immune cells, thus playing a crucial role in host immune defence.

Gene recombination, somatic hypermutation, and the varied pairing of heavy and light chains give rise to extensive sequence diversity exhibited in antibodies. This diversity makes sequence information essential for understanding antigen binding,

affinity maturation, and adaptive immunity [35]. Furthermore, antibodies serve as widely used research tools and therapeutic agents, requiring detailed primary structure and characterisation for validation, production, and functional optimisation [36,37].

Traditionally, antibody sequencing has involved isolating antibody-producing cells and cloning the relevant mRNAs [38]. However, such procedures can be time-consuming and come with low success rates, particularly when only limited numbers of cells remain or when they are entirely inaccessible. Secreted antibodies frequently exist independently of their original B-cells, further complicating cell-based sequencing methodologies. By contrast, methods like *de novo* MS sequencing dispenses with the need for source cells by directly investigating the antibodies as end-product proteins. This approach complements standard RNA-based sequencing by capturing sequence variants that may be overlooked in cell-dependent workflows [39].

Although T-cell responses are vital for long-term immunity against GAS, IgG antibodies remain among the most extensively characterised components of the adaptive response. Studies of antibodies targeting both serotype-specific (M protein) and conserved (non-M-protein) GAS antigens suggest these responses confer protection, paving the way for multicomponent vaccine design [40]. Emerging evidence highlights IgG's key functions in mediating opsonisation, complement activation, and prolonged immunity against GAS. In particular, IgG binding to conserved regions of the M protein has shown strong opsonic and cross-protective properties [41], while subclass-specific findings suggest that IgG3, with enhanced fragment crystallisable region (Fc) flexibility, provides robust protection [42]. Vaccine candidates directed against conserved antigens, such as the streptococcal haemoprotein receptor [43], have also demonstrated promising IgG-mediated immunity. However, the precise location of epitopes recognised by protective antibodies, and how epitope specificity relates to functional immune outcomes, remain largely unknow, mainly because the shortage of high-quality monoclonal antibodies (mAbs) and the complexity of multidomain key GAS antigens limit detailed analytical studies.

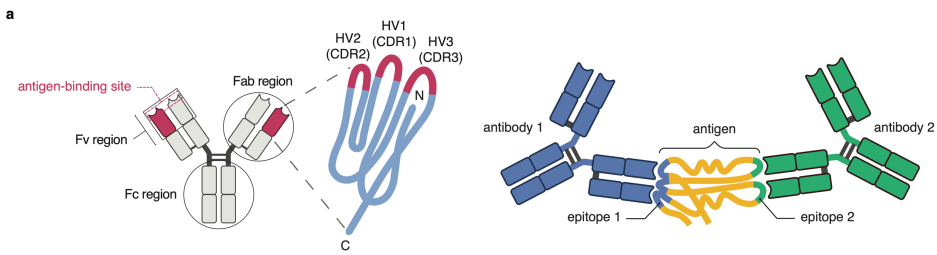


Figure 1 | Antibody and epitope of antigen. Figure Created in BioRender under publication license. Fv, variable domain; Fc: fragment crystallisable region; Fab: fragment antigen-binding; HV: hypervariable; CDR, complementarity-determining region. N, N-terminus; C, C-terminus.

Antibody-guided epitope mapping and vaccine design

A powerful way to pinpoint B-cell epitopes is through structural determination of the antigen-antibody complex by X-ray crystallography or cryogenic electron microscopy (cryo-EM), which is able to offer high-resolution insights into precise interaction sites [44–47]. However, these methods are not always feasible or accessible, prompting researchers to turn to established function-based techniques such as domain exchange [48], peptide array analysis [49], and alanine-scanning mutagenesis [50]. While each of these methods can resolve epitopes at the domain, peptide, or residue level, respectively, considerable time and resources are often required, and these methods may still fail to identify more complex or discontinuous epitopes [51].

There are many ways to extract the epitope of relevance (such as synthetic peptide, phage display, *in silico* prediction), and one direct approach for determine an epitope of function centres on isolating a monoclonal antibody that robustly blocking a key GAS pathomechanism. Such a mAb embodies the protective mechanism that researchers hope to replicate through vaccination, namely "reverse immunology". If this particular mAb, at the same time, demonstrates potent neutralising activity across multiple GAS strains, immunising with its target epitope may induce comparably broad and effective immunity *in vivo* [52]. This "reverse immunology" strategy maps an mAb's functional capacity to its corresponding epitope and subsequently reproduces it in an optimised vaccine formulation. In the context of GAS, effective implementation would enhance existing multicomponent vaccines by ensuring the inclusion of critical, functionally relevant epitopes that elicit potent, cross-protective antibody responses. Naturally, systematically characterising these protective epitopes remains essential for advancing vaccine design and deepening our understanding of GAS immunity.

Table 1. Pro and cons of each technique.

Technique	X-Ray	HDX	HRF	XL	ALN	PepArr	DomX
Speed	Slow (months)	Fast (2–4 weeks)	Fast (4–6 weeks)*	Fast (3–5 weeks)	Fast (3–6 weeks)**	Slow (1–3 month)	Fast (4–6 weeks)
Material Used	~20 mg	0.5–1 mg	0.5–1 mg	~0.2 mg	Transient transfection	20–100 µg antibody	Transient transfection
Resolution	Atomic	Peptide	Residue	3 AA	Residue	Peptide /Residue	Domain /Residue
PTM tolerance	Very low	Moderate-High	High	High	Native PTMs	Absent	Native PTMs
Cost	$$$	$$	$$	$$	$	$	$
Throughput	Single	Several	Several	Several	Several	Multiple	Multiple
Access to technology	Specialist	Broad	Specialty CRO	Broad/CRO	Broad	Broad/ CRO	Broad
Paratope Mapping	Yes	Low res	Yes	Yes	No	No	No
Linear or conformational Epitope	Conformational	Conformational	Conformational	Conformational	Conformational	Linear	Conformational
Structural or functional Epitope	Structural	Structural	Structural	Structural	Functional	Functional	Functional
Notes	No guaranteed output	Measures "protection" upon binding	Measures "protection" upon binding. 12+ out of 20 AA can be targeted.	Only crosslinks 6 out of 20 AA	Mutations can compromise antigen folding and function	Linear epitope mostly	Structure should be available

*Subject to synchrotron availability.
**Longer for soluble antigen.

Table 1 | Summary of current antibody-guided epitope mapping methods. HDX, hydrogen/deuterium exchange mass spectrometry; HRF, hydroxyl radical footprinting mass spectrometry; XL: cross-linking mass spectrometry; ALN, alanine-scanning mutagenesis; PepArr, peptide array; DomX, domain exchange. Figure taken from Dang, X., Guelen, L., Lutje Hulsik, D., Ermakov, G., Hsieh, E. J., Kreijtz, J., … van Eenennaam, H. (2023). *Epitope mapping of monoclonal antibodies: a comprehensive comparison of different technologies.* ***mAbs***, 15(1) under open access CC BY license.

2.2 Novel strategies against bacterial infections

2.2.1 Current vaccine landscape for GAS

In 1928, the Scottish physician Alexander Fleming observed a contaminating fungal colony producing a substance capable of lysing *Staphylococcus* cells; this discovery, later termed penicillin (Fleming, 1929), heralded the advent of antibiotic therapy. Despite decades of clinical use, Group A *Streptococcus* (GAS) remains largely sensitive to penicillin-based antibiotics, although reduced susceptibility has been reported in recent years [53,54].

Developing a safe and broadly efficacious vaccine against GAS has proved highly challenging, owing to the pathogen's extensive genetic diversity, the risk of autoimmune complications, and the practical difficulties in assessing immunity against a strictly human-adapted organism with diverse clinical manifestations [55]. Regulatory and commercial obstacles, notably a 25-year ban in the United States (imposed by the Food and Drug Administration, FDA) on the administration of GAS-derived products to humans [56], further impeded progress. Although this ban was lifted in 2005, only five vaccine candidates have subsequently progressed to early phases of clinical trials[57] .

All present clinical candidates target the GAS M protein, either by formulating multiple N-terminal fragments or focusing on conserved C-repeat epitopes. One such approach is StreptAnova, a 30-valent N-terminal peptide vaccine that demonstrated robust immunogenicity in a recent phase I trial [58], alongside evidence of cross-opsonic activity suggesting wider serotype coverage [59]. Nonetheless, it may still offer insufficient protection in certain geographic areas with high disease incidence [60]. By contrast, strategies aiming at the conserved C-repeat region, exemplified by MJ8VAX which induces antibodies against J8 (a 12-residue peptide from the C-repeat domain embedded in a non-M-protein, helix-forming sequence), have shown promise for more global protection [61]. MJ8CombiVax incorporates J8 together with an epitope from SpyCEP, conferring enhanced protection against hypervirulent CovR/S mutant strains [62]. Other C-repeat-focused vaccines, including StreptInCor and P*17, have displayed encouraging efficacy in preclinical models [63], while extensive safety assessments have revealed no signs of autoimmunity [64].

In addition to M protein-based strategies, numerous "non-M" vaccine formulations combine various antigens with both high gene carriage and minimal sequence variability [65]. Notable examples are GlaxoSmithKline's three-component vaccine (comprising SLO, *Streptococcus pyogenes* Adhesion and Division protein [SpyAD], and SpyCEP) [34], Vaxcyte's VAX-A1 (encompassing ScpA, SLO, SpyAD-Group A carbohydrate polyrhamnose backbone) [66], and multi-antigen blends including Combo4 [67], Combo5 [68,69], 5CP [70], and Spy7 [71], all of which have

protected animals against GAS challenge. Another promising candidate, TeeVax, targets various T antigens [72].

Despite considerable efforts in the development of diverse GAS vaccines encompassing both M protein-based constructs to non-M antigen formulations, the immunological mechanisms that underlie their protective efficacy are still not fully resolved. Recent investigations into IgG subclass responses underscore their importance: for instance, a cyclic lipopeptide J8-derived vaccine stimulated robust IgG1 and IgG3 production, key for opsonisation and bacterial clearance, alongside a balanced Th1/Th2 immune profile that enhanced antibody functionality [73]. Sublingual and intranasal immunisation with appropriate adjuvants has also been shown to induce strong systemic IgG, including subclasses particularly associated with protective immunity [74]. At the same time, T-helper cell polarisation studies reveal that conserved M protein epitopes can drive robust Th1 and Th17 responses, conferring broad-spectrum protection [75]. Cytokine profiling further emphasises the complexity of these immune responses. Cyclic GAS vaccines appear to elicit both Th1- (IFN-γ) and Th2- (IL-4) associated cytokines, correlating with durable IgG-mediated protection [76], while mucosal GAS vaccination enhances IFN-γ, IL-17, and IL-4, reflecting a combination of Th1, Th17, and Th2 immunity [74]. Integrated approaches that measure opsonic antibody activity and T cell responses indicate that epitopes in conserved M protein regions can promote cross-reactive, protective immunity [75].

Nevertheless, critical uncertainties remain about the epitope-level correlates of protection, particularly for complex GAS antigens. It is often unclear precisely which epitopes vaccine-induced antibodies target, how PTMs might influence IgG function, or whether epitope recognition differs among individuals. Addressing these gaps in our understanding of antibody-antigen interactions will be essential for optimising vaccine design and fully harnessing the power of IgG-mediated immunity against GAS.

Recent renewed focus on GAS vaccine research follows the World Health Organisation's (WHO) designation of rheumatic heart disease as a global priority, coupled with newly published economic analyses predicting substantial cost savings from an effective and safe GAS vaccine [77]. Advances in formulation, such as using adjuvants that stimulate both humoral and cellular immunity [63,69], and novel delivery systems like microarray patches [78] are under active exploration. Additionally, refinements in animal modelling [68,79] and the establishment of a human challenge model [80] offer more definitive insights into correlates of protection, thus holding potential to accelerate clinical translation.

2.2.2 Antibody-guided vaccinology

Epitope-focused vaccine strategies involve engineering immunogens that reliably replicate a protective epitope while minimising extraneous antigenic features, thereby guiding the immune system towards a recognised site of vulnerability [81]. A recent investigation showcases this concept by demonstrating that compact, thermally stable protein scaffolds, designed to mimic a protective epitope of respiratory syncytial virus [81], elicit neutralising antibodies in immunised macaques. This outcome verifies the principle that maintaining a high-fidelity epitope structure is critical for inducing a functional (*i.e.* protective) antibody response, thus providing a promising framework for the development of next-generation vaccines aimed at pathogens with conserved, yet immunologically recessive, epitopes.

From an immunobiological viewpoint, the capacity of a meticulously stabilised epitope to elicit potent neutralising antibodies underscores the significance of epitope structural rigidity for promoting high-affinity B-cell responses [82,83]. In numerous viral infection studies, especially those involving highly variable pathogens like HIV or influenza, broad neutralisation hinges on recognising conserved epitopes, which are frequently obscured by more variable, immunodominant segments [84,85]. By concentrating the immune response on these less common but essential targets, scaffold-based immunogens shift the typical immunodominance profile [86,87]. This approach may also mitigate the risk of disease enhancement by restricting exposure to "off-target" viral fragments that could trigger non-neutralising or otherwise harmful immune responses [88].

Such insights prove central for rational vaccine design. The correlation between scaffold-induced epitope mimicry and the generation of high-affinity, protective antibodies supports computational design strategies that incorporate backbone flexibility and precise epitope presentation [81]. This further highlights the potential for iterative improvements to scaffolds (*e.g.*, through modifications to adjuvants, dosage, or delivery methods) to optimise neutralisation titres and the durability of immune protection. In addition, incorporating multiple epitopes, each scaffolded to conserve vital functional elements, could expand the breadth of protection and address variability across different host B-cell repertoires.

Notably, outcomes from the same scaffold-based immunogens sometimes differ between mice and macaques, indicating that species-dependent differences in B-cell repertoire may affect the magnitude and frequency of protective responses [81]. This observation emphasises the importance of multi-epitope formulations and extensive preclinical testing. Ultimately, the confirmation that monoclonal antibodies isolated from an immunised macaque preserved precise specificity for the scaffold epitope provides strong evidence that epitope-focused vaccines can successfully "immuno-focus" antibody production on conserved regions [81]. Such an approach represents a powerful tactic for confronting difficult pathogens, particularly in scenarios where

conventional whole-antigen vaccines do not generate broad or long-lasting immunity.

2.2.3 Protein design and emerging vaccine modalities

De novo protein design

ProteinMPNN is a deep-learning-based approach for de novo protein sequence design that predicts the most probable amino acid residue at each position in a given backbone, drawing upon an extensive range of structures deposited in the Protein Data Bank. Unlike physics-based software such as Rosetta, which relies on large-scale sidechain packing calculations and typically demands substantial parameter tuning, ProteinMPNN operates with minimal adjustments. This capability allows it to excel in recovering native amino acid sequences across diverse backbone scaffolds, displaying notably higher sequence recovery (up to 52.4%) compared with conventional approaches (approximately 32.9%). Moreover, it often rescues designs that have failed in Rosetta or AlphaFold for protein monomers, assemblies, and interfaces [89,90].

Unlike physically driven algorithms that exclusively optimise for low-energy conformations, ProteinMPNN circumvents challenges such as unintended oligomer formation or ambiguous side-chain placement in boundary regions. Using a deep-learning model trained directly to predict amino acid residues from three-dimensional structures, this can address instances where minor residue substitutions might otherwise disrupt protein folding. By virtue of the high success rate in experimental validation, computational efficiency, and adaptability, ProteinMPNN has become instrumental in designing novel proteins [91] and in refining existing ones (*e.g.*, enhancing stability or crystallisability). These improvements accelerate structure determination efforts and open new avenues for generalised, robust protein engineering.

Nanomaterial as an emerging vaccine modality

Nanoparticles have concurrently emerged as a promising platform in vaccine development, enhancing immune responses through adjuvant-like effects, carrier functions, or antigen-display properties [92]. Among the various formats available, protein-based nanoparticles stand out for their biocompatibility, genetic flexibility, and built-in capacity to display multiple antigens in an organised, repetitive manner [92,93]. Unlike inorganic or micellar systems, which may pose concerns regarding toxicity, degradation, or restricted antigen loading, protein-based nanoparticles can be uniformly assembled and tailored for a broad spectrum of pathogens [94,95].

A notable advantage of protein-based nanoparticles is their capacity to support antigen uptake by antigen-presenting cells (APCs) and to provide high-density antigen array display that drives robust B-cell activation [92,96]. By enlarging the

antigen construct, typically to the recommended 20-50 nm range, and incorporating numerous antigen copies onto the nanoparticle surface, these platforms improve retention within lymphoid tissues, facilitate enhanced opsonisation, and promote efficient cross-linking of B-cell receptors [97–99]. Studies indicate that higher-valency antigen display generally correlates with improved humoral responses, with multiple binding events per nanoparticle reinforcing signals to both APCs and B-cells [100,101].

Efforts to optimise nanoparticle size, morphology, and antigen density increasingly rely on computational protein design to produce fully synthetic, self-assembling architectures [101–103]. These designed nanoparticles may derive inspiration from natural scaffolds (*e.g.*, ferritin or lumazine synthase) but often adopt symmetrical patterns uncommon in nature. Icosahedral platforms, for instance, have received particular attention due to their potential to form large oligomeric assemblies that match or exceed viral capsids in size and antigenic capacity [104]. By leveraging icosahedral geometry, composing 20 trimeric or 12 pentameric vertices, researchers have created new protein building blocks that self-assemble into structures resembling dodecahedrons, rhombic triacontahedrons, or standard icosahedra [105,106].

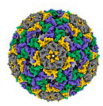
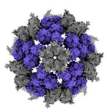
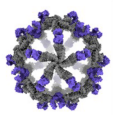
A key milestone of this approach is the use of two- or multi-component co-assembling strategies, wherein separate monomeric, dimeric, or trimeric modules are engineered to interact precisely at specified interfaces [107]. When assembled correctly, these subunits give rise to highly ordered nanoparticles containing cavities or channels capable of accommodating a variety of antigen conformations by genetic fusion. Researchers can adjust nanoparticle diameter, antigen valency, and epitope spacing by systematically altering the interface geometry [108]. This designed control allows for "epitope-focused strategy," in which immunodominant yet non-neutralising regions are effectively downplayed while critical protective epitopes are highlighted on the nanoparticle surface [81].

Importantly, vaccine efficacy hinges on both B-cell and T-cell responses. Some synthetic nanoparticle formulations incorporate internal T-cell epitopes which are universal $CD4^+$ or $CD8^+$ epitopes, for example, to bolster cellular immunity [109–111]. By simultaneously eliciting potent humoral and cell-mediated responses, protein nanoparticles may overcome the immune evasion mechanisms of persistent pathogens like HIV, dengue virus, and certain influenza variants [112]. Moreover, designing the optimum nanoparticle platform entails a fine balance of multiple parameters, particle size, surface charge, epitope spacing, antigen valency, and manufacturing feasibility, while preserving overall stability [97,113]. Although there are concerns that immunised individuals might generate antibodies against the nanoparticle scaffold components, several studies suggest that careful selection of building blocks and considerate epitope arrangement can help mitigate this issue [88].

In summary, protein-based icosahedral nanoparticles hold significant promise in modern vaccinology, combining safety with strong immunogenic potential by

replicating the repetitive surface structures of pathogens and providing extensive scope to integrate neutralising epitopes and T-cell determinants. Guided by further refinements in computational design, structural biology, and immunological methods, these platforms are poised to enable next-generation vaccines against a diverse range of infectious diseases.

Table 1. Selected Examples of Non-Viral Protein Nanoparticle Platforms[a]

Name	Model[62]	Global Symmetry	Number of Chains and Oligomers in Assembly	Diameter	PDB/ EMDB	Ref.
***H. pylori* ferritin**		Octahedral	24 monomers arranged as 12 dimers	12 nm	1FHA	63
***T. maritima* encapsulin**		Icosahedral	60 monomers arranged as 12 pentamers	24 nm	3DKT	64
***A. aeolicus* lumazine synthase (AaLS)**		Icosahedral	60 monomers arranged as 12 pentamers	15 nm	1HQK	65
Evolved lumazine synthase NC-4		Icosahedral	240 monomers arranged as 12 pseudopentamers and 20 trimers	30 nm	7A4J	56
Computationally designed two-component nanoparticle		Icosahedral	120 monomers arranged as 20 trimers and 12 pentamers	27 nm	6P6F	18,55
Computationally designed antibody nanoparticle		Icosahedral	36 monomers arranged as 12 *de novo* trimers and 12 Fc dimers	25 nm	In Ref.	21

[a]Nanoparticle models were made using UCSF ChimeraX.[62]

Table 2 | Some examples of established protein-based nanoparticle platforms. Figure taken from Audrey Olshefsky, Christian Richardson, Suzie H. Pun, and Neil P. King (2022). *Engineering Self-Assembling Protein Nanoparticles for Therapeutic Delivery*. ***Bioconjugate Chemistry***, 33 (11), 2018-2034 under open access CC-BY-NC-ND 4.0 license.

2.2.4 Streptolysin O: an important GAS vaccine target

SLO is a 69 kDa pore-forming exotoxin universally expressed by GAS [114]. Structurally, the mature secreted toxin contains multiple domains: the first 69 N-terminal residues form an unstructured region, while domains 1 to 3 are

discontinuous and culminate in a C-terminal domain 4 that is essential for cell membrane binding [115]. Like other cholesterol-dependent cytolysins (for example, pneumolysin [PLY] from *Streptococcus pneumoniae*), SLO undergoes major conformational rearrangements in secondary structures, notably within domain 3, to form pores and insert into host cell membranes.

