Structure-Function of *N*- and *O*-Glycosylated Human Corticosteroid-Binding Globulin

Anastasia Chernykh

A thesis submitted in partial fulfilment of the degree of Master of Research

Supervisor: Dr. Morten Thaysen-Andersen

Department of Molecular Sciences
Faculty of Science and Engineering
Macquarie University, Sydney, Australia

25th October 2019

Table of Contents

Statement of originality

The research presented within this thesis has been conducted between January 2019 and September 2019 for the completion of Master of Research degree in the Department of Molecular Sciences in the Faculty of Science and Engineering (FSE) at Macquarie University, New South Wales, Australia. This research thesis is certified to be an original work by the author, unless otherwise referenced in the literature and/or acknowledged in the text as personal advice or collaboration.

This thesis entitled "**Structure-Function of *N*- and *O*-Glycosylated Human Corticosteroid-Binding Globulin**" is formatted according to Master of Research Thesis guidelines prescribed by the Faculty of Science and Engineering and the Department of Molecular Sciences and has not been submitted for qualification or assessment to any other institution.

Thank You,

_____

Anastasia Chernykh, 25th Oct 2019

Student ID: 44474938

Acknowledgements

First of all, I would like to thank my supervisor Dr Morten Thaysen-Andersen for the great support throughout this year. Your extensive knowledge, deep curiosity and passion for science, and patience and calmness are the most encouraging things accompanying my experience during this year. Thank you for this extraordinary journey to the world of glycoimmunology. I am also thankful to Prof. Nicolle Packer for the guidance and all recommendations regarding my project. It is a great pleasure to work alongside you. I learned a lot from your creative vision and outstanding expertise.

Next, I would like to thank all members of the Glyco@MQ group for helping me this year. To Dr Rebeca Kawahara, Dr Zeynep Sumer-Bayraktar and Dr Hannes Hinneburg, for the valuable support in the experimental work. To Dr Krishnatej Nishtala and Dr Edward, for all extensive guidance and great assistance in my research. To Sayantani Chatterjee and Harry C. Tjondro for valuable suggestions and all consultation during this year. To Julian Ugonotti and Nick Debono, for words of support and help all the way through. To Kevin Chau, for all encouragements and conversations. To Dr Liisa Kautto, for the suggestions and training in the laboratory. To Dr Nima Sayyadi, for your extensive knowledge and discussions throughout the year.

I am also grateful to Dr Emily J. Meyer, Prof. David J. Torpy and Prof. Geoffrey L. Hammond for the great support and extensive collaboration throughout the year.

I would like to thank Macquarie University for the offering me the MQRES-MRES scholarship.

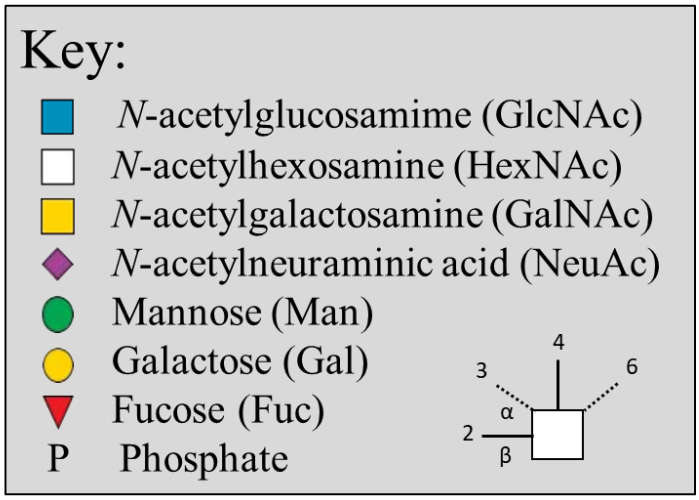
Lastly, I would like to thank my husband, Konstantin Madorskiy, and my parents, Iurii Chernykh and Tatiana Chernykh for the great support.

Abbreviations and symbol nomenclature

ACN	Acetonitrile
ACTH	Adrenocorticotropin
AGC	Automatic gain control
CBG	Corticosteroid-binding globulin
CID	Collision induced dissociation
CRH	Corticotropin-releasing hormone
Da	Dalton
DTT	Dithiothreitol
EIC	Extracted ion chromatogram
ER	Endoplasmic reticulum
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
EThcD	Electron-transfer/higher-energy collision dissociation
FA	Formic acid
FDR	False discovery rate
FWHM	Full width at half maximum
GalNAc-T	Polypeptide *N*-acetylgalactosamine transferase
GR	Glucocorticoid receptor
HCD	Higher-energy collision dissociation
HNF	Hepatocyte nuclear factor
HPA	Hypothalamic-pituitary-adrenal
HPLC	High performance liquid chromatography
IL	Interleukin
LC	Liquid chromatography
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NCE	Normalised collision energy
NE	Neutrophil elastase
OST	Oligosaccharyltransferase
PAE	*P. aeruginosa* elastase
PAGE	Polyacrylamide gel electrophoresis
PGC	Porous graphitised carbon
PNGase F	Peptide-*N*-glycosidase F
PVDF	Polyvinylidene fluoride
RCL	Reactive centre loop
SAX	Strong anion exchange
SCX	Strong cation exchange
SDS	Sodium dodecyl sulfate
SNP	Single nucleotide polymorphism
SPE	Solid-phase extraction
TBG	Thyroxin-binding globulin

Glycan Symbol Nomenclature

The monosaccharide symbols and the glycosidic linkage representation used to depict glycans in this thesis followed the convention outlined by the 3rd edition of the Essentials of Glycobiology (1).



Abstract

Corticosteroid-binding globulin (CBG), a heavily glycosylated liver-derived protein, is a principal carrier of anti-inflammatory cortisol in plasma. CBG impacts the cortisol bioavailability and facilitates tissue-specific and timely hormone delivery, and is, consequently, a recognised immunomodulatory glycoprotein with an untapped therapeutic potential. The intriguingly complex glycosylation covering the CBG surface is known to modulate its function, but the underpinning glycobiology remains incompletely understood. To this end, this thesis has employed advanced mass spectrometry-based glycan and glycopeptide profiling and functional assays to characterise the structure-function of two biologically-relevant forms of HEK293-produced recombinant human CBG (rhCBG) with a focus on documenting the glycosylation covering the functionally-important reactive centre loop (RCL). rhCBG exhibited strong site-specific *N*-glycosylation and occupancy differences across the six *N*-glycosylation sites. Key glyco-features for future therapeutic applications of rhCBG including mannose-6-phosphorylation, were identified. The RCL-localised Asn347 displayed very low *N*-glycan occupancy relative to the native counterpart. Excitingly, the proximal Thr345 was instead occupied by core 1/2-type *O*-glycosylation, a previously unknown feature of human CBG. Importantly, the novel Thr345-glycosylation was shown to strongly inhibit neutrophil elastase-mediated RCL proteolysis, thereby implying its involvement in the cortisol delivery mechanisms. This thesis has generated novel data that advances our knowledge of the fascinating CBG glycobiology.

Chapter 1: Introduction and thesis aims

This section aims to succinctly introduce the main scientific topics and concepts underpinning this thesis. Discipline-specific terminology, nomenclature and background literature useful to understand, interpret and reflect on the data presented in the Results and Discussion are introduced. Finally, this section provides a rationale for the research project and outlines the three thesis aims.

1.1 Tissue expression and regulation of CBG, the principal cortisol carrier

1.1.1 Gene expression of CBG and related proteins

Corticosteroid-binding globulin (CBG), the focus of this thesis, belongs to the clade A of the serine protease inhibitor (serpin) superfamily. Human CBG is encoded by *SERPINA6*, a gene that comprises five exons spanning more than 19 kb of genomic DNA and is located on chromosome 14q32.1 (2). *SERPINA6* is a part of the serpin gene cluster encompassing 11 serpin genes including key genes that encode α1-antitrypsin, α1-antichymotrypsin, kallikrein inhibitor, protein C1-inhibitor, thyroxin-binding globulin (TBG), angiotensinogen, centerin (serpin A9), and protein Z-dependent protease inhibitor (3–5). The serpin gene cluster is controlled by locus control regions consisting of multiple regulatory elements that are known to recruit different transcriptional factors. Of interest to the liver-derived serpin superfamily, the gene expression in hepatic tissues is regulated by a number of liver-specific transcription factors such as hepatocyte nuclear factor-1 (HNF-1), HNF-3, HNF-6 and the CAAT/enhancer-binding protein (C/EBP) (4). CBG is widely distributed amongst vertebrate species and is expressed in mammals, birds, reptiles, amphibians and fish (6).

