Accepted Manuscript


Title: Cryptic Epitopes and functional diversity in Extracellular Proteins

Author: Gysell M. Mortimer Rodney F. Minchin

PII: S1357-2725(16)30330-2
DOI: http://dx.doi.org/doi:10.1016/j.biocel.2016.10.020
Reference: BC 5021

To appear in: *The International Journal of Biochemistry & Cell Biology*

Received date: 1-8-2016
Revised date: 24-10-2016
Accepted date: 25-10-2016

Please cite this article as: Mortimer, Gysell M., & Minchin, Rodney F., Cryptic Epitopes and functional diversity in Extracellular Proteins.*International Journal of Biochemistry and Cell Biology* http://dx.doi.org/10.1016/j.biocel.2016.10.020

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

**Review article**

# Cryptic Epitopes and functional diversity in Extracellular Proteins

Gysell M. Mortimer and Rodney F. Minchin*

*Laboratory for Molecular and Cellular Pharmacology, School of Biomedical Sciences, The University of Queensland, St. Lucia, Brisbane, Queensland, 4072, Australia*

Corresponding author:
Prof Rod Minchin
Email: r.minchin@uq.edu.au
Phone: 61-7-3346-1620
Fax: 61-7-3365-1766

**Running title:** Cryptic epitopes and functional diversity in proteins

**Abstract**

The functional diversity of proteins is a major factor determining the complexity of cells and tissues. Both translational and post-translational modifications contribute to this diversity. Recently, protein unfolding and refolding has been recognised as another mechanism for diversity by unmasking buried or cryptic sequences (epitopes) that possess physiological functions. In the current review, we focus on extracellular proteins where folding dynamics can be influenced by mechanical forces, protein-protein interactions and denaturation. Many cryptic epitopes in these proteins are exposed following proteolytic cleavage, but recent data indicate that unfolding/refolding play an important role in regulating the physiological behaviour of extracellular proteins. By understanding how and when hidden sequences are exposed, novel techniques for manipulating the function of these proteins may be uncovered.

**Keywords:** Protein unfolding, protein conformation, cryptic epitopes

## 1. Introduction

Proteins are essential to all aspects of cellular function. They catalyse enzymatic reactions, regulate gene expression and ion transport, communicate between the extracellular environment and the intracellular space, control the structure, shape and mobility of cells, and modulate the movement of cells within tissues. A single gene can encode a protein with diverse functions depending on various events such as alternative splicing, post-translational modifications and compartmentalisation. Some proteins perform multiple tasks that are mechanistically distinct (moonlighting). For example, pyruvate kinase carboxylates pyruvate but, in some species, it also

ACCEPTED MANUSCRIPT

acts as a transporter for other proteins (Ozimek et al., 2003). Similarly, cytochrome C participates in mitochondrial electron transport but, in the cytosol, it forms a complex with Apaf-1 to initiate apoptosis (Hao et al., 2005). There are numerous examples of these multi-purpose proteins (Huberts and van der Klei, 2010). The different translational and post-translational modifications that a protein can undergo gives rise to functional diversity that contributes to the biological complexity of cells and tissues. In recent years, subtle changes in protein structure due to induced unfolding and refolding has been recognised as a further contributor to this functional diversity (Uversky, 2015).

The vast array of functions that proteins perform is enabled by the precise folding of an amino acid sequence into a specific but dynamic three-dimensional conformation. Differences between structural elements distinguish one protein from another and represent the most fundamental property affecting the biological activity of proteins. Correct biological functionality requires a protein to exist in its native conformation. Misfolding can lead to proteopathy, diseases caused by protein aggregation, which include amyloidosis, Alzheimer's disease, amyotrophic lateral sclerosis, tauopathies, retinitis pigmentosa, Huntington's disease and Parkinson's disease.

Proteins are structurally dynamic with ordered and disordered regions that unfold and refold over time and in response to various stimuli (Jakob et al., 2014). It has been widely reported that changes in the folding state of a protein can expose functional domains normally hidden or shielded within the protein. This has been best illustrated by antibody recognition of various antigenic determinants (epitopes) in misfolded proteins (Sela-Culang et al., 2013). Epitopes can be specific sequences (linear peptides) or motifs comprised of several structural elements in close proximity. Upon unfolding of a protein, the antigenic site (cryptic epitope) is exposed and elicits an immunological reaction (Laver et al., 1990). This is the basis for many auto-immune diseases. For example, the plasma transport protein transthyretin is involved in several forms of amyloidosis. Following unfolding and aggregation of transthyretin (Goldsteins et al., 1999; Phay et al., 2014), a normally buried β-strand is exposed (Gustavsson et al., 1994). Moreover, antibodies to this β-strand are common in patients with transthyretin amyloidosis but not in controls. The influenza coat protein hemagglutinin is another well-characterised example of a molecule that partially unfolds in order to elicit a biological effect, in this case, fusion with the endosomal membrane during infection (Huang et al., 2003).

While our understanding of cryptic epitope recognition by antibodies is well established, recently evidence has begun to emerge that similar hidden sequences responsible for non-antigenic biological activities are present in some proteins. This is particularly evident in extracellular matrix proteins (Schenk and Quaranta, 2003) where exposure of these buried motifs, referred to as 'matricryptic' sites, involves conformational changes and/or proteolytic cleavage (Davis et al., 2000). Some of these are central to the primary role of the protein while others create novel functions that are not obvious unless the buried peptide sequence is exposed. These cryptic epitopes increase the functional diversity of many extracellular and intracellular proteins (Fig 1). The difficulties in identifying functional cryptic epitopes are well reported but recent studies into the effects of protein binding to nanomaterials suggest a model for partially unfolding proteins that can be used to study their altered biological effects (Deng et al., 2011; Mortimer et al., 2014). The current review briefly summarises recent advances in our understanding of functional diversity focussing on key extracellular proteins.

## 2. Learning from extracellular matrix proteins: Matricryptic sites

Extracellular matrix proteins commonly carry out multiple and diverse biological functions, many of which are the result of cryptic site exposure (Maquart et al., 2004). Although these masked sequences can be activated following either proteolytic cleavage or conformational modifications (Davis et al., 2000; Tran et al., 2005), discussion here will focus on the latter. Readers are directed to other reviews that cover examples of cryptic sites in extracellular matrix proteins revealed by enzymatic digestion (Davis et al., 2000; Schenk and Quaranta, 2003). Exposure of bioactive sites by conformational modifications can be induced by several different mechanisms including mechanical

ACCEPTED MANUSCRIPT

force (Hocking et al., 2008), heterotypic binding to other molecules (Dardik and Lahav, 1999), self-assembly reactions (Mao and Schwarzbauer, 2005; Wierzbicka-Patynowski and Schwarzbauer, 2003), and denaturation. The unmasking of cryptic epitopes by molecular stretching has been well characterised for intracellular proteins such as talin and alpha-catenin (del Rio et al., 2009; Yonemura et al., 2010).

An important point for consideration regarding cryptic sites in extracellular matrix proteins is the gain in functionality. This is highlighted in work linking matricryptic site exposure and cell responses following tissue injury (Davis, 2010). Matricryptic sites can also be similar to those of pattern recognition receptors. For instance, DMBT1 (a basement membrane matrix protein also referred to as hesin) binds polyanionic ligands including LPS in a similar manner as scavenger receptors (End et al., 2009). It has been proposed that unmasking of these site that engage pattern recognition receptors helps regulate cell responses during tissue injury (Davis, 2010).

### 2.1. Collagen

Collagens are the most abundant proteins in the extracellular space involved in the organisation of the structural support to surrounding cells throughout the body (Di Lullo et al., 2002). Collagen has multiple diverse biological effects such as inhibition of angiotensin 1-converting enzyme (ACE), anti-angiogenesis activity, tumor growth inhibition, chemotaxis and oxidative stress relief (Banerjee et al., 2015). Cryptic binding sites contribute substantially to this functional diversity. The majority of these sites require proteolytic activation (Marneros and Olsen, 2001; Sasaki et al., 2002). However, conformational changes to collagen can also expose hidden binding domains. For instance, the pro-angiogenic HUIV26 cryptic epitope within the triple helix of collagen IV can be exposed by temperature-induced unfolding (Hangai et al., 2002; Xu et al., 2001) as well as ionizing radiation (Brooks et al., 2002). This motif has been linked to several biological effects of collagen, including cell migration, adhesion and angiogenesis (Favreau et al., 2014; Xu et al., 2001). Despite attempts to identify the HUIV26 epitope, the amino acid sequence remains unknown suggesting it may represent a non-contiguous binding domain. Importantly, this epitope has been shown to have a role in tumour metastasis and may be a novel target for drug development (Roth et al., 2006).

Mechanical forces transmitted across the collagen molecule can lead to localized unfolding (micro-unfolding) (Bourne et al., 2014). In one domain of the protein, located in the N-terminus of the alpha1 chain, mechanical force exposes a chymotrypsin cleavage site. Mutations in this region, as seen in Ehlers-Danlos syndrome and osteogenesis imperfecta, significantly alter the unfolding kinetics of the collagen molecule (Makareeva et al., 2006). Tissue injury and remodelling of extracellular proteins can also result in conformational changes in this region of the collagen protein (Leikina et al., 2002; Persikov and Brodsky, 2002).

Collagen is known to bind to cellular integrins via a cryptic epitope. This tripeptide sequence (arginine-glycine-aspartate or RGD) is exposed when collagen unfolds or is denatured (Engvall et al., 1978; Ingham et al., 1985). The RGD epitope plays a central role in the function of collagen as a cell signalling molecule and regulates cell proliferation and differentiation (Taubenberger et al., 2010). In MC3T3-E1 cells, partially denatured collagen engages $\alpha_5\beta_1$- and $\alpha_v$-integrins to promote phosphorylation of focal adhesion kinase (FAK), which eventually stimulates cell differentiation. Unfolding has also been shown to transform collagen type I from a β1 integrin-associated ligand to an αvβ3-dependent ligand due to exposure of an RGD epitope (Davis, 1992; Montgomery et al., 1994). A similar pathway has been identified for collagen-stimulated endothelial cell proliferation in angiogenesis (Ames et al., 2015).