Beyond its primary role in lysing host cells, SLO also contributes to bacterial immune evasion and intracellular survival. Consistent with this, SLO-deficient strains exhibit reduced virulence relative to wild-type GAS [116–118]. Likewise, immunisation with SLO, whether enzymatically inactivated or in its active form—confers protection in murine challenge studies [34,119], and nanoparticle-based treatments that neutralise SLO activity demonstrate therapeutic promise against invasive GAS infections [120]. SLO's cytotoxic effects target a broad range of cells, including erythrocytes, platelets, leukocytes, macrophages, and epithelial cells [116,121,122]. The mechanism involves successive steps: binding to the cell membrane, oligomerisation, and subsequent pore formation [123,124].

Notably, SLO promotes the entry of NAD^+ glycohydrolase into the host cytosol, thereby enhancing cytotoxicity [16,125]. NAD^+ glycohydrolase can induce cell death through NAD^+ depletion or programmed necrosis [126], and the stabilisation of both SLO and NAD^+ glycohydrolase upon interaction further broadens the range of targeted cells while enhancing GAS virulence [127–130]. Recent research also indicates that SLO cooperates with the streptococcal superantigen (SSA): SLO-mediated pore formation releases glutathione, which converts the inactive SSA dimer into an active monomer [32]. In addition to direct cell damage, SLO and NAD^+ glycohydrolase modulate host immunity by activating the inflammasome and inducing IL-1β secretion in some contexts [131,132], but inhibiting IL-1β, IL-8, and other cytokines via Golgi disruption in others [133]. These effects undermine neutrophil and macrophage function [134], while intracellularly SLO enables bacterial escape from endosomes and avoids autophagy-mediated destruction [135]. Overall, these multifaceted roles identify SLO as a prominent therapeutic and vaccine target in both preclinical and clinical research [55].

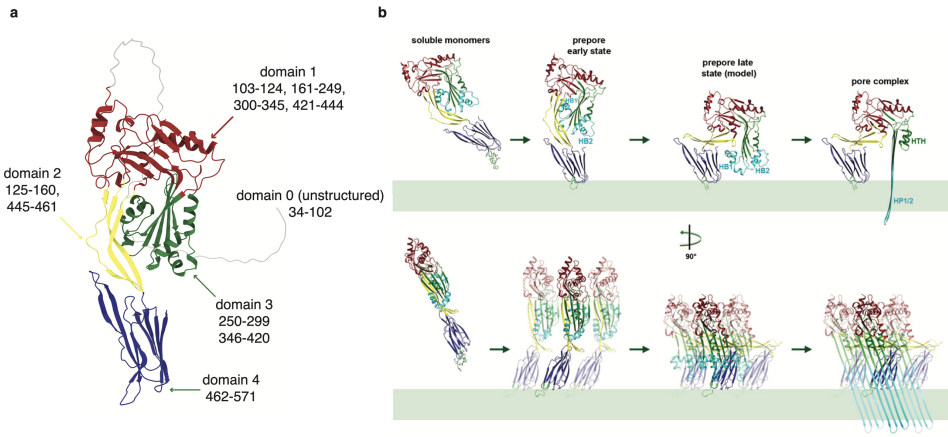


Figure 2 | (**a**) SLO domain annotation and residue numbering. (**b**) A typical CDC (using PLY as a showcase here) pore formation process. HB, helix bundle; HTH, helix-turn-helix; HP, β-hairpin. Figure taken from Katharina van PeeAlexander NeuhausEdoardo D'ImprimaDeryck J MillsWerner KühlbrandtÖzkan Yildiz (2017). *CryoEM structures of membrane pore and prepore complex reveal cytolytic mechanism of Pneumolysin*. ***eLife***, 6:e23644 under open access CC BY license.

2.3 Multimodal MS and integrated biology

2.3.1 Protein mass spectrometry

Mass spectrometry (MS) is a technique that measures the mass-to-charge ratio of gas-phase ions to simultaneously identify and quantify multiple molecules in both simple and complex samples. Since proteins constitute the majority (approximately 90%) of a cell's dry mass, they underpin the sophisticated chemistry that sustain life. A widely employed strategy in MS-based workflows is bottom-up proteomics, wherein proteins are enzymatically cleaved into shorter peptides, typically using trypsin, to facilitate analysis [136]. These peptides are then separated via liquid chromatography (LC), ionised, and subsequently subjected to MS analysis [137,138]. Tandem MS (MS/MS) generates peptide sequencing data by fragmentation, which can be compared against protein databases to identify the source proteins [139], or deduces peptides directly from experimental MS/MS spectra via a *de novo* sequencing manner [140]. Over recent decades, MS-based proteomics methodologies have evolved beyond merely global or in-depth protein identification and quantification, extending to the study of large-scale interactome, protein structures, functions, and solution dynamics[141,142] .

Although proteomics has yet to reach the same breadth of coverage as genomic methods, directly analysing proteins via MS excels at uncovering functional associations under near-native conditions on a large scale [143], particularly through affinity-enrichment/purification MS (AE/AP-MS) [144], cross-linking MS (XL-MS) [145], and co-fractionation MS [146]. These approaches also inform mechanistic follow-up studies, where structural proteomics including XL-MS and hydrogen/deuterium exchange MS (HDX-MS) coupled with integrated structural biology are invaluable [147]. In the following sections, I intend to highlight the main MS-based proteomics techniques (*i.e.* AE-MS, XL-MS, HDX-MS, and *de novo* MS) applied in this thesis.

Affinity-enrichment mass spectrometry

Pathogens often reprogramme and exploit host protein networks, where the extent of host-pathogen interplay is closely related to pathogen fitness during infection [148]. Unravelling protein-protein interactions (PPIs) therefore offers essential insights into disease pathogenesis and provides promising avenues for novel drug and vaccine development.

Proteins assemble into "protein machines" to accomplish the diverse biological functions required and to serve key structural roles. Generally, AE-MS procedure involves isolating specific biological targets from a complex mixture using a ligand linked to a solid support, commonly agarose or magnetic beads [149]. This can be achieved by inserting an epitope tag, such as the streptavidin-derived Strep-tag [150], which is recognised by the support, or through immobilising an antibody that

recognises the tag. When coupled with MS, the AE-MS approach becomes a cornerstone of functional proteomics, offering a high-throughput route to explore PPIs in diverse biological settings [151,152]. By genetically fusing a tag to a bait protein of interest, AE-MS helps maintain the native conformation and PTMs of both isolated interacting partners, enabling biologically relevant delineation of interaction networks. This method is highly valuable in areas ranging from fundamental research to translational applications, with notable significance in studying host-pathogen interactions [153–157].

Given the therapeutic potential of enhancing or inhibiting PPIs, it is frequently crucial to distinguish direct from indirect associations, those potentially mediated by intermediate proteins or other macromolecules, on the basis of AE-MS data [158–160]. Beyond the recombinant production of proteins for *in vitro* validation of biochemical interactions, pinpointing interacting surfaces that constitute the protein-protein interactions (PPIs) is similarly important. The polypeptide segments in close spatial proximity within protein assemblies isolated by AE can be determined using cross-linking MS (XL-MS) [161]. More importantly, combining these methods can help capture weak or transient PPIs that might be lost during the wash steps required in AE part. Integrative efforts that unite AE-MS and XL-MS are therefore under active investigation for advanced PPI studies [157].

Scalability and robustness render AE-MS particularly suitable for constructing interactome maps. For example, AE-MS experiments have revealed thousands of previously unidentified interactions in large-scale datasets, with software tools such as SAINTexpress, CompPASS, and MiST employed to differentiate genuine interactors from contaminants introduced during sample preparation [151,162,163]. Consequently, AE-MS has emerged as a central approach for elucidating the modularity and dynamism of the interactomes, including the characterisation of other types of macromolecular complexes (*e.g.*, protein-DNA, protein-RNA, protein-glycan) beyond solely protein interactions[164].

Cross-linking mass spectrometry

XL-MS has gained significant traction as a core methodology for investigating protein-protein interaction (PPI) networks, structural conformations, and the dynamic attributes of large biomolecular complexes [165]. By creating covalent linkages between residues in close proximity, XL-MS offers important distance constraints that elucidate the overall topology of protein assemblies under conditions close to their native states [165]. In contrast to conventional structural biophysics techniques, XL-MS can accommodate sample heterogeneity and requires comparatively small protein amounts, making the method particularly valuable for examining complexes in multiple conformational states. These features have placed XL-MS at the forefront of structural systems biology, where understanding protein behaviour in solution is vital for clarifying roles in both health and disease.

Typically, XL-MS begins with the application of chemical cross-linkers that selectively react with specific amino acid side chains, commonly lysine, or by employing photoactivated probes to map solvent-inaccessible regions [166]. Following cross-linking, the sample is proteolytically digested, generating a mixture of cross-linked and regular (non-cross-linked) peptides. Because cross-linked peptides often occur at sub-stoichiometric levels, specialised enrichment methods and refined liquid chromatography separation steps are used. Subsequent tandem mass spectrometry (MS/MS) detects cross-linked species by matching their distinct fragmentation signatures to the relevant peptide pairs. Such cross-link data confirm residue proximity and can be employed to model or refine protein structures, highlight binding interfaces, and unveil potential allosteric relationships, even within highly dynamic or transient assemblies.

A key benefit of XL-MS lies in the ability to capture diverse protein states, including partially disordered or ligand-induced conformations. By conducting cross-linking experiments under multiple conditions, researchers can compare cross-link patterns to identify conformational changes elicited by small molecules, binding partners, or PTMs [167]. Unlike techniques that yield only an averaged structure, XL-MS can detect multiple conformers coexisting in solution, thereby offering snapshots of the functional flexibility inherent to numerous protein complexes [168]. These findings can then be integrated with other structural data, such as cryo-electron microscopy (cryo-EM) or small-angle X-ray scattering, to generate a more comprehensive understanding of intricate protein architectures [169,170].

Within integrated structural biology, XL-MS has proved indispensable for refining models and verifying protein arrangements that cannot be fully captured by a single technique. XL-MS has contributed greatly to the characterisation of systems ranging from full-length proteins containing large intrinsically disordered regions to vast macromolecular machines, such as the COP9 signalosome and transcriptional complexes [171,172]. One recent illustration concerns the replication-transcription cofactors nsp7 and nsp8 in SARS-CoV-2, where XL-MS combined with hydrogen/deuterium exchange (HDX) showed that these proteins predominantly form linear hetero-tetramers in solution, rather than the cubic structures suggested by crystallography [173]. Likewise, XL-MS data have identified binding regions missing from existing models, showcasing the technique's ability to detect overlooked or elusive contact sites in extensive assemblies [174,175].

Additionally, XL-MS manages conformational diversity by capturing ensembles of structures rather than relying exclusively on static, crystallised forms [176]. This capability is crucial for understanding how protein subunits rearrange around regulatory domains or shift upon binding partners. The method has excelled at investigating molecular machines, large assemblies that undergo extensive conformational transitions as part of their functional cycles [177]. By pinpointing cross-links that change when a particular subunit reorients or a regulatory factor binds, XL-MS provides a molecular perspective on how complexes alternate

between active and inactive states [178]. As the technique advances further, aided by new cross-linking reagents, peptide enrichment strategies, improved instrumentation, and computational innovations, the capacity to map intricate protein architectures continues to expand. Through integrative approaches that merge diverse structural datasets, XL-MS remains a keystone for illuminating how proteins carry out their functions in both normal physiological and pathological contexts [165].

Hydrogen/deuterium exchange mass spectrometry

Researchers have long focused on clarifying how a linear amino acid sequence adopts a distinct three-dimensional conformation, how protein motion adapts during interactions, and how such processes adjust in both healthy and diseased contexts [179,180]. Although atomic-resolution snapshots of protein structures can be acquired using nuclear magnetic resonance (NMR) spectroscopy, X-ray crystallography, or cryo-EM, capturing protein dynamics in solution remains persistently difficult. Equally crucial is understanding how proteins engage with surrounding solvent and how hydrogen bonding, as inferred from the structure, factors into protein function [181]. Recent progress in both measuring and interpreting these hydrogen-exchange "signals" provides a unique opportunity to connect structural features with function and to guide protein engineering [182].

Hydrogen/deuterium exchange (HDX) refers to the phenomenon where, once a protein is diluted into deuterated buffer, certain hydrogens spontaneously substitute with deuterium in solution. In regions that lack defined folding, the exchange rate largely depends on buffer properties such as temperature, pH, and ionic strength, along with residue-specific acid dissociation constants. However, when a protein adopts a distinctive conformation, intramolecular hydrogen bonding and limited solvent contact reduce the rate of this exchange in a "protective" manner, mirroring the corresponding structural and dynamic traits [181]. Over the past two decades, HDX coupled with mass spectrometry (MS) has emerged as a versatile alternative to HDX-NMR [183], offering more automation, enhanced statistical robustness, and greater flexibility regarding protein size range and sample preparation strategies.

HDX-MS quantifies the mass shift of a protein or peptide attributable to deuterium incorporation. Typically, an ample dilution ratio in deuterated buffer is chosen, and the protein is exposed for varying lengths of time intervals. HDX patterns can then be evaluated globally (observing the intact protein) or at higher spatial resolution by using bottom-up proteomics, where the deuterated protein is digested, and the mass shifts of the resulting peptides are measured. This spatial specificity may be improved further by employing innovative fragmentation techniques, alternative proteases, or dual-enzyme protocols [184].

The analysis of HDX-MS data typically compares at least two scenarios (for example, a protein in different buffers or a protein in the presence versus the absence

of a binding partner). For each peptide, statistical methods (*e.g.*, Student's t-test [185] or linear regression models [186,187]) assess whether observed differences in deuterium uptake are significant. Any notable variations often point to structural changes such as protein unfolding, binding-site alterations, or allosteric effects.

HDX-MS has shown considerable promise for investigating antigen-antibody systems and mapping the binding interfaces. Even antigens of moderate purity can be used, and recent work has illustrated success in identifying epitopes recognised by multiple antibody types, accentuating the method's uniqueness, throughput, and broad applicability [188–190]. Furthermore, this technique excels at detecting conformational epitopes, which are nearly impossible to pinpoint through standard B-cell epitope mapping strategies (*e.g.*, peptide array, domain exchange, or scanning mutagenesis) [51].

De novo MS sequencing

Most MS-focused strategies employ a bottom-up liquid chromatography–tandem mass spectrometry (LC-MS/MS) pipeline [191]. Multiple proteases are used to produce overlapping peptides that enhance sequence coverage, while software platforms either reconstruct peptide reads *de novo* or refine them against germline-based models to identify somatic mutations [192]. Principal applications include rescuing 'lost' hybridomas and investigating polyclonal mixtures or entire serum repertoires [193,194].

Practically, MS-based antibody sequencing needs only modest sample quantities and can be achieved rapidly. For instance, combining multiple proteases with dual-fragmentation MS/MS approaches, specifically electron-transfer high-energy collision dissociation and stepped higher-energy collision dissociation, yields thorough coverage of variable domains, including CDRs Such methods have been successfully applied to widely used monoclonal antibodies, confirming both their sequences and their capacity to bind target antigens by reverse expression of sequenced mAb [195–197]. Overall, *de novo* MS sequencing represents a robust and complementary strategy for profiling antibody diversity, expediting therapeutic antibody advances, and driving fundamental immunological research. This holds particularly true when only specific functional antibodies (*e.g.*, those produced and purified from hybridoma techniques) are available, permitting relatively high-throughput and precise characterisation.

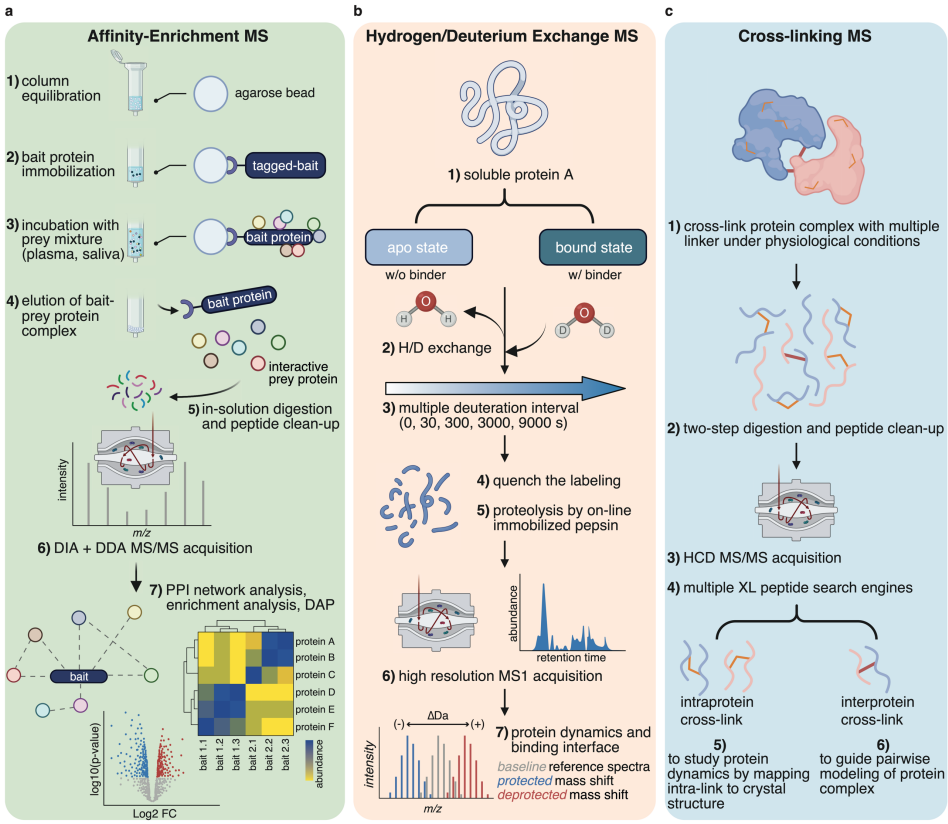


Figure 3 | Representative workflows for (**a**) AE/AP-MS, (**b**) HDX-MS, and (**c**) XL-MS. Figure taken from ***Paper I***, licensed under CC BY 4.0**.**

2.3.2 Data-driven modelling

Data-driven modelling has emerged as a central approach in structural bioinformatics, driving notable progress in predicting protein structures, including those antigen-antibody complexes. The introduction of deep learning tools such as AlphaFold2, RoseTTAFold, and related programmes has enabled high-accuracy prediction of individual protein architectures directly from sequence data input [198–202]. These developments have addressed critical shortcomings in protein structure determination, particularly for static or monomeric proteins, where experimental techniques may prove either unfeasible or prohibitively resource-intensive [201].

On the other hand, modelling protein complexes, particularly antigen-antibody interactions, continues to present a considerably greater challenge. While individual

antigens and antibodies can frequently be accurately modelled separately, the complexity of antigen-antibody assemblies involves numerous potential variations and uncertainties, particularly in anticipating interaction sites and conformational rearrangements upon binding [203,204]. This inherent complexity frequently results in notable disparities between predicted and actual binding interfaces and orientations, underscoring the limitations of exclusively computational strategies for modelling highly intricate biomolecular complexes with scarce co-evolutional features [204].

One key shortcoming of current data-driven modelling is an over-reliance on computational predictions without sufficient incorporation of empirical data or constraints. Although AI-based tools can produce credible structural folding, they frequently do not capture the dynamic and context-sensitive qualities of protein-protein interactions, which are influenced by factors such as binding-induced structural rearrangements, PTMs, and steric hindrances [205]. For instance, antibody CDRs, notably the 3rd loop of heavy chain CDR crucial for antigen recognition, display considerable structural variability, complicating accurate modelling [206,207].

Introducing experimental constraints, such as loosely defined binding interface information or structural insights from given complexes, can substantially bolster both the accuracy and consistency of complex modelling efforts. As an example, information-guided docking methods such as HADDOCK and AlphaLink, which utilise experimental findings to improve and constrain computational predictions, have shown superior outcomes compared with purely *ab initio* modelling methods [208,209]. In the absence of such constraints, attempts at modelling antigen-antibody complexes often produce inconsistent and biologically irrelevant outcomes, especially when both antigen and antibody models are generated computationally [203,210].

Therefore, although data-driven modelling has achieved remarkable success in the prediction of primary protein structures, the integration of experimental validation with computational methods is fundamental to accurately capture pairwise/multi-component and antigen-antibody protein complexes in both structural and functional terms.

2.3.3 Single particle cryo-EM

Cryo-EM is a transformative technique for determining three-dimensional structures of large macromolecular complexes at near-atomic resolution [211,212]. The workflow includes rapid vitrification of samples, imaging in cryogenic conditions, and computational three-dimensional (3D) reconstruction, thereby reducing radiation damage while preserving the molecular integrity of the specimen. Cryo-EM is particularly suited to highly symmetric assemblies, as single-particle averaging capitalises on multiple, repeating subunits to boost the signal-to-noise ratio [211].

However, mismatches in perfect symmetry, such as sub-stoichiometric occupancy of receptors or antibodies, may produce "symmetry mismatches", causing the loss of coherent signal if a global symmetry is incorrectly enforced [213]. Notable cases include trimeric fibres on icosahedral viruses or internal chaperonin-substrate complexes misaligned with the broader cage structure [214,215]. Historically, these discrepancies have hampered the achievable resolution in deviating regions, as they become averaged out or are incorrectly aligned [216]. To address these issues, localised reconstruction methodologies have been developed, treating these mismatched or flexible subunits as independent entities for classification and refinement. By centring on these specific subunits rather than the overall particle, this approach is now possible to obtain higher-resolution detail while retaining contextual information about the complex [217]. This approach has substantially improved reconstructions, revealing partial occupancies and local variations in binding angles [213].

