Defining the Structural Features and Functional Characteristics of Complement-Active IgM

Michael Jeremy Watson

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Miklos Guttman, Chair

William M. Atkins

Kelly K. Lee

Program Authorized to Offer Degree:

Medicinal Chemistry

University of Washington

Abstract

Defining the Structural Features and Functional Characteristics of Complement-Active IgM

Michael Jeremy Watson

Chair of the Supervisory Committee:

Miklos Guttman

Department of Medicinal Chemistry

Immunoglobulin M (IgM) is an extremely important, yet surprisingly under-studied glycoprotein that is shared among all vertebrates on an evolutionary timeline spanning more than 500 million years. In addition to the genetically conserved subset of 'natural' IgM that conveys an innate degree of protection before birth, the adaptive immune system also produces 'induced' IgM as the initial response to all foreign pathogens. IgM is the most potent known activator of the classical complement cascade, making it an attractive therapeutic platform for the potential treatment of a wide array of disease states; while multiple IgG molecules must bind with coordination of their Fc regions together at the membrane of a pathogenic target, a single molecule of IgM is sufficient for the full-scale activation of complement, owing to the multivalency of its secreted pentameric and hexameric isoforms. Although the relationship between IgM and complement has been appreciated for decades, the biophysical details that underlie its activation mechanism still remain largely unknown. Many past attempts have been made to uncover the structural details surrounding the IgM-mediated activation of complement, but the same physical traits that give IgM its unique properties and relationship with the immune system have also earned it a notorious reputation for being technically difficult to handle and investigate experimentally. The large size, extensive glycosylation, and high degree of flexibility has rendered IgM largely refractive to many established structural techniques, and to date no fully intact IgM structures have been resolved; all structural characterizations and protein:protein interfaces have so far been inferred from combinations

of fragmented structures, homology modeling, electron microscopy, X-ray scattering, and molecular dynamics simulations. While a number of current IgM drug candidates have shown promise for the treatment of several different cancers, the biophysical details of IgM-mediated complement activation still remain poorly understood.

The objective of this dissertation is to establish a baseline set of structural features and functional characteristics that define the complement-active, antigen bound form of IgM – commonly referred to as the 'staple' conformation – and differentiate it further from other sub-active and inactive forms. Chapter 1 provides a historical review and outlines the contextual background that currently surrounds some of the major gaps in knowledge regarding the structure and function of IgM as it relates to the activation of complement. In Chapter 2, background work is presented on the development of a low-cost, automated LC-MS system designed to improve both the throughput and consistency of decoupled hydrogen/deuterium exchange mass spectrometry (HDX-MS) samples, which is the primary structural technique used to assess the defining features of the staple conformation. In Chapter 3, the functional aspects of complement-active IgM are explored through a combination of novel ligand binding and activity assay approaches, revealing a number of key variables that directly impact the binding kinetics and activation rates of the initiating molecules of the classical cascade pathway, C1q and C1. The mechanistic role of IgM binding valency/avidity was also investigated through a prebinding approach that, to our surprise, indicated there to be no hard stoichiometric cutoff for the number of surface bound arms of IgM required to bind/activate C1q/C1. The functional characteristics determined in Chapter 3 pair directly with the structural features found in Chapter 4 (via HDX-MS) in order to complete the defining structure/function relationship that underlies the active staple conformation of IgM. Strong evidence of a mechanistic hinge, compensatory rearrangement of domains within the Fc core, and the identity of previously unreported residues that may comprise a portion of the C1q binding site (in addition to the longstanding putative location) were discovered within the multivalent surface bound

form of complement-active IgM. Additionally, the existence of an unexpectedly extensive network of allosterism was observed within solution bound IgM that could add much needed clarity to the litany of confusing reports found throughout the classical literature. Finally, in Chapter 5, a brief project conclusion is provided along with a discussion of future directions. Collectively, this work establishes key structural and functional qualities that define the complement-active form of IgM, along with a number of newly developed methods, improved protocols, and a foundation for IgM structural work upon which future research can be built.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I want to first express my sincerest gratitude to all of the kindhearted staff and brilliant faculty of the UW department of Medicinal Chemistry for fostering such a uniquely warm, supportive, and collaborative learning environment. I feel incredibly fortunate to have had the opportunity to interact and learn from all of you.

I will be forever thankful to my advisor – Mike Guttman – not only for all of the time, support, and knowledge that he has extended to me throughout the years, but also for the kindness, patience, and incredibly steadfast example that he set for me as a professional scientist. You were an amazing boss, but more importantly, you were a true role model for me personally and I hope that I can live up to the example you set for me as I go forward in life. While a mere 'thank you' feels hilariously insufficient in response to all that you've done for me, I'm eternally thankful that I had the opportunity to join your lab, and I'm also truly honored to have been your first student. Some of my favorite memories come especially from the early days, when we would lose complete track of time and find ourselves working late into the night on some Frankenstein-like instrument, or even when simply spit-balling ideas during one of our many world-famous whiteboard brainstorming sessions. You always heard me out and listened to all of my ideas (even the bad ones!), and I really appreciate the freedom you allowed me to have along with all of the support you gave me when things didn't go according to plan. I feel a real sense of accomplishment from the work that we did, and I look forward to seeing where it goes in the future!

To all of my fellow lab members, past and present – thank you for all the fun moments we shared, and good luck! I'd like to give a special shoutout to Abby Mookherjee and Sunny Uppal for allowing me to waltz right into that office and endlessly distract you both, and I wish the both of you nothing but continued success in your future endeavors. To those of you still 'sending it' – Ellie James, Jonathan Palmer, Alesi Escobedo, and especially Charlie Mundorff – keep up the good work, and keep in touch! Grad school is tough at times, sure, but you've got a great boss and I'm confident that all of you will continue to be incredibly successful!

To my cohort – Eddie Hodge and Amy Li – it's been an honor. I'm so thankful to you both for being such great friends and for all of your support throughout the years. You were always there to help keep me going, and I hope I provided some of that same support to you along the way.

Finally, to my incredibly beautiful, extremely loving, and endlessly patient wife – Siobhan – I love you, and I would never have been able to accomplish this without you in my corner. I hope I've made you proud, and I can't wait for the next chapter of our lives. After all the love you've shown me, along with the love we both continue to get from our amazing dog, Bella, I feel like the luckiest man alive. To my late dog – Sammy – who I lost unexpectedly while writing the final chapter of this dissertation: rest in peace, and I will forever miss you. You were my best friend, and I will never forget the last 11 years spent with you by my side.

– Chapter 1 –

Introduction to IgM and the Activation of Complement

1.1 Introduction to the Mammalian Antibody System

"…More often the sick and the dying were tended by the pitying care of those who had recovered, because they knew the course of the disease and were themselves free from apprehension. For no one was ever attacked a second time, or not with a fatal result." – Thucydides, *History of the Peloponnesian War*, 430 BCE, Athens, Greece

For millennia, we as humans have understood that if one was fortunate enough to survive the spread of a pathogenic infection or 'disease' that was otherwise ravaging those around them, they would no longer have to fear the illness themselves and could begin to participate in the caretaking of others. Although the basic concept of acquired immunity had already begun to take root for those living in the times of ancient Greek historian Thucydides (who provides us with one of the earliest known eyewitness accounts of the spread of a pandemic plague), the term "antibody" would not appear for another 2,000+ years. Indeed, the very concept of our immunity being comprised of two independent-yet-cooperative systems is a relatively new idea in human history, despite how quickly we've grown accustomed to the idea of antibodies and antibody-based therapeutics today.

Since the first FDA-approved monoclonal antibody (mAb) therapy was released to the public in 1985, antibody-based drugs have continued to grow substantially in both their engineering and range of utilization, with more than 170 mAb therapies approved or under regulatory review by the end of 2022, with a further 1,200+ drug candidates registered into clinical trials.[1,2] All of our currently approved antibody therapeutics are understandably based on the framework of Immunoglobulin Gamma (IgG), as it is by far the most well-established and intimately understood class of antibody within the literature. However, there are five mammalian antibody classes (isotypes) in total – IgA (α), IgD (δ), IgE (ε), IgG (γ),

and IgM (μ) – and each of them plays a unique role within the immune system that might also be utilized therapeutically in future antibody-based drug platforms (**Figure 1.1.1**).[3]

Discovered by Tiselius & Kabat in the late 1930s, IgG was the first official antibody class to be established within the mammalian system, and due to its prominent role and abundance *in vivo* along with its favorable physical properties – such as its relatively small size (~150 kDa), minimal glycosylation (only a single conserved N-linked glycan), and extended circulatory half-life (~21 days) – has justifiably led to it becoming the most extensively researched and well-understood antibody class within our collective repertoire of knowledge regarding the adaptive immune system.[4] The classic 'Y' shape structure of the IgG molecule is constructed from two identical heavy (γ) and light (κ or λ) chain pairings that can be further divided into three functionally-distinct regions. At the N-terminal side of the molecule are two identical domains that mediate antibody binding to antigens, otherwise known as the Fab regions. This portion of the molecule contains the full light chain along with the first heavy domain (Cγ1) that is conserved across all four primary isoforms, or subclasses of IgG (IgG1 – IgG4). At the C-terminal base of each Fab domain lies a typically proline-rich 'hinge' region, which affords a significant degree of segmental flexibility to the molecule and varies in both length and composition between each isoform. The hinge region physically bridges the two N-terminal Fab domains to the tightly packed C-terminal portion of the antibody, called the Fc region, which derives its name from being the fragment of the protein that is readily crystallizable. The Fc contains all of the conserved domains and amino acids necessary to provide each class of antibody their own unique effector functions within the immune system.

While all of the mammalian antibodies share a generally similar architecture and function within the Fabs, the structure and composition of their Fc region is what defines each class of antibody, and every isoform of the same Ig class will share the conserved residues and common secondary/tertiary domain arrangements that make up the Fc. The knowledge we have gained throughout decades of basic

IgG research has afforded a substantial amount of insight into the structure/function relationship of all antibodies in general, and researchers today are applying the advanced knowledge to further the engineering and therapeutic potential of IgG into new and exciting drug modalities. For example, antibody drug conjugates (ADCs) have been developed to effectively combine small and large molecule agents into potent complexes toward the treatment of various disease states. There are also now bispecific, and even some trispecific molecules that combine multiple complementary effector functions together. Some of the newest drug candidates also include small, single chain antibodies derived from exotic species such as camelids and sharks that provide increased access to tight junctional regions, such as tumor microenvironments, while also improving upon the immunogenicity and tolerance profile of future biologics in general.[1,5,6]

The IgA antibody was discovered in 1959 by Heremans et. al., and its two isoforms have been observed to exist naturally as secreted homodimers in complex with a small (~15 kDa) covalently linked peptide known as the joining chain (J chain).[7] While IgA molecules can be found in systemic circulation, the J chain allows for transport of the vast majority of secreted IgA to the surface of mucosal tissues via the polymeric immunoglobulin receptor (pIgR), where it is regarded as the primary defensive antibody of the mucous membranes and their surrounding environments such as saliva, tears, and breast milk.[8] It is by far the most abundantly produced antibody *in vivo* (upwards of 60 mg/kg per day), which is required to adequately protect the vast surface area and fast-paced regeneration of the cells that compose sensitive mucosal membranes such as the linings of the respiratory and digestive tracts.[9,10] The therapeutic potential of IgA has been demonstrated within the literature towards the treatment of various cancers, and was even submitted to phase-II clinical trials as a protective dental agent, formulated as a monoclonal plant-derived IgA therapeutic for the prevention of tooth decay.[11,12]

Despite being discovered as early as 1965 by Rowe & Fahey, IgD still remains the least well-understood antibody within the mammalian system today, partly due to low expression and serum

concentration (~0.25% of total serum Ig) along with a short reported half-life of only 2-3 days, which together make it a practically difficult target for basic research.[13,14] While little is known about either its secreted or membrane bound biological function, recent studies have indicated its involvement within a small subset of B cells that express membrane bound versions of IgD-based receptors, and the secreted form of IgD is thought to play a supportive role in the defense of mucosal tissues that line the airway and digestive tracts, which are predominantly guarded by IgA.[8] Although the current gaps in knowledge surrounding IgD largely preclude it from being pursued as a therapeutic drug candidate, it likely plays a role in many fundamental aspects of the adaptive immune system and will undoubtedly continue to be a primary target of basic research in future studies.

IgE was the last mammalian antibody class to be discovered in 1966 by Ishizaka et. al., and in 2002 its X-ray crystal structure revealed it to be a single monomeric unit composed of four distinct heavy domain (ε) dimers, arranged into a more tightly compacted conformation than had previously been expected.[15–17] The antibody was also observed to lack a clearly defined (based on IgG) hinge region entirely, which is an important structural feature that will be discussed in further detail throughout the following chapters of this dissertation. While it primarily defends against parasitic infections such as roundworms and tapeworms, IgE has also become notorious for its well-understood role in the mediation of adverse immune reactions due to a high binding affinity with FcεRI, which is expressed along both lung and intestinal epithelial cells along with other potent immune cell surfaces such as mast cells and basophils.[16,18] Despite the high degree of structural homology between FcεRI and IgG-based receptors, such as FcRn, the IgE binding affinity is ~2-3 orders of magnitude stronger (~$1E^{-10}$ M) to its receptor than IgG. This high degree of binding affinity leads to a relatively low triggering threshold of crosslinked IgE molecules required for the induction of severe hypersensitivity reactions that, in the absence of immediate medical care, can quickly lead to severe anaphylaxis and death. IgE is also able to bind loosely to FcεRII, otherwise known as CD23, which is a polyvalent member of the calcium-

dependent lectin superfamily that is also capable of contributing towards increased hypersensitivity to antigen.[18] The mediation of severe immune reactions by IgE has led to the recent development of a handful of IgG-based drug therapies that target IgE effector function via competitive inhibition to major allergens such as peanuts and shellfish.[19,20]

Finally, the last remaining mammalian antibody class, and the primary subject of the work presented throughout this dissertation – IgM – was actually discovered only shortly after IgG by Waldentsröm and Pedersen, in 1944. The advent of newly developed electrophoresis and ultracentrifugation techniques led them to unexpectedly find "[sic]… *a globulin fraction with a very large molecule (mol. weight about 1,000,000) and a migration that did not correspond to any of the known globulin fractions*.", with a sedimentation profile that greatly exceeded that of IgG, or "β-*globulin*" as it was known at the time.[21] This large myeloma-associated 'macroglobulin' would later be relabeled in the 1960s by an updated nomenclature to the name IgM.[22] Despite being known for more than half a century, the IgM antibody remains an extremely important, yet surprisingly under-studied glycoprotein that plays a fundamental role within many key processes throughout both the innate and adaptive immune systems. The following sections of this introductory chapter are intended to provide the contextual background that currently surrounds some of the major gaps in knowledge concerning the structure and function of the IgM antibody, particularly within its role of immune system activation via interaction with the complement system.

1.2 IgM Structure

The IgM monomer (**Figure 1.2.1**) is constructed in a fashion similar to IgG and the rest of the mammalian antibodies, with two identical heterodimers composed of heavy (μ) and light (λ or κ) chain pairings associating to form the typical 'Y'-shaped molecule that can be further divided into common functional subregions, with the N-terminal Fab portions responsible for mediating antibody binding to

antigens, and the C-terminal Fc region containing the defined structural features necessary for mediating the unique effector functions of each antibody class. The two Fab regions of IgM are composed of the first constant heavy domain (Cμ1) along with the typical Ig variable fragment (Fv) that contains the hypervariable regions of the heavy and light chains that comprise the six complementarity-determining regions (CDRs), which creates the interface (i.e. paratope) that binds directly to antigen (i.e. epitope). Unlike IgG, IgA, and IgD, which feature a distinctly flexible hinge region between the Fab domains and the two remaining constant heavy domains of the Fc (CH2/CH3), the IgM antibody is constructed from a total of four constant heavy domains (analogous to IgE) and connected directly to the Fc region comprised of the Cμ2, Cμ3, and Cμ4 domains, with no classically defined hinge region (see also **Figure 1.2.2** for primary sequencing information).[23–26] When membrane bound, the IgM monomeric subunit functions as the primary B cell receptor (BCR). When secreted, the C-terminal membrane bound portion of the Cμ4 domain is replaced by a smaller soluble polypeptide, known as the 'μ-tailpiece', that directly mediates the oligomerization of monomeric subunits into the secreted pentameric (IgMp) and hexameric (IgMh) polyvalent isoforms.[24,27,28]

While the mechanisms that govern the production ratio of the two secreted subclasses of polyvalent IgM remain largely unknown, it is now well-understood that IgM is predominantly (~95%) secreted within humans in the pentameric form. The IgM pentamer is also typically associated with the same ~15 kDa joining chain (J chain) as IgA, which acts as a placeholder for the sixth potential monomeric 'arm' of the IgM hexamer.[29,30] Several studies have demonstrated that, while the J chain is not strictly required for functional oligomerization, its presence and availability to the B cell will impact the relative proportions of IgMp and IgMh that are secreted.[29,31] The ratio of isoform secretion has also been observed to vary across different species, and some organisms such as teleost fish have even been observed to exclusively produce tetrameric versions of IgM, with a further variety of oligomeric states reported in several bird and reptile species.[31–33] As with IgA, the J chain allows IgMp to be transported

from circulation to the mucosal tissues via the pIgR, which augments the immune functions that are vital to the vulnerable mucosal linings (e.g. eyes, nose, throat) that are predominantly protected by IgA.[34] IgMh, on the other hand, lacks the J chain entirely and appears to be better suited for mediation of the effector function that distinguishes the IgM antibody class – initiation of the classical complement cascade.[35]

IgM has gained notoriety throughout the literature for being difficult to handle experimentally, and detailed structural analysis has largely been hindered historically due to the combination of intrinsic physical attributes that have proven to be technically challenging to overcome, such as its large size, extensive flexibility, and high degree of glycosylation and glycan heterogeneity. When fully oligomerized, IgMp (~960 kDa) and IgMh (~1.1 MDa) measure approximately 40 nm in diameter, making it by far the largest antibody class amongst all currently known vertebrate species.[32,36,37] Despite lacking the typical proline-rich hinge region seen within the flexible IgG isotypes, IgM has been observed to display an unusually high range of segmental flexibility, with some suggesting it to actually be the most flexible of all the mammalian antibody classes.[38–41] Human IgM contains five conserved N-linked glycans on each individual heavy chain (with an additional N-linked glycan provided by the J chain in IgMp) which comprises roughly 25% of the total molecular weight, and is further complicated by a high degree of species-dependent heterogeneity in both glycan type and occupancy rate.[42,43] Still to date, no fully intact IgM crystal structures have been resolved; all structural characterizations have been inferred from combinations of fragmented structures, homology modeling, X-ray scattering, molecular dynamics (MD) simulations, and electron microscopy (EM).[24,27,44–51]

While the void caused by the lack of structural knowledge has lingered for decades, recent work with advanced cryo-EM methods are starting to close the gap. Each of the four conserved heavy chain domains in IgM had so far been resolved individually through a combination of crystallography and NMR until just recently, when two separate groups submitted high-resolution cryo-EM structures of the

human IgM 'Fc core' comprised of the J chain, Cμ4, and majority of the Cμ3 domains oligomerized as an intact pentamer.[23,24,52] While all recent attempts to resolve the fully intact IgM molecule have failed, it is still intriguing and perhaps nonetheless informative to consider that all recent attempts have failed in the exact same fashion, with resolution dropping off precisely at the N-terminus of the Cμ3 domain (**Figure 1.2.3**).[23,45,52] In every case, the flexibility that originates in or near the residues spanning the Cμ3/Cμ2 domain junctions have rendered the rest of the molecule unresolvable, which could be viewed as a practical consequence of a highly flexible hinge that exists within that region. While there still has yet to be any structural determination of the human Cμ2 domain itself, it is expected to follow the general architecture of mouse IgM and exist as a natural homodimer with an intramolecular disulfide bridge at C337. The Cμ3 domain, on the other hand, exists only in monomeric form and is loosely stabilized within the secreted IgM polymer by neighboring domain contacts along a single intermolecular disulfide bridge at C414. The Cμ4 domain is composed into a tightly packed homodimer and is bridged to neighboring Cμ4' domains of adjacent arms by an intermolecular disulfide bridge located within the C-terminal 'μ-tailpiece' subregion, at C575. The two recently deposited high-resolution EM structures have also revealed the human Cμ4 domain to be structurally unique compared to that of murine IgM, featuring an antiparallel architecture that positions the N-termini of each subunit onto opposite sides of the dimer, whereas mouse Cμ4 is arranged in a parallel fashion in which opposing N-termini are on the same side of the dimer. While the significance of the differing Cμ4 domain arrangements is still unclear, it should be noted that the IgM material (MD4) used throughout the studies presented within this dissertation is murine-derived, which differs notably from humans in the number, type, and location of conserved N-linked glycans, as illustrated in **Figure 1.2.4.** [42,53,54]

1.3 IgM Function

Even before birth, an entire subset of 'natural' IgM develops within each individual that confers an inherent, basal degree of generalized protection that is genetically encoded and evolutionarily conserved.[55] These polyreactive IgM are present in circulation prior to any direct contact with foreign antigen, and all mammals will develop this common baseline repertoire of natural protection against basic antigenic epitope patterns.[56,57] In addition to providing a first line of defense, natural IgM also helps to regulate basic cell function and is involved in a number of maintenance processes such as the clearance of healthy apoptotic cells, as well as the early detection and removal of cells that become cancerous.[58–61] At the time of birth, infants are equipped both with natural IgM antibodies as well as the maternal IgG and IgA antibodies received throughout early development, which together provide a generally sufficient buffer against the majority of common environmental pathogens and affords enough time (~6 months) for the adaptive immune system to sufficiently mature.[62]

Throughout an individual's lifetime the adaptive immune system will continue to grow, diversify, and ultimately strengthen the overall level of protection against the bulk of common foreign pathogens. However, due to the encoding of the genes that control B cell activation and the process of class-switching, the antibody that is always produced first in response to novel pathogens, foreign antigens, or even exogenous drug compounds (often of abuse, such as cocaine), is 'induced', or secreted IgM.[63–66] Because it is always the first antibody to be secreted by the adaptive immune system, induced IgM is typically precluded from the benefits of affinity maturation processes, which leads to the production of an antibody with relatively low specificity and monovalent Fab:Ag binding affinity. Despite this, the structural framework of both secreted polyvalent isoforms still enables IgM to play a fundamental role within the immune system that is nonetheless well-suited for carrying out both passive and active forms of immune function. With 10-12 Fab domains per IgM molecule, direct neutralization and crosslinking

often emerge as the dominant mechanisms of handling soluble antigen, leading to agglutination/precipitation events that further expand recruitment of the immune system via large and aggregated Ag:Ab immune complexes.[67] This extreme level of binding valency is the key aspect of IgM that allows it to play such a pivotal role within the early immune response, despite only comprising 5-10% of total serum antibodies.[68] In addition to excelling at the standard clearance mechanisms of direct neutralization and precipitation of soluble antigen, IgM also participates in opsonization, which is a mechanism that effectively marks a target for future clearance by other elements and pathways of the immune system. What sets IgM apart from the rest of the mammalian antibodies is its ability to efficiently opsonize targets via recruitment of the classical complement cascade, which is an incredibly powerful signaling pathway of the innate immune system that ultimately leads to the total destruction (via cell lysis) of pathogenic targets.

By the mid-1980s it was largely agreed that a single IgM molecule is capable of activating complement, owing to the fact that the antibody is inherently capable of producing its own multivalent (pentameric/hexameric) display of binding sites for the initiating molecule of the cascade, C1q.[48,69] In the case of IgG, however, it's clear that the activation process is much less efficient and relies solely on the probability of multiple IgG molecules binding in sufficient proximity so as to coordinate their Fc regions into a sufficiently multivalent display of C1q binding sites.[70–72] For IgM, it has also been demonstrated that the rare IgMh isoform is vastly more potent than the predominant IgMp isoform, with some reports estimating it to be ~20-100X more efficient at activating complement.[73] Besides the obvious addition of a sixth binding 'arm', the only physical difference between IgMp and IgMh is the incorporation of the J chain, which takes the place of the sixth monomeric subunit in IgMp. To add even more complexity, complement activity has been shown to differ further still between J^+ and J^- pentamers such that J^+ IgMp < J^- IgMp << IgMh in terms of activation efficiency, which gives added credence to the notion that the J chain itself could potentially be disruptive to complement activation.[74–76]

In the following sections, further background information is provided about the complement cascade signaling pathway and its initiating molecular components, along with the current leading hypotheses of its IgM-mediated activation process and some of the major gaps in knowledge that this project is aimed towards.

1.4 The Complement Cascade

IgM is the oldest immunoglobulin component of the adaptive immune system and has been shared amongst all vertebrates for more than roughly 500 million years, but its counterpart within the innate immune system – complement – is thought to predate IgM, and possibly even predates the entirety of the adaptive immune system itself.[77,78] First discovered by Bordet in 1895 and described as a "*heat-labile serum component*", the complement family has now grown to include more than 30 individual solution and membrane bound proteins.[79–81]

The initiation of the classical cascade pathway requires antibody-mediated activation of the C1 complex, which is itself a non-covalent heterotrimer composed of C1q (410 kDa) along with serine proteases C1r (92 kDa) and C1s (86 kDa), arranged in a 1:2:2 molar ratio, respectively.[82] C1q is the primary molecule that makes initial contact and binds directly to antibodies and, perhaps more interestingly, is also arranged into a globally hexameric architecture (**Figure 1.4.1**).[83] Often referred to as a "*bouquet of tulips*", the C1q glycoprotein is built from six individual trimers of 18 total A,B, and C polypeptide chains that are bundled together at the N-terminus (the 'stalk' portion) that eventually split apart midway to form six individual, C-terminal globular heads (the 'tulip' portion).[84,85] The N-terminal stalk is collagen-like and imparts a significant degree of structural stability and rigidity, while the C-terminal globular heads resemble Ig-like folding and interact with the Fc region of antibodies (along with many other endogenous ligands) primarily through charge-based pairings.[80,85,86] The C1r and C1s proteases have been observed to associate together in their 'zymogen' (i.e. enzymatically inactive) forms

as a stable heterotetramer ($C1r_2C1s_2$), although the finer details of their arrangement and association mechanisms with C1q are still unclear.[87]

There are three main converging pathways of the complement cascade – the classical, mannose-binding lectin (MB-Lectin), and alternative pathways – each of which feature a unique mechanism of initiation (**Figure 1.4.2**). While it's possible that any one or more of these pathways can be activated in response to pathogens, it's only the classical signaling cascade that bridges the innate and adaptive immune systems together through antibody-mediated opsonization of targets via formation of Ag:Ab immune complexes. Much like other major signaling pathways found throughout the body, the cascade consists of many auto-catalytic steps carried out by a number of proteolytic intermediaries that act on nearby ligands and other downstream protein components. The majority of the participating molecules can be classified as 'zymogens', or pro-enzymes that are activated in a largely sequential order and produce cleavage products that serve to both amplify the signal as well as advance the cascade towards the ultimate formation of the membrane attack complex (MAC), which functions only to destroy targeted cells by creating physical pores in the membrane, leading to complete cell lysis.[88]

While many of the details surrounding the C1 activation mechanism are also poorly understood, the current leading hypothesis is that, upon binding, C1q undergoes an intramolecular activation mechanism (via conformational shifting/twisting of its two 'halves') that ultimately induces the auto-activation of the previously stable and otherwise inactive proC1r dimer into its active proteolytic form. In turn, C1r then functions as the activating protease of the ProC1s dimer and cleaves it to form the active C1s protease, thereby completing the first major step of the classical activation pathway.[47,87] Activated C1s can then proceed as a protease in its own interactions with the next two major components of the classical pathway, C4 and C2. Cleavage of C4 results in two fragments, C4a and C4b, the latter of which becomes the first component of the pathway to physically anchor to the target membrane via covalent attachment. Active C1s also interacts with C2, which is again cleaved into fragment products C2a and

C2b. The resulting C2a fragment is then able to combine with the attached C4b anchor, resulting in yet another key protein complex known as C3 convertase (C4bC2a). The activation of C3 (by C3 convertase) is the crucial point at which all three activation pathways of the complement cascade converge.[89] From this point onward there are still, however, a number of reactions involving multiple cascade components that must take place before successful MAC formation, as illustrated in **Figure 1.4.2**.[90]

While there are many regulatory components of the cascade that present throughout various stages of complement pathway progression, there is only one molecule that is known to directly regulate the initiation of the classical cascade itself – the aptly named C1-inhibitor (C1-INH).[91] C1-INH is a member of the serpin family of proteins and is present in serum at roughly 4-5X the molar abundance of C1r and C1s, and has been observed to act quickly as a potent suicide inhibitor of both proteases, limiting the estimated half-life of activated C1 to less than 15 seconds under physiological conditions.[92] Again, while much of the nuance underlying the activation and inactivation mechanisms of C1 remains unclear, it is believed that C1-INH does impart an added degree of stability to the inactive C1 molecule and, upon C1 activation, the inhibitory mechanism of C1-INH induces dissociation of the $C1r_2C1s_2$ tetramer from C1q, enabling the free C1q molecule to continue participating in immune function at the targeted cell surface.[93,94] In Chapter 3, we explore some of the functional effects of C1-INH presence on the binding kinetics and stability of C1.

1.5 The Activation of Complement by IgM

In 1981, a conference entitled "*Protein Conformation as an Immunological Signal*" was held in Porto Venere, Italy, where researchers from around the world convened to share and discuss their findings throughout many key areas of the field of immunology. In a section dedicated to the structure/function relationship of IgG and IgM, Arthur Feinstein and his colleagues outlined their newly developed concept of a particular IgM conformational change, mediated by the physical distortion

derived from multivalent surface binding, as the means by which the antibody recruits and activates the initiating molecular complex of the classical complement cascade, C1.[48] In this proposed surface bound conformation, which they termed the "staple" conformation based on the profile of the molecule observed in early electron micrographs (**Figure 1.5.1**), the Fab domains of IgM were bent down and away from the central Fc core of the Cμ3/Cμ4 Fc region in order expose the binding site(s) for C1q, which is the primary molecule of C1 that makes direct contact with antibody immune complexes.[95,96] In comparison, the C1q binding site on IgG was already understood to be fully exposed natively and, while IgG is also capable of activating complement, its monovalent nature demands precise Fc multimerization at the antigenic surface and is therefore not the most efficient means of activating the cascade; it is currently thought that upwards of six individual IgG molecules may be required to locally coordinate/aggregate their Fc regions into a combined multivalent display of C1q binding sites.[48,69,97,98] In further demonstration of this principle, mutations of the C-terminus that induce the self-assembly of multiple IgG molecules into hexameric polymers (via addition of the Cμ4 μ-tailpiece; more details on this region are discussed in Chapter 4) have recently been demonstrated to greatly enhance complement activation by IgG. In fact, mutant IgG hexamers – even without their Fab domains attached – were also shown to bind and activate the classical cascade through the sole display of their hexameric Fc domain 'rings', even in the complete absence of antigen.[27,70,99] IgM, on the other hand, was understood to function as a "*built-in aggregate*" that comes prepackaged to activate C1 on a unimolecular, 1:1 molar basis.[69,97,100]

While an exact geometric match to C1q globular head spacing would theoretically be the most ideal antibody architecture, there's plenty of evidence throughout the literature indicating that C1/C1q does not actually require a strictly symmetrical hexameric display in order to become activated. The IgM pentamer was recently observed to display a gap (~50°) where the J chain acts as placeholder for the sixth monomeric subunit, which adds further bias to the rest of the arm spacing and causes them to

deviate from perfect hexameric radial symmetry (60°).[45] Additionally, IgMp formed in the absence of the J chain (J^- IgMp) displayed a smaller gap (~40°) between monomers, which could help to explain the differences observed in activity between the two forms of IgMp – a smaller gap between monomers effectively pulls the Fc domains closer together, and the monomers are packed as tightly as they can be within IgMh which has drastically more activity than either J^+ or J^- pentameric forms. Considering that IgG hexamers are also able to efficiently activate complement, it begs to question, then, for the case of IgM: *How many 'arms' are required to bind antigen in order to activate C1*? *Is staple a local quality rather than a distinct, globally-concerted conformation?*

Much has changed regarding our general understanding of antibody biophysics over the 40+ years that have followed since the initial staple proposal from Feinstein, but the mechanistic details that underlie the activation of complement by IgM remain almost as poorly understood today as they were in the 1980s. The longstanding hypothesis has continued to be that, in order to activate complement, IgM must first bind multivalently to a sufficiently sized antigen or antigenic surface in order to induce a prerequisite conformational change that effectively uncovers the necessary binding site(s) for C1q, transitioning the molecule from an otherwise C1-inert to C1-activating form (**Figure 1.5.2**). While this prerequisite conformational change has now been generally agreed upon by those in the field, the literature is nonetheless full of decades' worth of disagreement and confusion, stemming from a wide array of phenotypic reports with very little structural knowledge. Due mainly to the combination of 1) technical limitations and 2) a severe lack of structural understanding, the mechanistic forces that drive the conformational change, along with the involved subdomains and identity of the residues that ultimately comprise the C1q binding site(s) on IgM, still remain largely unknown.

1.5.1 The Original Driving Force Debate: Allosterism vs Physical Distortion

From the very start, two major camps of thought were formed by early immunologists who sought to explain the driving force behind the prerequisite conformational change of IgM into its complement-active form. One side concluded that allosteric changes, caused by multivalent Ag binding at the Fabs, was sufficient to drive a mostly planar molecule towards the acutely distorted staple conformation.[100–102] The other side insisted that allosterism was not enough, and instead concluded that an induced physical distortion, created by the act of IgM binding multivalently to a fixed antigenic surface, was the only way to explain such a dramatic and global conformational change.[48,95,103,104]

While much of this debate stemmed from an understandable lack of information, the confusion can likely be distilled down to interpretations regarding the relative complement activity of solution bound IgM. Unbound (apo) IgM does not demonstrate any appreciable amount of C1 activity in solution, although some of the earliest reports concluded it to be a weak activator with high micromolar (~$1E^5$ M) affinity.[105,106] The literature mainly presents only three general approaches for measuring the activation of the classical cascade, all of which still remain objectively difficult to control for and interpret with accuracy: 1) C1 'fixation' (i.e. binding) assays, 2) complement-dependent cell lysis assays (measured by EC_{50} values), and 3) ELISA-based detection of various downstream cascade components.[50,73,97,99,107–110] The complement cascade is immensely complex and is also notoriously difficult to study as it tends to readily self-activate, which is an almost paradoxical situation given the destructive power of the cascade itself. In fact, self-activation of complement is so inherent to the system that it warrants its own separate pathway – the alternative pathway – which is indeed initiated by spontaneous self-hydrolysis of C3, with further contributions from C1 auto-activation of C1r and C1s proteases.[111–113] While there are endogenous factors that act to largely repress the activation of the alternative pathway itself, the other two initiation pathways are also tightly regulated by a number of inhibitory components and mechanistic

feedback loops, as unrestricted complement activity would be detrimental to healthy host cells and tissues.

1.5.2 The Structural Features of IgM Staple Formation

Regardless of the driving force(s) that trigger the mechanism of conformational change, the next major endeavor is understanding the structural features that define the biophysical adoption of the staple conformation itself. This includes identifying the local regions and subdomains that are significantly involved in the transition, along with garnering any potential mechanistic insight towards the physical requirements of adopting the multivalent, surface bound form of IgM staple. In the case of IgG, the length and composition of the proline-rich hinge region largely dictates the ability of a given isotype to activate the classical complement pathway. The hinge increases flexibility and confers an enhanced range of motion to the Fab domains, which must be moved out of the way to provide steric clearance for the C1q globular heads to bind at the periphery of the Cγ2 domain.[98] The N-terminal segment of the hinge confers Fab-Fab flexibility along with rotational freedom, and its length is often a strong predictor of complement activation. IgG3 has a significantly extended hinge region (~60 residues) and is accordingly the most flexible and potent IgG activator of complement, followed by IgG1 with a moderately sized (~15 residues) hinge region and degree of flexibility. At the other end of the spectrum are IgG2 and IgG4, both of which have shorter hinges with more restrictive sequence compositions and display little to no complement activity.[38–40]

Homology modeling of IgM with IgG reveals analogous domain pairings that match the Cγ1, Cγ2, and Cγ3 heavy domains of IgG to the Cμ1, Cμ3, and Cμ4 heavy domains of IgM, respectively.[54,114] Since IgM lacks a formal hinge region altogether, it follows that the unmatched Cμ2 domain would either contain a hinge-like region or act as the hinge itself to provide flexibility between the Fabs and Fc core. From the classical, low resolution negative-stain EM studies to the modern, high resolution cryo-

EM models, it's clear that IgM is able to globally adopt a pronounced dome-like structure when bound multivalently to a surface, and the bending of the Fabs down and away from the Fc core is so severe that its 2D profile resembles a metal staple, with sharp (~90°) angles formed between the raised Fc platform and the opposing arms of the surface bound Fab domains.[37,38,47,51,95,98,115]

While it would seem reasonable, then, to predict that the Cμ2 domain acts at least to some extent as a functional hinge, there are no structural models available that directly confirm this hypothesis. The only other mammalian antibody that is composed of four heavy chain domains is IgE, which has a distinctly bent and compacted Cε2 domain that exhibits relatively low segmental flexibility.[15,39,44] However, IgM displays a significantly higher degree of segmental flexibility than would be predicted based on its lack of a formal hinge region as defined by the properties of IgG. Furthermore, the X-ray crystal structure of the IgE Fc domain was published in 2002, and subsequent predictions of IgM based on its IgE homology would naturally have assumed a similar restriction in segmental flexibility. In retrospect, the fact that an intact structure of IgE was ever resolved in the first place should have been a clue as to the degree of flexibility that IgM is actually capable of, since to date there still has yet to be a highly resolved, fully intact IgM structure.

1.5.3 Identification of the C1q binding site(s)

Despite the confusion surrounding the driving force and structural features of staple formation, many groups have tried to circumvent those issues and instead look to determine the identity of the C1q binding site(s) on IgM. The establishment of C1q binding sites located in the Cγ2 domain of IgG led many groups to look for similar motifs and charged residues within the homologous Cμ3 domain of IgM.[71,116] In 1994, a detailed and systematic mutational study of charged residues in the Fc domain of IgM was published by Arya et al., based on earlier work done by Shulman and others that implicated a handful of charged Cμ3 residues that had a severe impact on complement activity as measured by C1 fixation and

cell lysis assays.[117–120] Studies that employ mutagenesis to investigate protein:protein interactions have often led to promising results, and in the case of IgG and other proteins that act in a monomeric nature the approach is often well-suited.[121–123] However, caution should be taken when considering the results and implications of mutational studies on proteins like IgM that are constructed in a polyvalent manner and function multivalently. Of the 16 total residue mutations presented by Arya et. al., 12 were found to disrupt the polymerization state by 20% or more, with 10 found to severely impact both the polymerization state and subsequent complement activity. The remaining mutations had little to no effect, except for one particularly intriguing case in which the mutation greatly *enhanced* complement activity.[117] Recently (2019), Sharp et. al. used cryo-EM with tomography to elucidate structures of C1 bound to IgM that were in-turn multivalently bound to a surface by all of its Fabs, adopting the expected dome-like conformation (**Figure 1.5.3**). Their data clearly show association of the C1q globular heads to the periphery of the Fc platform in regions that correspond to the five (or six) surface-exposed DLPSP residues (432-436) within the Cμ3 domain, which has remained the putative C1q binding site for the past 30 years.[47,51,118] A basic 3D model of the surface bound conformation was also proposed in which the Fabs are sharply bent down and away from the Fc core platform at the Cμ3/Cμ2 junction, further supporting the original staple conformation hypothesis along with the notion that the Cμ2 domain acts as a hinge to provide steric clearance for the C1q globular heads.

Based on IgG homology, mutational studies, and recent EM evidence, it would seem that the Cμ3 likely houses many of the residues critical for complement activation. However, at the same time that much of the classical work on Cμ3 was being performed, another significant group of researchers were focused on the Cμ4 domain of IgM, and were <u>also</u> able to provide evidence of a C1q binding site located somewhere within it.[124–128] In fact, the evidence was so strong that Hurst et. al. delineated the relationship to a simple protein fragment, composed of Cμ4 residues 468-491 fused with 515-546, and successfully demonstrated that it was able to bind/fix C1 to a significant degree.[127] These reports were

further corroborated by Johnson and Thames in 1976, when they used synthetic linear peptide fragments that mimicked the sequences of various portions of the Cμ4 domain, in conjugation with human-derived complement, to demonstrate that arranging arginine and tryptophan residues into an order that mimicked R491 and W488 of the Cμ4 domain produced peptide fragments capable of complement activity via fixation and hemolytic assays.[128] Finally, in 1978, Bubb & Conradie explicitly reported the inability of the Cμ3 domain itself to interact with complement, and instead showed that only portions of the Cμ4 domain were capable of binding to C1. The authors also emphasized that fixation of C1 may differ entirely from full-scale activation, which speaks to the aforementioned complexity of accurately measuring and interpreting complement activity.[124]

While the majority of classical studies were understandably focused towards the Cμ3 and Cμ4 domains, there were also a few key studies done by groups looking at the role of the Cμ2 domain within the context of complement activation. Although no evidence was gathered to suggest the potential location of a C1q binding site within the Cμ2 domain itself, in 1981 Siegel and Cathou determined that disruption of native Cμ2 structural integrity had a directly negative impact on complement activity, and even demonstrated the activity to be recoverable so long as the Cμ2 domain was allowed to correctly re-fold.[129] They were also able to conclude that binding of multiple Fab domains to the same Ag particle was a critical step in complement activation. With Fab spacing largely dictated by the hinge region of other antibodies, they reasoned that the Cμ2 domain was critical for maintaining the correct Fab spacing of surface bound IgM and that it also could facilitate C1 binding by acting as a hinge, resulting in the overall stabilization of a multivalently bound conformation.[103]

1.5.4 The Role of Antigen Presentation

As evidenced by the many observations of its global and dramatic conformational change, it seems apparent that IgM is an inherently flexible molecule capable of binding to a surface with maximal

valency, which requires bending of F(ab)'$_2$ units (Cμ2 + both Fabs) down and away from the plane of the Fc platform, with even further adjustments needed for individual Fab lobe positioning and spatial accommodation, as evidenced by recent Cryo-EM studies (**refer to Figures 1.5.1 & 1.5.3**). Furthermore, given the apparently drastic increase in segmental flexibility that, by all recent accounts, appears to originate near the Cμ3/Cμ2 junctional border and continues to preclude structural determination by modern Cryo-EM techniques (**refer to Figure 1.2.3**), it's plausible that transient, local staple formation within individual arms of IgM could be within the natural energy landscape of Fc/Fab flexibility and could therefore be reasonably predicted to occur randomly as a result of natural fluctuations in solution.[41,115] However, the probability that all of the arms of IgM would correctly fluctuate both transiently and simultaneously is low, and even if it did occur, the transitory conformation would most likely not be stabilized long enough for an immunologically meaningful interaction with C1.[41,44,98,115] Furthermore, the entire purpose of MAC formation would be effectively nullified, as there would be no prerequisite membrane in which to attack. This is where the role of surface Ag comes into play, as the Ag presentation itself will dictate the binding modality of IgM.[103,130–133] Provided sufficient density and spacing, we hypothesize that antigenic surfaces inherently act to stabilize and mediate the structural transitions required to form an *ensemble* of local/discrete staple conformations from concurrent surface binding of all the available arms within a single IgM molecule.

As previously mentioned, transient staple formation in solution bound IgM could likely be a major root-cause of the original driving force debate that arose from conflicting phenotypic reports throughout the classical literature, and could also be an explanation as to why solution state IgM was considered to have, as Feinstein put it: "[sic]…*only a single C1q binding-site, with a binding constant of* $5x10^5\ M^{-1}$*, leading to a bound half-life of only a few seconds*." [98] The probability that all, or even most of the arms of solution bound IgM would readily form staple in both a transient and immunologically-meaningful manner is probably quite low – evidenced further by the fact that many autoimmune

disease states, such as systemic lupus erythematosus (SLE), occur in part as a result of the spontaneous overactivation of complement.[134,135] Factors such as the temporal range of the interaction, the distance from the targeted membrane, and the degree of binding site valency that is ultimately displayed by surface bound IgG/IgM antibodies will greatly impact the net activation rate of complement, which is highly regulated by the body due to the inflammatory and highly destructive nature of the cascade and MAC formation, respectively.

Since antigen presentation dictates the binding conformation of IgM, and the binding conformation of IgM dictates subsequent C1 activation, it follows that a strong correlation should exist between antigen presentation and C1 activation. In 1979 Karush et. al. demonstrated that small, otherwise inactive monomeric Ag could be engineered into a bivalent display that successfully induced complement activity, so long as the epitope spacing was appropriate for the average distance between Fab domain pairs of IgM (~20 nm). Two years later, Borsos et. al. expanded on the results of Karush and others and posited three main categories in which an antigen bound IgM molecule could be placed – C1 activating (fixation and cell lysis), C1 sequestering (fixation without cell lysis), and C1 inert (neither C1 fixation nor cell lysis).[108] Manipulation of antigen density alone was enough to shift the population of each bound IgM category, effectively demonstrating that antigen presentation directly impacts the average surface bound conformation of IgM which, in turn, affects the average extent of C1 activity. Antigen presentation is an intrinsic variable that heavily impacts every study in which complement activity is measured, regardless of the method used. Investigating the relationship between Ag presentation and antibody-mediated effector functions still remains a technologically challenging endeavor. Nevertheless, new and interesting studies continue to emerge that shed further light on the nuanced effects that antigen spacing, surface curvature, and antibody flexibility have on complement activation.[130,131,133]

1.6 Dissertation Overview and Chapters Preview

The foundation that has been laid throughout our collective history of IgG-based research has paved the way for new forms of antibody therapies that could one day utilize the rest of the mammalian Ig classes. One of the most promising examples can be found today in the IgM-derived drug candidates that are currently spread throughout various phases of clinical trials. One of the most advanced candidates, Imvotomab (IgM Biosciences; Phase II), combines the extreme binding valency/avidity of the IgMp framework into a potent bispecific antibody indicated for the treatment of non-Hodgkin's lymphoma, by combining the deca-valency of ten anti-CD20 Fab domains with an anti-CD3 T-cell engaging subunit that is substituted in place of the J chain. However, because IgM is the most potent activator of the classical complement cascade, it's vital that an understanding of the structural and functional features of the complement-active form of IgM be well-established, as complement activation may not always be a desired outcome of certain biotherapeutics and disease state indications. Many attempts have been made to uncover the structural details surrounding the formation of the active staple conformation, but the same physical traits that give IgM its unique properties and relationship with the immune system have also earned it a notorious reputation for being technically difficult to handle and investigate experimentally. Consequently, we decided to approach some of the fundamental questions surrounding the IgM-mediated activation of complement through modern structural and functional techniques that are highly amenable to the investigation of large and flexible glycoproteins. The primary objective of the work presented within the following chapters of this dissertation was to identify the major structural features and functional characteristics that define the complement-active staple conformation of IgM.

In Chapter 2, background work is presented on the development of a low-cost, automated LC-MS system designed to improve both the throughput and consistency of decoupled hydrogen/deuterium

exchange – mass spectrometry (HDX-MS) samples, which is the primary technique used throughout the structural analysis of complement-active IgM presented in Chapter 4. HDX-MS is capable of providing unique insights into complex biological systems that are often difficult to study by other techniques. However, due to arduous sample handling requirements, automating HDX experimentation for higher throughput requires specialized equipment. While recent advances have started to normalize the hardware and instrumentation needed to automate some forms of HDX-MS sample preparation and analysis, the complexity and cost of the hardware needed is often prohibitive, and several proteins of interest along with certain types of HDX approaches remain incompatible with automated workflows and require manual sample handling that greatly limits experimental throughput. In this chapter we provide background into the HDX methodology itself, along with the design details and performance of the LC-MS systems used to structurally analyze the complement-active form(s) of IgM.

In Chapter 3, the phenotypic traits of complement-active IgM are investigated using a combination of novel Biolayer Interferometry (BLI) and C1 activity assay approaches in order to differentiate the kinetic and functional traits that define staple IgM from that of less active and inactive forms of Ag bound and unbound forms, respectively. The primary goal of BLI was to establish a reliable kinetic profile of the 'nominal', or most efficient form of C1-activating surface bound IgM that we could form consistently within our surface Ag and C1q/C1 binding experiments, which we hypothesize to be reflective of the most optimal conformational arrangement of the staple ensemble that we could achieve consistently throughout the course of this project. In order to verify biological relevance and further compare the activation potential of the various forms of Ag bound IgM that were investigated throughout this project, we developed a simplified and relatively straightforward activity assay that measures the conversion rate of ProC1s into its activated form, which is caused directly by C1 activation and is the primary objective of the first major step of the classical signaling pathway. Combined with the binding characterizations determined by BLI, the results from our C1 activity assays help to provide a

critical functional context that can be further tied to the unique structural features of the various IgM binding modalities, as determined by HDX-MS.

In Chapter 4, the structural features of IgM are investigated in detail by HDX-MS using the instrumentation outlined in Chapter 2, and the phenotypic traits determined in Chapter 3 are also used to complete the structure/function relationship that defines the complement-active, multivalent surface bound conformation of the antibody. Multiple HDX-MS experiments, each with their own uniquely iterative approach and measured endpoints, are ultimately combined in order to confidently assign major structural features that define staple IgM and provide additional insights into the mechanistic details that underlie the required structural transition and exposure of potential C1q binding sites. Finally, a summary of the major conclusions and future directions of the project are presented in Chapter 5.

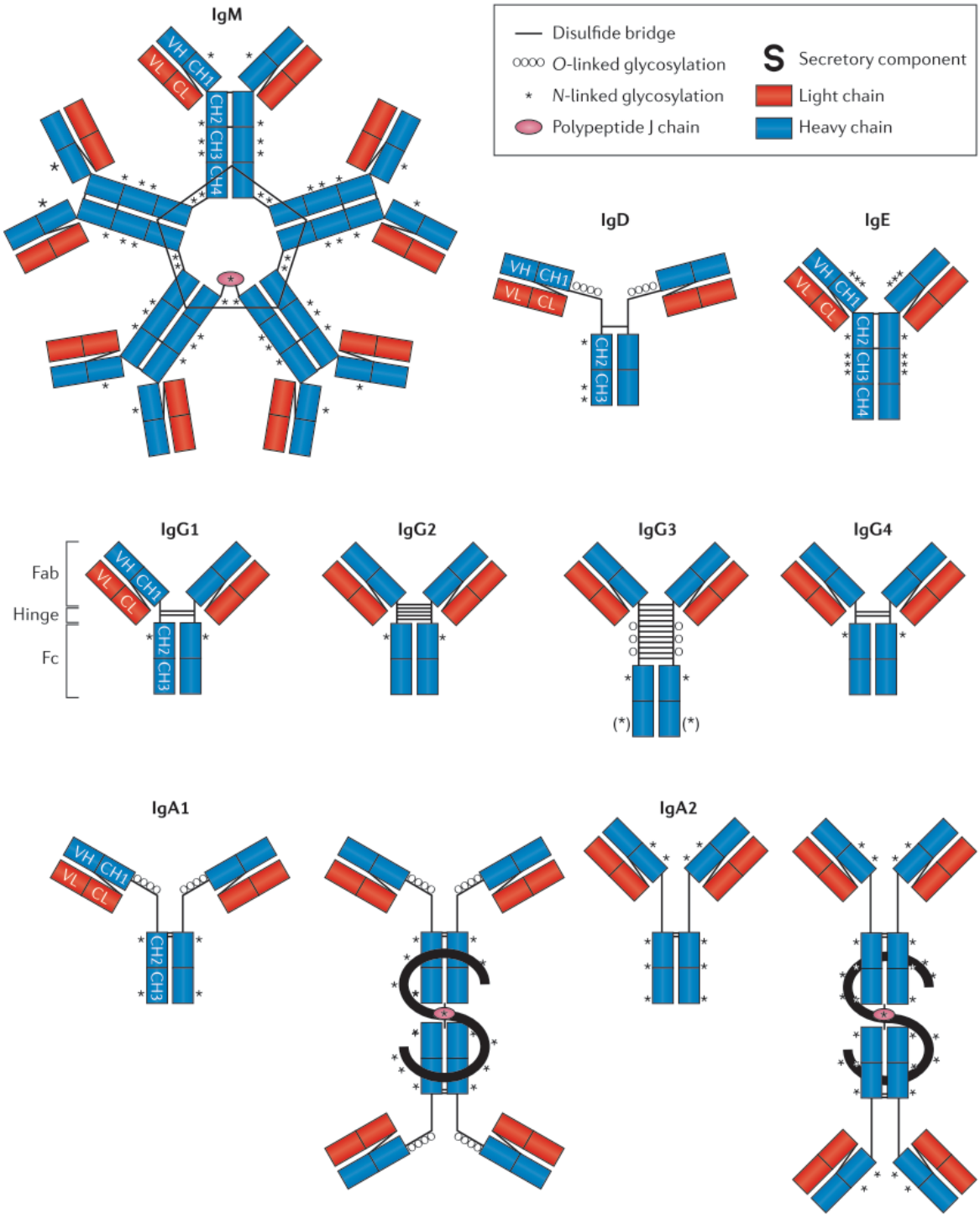


Figure 1.1.1 | The mammalian antibody system. ***Adapted from Nature Reviews: Immunology (2018)***

The basic structural architecture of the five mammalian antibody isotypes – IgA, IgD, IgE, IgG, and IgM – are displayed along with the indicated structural features. The hinge region that distinguishes all four primary IgG isoforms has long been considered to exist only within antibody classes composed of three primary heavy domains (IgA, IgD, and IgG), while those with four-domain Fc architectures (IgE and IgM) lack the classically defined hinge region entirely.

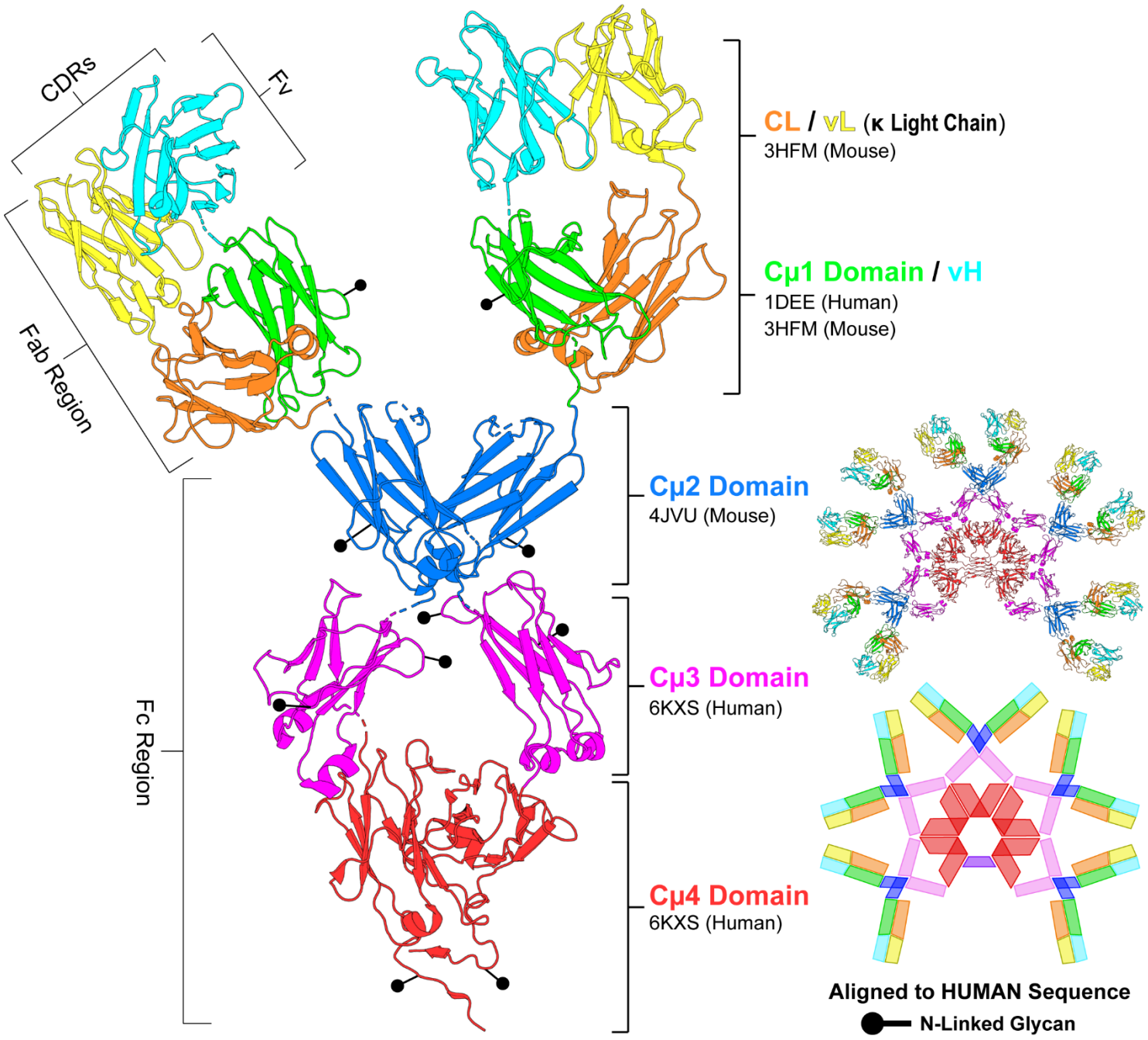


Figure 1.2.1 | Structural model of the IgM monomeric subunit

The color-coded IgM domain arrangement of a single monomeric IgM subunit, or 'arm' that is polymerized to form the pentameric (right) and hexameric structure of secreted IgM isoforms are displayed. The relevant PDB IDs of the domain structures used for modeling are listed. The structure, primary sequence (see Fig.1.2.2), and relative locations of conserved N-linked glycans are aligned to the canonical sequence of the human μ heavy chain (P01871).

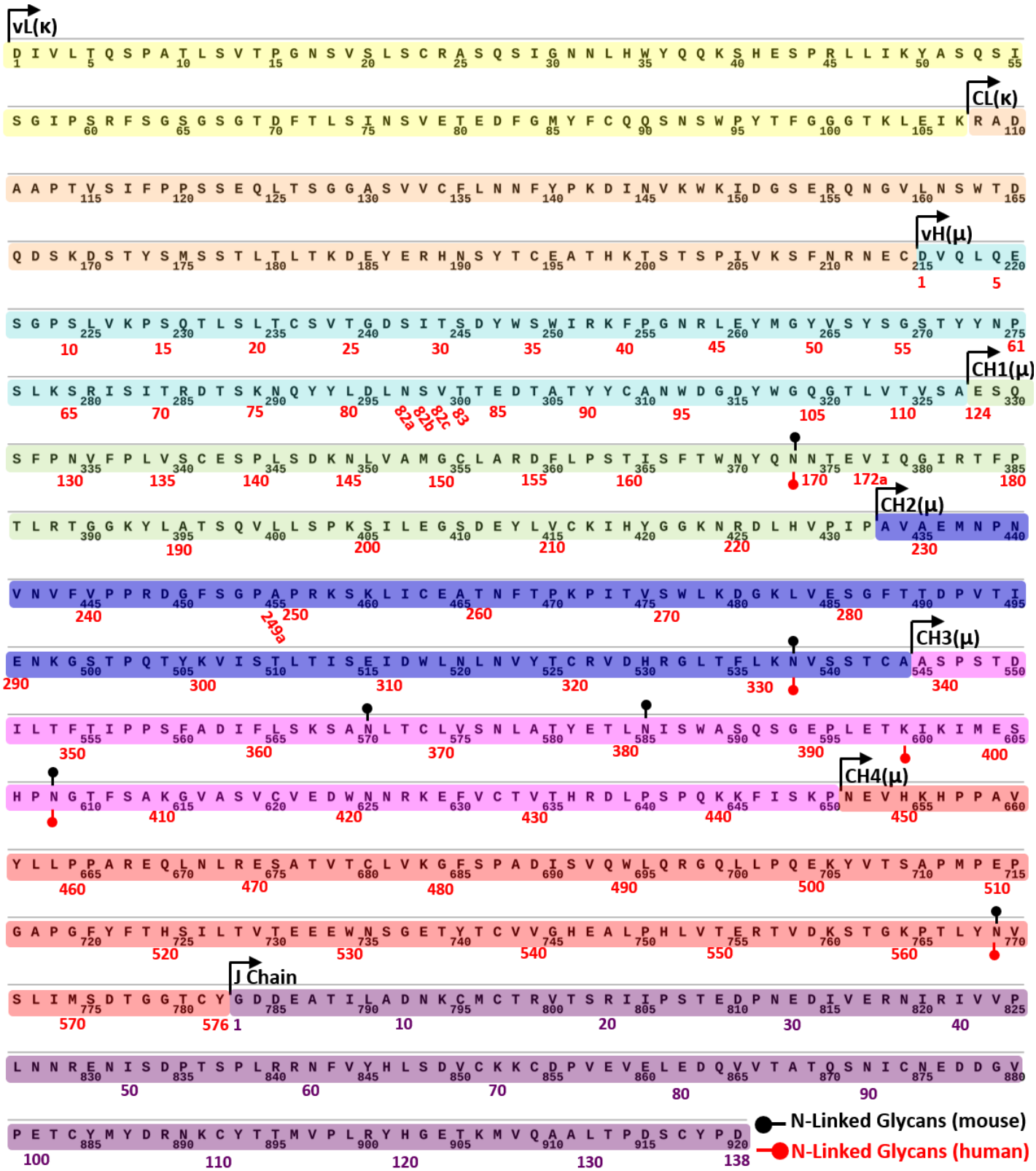
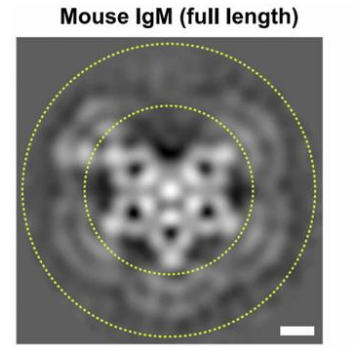
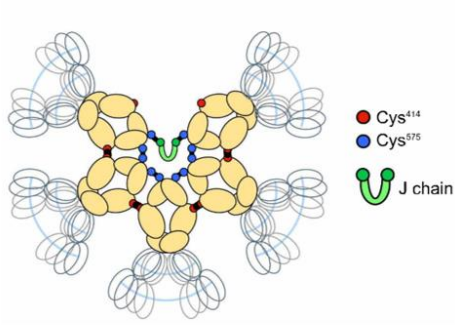


Figure 1.2.2 | Primary amino acid sequence of IgM MD4 (mouse)

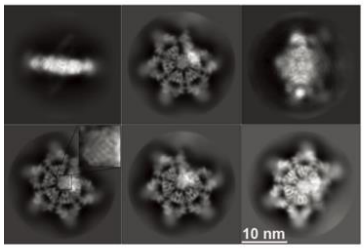
The continual sequence of IgM MD4 is provided, based on the sequences of the secreted forms of the canonical κ Ig light chain (P01834), the canonical murine μ heavy chain (P01872), and the canonical murine J chain (P01592). The red numbering beneath corresponds to the canonical human μ heavy chain (P01871) with the Kabat numbering scheme applied to the sequence starting at the heavy μ variable region (1-123).

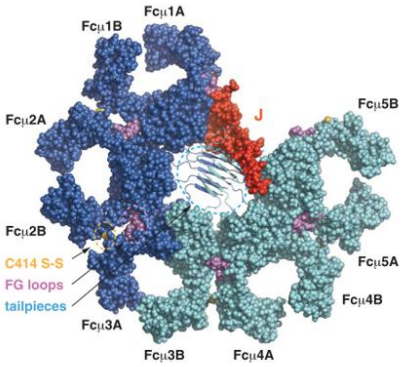




Hiramoto et. al. (2018)

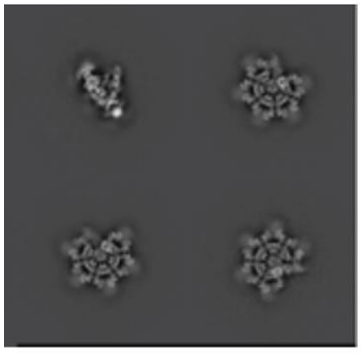
"The peripheral region corresponds to the Fab…which appears to move flexibly and structurally unlocked and thus could not be observed clearly."

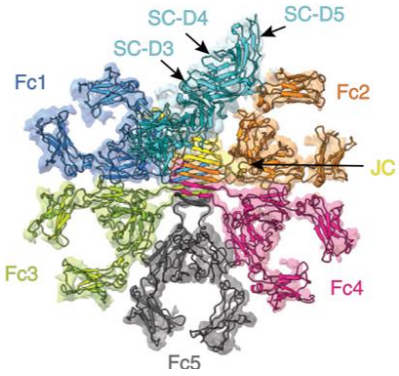




Li et. al. (2020)

"The IgM-Cμ2 domains were not modeled due to weak density."





Kumar et. al. (2021)

"Despite Cμ2 being included in the construct, the flexibility of this domain precluded structure determination."

Figure 1.2.3 | Impact of IgM flexibility on recent structural determinations by Cryo-EM

Adapted figures from recent high resolution cryo-EM studies show a similar loss of resolution towards the N-terminal region of the Cμ3 domains of polyvalent IgM, near the Cμ3/Cμ2 domain junctions. These recent attempts provide tangible evidence of a significant change in flexibility that originates in or near the periphery of the Cμ4/Cμ3 Fc core, which could be explained by the existence of a non-classical functional hinge located within that region. A detailed discussion of this structural feature is provided in Chapter 4.

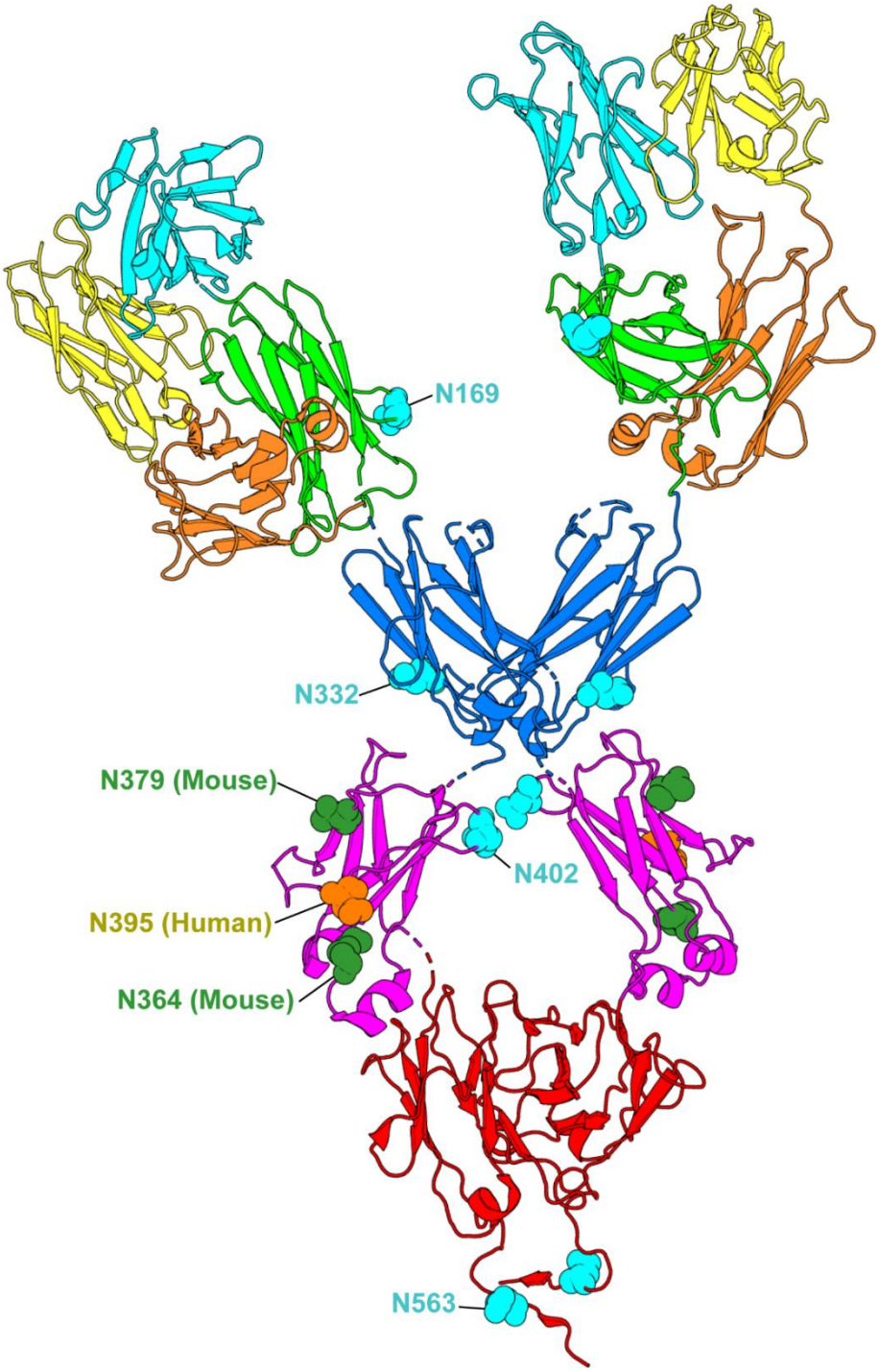


Figure 1.2.4 | Comparison of mouse and human conserved N-linked glycosylation

While some homology exists between conserved mouse and human IgM N-linked glycosylation sites (cyan spheres), there are key differences in both the number and location of some conserved N-linked glycans within the Cμ3 domain. Mouse IgM (green) exhibits an additional N-linked site within the Cμ3 for a total of six conserved N-linked glycans, while humans (orange) contain only two within the Cμ3 domain, for a total of five. See Chapter 4 for further glycosylation details observed during the structural analysis of IgM MD4, which is murine-derived.

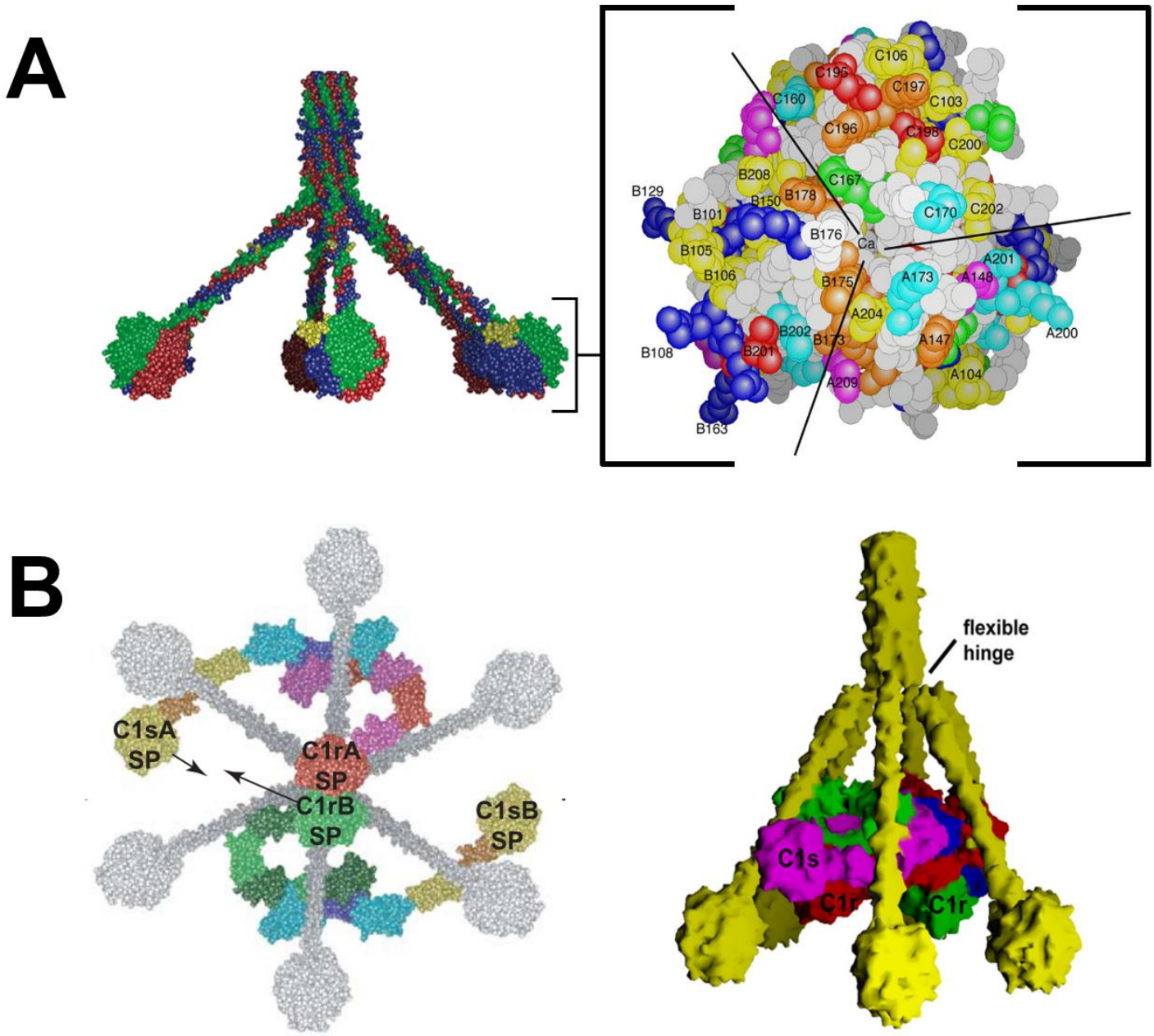


Figure 1.4.1 | Structural details of C1q and C1. ***Adapted from Gaboriaud et. al. (2004 & 2012)***

A) Structural model of C1q and the surface residue composition of its globular heads (inset; right) are shown, where a wide array of basic (blue), acidic (red), and hydrophobic (yellow) surface exposed residues are derived from the A, B, and C, polypeptide chain 'modules', as indicated. **B)** Proposed model of C1 complex arrangement of C1q, C1r, and C1s proteins that are known to assemble in a 1:2:2 ratio (respectively) to form C1, which is the initiating molecule of the classical complement cascade. C1q is the primary molecule that is thought to make direct contact with surface bound IgG and IgM immune complexes to initiate activation of the classical signaling pathway.

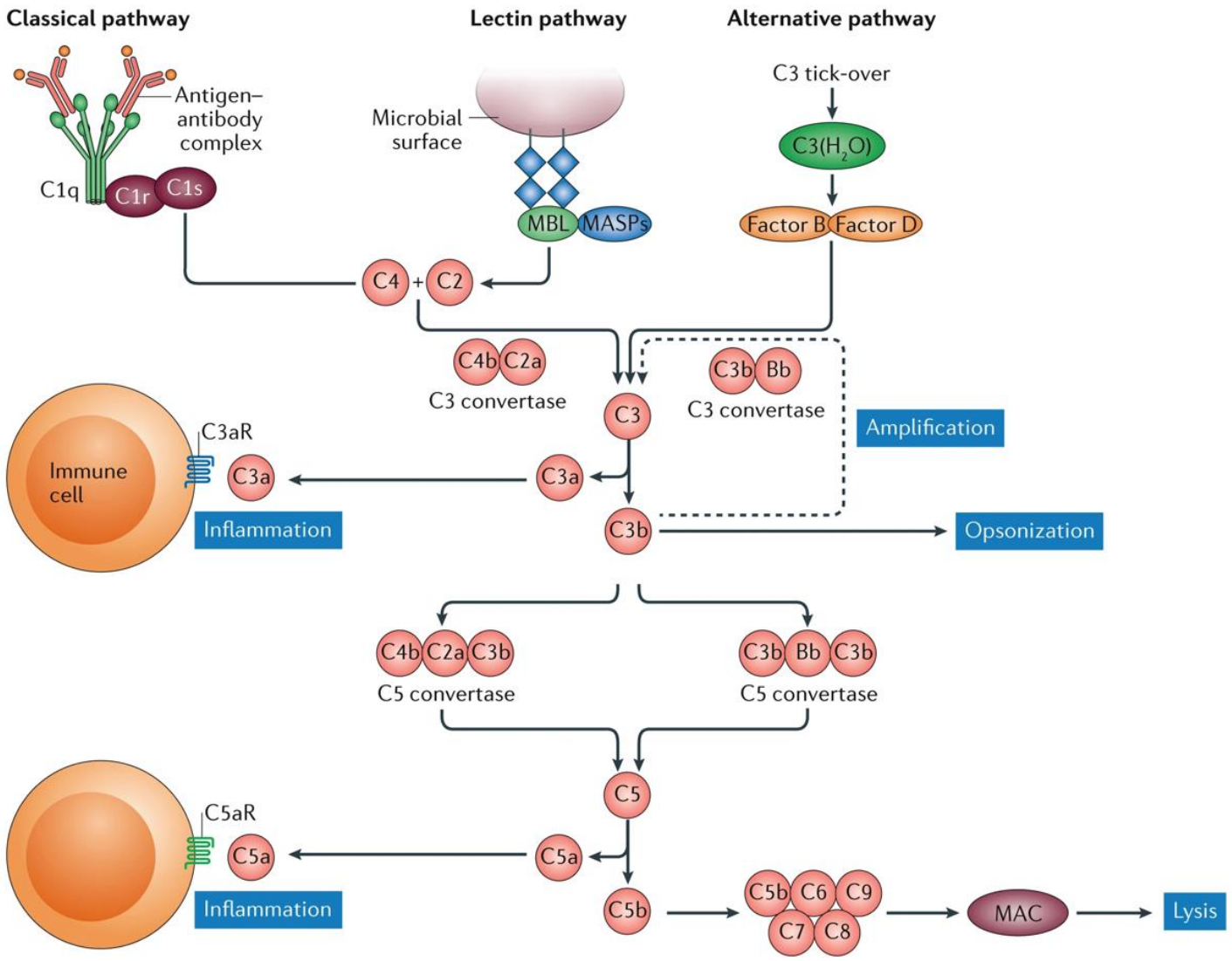


Figure 1.4.2 | Overview of the complement cascade. ***Adapted from Nature Reviews: Rheumatology (2017)***

There are three main converging pathways, each with its own unique form of initiation. The classical cascade pathway (left) is the only antibody-mediated pathway and serves as a bridge between the innate and adaptive immune systems, while the other two are confined to the machinery of the innate immune system only. The lectin pathway is initiated by a number of molecules that predominantly recognize and bind to pathogenic carbohydrate moieties, while the alternative pathway is reserved for self-hydrolysis of C3, which is further augmented by the auto-activation of C1. All three initiating pathways merge at the formation of C3, and the ultimate goal of the cascade is the formation of a viable membrane attack complex (MAC), which is physically anchored to the targeted cell membrane and destroys the cell via pore formations that cause total cell lysis. Each major step of the cascade also results in the propagation of pro-inflammatory responses, which lowers the local pH and also serves to recruit a variety of other immune cells, such as macrophages.

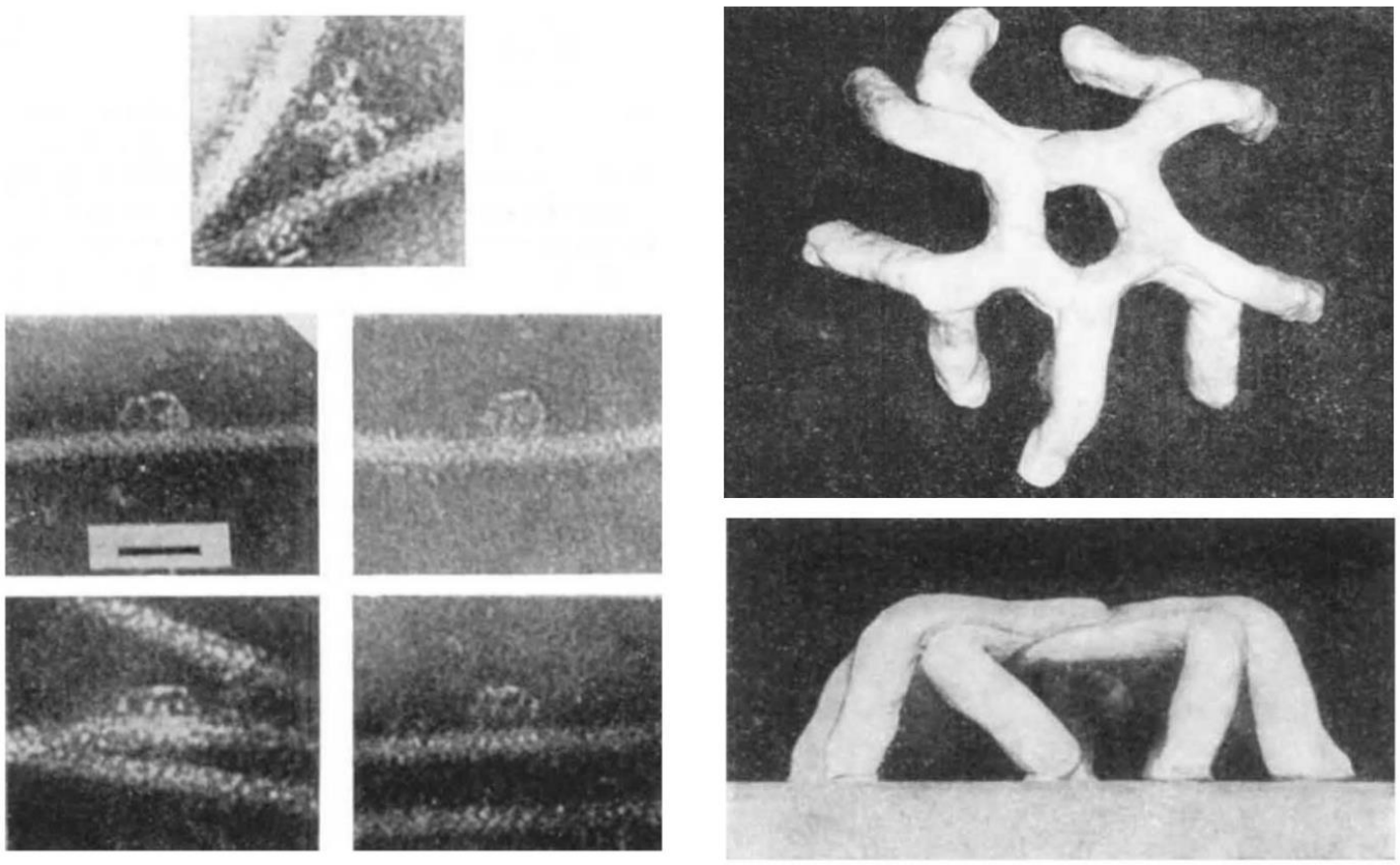


Figure 1.5.1 | The IgM "staple" conformation. ***Adapted from Feinstein & Munn (1969)***

Early negative-stain EM images (left) show surface bound IgM in both a crosslinked/planar form (top), along with the multivalent surface bound "staple" conformation (bottom four panels). Three-dimensional clay models of pentameric IgM staple (right) are also shown. In order to activate complement, IgM arms are predicted to fold sharply along the Cμ3/Cμ2 junction at the periphery of the Fc platform formed by the Cμ3 and Cμ4 domains, creating a two-dimensional profile that resembles the distinctly sharp (~90°) angles of a metal staple when bound multivalently to an antigenic surface, which is thought to be required for exposure of the C1q/C1 binding site(s).

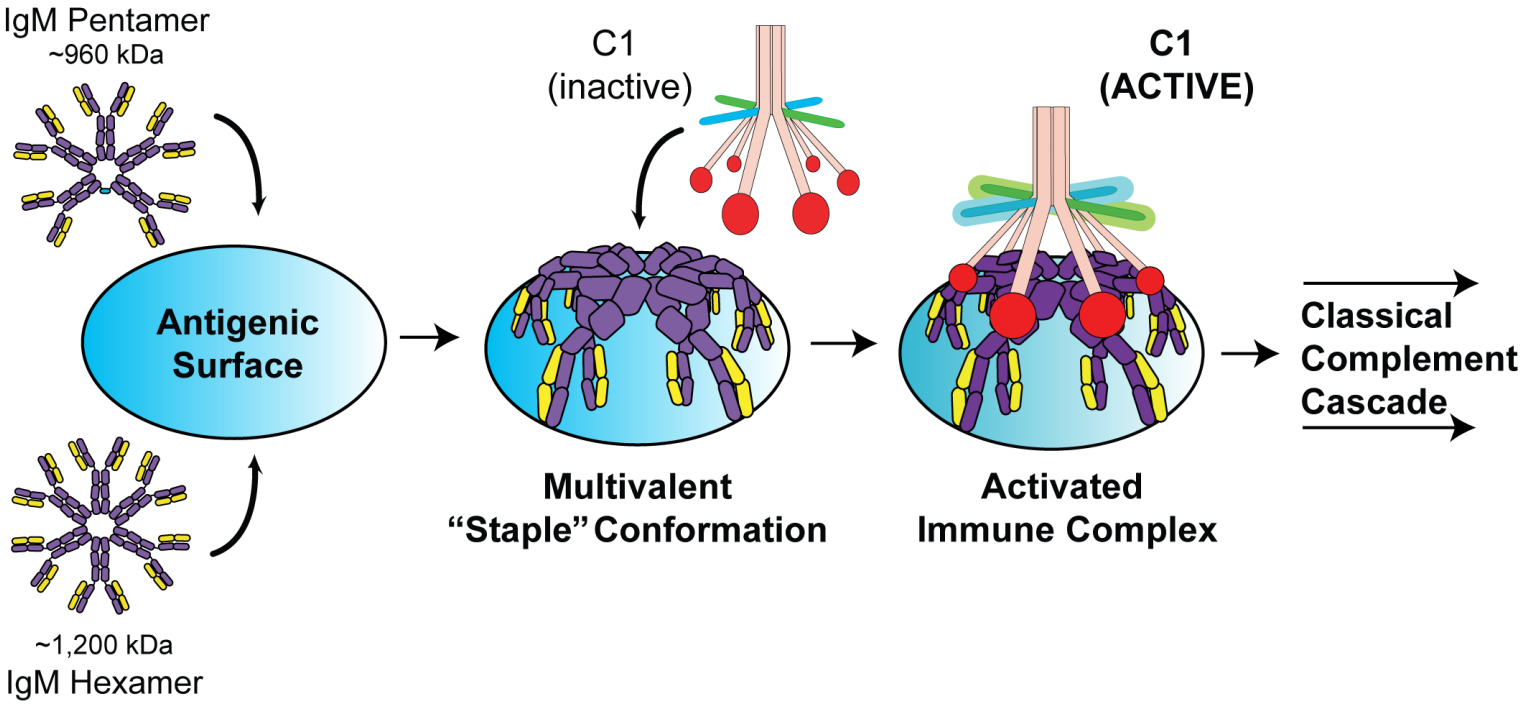


Figure 1.5.2 | The IgM-mediated activation of complement

The multivalent binding of pentameric (IgMp) and/or hexameric (IgMh) isoforms of IgM to a sufficiently dense Ag surface is thought to be required to adopt the complement-active 'staple' conformation, which results in the presentation of residues that comprise the C1q binding site(s) of C1. It is believed that widening/spreading of the hexameric arms of C1q leads to the mechanically based auto-activation of the C1r serine protease, which directly cleaves proC1s into its active proteolytic form, thereby completing the initiating step of the classical cascade pathway. From this point onward, active C1s protease will continue to interact with additional complement components in order to complete the next stage(s) of the classical pathway (refer to overview of the complement cascade illustrated in Fig. 1.4.2)

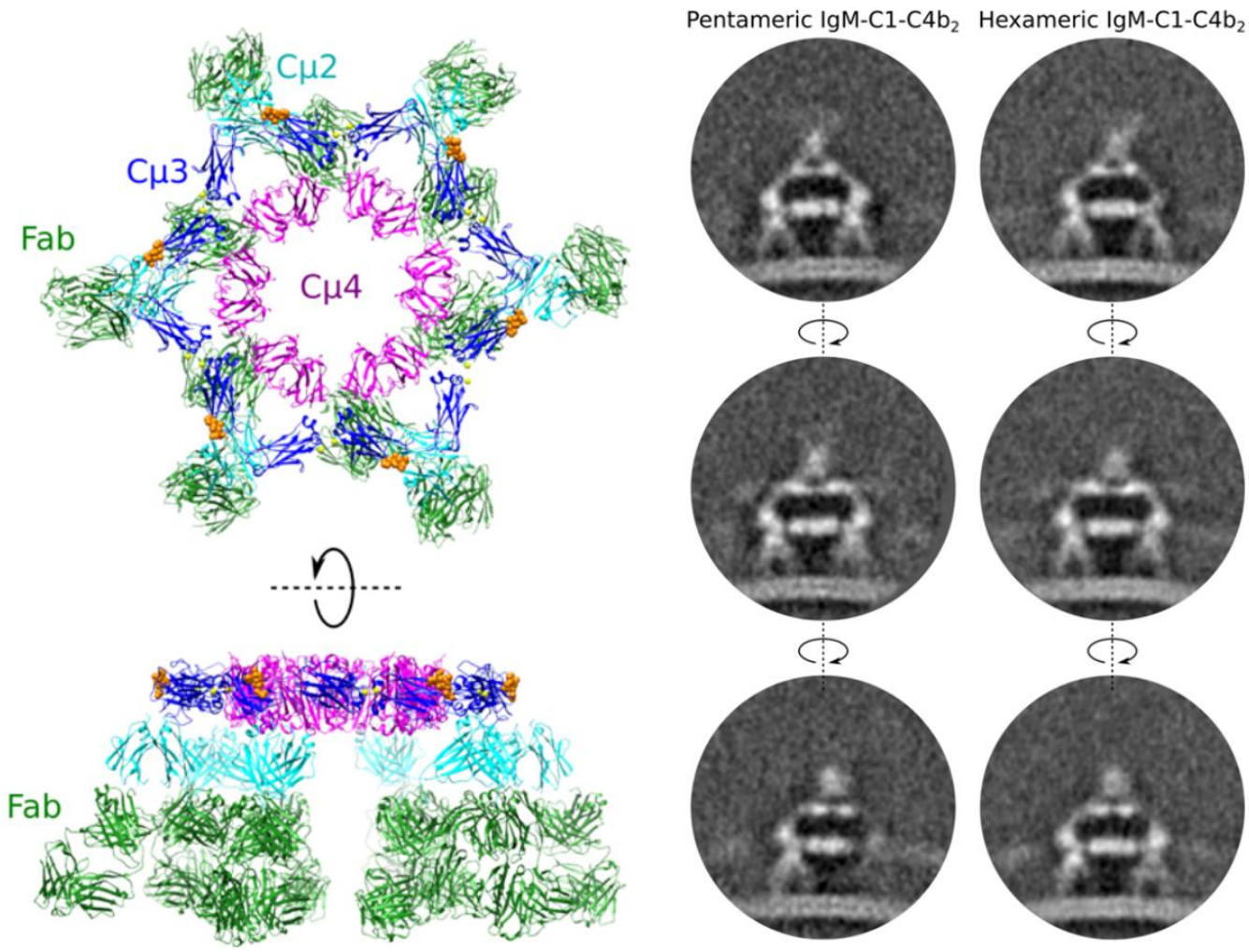


Figure 1.5.3 | 3D model of staple derived from Cryo-EM with Tomography. ***Adapted from Sharp et. al. (2019)***

In the predicted three-dimensional model (left), the Fab domains of surface bound IgM are twisted such that one Fab lobe is tucked slightly underneath the Fc platform in order to accommodate surface binding with maximal valency. The model was derived from data generated by cryo-EM with tomography that show profile slices of C1 bound to the top of the IgM pentamer and hexamer Fc core platform (right), with the globular heads of C1q associated near the Cμ3 domain periphery, along with additional densities of complement component C4b which is downstream of initial C1 activation.