Clearly, the collagens are multi-function proteins that play an important role in the remodelling of the extracellular matrix. Much of the diversity seen with the protein originates from

ACCEPTED MANUSCRIPT

cryptic epitopes that bind to a variety of proteins and receptors once the collagen molecule is partially unfolded.

### 2.2. Fibronectin

Two forms of fibronectin mediate a diverse array of biological effects. While the plasma form is involved in coagulation, most fibronectin functions are attributed to the less-soluble cellular form that interacts as part of the extracellular matrix. These functions include cell adhesion, growth, migration and differentiation as well as tissue responses to wound healing, malignant transformation, inflammation and haemostasis. There are various matricryptic sites within fibronectin, the majority of which are exposed following proteolysis (see Table 2 in Schenk *et al.*) (Schenk and Quaranta, 2003). However, there is evidence for unmasking of cryptic epitopes by partial unfolding. For example, incorporation of fibronectin in the extracellular matrix occurs by fibrillogensis, which involves self-association (polymerisation) of fibronectin into fibrils. Some of the sites mediating self-association are cryptic or partially cryptic and require conformational changes (such as mechanical stretching) to enable accessibility. Cell-driven increase in tension leads to unfolding and extension of the Fn III modules of the protein exposing cryptic polymerisation sites (Hao et al., 2005). Controlled strain experiments on fibronectin fibers show that stretch-induced exposure of cryptic sites increases gradually over the full range of fiber extension. This is accompanied by the onset of Fn III module unfolding (Klotzsch et al., 2009). The ability of fibronectin to unfold and reveal these sites is attributed to a lack of disulphide bonds in the Fn III modules, which imparts structural flexibility (Ohashi and Erickson, 2011). A partially cryptic disulphide isomerase activity has also been suggested in the C-terminal $FnI_{12}$ module that may catalyse the cross-linking of disulphide bonds. This intrinsic rearrangement then facilitates the incorporation of fibronectin into the extracellular matrix (Langenbach and Sottile, 1999).

### 2.3 Laminin

The laminins are a large family of matrix proteins that regulate cell adhesion, migration, proliferation and differentiation. They are heterotrimeric proteins with at least 15 different isoforms (DeBiase et al., 2006). Like many other extracellular matrix proteins, laminin is prone to unfolding under mechanical stress. Using atomic force microscopy, Nemes and colleagues demonstrated a stepwise progression in laminin unfolding indicative, but not proof, of a modular structure to the protein (Nemes et al., 2002). Temperature-dependent unfolding of laminin has been widely reported and used to study its secondary structure. Brooks' group used this approach to map cryptic epitopes within the molecules by screening for synthetic peptides that bound to the unfolded protein (Akalu et al., 2007). The epitope recognised by the peptide was selectively exposed in melanoma tumor tissue but not in normal tissue. Importantly, blockade of the cryptic site in the tumor tissue led to attenuation of angiogenesis, tumor growth and metastasis. These findings not only suggested that laminin unfolds under physiological conditions but that the exposed cryptic epitopes following unfolding are central to tumorigenesis, at least for the melanoma model used in the study.

## 3. Proteins not associated with the Extracellular Matrix

### 3.1 Vitronectin

The glycoprotein vitronectin is found in the circulation and has been associated with a range of functions including complement activation, coagulation and cell adhesion. The molecule exists is several conformational states depending on its degree of unfolding. Native monomeric vitronectin interacts with thrombin-antithrombin complexes and glycoprotein IIb/IIIa while the unfolded

ACCEPTED MANUSCRIPT

multimeric form binds to β-endorphin, PAI-1, urokinase receptor and collagen (Seiffert and Smith, 1997). This appears to be mediated through the somatomedin B domain in the N-terminal of the protein (Zhou, 2007). The native form does not bind integrins because the required RGD domain of vitronectin is buried adjacent to its N-terminal. Thus, perturbation of this region of the protein may expose the cryptic RGD site. Indeed, Seiffert and Smith have reported that PAI-1 binding to vitronectin activates the protein to recognise integrins (Seiffert and Smith, 1997).

Understanding cryptic epitopes not only reveals the intriguing functional diversity of proteins but also may identify potential drug targets. This is demonstrated with both laminin as described above and with vitronectin. Subtractive phage display technology has been used to identify antibodies that selectively bind to the unfolded active form of vitronectin but not to the native form (Bloemendal et al., 2004). Selectivity was demonstrated with intense immunostaining for the unfolded vitronectin in tumor tissue but not in normal tissue.

### 3.2 Fibrinogen

Fibrinogen is another extracellular protein that illustrates the role of cryptic sites. Of its multiple biological functions, clot formation is considered its primary role *in vivo*. Secondary physiological processes involve initiating, monitoring and managing inflammatory responses. Fibrinogen is a dimer comprised of three protein chains (α, β, and γ) containing three major structural regions: two D domains at the terminal ends of the chains and a centrally located E domain (Fig 2A). Conversion of fibrinogen to fibrin, a crucial event leading to exposure of these sites, occurs by thrombin-mediated removal of two pairs of fibrinopeptides (fibrinopeptides A & B) from the central region of the protein. This creates new N-terminal regions with exposed polymerisation sites required for the assembly of fibrin polymers in clot formation (Fig 2B).

Thrombin cleavage alters the folded state, or secondary structure of the protein, which results in exposure of multiple sites that bind proteins and cells. Binding sites in the D region for plasminogen and tPA (Aα148–160 and γ312–324) have been described that are functional in fibrin but not fibrinogen. This was shown by the lack of binding by monoclonal antibodies raised specifically for each site (Schielen et al., 1991; Schielen et al., 1989). The sites were masked in fibrinogen and were exposed as a result of intermolecular interaction between the D and E region upon fibrin formation. Using a number of fibrinogen and fibrin derived fragments, it was shown that only D region fragments corresponding to their fibrin converted form bound plasminogen, tPA and monoclonal antibodies for these sites. In addition, binding occurred in fragments forming the intermolecular interaction between the D and E regions, but not upon their dissociation. This showed that a change in conformation of the D region upon interaction with region E made the binding sites accessible (Yakovlev et al., 2000).

Conformational dependent cryptic sites are also found in other areas of the protein. As well as the three major domains comprising fibrinogen, there are also two compact surface associated regions called the αC-domains that interact with each other intramolecularly (Medved et al., 1985). Electron microscopy experiments showed that, upon fibrin formation, the conformation of the αC-domains changes so that they dissociate and interact intermolecularly rather than intramolecularly (Gorkun et al., 1994; Veklich et al., 1993). The intermolecular interaction of the αC-domains then promotes the formation of thicker aggregated fibers required for clotting. This inter to intra-molecular switch has also been shown to unmask sites within the αC-domains that bind plasminogen and tPA. As described for the D region, these sites only become accessible upon fibrinogen conversion to fibrin (Medved et al., 2001). Since fibrinogen does not bind tPA or plasminogen, these sites are cryptic in the fibrinogen αC-domains and only become available when the domains switch to their physiologically active conformation upon fibrin assembly (Tsurupa et al., 2011).

ACCEPTED MANUSCRIPT

The list of cryptic binding sites within fibrinogen continues to expand. Similar binding studies using recombinant fragments along with ELIZA and surface plasmon resonance studies have more recently identified cryptic sites for alpha(2)antiplasmin in the D-region and C-terminal region of the αC-domain (Tsurupa et al., 2010), as well as a cryptic site for apolipoprotein(a) in the αC-domain of fibrinogen (Tsurupa et al., 2003). Fibrinogen also contains a cryptic integrin binding site located in the γC region of the D domain (Altieri et al., 1993; Lishko et al., 2004). Removal of the C-terminal tail of this region exposes the $\alpha_M\beta_2$ site (Mac-1) for binding to the Mac-1 receptor (Lishko et al., 2002). The interaction of the Mac-1 site with its corresponding receptor on cells induces a pro-inflammatory response. Importantly, access to the $\alpha_M\beta_2$ binding site has been shown to occur upon proteolytic cleavage of this tail or upon surface-induced immobilisation of fibrinogen (Lishko et al., 2002). This suggests that the masking of the $\alpha_M\beta_2$ binding site by the C-terminal tail can be removed by proteolysis or by conformational alteration. These observations together with the conformational flexibility exhibited by this C-terminal region (Mosesson et al., 2001; Yee et al., 1997), has lead to a C-terminal "pull out" hypothesis (Yakovlev et al., 2000). Yakovlev and colleagues' proposal articulates how this conformational transition could control access of ligands to several fibrinogen binding sites and thus regulate the functional interactions of fibrinogen.

Many nanomaterials bind fibrinogen including metal oxides (Canoa et al., 2015; Deng et al., 2009) silicates (Marucco et al., 2014), gold nanoparticles (Deng et al., 2013; Deng et al., 2012), carbon black (Kendall et al., 2011), polymeric spheres (Cedervall et al., 2007), and single-walled carbon nanotubes (Song et al., 2006). In addition, fibrinogen is a commonly reported constituent of the 'hard corona' of nanoparticle-protein complexes following exposure to biological environments (Aggarwal et al., 2009; Karmali and Simberg, 2011). Binding of proteins to nanomaterials is often associated with partial unfolding. This has been demonstrated for fibrinogen where its interaction with polymer-coated gold nanoparticles led to the unfolding of the γC region of the D-domain resulting in the exposure of the peptide sequence that binds to the Mac-1 receptor (Deng et al., 2011). The extent of unfolding was dependent on the size of the nanoparticle as well as its charge (Deng et al., 2013). Moreover, other nanomaterials have been shown to elicit the same response (Deng et al., 2011). These observations are important as they provide a possible mechanistic basis for the inflammatory properties of some nanomaterials. However, they also suggest that precisely engineered nanoparticles may be a useful tool to identify and characterise cryptic epitopes in proteins associated with their surfaces. This principle has also been demonstrated with albumin, as outlined below.