One crucial factor in elucidating the mechanisms of potential biotherapeutics is accurate epitope mapping for mAbs. Recently, negative-stain electron microscopy (EM) and cryo-EM have been implemented for both linear and conformational epitope identification. These methods require only microgram-scale quantities of the antibody-antigen complex and can be integrated with other analytical strategies, like HDX-MS or fast photochemical oxidation of proteins (FPOP), to determine epitopes at the residue level. Long and colleagues, for example, used cryo-EM to clarify how two potent, chikungunya virus-neutralising human mAbs bind virus-like particles and block fusion between the viral and host membranes [218]. Likewise, negative-stain EM has proven valuable in initial exploratory studies; for instance, Ciferri and colleagues determined the structure of an antibody complex with HTRA1, a serine protease implicated in age-related macular degeneration, highlighting the cage-like arrangement and inhibition mechanism[219].

Cryo-EM has additionally delivered key insights for vaccine design, especially in relation to dynamic and intricate antigens like the HIV-1 envelope glycoprotein gp160 trimer (Env). As the principal target for developing broadly neutralising antibodies (bnAbs), Env is a focal point of HIV vaccine research. Lee and colleagues revealed how Env rearranges the conformation upon binding MPER (membrane-proximal external region)-directed antibodies, emphasising the antigen's flexible nature and underscoring challenges in vaccine formulation [220]. Moreover, recent efforts have deciphered the initial quaternary contact between Env and the CD4 receptor, providing vital structural context to guide rational vaccine and antibody engineering [221].

In summary, cryo-EM, augmented by single particle analysis, localised reconstruction techniques, and used in tandem with other complementary analytical approaches, enables exceptional insights into functional heterogeneity in large biomolecular assemblies. These advances have direct relevance for vaccine development, therapeutic design, and antibody engineering [222].

2.3.4 Integrated structural biology

Structural insights of biomacromolecules provide the fundamental basis for mechanistically understanding the full spectrum of transformations occurring in organisms across all domains of life. Integrated structural biology merges and interprets data from multiple complementary orthogonal techniques (*e.g.*, X-ray crystallography, small-angle X-ray scattering, NMR, cryo-EM, mass spectrometry, molecular dynamics, and artificial intelligence [AI]-based methods), each grounded in diverse physical and chemical principles [223]. This integrated approach addresses the major challenge of correlating structural information with biological function and is essential for elucidating the roles of macromolecules and their assembled complexes in living systems.

The success of AI-driven protein structure prediction and generative tools has prompted a renewed perspective on how integrated structural biology should evolve. Adding another dimension as temporal scale to the classic 3-D information gives rise to the term "4-D structural biology," which seeks to capture how biomolecules adapt to changing environments or biological processes, thereby offering deeper insights into function and regulation [223]. In particular, cryo-EM with structural MS techniques, particularly XL-MS, has demonstrated success in uncovering the architecture of multimeric complexes [169,224,225]. Nonetheless, gaps persist in structural biology field, notably the incomplete representation of PTMs and RNA structures, posing significant hurdles to achieving a more comprehensive structural atlas of biomolecular function [223].

3 Aim of the thesis

3.1 Problem statement

Despite decades of intensive research on GAS pathogenesis and host immunity, no licensed vaccine targeting this bacterial pathogen of concern is currently available. This shortfall persists despite mounting evidence that virulence factors beyond the highly variable M protein, most notably SLO, could function as promising vaccine antigens. SLO is a multifunctional CDC that damages host cells through pore formation while also modulating immune defences. Yet its interactions with various host factors remain incompletely characterised, and definitive mapping of protective SLO epitopes has not been achieved. Current efforts to incorporate SLO into vaccine strategies have yielded only limited immunogenicity and protection, likely due to the dominance of non-protective regions overshadowing functionally critical epitopes.

Two main gaps warrant attention:

1. **SLO pathomechanisms and host-pathogen interaction.**

   While the pore-forming function of SLO is well documented, the broader landscape of SLO's PPI interactions with human proteome, particularly under native conditions such as in plasma or saliva, remains poorly explored. Understanding these potential PPI of relevance is crucial for deciphering how GAS exploits the host protein networks and pathways to promote bacterial invasiveness.

2. **Epitope-specific immunity and the need for integrated MS approaches.**

   Precisely which SLO regions (*i.e.* epitopes) elicit neutralising or protective antibody responses remains unclear, and identifying these regions on native antigen represents a significant analytical challenge. Recent advances in functional and structural proteomics (including AE-MS, HDX-MS, XL-MS, and *de novo* MS sequencing) provide powerful tools for mapping protein interaction interfaces and discovering epitopes of importance. However, their combined usage as a multimodal MS workflow is still underexploited. Leveraging these complementary orthogonal techniques in sequence could possibly clarify the structural basis of SLO-mediated

adaptive immunity, laying the groundwork for next-generation, epitope-focused vaccine designs.

By applying this integrated mass spectrometry framework to SLO, one of GAS's major virulence determinants, it becomes possible to systematically study host-pathogen interactions and pinpoint the precise structural regions that underpin protective antibody responses. Such insights will not only advance our fundamental understanding of GAS pathogenesis but also facilitate rational vaccine engineering. In particular, epitope-focused immunogen design can be optimised to present the most functionally relevant regions of SLO, potentially through innovative vaccine modality or conjugate platforms that display these epitopes in a manner conducive to robust, long-lasting immune protection. Ultimately, bridging these biological and technological gaps is crucial for realising a safe and effective vaccine against GAS.

3.2 Aim

The overarching aim of this thesis is to develop a multimodal protein MS strategy that unifies functional and structural proteomics to achieve an increased understanding of SLO, thereby contributing to next-generation vaccine initiatives against severe GAS infections. This goal comprises five interlinked objectives:

1. **Map host-pathogen interactome centred on SLO**

   To fully profile and interpret protein-protein interaction networks between SLO and human plasma proteins. By leveraging functional proteomics (*e.g.*, AE-MS) and network analysis, we will identify host factors that bind SLO and investigate how these interactions contribute to GAS virulence.

2. **Determine the biological relevance of SLO-interactive partners**

   To analyse the physiological/pathological impacts of SLO's interactions with key human proteins and elucidate the molecular bases of these interactions using HDX-MS, XL-MS, and computational modelling. These findings will clarify unrecognised mechanistic pathways related to bacterial dissemination and immune modulation exerted by SLO.

3. **Develop a structural antibody-guided epitope-mapping workflow**

   To establish a pipeline that integrates *de novo* MS sequencing (for determining the antibody primary structures), HDX-MS (for probing protein dynamics in solution), and XL-MS (for identifying proximal distance constraints within antigen-antibody complexes), in combination

with data-driven modelling. This toolkit aims to confidently identify bacterial epitopes of protective significance, particularly on SLO, in their native three-dimensional conformations.

4. **Design, engineer, and evaluate an epitope-focused immunogen**

   To design a novel immunogen incorporating well defined protective SLO epitopes, and present such on an emerging vaccine modality inspired from self-assembling symmetric protein-based nanoparticles. We will assess whether this epitope-focused construct can elicit a more targeted immune response, offering enhanced and safer protection profile in animal models compared to community-standard SLO-derived vaccines.

5. **Generalise the multimodal protein MS workflow**

   To broaden the applicability of the previously established multimodal protein MS approach by investigating more diverse monoclonal antibodies (mAbs) and mapping epitopes in other homologous cytolysins (*e.g.*, PLY). We will thus evaluate the feasibility of extending these methods to other clinically significant bacterial antigens, enabling high-throughput accurate characterisation of structural epitope landscapes, demonstrating uniqueness and advantages of MS-based epitope mapping methods over conventional methods, clarifying how epitopes relate to mAb-directed functionalities, and updating rational immunogen design beyond streptococcal infections.

By pursuing these five objectives, the thesis aims to exemplify the versatilities of multimodal protein MS, extending from initial host-pathogen interaction mapping, protein interface studies, to rational immunogen engineering. These will also demonstrate how such integrated approaches can deepen our understanding of key immunogens and facilitate optimised vaccine strategies against GAS. Ultimately, these efforts may provide a blueprint for vaccine design targeting other difficult pathogens of public health concerns.

4 Present investigations

4.1 Overview

The four papers presented here collectively demonstrate an integrative MS proteomics strategy for deciphering and combating bacterial pathogens, focusing on GAS and *Streptococcus pneumoniae*. ***Paper I*** reveals a previously unrecognised multifunctional role of SLO in binding human plasminogen (PLG), thereby enhancing GAS-acquired plasmin activity to facilitate bacterial tissue invasion and immune evasion [226]. ***Paper II*** introduces a multimodal protein MS workflow, encompassing *de novo* antibody sequencing, XL-MS, HDX-MS and modelling, that identifies a protective conformational epitope of SLO in three-dimensional conformation, enabling an informative antibody-guided "reverse vaccinology" for severe bacterial infections [227].

Building upon these findings, ***Paper III*** describes epitope-focused vaccine design through the *de novo* protein design of a protective SLO epitope (D3m) and the multivalent display on a protein nanoparticle scaffold. The resulting vaccine construct (D3m-NP) induces potent immunogenicity and confers strong protection in a murine challenge model. ***Paper IV*** applies the updated multimodal protein MS framework to a collection of monoclonal antibodies raised from immunisation of. PLY, a homologous cytolysin produced by *S. pneumoniae*. This study successfully delineates the relationship between epitope recognition and functional outcomes, also identifies a cross-reactive conserved epitope of protection shared among other CDCs.

In addition to these four main papers, I have also contributed to completed projects on M protein-specific antibody research [42,228] and multienzyme *de novo* MS sequencing methodology [229]. Readers seeking broader context may refer to these publications for further information.

4.2 Paper I: Streptolysin O as a new plasminogen-interacting protein

4.2.1 Background

Overview of GAS

GAS is a major human pathogen responsible for a broad spectrum of diseases, from superficial infections such as pharyngitis and impetigo to life-threatening conditions including necrotising fasciitis and streptococcal toxic shock syndrome [2]. Globally, GAS causes more than 700 million non-invasive infections and approximately 1.78 million severe cases each year, leading to over 500,000 deaths [3]. This prevailing disease burden arises not only from direct outcomes of invasive infection but also from post-streptococcal autoimmune sequelae, including rheumatic fever and acute glomerulonephritis [3].

Plasminogen as a host pathogenic factor in GAS pathogenesis

A central element of host sensing and defence against bacterial invasion is the coagulation-fibrinolysis axis, wherein protective fibrin immobilise and entrap pathogens [230]. GAS interferes with this system by exploiting plasminogen (PLG), a zymogen essential for fibrinolysis. PLG predominantly circulates in a "closed" conformation, thereby resisting unnatural activation. Upon binding to fibrin or cellular receptors, PLG partially relaxes, exposing an activation loop recognised by tissue plasminogen activator (tPA) or other plasminogen activators (*e.g.*, SKA [231]). Subsequent site-cleavage converts PLG into the active protease plasmin, which degrades fibrin networks and facilitates bacterial dissemination [232]. Multiple GAS factors, such as PAM (*i.e.* plasminogen-binding group A streptococcal M-like protein), α-enolase, glyceraldehyde 3-phosphate dehydrogenase, extracellular protein factor, and SKA, can recruit PLG to the bacterial surface, thereby enhancing plasmin activity and facilitating tissue invasion [233].

SLO as a multifunctional cytolysin

SLO is a major GAS virulence factor and part of the CDC superfamily. The classical biological function involves forming pores in cholesterol-rich host cellular membranes, yet SLO also exerts immunomodulatory effects, helps evade phagocytosis [234], modulates host inflammatory reactions [235], and augments the activity of certain streptococcal superantigens [32].

4.2.2 Results

PLG-SLO interaction determined by AE-MS

AE-MS was applied to identify plasma proteins that bind recombinant affinity-tag SLO, compared with two controls: the C5α peptidase and green fluorescent protein. Across multiple plasma dilutions and statistical analyses, PLG consistently stood out as the most specific and significant SLO-enriched plasma binder. This interaction persisted despite the presence of endogenous anti-SLO immunoglobulins contained in pooled plasma, suggesting a robust association between SLO and PLG. Other SLO-interacting host factors included proteins involved in complement regulation, lipid transport, and coagulation modulation, highlighting the multifunctional strategies SLO employs to adapt within the host microenvironments.

Enhanced plasminogen activation by host and bacterial activators

A key finding of this study is that although SLO does not directly cleave PLG, SLO binding markedly increases the efficiency of PLG activation by both host tPA and the bacterial activator SKA. Chromogenic assays revealed a dose-dependent accelerated rate in PA catalysing PLM production when PLG was bound to SLO, compared to free PLG alone (**Figure 4a**). Once PLG had been converted to PLM, further addition of SLO did not affect the proteolytic activity of PLM, indicating that SLO's role is confined to facilitating the initial conversion process. Importantly, SLO retained the pore-forming capacity in parallel, suggesting that haemolysis and PLG binding are mutually compatible functions.

Structural insights into PLG-SLO complexes: HDX-MS and XL-MS

To investigate how SLO binding promotes PLG activation, HDX-MS and XL-MS were used in combination to obtain complementary orthogonal data:

1. HDX-MS revealed protected regions in PLG (notably within kringle 2 [K2], peptidase S1 [PSD], and portions of the pan-apple N-terminus domain) upon complex formation with SLO, implying either direct binding-induced interfaces or global stabilisation effects. Although no relaxation pattern equivalent to fibrin binding was observed, the data suggest a local conformational shift that renders PLG more susceptible to activation exerted by PAs.
2. XL-MS identified specific lysine-lysine cross-links between domains 1, 3, and 4 of SLO and the K2, PSD, and PAN domains of PLG. Over-length intra-protein cross-links within SLO further pointed to a dynamic or "bent" conformation that accommodates PLG binding, corroborated by computational simulation that places SLO domain 4 in an inward orientation.

Computational docking models of PLG-SLO complex

Integrative modelling integrating XL-MS data generated a plausible architecture of the PLG-SLO complex. The highest-confidence models placed a bent conformer of SLO in contact with the K2 and PSD domains of PLG, supporting experimental HDX-MS data that implicate these regions involvement in complex formation. Independent experimental evidence further suggests the assemblies of transient hetero-trimeric complexes comprising either host or bacterial PLG activator that can cleave PLG more efficiently when PLG is bound to SLO.

SLO-centred interaction networks and implications

Functional enrichment analyses of other plasma proteins associated with SLO (as derived from AE-MS) revealed associated pathways and networks enclosing immunoglobulins, complement factors, and coagulation-related proteins. These results underscore SLO's multifaceted strategy for bacterial adaption and survival within the host: in addition to triggering cell lysis, SLO can engage or counteract various plasma components, eventually producing a microenvironment conducive to GAS colonisation and dissemination.

4.2.3 Conclusion

A novel moonlighting function of SLO

This research demonstrates for the first time that SLO, well known primarily for pore formation and immunomodulation, undertakes an additional critical biological function: binding to plasminogen and facilitating PLG activation by both endogenous and exogenous activators. Instead of directly catalysing PLG (as SKA does), SLO stabilises an intermediate conformation of PLG, rendering SLO-bound PLG molecule more susceptible to activator-mediated cleavage. This cooperative process likely reinforces GAS's capacity to penetrate host tissues by dismantling protective fibrin barriers and circumventing immune defences by targeting non-fibrin substrates (**Figure 4b**).

Pathomechanism and therapeutic potentials

The integrative network and structural analyses (AE-MS, HDX-MS, XL-MS, and computational modelling) converge on a mechanism wherein a dynamic SLO conformation principally interacts with the K2 and PSD domains of PLG. SLO retains the haemolytic function while effectively priming PLG for enhanced activation, exemplifying the shared or synergistic roles of GAS virulence factors. Therapeutically, disrupting the PLG-SLO interface may provide a novel strategy to damper GAS invasiveness by reducing added plasmin activity without compromising the host physiological fibrinolytic processes. Given that SLO is highly conserved and expressed among nearly all *S. pyogenes* strains, such a targeted intervention may hold broad therapeutic value.

Readers seeking additional context detail can refer to the ***Paper I*** [226] attached.

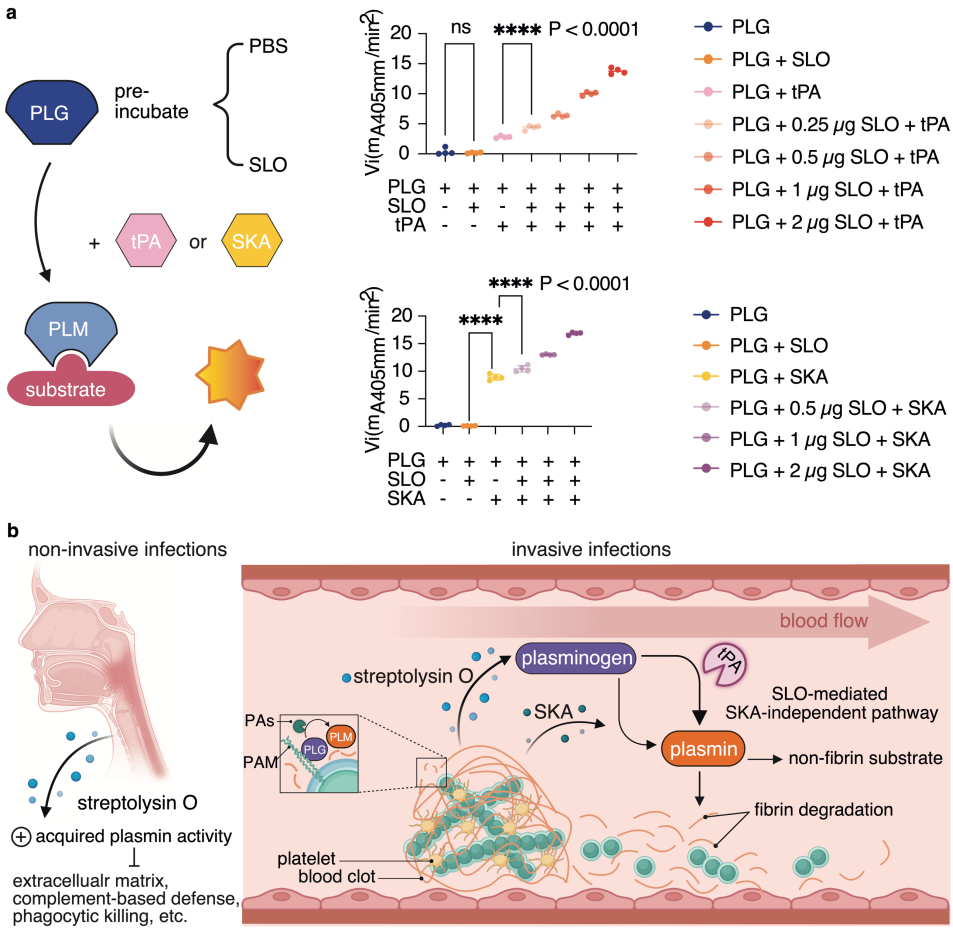


Figure 4 | (**a**) SLO enhances plasminogen (PLG) activation mediated by both tPA and streptokinase (SKA). (**b**) Proposed *in vivo* mechanism of SLO-mediated acquisition of plasmin (PLM) activity in GAS pathogenesis. Figure taken from ***Paper I***, licensed under CC BY 4.0.

4.3 Paper II: Conformational epitope mapping via multimodal MS

4.3.1 Background

Antibody-guided vaccine design

Antibody-guided vaccine design capitalises on knowledge gained from protective antibodies to develop immunogens capable of eliciting similarly potent immune responses *in vivo*. Such "reverse vaccinology" approaches have proven highly beneficial for viral pathogens such as influenza and HIV [236], where structural insights into neutralising antibodies help guide epitope selection and antigen engineering.

Structural MS for epitope mapping

Although traditional structural techniques (*e.g.*, X-ray crystallography, cryo-EM) remain the gold standard for elucidating antigen-antibody complexes, they can be constrained by sample requirements or molecular size limits. In contrast, structural protein MS methods, particularly XL-MS and HDX-MS, provide valuable and unique experimental data information when a smaller antigen or heterogenous antibody samples are indicated:

1. HDX-MS monitors change in deuterium uptake in antigen-derived peptides to identify epitope regions shielded from solvent exchange upon antibody binding.
2. XL-MS employs chemical cross-linkers to connect proximate residues at the antibody-antigen binding interface, generating informative distance constraints that guide reliable computational modelling.

Used in tandem, these techniques yield complementary information with modest sample quantities and experimental demands.

Streptolysin O (SLO) and the vaccine potential

GAS continuously causes significant morbidity and mortality around the globe, largely attributed to the potent, pore-forming toxin SLO, expressed in nearly all sequenced clinical GAS isolates, is highly conserved and immunologically important:

1. Current SLO-derived vaccine landscape: Existing multicomponent vaccine candidates (*e.g.*, Combo4, Combo5, VAX-A1, 5CP) incorporate a detoxified SLO construct (non-lytic variants). However, a lack of well-defined neutralising mAbs has greatly hindered epitope-focused vaccine research and development.

2. A more rational approach in need: Isolating a SLO-specific, functional mAb and determining the respective epitope could support next-generation GAS vaccine design. This study introduces a multimodal MS workflow, including *de novo* antibody sequencing, epitope mapping via XL-MS, pairwise modelling, and HDX-MS, and validation of the epitope's conservation level, to showcase the antibody-guided vaccinology against SLO.