1.7 References

1. Lu, R. *et al.* Development of therapeutic antibodies for the treatment of diseases. *J. Biomed. Sci.* **27**, 1–30 (2020).
2. Kaplon, H. *et al.* Antibodies to watch in 2023. *MAbs* **15**, 1–40 (2023).
3. Lu, L. L., Suscovich, T. J., Fortune, S. M. & Alter, G. Beyond binding: Antibody effector functions in infectious diseases. *Nat. Rev. Immunol.* **18**, 46–61 (2018).
4. Tiselius, B. Y. A., D, P. & Kabat, E. A. AN ELECTROPHORETIC STUDY OF IMMUNE SERA AND PURIFIED ANTIBODY PREPARATIONS. *J. Exp. Med.* **87**, 119–131 (1939).
5. Goulet, D. R. *et al.* Kinetic mechanism of controlled Fab-arm exchange for the formation of bispecific immunoglobulin G1 antibodies. *J. Biol. Chem.* **293**, 651–661 (2018).
6. Laursen, N. S. *et al.* Functional and Structural Characterization of a Potent C1q Inhibitor Targeting the Classical Pathway of the Complement System. *Front. Immunol.* **11**, 1–15 (2020).
7. Heremans, J. ., Heremans, M. & Schultze, H. . Isolation and Description of a few Properties of the B2a-Globulin of Human Serum. *Clin. Chim. Acta* **4**, 96–102 (1959).
8. Chen, K., Magri, G., Grasset, E. K. & Cerutti, A. Rethinking mucosal antibody responses: IgM, IgG and IgD join IgA. *Nat. Rev. Immunol.* (2020) doi:10.1038/s41577-019-0261-1.
9. Brandtzaeg, P. & Pabst, R. Let's go mucosal: communication on slippery ground. *Trends Immunol.* **25**, 570–577 (2004).
10. Sousa-pereira, P. De & Woof, J. M. IgA: Structure, Function, and Developability. *Antibodies* **2**, 1–30 (2019).
11. Wycoff, K. Secretory IgA Antibodies from Plants. *Curr. Pharm. Des.* **11**, 2429–2437 (2005).
12. Boross, P. *et al.* IgA EGFR antibodies mediate tumour killing in vivo. *EMBO Mol. Med.* **5**, 1213–1226 (2013).
13. Vladutiu, A. O. Immunoglobulin D : Properties, Measurement, and Clinical Relevance. *Clin. Diagn. Lab. Immunol.* **7**, 131–140 (2000).
14. Rowe, D. S. & Fahey, J. L. A New Class of Human Immunoglobulins II. Normal Serum Igd. *J. Exp. Med.* **121**, 185–199 (1965).
15. Wan, T. *et al.* The crystal structure of IgE Fc reveals an asymmetrically bent conformation. *Nat. Immunol.* **3**, 681–686 (2002).
16. Sutton, B., Davies, A., Bax, H. & Karagiannis, S. IgE Antibodies: From Structure to Function and Clinical Translation. *Antibodies* **8**, 19 (2019).

17. Ishizaka, K. & Ishizaka, T. Physicochemical properties of reaginic antibody. *J. Allergy Clin. Immunol.* **38**, (1966).

18. Gould, H. J. & Sutton, B. J. IgE in allergy and asthma today. *Nat. Rev. Immunol.* **8**, 205–217 (2008).

19. Croote, D., Darmanis, S., Nadeau, K. C. & Quake, S. R. High-affinity allergen-specific human antibodies cloned from single IgE B cell transcriptomes. *Science (80-.).* **1309**, 1306–1309 (2018).

20. Orengo, J. M. *et al.* Treating cat allergy with monoclonal IgG antibodies that bind allergen and prevent IgE engagement. *Nat. Commun.* **9**, (2018).

21. Waldenström, J. Incipient myelomatosis or essential hyperglobulinemia with fibrinogenopenia - a new syndrome? *Acta Med. Scand.* **117**, 216–247 (1944).

22. Black, C. A. A brief history of the discovery of the immunoglobulins and the origin of the modern immunoglobulin nomenclature. *Immunol. Cell Biol.* **75**, 65–68 (1997).

23. Li, Y. *et al.* Structural insights into immunoglobulin M. *Science (80-.).* **367**, 1014–1017 (2020).

24. Müller, R. *et al.* High-resolution structures of the IgM Fc domains reveal principles of its hexamer formation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 10183–8 (2013).

25. Graille, M. *et al.* Crystal structure of a Staphylococcus aureus protein a domain complexed with the Fab fragment of a human IgM antibody: Structural basis for recognition of B-cell receptors and superantigen activity. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5399–5404 (2000).

26. Padlan, E. A. *et al.* Structure of an antibody-antigen complex: Crystal structure of the HyHEL-10 Fab-lysozyme complex. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 5938–5942 (1989).

27. Pasalic, D. *et al.* A peptide extension dictates IgM assembly. *Proc. Natl. Acad. Sci.* 201701797 (2017) doi:10.1073/pnas.1701797114.

28. Melis, J. P. M. *et al.* Complement in therapy and disease. Regulating the complement system with antibody-based therapeutics. *Mol. Immunol.* **67**, 117–130 (2015).

29. Brewer, J. W., Randall, T. D., Parkhouse, R. M. E. & Corley, R. B. Mechanism and Subcellular Localization of Secretory IgM Polymer Assembly. *J. Biol. Chem.* **269**, 17338–17348 (1994).

30. Brewer, J. W., Randall, T. D., Parkhouse, R. M. E. & Corley, R. B. IgM hexamers? *Immunol. Today* **15**, 165–168 (1994).

31. Randall, T. D., Parkhouse, R. M. E. & Corley, R. B. J chain synthesis and secretion of hexameric IgM is differentially regulated by lipopolysaccharide and interleukin 5. *Proc Natl Acad Sci U S A* **89**, 962–966 (1992).

32. Oreste, U., Ametrano, A. & Coscia, M. R. On origin and evolution of the antibody molecule. *Biology (Basel).* **10**, 1–18 (2021).

33. Kaattari, S., Evans, D. & Klemer, J. Varied redox forms of teleost IgM: an alternative to isotypic

diversity? *Immunol. Rev.* **166**, 133–142 (1998).

34. Cho, Y. *et al.* Molecular characteristics of IgA and IgM Fc binding to the Fc(alpha)/(mu)R. *Biochem. Biophys. Res. Commun.* **345**, 474–478 (2006).

35. Azuma, Y. *et al.* Recombinant human hexamer-dominant IgM monoclonal antibody to ganglioside GM3 for treatment of melanoma. *Clin. Cancer Res.* **13**, 2745–2750 (2007).

36. Pettinello, R. & Dooley, H. The Immunoglobulins of Cold-Blooded Vertebrates. *Biomolecules* **4**, 1045–1069 (2014).

37. Parkhouse, R. M., Askonas, B. A. & Dourmashkin, R. R. Electron microscopic studies of mouse immunoglobulin M; structure and reconstitution following reduction. *Immunology* **18**, 575–584 (1970).

38. Roux, K. H. Immunoglobulin structure and function as revealed by electron microscopy. *Int. Arch. Allergy Immunol.* **120**, 85–99 (1999).

39. Roux, K. H., Strelets, L., Brekke, O. H., Sandlie, I. & Michaelsen, T. E. Comparisons of the Ability of Human IgG3 Hinge Mutants, IgM, IgE, and IgA2, to Form Small Immune Complexes: A Role for Flexibility and Geometry. *J. Immunol.* 4083–4090 (1998).

40. Roux, K. H., Strelets, L. & Michaelsent, T. E. Flexibility of human IgG subclasses. *J. Immunol.* **159**, 3372–3382 (1997).

41. Samsudin, F., Yeo, Y., Gan, S. K. & Bond, P. J. Not all therapeutic antibody isotypes are equal: the case of IgM versus IgG in Pertuzumab and Trastuzumab. *Chem. Sci.* 2843–2854 (2020) doi:10.1039/c9sc04722k.

42. Moh, E. S. X., Lin, C. H., Thaysen-Andersen, M. & Packer, N. H. Site-Specific N-Glycosylation of Recombinant Pentameric and Hexameric Human IgM. *J. Am. Soc. Mass Spectrom.* **27**, 1143–1155 (2016).

43. Leibiger, H., Kersten, B., Albersheim, P. & Darvill, A. Structural characterization of the oligosaccharides of a human monoclonal anti-lipopolysaccharide immunoglobulin M. *Glycobiology* **8**, 497–507 (1998).

44. Czajkowsky, D. M. & Shao, Z. The human IgM pentamer is a mushroom-shaped molecule with a flexural bias. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14960–5 (2009).

45. Hiramoto, E. *et al.* The IgM pentamer is an asymmetric pentagon with an open groove that binds the AIM protein. *Sci. Adv.* **4**, eaau1199 (2018).

46. Akhouri, R. R., Goel, S., Furusho, H., Skoglund, U. & Wahlgren, M. Architecture of Human IgM in Complex with Article Architecture of Human IgM in Complex with P . falciparum Erythrocyte Membrane Protein 1. *CellReports* **14**, 723–736 (2016).

47. Sharp, T. H. *et al.* Insights into IgM-mediated complement activation based on in situ structures

of IgM-C1-C4b. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 11900–11905 (2019).

48. Feinstein, A., Richardson, N. E., Gorick, B. D. & Hughes-Jones, N. C. Immunoglobulin M conformational change is a signal for complement activation. *Protein Conform. as an Immunol. Sign* 47–57 (1981).

49. Zlatarova, A. S. *et al.* Existence of different but overlapping IgG- and IgM-binding sites on the globular domain of human C1q. *Biochemistry* **45**, 9979–9988 (2006).

50. Gadjeva, M. G. *et al.* Interaction of human C1q with IgG and IgM: Revisited. *Biochemistry* **47**, 13093–13102 (2008).

51. Perkins, S. J., Nealis, A. S., Sutton, B. J. & Feinstein, A. Solution structure of human and mouse immunoglobulin M by synchrotron X-ray scattering and molecular graphics modelling. *J. Mol. Biol.* **221**, 1345–1366 (1991).

52. Kumar, N., Arthur, C. P., Ciferri, C. & Matsumoto, M. L. Structure of the human secretory immunoglobulin M core. *Structure* **29**, 564-571.e3 (2021).

53. Ramsland, P. A. *et al.* Crystal structure of a glycosylated Fab from an IgM cryoglobulin with properties of a natural proteolytic antibody. *Biochem. J.* **481**, 473–481 (2006).

54. Putnam, F. W., Florent, G., Paul, C., Shinoda, T. & Shimizu, A. Complete Amino Acid Sequence of the Mu Heavy Chain of a Human IgM Immunoglobulin. *Science (80-.).* 287–291 (1973).

55. Wang, H., Coligan, J. E. & Morse, H. C. Emerging functions of natural IgM and its Fc receptor FCMR in immune homeostasis. *Front. Immunol.* **7**, 1–7 (2016).

56. Rasche, L. *et al.* The Natural Human IgM Antibody PAT-SM6 Induces Apoptosis in Primary Human Multiple Myeloma Cells by Targeting Heat Shock Protein GRP78. *PLoS One* **8**, (2013).

57. Haury, M. *et al.* The repertoire of serum IgM in normal mice is largely independent of external antigenic contact. *Eur. J. Immunol.* **27**, 1557–1563 (1997).

58. Holodick, N. E., Rodriguez-Zhurbenko, N. & Hernández, A. M. Defining Natural Antibodies. *Front. Immunol.* **8**, 2–9 (2017).

59. Manson, J. J., Mauri, C. & Ehrenstein, M. R. Natural serum IgM maintains immunological homeostasis and prevents autoimmunity. *Springer Semin. Immunopathol.* **26**, 425–432 (2005).

60. Shaw, P. X. *et al.* Natural antibodies with the T15 idiotype may act in atherosclerosis , apoptotic clearance , and protective immunity. *J. Clin. Invest.* **105**, 1731–1740 (2005).

61. Boes, M. Role of natural and immune IgM antibodies in immune responses. *Mol. Immunol.* **37**, 1141–1149 (2001).

62. Siegrist, C. & Aspinall, R. B-cell responses to vaccination at the extremes of age. *Nat. Rev. Immunol.* **9**, 185–194 (2009).

63. Stavnezer, J. & Amemiya, C. T. Evolution of isotype switching. *Semin. Immunol.* **16**, 257–275 (2004).

64. Das, S., Hirano, M., Tako, R., Mccallister, C. & Nikolaidis, N. Evolutionary Genomics of Immunoglobulin-Encoding Loci in Vertebrates. 95–102 (2012).

65. Orson, F. M. *et al.* Spontaneous Development of IgM Anti-Cocaine Antibodies in Habitual Cocaine Users: Effect on IgG Antibody Responses to a Cocaine Cholera Toxin B Conjugate Vaccine. *Am. J. Addict.* **22**, 169–174 (2013).

66. Künzel, C. *et al.* Assay concept for detecting anti-drug IgM in human serum samples by using a novel recombinant human IgM positive control. *Bioanalysis* **13**, 253–263 (2021).

67. Blandino, R. & Baumgarth, N. Secreted IgM: New tricks for an old molecule. *J. Leukoc. Biol.* **106**, 1021–1034 (2019).

68. Goldstein, M. F. *et al.* Selective IgM immunodeficiency: Retrospective analysis of 36 adult patients with review of the literature. *Ann. Allergy, Asthma Immunol.* **97**, 717–730 (2006).

69. Borsos, T. & Rapp, H. J. Hemolysin Titration Based on Fixation of the Activated First Component of Complement : Evidence That One Molecule of Hemolysin Su ces to Sensitize an Erythrocyte. *J. Immunol.* **95**, 559–566 (1965).

70. Diebolder, C. A. *et al.* Complement Is Activated by IgG Hexamers Assembled at the Cell Surface. *Science (80-.).* **343**, 1260–1263 (2014).

71. Alexander R. Duncan & Greg, W. The binding site for C1q on IgG. *Nature* **332**, 738–740 (1988).

72. van Osch, T. L. J. *et al.* Fc Galactosylation Promotes Hexamerization of Human IgG1, Leading to Enhanced Classical Complement Activation. *J. Immunol.* **207**, 1545–1554 (2021).

73. Collins, C., Tsui, F. W. L. & Shulman, M. J. Differential activation of human and guinea pig complement by pentameric and hexameric IgM. *Eur. J. Immunol.* **32**, 1802 (2002).

74. Wiersma, E. J., Collins, C., Fazel, S. & Shulman, M. J. Structural and functional analysis of J chain-deficient IgM. *J. Immunol.* **160**, 5979–89 (1998).

75. Randall, T. D., King, L. B. & Corley, R. B. The biological effects of IgM hexamer formation. *Eur. J. Immunol.* **20**, 1971–1979 (1990).

76. Davis, A. C., Roux, K. H. & Shulman, M. J. On the structure of polymeric IgM. *Eur. J. Immunol.* **18**, 1001–1008 (1988).

77. Elvington, M., Liszewski, M. K. & Atkinson, J. P. Evolution of the complement system: from defense of the single cell to guardian of the intravascular space. *Immunol. Rev.* **274**, 9–15 (2016).

78. Matsushita, M. *et al.* Origin of the classical complement pathway : Lamprey orthologue of mammalian C1q acts as a lectin. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 10127–10131 (2004).

79. Cavaillon, J., Sansonetti, P. & Goldman, M. 100th Anniversary of Jules Bordet ' s Nobel Prize : Tribute to a Founding Father of Immunology. *Front. Immunol.* **10**, 1–8 (2019).

80. Gaboriaud, C., Ling, W. L., Thielens, N. M., Bally, I. & Rossi, V. Deciphering the fine details of C1 assembly and activation mechanisms: 'Mission impossible'? *Front. Immunol.* **5**, 3–9 (2014).

81. Dunkelberger, J. R. & Song, W. C. Complement and its role in innate and adaptive immune responses. *Cell Res.* **20**, 34–50 (2010).

82. Gaboriaud, C. *et al.* Structure and activation of the C1 complex of complement: Unraveling the puzzle. *Trends Immunol.* **25**, 368–373 (2004).

83. Gaboriaud, C., Frachet, P., Thielens, N. M. & Arlaud, G. J. The human C1q globular domain: Structure and recognition of non-immune self ligands. *Front. Immunol.* **2**, 1–8 (2012).

84. Reid, K. B. M. & Porter, R. R. Subunit composition and structure of subcomponent C1q of the first component of human complement. *Biochem. J.* **155**, 19–23 (1976).

85. Gaboriaud, C. *et al.* The Crystal Structure of the Globular Head of Complement Protein C1q Provides a Basis for Its Versatile Recognition Properties. *J. Biol. Chem.* **278**, 46974–46982 (2003).

86. Païdassi, H. *et al.* C1q binds phosphatidylserine and likely acts as a multiligand-bridging molecule in apoptotic cell recognition. *J. Immunol.* **180**, 2329–38 (2008).

87. Almitairi, J. O. M., Venkatraman, U., Furze, C. M. & Simpson-gray, X. Structure of the C1r – C1s interaction of the C1 complex of complement activation. *Proc. Natl. Acad. Sci.* **115**, 768–773 (2018).

88. Murugaiah, V. *et al.* Complement proteins as soluble pattern recognition receptors for pathogenic viruses. *Viruses* **13**, 1–30 (2021).

89. Wouters, D. & Zeerleder, S. Complement inhibitors to treat igM-mediated autoimmune hemolysis. *Haematologica* **100**, 1388–1395 (2015).

90. Trouw, L. A., Pickering, M. C. & Blom, A. M. The complement system as a potential therapeutic target in rheumatic disease. *Nat. Rev. Rheumatol.* **13**, 538–547 (2017).

91. Wouters, D., Wagenaar-bos, I., Ham, M. Van & Zeerleder, S. C1 inhibitor: just a serine protease inhibitor? New and old considerations on therapeutic applications of C1 inhibitor C1 inhibitor. *Expert Opin. Biol. Ther.* **2598**, (2008).

92. Ziccardi, R. J. The First Component of Human Complement (C1): Activation and Control. *Springer Semin. Immunopathol.* **6**, 213–230 (1983).

93. Wallis, R., Mitchell, D. A., Schmid, R., Schwaeble, W. W. & Keeble, A. H. Paths reunited : initiation of the classical and lectin pathways of complement activation. *Immunobiology* **215**, 1–11 (2010).

94. Zwarthoff, S. A. *et al.* C1q binding to surface-bound IgG is stabilized by C1r2s2 proteases. *Proc. Natl. Acad. Sci.* **118**, (2021).

95. Feinstein, A. & Munn, E. A. Conformation of the free and antigen-bound IgM antibody molecules. *Nature* **224**, 1307–1309 (1969).

96. Thielens, N. M., Tedesco, F., Bohlson, S. S., Gaboriaud, C. & Tenner, A. J. C1q: A fresh look upon an old molecule. *Mol. Immunol.* **89**, 73–83 (2017).

97. Borsos, T. & Rapp, H. J. Complement Fixation on Cell Surfaces by 19S and 7S Antibodies. *Science (80-.).* **150**, 505–506 (1965).

98. Feinstein, A., Richardson, N. & Taussig, M. J. Immunoglobulin flexibility in complement activation. *Immunol. Today* **7**, (1986).

99. Wang, G. *et al.* Molecular Basis of Assembly and Activation of Complement Component C1 in Complex with Immunoglobulin G1 and Antigen. *Mol. Cell* **63**, 135–145 (2016).

100. Ishizaka, T., Tada, T. & Ishizaka, K. Fixation of C ' and C ' la by Rabbit γ G- and γ M-Antibodies with Particulate and Soluble Antigens. *J. Immunol.* **100**, 1145–1153 (1968).

101. Brown, J. C. & Koshland, M. E. Activation of antibody Fc function by antigen-induced conformational changes. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 5111–5 (1975).

102. Pecht, I. Recognition and Allostery in the Mechanism of Antibody Action (Springer book chapter).pdf. 41–54 (1976).

103. Siegel, R. C. & Cathou, R. E. Conformation of Immunoglobulin M . III . Structural requirements of antigen for complement fixation by equine IgM . CONFORMATION OF IMMUNOGLOBULIN M 111 . Structural Requirements of Antigen for Complement Fixation by Equine IgM '. *J. Immunol.* **125**, (1980).

104. Kratz, H. J., Borsos, T. & Isliker, H. Mouse Monoclonal Antibodies at the Red Cell Surface - II. effect of hapten density on complement fixation and activation. *Mol. Immunol.* **22**, 229–235 (1985).

105. Poon, P. H., Phillips, M. L. & Schumaker, V. N. Immunoglobulin M possesses two binding sites for complement subcomponent C1q, and soluble 1:1 and 2:1 complexes are formed in solution at reduced ionic strength. *J. Biol. Chem.* **260**, 9357–9365 (1985).

106. WEINER, E. M. On the Interaction of the First Complement Component C1 and its Subunit Clq with Solid-Phase IgM Immune Complexes. *Scand. J. Immunol.* **28**, 425–430 (1988).

107. Mayer, M., Osler, A., Bier, O. & Heidelberger, M. Quantitative Studies of Complement Fixation : I . A Method. *J. Immunol.* (1948).

108. Borsos, T., Chapuis, R. M. & Langone, J. J. Distinction between fixation of C1 and the activation of complement by natural IgM anti-hapten antibody: Effect of cell surface hapten density. *Mol. Immunol.* **18**, 863–868 (1981).

109. Plaut, A., Cohen, S. & Tomasi, T. Immunoglobulin M: Fixation of Human Complement by the Fc

Fragment. *Science (80-.).* 7–8 (1972).

110. Diseases, I. & Institutes, N. An Improved Method for Performance of the C1 Fixation and Transfer Test. **55**, 115–122 (1982).

111. Ziccardi, R. J. Spontaneous activation of the first component of human complement (C1) by an intramolecular autocatalytic mechanism. *J. Immunol.* **128**, 2500–2504 (1982).

112. Folkerd, E. J., Gardner, B. & Hughes, J. N. The relationship between the binding ability and the rate of activation of the complement component C1. *Immunology* **41**, 179–185 (1980).

113. Merle, N. S., Church, S. E., Fremeaux-Bacchi, V. & Roumenina, L. T. Complement system part I – molecular mechanisms of activation and regulation. *Front. Immunol.* **6**, 1–30 (2015).

114. Florent, G., Lehman, D., Lockhart, D. & Putnam, F. W. Identity of the Fc Fragments of Pathological and Normal Human Immunoglobulin Mt. **13**, 3372–3381 (1974).

115. Feinstein, A., Munn, E. A. & Richardson, N. E. The three-dimensional conformation of γM and γA globulin molecules. *Ann. N. Y. Acad. Sci.* **190**, 104–121 (1971).

116. Idusogie, E. E. *et al.* Mapping of the C1q Binding Site on Rituxan, a Chimeric Antibody with a Human IgG1 Fc. *J. Immunol.* **164**, 4178–4184 (2000).

117. Arya, S. *et al.* Mapping of amino acid residues in the C mu 3 domain of mouse IgM important in macromolecular assembly and complement-dependent cytolysis. *J. Immunol.* **152**, 1206–12 (1994).

118. Shulman, M. J., Collins, C., Pennell, N. & Hozumi, N. Complement activation by IgM: evidence for the importance of the third constant domain of the μ heavy chain. *Eur. J. Immunol.* **17**, 549–554 (1987).

119. Shulman, M. J., Heusser, C., Filkin, C. & Kohler, G. Mutations Affecting the Structure and Function of Immunoglobulin M. *Mol. Cell. Biol.* **2**, 1033–1043 (1982).

120. Wright, J. F., Shulman, M. J., Isenman, D. E. & Painter, R. H. C1 binding by mouse IgM. The effect of abnormal glycosylation at position 402 resulting from a serine to asparagine exchange at residue 406 of the μ-chain. *J. Biol. Chem.* **265**, 10506–10513 (1990).

121. Nadeau, J. H. & Frankel, W. N. The roads from phenotypic variation to gene discovery: Mutagenesis versus QTLs. *Nat. Genet.* **25**, 381–384 (2000).

122. Monnet, C. *et al.* Selection of IgG variants with increased FcRn binding using random and directed mutagenesis: Impact on effector functions. *Front. Immunol.* **6**, 1–14 (2015).

123. Planchais, C. *et al.* Oxidized hemoglobin triggers polyreactivity and autoreactivity of human IgG via transfer of heme. *Commun. Biol.* **6**, 1–15 (2023).

124. Bubb, M. O. & Conradie, J. D. Studies on the structural and biological functions of the CH3 and CH4 domains of IgM. *Immunology* (1978).

125. Hurst, M. M., Volanakis, J. E., Stroud, R. M. & Bennett, J. C. A Comparative Analysis of the C1-Binding Ability of Fragments Derived from Complement-Fixing and Noncomplement-Fixing IgM Proteins. *J. Clin. Invest.* **58**, 16–21 (1976).

126. Hurst, M. M., Volanakis, J. E., Hester, R. B., Stroud, R. M. & Bennett, J. C. THE STRUCTURAL BASIS FOR BINDING OF COMPLEMENT BY IMMUNOGLOBULIN M. *Exp. Med.* **140**, 1117–1121 (1974).

127. Hurst, M. M., Volanakis, J. E., Stroud, R. M. & Bennett, J. C. C1 fixation and classical complement pathway activation by a fragment of the CH4 domain of IgM. *Exp. Med.* **142**, (1975).

128. Johnson, B. J. & Thames, K. E. Investigations of the Complement-Fixing sites of Immunoglobulins. *J. Immunol.* **117**, 1491–1494 (1976).

129. Siegel, Richard C; Cathou, R. E. Effects of Limited Denaturation by Heat on the Dynamic Conformation of Equine Immunoglobulin M Antibody and on Interaction with Antigen and Complement. *Biochemistry* 192–198 (1981).

130. Pedersen, M. B. *et al.* Curvature of Synthetic and Natural Surfaces Is an Important Target Feature in Classical Pathway Complement Activation. *J. Immunol.* **184**, 1931–1945 (2010).

131. Oskam, N. *et al.* At Critically Low Antigen Densities, IgM Hexamers Outcompete Both IgM Pentamers and IgG1 for Human Complement Deposition and Complement-Dependent Cytotoxicity. *J. Immunol.* **209**, 16–25 (2022).

132. Karush, F., Chua, M. M. & Rodwell, J. D. Interaction of a Bivalent Ligand with IgM Anti-Lactose Antibody. *Biochemistry* **18**, 2226–2232 (1979).

133. Shaw, A. *et al.* Binding to nanopatterned antigens is dominated by the spatial tolerance of antibodies. *Nat. Nanotechnol.* **14**, 184–190 (2019).

134. Song, D. *et al.* The spectrum of renal thrombotic microangiopathy in lupus nephritis. *Arthritis Res. Ther.* **15**, 1–12 (2013).

135. Tichaczek-Goska, D. Deficiencies and Excessive Human Complement System. *Adv Clin Exp Med 2012* **21**, 105–114 (2012).

– Chapter 2 –

Simple Platform for Automating Decoupled LC-MS Analysis of HDX Samples

Portions of the text in this chapter have been modified and reproduced with permissions from:

Watson, M. J. et al. Simple Platform for Automating Decoupled LC-MS Analysis of Hydrogen/Deuterium Exchange Samples. J. Am. Soc. Mass Spectrom. (2020) doi:10.1021/jasms.0c00341.

2.1 Introduction

Solution state proteins are constantly in motion, with movements ranging from small-scale fluctuations in local secondary structure to large-scale tertiary and quaternary domain rearrangements that can lead to an array of distinct global conformational states.[1] While classical structural techniques such as NMR and X-ray crystallography are often able to provide high resolution structures of distinct protein conformations, the data produced by these methods are inherently static and therefore limited in the extent to which they can inform on protein dynamics or be used for mechanistic interpretations of protein conformational change. Hydrogen/deuterium exchange coupled with mass spectrometry (HDX-MS) is a versatile structural technique that allows proteins to be investigated in their physically native state, and can often be used to unveil subtle conformational changes within challenging biological systems that would otherwise be unobservable by many other established structural techniques.[2] The investigation and understanding of a protein's conformational dynamics are critical to the development of biotherapeutics, as the structural range of a protein will inherently dictate its range and efficacy of biological function.[3] The flexibility and application variety of HDX-MS methodology has led to it becoming a routine tool in both academic and industrial pursuits of detailing protein-protein interactions, identifying allosteric effects, mapping folding pathways, and investigating mechanisms of

protein structure/function relationships across many complex and otherwise inaccessible systems, including highly flexible, extensively glycosylated, and intrinsically disordered proteins.[4–7]

The key aspect of HDX-MS that underlies its utility in the study of protein structure is the exploitation of exchange rate variability amongst protein backbone amides, as regions with different degrees of solvent accessibility will be labeled differentially with solvated deuterium over fixed periods of time; while the natural exchange between hydrogen and deuterium is chemically dependent on solution variables such as pH and temperature, the real utility of the method lies in the differential uptake of the deuterium label caused by differences in solvent accessibility to the protein backbone, which is heavily influenced by the local structural arrangement and surrounding degree of hydrogen bonding.[5,8,9] For example, residues buried within the core of a protein domain will exchange at a slower rate compared to those positioned along the surface that are exposed directly to deuterated solvent, and backbone amides that are arranged into stable secondary structures such as alpha helices or beta sheets will generally exchange much slower (several orders of magnitude) than those found in disordered loops or the junctional gaps between major protein domains.[10] Solvent exposure is a fundamental aspect that enables many types of differential analyses, and conditional restrictions to solvent exposure can point directly to protein regions of functionality, such as the binding interface of an antibody Fab domain determined through the straightforward comparison of its unbound and ligand bound states. In many cases the resolution of the generated data can be sufficient to quickly pinpoint the functional involvement of specific residues, with the resolving power limited ultimately by the total coverage, number, and degree of overlap between afforded peptide fragments (**Fig. 2.1.1**).

By monitoring the exchange of a protein over fixed periods of time, the resulting extent of deuterium uptake can then be used to infer information about biologically-relevant changes to local secondary structure, as well as the mechanistic qualities of large scale conformational changes that underlie many key physiological functions.[11] The relative rate at which a peptide subregion takes up the

deuterium label within any given condition is based on the distribution of its conformational subpopulations, which will be inherently reflected in the isotopic distribution of its mass spectral envelopes. Because the vast majority of protein motion occurs on shorter timescales than the chemical rate of H/D exchange, labeling will often be observed to occur gradually over time as the protein natively experiences structural fluctuations, in a scenario often referred to as "EX2 kinetics" that is hallmarked by unimodal mass spectral envelopes that gradually shift to higher m/z ranges with each progressive time point that is analyzed (**Fig 2.1.2 A**).[9,10,12] While EX2 kinetics are the most commonly observed in practice, there are some cases in which a protein will undergo larger structural transitions that can last for relatively extensive periods of time (multiple seconds), which causes collective regions comprised of multiple backbone amides to become fully labeled by the time the transition is complete. In these "EX1" scenarios (**Fig. 2.1.2 B**), the resulting isotopic distribution is often observed to be multimodal, which is reflective of distinct conformational subpopulations present within a given protein sample, as the rate and degree to which each distinct protein state is labeled will be inherently different.[13,14] For example, some catalytic enzymes are known to sample distinctly 'closed' or 'open' conformational states which, under constant HDX conditions, results in regions that become relatively 'protected' or 'de-protected'(i.e. 'exposed') from uptake of the deuterium label, compared to an undeuterated reference.[15]

While the bimodal profiles of mass spectral envelopes can sometimes provide an added layer of nuance to the structural dynamics of differing protein conditions and conformational states, it should be noted that kinetic interpretations of some HDX-MS experiments can be challenging due to the complexity and dynamic range of the proteins involved, along with the technical challenges of performing HDX-MS labeling and analysis in an analytically robust manner.[16] For example, there are some proteins that exist naturally in a variety of stable conformational subpopulations, such as the protein shell capsid domains of viruses, that can lead to HDX-MS peptides with varying degrees of uptake that take on a mixture of both EX1 and EX2 kinetic profiles (**Fig. 2.1.2 C**).[10,17,18] This sort of isotopic

distribution mixture can, at the same time, also result from sample impurities or the presence of aggregated and/or degraded protein material, along with heterogeneity derived through inadequate preparation of ligand bound conditions that don't reach equivalent degrees of intended binding saturation.[5] Additional complexity can be found further still in proteins that happen to naturally undergo motions at truly mixed time scales (**Fig. 2.1.2 D**), along with the ever-present chance of encountering technical errors during sample analysis, such as analytical carryover (**Fig 2.1.2 E**) or the inadvertent and spontaneous loss of the deuterium label itself, commonly referred to as 'back-exchange'.[19–24]

While all peptide segments of a protein will eventually become fully deuterated if given enough time, the use of differential time point comparisons between unique protein states can lead directly to major structural insights. In a typical differential HDX-MS experiment, protein sample conditions are rapidly diluted with a D_2O buffer solution within a series of fixed and predetermined time periods that can range anywhere from several days to only a few seconds, with some groups even reporting new approaches that are capable of sampling sub-second timescales.[25] At the end of each time point, the exchange reaction is halted rapidly by the addition of a 'quenching' solution that drops the solution pH to 2.5 and quickly lowers the temperature to 0° C, whereupon samples can either be immediately analyzed by LC-MS, or flash-frozen and stored at -80° C for long-term storage and decoupled LC-MS analysis.[4] In the commonly utilized bottom-up approach, protein samples are often passed through inline protease columns where they can be quickly digested into smaller peptides before separation and analysis by LC-MS. While viable protease options are still severely limited due to the temperature and pH constraints of HDX quenching conditions, columns can often still be custom made with purified proteases (most often pepsin) or purchased from various commercial vendors.[26–28]

Historically, HDX experimentation has suffered from low reproducibility along with low sample throughput, which can both be tied to directly to back-exchange of the deuterium label.[20,29,30] As previously mentioned, after each exchange, every sample must immediately be quenched to low pH,

kept cold, and analyzed rapidly – all to slow the spontaneous loss of deuterium after labeling. Equal and consistent treatment of every sample replicate, across all experimental conditions, is vital to data integrity and will heavily influence the impact of interpretations that can potentially be made. An advantage of using a decoupled approach is that samples are run in a continuous queue, rather than waiting as various time points are prepared, thereby promoting the consistency of analytical variables along with making the most efficient use of the MS instrument time. Recent commercial systems have been developed to automate nearly every step in HDX-MS analysis, including automated deuterium labeling with in-line injection, protease digestion, and LC-MS analysis.[29,31–35] While this technology has greatly expanded the throughput and precision of HDX-MS, limitations still remain in terms of sampling different exchange conditions along with post-quench sample manipulation(s). For example, current robotic systems are not typically well-suited for sampling rapid timescales (ms) or pulsed labeling experiments, and are not equipped to perform complicated cleanup steps that are often required when working with complex systems, such as membrane bound proteins.[36–38]

While some limitations for automated in-line HDX-MS are starting to be addressed, such as the recent development of filtration columns designed for post-quench removal of lipids, there are still many types of samples and experimental parameters that cannot be automated and require manual sample preparation.[39] While not as convenient, preparing samples manually does offer notable advantages: 1) all samples can be prepared in a short time window minimizing potential protein degradation; 2) exchanges can be performed on immobilized proteins, lyophilized proteins, or colloidal particles; 3) a wider range of exchange timepoints can be sampled; and 4) samples are amenable to complicated post-quench cleanup steps. The major disadvantage, however, comes with having to minimize and maintain consistent levels of back-exchange. In a decoupled approach, samples remain frozen until the moment they are ready for LC-MS analysis, and every sample must be consistently thawed before injection. On top of this, reproducibility suffers when samples are analyzed on different days, meaning that every

sample within a given experiment should be analyzed back-to-back for optimal results.[3] Without the access to an autosampler capable of maintaining subfreezing temperatures (-80°C), these constraints require users to be physically present for the manual thawing and injection of every sample within a given experiment.[33,35] In this chapter we describe a simple and inexpensive sample manager capable of alleviating this inherent bottleneck with a decoupled HDX-MS pipeline.

2.2 Methods

2.2.1 HDX-MS Sample Preparation

10 µL of protein samples at concentrations of 0.1 - 1 mg/mL were diluted ten-fold into deuterated buffer (85% D2O final) for timepoints ranging from 3 seconds up to 20 hours at room temperature. Samples were immediately quenched with an equal volume of ice-cold quench buffer (containing up to 8 M urea and 200 mM TCEP with 0.2% formic acid) to bring the pH to 2.5 and the final volume to 200 µL. For manual injections, samples were snap frozen in liquid nitrogen and stored at -80° C until LC-MS analysis. Prior to manual injection, each sample was removed from a liquid nitrogen reservoir and allowed to sit on ice for 5 mins. The sample was then placed on the countertop at room temperature for exactly 1 minute to reach a slurry-like state, and immediately injected with a 250 µL gastight syringe (Hamilton) wrapped in parafilm for insulation and precooled on ice. For automated injections, quenched samples were rapidly transferred to 1 mL autosampler glass vials (Total Recovery, Waters) pre-frozen in a 54-vial sample tray within a bath of ethanol and dry ice. The bath was treated with the same precautionary measures used for cold traps employing organic solvents and dry ice, including necessary precautions for flammable solvents and the use of cryogenic gloves. Samples were capped with magnetic screw caps, and the entire 54-vial tray was stored at -80°C until LC-MS analysis. Samples were thawed by moving the vials from the dry ice/ethanol compartment to a sample block held at 4°C and waiting for a specified time before drawing the sample and injecting it into the injection port

on the HDX system. Thaw times were in the range of 3 to 10 minutes dependent on the composition of the quenched solution.

2.2.2 Data Analysis

Peptide assignments were made using MS/MS data obtained from undeuterated samples analyzed with Protein Prospector and Byonic (Protein Metrics). Deuterium uptake was analyzed and summarized using HDExaminer v2 software. All timepoints and replicates from available data sets were included unless the sample encountered a known technical problem (e.g. leaks or clogs). Variability statistics (standard deviations) were calculated from each discreet sample set using the deuterium uptake of a reporter peptide known to be fully deuterated by the earliest exchange timepoint (3 sec).

2.2.3 Subfreezing Sample Manager

The main body for the dry ice/ethanol sample manager was machined from a single slab of high density polyethylene plastic (HDPE, McMaster-Carr, Los Angeles, CA, USA), such that a thick block of dry ice (approx. 2” X 8” X 8”) can be placed in the rear dry ice cavity and a 54 sample vial tray can be securely seated in the front cavity (**Fig. 2.2.1 A-B**). A separately attached flow-through divider machined from 0.5” thick HDPE (McMaster-Carr) separates the rear and front cavities but is slotted to allow for the free circulation of dry ice cooled ethanol between the two cavities (**Fig. 2.2.1 C**). The rear dry ice cavity contains a slight downward slope to maintain contact between the dry ice/ethanol bath as the dry ice depletes over time. An insulated and slidable tongue-and-groove lid was fabricated from HDPE, covering the sample tray holder cavity, and is opened/closed by a LEAP/PAL HTX-XT robotic arm (CTC Analytics, Zwingen, Switzerland). The main body is insulated by surrounding all sides (including top and bottom) with 1” thick impact resistant low temp PVC foam insulation sheet (McMaster-Carr) and then fully enclosing the insulation with a framing of 0.5” thick HDPE. The entire sample manager was fastened to a flat surface with L brackets.

A simple circuit was designed and attached to the side of the sample controller allowing for the temperature of the ethanol/dry ice bath to be monitored remotely. It incorporates a microcontroller (ARDUINO UNO WIFI-REV2 microcontroller, Arduino LLC, Iverea, Italy), a thermocouple amplifier (PN: 485-269, Mouser Electronics, Mansfield, TX) and a type K-thermocouple wire (PN: 485-270, Mouser Electronics, Mansfield, TX). The thermocouple can accurately monitor temperatures as low as -200 °C. The microcontroller has a built-in Wi-Fi antenna which can connect to the internet and post data that is sent to a MATLAB/ThinkSpeak account plotting temperature over a designated time period.

2.2.4 Analytical HDX Cold Box

The analytical HDX cold box is a simple in-house designed and fabricated system that allows for micro-flow, high pressure 2-way computer-controlled switching valves, chromatography columns and plumbing lines to be cooled and maintained at 0° C, similar to that described previously.[31] Both stainless steel and PEEK tubing buffer transport lines are plumbed into the box through multiple bulkhead unions (VICI-VALCO Instruments, PN: ZBU1T – 0.062" internal diameter; PN: ZBU1L – 0.040" internal diameter; PN: ZBUXC – 0.006" internal diameter). A sample injection port was attached to the remaining external side of the box. An independently controlled cooler/heater for a pepsin column was mounted on the outside of the cold box, allowing the pepsin column to operate at a temperature range of 5° to 55° C (Analytical Sales, Flanders, NJ, USA, PN: Cold-Sleeve-30). Stainless steel buffer cooling loops inside the cold box (5 feet long, internal diameter 0.020") allow for ambient temperature liquid buffers entering the box to reach the lower internal box temperature.

2.2.5 Automation of Thaw/Inject Cycles

A LEAP/PAL 2DW two-drawer stack cooler (CTC Analytics, Zwingen, Switzerland) was employed as a 'thaw shelf' with a programmable temperature setting of 4° C. Thaw/inject cycles were created by programming a series of events in Cycle Composer software (LEAP Technologies), using the LEAP robotic

arm equipped with a magnetic tip and 250 µl sampling syringe to: (1) slide open the thaw shelf door, (2) move to the frozen sample holder and slide open the access lid, (3) pick up designated sample via magnetic HPLC vial cap, (4) place the sample in the thaw shelf, (5) return to the frozen sample holder and slide access lid closed, and (6) move back to thaw shelf and slide the door closed. Optimal thaw times were empirically determined to be: 10 minutes for samples without any urea or TCEP, 4 minutes for samples containing 4 M urea (final), 8 minutes for samples containing 500 mM TCEP, and 3 minutes for samples containing both 4 M urea and 500 mM TCEP. As the programmed thawing time expires, the LEAP arm (7) moves to the injection port and draws a full syringe of cold loading buffer, (8) dispenses all but 50 µL of the cold loading buffer, (9) opens the thaw shelf, (10) injects 50 µL of cold loading buffer into the thawed sample, (11) aspirates and dispenses the sample twice for mixing, (12) draws the entire sample (250 µL total) into the syringe, and (13) moves to injection port and injects sample into the analytical HDX analytical cold box before finally (14) returning to and closing the thaw shelf door (see supporting video).

The syringe is then rinsed three additional times with 5% methanol, 80% methanol, and 5% methanol, respectively, through the LEAP syringe cleaning stations. During trapping and LC separation, the LEAP robotic arm is used to inject a series of wash solutions as outlined in the 'automated system washes' section. Each major system component, namely the sample manager, analytical HDX cold box, and mass spectrometer are linked together via the PAL LEAP system and coordinated by Cycle Composer. We note that earlier generations of PAL LEAP systems, such as the PAL CTC, are equally capable of performing these functions, and that much of the cost associated with the construction of this system went towards the iterative design and machining process for the major structural components, which are provided as supporting files.

2.2.6 Protocol for running LC-MS analysis

Injected samples were digested at 0°, 5°, or 10° C through a custom-packed inline pepsin column (2.1 x 50 mm) with a flow of loading buffer (0.1% TFA, 2% ACN in water) at a rate of 200 uL/min. Peptides were trapped onto a Waters XSelect CSH trap cartridge (2.1 x 5 mm, 1.7 µm, 130Å) and resolved over a Waters CSH C18 column (1 x 100 mm, 1.7 µm, 130Å) using a 15 min gradient of 3% to 40% mobile phase B (A: 2% ACN, 0.1% FA, 0.025% TFA in H2O; B: 0.1% FA in ACN). MS analysis was performed using either a Thermo LTQ-Orbitrap (capillary temperature 120°C, mass resolution 60k) or a Waters Synapt G2-Si Q-TOF (source and desolvation temperature set to 70° and 130°C, respectively) with ion mobility enabled.

2.2.7 Automated System Washes

Between all analytical samples, automated cleaning protocols are executed to wash the protease and analytical/trap columns to limit carryover. The protease column wash is a four-part injection series (250 µL) in the following order: 1) 0.1% Fos-12 in 0.1% FA; 2) 2M guanidine HCl in 0.1% FA; 3) 10% acetic acid, 10% ACN, 5% IPA; 4) 10 % FA.[19,40] The analytical trap/column wash is also a four-part injection (250 µL) series and proceeds in the following order: 1) 10 % FA; 2) 30 % trifluoroethanol; 3) 80 % MeOH; 4) 66% isopropanol, 34% ACN.[41] During the trap washing the main LC column is cleaned with a series of three rapid gradients from 5 to 95% B and equilibrated with 2% B prior to the next sample run. All wash solutions were loaded into 10 mL glass bottles on the standard LEAP wash station and configured as pots within Cycle Composer. In addition to these wash protocols, the configuration of the valves allows for back-flushing of the loading, pepsin, and trapping LC lines which further help to minimize sample carryover.

2.2.8 3D Printed LC-MS Sample Trays

Custom 54-well sample trays were 3D printed in standard polylactic acid (PLA) using an Original Prusa I3 MK3S printer as illustrated in **Figure 2.2.2** (Prusa Research, Prague, Czech Republic). Extruder temperature was set to 210° C for the first layer before being lowered to 200° C for the duration of the print, and bed temperature remained constant at 60° C. Objects were created using Autodesk Fusion 360 and slicing was done using PrusaSlicer software.

2.3 Results & Discussion

2.3.1 System Design and Overview

The influence that temperature, pH, and amino acid identity have on exchange rates are well-established and allow for accurate prediction of back-exchange under different conditions.[8] These rates suggest that a temperature of -60° C is sufficient for reducing deuterium loss in quenched samples (pH 2.5) to less than 1% over a period of 24 hours (**Fig. 2.3.1**). However, longer storage times at -60°C would be detrimental, with more than 5% deuterium loss after 1 week. While these predictions overlook the possibility that solid-state exchange kinetics may significantly deviate from that predicted in solution, they do indicate that, while long-term storage of HDX-MS samples at -80°C temperatures is warranted, samples may not necessarily need to be kept at such low temperatures to mitigate back-exchange for sample queues less than 24 hours.

Baths composed of dry ice and ethanol are commonly utilized in chemistry and offer a simple and inexpensive way to achieve a stable temperature under -60°C. We therefore devised a subfreezing sample manager capable of maintaining frozen HDX-MS samples using dry ice/ethanol for integration into a cost-effective automated LC-MS platform (**Fig 2.3.2 A**). The sample manager is an in-house fabricated unit designed to house a single 54-well sample tray and dry ice is stored in a large peripheral reservoir while the ethanol conducts heat to keep the sample tray at -65° C as measured by a calibrated

K-type thermocouple (**Fig. 2.3.2 B-C**). An insulated sliding lid above the sample tray allows for access to the sample manager via the integrated LEAP robot. A standard LEAP stack cooler is maintained at 4° C and used for controlled thawing of frozen samples. The full description, schematics, and designs for the sample manager and 54-well sample tray, along with the detailed thaw/injection protocols, are outlined in methods section 2.2.

2.3.2 Increasing Throughput

The constraints of HDX-MS sample handling have greatly restricted the throughput of LC-MS analysis. Accounting for the time required for wash steps to minimize carryover, we can only achieve an average sampling rate on the order of 2 samples per hour, or 24 samples within a 12-hour shift. Furthermore, to minimize variability, every sample within a given dataset should be collected by the same analyst within a single continual session, making the manual LC-MS analysis of HDX samples a long and tedious process. By creating a sample manager capable of being accessed robotically, we aimed to improve throughput by directly lifting the burden of a manual thaw/inject approach.

To assess our viable range of sample storage time we set up an extended queue of HDX-MS samples that contained fully deuterated bradykinin peptide to serve as an internal reporter for back-exchange. After filling the dry ice reservoir, the queue was initiated, and data was collected every 40 minutes over the course of 20 continuous hours without any manual intervention or resupply of dry ice. Consistent with our initial estimation of back-exchange at low temperatures, bradykinin deuteration levels showed no systematic change throughout the course of 18 hours, after which a sharp drop in deuterium retention was observed (**Fig. 2.3.3 A**). This drop in deuterium level is a direct consequence of accelerated back-exchange, occurring as the sample manager begins to warm after depletion of the dry ice reservoir. The design of our sample manager addresses this issue by allowing for the resupply of dry ice at any point without halting data collection. As a queue nears completion, the entire 54-well sample

tray may be quickly swapped with another tray for a seamless and virtually unlimited continuation of LC-MS analysis with little manual interaction required.

2.3.3 Reducing Variability

To gauge the precision of our automated system we compared an extensive series of decoupled HDX-MS data sets to assess the general variability between manual and automated thaw/inject methods. Using a fully deuterated peptide, we determined that the automated platform reduces overall variability by more than 15% (**Fig 2.3.3 B**). While the study included roughly the same number of discreet sample sets between the two methods, we note that the automated sample sets are considerably larger due to the increased throughput afforded by the system (Table S1). While sample preparation (labeling and quenching) remains a large contributor to variance within any given experiment, our system shows an overall reduction in variability due to the precise control of timing and temperature throughout every thaw/inject cycle.

2.4 Conclusion

We present an inexpensive sample manager integrated with a robotic autosampler for high throughput decoupled HDX-MS. The platform enables collection of frozen HDX-MS samples with superior precision over manual injections and enables collection of large data sets for complex systems that are currently unattainable by fully automated HDX-MS systems.

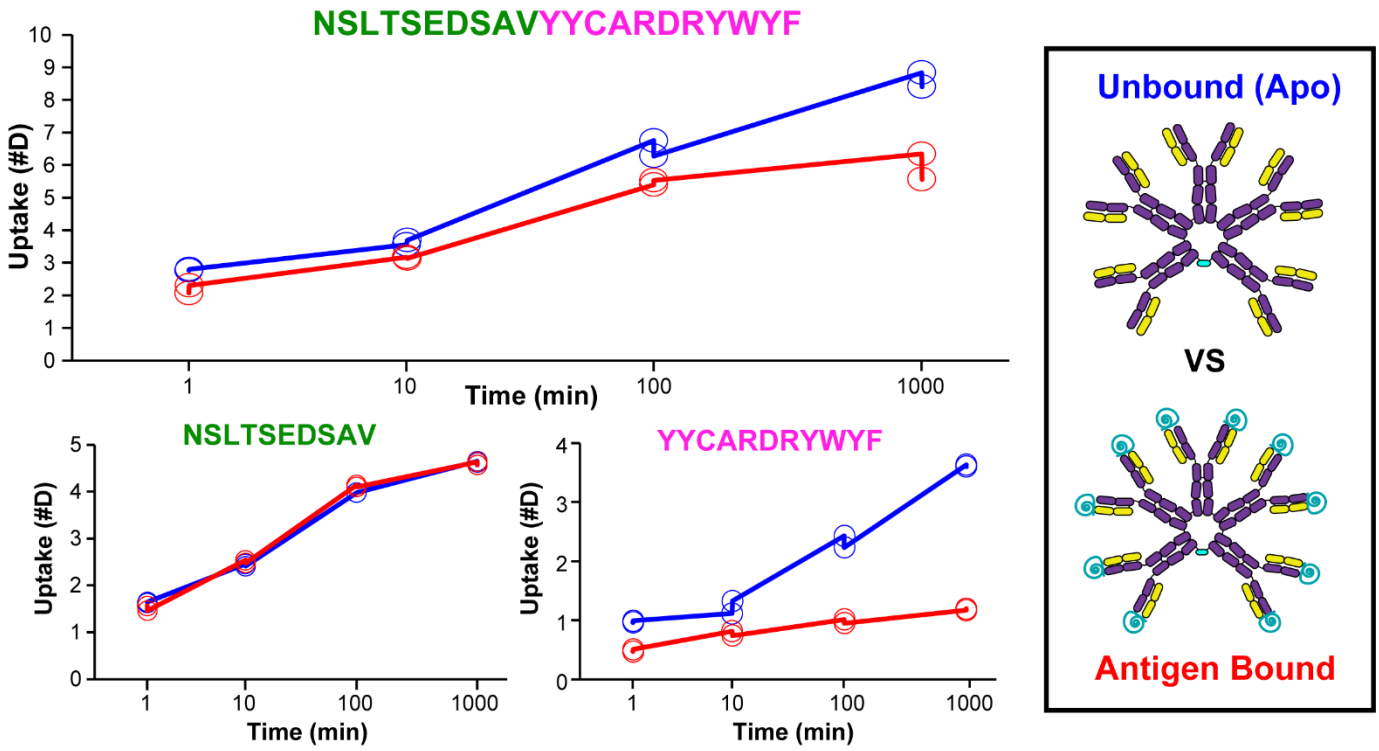


Figure 2.1.1 | Example differential HDX-MS analysis of an antibody binding interface

HDX-MS is sensitive to changes in solvent accessibility along the backbone amides of a protein, which is a fundamental aspect of the method that can be applied towards the determination of regions that contribute toward protein functionality, such as the identification of residues that comprise antibody binding interfaces. The exclusion of solvent via binding interactions at the Fab domain of an antigen bound antibody, such as IgM (shown above), can be used to pinpoint the residues that make direct protein:ligand contact via differential analysis of deuterium uptake between bound and unbound states. In many cases the resolution of the data can be sufficient for narrowing down structural differences to small local regions (<10 residues) of the primary sequence, and in some cases, with enough fragmentation and sequence coverage, the sight-specific information of single amino acids can also be achieved.

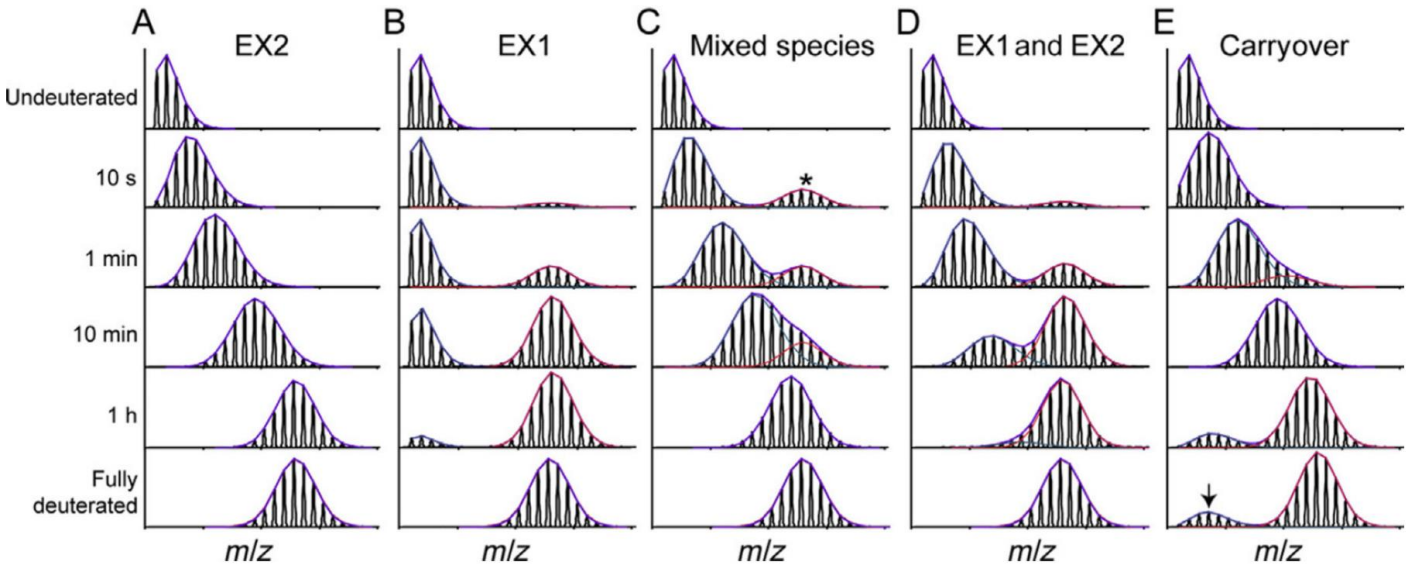


Figure 2.1.2 | Example mass spectra of HDX-MS kinetic profiles. ***Adapted from Guttman & Lee (2016)***

Example data of mass spectral envelopes that arise from various HDX kinetics that show **A)** the common EX2 unimodal envelope progression with time, and **B)** the less common EX1 bimodal spectra that can arise for certain peptides that undergo larger structural rearrangements over relatively extended timescales. Complexities in the data can arise when analyzing samples containing **C)** mixed species or heterogenous protein conditions that experience differential uptake of the deuterium label (minor population indicated with asterisk), along with **D)** protein regions that truly undergo motions at mixed timescales resulting in the joint appearance of both EX1 and EX2 exchange kinetics over time. Data can be impacted further by technical challenges as well, such as back-exchange of the deuterium label or **E)** significant analytical carryover that can degrade the quality of resulting interpretations if warning signs are missed, such as the presence of a smaller and less deuterated subcomponent within a fully deuterated sample condition (arrow) that should instead be completely and homogeneously exchanged.

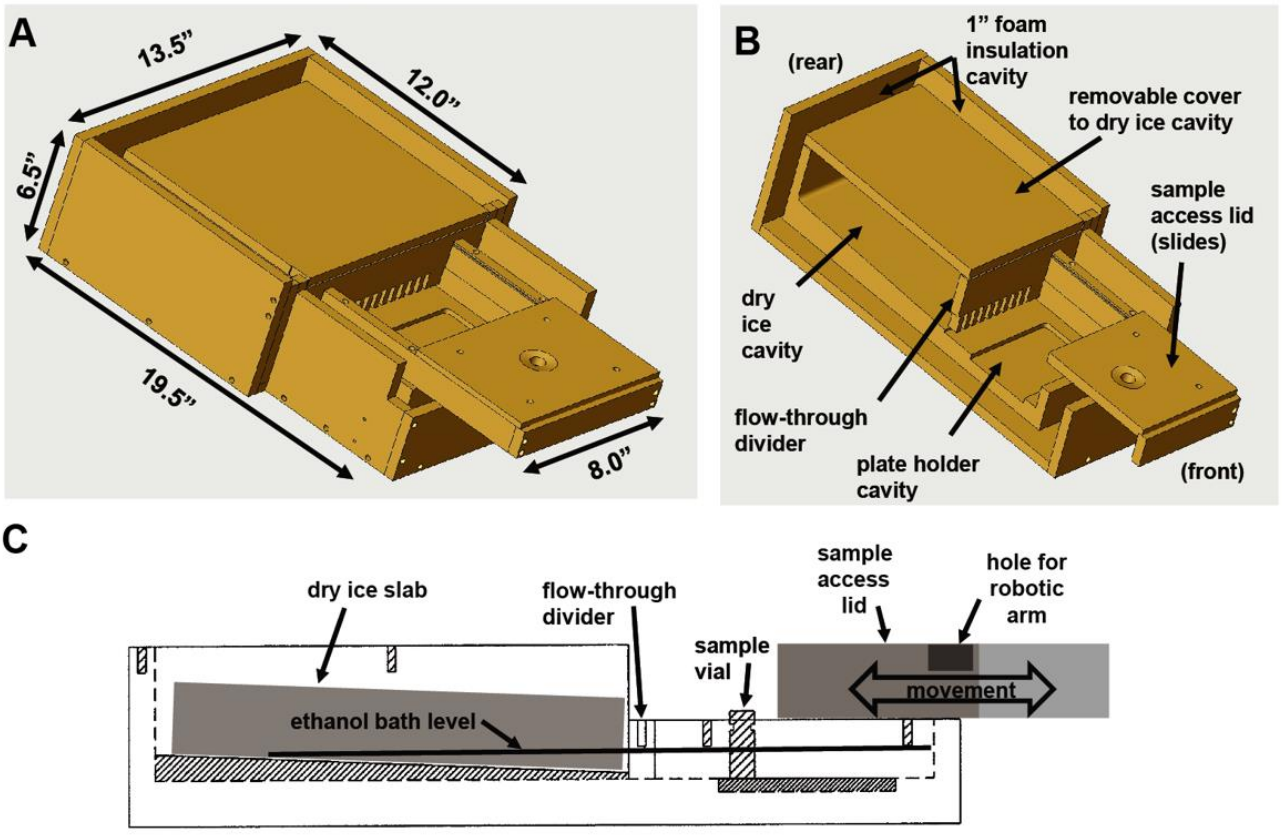


Figure 2.2.1 | Design details of the subfreezing sample manager

A) Size dimensions and **B)** cut-away view of internal compartments and features of the sample manager. **C)** The sample manager utilizes a dry ice/ethanol bath housed in a slanted chamber with a flow-through divider to ensure that HDX samples are kept frozen as the dry ice reservoir depletes over time.

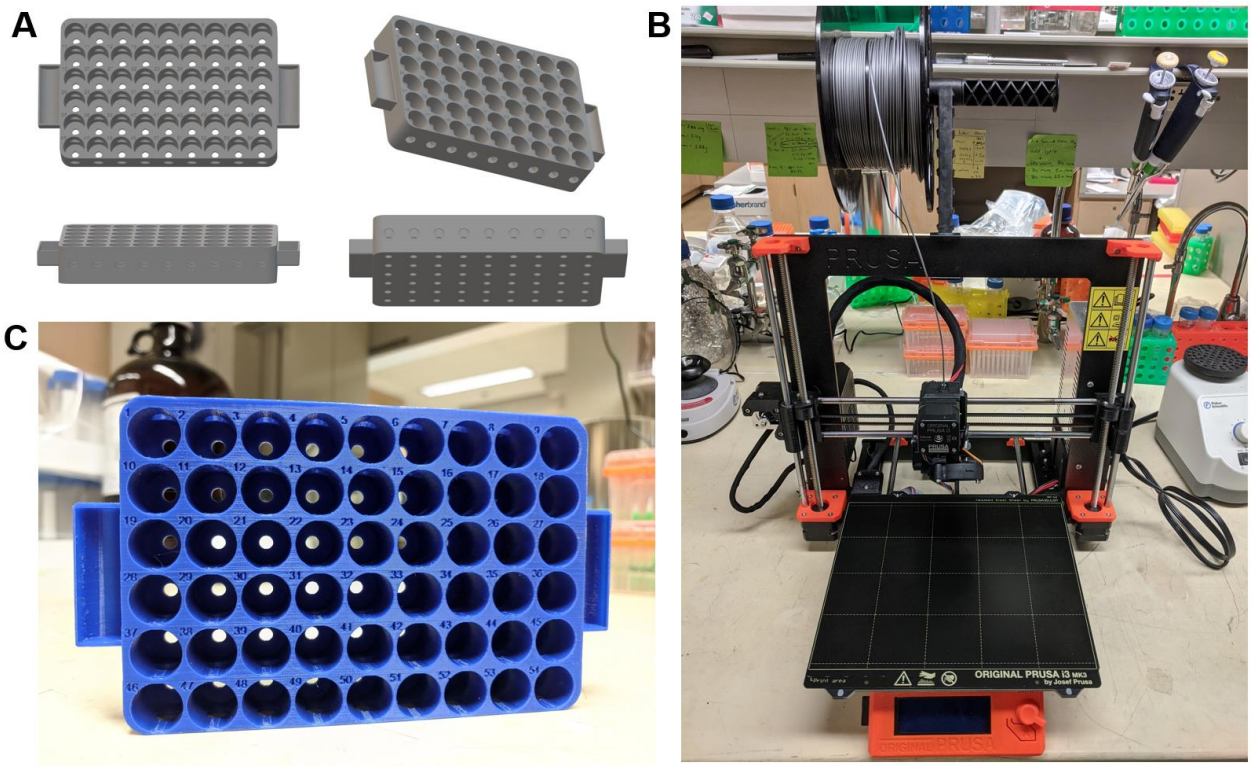


Figure 2.2.2 | Design and 3D printing of custom 54-well LC-MS sample trays

A) The design features a series of vertical and horizontal holes that act as fluid channels to ensure sample vials are in contact with the dry ice/ethanol bath. Additionally, two slots designed to hold steel dowel pins were added to the exterior of each side as weights to ensure the tray remains fully submerged. **B)** Prusa i3 MK3S 3D printer. **C)** closeup view of printed 54-well tray.

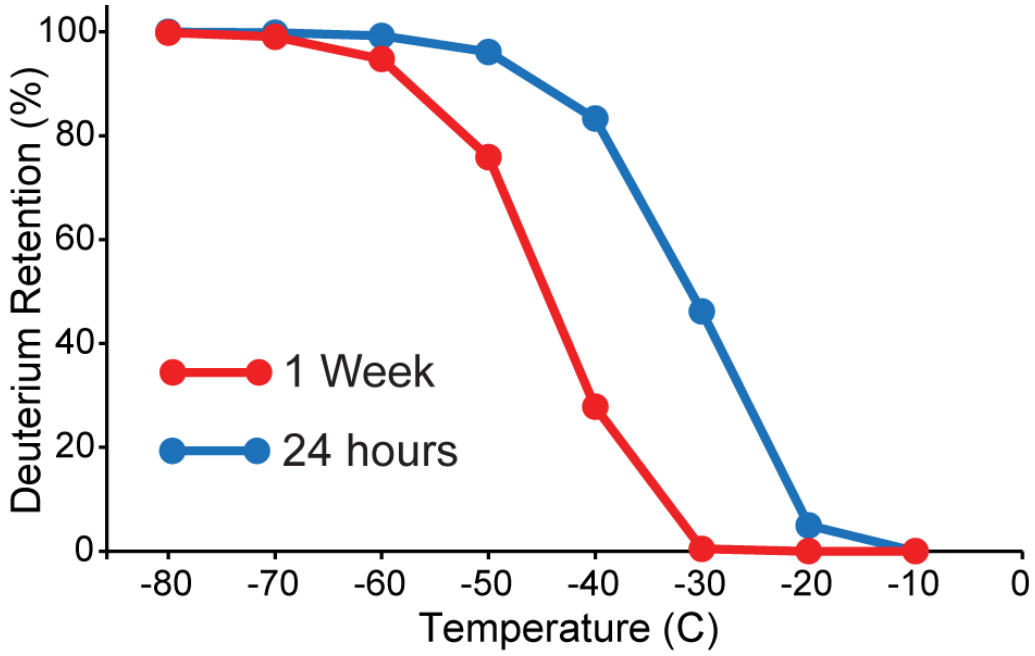


Figure 2.3.1 | Theoretical effects of temperature on back-exchange of frozen HDX samples

The deuterium retention for an average protein backbone amide is shown after 24 hours (blue) or 1 week (red) when stored at different temperatures. Notably, -60° C is sufficient for maintaining >99% of deuterium after 24 hours (blue).

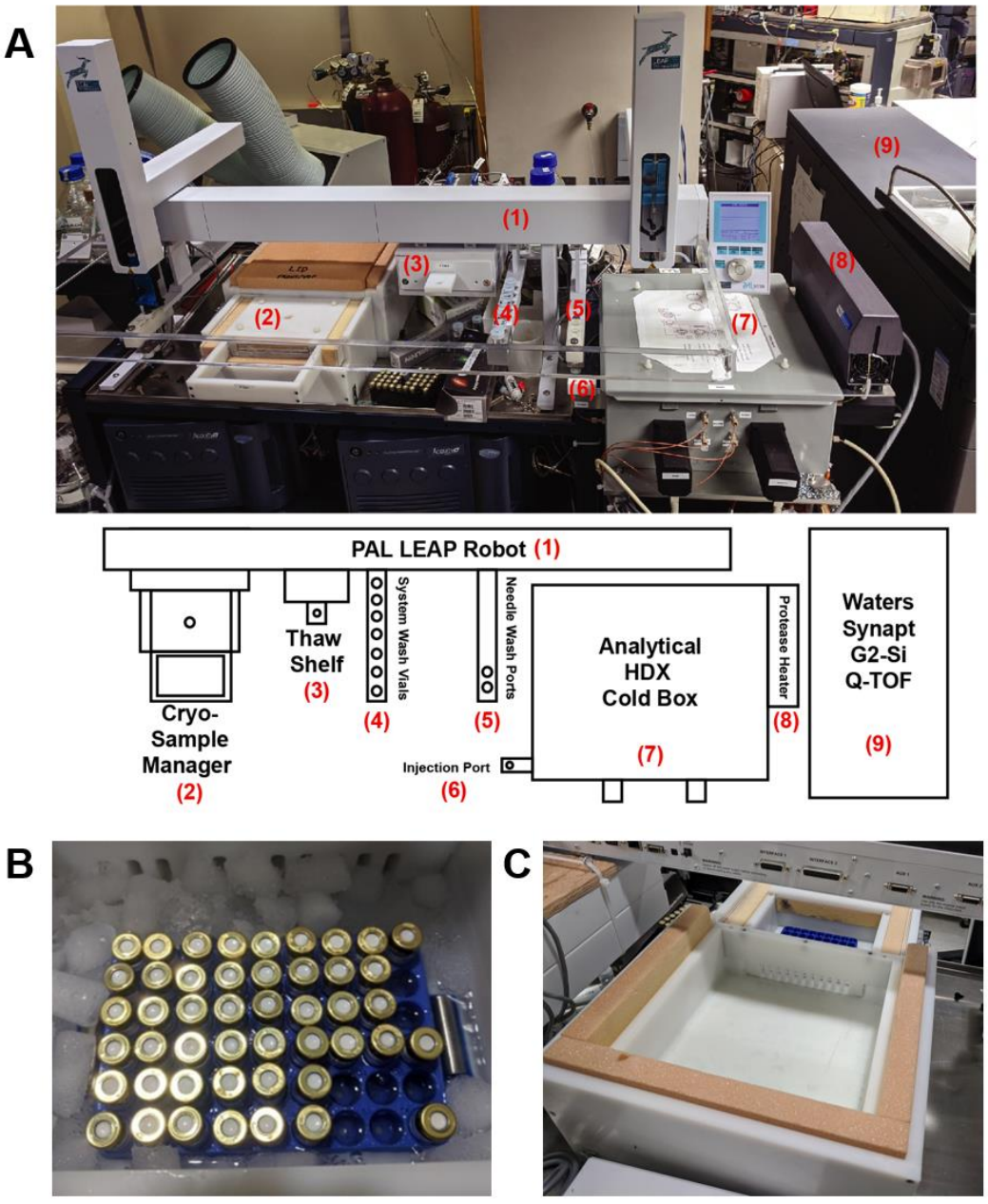


Figure 2.3.2 | PAL-integrated HDX-MS system for decoupled sample analysis

A) An overview of the automated HDX LC-MS system with major system components labeled. **B)** Closeup view of 3D printed sample tray placed within the front compartment of the sample manager **(2)**. **C)** Rear view of sample manger **(2)** showing the dry ice reservoir compartment that houses large single blocks and/or pellets of dry ice in order to keep sample tray at -60 °C or below, via the flowthrough divider with integrated solvent channels that separate the two compartments.

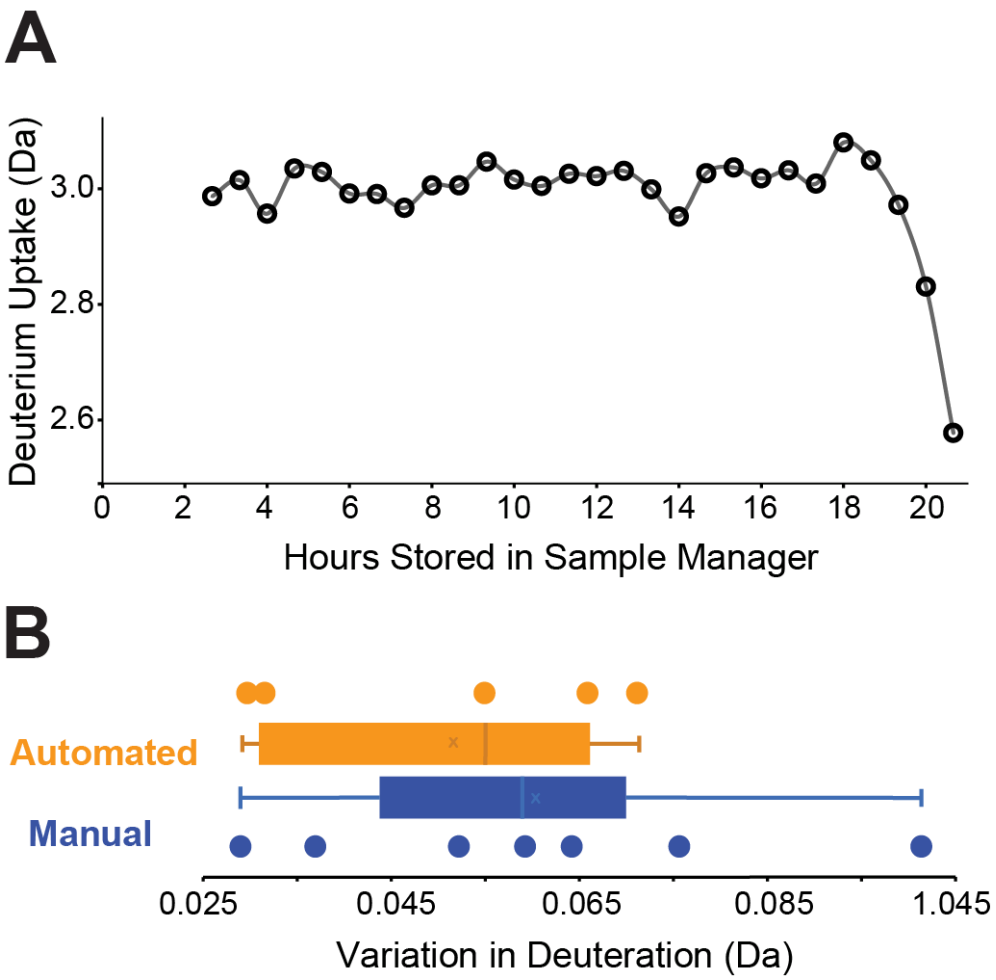


Figure 2.3.3 | Performance of subfreezing sample manager and automated HDX platform

A) Deuteration levels of internal standard Bradykinin remain consistent for each sample stored in the dry ice/ethanol subfreezing sample manager, until total depletion of the dry ice reservoir occurs after roughly 18 continuous hours. **B)** A collection of discreet data sets were analyzed to compare variability between automated and manual thaw/inject methods. Within each dataset, a fully deuterated reporter peptide was used to assess variability in back-exchange. The average standard deviation (x) for the automated and manual method was determined to be 0.051 and 0.060 Da, respectively. Sample set details are shown in Table 2.3.3.

Method	Sample Set	Sample Size	Standard Deviation	Avg. SD	Weighted Avg. SD
Robot	A	24	0.066	0.051	0.048
	B	26	0.030		
	C	24	0.071		
	D	18	0.055		
	E	73	0.032		
Manual	A	8	0.059	0.060	0.064
	B	8	0.029		
	C	7	0.037		
	D	8	0.052		
	E	8	0.101		
	F	8	0.076		
	G	8	0.064		

Table 2.3.3 | Variance comparison of robotic and manual HDX thaw/inject methods

The standard deviation of each discreet set was calculated using a single reporter peptide that becomes fully deuterated by the earliest time point (3 sec). The robotic method shows a decrease in variability by more than 16% using a simple raw average, but more than 28% when factoring in the relative size of each sample set (number of continuous injections) included. Weighted average standard deviations were calculated using the following general formula: $Avg\ SD_{weighted} = \sqrt{\frac{(n_1-1){s_1}^2+(n_2-1){s_2}^2+\ldots+(n_k-1){s_k}^2}{n_1+n_2+\ldots+n_k-k}}$, where n is the sample size and s is the standard deviation for each discreet sample set.

2.5 References

1. Yang, L. *et al.* Protein dynamics and motions in relation to their functions: several case studies and the underlying mechanisms. *J. Biomol. Struct. Dyn.* **32**, 372–393 (2014).

2. 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).

3. Houde, D., Berkowitz, S. A. & Engen, J. R. The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. *J. Pharm. Sci.* **100**, 2071–86 (2011).

4. 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).

5. Hodge, E. A., Benhaim, M. A. & Lee, K. K. Bridging protein structure, dynamics, and function using hydrogen/deuterium-exchange mass spectrometry. *Protein Sci.* **29**, 843–855 (2020).

6. Chalmers, M. J., Busby, S. A., Pascal, B. D., West, G. M. & Griffin, P. R. Differential hydrogen/deuterium exchange mass spectrometry analysis of protein-ligand interactions. *Expert Rev. Proteomics* **8**, 43–59 (2011).

7. Wei, H. *et al.* Hydrogen/deuterium exchange mass spectrometry for probing higher order structure of protein therapeutics: methodology and applications. *Drug Discov. Today* **19**, 95–102 (2014).

8. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary structure effects on peptide group hydrogen exchange. *Proteins Struct. Funct. Genet.* **17**, 75–86 (1993).

9. Englander, W. S. & Kallenbach, N. R. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **16**, 521–655 (1984).

10. Guttman, M. & Lee, K. K. Isotope Labeling of Biomolecules: Structural Analysis of Viruses by HDX-MS. *Methods Enzymol.* **566**, 405–426 (2016).

11. Ramirez-sarmiento, C. A. & Komives, E. A. Hydrogen-deuterium exchange mass spectrometry reveals folding and allostery in protein-protein interactions. *Methods* **144**, 43–52 (2023).

12. Konermann, L., Tong, X. & Pan, Y. Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. *J. Mass Spectrom.* **43**, 1021–1036 (2008).

13. Weis, D. D., Wales, T. E., Engen, J. R., Hotchko, M. & Eyck, L. F. Ten. Identification and Characterization of EX1 Kinetics in H/D Exchange Mass Spectrometry by Peak Width Analysis. *J. Am. Soc. Mass Spectrom.* **17**, 1498–1509 (2006).

14. Ferraro, D. M., Lazo, N. D. & Robertson, A. D. EX1 Hydrogen Exchange and Protein Folding †. *Biochemistry* **43**, 587–593 (2004).

15. Narang, D., Lento, C. & Wilson, D. J. HDX-MS : An Analytical Tool to Capture Protein Motion in Action. *Biomedicines* **8**, 1–20 (2020).

16. Trabjerg, E., Nazari, Z. E. & Rand, K. D. Conformational analysis of complex protein states by hydrogen / deuterium exchange mass spectrometry (HDX-MS): Challenges and emerging solutions. *Trends Anal. Chem.* **106**, 125–138 (2018).

17. Chamberlain, A. K., Handel, T. M., Marqusee, S. & Hall, S. Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH. *Nat. Struct. Biol.* **3**, 782–787 (1996).

18. Gertsman, I., Komives, E. A. & Johnson, J. E. HK97 Maturation Studied by Crystallography and H/2H Exchange Reveals the Structural Basis for Exothermic Particle Transitions. *J. Mol. Biol.* **397**, 560–574 (2010).

19. Hamuro, Y. & Coales, S. J. Optimization of Feasibility Stage for Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **29**, 623–629 (2018).

20. Hudgens, J. W. *et al.* Interlaboratory Comparison of Hydrogen-Deuterium Exchange Mass Spectrometry Measurements of the Fab Fragment of NISTmAb. *Anal. Chem.* **91**, 7336–7345 (2019).

21. Guttman, M., Weis, D. D., Engen, J. R. & Lee, K. K. Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. *J. Am. Soc. Mass Spectrom.* **24**, 1906–1912 (2013).

22. Zhang, J., Ramachandran, P., Kumar, R. & Gross, M. L. H/D Exchange Centroid Monitoring is Insufficient to Show Differences in the Behavior of Protein States. *J. Am. Soc. Mass Spectrom.* **24**, 450–453 (2013).

23. Wu, Y., Kaveti, S. & Engen, J. R. Extensive deuterium back-exchange in certain immobilized pepsin columns used for H/D exchange mass spectrometry. *Anal. Chem.* **78**, 1719–1723 (2006).

24. Hudgens, J. W., Huang, R. Y.-C. & D'Ambro, E. Method Validation and Standards in Hydrogen Exchange Mass Spectrometry. in *Hydrogen-Deuterium Exchange Mass Spectrometry: Fundamentals, Techniques, and Applications* 55–72 (John Wiley & Sons, Chichester, -1, 2016). doi:10.1002/9781118703748.ch4.

25. Anacleto, J. *et al.* Apparatus for Automated Continuous Hydrogen Deuterium Exchange Mass Spectrometry Measurements from Milliseconds to Hours. *Anal. Chem.* **95**, 4421–4428 (2023).

26. Cravello, L., Lascoux, D. & Forest, E. Use of different proteases working in acidic conditions to improve sequence coverage and resolution in hydrogen / deuterium exchange of large proteins. *Rapid Commun. Mass Spectrom.* **17**, 2387–2393 (2003).

27. Vorauer, C. *et al.* Rapid Assessment of Pepsin Column Activity for Reliable HDX-MS Studies. *J. Am. Soc. Mass Spectrom.* **32**, 2386–2390 (2021).

28. Zheng, J. *et al.* Comparative Analysis of Cleavage Specificities of Immobilized Porcine Pepsin and Nepenthesin II under Hydrogen/Deuterium Exchange Conditions. *Anal. Chem.* **92**, 11018–11028 (2020).

29. Cummins, D. J. *et al.* Two-Site Evaluation of the Repeatability and Precision of an Automated Dual-Column Hydrogen/Deuterium Exchange Mass Spectrometry Platform. *Anal. Chem.* **88**, 6607–6614 (2016).

30. Burkitt, W. & O'Connor, G. Assessment of the repeatability and reproducibility of

hydrogen/deuterium exchange mass spectrometry measurements. *Rapid Commun. Mass Spectrom.* **22**, 3893–3901 (2008).

31. Keppel, T. R., Jacques, M. E., Young, R. W., Ratzlaff, K. L. & Weis, D. D. An efficient and inexpensive refrigerated LC system for H/D exchange mass spectrometry. *J. Am. Soc. Mass Spectrom.* **22**, 1472–1476 (2011).

32. Chalmers, M. J. *et al.* Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. *Anal. Chem.* **78**, 1005–1014 (2006).

33. Woods, V. L. & Hamuro, Y. High resolution, high-throughput amide Deuterium Exchange-Mass Spectrometry (DXMS) determination of protein binding site structure and dynamics: Utility in pharmaceutical design. *J. Cell. Biochem.* **84**, 89–98 (2001).

34. Englander, J. J. *et al.* Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 7057–7062 (2003).

35. Espada, A. *et al.* A Decoupled Automation Platform for Hydrogen/Deuterium Exchange Mass Spectrometry Experiments. *J. Am. Soc. Mass Spectrom.* **30**, 2580–2583 (2019).

36. Keppel, T. R. & Weis, D. D. Analysis of disordered proteins using a simple apparatus for millisecond quench-flow H/D exchange. *Anal. Chem.* **85**, 5161–5168 (2013).

37. Walters, B. T., Mayne, L., Hinshaw, J. R., Sosnick, T. R. & Englander, S. W. Folding of a large protein at high structural resolution. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18898–18903 (2013).

38. Astorga-Wells, J., Landreh, M., Johansson, J., Bergman, T. & Jörnvall, H. A membrane cell for on-line hydrogen/deuterium exchange to study protein folding and protein-protein interactions by mass spectrometry. *Mol. Cell. Proteomics* **10**, 1–7 (2011).

39. Anderson, K. W., Gallagher, E. S. & Hudgens, J. W. Automated Removal of Phospholipids from Membrane Proteins for H/D Exchange Mass Spectrometry Workflows. *Anal. Chem.* **90**, 6409–6412 (2018).

40. Majumdar, R. *et al.* Minimizing carry-over in an online pepsin digestion system used for the H/D exchange mass spectrometric analysis of an IgG1 monoclonal antibody. *J. Am. Soc. Mass Spectrom.* **23**, 2140–2148 (2012).

41. Fang, J., Rand, K. D., Beuning, P. J. & Engen, J. R. False EX1 signatures caused by sample carryover during HX MS analyses. *Int. J. Mass Spectrom.* **302**, 19–25 (2011).

– Chapter 3 –

Defining the Functional Characteristics of Complement-Active IgM

3.1 Introduction

As outlined in Chapter 1, many classical studies in the literature have touched on the role played by the display of antigen (Ag) within the context of classical cascade activation. Past reports on IgM functionality as it relates to Ag presentation have varied in regard to physical variables such as the relative size of monomeric Ag and degree of valency displayed in polyvalent Ag. For example, the extent and particular branching arrangements of some dextran chains commonly found in bacterial polysaccharides have been demonstrated to modulate the resulting degree of IgM-mediated complement activation.[1,2] While there is now a general consensus that monomeric Ag alone is not sufficient for the induction of complement (at least not to an immunologically meaningful degree *in vivo*), many early studies have demonstrated that polyvalent antigen is capable of significant cascade activation as long as the relative densities and IgM:Ag ratios are sufficiently within an optimal range that leads to efficient C1 activation.[3] With an excess of polyvalent Ag, each Fab domain of IgM is able to bind to its own discrete Ag molecule (i.e. crosslinking), which effectively recapitulates the binding to multiple monomeric Ag particles. Conversely, if polyvalent Ag density is too low, the lack of available binding sites will lead to a significant subpopulation of IgM that is bound monovalently (at only 1 arm) and therefore unable to adopt the multivalent, complement-active 'staple' conformation that is thought to be required for sufficient recruitment of C1q and activation of C1 (refer to Chapter 1, **Figure 1.5.2**). While an immobilized Ag surface clearly remains the most potent form of Ag presentation, there are still many nuances surrounding the physical presentation and solution binding environment in which surface Ag displays are used experimentally.

To investigate the binding between IgM and surface Ag along with subsequent C1q/C1 binding interactions we used biolayer interferometry (BLI), which is a real-time, label-free instrumentation platform that uses the interference of reflected light to measure binding via the change in thickness along a biocompatible tip surface, relative to an internal reference layer. The resulting data are similar to that of other well-established ligand binding techniques such as surface plasmon resonance (SPR).[4] The binding of analyte molecules to the immobilized ligands that line the tip surface results in a direct wavelength shift that can be used to characterize the association and dissociation kinetics of antibody-antigen binding interactions, all within a convenient 'dip-and-read' format that uses microwell plates rather than the microfluidic hardware needed for SPR experiments (**Figure 3.1.1**). There are multiple tip chemistries and chemical linker options commercially available, and in our experiments we employed standard streptavidin (SA) coated tips combined with biotin Ag linkers to create uniform binding surfaces. This allowed us to investigate a number of key variables that are thought to influence the activation of complement, such as the pH and ionic strength of the solution environment, along with the impacts of attenuated surface binding valency on the IgM/C1q interaction and the role of IgM binding avidity in the activation mechanism of C1.

While binding to C1 is an obvious prerequisite to classical cascade activation, Ag binding alone is not enough to fully-inform on the potency of complement-active IgM, nor adequately differentiate the staple conformation from other potentially less-active forms of Ag bound IgM. In order to verify the biological relevance of our kinetic observations by BLI and further expand on the details surrounding the functional aspects of complement-active IgM, we also developed our own simplified complement activity assay that is capable of directly monitoring the activation rate of C1. While there are many established activity assays and experimental approaches published throughout the literature, all of them come with specific caveats and limitations that must be carefully considered before interpretations can be made with confidence. We designed our activity assay to 1) pair directly with the kinetic binding data

presented throughout this chapter, and 2) also be used throughout the structural characterization of IgM that is presented in the following chapter (Ch.4), thereby bridging the gap between the functional features and structural details of complement-active IgM.

3.2 Materials and Methods

3.2.1 Materials

Purified IgM MD4 antibody material (murine; pentameric) was generously provided by Dr. John Kearney at the University of Alabama Birmingham, and IgG H10 material (murine) was obtained in-house from the Atkins Lab at the University of Washington. Antibody material was checked for homogeneity and purity through a combination of SDS-PAGE, negative-stain EM, and SEC-MALS. Purified mouse and human complement protein components (C1q, C1r, C1s, C1, and C1-INH) were all purchased from CompTech (Complement Technology Inc., Texas, United States). C1/C1-INH combined material was prepared for use in BLI and activity assays by co-dialyzing stock C1-INH and C1 material overnight into 1X activity buffer (10 mM HEPES, 150 mM NaCl, 5mM $CaCl_2$, 1.5 mM $MgCl_2$, pH 7.3) using 5K MW CO dialysis cassettes . Hen egg lysozyme (HEL) antigen was purchased from ThermoFisher (catalog #89833) and was weighed out as a powder and resuspended in 1X HBS, followed by an initial high-speed (13 x G) spin to minimize the presence of aggregated protein material from the retained supernatant. IgG H10 Fab material was obtained through ficin digest using the materials and protocols provided by the manufacturer (Pierce Mouse IgG1 Fab and F(ab')2 Preparation Kit Catalog #44980). Western blot primary antibody against C1s (goat anti-human C1s; catalogue #A204) was purchased from CompTech, and secondary antibody (IRDye 680RD Donkey anti-goat; catalogue #923-68074) was purchased from LI-CORE and prepared according to the manufacturer's suggested protocols.