### 3.3 Albumin

Albumin consists of a single polypeptide chain of 585 amino acids and is produced by the parenchymal cells of the liver. It is one of the few non-glycosylated serum proteins and the most abundant, constituting ~50% of the blood plasma volume. Albumin has been studied extensively due to its ability to reversibly bind and transport a vast array of molecules throughout the body, and hence influence multiple biological systems (Peters and Stewart, 2013). Albumin has several physiological functions, driven principally by its colloidal properties (Kragh-Hansen, 2013). Its primary function is to regulate blood volume by maintaining the colloid osmotic pressure of the blood compartment. It also influences fluid distribution of tissues providing 75% of the intravascular oncotic pressure (Doweiko and Nompleggi, 1991). Additional functions include a role in acid-base physiology owing to its 16 histidine imidazole residues (Agrafiotis et al., 2014; Watson, 1999), anti-oxidant properties by neutralising oxygen radicals (Anraku et al., 2013; Bruschi et al., 2013; Iwao et al., 2012), anticoagulation properties (Galanakis, 1992; Jorgensen and Stoffersen, 1979, 1980), and the binding and transport of ligands (Anand and Mukherjee, 2013; Dalvit et al., 2002; Sarver et al., 2005).

ACCEPTED MANUSCRIPT

The ability to unfold and refold (i.e conformational flexibility) is built into the structure of albumin so that it may alter its shape in response to changes in its environment (Kragh-Hansen et al., 2013). This allows albumin to bind many substances and operate as an excellent molecular chaperone (Anand and Mukherjee, 2013). The three-dimensional structure of albumin is often referred to as a globular heart-shape consisting of 3 homologous domains. These are subdivided into 2 contiguous helical subdomains A and B (Kragh-Hansen, 2013). The folding flexibility of albumin has been attributed to proline residues located between the subdomains that allow movement of one subdomain relative to another (Brown, 1976). Molecular flexibility also extends to the hydrophilic surface of the protein allowing a vast array of small molecules to bind. Because of the flexibility between subdomains, small molecule binding alters the spatial orientation between the subdomains which in turn affects the accessibility of other molecules for their binding sites (Quinlan et al., 2005).

The high internal mobility of albumin also controls exposure of a cryptic epitope that provides anti-apoptotic protection to the endothelium (Bolitho et al., 2007). CNBr fragmentation of human serum albumin (HAS) was shown to significantly increase this anti-apoptotic activity, which is distinct from the known $Cys^{34}$ radical scavenging mechanism. Conformational change of native albumin is required for exposure of the cryptic site, which is demonstrated by the ability of glycosylated HSA to inhibit the anti-apoptotic activity (Zoellner et al., 2009). The structural modifications produced in glycosylated HSA limit the high intramolecular flexibility of albumin, thus preventing cryptic site exposure.

Chemically-modified albumin is targeted to cell surface scavenger receptors where it is cleared from the circulation (Mikulikova et al., 2005). The major scavenger receptor involved appears to be SR-B1 (Binder et al., 2012; Miyazaki et al., 2002). The exact mechanism remains to be determined, but may involve the modified region of the protein or unfolding induced by the covalent binding of the chemical to the protein. Native albumin is not a ligand for the scavenger receptor family. By contrast, native albumin bound to the surface of various nanomaterials is recognised by scavenger receptors. It was initially suspected that nanoparticle bound albumin was ‘structurally damaged’ (Dutta et al., 2007). However, more recent studies have shown the unfolding of albumin is a critical step. Studies by Mortimer *et al.* (Mortimer et al., 2014) and Fleisher *et al.* (Fleischer and Payne, 2014) demonstrate that changes to the secondary structure of albumin due to nanoparticle binding directs recognition by receptors and hence uptake into cells. Anionic polystyrene nanoparticles bound with albumin were internalised by albondin expressing cells, while cationic nanoparticles bound albumin and were taken up by scavenger receptor expressing cells (Fleischer and Payne, 2014). Mortimer and colleagues show that the binding of albumin to anionic layered silicate nanoparticles targets the nanoparticle-albumin complex to scavenger receptors (Mortimer et al., 2014). Circular dichroism demonstrated how the loss of helical content (i.e change in secondary structure) controls the recognition of complexes by scavenger receptors. Much work has been done at the molecular level to characterise the binding sites of albumin for various molecules as well as how the binding of these molecules alters structural conformation and functionality of the protein (Kragh-Hansen et al., 2013). Changes in protein conformation can be accompanied by the exposure of new peptide sequences. As pointed out by Fleisher *et al.*, separating epitope exposure from altered protein conformation is difficult. Mortimer *et al.* addressed this by showing that an alternate process to nanoparticle-induced unfolding of albumin, heat denaturation, could similarly expose an epitope that bound to scavenger receptors. Furthermore, separate expression of each of the three albumin domains revealed that domains I and II were recognised by scavenger receptors to a greater extent following either heat denaturation or binding to nanoparticles. These insights suggest that the well-known clearance of structurally damaged albumin from the circulation by macrophages is a result of cryptic epitopes designed to be detected by specific changes in protein structure. That the structural changes must be specific are highlighted by the fact that not all foreign material that binds albumin leads to scavenger receptor

ACCEPTED MANUSCRIPT

recognition, as Fleisher *et al.* demonstrated by the unaltered secondary structure of albumin following binding to anionic polystyrene particles.

These findings indicate that the unfolding dynamics of albumin upon interaction with various surfaces serves to expose hidden epitopes, expanding the functionality of albumin to include a means of clearing various particles from the circulation via the scavenger receptor system. This example also demonstrates how cryptic site exposure operates in conjunction with related pattern recognition receptors to regulate cell responses - a proposal put forward by Davis *et al.* for extracellular matrix proteins and the recognition of matricryptic sites by pattern recognition receptors involved in tissue injury responses (Davis, 2010).

### 3.4 Complement Factor H

Complement factor H is a 150 kDa glycoprotein of the complement control protein superfamily (Perkins et al., 2010). It is included in this review because it has several binding sites for different ligands whose accessibility is dependent on unfolding of the protein. The molecule consists of 20 complement control protein domains connected by short linkers of 3-8 amino acid residues. Complement control protein domains enable folding flexibility that control interactions with other proteins and hence functionality (Okemefuna et al., 2009; Oppermann et al., 2006). Factor H functions as a complement regulator essential for controlling complement activation in plasma and on cell surfaces (Perkins et al., 2010). Complement is part of the innate immune system that provides a defence mechanism against pathogens such as bacteria. Complement activation can be initiated by three pathways: the classical, lectin or alternative pathways. The cascading events in each of these pathways ultimately generate the central complement component, C3 convertase, which produces C3b molecules (Ricklin et al., 2010). Factor H has the critical role of regulating C3b so that host protection against pathogens occurs without C3b over-activated damage to host cells (Bajic et al., 2015). Factor H controls C3b activity by inhibiting binding of components with C3b and by participating in its inactivation through proteolytic cleavage. Factor H presents five major ligand interactions that influence its complement regulatory activity, recently reviewed by Miller and colleagues (Perkins et al., 2012). The ligands include C3b, C3d, heparin, C-reactive protein, zinc and itself by self-association. There are 2 binding sites for each ligand except zinc which has several weak interaction sites. In addition, a third binding site for C3b has been proposed (Schmidt et al., 2008). Binding of heparin, C3d and C-reactive protein to Factor H regulates complement activation at host cell surfaces whereas zinc binding inhibits the regulatory activity of Factor H by promoting its aggregation (Perkins et al., 2012). Importantly, these binding sites are cryptic within Factor H, requiring specific conformational arrangement of the protein to enable accessibility (Morgan et al., 2011). A model of Factor H unfolding depicts the native protein as a compact structure with only the C-terminal domain binding sites for C3b, C3d and heparin accessible. Binding of ligands to the C-terminal domain initiates unfolding, exposing additional sites that facilitate further ligand binding and alteration in conformation (Oppermann et al., 2006).

Bacterial proteins such as PspC from Streptococcus pneumoniae bind with high affinity to Factor H and induce a conformational change that exposes a second cryptic binding domain for the C3b fragment of C3 (Herbert et al., 2015). The resulting complex enhances the functionality of Factor H as a suppressor of the complement system.

## 4 Cryptic epitopes and evolutionary conservation

Functional epitopes, whether explicitly displayed on the surface of a protein or hidden within its structure, may show evolutionary conservation if they are important for the physiological role of the protein. However, identifying these epitopes can be challenging depending on whether they are represented by a linear peptide sequence or a motif involving structural elements in close proximity. An excellent example of cryptic structural motifs with evolutionary conservation is the regions

ACCEPTED MANUSCRIPT

within the HIV-1 envelope glycoprotein that mediate cellular interactions required to infect target cells. These regions or epitopes provide the attachment and fusion points to target cells and hence are resistant to variation by selective pressure (Wyatt et al., 1998). Therefore the virus evolved the cryptic nature of these epitopes as a protective mechanism. This is shown by the conformational flexibility of the structures containing the binding regions that keep the sites hidden and only become accessible following certain sequential interactions (Rizzuto et al., 1998; Sullivan et al., 1998; Wu et al., 1996). Additional crypticity is provided by variable loops that mask these regions (Lusso et al., 2005; Wyatt et al., 1995) as well as considerable surface glycosylation to evade epitope recognition by the immune system (Wyatt et al., 1998).