4.3.2 Results

Discovery and characterisation of a neutralising mAb

Initially, available mAbs were screened for SLO binding and neutralisation of SLO-mediated cytolysis. One murine antibody effectively inhibited SLO-induced sheep erythrocyte lysis. *De novo* MS sequencing (involving multi-enzyme proteolytic digestion) confirmed this mAb's primary structure, including the CDRs. A predicted model of the corresponding fragment antigen-binding (Fab) domain served as the basis for subsequent epitope mapping.

Multimodal MS approach reveals key binding interfaces

A dual structural protein MS approach, combining XL-MS and HDX-MS, was applied to the investigate native SLO-mAb immune complex:

1. XL-MS and data-driven pairwise modelling
    - Cross-linkers of various spacer arm lengths captured in-solution proximity constraints between SLO and the Fab heavy/light chains.
    - Analysis of the generated cross-link distance constraint inferred the possible interaction interface on SLO to a region primarily located in domain 3.
    - Computational tools (DisVis [237] and HADDOCK [205]) further refined these data, excluding distance-violating structural models and highlighting SLO surface residues implicated in Fab targeting.
2. HDX-MS
    - Comparisons of deuterium uptake between the complex state (mAb-bound) and apo (unbound) state of SLO identified significantly decreased solvent exchange in a domain 3 region.
    - This protection pattern independently corresponded to the XL-MS and docking derived interface (**Figure 5a**) and also revealed that SLO's overall conformation remained largely unchanged. Thus, neutralisation

effect appears to result from direct interference with critical pore-formation steps.

Epitope validation via construct reengineering

To confirm that the identified region alone was sufficient for antibody binding, a smaller construct (d3_m), encompassing the relevant conformational epitope, was redesigned and expressed recombinantly. The mAb bound this construct with an affinity comparable to that of full-length SLO, reinforcing the correctness of the multimodal protein MS approach.

Conservation and plausible protective mechanism

Bioinformatic analysis of over 2,000 GAS genomes revealed that the identified mAb-directed epitope of SLO is nearly 100% conserved, emphasising the broad clinical relevance and efficacy potential. As domain 3 of SLO undergoes substantial secondary structure rearrangements during pore formation, mAb binding in this particular region likely impedes oligomerisation or sequential membrane insertion steps (**Figure 5b**). Thus, targeting domain 3 could possibly provide a robust, widely effective immune defence against severe GAS infections.

4.3.3 Conclusion

This work exemplifies how a multimodal protein MS approach encompassing *de novo* MS sequencing, XL-MS for interface analysis, HDX-MS for epitope mapping, and computational modelling, can accurately pinpoint conformational epitopes of protection on bacterial antigens in near-native conditions. The identified neutralising mAb recognises a conserved region located in domain 3 of SLO, affording important insights for future antibody-guided, epitope-focused vaccine design against GAS infections. By elucidating the relevant epitope involvement in SLO's pore-forming mechanism, the study provides both a conceptual and methodological framework for refining current subunit GAS vaccine strategies and highlights how integrated proteomics coupled with latest computational biology can accelerate epitope-targeting immunogen refinement against challenging pathogens (**Figure 5c**).

Readers seeking additional contextual information should refer to ***Paper II*** [227] attached.

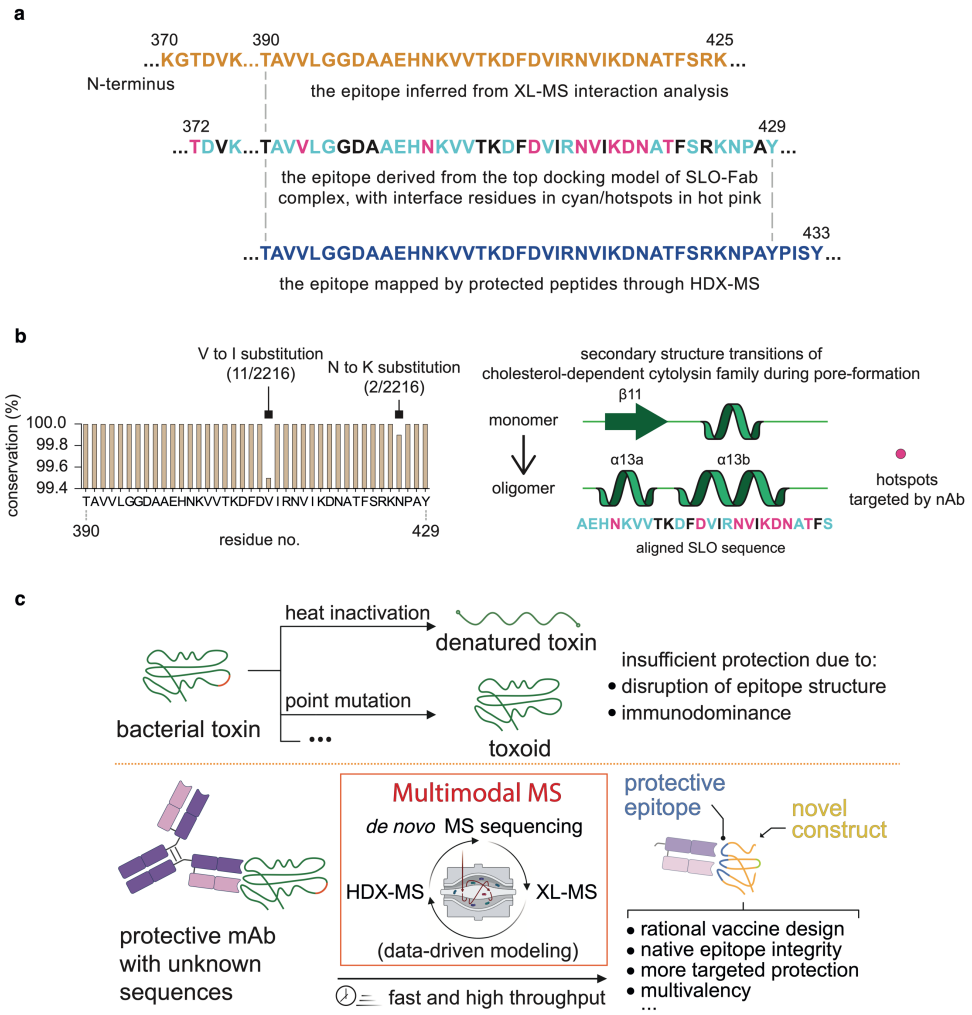


Figure 5 | (**a**) Epitope alignment and comparison of the three protein MS-based epitope-mapping methods. (**b**) Conservation level of the protective epitope and proposed mechanism of mAb neutralisation on SLO haemolysis. (**c**) Multimodal protein MS workflow advances next-generation vaccine design against challenging bacterial infections. Figure taken from ***Paper II***, licensed under CC BY 4.0.

4.4 Paper III: D3m-NP, an epitope-focused nanoparticle vaccine

4.4.1 Background

Current GAS vaccine development

GAS is responsible for a diverse range of human infections, from superficial lesions to life-threatening conditions such as sepsis. Despite extensive research efforts, a licenced vaccine remains unavailable. Traditional strategies emphasising the highly variable M protein have met with issues of limited serotype coverage and possible autoimmune complications. Consequently, more conserved antigens, particularly SLO, a CDC exotoxin expressed by GAS, are increasingly of interest. Existing SLO-based vaccine concepts include full-length toxoids and mRNA-LNP formulations, although such approaches risk eliciting antibodies targeting irrelevant immunodominant segments of the SLO antigen. Furthermore, variant constructs that have been modified or fragmented in an *ad hoc* manner might lose important conformational epitopes essential for protective immunity.

Novel immunogen design and emerging vaccine modalities

Mapping functionally relevant epitopes in bacterial antigens often presents considerable challenges, particularly because conventional methods such as alanine scanning or peptide arrays may fail to define discontinuous binding interfaces (*i.e.* conformational epitope) dependent on the protein's native three-dimensional structure. However, recent advance in structural proteomics (*e.g.*, HDX-MS, XL-MS) and computational prediction and design (*e.g.*, AlphaFold, ProteinMPNN) now allow for the accurate identification and engineering of "epitope-focused" constructs that retain these critical regions in near-native conformations. Building on these advances, new vaccine platforms, especially self-assembling protein nanomaterials, provide an opportunity for displaying surface antigen array with precise control over presentation, orientation, valency, and stability. By organising multiple copies of one or more fine-tuned epitopes in a spatially defined manner, such nanoparticle immunogens greatly enhance B-cell receptor cross-linking and immune responses within germinal centres. Collectively, these emerging strategies minimise immunodominant yet non-protective elements while selectively promoting effective immune responses against functionally important epitopes, thus laying a robust foundation for novel epitope-targeting vaccines against pathogens like GAS based on acquired knowledge regarding the virulence factor SLO.

4.4.2 Results

Design of the D3m nanoparticle immunogen

We began by creating a *de novo* refined version of SLO domain 3 (D3m) using a previously identified protective epitope as the starting point. In the native form, the primary sequence of SLO domain 3 is interrupted by a subsequence from domain 1, producing a discontinuous domain. Employing ProteinMPNN and AlphaFold, we replaced this segmented region with a short four-residue linker, generating a new construct aimed at recapitulating the native conformation of domain 3 located onto the whole SLO antigen (**Figure 6a**).

To validate structural fidelity, we then used HDX-MS to compare deuterium uptake between D3m and the domain 3 region in full-length SLO. Across multiple labelling intervals, D3m closely mirrored the solvent accessibility profile of native domain 3, suggesting that the folding and key structural motifs remained intact. Although minor differences were observed near the N- and C-termini, the data indicated that D3m closely replicates the intended domain.

Multivalent presentation on I53_dn5 nanoparticle

Following the successful redesign of domain 3, we sought to enhance immunogenicity by displaying D3m on a self-assembling protein nanoparticle scaffold. The two-component I53_dn5 system comprises I53_dn5A pentamers and I53_dn5B trimers, which together assemble into a 120-subunit icosahedral nanoparticle. By genetically fusing D3m to I53_dn5B (forming a D3m-I53_dn5B chimera), we enabled multivalent presentation of the immunogen on the nanoparticle surface.

Size-exclusion chromatography coupled with multi-angle light scattering (SEC-MALS) produced a prominent peak consistent with the expected molecular weight of the assembled construct, without signs of protein aggregation. Native-PAGE and negative-stain electron microscopy also confirmed the assembly and integrity of both the I53_dn5 nanoparticle (NP) and the D3m-displaying variant (D3m-NP). In cryo-EM single particle analysis, the I53_dn5 core resolved to a global resolution of approximately 4.06 Å; however, the flexible linker connecting D3m to the trimeric component resulted in lower-resolution densities for the displayed immunogen (**Figure 6b**). Nevertheless, up to 60 copies of D3m appeared consistently loaded on the NP surface, highlighting successful assembly of targeted D3m nanoparticle immunogen.

HDX-MS of monomeric, fused trimeric, and nanoparticle states of D3m

We next employed HDX-MS to examine how fusion and nanoparticle configurations influence D3m's overall stability and protein dynamics in solution. Three states were included and compared: (i) monomeric D3m, (ii) the D3m-I53_dn5B fused trimer, and (iii) the fully assembled D3m-NP. Almost complete peptide coverage enabled a robust and detailed analysis of deuterium uptake pattern across the entire immunogen.

Overall, monomeric D3m showed the smallest degree of deuterium exchange, implying a relatively compact fold in solution. The nanoparticle-displayed state displayed slightly higher solvent exposure in certain regions, while the trimeric construct showed the largest deuterium uptake. Compared with the monomeric state, the most deprotected segment in D3m-NP was at the C-terminus, where D3m was connected to I53_dn5B through a short linker. A hybrid significance test revealed that β11 and α13 were the motifs most affected in the nanoparticle-displayed state versus the monomer. Nonetheless, these alterations did not disrupt scaffolded D3m's overall structural integrity, indicating that, despite modest changes in solvent accessibility or dynamics, the presented epitope largely retained the intended conformation.

Mouse immunisation with D3m-NP

To assess and compare the immunogenicity of these SLO-derived constructs (*i.e.* D3m and D3m-NP), we injected groups of nine BALB/c mice with either one of three immunogens: (i) monomeric D3m, (ii) detoxified full-length SLO (dSLO, state-of-the-art construct), or (iii) D3m-NP, each formulated separately in AddaVax adjuvant. Equivalent 2 μg of domain 3 was administered to the both D3m and D3m-NP immunisation groups, whereas 10 μg of dSLO served as a community-standard comparison. Each group received three doses over two-week intervals, with five mice designated for immunogenicity studies (cohort A) and four mice reserved for a GAS infection challenge trial (cohort B).

Sera collected from immunised mice were tested for antibody titres against both D3m and full-length SLO. All constructs gave rise to domain 3-specific antibodies, meanwhile D3m-NP induced significantly higher anti-D3 titres than the other vaccines, reflecting the advantage of multivalent nanoparticle in stimulating B-cell responses (**Figure 6c**). While dSLO also elicited strong humoral immunity, likely directed against multiple regions of the toxoid, some of which may be immunodominant but not necessarily protective.

In a haemolysis inhibition assay, sera from all vaccine regimens demonstrated measurable neutralisation of SLO-induced lysis, albeit to varying degrees. The strong anti-D3 IgG induced by D3m-NP did not necessarily surpass dSLO-induced IgG in haemolysis neutralisation. These finding hints that full-length toxoids may include additional neutralising epitopes or that certain D3m-NP-induced antibodies hold alternative immune functions beyond mere cytolysis inhibition.

Protection against lethal GAS challenge

The *in vivo* protective efficacy of these vaccine constructs was further evaluated using a lethal challenge pilot model with the virulent GAS AP1 strain. Three mice per group (cohort B) received subcutaneous inoculations of a lethal dose and were monitored for 48 hours. Body weight changes and murine sepsis scores (MSS) served as key clinical indicators. Mice immunised with D3m-NP showed improved weight loss and lower MSS, suggesting that disease severity was mitigated.

Bacterial loads revealed a notable reduction in local bacterial counts for the D3m-NP group, though disease burdens in distal organs were similar across the three immunised groups. Flow cytometry showed that D3m-NP-treated mice had elevated white blood cell and platelet counts, suggesting strong yet balanced immune activation. Concurrently, lower levels of pro-inflammatory markers (*e.g.*, IL-6, lactate dehydrogenase [LDH]) and reduced CD11b expression on neutrophils and monocytes pointed to a more controlled inflammatory response, correlating with improved clinical manifestations in D3m-NP group (**Figure 6c**).

Overall, D3m-NP immunisation offered improved protection against lethal GAS infection, outperforming the detoxified SLO construct in immune activation while retaining a favourable safety immunity profile. By focusing the immune response on key antigenic determinants in domain 3, the nanoparticle-based vaccine appears to curb the excessive inflammatory activity often associated with bacterial infection while still able to promote effective immunity to mitigate GAS infection.

4.4.3 Conclusion

This study demonstrates, for the first time, the potential of an epitope-focused strategy for GAS vaccine development by targeting a conserved protective region within the streptolysin O (SLO) domain 3. Using integrative protein mass spectrometry and deep-learning-based structural tools, we designed a targeted immunogen (D3m) that preserves native-like structural motifs and protective epitopes, subsequently displaying these immunogens on an icosahedral self-assembling nanoparticle scaffold.

Biophysical characterisation, including SEC-MALS, native-PAGE, negative-staining EM, cryo-EM, and HDX-MS, collectively confirmed the correct assembly and structural fidelity of scaffolded D3m immunogen. Although the flexible linker limited D3m's resolution in cryo-EM reconstruction, indirect measurements still verified consistent, homogeneous loading of 60 copies of the immunogen per particle. HDX-MS comparisons across monomeric, fusion-trimeric, and nanoparticle-displayed states indicated that D3m's essential conformation remained intact, despite minor variations in solvent accessibility near the fusion junction.

Critically, immunisation studies in mice showed that D3m-NP elicited significantly highest domain 3-specific antibody titres among all and conferred clearly improved protection relative to detoxified, full-length SLO in a lethal GAS challenge model. The D3m-NP construct stimulated robust immune activation while minimising inflammatory damage, reflecting a favourable safety immunity profile. Taken together, these findings underscore the considerable potential of rationally engineered, epitope-focused vaccines to address longstanding hurdles in bacterial vaccine research. By preserving pivotal structural epitopes and delivering them via emerging modalities, such as multivalent protein-based nanoparticles, we can more

effectively boost targeted immune responses and limit immunodominance hierarchy stemming from non-functional epitopes.

Moreover, this work not only exemplifies a promising avenue for next-generation GAS vaccines but also highlights how cross-disciplinary techniques, from advanced structural protein MS to computational protein design and nanotechnology, can be harnessed to tackle widespread concerning pathogens equipped with sophisticated virulence determinants. Future refinements could involve *de novo* modifications to further stabilise domain 3's native-like conformation and the incorporation of additional protective epitopes of other antigens to broaden the scope of functional immunity. Nonetheless, the concepts illustrated here are widely applicable and may accelerate the development of safe and effective vaccines against severe GAS infection.

Readers interested in more detailed context can refer to the corresponding manuscript ***Paper III*** as attached.

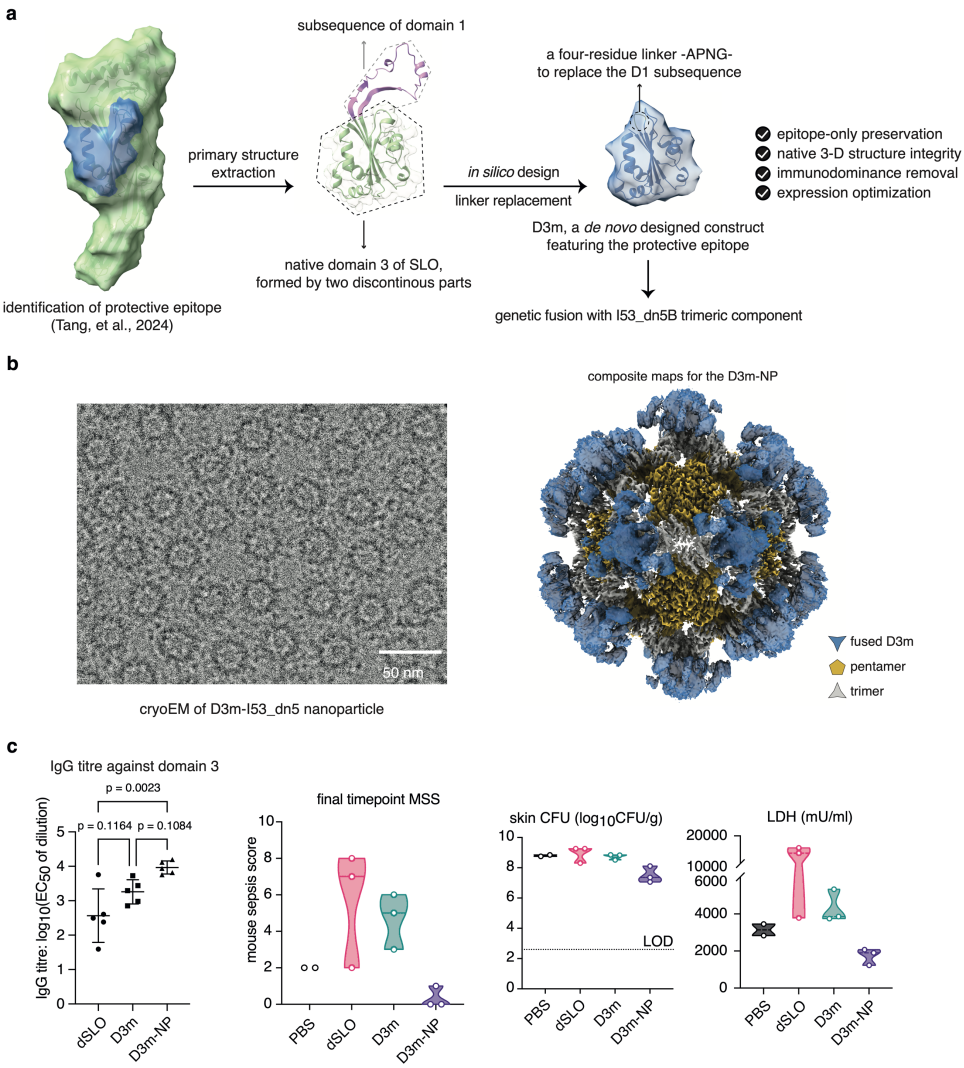


Figure 6 | (**a**) Antibody-guided epitope mapping, extraction, redesign through structural protein MS and *de novo* protein design. (**b**) Cryo-EM micrograph and reconstruction of D3m-NP immunogen. (**c**) Immunisation and challenge studies with three different SLO-derived vaccine constructs. Figure taken from ***Paper III*** (unpublished manuscript), © by the Authors.

4.5 Paper IV: Structural epitope landscape of pneumolysin

4.5.1 Background

Streptococcus pneumoniae and current pneumococcal vaccines

S. pneumoniae remains a prominent Gram-positive pathogen that contributes substantially to global morbidity and mortality, particularly among young children, older adults, and immunocompromised individuals. Although capsular polysaccharide (CPS)-based vaccines have greatly reduced the incidence of pneumococcal diseases, these vaccines are constrained by their serotype coverage and suboptimal immunogenicity. Meanwhile, non-vaccine serotypes have risen in prevalence, and in parallel, antimicrobial resistance has intensified, underscoring the need for alternative or complementary vaccination strategies.