1.1.2 Regulation of CBG expression by hormones and cytokines

Expression of human CBG is regulated via three main mechanisms (i-iii) discussed briefly below.

i. Glucocorticoids, which span a large class of steroid hormones, including cortisol, are key regulators of CBG production. Prolonged exposure to natural and synthetic glucocorticoids inhibits the production and secretion of CBG (7). The glucocorticoid-related regulation of CBG is mechanistically mediated by glucocorticoid receptors (GRs). GRs can bind directly to glucocorticoid response elements (GREs) at regulatory regions of the target gene, tether to other DNA-bound transcription factors regulating the expression of target genes or facilitate signals that may influence the transcription via non-genomic pathways (8). Specific GREs have not been identified for the human *SERPINA6* proximal promoter which indicates that glucocorticoids may downregulate the transcription of CBG not by means of direct binding to its promoter, but rather by using two other

mechanisms (9). Recently, the transcription factor C/EBPβ was shown to interact with GRs and thereby downregulate the CBG expression upon stimulation with glucocorticoids (10).

ii. Interleukin-6 (IL-6), a major inflammatory mediator, was also shown to decrease the synthesis of liver-derived CBG. Increased serum concentrations of IL-6 are associated with many immune-related disorders. During severe stress, for example, the body-wide stress associated with septic shock, surgery and in burn patients, the concentrations of serum IL-6 are significantly elevated; coincidently, these conditions are often associated with low CBG levels (11, 12). In support of the relationship between IL-6 and CBG, an *in vitro* study utilising human hepatoblastoma-derived HepG2 cells showed a dose-dependent inhibitory effect of IL-6 on CBG synthesis (13). The observed effect was associated with a posttranscriptional level regulation by IL-6, leading to a decrease in the CBG mRNA stability. Others showed an IL-6-mediated dose-dependent decrease in the CBG mRNA level, indicating an inhibitory role of IL-6 in the transcriptional regulation of CBG synthesis again (14).

It was also shown that stimulation of HepG2 with thyroxin, a non-glucocorticosteroid hormone produced by the thyroid gland, reduces the CBG secretion in a dose-independent manner (15). In the same study, stimulation of HepG2 with insulin resulted in a dose-dependent decrease in CBG secretion. In support, an earlier study demonstrated that insulin and insulin-like growth factor I decrease the CBG production and secretion from HepG2 cells by reducing the steady-state level of CBG mRNA in a dose-dependent manner (16).

iii. The synthesis of CBG is known to be significantly elevated in response to female sex hormones in particular to estrogen. For example, combined hormonal contraceptives which include estrogen and progestin components administrated orally or transdermally are known to increase the production of several estrogen-dependent hepatic proteins including sex hormone-binding globulin, CBG, TBG, C-reactive protein, AGT and apolipoprotein A1 (17, 18). Further support for the involvement of estrogen in the CBG expression was provided by the observation that circulating CBG levels are increased by a two-to-three fold in women during their second and third trimesters of pregnancy (19). The increased CBG synthesis by hepatocytes in pregnant women is most likely mediated by the hepatic estrogen receptor α (ERα) (20), a well-known transcription factor that together with estrogen binds to a target element of DNA and regulate the gene transcription (21).

1.1.3 Hepatic and extra-hepatic CBG expression across (patho)physiological conditions

CBG is primarily synthesised by hepatocytes, but can also be expressed in a variety of other tissues (22). Specifically, it was shown that CBG mRNA is present in lung, testis, kidney, heart, fallopian

tubes, endometrium and placenta indicating that these tissues may also produce CBG albeit presumably in much lower quantities than the liver (23–27).

The CBG production varies significantly with age and differs between genders. The CBG concentration in foetal blood is low. At full term, the neonatal CBG levels only amount to approximately half of the normal adult level of CBG. The plasma CBG levels in healthy adult men are 32.5 ± 9.1 mg/L, which are lower than the observed range for non-pregnant healthy women, i.e. 39.2 ± 13.9 mg/L, demonstrating an intriguing gender difference underpinning CBG biology (28). Notably, a significant two-to-three-fold increase in the plasma CBG levels occurs in pregnancy (29).

CBG levels differ significantly across a range of physiologies experienced by healthy individuals and in pathological states. CBG is a negative acute-phase protein playing a crucial role in the glucocorticoid response to severe stress (see above). In sepsis, plasma CBG and albumin levels are dramatically reduced by ~50%, which lead to an increased level of free cortisol (30). Thus, a significant decrease in the CBG concentration and a concomitant increase of free cortisol are found in early phases of septic shock and trauma (11). In patients with burn injuries, the serum CBG is decreased within two weeks post-injury with a concomitant increase in the free cortisol level (31).

1.2 Cortisol and CBG, key anti-inflammatory components of the HPA axis

1.2.1 The hormone carrier function of CBG and effects of CBG mutations

CBG is a carrier of the hydrophobic, and hence lowly soluble, glucocorticoids (primarily cortisol) and progesterone in human plasma. Under normal conditions, CBG binds with high affinity to ~80% of the serum cortisol in a 1:1 molar stoichiometry. Approximately 10-14% of serum cortisol is bound unspecifically to albumin, and less than 6% of serum cortisol is circulating as free cortisol, reflecting its low aqueous solubility (32). CBG binds hormones in a dedicated binding site; higher affinities have been demonstrated for cortisol (K_a 76 x 10^{-7} M^{-1}) than for progesterone (K_a 59 x 10^{-7} M^{-1}) (33).

Interestingly, a raised body temperature (up to 39°C) results in a decreased CBG binding affinity for cortisol, which, in turn, leads to at least a threefold increase in the free cortisol level in circulation (34). Upon returning to normal body temperature (37°C), the free cortisol concentration also returns to its normal level. Thus, this “thermo-couple” effect of CBG should be considered when determining the free cortisol concentrations and the underlying causes in donor and patient blood (34).

The relatively rarely occurring genetic point mutations in *SERPINA6* are also known to affect CBG production, secretion, steroid-binding affinity and capacity (35). For example, the single nucleotide

polymorphism (SNP) such as Leu93His (Leuven) and Asp367Asn (Lyon) are known to lead to a significant reduction in the steroid-binding affinity of CBG by a three- and four-fold, respectively, compared to native CBG (36, 37). Further, the Gly237Val and Trp371Ser substitutions were reported to lead to a complete loss of steroid binding activity of CBG resulting in generally increased concentrations of free cortisol, but low morning levels of serum cortisol (38, 39). It was also reported that a Glu102Gly substitution of CBG is associated with a decreased steroid-binding capacity (40).

Further, a genetic null mutation occurring in exon 2 of *SERPINA6* (Gly121Ala) was shown to lead to premature termination of the transcription of the gene (41). The resulting CBG deficiency was surprisingly not lethal but resulted in low total plasma level of cortisol; a condition referred to as secondary hypoadrenalism. It was shown that individuals homozygous and heterozygous for this deleterious CBG mutation have a high prevalence of fatigue and hypotension (41).

Some SNPs of CBG are particularly enriched in certain ethnic or geographically-isolated groups. For example, CBG Ala51Val is found in 1:35 Han Chinese (equal prevalence in males and females), a significantly higher prevalence than in other populations (40). This mutation leads to a secretion defect resulting in a 50% reduction of the plasma CBG level in heterozygous individuals.

Individuals suffering from the above CBG mutations are particularly valuable study subjects for CBG research. Understanding the genotype/phenotype of these CBG mutants relative to the wild type CBG is one commonly used pathway to advance our knowledge of the CBG function and its relevance in human health and disease. The above-mentioned mechanisms causing alterations in the CBG secretion and function are important to consider when diagnosing and treating individuals with disorders in the hypothalamic-pituitary-adrenal axis (HPA), a three-component axis that plays crucial regulatory roles related to stress response, and, hence, of high relevance to CBG biology.

1.2.2 Synthesis of glucocorticoids and their delivery to the sites of inflammation

The primary function of CBG is to transport and deliver cortisol to the target tissues. Under normal physiological conditions, glucocorticoid hormones are synthesised and released in a circadian manner (42). Light and food intake activate the HPA axis by stimulating the paraventricular nucleus of the hypothalamus, which, in turn, leads to the release of the corticotropin-releasing hormone (CRH), a 41 amino acid neuropeptide (**Figure 1**). CRH then initiates the cleavage of the pituitary pro-opiomelanocortin precursor polypeptide into several peptides including adrenocorticotropin (ACTH) with a subsequent release of these peptides into the blood. The release of ACTH triggers the synthesis

and secretion of glucocorticoids from the adrenal cortex (43, 44). The released free glucocorticoids, in turn, regulate the activity of the HPA axis and reduces the production of CRH and ACTH (45).