3.2.2 Antigen Preparation

Monomeric HEL antigen was prepared for use in BLI and C1 activity assays by minimal biotinylation (1:1 molar ratio) of free HEL protein using either NHS-biotin, NHS-LC-LC-biotin, or NHS-PEG_{12}-Biotin EZ-link reagents (ThermoFischer; catalogue #20217, #21343, and #21313, respectively) at room temperature for at least 30 min. The biotinylated HEL material was then filtered 3X using a 2mL VivaSpin centrifuge column (5K MW CO) at 4,000 x g for 20 minutes each to remove any loose/remaining biotin linker molecules before being combined and measured for concentration by UV-Vis using a ThermoFisher Scientific Nanodrop OneC spectrophotometer. For SA bead activity, agarose bead stock (50% slurry) was initially aliquoted into individual tubes and washed with fresh buffer 3X prior to being combined with biotinylated HEL protein Ag. Surface conjugation reactions were then performed at room temperature (RT) for 30 min on a rotator, followed by 3X max volume washes with fresh buffer, and the maximum supernatant volume was then removed carefully with a pipette in order to reach the top of the settled bed height of the SA beads.

3.2.3 Prebound IgM Preparation

IgM MD4 was preincubated with excess molar ratios of free/loose non-biotinylated HEL protein that was previously resuspended and quantified by UV absorbance (A280) for concentration prior to combining with IgM material. We refer to this condition as “prebound” IgM. Free HEL protein antigen solutions were prepared separately to their intended final concentration, based on each respective prebound ratio (Ag:IgM), before being combined 1:1 volumetrically with IgM MD4. Incubations were then conducted at RT for at least 60 minutes prior to being utilized in BLI and C1 activity assays.

3.2.4 Biolayer Interferometry (BLI)

All BLI experiments were performed using an Octet Red system (FortèBio) in combination with standard 96-well BLI microplates and streptavidin (SA) coated biosensors purchased from Sartorius. Streptavidin (SA) biosensors were pre-soaked in a standardized HBS run buffer with added calcium (10 mM HEPES, 150 mM NaCl, 5 mM $CaCl_2$, 1.5 mM $MgCl_2$ pH 7.3, supplemented with 0.1% BSA and 0.005% Tween 20) for at least 10 minutes prior to testing. A baseline reading in blank run buffer was taken for 1 minute before the hydrated SA tips were loaded with biotinylated HEL antigen (prepared at ~1 ug/mL) for 3-5 minutes in order to reach tip saturation. After recording a stable buffer baseline for 1 minute, the HEL-immobilized sensors were then dipped into wells containing specific IgM MD4, IgG H10, or IgG H10 Fab material conditions, followed by another 1 minute blank buffer baseline, at which point tips were moved into wells containing prepared hC1, hC1q, mC1q, and/or C1/C1-INH co-dialyzed material and monitored for association over the course of ~3-5 min on average. After the association phase, tips were moved into separate, fresh buffer wells for monitoring dissociation over the course of ~3-20 min on average. Binding responses were corrected for signal drift whenever possible using either a blank run buffer reference channel or negative control channel that was always collected in parallel with conditional sample. The data were analyzed using FortéBio Data Analysis 11.0 software and processed using Savitzky-Golay filtering and fitted with a basic 1:1 or 2:1 fitting model applied. Analysis of linear burst-phase binding kinetics was determined using only the initial 30 sec and 10 sec linear slope portions of IgM and C1q binding responses, respectively.

3.2.5 C1 Activity Assays

Activity assay buffer and SA beads were prepared as described in section 3.2.1 and 3.2.2, respectively. After conjugation with biotinylated PEG_{12}-HEL antigen, the SA beads were washed in triplicate with 10X volume equivalents of 1X activity buffer in order to remove any

unconjugated/unbound antigen. IgM material was then added to the tubes and again placed on a rotator for loading and equilibration (60 min), followed by a final 3X series of washes and total removal of any remaining supernatant, to the top of the settled bed height of the SA beads. Activity assays were started upon the addition of C1/C1-INH material and immediately placed into a heat block at 37° C, and aliquots were taken at the designated timepoints with gentle tube inversion dispersed regularly throughout the assay in order limit settling of the beads. For BLI-based activity, BLI biosensors were loaded normally (to saturation) with PEG_{12}-HEL antigen followed by IgM material loading. After IgM loading the instrument method was stopped, and the tips were then manually removed and placed directly into tubes prepared for C1-activity assays. Activity assays were performed as normal at 37° C, but for extended periods of time (~2-4 hours) due to lower available IgM concentrations on the BLI tip surface caused by a reduced binding capacity relative to the SA beads.

Upon removal of time point samples, aliquots were transferred into tubes prepared for reduced SDS-PAGE. After electrophoresis, gels were then prepared for western blot analysis using PVDF transfer membranes that were presoaked in methanol (1 min) and water (2 min) baths prior to sandwich assembly. Transfer solution was made from 1X Tris-Glycine buffer with 10% MeOH included, and blocking solution was made from TBSt (0.1% Tween-20) with 5% BSA added and mixed thoroughly with a magnetic stir bar for at least 2 hours prior to use. Primary antibody (α-C1s) was added in a 1:2000 dilution to blocking buffer and membranes were incubated with primary solution on a rocker for at least 2 hours (or overnight at 4 °C) prior to a series of 5 min washing steps with fresh blocking buffer on the rocker, at which point secondary antibody (~1:15,000 dilution) was added and membranes were incubated again at RT for 1 hour. Western blots were imaged at 700 nm using an Odyssey DLx (LI-COR) imaging system, and data was integrated via ImageJ gel analysis software tools.

3.3 Results

The work in this chapter aimed to characterize the binding and activation phenotype of multivalent surface bound IgM (staple) in relation to C1q and C1, which can then be paired with detailed structural insights (Ch.4) to define the structure/function relationship of complement-active IgM. Our investigational approach generally relied on the establishment of a baseline set of consistently observed kinetic features for the staple-optimized, or 'nominal' surface bound condition of IgM, in order to make comparisons against variable conditions that are thought to have significant impacts on the activation of complement. As all of our assays relied on the initial preparation and presentation of protein antigen to influence the bound conformation of IgM and subsequent C1q/C1 binding responses, the 'nominal' condition represents the maximally performing version of complement-active IgM that we could consistently reproduce experimentally.

3.3.1 – Optimization of BLI Binding Surface

IgM MD4 is based on the well-documented murine HyHEL-10 IgG1 antibody system ("H10") and likewise binds with high affinity to the same hen egg lysozyme (HEL) protein antigen, which is monomeric and highly amenable to the biotinylation reaction needed for conjugation to streptavidin (SA) coated BLI biosensors.[5–10] The quality of the resulting antigen display is largely dependent on both the location and number of biotin linkers attached, as the epitope needs to be sterically accessible to the IgM Fab domains in order to facilitate multivalent binding. Our conjugation protocol (see section 3.2 for further detail) was designed to err on the side of caution, aiming only for the *minimal* extent of biotinylation (max 1:1 molar ratio), with the understanding that any remaining HEL molecules that were not successfully biotinylated and/or conjugated to the surface would eventually be removed through initial spin filtration and wash steps. MALDI-TOF mass spectrometry was used to verify the resulting stoichiometry and indicated the majority of HEL to be biotinylated with only a single copy of the biotin

linker (**Figure 3.3.1.1 A**). Additionally, peptide mapping experiments using MS/MS data generated from a high resolution Fusion Orbitrap indicated the predominant biotinylation site to be K116, which is not obstructive to the binding epitope of the HyHEL-10 Fab domain (**Figure 3.3.1.1 B**).[5,7] As with all ligand binding assays, the Ag loading step is a crucial element in BLI as it directly impacts the rates/magnitudes of association and dissociation for every analyte that subsequently binds the tip surface. To the best of our knowledge, this is the only reported BLI binding assay with IgM and C1q that is based on the manual loading and surface presentation of a soluble monomeric protein antigen that is otherwise not pre-immobilized to the BLI tip surface by the manufacturer.

While the 'dip-and-read' format of BLI is convenient in a number of ways, there are inherent technical limitations to the precision and fine control of antigen density and spacing. For this reason, Ag loading was always allowed to reach tip saturation before advancing to subsequent analyte binding phases, which enabled us to create a reproducible binding surface for the investigation of a number of other major variables. Because the biotin binding site(s) within tetrameric streptavidin are partially buried, we first looked to determine if extending the length of the surface Ag biotin linker would improve epitope accessibility, which we hypothesized would increase the average binding valency/avidity of IgM and ultimately bolster the strength of subsequent IgM:C1q interactions.[11] With the goal of increasing the BLI binding magnitudes of both IgM and C1q/C1, we first conducted side-by-side comparisons of the standard biotin-NHS linker (1.4 nm), LC-LC-biotin (3.05 nm) and PEG_{12}-Biotin (5.6 nm). As shown in **Figure 3.3.1.2,** both of the extended linkers effectively doubled the observed IgM binding response, but the PEG_{12} linker consistently produced the highest overall hC1q binding magnitudes and was therefore chosen to be the nominal Ag linker for all surface binding experiments presented throughout this dissertation.

3.3.2 – Variables of the BLI Binding Environment

While the switch to an extended linker did enhance the binding responses of both IgM and C1q, there are still many key variables present within the surrounding solvent environment that have been reported throughout the literature to have a major impact on C1q binding and C1 activation potential. Although some major gaps in knowledge concerning C1q functionality still remain, the promiscuity of C1q *in vivo* has been well-established by a multitude of reports regarding C1q binding interactions with a wide array of endogenous ligands. The diversity of C1q interactions is due largely in part to the high degree of binding valency avidity afforded by its hexameric architecture, along with the composition of its six globular heads that feature a wide variety of surface residues that are considered to be both modular and complementary in nature. Although the molecule is arranged into a manner resembling that of many similar homotrimers, C1q is technically a heterotrimer composed of three distinct polypeptide chains (A, B, and C) that each provide a unique (i.e. 'modular') composition of surface residues to the C-terminal globular heads. The combination of all three chains together, however, enables a wide array of binding partners even under diverse solution state conditions, such as the acidic and basic pH environments found in inflammatory and apoptotic environments, respectively (refer to Chapter 1, **Figure 1.4.1**).[12]

C1q binding interactions are considered to be predominantly governed by electrostatic forces.[12,13] We confirmed this to be the case through BLI by monitoring the binding interaction of IgM and human C1q (hC1q) within a series of run buffers of decreasing ionic strength, where we observed a significant and step-wise increase in max hC1q binding magnitudes with each decrease in run buffer salt concentration, starting from physiological conditions (~150 mM NaCl, pH 7.3) which was the highest ionic strength tested and weakest subsequent binding environment observed (**Figure 3.3.2.1 A**). We were also able to 'ionically lock' the IgM-bound C1q into place by conducting dissociation phases within

buffers of decreased ionic strength, as illustrated in **part B of Figure 3.3.2.1**. Decreasing the ionic strength of the C1q buffer was also intentionally used in other experiments to increase the separation between conditions and amplify the effects of other tested variables. The varying A, B, and C chain sequences that compose the surface-exposed residues of the C1q globular heads give rise to optimal pH ranges that promote a variety of protein:protein interactions and modify the extent of many C1q effector functions. We observed the C1q molecule as a whole to display optimal binding to IgM within a pH range of 6-6.5, which follows a similar pattern to the reported preferential binding trends of the A and C chains (**Figure 3.2.2.2**).[13]

Besides pH and ionic strength, another significant solution variable is the presence of calcium ions (Ca^{2+}). While some confusion exists within the literature surrounding the mechanistic details and functional role of calcium within the context of C1 activation, it's clear from the C1q crystal structure that a binding pocket for Ca^{2+} exists within a solvent channel at the N-terminus of the molecule, which is also the case for other similar family proteins, such as collagen X which contains four Ca^{2+} binding pockets buried within in its core.[12,14] It is currently unclear whether calcium is required for the mechanism of C1 activation, or if its presence is related more to target recognition by C1q only; while there appears to be some cases in which the presence of calcium is strongly associated with complement activity, such as the interaction between C1q and HIV envelope protein gp41, removal of calcium via EDTA treatment does not appear to have any effect on interactions between C1q and other endogenous ligands such as fibromodulin.[14–16] To assess the role of calcium and its effects within our IgM:Ag (MD4:HEL) system we used BLI to monitor hC1q binding to IgM within run buffers spiked with EDTA, and did not observe any attenuation to binding magnitudes or changes in binding profiles (**Figure 3.3.2.3**). This suggests that, while calcium may indeed be required for the C1 activation mechanism itself, it is not necessary for initial target recognition and binding by C1q, at least in the case IgM.

3.3.3 – Kinetic Analysis of Staple IgM

To gain a better understanding of the complex kinetics of our IgM system we next used BLI to 1) quantify the total surface binding affinity (i.e. avidity) of IgM MD4 to our optimized PEG_{12}-HEL Ag surface, and 2) further characterize the highly multivalent binding interaction between surface bound IgM and C1q/C1. Upon initial testing it became clear that the binding strength of our IgM material was extremely high, as off-rates could not be practically observed. While this result was somewhat expected due to the transgenic nature of our IgM material (which combines the binding potency and specificity of IgG1 Fab domains with the highly polyvalent Fc architecture of pentameric IgM), the combined avidity of decavalent MD4 surface binding could not be adequately fit in order to obtain accurate K_{on} and K_{off} values. Consequently, the resulting affinity (K_D) could only be approximated to the sub-picomolar range ($K_D < 1.0E^{-12}$ M ± $1.39E^{-11}$), as shown in part A of **Figure 3.3.3.1.**

To circumvent this issue and improve our grasp on the global affinity of our system we turned to the IgG1 equivalent of IgM MD4, IgG "H10", and used it for further analysis to measure the kinetics of bivalent (whole IgG) and monovalent (Fab only) Ag binding interactions.[6,9] The bivalent affinity was first evaluated by BLI using the fully-intact H10 IgG antibody, and the global K_D was calculated (1:1 fitting) to be $9.53E^{-11}$ M ± $2.44E^{-11}$ (**Figure 3.3.3.1 B**), which is a close match to previously reported values generated from particle concentration fluorescence immunoassay (PCFIA) and surface plasmon resonance (SPR) experiments.[7,9] In order to measure the monovalent Fab:Ag binding affinity individual Fab domains were first isolated and purified from IgG H10 via ficin digest, which is a sulfhydryl protease (similar to papain) that cleaves the hinge regions of IgG1 antibodies and is often used specifically for mouse antibodies to liberate Fab and/or $F(ab')_2$ fragments.[17,18] After purification via protein A spin columns to remove the Fc byproduct, the monovalent binding affinity (1:1 fitting) of H10 Fab to our PEG_{12}-HEL Ag surface was then tested by BLI, and the global K_D was calculated to be $1.83E^{-09}$ M ± $1.92E^{-11}$

(**Figure 3.3.3.1 C**). With the drastic change in binding affinity observed between monovalent H10 Fab and bivalent H10 IgG (~100X increase), it follows that a decavalent (10 Fab) IgM version of the molecule would exhibit extremely tight binding in practice. We note that, while the avidity of our IgM MD4 material is likely not representative of the relatively-low binding affinities of induced IgM found *in vivo*, (due to the enhanced specificity/potency of the IgG Fabs) the functional characterization of its interaction with C1q and C1 is still biologically relevant and mechanistically informative.

The complexity of fitting high-valency binding models is even more pronounced when it comes to characterizing the IgM:C1q interaction, as C1q is hexameric and it is unclear if IgMp displays five or ten (six or twelve for IgMh) total copies of the C1q binding site. Based on the most recent Cryo-EM structures of the human IgMp Fc core, however, it seems likely that only five sites are actively displayed as the μ heavy domain Fc dimers appear to be arranged in an alternating '1 up/1 down' configuration when the molecule is bound multivalently to an immobilized surface.[19–22] To characterize the IgM:C1q interaction within our BLI system we first performed dilution series experiments with purified mouse and human C1q proteins (mC1q and hC1q, respectively). Despite the relatively 'fast on/fast off' nature of the binding interactions, we approximated the affinity to still be within the nanomolar range for human C1q (hC1q; $6.98E^{-07}$ M ± $1.92E^{-06}$) and mouse C1q (mC1q; $7.97E^{-09}$ M ± $3.06E^{-09}$) binding to mouse IgM MD4 (using 2:1 heterogeneous ligand fitting). These kinetic estimations, along with the general 'fast on/fast off' nature of the binding interaction as a whole, are also in close agreement with recent reports of human recombinant IgM binding to purified hC1q material.[23] We emphasize, however, that mC1q displayed ~100X greater affinity than hC1q (**Figure 3.3.3.2**) to our murine IgM system, which could be indicative of species' mismatch effects and corroborative of past observations that have reported varying degrees of complement activity when IgM and complement proteins are derived from different mammalian sources.[24]

After testing the basic interaction between surface bound IgM and C1q we then looked for qualitative changes in binding profiles between C1q and the intact C1 complex, which is the initiating molecule of the classical complement pathway formed by the non-covalent association of C1q, C1r, and C1s in a 1:2:2 ratio, respectively. To improve the physiological relevance of our binding assays, we also included comparisons with C1 material that was co-dialyzed with a slight excess (~2:1) molar ratio of C1 inhibitor protease (C1-INH), which is the only known protease inhibitor that directly regulates classical cascade activity and is naturally present in excess molar abundance (~4-5X) to C1r and C1s.[25] While we observed the binding profile of the C1 complex to generally resemble the 'fast on/fast off' nature of C1q itself (**Figure 3.3.3.3 A**), the added presence of the C1-INH also induced significant changes to the overall kinetic profile of the IgM:C1 binding interaction, as illustrated in part B of **Figure 3.3.3.3**. Although initial burst-phase association rates appeared to be only slightly reduced, the addition of C1-INH created an unexpected secondary binding event that appeared shortly after the initial burst-rate phase, and also greatly reduced the rate of C1 dissociation. While deconvolution of this additional interaction is untenable by BLI alone, we hypothesize the secondary binding event to be reflective of surrounding C1-INH molecules binding to activated C1r/C1s proteases that have yet to be released from the surface of the tip, which is an event that would be likely observable by BLI. While this specific change in the binding phenotype remains outside the scope of the studies currently presented, we do plan on re-visiting this concept in detail within future experiments.

3.3.4 – C1 Binding is Dictated by IgM Surface Binding Valency

The binding avidity of a multivalent antibody is typically defined as the sum of its individual binding affinities. In the case of IgM, high surface binding valency and combined interaction avidity is likely fundamental to its functional role within the immune system. Utilizing the enhanced binding strength of our transgenic IgM MD4 system, we next employed a prebinding strategy to assess the impact of

reduced IgM surface binding valency by measuring the change in HEL:IgM and IgM:C1q binding kinetics. Without having fine control over the precision of antigen spacing, and further constrained by the need to keep IgM and C1q in their physically-native states, we decided to alter the IgM in a physically non-invasive manner by incubating it with excess molar ratios of free, non-biotinylated monomeric HEL antigen prior to testing. The nanomolar Fab:Ag monovalent binding affinity afforded a sufficient temporal range in which we could effectively remove discrete proportions of participating IgM Fab domains and measure the gross changes in both IgM and subsequent C1q binding kinetics compared with the nominal, non-prebound (0:1, Ag:IgM) form of MD4 that is able to bind the BLI tip surface with maximum valency/avidity (also referred to as "max-avid"). We do note, however, that each prebound condition likely contained an inherent degree of heterogeneity. As illustrated in **Figure 3.3.4.1,** each prebound ratio is likely to have varying distributions of bound subpopulations, differing further still in the final configurations of prebound HEL within individual IgM molecules.

This prebinding approach allowed us to further probe the functional impacts of reduced IgM surface binding valency/avidity within the context of C1 activation, without having to invoke any physically invasive alterations such as residue mutations or chemical cross-linking. We hypothesized that the inherent flexibility of IgM would still allow for measurable BLI surface binding at excess Ag:IgM ratios up to and including 9:1, in which only one of the ten Fab domains would be theoretically available on average to bind the tip surface, as illustrated in **part A of Figure 3.3.4.2**. While each condition likely contained smaller subpopulations of varying prebound states, the incubations were performed at room temperature and given excess time for equilibration prior to testing (>30 min RT; based on H10 Fab binding measurements of K_{on} = 1.07E^{05} Ms^{-1}). Furthermore, any IgM molecules that had become fully saturated during prebinding (10:1) were expected to be precluded from participation, as they would be unable to bind to the BLI tip during IgM loading phases.

We observed IgM to be readily able to bind the BLI tip surface when prebound at excess Ag ratios up to and including 10:1 (**Figure 3.3.4.2 B, left panel**), at which point the average interaction between participating IgM molecules and surface Ag at the BLI tip is expected to be predominantly monovalent, evidenced further by the sharp drop in binding signal observed between the 9:1 and 10:1 prebound ratio conditions. We also observed a notable degree of IgM dissociation during the 60 sec buffer baseline that followed immediately after IgM loading, which became noticeable at 6:1 and continued increasing in magnitude up to the 10:1 prebound condition, providing additional evidence of a step-wise decrease in IgM surface binding valency/avidity. Despite the continual increase in total IgM binding magnitudes, the responding hC1q binding magnitudes continually decreased (**Figure 3.3.4.2 B, right panel**), and linear burst-phase association rates for both IgM and hC1q were attenuated within each increasing prebound IgM ratio (**Figure 3.3.4.2 C**). Neither IgM binding nor hC1q association was observed within the 20:1 condition, which was used as a negative control. The drop in IgM binding efficiency (determined by the ratio of max C1q to IgM binding magnitudes for each condition) was progressive in nature, and there was no indication of a 'hard cutoff' observed for the number of IgM arms required to be bound to the BLI tip surface in order to induce measurable hC1q binding (**Figure 3.3.4.2 D**).

To our surprise, the progressive decrease in hC1q binding efficiency was not offset by the progressive increase in the total IgM binding magnitude seen within each increasing prebound ratio. In addition to the mass contributed by the prebound HEL molecules themselves, we hypothesized the rising trend in IgM binding signal to be caused primarily by net increases in the total number of IgM molecules bound to the tip within each increasing prebound ratio. We predicted that as IgM binding valency was progressively reduced through prebinding, the average IgM molecule would consequently take up lesser amounts of available surface Ag which, in turn, would then allow for more IgM molecules to bind to the tip surface in total. To ensure that 'avidity effects' were not the cause of optical artifacts that may have influenced the signal within our prebound IgM experiments, we designed a follow-up BLI quantitation

experiment to assess the relative amounts of IgM molecules bound to the tip, at saturation, within a stepped series of increasing prebound ratios. After loading three distinct prebound IgM ratios (3,6, and 9:1) along with three nominal/non-prebound conditions at different starting concentrations (15, 20 and 25 nM), the tips were moved into wells containing α-IgM IgG antibody and the binding responses were then recorded in parallel over the course of 5 minutes. As shown in **part E of Figure 3.3.4.2**, the α-IgM IgG binding response was found to be significantly increased within each increasing prebound ratio, and was completely overlapped for all three nominal (max-avid) conditions, which provided direct evidence in support of our hypothesized binding valency concept. However, because the BLI detector technically measures the change in optical *thickness* across the entire 2D tip surface, we note that the bound IgM conformation itself should also be considered, since IgM bound by only a single arm is likely more laterally extended from the tip surface than IgM that is bound by multiple arms.

3.3.5 – Measuring Complement Activity

The inherent complexities of the complement cascade – the convergence of its multiple initiation pathways, the number of required steps and total protein components, along with the added complications of working with serum as a complement source – has historically made the measurement and interpretation of complement activity a technically challenging endeavor. While a handful of general approaches have been described throughout the literature, all of them come with technical caveats that are often difficult and/or tedious to control for reliably. For example, complement activity assays that rely on endpoints such as cell lysis of sensitized sheep erythrocytes (determined via EC_{50} measurements) depend on the removal of many non-specific factors contained in serum that might otherwise contribute directly toward cell lysis and/or non-specific activation of parallel complement pathways.[25] Controlling the non-specific activation of complement can be difficult to achieve in serum, and reconstitution of a

given pathway from purified protein components is challenging as many of the proteins themselves are difficult to extract and purify from serum.[25,26]

To further expand on the functional characterization of complement-active IgM and verify the biological relevance of our BLI binding experiments, we also sought to create our own complement activity assay that would accurately and specifically report on C1 activation by surface bound IgM. To that end we developed a simplified C1 activation assay (**Figure 3.3.5.1**) that reports directly on the rate of C1 activity by utilizing a straightforward western blot approach to measure the conversion of the inactive 'zymogen' form of the C1s protease (also referred to as "ProC1s") into its active proteolytic fragments (C1sα and C1sβ), which is an endpoint that relies specifically on the activation of ProC1r, mediated by C1 binding (via C1q) to IgM. While the added complexity of working with serum is not needed for the setup of our assay, minor optimizations of commercially available C1 material were still required. Because C1 is prone to self-activation, most commercial vendors have formulated their purified C1 products to contain additives (such as protease inhibitors) that slow or prevent self-activation of C1, even when kept at 4 °C. As previously mentioned, the C1-INH is present *in vivo* in excess abundance to that of C1r and C1s, and while it is known to act as a potent suicide inhibitor of both active proteases, its relationship with inactive C1 appears to be supportive in nature and is actually thought to confer increased stability to the C1 complex itself, slowing its auto-activation rate.[25,27,28] We therefore chose to co-dialyze our C1 material with a two-fold molar excess of purified C1-INH (also commercially available), which proved to be successful at minimizing the self-activation rate of C1, even when placed at 37 °C for multiple hours during the course of an activity assay.

After initially developing the method using a pre-established approach of heat-aggregated IgG1 material (HAGG) as a catalyst for C1 complement activation, we next looked to incorporate IgM and expand the assay to include SA-coated agarose beads, which are also the antigen displays used throughout the structural work presented in the following chapter.[25] The SA beads were prepared in an

analogous fashion to the SA biosensors used in BLI (see section 3.2 for more detail), and were also found to be a highly-active and specific Ag display mediator of IgM induced C1 activation (**Figure 3.3.5.2 A**). The relative C1 activation rates of both highly prebound (~9:1) surface and solution bound IgM were also investigated, and both were observed to activate C1 at a reduced rate to that of the nominal surface bound form, with solution bound IgM displaying the lowest activity overall (**Figure 3.3.5.2 B**). While the activation rate of prebound IgM was higher than we expected, allowing the prebinding incubations to equilibrate for an extended period of time (overnight at 4° C) did appear to result in better separation between the nominal and prebound surface IgM conditions (**Figure 3.3.5.2 C**). While it's possible that sufficient equilibration of the prebound condition had not been reached within the initial experiment, we still expect 30 minutes at room temperature to be sufficient for reaching binding equilibrium given the fast on-rates observed during analysis of the monovalent H10 Fab material. Instead, we find it more likely that our assay itself is inherently prone to significant day-to-day variability, given the relatively imprecise nature of gel-based quantitation along with other caveats that come from the western blot assay format, such as incomplete membrane transfers and/or optical artifacts that arise during imaging. We therefore repeated the activity assay with a focus on solution state forms of IgM, in order to ensure that 1) the antibody itself was relatively inactive in the unbound (apo) state, and that 2) our prior observations of a low-level, basal degree of activity in solution bound IgM would continue to persist and remain significant in comparison to the unbound form and negative controls. As expected, unbound IgM was repeatedly observed to be devoid of any measurable C1 activity relative to positive controls. Importantly, solution bound IgM also continued to exhibit a low, but significant degree of C1 activation potential (**Figure 3.3.5.2 D**).

Finally, to validate the biological relevance of our BLI binding experiments, we developed a novel hybrid approach that combines the real-time binding data of BLI with our newly developed C1 activity assay. As an initial proof-of-concept, the standard BLI SA tips were loaded normally with saturating

amounts of PEG_{12}-HEL antigen before being dipped into wells with and without IgM present. After reaching saturation, the BLI instrument was halted, and the tips were then physically removed and transferred immediately into tubes prepared for activity assays at 37 °C, and sampled with regular time point aliquots over the course of 4 continuous hours. As illustrated in **part A of Figure 3.3.5.3**, this experiment successfully demonstrated the specific C1 activity of nominal (max-avid) IgM bound to the BLI tip surface, adding further confidence to the biological relevance of the BLI data generated throughout our kinetic characterizations of surface bound IgM. After validating the activity of the nominal IgM condition formed within our BLI system, we next looked to assess the effects of reduced IgM binding valency/avidity on C1 activation by repeating the hybrid activity assay on BLI tips loaded with highly prebound (9:1) IgM. Based on our previous prebinding experiments, we expected the change in C1 activity to trend in a similar fashion to that of the binding kinetics, with reduced surface binding valency ultimately leading to diminished C1 activation rates. As expected, prebound surface IgM displayed significantly reduced C1 activity compared to that of the nominal surface bound condition (**Figure 3.3.5.3 B**). However, despite the clear reduction in C1 activity, there was still nonetheless a measurable amount of activity observed within the prebound condition, relative to the negative (-IgM) control.

3.4 Discussion

3.4.1 – Optimization of BLI binding Surface and Buffer Environment

The presentation of surface antigen has been demonstrated throughout the literature to be fundamental to the binding conformation of IgM and subsequent activation of complement. For this reason we hypothesized the method format of BLI to be particularly well-suited for investigating the nuanced interaction of IgM with C1q and C1, as it allowed us to create a biologically relevant binding surface from monomeric protein antigen, under controlled solution environments that could be further

manipulated to study the functional characteristics of IgM staple formation – without having to make any physical alterations to IgM or C1q/C1 that may potentially impact their physical interaction(s) with one another.

The overarching goal throughout BLI method development was to optimize parameters towards the promotion of IgM staple, with the leading field hypothesis being that only the staple conformation is capable of fully binding/activating C1. While the switch to an extended linker did result in stronger IgM and subsequent C1q binding signals, we note that extension of the surface Ag linker could also lead to unintended consequences within the context of cascade completion (rather than merely pathway activation). A recent study from Oskam et. al. (2022) specifically examined the impacts of surface Ag linker length on the IgM/IgG interaction with C1, and demonstrated that linker extension actually reduced complement activity as measured by cell lysis end points.[29] In 2019, Sharp et al. performed cryo-EM with tomography on surface bound IgM bound to C1 and observed a number of potential interactions between the Fab domains of IgM and portions of the active C4b fragment, which is the first protein component of the classical cascade that attaches to the targeted membrane for the base of MAC formation.[19] While IgM staple formation may be significantly promoted with the use of extended Ag linkers that improve epitope accessibility, cascade completion itself may suffer if the required complement proteins, especially those destined for anchoring to the target membrane (like C4b or C2a), are simply too far away from the membrane itself. Interpretations of activity, therefore, may be skewed depending on how complement activity is measured; if the monitored endpoint is sufficiently downstream (e.g. cell lysis or C5b9 deposition), then it may inadvertently appear as if the cascade were not initiated at all (a false negative result). On the other hand, if C1q fixation (i.e. 'sequestering') is the only endpoint measured, extrapolation to cascade completion or even comparisons with orthogonal cell lysis assays could unintentionally paint a false picture of high activity when, in reality, C1q binding may not have successfully led to downstream MAC formation.[27,30,31]

The A and C polypeptide chain 'modules' of C1q are similar in that they both display a prominent mix of acidic and basic residues along their surfaces, while the global charge of the B module is distinctly net-positive and features significant patches of surface-exposed arginine residues that have previously been implicated in mediating the binding interaction with IgG (**refer to Figure 1.4.1**).[12–14,32–34] The range of electrostatic potential across each of the three primary chains of C1q contributes largely to the range of endogenous ligands that C1q is known to interact with, and it is now well-established that C1q interactions are mediated predominantly through ionic pairings. Our BLI testing of the IgM:C1q interaction adds further support to this idea, and demonstrates the significant role that salt concentration plays in modulating the strength of the binding interaction, as the weakest binding interaction we observed in BLI occurred at the highest salt buffer concentration tested, which was at physiological (~150 mM NaCl) conditions. The sequence variety of the globular head surface also gives rise to optimal pH ranges that vary depending on the specific ligand and local environment in which the interaction takes place. This is especially relevant within the biological context of the complement cascade itself, as the inflammation that propagates throughout cascade progression also lowers the local pH of the C1 binding environment.[35,36] In addition to functioning within the acidic conditions of pro-inflammatory environments, C1q is also known to play a role in the recognition and clearance of apoptotic cells, wherein the local pH environment is increased to alkaline conditions.[12,13,34,37]

In a 2006 study published by Zlatarova et. al., the A, B, and C globular head modules of C1q were each produced recombinantly and found to bind optimally at different pH ranges when binding to IgG and IgM antibodies. They also found the C1q molecule as a whole to bind optimally to IgG only under relatively acidic conditions (pH ~5.0), while binding to IgM was observed to be generally high throughout the entire pH range tested, with significant binding seen at pH ~4.5 and maximal binding found at pH ~8.5. As with the case of electrostatic potential, the A and C modules were both found to bind optimally within lower pH ranges (max binding at pH ~5.0) while the B module uniquely displayed maximal binding

at a higher pH ranges (max binding at pH ~8.5).[13] While we observed the C1q molecule as whole to reflect the preferential binding trends of the A and C chains, with optimal binding observed in slight acidic run buffers (pH ~6.0 and ~6.5), the B chain is expected to add a significant range of diversity to the molecule and is also thought to be the primary mediator of the IgG:C1q binding interaction, with its distinct concentration of surface arginine residues found to contribute at least three major points of contact (ArgB101, ArgB114, and ArgB129) with the Cγ2 domain.[13] We found pH effects on the IgM:C1q interaction to be generally minor (especially compared to the impact of ionic strength), which is corroborative of past reports and indicative of the tolerance range of the C1q globular heads. While the binding site(s) involved in the IgM-C1q binding interaction have yet to be elucidated, data presented in a 2012 follow-up paper from the same group suggested a possible overlap in the IgG/IgM binding sites on C1q, in addition to unique locations in the globular head residues that are not shared between the two antibody classes.[32] Structural details concerning the identity of potential C1q binding site(s) on IgM will be discussed further in the following chapter (Ch. 4).

3.4.2 – Kinetic Assessment of IgM Binding to Surface Ag and C1q/C1

Although the 'fast on/fast off' nature of the IgM:C1q binding interaction came as a surprise to us initially, we hypothesize this type of interaction to be advantageous within the general context of biochemical signaling pathways *in vivo*. The complement cascade, much like other immune signaling pathways such as the T cell response to antigen-presenting cells (APCs), appears to be largely dependent on the speed of initiation, degree of signal amplification, and extent to which recruitment signals are propagated – all of which are critical to overcoming regulatory thresholds and ultimately advancing the cascade towards successful completion.[38–41] The high binding valency of C1q likely increases the speed of burst-phase association while also extending its bound half-life, and it is also possible that C1q could

mimic the 'antibody walking' phenomenon that is known to take place for IgG antibodies bound along antigenic surfaces of sufficient density, which also readily applies to IgM surface binding.[42]

Despite the relatively 'fast on/fast off' nature of the binding interaction, we approximated the global binding affinities to still be within the nanomolar range for both human C1q (hC1q) and mouse C1q (mC1q) binding to IgM. It should be emphasized, however, that mC1q displayed a roughly 100-fold greater affinity for MD4 (which is also murine-based) compared to that of hC1q, which provides kinetic evidence that may help to explain the differential activity reported in the literature stemming from species mismatches between antibody and complement sources.[24,32] In the case of IgG, the exact residues and locations of the known C1q binding sites differ between mouse and human IgG isoforms, and while the binding sites are still unknown for IgM, one factor that may contribute towards its species-dependent binding effects with C1 is heterogeneity of the conserved N-linked glycan locations within the Fc. This includes differences in the compositions of glycans that are commonly expressed at each location, along with their relative occupancy rates.[43–48] Furthermore, deglycosylation of IgG via PNGase-F treatment was recently shown to abrogate C1q binding entirely, indicating that the sugar moiety itself may be significantly involved within the IgG-mediated activation of complement.[49] Structural insights regarding the potential involvement of the six conserved N-linked glycans within our mouse IgM will be discussed in detail in the following chapter (Chapter 4).

Depending on the strength of the specific Ag:IgM interaction involved, the 'fast on/fast off' kinetic quality of the IgM:C1q/C1 binding interactions that we observed lead us to envision a potential biological scenario in which multiple 'rounds' of activation take place, effectively catalyzed by the presence of IgM staple. The combined strength of IgM binding avidity could perhaps enable a scenario in which several C1 molecules are individually recruited, activated, and released from a single IgM molecule within a sufficient amount of time to amplify the biochemical signal and overcome initial regulatory thresholds, such as the proximal abundance of the C1 inhibitor (C1-INH). The C1-INH

molecule is the only known protease inhibitor that directly regulates complement activity, and is present in human plasma at roughly 4-5X the molar concentrations of C1r and C1s (~240 ug/mL; ~3 uM).[50–53] During BLI analysis we observed a significant change to the binding profile of C1 in the presence of excess C1-INH, which we predict to be reflective of the high degree of inhibitory potency that has been reported throughout the literature.

While the majority of details surrounding the classical activation mechanism are poorly understood, the current understanding within the field of immunology is that C1 becomes fully-activated upon a mechanically-induced cleavage/activation of the proC1r dimer, which only specifically takes place after C1q interacts with a viable Ag:Ab immune complex. The only currently known function of the C1r protease is the cleavage and activation of proC1s, and it is believed that both activated fragments of C1r and C1s are able to dissociate from the C1q 'housing' upon fragmentation, albeit details surrounding the post-activation clearance mechanisms are mostly still missing.[53] What is known, however, is that the C1-INH molecule is highly abundant *in vivo* relative to C1r and C1s, and that it acts quickly as a potent suicide inhibitor to halt the activity of both proteases, with reported measurements placing the functional half-life of C1 at only 13 seconds under physiological conditions.[41] The binding of C1-INH to active C1 may actually be the primary mechanistic driving force behind dissociation of C1r and C1s from antibody-bound C1q, with the resulting inhibition products containing upwards of four molecules in total (C1-INH–C1r–C1s–C1-INH).[41,54] Our BLI results appear to corroborate the fast-acting nature of C1-INH, and we plan to revisit this topic in future experiments to further explore the role of C1-INH within the initial activation mechanism of C1 by IgM.

3.4.3 – The Mechanistic Role of Avidity Within the IgM-Mediated Activation of C1

Even if the assumption is made that all five arms in pentameric IgM are equal in terms of their functional binding capacity, there still remains a significant number of permutations within the total set

of possible configurations of each discrete prebound state, as illustrated previously in **Figure 3.3.4.2**. In reality, it's possible that the arms in IgMp are not all functionally equivalent, as at least one of the two arms that connect to the J chain may be hindered in its relative flexibility compared to the other arms, which was suggested recently by Chen et. al. (2022).[55] Based on our binding and activity assay experiments using prebound IgM, we hypothesize that IgM surface binding valency/avidity is mechanistically linked to the conformational changes required to bind and activate C1q/C1. Our results suggest that IgM staple formation is likely not a globally-concerted conformational change. Rather, the staple conformation of IgM may be better viewed instead as an *ensemble* of local transitions that each arm contributes to in a semi-independent manner, with local formations at each arm leading to a combined and multivalent display of C1q binding sites. The activation potency of surface bound IgM molecules would therefore be expected to decrease as more arms are prevented from participating in binding to the Ag surface, which is what we observed both kinetically and functionally throughout our testing with prebound IgM material through our BLI experiments and C1 activity assays, respectively.

C1 appears to be an easily triggered protein complex, which could ultimately be a consequence of the high binding valency and promiscuity of the hexameric C1q molecule. In 1975, Hurst et. al. concluded that "*tight*" binding of C1 was not a prerequisite for effective complement activation as measured by C4 consumption, which is the next major stage of classical cascade progression following the initial production of active C1s protease.[56] After demonstrating that simple (and even chemically-reduced) peptide fragments of the Cμ4 domain were capable of activating complement, they reasoned that transient binding of C1 may be sufficient to induce activation. This conclusion matches well with our own observations and can potentially help to explain the various levels of C1 activity that we observed from all of our tested forms of Ag bound IgM, including the minimally-avid prebound IgM conditions along with the strictly monovalent, solution bound form of IgM.

While binding to a surface could induce changes amenable to the exposure of C1q binding sites, the activity of solution bound IgM is harder to explain given the leading field hypothesis that a fixed Ag surface is required for staple formation. Although the potential presence of aggregated HEL molecules in our activity experiments could inadvertently add to the effective C1 activity rates that we observed for the solution bound state, the enhanced binding affinity of our IgM system is likely a much more significant caveat. We emphasize that, while the C1 activity rates we observed for solution bound IgM are likely *pronounced* by the enhanced binding affinity of our transgenic IgM system, the activity itself appears to nonetheless be a real biophysical consequence of Ag binding. We hypothesize this activity to ultimately stem from allosteric effects, which will be discussed in further detail throughout the following chapter (Ch. 4). Binding to soluble antigen could potentially disrupt pre-existing Fab-Fab and/or Fab-Fc interactions that help to stabilize the unbound form of IgM and prevent it from transiently sampling local staple formations due to an increase in segmental flexibility.[54]

Since prebound IgM molecules are unable to bind to the Ag surface with maximal valency, the fact that activity was still measurable ultimately infers that the staple conformation of IgM is <u>not</u> a globally distinct or cooperative conformational change. We also did not observe any hard cutoff in the number of arms/Fabs required to bind the surface for measurable C1q binding, which further motivates us to consider staple formation as a local structural quality, rather than a globally-concerted conformational change; while the max-avid nominal form of surface bound IgM was clearly the most efficient modality we observed, we could not find a sufficient trend in the data that fit with any hypothetical cutoff of four, three, or even two available binding arms. While this result could be attributed to the inherent heterogeneity of prebound IgM population states (e.g. activity contributions from 7:1 or 8:1 prebound IgM molecules), it could also be indicative of the significant flexibility inherent to IgM and/or the low interaction threshold that is physically required for C1 activation.

Because each arm has a certain degree of independence in terms of segmental flexibility and relative Fab-Fc motion, we propose that the complement-active form of surface bound IgM is better viewed as an *ensemble* of local staple formations that are contributed by each arm in a semi-independent manner, rather than an absolute/globally-cooperative transition induced by a single mechanistic driving force. If allosterism alone was enough of a driving force to induce the full staple ensemble in solution bound IgM, there would likely not be a significant difference in activity between surface and solution bound forms, and the *in vivo* consequences of binding to soluble monomeric Ag would likely be catastrophic. In the same way that IgM depends on the multivalent display of antigen to form an effective Ag:Ab immune complex, C1q depends on the multivalent display of its binding sites within IgM in order to efficiently bind and activate the classical cascade. In a sense, the entire activation mechanism of complement lies solely on the original display of antigen, and it's interesting to consider the clinical implications of the relationship between specific pathogen types, the density of relevant surface antigen(s) displayed, and the degree of responding complement activity.

While prebound IgM is probably not a biologically relevant species within the *in vivo* context of cascade activation, we found the effects of reduced IgM surface binding valency to clearly result in the reduction of C1q binding affinity and C1 activation potential. We hypothesized that a 9:1 prebound IgM molecule would serve as a close approximation to solution bound IgM, which was not expected to display significant activity relative to the nominal surface bound IgM condition. The complement activity of solution bound IgM has been a major point of contention within the literature, as the hypothesized staple conformation would be effectively nullified if there were no functional differences between surface and solution bound IgM. While the concept of C1 activation can be viewed as a simple binary system, where C1 exists in either an 'active' or 'inactive' state, it may be hindering to view it that way when trying to decipher the nuance of the mechanism itself. Given the aforementioned complexities of measuring complement activity, combined with the historically underappreciated flexibility of IgM in

solution, it's possible that much of the confusion surrounding the C1 activation mechanism stems from non-specific, intermittent/transient staple formation by solution bound IgM.

The transient nature of C1q binding site exposure in solution bound IgM would indeed lend itself toward contradictory phenotypic observations. While C1 activation itself is a binary system, the mechanism of IgM staple formation is not, and IgM is likely naturally capable of transient staple formation within any given Ag bound arm due its high degree of segmental flexibility. The probability, however, that all or even most of the arms would be able to adopt staple both transiently and simultaneously in solution is very low, and even if it were to occur it would likely not last long enough to trigger a full-scale response of the immune system. For the conformational change to be truly meaningful (in both a temporal and scalar sense), a fixed antigenic surface is likely required to stabilize the simultaneous adoption of multiple local staple formations. Within the context of a biochemical signaling cascade, where certain regulatory thresholds must be surpassed in order to progress along a given pathway, IgM staple formation can perhaps be analogously viewed as an 'intramolecular signaling threshold' for C1 activation. While we did not observe any indication of a hard cutoff for the discrete number of surface bound IgM arms required to bind and activate C1, the significance of the interaction will likely only be increased with an increasing number of available C1q binding sites.

3.5 Conclusion

In this chapter we explored the kinetic and functional binding characteristics that define complement-active IgM using novel BLI and C1 activity assay approaches. To our knowledge, this is the first report of a hybrid assay format that combines the real-time binding information of BLI with a direct functional measurement of C1 activity, all of which is built upon the initial display of surface Ag. To reiterate, the activity assay designed for this project is tuned only towards measuring classical cascade activation, and does not report on the quality of cascade progression or formation of a viable membrane

attack complex (MAC). Provided a sufficient binding environment and presence of required downstream components, it is presumed that specific cleavage of the zymogen form of C1s – via C1q binding to multivalent surface bound IgM, followed by subsequent C1r activation – is an endpoint that adequately reflects a biologically relevant event. Together, the results from the activity assays validate the biological relevance of our BLI- and SA bead-based surface Ag displays, and further demonstrate the mechanistic role of IgM binding valency/avidity within the activation of the classical complement cascade.

We also tested the major environmental variables that affect the binding interaction of IgM with C1/C1q, and took a deeper look at the role of IgM surface binding avidity in the activation mechanism of C1. Based on our observations, reductions in effective IgM binding valency led directly to a reduction in C1 binding efficiency, but only in a surprisingly gradual nature, with no hard cutoffs found for the number of surface bound IgM arms required to register C1q binding and/or C1 activation. In order for a biological definition of complement-active IgM to be complete, however, it's crucial that the functional characterizations from this chapter be paired with detailed structural insight of the conformational features that define the staple conformation itself, which we examine in detail throughout the following chapter.

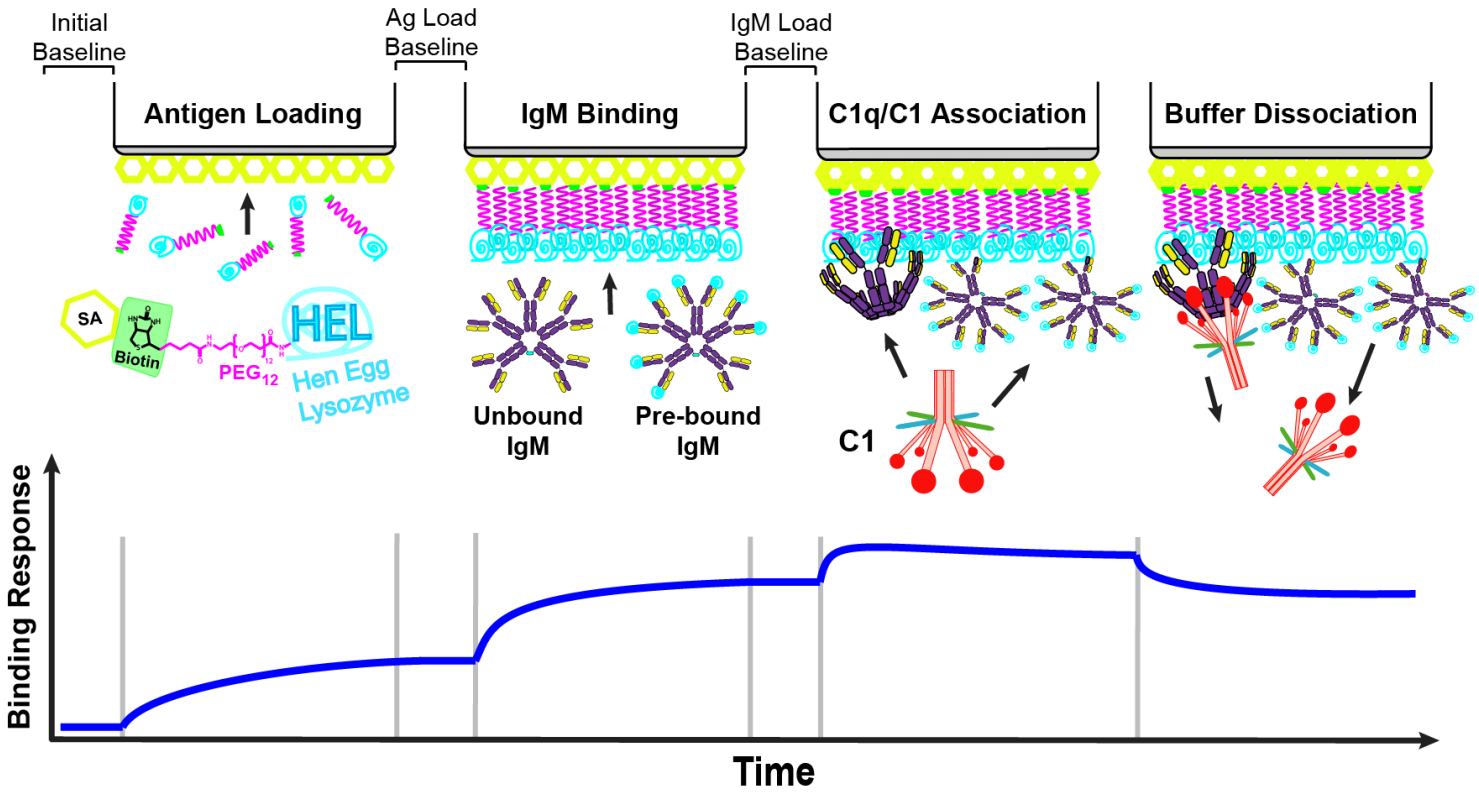


Figure 3.1.1 | Biolayer interferometry (BLI) method overview

Depictions of each major stage of the BLI experimental format are illustrated along with an example data trace of the binding response signal. After taking an initial baseline in blank running buffer, streptavidin (SA) coated BLI tips are dipped into wells containing biotinylated antigen until reaching tip saturation, followed by another short buffer baseline after loading is complete. Tips are then moved into wells containing IgM conditions, followed by a final baseline before association is measured in wells containing C1q/C1, with final dissociation monitored in wells containing fresh/blank run buffer. Up to 8 separate channels can be monitored simultaneously.

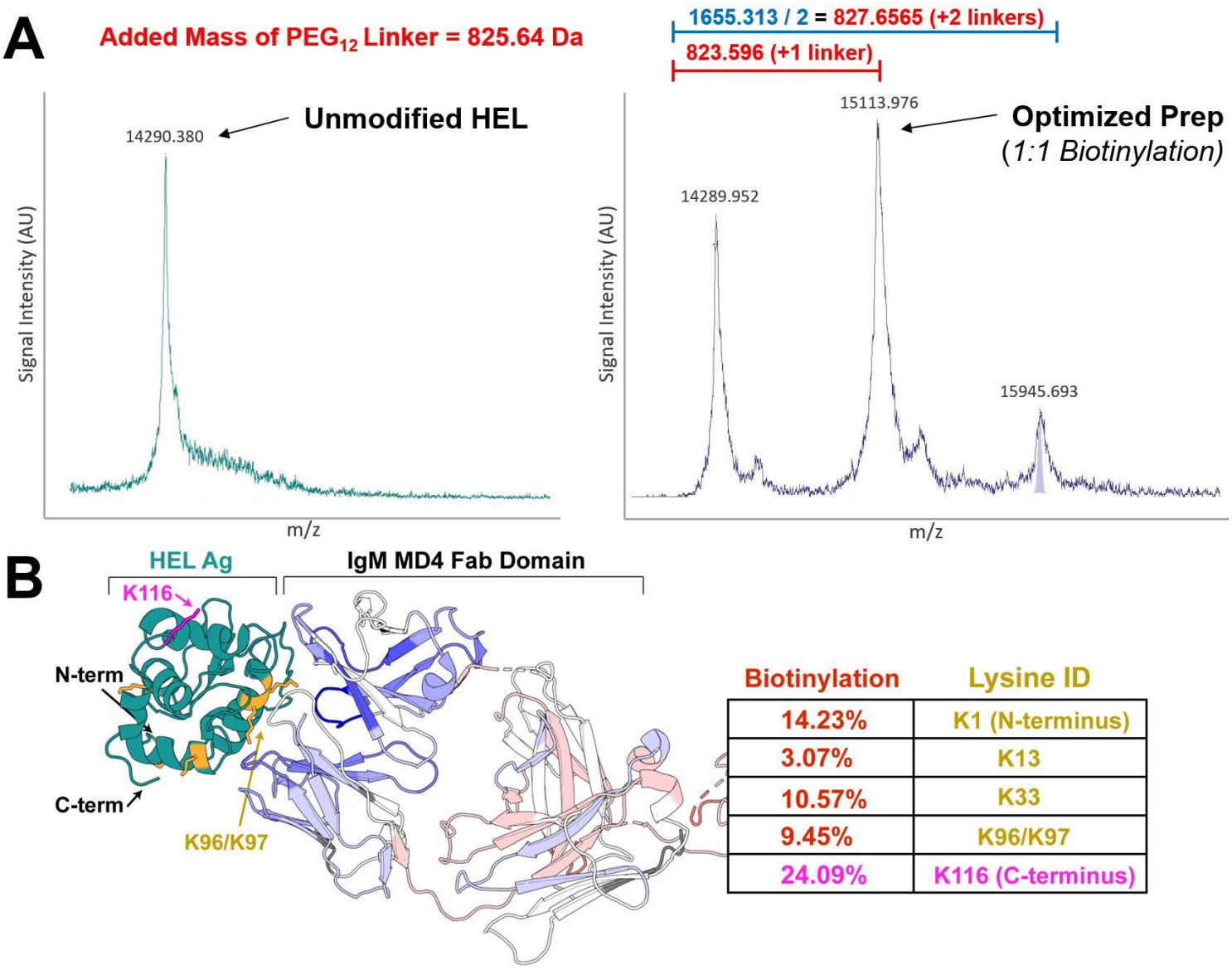


Figure 3.3.1.1 | Optimized biotinylation of hen egg lysozyme (HEL) Ag with nominal PEG_{12} linker

The biotinylation protocol of soluble monomeric HEL protein antigen was assessed by **A)** MALDI mass spectrometry for confirmation of resulting stoichiometry, with the goal of *minimal* (1 biotin linker per HEL molecule) biotinylation. High resolution peptide mapping using a Fusion Orbitrap was also used to determine **B)** the extent and locations of biotin linker conjugation, which was found to be predominantly located at K116 (colored magenta) which is distal from the binding epitope and therefore not expected to interfere with IgM binding. Less than 10% of the total linker conjugation is predicted to occur within the binding interface of IgM MD4 (located at K96/K97), and is therefore not expected to have a significant impact on experiments involving surface Ag binding.

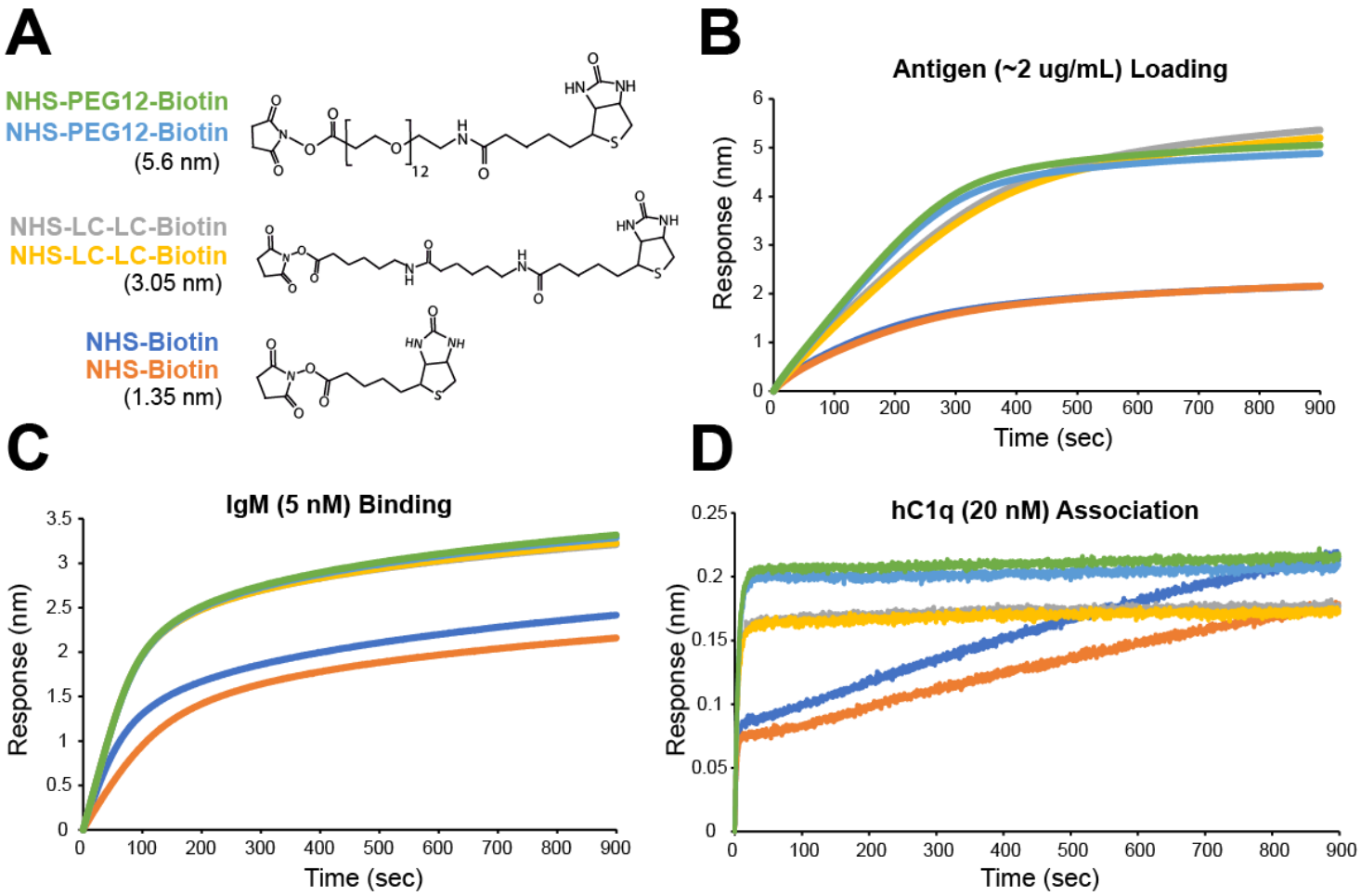


Figure 3.3.1.2 | Optimization of surface Ag biotin linker

A) The short commercial standard biotin-NHS linker reagent (dark blue/orange) was compared against extended LC-LC-biotin-NHS (gold/grey) and PEG_{12}-biotin-NHS (light blue/green) linkers for optimization of surface Ag presentation. **B)** After loading biotinylated Ag to saturation onto SA coated biosensors, the differences in **C)** IgM MD4 binding magnitudes and **D)** subsequent human C1q binding magnitudes were evaluated. While the IgM binding responses were similar between the two extended linkers, the PEG_{12} biotin linker displayed the greatest subsequent hC1q binding response and was therefore chosen as the nominal surface Ag linker for all experiments involving IgM surface binding. The increase to C1q binding magnitudes is thought to be a direct consequence of improved IgM surface binding multivalency, which is fundamental to the hypothesized formation of IgM staple.

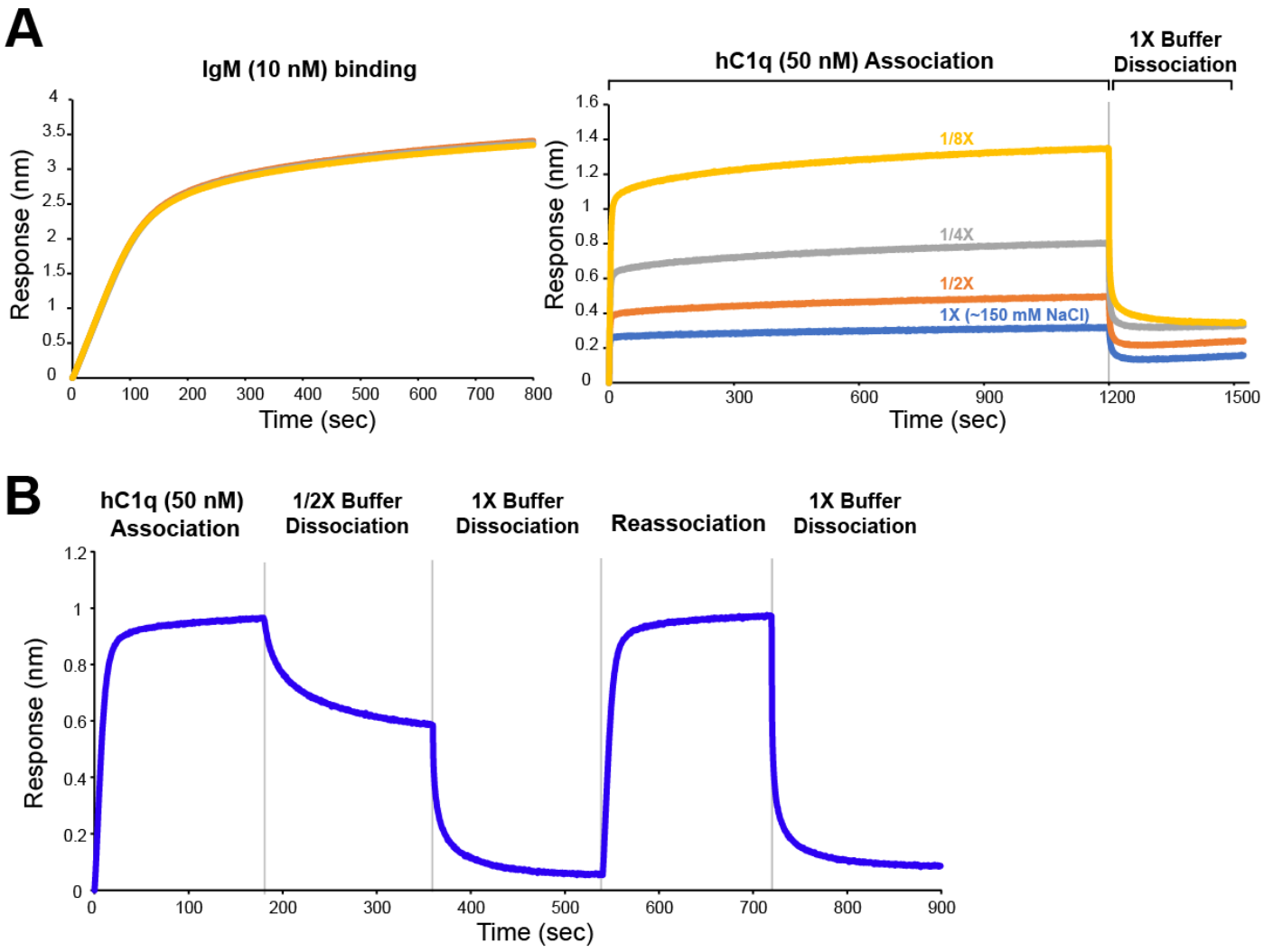


Figure 3.3.2.1 | Effects of ionic strength on the C1q binding interaction with IgM

A) The predominant effects of ionic strength of the solution environment is demonstrated via BLI within a series of run buffers of decreasing salt concentration, starting at 1X physiological (150 mM NaCl, pH 7.3; blue trace) conditions. The hC1q binding magnitudes were observed to increase in response to each reduction in ionic strength of the buffer environment, owing to the fact that C1q binding is predominantly based on charge-based pairings between charged (acidic /basic) surface residues. **B)** In further demonstration of this well-established principle, an experiment was performed in which hC1q was repeatedly dissociated into fresh running buffer prepared with 1X and 0.5X ionic strength, which resulted in a significant reduction to the observed dissociation rate of C1q bound to IgM within the 0.5X buffer.

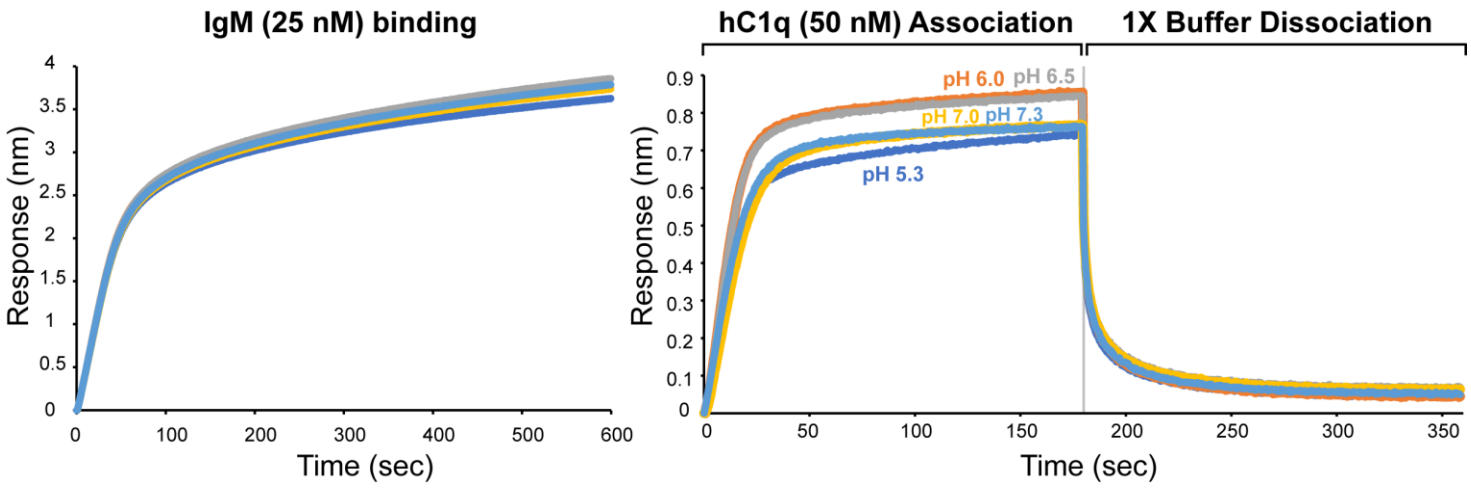


Figure 3.3.2.2 | Effects of pH on the C1q binding interaction with IgM

The effects of solution pH were tested with human C1q (hC1q) prepared within a range of buffer pH values, held at a constant 0.5X ionic strength (~75 mM NaCl). The hC1q on-rates and total binding magnitudes are maximized within the pH 6.0 – 6.5 range, which matches the reported binding trends of the A and C globular head modules. The overall impact of changes to pH appear to be only minor in effect, especially compared to that of ionic strength effects (refer to Figure 3.3.2.1).

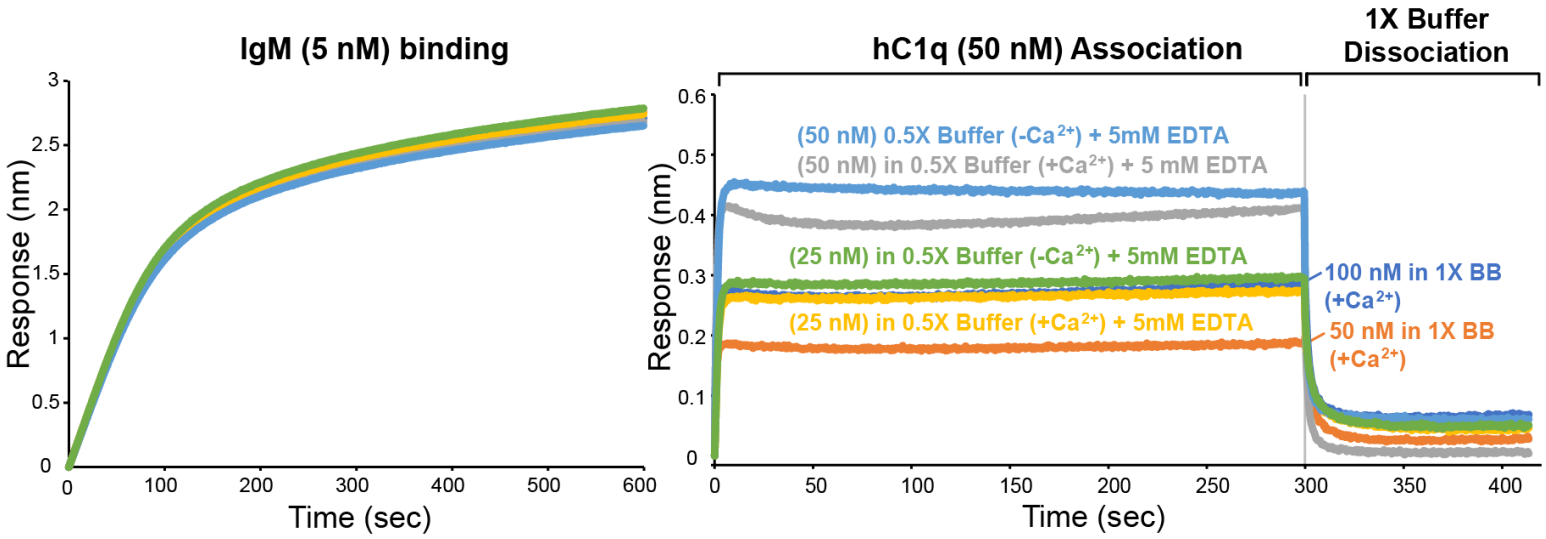


Figure 3.3.2.3 | Effects of EDTA treatment on the C1q binding interaction

Human C1q (hC1q) was prepared at various concentrations and compared alongside nominal (1X binding buffer) concentration conditions. Despite the binding magnitudes of EDTA-containing conditions increasing due to the drop in ionic strength of the buffer from 1X (~150 mM NaCl) to 0.5X (~75 mM NaCl), no differences appear in the profiles or binding kinetics (on/off rates) across any of the tested conditions, indicating that the presence of EDTA had no observable impact on hC1q binding. While calcium does not appear necessary for C1q target recognition or binding to IgM, the presence of calcium is likely required for complete functionality of the whole C1 complex.

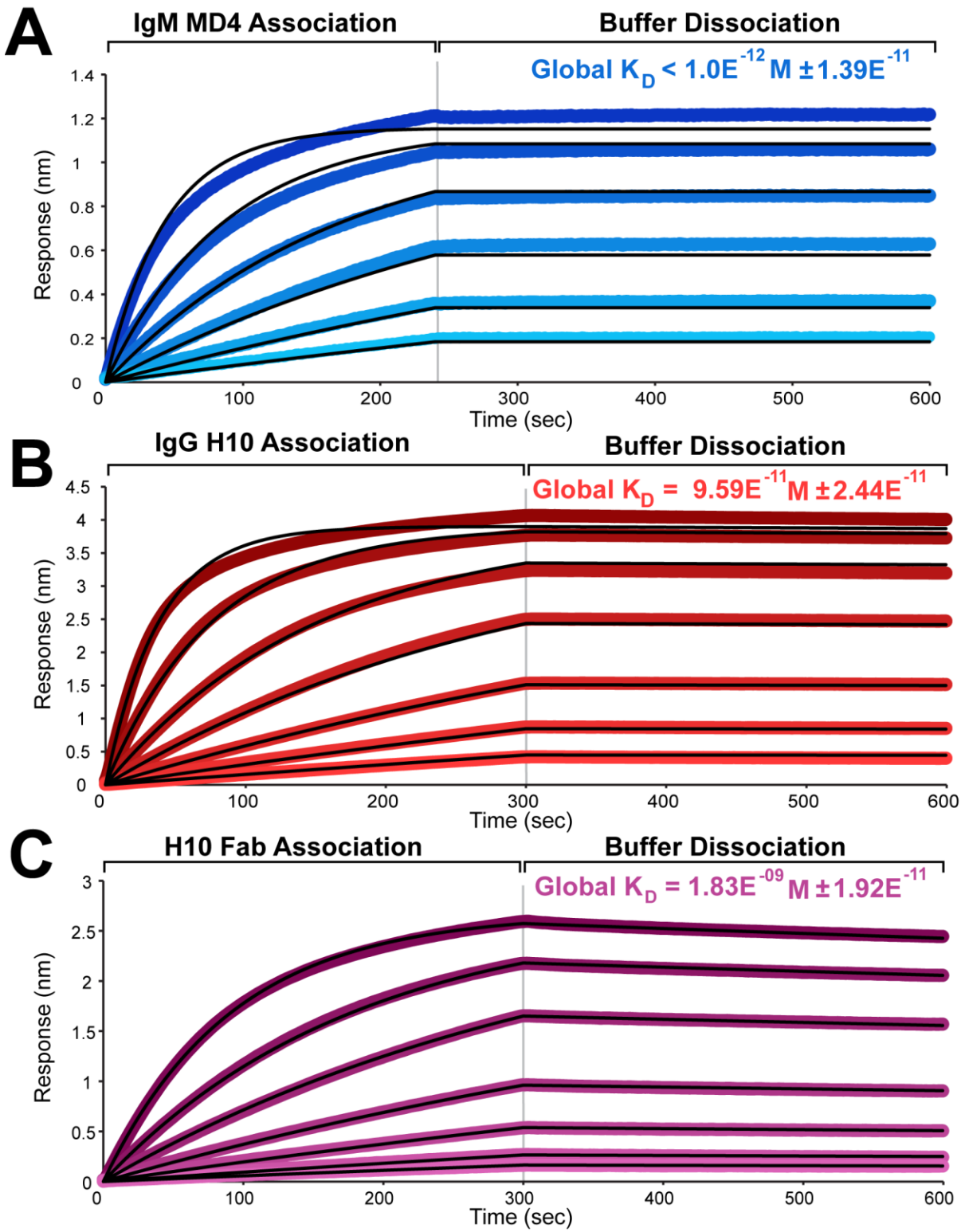


Figure 3.3.3.1 | BLI binding kinetics of IgM MD4 and IgG H10

Serial dilutions were performed to provide estimations of: **A)** IgM MD4 binding kinetics. The off-rate was too slow to measure practically, and the global K_D was approximated to be in the femtomolar range or lower (**<** $1.0E^{-12}$ M). Additional testing was done to approximate **B)** IgG H10 whole antibody binding kinetics, and **C)** IgG H10 Fab only binding kinetics. All kinetic estimates were calculated by the data analysis software with 1:1 fitting (black traces) applied.

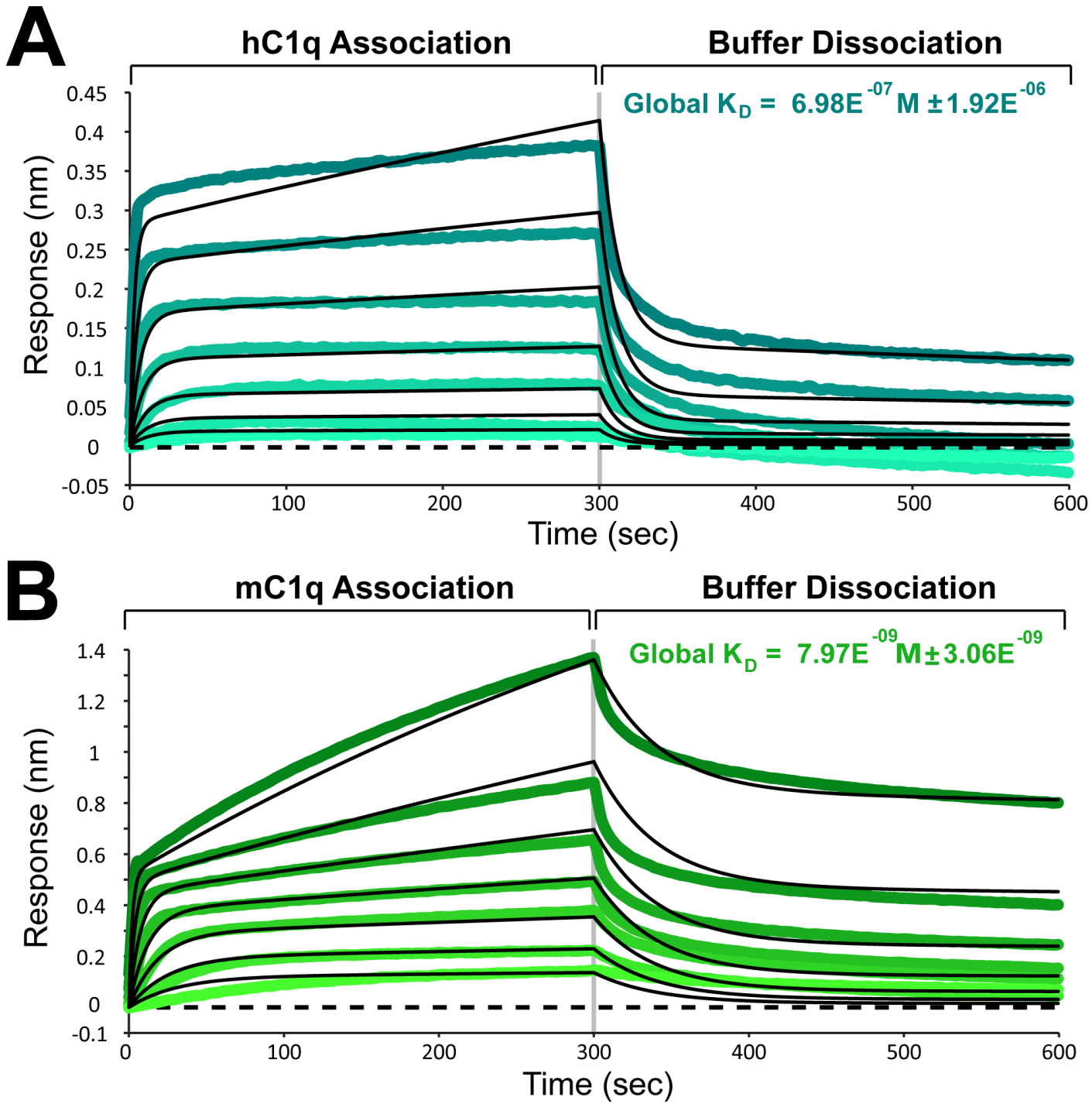


Figure 3.3.3.2 | Species effects on the C1q binding interaction with IgM

The BLI binding profiles of both **A)** human C1q (hC1q) and **B)** mouse C1q (mC1q) to nominal surface bound IgM displayed 'fast on/fast off' binding interaction phenotypes. Kinetic values were approximated from serial dilution experiments by applying a 2:1 (heterogeneous ligand) fitting model (black traces), revealing an average global K_D of approximately 698 nM and 7.97 nM for hC1q and mC1q binding to mouse IgM MD4, respectively, demonstrating the effects of species' mismatches between antibody and complement sources.

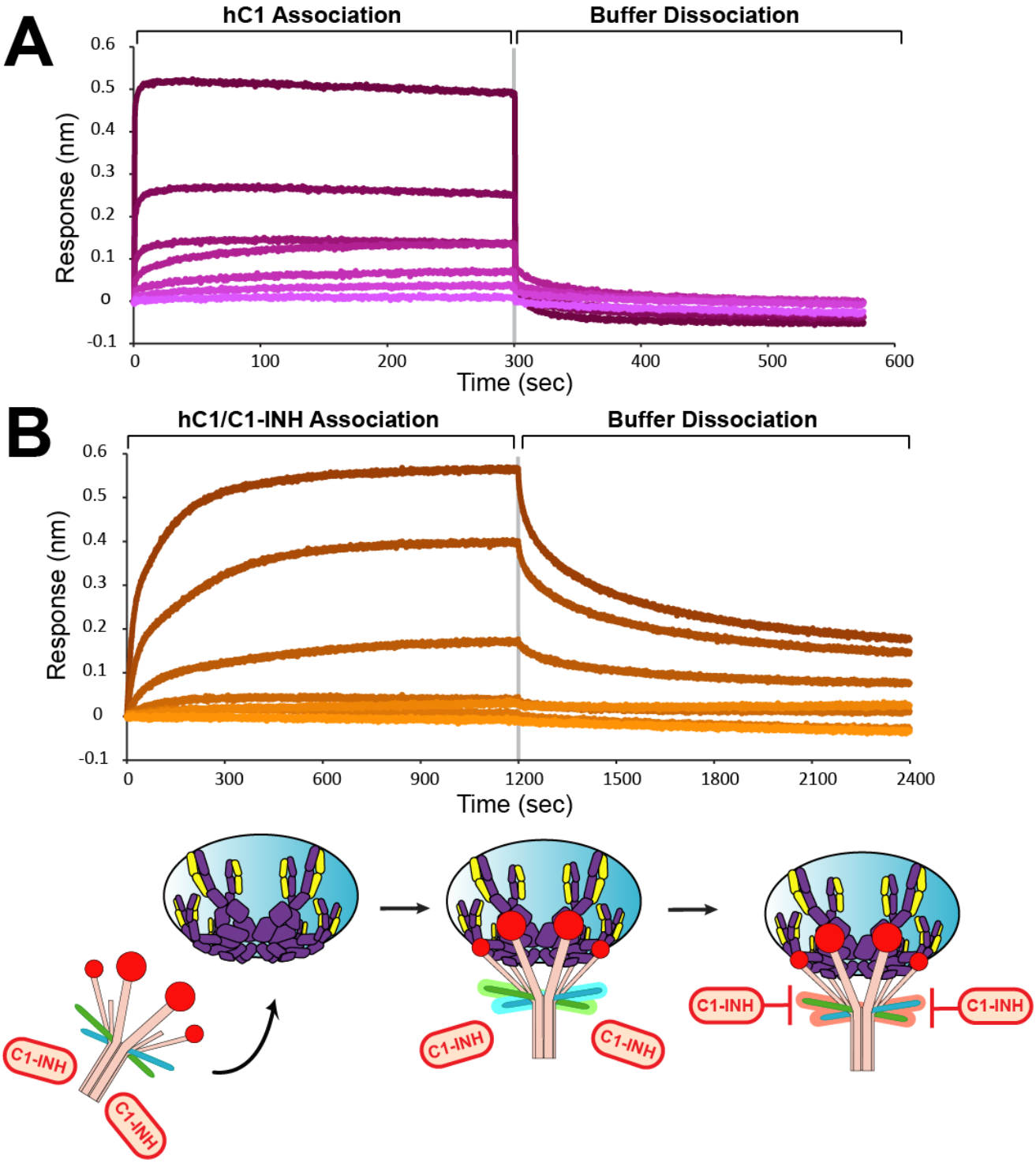


Figure 3.3.3.3 | Effects of C1-INH presence on the C1 binding interaction with IgM

The binding phenotype of **A)** human C1 (hC1) is significantly altered by **B)** the presence of co-dialyzed C1 inhibitor protein (C1-INH), even at a slight molar excess of ~2:1 (C1-INH:C1). The secondary binding event observed after initial burst-phase binding is thought to be indicative of C1-INH binding to activated C1r/C1s molecules that have yet to be released from the tip surface, as illustrated in the cartoon scheme below.

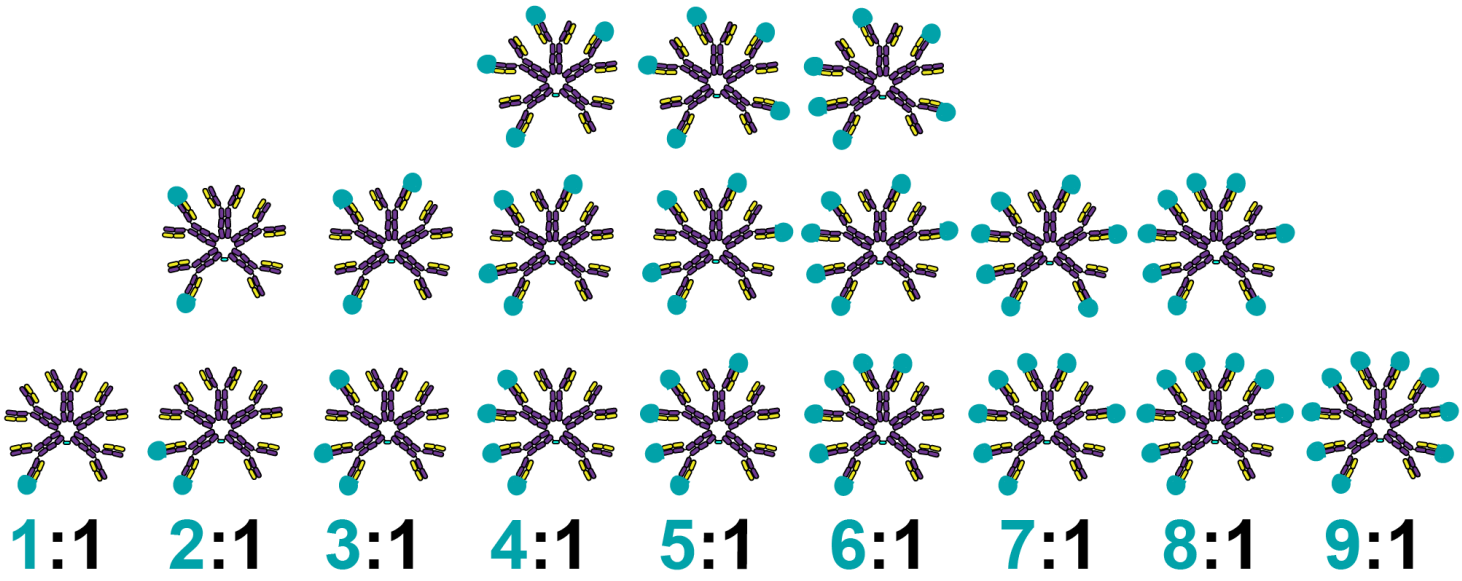


Figure 3.3.4.1 | Simplified distribution of prebound IgM configuration subpopulations

Even when assuming all IgM arms are equal in terms of binding capacity, there are still multiple bound configurations possible within each theoretically ideal prebound IgM condition. In all likelihood, the arms that are attached to the J chain may indeed exhibit limited flexibility and binding capacity, further complicating the kinetics and distributions of each sub-population.

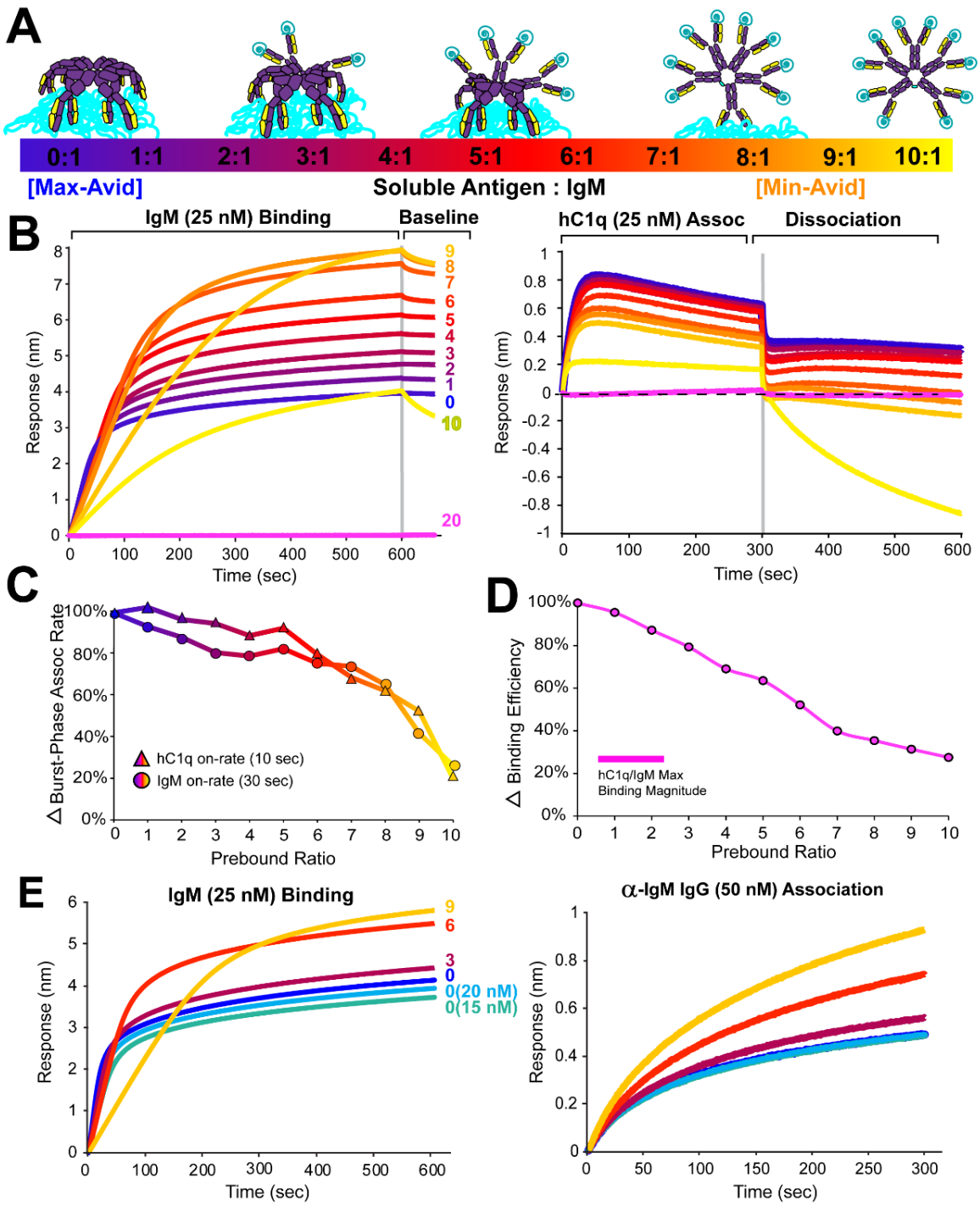


Figure 3.3.4.2 | Effects of reduced IgM surface binding valency

A) Prebinding legend displaying the average reduction in IgM binding valency as prebound ratio increases. **B)** BLI analysis of a complete prebound series from the nominal (0:1,"max-avid") form of IgM staple to the saturating ratio of 10:1, which resulted in a continual decrease in hC1q binding responses (right panel). The kinetic effects of reduced IgM binding valency appear to gradually reduce **C)** the burst-phase binding rates of both IgM and hC1q, as well as continually lower **D)** the total efficiency of the IgM to recruit C1, determined by the intensity ratio of maximum C1q/IgM binding signal magnitudes. **E)** Relative IgM quantitation within each prebound condition was performed by dipping into wells containing α-IgM IgG antibody material, demonstrating the hypothesized increase in the total number of IgM molecules bound to the BLI tip within each increasing prebound ratio, which largely contributes to the increasing IgM binding signal observed in part B.