A more broadly protective approach centres on *S. pneumoniae* protein antigens. PLY, a CDC produced by nearly all pneumococcal strains, is highly conserved and plays multiple roles during infection, ranging from pore formation on host cell membranes to immunomodulation, highlighting the biological relevance in pneumococcal pathogenesis. Therefore, PLY is gaining attention as a promising therapeutic and vaccine target. However, conventional vaccine designs typically rely on full-length, detoxified antigens. Such methods may modify native epitopes and accentuate immunodominant yet non-protective sites, potentially diverting the immune response from effective epitopes.

Antibody-guided vaccine strategies address these issues by isolating monoclonal antibodies (mAbs) with protective or neutralising activity, then pinpointing the precise epitopes underpinning those effects. These defined epitopes can serve as the starting point for subunit vaccines or novel immunogens that preferentially generate protective immune responses *in vivo*. This approach has proven invaluable in viral infection fields, where in-depth characterisation of antibody-antigen complexes has steered the design of innovative immunogens.

Pneumolysin: structure and function

PLY is composed of four major domains (D1 to D4), which collectively mediate the pneumococcal ability to bind host cell membranes, form pores, and induce cell death. Domain 4 (D4), located at the C-terminus, is critical for membrane binding and interaction with non-cholesterol cellular receptors. The structural intricacies of PLY, alongside the conformational shifts PLY undergoes during pore formation, render epitope mapping particularly difficult. Traditional techniques like peptide array and alanine-scanning mutagenesis often fall short in identifying discontinuous or conformational epitopes, similar to other multidomain bacterial antigens.

Advances in structural protein MS now enable detailed investigation of native antigen-antibody complexes under near-physiological conditions. In particular, three complementary MS-based methods, *de novo* MS sequencing, XL-MS, and HDX-MS, offer orthogonal insights into binding modes and protein dynamics. *De novo* MS sequencing elucidates the primary structures of unknown mAbs, XL-MS uses chemical cross-linkers to reveal residue-level proximities within antigen-antibody complexes, and HDX-MS highlights induced protection or deprotection pattern upon antibody binding. Integrating these techniques not only clarifies the specific binding sites of protective mAbs but also illuminates how antibody engagement influences antigen conformation, ultimately linking distinct epitopes to varied biological outcomes.

4.5.2 Results

Monoclonal antibody binding kinetics and neutralisation potency

A panel of ten mAbs raised against full-length PLY exhibited varying degrees of protection against PLY-induced cytolysis. Although all ten bound strongly to PLY suggested by indirect ELISA assays, not all conferred neutralising protection. Specifically, seven mAbs effectively reduced or abolished haemolysis, whereas three offered almost no protection at all (**Figure 7a**).

Moreover, strong binding affinity did not necessarily translate into potent neutralisation, indicating that the site of epitope engagement, rather than overall affinity, was central to blocking PLY's primary cytolytic mechanism. Even among the neutralising mAbs, neutralisation potency differed considerably, suggesting that multiple binding modes and epitope diversities exist on the PLY antigen.

Clonotypes and B-cell lineage reconstruction

To better understand these diverse binding profiles, we determined the primary structures of all ten mAbs by *de novo* sequencing their heavy (H) and light (L) chains. Comparative analyses revealed substantial somatic hypermutation, with varying levels of mutational "depth" in the CDRs. Notably, some clones sharing the same V-J gene usage nonetheless harboured distinct point mutations in their CDRs, implying that even minor changes in key residues could alter how each antibody engages target antigen PLY. In other words, multiple B-cell lineages might converge on overlapping regions of PLY yet yield different functional outcomes.

Lineage tree reconstruction analysis indicated that the heavy-chain CDRs, in particular, were highly diversified, aligning with their typical prominence in recognising antigen. Concurrently, certain mAbs displayed evident "family" relationships, hinting that they originated from an identical germline precursor but diverged into unique maturation trajectories culminating in distinct protective potencies. This variation underscores how *S. pneumoniae* may leverage

immunodominance hierarchy to induce suboptimal or non-protective antibody responses; nevertheless, this result also reveals that rare but highly effective neutralising clones can still emerge. By exploring these lineage relationships, we gain insights into how specific combinations of somatic mutations correlate with functional changes in epitope recognition, ultimately shaping the spectrum of protective efficacy.

Structural epitope landscape

XL-MS of each mAb-PLY complex mapped cross-linked sites and revealed that domain 4, the region responsible for binding to cellular and membrane insertion, was frequently targeted by neutralising clones. Some mAbs also interacted with other adjacent domains, reflecting PLY's conformational flexibility and domain movements. Nonetheless, domain 4 consistently emerged as a "hot zone" for functionally important interactions (**Figure 7b**).

Multiple cross-links involved residues near motifs previously implicated in membrane binding. Among neutralising clones, distinct cross-link patterns indicated that small differences in epitope recognition could yield substantial variations in protective outcome. Some mAbs preferentially cross-linked to a specific lysine residue in domain 4, whereas others displayed a broader cross-link distribution, possibly reflecting secondary Fab interactions or transient binding events across neighbouring domains.

To complement these data, HDX-MS revealed solvent accessibility patterns upon antibody binding. Potent neutralisers consistently protected domain 4 residues critical to PLY's cell membrane interactions. Other clones showed protection or deprotection patterns in domain 3, suggesting alternative protection mechanisms that may destabilise PLY or hinder oligomerisation. In each case, the region of protection correlated with neutralisation potency, illustrating how epitope targeting and engaging directly influences functional outcome.

Identification of a broadly conserved epitope

One intriguing protective mAb, clone 6E5, bound a highly conserved region in domain 4 that overlaps with the undecapeptide loop, an element known to be vital for membrane recognition in PLY and other CDCs. Further tests indicated that 6E5 recognised not only PLY but also several other CDCs secreted by diverse bacterial species, inhibiting corresponding haemolysis across multiple CDC families. This broad-spectrum reactivity underscores the existence of a structurally and evolutionarily conserved motif within CDCs, pointing to a promising avenue for developing therapeutic or vaccine strategies that offer cross-species protection (**Figure 7c**).

4.5.3 Conclusion

This study showcases the robustness and generalisability of the multimodal protein MS workflow, namely *de novo* MS sequencing, XL-MS, and HDX-MS, to elucidate the spectrum of antigenic determinants on PLY. By examining ten mAbs exhibiting varied neutralising potencies and biological functionalities, we demonstrate that high affinity alone does not necessarily translate into the desired functional outcomes.

Among the notable findings is the identification of domain 4 as a crucial hub for protective epitope engagement, underscoring the significance in pneumococcal pathogenesis. Equally noteworthy is the ability of the 6E5 clone to recognise a conserved loop in domain 4, a region that provides broad cross-species reactivity and protection against other CDCS. This discovery highlights the possibility of developing universal therapeutics or next-generation vaccines that exploit common structural motifs in virulence factors originating from diverse bacterial species background.

The methodologies and conceptual framework detailed here can also be adapted for interrogating other challenging bacterial or viral pathogens. Rather than relying on full-length antigens inevitably subjected to immunodominance, future vaccine or therapeutic antibody designs may focus on functionally critical epitopes to block or interrupt key disease mechanisms. In this manner, improved structural knowledge of epitope-paratope interactions offers a path to more targeted and safer strategies for combating infectious agents.

From a pneumococcal vaccine standpoint, our findings underscore how selective epitope targeting can strengthen protective responses, a fundamental consideration given the rise in antimicrobial resistance and the increasing prevalence of non-vaccine serotypes. By applying robust, epitope-focused designs to clinically significant virulence proteins, particularly those reliant on CDC mechanisms, researchers can pursue broader-spectrum interventions with superior efficacy.

Readers seeking further details are encouraged to consult the attached manuscript ***Paper IV***.

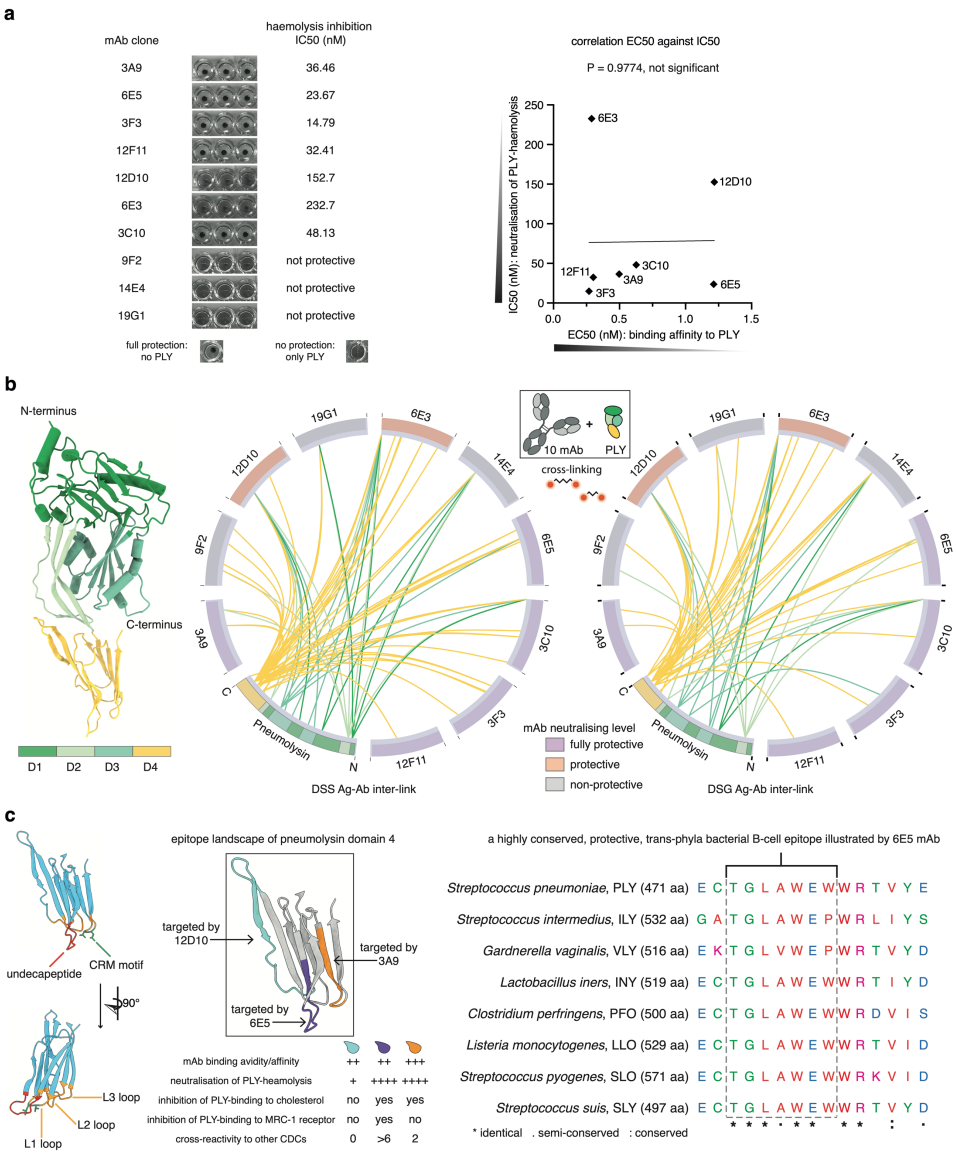


Figure 7 | (**a**) Antibody neutralising potency on PLY haemolysis and correlates to PLY-binding avidities. (**b**) XL-MS linkage maps between PLY and 10 mAbs, annotated by functional outcomes. (**c**) Structural epitope landscape of PLY and alignment of 6E5 clone-recognised epitope across different CDCs. Figure taken from ***Paper IV*** (unpublished manuscript), © by the Authors.

5 Discussion

The collective body of work presented in these four papers emphasises a unifying theme: how a multimodal mass spectrometry (MS) approach can unravel new dimensions of bacterial virulence and guide the rational design of optimised vaccine candidates. Focusing on two clinically important yet complex bacterial antigens as SLO and PLY, these studies combine orthogonal multimodal protein MS with integrated structural biology to systematically profile host-pathogen interactions, accurately identify immunologically relevant epitopes in 3-D space, and critically evaluate epitope-focused immunogens and deciphering functional correlates of B-cell immunity. This section will start with novel aspects/highlights of the present investigations, followed by a critical discussion highlighting the limitations of the present work.

5.1 Highlights

5.1.1 Uncovering multifunctional roles of streptolysin O

Historically, SLO has been viewed primarily as a CDC exotoxin that disrupts host cells primarily through pore formation. ***Paper I***, however, reveals a previously underappreciated function: SLO can also bind human PLG, thereby enhancing PLG conversion to plasmin by endogenous (*e.g.*, tPA) and exogenous activators (e.g., SKA). By stabilising an intermediate SLO-PLG conformation, SLO effectively renders PLG more susceptible for proteolytic cleavage, resulting in increased plasmin production. This moonlighting function of SLO provides new insight on how GAS navigates the host coagulation-fibrinolysis axis to promote bacterial spread and immune evasion. The fact that SLO is highly carried across GAS strains and highly conserved PLG-binding motifs underscores the broad relevance of this pathomechanism.

5.1.2 Multimodal MS in antibody-guided epitope-mapping

Paper II extends from exploring generic host-pathogen interactions to focusing defined immune complexes by mapping protective epitopes on SLO via a

multimodal protein MS workflow. By combining HDX-MS for surface solvent accessibility patterns, XL-MS for generating distance constraints, and *de novo* MS sequencing for monoclonal antibody (mAb) characterisation, the study shows how orthogonal techniques synergise to identify an unreported epitope critical for neutralising SLO-mediated haemolysis. In contrast to state-of-the-art structural epitope mapping techniques that demand large amounts of purified antigen-antibody complexes or strictly sized target molecules, the MS-based approach can operate with relatively small sample volumes, accommodate a broad range of molecular sizes, and analyse under near-physiological solution conditions. This methodological versatility enables unique interpretation and measurement into antibody-binding interfaces, including conformational or discontinuous epitopes that might be invisible to conventional tools like peptide array or Alaine-scanning mutagenesis.

5.1.3 Epitope-focused bacterial vaccine design

Building on epitope-targeting structural vaccinology, ***Paper III*** demonstrates the practical utilisation of these structural data by engineering an optimised immunogen (D3m) encompassing a crucial protective epitope of SLO. Rather than relying on full-length toxoids or uncharacterised fragments of SLO, this epitope-focused strategy is achieved with an emerging self-assembling nanoparticle presentation system (I53_dn5). We show that multivalent display of D3m on the nanoparticle surface greatly boosts antigen-specific antibody titres and confers improved protection in a pilot mouse model of invasive GAS infection challenge. Notably, while high antibody titres against domain 3 of SLO do not always equate to maximal neutralisation of haemolysis, the overall protective benefit stands out among all tested SLO-derived vaccine constructs. The study illustrates that structural fidelity of the relevant epitope, combined with enhanced immunogen presentation, can elicit a more targeted functional immune response, in this instance, a balanced yet protective inflammatory profile that effectively limits disease severity.

5.1.4 Expanding the paradigm to *S. pneumoniae*

Turning to another major pathogen of public health concern, ***Paper IV*** adapts the established multimodal MS pipeline to PLY, a significant cytolysin of *S. pneumoniae*. By examining a panel of monoclonal antibodies with diverse functionalities derived from PLY immunisation, the work illuminates how strong binding affinity alone, though benefitting from extensive affinity maturation process *in vivo*, is insufficient to guarantee robust protective activity. Instead, epitope recognition and engagement, particularly in domain 4 of PLY, appears crucial in interfering with membrane binding and other receptor recognition. Additionally, the discovery of a trans-species, highly conserved epitope targeted by a potent

neutralising mAb further highlights the potential for wide-spectrum protection targeting other CDC-family toxins. This suggests a blueprint for extending specific epitope-focused immunogen design to a wide range of microbial pathogens that rely on homologous virulence determinants and similar virulence mechanism.

5.1.5 Overarching Significance

Collectively, these investigations illustrate both the technical versatility and the necessity of combining complementary MS-based techniques. Functional proteomics (affinity-enrichment paired with network analysis) reveals unexpected roles of bacterial toxins in host systems, while structural proteomics (HDX-MS, XL-MS, and computational modelling) pinpoints discrete interfaces or epitopes critical to bacterial pathogenicity and host immunity. The resulting insights not only deepen our understanding of SLO's human-adapted virulence and diverse epitopes in CDCs but also highlight how rationally targeting well-defined structural epitopes can lead to more focused, and potentially more effective, vaccine development.

Moreover, these findings speak to a larger theme in molecular immunology and structural vaccinology: success often centres on identifying functionally relevant, conformationally correct epitopes rather than indiscriminately utilising entire proteins as antigen vaccine. By localising protective regions of true biological relevance within toxins such as SLO or PLY, it becomes feasible to engineer and refine vaccine constructs that elicit high titres of functional antibodies, while minimising the deviation of immune responses to non-protective or immunodominant decoy regions.

5.1.6 Bridging pathogenesis and translational research

An integrated takeaway is that investigating bacterial toxins at high spatial resolution can converge fundamental pathogenesis research with translational goals, including the design of therapeutic drugs and novel vaccine prototype development. The synergy of functional and structural proteomics fosters a twofold gain: (i) a systematic map of a host-pathogen PPI network that features critical molecular events, and (ii) a route to next-generation immunogens that are less prone to pitfalls like immunodominance hierarchy effect and suboptimal immunity.

In essence, these four papers jointly set an example for how the relevant research communities might further combine omics-level research, analytical advance in MS, computational modelling, and immunobiology to tackle with pressing challenges in bacterial infectious diseases. The framework provided herein, comprising identifying understudied host-pathogen association (***Paper I***), mapping structural epitopes (***Papers II*** and ***IV***), and translating these biological insights into epitope-focused vaccines (***Paper III***), demonstrates a sequential process of discovery,

validation, and application. It's a roadmap that can be extrapolated beyond SLO or PLY to other important streptococcal virulence factors, as well as to entirely different classes of complex pathogens.

5.2 Critical discussion

5.2.1 Limitations of a multimodal MS approach

While a multimodal protein MS strategy offers a powerful toolkit for delineating protein-protein interfaces and epitopes in general, several challenges remain. First, data interpretation from experiments like HDX-MS or XL-MS can be confounded by in-solution protein dynamics, such as flexible linkers often found in multi-domain molecule or transient conformations. Although the integration of computational docking can help reconcile conflicting data, the approach heavily relies on data quality and robustness of restraint parameters. An incomplete or noisy dataset (*e.g.*, low coverage in HDX-MS or low-confidence/conflicting cross-link identification) can skew the resulting models, leading to uncertainties in epitope delineation or protein complex architecture building.

Second, designated experiments prior to MS analysis pose another practical constraint. AE-MS approaches depend on the relative abundance of prey proteins in mixture and the absence of interfering factors. On the other hand, HDX-MS typically demands rather high protein purity and protein stability to preserve the structural integrity of investigated proteins/protein partners. These experimental conditions may not perfectly mimic native physiological condition *in vivo*, where concentration gradients, competition by other host factors, and local microenvironments can alter protein binding profiles.

Moreover, chemical reactivity and residue specificity of cross-linker can hamper or obscure the identification of critical interactions. If reactive sites are not suitably positioned within/close to the binding interface, XL-MS might fail to capture the relevant cross-links, leading to absence of informative distance constraints. Conversely, promiscuous cross-links outside the actual epitope can lead to false positives. Thus, our studies have managed to mitigate these obstacles through incorporating more options in each step from experiment (replicates, multiple linkers) to data analysis (multiple search engines) and interpretation (computational tools), but the approach remains technically nontrivial and must be carefully optimised for each protein system.

In the future, incorporating more established structural protein MS approaches (*e.g.*, protein footprinting MS, native MS, limited proteolysis MS) could enhance the

reliability and provide additional complementary data, ultimately leading to more definitive conclusions backed by solid experimental evidence.

5.2.2 Interpreting *in vivo* relevance

A common question arising from functional and structural proteomics is the extent to which *in vitro* findings translate to *in vivo* complexity. The demonstration in ***Papers II*** and ***III*** that rationally determined epitopes confer protection in a pilot experimental mouse model represents a step toward bridging this gap. However, immune responses in humans, with diverse genetic backgrounds and immune exposure histories, can exhibit unanticipated complexities: individuals may generate distinct sets of immunoglobulins, or the same epitope might engage different B-cell lineages with variable protective capacities. Additionally, the interplay between humoral and cellular immunity can differ profoundly in real-word scenarios from the tightly controlled animal models frequently used in preclinical trials.

In the case of SLO, for instance, the synergy between plasminogen manipulation and pore-forming activity (***Paper* I**) could play out differently in microenvironments where competing host proteins or other immune factors intervene. Similarly, the structural findings in PLY epitope mapping (***Paper IV***) might not encapsulate all constraints that arise in advanced pneumococcal infections, such as biofilm contexts or partial immunity originating from prior exposure. These considerations underscore the need for iterative validation of cautions in more sophisticated or more refined models that better approximate human infection physiology.