Cortisol regulates many biological processes including, but not limited to, the metabolic regulation, immunosuppression, development, modulation of biological functions of the reproductive system, brain, liver, heart, lungs, adipose tissue, stomach and muscles (8, 46). Taking into consideration the variety of physiological processes that are controlled by the glucocorticoids, it is exceedingly important that their concentrations are maintained at appropriate levels at all times. Excess levels of glucocorticoids are known to result in severe pathological conditions, including Cushing's syndrome, whereas the deficiency in glucocorticoid production is a cause of Addison's disease (8). Imbalance in the glucocorticoid production and turnover is a direct cause of altered energy metabolism, common symptoms underpinning the Cushing's syndrome and Addison's disease (47).

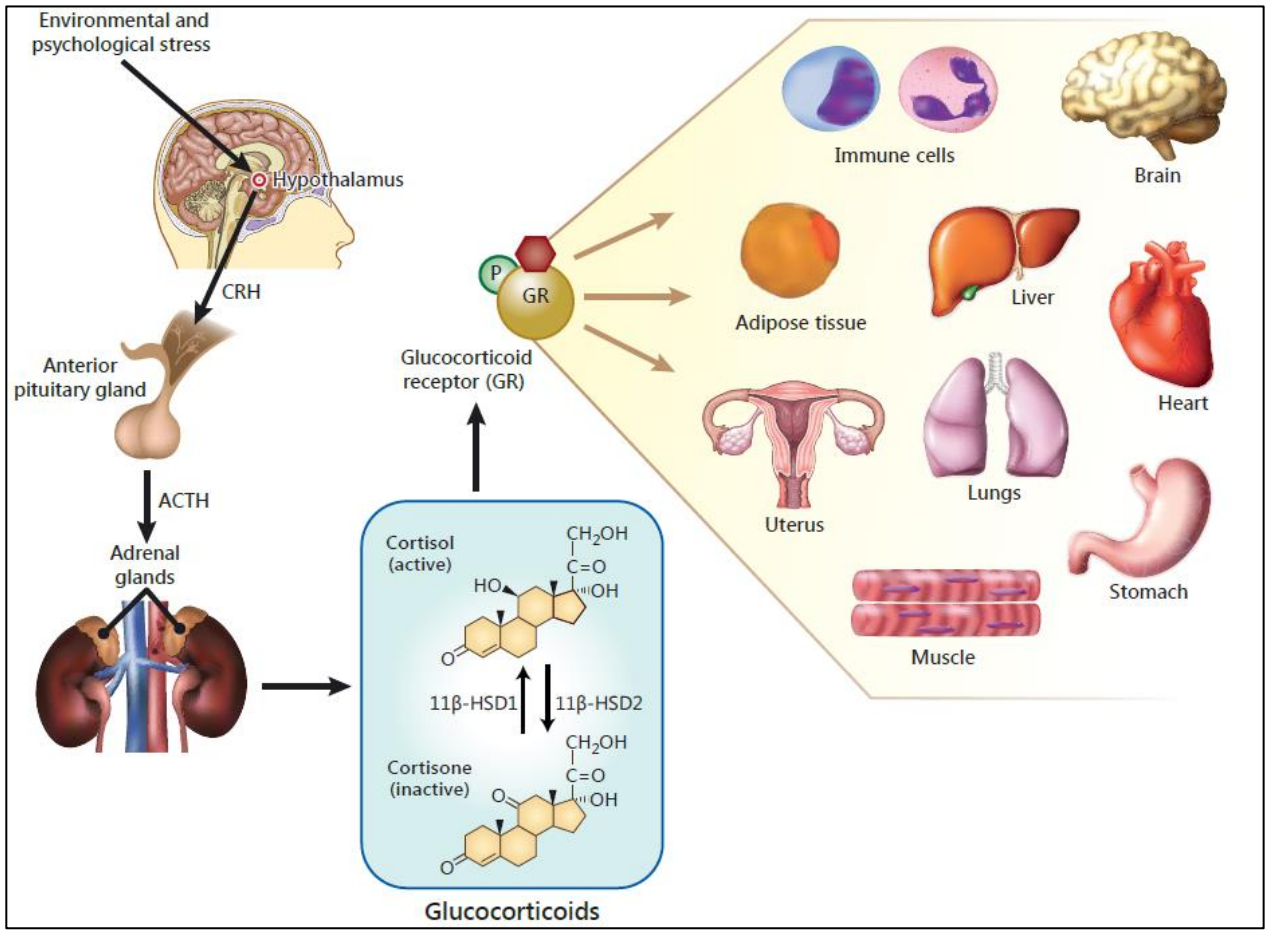


Figure 1. Secretion and effects of glucocorticoids (e.g. cortisol) produced by the hypothalamic-pituitary-adrenal (HPA) axis. Upon exposure to environmental or psychological stress, the hypothalamus is stimulated to release CRH, which stimulates the anterior pituitary gland to secrete ACTH. In turn, ACTH targets the cortex of the adrenal glands to release cortisol into the bloodstream. In circulation, cortisol can be converted to the inactive form, cortisone, by 11β-hydroxysteroid dehydrogenase type 2 (11β-HSD2). Conversely, 11β-HSD1 converts cortisone to cortisol. Upon binding to glucocorticoid receptors (GRs), glucocorticoids modulate the activity of several tissue types and organs and thereby impact many biological processes. From (8).

CBG plays a crucial role in cortisol delivery by regulating the cortisol bioavailability to target tissues under normal and stressed conditions (48, 49). According to a free-hormone hypothesis, only free cortisol can diffuse through target cell membranes and interact with dedicated GRs (50). Supporting this hypothesis, CBG has been shown to deliver cortisol in a timely manner to the extracellular space at sites of inflammation by means of a fascinating protease-mediated mechanism; the cortisol release from CBG is dependent on the proteolytic cleavage of the reactive centre loop (RCL) of CBG by a potent serine protease, neutrophil elastase (NE) (**Figure 2**) (51). After release, the free hormone can diffuse across the cell membrane and induce its immunomodulatory effects via interactions to the GRs.