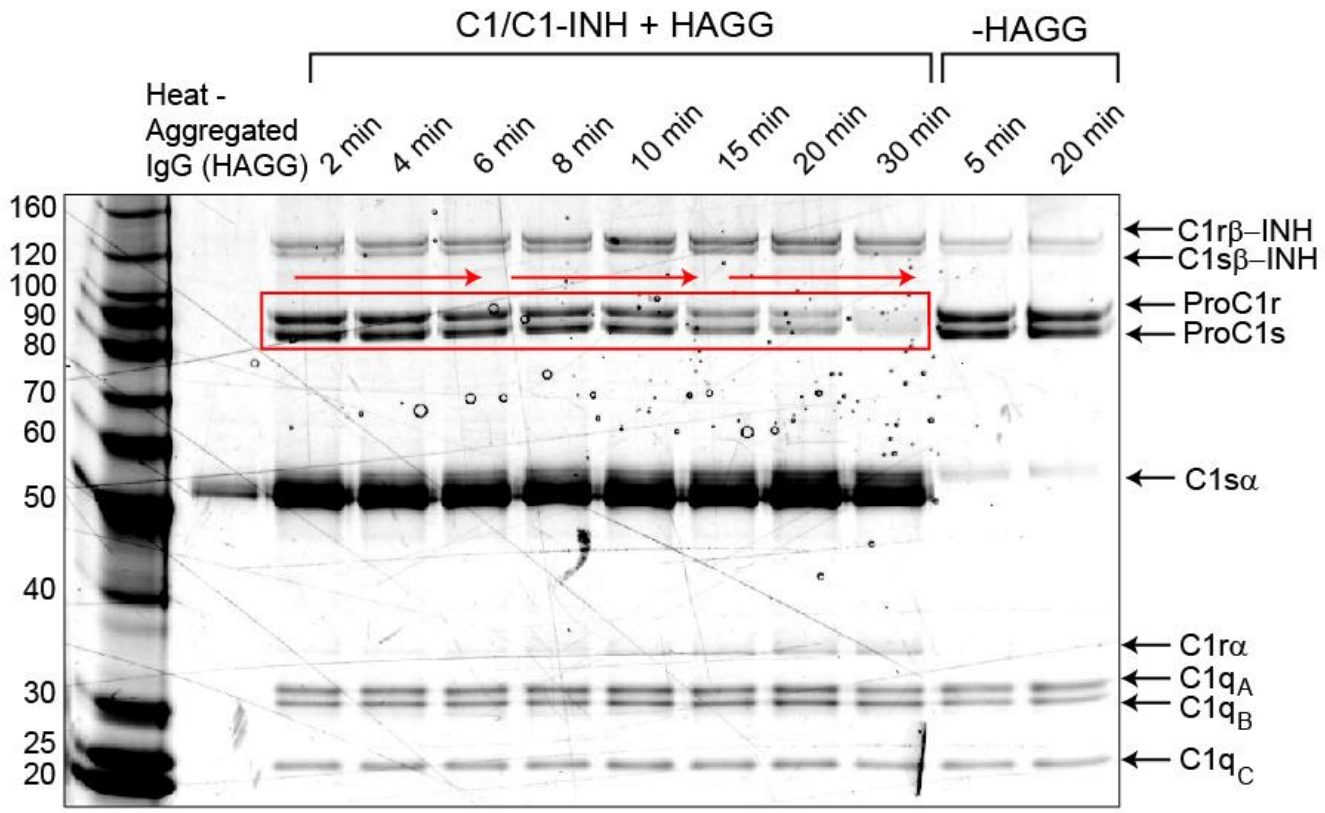


Figure 3.3.5.1 | C1 activity assay method development

Initial method development of a gel-based C1 activity assay that utilized heat-aggregated IgG (HAGG) as a positive control for C1 activation is presented.[25] In this reduced SDS-PAGE experiment, protein bands that correspond to the inactive/pro-C1s and C1r proteases begin to fade over time as they both become cleaved/activated. At the same time, other bands that correspond to bound C1-INH fragment products of C1r and C1s are observed to appear over time (top), along with the proteolytically active C1r and C1s fragments, the latter of which proceeds to onward to subsequent stages of the cascade pathway. At the bottom are three protein bands that correspond to the primary A,B, and C polypeptide chains of C1q which are not cleaved upon activation and remain constant over time. The large shadow that appears over the top of C1sα (~55 kDa) is contributed from an aggregated excess of the heavy chain of the IgG1 material used to form the HAGG positive control, and does not appear within the western blot analysis of C1s due to the specificity of the primary and secondary antibodies used.

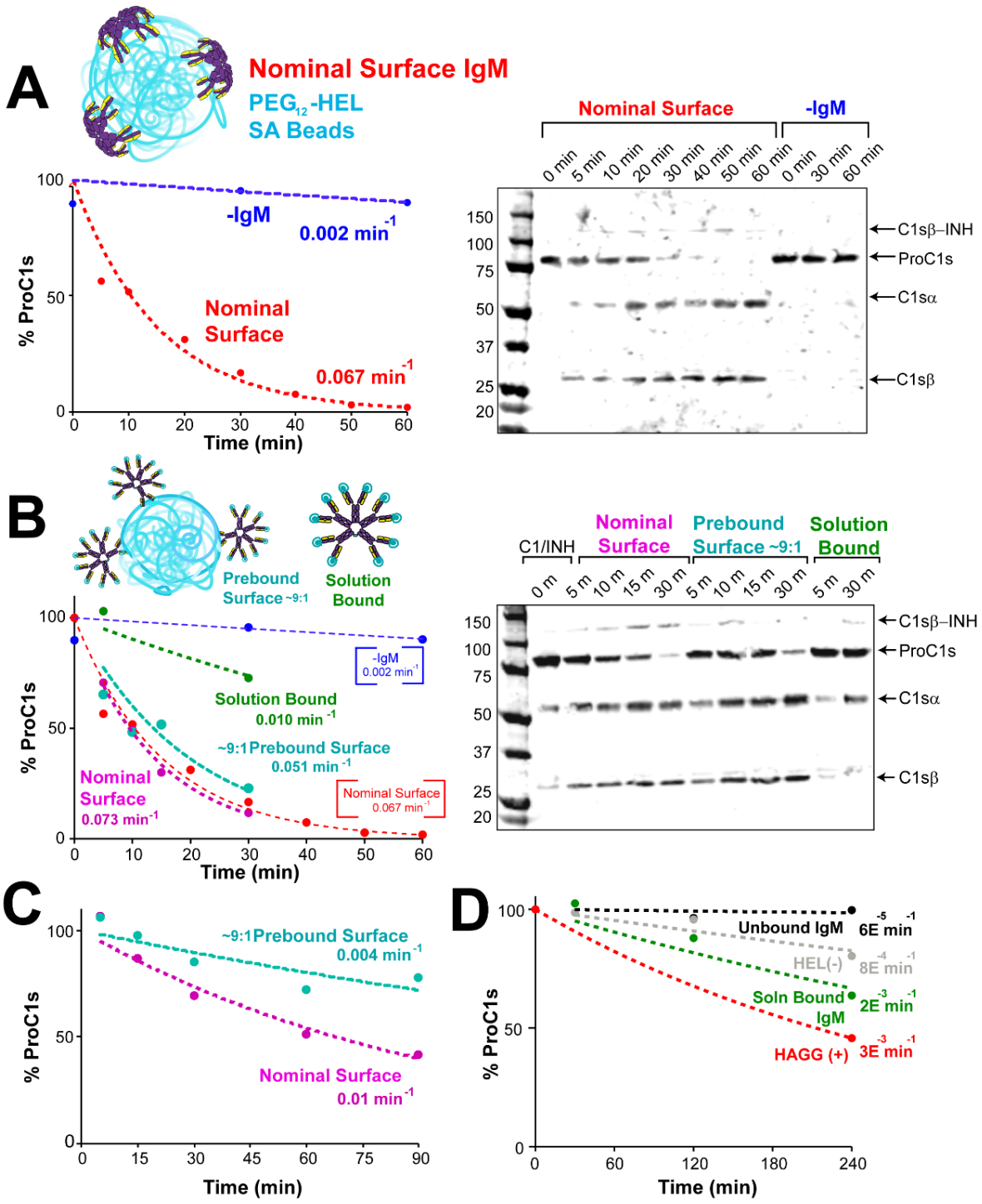


Figure 3.3.5.2 | C1 activity assessment of SA bead antigen display

A) Validation of nominal surface bound activity using high capacity SA beads for solution based activity assays. **B)** Additional bound IgM comparison experiment with nominal curves from part A overlayed (in brackets). **C)** Direct comparison of nominal and prebound (9:1) surface IgM bound to SA bead antigen displays. Prebinding incubations were set up the day prior to testing and allowed to incubate at room temperature for 3 hours before being placed overnight at 5° C to give excess time for binding equilibration, leading to a more pronounced (compared to part B) reduction in prebound IgM C1 activation rate compared to the nominal, max-avid form of IgM. **D)** Solution state IgM comparison of C1 activity was performed separately, confirming unbound (apo) IgM to be C1-inert while also confirming the low level activity of the solution bound form. Heat aggregated IgG (HAGG) material was used as a positive control for C1 activation.

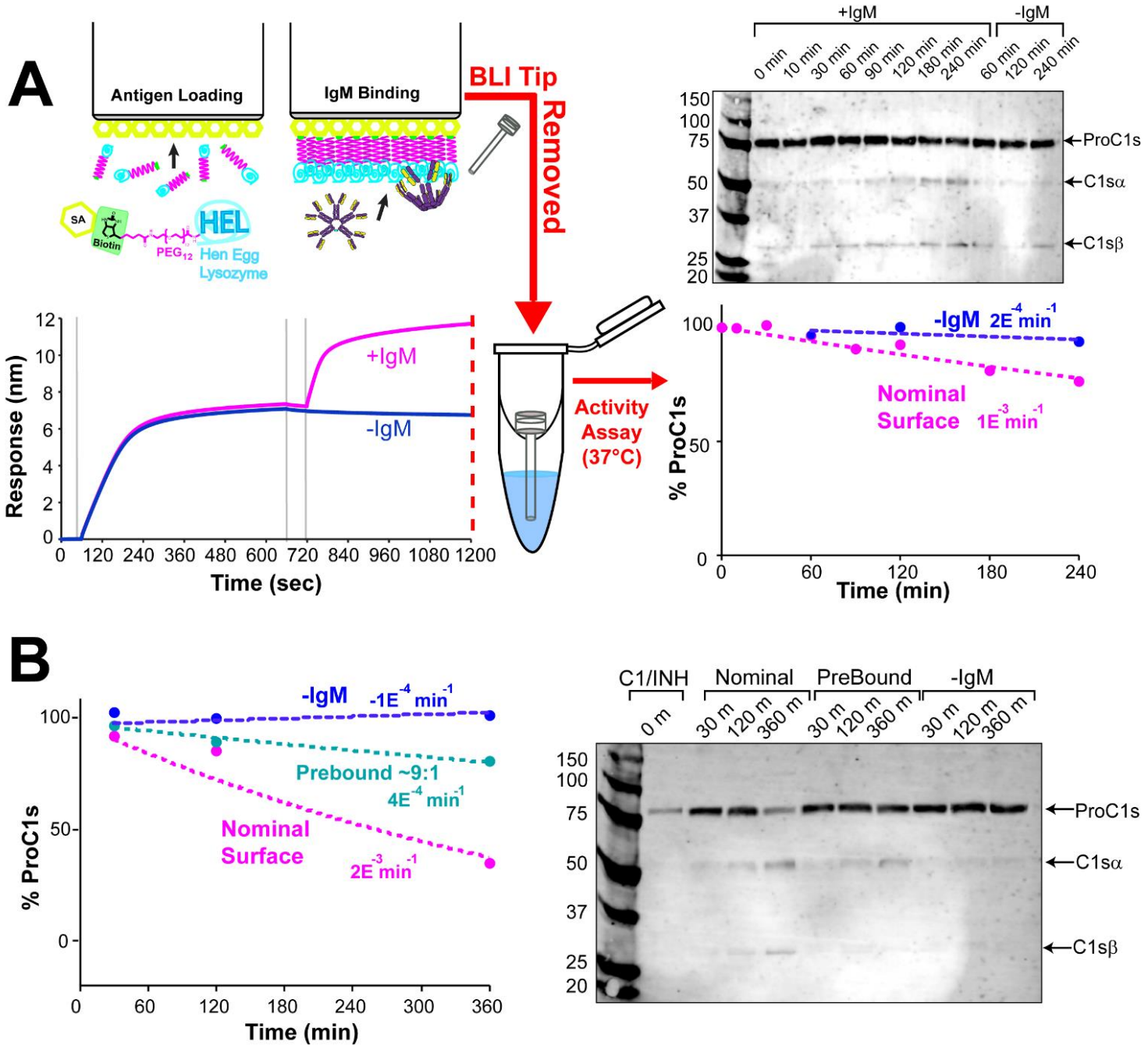


Figure 3.3.5.3 | Hybrid BLI & C1 activity assay approach

A) BLI tips were loaded normally with saturating levels of PEG_{12}-HEL antigen and baselined in blank run buffer for 60 sec prior to dipping into wells containing either buffer (blue) or ~10 nM of nominal IgM MD4 (magenta). After loading, the instrument was halted, and tips were transferred directly into tubes containing co-dialyzed C1/C1-INH and monitored for activity at 37° C over the course of 4 continuous hours, with aliquots of equal volume removed from each condition at the indicated time points. Activity was measured normally by monitoring the conversion of proC1s into its two activated fragments and plotted as a function of the percentage of remaining ProC1s abundance over time. **B)** BLI-based activity assay comparison of nominal (0:1; magenta) to highly prebound (9:1; green) IgM, along with a negative control (-IgM; blue), displaying a clear reduction in the C1 activation potential of prebound IgM bound to the BLI tip surface.

3.6 References

1. Bhoopalam, N., Heller, P., Meyerstein, N. & Hall, L. Effect of dextran-S (alpha , 1-3 dextran) on the growth of plasmacytomas MOPC-104E and Information about subscribing to The Journal of Immunology is online at : EFFECT OF DEXTRAN-S (a , l-3 DEXTRAN) ON THE GROWTH OF PLASMACYTOMAS MOPC-104E AND J558l. 1–3 (2018).

2. Zidovetzki, R., Blatt, Y., Schepers, G. & Pecht, I. Thermodynamics of oligosaccharides binding to a dextran-specific monoclonal IgM. *Mol. Immunol.* **25**, 379–383 (1988).

3. Valim, Y. M. L. & Lachmann, P. J. The effect of antibody isotype and antigenic epitope density on the complement-fixing activity of immune complexes: a systematic study using chimaeric anti-NIP antibodies with human Fc regions. *Clin. Exp. Immunol.* **84**, 1–8 (1991).

4. Concepcion, J. *et al.* Label-Free Detection of Biomolecular Interactions Using BioLayer Interferometry for Kinetic Characterization. *Comb. Chem. High Throughput Screen.* **12**, 791–800 (2009).

5. Padlan, E. A. *et al.* Structure of an antibody-antigen complex: Crystal structure of the HyHEL-10 Fab-lysozyme complex. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 5938–5942 (1989).

6. Kam-morgan, L. N. W. *et al.* High-resolution mapping of the HyHEL-10 epitope of chicken lysozyme by site-directed mutagenesis. *Proc Natl Acad Sci U S A* **90**, 3958–3962 (1993).

7. Xavier, K. A. & Willson, R. C. Association and dissociation kinetics of anti-hen egg lysozyme monoclonal antibodies HyHEL-5 and HyHEL-10. *Biophys. J.* **74**, 2036–2045 (1998).

8. Pons, J., Stratton, J. R. & Kirsch, J. F. How do two unrelated antibodies, HyHEL-10 and F9.13.7, recognize the same epitope of hen egg-white lysozyme? *Protein Sci.* **11**, 2308–2315 (2002).

9. Acchione, M., Kwon, H., Jochheim, C. M. & Atkins, W. M. Impact of linker and conjugation chemistry on antigen binding, Fc receptor binding and thermal stability of model antibody-drug conjugates. *MAbs* **4**, 362–372 (2012).

10. Jensen, P. F., Jørgensen, T. J. D., Koefoed, K., Nygaard, F. & Sen, J. W. Affinity capture of biotinylated proteins at acidic conditions to facilitate hydrogen/deuterium exchange mass spectrometry analysis of multimeric protein complexes. *Anal. Chem.* **85**, 7052–7059 (2013).

11. Gitlin, G., Khait, I., Bayer, E. A., Wilchek, M. & Muszkat, K. A. Studies on the biotin-binding sites of avidin and streptavidin. A chemically induced dynamic nuclear polarization investigation of the status of tyrosine residues. *Biochem. J.* **259**, 493–498 (1989).

12. Gaboriaud, C. *et al.* The Crystal Structure of the Globular Head of Complement Protein C1q Provides a Basis for Its Versatile Recognition Properties. *J. Biol. Chem.* **278**, 46974–46982 (2003).

13. Zlatarova, A. S. *et al.* Existence of different but overlapping IgG- and IgM-binding sites on the globular domain of human C1q. *Biochemistry* **45**, 9979–9988 (2006).

14. Roumenina, L. T. *et al.* Role of Ca2+ in the electrostatic stability and the functional activity of the globular domain of human C1q. *Biochemistry* **44**, 14097–14109 (2005).

15. Stoiber, H., Ebenbichler, C. F., Thielens, N. M., Arlaudt, G. J. & Dierich, M. P. HIV-l rsgp41 DEPENDS ON CALCIUM FOR BINDING OF HUMAN Clq BUT NOT FOR BINDING OF gp120. *Mol.*

Immunol. **32**, 371–374 (1995).

16. Sjoberg, A., Onnerfjord, P., Morgelin, M., Heinegard, D. & Blom, A. M. The Extracellular Matrix and Inflammation FIBROMODULIN ACTIVATES THECLASSICAL PATHWAYOFCOMPLEMENTBY DIRECTLYBINDING C1q. *J. Biol. Chem.* **280**, 32301–32308 (2005).

17. Baeyens-volant, D., Matagne, A., El, R., Wattiez, R. & Azarkan, M. A novel form of ficin from Ficus carica latex : Purification and characterization. *Phytochemistry* **117**, 154–167 (2015).

18. Mariani, M., Camagna, M., Tarditi, L. & Seccamani, E. A new enzymatic method to obtain High-Yield F(ab)2 suitable for clinical use from mouse IgG1. *Mol. Immunol.* **28**, 69–77 (1991).

19. Sharp, T. H. *et al.* Insights into IgM-mediated complement activation based on in situ structures of IgM-C1-C4b. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 11900–11905 (2019).

20. Li, Y. *et al.* Structural insights into immunoglobulin M. *Science (80-.).* **367**, 1014–1017 (2020).

21. Kumar, N., Arthur, C. P., Ciferri, C. & Matsumoto, M. L. Structure of the human secretory immunoglobulin M core. *Structure* **29**, 564-571.e3 (2021).

22. Müller, R. *et al.* High-resolution structures of the IgM Fc domains reveal principles of its hexamer formation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 10183–8 (2013).

23. Chouquet, A. *et al.* Biophysical Characterization of the Oligomeric States of Recombinant Immunoglobulins Type-M and Their C1q-Binding Kinetics by Biolayer Interferometry. *Front. Bioeng. Biotechnol.* **10**, 1–15 (2022).

24. Collins, C., Tsui, F. W. L. & Shulman, M. J. Differential activation of human and guinea pig complement by pentameric and hexameric IgM. *Eur. J. Immunol.* **32**, 1802 (2002).

25. Mutti, M., Ramoni, K., Nagy, G., Nagy, E. & Szijártó, V. A New Tool for Complement Research : In vitro Reconstituted Human Classical Complement Pathway Bacterial Strains and Media. **9**, 1–9 (2018).

26. Zimmer, J., Hobkirk, J., Mohamed, F., Browning, M. J. & Stover, C. M. On the functional overlap between complement and anti-microbial peptides. **5**, 1–10 (2015).

27. Folkerd, E. J., Gardner, B. & Hughes, J. N. The relationship between the binding ability and the rate of activation of the complement component C1. *Immunology* **41**, 179–185 (1980).

28. Davis III, A. E., Mejia, P. & Lu, F. BIOLOGICAL ACTIVITIES OF C1 INHIBITOR. *Mol. Immunol.* **45**, 4057–4063 (2008).

29. Oskam, N. *et al.* At Critically Low Antigen Densities, IgM Hexamers Outcompete Both IgM Pentamers and IgG1 for Human Complement Deposition and Complement-Dependent Cytotoxicity. *J. Immunol.* **209**, 16–25 (2022).

30. Borsos, T., Chapuis, R. M. & Langone, J. J. Distinction between fixation of C1 and the activation of complement by natural IgM anti-hapten antibody: Effect of cell surface hapten density. *Mol. Immunol.* **18**, 863–868 (1981).

31. Hurst, M. M., Volanakis, J. E., Stroud, R. M. & Bennett, J. C. A Comparative Analysis of the C1-Binding Ability of Fragments Derived from Complement-Fixing and Noncomplement-Fixing IgM Proteins. *J. Clin. Invest.* **58**, 16–21 (1976).

32. Gadjeva, M. G. *et al.* Interaction of human C1q with IgG and IgM: Revisited. *Biochemistry* **47**, 13093–13102 (2008).

33. Gaboriaud, C., Ling, W. L., Thielens, N. M., Bally, I. & Rossi, V. Deciphering the fine details of C1 assembly and activation mechanisms: 'Mission impossible'? *Front. Immunol.* **5**, 3–9 (2014).

34. Thielens, N. M., Tedesco, F., Bohlson, S. S., Gaboriaud, C. & Tenner, A. J. C1q: A fresh look upon an old molecule. *Mol. Immunol.* **89**, 73–83 (2017).

35. Chen, L., Deng, H., Cui, H., Fang, J. & Zuo, Z. Inflammatory responses and inflammation-associated diseases in organs. *Oncotarget* **9**, 7204–7218 (2018).

36. Rajamäki, K. *et al.* Extracellular Acidosis Is a Novel Danger Signal Alerting. *J. Biol. Chem.* **288**, 13410–13419 (2013).

37. Gaboriaud, C., Frachet, P., Thielens, N. M. & Arlaud, G. J. The human C1q globular domain: Structure and recognition of non-immune self ligands. *Front. Immunol.* **2**, 1–8 (2012).

38. Cohen-khait, R. & Schreiber, G. Selecting for Fast Protein − Protein Association As Demonstrated on a Random TEM1 Yeast Library Binding BLIP. *Biochemistry* **57**, 4644–4650 (2018).

39. Wang, C. J., Bergmann, A., Lin, B., Kim, K. & Levchenko, A. Diverse Sensitivity Thresholds in Dynamic Signaling Responses by Social Amoebae. *Sci. Signal.* **5**, 1–26 (2018).

40. Gallagher, M. P., Conley, J. M., Vangala, P., Garber, M. & Reboldi, A. Hierarchy of signaling thresholds downstream of the T cell receptor and the Tec kinase ITK. *Proc. Natl. Acad. Sci.* **118**, 1–12 (2021).

41. Ziccardi, R. J. The First Component of Human Complement (C1): Activation and Control. *Springer Semin. Immunopathol.* **6**, 213–230 (1983).

42. Preiner, J. *et al.* IgGs are made for walking on bacterial and viral surfaces. *Nat. Commun.* (2014) doi:10.1038/ncomms5394.

43. Shulman, M. J., Heusser, C., Filkin, C. & Kohler, G. Mutations Affecting the Structure and Function of Immunoglobulin M. *Mol. Cell. Biol.* **2**, 1033–1043 (1982).

44. Wright, J. F., Shulman, M. J., Isenman, D. E. & Painter, R. H. C1 binding by mouse IgM. The effect of abnormal glycosylation at position 402 resulting from a serine to asparagine exchange at residue 406 of the μ-chain. *J. Biol. Chem.* **265**, 10506–10513 (1990).

45. Wang, F., Nakouzi, A., Angeletti, R. H. & Casadevall, A. Site-specific characterization of the N-linked oligosaccharides of a murine immunoglobulin M by high-performance liquid chromatography/electrospray mass spectrometry. *Anal. Biochem.* **314**, 266–280 (2003).

46. Moh, E. S. X., Lin, C. H., Thaysen-Andersen, M. & Packer, N. H. Site-Specific N-Glycosylation of Recombinant Pentameric and Hexameric Human IgM. *J. Am. Soc. Mass Spectrom.* **27**, 1143–1155 (2016).

47. Hennicke, J. *et al.* Glycan profile of CHO derived IgM purified by highly efficient single step affinity chromatography. *Anal. Biochem.* **539**, 162–166 (2017).

48. van Osch, T. L. J. *et al.* Fc Galactosylation Promotes Hexamerization of Human IgG1, Leading to Enhanced Classical Complement Activation. *J. Immunol.* **207**, 1545–1554 (2021).

49. Zhou, W., Lin, S., Chen, R., Liu, J. & Li, Y. Characterization of antibody-C1q interactions by Biolayer Interferometry. *Anal. Biochem.* **549**, 143–148 (2018).

50. Karnaukhova, E. C1-Esterase Inhibitor : Biological Activities and Therapeutic Applications Thromboembolic Diseases C1-Esterase Inhibitor : Biological Activities and Therapeutic Applications. *J. Hematol. Thromboembolic Dis.* **1**, (2014).

51. Nielsen, E. *et al.* Effect of supraphysiologic levels of C1-inhibitor on the classical , lectin and alternative pathways of complement. *Mol. Immunol.* **44**, 1819–1826 (2007).

52. Davis, A. E., Lu, F. & Mejia, P. C1 inhibitor, a multi-functional serine protease inhibitor. *Thromb. Haemost.* **104**, 886–893 (2010).

53. Amano, M. T., Farah, C. S., Jensenius, J. C. & Isaac, L. Genetic analysis of complement C1s deficiency associated with systemic lupus erythematosus highlights alternative splicing of normal C1s gene Genetic analysis of complement C1s deficiency associated with systemic lupus erythematosus highlights alternative. *Mol. Immunol.* **45**, 1693–1702 (2008).

54. Jiang, H., Wagner, E., Zhang, H. & Frank, M. M. Complement 1 Inhibitor Is a Regulator of the Alternative Complement Pathway. *J. Exp. Med.* **194**, 1609–1616 (2001).

55. Chen, Q., Menon, R., Calder, L. J., Rosenthal, P. B. & Tolar, P. Cryomicroscopy reveals the structural basis for a flexible hinge motion in the immunoglobulin M pentamer. *Nat. Commun.* **13**, 1–11 (2022).

56. Hurst, M. M., Volanakis, J. E., Stroud, R. M. & Bennett, J. C. C1 fixation and classical complement pathway activation by a fragment of the CH4 domain of IgM. *Exp. Med.* **142**, (1975).

– Chapter 4 –

Defining the Structural Features of Complement-Active IgM

4.1 Introduction

To activate the classical complement cascade, it is believed that IgM must first bind multivalently to surface-immobilized antigen (Ag) in order to adopt the "staple" conformation that effectively uncovers the putative C1q binding site located within the Cμ3 domain of the Fc core (refer to Chapter 1, **Figures 1.5.1 & 1.5.2**). While this structural transition has been visualized at low resolution since the 1970s, the mechanistic details that underlie its formation have yet to be determined. To date, no intact IgM crystal structures have been resolved; all structural characterization and protein:protein interfaces have been inferred from combinations of fragmented structures, homology modeling, electron microscopy, X-ray scattering, and molecular dynamics simulations.[1–5] From a technical standpoint, IgM antibodies still remain a challenging system to investigate due to their large size, significant flexibility, and dense glycosylation (20-25% by mass).[6–9] To complicate matters even further, the general behavior and potency of effector function differs vastly between the pentameric and hexameric isoforms, adding a further complication to an otherwise confusing collection of past observational reports and conflicting phenotypic characterizations made throughout the literature.[10,11]

While classical techniques such as X-ray crystallography and NMR have proven useful for resolving the structural details of individual IgM domains, the high degree of flexibility and conformational heterogeneity of intact polyvalent IgM has ultimately limited the resolving power of these methods. Only recently has meaningful progress started to be made, with two partially intact high-resolution cryo-EM structures deposited in the last two years of the human IgM 'Fc core', which contains the pentameric Fc ring made by the Cμ3/Cμ4 domains along with an attached J chain and portion of the polymeric immunoglobulin receptor (pIgR) known as secretory component (refer to

Chapter 1, **Figure 1.2.3**).[12,13] To investigate the structural features that define the staple conformation of IgM we employed hydrogen/deuterium exchange coupled with mass spectrometry (HDX-MS), which is a relatively modern approach that is also highly amenable to the investigation of large and complex glycoproteins.[14,15] As outlined previously in Chapter 2, HDX-MS provides local structural information (peptide-level resolution) by monitoring the exchange of hydrogen for solvated deuterium within protein backbone amides over fixed periods of time, and is an excellent tool for differential comparisons of protein conformations, such as epitope mapping via differential comparison of unliganded vs liganded antibodies.[16]

In this chapter we aim to complete our characterization of the features that define the structure/function relationship of complement-active IgM by employing a series of differential HDX-MS analyses on the nominal, or maximally performing surface bound IgM condition that was functionally characterized throughout Chapter 3 using a combined approach of biolayer interferometry (BLI) and novel C1 activity assays. By pairing the structural features found in this chapter via HDX-MS, with the corresponding functional phenotypes of distinct IgM states and conformations, the defining characteristics of the staple conformation begin to emerge.

4.2 Materials and Methods

4.2.1 HDX with IgM MD4 and IgG H10

Mouse IgM MD4 was generously provided by Dr. John Kearney at the University of Alabama-Birmingham. IgG H10, originally produced by Dr. Sandra Smith-Gill from the National Cancer Institute in Frederick, MD, was provided by Dr. William Atkins. Hen egg lysozyme (Cat# 89833) antigen material and Pierce high-capacity streptavidin (SA) agarose (Cat# 20359) beads were purchased directly from ThermoFisher Scientific. All proteins were prepared in a standardized HBS buffer with added calcium (10 mM HEPES, 150 mM NaCl, 5 mM $CaCl_2$, 1.5 mM $MgCl_2$, pH 7.3).

4.2.2 Solution HDX reactions

Solution state HDX reactions were prepared by diluting the protein mixture 10-fold into deuterated buffer (85% D_2O final; pH* = 7.3) for timepoints ranging from 1 to 100 minutes, with each sample replicate containing ~10 ug of IgM at a total labeling volume of 100 uL. At the end of each timepoint exchange was immediately halted by the addition of an equal volume of ice-cold quench buffer (containing 8 M urea and 200 mM TCEP with 0.2% formic acid), bringing down the pH to ~2.5 with a final sample volume of 200 uL. Quenched samples were then quickly placed on ice for 1 minute before finally being transferred to pre-frozen HPLC vials (High Recovery, Waters) contained within a bath of dry ice/ethanol, where samples were rapidly frozen on contact and capped with magnetic screw caps (PAL #09151907) before being stored at -80°C until LC-MS analysis. For further details on HDX-MS protocols, including the instrumentation used for analysis, see Chapter 2.

4.2.3 Surface HDX Reactions

Surface HDX reactions were prepared by adapting the standard HDX protocol and scaling it to a large volume (1mL) reaction containing surface bound antigen displays created from high capacity (SA) agarose beads (**Figure 4.2.1**). Briefly, surface-Ag displays were created by first *lightly* biotinylating (1:1 molar ratio) hen egg lysozyme (HEL) using either NHS-biotin or NHS-PEG_{12}-Biotin linkers at room temperature for 30 min (refer to Chapter 3, **Figure 3.3.1.1**). The biotinylated HEL material was then purified/buffer exchanged using a 2mL VivaSpin centrifuge column (5,000 MW CO) at 4,000 x g for 20 minutes before being incubated with high-capacity SA beads for at least 1 hour on a rotator at room temperature. After washing the beads three times with fresh buffer, the beads are equally distributed such that each surface HDX sample contains ~100 uL of 50/50 surface-Ag bead slurry (~50 uL dry bed volume). Samples were then prepared by adding IgM (25 ug) to a total incubation volume of 1 mL before being placed on a rotator for 1 hour at room temperature. After incubation, the samples are washed again with 10X volume equivalents of fresh buffer three times, followed by removal of the supernatant to the bed-level of the beads (leaving ~20 uL

residual liquid volume behind). Samples are finally prepared for labeling by adding 80 uL of buffer (containing internal standards) to bring the total sample solution volume to 100 uL. Surface-HDX timepoints were started by adding 900 uL of deuterated buffer (85% D_2O final) and immediately placed on a rotator angled for gentle side-to-side rocking at room temperature. Before the end of each timepoint (~T-2 min) each sample is taken off the rotator and gently spun down with quick pulses in a micro table-top centrifuge in order to settle the beads to the bottom of the tube. Just before (~T-30 sec) quenching the timepoint, 875 uL of solution is carefully removed from the tube without disturbing the beads. Samples are then quenched by adding an equal volume (125 uL) of ice-cold quench solution for a total volume of 250 uL. After an additional 1 min placed on ice, samples were quickly spun down a final time, and 200 uL of solution was quickly removed (leaving ~50 uL of residual volume above beads) and transferred to pre-frozen autosampler vials contained within a bath of dry ice and ethanol, where samples are rapidly frozen and capped with magnetic HPLC screw-on vial caps before being stored at -80°C for later decoupled LC-MS analysis. Prebound IgM conditions were prepared by first co-incubating MD4 with unmodified HEL in 1X HBS buffer at various excess molar ratios for 1 hour at room temperature prior to being incubated with the beads as described above for all surface-HDX experiments.

4.2.4 Automated Reverse-Phase Liquid Chromatography Mass Spectrometry

All HDX time point samples were analyzed by RPLC-MS using a custom dual-column robotic HDX system, which is detailed in Chapter 2. Briefly, samples are temporarily housed in a cryo-sample manager at <60°C and are robotically transported to a temperature-controlled thaw block compartment held at 5°C for 3 minutes before being injected. Injected samples were digested at 20°C through an Nepenthesin II (Nep-II) protease column (2.1 x 20 mm; Affipro) with loading buffer (0.1% TFA, 2% ACN in optima water) at a flow rate of 400 µL/min. Peptides were trapped onto Waters Xselect CSH trap cartridges (2.1 x 5 mm, 1.7 µm, 130 Å) and resolved over Waters CSH C18 columns (1 x 100 mm, 1.7 µm, 130 Å) using an 18 minute gradient of 3% to 40% mobile phase B (MPA: 2% ACN, 0.1% FA, 0.025% TFA in optima-grade H2O;

MPB: 0.1% FA in optima-grade ACN). MS analysis was performed using a Waters Synapt G2-Si Q-TOF with ion-mobility separation (source and desolvation temperature set to 80° and 150°C, respectively). Between all analytical samples, an automated cleaning protocol containing a series of various wash solutions are executed to clean the protease and analytical/trap columns in order to preserve performance and limit carryover. The protease column is washed with a four-part injection series (250 uL each) in the following order: 1) 0.1% Fos-12 in 0.1% FA; 2) 2M guanidine HCl in 0.1% FA; 3) 10% acetic acid, 10% ACN, 5% IPA; 4) 10% FA. The analytical traps/columns are also washed using a four-part injection (250uL each) series and proceeds in the following order: 1) 10% FA; 2) 30% trifluoroethanol; 3) 80% MeOH; 4) 66% isopropanol, 34% ACN. In addition to these wash protocols, back-flushing of the LC-MS lines is also programmed to further clear out any large particulates or accumulating matter.

4.2.5 HDX-MS Data Analysis

The working database of Nep-II generated peptides was created by first passing an undeuterated MD4 sample replicate through the in-line protease column and LC-MS system and collecting the flow-through, followed by drying overnight via speed vac and resuspension in optima-grade water containing 0.1% FA to a total volume of 20 µL. Re-suspended Nep-II peptide mapping samples were then analyzed at the UW Proteomics Resource (UWPR; Seattle, WA) facility using high resolution MS/MS data generated via LC-MS using a Waters Acquity M-class nano-flow LC pump system coupled to a Thermo Fusion Orbitrap mass spectrometer. Commercially purchased trapping columns were used in conjunction with a self-made RP analytical column (~30 cm length; laser-pulled tip) packed with Reprosil-Pur 120 C18-AQ beads (5 µm; ESI Source Solutions) via 'bomb-loading' technique through a pressure cell provided by the UWPR. 6 µL of sample was injected and run using a 60 minute linear gradient (90 minute total run time) from 2% to 30% B (A: 0.1% formic acid; B: acetonitrile with 0.1% formic acid) with a flow rate of 300 nL/min. Data was acquired in data-dependent mode using EThcD fragmentation (15% collision energy) with dynamic exclusion enabled along with a targeted m/z trigger list for basic glycan fragment masses

(HexNAc: 204.0827 m/z; HexHexNAc: 366.14 m/z; Hex2HexNAc: 528.1928 m/z; mass tolerance ± 5 ppm), along with the following instrument parameters: ion source: 2.1 kV for positive mode; ion transfer tube temperature: 350 °C; scan range (m/z): 200 – 2000; MS resolution: 120000, MS2 = 30000. Peptide assignments were made using EThcD data with searches against the primary sequence of MD4 using Byonic v3.8 software (Protein Metrics), with results filtered by score and/or delta mod score (≥100) categories for searches of occupied glycopeptides. The retention time, IMS drift time, and signal strength suitability for all top scoring peptides were verified manually in Driftscope v2.5 (Waters) software using undeuterated sample data afforded from the Synapt G2. Deuterium uptake of the single best charge state (chosen based on signal strength and quality across all time point replicates) for each peptide was analyzed and summarized using HDExaminer v3 software, and in cases where multiple glycoforms were found (including unoccupied versions), multiple HDExaminer files were created. Significance was determined automatically by HDExaminer using a 95% confidence interval.[17] Internal imidazolium compound exchange standards were included in data sets to assess equivalence of labeling conditions along with bradykinin peptide used to report on back-exchange.[18,19] The combined tables of all relevant statistics are included at the end of this chapter and precede the uptake curves of each peptide generated by HDExaminer for each discrete experimental set.

4.2.6 HDX-MS Data Modeling

A solution state model of pentameric IgM (refer to Chapter 1, **Figure 1.2.1**) was created in Pymol v2.8 using the following PDB IDs: 6KXS[12] (Human Cμ3/Cμ4/Jchain), 4JVU[20] (mouse Cμ2), 1DEE[21] (human Cμ1), 3HFM[22] (Human vH and kappa light chain vL/cL). A model of staple IgM was created by manual displacement of the Fab'$_2$ units (Cμ2+both Fab domains) approximately 90° downward from plane of the Cμ3/Cμ4/J chain Fc core using an intact human IgM pentamer model generated by AlphaFold with the arms aligned to the Cμ3/Cμ4 domains of PDB ID 6KXS. Deuterium uptake percentages determined through HDExaminer (using theoretical max of each peptide; first two N-term residues excluded) were

summed across all timepoints within each experimental condition. Peptide uptake differentials were then calculated and rounded to the nearest 5% using Microsoft Excel, and the color scale was adjusted accordingly with evenly spaced 5% intervals on ± 40% change scale. All experimental sets were mapped using the same coloring scale, and the mapping process was executed in Pymol using scripts with the coloring order arranged from N-term to C-term by peptide length (largest to smallest).

4.3 Results

Guided by the working hypothesis that only the multivalent surface bound conformation of IgM (i.e. 'staple') is capable of activating complement, we looked to determine the structural features that are both unique and fundamental to the active form through a series of differential HDX-MS analyses of the 'nominal', or maximally-performing surface bound condition, against other sub-active (bound) and inactive (unbound) forms. The functional assessment of these conditions were previously assessed in Chapter 3 through kinetic (BLI) and functional (C1-activity) characterization assays. The structural features of the nominal max-avid surface bound form of IgM is the primary state by which we make the following HDX differential comparisons:

4.3.1 HDX-MS Analysis of Surface Ag Linker Length Effects (Extended vs Short Linker)

As presented in Chapter 3, our surface Ag display engineered from high-capacity SA agarose beads and lightly biotinylated monomeric HEL antigen was demonstrated to be biologically relevant, and successfully facilitated the multivalent surface binding that is hypothesized to be a mechanistic prerequisite to the formation of complement-active staple IgM. Compared to the relatively short (1.35 nm) standard commercial NHS-Biotin linker, the extended (5.6 nm) NHS-PEG_{12}-Biotin linker was demonstrated via BLI to significantly improve IgM binding magnitudes along with subsequent hC1q binding magnitudes (refer to Chapter 3, **Figure 3.3.1.2**), which is indicative of promoted IgM staple

formation, as a larger population of staple IgM is predicted to induce a greater C1q/C1 binding response. We also validated the biological relevance of the SA bead Ag surface display through an activity assay that monitors the initiation of the classical pathway by measuring the conversion rate of proC1s into its active proteolytic fragments, which is the primary endpoint of C1 activation (refer to Chapter 3, **Figure 3.3.5.2**). Since complement activation depends on the bound conformation of IgM, and the bound conformation of IgM is strictly contingent on the initial display of antigen, we first looked for potential structural differences induced by extension of the surface Ag biotin linker (NHS-PEG_{12}-Btn vs NHS-Btn) of the SA beads. Based on the improved BLI binding magnitudes that we observed upon switching to the PEG_{12} linker, we hypothesized that linker extension directly improved surface Ag accessibility, which further promoted the formation of complement-active IgM. We reasoned, therefore, that the net increase in surface binding valency would be reflected within this initial HDX comparison, and that structural differences would be found primarily within regions of the antibody that are responsible for mediating the formation of staple itself.

The resulting peptide coverage of IgM via inline Nepenthesin-II (NepII) digestion was extensive, resulting in a total of 145 unique peptides resolved and tracked by HDX-MS that covered approximately 95%, 98%, and 44% of the light chain (κ), heavy chain (μ), and joining chain (J), respectively, for a total protein sequence coverage of ~89.5% (see **Supplemental Table S4.1** and related uptake plots for further detail). Because the J chain is present at a much lower molar ratio within the pentameric IgM molecule, the signal strength of J chain peptides was often weak and difficult to detect, which resulted in notably lesser coverage compared to that of the light and heavy chains.

The differential HDX analysis between the short and extended surface Ag biotin linkers revealed a number of changes throughout every major heavy domain of IgM, with the largest differences observed primarily within the Cμ2 – Cμ4 domains of the Fc region (**Figure 4.3.1 A**). The net effects of binding to surface Ag conjugated with the extended PEG_{12} biotin linker are shown in **part B of figure**

4.3.1, where the sum differences in uptake percentage (from theoretical max) across the 1 min and 30 min time points were calculated for each peptide and mapped to a 3D model of pentameric IgM (see section 4.2.6 for further modeling details). Because this comparison was confined to surface bound IgM only, with the sole variable being the linker used to create the surface Ag displays within each condition, the structural features identified from this experiment are thought to point primarily towards regions of IgM that are fundamentally involved in staple formation. The most striking differences were observed within regions near the Cμ2/Cμ3 domain junction (328-337 of Cμ2 domain and 379-393 of Cμ3 domain), where a significant increase in exchange protection (~35%) was observed within the extended PEG_{12} biotin linker condition. Immediately preceding this Cμ2 junctional peptide was a region (316-327) of significant de-protection/exposure, along with a moderate degree of exposure observed near the Cμ3/Cμ4 junctional 'elbow' regions (534-553). While this initial look was not enough to establish any definitive patterns or mechanistic interpretations, it did provide an initial glimpse into regions that are likely to be involved in staple formation, especially if the exchange trends were upheld throughout subsequent experiments.

In addition to the structural differences we observed, it should be noted that we also observed a marked decrease in the MS signal strength for peptides generated from the extended PEG_{12} biotin linker condition, which is another practical indication of increased IgM binding avidity. With the average degree of binding multivalency increased through extension of the biotin linker (via increased accessibility to surface Ag), the average IgM binding avidity is also increased, which results in a significant decrease to the average dissociation rate. This principle was demonstrated in the previous chapter through BLI testing of a full prebound ratio series (refer to Chapter 3, **Figure 3.3.4.2**), where increasing dissociation rates were observed as IgM surface binding valency was progressively decreased. To reiterate, our IgM system ("MD4") was originally expressed as a transgenic hybridoma that combined the specific and potent binding affinity of the HyHEL-10 IgG ("H10") Fab domains with the highly

multivalent architecture of pentameric IgM, which together results in an extremely tight binding Ag:Ab system with an effective combined affinity (i.e. avidity) estimated to be at least in the femtomolar range ($K_D < 1E^{-12}$ M).[22–25] For context, the change in binding affinity from the monovalent, H10 Fab-only material to the bivalent, whole H10 IgG was ~100X (rather than merely 2X or 4X), so the extreme affinity (avidity) of IgM MD4 is not surprising. During method development we discovered that the strength of multivalent surface binding by IgM MD4 to the optimized/extended PEG_{12}-HEL surface antigen was so strong in practice that it remained resistant to dissociation, *even when combined with quench buffers containing 200 mM TCEP and the maximum soluble amounts of urea (8 M), guanidine HCL (6 M), and/or combinations of both* (**Figure 4.3.2**).

Despite our attempts throughout method development to optimize the quench solution for maximal dissociation of IgM from the PEG_{12}-HEL surface Ag beads, we could not find a feasible chemical solution that would afford enough MS signal strength from the nominal (max-avid) surface bound IgM condition prepared by freezing immediately upon quenching. It is for this reason that adjustments had to be made to the surface HDX labeling protocol, with post-quench exchange samples being thoroughly mixed and placed on ice for an additional minute before being flash-frozen in order to maximize dissociation and achieve adequate MS signal strength (refer to **Figure 4.2.1**). While none of the other IgM conditions that we tested required this additional step (including surface bound IgM to SA beads conjugated with the short/standard biotin-NHS linker), they were nonetheless treated equally with the nominal surface condition and were likewise placed on ice for an additional minute after addition of the quench solution.

4.3.2 Three-State HDX (Unbound vs Solution Bound vs Surface Bound)

As previously outlined in Chapter 1, one of the biggest issues presented within the classical literature stems from the original debate over the mechanistic driving force behind the formation of

staple. While one side hypothesized that allosteric changes induced through binding of soluble monomeric Ag were sufficient to propagate the necessary structural changes to the Fc core, the other side maintained that the global conformation could only be induced (and further stabilized) through a severe physical distortion induced by the multivalent binding of IgM to a fixed antigenic surface.[2] To address this issue directly we next performed a 3-state HDX-MS analysis of unbound, solution bound (to saturation, using monomeric HEL protein Ag), and nominal (max-avid) surface bound conditions, with two separate differential analyses performed in tandem in order to 1) search for and catalogue any potential allosteric changes that occurred as a result of binding monomeric HEL Ag in solution, and 2) differentiate any potential allosteric changes from those unique to surface bound IgM.

The resulting NepII peptide coverage was similar to that of prior experiments, with a total of 143 unique peptides observed that covered ~95%, ~99%, and ~37% of the light chain, heavy chain, and J chain, respectively, for a total protein sequence coverage of approximately 88.6% (see **Supplementary Tables S4.2** for complete HDX set statistics).

In order to first look for any potential allosteric changes induced by Fab binding in solution, an initial differential analysis was performed in which solution bound IgM was compared against the unbound state (**Figure 4.3.2.1**). While the largest differences were observed to be localized to the light and heavy chain complimentary determining regions (CDRs) of the Fab domain, which was expected for a comparison to unliganded IgM, a surprising degree of allosteric change was also observed, with structural differences extending all the way from the Fab domains to the Fc core of the Cμ3/Cμ4 domains. While the magnitudes of the allosteric Fc changes were relatively minor across the board (only 5-10% avg), we are confident that these changes are nonetheless significant and indicative of an unexpectedly extensive network of allosterism. The only major regions of exchange protection were seen within the Fab domains where protein-protein binding interactions occur between the paratope of the IgM Fab domains and the epitope of the HEL protein antigen, while the rest of the allosteric changes

throughout the Fc were observed to be predominantly de-protected/exposed to solvent, which could be indicative of a global increase in segmental flexibility induced throughout the heavy chain domains as a result of antigen binding.

In the second differential analysis, the nominal surface bound IgM condition was compared against the solution bound form in order to differentiate staple-related structural features from those potentially caused by allosterism (**Figure 4.3.2.2**). There were no significant differences found within the CDRs of the Fab regions, which indicated that both conditions were bound equivalently to antigen. We also observed a global reduction in the MS signal for surface bound IgM peptides (despite adding ~1.5X more IgM to the surface samples), which matched our previous observations during the linker comparison and further indicated the nominal surface bound condition to be bound in a highly multivalent fashion, which is fundamental to the underlying hypothesis of staple formation. The trends observed throughout this second differential comparison (surface vs solution bound) varied substantially in terms of relative protection from and/or exposure to deuterium uptake trends compared to the allosteric analysis, which showed predominantly de-protected/exposed regions throughout the entirety of the Fc. While the magnitudes of the HDX differences were also relatively minor in scale, many of the surface-unique trends resembled those found throughout previous experiments, with a notable degree of protection observed again near the Cμ2/Cμ3 domain junctions. A significant degree of protection was also observed within the C-terminal 'μ-tailpiece' regions of the Cμ4 domain, along with some distinct areas of exposure within the Fc core domains (Cμ3/Cμ4) that mirrored the allosteric changes observed in solution bound IgM. While we currently lack the high resolution structures needed to fully interpret the HDX differences observed within surface bound IgM, the regions of overlap between surface and solutions bound forms may help to explain the C1 activity we observed for solution bound IgM within the activity assays presented in Chapter 3.

4.3.3 Nominal vs Prebound Surface HDX

Although the conditions of the 3-state HDX experiment were relatively well-matched, we experienced some technical difficulties that resulted in the loss of a number of time point replicates, resulting in a statistically weaker set than had hoped for. On top of this, there was a concern that the comparison between surface and solution bound IgM could have been potentially skewed due to the physical presence of the agarose beads themselves within the surface bound condition. While care was taken to ensure that all of the conditions were treated as equally as possible in terms of both sample handling and labeling/quenching protocols, we felt that an even more robust comparison could still potentially be made that would add a final layer of confidence to the list of structural features and regions that we had thus far categorized as being unique to staple IgM.

In Chapter 3 we employed a prebinding strategy to attenuate the relative surface binding valency and avidity strength of IgM prior to testing by BLI and C1 activity assays, which takes advantage of the extreme multivalent binding strength of IgM MD4 to a surface of immobilized PEG_{12}-HEL antigen. While the biological relevance of a prebound surface IgM molecule is dubious, we hypothesized that highly prebound (~9:1) IgM molecules bound to a surface could still serve as a useful analytical surrogate to that of solution bound IgM, despite the obvious caveat of the one arm that is ultimately bound to the surface-Ag display**.** To ensure that the major HDX trends and regions of surface-unique structural changes were real we conducted a final large-scale, surface-only HDX experiment designed to compare the nominal, maximally avid surface bound IgM condition to a highly prebound, minimally avid surface bound form. The two conditions were identical in SA bead concentration, reaction volume, sample handling, and labeling/quench protocols. Furthermore, any free/loose HEL antigen particles or fully saturated (10:1) IgM molecules that may have persisted through initial prebinding reactions were naturally removed throughout a series of extensive wash steps, prior to the start of exchange time points. Because the final readout of HDX-MS is averaged across all of the various conformational sub-

populations present within any given sample, we hypothesized that any major structural features unique to staple IgM would still be differentiable from the excess ~20% (1/5) of signal for each peptide that would potentially be altered by the surface bound arm(s) within the prebound condition**.** While this experiment does not perfectly recapitulate the comparison between surface and solution bound forms of IgM, the two conditions were inherently identical in terms of labeling conditions and agarose bead presence, and the only physical variable between the two forms was the decreased surface binding valency of the highly prebound condition.

The resulting NepII peptide coverage was similar to that of prior experiments, with a total of 141 unique peptides observed that covered ~95%, ~97%, and ~45% of the light chain, heavy chain, and J chain, respectively, for a total protein sequence coverage of approximately 88.9% (see **Supplemental Table S4.3** for complete HDX set statistics). To ensure an increased statistical relevance, time point samples were prepared in quadruplicate, and the resulting confidence interval (95%) was less than 0.2 (D) per timepoint. The butterfly and residual plots, along with the 3D exchange heatmaps are **shown in Figure 4.3.3.**

While a few unanticipated regions of significant exposure were observed within the Cμ1 domain (near the Cμ1/Cμ2 junction), these changes could likely be the result of the aforementioned caveat of prebound IgM, which was heterogeneously bound, by default, due to the physical differences in the prebound arms vs the arm(s) that were bound to the SA bead surface. As discussed previously in Chapter 3, the prebinding strategy inherently leads to heterogeneity in both the final bound configurations and subpopulation of IgM within any given prebound ratio. However, in this experiment we were aiming only for a highly prebound state (9:1) which likely resulted more in the loss of fully saturated (10:1) IgM molecules than an increased prevalence of lesser (i.e. 7:1 or 8:1) prebound subpopulations. While some heterogeneity is likely to still have persisted within the data from this experiment, the uptake results of the prebound condition were largely reminiscent of solution bound

IgM. More importantly, the differential analysis between the nominal and highly prebound (~9:1) surface conditions confirmed the major trends and locations of key structural change that we predicted to be unique to staple IgM, based on the portfolio of results generated throughout our previous HDX-MS experiments. We once again observed a significant degree of protection within the Cμ2/Cμ3 domain junction, which was one of the most prominent surface-unique features that we had observed throughout all of our HDX-MS experiments. There was also exposure observed near the outer 'elbow' regions of the Cμ3/Cμ4 domain junctions, and the C-terminal μ-tailpiece regions were once again found to be significantly protected for the nominal surface bound form of IgM.

The exchange conditions were confirmed to be highly matched by the overlapping uptake observed in a number of internal exchange standards (see **tables S4.1, S4.2, and S4.3** that preface the related uptake plots of each experiment). The degree of antigen binding was also found to be equivalent, as evidenced by the overlap in uptake across the light and heavy chain CDR peptides of the Fab regions. The back-exchange was also even between conditions, and minimized to <20% across all of the time point samples, as indicated by the retention of the deuterium label in our back-exchange reporter peptide (bradykinin; refer to Chapter 2, **Figure 2.3.3**). Together these conditional criteria indicate the comparison to be analytically robust, and the observed changes can therefore be confidently inferred to result directly from structural differences and conformational dynamics of the antibody itself.

4.4 Discussion

While HDX methodology has been in use since the 1950s, its combination with mass spectrometry is relatively new (late 1990s), and the added versatility has continued to bring about many novel approaches and modernized applications reported throughout the recent literature. Despite the continued expansion of its utility, there are still inherent limitations and important caveats that must be

considered when attempting to translate HDX data into physical interpretations of protein conformational dynamics, especially if structural information is limited or outright missing, as in the case of IgM. Although we observed a number of significant and consistent surface-unique trends over the course of multiple HDX-MS experiments, our mechanistic interpretations of staple formation remain inherently limited due to the complete lack of a resolved structure of IgM in any of its fully intact monovalent or polyvalent forms. This hurdle is what motivated us to approach the project iteratively, with the goal of building up enough historical results from varying differential comparisons against the nominal surface bound condition, in order to elucidate the strongest emerging patterns that might provide us with added confidence and further statistical rigor.

IgM has gained notoriety for being a challenging target of basic research due largely to physical characteristics such as its extensive size, flexibility, and glycosylation. Nonetheless, we set out to study IgM in its physically native form, under physiologically-native conditions, in order to elucidate the structural features that define its biologically and clinically relevant complement-active form. In general we did not observe any changes by HDX-MS that would indicate the breaking of secondary structure (no high magnitude differences), which indicates the mechanism of staple formation to be likely based on tertiary/quaternary domain rearrangements only, with the bulk of the conformational changes mediated by junctional regions between the conserved heavy domains of the Fc. We did, however, find a number of consistently significant structural changes throughout each conserved heavy μ domain of the Fc that were unique to the nominal surface bound form of IgM that could point towards mechanistic features of staple formation as well as provide potential insights into the identity of residues that comprise the C1q binding interface.

4.4.1 The Allosteric Network of Solution Bound IgM Extends Deep into the Fc Core

"If we accept the concept that the structure of a protein in solution is not static but rather a dynamic equilibration between all of its kinetically accessible forms, then the sites which interact with complement may flicker into and out of existence prior to the interaction of IgM with antigen." – Arthur Feinstein, 1981[2]

While it is now appreciated that multivalent surface binding of IgM is required for C1 activation at a scale that is biologically meaningful, the potential effects of allosterism should not be disregarded entirely. In Chapter 3 we observed a low but consistent degree of C1 activity in solution-state IgM MD4 bound to saturation with monomeric HEL antigen, which may be the reasoning behind the disparate phenotypic observations and confusion that emerged early-on throughout the classical literature concerning intermittent complement activation by unbound and solution bound forms of IgM. We maintain that a unique set of specific structural rearrangements (i.e. staple formations) underlie the relatively high activity rates that we observed for the nominal surface bound condition, but it is still nonetheless possible, however, that some of the allosteric changes we observed by HDX-MS (**Figure 4.4.1**) could *promote* the transient formation of local staple formation within any given arm of solution bound IgM. It would appear that binding to monomeric Ag in solution, at least within our IgM system, does induce an extensive degree of allosteric structural change, as well as convey a low degree of C1 activity. If we presume that structural changes lead directly to changes in functionality, then together these results may provide a biophysical basis for why the early proponents of allosterism were likely to be at least *partially* correct about the driving force of staple formation.

Multiple key studies were published throughout the mid-1970s by proponents of allosterism, and in each case the presented evidence led to the conclusion that solution bound IgM (and in some cases even unbound IgM) was readily capable of activating complement. In 1975, Brown & Koshland reported that monomeric presentations of small antigenic molecules bound to a larger carrier protein (i.e. 'haptens'), were able to promote the binding (i.e. 'fixation') of IgM to guinea pig complement at an equivalent rate

to that of multivalent Ag displays, which supported the notion that multivalent surface binding was not a prerequisite to meaningful complement activation.[26] They could not reach this conclusion, however, without also disclosing that 1) IgM bound to the unconjugated version of the antigen itself was devoid of complement activity, and that 2) the Ag molecule by itself was also capable of inhibiting the activity induced by their larger hapten displays created via conjugation to a carrier protein. On top of this, they also concluded that C1 binding activity was largely dependent on the size of the carrier protein used to create their Ag displays. Despite these gaps in their case to support an allosteric driving force behind IgM staple formation, the study was nevertheless followed up by further corroboration; the conclusions of Brown & Koshland were directly supported the following year (1976) by Pecht, who also concluded that allosteric change was sufficient for triggering the activation of complement, as long as the antigen particle was "*of significant size and conformation*".[27]

While we cannot at this time offer a concrete explanation for why the size of a monomeric antigen particle would change the functional outcome of a purely allosteric mechanism, we were nonetheless surprised to find that not only does an allosteric network exist within IgM, but the structural changes that originate at the Fab domains also appear to translate all the way into the Fc core. In Chapter 3 we repeatedly saw a low degree of C1-activity in solution bound IgM, and there were also a number of structural changes that we found by HDX-MS that appear to be shared between the solution bound and nominal surface bound forms. While it's possible that allosteric changes promote the formation of staple, it's also possible that the act of binding is indirectly disruptive of pre-existing interactions that stabilize the arms of IgM in its unbound form, which would be supportive of recently published evidence that indicated a stabilized association between the Cμ2 and Cμ1 domains that confers a degree of rigidity to the entire F(ab)'$_2$ unit, which is defined as Cμ2 and both attached Fab domain lobes of a single arm.[28] If allosterism is enough of a force to disrupt an otherwise naturally flexible F(ab)'$_2$ unit, then perhaps the explanation lies in the propensity to form a localized and temporary/transient staple conformation at any

given Ag bound arm of solution bound IgM. This could explain why we observed unbound IgM to be relatively C1-inert, while solution bound IgM repeatedly showed a low but significant degree of C1-activation potential (refer to Chapter 3, **Figure 3.3.5.2**).

While our functional C1 activity assays clearly demonstrated the nominal surface bound form of IgM to consistently be the most potent binder/activator of C1q/C1, we also observed unbound IgM to be completely devoid of C1 activity, which forces us to conclude that allosteric changes do, at least to some extent, convey a degree of C1 functionality to the solution bound form of IgM. We also observed a significant degree of binding and C1 activation in all of the prebound forms of IgM we tested, which indicates that 1) IgM is indeed a highly flexible molecule capable of binding both free monomeric Ag in solution as well as surface-immobilized Ag, and 2) staple formation is likely not a globally-concerted mechanism or cooperative structural transition. If allosterism induces structural changes that liberate the Cμ2 domains and increase the segmental $F(ab)'_2$ flexibility of the arms, then staple could perhaps be formed locally and transiently in solution along any given Ag bound arm of IgM. The size of the Ag could perhaps be involved in the degree to which the pre-existing stabilizing interactions of unbound IgM are disrupted, which would explain the classical findings of Brown & Koshland, among others.

Regardless of the extent to which allosterism contributes towards the activation mechanism itself, the activity of solution bound IgM still paled in comparison to the potency observed in the nominal, max-avid surface bound form of IgM. If transient staple formation does occur more frequently in solution bound IgM, it is still highly unlikely that all five or six arms of pentameric or hexameric IgM would both transiently and simultaneously adopt the staple conformation while bound to monomeric antigen in solution. Even if it did occur, the time scale at which those transitions would last is likely to be insufficient to activate complement to any biologically-meaningful degree. The complement cascade is a highly destructive, pro-inflammatory signaling pathway that is also highly-regulated by various inhibitory elements and feedback loops. If activation *in vivo* were to occur frequently enough by solution bound

IgM to significantly activate the cascade, the complement system would probably not be as widely conserved as it is, throughout all vertebrate species.

While some of the allosteric HDX trends that we observed point to regions that overlap with the nominal surface bound form, there were many locations and significant changes found that were unique only to surface bound IgM. In the following sections we discuss the major HDX trends that were observed throughout each of the constant μ heavy domains that were found to be upheld throughout the combined history of all of our nominal surface HDX-MS experiments.

4.4.2 Surface-Unique Changes Observed within the IgM Fab Domains

In order for differential comparisons of distinct bound forms of IgM to be fair, it is critical that the IgM within each condition be bound to an equal degree – ideally to saturation in all cases. Just as labeling conditions and back-exchange can negatively impact the comparative power and statistical relevance of differential HDX-MS experiments, the multivalency of IgM requires that all of the arms of a single condition be homogenously bound to Ag, and that structural comparisons against bound arms of other distinct conformational states be made truly among bound arms only, as unbound Fab domains could likely add a significant degree of structural heterogeneity, especially if the act of binding changes the flexibility range of the $F(ab)'_2$ units. In demonstration of this concept, a preliminary test was performed against a surface bound condition prepared from low concentration (0.1X) PEG_{12}-HEL Ag solution with the resulting uptake curves in the CDR3 of the heavy chain (HCDR3; 94-102 NWDGDY) found to be a good reporter of the relative difference in unbound Fab populations, which is corroborative of past reports suggesting the HCDR3 to be the primary point of binding contact for most antibodies (**Figure 4.4.2 A**).[29,30] After initial screening of the various internal standards that monitor labeling conditions and back exchange for each major HDX-MS experimental set, we first looked to the IgM Fab domains (primarily the CDR3 of the heavy chain; HCDR3) to verify that deuterium uptake within

the binding interfaces was overlapping throughout all sampled time points and between differing bound conditions.

The Fabs were found to generally be the most structurally heterogeneous region in terms of HDX trends over the course of our experiments. To reiterate, all of the experiments presented throughout this dissertation were based on the initial presentation of antigen, which increases the relevance of the data, but also adds an inherent degree of variability that required thorough optimization and rigorous sample preparation in order to maintain consistency in IgM binding and subsequent hC1q interactions throughout all surface HDX experiments. The consistency of antigen preparation and SA bead surface conjugation was vital to maintain throughout all of our experiments, as variance in surface Ag displays (e.g. non-uniformity of Ag biotinylation, loading density, and individual spacing) would likely add a high degree of variance to the structural results obtained by HDX-MS. While it can be difficult to envision the exact arrangement of surface bound Fab domains within each experiment (especially in prebound surface comparisons), we nonetheless observed a few regions of structural change within the Fabs that may be insightful for multivalent surface binding of IgM.

For example, in the nominal vs prebound surface differential experiment, we observed a strong degree of de-protection/exposure near the C-terminus of the Cμ1 domains that may be indicative of the strain conferred to the Fab domains upon multivalent surface binding (**Figure 4.4.2 B**). In the prebound condition, the majority of the arms are bound to free monomeric HEL Ag molecules and, while we cannot be certain of exactly which arms are precluded from binding to the immobilized PEG_{12}-HEL surface, we predict the prebound Fabs/arms to be generally less physically strained compared to the arms in the nominal surface condition that appear to be fully bound to the antigenic surface. To accommodate such a high degree of multivalent binding, the Fabs are likely to be physically forced into an arrangement that can only be stabilized by the surface itself, resulting in more deuterium uptake than

prebound Fabs that are not forced to interact with the surface. In the 2019 EM data published by Sharp et. al., the Fab domains appear to be bound to the surface in a staggered arrangement that places one Fab domain of each arm underneath the Fc platform itself, while the other is angled outward (refer to Chapter 1, **Figure 1.5.3**). If this arrangement is also adopted by the IgM molecules in the nominal, max-avid surface bound condition within our experiments, then it would be reasonable to predict that HDX differences would arise from comparisons to the majority of Fab domains that are not bound to the surface, resulting in the high exchange differences we observed.

4.4.3 IgM Staple Formation is Facilitated by Hinge Motion of the Cμ2 Domain

In 1971 Arnold Feinstein and colleagues described their newly coined "staple" conformation of IgM as a state in which "[sic]…*many, or sometimes possibly all ten, arms of the yM molecule are no longer extended radially as in the free yM molecule but are folded down from the central disc…*", with the molecule forming a three dimensional structure at the surface that resembled a *"…multilegged table, seen in profile as a staple*." [4]

While much has changed within our collective knowledge of immunology and structural biology throughout the half-century of research that has since followed Feinstein's original description of the IgM staple conformation, our understanding of the mechanistic details behind its formation remains almost just as limited today. The elucidation of a highly resolved, intact structure of IgM has been hindered by its large size, prominent flexibility, and extensive degree of glycosylation and glycan heterogeneity. Recent attempts using advanced cryo-EM approaches have started to gain significant ground, however, with high resolution structures of the partially intact IgM pentamer being published in the past few years that provide deeper insight into architecture of the Fc 'core', which is the central ring composed of the Cμ3/Cμ4 domains and J chain that acts as a placeholder for the sixth potential arm of the hexameric isoform.[12,13,31] While it is unfortunate that these modern approaches have thus far failed

to produce a completely intact structure of IgM**,** it is still intriguing and perhaps nonetheless informative to consider that all recent attempts have failed in the exact same fashion, with resolution dropping off precisely at the N-terminus of the Cμ3 domain (refer to Chapter 1, **Figure 1.2.3**). In every case, the flexibility that originates in or near the residues that comprise the Cμ3/Cμ2 domain junction has rendered the rest of the molecule unresolvable which, in our view, is a practical consequence of a highly flexible hinge that exists within that region.

Although a definitive agreement has still yet to be reached in the literature, it should be noted that hinge function of the Cμ2 domain is not itself a new concept. Evidence of Cμ2 domain involvement in the activation of complement was published as early as 1981, where Siegel & Cathou demonstrated that heat denaturation of the Cμ2 domain abrogated complement activity entirely, and that the activity could be restored upon successful refolding of the domain.[32] Bending of the $F(ab')_2$ segments (Cμ2 + both Fabs) down and away from the plane of the Fc core 'platform' was also central to the predicted mechanism of staple formation as outlined in the 1991 Perkins model of solution-state IgM, where the authors (Feinstein included) ultimately proposed a "*purely steric accessibility model*" for the exposure of C1q binding sites along the periphery of the Cμ3, which contrasted from the highly debated notion that allosterism and/or physical distortion might cause a conjunction of residues to form a C1q binding site that was previously unformed or otherwise inaccessible.[5]

In the case of IgG, the length and composition of the hinge motif is a strong predictor of the capacity of each IgG isoform to activate complement, as steric clearance of the Fab domains is fundamentally required to make room for the globular heads of C1q to bind at the Cγ2 domain.[33–35] Comparisons with IgG based on domain homology leaves the Cμ2 of IgM unmatched, and while IgM has not historically been considered to contain a hinge, the data collected over decades' worth of EM observations (low-resolution notwithstanding) has provided enough evidence to make the case of hinge function originating somewhere between the Fc platform and Fab domains which, by default, further implicates

involvement of the Cμ2 domain.[1,4,36] A model of surface bound IgM based on cryo-EM with tomography was published in 2019 by Sharp et. al. in which the Cμ2 domains were fit at an angle of 100° relative to the plane of the Fc core disk (80° inward bend), as illustrated earlier in Chapter 1 (refer to Chapter 1, **Figure 1.5.3**).[1] Furthermore, in a very recent (Oct 2022) report published in Nature Communications, Chen et. al. presented data from single particle cryo-EM and FRET-based analyses of solution state IgM, and reached the conclusion that a hinge was located within the Cμ3/Cμ2 domain interface based on observations that the Cμ2 domain was able to pivot "*both in-plane and out-of-plane*" relative to the Fc core.[28] The authors also noted that the Cμ1/Cμ2 interface could experience a degree of flexion as well, and we observed a corresponding degree of structural change within the C-terminal region of the Cμ1 domain that extends into the N-terminus of the Cμ2, with a trend in protection analogous to the Cμ2/Cμ3 domain junction.

The most consistent and prominent exchange trend that we observed for surface bound IgM, throughout all of our HDX-MS experiments, was widespread protection from exchange throughout the entirety of the Cμ2 domain, with an especially striking region of protection localized to the residues that comprise the Cμ2/Cμ3 domain junction. This region, which begins at the C-terminal glycopeptide of the Cμ2 domain (328 – 337; TFLKN*VSSTC), was significantly protected for the nominal surface bound condition relative to every other form of IgM we tested, including the surface-only comparison against IgM bound to the surface immobilized HEL conjugated with the short biotin-NHS linker (**Figure 4.4.3.1**). In the allosteric analysis of solution bound vs unbound IgM we observed a distinct lack of protection throughout the Cμ2 (and in fact, throughout all of the Fc), but within this specific junctional region we observed an overlap in deuterium uptake between the two conditions, indicating that either 1) allosteric changes do not occur within this region to any significant degree, or 2) exchange is inherently rapid in this region due to a high degree of structural flexibility, which causes both the unbound and solution bound forms of IgM to become fully-exchanged before the end of the first time point that we sampled,

which could only be reliably shortened to 1 min due to the nature of handling the bead containing samples. The Cμ2/Cμ3 domain junction was, however, consistently preceded by a prominent region of de-protection/exposure within the ~8-10 residues that come directly before it (316-327; NVYTCRVDHRGL), which is an intriguing allosteric feature that appears to be shared to some extent between both surface and solution bound IgM. If the exposure at this preceding region leads mechanistically to staple formation, then the fact that both solution and surface bound IgM share HDX trends here may point towards an explanation of how solution bound IgM is able to activate C1. Because the nominal surface bound form of IgM is inherently more rigid than either solution bound or the highly prebound surface conditions, the protection we see at the junctional hinge may be a direct consequence of staple formation that persists to at least 100 min. The exposure in the region that comes before it would therefore be shared amongst the two conditions, while the protection at the hinge itself would only appear in the nominal surface bound form, due to the high degree of flexibility in solution bound IgM which enables it to rapidly pick up the deuterium label.

The significant protection that we consistently observed in the Cμ2/Cμ3 junction for the nominal surface bound condition could likely be reflective of a highly stabilized transition mediated by the folding of a mechanistic hinge. The surface-unique changes that we observed in the H/D exchange rate would be reasonably expected for a protein region that undergoes quaternary/tertiary domain compression that may physically reduce solvent accessibility and/or increase the local structural stability via increased hydrogen bond formation. If anything, the fact that no other forms of IgM displayed protection in this region would indicate this result to be a unique and stable structural feature of surface bound IgM. Despite the current lack of available structural information throughout all of the interdomain regions of Ig M, we are confident that our sequence/peptide IDs are accurate and that the changes we observed are reflective of a structural transition that is unique to the staple conformation of surface bound IgM.

While allosteric effects may have promoted the transient formation of staple, which we hypothesize to be the cause of the C1 activity observed in solution bound IgM, the activity itself was still minimal compared to that of the nominal surface bound condition (refer to Chapter 3, **Figure 3.3.5.2**). This result indicates that allosterism alone is likely not enough of a driving force to trigger the full-scale activation of complement. On the other hand, the fact that solution bound IgM had any activity at all (relative to the C1-inert, unbound form of IgM) would also indicate that surface-induced physical distortion is not a mechanistic prerequisite to C1 activity, either. What is clear, however, is that 1) a multivalent display of Ag is required for maximal C1 activation, and 2) IgM must bind to it with maximal valency/avidity, in order to display what we predict to be an *ensemble* of discrete local staple conformations adopted by all of the available binding arms, each of which likely presents a copy of the binding site(s) needed to facilitate yet another multivalent binding interaction with the hexameric C1q molecule. Due to the relatively low magnitudes of HDX differences observed throughout the entirety of the antibody, we are in favor of the view that staple formation is governed by steric accessibility and determined strictly by the spacing and availability of the antigen display, rather than the sole result of allosterism or any significant surface-induced physical distortion that would otherwise be conformationally unobtainable naturally in solution.

In a way, both sides of the original driving force debate were partially correct, as 1) allosteric changes do appear to extend deep into the Fc core and induce an inherent degree of activity within an otherwise C1-inert unbound IgM molecule, and 2) a fixed antigenic surface does greatly enhance the activation rate of C1 and is likely the most potent form of antigen display. We note also that complement activity has been documented to stem from soluble forms of polyvalent antigen particles (such as branched dextran chains), so long as the Ag:IgM density ratio lies within the optimal range to facilitate the maximal degree of multivalent surface binding and subsequent staple formation.[37] What

appears to matter most, however, is the correct presentation and spacing of a fixed array of available Ag epitopes – regardless of the carrier protein, bead surface, or biosensor used to display them.

One major question that still remains is if the Cμ2 hinge is able to pivot in both vertical directions (+/- 90° relative to the plane of the Fc platform), or if staple can only be formed in a single direction due to a physical asymmetry inherent to the IgM domain architecture (**Figure 4.4.3.2**). In 2009, Czajkowsky & Shao used atomic force microscopy (cryo-AFM) to show that IgM is not a strictly flat/planar molecule in solution, as it is often depicted, but rather they described the three-dimensional structure as being "mushroom shaped" due to a raised portion of centralized density that they attributed to the Cμ4 domains. Using IgE homology and free energy calculations they predicted the molecule to have a flexural bias in terms of its Fab rotation relative the plane of the Fc disk, and indicated that it was likely more energetically favorable for the IgM pentamer to engage surface antigen by folding of the arms in a single direction (the "–β side"), with a C1q binding interface formed across the opposite face of the Fc disk (the "+β side").[38] While these predictions were based on the structure of IgE, which imposed excess restrictions on segmental flexibility and precluded the consideration of the J chain entirely, the premise of a physically asymmetrical IgM molecule still remains valid – at least in the case of pentameric IgM.

The pentameric isoform of IgM has been shown in recent studies to be impacted asymmetrically by the physical presence and attachment of the J chain.[28,31] The J chain may ultimately be responsible for the disparity in activation potential between J^+ and J^- pentameric (IgMp) and hexameric IgM (IgMh) that has been observed throughout the literature (where J^+ IgMp < J^- IgMp << IgMh in terms of activation potential), but the physical capabilities of the hexamer will have to be studied in further detail in order to make this distinction clear. The J chain could inhibit pentameric IgM in both a steric and mechanistic sense, as it could 1) hinder the complete ensemble of staple formations and/or 2) sterically clash with incoming C1q globular heads. It's also possible that the symmetry of the hexamer allows it to form staple in either direction, which could infer a greater total activation potential as the probability of

forming staple is effectively doubled, assuming an evenly and sufficiently distanced array of available Ag epitopes.

Regardless of the flexibility and directional biases of the hinge(s) involved, an IgM immune complex formed along a 2D surface will most likely result in the same relative presentation of C1q binding sites, so long as all of the arms are able to sufficiently adopt the staple conformation. C1q can only bind to the side of the IgM Fc platform that faces the solution environment, meaning that a maximum of 5-6 copies of C1q binding sites can be displayed in unison. The other half are presumed to be underneath the Fc platform facing the Ag surface, or occluded outright by the densities of the Cμ2 and Fab domains. The symmetry and flexibility range of IgM in solution, at least for the pentameric isoform, are also unclear

4.4.4 Surface-Unique Changes in the Putative C1q Binding Site and Nearby Cμ3 Residues

Some of the HDX differences we observed may also provide insights pertaining to the exposure of the C1q binding site residues that are widely believed to be located somewhere within the Cμ3 domain. The binding site for C1q on IgG was first presented in 1988 by Duncan & Winter, and is composed of residues Glu318, Lys320 and Lys322 located in the Cγ2 domain. While there does now appear to be slight variations across species and different isoforms of IgG, the homologous residues on IgM are still considered to be the DLPSP motif of residues 432-436 within the Cμ3 domain, which has remained the putative C1q binding site for more than 30 years.[5,39,40] While this patch of surface-exposed residues is indeed located at the periphery of the Fc core platform, and recent work with cryo-EM has provided further low-resolution evidence of C1q globular head association within the same general region of surface bound IgM (refer to Chapter 1, **Figure 1.5.3**), it should be noted that direct proof of a Cμ3 binding site is still surprisingly lacking. In fact, evidence that supports the complete opposite viewpoint was presented in the late 1970s by Bubb & Conradie, where they observed the isolated Cμ3

domain and its related peptide fragments to explicitly not facilitate any significant C1 fixation.[41,42] Instead, they found a high degree of C1 and C1q binding activity within the Cμ4 domain along with a number of its smaller peptide fragments, which was also in close agreement with a series of previous studies performed by Hurst et. al. (1974, 1975, and 1976) along with Johnson & Thames (1976).[43–46]

Conversely, the investigation of the IgM binding interface within the A,B, and C chains that compose the C1q globular heads (refer to Chapter 1, **Figure 1.4.1**) has also proven to be historically difficult, and there appears to be only a partial overlap with IgG-specific regions.[47] While the true identity of the IgM C1q binding sites have still yet to be proven outright, the Cμ3 domain of IgM has nonetheless been thoroughly demonstrated as fundamental to the activation of complement, with the evidence commonly derived through mutational approaches such as alanine screening.[48] Studies that employ mutagenesis to investigate protein:protein interactions have often led to promising results, and in the case of IgG and other proteins that act in a monomeric nature, the approach is often well-suited. However, caution should be taken when considering the results and implications of mutational studies on proteins like IgM that are constructed in a polyvalent manner and act multivalently in fashion.

In 1994, a detailed and systematic mutational study of charged residues in the IgM Fc was carried out by Arya et al., based on earlier work done by Shulman and others that implicated a few charged Cμ3 residues that had severe impacts on complement activity as measured by C1 fixation and cell lysis assays.[39,40,49,50] While they were able to show some evidence that corroborates the involvement of the putative C1q binding site, only a single mutation was found to disrupt complement activity without causing disruption to the oligomerization state of IgM. Of the 16 total residue mutations presented by Arya et. al, 12 were found to disrupt the oligomerization state by 20% or more, with 10 found to severely disrupt both oligomerization and complement activity simultaneously. The remaining mutations had little to no effect, except for one particularly intriguing case in which the mutation greatly

enhanced complement activity (E391A; Cμ3 domain).[49] SDS-PAGE was the only technique utilized to assess the oligomeric status of the IgM mutants, which would not have been able to inform on any potential conformational side effects that might hinder the ability of the molecule to correctly adopt the active staple conformation, which further complicates interpretation of the data.

As with the rest of the IgM mutational studies found throughout the literature, it is often difficult to delineate mutational effects that remove a potential binding site from those that disrupt the mechanism of its formation, which was the primary reasoning behind our decision to leave the IgM physically untreated throughout all of our IgM experiments. This includes choosing to forgo PNGase-F treatment in order to deglycosylate the molecule, as it has been shown recently in the case of IgG to abrogate Fc effector functions.[51] While de-glycosylation would have simplified the extent of data analysis required, the risk of disrupting the ability to form staple and/or occluding the presentation of the C1q binding sites warranted our pursuit of binding valency attenuation via prebinding with monomeric HEL antigen, in a physically non-invasive manner. It should be noted that glycosylation differences between species do indeed exist, and that our IgM material derived from murine sources differs from human IgM in both the total number and locations of its conserved N-linked glycans, as outlined in **Table 4.4.4**.[8,52,53]

The Cμ3 domain of our mouse IgM system contains three conserved N-linked glycans within a concentrated region of less than 40 residues, and while we observed differences unique to surface bound IgM in almost all of them, there was one semi-occupied glycopeptide in particular (N379) that displayed a notably intriguing pattern of uptake by HDX-MS. As confirmed by the recently deposited high-res EM structures of the IgM Fc core, this particular glycopeptide ("YETL") is located at the periphery of the Fc platform, appears to be highly solvent-accessible, and also happens to be located in close 3D proximity (approx. 8-10 Å) to the putative C1q binding site (432-436;DLPSP), as shown in **Figure 4.4.4.** The partially-occupied glycan is also immediately preceded by an acidic/charged residue (E376)

which also appears to be highly conserved (**Figure 4.4.4 D**). While the charged residue appears to be more commonly an aspartic (D) rather than glutamic (E) acid, we find it nonetheless noteworthy that this peripheral Cμ3 residue appears to be located in a region of prominent de-protection/exposure unique to surface bound IgM, as the interaction between C1q and the vast majority of its endogenous ligands is primarily mediated through charge-based pairings.[48,54–57] In Chapter 3 we confirmed ionic strength to be a major solution variable that heavily impacts the binding interaction between C1q and IgM, where we observed the BLI binding magnitudes to be significantly and directly impacted by the salt concentration of the buffer environment (refer to Chapter 3, **Figure 3.3.2.1**).

While we cannot at this time confirm the involvement of these Cμ3 regions in the interaction with C1q, the HDX-MS methods that we have developed here will allow us to directly investigate structural changes involved in the IgM:C1/C1q interaction in future studies. The collection of regions and HDX trends that we have thus far compiled for complement-active IgM will help to guide our future experiments, as well as provide valuable dynamic context to any high-resolution structures of the fully intact, pentavalent form of secreted IgM that are resolved in the future.

4.4.5 Staple-Induced Changes are Likely Translated into the Cμ4 Domains

Finally, in addition to the unique protection we observed in the implicated hinge located within the Cμ2/Cμ3 domain interface, the other most prominent and consistent patterns of surface-unique structural change that we observed were localized to two primary regions of the Cμ4 domain. The first region is composed of a series of residues located toward the Cμ3 domain in the peripheral 'elbow' of the Cμ4 domain ring, where the Cμ4 dimers of neighboring arms come into close proximity with one another. Here we observed a strong degree of de-protection/exposure that started at the solution-exposed exterior of the Cμ4 'elbow' (534 - 553; YTCVVGHEALPHLVTERTVD) and appeared to 'radiate' inward, with lesser magnitudes of HDX exposure also observed within residues that are packed further

into the interior core of the Cμ4 domain (476-483; VKGFSPAD) as illustrated in **Figure 4.4.5 A-C**. While exposure in the outer 'elbow' region was also a notable site of allosteric change that we observed in solution bound IgM, the changes at the inner core of the domain were unique only to the surface bound condition, which could be indicative of a more severe and pronounced structural change that occurs as a consequence of the increased stress imposed on the molecule when it is bound multivalently to a surface.

Because the Cμ3 domains of IgM are not dimerized and only partially stabilized by interdomain interactions, the allosteric changes we observed at the exterior of the Cμ4 'elbow' could also be involved in transient staple formation by solution bound IgM, which would help to explain the mixed reports throughout the literature along with the low level C1 activity that we observed for the solution bound IgM in Chapter 3. The proline in this region (P544) was also one of the residues examined specifically by Arya et. al. in 1994, where mutating the residue to a glycine resulted in the near total abolition (96% loss) of the polyvalent state of IgM, which indicates that this region is at least vital for proper formation of pentameric and hexameric IgM.[49] This peptide region was also reported to potentially contain a C1 binding site by Hurst et. al., where they demonstrated the ability of a peptide fragment containing Cμ4 residues 515-550 to efficiently bind and activate C1 throughout a series of papers published in the mid-1970s.[44,45]

In addition to the Cμ2/Cμ3 junctional hinge region discussed previously, the other most prominent region of structural change that we consistently observed throughout all of our HDX-MS experiments was located in the Cμ4 domain, at the very C-terminal residues of the protein itself (568-576;SDTGGTCY), in a region commonly referred to as the 'μ tailpiece' that is known to mediate the polymerization of pentameric and hexameric IgM (**Figure 4.4.5 D**).[20,58] As proof of this structural principle, IgG mutants that bear the μ tailpiece have recently been demonstrated to readily self-assemble into hexamers that

are also potent activators of complement.[59–62] Here we observed a striking pattern of protection from H/D exchange throughout all of our experiments with surface bound IgM, as well as overlapping deuterium uptake curves between unbound and solution bound IgM conditions, which indicates that the tailpiece region undergoes significant structural change only as a result of multivalent surface binding. This region was specifically implicated in the formation of staple by Feinstein himself in 1986, where he hypothesized the μ-tailpiece regions could function as "*flexible spacers*" in order to compensate for the spatial rearrangement of neighboring Cμ4 domains as they move relative to one another.[34] While we cannot confirm the role of the tailpiece regions without a resolved structure of IgM staple, we do find it plausible that the physical stresses of multivalent surface binding would ultimately be translated radially inwards toward the center of the molecule which, by default, would leave the tailpiece regions to ultimately compensate for the compression of the major heavy domains. The Cμ4 domain of the IgM monomer was recently studied by cryo-EM and FRET-based analysis and was reported to be "*inherently dynamic*", with stability being conferred via polymerization to the other arms along with the attachment of the J chain, which adds further support to the notion that structural rearrangement of the C-terminal tailpiece regions may be necessary to adopt the multivalent surface bound form of staple IgM.[28]

As previously discussed, it's also possible that the Cμ4 domain contains at least a portion of the C1/C1q binding site(s). If staple formation is truly required to expose residues that are otherwise inaccessible, then some of the changes we observed in the Cμ4 could point toward regions that may be involved directly in the C1q binding interaction. We plan to address this along with other questions generated throughout this project in future experiments.

4.5 Conclusion

In this chapter we completed our biological characterization of complement-active IgM by adding structural features that pair directly to the functional phenotypes observed in Chapter 3 that define the multivalent, surface bound "staple" form of IgM. While many structural changes were observed throughout the entire antibody, we discovered a handful of prominent features that likely reflect some of the key mechanistic qualities of staple formation itself, such as evidence pointing directly towards the existence of a functional hinge within the Cμ2/Cμ3 domain junction. We also found promising signs of a previously unreported glycopeptide region that may comprise a portion of the C1q binding site, along with surface-unique changes within the putative DLPSP motif located in the Cμ3 domain. Finally, a number of unique Cμ4 features were observed that could be indicative of the physical stresses induced by multivalent binding to a fixed antigenic surface, that may ultimately be compensated for globally by the μ-tailpiece regions located in the center of Fc core.

While the potential for allosterically-induced complement activity in solution bound IgM was previously unclear from the classical literature, we observed direct evidence of a surprisingly extensive allosteric network that exists within solution bound IgM that may indeed confer a low degree of C1 activity to an otherwise inactive, unbound IgM molecule. Furthermore, while many had previously hypothesized surface-induced physical distortions to be a prerequisite to the complement-active staple conformation of IgM, we did not observe any high magnitude structural changes by HDX-MS that would suggest the breaking of secondary structure, which we believe limits the conformational changes to quaternary/tertiary rearrangements only. While it is clear that antigenic surfaces are required for the maximal activation of C1, we hypothesize the mechanism of staple formation to be governed primarily by steric accessibility, dependent on the multivalent display of antigen, rather than the direct result of allosterism or any major surface-induced physical distortion that would otherwise be naturally unobtainable in solution.

While solution bound IgM may be flexible enough to transiently form staple along any given Ag bound arm, we hypothesize that allosteric changes are unlikely to be enough of a driving force to incur full-scale complement activation *in vivo*. Instead, we predict staple formation to be a localized structural transition (rather than a globally concerted change) that is ultimately dependent on a sufficiently accessible and multivalent display of surface antigen, which is required for the stabilization of a complete *ensemble* of individual staples adopted by each of the five or six binding arms of pentameric or hexameric IgM. Although many questions still remain to be answered, the HDX-MS data generated here along with the methods created throughout this project will continue to provide insights and guidance for future experiments.