Between species, protein homology is highest in those regions of functional importance as structural homology often parallels sequence homology (Schmitt et al., 2002). By contrast, regions of proteins that are buried tolerate lower sequence similarity so long as structure is maintained. Thus, buried sequences of high homology infer a physiological function. For cryptic epitopes, or matricryptic sites, that are linear and physiologically relevant, evolutionary conservation would be expected. This is demonstrated with the amino acid sequence in the γC region of the fibrinogen molecule that engages the Mac-1 receptor (Deng et al., 2011). The epitope is exposed following local unfolding. Alignment of the fibrinogen C chain from 15 vertebrate species shows greatest homology (>90%) at the sequence for the Mac-1 receptor binding epitope (Fig 3A). By contrast, homology is markedly lower in the flanking regions. This is an interesting observation that supports a physiological role for the cryptic epitope, as opposed to an aberrant response to protein unfolding. While evolutionary conservation does not prove that a buried region in a protein may have an effect if exposed, it does provide leads when the tools for discovering such sequences are lacking. When it was discovered that albumin could unfold and recognise the scavenger receptor, the binding epitopes were not specifically identified but several regions were suggested based on their anionic properties – amino acids 56-63, 249-256 and 292-301 (Mortimer et al., 2014). Alignment of these regions from 17 different vertebrate species (Fig 3B) shows good homology in the first sequence (85-90%), excellent homology in the second sequence (95-100%), but poor homology in the third sequence (~ 70%). This bioinformatics approach, in combination with site-directed mutagenesis, may assist in identifying potential cryptic epitopes in the future.

## 5 Nanoparticles and protein unfolding

Nanoparticles possess many unique properties that are the basis for their novel uses in engineering, product development, material science and medicine. For nanoparticles intended for medical applications, protein binding to their surface is well recognised as a major limitation. Surface-bound proteins not only alter the essential properties of the nanoparticles but can also determine their fate *in vivo* (Butcher et al., 2016; Hadjidemetriou et al., 2015; Sahneh et al., 2015; Schottler et al., 2016). However, as discussed above with fibrinogen and albumin, nanoparticles may provide a useful tool to identify and characterise cryptic epitopes in many different proteins. This concept was first proposed by Lynch et al (Lynch et al., 2006) and is dependent on the surface curvature and composition of the nanoparticle. As proteins bind, they can spread across the nanoparticle surface. While this results in a corona of protein adhering to the nanoparticle surface, it can also result in protein unfolding and exposure of cryptic epitopes. Proteins bound to nanoparticles often exhibit characteristics not normally associated with them, such as aggregation, receptor interaction and immunogenicity (Monopoli et al., 2012). When bound proteins are transformed into ligands for specific cell-surface receptors to which they do not normally interact, then the unfolding process on the surface of the nanoparticle is likely to have revealed a cryptic epitope. This has been the explanation for the binding of fibrinogen to the Mac-1 receptor and albumin to the scavenger receptor when associated with nanoparticles (Deng et al., 2011; Mortimer et al., 2014).

ACCEPTED MANUSCRIPT

Aggressive unfolding of many proteins exposes internal hydrophobic domains that cause rapid aggregation in an aqueous environment. When this occurs *in vivo*, toxic amyloidosis can result. Proteins such as b2-microglobulin aggregate following unfolding on various nanoparticle (Linse et al., 2007). However, with precisely engineered particles with different sizes and surface potentials, it may be possible to screen for epitopes by inducing more subtle unfolding. This was elegantly demonstrated with the partial unfolding of transferrin bound to polystyrene nanoparticles (Kelly et al., 2015). The challenge is to develop appropriate screening techniques that can identify functional epitopes following unfolding.

## 6 Conclusions

The identification of physiologically functional cryptic epitopes is increasing in the literature both for extracellular and intracellular proteins. The unfolding or conformational change required to unmask these epitopes is a process distinct from the more commonly described proteolytic cleavage mechanism. Unfolding is an important feature relevant to protein functionality because it can allow for refolding dynamics that cleavage events cannot. Furthermore, protein unfolding/refolding provides a crucial way of regulating the functionality of proteins and hiding biological activity until it is required. The extent to which a cryptic site is masked (i.e. the amount of folding required to keep it hidden) may also control biological activity. These factors may control the conformation constraints upon the secondary structure thus contributing to the extent to which a hidden site may be buried. There is also evidence to suggest that binding sites can be hidden as opposed to exposed upon unfolding.

There are many questions that need answering to further our knowledge in this field. Importantly, the experimental tools that will allow for the identification of important sites still need to be developed. The examples of nanoparticle induce-unfolding discussed in this review present the opportunity for using nanoparticles as a tool to study proteins and how they unfold in different environments. Using nanomaterials in this way is a more sophisticated and subtle method of providing an environment to induce unfolding than surface immobilisation or heat denaturation which can influence the entire protein.

**Acknowledgements.** The research of the authors is supported by grants from the Australian National Health and Medical Research council.

## References

Aggarwal, P., Hall, J.B., McLeland, C.B., Dobrovolskaia, M.A., McNeil, S.E., 2009. Nanoparticle interaction with plasma proteins as it relates to particle biodistribution, biocompatibility and therapeutic efficacy. Adv Drug Deliv Rev 61(6), 428-437.

Agrafiotis, M., Keklikoglou, I., Papoti, S., Diminikos, G., Diplaris, K., Michaelidis, V., 2014. Effect of the independent acid base variables on anion gap variation in cardiac surgical patients: a Stewart-Figge approach. ScientificWorldJournal 2014, 907521.

Akalu, A., Roth, J.M., Caunt, M., Policarpio, D., Liebes, L., Brooks, P.C., 2007. Inhibition of Angiogenesis and Tumor Metastasis by Targeting a Matrix Immobilized Cryptic Extracellular Matrix Epitope in Laminin. Cancer Research 67(9), 4353-4363.

Altieri, D.C., Plescia, J., Plow, E.F., 1993. The structural motif glycine 190-valine 202 of the fibrinogen gamma chain interacts with CD11b/CD18 integrin (alpha M beta 2, Mac-1) and promotes leukocyte adhesion. The Journal of Biological Chemistry 268(3), 1847-1853.

Ames, J.J., Contois, L., Caron, J.M., Tweedie, E., Yang, X., Friesel, R., Vary, C.P.H., Brooks, P.C., 2015. Identification of an endogenously generated cryptic collagen epitope (XL313) that may

ACCEPTED MANUSCRIPT

selectively regulate angiogenesis by an integrin-YAP mechano-transduction pathway. Journal of Biological Chemistry.
Anand, U., Mukherjee, S., 2013. Binding, unfolding and refolding dynamics of serum albumins. Biochimica et biophysica acta 1830(12), 5394-5404.
Anraku, M., Chuang, V.T., Maruyama, T., Otagiri, M., 2013. Redox properties of serum albumin. Biochimica et biophysica acta 1830(12), 5465-5472.
Bajic, G., Degn, S.E., Thiel, S., Andersen, G.R., 2015. Complement activation, regulation, and molecular basis for complement-related diseases. Embo J 34(22), 2735-2757.
Banerjee, P., Suguna, L., Shanthi, C., 2015. Wound healing activity of a collagen-derived cryptic peptide. Amino Acids 47(2), 317-328.
Binder, V., Holzer, M., El-Gamal, D., Dirnberger, S., Marsche, G., 2012. Uremia-induced lysine modifications transform plasma albumin into a high-density lipoprotein receptor inhibitor. BMC Pharmacology & Toxicology 13(Suppl 1), A40-A40.
Bloemendal, H.J., de Boer, H.C., Koop, E.A., van Dongen, A.J., Goldschmeding, R., Landman, W.J., Logtenberg, T., Gebbink, M.F., Voest, E.E., 2004. Activated vitronectin as a target for anticancer therapy with human antibodies. Cancer immunology, immunotherapy : CII 53(9), 799-808.
Bolitho, C., Bayl, P., Hou, J.Y., Lynch, G., Hassel, A.J., Wall, A.J., Zoellner, H., 2007. The anti-apoptotic activity of albumin for endothelium is mediated by a partially cryptic protein domain and reduced by inhibitors of G-coupled protein and PI-3 kinase, but is independent of radical scavenging or bound lipid. J Vasc Res 44(4), 313-324.
Bourne, J.W., Lippell, J.M., Torzilli, P.A., 2014. Glycation cross-linking induced mechanical-enzymatic cleavage of microscale tendon fibers. Matrix biology : journal of the International Society for Matrix Biology 34, 179-184.
Brooks, P.C., Roth, J.M., Lymberis, S.C., DeWyngaert, K., Broek, D., Formenti, S.C., 2002. Ionizing radiation modulates the exposure of the HUIV26 cryptic epitope within collagen type IV during angiogenesis. Int J Radiat Oncol Biol Phys 54(4), 1194-1201.
Brown, J.R., 1976. Structural origins of mammalian albumin. Federation Proceedings 35(10), 2141-2144.
Bruschi, M., Candiano, G., Santucci, L., Ghiggeri, G.M., 2013. Oxidized albumin. The long way of a protein of uncertain function. Biochimica et biophysica acta 1830(12), 5473-5479.
Butcher, N.J., Mortimer, G.M., Minchin, R.F., 2016. Drug delivery: Unravelling the stealth effect. Nature nanotechnology 11(4), 310-311.
Canoa, P., Simon-Vazquez, R., Popplewell, J., Gonzalez-Fernandez, A., 2015. A quantitative binding study of fibrinogen and human serum albumin to metal oxide nanoparticles by surface plasmon resonance. Biosens Bioelectron 74, 376-383.
Cedervall, T., Lynch, I., Foy, M., Berggard, T., Donnelly, S.C., Cagney, G., Linse, S., Dawson, K.A., 2007. Detailed identification of plasma proteins adsorbed on copolymer nanoparticles. Angewandte Chemie 46(30), 5754-5756.
Dalvit, C., Flocco, M., Stockman, B.J., Veronesi, M., 2002. Competition binding experiments for rapidly ranking lead molecules for their binding affinity to human serum albumin. Combinatorial Chemistry & High Throughput Screening 5(8), 645-650.
Dardik, R., Lahav, J., 1999. Functional changes in the conformation of thrombospondin-1 during complexation with fibronectin or heparin. Experimental cell research 248(2), 407-414.
Davis, G.E., 1992. Affinity of integrins for damaged extracellular matrix: alpha v beta 3 binds to denatured collagen type I through RGD sites. Biochem Biophys Res Commun 182(3), 1025-1031.
Davis, G.E., 2010. Matricryptic sites control tissue injury responses in the cardiovascular system: relationships to pattern recognition receptor regulated events. J Mol Cell Cardiol 48(3), 454-460.
Davis, G.E., Bayless, K.J., Davis, M.J., Meininger, G.A., 2000. Regulation of tissue injury responses by the exposure of matricryptic sites within extracellular matrix molecules. The American journal of pathology 156(5), 1489-1498.