5.2.3 Balancing neutralising effect versus immune modulation

Another interpretative challenge is dissecting which immunological mechanisms are primarily responsible for protection, within the context ranging from a specified virulence factor to the whole pathogen. ***Papers II*** and ***III*** illustrate that high antibody titres directed at domain 3 of SLO do not consistently translate to stronger haemolysis inhibition *in vitro*. These findings suggest that some elements of protective immunity induced by domain 3 immunisation may not be entirely reliant on neutralisation. Instead, corresponding opsonisation, enhanced phagocytic clearance, or cooperative T-cell responses could play a critical role here. Conversely, sole neutralisation of pore formation might not suffice if the pathogen employs additional virulence factors that circumvent immune responses, or simply that SLO's other critical pathomechanisms (known or yet unexplored) are not affected. Thus, a purely epitope-centric perspective, while valuable, cannot fully capture the nuanced immunological interplay at the bacterial organism level.

Similarly, domain 4 in PLY proved a prevalent target for neutralising antibodies in ***Paper IV***, but other tested clones could recognise epitopes that may reduce toxin

oligomerisation or imply other protection mechanisms rather depriving PLY-binding to the membrane. Even potent neutralisers might fail to achieve full protection if they do not promote effective bacterial clearance. As the field leans more heavily on structural methods to rationalise immunogen design, it becomes critical to integrate more functional assays that capture the broader range of immune effectors beyond merely straightforward toxin neutralisation. It also underscores the need for deeper investigation into how antibody-mediated functions correlate with factors beyond specific paratope structure, such as antibody isotype, subclass, and Fc glycosylation.

5.2.4 Implications for vaccine composition

Results across these four papers highlight a tension between epitope-focused vs. full-length/fragments of antigen approaches, especially in bacterial infection field. Epitope-focused constructs (*e.g.*, the D3m domain in ***Paper III***) are attractive for precisely guiding the immune system toward functionally relevant targets, potentially minimising immunodominance of nonprotective sites. Yet, eliminating large segments of the antigen might exclude beneficial T-cell epitopes or synergy with other protective B-cell epitopes. Indeed, full-length toxoids are almost inevitably prone to generate an overabundance of non-functional antibodies, as an evolutionary host-adapted bacterial defence mechanism.

A nuanced resolution could involve combining carefully curated epitopes spanning either one or multiple virulence determinants, each carefully validated by structural and functional methods, into fine-tuned multicomponent vaccine formulations. This approach may recapitulate the coverage of an effective broad-spectrum subunit-based vaccine while preserving structural integrity for each critically relevant epitope. Nevertheless, the complexity and R&D costs for such next-generation vaccines can rise rapidly.

5.2.5 Technological and conceptual advance

Finally, the synergy between omics-driven analytics (*e.g.*, proteomics, immunopeptidomics) and structural analytical techniques (*e.g.*, cryo-EM, advanced protein MS) is rapidly evolving. Tools like deep learning-based protein structure prediction (AlphaFold) and generative protein design (ProteinMPNN) substantially accelerate epitope construct optimisation and *de novo* designed immunogen engineering. Still, these computational models can miss subtleties inherent to real-world infection and *in vivo* immunogenicity. Thus, the future likely hinges on an iterative workflow: predict, screen, validate via MS-based proteomics, refine the design, and revalidate *in vivo*. As technology evolves, the field must maintain a balance between the high-throughput and precision offered by advanced algorithms

and the biological realism mandated by functional profile assays in relevant host environments.

In summary, while these four papers collectively demonstrate the promise of multimodal MS to elucidate host-pathogen interactions and to guide epitope-focused vaccine design, the road ahead involves refining these methods for broader application, ensuring that *in vitro* gained structural insights correlate accurately with effective *in vivo* immunity, and navigating the complexities of real-world human immune systems to achieve desirable functional outcomes.

6 Conclusion and future perspectives

6.1 Conclusion

Across the four papers, a recurring narrative emerges: functional proteomics, structural proteomics, *de novo* protein sequencing, and data-driven computational modelling enables detailed characterisation of understudied bacterial virulence mechanisms and opens new avenues for epitope-targeting vaccine design. For GAS, the discovery of a plasminogen-binding role for SLO challenges the conventional view of CDC toxins as mere pore formers on cellular membranes, while integrated multimodal MS workflows reveal critical, conformation-dependent epitopes. Translating these insights, engineering an optimised epitope and displaying redesigned epitopes on a nanoparticle scaffold yields a protective immunogen that holds promise for further refinement and broader implementation. Extending the multimodal MS methodology to PLY underscores the universality of this approach and hints at trans-species pathogen applicability.

Looking ahead, the synergy of advanced protein MS, computational biology, and immunology stands to redefine how bacterial vaccines are strategically conceived and optimised. By combining accurate epitope mapping with novel immunogen presentation/delivery system, we can further accelerate the development of next-generation immunotherapies that specifically targets key bacterial pathomechanisms, affording a promising solution to both persistent and emerging infectious diseases.

6.2 Future perspectives

6.2.1 Toward enriched MS and integrated structural biology

A logical next step is to broaden the application of these integrative MS workflows. While the current papers concentrate on SLO, pneumolysin, and their respective antibodies, other CDC-related proteins or unrelated other GAS virulence factors could be explored in the same manner. By optimising steps such as automating XL-MS cross-linking workflows or adopting newly developed HDX-MS DIA workflow, we could handle higher-throughput projects encompassing multiple

toxins, more diverse antibody panels featuring monoclonal or polyclonal antibodies, or clinically derived samples. Realising this vision requires closer manipulation among advanced automation, analytical upgrade and data analysis. Another key avenue is further integration with classical biophysical techniques such as X-ray crystallography, small-angle X-ray scattering, NMR, and cryo-EM, to broaden the scope of integrated structural biology.

6.2.2 Incorporating T-cell epitope analysis

Although these studies shed light on B-cell focused protective epitopes, effective vaccines often rely on T-cell involvement for establishing long-term immune protection. Incorporating T-cell epitope mapping, possibly via emerging MS-based immunopeptidomics and prediction algorithms (*e.g.*, NetMHC) validated against various MHC molecule binding assays, could highlight linear peptides of significance that enhance immunogenicity and immunological memory. A combined B- and T-cell synergy might resolve certain challenges, such as why some constructs yield high antibody titres but relatively modest effects. The next wave of epitope-focused vaccines could thus harness both arms of adaptive immunity in a purposefully designed manner.

6.2.3 Pushing vaccine platforms

Beyond the I53_dn5 nanoparticle system highlighted in ***Paper III***, numerous self-assembling scaffolds (e.g., ferritin, E2 protein, virus-like particles) could be tested for presenting defined epitopes of interest. Each platform offers distinct advantages in geometry, thermostability, and valency choice. Tailoring a nanoparticle to the orientation demands of a specific epitope domain might improve vaccine efficacy. Future designs might also include modular compartments for combining multiple epitopes, spanning different domains of a toxin or even bridging other toxins and important immune determinants. Simultaneously, rational adjuvant selection could further optimise T-helper responses and immunological memory formation.

6.2.4 Exploring cross-pathogen protection

The cross-reactivity data from ***Paper IV***, where an anti-PLY mAb recognised a physiochemically and structurally similar region present in multiple CDC toxins, prompts the exciting possibility of a universal vaccine targeting CDC families. Identifying the functionally crucial conformational epitope that share among the CDC family could be the springboard for an immunogen with protective potential against multiple pathogens (*e.g.*, *S. pyogenes, S. pneumoniae, Listeria monocytogenes, or Clostridium perfringens*) simultaneously. To achieve this would require systematic screening of relevant toxins, verifying that each share

conformationally similar structural motifs for sufficient cross-species neutralisation, and designing an immunogen robust enough to account for subtle species-specific variations in the backbone residues.

6.2.5 Clinical translation and regulatory pathways

To progress from lab bench findings to real-world applications, pipeline optimisation is paramount. Preclinical immunogenicity tests must scale to more diverse models, possibly including non-human primates or genetically modified mouse strains (*e.g.*, human transgenic strain). Efficacy must be evaluated against diverse clinical isolates, representing epidemically dominant strains. Moreover, a successful transition to clinical trials demands rigorous assessment of vaccine safety, immunological correlates of protection, and manufacturing feasibility. Regulatory bodies often require extensive data on batch consistency, adjuvant safety, and universally arised immunogenicity, challenges that become especially pronounced for novel protein-based nanoparticle platforms or redesigned epitope constructs.

6.2.6 Hybrid approaches: computation meets experiment

Contemporary advances in protein structure prediction (*e.g.*, AlphaFold and Rosetta) and generative protein design (*e.g.*, ProteinMPNN) streamline the *de novo* creation of stable epitope and the supporting scaffolds. The synergy between these computational methods and MS-based validation holds great potential. In an iterative loop, a hypothesised immunogen structure can be swiftly generated *in silico*, expressed recombinantly, and then probed by HDX-MS and XL-MS to verify actual folding and epitope solvent accessibility (exposure). Incorrect folding or undesired instability can be corrected in subsequent iteration loops of design-production-validation. This cycle could drastically accelerate optimal immunogen engineering compared to trial-and-error methods of the past.

6.2.7 Potential impact beyond bacterial toxins

While the immediate work centres on bacterial cytolysins, the conceptual underpinnings, distilling a protein's functional "Achilles' heel" into a tractable epitope, could apply equally to biologically relevant bacterial, viral, or parasitic antigens. The workflow for epitope identification and structural confirmation is readily adaptable. If the critical neutralisation site of a given pathogenic antigen is revealed, a stable immunogen preserving that site can be computationally designed, validated by MS and other biophysics methods, and then tested for *in vivo* efficacy in immunisation models. Over time, these multi-epitope or multi-pathogen vaccine constructs could lessen the need for repeated immunisation in response to emerging

variants, provided that the chosen epitopes are sufficiently conserved, a feature readily verified through high-throughput genomic approaches.

6.2.8 Conclusion of prospective vision

To summarise, the future of bacterial vaccinology appears poised to shift from broad-spectrum guesses to meticulously engineered constructs guided by robust structural data. The four papers featured here illustrate a miniature of that transition, revealing uncharacterised functions of a classic toxin, delineating precise neutralising epitopes, and validating novel vaccine modalities *in vivo*. As more groups adopt and refine these MS-based and computational strategies, one can anticipate a cascade of improved immunogens tailored to inhibiting key virulence functions with minimal collateral issues such as immunodominance hierarchy or autoimmunity risks. Ultimately, these developments promise to accelerate the delivery of safer, more effective bacterial vaccines, addressing a global need for innovative disease prevention in the face of rising antibiotic resistance.

Acknowledgements

Johan, first and foremost, I want to thank you, my main supervisor. You have been my biggest inspiration since day one, teaching me how to think strategically like a scientist, how to initiate and sustain collaboration, and how to stay open-minded to new ideas. Your kindness, creativity, and genuine care have helped me grow as both an independent researcher and a person. I really appreciate the time you always made for me, your support in my attending conferences and applying for grants, and the way you made sure my contributions and works were fairly recognised and valued.

Lotta, I would like to thank you from the bottom of my heart, who have been closely supervising and mentoring me since my master's degree project. You were always there whenever I needed help, whether it was troubleshooting confusing experiments or discussing complex data. Your patience, attention to detail, and daily support greatly shaped my lab skills and guided me through many challenges. There is just so much to learn from you and, honestly, such a privilege to work with you as a junior researcher.

Oonagh, **Anders**, and **Magnus**, I want to express my genuine gratitude to each of you as my co-supervisors for bringing a unique perspective to my research, from broad biology/immunology knowledge to microbiological expertise to clinical insights. Your support, encouragement, and belief in my work inspired me to explore new directions and helped shape this thesis project into what it is today. Thank you for being an important part of my PhD journey.

Anita, I cannot imagine how I would have handled all the administrative details without you. From anything financial to residence permits, you managed everything with professionalism, calmness, efficiency and kindness, even during the toughest times.

Anahita, thank you for being such a wonderful friend and colleague. Your positive attitude and team spirit made coming to work every day something I truly looked forward to. You helped create a caring and fun working environment and led a organised mode in both IMP and BioMS labs, for which we should all be grateful.

Simon, working with you for more than two years has been one of the best parts of my PhD. You were not only an excellent scientist but also a mentor who helped me see problems through a critical lens. I really enjoyed those hours in your office, brainstorming and interpreting results. I'm glad our collaboration turned out quite productive and great.

Lars M., your help and discussions have been invaluable. I admire your expertise and knowledge in computational biology, tremendously boosting our current work. Thank you for always being on board and responsive.

Hamed, I'm grateful for your collaborative spirit and for showing me how research can flourish when different fields unite. Thank you for broadening my horizons.

Tirthankar, your open-mindedness and friendliness make teamwork effortless. I appreciate your sincere support and generous sharing of insights.

Tommaso, thank you for going the extra mile in maintaining IMP instruments for years. Our everyday chats brightened my day, and I value your kindness.

Christofer, everyone in the lab truly relies on your guidance and help including me, and you never fail them. Your humility and dedication truly made a difference, thank you, CK!

Carlos, I appreciate your humour. Your mass spectrometry and bioinformatics expertise showed me new ways to view data. Every discussion sparked fresh ideas.

Yashuan (**Yosh**), we share many traits, and your generosity never goes unnoticed. Your bakery and kind spirit lift everyone including me, thank you for always caring.

Aaron, your skill with high-dimensional data and out-of-the-box thinking impressed me. Our chats, from research to basketball, were always refreshing.

Erik H., our year as office mates has been wonderful; your diligence, critical thinking and positivity stood out. Thank you for being an inspiring presence.

Erik M., you balance clinical duties with research seamlessly, and your dedication motivates me. Thanks for motivating me and making collaboration so welcoming.

Elisabeth, your involvement in multiple projects pushed our work forward. Thank you for always collaborating with kindness and expertise.

Emma N., from day one, your warm greeting made me feel welcomed in the lab. Thank you for your kindness and hospitality.

Ale, your knowledge and ability to spark curiosity constantly amazed me. Thank you for revealing overlooked points and being insightful.

Joel, your humble approach to data analyses and discussions proved instrumental to our projects. Thank you for being dependable and advancing our research.

Filip, whether discussing research, travel or sharing cat pictures, you brighten my day. Thank you for making everyday life at work more joyful.

Sven, managing BioMS and hosting the farewell party truly united us. I appreciate your generosity and warmth as a senior colleague.

Lucas, your diligence in instrumentation and idea-sparking discussions always impress me. Thank you for your perseverance and motivation in helping my projects.

Mattias, your leadership and knowledge impressed me back in master's study, the lab week you organised guiding me to Johan's group. Thank you for shaping my career.

Sebastian, thank you for being a wonderful friend, making work lunches and dinners memorable. Our conversations and your insightful advice truly made a difference.

Hong, since my earliest days, your mass spectrometry expertise was crucial to my learning. Like everyone, I still miss your friendly presence on D13.

Axel, thank you for bringing an engineering perspective, broadening our lab's horizons. I have been being impressed by your unique insights and research.

Ariane, your guidance, and kindly lending me your lab bench made my work on B14 smoother. I truly appreciate your help and support as always.

Berit, your experimental spirit, diligence, and kindness have been invaluable. Thank you for your vital contribution to many projects I was involved in or led.

Wael, **Martin**, **Arman**, **Pontus**, your contribution and support enriched my research greatly. I'm grateful for our established project collaborations.

To **Wolfgang**, **Celeste**, **Annika**, **Maria**, **Mikael**, **Sara S.**, and **Tanvir**, huge thanks for the critical support you provided through core facilities and the national infrastructure. Your help and assistance significantly advanced my projects.

Collaborators, alumni of IMP/BioMS, and former classmates, **Moritz**, **Indrė**, **Eleni, Sounak**, **Emma A.**, **Paul**, **Lina**, **Dongnian**, **Charlotte**, **Frida**, **Nejc**, **Elin**, **Tilde**, thank you for the time we spent together spanning from projects to lab work to classes. Your support and contributions continue to motivate me

To **Roger**, **Henriett**, **Melinda**, **Lazaro**, **György**, **Thomas**, and colleagues from BMC D13; **Congyu**, **Isabella**, **Frida**, **Sara W.**, **Lars B.**, and colleagues from BMC B14: thank you for fostering a supportive environment around mass spectrometry and infection medicine.

Ying and **Yufang**, you inspired humility and integrity in me as high-profile researchers, shaping my initial scientific approach. Thank you for your mentorship and critical guidance.

Zhengkang (**Lucas**), **Jiajun**, **Chang**, **Yuang**, and **Jiancheng**, across 15 years, our friendship kept me strong and optimistic. Thank you for the laughter, memories, and support.

Meng (**Jasper**), who was once my inspiring English teacher and has become a dear friend, you nurtured my language skills from an early age. That decade of guidance and mentoring opened doors to worlds beyond my own. You have my most sincere gratitude.

The Royal Physiographic Society in Lund and the Sigurd & Elsa Goljes Memory Foundation: your grants enabled me to attend international conferences and pursue innovative projects. Thank you for your generous support.

爸妈，感谢你们从小到大对我的教导、照顾和信任，成为我追逐梦想的力量。你们的不懈努力和初心，是我最好的榜样。现今践行“好男儿志在四方”，不能常相伴身旁，但你们的理解和支持让我更有动力去拼搏，继续前行。谢谢你们，我爱你们。

Mi querida **Zhicang** (**Alba**), a pesar de ser dos años mayor que tú, termino mi doctorado más de dos años después. Estoy agradecido por tu apoyo incondicional, tu paciencia y la confianza que has depositado en mí. Prometo esforzarme al máximo para que alcancemos nuestras metas, porque tu presencia significa todo para mí.

References

1. Richards, V. P. *et al.* Phylogenomics and the Dynamic Genome Evolution of the Genus Streptococcus. *Genome Biol. Evol.* **6**, 741–753 (2014).
2. Castro, S. A. & Dorfmueller, H. C. A brief review on Group A Streptococcus pathogenesis and vaccine development. *Roy Soc Open Sci* **8**, 201991 (2021).
3. Carapetis, J. R., Steer, A. C., Mulholland, E. K. & Weber, M. The global burden of group A streptococcal diseases. *Lancet Infect Dis* **5**, 685–694 (2005).
4. Walker, M. J. *et al.* Disease Manifestations and Pathogenic Mechanisms of Group A Streptococcus. *Clin Microbiol Rev* **27**, 264–301 (2014).
5. Smeesters, P. R., McMillan, D. J. & Sriprakash, K. S. The streptococcal M protein: a highly versatile molecule. *Trends Microbiol* **18**, 275–282 (2010).
6. Cywes, C. & Wessels, M. R. Group A Streptococcus tissue invasion by CD44-mediated cell signalling. *Nature* **414**, 648–652 (2001).
7. Wierzbicki, I. H. *et al.* Group A Streptococcal S Protein Utilizes Red Blood Cells as Immune Camouflage and Is a Critical Determinant for Immune Evasion. *Cell Rep.* **29**, 2979-2989.e15 (2019).
8. Edwards, R. J. *et al.* Specific C-Terminal Cleavage and Inactivation of Interleukin-8 by Invasive Disease Isolates of Streptococcus pyogenes. *J. Infect. Dis.* **192**, 783–790 (2005).
9. Cleary, P. P., Prahbu, U., Dale, J. B., Wexler, D. E. & Handley, J. Streptococcal C5a peptidase is a highly specific endopeptidase. *Infect. Immun.* **60**, 5219–5223 (1992).
10. Buchanan, J. T. *et al.* DNase Expression Allows the Pathogen Group A Streptococcus to Escape Killing in Neutrophil Extracellular Traps. *Curr. Biol.* **16**, 396–400 (2006).
11. Khil, J. *et al.* Plasminogen Enhances Virulence of Group A Streptococci by Streptokinase-Dependent and Streptokinase-Independent Mechanisms. *J. Infect. Dis.* **188**, 497–505 (2003).
12. Vincents, B., Pawel-Rammingen, U. von, Björck, L. & Abrahamson, M. Enzymatic Characterization of the Streptococcal Endopeptidase, IdeS, Reveals That It Is a Cysteine Protease with Strict Specificity for IgG Cleavage Due to Exosite Binding †. *Biochemistry* **43**, 15540–15549 (2004).
13. Naegeli, A. *et al.* Streptococcus pyogenes evades adaptive immunity through specific IgG glycan hydrolysisImmune evasion through IgG glycan hydrolysis. *J Exp Medicine* **216**, 1615–1629 (2019).

14. Sumitomo, T., Nakata, M., Higashino, M., Terao, Y. & Kawabata, S. Group A Streptococcal Cysteine Protease Cleaves Epithelial Junctions and Contributes to Bacterial Translocation*. *J. Biol. Chem.* **288**, 13317–13324 (2013).

15. LaRock, D. L. *et al.* Group A Streptococcus induces GSDMA-dependent pyroptosis in keratinocytes. *Nature* **605**, 527–531 (2022).

16. Michos, A. *et al.* Enhancement of Streptolysin O Activity and Intrinsic Cytotoxic Effects of the Group A Streptococcal Toxin, NAD-Glycohydrolase*. *J Biol Chem* **281**, 8216–8223 (2006).

17. McCormick, J. K., Yarwood, J. M. & Schlievert, P. M. TOXIC SHOCK SYNDROME AND BACTERIAL SUPERANTIGENS: An Update. *Annu. Rev. Microbiol.* **55**, 77–104 (2001).

18. Frost, H., Excler, J.-L., Sriskandan, S. & Fulurija, A. Correlates of immunity to Group A Streptococcus: a pathway to vaccine development. *npj Vaccines* **8**, 1 (2023).

19. LaRock, D. L., Russell, R., Johnson, A. F., Wilde, S. & LaRock, C. N. Group A Streptococcus Infection of the Nasopharynx Requires Proinflammatory Signaling through the Interleukin-1 Receptor. *Infect. Immun.* **88**, 10.1128/iai.00356-20 (2020).