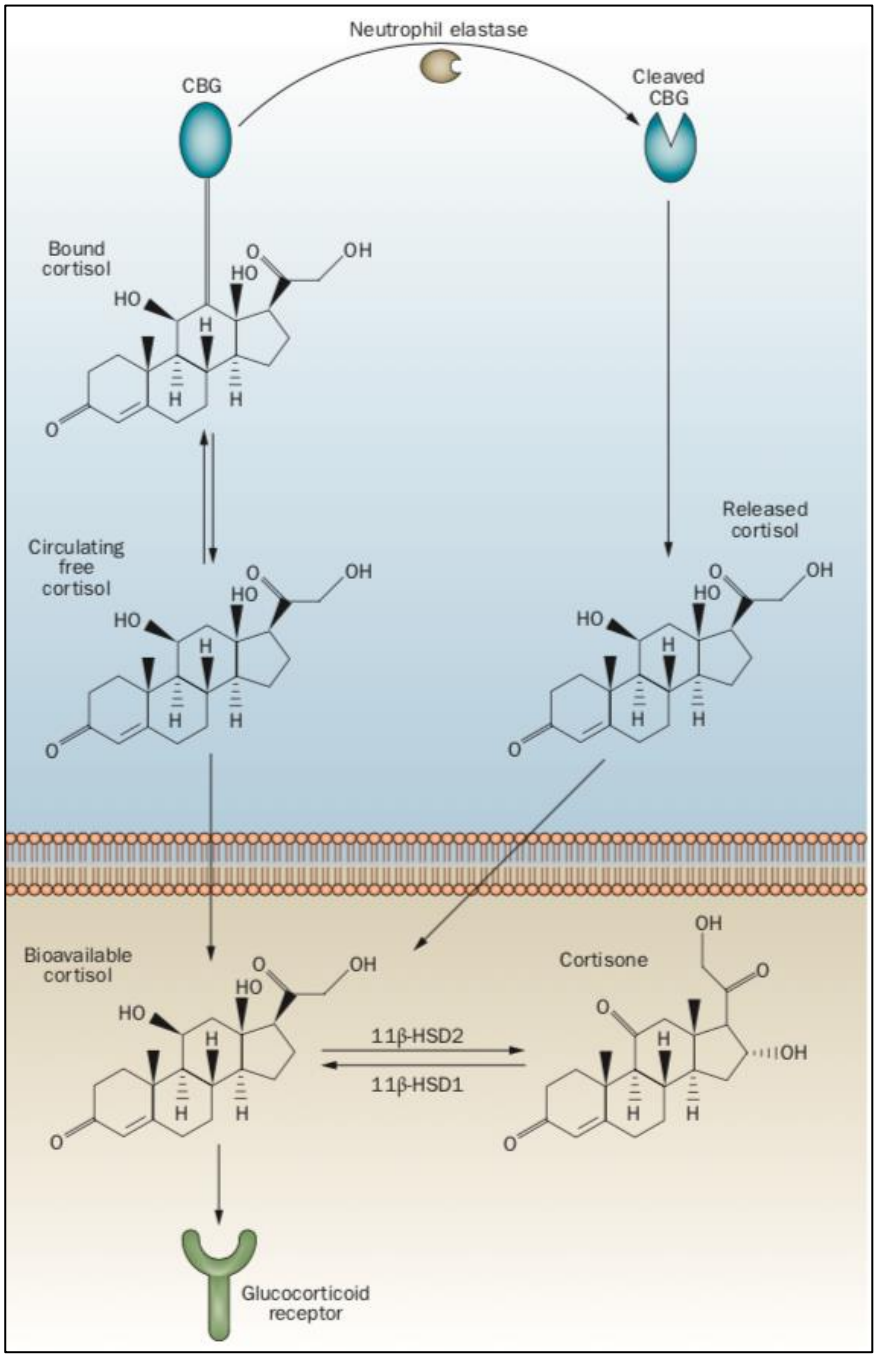


Figure 2. Regulation of the cortisol bioavailability at the site of inflammation. The circulating free cortisol is in equilibrium with CBG-bound cortisol. At sites of inflammation, CBG is cleaved by neutrophil elastase (NE) or other proteases released by neutrophils homing to this area, which results in the release of cortisol. Only free cortisol can diffuse across the plasma membrane of the cell and interact with dedicated glucocorticoid receptors (GRs). In tissues with 11β-HSD2 activity, cortisol may be reversibly inactivated by the partial conversion to cortisone. From (46).

The anti-inflammatory activity of corticosteroids is associated with their interaction with intracellular GRs and the subsequent inhibition of the synthesis and expression of pro-inflammatory proteins (8). GRs can also physically interact with other pro-inflammatory transcription factors, e.g. nuclear factor-kB and activator protein-1, which leads to the suppression of the activity of these factors and, consequently, inhibits the transcription of the genes of pro-inflammatory proteins (52).

It is suggested that glucocorticoids have both pro- and anti-inflammatory roles (53). The balance of these opposite effects allows the organism to elicit a proper response to stress and quick restoration of homeostasis. In this regard, both natural and synthetic glucocorticoids including hydrocortisone are utilised for the treatment of inflammation, autoimmune disorders and severe lymphoid malignancies including leukaemia, lymphoma, myeloma, and in organ transplantation (53). The therapeutic potential of recombinant CBG-cortisol will be discussed later in the thesis.

1.3 Structural features of CBG, an unusual serpin member

1.3.1 Primary, secondary and tertiary structure of CBG

CBG is a soluble monomeric glycoprotein that contains six utilised *N*-linked glycosylation sites (discussed in Section 1.4). CBG is synthesised as a 405 amino acid precursor polypeptide containing an N-terminal signal peptide of 22 amino acid residues. Mature CBG contains 383 amino acid residues following the removal of the signal peptide. The two cysteine residues present in the polypeptide chain are not linked via a disulphide bridge leaving open the question of whether these reactive functional groups are engaged in any yet-to-be-elucidated inter-molecular disulphide bonding (23, 54). A slight over-representation of negatively charged amino acid residues contributes to the acidic isoelectric point (pI) of CBG (pI 5.6), which further is modulated by the acidic oligosaccharides decorating CBG (55).

Being a member of the serpin family, CBG adopts the common three-dimensional structure characteristic of serpins, but, importantly, does not confer protease inhibition (56). CBG has three β-sheets, ten α-helices and a surface-exposed RCL (**Figure 3**). In concert, the β-sheet B and the α-helices A and H create a hydrophobic steroid-binding pocket.

The crystal structure of human cleaved CBG in complex with progesterone solved at high resolution (PDB: 4BB2, 2.48 Å) has revealed many important details of the hormone-binding mechanism (**Figure 4**) (57). A molecular hallmark of the hormone binding is the stacking of the indole ring of Trp371 against the cyclohexane ring of progesterone. The hydrogen bond between progesterone and CBG is formed by the carbonyl oxygen at C-20 and the side chain of Gln232. The

additional hydroxyl groups attached to C-11, C-17 and C-21 of cortisol form additional hydrogen bonds with Asn264, His368, and a water molecule coordinated between Gln224 and Gly259 of CBG. The multiple hydrogen bonds between cortisol and CBG explain the higher binding affinity for cortisol as compared to progesterone (57).

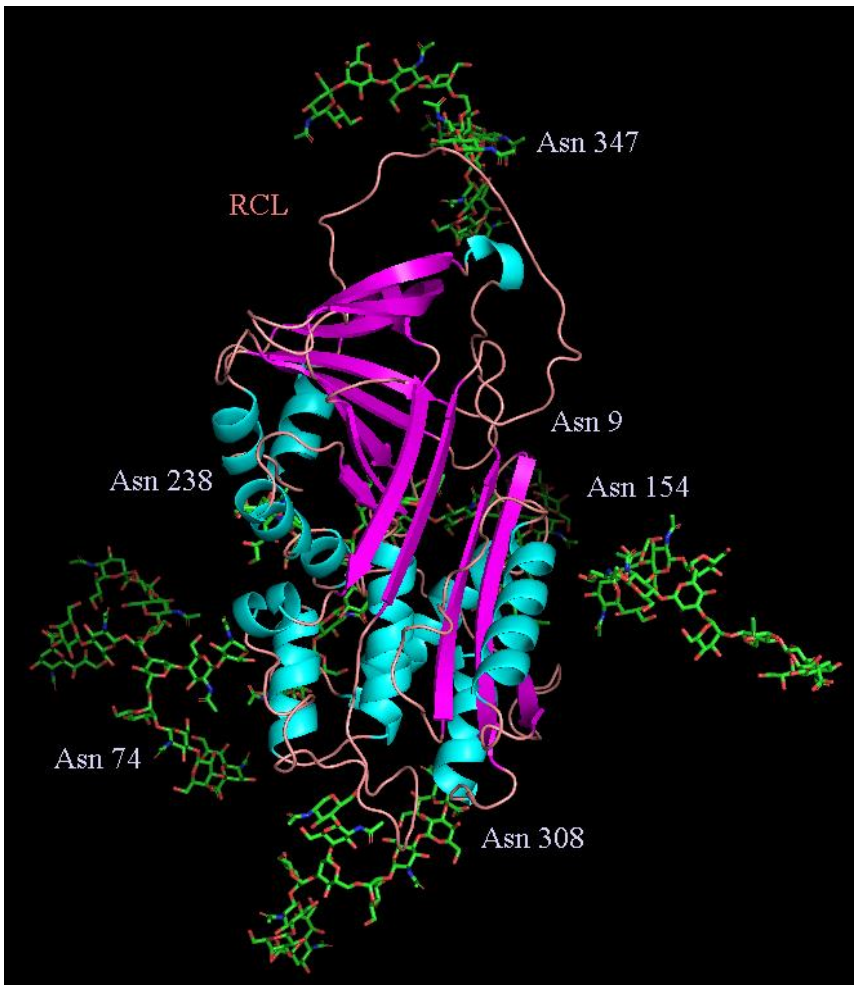


Figure 3. Homology model of uncleaved (stressed) human CBG based on human TBG (PDB: 2CEO). Human CBG and TBG are close structural homologs that display 45% sequence identity. The β-sheets are shown in magenta, α-helices are in light blue. The exposed reactive centre loop (RCL) is indicated. The locations of the six *N*-glycosylation sites have been labelled, and the most common *N*-glycans (shown in green/red sticks) have been attached to the six sites using Glycam (58). The visualisation was performed using PyMOL.

Site-directed mutagenesis studied have revealed important information of the ligand binding characteristics of CBG and showed that CBG may also carry non-physiological ligands. Interestingly, the Trp371Arg mutation resulted in a near-complete loss of cortisol binding, but, in parallel, acquired a low affinity for thyroxine, the hormone that is normally transported by TBG (K_d, 25°C = 840 nM vs. 0.73 nM for TBG) (59).

Advancing our knowledge of the intriguing and often delicate hormone:CBG interactions may provide so far untapped opportunities for the engineering of CBG with desired functions in the future, for example, designing CBG variants that display a higher hormone-binding affinity or a longer circulation half-life. Moreover, the prospects of altering the RCL by creating new recognition sites

for tissue-specific proteases may allow the release of ligands (drugs) at defined locations at the appropriate time in the body, making CBG an interesting candidate for *in vivo* drug delivery (59).