ACCEPTED MANUSCRIPT

DeBiase, P.J., Lane, K., Budinger, S., Ridge, K., Wilson, M., Jones, J.C.R., 2006. Laminin-311 (Laminin-6) Fiber Assembly by Type I-like Alveolar Cells. Journal of Histochemistry & Cytochemistry 54(6), 665-672.
del Rio, A., Perez-Jimenez, R., Liu, R., Roca-Cusachs, P., Fernandez, J.M., Sheetz, M.P., 2009. Stretching single talin rod molecules activates vinculin binding. Science (New York, N.Y.) 323(5914), 638-641.
Deng, J., Sun, M., Zhu, J., Gao, C., 2013. Molecular interactions of different size AuNP-COOH nanoparticles with human fibrinogen. Nanoscale 5(17), 8130-8137.
Deng, Z.J., Liang, M., Monteiro, M., Toth, I., Minchin, R.F., 2011. Nanoparticle-induced unfolding of fibrinogen promotes Mac-1 receptor activation and inflammation. Nature nanotechnology 6(1), 39-44.
Deng, Z.J., Liang, M., Toth, I., Monteiro, M.J., Minchin, R.F., 2012. Molecular interaction of poly(acrylic acid) gold nanoparticles with human fibrinogen. ACS nano 6(10), 8962-8969.
Deng, Z.J., Mortimer, G., Schiller, T., Musumeci, A., Martin, D., Minchin, R.F., 2009. Differential plasma protein binding to metal oxide nanoparticles. Nanotechnology 20(45), 455101.
Di Lullo, G.A., Sweeney, S.M., Korkko, J., Ala-Kokko, L., San Antonio, J.D., 2002. Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. The Journal of Biological Chemistry 277(6), 4223-4231.
Doweiko, J.P., Nompleggi, D.J., 1991. Role of albumin in human physiology and pathophysiology. Journal of Parenteral and Enteral Nutrition 15(2), 207-211.
Dutta, D., Sundaram, S.K., Teeguarden, J.G., Riley, B.J., Fifield, L.S., Jacobs, J.M., Addleman, S.R., Kaysen, G.A., Moudgil, B.M., Weber, T.J., 2007. Adsorbed proteins influence the biological activity and molecular targeting of nanomaterials. Toxicol Sci 100(1), 303-315.
End, C., Bikker, F., Renner, M., Bergmann, G., Lyer, S., Blaich, S., Hudler, M., Helmke, B., Gassler, N., Autschbach, F., Ligtenberg, A.J., Benner, A., Holmskov, U., Schirmacher, P., Nieuw Amerongen, A.V., Rosenstiel, P., Sina, C., Franke, A., Hafner, M., Kioschis, P., Schreiber, S., Poustka, A., Mollenhauer, J., 2009. DMBT1 functions as pattern-recognition molecule for poly-sulfated and poly-phosphorylated ligands. Eur J Immunol 39(3), 833-842.
Engvall, E., Ruoslahti, E., Miller, E.J., 1978. Affinity of fibronectin to collagens of different genetic types and to fibrinogen. The Journal of Experimental Medicine 147(6), 1584-1595.
Favreau, A.J., Vary, C.P., Brooks, P.C., Sathyanarayana, P., 2014. Cryptic collagen IV promotes cell migration and adhesion in myeloid leukemia. Cancer medicine 3(2), 265-272.
Fleischer, C.C., Payne, C.K., 2014. Secondary structure of corona proteins determines the cell surface receptors used by nanoparticles. The Journal of Physical Chemistry B 118(49), 14017-14026.
Galanakis, D.K., 1992. Anticoagulant albumin fragments that bind to fibrinogen/fibrin: possible implications. Seminars in Thrombosis and Hemostasis 18(1), 44-52.
Goldsteins, G., Persson, H., Andersson, K., Olofsson, A., Dacklin, I., Edvinsson, A., Saraiva, M.J., Lundgren, E., 1999. Exposure of cryptic epitopes on transthyretin only in amyloid and in amyloidogenic mutants. Proceedings of the National Academy of Sciences of the United States of America 96(6), 3108-3113.
Gorkun, O.V., Veklich, Y.I., Medved, L.V., Henschen, A.H., Weisel, J.W., 1994. Role of the alpha C domains of fibrin in clot formation. Biochemistry 33(22), 6986-6997.
Gustavsson, A., Engstrom, U., Westermark, P., 1994. Mechanisms of transthyretin amyloidogenesis. Antigenic mapping of transthyretin purified from plasma and amyloid fibrils and within in situ tissue localizations. The American journal of pathology 144(6), 1301-1311.
Hadjidemetriou, M., Al-Ahmady, Z., Mazza, M., Collins, R.F., Dawson, K., Kostarelos, K., 2015. In Vivo Biomolecule Corona around Blood-Circulating, Clinically Used and Antibody-Targeted Lipid Bilayer Nanoscale Vesicles. ACS nano 9(8), 8142-8156.
Hangai, M., Kitaya, N., Xu, J., Chan, C.K., Kim, J.J., Werb, Z., Ryan, S.J., Brooks, P.C., 2002. Matrix metalloproteinase-9-dependent exposure of a cryptic migratory control site in collagen is required before retinal angiogenesis. The American journal of pathology 161(4), 1429-1437.

ACCEPTED MANUSCRIPT

Hao, Z., Duncan, G.S., Chang, C.C., Elia, A., Fang, M., Wakeham, A., Okada, H., Calzascia, T., Jang, Y., You-Ten, A., Yeh, W.C., Ohashi, P., Wang, X., Mak, T.W., 2005. Specific ablation of the apoptotic functions of cytochrome C reveals a differential requirement for cytochrome C and Apaf-1 in apoptosis. Cell 121(4), 579-591.
Herbert, A.P., Makou, E., Chen, Z.A., Kerr, H., Richards, A., Rappsilber, J., Barlow, P.N., 2015. Complement Evasion Mediated by Enhancement of Captured Factor H: Implications for Protection of Self-Surfaces from Complement. Journal of immunology (Baltimore, Md. : 1950) 195(10), 4986-4998.
Hocking, D.C., Titus, P.A., Sumagin, R., Sarelius, I.H., 2008. Extracellular matrix fibronectin mechanically couples skeletal muscle contraction with local vasodilation. Circ Res 102(3), 372-379.
Huang, Q., Sivaramakrishna, R.P., Ludwig, K., Korte, T., Bottcher, C., Herrmann, A., 2003. Early steps of the conformational change of influenza virus hemagglutinin to a fusion active state: stability and energetics of the hemagglutinin. Biochimica et biophysica acta 1614(1), 3-13.
Huberts, D.H.E.W., van der Klei, I.J., 2010. Moonlighting proteins: An intriguing mode of multitasking. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1803(4), 520-525.
Ingham, K.C., Landwehr, R., Engel, J., 1985. Interaction of fibronectin with C1q and collagen. Effects of ionic strength and denaturation of the collagenous component. European Journal of Biochemistry 148(2), 219-224.
Iwao, Y., Ishima, Y., Yamada, J., Noguchi, T., Kragh-Hansen, U., Mera, K., Honda, D., Suenaga, A., Maruyama, T., Otagiri, M., 2012. Quantitative evaluation of the role of cysteine and methionine residues in the antioxidant activity of human serum albumin using recombinant mutants. Iubmb Life 64(5), 450-454.
Jakob, U., Kriwacki, R., Uversky, V.N., 2014. Conditionally and transiently disordered proteins: awakening cryptic disorder to regulate protein function. Chemical reviews 114(13), 6779-6805.
Jorgensen, K.A., Stoffersen, E., 1979. Heparin like activity of albumin. Thrombosis Research 16(3-4), 569-574.
Jorgensen, K.A., Stoffersen, E., 1980. On the inhibitory effect of albumin on platelet aggregation. Thrombosis Research 17(1-2), 13-18.
Karmali, P.P., Simberg, D., 2011. Interactions of nanoparticles with plasma proteins: implication on clearance and toxicity of drug delivery systems. Expert Opin Drug Deliv 8(3), 343-357.
Kelly, P.M., Aberg, C., Polo, E., O'Connell, A., Cookman, J., Fallon, J., Krpetic, Z., Dawson, K.A., 2015. Mapping protein binding sites on the biomolecular corona of nanoparticles. Nature nanotechnology 10(5), 472-479.
Kendall, M., Ding, P., Kendall, K., 2011. Particle and nanoparticle interactions with fibrinogen: the importance of aggregation in nanotoxicology. Nanotoxicology 5(1), 55-65.
Klotzsch, E., Smith, M.L., Kubow, K.E., Muntwyler, S., Little, W.C., Beyeler, F., Gourdon, D., Nelson, B.J., Vogel, V., 2009. Fibronectin forms the most extensible biological fibers displaying switchable force-exposed cryptic binding sites. Proceedings of the National Academy of Sciences of the United States of America 106(43), 18267-18272.
Kragh-Hansen, U., 2013. Molecular and practical aspects of the enzymatic properties of human serum albumin and of albumin-ligand complexes. Biochimica et biophysica acta 1830(12), 5535-5544.
Kragh-Hansen, U., Minchiotti, L., Galliano, M., Peters, T., Jr., 2013. Human serum albumin isoforms: genetic and molecular aspects and functional consequences. Biochimica et biophysica acta 1830(12), 5405-5417.
Langenbach, K.J., Sottile, J., 1999. Identification of protein-disulfide isomerase activity in fibronectin. J Biol Chem 274(11), 7032-7038.
Laver, W.G., Air, G.M., Webster, R.G., Smith-Gill, S.J., 1990. Epitopes on protein antigens: misconceptions and realities. Cell 61(4), 553-556.
Leikina, E., Mertts, M.V., Kuznetsova, N., Leikin, S., 2002. Type I collagen is thermally unstable at body temperature. Proceedings of the National Academy of Sciences of the United States of America 99(3), 1314-1318.