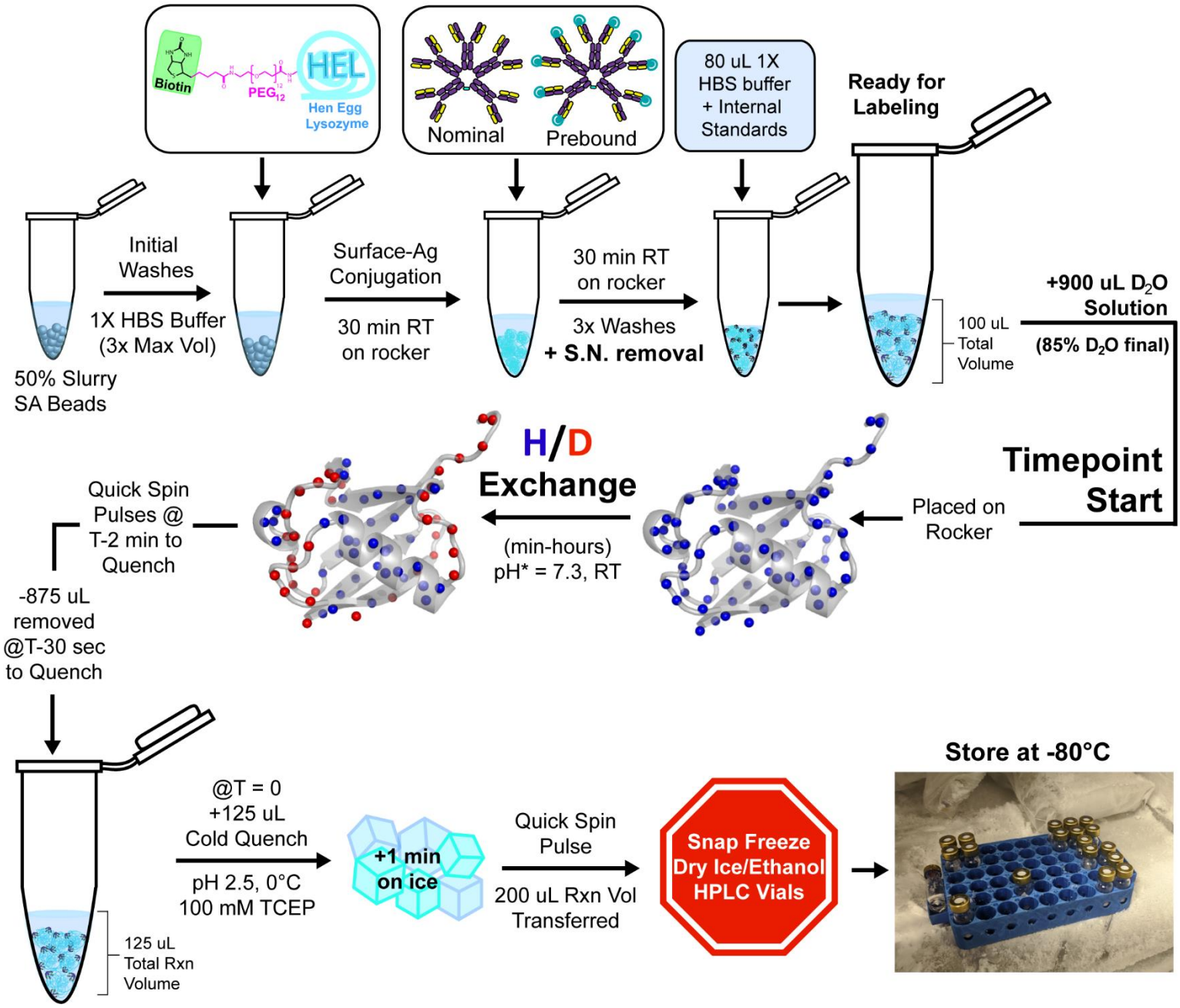


Figure 4.2.1 | Overview of surface HDX method protocol

High-capacity SA agarose beads were first aliquoted into individual sample tubes in order to maintain equal surface Ag distributions across time points and between discrete experiments. Thorough washing steps were performed with fresh 1X HBS buffer after each step in order to remove unconjugated Ag and/or unbound IgM molecules, and care was taken to minimize bead loss within each individual sample tube during final removal of the supernatant. For 1 min time points, the rocker was not used, and spin pulses were done 15 seconds prior to the removal of 875 uL of reaction volume, which occurred 30 seconds prior to quenching. A complete description of this surface-HDX protocol is outlined in methods section 4.2.3.

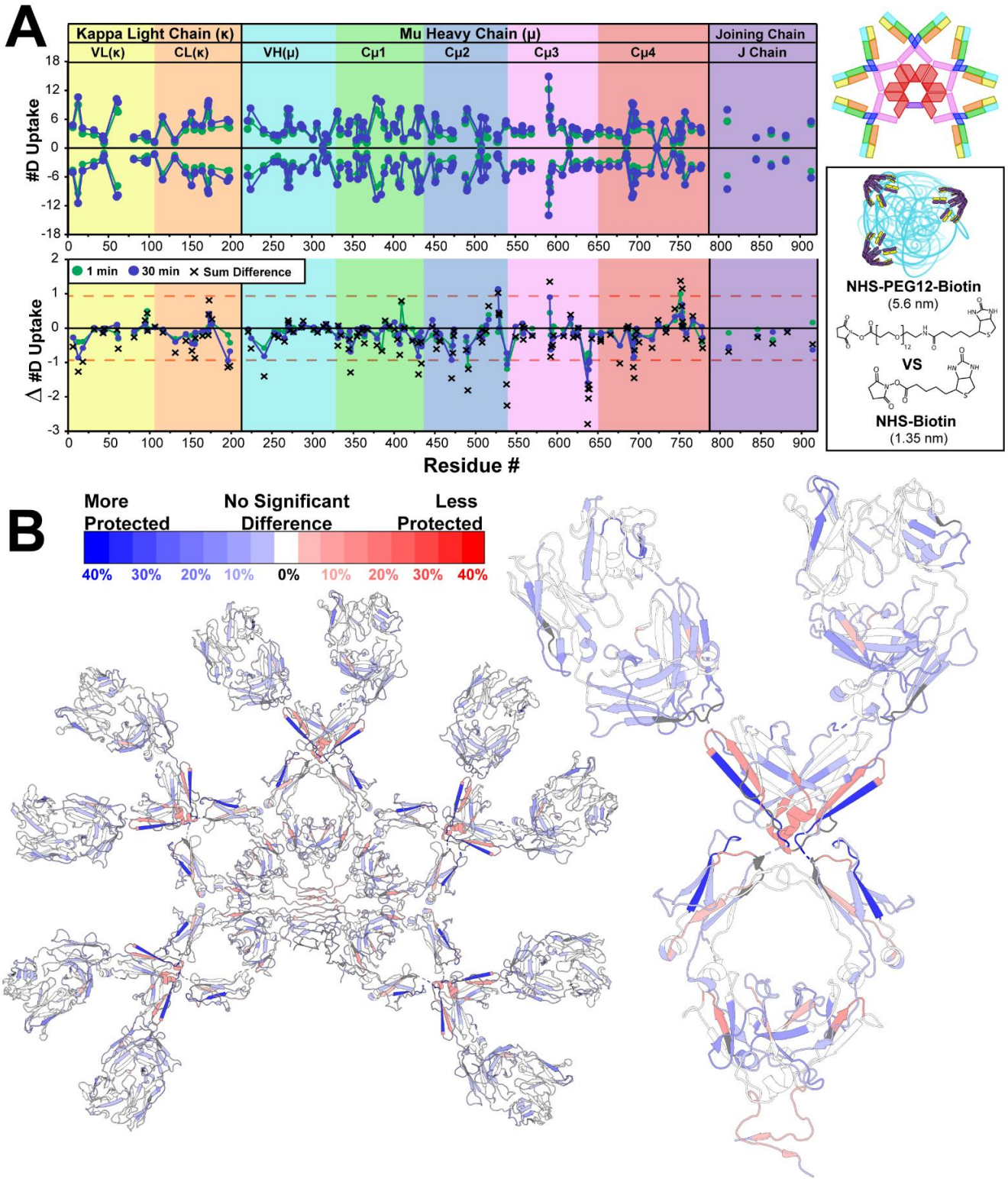


Figure 4.3.1.1 | Structural effects of surface Ag linker extension on surface bound IgM

A) Butterfly plot of extended linker vs short linker H/D uptake, with the uptake differential (#D) plotted below (PEG_{12}-btn-NHS vs biotin-NHS). Domains are color-coded to match the structural legend on the right. **B)** Differential heat map of uptake, calculated by total deuteration percentage of the theoretical maximum. Because this comparison is between surface bound IgM only, the differential analysis predominantly reveals staple-unique regions with trends that become more pronounced in magnitude as a result of optimized surface binding and promotion of IgM staple formation using the PEG_{12} linker. The largest differences were observed in protected regions located near the Cμ2/Cμ3 domain junction.

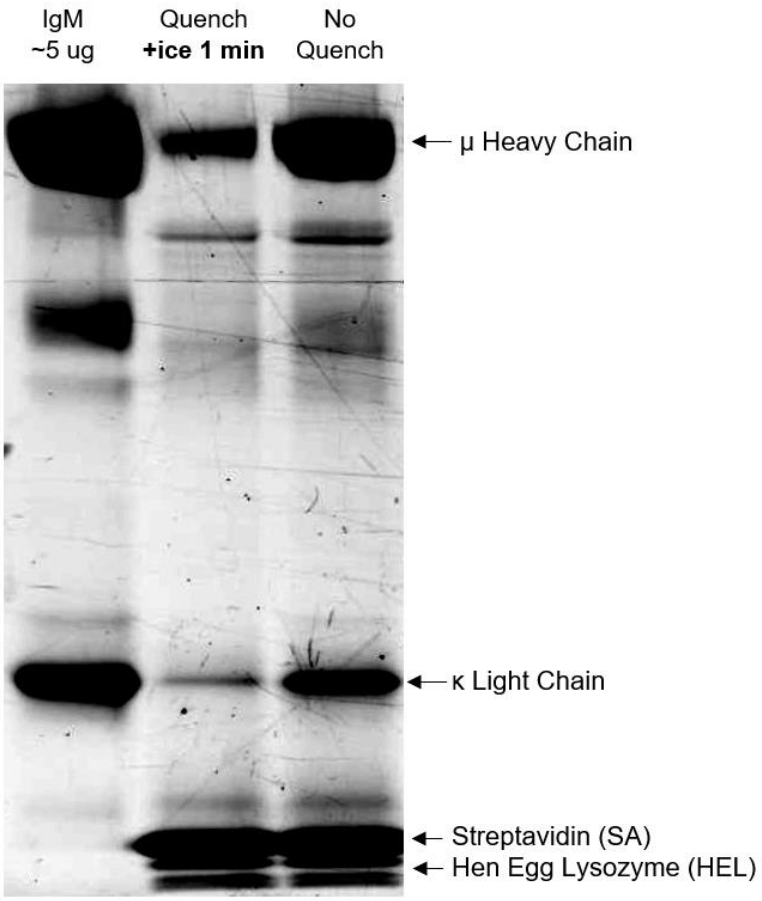


Figure 4.3.1.2 | Troubleshooting dissociation with nominal (PEG_{12}-HEL) surface HDX samples

With the improved binding strength afforded by the nominal PEG_{12}-HEL surface Ag linker, dissociation from the SA beads became an issue that could not be fully resolved through adjustments to the composition of the quench solution. HDX samples were therefore given an extra minute on ice, after the addition of quenching solution, in order to increase the resulting peptide signal during MS analysis. Although 1 minute was not enough to fully dissociate the IgM from the SA beads, as shown in the above reducing SDS-PAGE analysis, it was enough time to afford adequate MS signal without incurring unnecessary amounts of back-exchange of the deuterium label, which was a balance that had to be struck during method development. Although this issue affected only the nominal (PEG_{12}-HEL) surface bound condition, all HDX-MS samples were given the same additional post-quench minute on ice in order to maintain consistency for fair differential analyses.

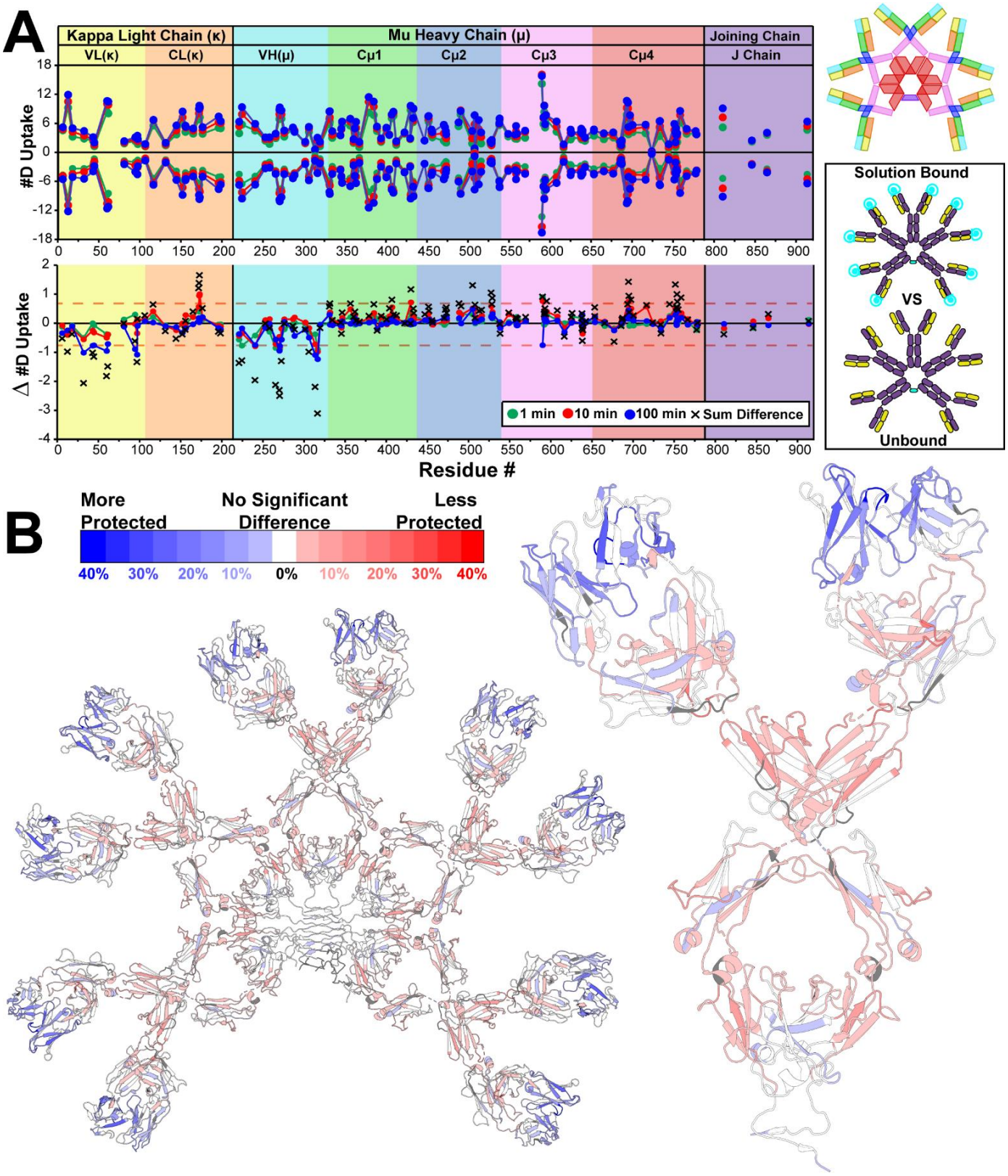


Figure 4.3.2.1 | Allosteric changes of solution bound IgM (3-State HDX)

A) Butterfly (top) and residual plot (bottom) with the results of the comparison of solution bound vs unbound IgM. **B)** An exchange heatmap of the %D uptake differential for each peptide, where there is a clear pattern of protection localized to the Fab domains, which is expected for solution bound IgM. There are, however, smaller changes that also extend deep into the Fc core domains.

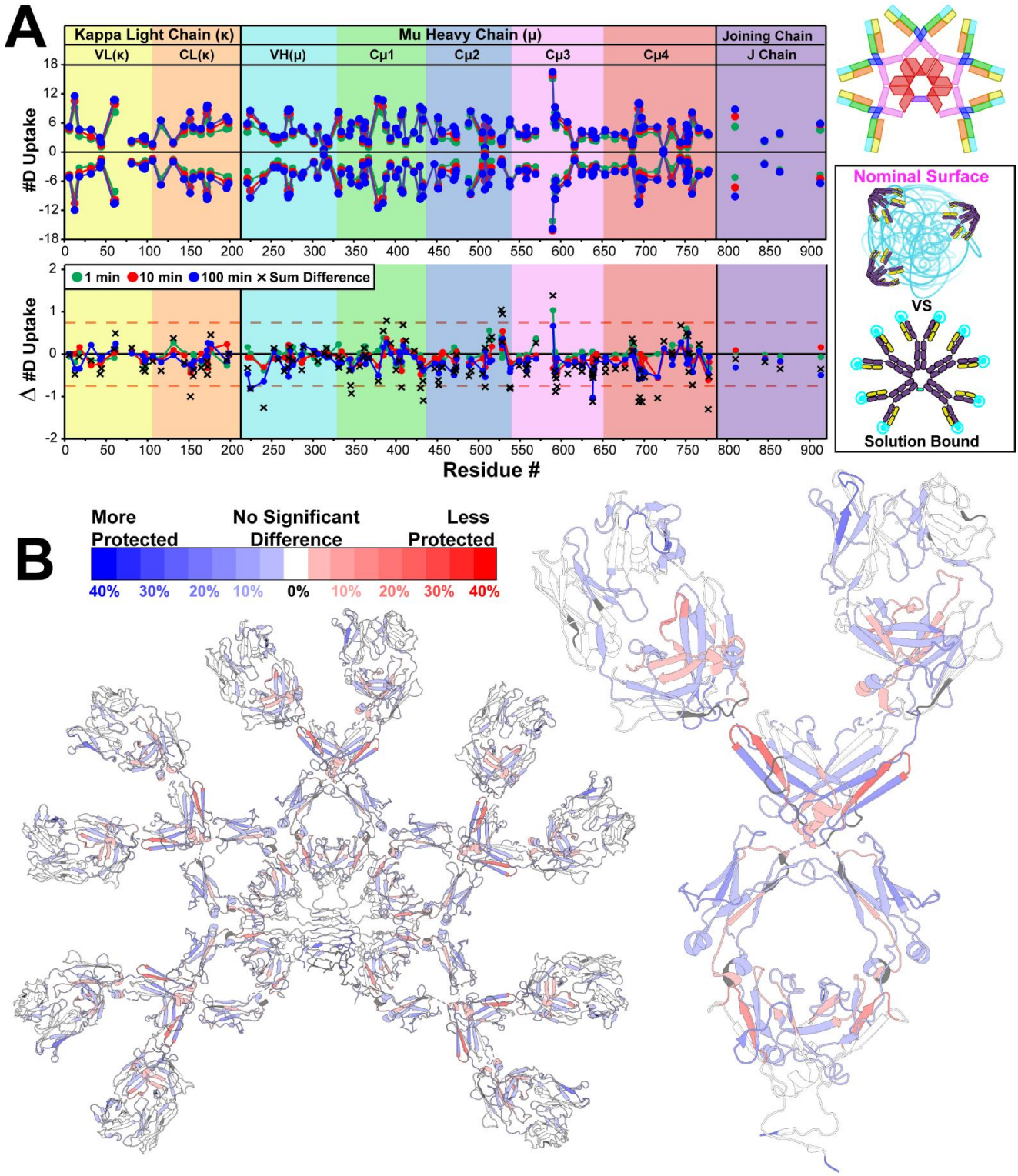


Figure 4.3.2.2 | Nominal surface-unique changes compared to solution bound IgM (3-State HDX)

A) Butterfly (top) and residual plot (bottom) with the results of the comparison between the nominal surface bound condition and solution bound IgM. **B)** An exchange heatmap of the %D uptake differential for each peptide, in which protection at the Fabs is no longer observed due to both conditions being bound to antigen to an equal extent. However, surface-unique changes were found throughout the Fc, with a significant dregee of protection again observed near the Cμ2/Cμ3 domain junction, as well as a prominent degree of protection at the C-termal 'μ-tailpiece' region.

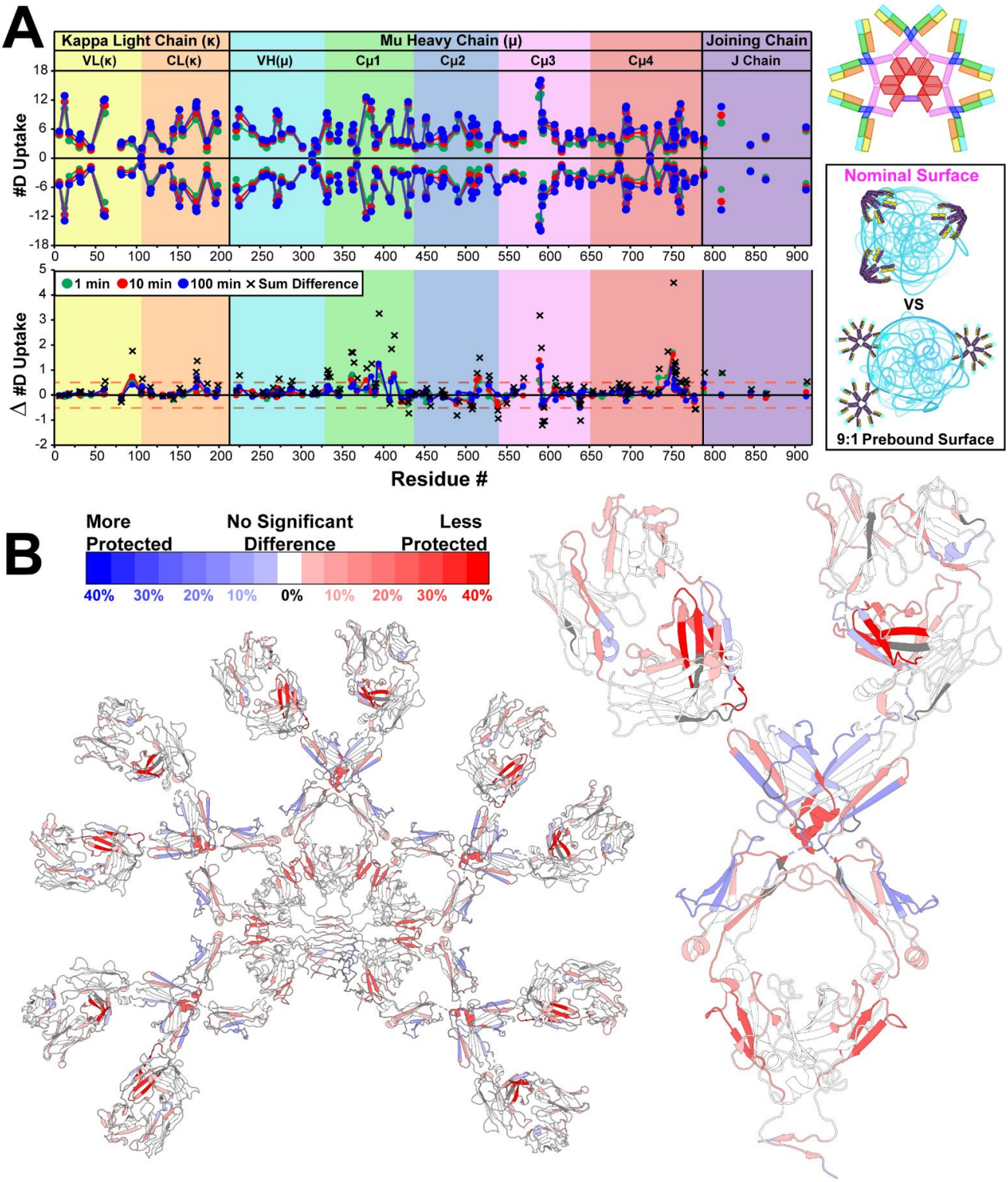


Figure 4.3.3 | Nominal surface-unique changes compared to 9:1 prebound surface IgM

A) Butterfly (top) and residual plot (bottom) with the results of the comparison between nominal surface bound IgM and a highly prebound (~9:1) surface condition that was used as an analytical 'surrogate' to approximate solution bound IgM. **B)** An exchange heatmap of the %D uptake differential for each peptide, in which protection at the Fabs is also not observed due to both conditions being bound to antigen to an equal extent. However, significant surface-unique changes were found throughout the Fc that corroborated key trends observed throughout past experiments, such the protection near the Cμ2/Cμ3 domain junction, as well as a prominent degree of protection at the C-termal 'μ-tailpiece' region.

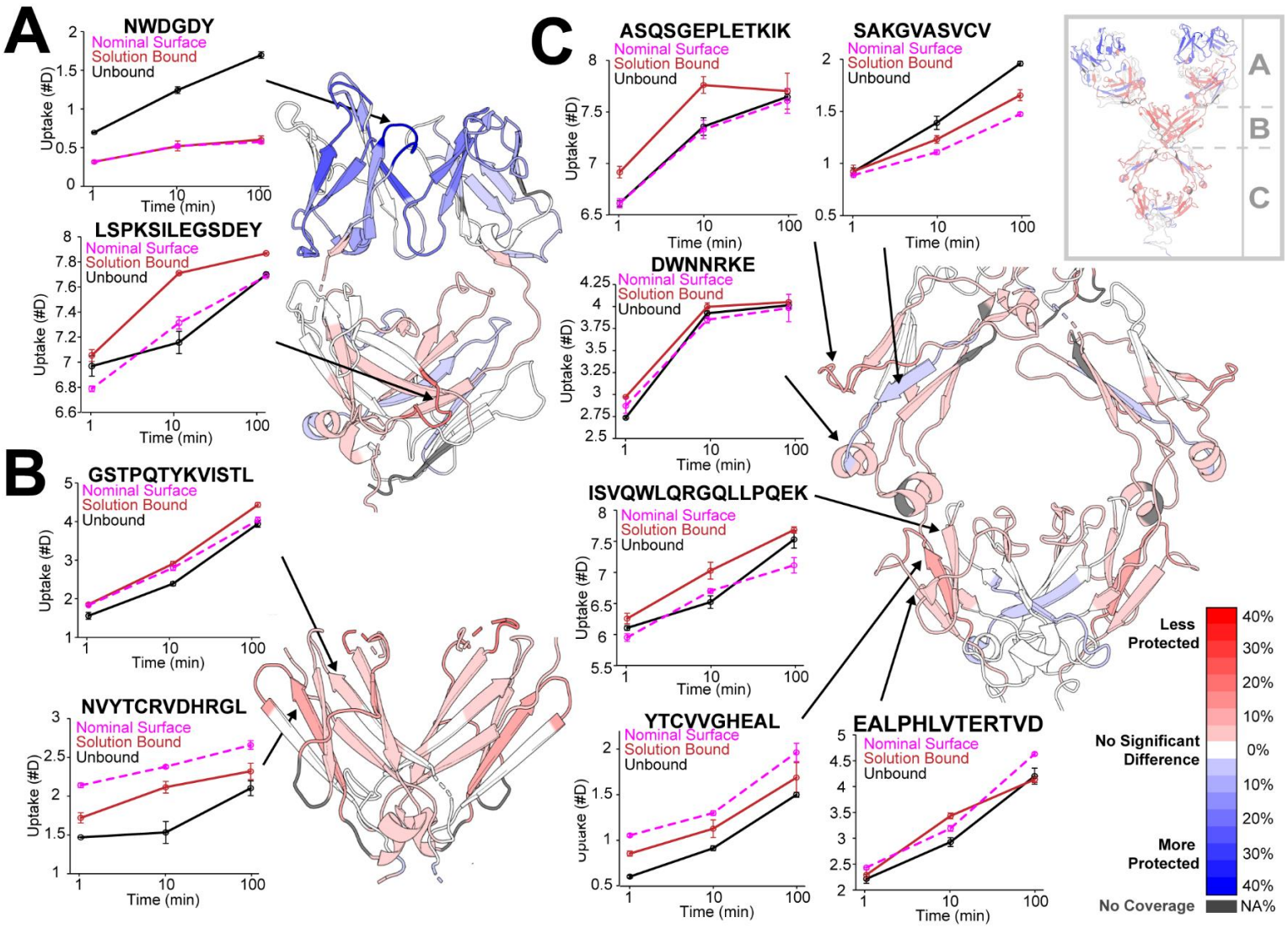


Figure 4.4.1 | The extensive allosteric network of solution bound IgM

Select uptake curves are shown for the 3-state HDX differential analysis of allosteric changes in solution bound IgM (solution bound vs unbound), throughout the **A)** Fab domains **B)** Cμ2 domain, and **C)** Cμ3 and Cμ4 'Fc core' domains. While some locations of allosteric change are completely unique to the solution bound state, there are some regions that overlap and appear similar to those observed in the nominal surface bound condition, which is shown as a dashed line (magenta) for reference. Some of these overlapping regions may contribute toward the C1 activity that we observed previously in the solution bound form of IgM (Chapter 3).

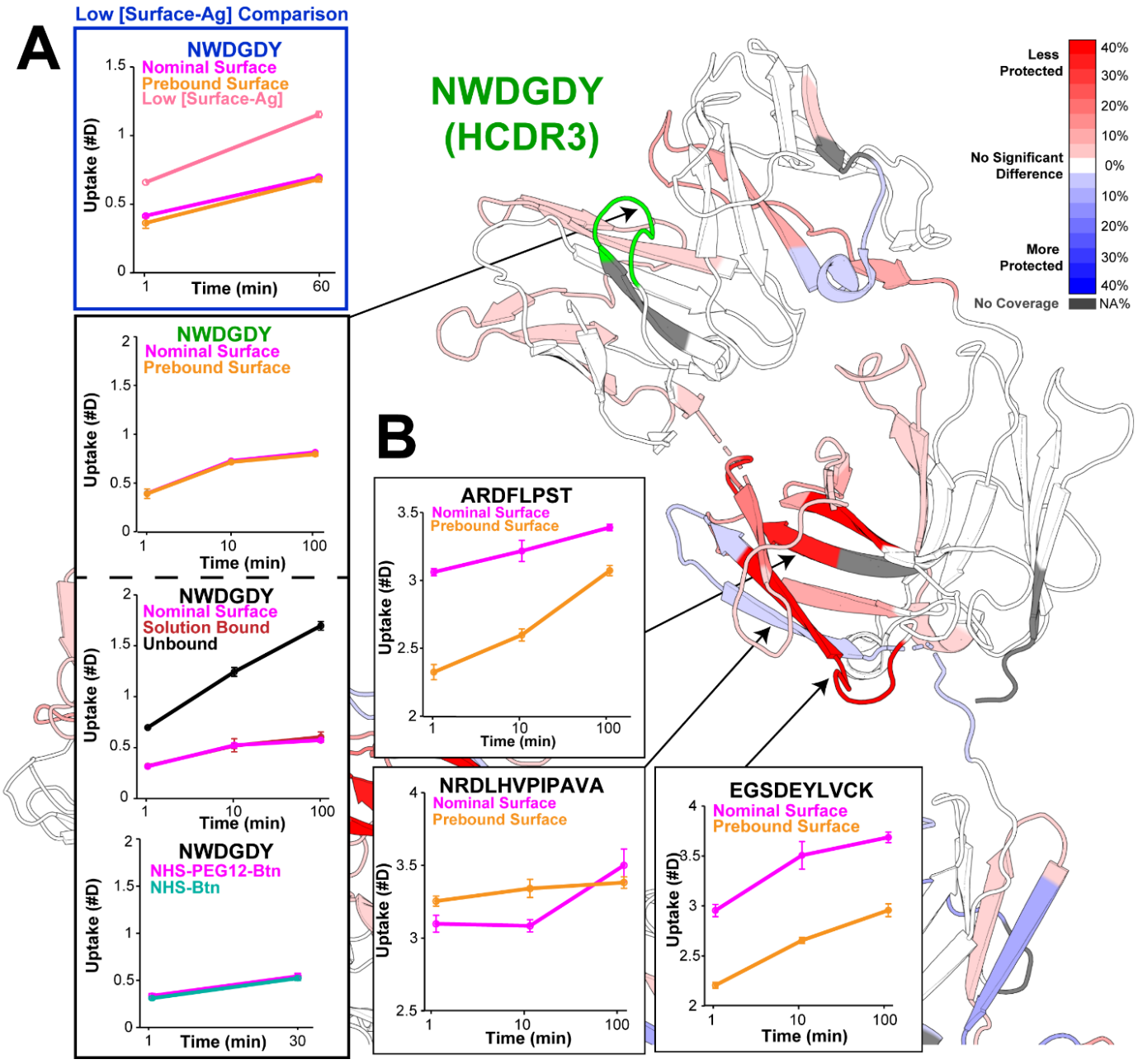


Figure 4.4.2 | Surface-unique Fab changes observed by HDX-MS

A) Monitoring of HDX at the HCDR3 peptide (NWDGDY) provides a direct readout of the extent to which IgM Fabs are bound within various Ag bound conditions/conformational states, which in the case of surface bound IgM is also largely dependent on the extent to which the antigen displays are conjugated with biotinylated HEL, as low densities of surface antigen will leave the majority of the Fabs unbound due to the lack of available surface antigen (top; blue). All three of the main HDX experiments showed equivalent binding between the Ag bound IgM conditions that were compared (bottom three panels) throughout this chapter. **B)** Surface-unique changes found at the N-terminus of the Cμ1 domain (near the Cμ1/Cμ2 junction) are likely reflective of the increased strain experienced at the Fabs due to spatial rearrangements required to accommodate multivalent surface binding of all ten available Fab domains. Visualized HDX data is mapped to the structure based upon the Nominal vs Prebound Surface HDX differential comparison (see section 4.3.3 for preparation details).

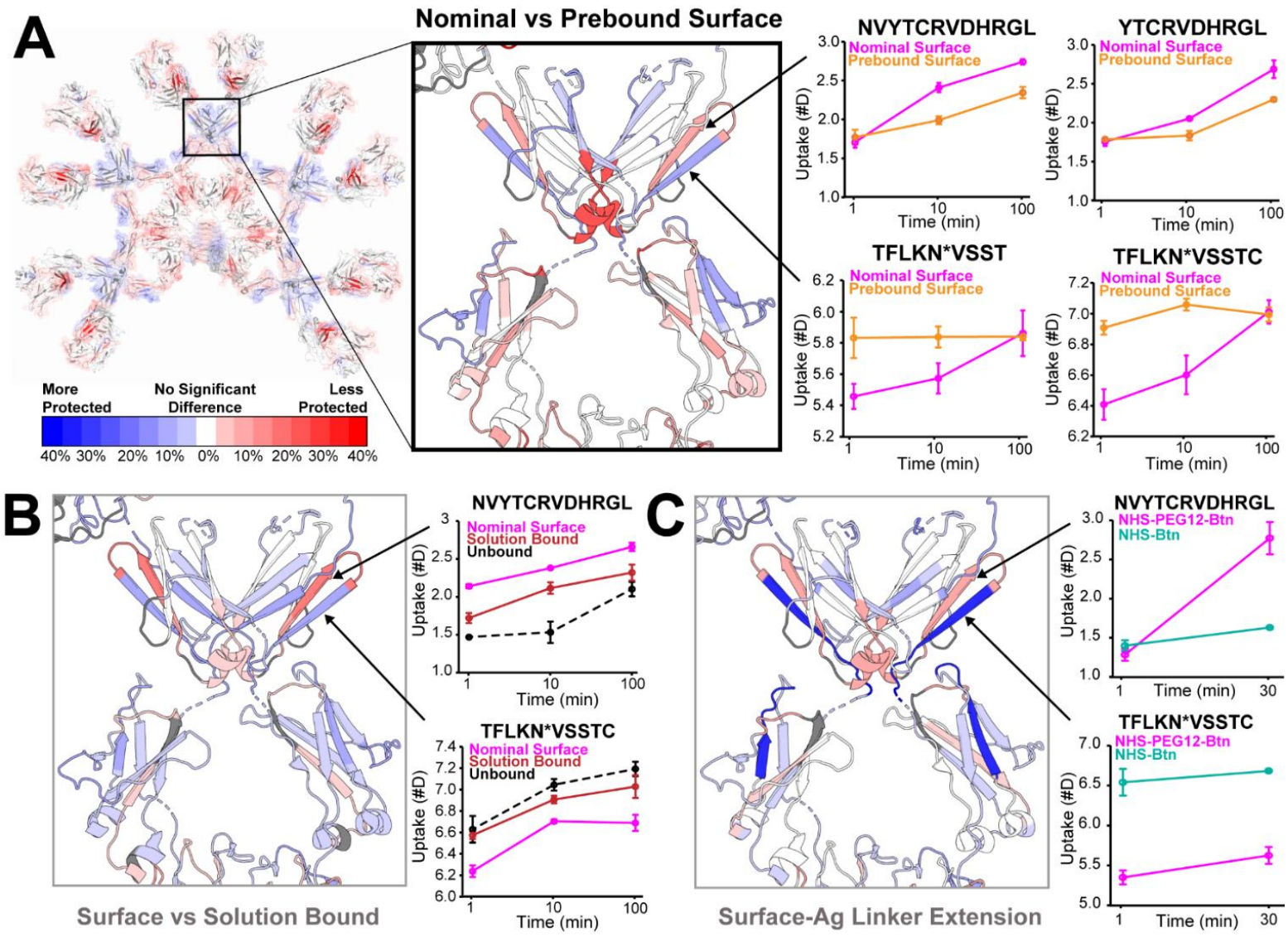


Figure 4.4.3.1 | IgM staple formation is mediated by hinge function of the Cμ2 domain

A prominent degree of protection from H/D exchange was consistently observed in the nominal (PEG_{12}-HEL) surface bound condition throughout differential comparisons with **A)** highly prebound (9:1) surface IgM, **B)** solution bound IgM, and **C)** IgM bound to surface Ag prepared with the short/unoptimized biotin linker (5.6 vs 1.35 nm). In all three cases, the junctional glycopeptide (TFLKN*VSST/C) was also directly preceded by a region of deprotection/exposure (NV/YTCRVDHRGL). While the exact physical motion is difficult to interpret without an intact structure, this clear and consistent exchange pattern is likely reflective of a mechanistic hinge that allows IgM to bind a surface via adoption of local staple formations at each participating binding arm.

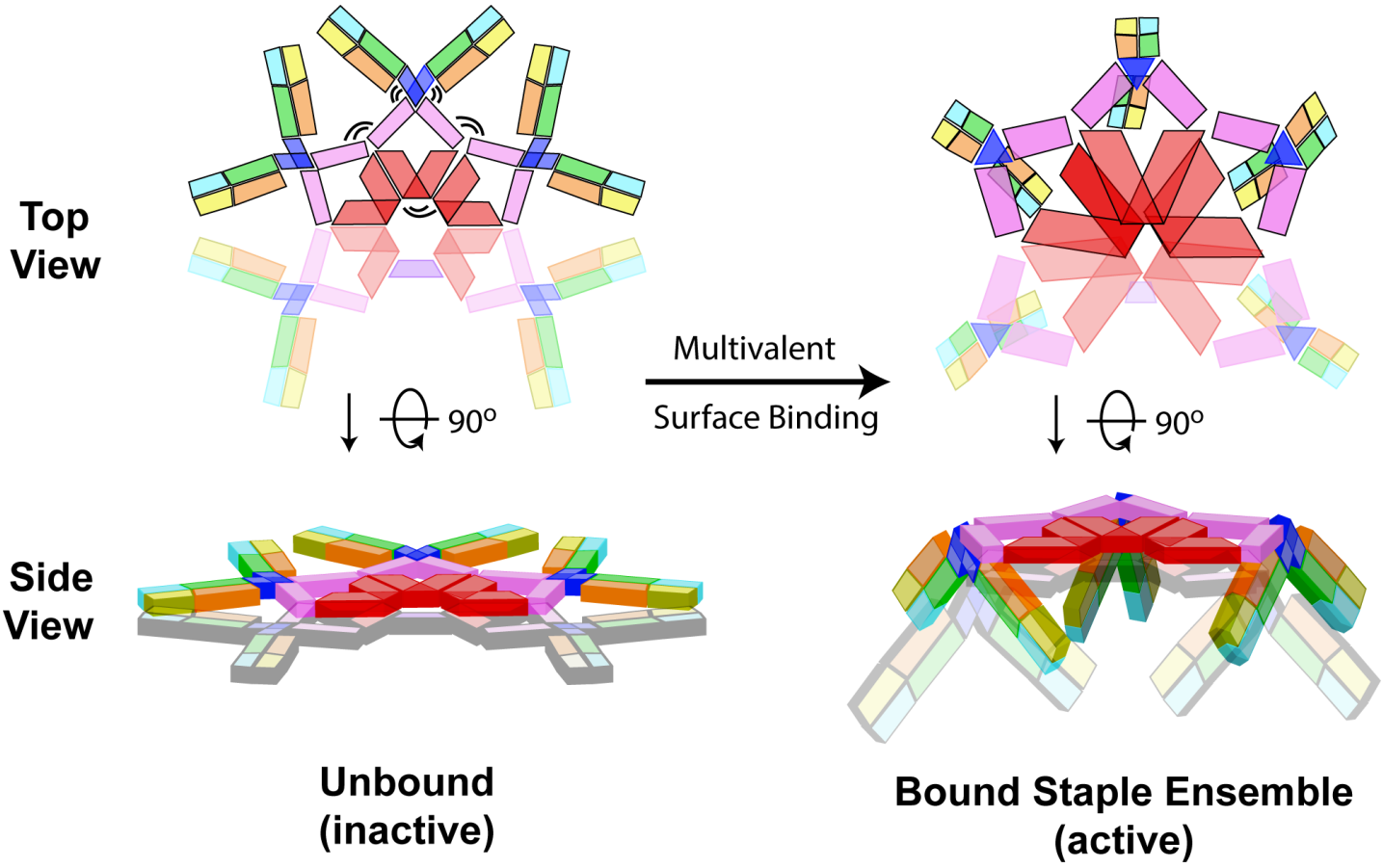


Figure 4.4.3.2 | Multivalent surface binding likely leads to compression of Fc core domains

Based on recent Cryo-EM evidence, along with our HDX-MS observations of long-lasting protection within the Cμ2/Cμ3 junctional region, we hypothesize that a mechanistic hinge motif exists within the junctional area that provides a significant degree of flexibility to the individual binding arms of IgM. Folding motion at the Cμ2 hinge allows bending of the arms down and away from the plane of the Fc core comprised of the Cμ3, Cμ4, and, in the case of pentameric IgM, the J chain. It is currently unclear if the hinge is able to bend in both vertical directions (up/down) relative to the Fc core plane, or if the asymmetry imposed by the J chain hinders its symmetric range of motion. The IgM hexamer may be mechanically symmetrical as it lacks a J chain, which could potentially enable it to form staple in either direction, which would likely add to its enhanced C1 activation potency. As the staple ensemble is formed along the targeted Ag surface, inward compression and further compensatory motion of the core domains is also likely to occur, in addition to further Fab rotation which may be necessary to sterically accommodate all of the individual Fab lobes at the surface.

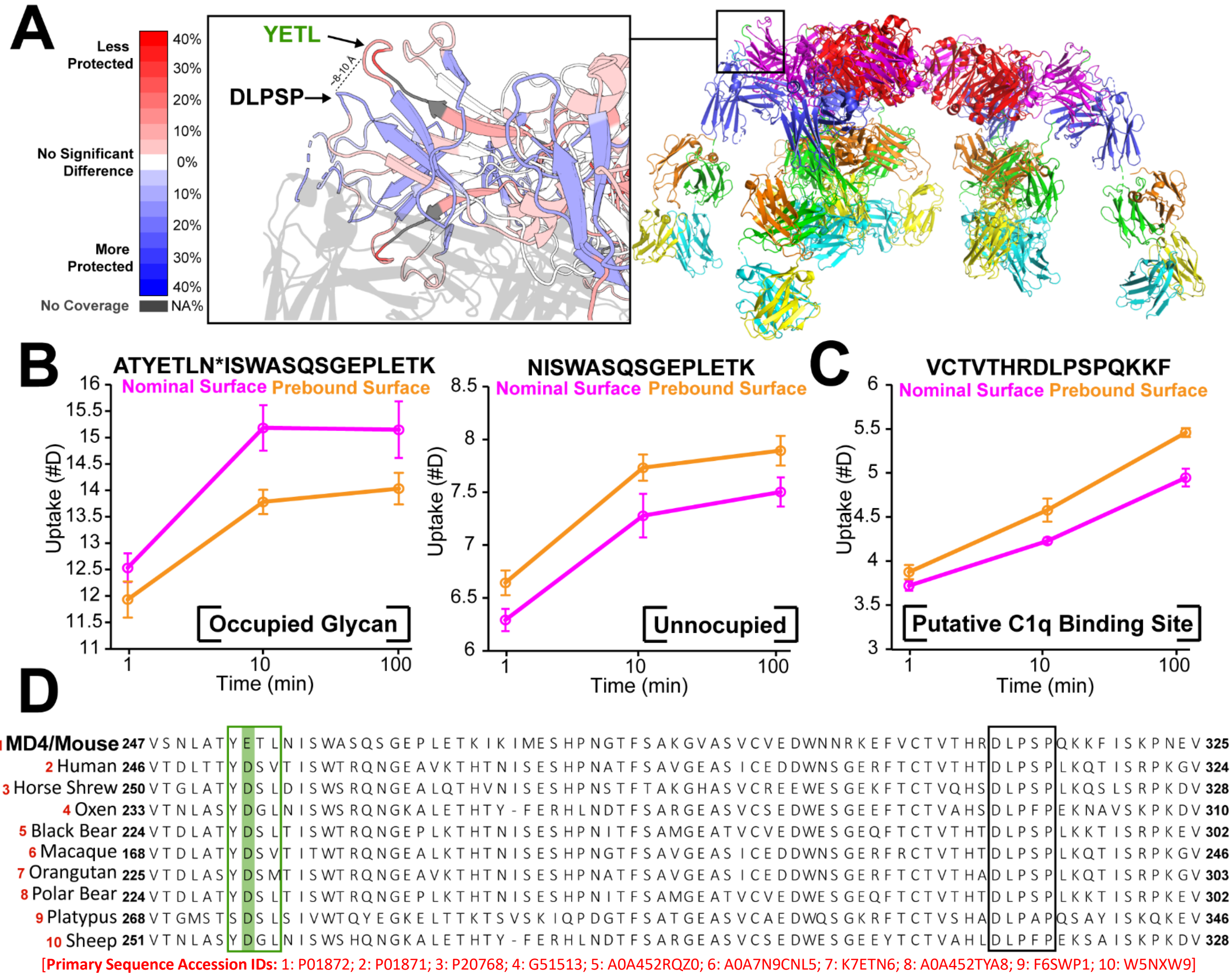


Figure 4.4.4 | Surface-unique Cμ3 domain changes in the putative C1q binding site and nearby residues

A) Zoomed in view of the Cμ3 periphery where the semi-occupied glycopeptide 'YETL' and the nearby (~8-10 Å apart) putative C1q binding site DLPSP regions were both observed to undergo surface-unique structural changes by HDX-MS. The Cμ2 domain (grey/transparent) is predicted to bend down and away from the Fc platform in order to expose the C1q binding sites, likely mediated by a hinge located at the Cμ3/Cμ2 junction (refer to Fig. 4.4.3.1). **B)** Uptake plots of the occupied (left) and unoccupied (right) versions of the YETL glycopeptide, with an observed trend flip from deprotection/exposure in the occupied form to significant protection in the unoccupied version of the peptide, which was consistently observed throughout all HDX-MS experiments with nominal surface bound IgM. The occupied form of the shorter peptide could not be resolved due to impaired protease digestion near the glycan of the occupied form. **C)** Uptake plot of a peptide containing the putative C1q binding site (DLPSP), which was consistently found to be significantly protected throughout all nominal surface bound IgM experiments. **D)** Alignment of related Cμ3 sequences between mouse IgM MD4 and other species. The charged/acidic residue in YETL is highly conserved across species (D more common than E), which implicates its potential involvement in binding to C1q, which is known to be predominantly mediated by charge-based pairings; putative DLPSP binding site motif is also shown (black box). For more details on structural modeling of the multivalent bound form of IgM, refer to methods section 4.2.6.

Species	Residue #	Position	Glycan(s) observed
Mouse (MD4)	N169	Cμ1 Domain	HexNaAc(4)Hex(5)Fuc(1)NeuGc(2)
	N332	Cμ2 Domain	HexNaAc(4)Hex(5)Fuc(1)NeuGc(2)
	N364	Cμ3 Domain	HexNAc(2)Hex(6)
			HexNAc(2)Hex(5)
	N379	Cμ3 Domain	Unoccupied
			HexNaAc(4)Hex(5)Fuc(1)NeuGc(2)
			HexNaAc(4)Hex(5)Fuc(1)NeuGc(1)
	N402	Cμ3 Domain	HexNAc(2)Hex(5)
			HexNAc(3)Hex(3)Fuc(1)
	N563	Cμ4 Domain	HexNAc(2)Hex(6)
Human	N169	Cμ1 Domain	Complex [8,47]
	N332	Cμ2 Domain	Complex [8,48]
	N395	Cμ3 Domain	Complex [8,48]
	N402	Cμ3 Domain	Simple (high mannose) [47,48]
	N563	Cμ4 Domain	Simple (high mannose) [47,48]

Table 4.4.4 | Conserved N-linked glycosylation of mouse (MD4) and human IgM

Occupied glycopeptides and glycan IDs observed in HDX-MS experiments using IgM MD4 (mouse) are provided, along with the locations and types of conserved N-linked glycans reported in the literature for human IgM, for comparison. A high degree of homology was found between the mouse and human glycoforms located at N169 and N332, with complex glycans found in both species. Heterogeneity exists within the Cμ3 domain however, with the high mannose N364 and complex N379 (YETL**N***IS) glycans both found to be unique to mouse IgM only. N402 is a shared location between the two species, and we identified both a high mannose and complex form at that location, which is mostly consistent with the high mannose glycans reported in the literature for human IgM. We also observed the N-linked glycan at N563 in the Cμ4 domain to be high mannose, which is homologous in both location and glycan type to human IgM and consistent with prior studies. Refer to Figs 1.2.1 and 1.2.2 (Chapter 1) for structural/sequence locations of human N-linked glycans (all sequence numbering and modeling is based on the canonical human IgM μ-heavy chain sequence; P01871).

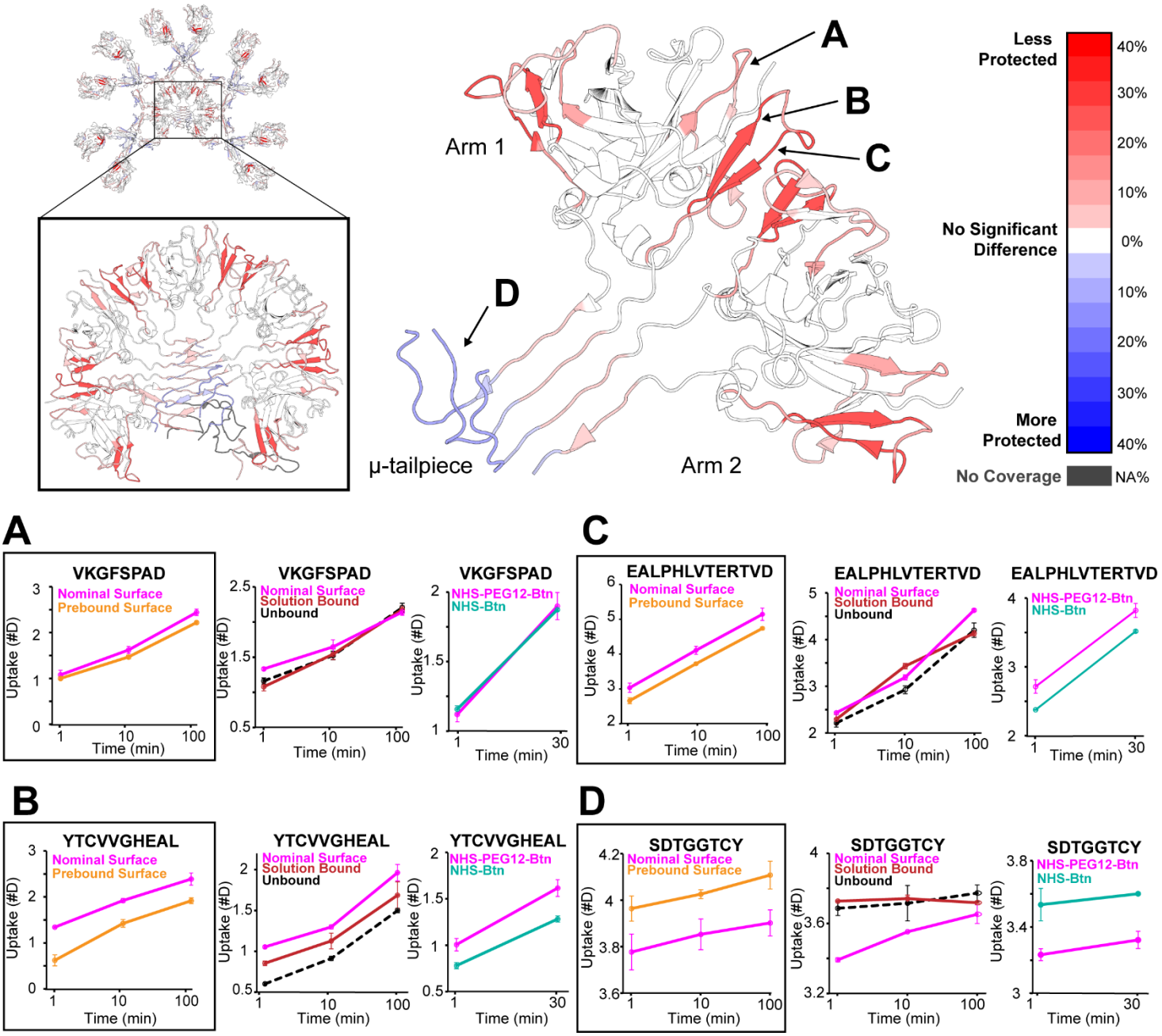


Figure 4.4.5 | Surface-unique Cμ4 domain changes observed by HDX-MS

At least two prominent regions of unique HDX changes were found within the Cμ4 domain of the nominal surface bound IgM condition. The first was located at the 'elbow' region near the N-terminal side of the domain, which is near the Cμ3 domain junction as well as neighboring Cμ4' domains of adjacent arms, where significant deprotection/exposure was observed at the outer regions **(B/C)** that appeared to extend deeper into the core of the Cμ4 dimer **(A)**. While exposure at the outer elbow region **(B/C)** was also observed in solution bound IgM (refer to Fig 4.4.1), the magnitudes of exposure here were found to be larger, and the changes within the interior of the domain **(A)** were unique only to the nominal surface bound condition. **D)** At the very C-terminus of the IgM heavy chain, in a region commonly known as the 'μ-tailpiece', a prominent degree of protection was consistently observed for the nominal surface bound condition, which may be indicative of the increased physical strain imposed by multivalent surface binding, for which the tailpieces are thought to structurally compensate for globally.

Supplementary Table S4.1 | Surface Ag Linker Comparison HDX - Experiment Statistics

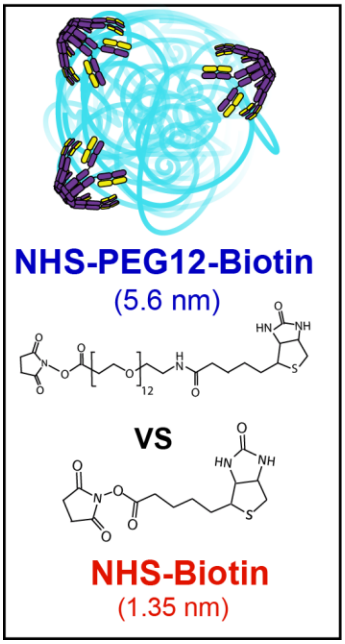
Data Set	NHS-PEG12-Biotin (Extended)†	NHS-Biotin (Short)‡
HDX reaction details	pH 7.3, 22°C, 85% D2O	
HDX time course	1 min, 30 min	
HDX controls	Undeuterated	
Back-exchange (mean) angiotensin II peptide internal standard	30.81%	29.65%
Number of peptides	**145 peptides**	
Sequence coverage	**89.5% Total Sequence Coverage (LC ~95%, HC ~98%, JC ~44%)**	
Average peptide length / Redundancy	12.8 +/- 5.2 residues	
	redundancy: 2.0 peptides/residue	
Replicates (biological or technical)	Triplicate biological replicates	Duplicate biological replicates
Repeatability	**0.100 (avg standard deviation)**	**0.074 (avg standard deviation)**
Significant differences in HDX	**0.4885 D (95% CI) ***	

† Condition is labeled as "Surface NEW" in generated uptake plots

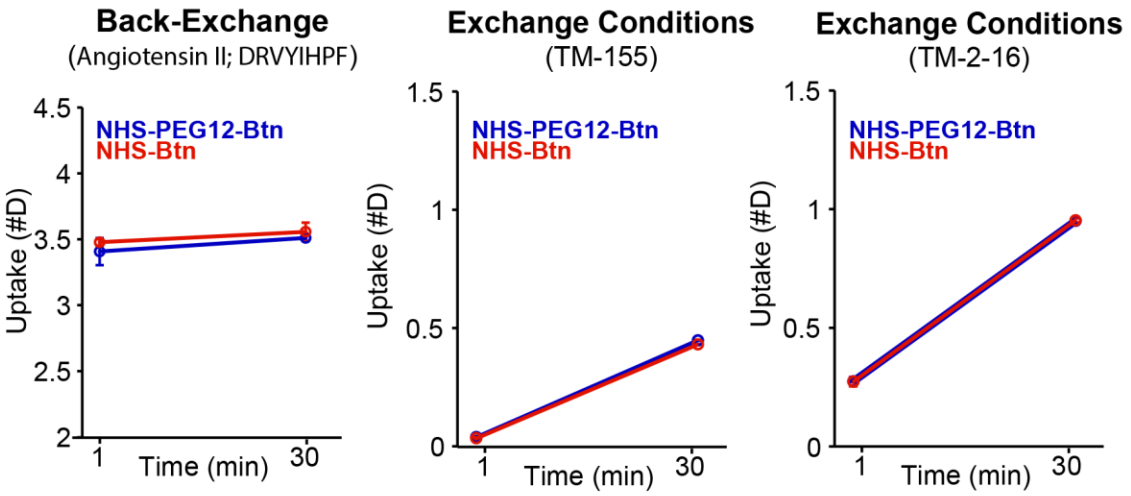
‡ Condition is labeled as "Surface OLD" in generated uptake plots

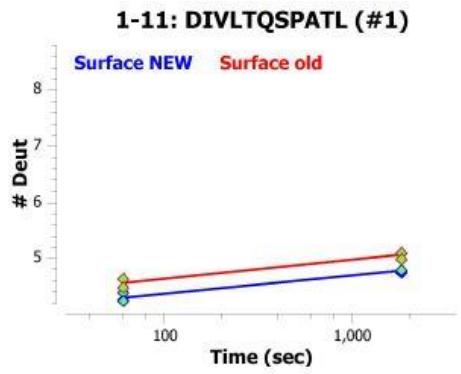
*This calculation is done by HDExaminer and does not include occupied YETL glycopep

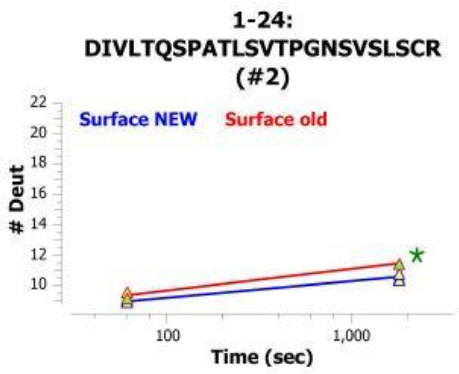
(Surface Only)
Sample Conditions

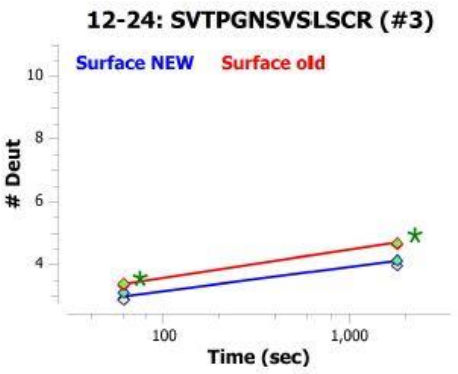


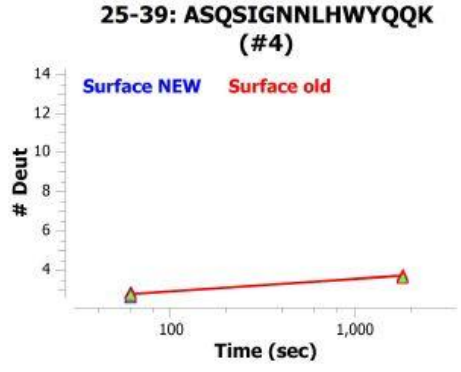
Internal Standards

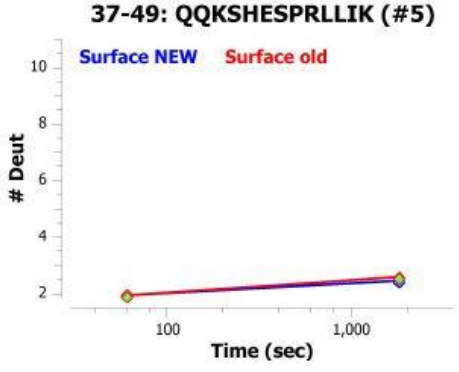


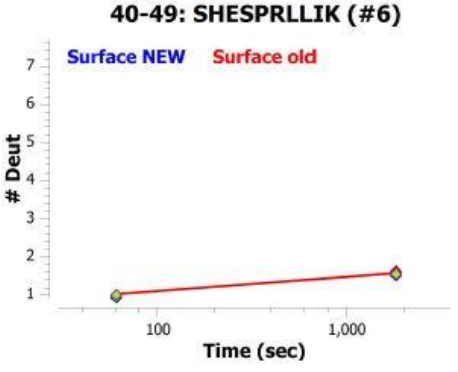

1-11: DIVLTQSPATL (#1)
Surface NEW
Surface old
Deut
Time (sec)

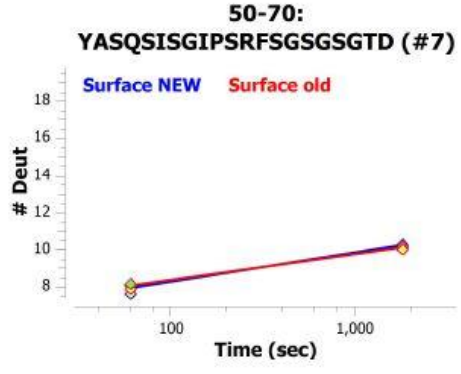

1-24: DIVLTQSPATLSVTPGNSVSLSCR (#2)
Surface NEW
Surface old
Deut
Time (sec)

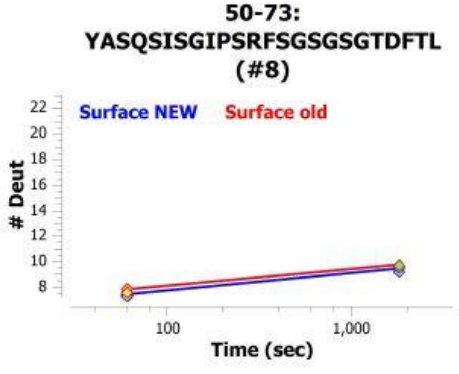

12-24: SVTPGNSVSLSCR (#3)
Surface NEW
Surface old
Deut
Time (sec)

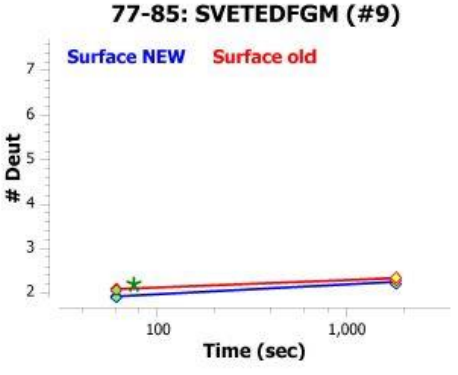

25-39: ASQSIGNNLHWYQQK (#4)
Surface NEW
Surface old
Deut
Time (sec)

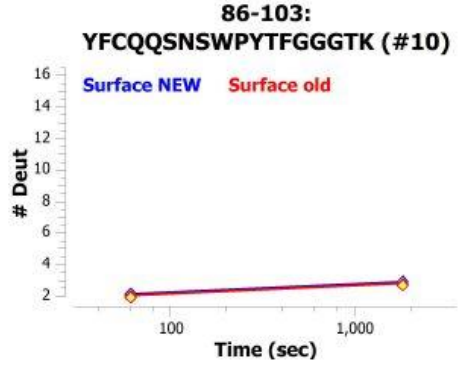

37-49: QQKSHESPRLLIK (#5)
Surface NEW
Surface old
Deut
Time (sec)

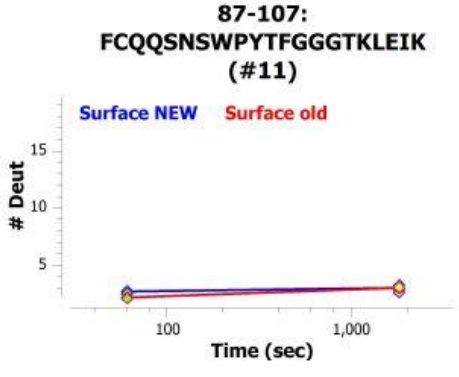

40-49: SHESPRLLIK (#6)
Surface NEW
Surface old
Deut
Time (sec)

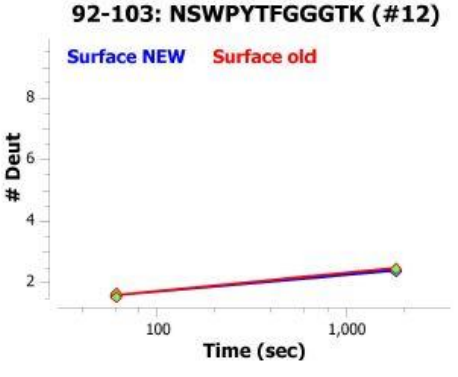

50-70: YASQSISGIPSRFSGSGSGTD (#7)
Surface NEW
Surface old
Deut
Time (sec)

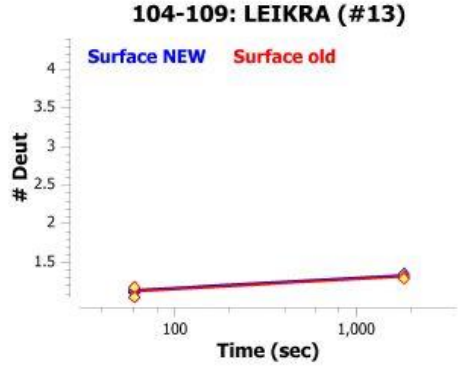

50-73: YASQSISGIPSRFSGSGSGTDFTL (#8)
Surface NEW
Surface old
Deut
Time (sec)

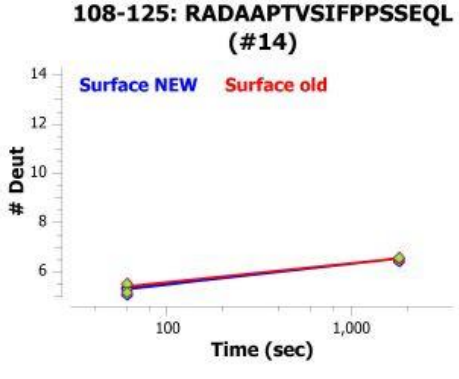

77-85: SVETEDFGM (#9)
Surface NEW
Surface old
Deut
Time (sec)

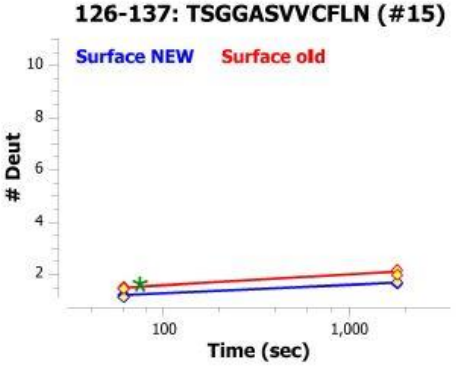

86-103: YFCQQSNSWPYTFGGGTK (#10)
Surface NEW
Surface old
Deut
Time (sec)

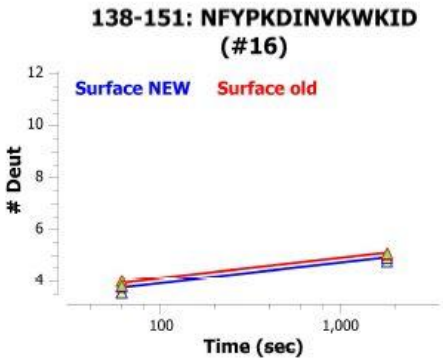

87-107: FCQQSNSWPYTFGGGTKLEIK (#11)
Surface NEW
Surface old
Deut
Time (sec)

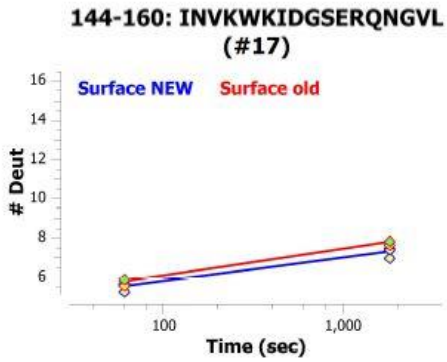

92-103: NSWPYTFGGGTK (#12)
Surface NEW
Surface old
Deut
Time (sec)

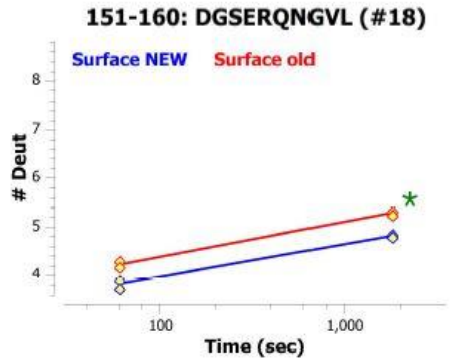

104-109: LEIKRA (#13)
Surface NEW
Surface old
Deut
Time (sec)

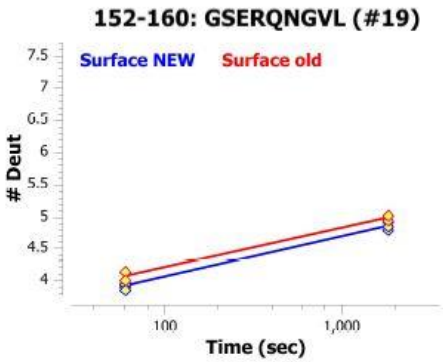

108-125: RADAAPTVSIFPPSSEQL (#14)
Surface NEW
Surface old
Deut
Time (sec)

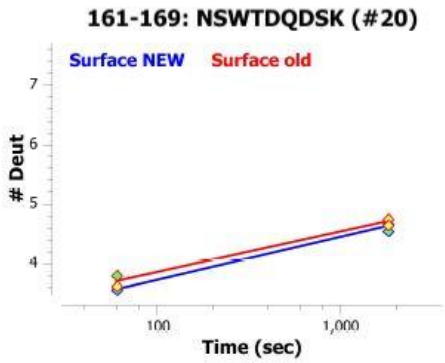

126-137: TSGGASVVCFLN (#15)
Surface NEW
Surface old
Deut
Time (sec)

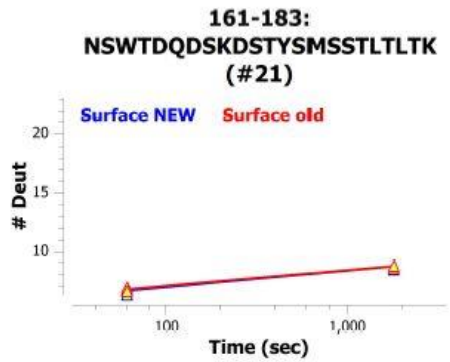

138-151: NFYPKDINVKWKID (#16)
Surface NEW
Surface old
Deut
Time (sec)

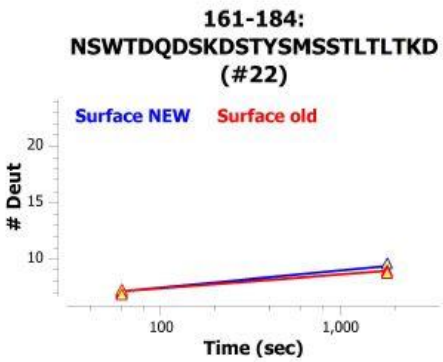

144-160: INVKWKIDGSERQNGVL (#17)
Surface NEW
Surface old
Deut
Time (sec)

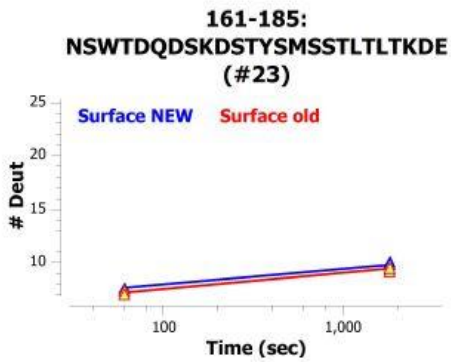

151-160: DGSERQNGVL (#18)
Surface NEW
Surface old
Deut
Time (sec)

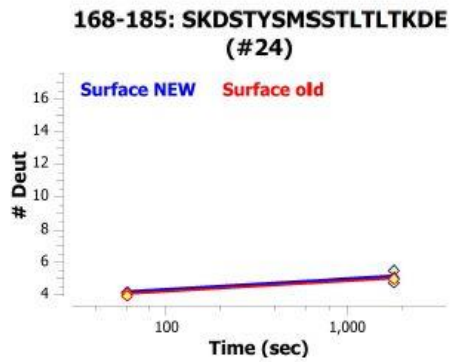

152-160: GSERQNGVL (#19)
Surface NEW
Surface old
Deut
Time (sec)

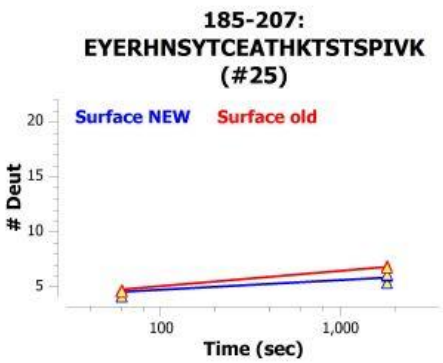

161-169: NSWTDQDSK (#20)
Surface NEW
Surface old
Deut
Time (sec)

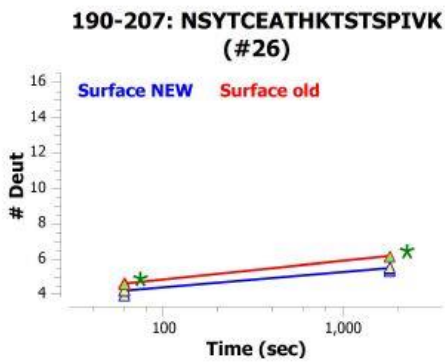

161-183: NSWTDQDSKDSTYSMSSTLTLTK (#21)
Surface NEW
Surface old
Deut
Time (sec)

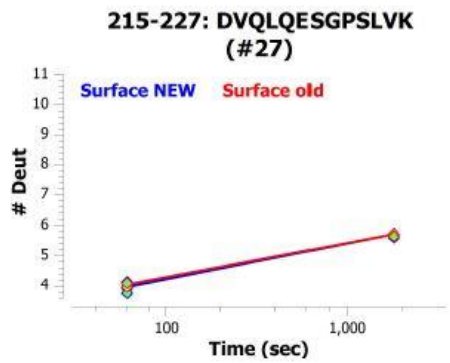

161-184: NSWTDQDSKDSTYSMSSTLTLTKD (#22)
Surface NEW
Surface old
Deut
Time (sec)

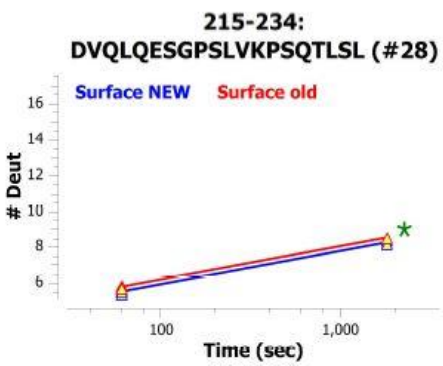

161-185: NSWTDQDSKDSTYSMSSTLTLTKDE (#23)
Surface NEW
Surface old
Deut
Time (sec)

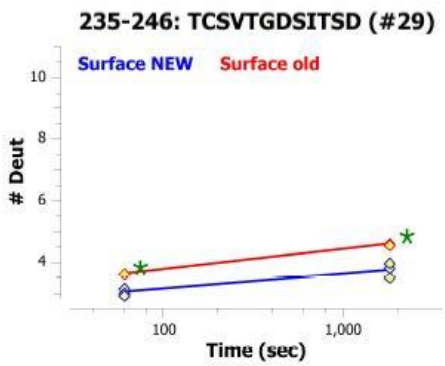

168-185: SKDSTYSMSSTLTLTKDE (#24)
Surface NEW
Surface old
Deut
Time (sec)

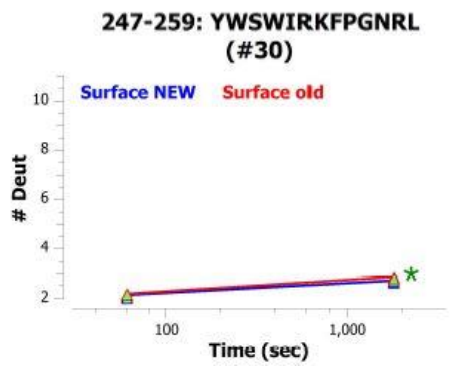

185-207: EYERHNSYTCEATHKTSTSPIVK (#25)
Surface NEW
Surface old
Deut
Time (sec)


190-207: NSYTCEATHKTSTSPIVK (#26)
Surface NEW
Surface old
Deut
Time (sec)

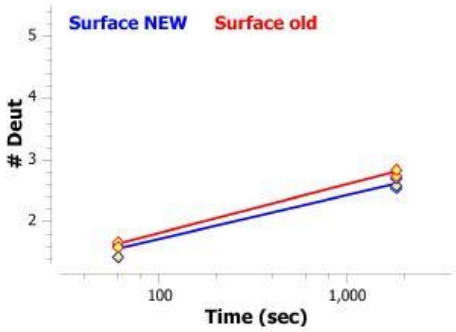

215-227: DVQLQESGPSLVK (#27)
Surface NEW
Surface old
Deut
Time (sec)


215-234: DVQLQESGPSLVKPSQTLSL (#28)
Surface NEW
Surface old
Deut
Time (sec)

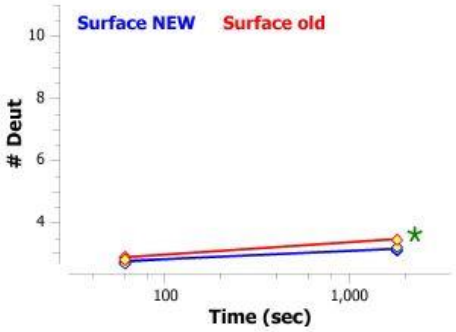

235-246: TCSVTGDSITSD (#29)
Surface NEW
Surface old
Deut
Time (sec)


247-259: YWSWIRKFPGNRL (#30)
Surface NEW
Surface old
Deut
Time (sec)

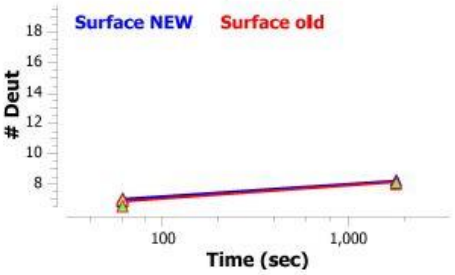
255-261: PGNRLEY (#31)



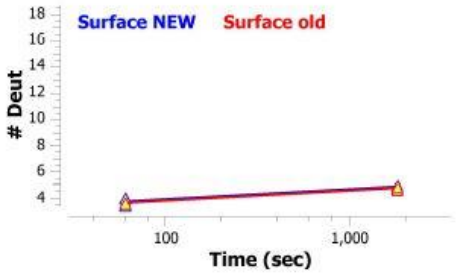
260-271: EYMGYVSYSGST (#32)



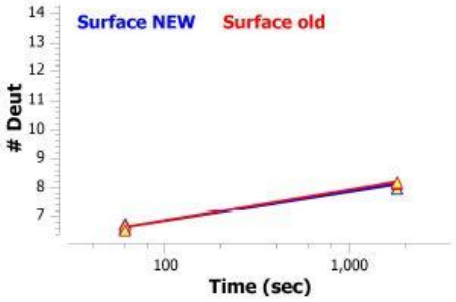
260-280: EYMGYVSYSGSTYYNPSLKSR (#33)



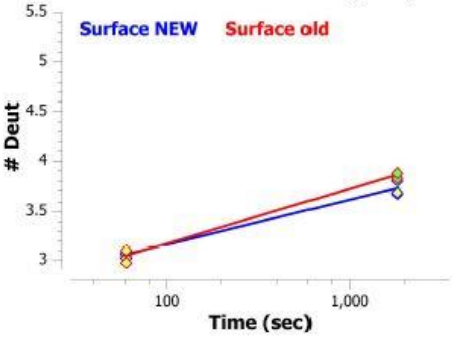
261-280: YMGYVSYSGSTYYNPSLKSR (#34)



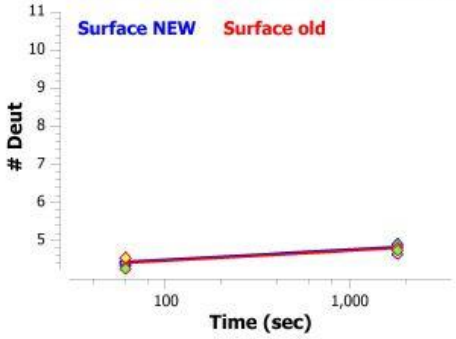
265-280: VSYSGSTYYNPSLKSR (#35)



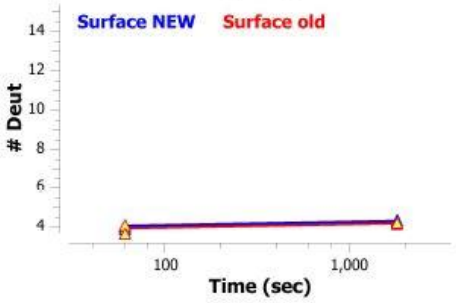
273-280: YNPSLKSR (#36)



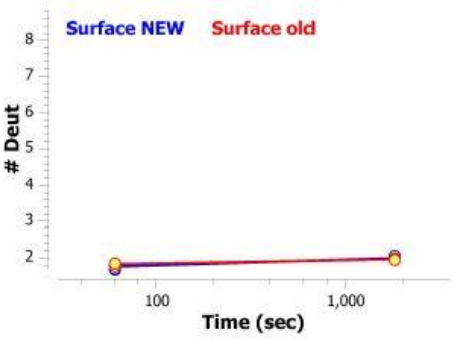
281-292: ISITRDTSKNQY (#37)



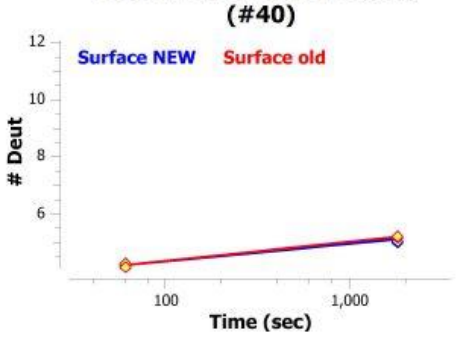
281-296: ISITRDTSKNQYYLDL (#38)



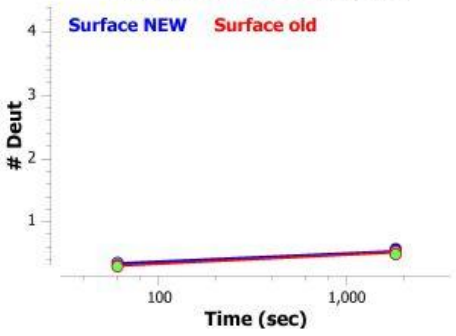
297-306: NSVTTEDTAT (#39)



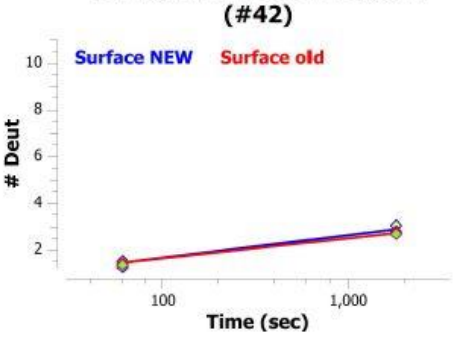
300-312: TTEDTATYYCANW (#40)



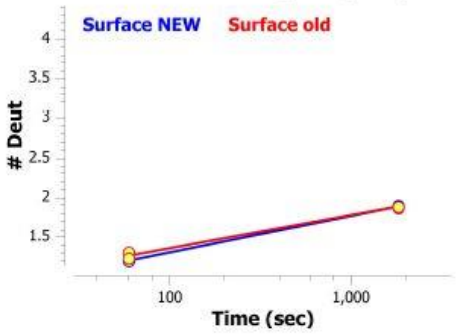
311-316: NWDGDY (#41)



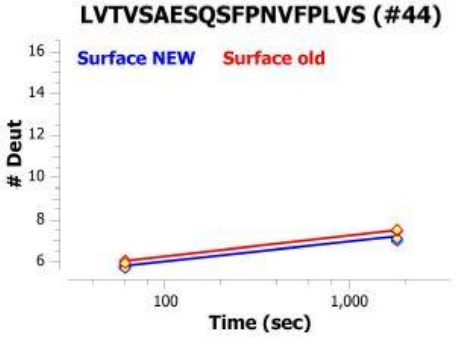
311-322: NWDGDYWGQGTL (#42)



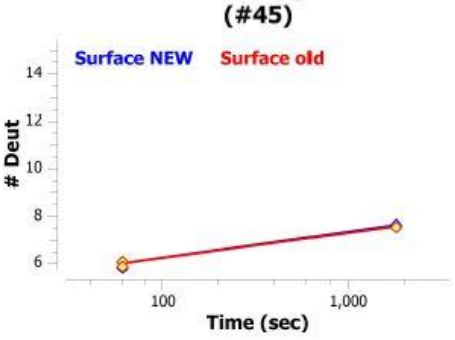
317-322: WGQGTL (#43)



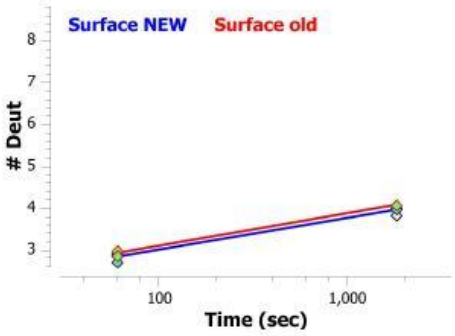
322-340: LVTVSAESQSFPNVFPLVS (#44)



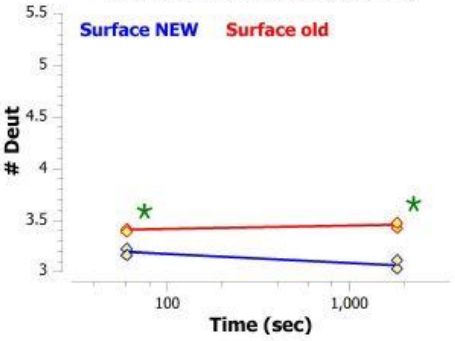
323-340: VTVSAESQSFPNVFPLVS (#45)



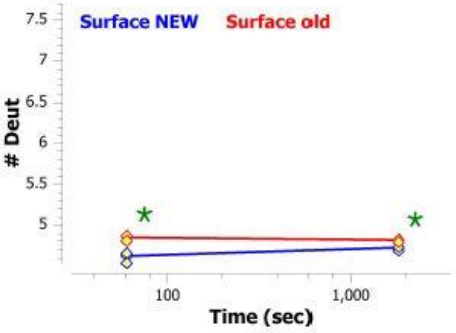
329-340: SQSFPNVFPLVS (#46)



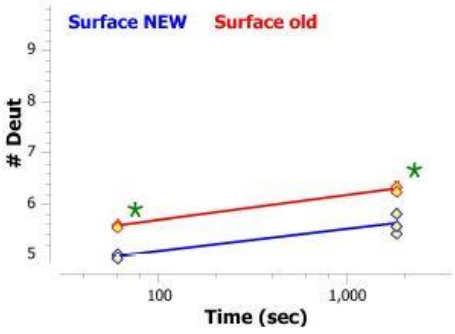
341-348: CESPLSDK (#47)



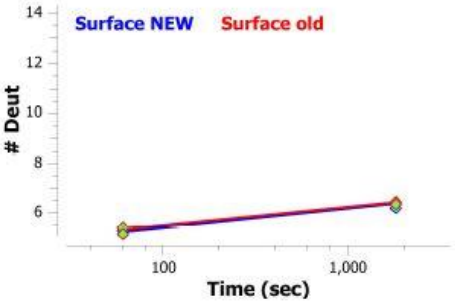
341-350: CESPLSDKNL (#48)



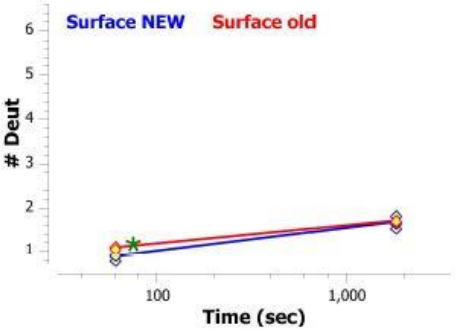
341-352: CESPLSDKNLVA (#49)



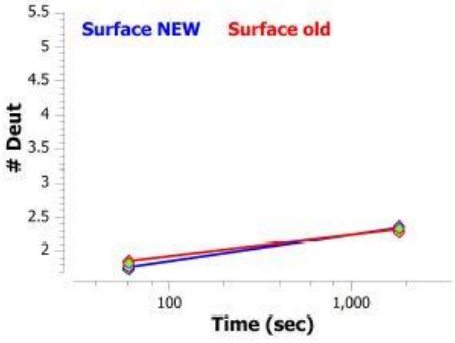
349-364: NLVAMGCLARDFLPST (#50)



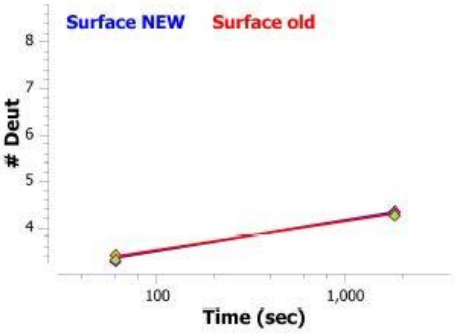
351-358: VAMGCLAR (#51)



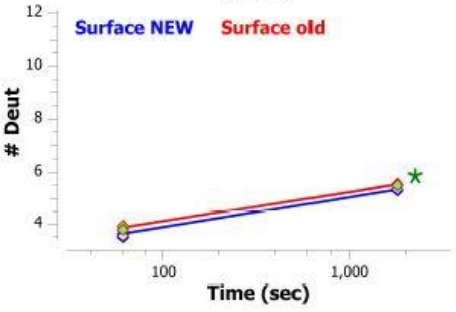
357-364: ARDFLPST (#52)