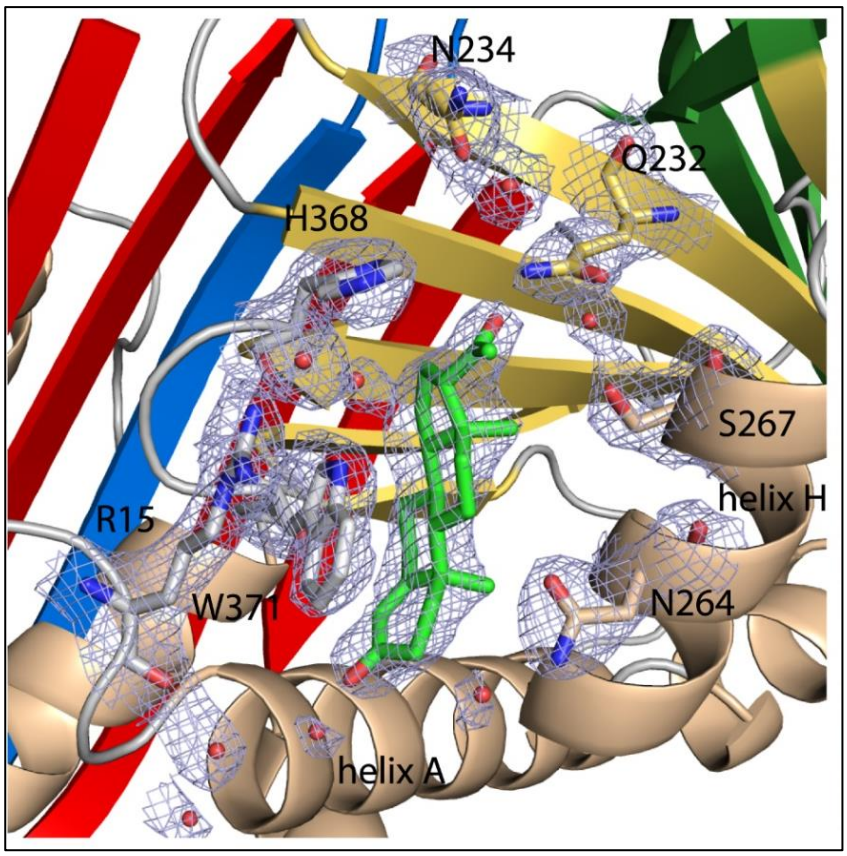


Figure 4. The crystal structure of the hydrophobic steroid-binding pocket of cleaved (relaxed) human CBG in complex with progesterone (PDB: 4BB2, 2.48 Å). The β-sheet B, α-helices A and H, as well as key amino acid residues for hormone binding, are indicated. From (57).

1.3.2 Cleavage of RCL and stressed-to-relaxed structural rearrangement of CBG

The reactive centre loop (RCL) (Glu333-Ile354) is a key structural feature of CBG. The RCL can be cleaved by several proteases including by the endogenously-expressed neutrophil elastase (NE) and chymotrypsin and by exogenous virulence factors (60–62). These proteases target specific sites within the RCL. NE released by neutrophils at sites of inflammation via degranulation was shown to rapidly and very specifically cleave the RCL at Val344-Thr345 (63). In contrast, the non-homologous zinc-metalloprotease LasB secreted by the bacterial pathogen *Pseudomonas aeruginosa* (hereafter referred to as *Pseudomonas aeruginosa* elastase, PAE) cleaves the RCL primarily at Asn347-Leu348 and less effectively at Thr345-Leu346, Leu346-Asn347, and Leu348-Thr349 (60). Finally, chymotrypsin cleaves the RCL at Leu346-Asn347 and Leu348–Thr349, although the physiological relevance of chymotrypsin, which is not present at the site of inflammation, is poorly understood (61). Interestingly, and perhaps somewhat counterintuitive given the crucial role of RCL in CBG biology, the RCL of CBG is only weakly conserved when comparing between species (60).

The RCL is fully exposed in the maturely folded CBG structure allowing for easy recognition by the above-mentioned target proteases (64). The cleavage of the RCL leads to a dramatic stressed-to-relaxed (S-to-R) conformational change, a unique characteristic shared by many serpins, including CBG (**Figure 5**). During the S-to-R transition, the cleaved loop is fully inserted in an antiparallel manner into the central β-sheet structure (57). The S-to-R transition leads to a significant increase in the thermo-stability of CBG. Uncleaved (stressed) human CBG is denatured at 55°C, whereas NE-cleaved (relaxed) CBG has an increased thermo-stability of up to 80°C (57). Important for the CBG biology, the S-to-R transition results in a 10-fold decrease in the steroid-binding affinity, a mechanism that drives the timely release of hormones when required at the site of inflammation (65).

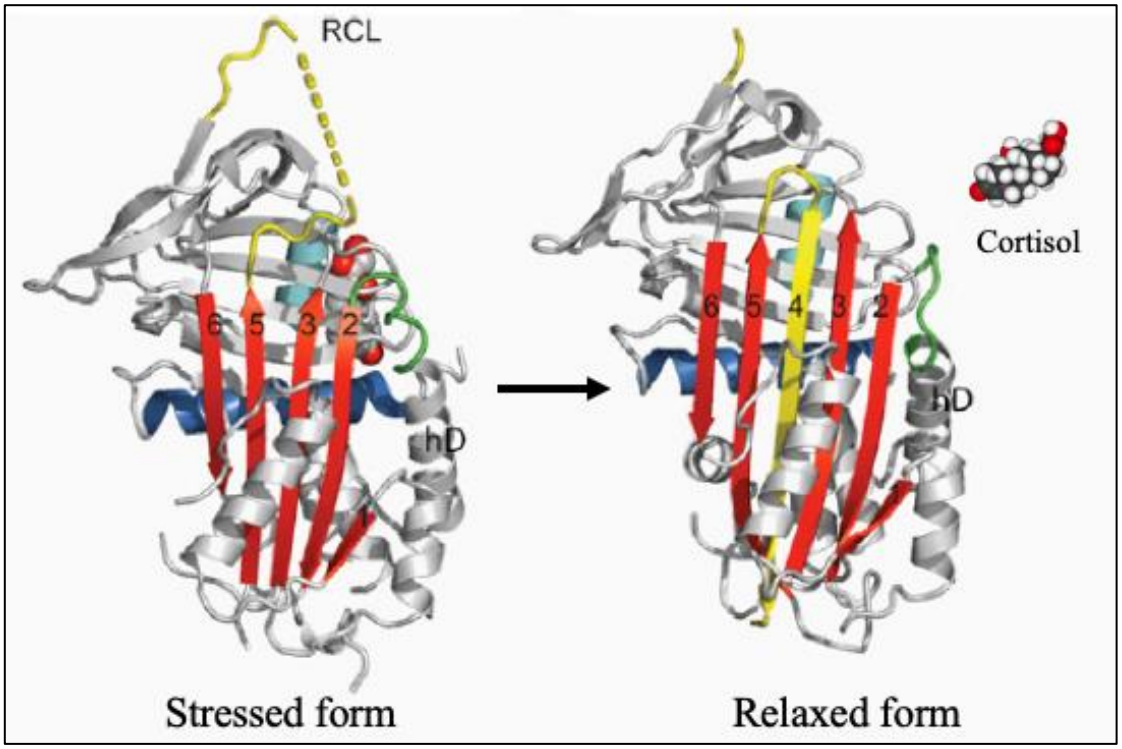


Figure 5. Stressed (S)-to-relaxed (R) rearrangement of the CBG structure upon RCL cleavage. The RCL cleavage mediated by NE (and other proteases) results in the insertion of the RCL (yellow) into a central β sheet (red) that drives an S-to-R transition, and, in turn, cortisol release. From (66).

1.4 CBG glycosylation

1.4.1 Protein glycosylation

Glycosylation is a common and the most complex type of post-translational modification of human proteins (67). Glycosylation is known to modulate the biological activity of proteins by altering the thermal stability, solubility, serum half-life, clearance rate, protein-protein interactions and immunogenicity amongst many other physicochemical properties of proteins, and, in turn, modulate and affect highly diverse molecular and cellular processes relevant to human biology (68).

The process of protein *N*-glycosylation is initiated in the rough endoplasmic reticulum by the action of an integral membrane protein, the oligosaccharyltransferase (OST). OST catalyses the transfer of a preassembled and immature oligosaccharide moieties to Asn residues of acceptor polypeptides located in well-defined consensus motifs, i.e. Asn-Xxx-Ser/Thr/Cys, where Xxx ≠ Pro (69). The glycoproteins subsequently undergo a glycan-processing pathway in the Golgi apparatus resulting in a great structural diversity of fully processed *N*-linked glycoproteins, a heterogeneity known as "glycoforms" (70). *N*-glycans comprise a common trimannosylchitobiose core ($Man_3GlcNAc_2$) and are often grouped into three main classes, i.e. oligomannosidic-, complex- and hybrid-type **(Figure 6)** (1). Recently, two non-conventional *N*-glycan classes, the paucimannosidic- and chitobiose core-type *N*-glycans, were found to add to the human *N*-glycan repertoire (71–73).