ACCEPTED MANUSCRIPT

Linse, S., Cabaleiro-Lago, C., Xue, W.-F., Lynch, I., Lindman, S., Thulin, E., Radford, S.E., Dawson, K.A., 2007. Nucleation of protein fibrillation by nanoparticles. Proceedings of the National Academy of Sciences 104(21), 8691-8696.
Lishko, V.K., Kudryk, B., Yakubenko, V.P., Yee, V.C., Ugarova, T.P., 2002. Regulated unmasking of the cryptic binding site for integrin alpha M beta 2 in the gamma C-domain of fibrinogen. Biochemistry 41(43), 12942-12951.
Lishko, V.K., Podolnikova, N.P., Yakubenko, V.P., Yakovlev, S., Medved, L., Yadav, S.P., Ugarova, T.P., 2004. Multiple binding sites in fibrinogen for integrin alphaMbeta2 (Mac-1). J Biol Chem 279(43), 44897-44906.
Lusso, P., Earl, P.L., Sironi, F., Santoro, F., Ripamonti, C., Scarlatti, G., Longhi, R., Berger, E.A., Burastero, S.E., 2005. Cryptic nature of a conserved, CD4-inducible V3 loop neutralization epitope in the native envelope glycoprotein oligomer of CCR5-restricted, but not CXCR4-using, primary human immunodeficiency virus type 1 strains. J Virol 79(11), 6957-6968.
Lynch, I., Dawson, K.A., Linse, S., 2006. Detecting cryptic epitopes created by nanoparticles. Sci STKE 2006(327), pe14.
Makareeva, E., Cabral, W.A., Marini, J.C., Leikin, S., 2006. Molecular mechanism of alpha 1(I)-osteogenesis imperfecta/Ehlers-Danlos syndrome: unfolding of an N-anchor domain at the N-terminal end of the type I collagen triple helix. J Biol Chem 281(10), 6463-6470.
Mao, Y., Schwarzbauer, J.E., 2005. Fibronectin fibrillogenesis, a cell-mediated matrix assembly process. Matrix Biol 24(6), 389-399.
Maquart, F.X., Pasco, S., Ramont, L., Hornebeck, W., Monboisse, J.C., 2004. An introduction to matrikines: extracellular matrix-derived peptides which regulate cell activity. Implication in tumor invasion. Critical Reviews in Oncology/Hematology 49(3), 199-202.
Marneros, A.G., Olsen, B.R., 2001. The role of collagen-derived proteolytic fragments in angiogenesis. Matrix Biol 20(5-6), 337-345.
Marucco, A., Turci, F., O'Neill, L., Byrne, H.J., Fubini, B., Fenoglio, I., 2014. Hydroxyl density affects the interaction of fibrinogen with silica nanoparticles at physiological concentration. J Colloid Interface Sci 419, 86-94.
Medved, L., Tsurupa, G., Yakovlev, S., 2001. Conformational changes upon conversion of fibrinogen into fibrin. The mechanisms of exposure of cryptic sites. Ann N Y Acad Sci 936, 185-204.
Medved, L.V., Gorkun, O.V., Manyakov, V.F., Belitser, V.A., 1985. The role of fibrinogen alpha C-domains in the fibrin assembly process. FEBS Letters 181(1), 109-112.
Mikulikova, K., Miksik, I., Deyl, Z., 2005. Non-enzymatic posttranslational modifications of bovine serum albumin by oxo-compounds investigated by chromatographic and electrophoretic methods. J Chromatogr B Analyt Technol Biomed Life Sci 815(1-2), 315-331.
Miyazaki, A., Nakayama, H., Horiuchi, S., 2002. Scavenger Receptors that Recognize Advanced Glycation End Products. Trends in Cardiovascular Medicine 12(6), 258-262.
Monopoli, M.P., Aberg, C., Salvati, A., Dawson, K.A., 2012. Biomolecular coronas provide the biological identity of nanosized materials. Nature nanotechnology 7(12), 779-786.
Montgomery, A.M., Reisfeld, R.A., Cheresh, D.A., 1994. Integrin alpha v beta 3 rescues melanoma cells from apoptosis in three-dimensional dermal collagen. Proceedings of the National Academy of Sciences of the United States of America 91(19), 8856-8860.
Morgan, H.P., Schmidt, C.Q., Guariento, M., Blaum, B.S., Gillespie, D., Herbert, A.P., Kavanagh, D., Mertens, H.D., Svergun, D.I., Johansson, C.M., Uhrin, D., Barlow, P.N., Hannan, J.P., 2011. Structural basis for engagement by complement factor H of C3b on a self surface. Nature structural & molecular biology 18(4), 463-470.
Mortimer, G.M., Butcher, N.J., Musumeci, A.W., Deng, Z.J., Martin, D.J., Minchin, R.F., 2014. Cryptic epitopes of albumin determine mononuclear phagocyte system clearance of nanomaterials. ACS nano 8(4), 3357-3366.
Mosesson, M.W., Siebenlist, K.R., Meh, D.A., 2001. The structure and biological features of fibrinogen and fibrin. Including results for Annals N Y Acad Sci 936, 11-30.

ACCEPTED MANUSCRIPT

Nemes, C., Ramsden, J.J., Rozlosnik, N., 2002. The unfolding of native laminin investigated by atomic force microscopy. Physica A: Statistical Mechanics and its Applications 313(3–4), 578-586.
Ohashi, T., Erickson, H.P., 2011. Fibronectin aggregation and assembly: the unfolding of the second fibronectin type III domain. J Biol Chem 286(45), 39188-39199.
Okemefuna, A.I., Nan, R., Gor, J., Perkins, S.J., 2009. Electrostatic interactions contribute to the folded-back conformation of wild type human factor H. Journal of molecular biology 391(1), 98-118.
Oppermann, M., Manuelian, T., Jozsi, M., Brandt, E., Jokiranta, T.S., Heinen, S., Meri, S., Skerka, C., Gotze, O., Zipfel, P.F., 2006. The C-terminus of complement regulator Factor H mediates target recognition: evidence for a compact conformation of the native protein. Clin Exp Immunol 144(2), 342-352.
Ozimek, P., van Dijk, R., Latchev, K., Gancedo, C., Wang, D.Y., van der Klei, I.J., Veenhuis, M., 2003. Pyruvate carboxylase is an essential protein in the assembly of yeast peroxisomal oligomeric alcohol oxidase. Molecular biology of the cell 14(2), 786-797.
Perkins, S.J., Nan, R., Li, K., Khan, S., Miller, A., 2012. Complement factor H-ligand interactions: self-association, multivalency and dissociation constants. Immunobiology 217(2), 281-297.
Perkins, S.J., Nan, R., Okemefuna, A.I., Li, K., Khan, S., Miller, A., 2010. Multiple interactions of complement Factor H with its ligands in solution: a progress report. Advances in experimental medicine and biology 703, 25-47.
Persikov, A.V., Brodsky, B., 2002. Unstable molecules form stable tissues. Proceedings of the National Academy of Sciences of the United States of America 99(3), 1101-1103.
Peters, T., Jr., Stewart, A.J., 2013. Albumin research in the 21st century. Biochimica et biophysica acta 1830(12), 5351-5353.
Phay, M., Blinder, V., Macy, S., Greene, M.J., Wooliver, D.C., Liu, W., Planas, A., Walsh, D.M., Connors, L.H., Primmer, S.R., Planque, S.A., Paul, S., O'Nuallain, B., 2014. Transthyretin aggregate-specific antibodies recognize cryptic epitopes on patient-derived amyloid fibrils. Rejuvenation Res 17(2), 97-104.
Quinlan, G.J., Martin, G.S., Evans, T.W., 2005. Albumin: biochemical properties and therapeutic potential. Hepatology 41(6), 1211-1219.
Ricklin, D., Hajishengallis, G., Yang, K., Lambris, J.D., 2010. Complement: a key system for immune surveillance and homeostasis. Nat Immunol 11(9), 785-797.
Rizzuto, C.D., Wyatt, R., Hernandez-Ramos, N., Sun, Y., Kwong, P.D., Hendrickson, W.A., Sodroski, J., 1998. A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. Science (New York, N.Y.) 280(5371), 1949-1953.
Roth, J.M., Caunt, M., Cretu, A., Akalu, A., Policarpio, D., Li, X., Gagne, P., Formenti, S., Brooks, P.C., 2006. Inhibition of Experimental Metastasis by Targeting the HUIV26 Cryptic Epitope in Collagen. The American journal of pathology 168(5), 1576-1586.
Sahneh, F.D., Scoglio, C.M., Monteiro-Riviere, N.A., Riviere, J.E., 2015. Predicting the impact of biocorona formation kinetics on interspecies extrapolations of nanoparticle biodistribution modeling. Nanomedicine (London, England) 10(1), 25-33.
Sarver, R.W., Gao, H., Tian, F., 2005. Determining molecular binding sites on human serum albumin by displacement of oleic acid. Analytical Biochemistry 347(2), 297-302.
Sasaki, T., Hohenester, E., Timpl, R., 2002. Structure and function of collagen-derived endostatin inhibitors of angiogenesis. Iubmb Life 53(2), 77-84.
Schenk, S., Quaranta, V., 2003. Tales from the crypt[ic] sites of the extracellular matrix. Trends in Cell Biology 13(7), 366-375.
Schielen, W.J., Adams, H.P., van Leuven, K., Voskuilen, M., Tesser, G.I., Nieuwenhuizen, W., 1991. The sequence gamma-(312-324) is a fibrin-specific epitope. Blood 77(10), 2169-2173.
Schielen, W.J., Voskuilen, M., Tesser, G.I., Nieuwenhuizen, W., 1989. The sequence A alpha-(148-160) in fibrin, but not in fibrinogen, is accessible to monoclonal antibodies. Proceedings of the National Academy of Sciences of the United States of America 86(22), 8951-8954.