20. Anderson, J. *et al.* Immune signature of acute pharyngitis in a Streptococcus pyogenes human challenge trial. *Nat Commun* **13**, 769 (2022).

21. Hurst, J. R. *et al.* The Streptococcus pyogenes hyaluronic acid capsule promotes experimental nasal and skin infection by preventing neutrophil-mediated clearance. *PLOS Pathog.* **18**, e1011013 (2022).

22. Tanaka, M. *et al.* Group A Streptococcus establishes pharynx infection by degrading the deoxyribonucleic acid of neutrophil extracellular traps. *Sci. Rep.* **10**, 3251 (2020).

23. Kasper, K. J. *et al.* Bacterial Superantigens Promote Acute Nasopharyngeal Infection by Streptococcus pyogenes in a Human MHC Class II-Dependent Manner. *PLoS Pathog.* **10**, e1004155 (2014).

24. Zeppa, J. J. *et al.* Nasopharyngeal infection by Streptococcus pyogenes requires superantigen-responsive Vβ-specific T cells. *Proc. Natl. Acad. Sci.* **114**, 10226–10231 (2017).

25. Hansen, M. B. *et al.* Association between cytokine response, the LRINEC score and outcome in patients with necrotising soft tissue infection: a multicentre, prospective study. *Sci. Rep.* **7**, 42179 (2017).

26. Eislmayr, K. *et al.* Nonredundancy of IL-1α and IL-1β is defined by distinct regulation of tissues orchestrating resistance versus tolerance to infection. *Sci. Adv.* **8**, eabj7293.

27. Zhu, C. *et al.* The gasdermin family: emerging therapeutic targets in diseases. *Signal Transduct. Target. Ther.* **9**, 87 (2024).

28. Deng, W. *et al.* Streptococcal pyrogenic exotoxin B cleaves GSDMA and triggers pyroptosis. *Nature* **602**, 496–502 (2022).

29. Finn, M. B., Ramsey, K. M., Dove, S. L. & Wessels, M. R. Identification of Group A Streptococcus Genes Directly Regulated by CsrRS and Novel Intermediate Regulators. *mBio* **12**, 10.1128/mbio.01642-21 (2021).

30. Gryllos, I. *et al.* Induction of group A Streptococcus virulence by a human antimicrobial peptide. *Proc. Natl. Acad. Sci.* **105**, 16755–16760 (2008).

31. Brouwer, S. *et al.* Streptococcus pyogenes Hijacks Host Glutathione for Growth and Innate Immune Evasion. *mBio* **13**, e00676-22 (2022).

32. Brouwer, S. *et al.* Prophage exotoxins enhance colonization fitness in epidemic scarlet fever-causing Streptococcus pyogenes. *Nat Commun* **11**, 5018 (2020).

33. Cunningham, M. W. Pathogenesis of Group A Streptococcal Infections. *Clin Microbiol Rev* **13**, 470–511 (2000).

34. Bensi, G. *et al.* Multi High-Throughput Approach for Highly Selective Identification of Vaccine Candidates: the Group A Streptococcus Case*. *Mol Cell Proteomics* **11**, M111.015693 (2012).

35. Watson, C. T., Glanville, J. & Marasco, W. A. The Individual and Population Genetics of Antibody Immunity. *Trends Immunol.* **38**, 459–470 (2017).

36. Grilo, A. L. & Mantalaris, A. The Increasingly Human and Profitable Monoclonal Antibody Market. *Trends Biotechnol.* **37**, 9–16 (2019).

37. Uhlen, M. *et al.* A proposal for validation of antibodies. *Nat. Methods* **13**, 823–827 (2016).

38. Georgiou, G. *et al.* The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat. Biotechnol.* **32**, 158–168 (2014).

39. Guthals, A. *et al.* De Novo MS/MS Sequencing of Native Human Antibodies. *J. Proteome Res.* **16**, 45–54 (2017).

40. Hysmith, N. D. *et al.* Prospective Longitudinal Analysis of Immune Responses in Pediatric Subjects After Pharyngeal Acquisition of Group A Streptococci. *J. Pediatr. Infect. Dis. Soc.* **6**, 187–196 (2017).

41. Brandt, E. R. *et al.* Functional analysis of IgA antibodies specific for a conserved epitope within the M protein of group A streptococci from Australian Aboriginal endemic communities. *Int. Immunol.* **11**, 569–576 (1999).

42. Izadi, A. *et al.* The hinge-engineered IgG1-IgG3 hybrid subclass IgGh47 potently enhances Fc-mediated function of anti-streptococcal and SARS-CoV-2 antibodies. *Nat. Commun.* **15**, 3600 (2024).

43. Huang, Y.-S., Fisher, M., Nasrawi, Z. & Eichenbaum, Z. Defense From the Group A Streptococcus by Active and Passive Vaccination With the Streptococcal Hemoprotein Receptor. *J. Infect. Dis.* **203**, 1595–1601 (2011).

44. Malito, E. *et al.* Defining a protective epitope on factor H binding protein, a key meningococcal virulence factor and vaccine antigen. *Proc. Natl. Acad. Sci.* **110**, 3304–3309 (2013).

45. King, M. T. & Brooks, C. L. Epitope Mapping Protocols. *Methods Mol. Biol.* **1785**, 13–27 (2018).

46. Barnes, C. O. *et al.* SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).

47. Nguyen, T. & Song, J. Direct IgG epitope mapping on bacterial AB toxins by cryo-EM. *STAR Protoc.* **2**, 100852 (2021).

48. Pérez-Martínez, D., Infante, Y. C., Ramírez, B. S. & Rojas, G. Domain-level epitope mapping of polyclonal antibodies against HER-1 and HER-2 receptors using phage display technology. *Sci. Rep.* **12**, 12268 (2022).

49. Reineke, U. Antibody Engineering, Methods and Protocols. *Methods Mol. Biol. (Clifton, NJ)* **248**, 443–463 (2003).

50. Weiss, G. A., Watanabe, C. K., Zhong, A., Goddard, A. & Sidhu, S. S. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc. Natl. Acad. Sci.* **97**, 8950–8954 (2000).

51. Dang, X. *et al.* Epitope mapping of monoclonal antibodies: a comprehensive comparison of different technologies. *mAbs* **15**, 2285285 (2023).

52. Gershoni, J. M., Roitburd-Berman, A., Siman-Tov, D. D., Freund, N. T. & Weiss, Y. Epitope Mapping. *BioDrugs* **21**, 145–156 (2007).

53. Hayes, A., Lacey, J. A., Morris, J. M., Davies, M. R. & Tong, S. Y. C. Restricted Sequence Variation in Streptococcus pyogenes Penicillin Binding Proteins. *mSphere* **5**, 10.1128/msphere.00090-20 (2020).

54. Chochua, S. *et al.* Invasive Group A Streptococcal Penicillin Binding Protein 2× Variants Associated with Reduced Susceptibility to β-Lactam Antibiotics in the United States, 2015–2021. *Antimicrob. Agents Chemother.* **66**, e00802-22 (2022).

55. Dale, J. B. & Walker, M. J. Update on group A streptococcal vaccine development. *Curr Opin Infect Dis* **33**, 244–250 (2020).

56. HHS, F. and D. A. Revocation of status of specific products; Group A streptococcus. Direct final rule. *Fed. Regist.* **70**, 72197–9 (2005).

57. Fan, J., Toth, I. & Stephenson, R. J. Recent Scientific Advancements towards a Vaccine against Group A Streptococcus. *Vaccines* **12**, 272 (2024).

58. Dale, J. B. *et al.* Potential coverage of a multivalent M protein-based group A streptococcal vaccine. *Vaccine* **31**, 1576–1581 (2013).

59. Dale, J. B., Penfound, T. A., Chiang, E. Y. & Walton, W. J. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. *Vaccine* **29**, 8175–8178 (2011).

60. Giffard, P. M., Tong, S. Y. C., Holt, D. C., Ralph, A. P. & Currie, B. J. Concerns for efficacy of a 30-valent M-protein-based Streptococcus pyogenes vaccine in regions with high rates of rheumatic heart disease. *PLoS Neglected Trop. Dis.* **13**, e0007511 (2019).

61. Sekuloski, S. *et al.* Evaluation of safety and immunogenicity of a group A streptococcus vaccine candidate (MJ8VAX) in a randomized clinical trial. *PLoS ONE* **13**, e0198658 (2018).

62. Pandey, M. *et al.* Combinatorial Synthetic Peptide Vaccine Strategy Protects against Hypervirulent CovR/S Mutant Streptococci. *J. Immunol.* **196**, 3364–3374 (2016).

63. Ozberk, V. *et al.* Prime-Pull Immunization with a Bivalent M-Protein and Spy-CEP Peptide Vaccine Adjuvanted with CAF®01 Liposomes Induces Both Mucosal and Peripheral Protection from covR/S Mutant Streptococcus pyogenes. *mBio* **12**, 10.1128/mbio.03537-20 (2021).

64. Postol, E. *et al.* Group A Streptococcus Adsorbed Vaccine: Repeated Intramuscular Dose Toxicity Test in Minipigs. *Sci. Rep.* **9**, 9733 (2019).

65. Davies, M. R. *et al.* Atlas of group A streptococcal vaccine candidates compiled using large-scale comparative genomics. *Nat Genet* **51**, 1035–1043 (2019).

66. Gao, N. J. *et al.* Site-Specific Conjugation of Cell Wall Polyrhamnose to Protein SpyAD Envisioning a Safe Universal Group A Streptococcal Vaccine. *Infect. Microbes Dis.* **3**, 87–100 (2020).

67. Benedetto, R. D. *et al.* Rational Design of a Glycoconjugate Vaccine against Group A Streptococcus. *Int. J. Mol. Sci.* **21**, 8558 (2020).

68. Rivera-Hernandez, T. *et al.* An Experimental Group A Streptococcus Vaccine That Reduces Pharyngitis and Tonsillitis in a Nonhuman Primate Model. *mBio* **10**, e00693-19 (2019).

69. Rivera-Hernandez, T. *et al.* Vaccine-Induced Th1-Type Response Protects against Invasive Group A Streptococcus Infection in the Absence of Opsonizing Antibodies. *mBio* **11**, e00122-20 (2020).

70. Bi, S. *et al.* A Multicomponent Vaccine Provides Immunity against Local and Systemic Infections by Group A Streptococcus across Serotypes. *mBio* **10**, 10.1128/mbio.02600-19 (2019).

71. Reglinski, M., Lynskey, N. N., Choi, Y. J., Edwards, R. J. & Sriskandan, S. Development of a multicomponent vaccine for Streptococcus pyogenes based on the antigenic targets of IVIG. *J. Infect.* **72**, 450–459 (2016).

72. Loh, J. M. S. *et al.* A multivalent T-antigen-based vaccine for Group A Streptococcus. *Sci. Rep.* **11**, 4353 (2021).

73. Madge, H. Y. R. *et al.* Physical mixture of a cyclic lipopeptide vaccine induced high titres of opsonic IgG antibodies against group A streptococcus. *Biomater. Sci.* **10**, 281–293 (2021).

74. Shaw, H. A., Remmington, A., McKenzie, G., Winkel, C. & Mawas, F. Mucosal Immunization Has Benefits over Traditional Subcutaneous Immunization with Group A Streptococcus Antigens in a Pilot Study in a Mouse Model. *Vaccines* **11**, 1724 (2023).

75. Spencer, J. A. *et al.* Cross-reactive immunogenicity of group A streptococcal vaccines designed using a recurrent neural network to identify conserved M protein linear epitopes. *Vaccine* **39**, 1773–1779 (2021).

76. Madge, H. Y. R. *et al.* Structure–Activity Analysis of Cyclic Multicomponent Lipopeptide Self-Adjuvanting Vaccine Candidates Presenting Group A Streptococcus Antigens. *J. Med. Chem.* **63**, 5387–5397 (2020).

77. Andrejko, K., Whittles, L. K. & Lewnard, J. A. Health-Economic Value of Vaccination Against Group A Streptococcus in the United States. *Clin. Infect. Dis.* **74**, 983–992 (2021).

78. Mills, J.-L. S. *et al.* M-protein based vaccine induces immunogenicity and protection from Streptococcus pyogenes when delivered on a high-density microarray patch (HD-MAP). *npj Vaccines* **5**, 74 (2020).

79. Sun, H. *et al.* Plasminogen Is a Critical Host Pathogenicity Factor for Group A Streptococcal Infection. *Science* **305**, 1283–1286 (2004).

80. Osowicki, J. *et al.* A controlled human infection model of Streptococcus pyogenes pharyngitis (CHIVAS-M75): an observational, dose-finding study. *Lancet Microbe* **2**, e291–e299 (2021).

81. Correia, B. E. *et al.* Proof of principle for epitope-focused vaccine design. *Nature* **507**, 201–206 (2014).

82. Tian, Y. *et al.* Structure−affinity relationships in the gp41 ELDKWA epitope for the HIV-1 neutralizing monoclonal antibody 2F5: effects of side-chain and backbone modifications and conformational constraints. *J. Pept. Res.* **59**, 264–276 (2002).

83. Ding, P. *et al.* Nanoparticle orientationally displayed antigen epitopes improve neutralizing antibody level in a model of porcine circovirus type 2. *Int. J. Nanomed.* **12**, 5239–5254 (2017).

84. Scheid, J. F. *et al.* Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* **458**, 636–640 (2009).

85. Bachmann, M. F., Mohsen, M. O., Zha, L., Vogel, M. & Speiser, D. E. SARS-CoV-2 structural features may explain limited neutralizing-antibody responses. *npj Vaccines* **6**, 2 (2021).

86. Sesterhenn, F. *et al.* De novo protein design enables the precise induction of RSV-neutralizing antibodies. *Science* **368**, (2020).

87. Morris, C. D. *et al.* Differential Antibody Responses to Conserved HIV-1 Neutralizing Epitopes in the Context of Multivalent Scaffolds and Native-Like gp140 Trimers. *mBio* **8**, 10.1128/mbio.00036-17 (2017).

88. Lainšček, D. *et al.* A Nanoscaffolded Spike-RBD Vaccine Provides Protection against SARS-CoV-2 with Minimal Anti-Scaffold Response. *Vaccines* **9**, 431 (2021).

89. Mu, J. *et al.* Graphormer supervised de novo protein design method and function validation. *Brief. Bioinform.* **25**, bbae135 (2024).

90. Dauparas, J. *et al.* Robust deep learning–based protein sequence design using ProteinMPNN. *Science* **378**, 49–56 (2022).

91. Haas, R. J. de *et al.* Rapid and automated design of two-component protein nanomaterials using ProteinMPNN. *Proc. Natl. Acad. Sci.* **121**, e2314646121 (2024).

92. Zhao, L. *et al.* Nanoparticle vaccines. *Vaccine* **32**, 327–337 (2014).

93. Shin, H. *et al.* Recent Advances in RNA Therapeutics and RNA Delivery Systems Based on Nanoparticles. *Adv. Ther.* **1**, (2018).

94. Morales-Hernández, S., Ugidos-Damboriena, N. & López-Sagaseta, J. Self-Assembling Protein Nanoparticles in the Design of Vaccines: 2022 Update. *Vaccines* **10**, 1447 (2022).

95. Roldão, A., Mellado, M. C. M., Castilho, L. R., Carrondo, M. J. & Alves, P. M. Virus-like particles in vaccine development. *Expert Rev. Vaccines* **9**, 1149–1176 (2010).

96. Wu, Y., Narum, D. L., Fleury, S., Jennings, G. & Yadava, A. Particle-based platforms for malaria vaccines. *Vaccine* **33**, 7518–7524 (2015).

97. Kato, Y. *et al.* Multifaceted Effects of Antigen Valency on B Cell Response Composition and Differentiation In Vivo. *Immunity* **53**, 548-563.e8 (2020).

98. Ols, S. *et al.* Multivalent antigen display on nanoparticle immunogens increases B cell clonotype diversity and neutralization breadth to pneumoviruses. *Immunity* **56**, 2425-2441.e14 (2023).

99. Diaz, D., Care, A. & Sunna, A. Bioengineering Strategies for Protein-Based Nanoparticles. *Genes* **9**, 370 (2018).

100. Pati, R., Shevtsov, M. & Sonawane, A. Nanoparticle Vaccines Against Infectious Diseases. *Front. Immunol.* **9**, 2224 (2018).

101. Yang, L. *et al.* Self-Assembly of Proteins: Towards Supramolecular Materials. *Chem. A Eur. J.* **22**, 15570–15582 (2016).

102. Kobayashi, N. & Arai, R. Design and construction of self-assembling supramolecular protein complexes using artificial and fusion proteins as nanoscale building blocks. *Curr. Opin. Biotechnol.* **46**, 57–65 (2017).

103. Luo, Q., Hou, C., Bai, Y., Wang, R. & Liu, J. Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. *Chem. Rev.* **116**, 13571–13632 (2016).

104. Bangaru, S. *et al.* Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. *Science* **370**, 1089–1094 (2020).

105. Hsia, Y. *et al.* Design of a hyperstable 60-subunit protein icosahedron. *Nature* **535**, 136–139 (2016).

106. Hempstead, P. D. *et al.* Comparison of the three-dimensional structures of recombinant human H and horse L ferritins at high resolution11Edited by R. Huber. *J. Mol. Biol.* **268**, 424–448 (1997).

107. Bale, J. B. *et al.* Accurate design of megadalton-scale two-component icosahedral protein complexes. *Science* **353**, 389–394 (2016).

108. King, N. P. *et al.* Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510**, 103–108 (2014).

109. Buonaguro, L. *et al.* Induction of neutralizing antibodies and cytotoxic T lymphocytes in Balb/c mice immunized with virus-like particles presenting a gp120 molecule from a HIV-1 isolate of clade A. *Antivir. Res.* **54**, 189–201 (2002).

110. Buonaguro, L., Tagliamonte, M., Tornesello, M. L. & Buonaguro, F. M. Developments in virus-like particle-based vaccines for infectious diseases and cancer. *Expert Rev. Vaccines* **10**, 1569–1583 (2011).

111. Kaba, S. A. *et al.* Protective Antibody and CD8+ T-Cell Responses to the Plasmodium falciparum Circumsporozoite Protein Induced by a Nanoparticle Vaccine. *PLoS ONE* **7**, e48304 (2012).

112. Priyanka *et al.* Nanovaccines: A game changing approach in the fight against infectious diseases. *Biomed. Pharmacother.* **167**, 115597 (2023).

113. Reddy, S. T. *et al.* Exploiting lymphatic transport and complement activation in nanoparticle vaccines. *Nat. Biotechnol.* **25**, 1159–1164 (2007).

114. Barnett, T. C. *et al.* Streptococcal toxins: role in pathogenesis and disease. *Cell Microbiol* **17**, 1721–1741 (2015).

115. Feil, S. C., Ascher, D. B., Kuiper, M. J., Tweten, R. K. & Parker, M. W. Structural Studies of Streptococcus pyogenes Streptolysin O Provide Insights into the Early Steps of Membrane Penetration. *J Mol Biol* **426**, 785–792 (2014).

116. Limbago, B., Penumalli, V., Weinrick, B. & Scott, J. R. Role of Streptolysin O in a Mouse Model of Invasive Group A Streptococcal Disease. *Infect Immun* **68**, 6384–6390 (2000).

117. Ato, M., Ikebe, T., Kawabata, H., Takemori, T. & Watanabe, H. Incompetence of Neutrophils to Invasive Group A streptococcus Is Attributed to Induction of Plural Virulence Factors by Dysfunction of a Regulator. *PLoS ONE* **3**, e3455 (2008).

118. Timmer, A. M. *et al.* Streptolysin O Promotes Group A Streptococcus Immune Evasion by Accelerated Macrophage Apoptosis. *J Biol Chem* **284**, 862–871 (2009).

119. Chiarot, E. *et al.* Targeted Amino Acid Substitutions Impair Streptolysin O Toxicity and Group A Streptococcus Virulence. *Mbio* **4**, e00387-12 (2013).

120. Escajadillo, T., Olson, J., Luk, B. T., Zhang, L. & Nizet, V. A Red Blood Cell Membrane-Camouflaged Nanoparticle Counteracts Streptolysin O-Mediated Virulence Phenotypes of Invasive Group A Streptococcus. *Front. Pharmacol.* **8**, 477 (2017).

121. Sierig, G., Cywes, C., Wessels, M. R. & Ashbaugh, C. D. Cytotoxic Effects of Streptolysin O and Streptolysin S Enhance the Virulence of Poorly Encapsulated Group A Streptococci. *Infect Immun* **71**, 446–455 (2003).

122. Bentley, C. C., Hakansson, A., Christianson, J. & Wessels, M. R. Extracellular group A Streptococcus induces keratinocyte apoptosis by dysregulating calcium signalling. *Cell. Microbiol.* **7**, 945–955 (2005).

123. Palmer, M. *et al.* Assembly mechanism of the oligomeric streptolysin O pore: the early membrane lesion is lined by a free edge of the lipid membrane and is extended gradually during oligomerization. *EMBO J.* **17**, 1598–1605 (1998).

124. Sekiya, K. *et al.* Ultrastructural analysis of the membrane insertion of domain 3 of streptolysin O. *Microbes Infect.* **9**, 1341–1350 (2007).

125. Madden, J. C., Ruiz, N. & Caparon, M. Cytolysin-Mediated Translocation (CMT) A Functional Equivalent of Type III Secretion in Gram-Positive Bacteria. *Cell* **104**, 143–152 (2001).