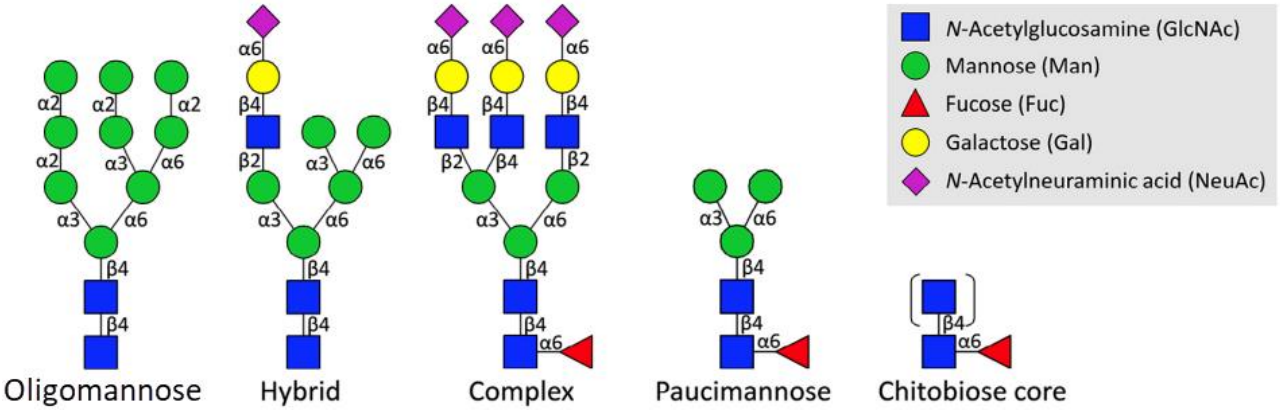


Figure 6. Human N-glycan types. N-glycans spanning the oligomannosidic-, complex-, and hybrid-type are linked to polypeptides via Asn-Xxx-Ser/Thr (Xxx ≠ Pro) and share a common trimannosylchitobiose core (Man3GlcNAc2). Two less studied types also exist, i.e. the paucimannosidic- and chitobiose core-type N-glycans. See key for symbols. Modified from (71).

Mucin-type *O*-glycans, i.e. GalNAc-core *O*-glycosylation, is another common type of protein modification, and a focus of this thesis. Mucin-type *O*-glycans are covalently-linked to the -OH group of Ser/Thr via α-GalNAc moieties and comprise eight main core types (core 1-8) (**Figure 7**). Protein *O*-glycosylation, which is initiated in the Golgi apparatus, involves the transfer of the first α-GalNAc residue from the nucleotide-sugar donor (UDP-GalNAc) to Ser/Thr of fully-folded proteins. *O*-glycosylation does not have predictable consensus motifs on their acceptor polypeptides. Remarkably, the first α-GalNAc transfer reaction is catalysed by 20 different polypeptide-*N*-acetyl-α-D-galactosaminyltransferases (GalNAc-Ts) displaying overlapping substrate specificities and tissue-specific expression (74).

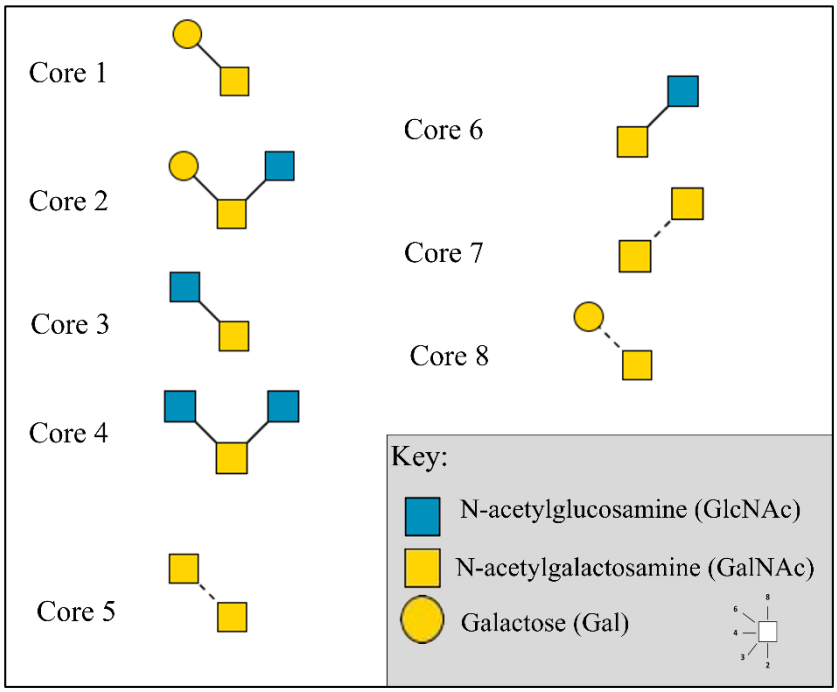


Figure 7. The eight different types of mucin-type *O*-glycans (core 1-8). Adapted from (75).

1.4.2 Site-specific *N*-glycosylation of serum CBG derived from donors (native-CBG)

Human CBG has six *N*-linked glycosylation sites, i.e. Asn9, Asn74, Asn154, Asn238, Asn308, and Asn347. Structural characterisation of the *N*-glycans decorating each of the six *N*-glycosylation sites of CBG isolated from blood obtained from a pool of donors has previously been performed using LC-MS/MS-based glycan and glycopeptide profiling (76). The observed *N*-glycans were complex-type structures predominantly showing bi- and tri-antennary branching patterns (**Figure 8)**. All CBG *N*-glycans were found to be capped with sialic acid residues; the glycans carried up to six terminal α2,3- and α2,6-linked neuraminic acid (NeuAc) residues. CBG showed strong site-specific *N*-glycosylation, which was related to the different solvent accessibilities of the six sites of maturely folded CBG. All sites were highly occupied (70.5-99.5%), and the conjugated glycans showed a low degree of core fucosylation (0-34.9%). The *N*-glycans of Asn74, Asn154 and Asn347 displayed a particular high level of heterogeneity.

Despite exciting advances in glycoanalytics (77), the deep structural characterisation of complex glycoproteins, including CBG, still requires a combination of analytical techniques and approaches to achieve both site-specific and glycan fine structure information. This is most commonly performed using site-unspecific glycomics in concert with site-specific glycopeptide profiling (76). Glycomics analysis using porous graphitised carbon (PGC)-LC tandem mass spectrometry (MS/MS), one of the analytical platform used in this thesis, makes use of the ability of the PGC-LC stationary phase to separate glycan isomers of reduced, but otherwise native *N*-glycans after their liberation from proteins

(78). The molecular mass and the glycan fragmentation pattern upon resonance-activation (ion trap) collisional-induced dissociation (CID) MS/MS in negative polarity mode (discussed further below) is a powerful identification method of glycans in complex mixtures (79). This approach, which may integrate linkage-specific exoglycosidases for an additional level of structural insight or certainty, allows the identification of monosaccharide compositions as well as provides information of the *N*-glycan fine structure albeit some structural ambiguity often remains when dealing with very complex glycans. The analysis of intact glycopeptides as it is often performed in glycoproteomics using high-resolution mass spectrometry allows the site-specific identification of monosaccharide compositions, the relative abundance of glycoforms and site-specific *N*-glycan occupancy (76). Multiple fragmentation techniques including higher-energy induced dissociation (HCD), electron-transfer/higher-energy collision dissociation (EThcD) and ion trap CID-MS/MS are often required (or at least highly beneficial) to extract complementary information from intact glycopeptides (77, 80). Glycan and glycopeptide fragmentation are discussed in Section 1.5.