ACCEPTED MANUSCRIPT

Schmidt, C.Q., Herbert, A.P., Kavanagh, D., Gandy, C., Fenton, C.J., Blaum, B.S., Lyon, M., Uhrin, D., Barlow, P.N., 2008. A new map of glycosaminoglycan and C3b binding sites on factor H. Journal of immunology 181(4), 2610-2619.
Schmitt, S., Kuhn, D., Klebe, G., 2002. A New Method to Detect Related Function Among Proteins Independent of Sequence and Fold Homology. Journal of Molecular Biology 323(2), 387-406.
Schottler, S., Becker, G., Winzen, S., Steinbach, T., Mohr, K., Landfester, K., Mailander, V., Wurm, F.R., 2016. Protein adsorption is required for stealth effect of poly(ethylene glycol)- and poly(phosphoester)-coated nanocarriers. Nature nanotechnology 11(4), 372-377.
Seiffert, D., Smith, J.W., 1997. The cell adhesion domain in plasma vitronectin is cryptic. J Biol Chem 272(21), 13705-13710.
Sela-Culang, I., Kunik, V., Ofran, Y., 2013. The structural basis of antibody-antigen recognition. Front Immunol 4, 302.
Song, L., Meng, J., Zhong, J., Liu, L., Dou, X., Liu, D., Zhao, X., Luo, S., Zhang, Z., Xiang, Y., Xu, H., Zhou, W., Wu, Z., Xie, S., 2006. Human fibrinogen adsorption onto single-walled carbon nanotube films. Colloids Surf B Biointerfaces 49(1), 66-70.
Sullivan, N., Sun, Y., Sattentau, Q., Thali, M., Wu, D., Denisova, G., Gershoni, J., Robinson, J., Moore, J., Sodroski, J., 1998. CD4-Induced conformational changes in the human immunodeficiency virus type 1 gp120 glycoprotein: consequences for virus entry and neutralization. J Virol 72(6), 4694-4703.
Taubenberger, A.V., Woodruff, M.A., Bai, H., Muller, D.J., Hutmacher, D.W., 2010. The effect of unlocking RGD-motifs in collagen I on pre-osteoblast adhesion and differentiation. Biomaterials 31(10), 2827-2835.
Tran, K.T., Lamb, P., Deng, J.S., 2005. Matrikines and matricryptins: Implications for cutaneous cancers and skin repair. J Dermatol Sci 40(1), 11-20.
Tsurupa, G., Ho-Tin-Noe, B., Angles-Cano, E., Medved, L., 2003. Identification and characterization of novel lysine-independent apolipoprotein(a)-binding sites in fibrin(ogen) alphaC-domains. The Journal of Biological Chemistry 278(39), 37154-37159.
Tsurupa, G., Mahid, A., Veklich, Y., Weisel, J.W., Medved, L., 2011. Structure, stability, and interaction of fibrin alphaC-domain polymers. Biochemistry 50(37), 8028-8037.
Tsurupa, G., Yakovlev, S., McKee, P., Medved, L., 2010. Noncovalent interaction of alpha(2)-antiplasmin with fibrin(ogen): localization of alpha(2)-antiplasmin-binding sites. Biochemistry 49(35), 7643-7651.
Uversky, V.N., 2015. Functional roles of transiently and intrinsically disordered regions within proteins. The FEBS journal 282(7), 1182-1189.
Veklich, Y.I., Gorkun, O.V., Medved, L.V., Nieuwenhuizen, W., Weisel, J.W., 1993. Carboxyl-terminal portions of the alpha chains of fibrinogen and fibrin. Localization by electron microscopy and the effects of isolated alpha C fragments on polymerization. The Journal of Biological Chemistry 268(18), 13577-13585.
Watson, P.D., 1999. Modeling the effects of proteins on pH in plasma. J Appl Physiol (1985) 86(4), 1421-1427.
Wierzbicka-Patynowski, I., Schwarzbauer, J.E., 2003. The ins and outs of fibronectin matrix assembly. J Cell Sci 116(Pt 16), 3269-3276.
Wu, L., Gerard, N.P., Wyatt, R., Choe, H., Parolin, C., Ruffing, N., Borsetti, A., Cardoso, A.A., Desjardin, E., Newman, W., Gerard, C., Sodroski, J., 1996. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. Nature 384(6605), 179-183.
Wyatt, R., Kwong, P.D., Desjardins, E., Sweet, R.W., Robinson, J., Hendrickson, W.A., Sodroski, J.G., 1998. The antigenic structure of the HIV gp120 envelope glycoprotein. Nature 393(6686), 705-711.
Wyatt, R., Moore, J., Accola, M., Desjardin, E., Robinson, J., Sodroski, J., 1995. Involvement of the V1/V2 variable loop structure in the exposure of human immunodeficiency virus type 1 gp120 epitopes induced by receptor binding. J Virol 69(9), 5723-5733.

ACCEPTED MANUSCRIPT

Xu, J., Rodriguez, D., Petitclerc, E., Kim, J.J., Hangai, M., Moon, Y.S., Davis, G.E., Brooks, P.C., 2001. Proteolytic exposure of a cryptic site within collagen type IV is required for angiogenesis and tumor growth in vivo. The Journal of cell biology 154(5), 1069-1079.
Yakovlev, S., Makogonenko, E., Kurochkina, N., Nieuwenhuizen, W., Ingham, K., Medved, L., 2000. Conversion of fibrinogen to fibrin: mechanism of exposure of tPA- and plasminogen-binding sites. Biochemistry 39(51), 15730-15741.
Yee, V.C., Pratt, K.P., Cote, H.C., Trong, I.L., Chung, D.W., Davie, E.W., Stenkamp, R.E., Teller, D.C., 1997. Crystal structure of a 30 kDa C-terminal fragment from the gamma chain of human fibrinogen. Structure 5(1), 125-138.
Yonemura, S., Wada, Y., Watanabe, T., Nagafuchi, A., Shibata, M., 2010. alpha-Catenin as a tension transducer that induces adherens junction development. Nature cell biology 12(6), 533-542.
Zhou, A., 2007. Functional structure of the somatomedin B domain of vitronectin. Protein science : a publication of the Protein Society 16(7), 1502-1508.
Zoellner, H., Siddiqui, S., Kelly, E., Medbury, H., 2009. The anti-apoptotic activity of albumin for endothelium is inhibited by advanced glycation end products restricting intramolecular movement. Cell Mol Biol Lett 14(4), 575-586.