126. Chandrasekaran, S. & Caparon, M. G. The Streptococcus pyogenes NAD+ glycohydrolase modulates epithelial cell PARylation and HMGB1 release. *Cell. Microbiol.* **17**, 1376–1390 (2015).

127. Tsai, W.-J. *et al.* Structural basis underlying the synergism of NADase and SLO during group A Streptococcus infection. *Commun. Biol.* **6**, 124 (2023).

128. Mozola, C. C. & Caparon, M. G. Dual modes of membrane binding direct pore formation by Streptolysin O. *Mol Microbiol* **97**, 1036–1050 (2015).

129. Velarde, J. J., O'Seaghdha, M., Baddal, B., Bastiat-Sempe, B. & Wessels, M. R. Binding of NAD+-Glycohydrolase to Streptolysin O Stabilizes Both Toxins and Promotes Virulence of Group A Streptococcus. *mBio* **8**, 10.1128/mbio.01382-17 (2017).

130. Zhu, L. *et al.* Contribution of Secreted NADase and Streptolysin O to the Pathogenesis of Epidemic Serotype M1 Streptococcus pyogenes Infections. *Am. J. Pathol.* **187**, 605–613 (2017).

131. Richter, J., Brouwer, S., Schroder, K. & Walker, M. J. Inflammasome activation and IL-1β signalling in group A Streptococcus disease. *Cell. Microbiol.* **23**, e13373 (2021).

132. Richter, J. *et al.* Streptolysins are the primary inflammasome activators in macrophages during Streptococcus pyogenes infection. *Immunol. Cell Biol.* **99**, 1040–1052 (2021).

133. Nozawa, T. *et al.* Intracellular Group A Streptococcus Induces Golgi Fragmentation To Impair Host Defenses through Streptolysin O and NAD-Glycohydrolase. *mBio* **12**, 10.1128/mbio.01974-20 (2021).

134. Uchiyama, S. *et al.* Streptolysin O Rapidly Impairs Neutrophil Oxidative Burst and Antibacterial Responses to Group A Streptococcus. *Front Immunol* **6**, 581 (2015).

135. Sakurai, A. *et al.* Specific Behavior of Intracellular Streptococcus pyogenes That Has Undergone Autophagic Degradation Is Associated with Bacterial Streptolysin O and Host Small G Proteins Rab5 and Rab7*. *J. Biol. Chem.* **285**, 22666–22675 (2010).

136. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).

137. Ho, C. S. *et al.* Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin. Biochem. Rev.* **24**, 3–12 (2003).

138. Karpievitch, Y. V., Polpitiya, A. D., Anderson, G. A., Smith, R. D. & Dabney, A. R. Liquid chromatography mass spectrometry-based proteomics: Biological and technological aspects. *Ann. Appl. Stat.* **4**, 1797–1823 (2010).

139. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976–989 (1994).

140. Ng, C. C. A., Zhou, Y. & Yao, Z.-P. Algorithms for de-novo sequencing of peptides by tandem mass spectrometry: A review. *Anal Chim Acta* 341330 (2023) doi:10.1016/j.aca.2023.341330.

141. Cox, J. & Mann, M. Quantitative, High-Resolution Proteomics for Data-Driven Systems Biology. *Annu. Rev. Biochem.* **80**, 273–299 (2011).

142. Low, T. Y. *et al.* Recent progress in mass spectrometry-based strategies for elucidating protein–protein interactions. *Cell. Mol. Life Sci.* **78**, 5325–5339 (2021).

143. Kustatscher, G. *et al.* Understudied proteins: opportunities and challenges for functional proteomics. *Nat Methods* **19**, 774–779 (2022).

144. Smits, A. H. & Vermeulen, M. Characterizing Protein–Protein Interactions Using Mass Spectrometry: Challenges and Opportunities. *Trends Biotechnol.* **34**, 825–834 (2016).

145. O'Reilly, F. J. & Rappsilber, J. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nat Struct Mol Biol* **25**, 1000–1008 (2018).

146. Salas, D., Stacey, R. G., Akinlaja, M. & Foster, L. J. Next-generation Interactomics: Considerations for the Use of Co-elution to Measure Protein Interaction Networks*. *Mol. Cell. Proteom.* **19**, 1–10 (2020).

147. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347–355 (2016).

148. Asensio, N. C., Giner, E. M., Groot, N. S. de & Burgas, M. T. Centrality in the host–pathogen interactome is associated with pathogen fitness during infection. *Nat. Commun.* **8**, 14092 (2017).

149. Dunham, W. H., Mullin, M. & Gingras, A. Affinity-purification coupled to mass spectrometry: Basic principles and strategies. *PROTEOMICS* **12**, 1576–1590 (2012).

150. Schmidt, T. G. & Skerra, A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protoc.* **2**, 1528–1535 (2007).

151. Jäger, S. *et al.* Global landscape of HIV–human protein complexes. *Nature* **481**, 365–370 (2012).

152. Morris, J. H. *et al.* Affinity purification–mass spectrometry and network analysis to understand protein-protein interactions. *Nat Protoc* **9**, 2539–2554 (2014).

153. Luo, Y. *et al.* HIV–host interactome revealed directly from infected cells. *Nat. Microbiol.* **1**, 16068 (2016).

154. Happonen, L. *et al.* A quantitative Streptococcus pyogenes–human protein–protein interaction map reveals localization of opsonizing antibodies. *Nat Commun* **10**, 2727 (2019).

155. Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).

156. Eckhardt, M., Hultquist, J. F., Kaake, R. M., Hüttenhain, R. & Krogan, N. J. A systems approach to infectious disease. *Nat. Rev. Genet.* **21**, 339–354 (2020).

157. Happonen, L. J. Bacterial Pathogenesis, Methods and Protocols. *Methods Mol. Biol.* **2674**, 181–200 (2023).

158. Meyer, K. *et al.* Mutations in Disordered Regions Can Cause Disease by Creating Dileucine Motifs. *Cell* **175**, 239-253.e17 (2018).

159. Frauenstein, A. *et al.* Identification of covalent modifications regulating immune signaling complex composition and phenotype. *Mol. Syst. Biol.* **17**, MSB202010125 (2021).

160. Gogl, G. *et al.* Quantitative fragmentomics allow affinity mapping of interactomes. *Nat. Commun.* **13**, 5472 (2022).

161. Leitner, A., Faini, M., Stengel, F. & Aebersold, R. Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem Sci* **41**, 20–32 (2016).

162. Teo, G. *et al.* SAINTexpress: Improvements and additional features in Significance Analysis of INTeractome software. *J. Proteom.* **100**, 37–43 (2014).

163. Sowa, M. E., Bennett, E. J., Gygi, S. P. & Harper, J. W. Defining the Human Deubiquitinating Enzyme Interaction Landscape. *Cell* **138**, 389–403 (2009).

164. Srivastava, G. P., Yadav, N., Yadav, B. N. S., Yadav, R. K. & Yadav, D. K. Pan-genomics: Applications, Challenges, and Future Prospects. 397–435 (2020) doi:10.1016/b978-0-12-817076-2.00021-4.

165. Piersimoni, L., Kastritis, P. L., Arlt, C. & Sinz, A. Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein–Protein Interactions□A Method for All Seasons. *Chem Rev* **122**, 7500–7531 (2022).

166. Holding, A. N. XL-MS: Protein cross-linking coupled with mass spectrometry. *Methods* **89**, 54–63 (2015).

167. Bartolec, T. K. *et al.* Cross-linking mass spectrometry discovers, evaluates, and corroborates structures and protein–protein interactions in the human cell. *Proc. Natl. Acad. Sci.* **120**, e2219418120 (2023).

168. Yu, C. & Huang, L. Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. *Anal Chem* **90**, 144–165 (2018).

169. Schmidt, C. & Urlaub, H. Combining cryo-electron microscopy (cryo-EM) and cross-linking mass spectrometry (CX-MS) for structural elucidation of large protein assemblies. *Curr. Opin. Struct. Biol.* **46**, 157–168 (2017).

170. Ogorzalek, T. L. *et al.* Small angle X-ray scattering and cross-linking for data assisted protein structure prediction in CASP 12 with prospects for improved accuracy. *Proteins: Struct., Funct., Bioinform.* **86**, 202–214 (2018).

171. Rappsilber, J. The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J Struct Biol* **173**, 530–540 (2011).

172. Gutierrez, C. *et al.* Structural dynamics of the human COP9 signalosome revealed by cross-linking mass spectrometry and integrative modeling. *Proc. National Acad. Sci.* **117**, 4088–4098 (2020).

173. Courouble, V. V. *et al.* Revealing the Structural Plasticity of SARS-CoV-2 nsp7 and nsp8 Using Structural Proteomics. *J. Am. Soc. Mass Spectrom.* **32**, 1618–1630 (2021).

174. Sarnowski, C. P., Bikaki, M. & Leitner, A. Cross-linking and mass spectrometry as a tool for studying the structural biology of ribonucleoproteins. *Structure* **30**, 441–461 (2022).

175. Yu, C. & Huang, L. New advances in cross-linking mass spectrometry toward structural systems biology. *Curr. Opin. Chem. Biol.* **76**, 102357 (2023).

176. Gong, Z. *et al.* Visualizing the Ensemble Structures of Protein Complexes Using Chemical Cross-Linking Coupled with Mass Spectrometry. *Biophys. Rep.* **1**, 127–138 (2015).

177. Tang, X., Wippel, H. H., Chavez, J. D. & Bruce, J. E. Crosslinking mass spectrometry: A link between structural biology and systems biology. *Protein Sci.* **30**, 773–784 (2021).

178. Gonzalez-Lozano, M. A. *et al.* Stitching the synapse: Cross-linking mass spectrometry into resolving synaptic protein interactions. *Sci. Adv.* **6**, eaax5783 (2020).

179. Sekhar, A. & Kay, L. E. An NMR View of Protein Dynamics in Health and Disease. *Annu. Rev. Biophys.* **48**, 1–23 (2019).

180. Toyama, Y., Harkness, R. W. & Kay, L. E. Structural basis of protein substrate processing by human mitochondrial high-temperature requirement A2 protease. *Proc. Natl. Acad. Sci.* **119**, e2203172119 (2022).

181. Englander, S. W., Mayne, L., Kan, Z.-Y. & Hu, W. Protein Folding—How and Why: By Hydrogen Exchange, Fragment Separation, and Mass Spectrometry. *Annu. Rev. Biophys.* **45**, 1–18 (2015).

182. James, E. I., Murphree, T. A., Vorauer, C., Engen, J. R. & Guttman, M. Advances in Hydrogen/Deuterium Exchange Mass Spectrometry and the Pursuit of Challenging Biological Systems. *Chem Rev* **122**, 7562–7623 (2022).

183. Masson, G. R. *et al.* Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods* **16**, 595–602 (2019).

184. Lermyte, F. Advanced Fragmentation Methods in Biomolecular Mass Spectrometry: Probing Primary and Higher Order Structure with Electrons, Photons and Surfaces. 15–32 (2020) doi:10.1039/9781839161056-00015.

185. Seetaloo, N., Kish, M. & Phillips, J. J. HDfleX: Software for Flexible High Structural Resolution of Hydrogen/Deuterium-Exchange Mass Spectrometry Data. *Anal. Chem.* **94**, 4557–4564 (2022).

186. Lau, A. M. C., Ahdash, Z., Martens, C. & Politis, A. Deuteros: software for rapid analysis and visualization of data from differential hydrogen deuterium exchange-mass spectrometry. *Bioinformatics* **35**, btz022- (2019).

187. Lau, A. M., Claesen, J., Hansen, K. & Politis, A. Deuteros 2.0: Peptide-level significance testing of data from hydrogen deuterium exchange mass spectrometry. *Bioinformatics* **37**, btaa677- (2020).

188. Ständer, S., Grauslund, L. R., Scarselli, M., Norais, N. & Rand, K. Epitope Mapping of Polyclonal Antibodies by Hydrogen–Deuterium Exchange Mass Spectrometry (HDX-MS). *Anal Chem* **93**, 11669–11678 (2021).

189. Grauslund, L. R. *et al.* Epitope Mapping of Human Polyclonal Antibodies to the fHbp Antigen of a Neisseria Meningitidis Vaccine by Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS). *Mol. Cell. Proteom.* **23**, 100734 (2024).

190. Vorauer, C. *et al.* Direct Mapping of Polyclonal Epitopes in Serum by HDX-MS. *Anal. Chem.* (2024) doi:10.1021/acs.analchem.4c03274.

191. Savidor, A. *et al.* Database-independent Protein Sequencing (DiPS) Enables Full-length de Novo Protein and Antibody Sequence Determination*. *Mol. Cell. Proteom.* **16**, 1151–1161 (2017).

192. Schulte, D., Peng, W. & Snijder, J. Template-Based Assembly of Proteomic Short Reads For De Novo Antibody Sequencing and Repertoire Profiling. *Anal. Chem.* **94**, 10391–10399 (2022).

193. Bondt, A. *et al.* Human plasma IgG1 repertoires are simple, unique, and dynamic. *Cell Syst* **12**, 1131-1143.e5 (2021).

194. Bihan, T. L. *et al.* De novo protein sequencing of antibodies for identification of neutralizing antibodies in human plasma post SARS-CoV-2 vaccination. *Nat. Commun.* **15**, 8790 (2024).

195. Tran, N. H. *et al.* Complete De Novo Assembly of Monoclonal Antibody Sequences. *Sci. Rep.* **6**, 31730 (2016).

196. Sen, K. I. *et al.* Automated Antibody De Novo Sequencing and Its Utility in Biopharmaceutical Discovery. *J Am Soc Mass Spectr* **28**, 803–810 (2017).

197. Peng, W., Pronker, M. F. & Snijder, J. Mass Spectrometry-Based De Novo Sequencing of Monoclonal Antibodies Using Multiple Proteases and a Dual Fragmentation Scheme. *J. Proteome Res.* **20**, 3559–3566 (2021).

198. Alford, R. F. *et al.* The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J Chem Theory Comput* **13**, 3031–3048 (2017).

199. Senior, A. W. *et al.* Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710 (2020).

200. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

201. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* eabj8754 (2021) doi:10.1126/science.abj8754.

202. Greenshields-Watson, A., Abanades, B. & Deane, C. M. Investigating the ability of deep learning-based structure prediction to extrapolate and/or enrich the set of antibody CDR canonical forms. *Front. Immunol.* **15**, 1352703 (2024).

203. Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. *Biorxiv* 2021.10.04.463034 (2022) doi:10.1101/2021.10.04.463034.

204. Yin, R., Feng, B. Y., Varshney, A. & Pierce, B. G. Benchmarking AlphaFold for protein complex modeling reveals accuracy determinants. *Protein Sci* **31**, e4379 (2022).

205. Ambrosetti, F., Jiménez-García, B., Roel-Touris, J. & Bonvin, A. M. J. J. Modeling Antibody-Antigen Complexes by Information-Driven Docking. *Structure* **28**, 119-129.e2 (2020).

206. Marks, C. & Deane, C. M. Antibody H3 Structure Prediction. *Comput. Struct. Biotechnol. J.* **15**, 222–231 (2017).

207. Abanades, B. *et al.* ImmuneBuilder: Deep-Learning models for predicting the structures of immune proteins. *Commun. Biol.* **6**, 575 (2023).

208. Ianni, A. D., Ianni, A. D., Cowan, K., Barbero, L. M. & Sirtori, F. R. Leveraging Cross-Linking Mass Spectrometry for Modeling Antibody–Antigen Complexes. *J. Proteome Res.* **23**, 1049–1061 (2024).

209. Stahl, K., Graziadei, A., Dau, T., Brock, O. & Rappsilber, J. Protein structure prediction with in-cell photo-crosslinking mass spectrometry and deep learning. *Nat Biotechnol* 1–10 (2023) doi:10.1038/s41587-023-01704-z.

210. Lensink, M. F. *et al.* Impact of AlphaFold on structure prediction of protein complexes: The CASP15-CAPRI experiment. *Proteins: Struct., Funct., Bioinform.* **91**, 1658–1683 (2023).

211. Cheng, Y. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell* **161**, 450–457 (2015).

212. Nakane, T. *et al.* Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152–156 (2020).

213. Ilca, S. L. *et al.* Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. *Nat. Commun.* **6**, 8843 (2015).

214. Briggs, J. A. G. *et al.* Classification and three-dimensional reconstruction of unevenly distributed or symmetry mismatched features of icosahedral particles. *J. Struct. Biol.* **150**, 332–339 (2005).

215. Morais, M. C. *et al.* Cryoelectron-Microscopy Image Reconstruction of Symmetry Mismatches in Bacteriophage φ29. *J. Struct. Biol.* **135**, 38–46 (2001).

216. Stagg, S. M. *et al.* Structure of the Sec13/31 COPII coat cage. *Nature* **439**, 234–238 (2006).

217. Abrishami, V. *et al.* Localized reconstruction in Scipion expedites the analysis of symmetry mismatches in cryo-EM data. *Prog. Biophys. Mol. Biol.* **160**, 43–52 (2021).

218. Long, F. *et al.* Cryo-EM structures elucidate neutralizing mechanisms of anti-chikungunya human monoclonal antibodies with therapeutic activity. *Proc. Natl. Acad. Sci.* **112**, 13898–13903 (2015).

219. Ciferri, C. *et al.* The trimeric serine protease HtrA1 forms a cage-like inhibition complex with an anti-HtrA1 antibody. *Biochem. J.* **472**, 169–181 (2015).

220. Lee, J. H., Ozorowski, G. & Ward, A. B. Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. *Science* **351**, 1043–1048 (2016).

221. Liu, Q. *et al.* Quaternary contact in the initial interaction of CD4 with the HIV-1 envelope trimer. *Nat. Struct. Mol. Biol.* **24**, 370–378 (2017).

222. Renaud, J.-P. *et al.* Cryo-EM in drug discovery: achievements, limitations and prospects. *Nat. Rev. Drug Discov.* **17**, 471–492 (2018).

223. Schwalbe, H. *et al.* The future of integrated structural biology. *Structure* **32**, 1563–1580 (2024).

224. Liu, F., Rijkers, D. T. S., Post, H. & Heck, A. J. R. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* **12**, 1179–1184 (2015).

225. Banerjee, A., Bhakta, S. & Sengupta, J. Integrative approaches in cryogenic electron microscopy: Recent advances in structural biology and future perspectives. *iScience* **24**, 102044 (2021).

226. Tang, D. *et al.* Streptolysin O accelerates the conversion of plasminogen to plasmin. *Nat. Commun.* **15**, 10212 (2024).

227. Tang, D. *et al.* Multimodal Mass Spectrometry Identifies a Conserved Protective Epitope in S. pyogenes Streptolysin O. *Anal. Chem.* **96**, 9060–9068 (2024).

228. Bahnan, W. *et al.* A human monoclonal antibody bivalently binding two different epitopes in streptococcal M protein mediates immune function. *Embo Mol Med* (2022) doi:10.15252/emmm.202216208.

229. Gueto-Tettay, C. *et al.* Multienzyme deep learning models improve peptide de novo sequencing by mass spectrometry proteomics. *Plos Comput Biol* **19**, e1010457 (2023).

230. Loof, T. G., Deicke, C. & Medina, E. The role of coagulation/fibrinolysis during Streptococcus pyogenes infection. *Front. Cell. Infect. Microbiol.* **4**, 128 (2014).

231. Urano, T., Castellino, F. J. & Suzuki, Y. Regulation of plasminogen activation on cell surfaces and fibrin. *J. Thromb. Haemost.* **16**, 1487–1497 (2018).

232. Fulde, M., Steinert, M. & Bergmann, S. Interaction of streptococcal plasminogen binding proteins with the host fibrinolytic system. *Front Cell Infect Mi* **3**, 85 (2013).

233. Walker, M. J., McArthur, J. D., McKay, F. & Ranson, M. Is plasminogen deployed as a Streptococcus pyogenes virulence factor? *Trends Microbiol* **13**, 308–313 (2005).

234. Bastiat-Sempe, B., Love, J. F., Lomayesva, N. & Wessels, M. R. Streptolysin O and NAD-Glycohydrolase Prevent Phagolysosome Acidification and Promote Group A Streptococcus Survival in Macrophages. *Mbio* **5**, e01690-14 (2014).

235. Hancz, D. *et al.* Streptolysin O Induces the Ubiquitination and Degradation of Pro-IL-1β. *J Innate Immun* **11**, 457–468 (2019).

236. Lanzavecchia, A., Frühwirth, A., Perez, L. & Corti, D. Antibody-guided vaccine design: identification of protective epitopes. *Curr Opin Immunol* **41**, 62–67 (2016).

237. Zundert, G. C. P. van *et al.* The DisVis and PowerFit Web Servers: Explorative and Integrative Modeling of Biomolecular Complexes. *J Mol Biol* **429**, 399–407 (2017).

About the author

Di Tang was born and raised in Guiyang, Guizhou, China. He earned his bachelor's degree at Shanghai Jiao Tong University School of Medicine before returning to Sweden to pursue further studies at Lund University Faculty of Medicine. Di has always been intrigued by the complexities of immunobiology and metabolism. He strives to tackle meaningful scientific questions with an experimental mindset, continuously refining and integrating cross-disciplinary methodologies. This thesis represents his efforts to expand his knowledge and practice in protein mass spectrometry and structural biology within the context of infection medicine.

Printed by Media-Tryck, Lund 2025 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

FACULTY OF MEDICINE

Division of Infection Medicine
Department of Clinical Sciences, Lund

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2025:24
ISBN 978-91-8021-677-7
ISSN 1652-8220