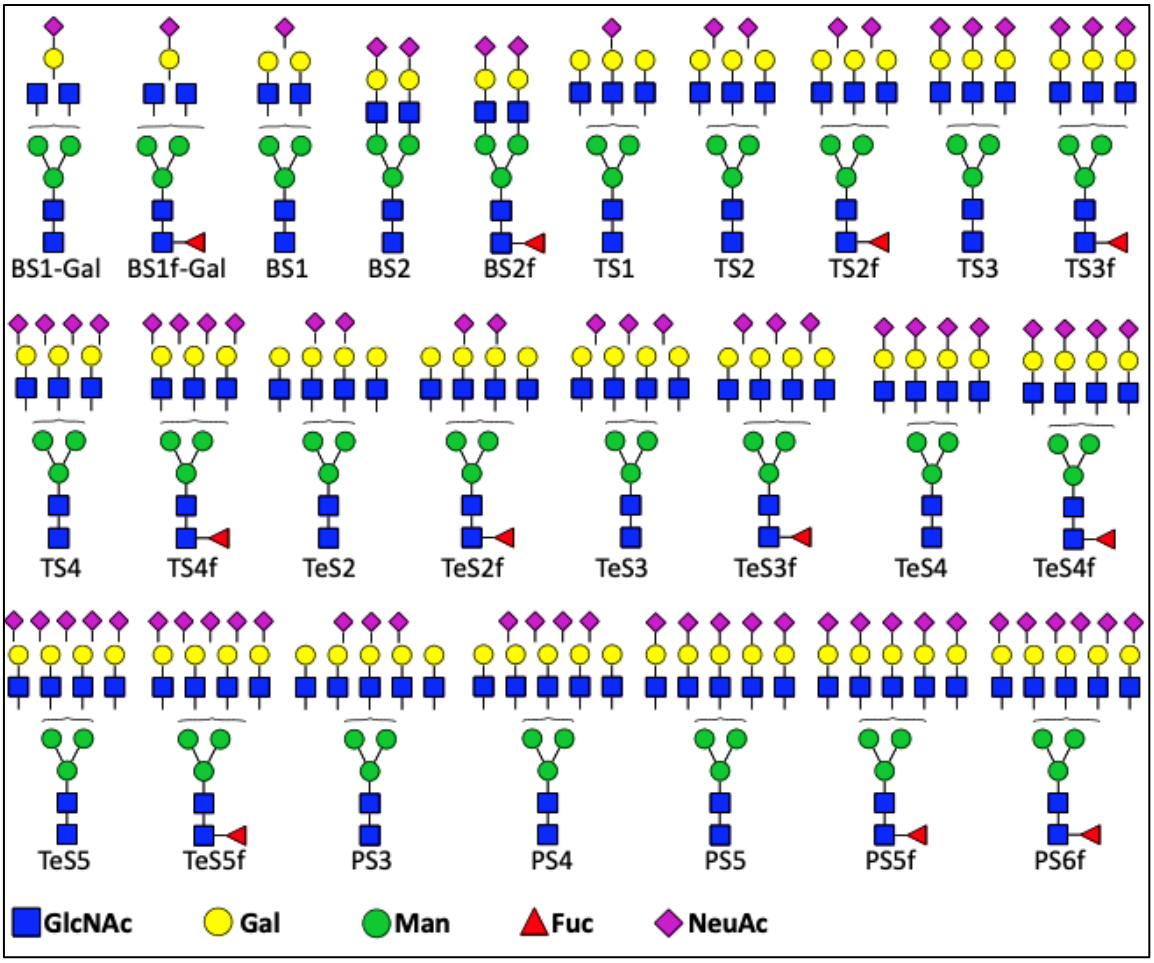


Figure 8. Map of the *N*-glycans identified from the six *N*-glycosylation sites of CBG from pooled plasma of donors. B, bi-antennary; T, tri-antennary; Te, tetra-antennary; P, penta-/tetra-antennary + one LacNAc repeat; S, NeuAc; f, fucose; Gal, galactose. Likely linkages and topologies are depicted based on experimental data and knowledge of the human *N*-glycosylation pathway. Brackets mean that the antennae position on the α1,3- or α1,6-mannose arm is unknown. Adapted from (76).

1.4.3 Functional roles of CBG *N*-glycosylation

As elaborated on below, the *N*-glycans known to abundantly decorate CBG are recognised to influence its physiochemical properties, and, as a consequence, play crucial roles in CBG biology.

Early observations showed that *N*-glycosylation affects the serum half-life of CBG in circulation (81). In that study, the presence of sialic acid residues in the non-reducing end of the CBG *N*-glycans (see above) was found to increase the circulation time of the protein supposedly by precluding removal via the hepatic asialo-receptor. This was supported by the observation that the removal of sialic acid via neuraminidase treatment resulted in enhanced protein clearance and a reduced circulation half-life. Another study showed that preventing *N*-glycosylation altogether via chemical inhibition of the *N*-glycosylation machinery did not prevent the biosynthesis or secretion of CBG, suggesting that *N*-glycans are not directly involved in those processes (82). Opposing this view, the secretion of unglycosylated recombinant CBG generated by genetic disruption of the glycosylation sites was found to be reduced as compared to the wild-type recombinant CBG carrying intact sequons indicating critical roles of the *N*-glycans in this secretion, folding and trafficking of the glycoprotein.

Less controversial, CBG *N*-glycosylation is known to play crucial roles in defining the steroid-binding capacity and ligand affinity. Specifically, the glycosylation of Asn238 was found to be essential for steroid binding since the disruption of the glycosylation of this site by the substitution of Asn238Gln or Thr240Ala resulted in a loss of cortisol-binding capacity (83). Another study revealed the importance of *N*-glycosylation of CBG in cortisol binding affinity (84). The authors showed that unglycosylated recombinant CBG has a nine-fold lower cortisol binding affinity as compared to glycosylated CBG at 37°C, and that cortisol dissociates at a 30-fold lower rate from glycosylated CBG relative to unglycosylated stressed CBG. Finally, relative to the unglycosylated counterpart, glycosylated CBG was found to be more sensitive to the elevated body temperature (37°C-42°C) that is associated with local or body-wide inflammation in the context of cortisol dissociation from CBG without an associated RCL cleavage (thermo-couple effect, see above).

Interestingly, a recent study showed that several structural features of the glycans conjugated to Asn347 located proximal to the RCL cleavage site affect the neutrophil elastase (NE) cleavage rate and thereby impacts the cortisol release process (introduced above) (85). In that study, it was shown that a high degree of branching, core fucosylation and site occupancy reduce the NE-mediated cleavage efficiency and the speed of cortisol release, which could be associated with a steric hindrance of the cleavage site and reduced accessibility of NE. Sialylation of the Asn347 glycans was shown to slightly increase the rate of NE-mediated cleavage, but an even greater benefit of sialylation on

Asn347 and other CBG sites on the RCL cleavage rate was observed for cleavage facilitated by the exogeneous PAE, an abundant virulence factor secreted by *P. aeruginosa* (see above). The benefit of sialic acid residues for RCL cleavage is likely explained by the electrostatic interactions formed between the negatively charged sialic acids and the positively charged surface of NE and PAE (85).

Despite providing new mechanistic insight into the cortisol delivery process, the above studies have left many unaddressed questions regarding the fascinating relationship between CBG *N*-glycosylation and cortisol release. For example, the details of the *N*-glycosylation have to date only been characterised using modern analytical methods for serum CBG obtained from a pool of donors. Thus, little is still known about the individual-specific structure and function of CBG *N*-glycosylation and how the decorating *N*-glycans may vary and play roles in different physiological conditions.

1.4.4 Altered CBG *N*-glycosylation in pregnancy

Several studies using classical biochemical techniques have indicated that CBG *N*-glycosylation is altered in pregnancy. The progressive increase in the concentration of maternal plasma CBG until week 36 (86), was reported to be accompanied with pregnancy-associated plasma CBG variants that display altered monosaccharide compositions, i.e. higher levels of *N*-acetylneuraminic acid (NeuAc), galactose (Gal) and *N*-acetylglucosamine (GlcNAc) residues relative to plasma CBG from healthy controls (87). These observations are in good agreement with another more recent finding indicating a shift in the CBG isoelectric point towards more acidic pI and heavier glycoforms (i.e. with more sialylation and a higher degree of branching) throughout the trimesters of the pregnancy (88).

The aberrant *N*-glycosylation of CBG associated with pregnancy may have biological importance for the delivery of cortisol to target tissues and the interaction of CBG with putative CBG receptors. Specifically, it was shown that a pregnancy-associated variant of CBG had different binding affinities for two binding sites on the membrane fraction derived from the placental syncytiotrophoblast obtained at 7-12 weeks of gestation as compared to a non-pregnant control cohort (89).

1.4.5 CBG *O*-glycosylation

Comprehensive *O*-glycoproteome characterisation of human plasma glycoproteins remains extremely challenging, a fact illustrated by the rather limited volume of literature available on the serum *O*-glycoproteome (90–92). Of great interest to the focus of this thesis, two large-scale *O*-glycopeptide studies however recently indicated that plasma CBG may carry *O*-glycosylation (93, 94), a feature not previously acknowledged by the large body of CBG literature generated over the

last 70 years. These putative modifications, which will be explored and discussed in greater details in this thesis, increase further the molecular and functional complexity and diversity of human CBG.

1.5 The art of glycan and glycopeptide destruction

1.5.1 Glycan and (glyco)peptide fragment nomenclature

The characterisation of glycoproteins can conveniently be divided into the analysis of glycans enzymatically or chemically released from their carrier protein (glycomics) and the analysis of intact glycopeptides (glycoproteomics, when performed at the system-wide level). Mass spectrometry (MS) remains a preferred analytical method for both glycan and glycopeptide profiling.