ACCEPTED MANUSCRIPT

Legend to Figures

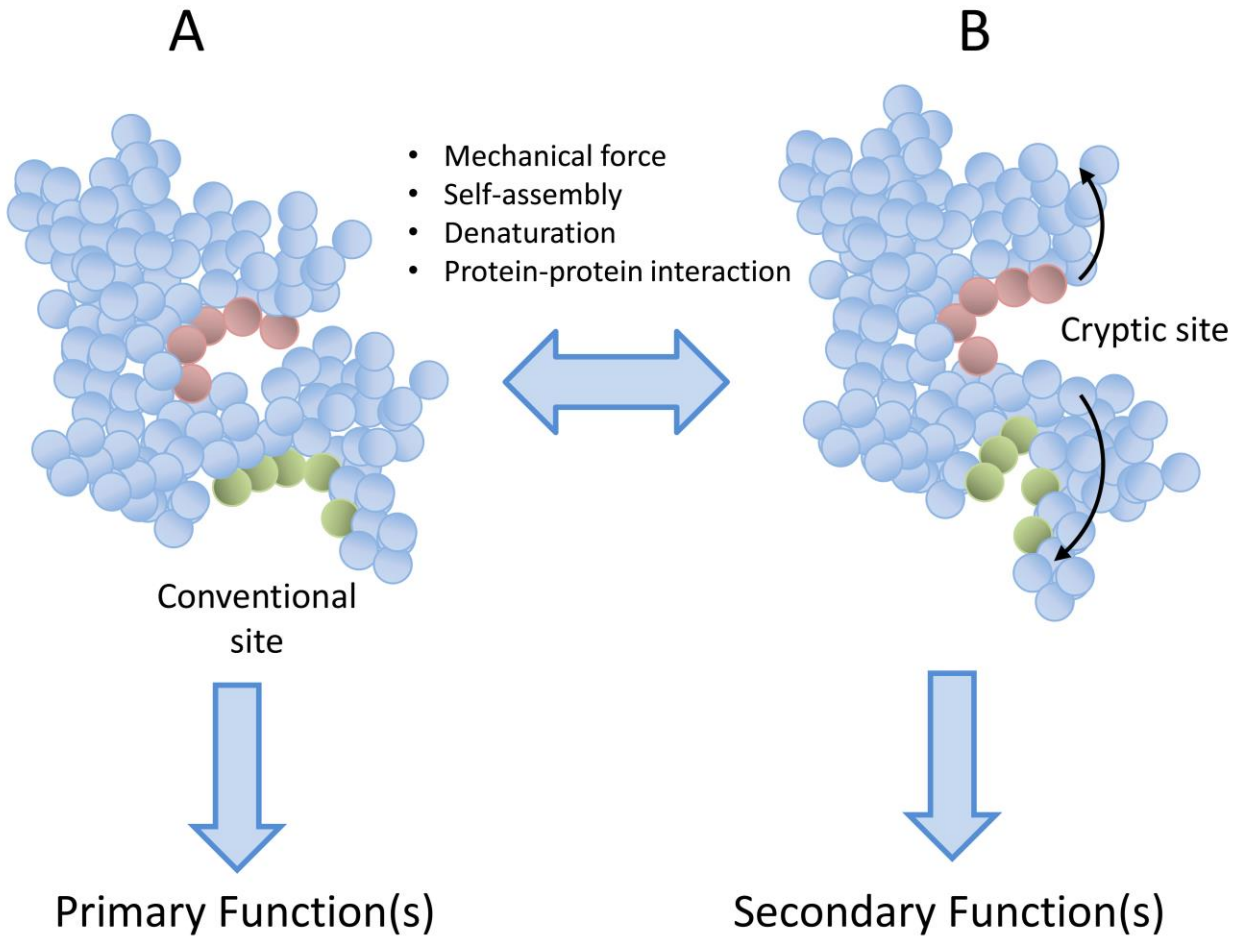


Figure 1. Representation of how buried sequences can be uncovered following unfolding induced by various factors such as mechanical forces or protein-protein interactions. Primary function of the protein is attributed to sites accessible when the protein is in its normal folded state. Secondary functions, which may be entirely different to the primary function, are only revealed following unfolding. Importantly, this process is reversible.

ACCEPTED MANUSCRIPT

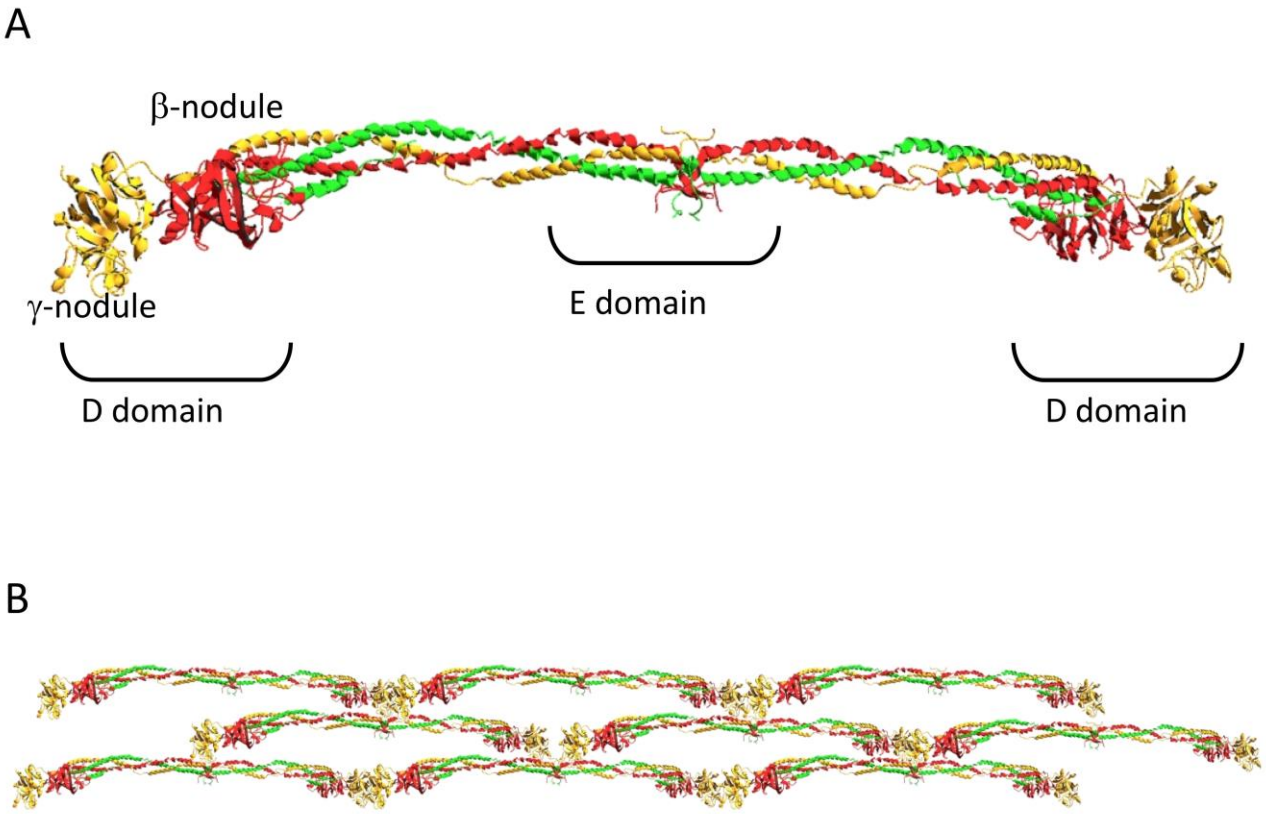


Figure 2. Cryptic epitopes in fibrinogen. A. The fibrinogen protein comprises 3 chains that are duplicated around a hinge region (E domain). A cryptic epitope that binds to the Mac-1 receptor is present in the γ-nodule. B. The protein polymerises by intermolecular interactions between the D domain and the E domain of the protein.

ACCEPTED MANUSCRIPT

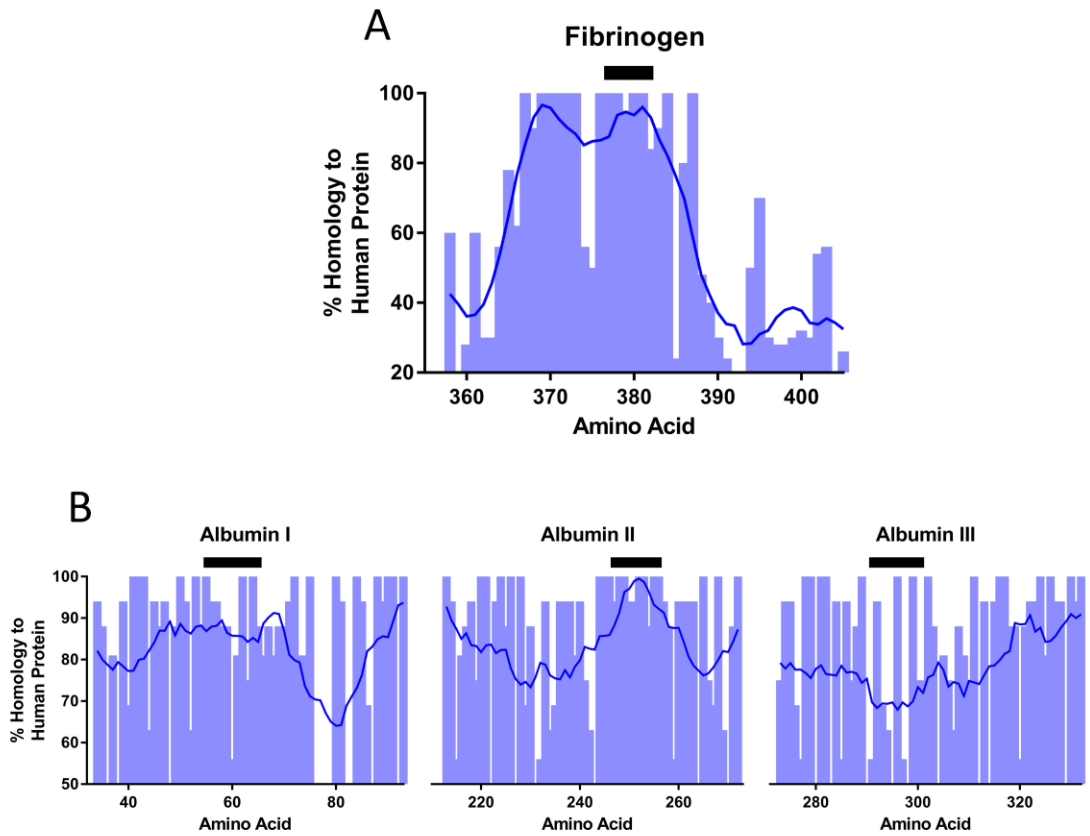


Figure 3. Conservation of buried sequences that comprise a cryptic epitope. A. The Mac-1 receptor binding sequence (black bar) in the C chain of fibrinogen is conserved across 15 different vertebrate species. This contrasts with flanking sequences where conservation is much less. B. Three anionic sequences buried in the albumin molecule and have been suggested as possible cryptic sites for scavenger receptor interactions (Mortimer et al., 2014). Two of these sequences (Albumin I: 56-63, Albumin II: 249-256) show high conservation between 17 vertebrate species while the third (Albumin III: 292-301) does not.