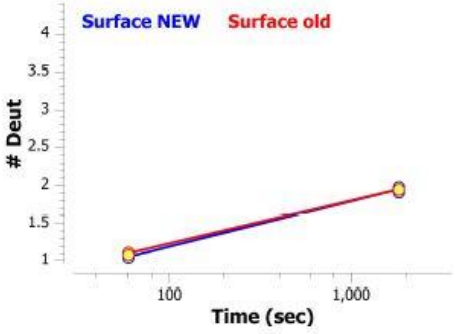
357-367: ARDFLPSTISF (#53)



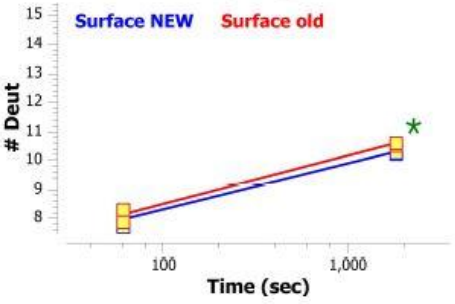
357-370: ARDFLPSTISFTWN (#54)



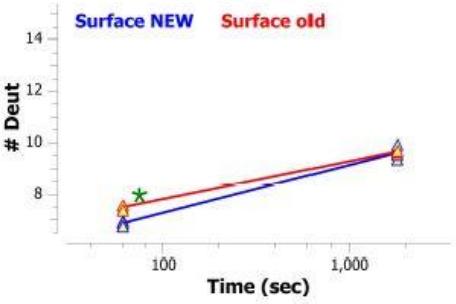
365-370: ISFTWN (#55)



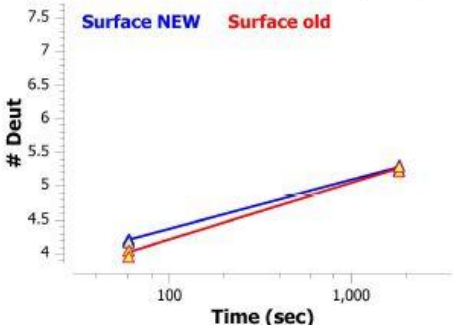
371-387: YQN*NTEVIQGIRTFPTL (#56)



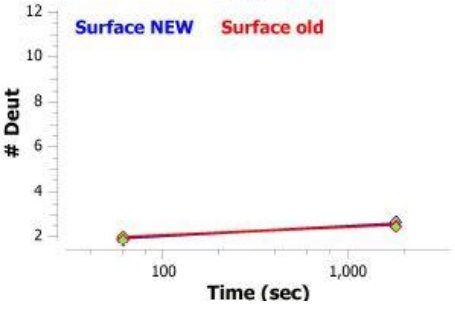
377-393: VIQGIRTFPTLRTGGKY (#57)



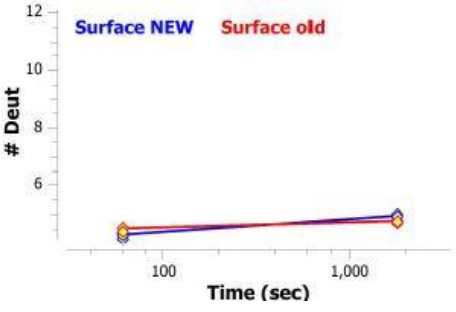
385-393: PTLRTGGKY (#58)



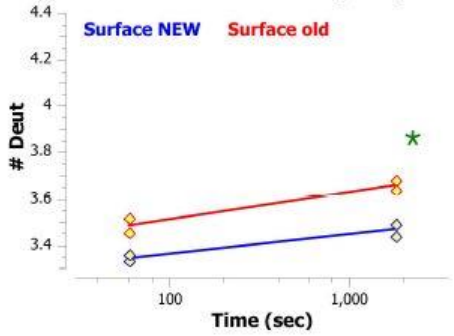
388-400: RTGGKYLATSQVL (#59)



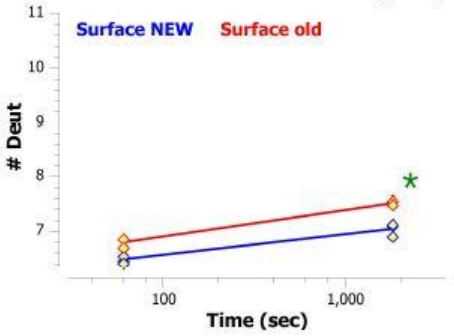
394-407: LATSQVLLSPKSIL (#60)



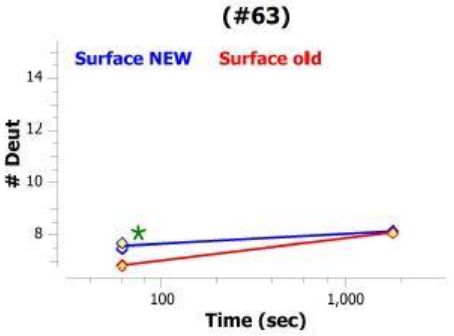
401-407: LSPKSIL (#61)



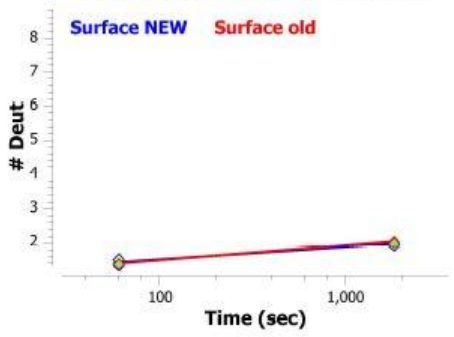
401-413: LSPKSILEGSDEY (#62)



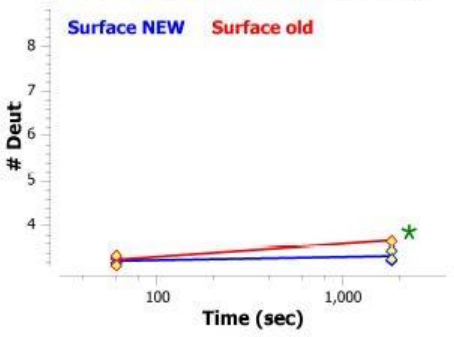
401-417: LSPKSILEGSDEYLVCK (#63)



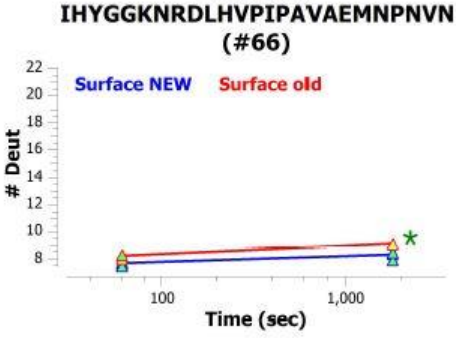
408-417: EGSDEYLVCK (#64)



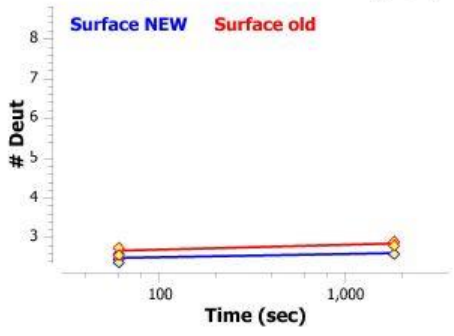
418-427: IHYGGKNRDL (#65)



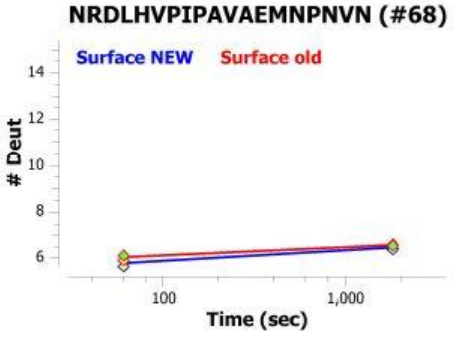
418-442: IHYGGKNRDLHVPIPAVAEMNPNVN (#66)



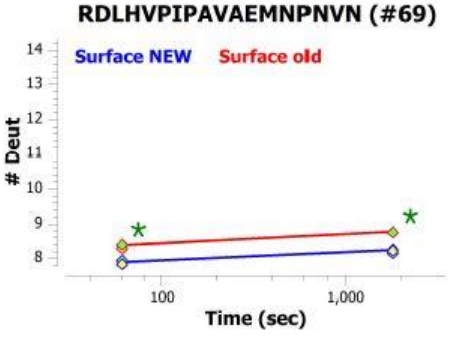
424-435: NRDLHVPIPAVA (#67)



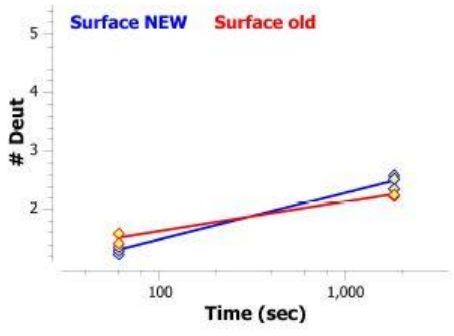
424-442: NRDLHVPIPAVAEMNPNVN (#68)



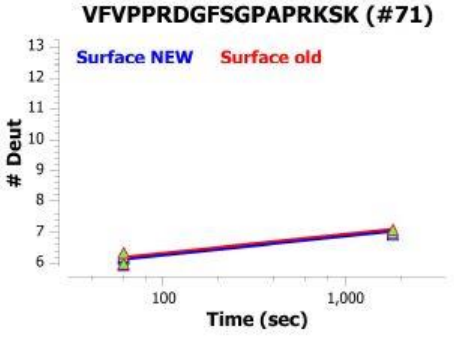
425-442: RDLHVPIPAVAEMNPNVN (#69)



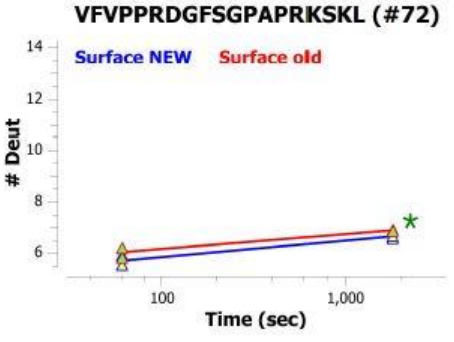
443-451: VFVPPRDGF (#70)



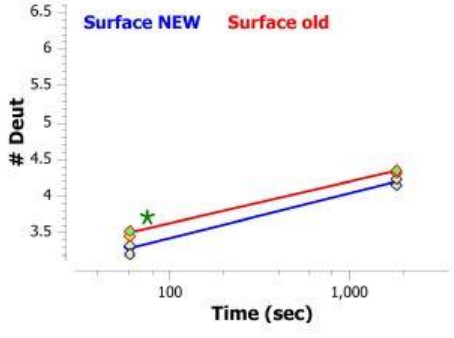
443-460: VFVPPRDGFSGPAPRKSK (#71)



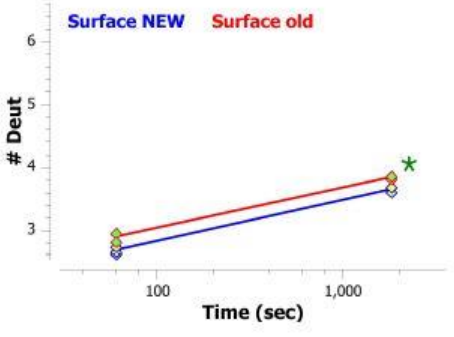
443-461: VFVPPRDGFSGPAPRKSKL (#72)



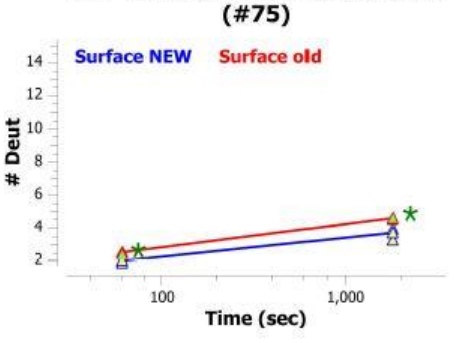
451-460: FSGPAPRKSK (#73)

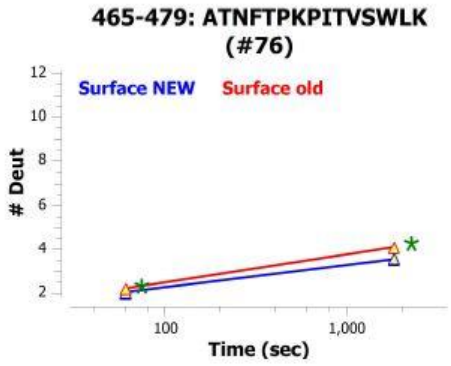


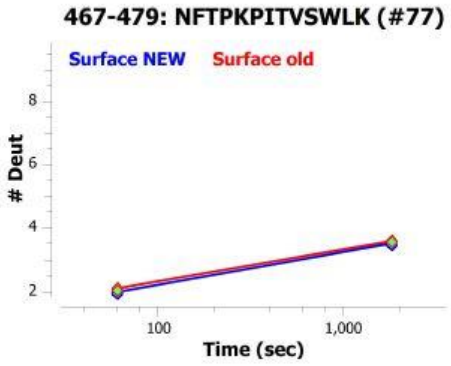
452-461: SGPAPRKSKL (#74)

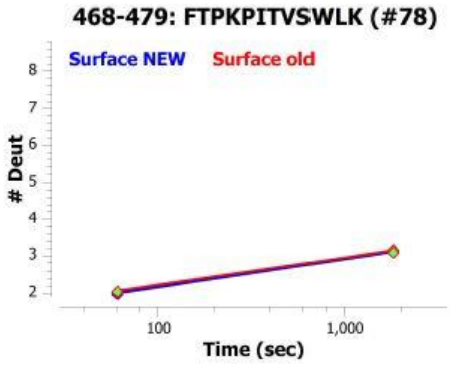


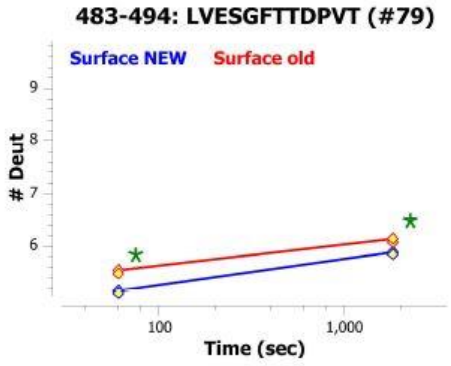
462-479: ICEATNFTPKPITVSWLK (#75)

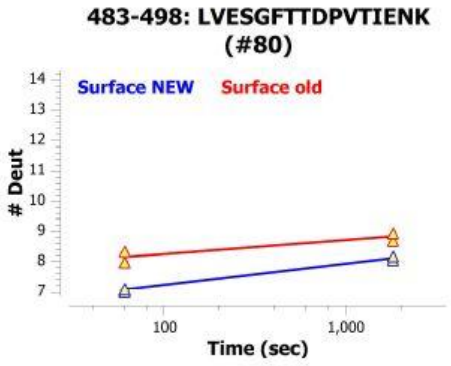


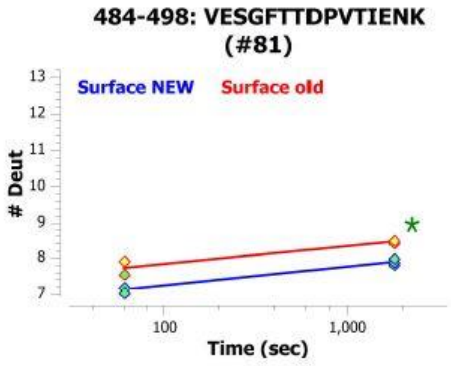

465-479: ATNFTPKPITVSWLK (#76)
Surface NEW
Surface old
Deut
Time (sec)

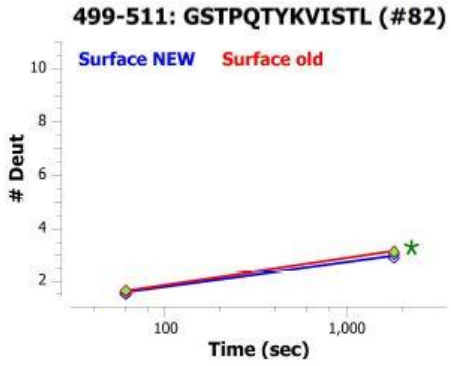

467-479: NFTPKPITVSWLK (#77)
Surface NEW
Surface old
Deut
Time (sec)

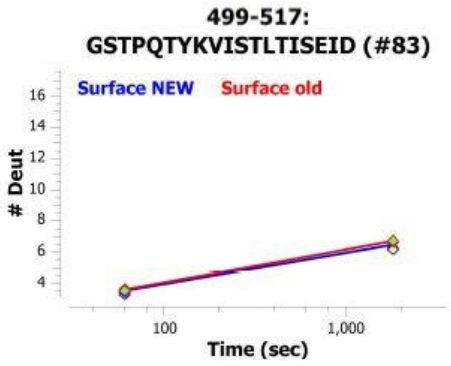

468-479: FTPKPITVSWLK (#78)
Surface NEW
Surface old
Deut
Time (sec)

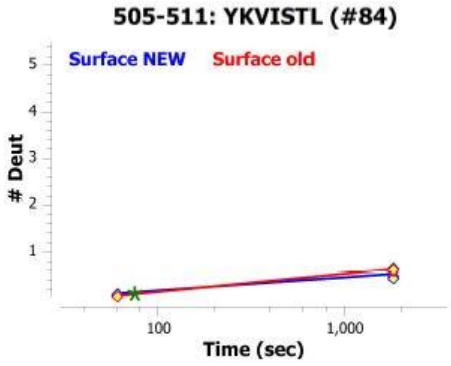

483-494: LVESGFTTDPVT (#79)
Surface NEW
Surface old
Deut
Time (sec)

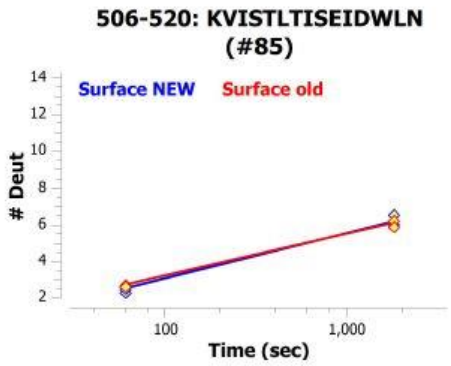

483-498: LVESGFTTDPVTIENK (#80)
Surface NEW
Surface old
Deut
Time (sec)

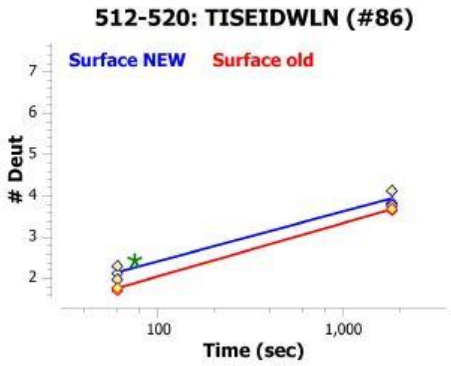

484-498: VESGFTTDPVTIENK (#81)
Surface NEW
Surface old
Deut
Time (sec)

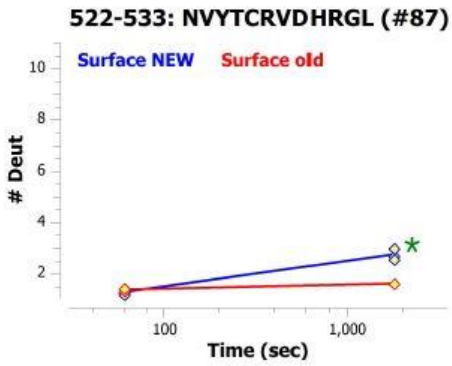

499-511: GSTPQTYKVISTL (#82)
Surface NEW
Surface old
Deut
Time (sec)

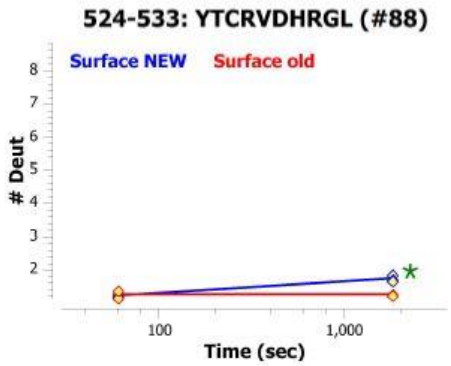

499-517: GSTPQTYKVISTLTISEID (#83)
Surface NEW
Surface old
Deut
Time (sec)

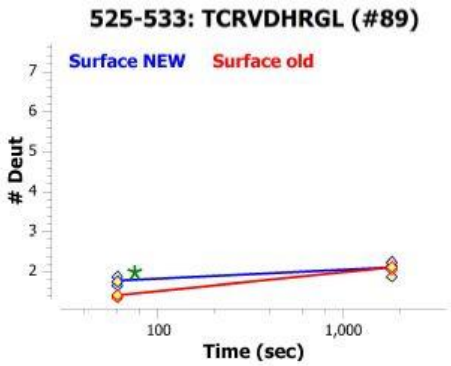

505-511: YKVISTL (#84)
Surface NEW
Surface old
Deut
Time (sec)

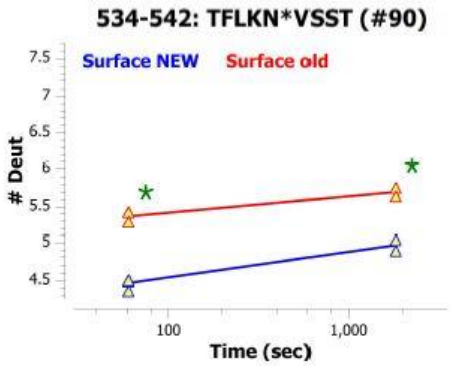

506-520: KVISTLTISEIDWLN (#85)
Surface NEW
Surface old
Deut
Time (sec)

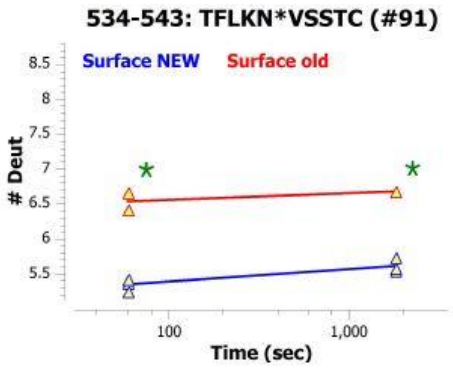

512-520: TISEIDWLN (#86)
Surface NEW
Surface old
Deut
Time (sec)

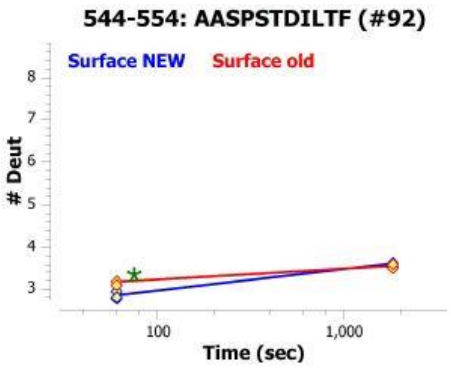

522-533: NVYTCRVDHRGL (#87)
Surface NEW
Surface old
Deut
Time (sec)

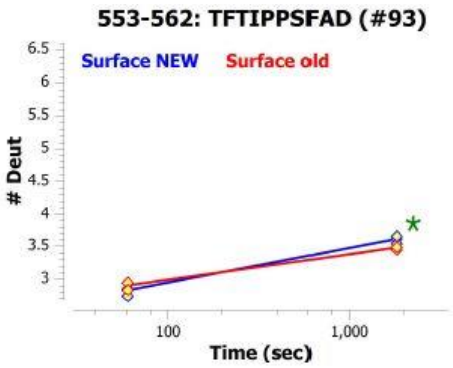

524-533: YTCRVDHRGL (#88)
Surface NEW
Surface old
Deut
Time (sec)

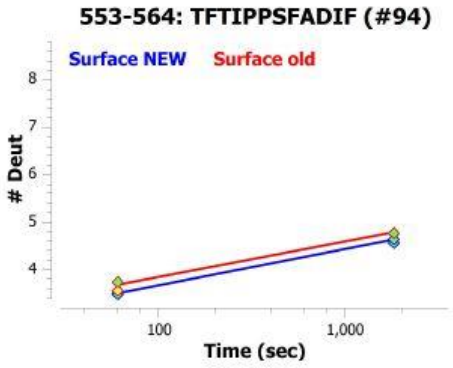

525-533: TCRVDHRGL (#89)
Surface NEW
Surface old
Deut
Time (sec)

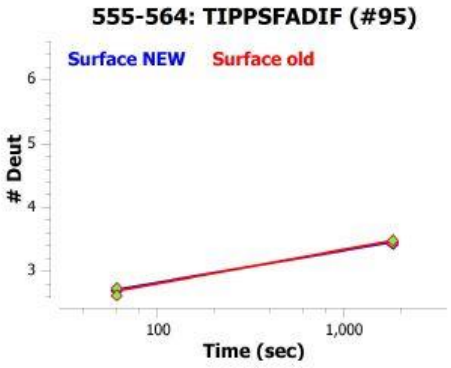

534-542: TFLKN*VSST (#90)
Surface NEW
Surface old
Deut
Time (sec)

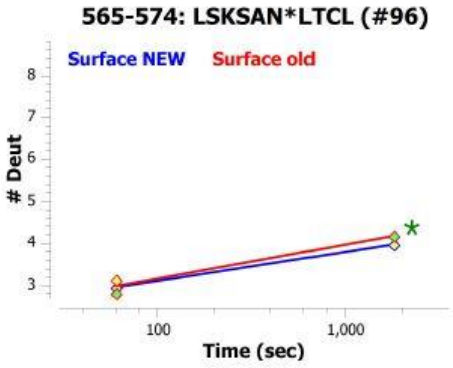

534-543: TFLKN*VSSTC (#91)
Surface NEW
Surface old
Deut
Time (sec)

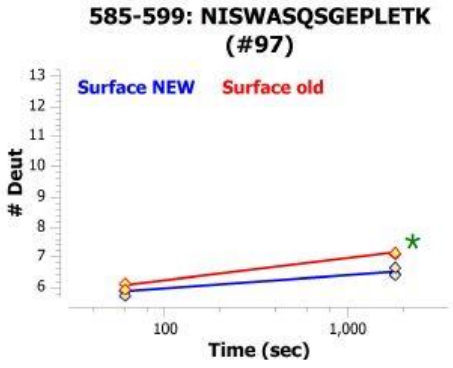

544-554: AASPSTDILTF (#92)
Surface NEW
Surface old
Deut
Time (sec)

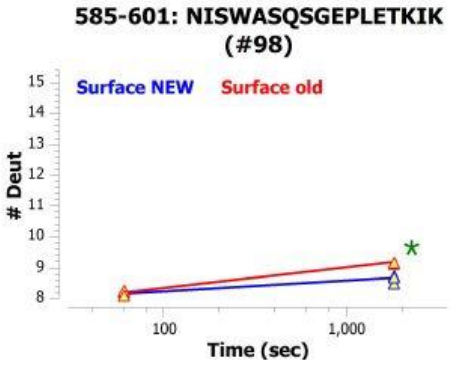

553-562: TFTIPPSFAD (#93)
Surface NEW
Surface old
Deut
Time (sec)

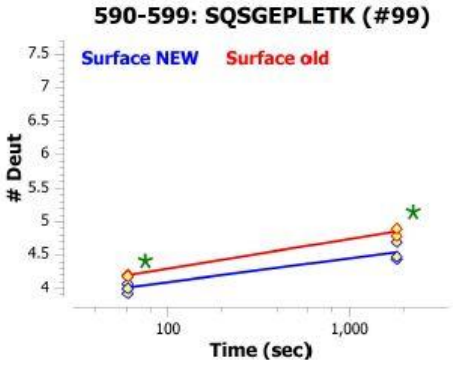

553-564: TFTIPPSFADIF (#94)
Surface NEW
Surface old
Deut
Time (sec)

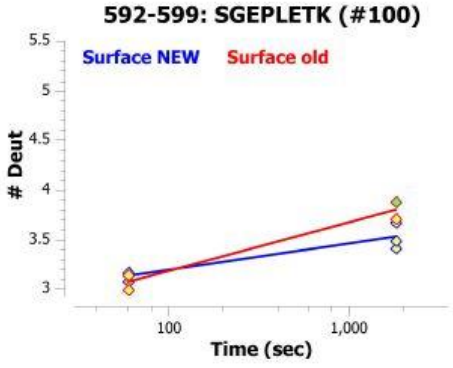

555-564: TIPPSFADIF (#95)
Surface NEW
Surface old
Deut
Time (sec)

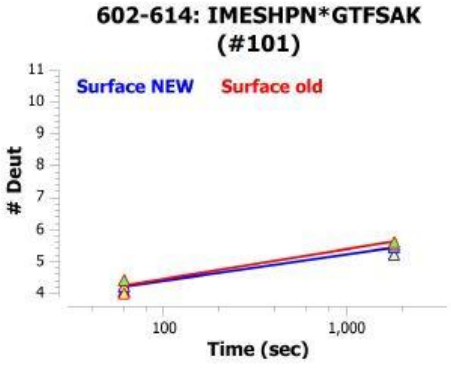

565-574: LSKSAN*LTCL (#96)
Surface NEW
Surface old
Deut
Time (sec)

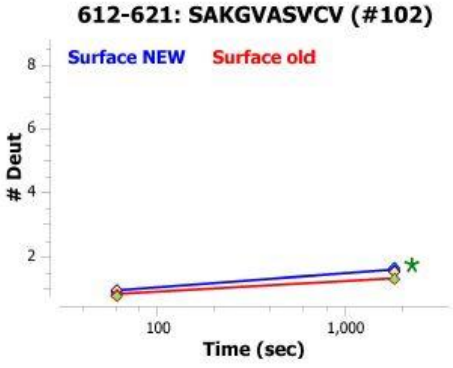

585-599: NISWASQSGEPLETK (#97)
Surface NEW
Surface old
Deut
Time (sec)

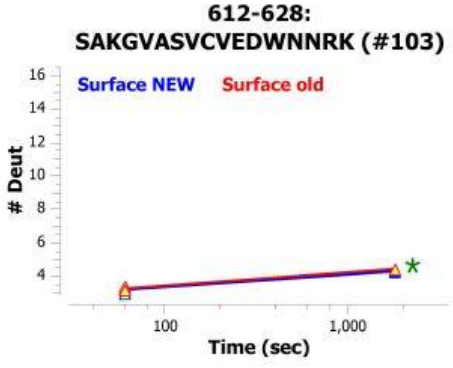

585-601: NISWASQSGEPLETKIK (#98)
Surface NEW
Surface old
Deut
Time (sec)

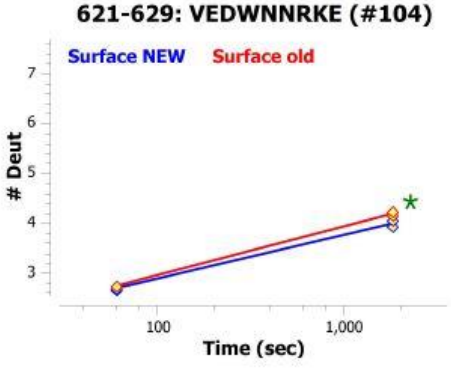

590-599: SQSGEPLETK (#99)
Surface NEW
Surface old
Deut
Time (sec)

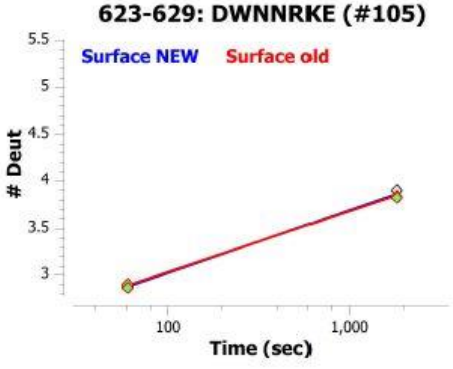

592-599: SGEPLETK (#100)
Surface NEW
Surface old
Deut
Time (sec)


602-614: IMESHPN*GTFSAK (#101)
Surface NEW
Surface old
Deut
Time (sec)

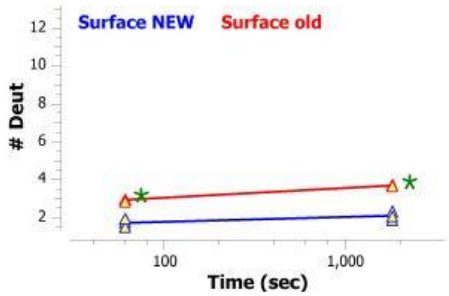

612-621: SAKGVASVCV (#102)
Surface NEW
Surface old
Deut
Time (sec)


612-628: SAKGVASVCVEDWNNRK (#103)
Surface NEW
Surface old
Deut
Time (sec)

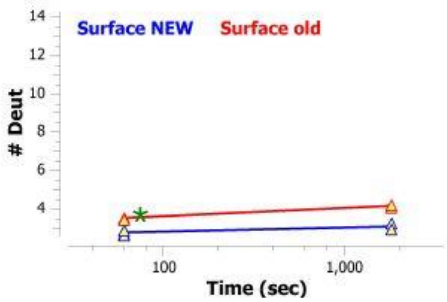

621-629: VEDWNNRKE (#104)
Surface NEW
Surface old
Deut
Time (sec)


623-629: DWNNRKE (#105)
Surface NEW
Surface old
Deut
Time (sec)

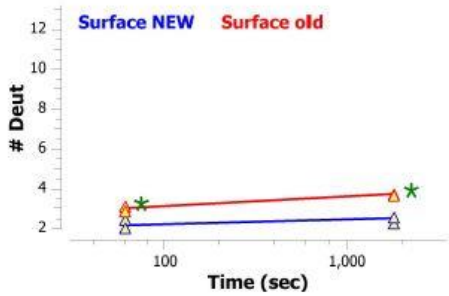
630-645: FVCTVTHRDLPSPQKK (#106)



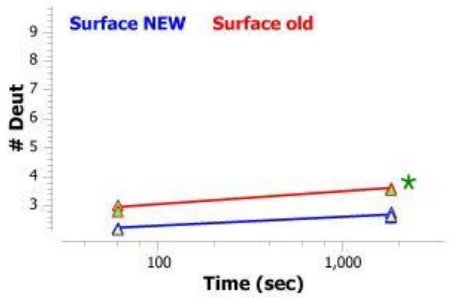
630-646: FVCTVTHRDLPSPQKKF (#107)



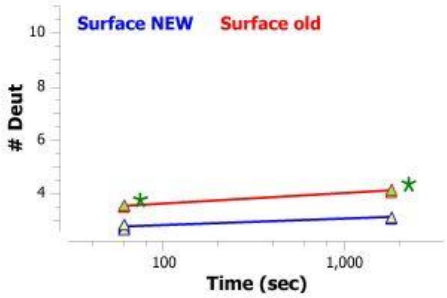
631-646: VCTVTHRDLPSPQKKF (#108)



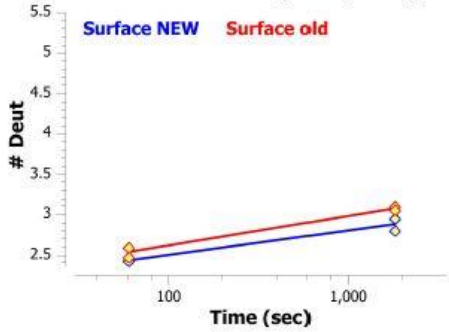
633-645: TVTHRDLPSPQKK (#109)



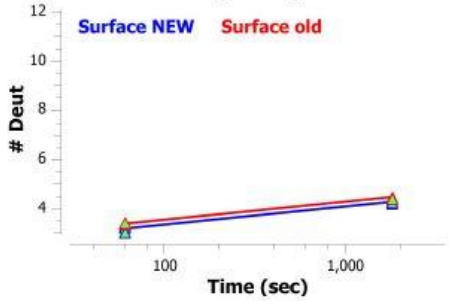
633-646: TVTHRDLPSPQKKF (#110)



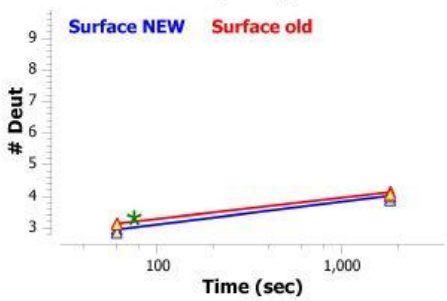
639-646: LPSPQKKF (#111)



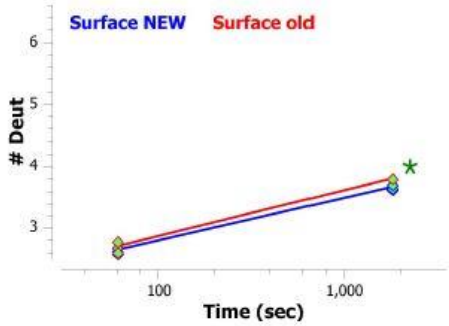
646-661: FISKPNEVHKHPPAVY (#112)



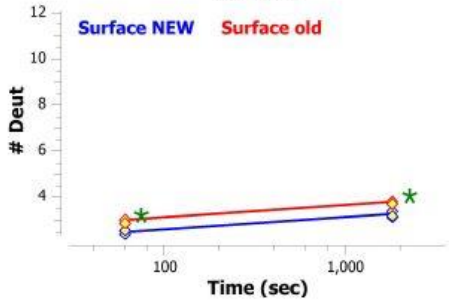
647-660: ISKPNEVHKHPPAV (#113)



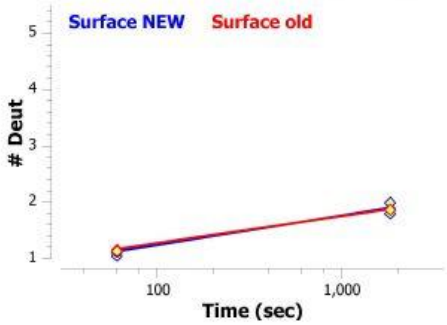
661-670: YLLPPAREQL (#114)



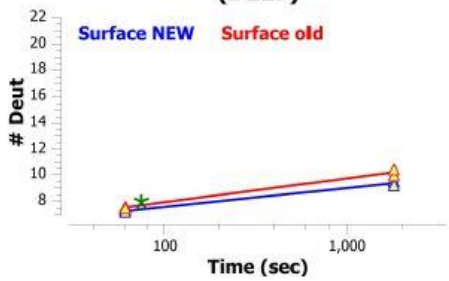
671-683: NLRESATVTCLVK (#115)



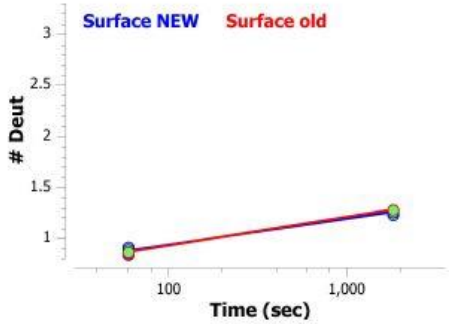
682-689: VKGFSPAD (#116)



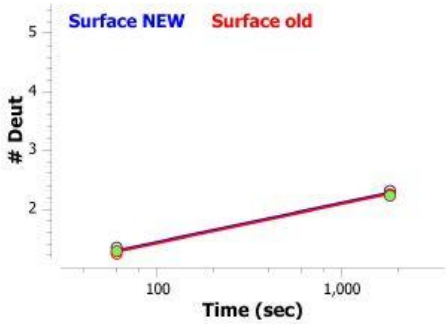
682-705: VKGFSPADISVQWLQRGQLLPQEK (#117)



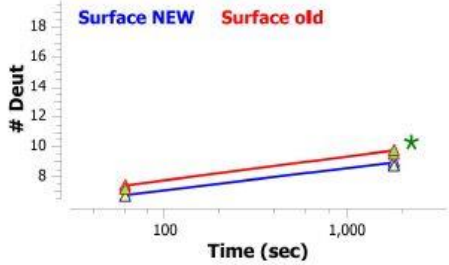
684-689: GFSPAD (#118)



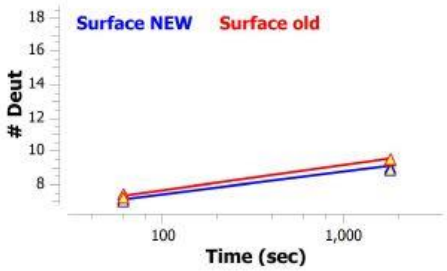
684-691: GFSPADIS (#119)



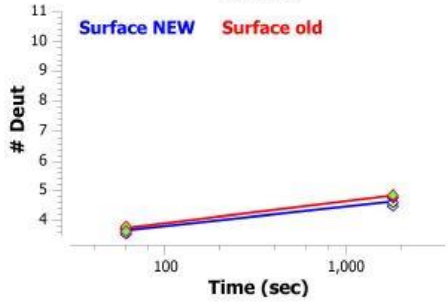
684-705: GFSPADISVQWLQRGQLLPQEK (#120)



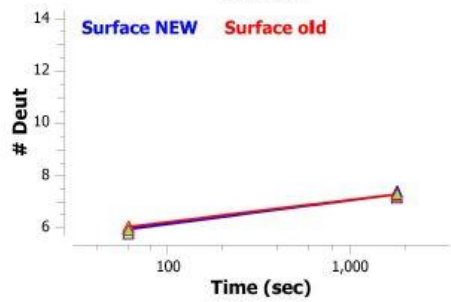
685-705: FSPADISVQWLQRGQLLPQEK (#121)



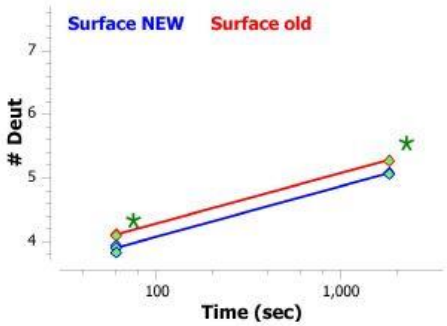
690-701: ISVQWLQRGQLL (#122)



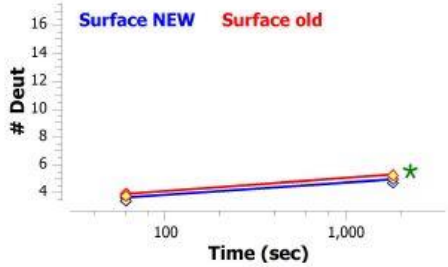
690-705: ISVQWLQRGQLLPQEK (#123)



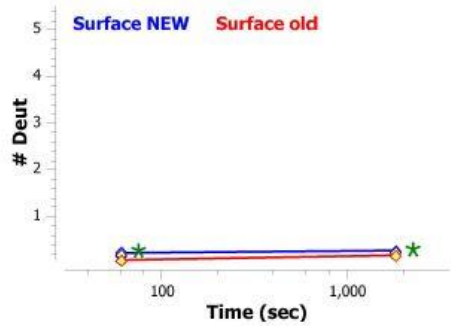
696-705: QRGQLLPQEK (#124)



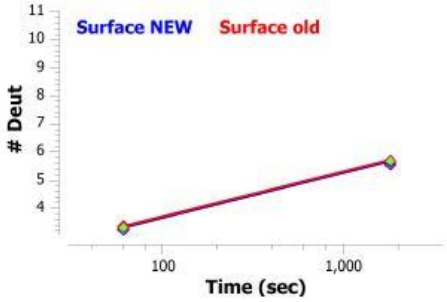
706-727: YVTSAPMPEPGAPGFYFTHSIL (#125)



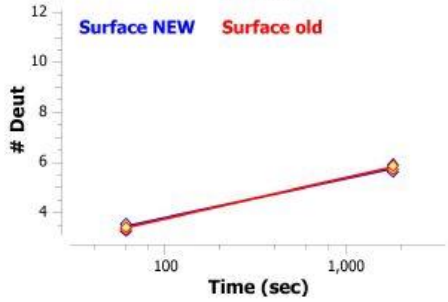
721-727: YFTHSIL (#126)



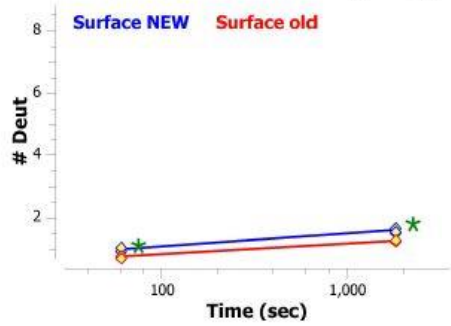
728-739: TVTEEEWNSGET (#127)



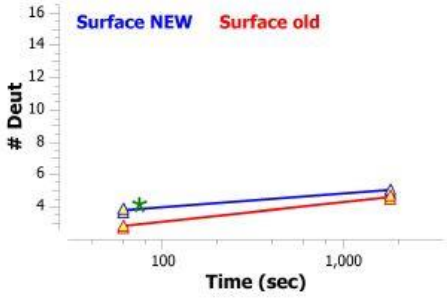
728-740: TVTEEEWNSGETY (#128)



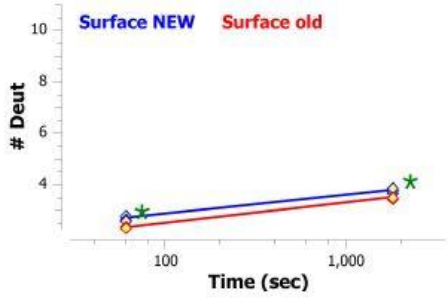
740-749: YTCVVGHEAL (#129)



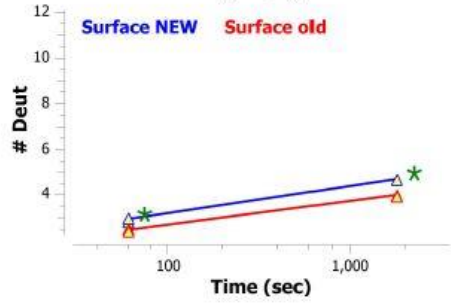
743-760: VVGHEALPHLVTERTVDK (#130)



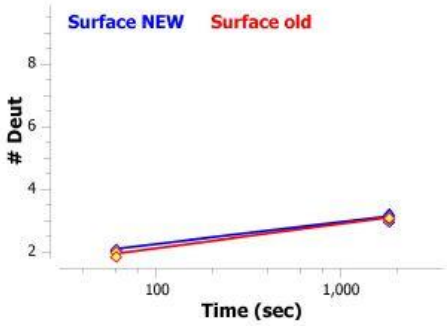
747-759: EALPHLVTERTVD (#131)



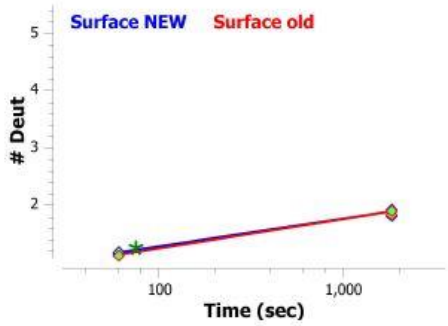
747-760: EALPHLVTERTVDK (#132)



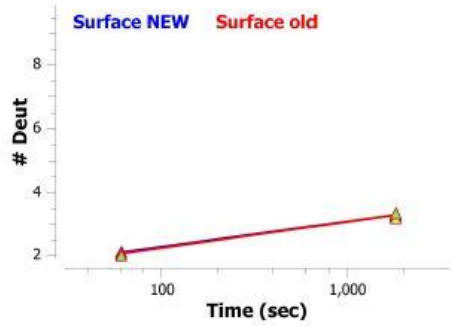
749-759: LPHLVTERTVD (#133)



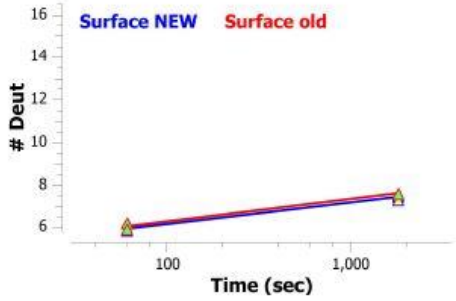
750-756: PHLVTER (#134)



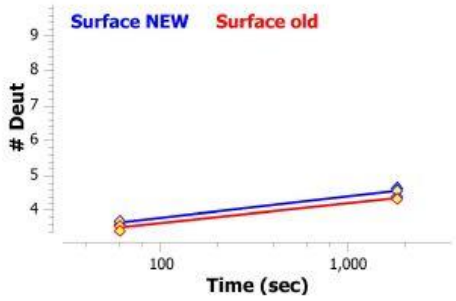
750-760: PHLVTERTVDK (#135)



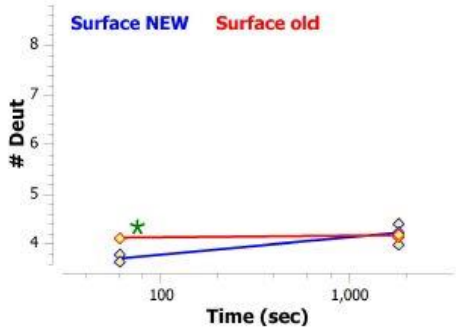
750-767: PHLVTERTVDKSTGKPTL (#136)



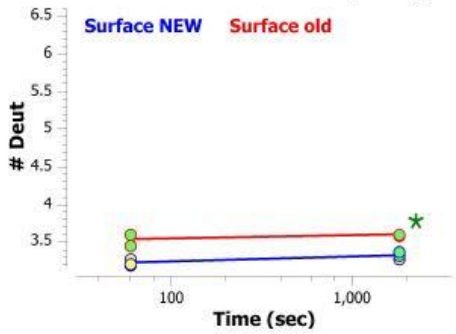
761-772: STGKPTLYN*VSL (#137)



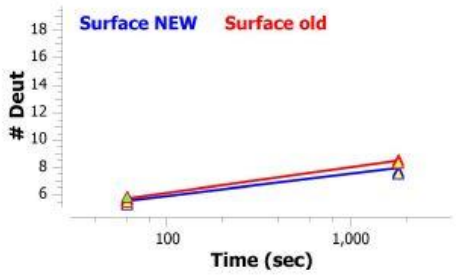
773-782: IMSDTGGTCY (#138)



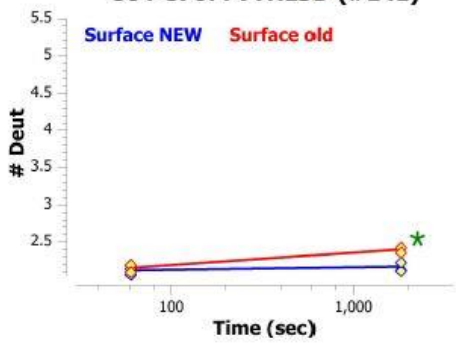
775-782: SDTGGTCY (#139)



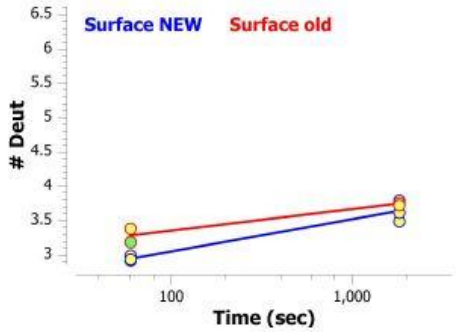
821-842: VTSRIIPSTEDPNEDIVERNIR (#140)



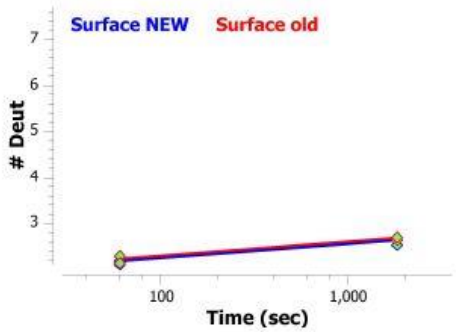
864-870: FVYHLSD (#141)



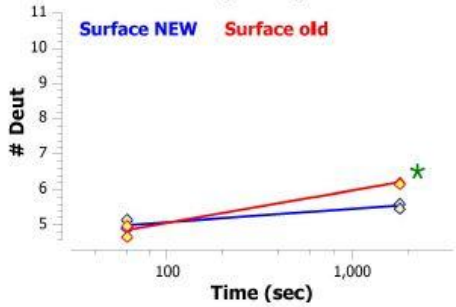
882-889: LEDQVVTA (#142)



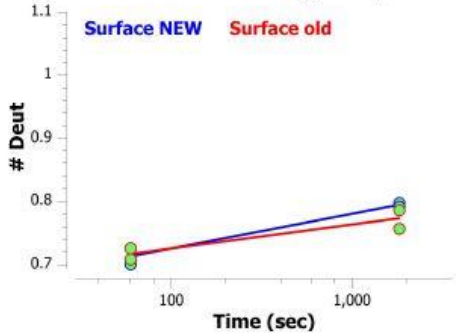
899-908: DGVPETCYMY (#143)



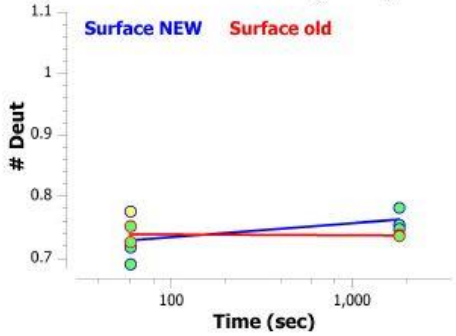
928-941: MVQAALTPDSCYPD (#144)



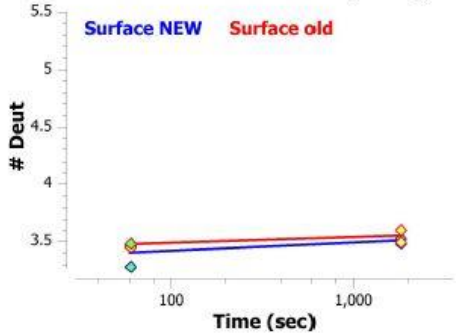
942-944: YPI (#145)



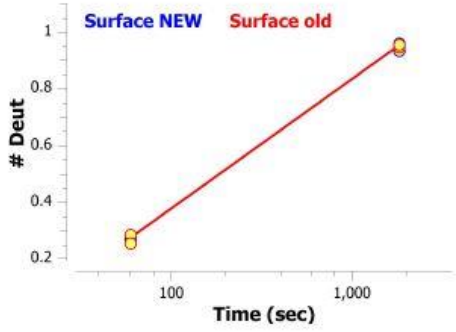
945-948: PPPF (#146)



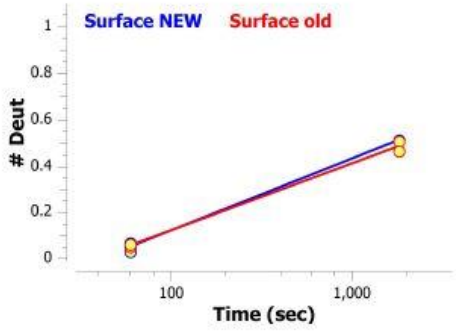
962-969: DRVYIHPF (#147)



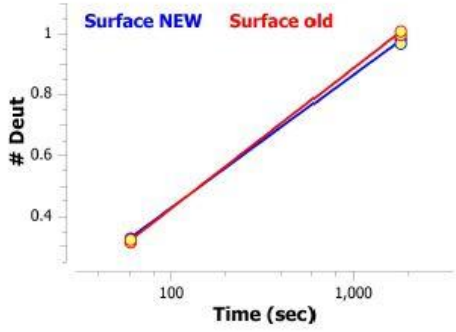
970-972: Y*W*R* (#148)

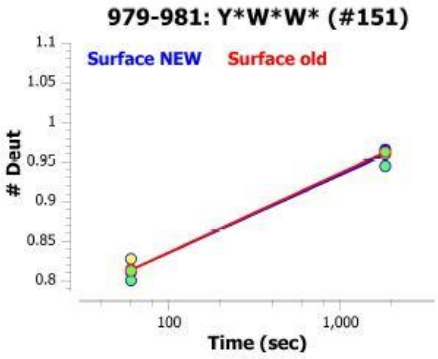


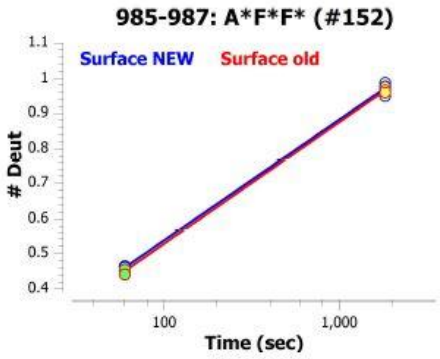
973-975: W*W*W* (#149)

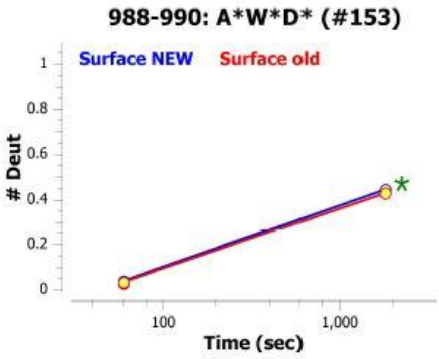


976-978: R*Y*W* (#150)




979-981: Y*W*W* (#151)
Surface NEW
Surface old
Deut
Time (sec)


985-987: A*F*F* (#152)
Surface NEW
Surface old
Deut
Time (sec)


988-990: A*W*D* (#153)
Surface NEW
Surface old
Deut
Time (sec)

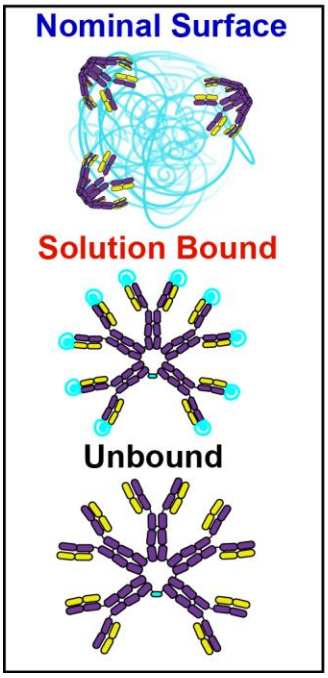
Supplementary Table S4.2 | Three-State HDX - Experiment Statistics

Data Set	Unbound	Solution Bound†	Nominal Surface
HDX reaction details	pH 7.4, 22°C, 85% D2O		
HDX time course	1 min, 10 min, 100 min		
HDX controls	Undeuterated		
Back-exchange (mean) bradykinin peptide internal standard	24.13%	24.69%	24.70%
Number of peptides	**143 peptides**		
Sequence coverage	**88.6% Total Sequence Coverage (LC ~95%, HC ~99%, JC ~37%)**		
Average peptide length / Redundancy	13.3 +/- 4.8 residues		
	redundancy: 1.9 peptides/residue		
Replicates (biological or technical)	Triplicate biological replicates	Triplicate biological replicates	Duplicate biological replicates
Repeatability	**0.084 (avg std deviation)**	**0.083 (avg std deviation)**	**0.076 (avg std deviation)**
Significant differences in HDX	**0.2405 D (95% CI) for Solution Bound vs Unbound***		
	0.2425 D (95% CI) for Surface Bound vs Solution Bound*		

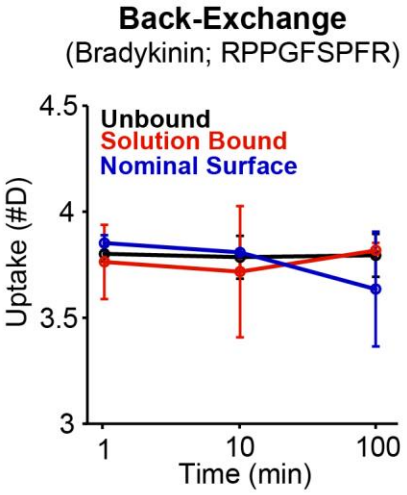
† Condition is labeled as "mono" in generated uptake plots

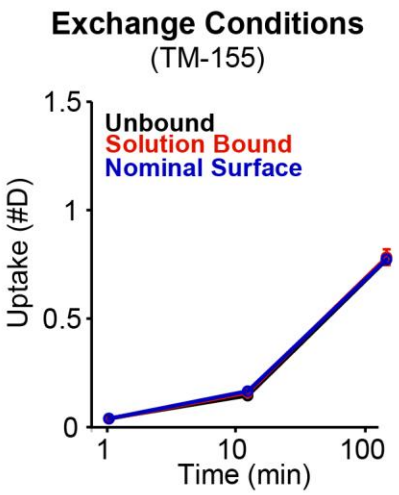
*This calculation is done by HDExaminer and does not include occupied YETL glycopep

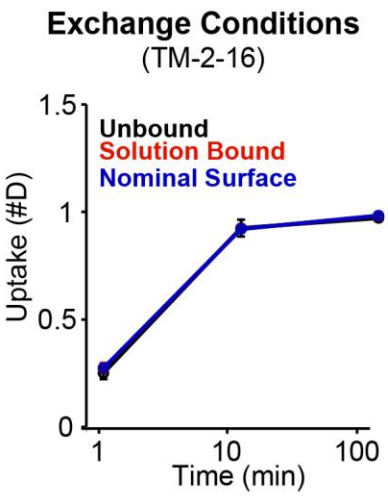
Sample Conditions

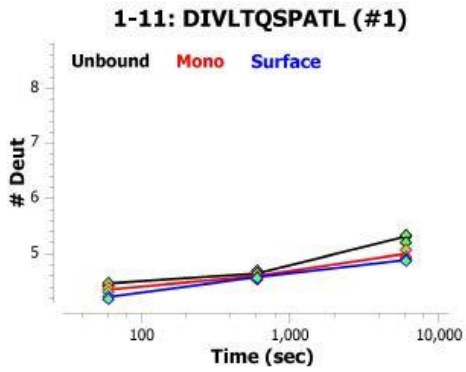


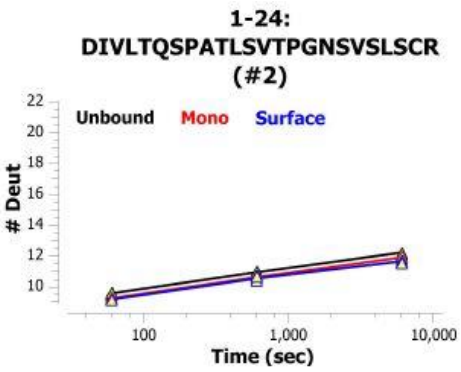
Internal Standards

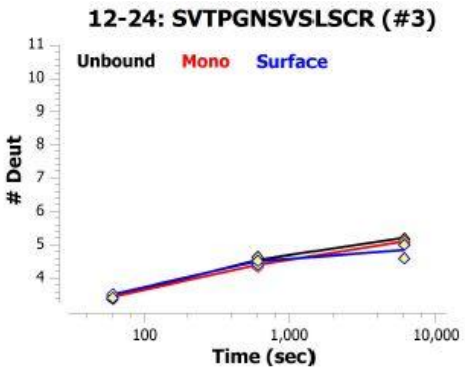


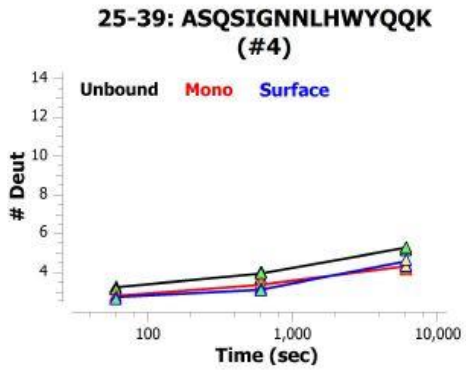


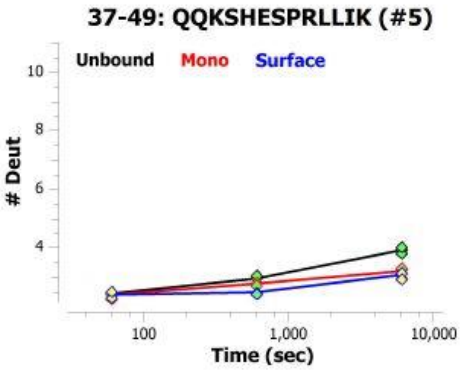


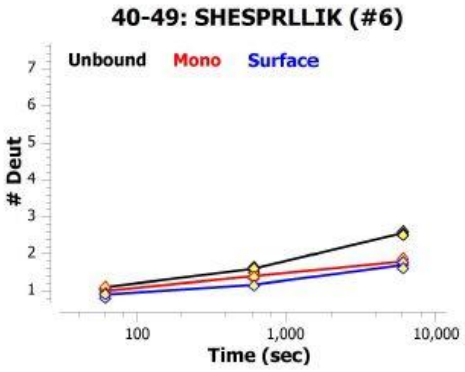

1-11: DIVLTQSPATL (#1)
Unbound Mono Surface
Deut
Time (sec)

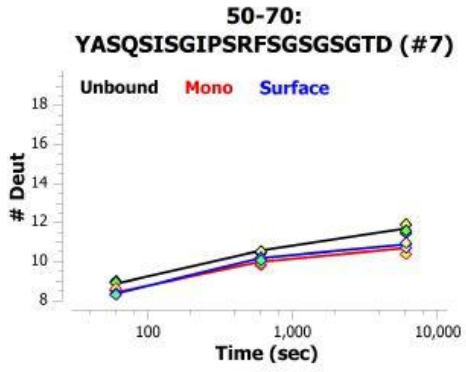

1-24: DIVLTQSPATLSVTPGNSVSLSCR (#2)
Unbound Mono Surface
Deut
Time (sec)

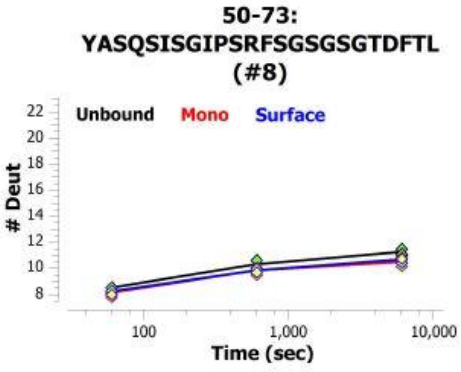

12-24: SVTPGNSVSLSCR (#3)
Unbound Mono Surface
Deut
Time (sec)

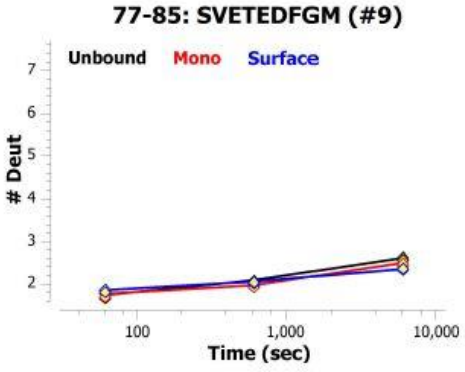

25-39: ASQSIGNNLHWYQQK (#4)
Unbound Mono Surface
Deut
Time (sec)

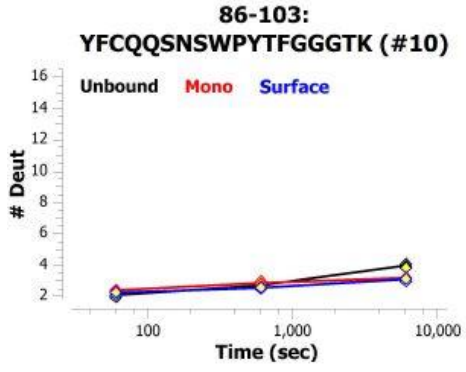

37-49: QQKSHESPRLLIK (#5)
Unbound Mono Surface
Deut
Time (sec)

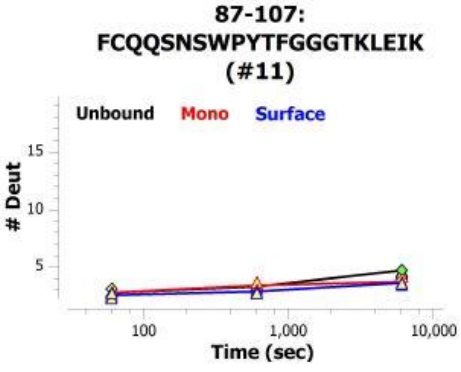

40-49: SHESPRLLIK (#6)
Unbound Mono Surface
Deut
Time (sec)

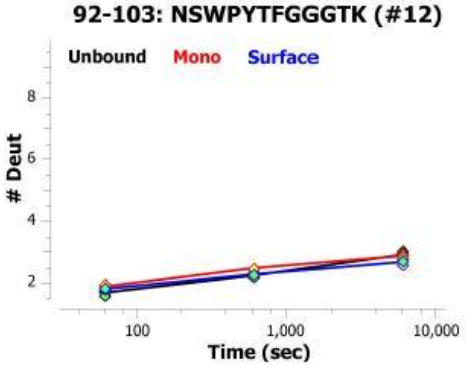

50-70: YASQSISGIPSRFSGSGSGTD (#7)
Unbound Mono Surface
Deut
Time (sec)

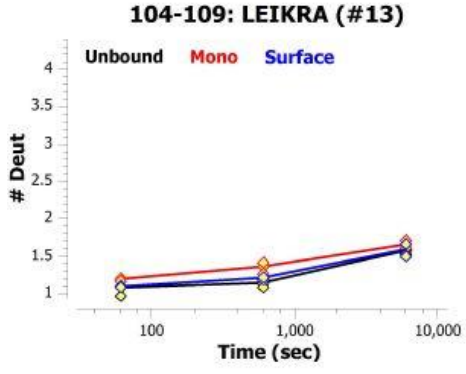

50-73: YASQSISGIPSRFSGSGSGTDFTL (#8)
Unbound Mono Surface
Deut
Time (sec)

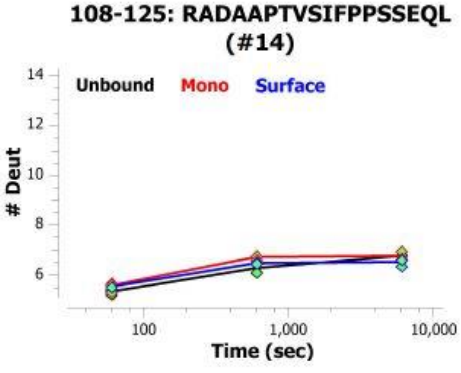

77-85: SVETEDFGM (#9)
Unbound Mono Surface
Deut
Time (sec)

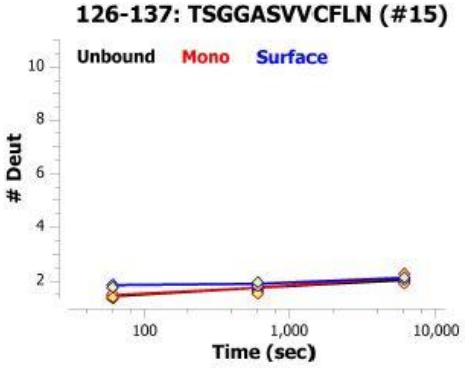

86-103: YFCQQSNSWPYTFGGGTK (#10)
Unbound Mono Surface
Deut
Time (sec)

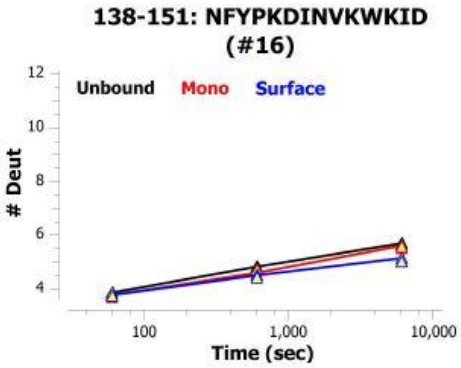

87-107: FCQQSNSWPYTFGGGTKLEIK (#11)
Unbound Mono Surface
Deut
Time (sec)

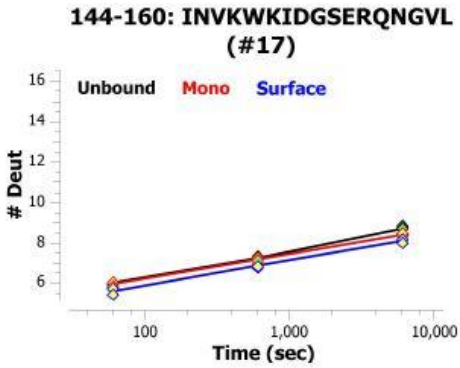

92-103: NSWPYTFGGGTK (#12)
Unbound Mono Surface
Deut
Time (sec)

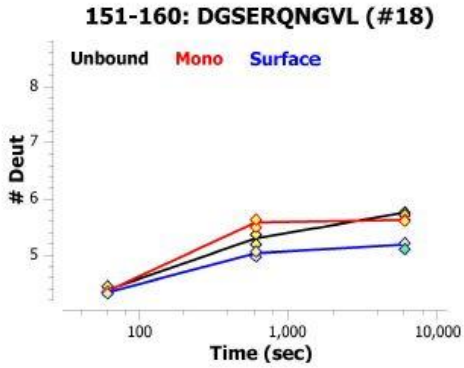

104-109: LEIKRA (#13)
Unbound Mono Surface
Deut
Time (sec)

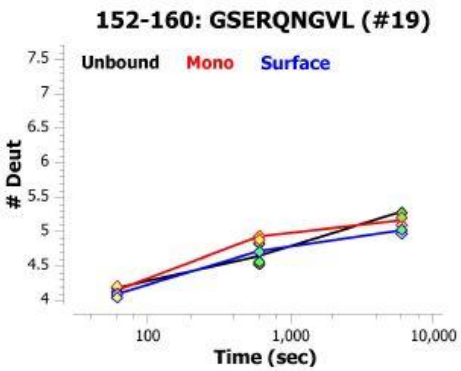

108-125: RADAAPTVSIFPPSSEQL (#14)
Unbound Mono Surface
Deut
Time (sec)

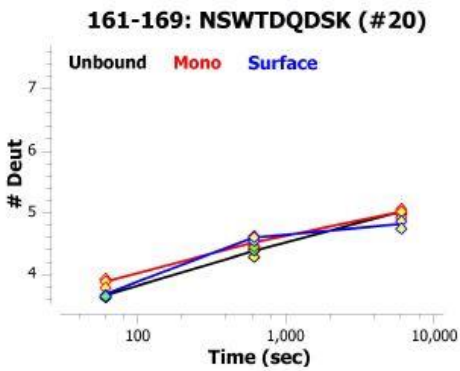

126-137: TSGGASVVCFLN (#15)
Unbound Mono Surface
Deut
Time (sec)

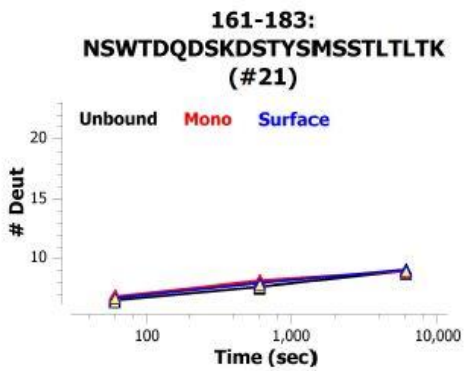

138-151: NFYPKDINVKWKID (#16)
Unbound Mono Surface
Deut
Time (sec)

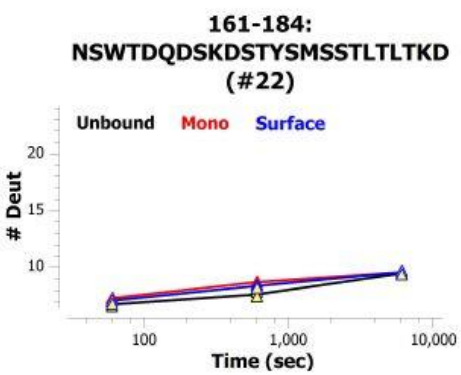

144-160: INVKWKIDGSERQNGVL (#17)
Unbound Mono Surface
Deut
Time (sec)

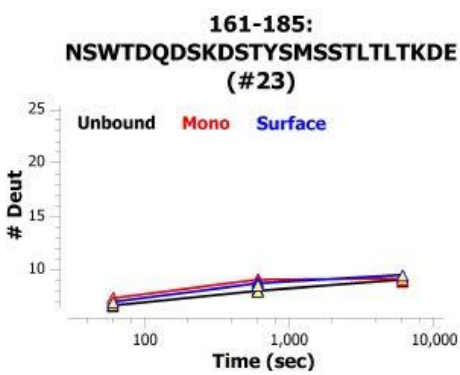

151-160: DGSERQNGVL (#18)
Unbound Mono Surface
Deut
Time (sec)

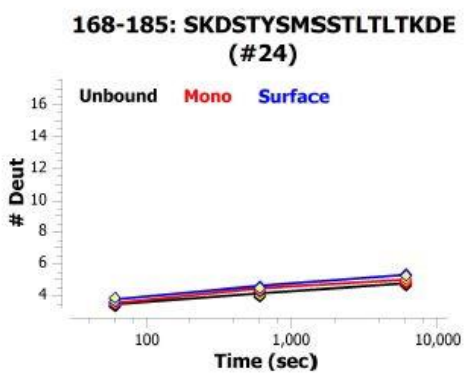

152-160: GSERQNGVL (#19)
Unbound Mono Surface
Deut
Time (sec)

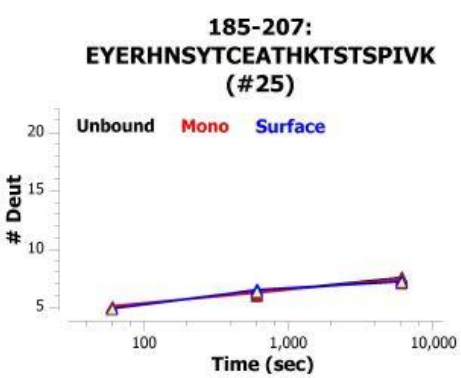

161-169: NSWTDQDSK (#20)
Unbound Mono Surface
Deut
Time (sec)

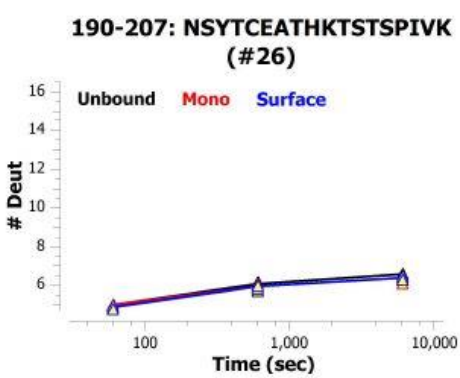

161-183: NSWTDQDSKDSTYSMSSTLTLTK (#21)
Unbound Mono Surface
Deut
Time (sec)

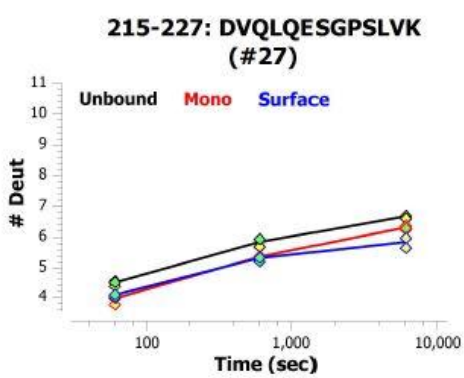

161-184: NSWTDQDSKDSTYSMSSTLTLTKD (#22)
Unbound Mono Surface
Deut
Time (sec)

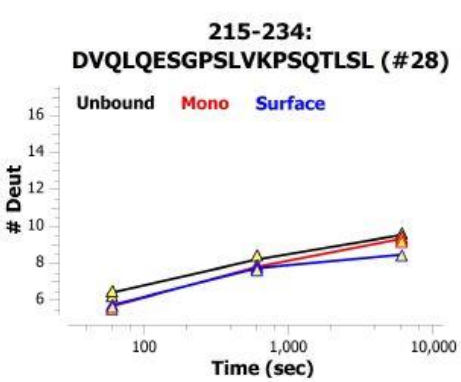

161-185: NSWTDQDSKDSTYSMSSTLTLTKDE (#23)
Unbound Mono Surface
Deut
Time (sec)

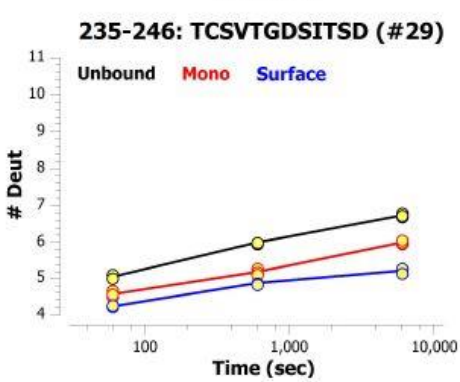

168-185: SKDSTYSMSSTLTLTKDE (#24)
Unbound Mono Surface
Deut
Time (sec)

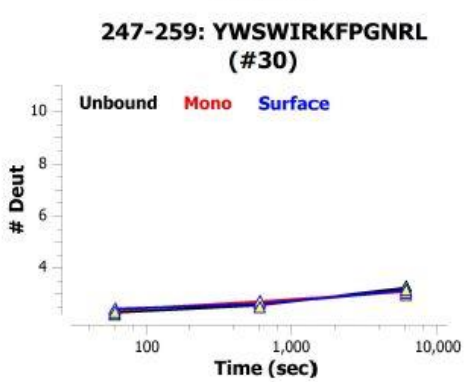

185-207: EYERHNSYTCEATHKTSTSPIVK (#25)
Unbound Mono Surface
Deut
Time (sec)

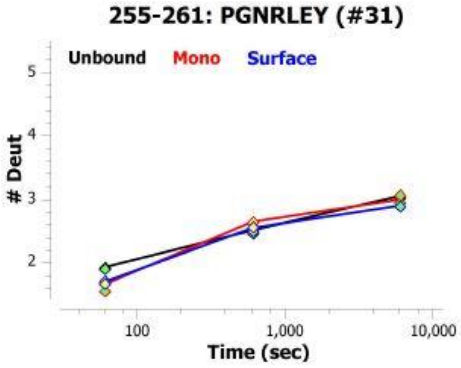

190-207: NSYTCEATHKTSTSPIVK (#26)
Unbound Mono Surface
Deut
Time (sec)

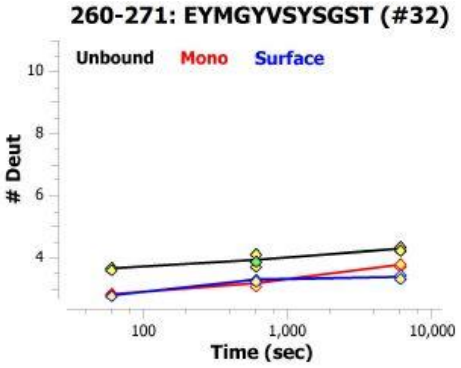

215-227: DVQLQESGPSLVK (#27)
Unbound Mono Surface
Deut
Time (sec)

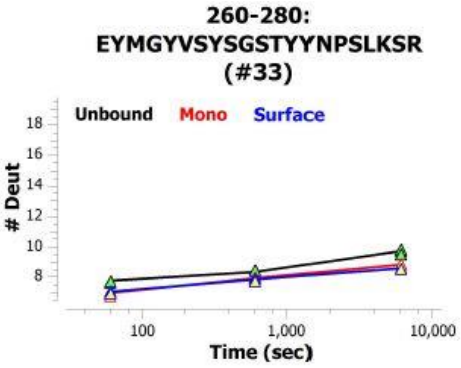

215-234: DVQLQESGPSLVKPSQTLSL (#28)
Unbound Mono Surface
Deut
Time (sec)

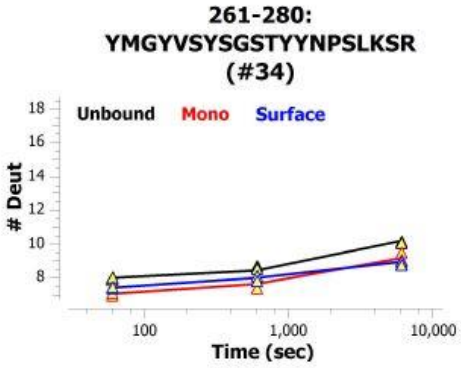

235-246: TCSVTGDSITSD (#29)
Unbound Mono Surface
Deut
Time (sec)

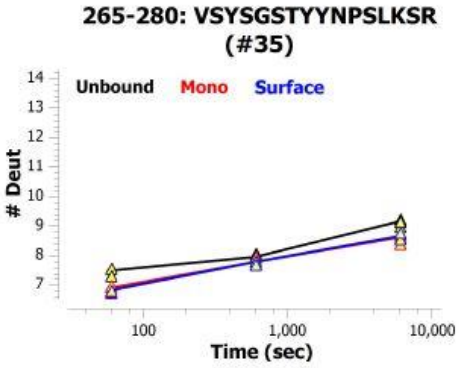

247-259: YWSWIRKFPGNRL (#30)
Unbound Mono Surface
Deut
Time (sec)

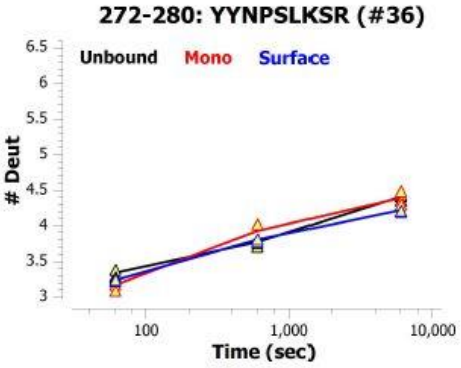

255-261: PGNRLEY (#31)
Unbound
Mono
Surface
Deut
Time (sec)

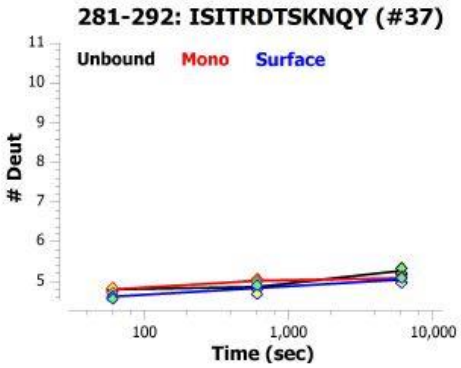

260-271: EYMGYVSYSGST (#32)
Unbound
Mono
Surface
Deut
Time (sec)

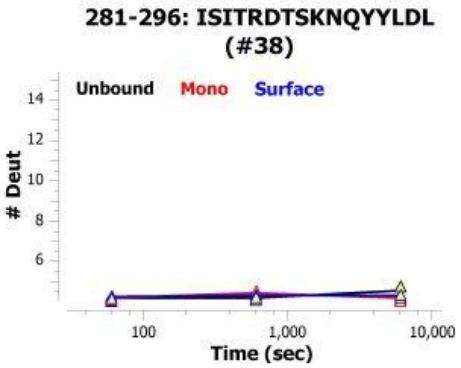

260-280: EYMGYVSYSGSTYYNPSLKSR (#33)
Unbound
Mono
Surface
Deut
Time (sec)

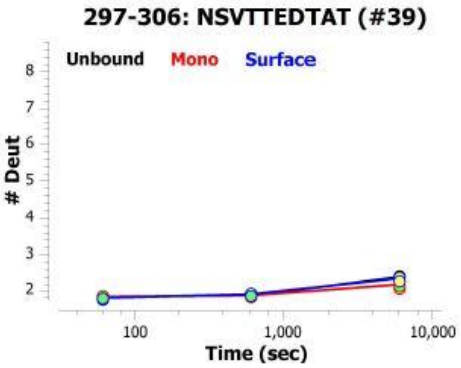

261-280: YMGYVSYSGSTYYNPSLKSR (#34)
Unbound
Mono
Surface
Deut
Time (sec)

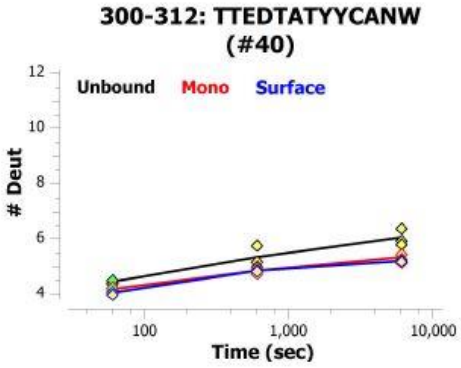

265-280: VSYSGSTYYNPSLKSR (#35)
Unbound
Mono
Surface
Deut
Time (sec)

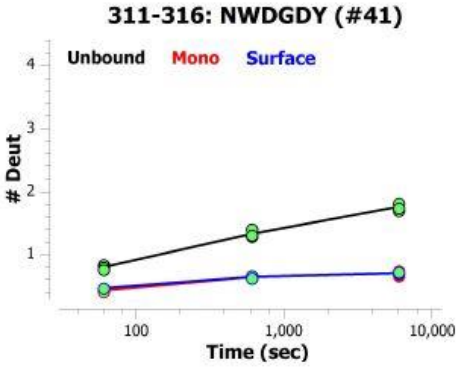

272-280: YYNPSLKSR (#36)
Unbound
Mono
Surface
Deut
Time (sec)

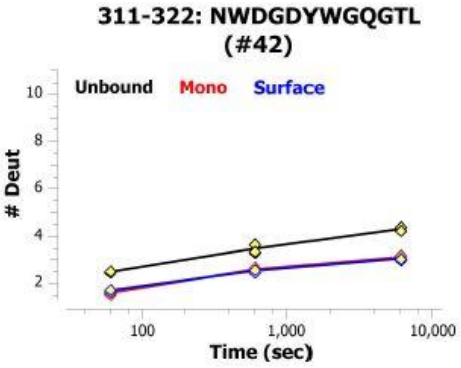

281-292: ISITRDTSKNQY (#37)
Unbound
Mono
Surface
Deut
Time (sec)

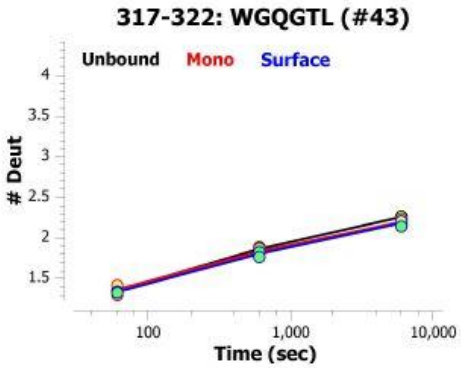

281-296: ISITRDTSKNQYYLDL (#38)
Unbound
Mono
Surface
Deut
Time (sec)

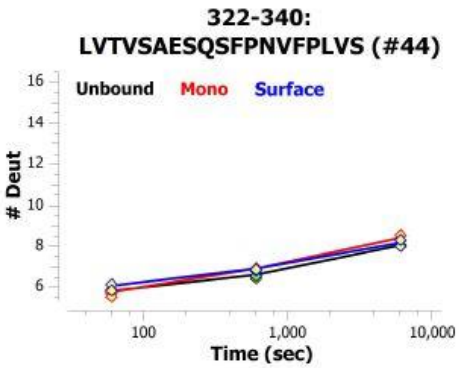

297-306: NSVTTEDTAT (#39)
Unbound
Mono
Surface
Deut
Time (sec)

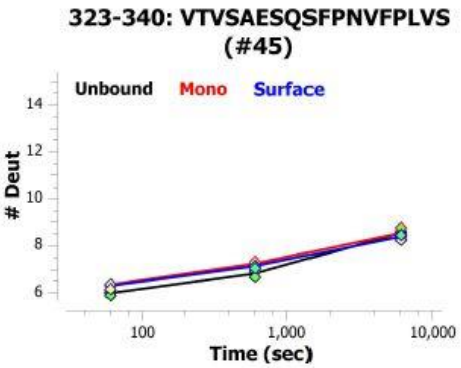

300-312: TTEDTATYYCANW (#40)
Unbound
Mono
Surface
Deut
Time (sec)


311-316: NWDGDY (#41)
Unbound
Mono
Surface
Deut
Time (sec)

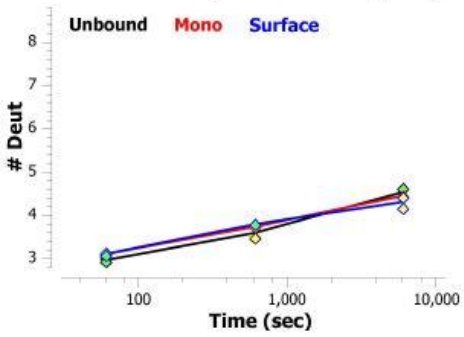

311-322: NWDGDYWGQGTL (#42)
Unbound
Mono
Surface
Deut
Time (sec)


317-322: WGQGTL (#43)
Unbound
Mono
Surface
Deut
Time (sec)

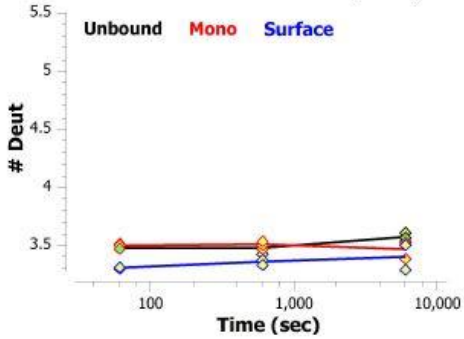

322-340: LVTVSAESQSFPNVFPLVS (#44)
Unbound
Mono
Surface
Deut
Time (sec)


323-340: VTVSAESQSFPNVFPLVS (#45)
Unbound
Mono
Surface
Deut
Time (sec)

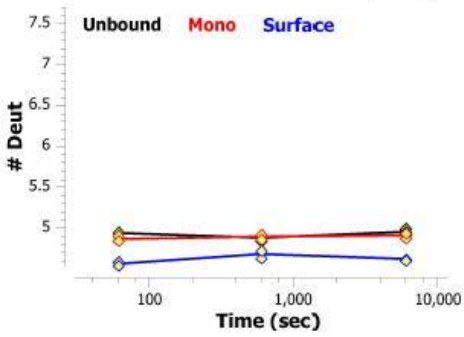
329-340: SQSFPNVFPLVS (#46)



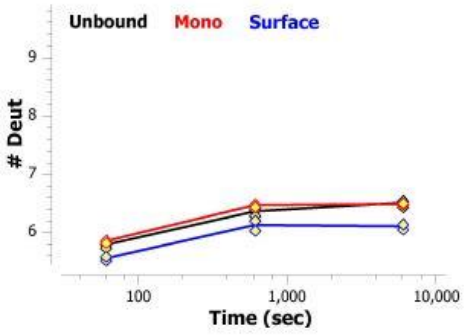
341-348: CESPLSDK (#47)



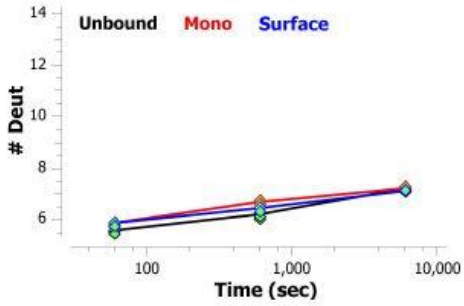
341-350: CESPLSDKNL (#48)



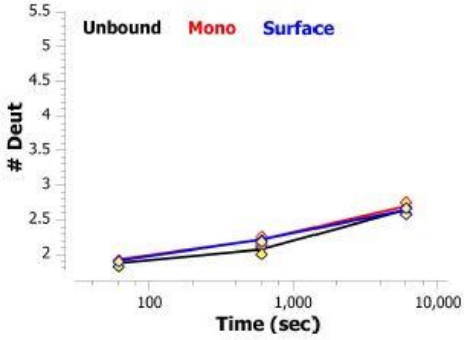
341-352: CESPLSDKNLVA (#49)



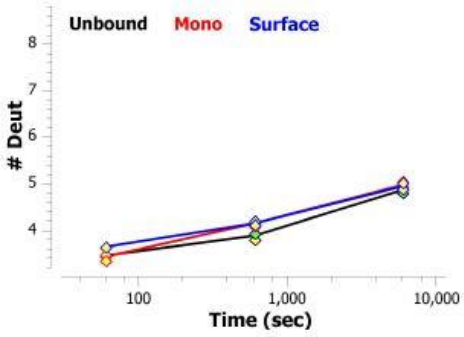
349-364: NLVAMGCLARDFLPST (#50)



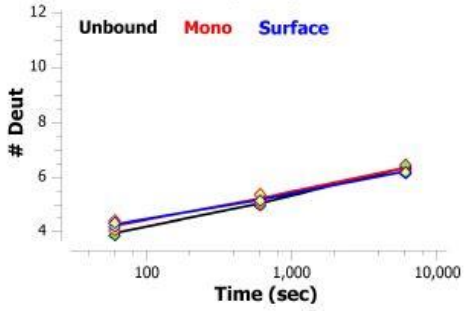
357-364: ARDFLPST (#51)



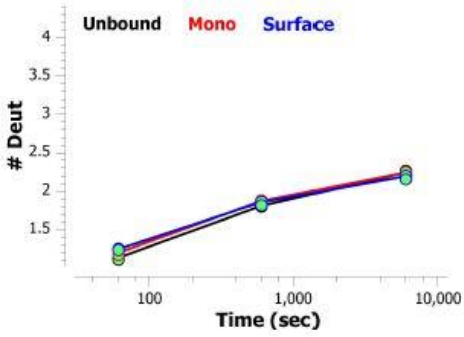
357-367: ARDFLPSTISF (#52)



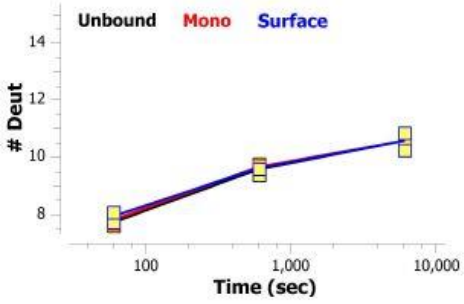
357-370: ARDFLPSTISFTWN (#53)



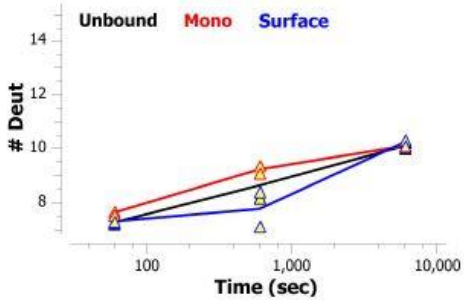
365-370: ISFTWN (#54)



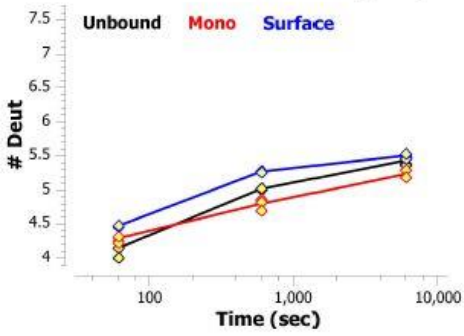
371-387: YQN*NTEVIQGIRTFPTL (#55)



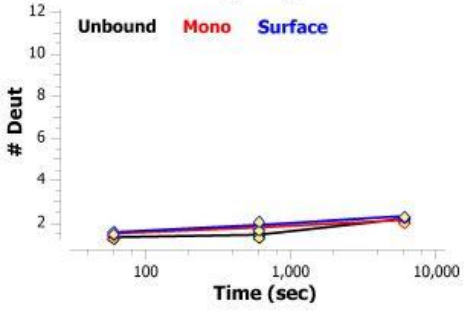
377-393: VIQGIRTFPTLRTGGKY (#56)



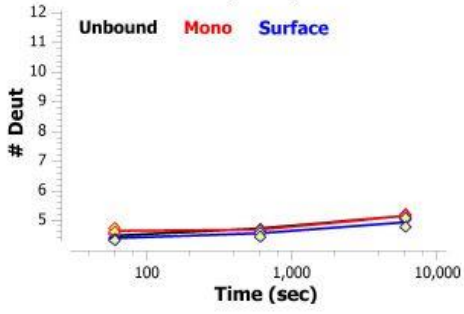
385-393: PTLRTGGKY (#57)



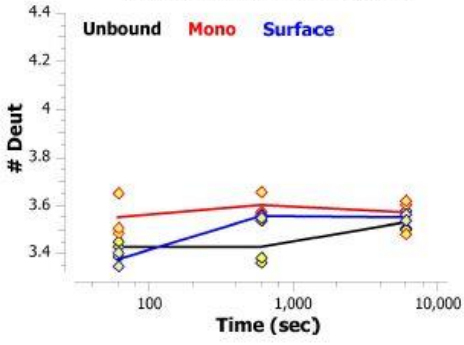
388-400: RTGGKYLATSQVL (#58)



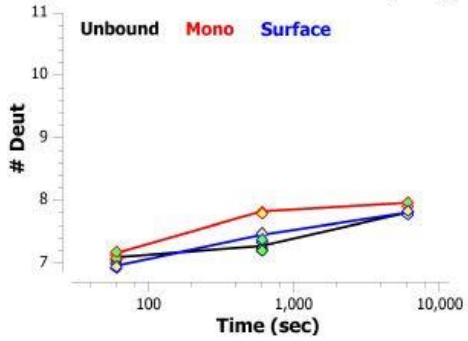
394-407: LATSQVLLSPKSIL (#59)



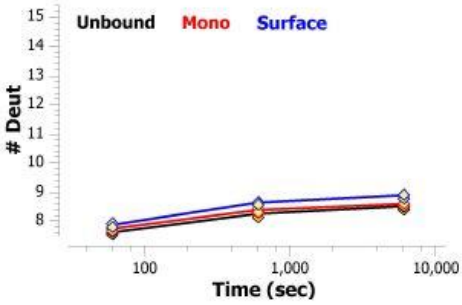
401-407: LSPKSIL (#60)



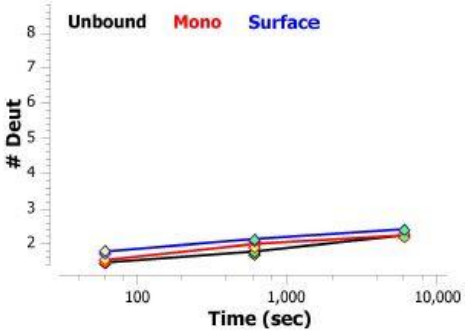
401-413: LSPKSILEGSDEY (#61)



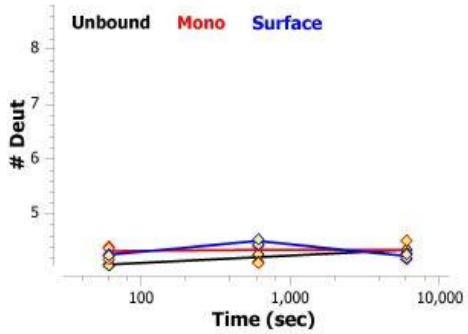
401-417: LSPKSILEGSDEYLVCK (#62)



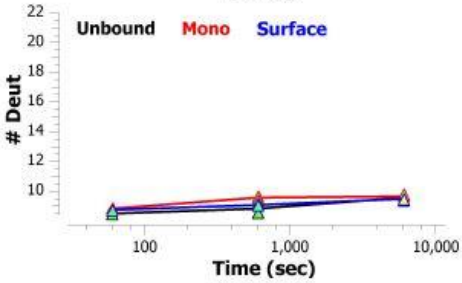
408-417: EGSDEYLVCK (#63)



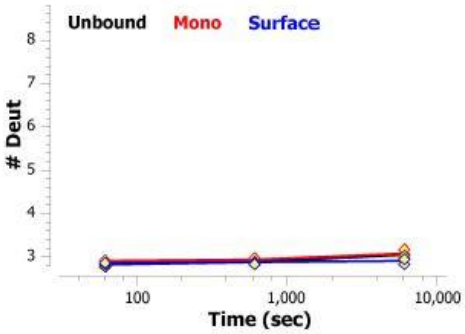
418-427: IHYGGKNRDL (#64)



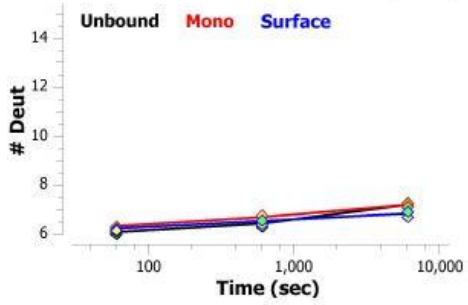
418-442: IHYGGKNRDLHVPIPAVAEMNPNVN (#65)



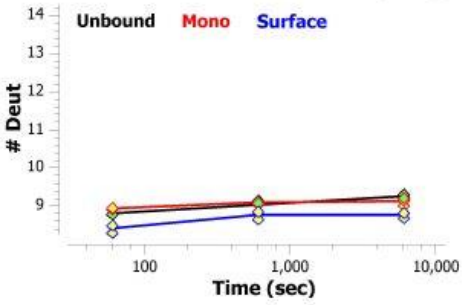
424-435: NRDLHVPIPAVA (#66)



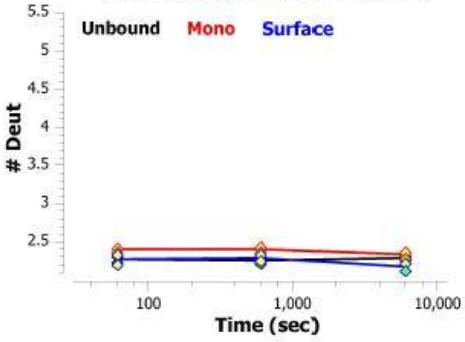
424-442: NRDLHVPIPAVAEMNPNVN (#67)



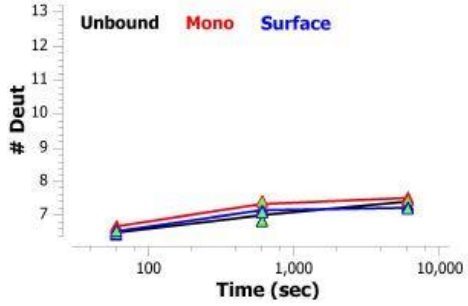
425-442: RDLHVPIPAVAEMNPNVN (#68)



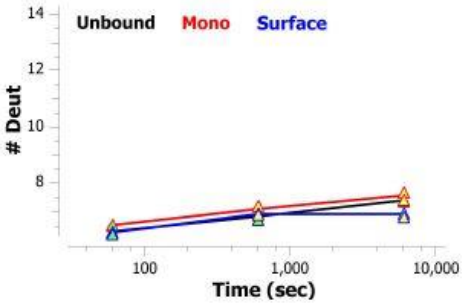
443-451: VFVPPRDGF (#69)



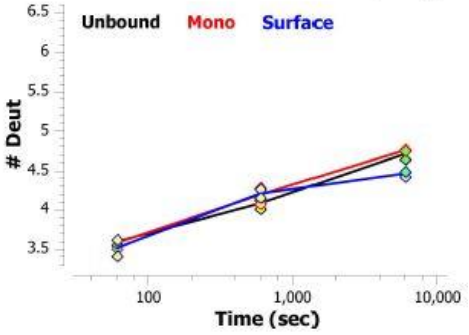
443-460: VFVPPRDGFSGPAPRKSK (#70)



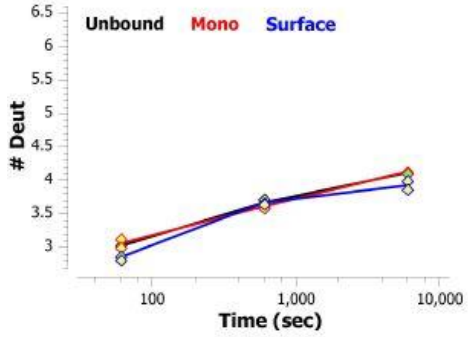
443-461: VFVPPRDGFSGPAPRKSKL (#71)



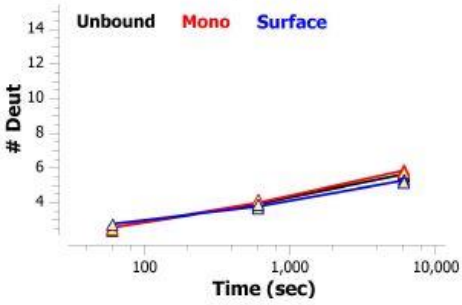
451-460: FSGPAPRKSK (#72)



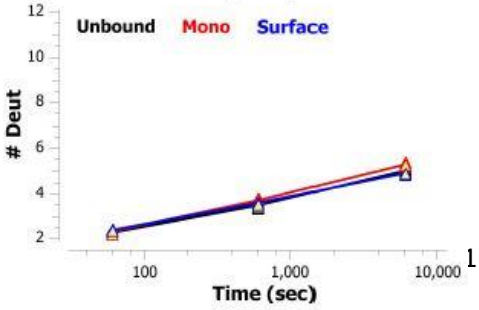
452-461: SGPAPRKSKL (#73)

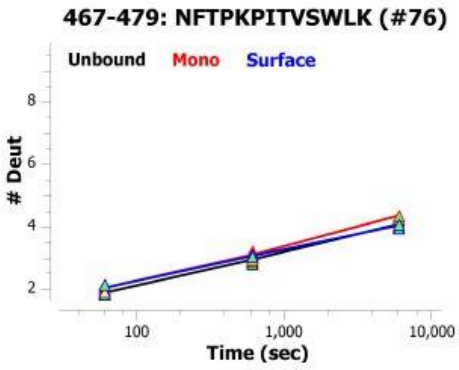


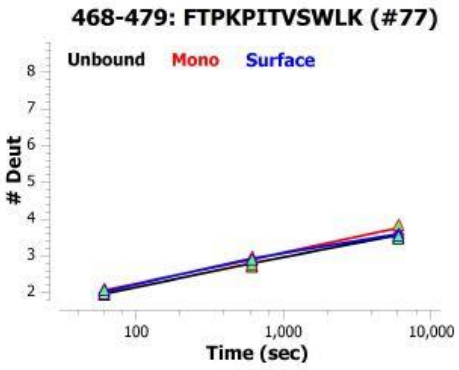
462-479: ICEATNFTPKPITVSWLK (#74)

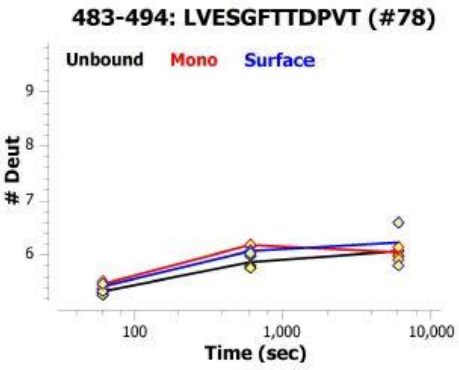


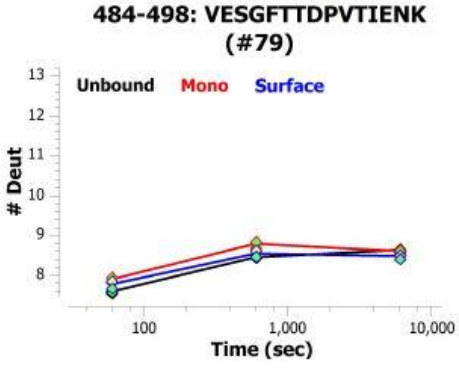
465-479: ATNFTPKPITVSWLK (#75)

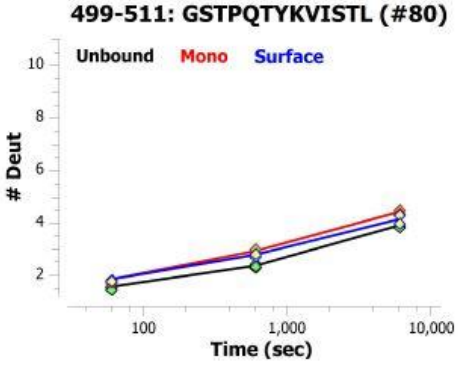


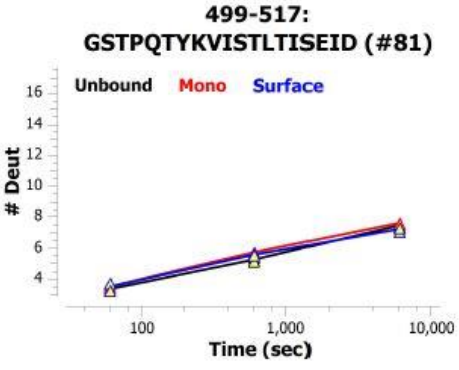

467-479: NFTPKPITVSWLK (#76)
Unbound Mono Surface
Deut
Time (sec)

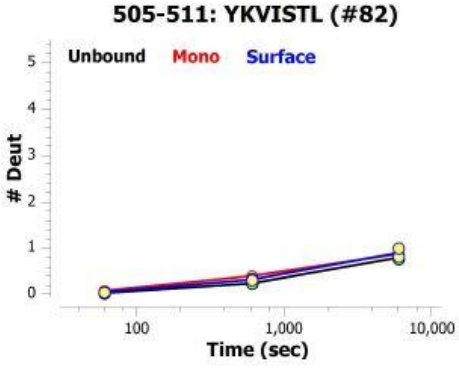

468-479: FTPKPITVSWLK (#77)
Unbound Mono Surface
Deut
Time (sec)

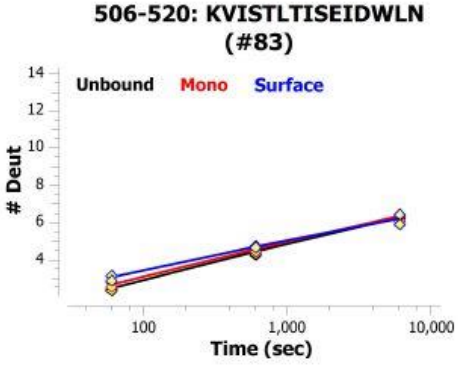

483-494: LVESGFTTDPVT (#78)
Unbound Mono Surface
Deut
Time (sec)

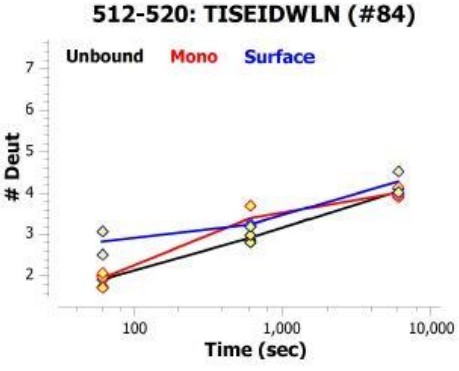

484-498: VESGFTTDPVTIENK (#79)
Unbound Mono Surface
Deut
Time (sec)

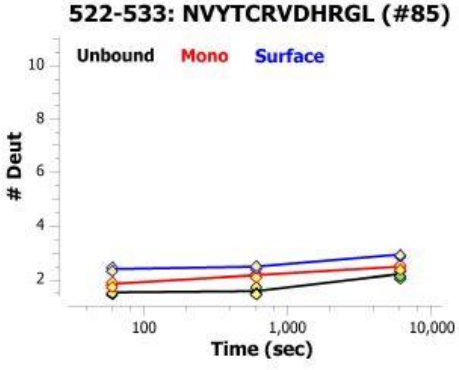

499-511: GSTPQTYKVISTL (#80)
Unbound Mono Surface
Deut
Time (sec)

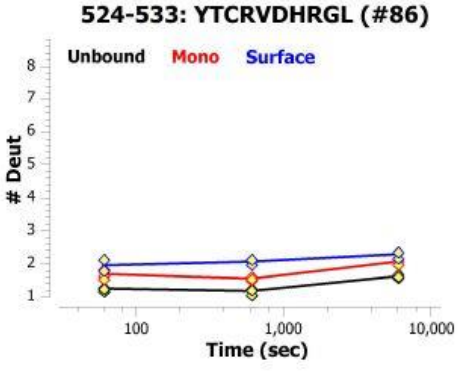

499-517: GSTPQTYKVISTLTISEID (#81)
Unbound Mono Surface
Deut
Time (sec)

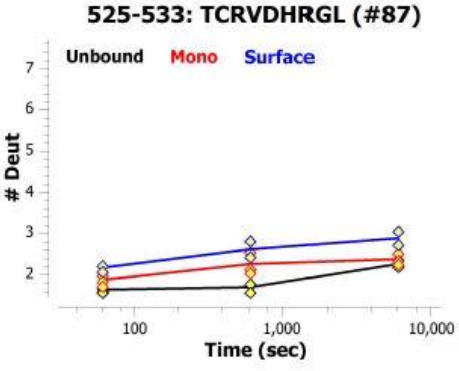

505-511: YKVISTL (#82)
Unbound Mono Surface
Deut
Time (sec)

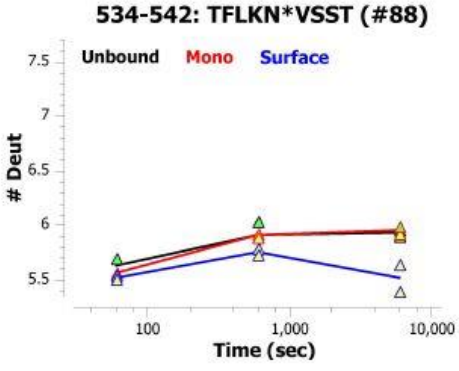

506-520: KVISTLTISEIDWLN (#83)
Unbound Mono Surface
Deut
Time (sec)

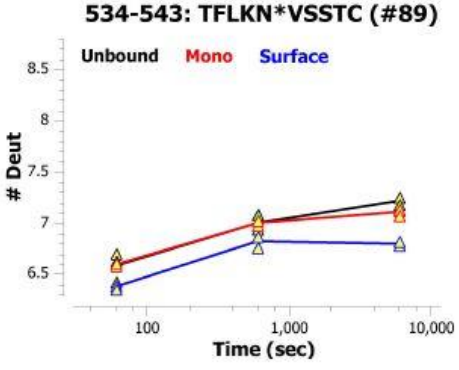

512-520: TISEIDWLN (#84)
Unbound Mono Surface
Deut
Time (sec)

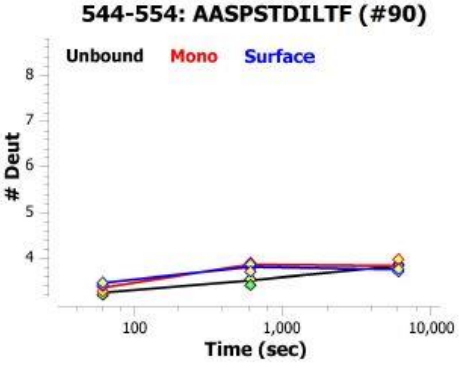

522-533: NVYTCRVDHRGL (#85)
Unbound Mono Surface
Deut
Time (sec)


524-533: YTCRVDHRGL (#86)
Unbound Mono Surface
Deut
Time (sec)

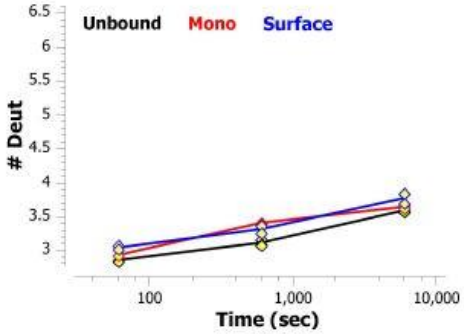

525-533: TCRVDHRGL (#87)
Unbound Mono Surface
Deut
Time (sec)


534-542: TFLKN*VSST (#88)
Unbound Mono Surface
Deut
Time (sec)

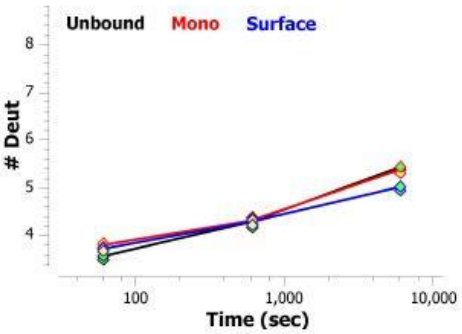

534-543: TFLKN*VSSTC (#89)
Unbound Mono Surface
Deut
Time (sec)


544-554: AASPSTDILTF (#90)
Unbound Mono Surface
Deut
Time (sec)

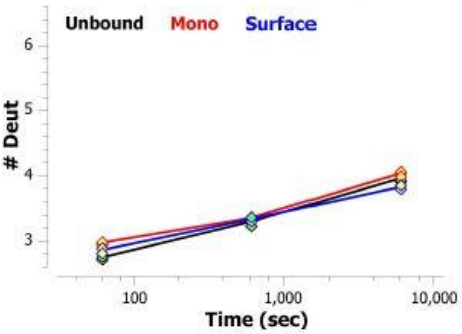
553-562: TFTIPPSFAD (#91)



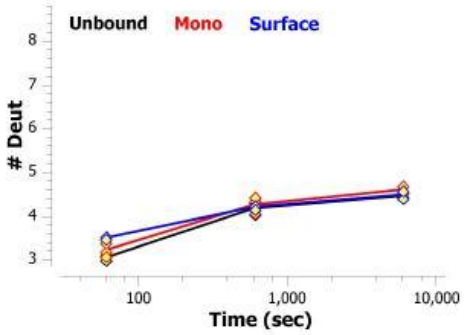
553-564: TFTIPPSFADIF (#92)



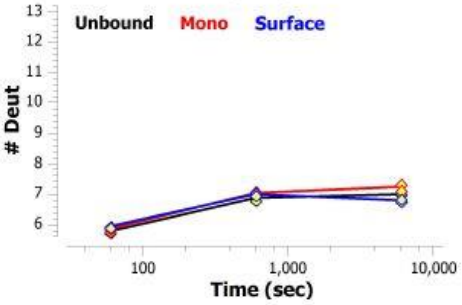
555-564: TIPPSFADIF (#93)



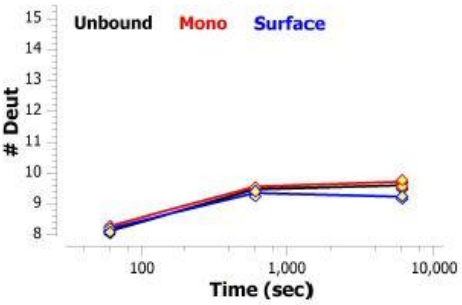
565-574: LSKSAN*LTCL (#94)



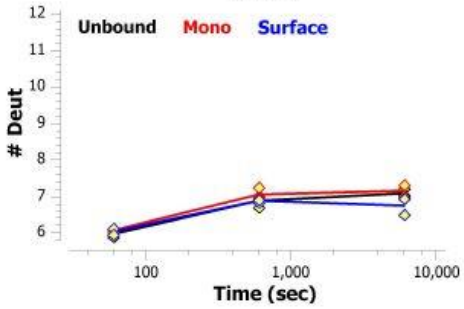
585-599: NISWASQSGEPLETK (#95)



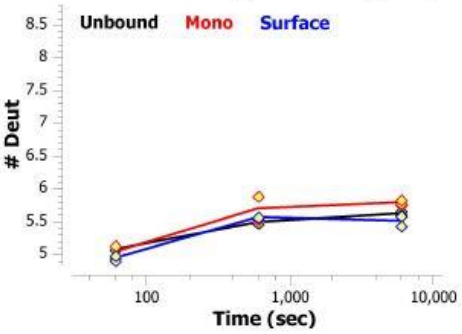
585-601: NISWASQSGEPLETKIK (#96)



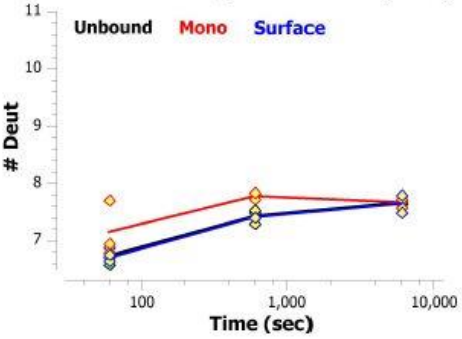
586-599: ISWASQSGEPLETK (#97)



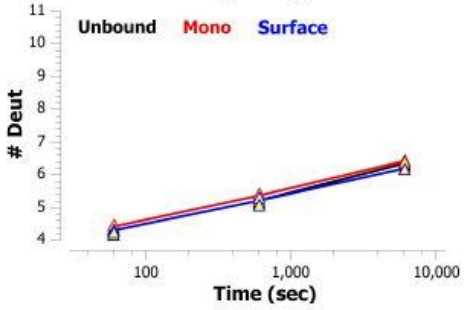
589-599: ASQSGEPLETK (#98)



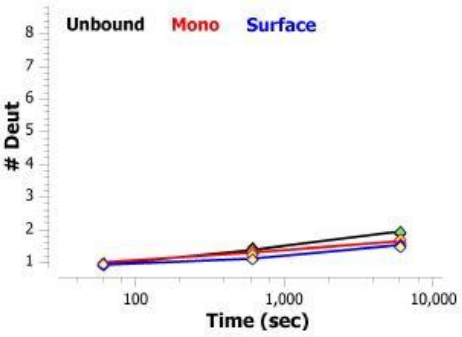
589-601: ASQSGEPLETKIK (#99)



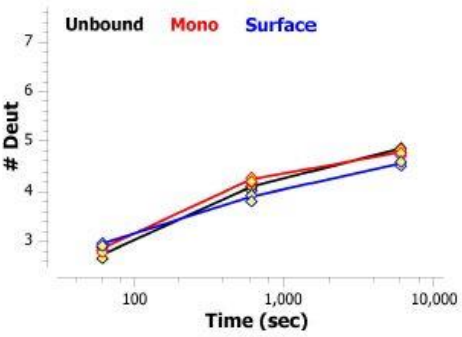
602-614: IMESHPN*GTFSAK (#100)



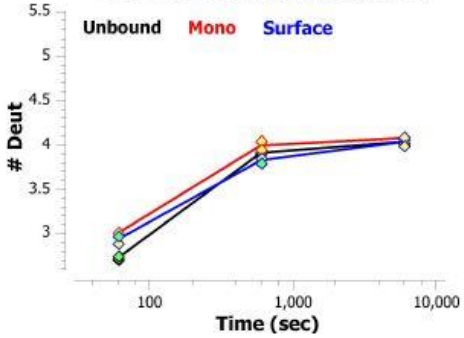
612-621: SAKGVASVCV (#101)



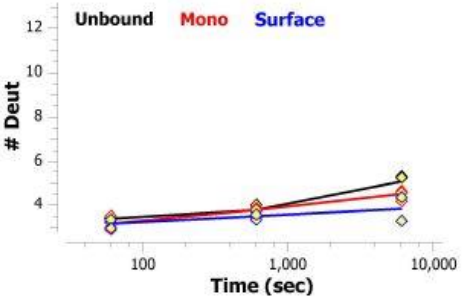
621-629: VEDWNNRKE (#102)



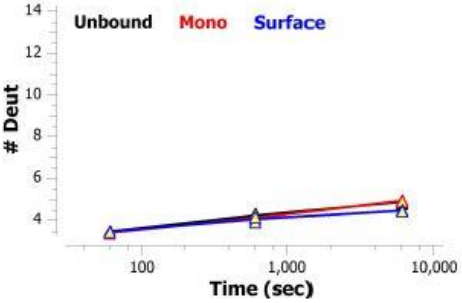
623-629: DWNNRKE (#103)



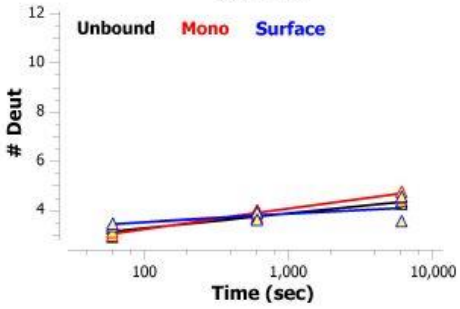
630-645: FVCTVTHRDLPSPQKK (#104)



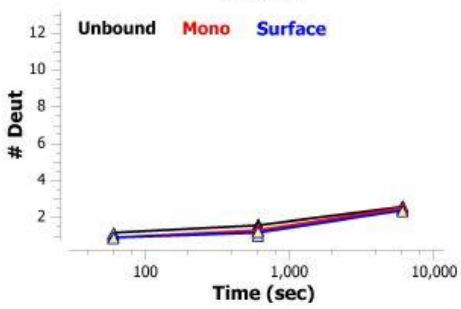
630-646: FVCTVTHRDLPSPQKKF (#105)



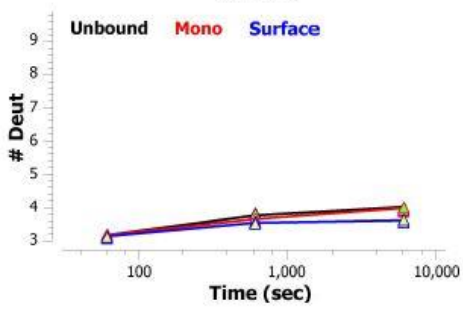
631-645: VCTVTHRDLPSPQKK (#106)



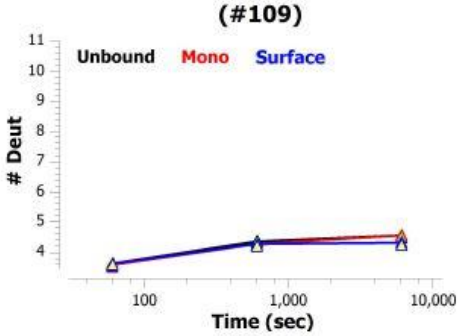
631-646: VCTVTHRDLPSPQKKF (#107)



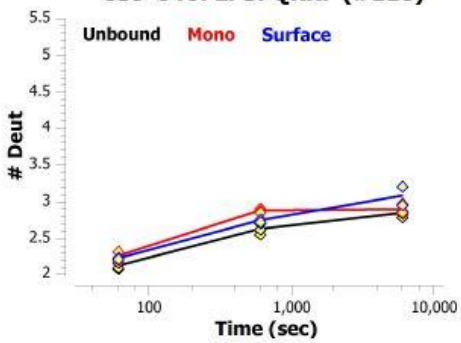
633-645: TVTHRDLPSPQKK (#108)



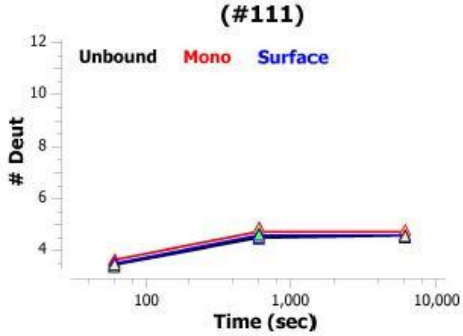
633-646: TVTHRDLPSPQKKF (#109)



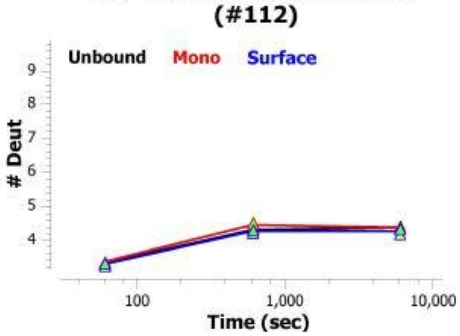
639-646: LPSPQKKF (#110)



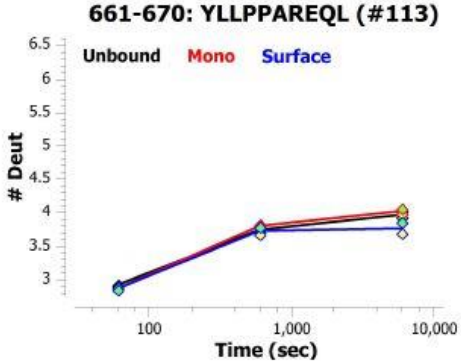
646-661: FISKPNEVHKHPPAVY (#111)



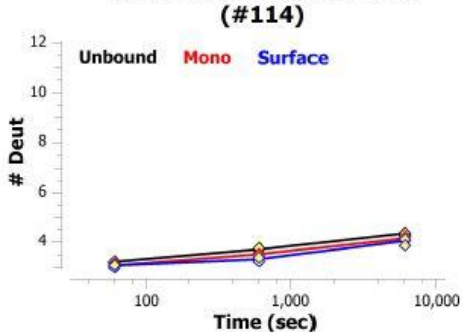
647-660: ISKPNEVHKHPPAV (#112)



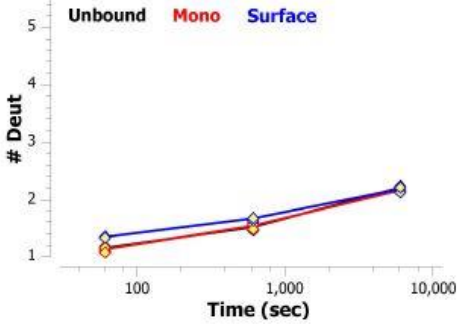
661-670: YLLPPAREQL (#113)



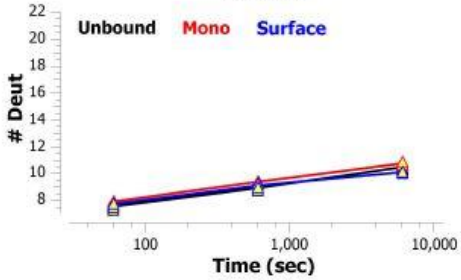
671-683: NLRESATVTCLVK (#114)



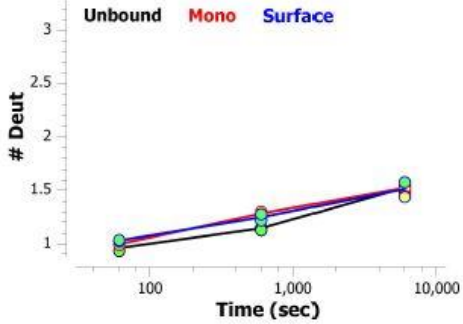
682-689: VKGFSPAD (#115)



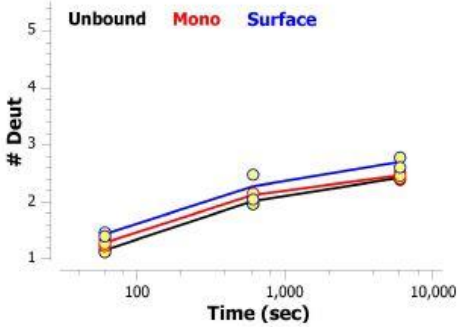
682-705: VKGFSPADISVQWLQRGQLLPQEK (#116)



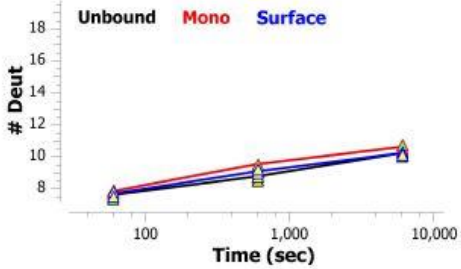
684-689: GFSPAD (#117)



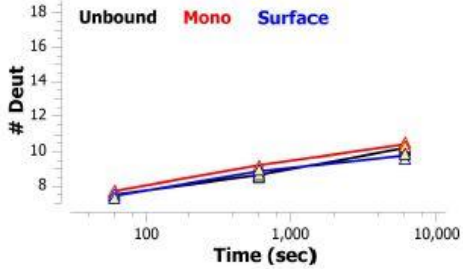
684-691: GFSPADIS (#118)



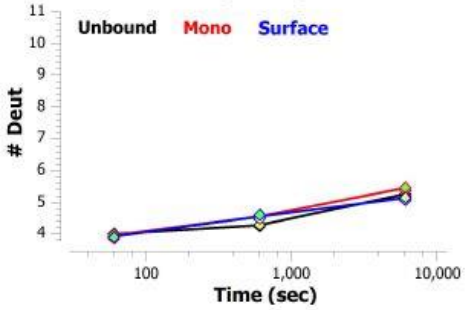
684-705: GFSPADISVQWLQRGQLLPQEK (#119)



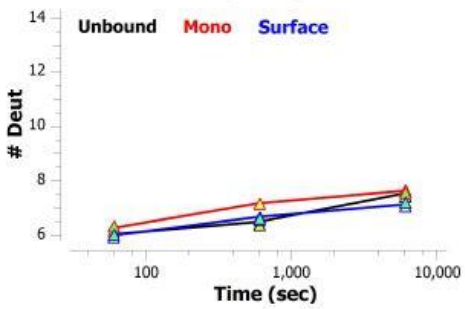
685-705: FSPADISVQWLQRGQLLPQEK (#120)



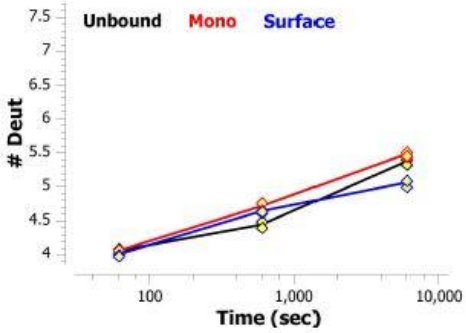
690-701: ISVQWLQRGQLL (#121)



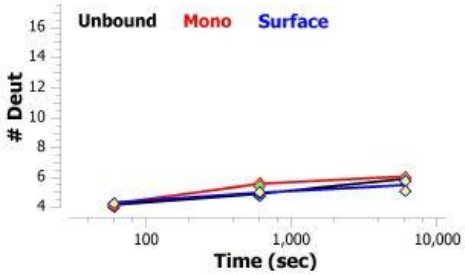
690-705: ISVQWLQRGQLLPQEK (#122)



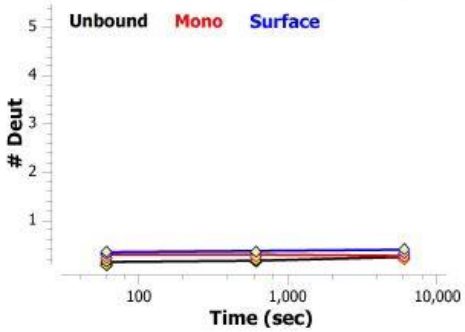
696-705: QRGQLLPQEK (#123)



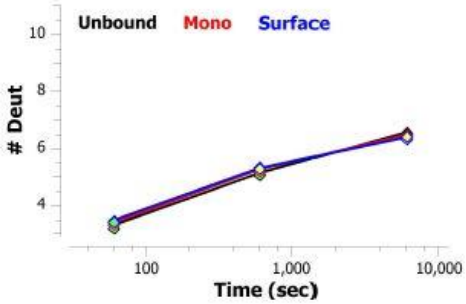
706-727: YVTSAPMPEPGAPGFYFTHSIL (#124)



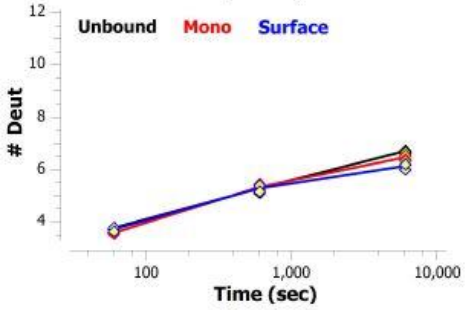
721-727: YFTHSIL (#125)



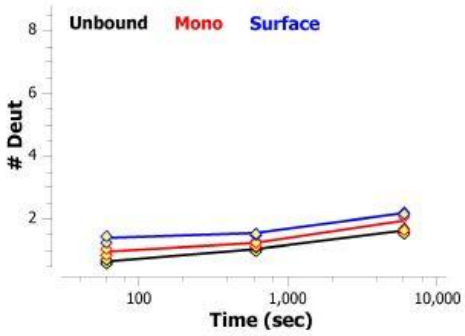
728-739: TVTEEEWNSGET (#126)



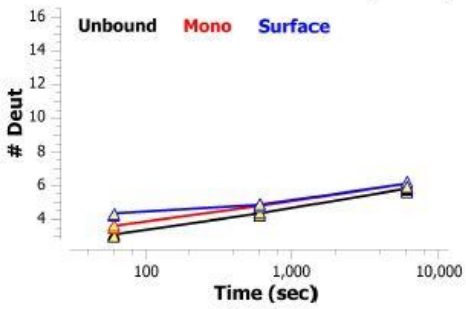
728-740: TVTEEEWNSGETY (#127)



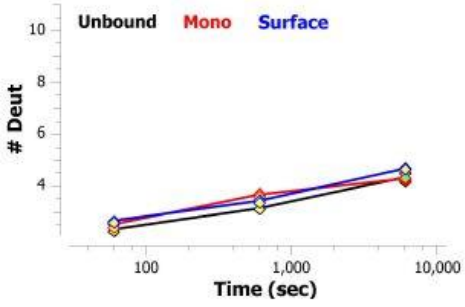
740-749: YTCVVGHEAL (#128)



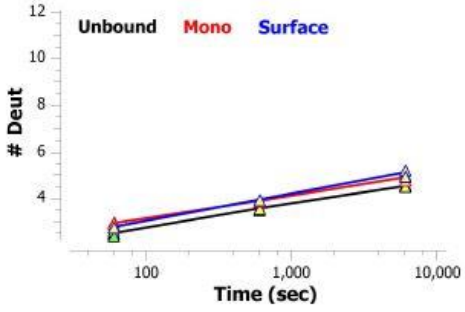
743-760: VVGHEALPHLVTERTVDK (#129)



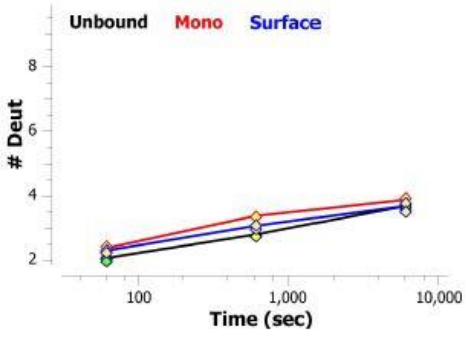
747-759: EALPHLVTERTVD (#130)



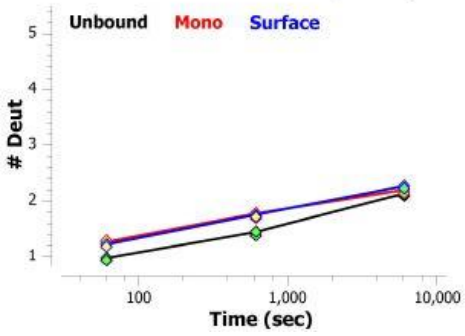
747-760: EALPHLVTERTVDK (#131)



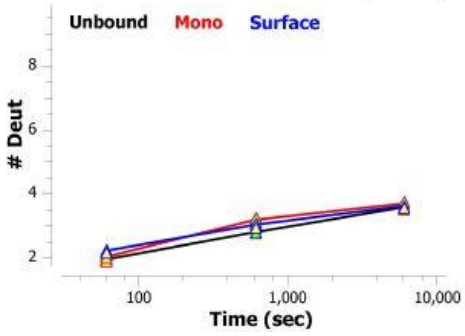
749-759: LPHLVTERTVD (#132)



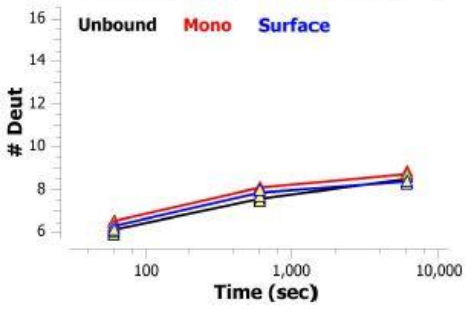
750-756: PHLVTER (#133)



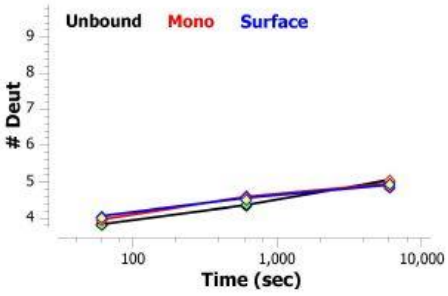
750-760: PHLVTERTVDK (#134)



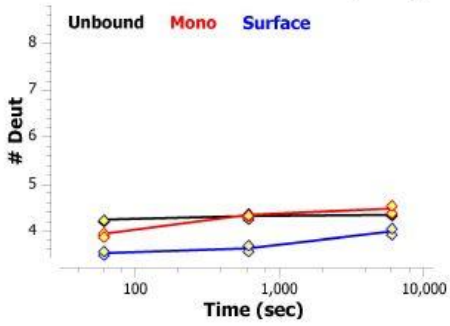
750-767: PHLVTERTVDKSTGKPTL (#135)



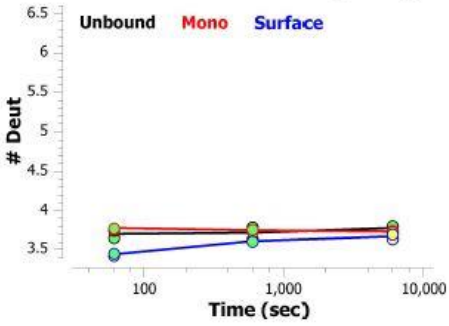
761-772: STGKPTLYN*VSL (#136)



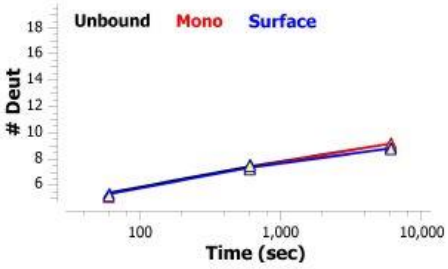
773-782: IMSDTGGTCY (#137)



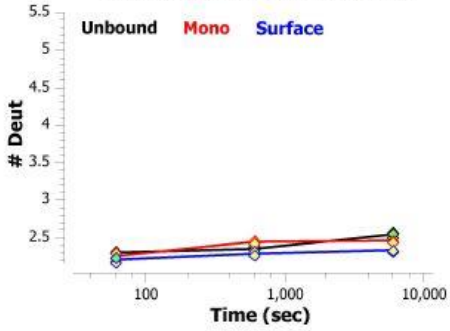
775-782: SDTGGTCY (#138)



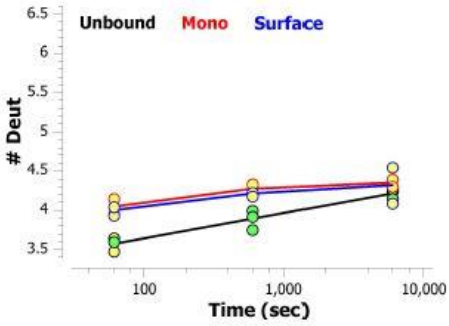
821-842: VTSRIIPSTEDPNEDIVERNIR (#139)



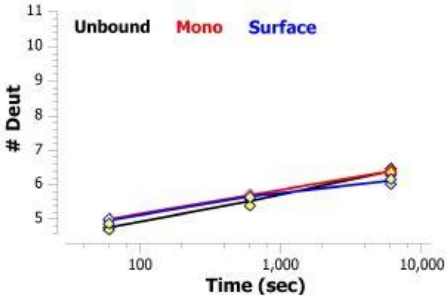
864-870: FVYHLSD (#140)



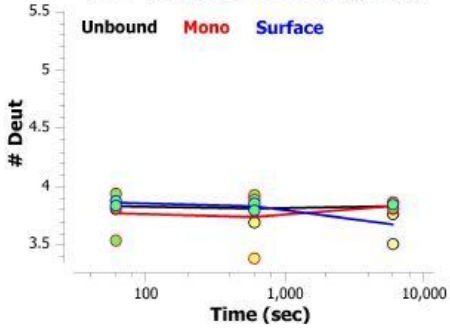
882-889: LEDQVVTA (#141)



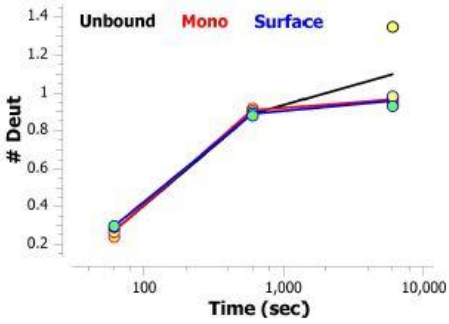
928-941: MVQAALTPDSCYPD (#142)



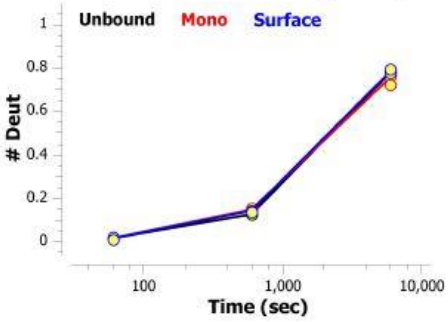
953-961: RPPGFSPFR (#143)



970-972: Y*W*R* (#144)



988-990: A*W*D* (#145)

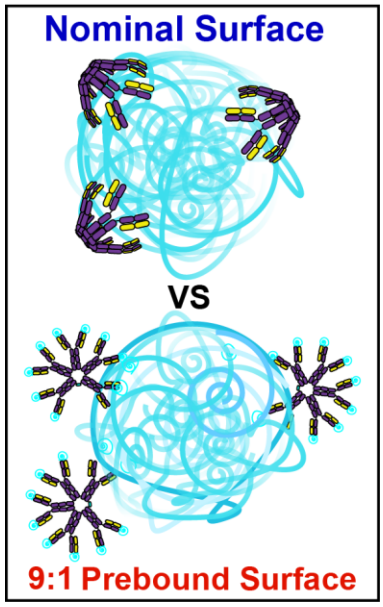


Supplementary Table S4.3 | Nominal vs Prebound Surface HDX - Experiment Statistics

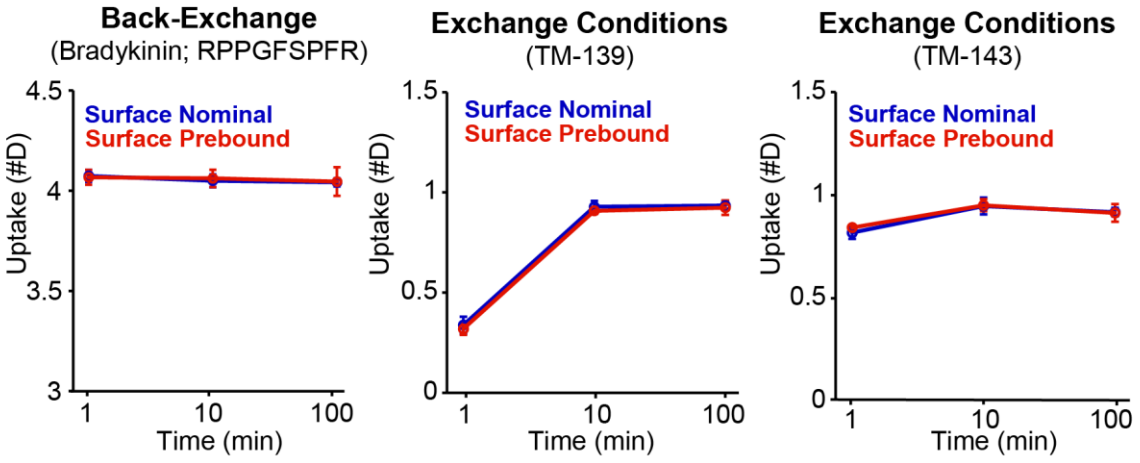
Data Set	Nominal Surface (0:1; Max-Avid)	Prebound Surface (~9:1; Min-Avid)
HDX reaction details	pH 7.3, 22°C, 85% D2O	
HDX time course	1 min, 10 min, 100 min	
HDX controls	Undeuterated	
Back-exchange (mean) bradykinin peptide internal standard	18.91%	18.85%
Number of peptides	**141 peptides**	
Sequence coverage	**88.9% Total Sequence Coverage; LC ~95%, HC ~97%, JC ~45%)**	
Average peptide length / Redundancy	12.9 +/- 5.0 residues	
	redundancy: 1.9 peptides/residue	
Replicates (biological or technical)	quadruplicate biological replicates	
Repeatability	**0.090 (avg standard deviation)**	**0.072 (avg standard deviation)**
Significant differences in HDX	**0.1681 D (95% CI)***	

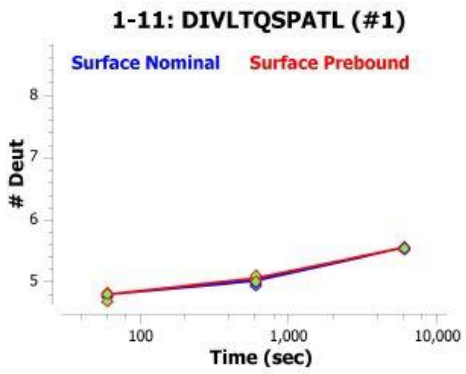
*This calculation is done by HDExaminer and does not include occupied YETL glycopep

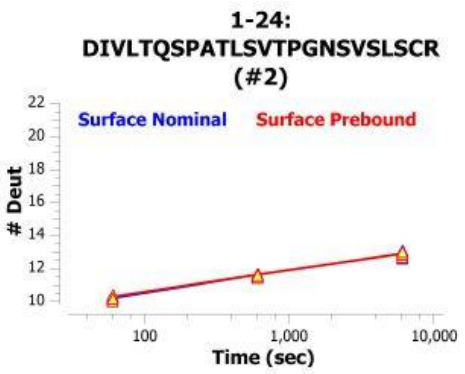
Sample Conditions

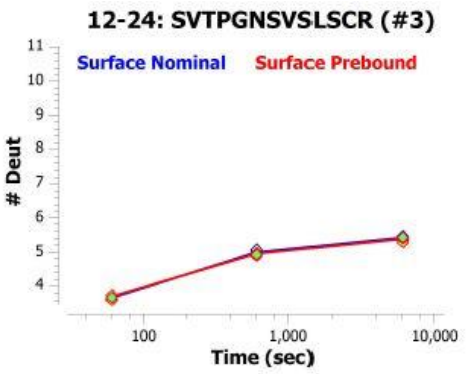


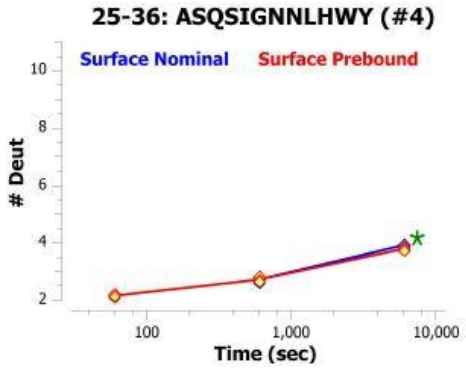
Internal Standards

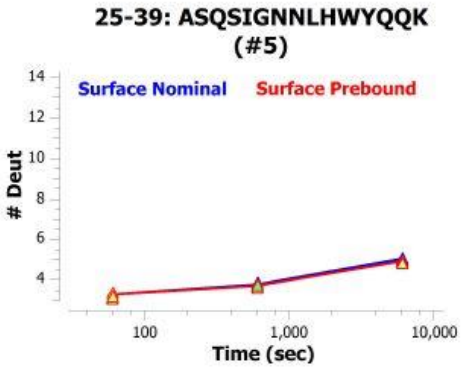


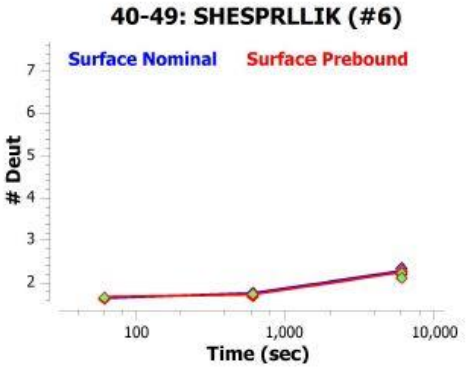

1-11: DIVLTQSPATL (#1)
Surface Nominal
Surface Prebound
Deut
Time (sec)

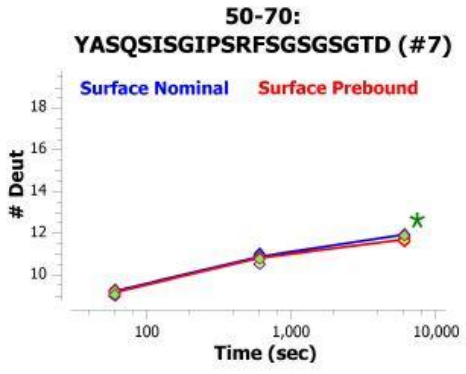

1-24: DIVLTQSPATLSVTPGNSVSLSCR (#2)
Surface Nominal
Surface Prebound
Deut
Time (sec)

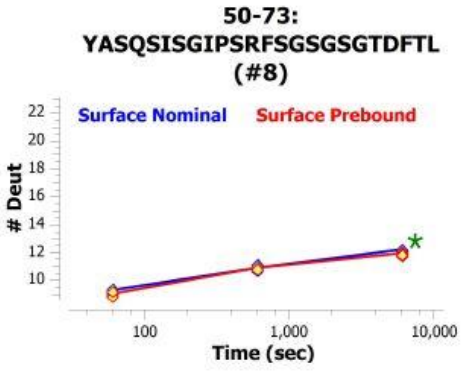

12-24: SVTPGNSVSLSCR (#3)
Surface Nominal
Surface Prebound
Deut
Time (sec)

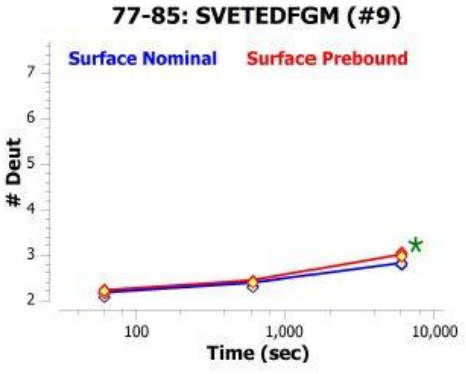

25-36: ASQSIGNNLHWY (#4)
Surface Nominal
Surface Prebound
Deut
Time (sec)

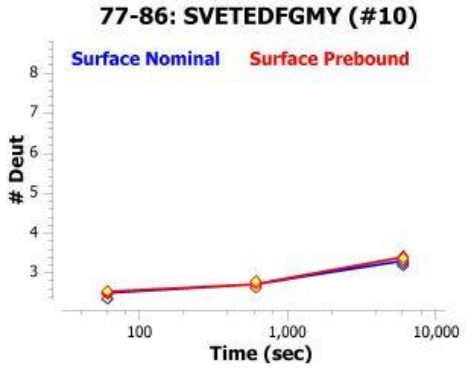

25-39: ASQSIGNNLHWYQQK (#5)
Surface Nominal
Surface Prebound
Deut
Time (sec)

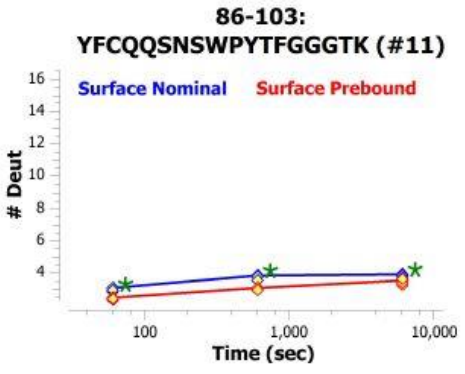

40-49: SHESPRLLIK (#6)
Surface Nominal
Surface Prebound
Deut
Time (sec)

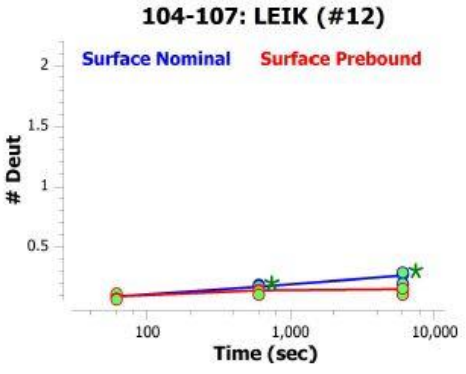

50-70: YASQSISGIPSRFSGSGSGTD (#7)
Surface Nominal
Surface Prebound
Deut
Time (sec)

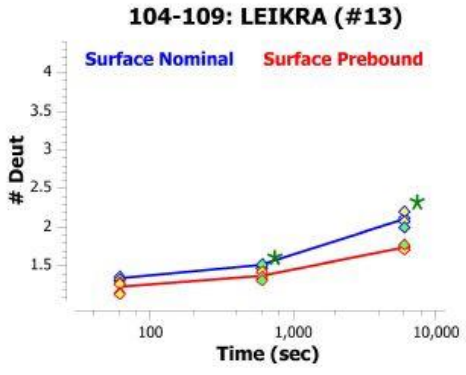

50-73: YASQSISGIPSRFSGSGSGTDFTL (#8)
Surface Nominal
Surface Prebound
Deut
Time (sec)

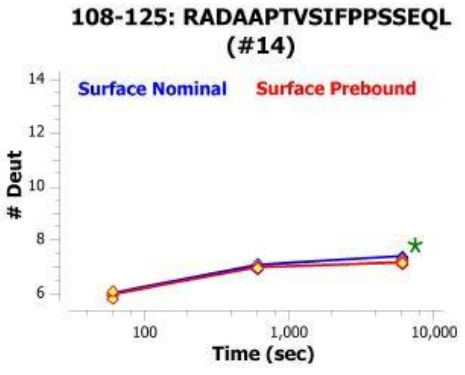

77-85: SVETEDFGM (#9)
Surface Nominal
Surface Prebound
Deut
Time (sec)

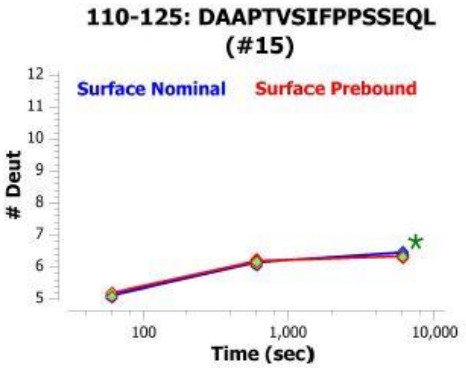

77-86: SVETEDFGMY (#10)
Surface Nominal
Surface Prebound
Deut
Time (sec)

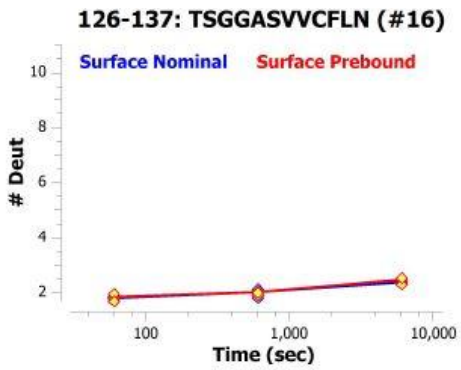

86-103: YFCQQSNSWPYTFGGGTK (#11)
Surface Nominal
Surface Prebound
Deut
Time (sec)

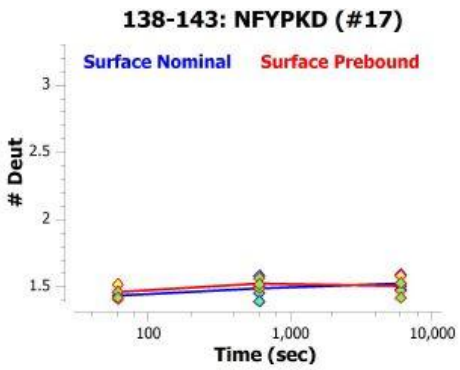

104-107: LEIK (#12)
Surface Nominal
Surface Prebound
Deut
Time (sec)

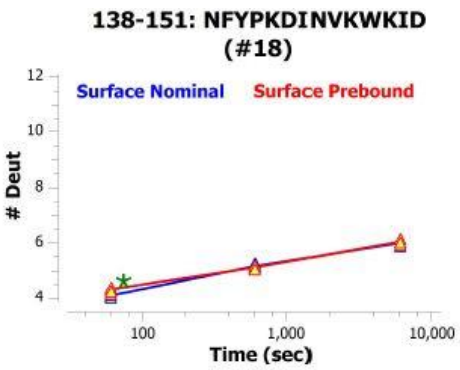

104-109: LEIKRA (#13)
Surface Nominal
Surface Prebound
Deut
Time (sec)

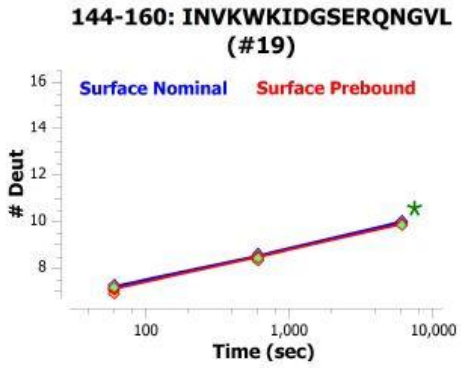

108-125: RADAAPTVSIFPPSSEQL (#14)
Surface Nominal
Surface Prebound
Deut
Time (sec)

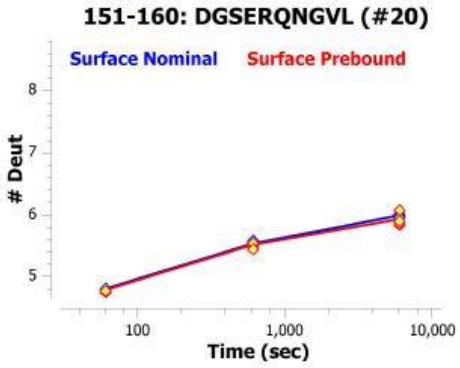

110-125: DAAPTVSIFPPSSEQL (#15)
Surface Nominal
Surface Prebound
Deut
Time (sec)

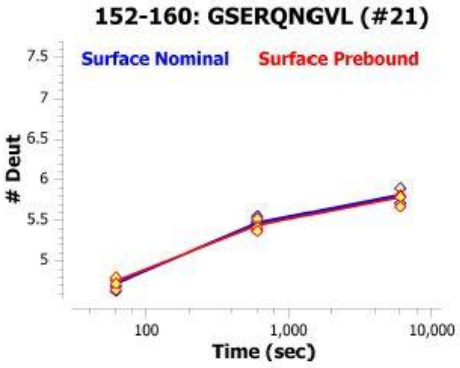

126-137: TSGGASVVCFLN (#16)
Surface Nominal
Surface Prebound
Deut
Time (sec)

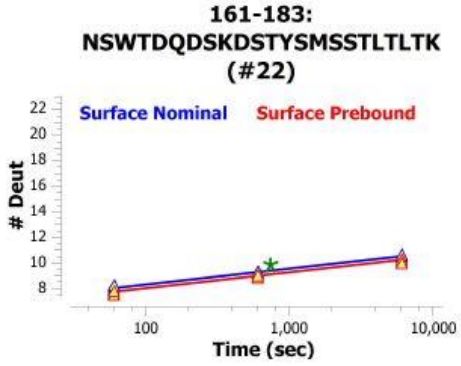

138-143: NFYPKD (#17)
Surface Nominal
Surface Prebound
Deut
Time (sec)

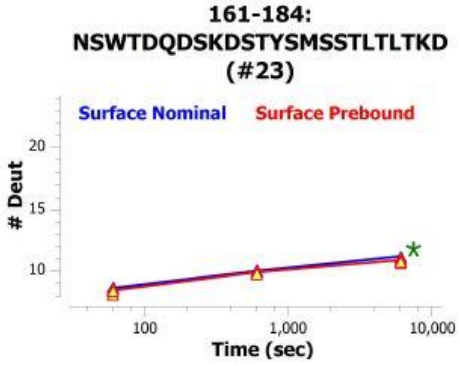

138-151: NFYPKDINVKWKID (#18)
Surface Nominal
Surface Prebound
Deut
Time (sec)

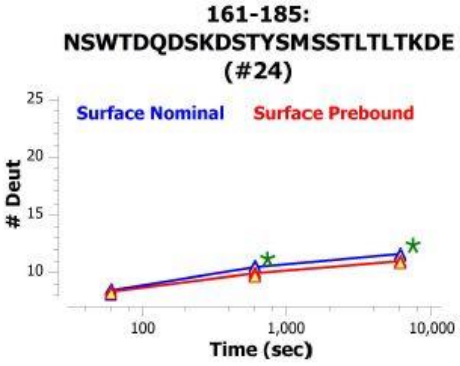

144-160: INVKWKIDGSERQNGVL (#19)
Surface Nominal
Surface Prebound
Deut
Time (sec)

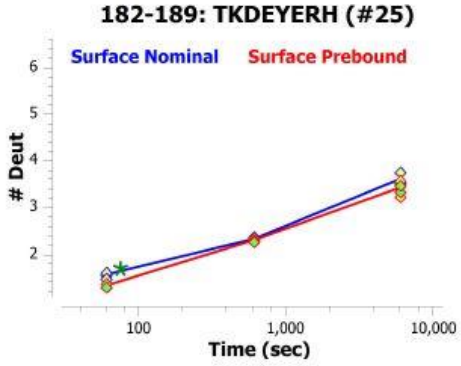

151-160: DGSERQNGVL (#20)
Surface Nominal
Surface Prebound
Deut
Time (sec)

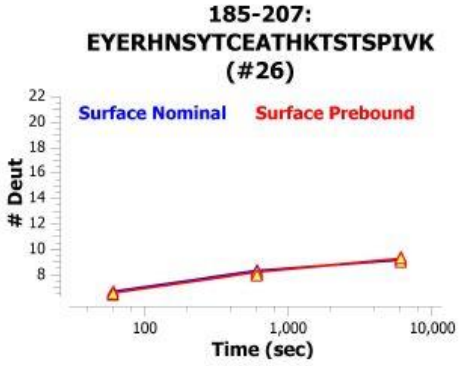

152-160: GSERQNGVL (#21)
Surface Nominal
Surface Prebound
Deut
Time (sec)

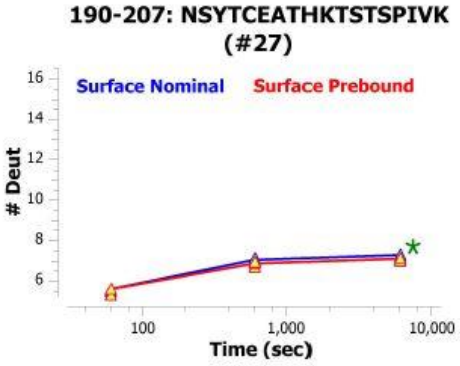

161-183: NSWTDQDSKDSTYSMSSTLTLTK (#22)
Surface Nominal
Surface Prebound
Deut
Time (sec)

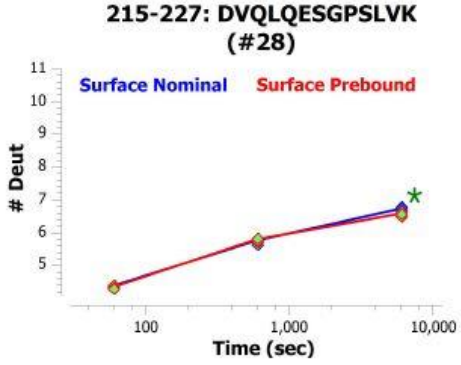

161-184: NSWTDQDSKDSTYSMSSTLTLTKD (#23)
Surface Nominal
Surface Prebound
Deut
Time (sec)

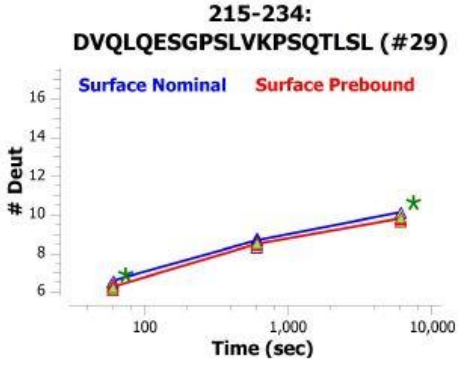

161-185: NSWTDQDSKDSTYSMSSTLTLTKDE (#24)
Surface Nominal
Surface Prebound
Deut
Time (sec)

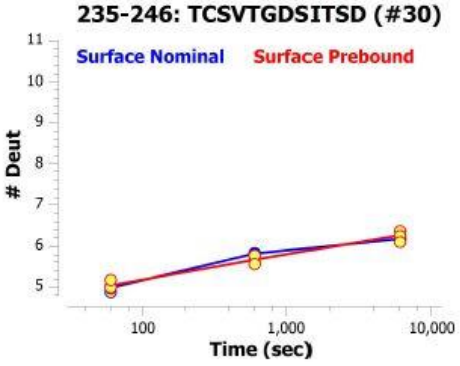

182-189: TKDEYERH (#25)
Surface Nominal
Surface Prebound
Deut
Time (sec)

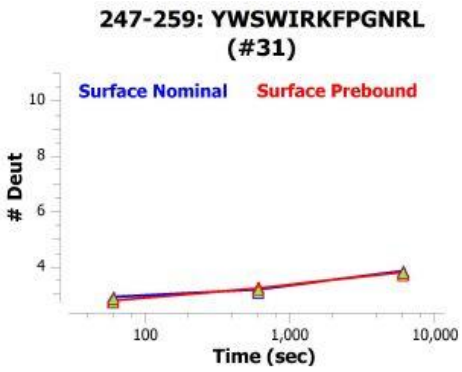

185-207: EYERHNSYTCEATHKTSTSPIVK (#26)
Surface Nominal
Surface Prebound
Deut
Time (sec)

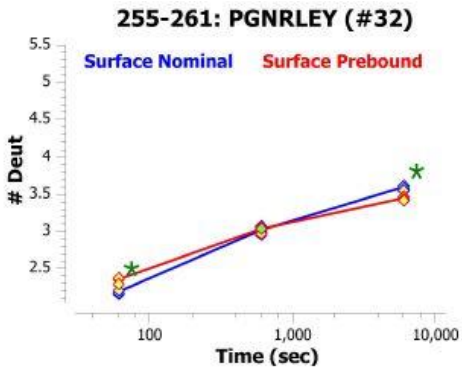

190-207: NSYTCEATHKTSTSPIVK (#27)
Surface Nominal
Surface Prebound
Deut
Time (sec)

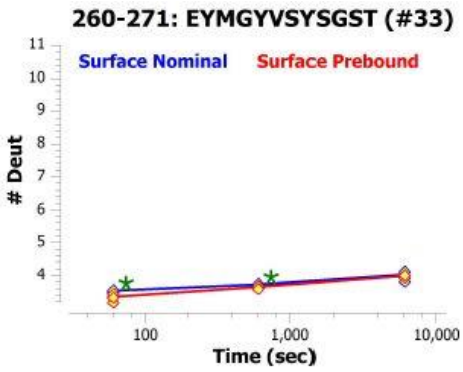

215-227: DVQLQESGPSLVK (#28)
Surface Nominal
Surface Prebound
Deut
Time (sec)

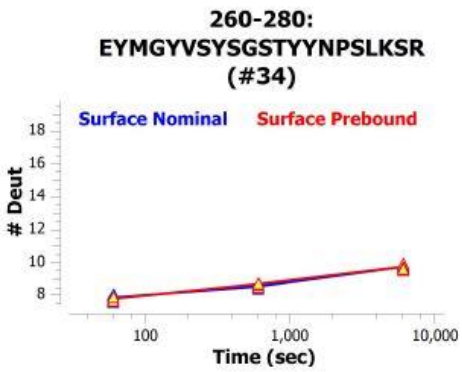

215-234: DVQLQESGPSLVKPSQTLSL (#29)
Surface Nominal
Surface Prebound
Deut
Time (sec)

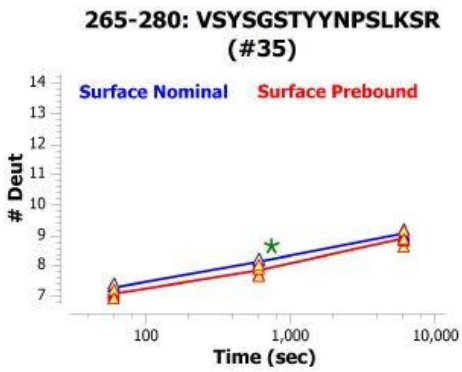

235-246: TCSVTGDSITSD (#30)
Surface Nominal
Surface Prebound
Deut
Time (sec)

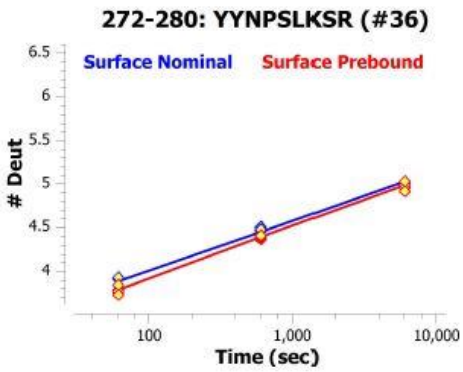

247-259: YWSWIRKFPGNRL (#31)
Surface Nominal
Surface Prebound
Deut
Time (sec)

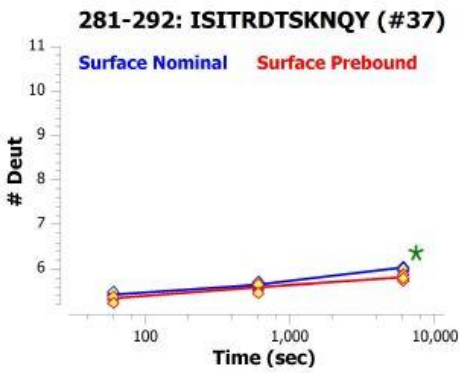

255-261: PGNRLEY (#32)
Surface Nominal
Surface Prebound
Deut
Time (sec)

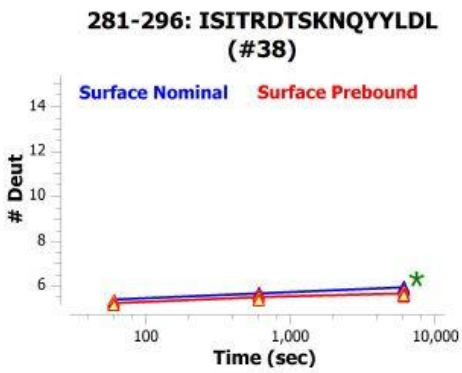

260-271: EYMGYVSYSGST (#33)
Surface Nominal
Surface Prebound
Deut
Time (sec)

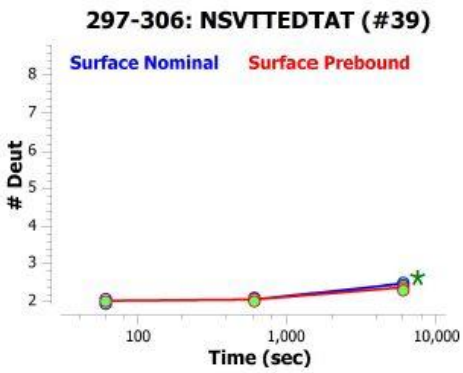

260-280: EYMGYVSYSGSTYYNPSLKSR (#34)
Surface Nominal
Surface Prebound
Deut
Time (sec)

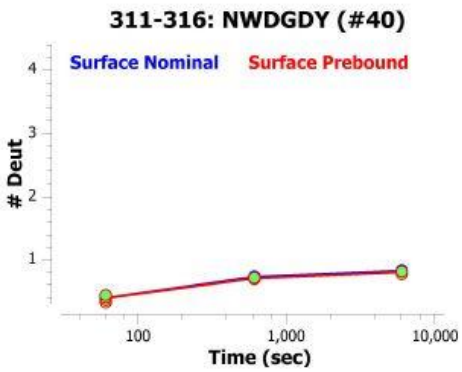

265-280: VSYSGSTYYNPSLKSR (#35)
Surface Nominal
Surface Prebound
Deut
Time (sec)

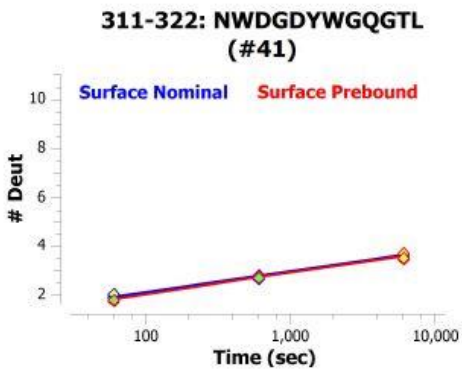

272-280: YYNPSLKSR (#36)
Surface Nominal
Surface Prebound
Deut
Time (sec)

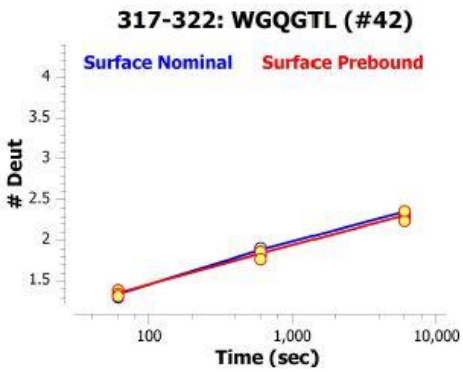

281-292: ISITRDTSKNQY (#37)
Surface Nominal
Surface Prebound
Deut
Time (sec)

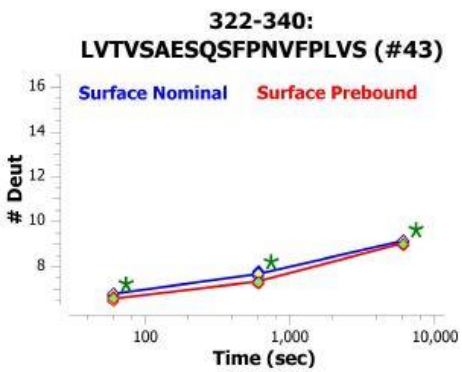

281-296: ISITRDTSKNQYYLDL (#38)
Surface Nominal
Surface Prebound
Deut
Time (sec)

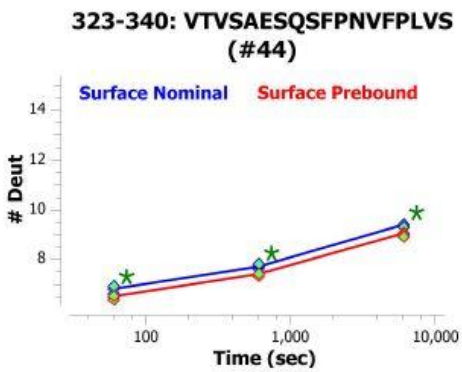

297-306: NSVTTEDTAT (#39)
Surface Nominal
Surface Prebound
Deut
Time (sec)

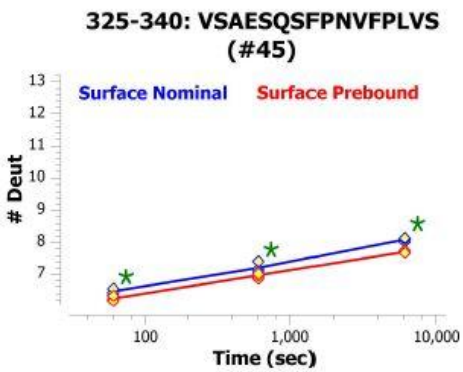

311-316: NWDGDY (#40)
Surface Nominal
Surface Prebound
Deut
Time (sec)


311-322: NWDGDYWGQGTL (#41)
Surface Nominal
Surface Prebound
Deut
Time (sec)

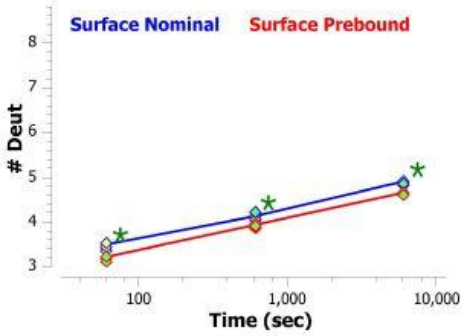

317-322: WGQGTL (#42)
Surface Nominal
Surface Prebound
Deut
Time (sec)


322-340: LVTVSAESQSFPNVFPLVS (#43)
Surface Nominal
Surface Prebound
Deut
Time (sec)

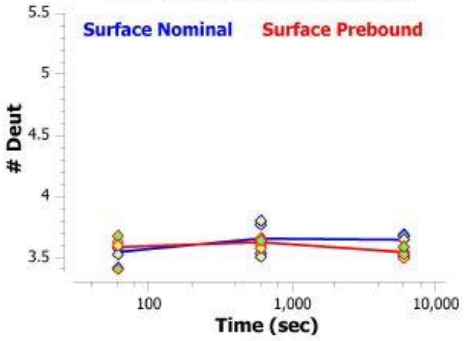

323-340: VTVSAESQSFPNVFPLVS (#44)
Surface Nominal
Surface Prebound
Deut
Time (sec)


325-340: VSAESQSFPNVFPLVS (#45)
Surface Nominal
Surface Prebound
Deut
Time (sec)

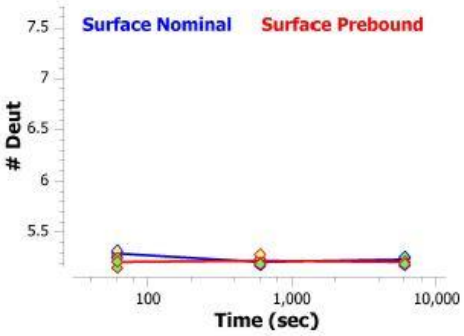
329-340: SQSFPNVFPLVS (#46)



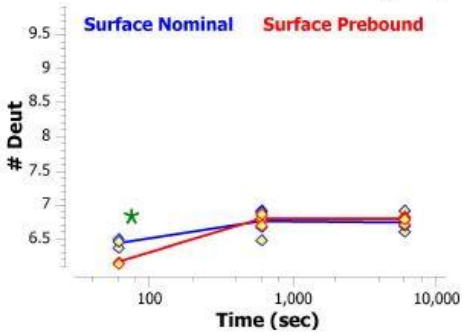
341-348: CESPLSDK (#47)



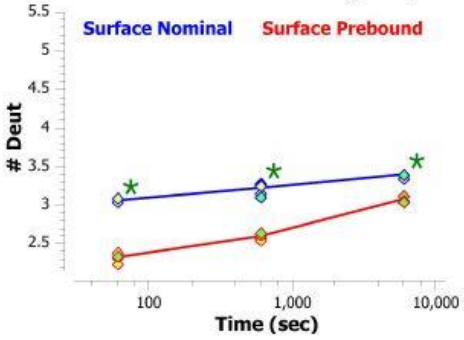
341-350: CESPLSDKNL (#48)



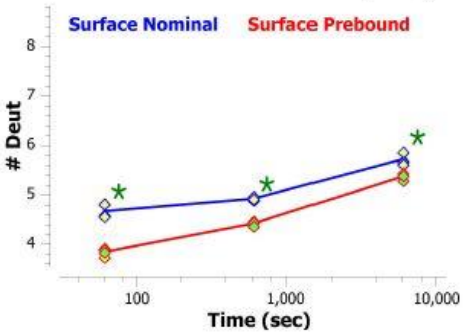
341-352: CESPLSDKNLVA (#49)



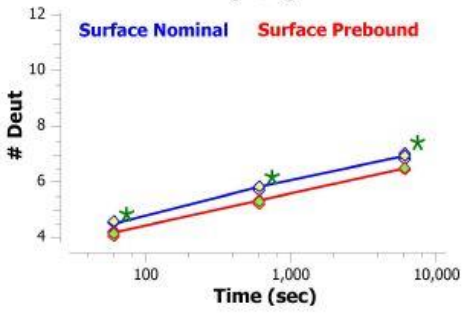
357-364: ARDFLPST (#50)



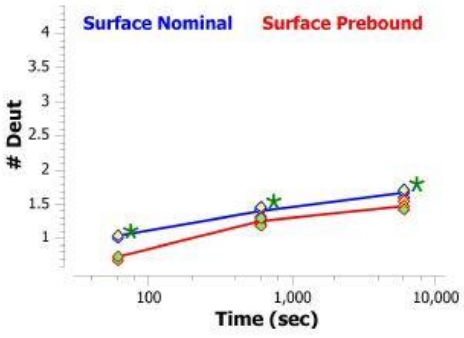
357-367: ARDFLPSTISF (#51)



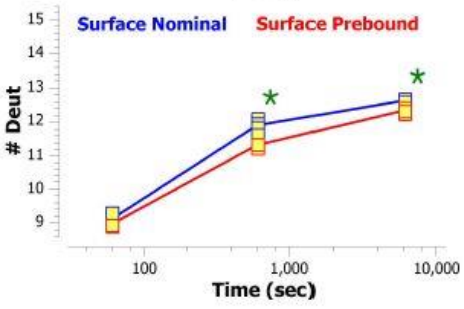
357-370: ARDFLPSTISFTWN (#52)



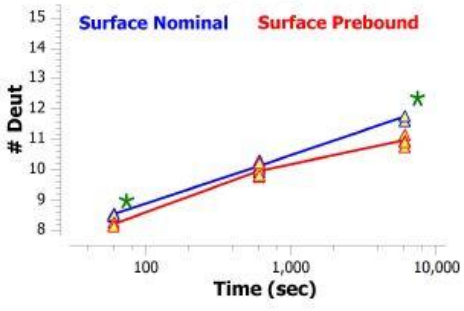
365-370: ISFTWN (#53)



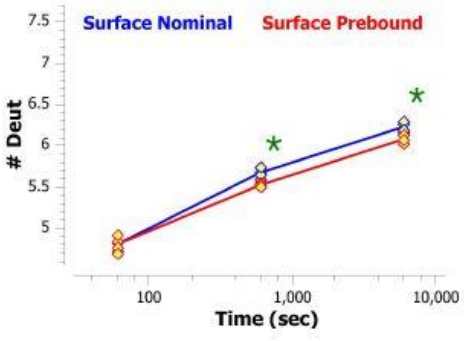
371-387: YQN*NTEVIQGIRTFPTL (#54)



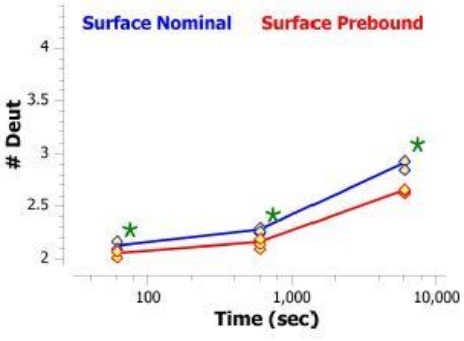
377-393: VIQGIRTFPTLRTGGKY (#55)



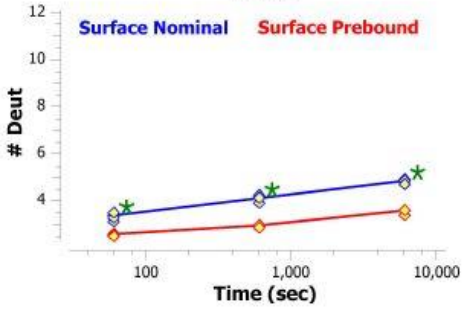
385-393: PTLRTGGKY (#56)



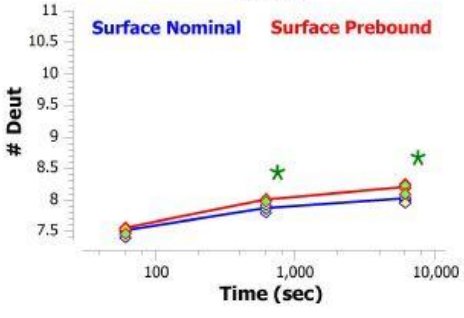
388-393: RTGGKY (#57)



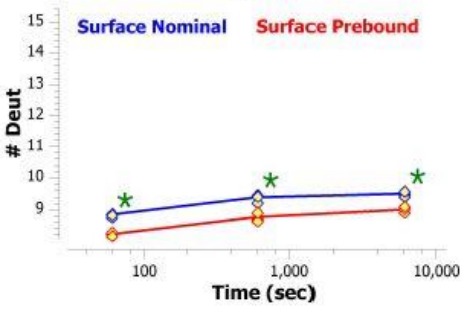
388-400: RTGGKYLATSQVL (#58)

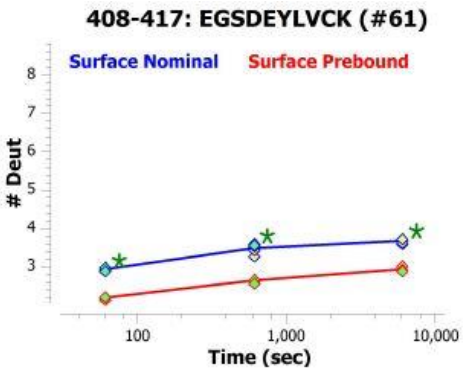


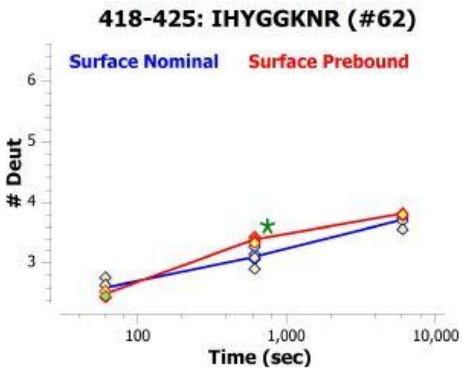
401-413: LSPKSILEGSDEY (#59)

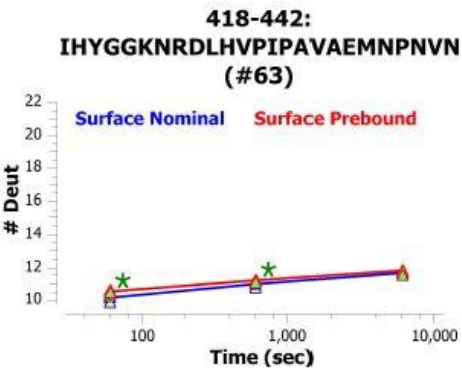


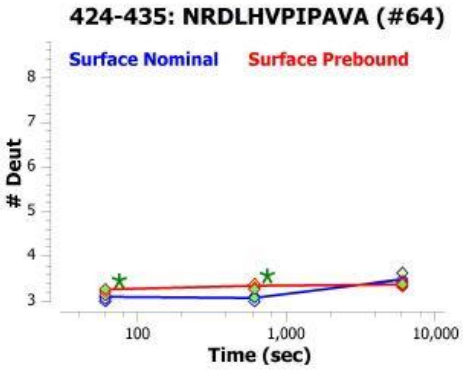
401-417: LSPKSILEGSDEYLVCK (#60)

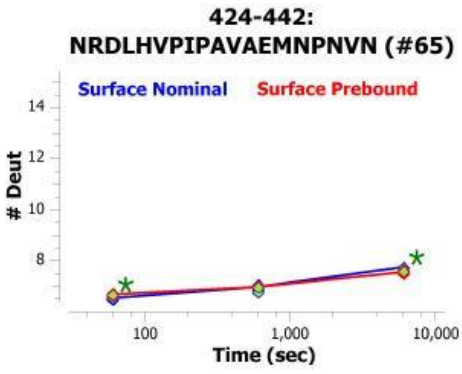


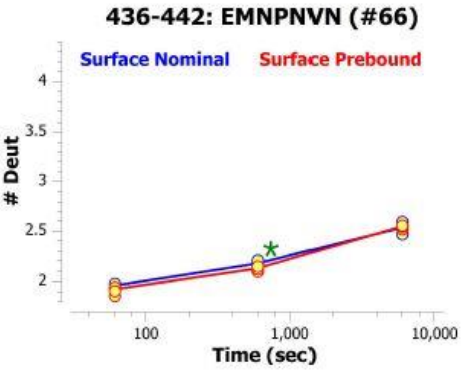

408-417: EGSDEYLVCK (#61)
Surface Nominal
Surface Prebound
Deut
Time (sec)

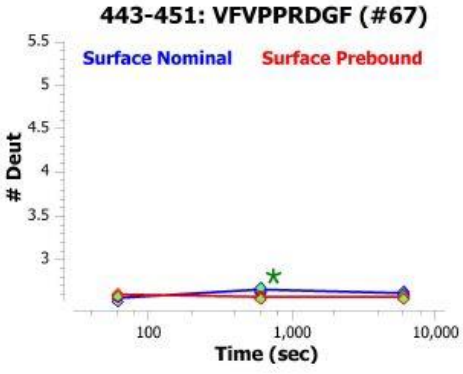

418-425: IHYGGKNR (#62)
Surface Nominal
Surface Prebound
Deut
Time (sec)

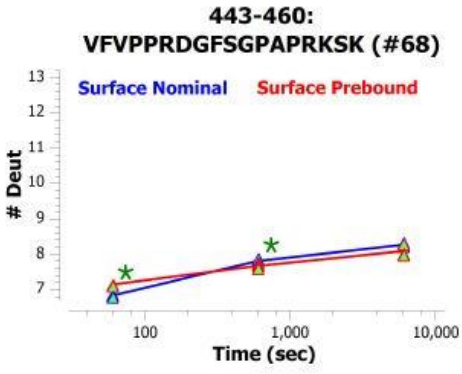

418-442: IHYGGKNRDLHVPIPAVAEMNPNVN (#63)
Surface Nominal
Surface Prebound
Deut
Time (sec)

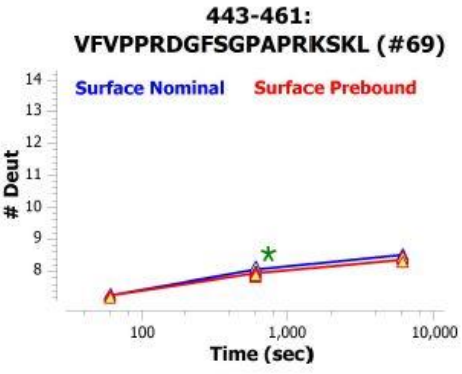

424-435: NRDLHVPIPAVA (#64)
Surface Nominal
Surface Prebound
Deut
Time (sec)

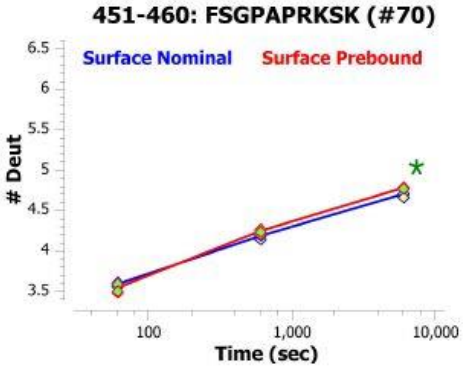

424-442: NRDLHVPIPAVAEMNPNVN (#65)
Surface Nominal
Surface Prebound
Deut
Time (sec)

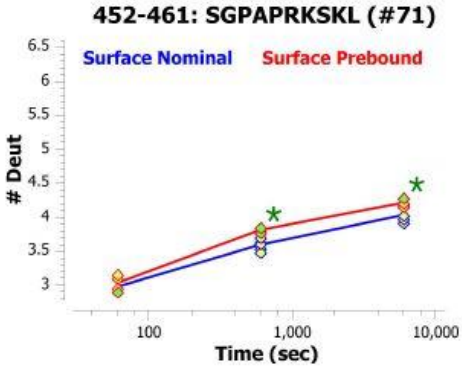

436-442: EMNPNVN (#66)
Surface Nominal
Surface Prebound
Deut
Time (sec)

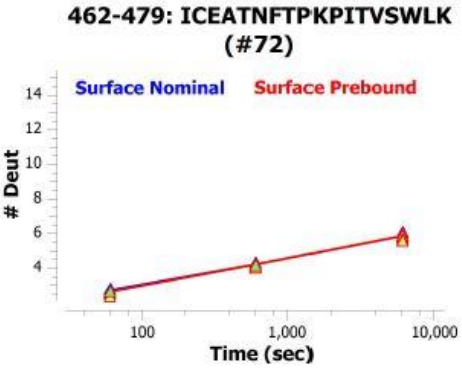

443-451: VFVPPRDGF (#67)
Surface Nominal
Surface Prebound
Deut
Time (sec)

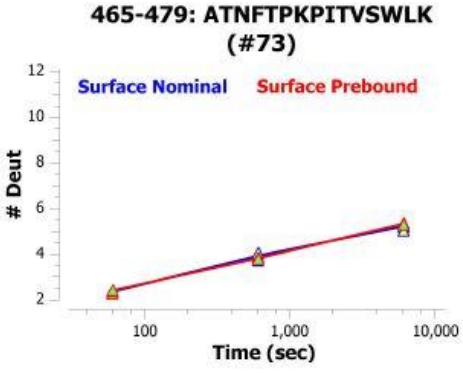

443-460: VFVPPRDGFSGPAPRKSK (#68)
Surface Nominal
Surface Prebound
Deut
Time (sec)

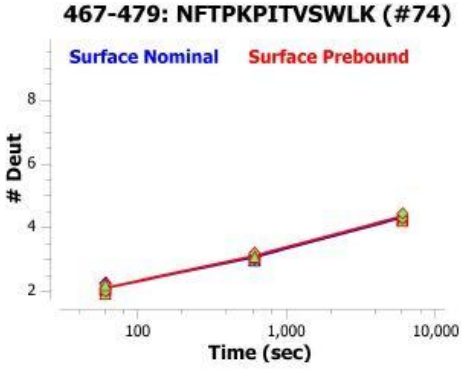

443-461: VFVPPRDGFSGPAPRKSKL (#69)
Surface Nominal
Surface Prebound
Deut
Time (sec)

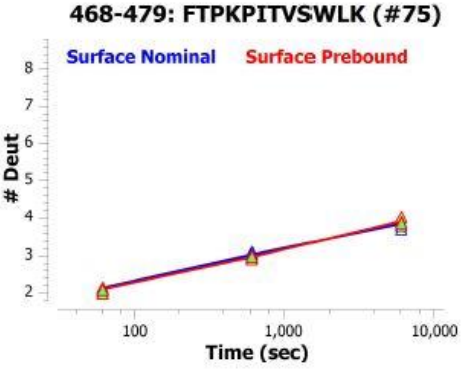

451-460: FSGPAPRKSK (#70)
Surface Nominal
Surface Prebound
Deut
Time (sec)


452-461: SGPAPRKSKL (#71)
Surface Nominal
Surface Prebound
Deut
Time (sec)

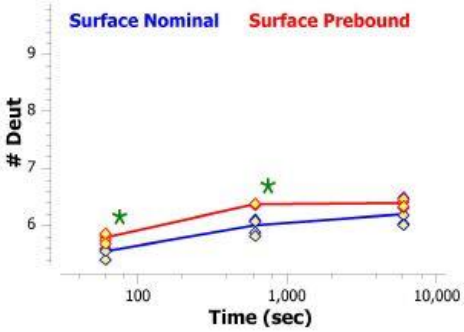

462-479: ICEATNFTPKPITVSWLK (#72)
Surface Nominal
Surface Prebound
Deut
Time (sec)


465-479: ATNFTPKPITVSWLK (#73)
Surface Nominal
Surface Prebound
Deut
Time (sec)

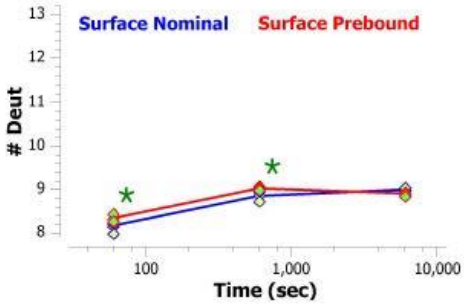

467-479: NFTPKPITVSWLK (#74)
Surface Nominal
Surface Prebound
Deut
Time (sec)


468-479: FTPKPITVSWLK (#75)
Surface Nominal
Surface Prebound
Deut
Time (sec)

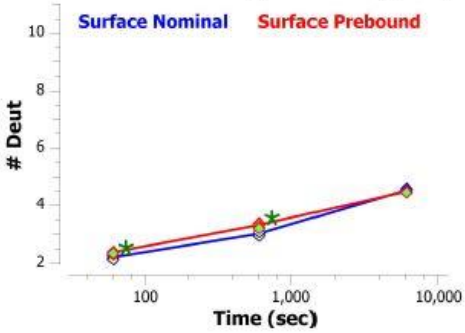
483-494: LVESGFTTDPVT (#76)



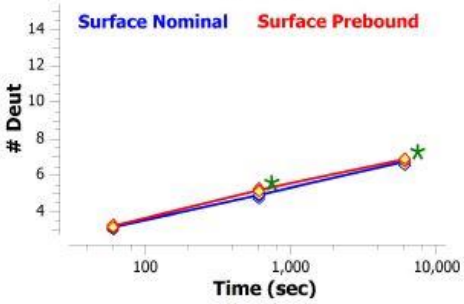
484-498: VESGFTTDPVTIENK (#77)



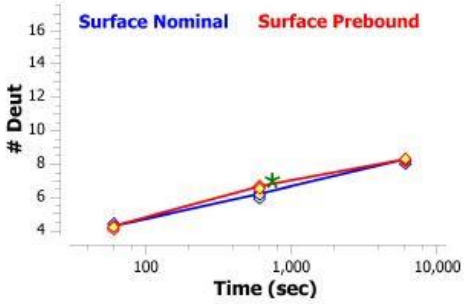
499-511: GSTPQTYKVISTL (#78)



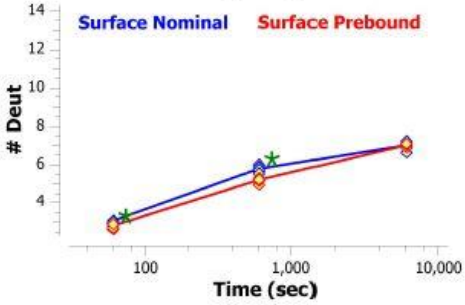
499-515: GSTPQTYKVISTLTISE (#79)



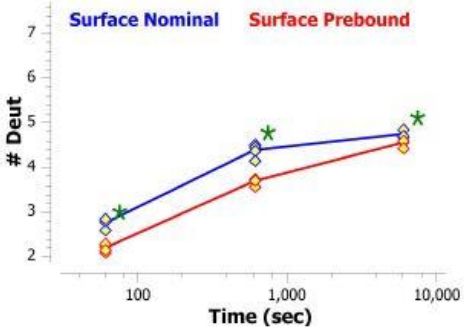
499-517: GSTPQTYKVISTLTISEID (#80)



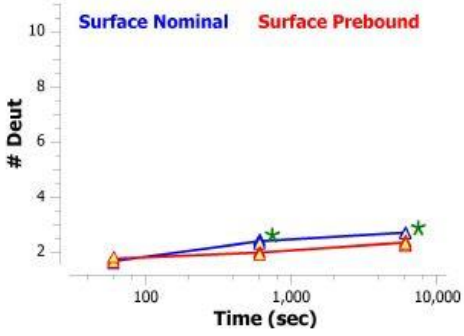
506-520: KVISTLTISEIDWLN (#81)



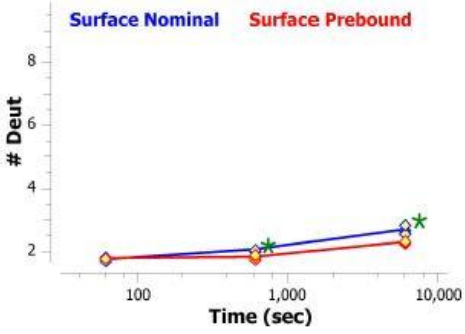
512-520: TISEIDWLN (#82)



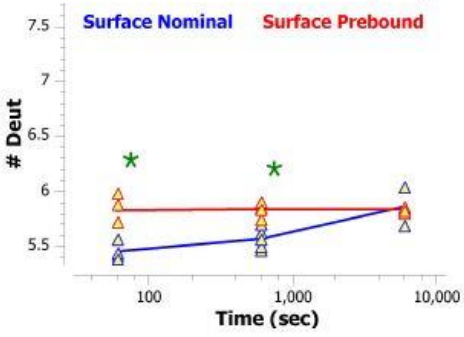
522-533: NVYTCRVDHRGL (#83)



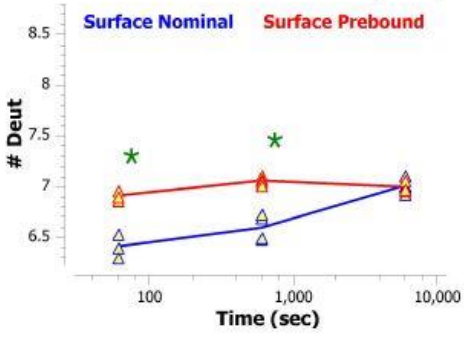
523-533: VYTCRVDHRGL (#84)



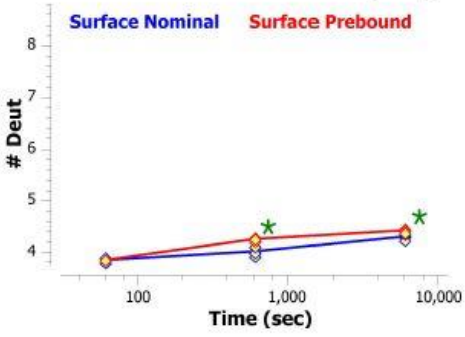
534-542: TFLKN*VSST (#85)



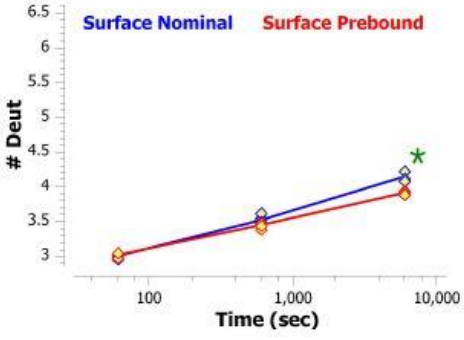
534-543: TFLKN*VSSTC (#86)



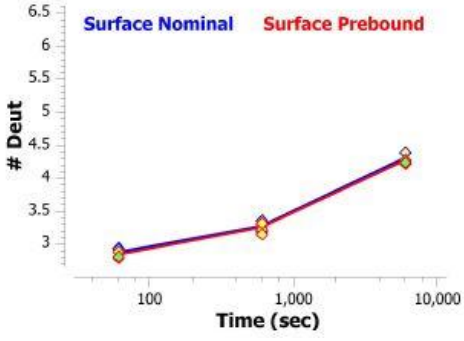
544-554: AASPSTDILTF (#87)



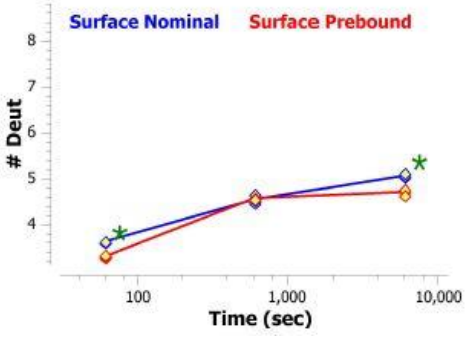
553-562: TFTIPPSFAD (#88)



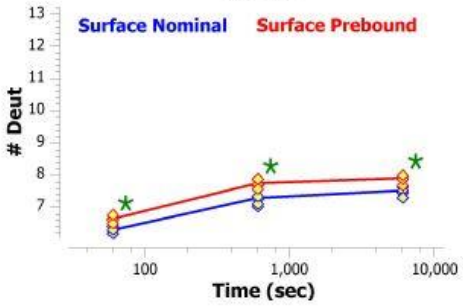
555-564: TIPPSFADIF (#89)



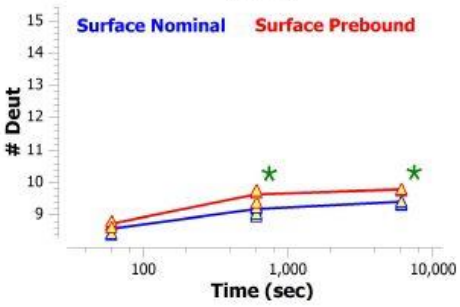
565-574: LSKSAN*LTCL (#90)



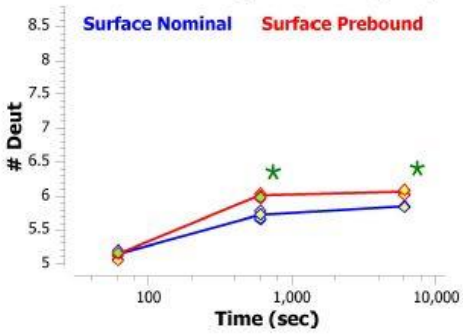
585-599: NISWASQSGEPLETK (#91)



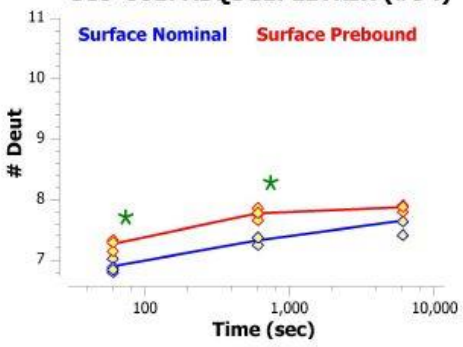
585-601: NISWASQSGEPLETKIK (#92)



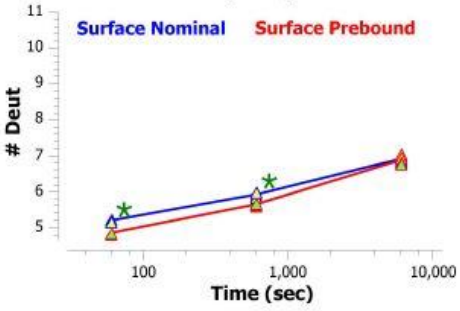
589-599: ASQSGEPLETK (#93)



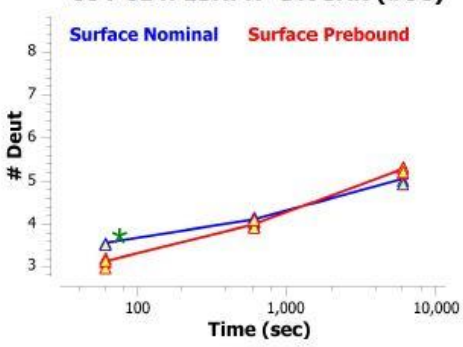
589-601: ASQSGEPLETKIK (#94)



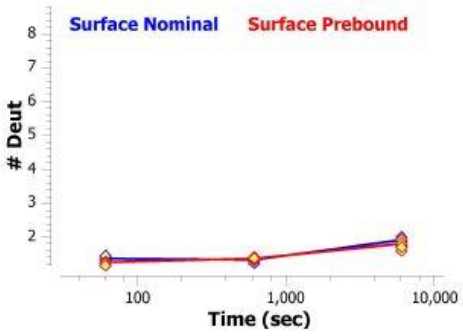
602-614: IMESHPN*GTFSAK (#95)



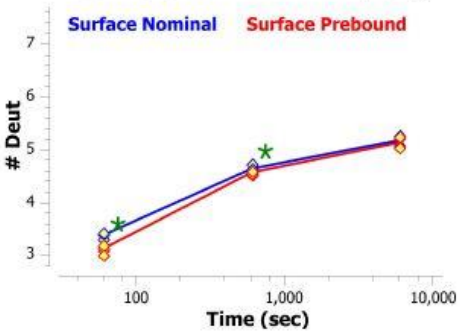
604-614: ESHPN*GTFSAK (#96)



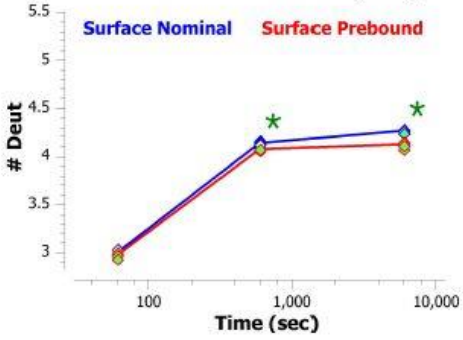
612-621: SAKGVASVCV (#97)



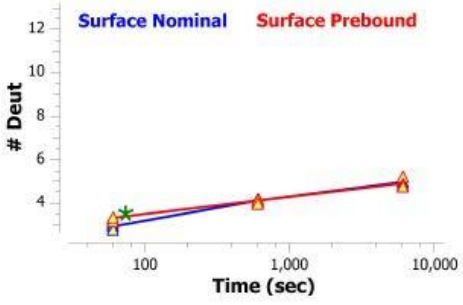
621-629: VEDWNNRKE (#98)



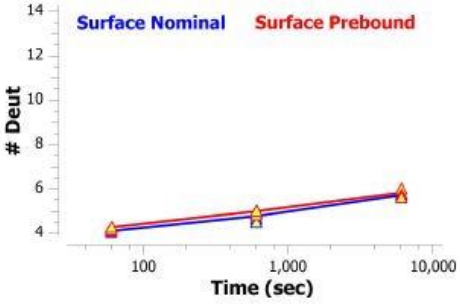
623-629: DWNNRKE (#99)



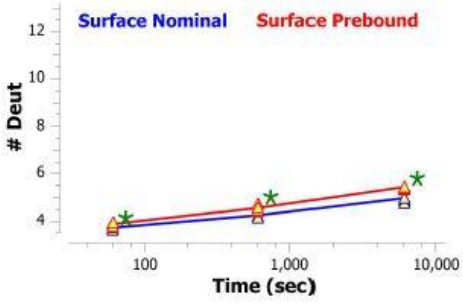
630-645: FVCTVTHRDLPSPQKK (#100)



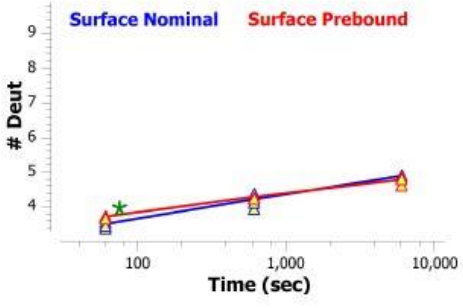
630-646: FVCTVTHRDLPSPQKKF (#101)



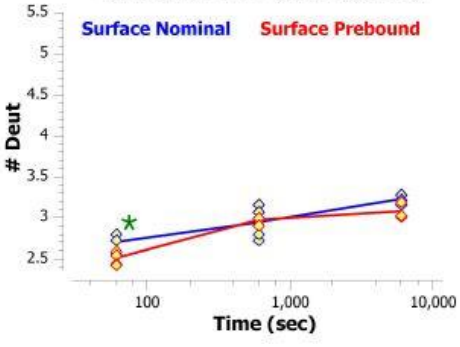
631-646: VCTVTHRDLPSPQKKF (#102)



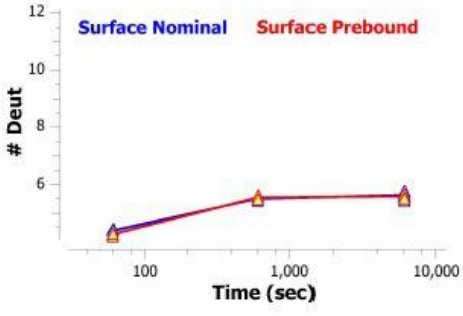
633-645: TVTHRDLPSPQKK (#103)

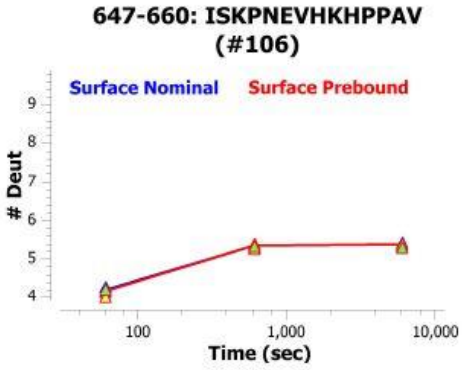


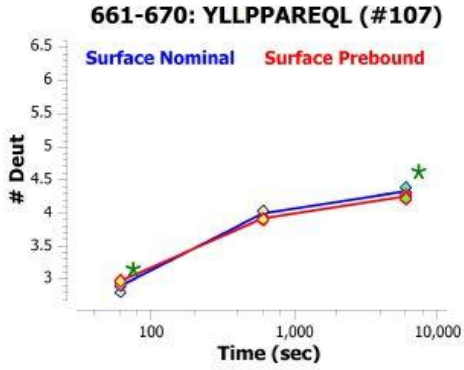
639-646: LPSPQKKF (#104)

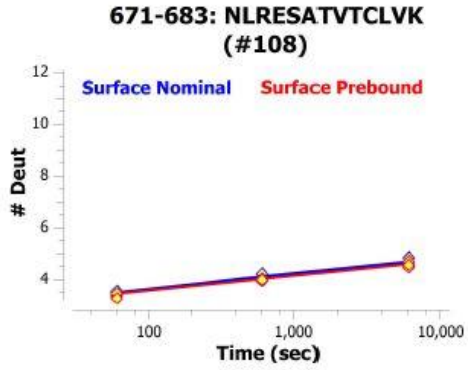


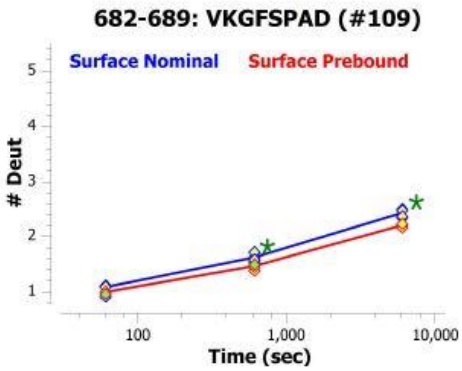
646-661: FISKPNEVHKHPPAVY (#105)

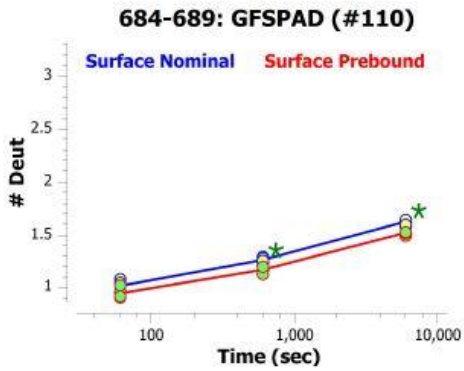


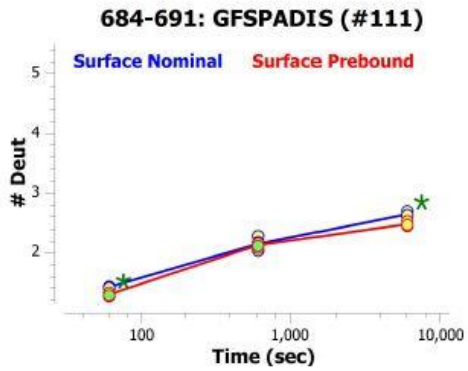

647-660: ISKPNEVHKHPPAV (#106)
Surface Nominal
Surface Prebound
Deut
Time (sec)

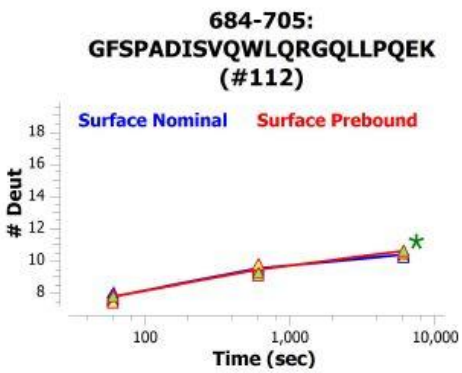

661-670: YLLPPAREQL (#107)
Surface Nominal
Surface Prebound
Deut
Time (sec)

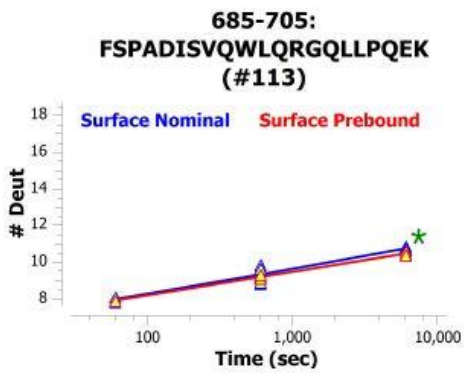

671-683: NLRESATVTCLVK (#108)
Surface Nominal
Surface Prebound
Deut
Time (sec)

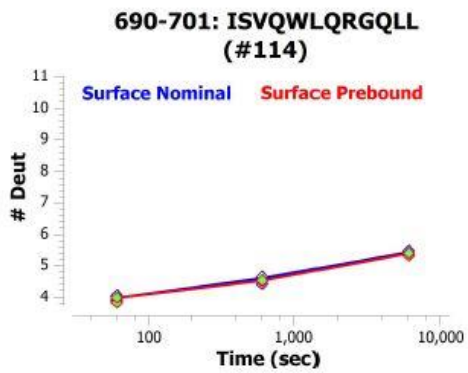

682-689: VKGFSPAD (#109)
Surface Nominal
Surface Prebound
Deut
Time (sec)

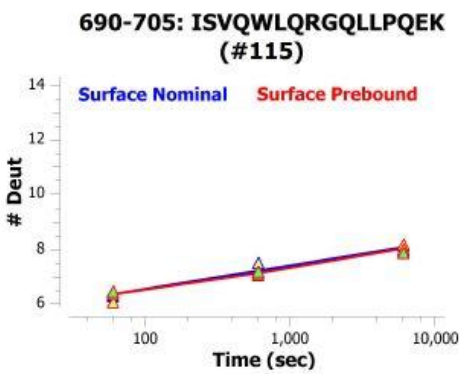

684-689: GFSPAD (#110)
Surface Nominal
Surface Prebound
Deut
Time (sec)

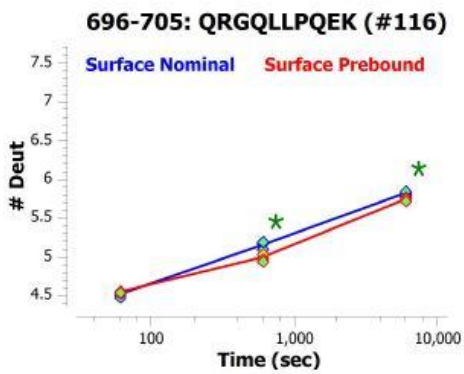

684-691: GFSPADIS (#111)
Surface Nominal
Surface Prebound
Deut
Time (sec)

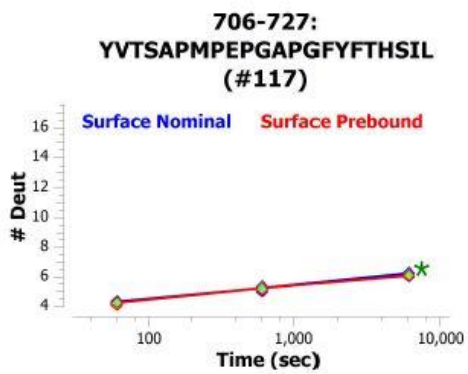

684-705: GFSPADISVQWLQRGQLLPQEK (#112)
Surface Nominal
Surface Prebound
Deut
Time (sec)

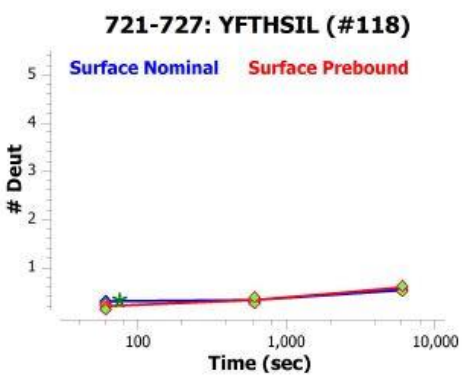

685-705: FSPADISVQWLQRGQLLPQEK (#113)
Surface Nominal
Surface Prebound
Deut
Time (sec)

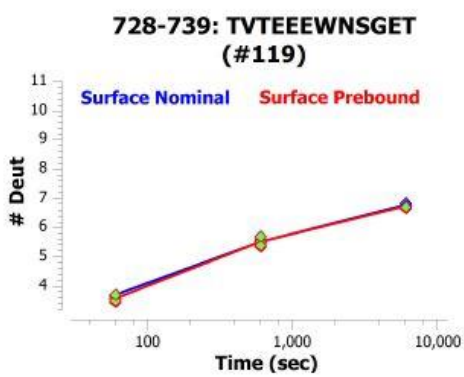

690-701: ISVQWLQRGQLL (#114)
Surface Nominal
Surface Prebound
Deut
Time (sec)

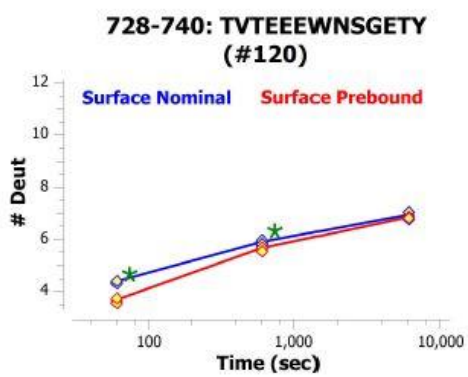

690-705: ISVQWLQRGQLLPQEK (#115)
Surface Nominal
Surface Prebound
Deut
Time (sec)

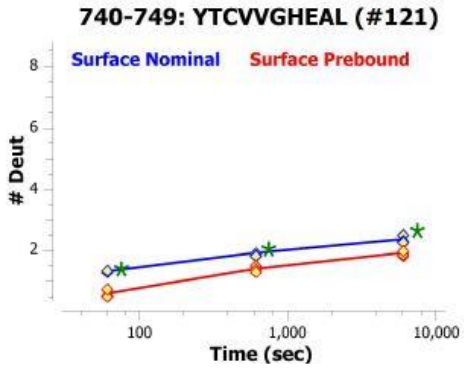

696-705: QRGQLLPQEK (#116)
Surface Nominal
Surface Prebound
Deut
Time (sec)

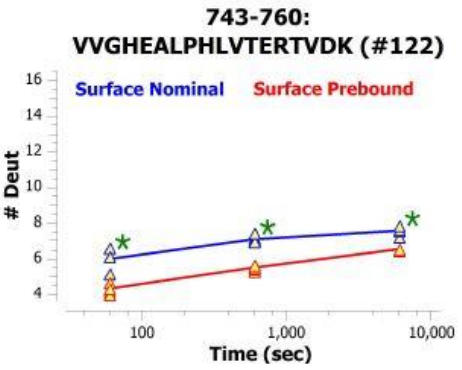

706-727: YVTSAPMPEPGAPGFYFTHSIL (#117)
Surface Nominal
Surface Prebound
Deut
Time (sec)

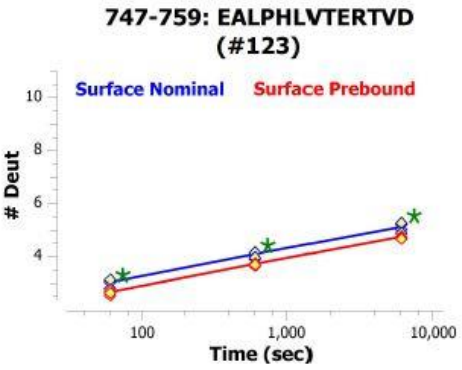

721-727: YFTHSIL (#118)
Surface Nominal
Surface Prebound
Deut
Time (sec)

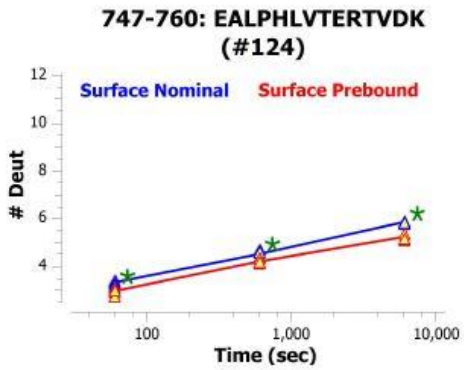

728-739: TVTEEEWNSGET (#119)
Surface Nominal
Surface Prebound
Deut
Time (sec)

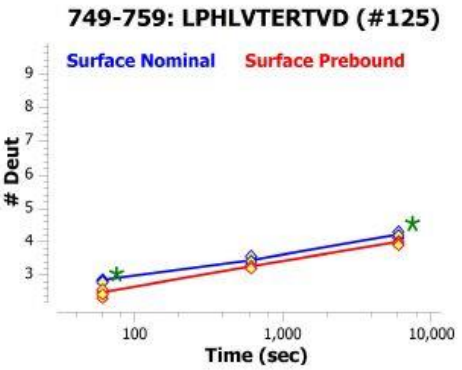

728-740: TVTEEEWNSGETY (#120)
Surface Nominal
Surface Prebound
Deut
Time (sec)

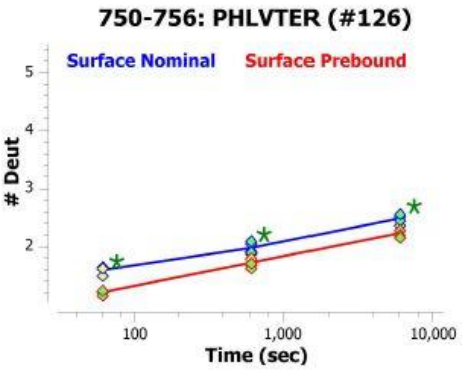

740-749: YTCVVGHEAL (#121)
Surface Nominal
Surface Prebound
Deut
Time (sec)

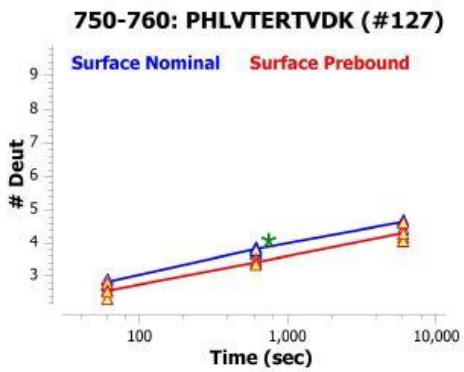

743-760: VVGHEALPHLVTERTVDK (#122)
Surface Nominal
Surface Prebound
Deut
Time (sec)

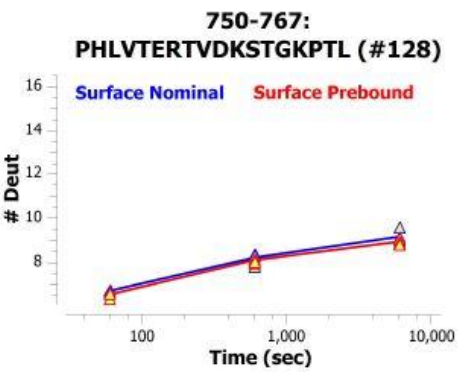

747-759: EALPHLVTERTVD (#123)
Surface Nominal
Surface Prebound
Deut
Time (sec)

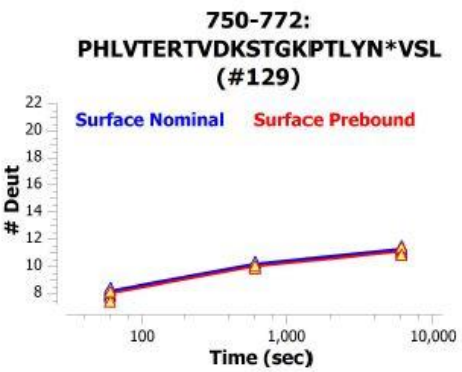

747-760: EALPHLVTERTVDK (#124)
Surface Nominal
Surface Prebound
Deut
Time (sec)

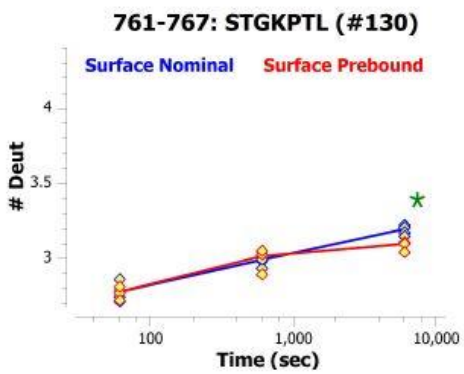

749-759: LPHLVTERTVD (#125)
Surface Nominal
Surface Prebound
Deut
Time (sec)

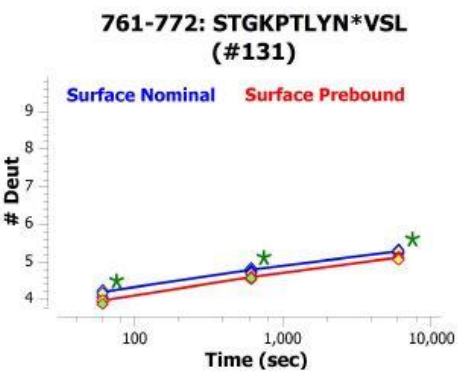

750-756: PHLVTER (#126)
Surface Nominal
Surface Prebound
Deut
Time (sec)

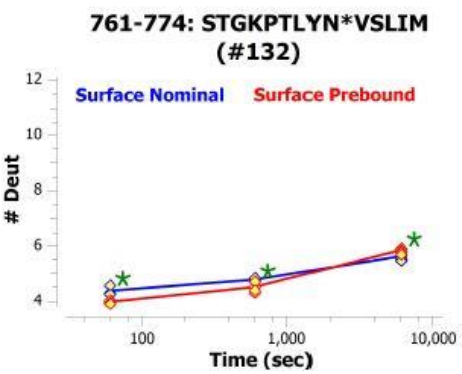

750-760: PHLVTERTVDK (#127)
Surface Nominal
Surface Prebound
Deut
Time (sec)

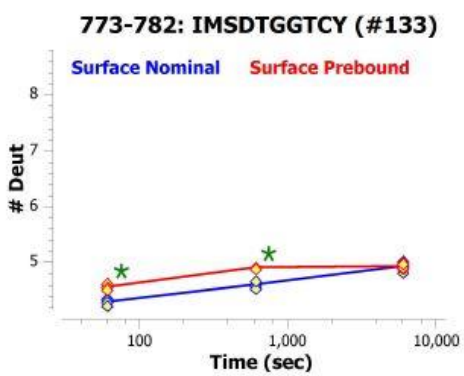

750-767: PHLVTERTVDKSTGKPTL (#128)
Surface Nominal
Surface Prebound
Deut
Time (sec)

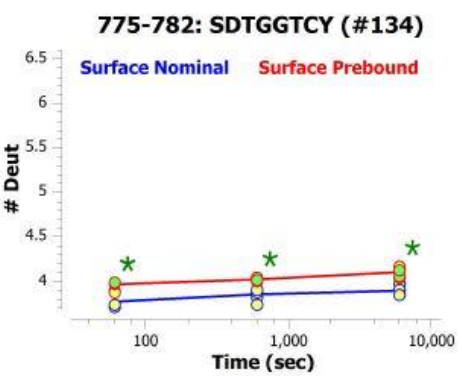

750-772: PHLVTERTVDKSTGKPTLYN*VSL (#129)
Surface Nominal
Surface Prebound
Deut
Time (sec)

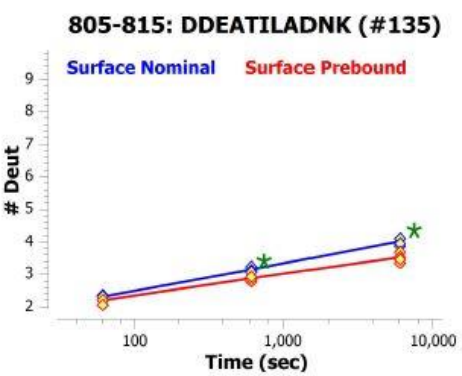

761-767: STGKPTL (#130)
Surface Nominal
Surface Prebound
Deut
Time (sec)

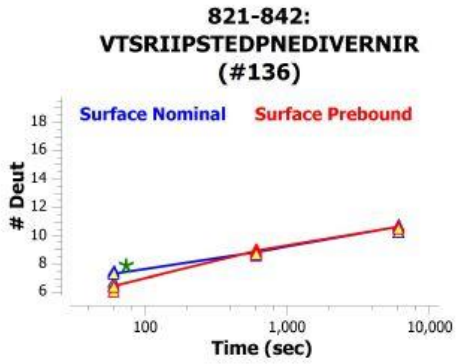

761-772: STGKPTLYN*VSL (#131)
Surface Nominal
Surface Prebound
Deut
Time (sec)

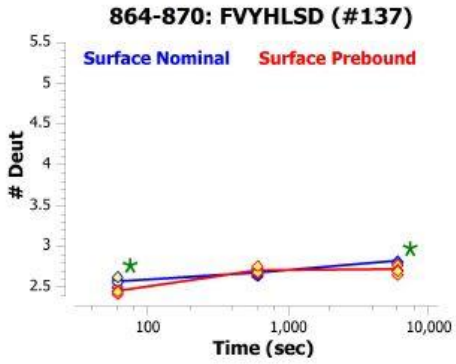

761-774: STGKPTLYN*VSLIM (#132)
Surface Nominal
Surface Prebound
Deut
Time (sec)

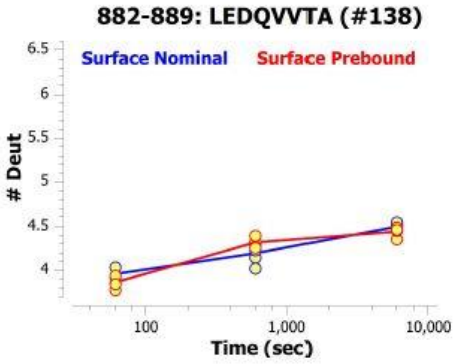

773-782: IMSDTGGTCY (#133)
Surface Nominal
Surface Prebound
Deut
Time (sec)

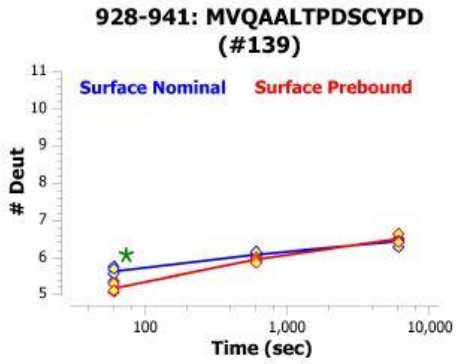

775-782: SDTGGTCY (#134)
Surface Nominal
Surface Prebound
Deut
Time (sec)

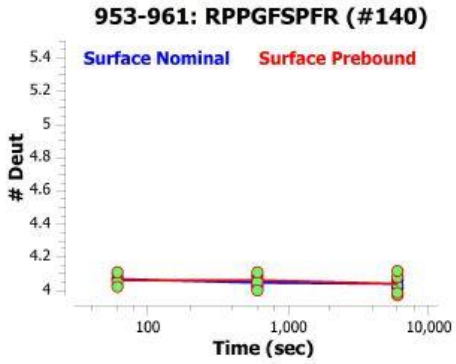

805-815: DDEATILADNK (#135)
Surface Nominal
Surface Prebound
Deut
Time (sec)

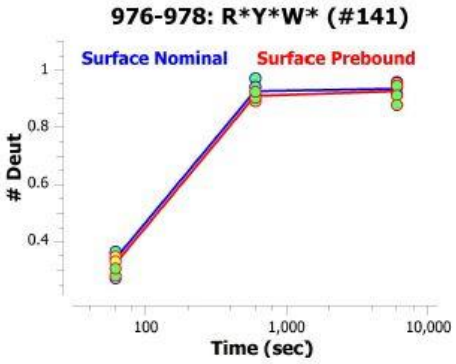

821-842:
VTSRIIPSTEDPNEDIVERNIR
(#136)
Surface Nominal
Surface Prebound
Deut
Time (sec)

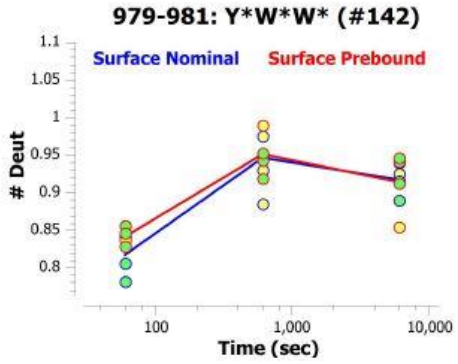

864-870: FVYHLSD (#137)
Surface Nominal
Surface Prebound
Deut
Time (sec)

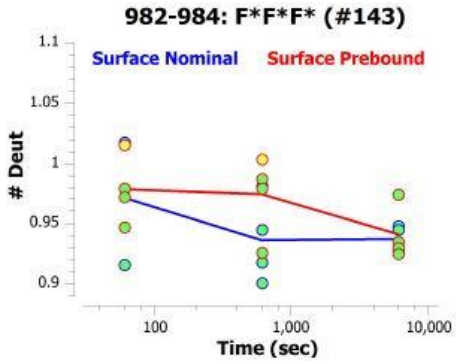

882-889: LEDQVVTA (#138)
Surface Nominal
Surface Prebound
Deut
Time (sec)


928-941: MVQAALTPDSCYPD
(#139)
Surface Nominal
Surface Prebound
Deut
Time (sec)


953-961: RPPGFSPFR (#140)
Surface Nominal
Surface Prebound
Deut
Time (sec)


976-978: R*Y*W* (#141)
Surface Nominal
Surface Prebound
Deut
Time (sec)


979-981: Y*W*W* (#142)
Surface Nominal
Surface Prebound
Deut
Time (sec)


982-984: F*F*F* (#143)
Surface Nominal
Surface Prebound
Deut
Time (sec)

4.6 References

1. Sharp, T. H. *et al.* Insights into IgM-mediated complement activation based on in situ structures of IgM-C1-C4b. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 11900–11905 (2019).

2. Feinstein, A., Richardson, N. E., Gorick, B. D. & Hughes-Jones, N. C. Immunoglobulin M conformational change is a signal for complement activation. *Protein Conform. as an Immunol. Sign* 47–57 (1981).

3. Feinstein, A. & Munn, E. A. Conformation of the free and antigen-bound IgM antibody molecules. *Nature* **224**, 1307–1309 (1969).

4. Feinstein, A., Munn, E. A. & Richardson, N. E. The three-dimensional conformation of γM and γA globulin molecules. *Ann. N. Y. Acad. Sci.* **190**, 104–121 (1971).

5. Perkins, S. J., Nealis, A. S., Sutton, B. J. & Feinstein, A. Solution structure of human and mouse immunoglobulin M by synchrotron X-ray scattering and molecular graphics modelling. *J. Mol. Biol.* **221**, 1345–1366 (1991).

6. Roux, K. H. Immunoglobulin structure and function as revealed by electron microscopy. *Int. Arch. Allergy Immunol.* **120**, 85–99 (1999).

7. Tobita, T., Oda, M. & Azuma, T. Segmental flexibility and avidity of IgM in the interaction of polyvalent antigens. *Mol. Immunol.* **40**, 803–811 (2004).

8. Moh, E. S. X., Lin, C. H., Thaysen-Andersen, M. & Packer, N. H. Site-Specific N-Glycosylation of Recombinant Pentameric and Hexameric Human IgM. *J. Am. Soc. Mass Spectrom.* **27**, 1143–1155 (2016).

9. Tchoudakova, A. *et al.* High level expression of functional human IgMs in human PER.C6® cells. *MAbs* **1**, 163–171 (2009).

10. Collins, C., Tsui, F. W. L. & Shulman, M. J. Differential activation of human and guinea pig complement by pentameric and hexameric IgM. *Eur. J. Immunol.* **32**, 1802 (2002).

11. Brewer, J. W., Randall, T. D., Parkhouse, R. M. E. & Corley, R. B. IgM hexamers? *Immunol. Today* **15**, 165–168 (1994).

12. Li, Y. *et al.* Structural insights into immunoglobulin M. *Science (80-.).* **367**, 1014–1017 (2020).

13. Kumar, N., Arthur, C. P., Ciferri, C. & Matsumoto, M. L. Structure of the human secretory immunoglobulin M core. *Structure* **29**, 564-571.e3 (2021).

14. Wei, H. *et al.* Hydrogen/deuterium exchange mass spectrometry for probing higher order structure of protein therapeutics: methodology and applications. *Drug Discov. Today* **19**, 95–102 (2014).

15. Hodge, E. A., Benhaim, M. A. & Lee, K. K. Bridging protein structure, dynamics, and function using hydrogen/deuterium-exchange mass spectrometry. *Protein Sci.* **29**, 843–855 (2020).

16. 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).

17. Hageman, T. S. & Weis, D. D. Reliable Identification of Significant Differences in Differential Hydrogen Exchange-Mass Spectrometry Measurements Using a Hybrid Significance Testing Approach. *Anal. Chem.* **91**, 8008–8016 (2019).

18. Murphree, T. A., Vorauer, C., Brzoska, M. & Guttman, M. Imidazolium Compounds as Internal Exchange Reporters for Hydrogen/Deuterium Exchange by Mass Spectrometry. *Anal. Chem.* **92**, 9830–9837 (2020).

19. Watson, M. J. *et al.* Simple Platform for Automating Decoupled LC-MS Analysis of Hydrogen/Deuterium Exchange Samples. *J. Am. Soc. Mass Spectrom.* (2020) doi:10.1021/jasms.0c00341.

20. Müller, R. *et al.* High-resolution structures of the IgM Fc domains reveal principles of its hexamer formation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 10183–8 (2013).

21. Graille, M. *et al.* Crystal structure of a Staphylococcus aureus protein a domain complexed with the Fab fragment of a human IgM antibody: Structural basis for recognition of B-cell receptors and superantigen activity. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5399–5404 (2000).

22. Padlan, E. A. *et al.* Structure of an antibody-antigen complex: Crystal structure of the HyHEL-10 Fab-lysozyme complex. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 5938–5942 (1989).

23. Acchione, M., Kwon, H., Jochheim, C. M. & Atkins, W. M. Impact of linker and conjugation chemistry on antigen binding, Fc receptor binding and thermal stability of model antibody-drug conjugates. *MAbs* **4**, 362–372 (2012).

24. Kam-morgan, L. N. W. *et al.* High-resolution mapping of the HyHEL-10 epitope of chicken lysozyme by site-directed mutagenesis. *Proc Natl Acad Sci U S A* **90**, 3958–3962 (1993).

25. Xavier, K. A. & Willson, R. C. Association and dissociation kinetics of anti-hen egg lysozyme monoclonal antibodies HyHEL-5 and HyHEL-10. *Biophys. J.* **74**, 2036–2045 (1998).

26. Brown, J. C. & Koshland, M. E. Activation of antibody Fc function by antigen-induced conformational changes. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 5111–5 (1975).

27. Pecht, I. Recognition and Allostery in the Mechanism of Antibody Action (Springer book chapter).pdf. 41–54 (1976).

28. Chen, Q., Menon, R., Calder, L. J., Rosenthal, P. B. & Tolar, P. Cryomicroscopy reveals the structural basis for a flexible hinge motion in the immunoglobulin M pentamer. *Nat. Commun.* **13**, 1–11 (2022).

29. Xu, J. L. & Davis, M. M. Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. *Immunity* **13**, 37–45 (2000).

30. Briney, B. S., Willis, J. R. & Crowe, J. E. Human peripheral blood antibodies with long HCDR3s are established primarily at original recombination using a limited subset of germline genes. *PLoS One* **7**, 1–13 (2012).

31. Hiramoto, E. *et al.* The IgM pentamer is an asymmetric pentagon with an open groove that binds the AIM protein. *Sci. Adv.* **4**, eaau1199 (2018).

32. Siegel, Richard C; Cathou, R. E. Effects of Limited Denaturation by Heat on the Dynamic Conformation of Equine Immunoglobulin M Antibody and on Interaction with Antigen and Complement. *Biochemistry* 192–198 (1981).

33. Roux, K. H., Strelets, L. & Michaelsent, T. E. Flexibility of human IgG subclasses. *J. Immunol.* **159**, 3372–3382 (1997).

34. Feinstein, A., Richardson, N. & Taussig, M. J. Immunoglobulin flexibility in complement activation. *Immunol. Today* **7**, (1986).

35. Alexander R. Duncan & Greg, W. The binding site for C1q on IgG. *Nature* **332**, 738–740 (1988).

36. Parkhouse, R. M., Askonas, B. A. & Dourmashkin, R. R. Electron microscopic studies of mouse immunoglobulin M; structure and reconstitution following reduction. *Immunology* **18**, 575–584 (1970).

37. Valim, Y. M. L. & Lachmann, P. J. The effect of antibody isotype and antigenic epitope density on the complement-fixing activity of immune complexes: a systematic study using chimaeric anti-NIP antibodies with human Fc regions. *Clin. Exp. Immunol.* **84**, 1–8 (1991).

38. Czajkowsky, D. M. & Shao, Z. The human IgM pentamer is a mushroom-shaped molecule with a flexural bias. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14960–5 (2009).

39. Shulman, M. J., Collins, C., Pennell, N. & Hozumi, N. Complement activation by IgM: evidence for the importance of the third constant domain of the μ heavy chain. *Eur. J. Immunol.* **17**, 549–554 (1987).

40. Shulman, M. J., Heusser, C., Filkin, C. & Kohler, G. Mutations Affecting the Structure and Function of Immunoglobulin M. *Mol. Cell. Biol.* **2**, 1033–1043 (1982).

41. Bubb, M. O. & Conradie, J. O. The Cmu3 Domain of IgM: ISOLATION ANO IDENTIFICATION OF THE INTACT FRAGMENT. *Biochem. Biophys. Res. Commun.* **77**, 613–620 (1977).

42. Bubb, M. O. & Conradie, J. D. Studies on the structural and biological functions of the CH3 and CH4 domains of IgM. *Immunology* (1978).

43. Hurst, M. M., Volanakis, J. E., Hester, R. B., Stroud, R. M. & Bennett, J. C. THE STRUCTURAL BASIS FOR BINDING OF COMPLEMENT BY IMMUNOGLOBULIN M. *Exp. Med.* **140**, 1117–1121 (1974).

44. Hurst, M. M., Volanakis, J. E., Stroud, R. M. & Bennett, J. C. C1 fixation and classical complement pathway activation by a fragment of the CH4 domain of IgM. *Exp. Med.* **142**, (1975).

45. Hurst, M. M., Volanakis, J. E., Stroud, R. M. & Bennett, J. C. A Comparative Analysis of the C1-Binding Ability of Fragments Derived from Complement-Fixing and Noncomplement-Fixing IgM Proteins. *J. Clin. Invest.* **58**, 16–21 (1976).

46. Johnson, B. J. & Thames, K. E. Investigations of the Complement-Fixing sites of Immunoglobulins. *J. Immunol.* **117**, 1491–1494 (1976).

47. Zlatarova, A. S. *et al.* Existence of different but overlapping IgG- and IgM-binding sites on the globular domain of human C1q. *Biochemistry* **45**, 9979–9988 (2006).

48. Gadjeva, M. G. *et al.* Interaction of human C1q with IgG and IgM: Revisited. *Biochemistry* **47**, 13093–13102 (2008).

49. Arya, S. *et al.* Mapping of amino acid residues in the C mu 3 domain of mouse IgM important in macromolecular assembly and complement-dependent cytolysis. *J. Immunol.* **152**, 1206–12 (1994).

50. Wright, J. F., Shulman, M. J., Isenman, D. E. & Painter, R. H. C1 binding by mouse IgM. The effect of abnormal glycosylation at position 402 resulting from a serine to asparagine exchange at residue 406 of the μ-chain. *J. Biol. Chem.* **265**, 10506–10513 (1990).

51. van Osch, T. L. J. *et al.* Fc Galactosylation Promotes Hexamerization of Human IgG1, Leading to Enhanced Classical Complement Activation. *J. Immunol.* **207**, 1545–1554 (2021).

52. Ramsland, P. A. *et al.* Crystal structure of a glycosylated Fab from an IgM cryoglobulin with properties of a natural proteolytic antibody. *Biochem. J.* **481**, 473–481 (2006).

53. Putnam, F. W., Florent, G., Paul, C., Shinoda, T. & Shimizu, A. Complete Amino Acid Sequence of the Mu Heavy Chain of a Human IgM Immunoglobulin. *Science (80-.).* 287–291 (1973).

54. Gaboriaud, C., Frachet, P., Thielens, N. M. & Arlaud, G. J. The human C1q globular domain: Structure and recognition of non-immune self ligands. *Front. Immunol.* **2**, 1–8 (2012).

55. Gaboriaud, C., Ling, W. L., Thielens, N. M., Bally, I. & Rossi, V. Deciphering the fine details of C1 assembly and activation mechanisms: 'Mission impossible'? *Front. Immunol.* **5**, 3–9 (2014).

56. Thielens, N. M., Tedesco, F., Bohlson, S. S., Gaboriaud, C. & Tenner, A. J. C1q: A fresh look upon an old molecule. *Mol. Immunol.* **89**, 73–83 (2017).

57. Levine, L., Wasserman, E. & Mills, S. Complement Fixation and Ionic Strength. *J. Immunol.* **86**, 675–680 (1961).

58. Pasalic, D. *et al.* A peptide extension dictates IgM assembly. *Proc. Natl. Acad. Sci.* 201701797 (2017) doi:10.1073/pnas.1701797114.

59. Sopp, J. M. *et al.* On-target IgG hexamerisation driven by a C-terminal IgM tail-piece fusion variant confers augmented complement activation. *Commun. Biol.* **4**, 1–14 (2021).

60. Ugurlar, D. *et al.* Structures of C1-IgG1 provide insights into how danger pattern recognition activates complement. *Science (80-.).* **359**, 794–797 (2018).

61. Wang, G. *et al.* Molecular Basis of Assembly and Activation of Complement Component C1 in Complex with Immunoglobulin G1 and Antigen. *Mol. Cell* **63**, 135–145 (2016).

62. Diebolder, C. A. *et al.* Complement Is Activated by IgG Hexamers Assembled at the Cell Surface. *Science (80-.).* **343**, 1260–1263 (2014).

– Chapter 5 –

Conclusions and Future Directions

5.1 Conclusions

The IgM antibody is perhaps one of the most fundamental components of the entire immune system, outside of the B cells that produce it. The genetically conserved repertoire of 'natural' IgM that protects all newborn mammals provides an innate degree of protection against basic pathogenic features, which helps in part to afford enough time for the adaptive immune system to sufficiently mature and expand.[1,2] Because 'induced' IgM is always secreted first in response to foreign pathogens, its antigen specificity is often relatively low, especially when compared to affinity-matured IgG that has undergone multiple rounds of somatic hypermutation and affinity maturation throughout the process of class-switching. Despite the inherent lack of antigen specificity, IgM has endured the test of time and is the oldest known Ig component shared amongst all vertebrate species, remaining essentially unchanged for more than 500 million years.[3] Its innate counterpart, C1q, along with other members of the complement family likely outdate the entirety of the adaptive immune system itself, evidenced by the recent discovery of an ancient protein homologue of C1q found in a descendant species of non-vertebrate chordates, that is also capable of binding to human IgG and inducing activation of the classical complement cascade.[4] Regardless of the evolutionary basis, the pronounced hexameric architecture of C1q is clearly an apt geometric match to the polyvalent structures of secreted pentameric (IgMp), and especially hexameric (IgMh) isoforms of IgM, which still remains the most potent known antibody activator of the classical complement cascade.[5–7]

The complement system is a 'cold' and 'unintelligent' signaling pathway that is governed by probability and driven by the binding interactions and conformational changes of its multitude of

otherwise-inactive protease components. The activated protease fragments that drive each major step of the pathway must fundamentally act against a number of regulatory thresholds, such as the highly abundant and potent C1-inhibitor molecule (C1-INH), that must be sufficiently and rapidly overcome in order for the cascade to continue its progression towards the ultimate formation of the membrane attack complex (MAC). While multivalent surface binding clearly remains the most potent modality of complement-active IgM, the ability of transient staple formation appears to be within the natural energy landscape of solution bound IgM as well. This leads us to hypothesize that the activation mechanism is driven neither by allosterism nor any form of significant physical distortion that would otherwise be naturally unobtainable in solution.

The forces that govern IgM staple formation are, from both an evolutionary and biophysical standpoint, likely based instead upon the same 'cold' and uncalculating nature that governs the complement cascade itself, which implies the activation mechanism to be driven primarily by steric accessibility and degree of surface binding multivalency. The role of the antigen display, therefore, is to facilitate and stabilize the multivalent *ensemble* of discrete staple formations that are formed by the participating binding 'arms' of IgM – each of which likely presents a copy of the C1q binding site(s) to facilitate the subsequent binding and activation of C1. In the same way that surface Ag presentation dictates the conformational range of the IgM molecule that binds it, the presentation of C1q binding sites via formation of the staple *ensemble* subsequently dictates the activation 'range' of C1. Allosterism, at least to some degree, does appear to favor and/or further promote the transient formation of staple when IgM is bound in saturation to soluble protein Ag – but there is little comparison between the potency of solution bound IgM and that of the multivalent surface bound form. This concept also follows logically, from an evolutionary standpoint, in which a targeted membrane (i.e. surface) would be fundamentally required for the selective pressure needed to drive evolution towards the intricate machinery that underlies the membrane attack complex (MAC).

From the functional characterizations of IgM staple observed in Chapter 3, to the structural features determined in Chapter 4 that appear to define the active conformation itself, it's now apparent that IgM is a highly flexible molecule that inherently acts as a pre-packaged, multivalent display for the recruitment/activation of C1q/C1. Attenuation to surface binding valency achieved by prebinding IgM with molar excesses of soluble monomeric Ag resulted in both the progressive loss of C1q binding efficiency as well as a marked reduction in C1 activity, which together demonstrates the critical role that binding valency and strength of avidity play within the IgM-mediated activation mechanism of complement. We observed the loss in binding efficiency to be surprisingly gradual in nature, which we believe is reflective of the flexibility and conformational range of the IgM antibody itself, and further demonstrates that IgM is capable of C1 activation even when the display of its surface bound arms is geometrically non-optimal – which could also be viewed as a primary consequence of an inherently low triggering threshold for the promiscuous C1q molecule. Solution bound IgM also displayed a consistently significant, albeit low degree of C1 activity, which was not predicted to occur based on Feinstein's original hypothesis that global staple formation might be a strict conformational prerequisite to the IgM-mediated activation of complement.

In general, all of the results that we have generated here *in vitro* should be viewed within the context of an *in vivo* biochemical signaling pathway whenever possible. Despite the consistent activity that we observed for solution bound IgM, these results do <u>not</u> necessarily infer that solution bound IgM is a significant contributor to complement activation *in vivo*. As discussed previously, the activation of C1 can be a potentially deceiving endpoint as it cannot be simply extrapolated to viable downstream MAC formation. Indeed, the complement system remains a highly complex and challenging topic of research.

5.2 Future Directions

While there are many fundamental questions that still need to be addressed before the therapeutic potential of IgM can become fully realized, there are at least three primary areas of research that we are planning to address in the near future using the instrumentation, methodologies, and experience that we have gained throughout the establishment of the IgM project presented within this dissertation. These areas include: 1) the continued assessment of further structural changes induced within IgM upon binding to C1q/C1, 2) a direct structure/function comparison between pentameric and hexameric IgM isoforms, and 3) a deeper look into the nuances of antigen presentation and variables that modulate the IgM-mediated activation of complement.

Studies are planned for the immediate future in which both C1q and the whole C1 complex will be incorporated into the HDX-MS methods that have now been established for surface bound IgM, in order to assess the regions that change further upon their binding interactions. Now that we have identified a number of key regions that appear to be structurally unique to the complement-active form of surface bound IgM, we can begin to proceed with further assessments that specifically target those areas. Additionally, owing to the versatility of HDX-MS itself, we can also orthogonally examine the binding interactions from the perspective of C1q, C1r, and C1s proteins, in addition to the assessment of changes within IgM itself. This will allow us to robustly address a number of structural questions that surround the activation of complement; not only will we be able to further elucidate the identity of the residues that comprise the C1q binding site(s) in IgM, but we will also be able to locate changes within C1q that could point toward the identity of its IgM binding site(s) as well. We can also immediately look for any structural differences induced by the binding of C1q compared to that of the whole C1 complex, which could also point us towards any potential physical involvement of the $C1r_2C1s_2$ tetramer, which has been implicated recently within the literature to potentially stabilize the binding interaction of C1 with IgG hexamer mutants.[8]

Despite being straightforward in concept, the actual incorporation of C1q/C1 into surface HDX-MS experiments is a non-trivial pursuit that will likely require further method optimizations. On top of the sheer increase in generated peptides that could require enhanced analytical separation techniques to resolve, the 'fast on/fast off' nature of C1q/C1 binding that we observed in Chapter 3 will also need to be addressed in order to achieve a sufficient temporal range of HDX labeling timepoints. Fortunately, HDX is a highly versatile method of differential structural analysis, and is also readily amenable to various exchange conditions. To that end, decreasing the ionic strength of the deuterated buffer and sample solutions could provide at least one analytical strategy within the short-term that might significantly extend the range of viable timepoints, as we demonstrated via BLI (Chapter 3) through 'ionic locking' of C1q onto surface bound IgM. Caution will have to be exercised, however, as there is likely a point in which reduced ionic strength may provide false-positive results by the induction of binding to charged residues that are not biologically relevant or involved in the activation mechanism.

Another major question that we plan to address in the near future concerns the structural and functional differences between pentameric and hexameric IgM isoforms that underlie their vast differences in complement activity reported throughout the literature. While an additional binding arm is at least one obvious advantage of IgMh that would likely contribute towards its potency, there are still many other confounding variables that warrant their own detailed investigations. For example, the asymmetry of the pentameric isoform caused by incorporation of the J chain has only recently become appreciated in highly-resolved structural detail, and we find it plausible that the J chain itself could negatively impact the flexibility of at least one, if not both of the binding arms in IgMp that are attached to it via disulfide bridges at C575. This subtle but potentially significant impact would be in addition to the steric effects that likely result from J chain incorporation, which, depending on the orientation of the awaiting IgM Fc platform, might also restrict the ability of C1q globular heads to access the staple-exposed binding sites of some arms within surface bound IgMp.[5,9,10] There are also potential kinetic

differences that warrant consideration, such as the probability and kinetics of staple formation along antigenic surfaces, that we are now in a direct position to investigate using the combinatorial approach of BLI and C1 activity assays. Any potential functional differences we find there can then be tied further to potential structural features found by HDX-MS. To begin addressing these questions we have recently established a collaboration with the Pepper lab from the UW Department of Immunology, who will assist us in the production of recombinant human IgMs against a wide variety of clinically-relevant disease states, along with added expertise and procedural guidance for prerequisite work, such as the analytical separation and purification of the two IgM subclasses.[11,12]

Finally, the underlying nuances of antigen presentation will be further probed using highly advanced, self-assembling protein nanoparticles (NPs) in place of the rudimentary and imprecise streptavidin agarose beads that we used here to create surface Ag displays capable of facilitating the multivalent formation of complement-active IgM. In collaboration with the King lab from the UW Department of Biochemistry and Institute for Protein Design (IDP), we are now able to utilize these extremely precise NPs that are capable of displaying antigen with computationally predetermined valencies and finely-controlled spacing down to the single nanometer. These Ag presenting particles have also been demonstrated recently to be a potent vaccine platform, capable of facilitating the production of broadly neutralizing antibodies against deadly common pathogens such as respiratory syncytial virus (RSV) and influenza.[13,14] The King lab will be able to assist with the design, cloning, expression, and purification of NPs fused to various antigens that we can then employ within our IgM experiments, which will vastly improve our experimental capability to investigate many of the biophysical details that underlie the critical role of the incident antigen display, which we hypothesize to be fundamental to the IgM-mediated activation mechanism of complement.

For example, NPs of various geometries and Ag densities could first be screened directly by our C1 activity assay for activation potency before being incorporated into HDX-MS experiments for the

determination of their structural effects on IgM, which will help to solidify the structure/function relationship and extent to which Ag presentation dictates the downstream activation of complement. We could also design NPs to indirectly dictate the effective binding valency of IgM by altering the spacing of individual Ag particles, which would essentially circumvent the need for a system of excessive Ag binding strength. While we were previously reliant on the strength of our transgenic IgM MD4 system to facilitate the prebinding approach, due to its greatly enhanced binding affinity derived from the highly-specific IgG H10 Fabs, we will now be able to test weaker IgM systems which will help to improve biological relevancy.[15] While each IgM system will require its own respective amount of preliminary testing and further method optimizations, the NPs will provide us with an entirely new avenue of research and allow us to probe nuances that would otherwise be experimentally unobtainable.

5.3 References

1. Haury, M. *et al.* The repertoire of serum IgM in normal mice is largely independent of external antigenic contact. *Eur. J. Immunol.* **27**, 1557–1563 (1997).

2. Boes, M. Role of natural and immune IgM antibodies in immune responses. *Mol. Immunol.* **37**, 1141–1149 (2001).

3. Das, S., Hirano, M., Tako, R., Mccallister, C. & Nikolaidis, N. Evolutionary Genomics of Immunoglobulin-Encoding Loci in Vertebrates. 95–102 (2012).

4. Gao, Z., Li, M., Ma, J. & Zhang, S. An amphioxus gC1q protein binds human IgG and initiates the classical pathway : Implications for a C1q-mediated complement system in the basal chordate. *Eur. J. Immunol.* **44**, 3680–3695 (2014).

5. Oskam, N. *et al.* At Critically Low Antigen Densities, IgM Hexamers Outcompete Both IgM Pentamers and IgG1 for Human Complement Deposition and Complement-Dependent Cytotoxicity. *J. Immunol.* **209**, 16–25 (2022).

6. Collins, C., Tsui, F. W. L. & Shulman, M. J. Differential activation of human and guinea pig complement by pentameric and hexameric IgM. *Eur. J. Immunol.* **32**, 1802 (2002).

7. Azuma, Y. *et al.* Recombinant human hexamer-dominant IgM monoclonal antibody to ganglioside GM3 for treatment of melanoma. *Clin. Cancer Res.* **13**, 2745–2750 (2007).

8. Zwarthoff, S. A. *et al.* C1q binding to surface-bound IgG is stabilized by C1r2s2 proteases. *Proc. Natl. Acad. Sci.* **118**, (2021).

9. Chen, Q., Menon, R., Calder, L. J., Rosenthal, P. B. & Tolar, P. Cryomicroscopy reveals the structural basis for a flexible hinge motion in the immunoglobulin M pentamer. *Nat. Commun.* **13**, 1–11 (2022).

10. Wiersma, E. J., Collins, C., Fazel, S. & Shulman, M. J. Structural and functional analysis of J chain-deficient IgM. *J. Immunol.* **160**, 5979–89 (1998).

11. Hale, M. *et al.* IgM antibodies derived from memory B cells are potent cross-variant neutralizers of SARS-CoV-2. *J. Exp. Med.* **219**, (2022).

12. Walls, A. C. *et al.* Elicitation of Potent Neutralizing Antibody Responses by Designed Protein Nanoparticle Vaccines for SARS-CoV-2. *Cell* **183**, 1367–1382 (2020).

13. Marcandalli, J. *et al.* Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus. *Cell* **177**, 1420–1431 (2019).

14. Boyoglu-barnum, S. *et al.* Quadrivalent influenza nanoparticle vaccines induce broad protection. *Nature* **592**, 623–628 (2021).

15. Kam-morgan, L. N. W. *et al.* High-resolution mapping of the HyHEL-10 epitope of chicken lysozyme by site-directed mutagenesis. *Proc Natl Acad Sci U S A* **90**, 3958–3962 (1993).