Tandem MS (MS/MS) is used to obtain sequence information of glycans by inducing cleavages on either side of the oxygen atom of the glycosidic bonds existing between neighbouring monosaccharide residues (generating B-/Y- and C-/Z- ions, respectively) (**Figure 9A**). Less abundant but equally informative cross-ring cleavages (A-/Z- ions) may also be generated (95, 96).

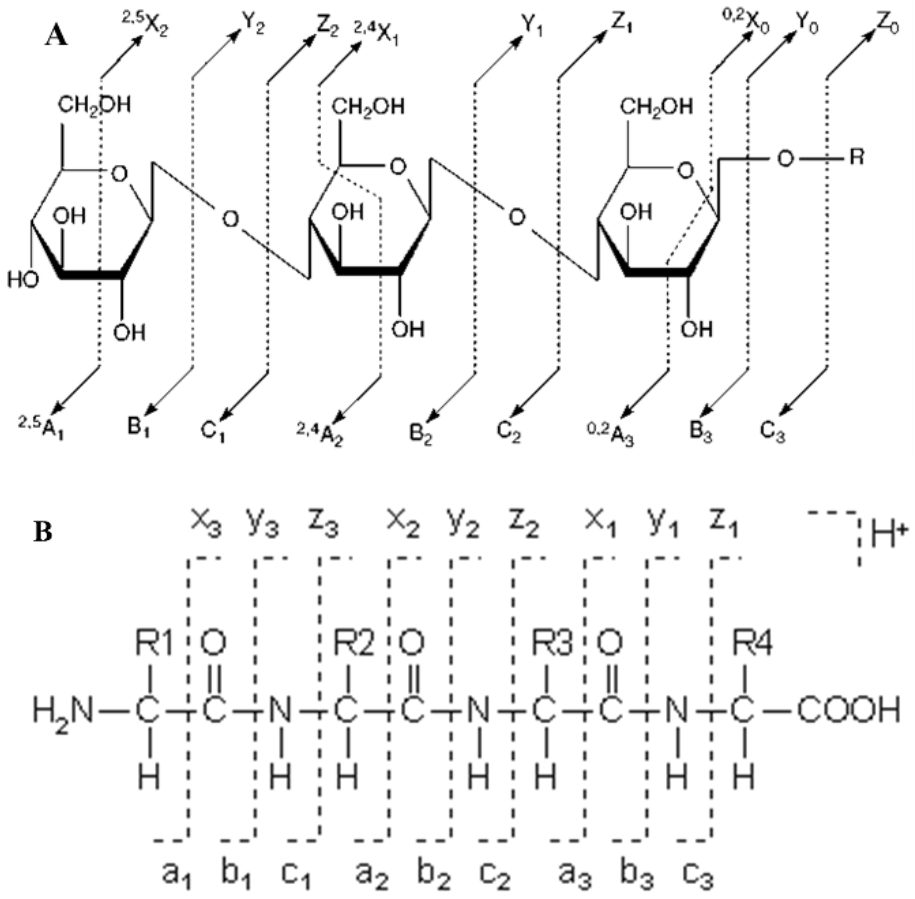


Figure 9. Overview of the established nomenclature of possible fragments generated from the gas-phase dissociation of (**A**) glycans (95) and (**B**) (glyco)peptides (97) in MS/MS. Fragments denoted with capital letters relate to glycan fragments (B-/Y-/C-/Z- ions, glycosidic fragments and A-/X-ions, cross-ring fragments) while fragments denoted with small letters relate to peptide fragments (a-/b-/c-/y-/c-/z-ions) arising from the dissociation of the polypeptide backbone.

By convention, the gas-phase dissociation of (glyco)peptides may generate a variety of peptide backbone fragments comprising a-/b-/c- and x-/y-/z-ions depending on which bonds within the

polypeptide backbone are broken and depending on which of the generated fragments the charge (often protons, H^+) is retained (**Figure 9B**) (97). Common for the glycan and glycopeptide fragmentation, the fragmentation pathways, which often remain incompletely understood, depend on the polarity, the amount of energy introduced into the analyte and on the derivatisation type or any adduct formation of the analyte of interest (98–100).

1.5.2 Brief overview of commonly used glycan and glycopeptide dissociation methods

Low-energy (ion trap) or high-energy (e.g. QTOF) collision-induced dissociation (CID) is a powerful method to obtain sequence information of glycans by inducing cleavages of the glycosidic bonds between monosaccharide residues of glycan thereby generating B-/Y- and C-/Z- ions, respectively (see Figure 9A). In short, the generated fragments that often cover both the non-reducing and reducing end of the glycan provide information about monosaccharide composition, sequence and branch point(s) of the studied glycan, but not the anomeric linkage type (α or β) of the glycosidic bonds.

To facilitate the identification of both the peptide sequence, glycan composition and site-specific localisation of an attached sugar moiety intact glycopeptide analysis is a preferred approach. Orthogonal dissociation methods are typically used to fragment glycopeptides (**Figure 10**) (77, 80, 101). Depending on the dissociation method applied, oxonium ions, often representing the non-reducing end glycan fragments (i.e. B1, B2, B3 ions) or internal fragments arising from two or more cleavages of the glycan sequence are often abundant fragments observed.

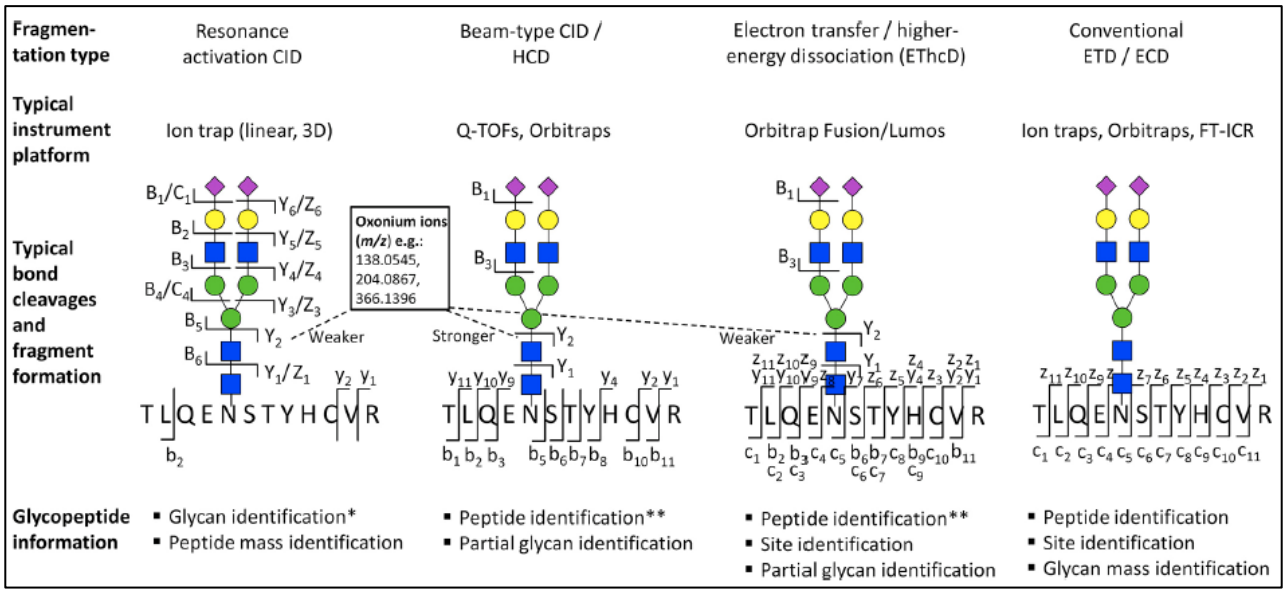


Figure 10. Common dissociation methods used in MS/MS-based analysis of intact glycopeptides. Typical bond cleavages and fragments arising from each dissociation method are depicted. From (102).

Further, the Y-ion series, in particular, the Y_1-ion denoting the intact peptide still conjugated with a HexNAc, are often prominent fragment ions that may be used to determine the glycan composition and inform on the peptide carrier.

Important to mention in this context, the different fragmentation methods generate very different information of the glycopeptide being fragmented, and are, thus, often used in concert in attempts to reach a more complete glycopeptide identification (103). For example, resonance-activation (low-energy as generated in ion trap instruments) CID predominantly generates B- and Y-ions from cleavages of the glycosidic bonds of the glycopeptide and only few b- and y-ions from the peptide part. In contrast, higher-energy collision dissociation (HCD) (high-energy as generated in the Orbitrap C-trap or in QTOF instruments) yield abundant oxonium ion formation and peptide backbone fragmentation. However, CID and HCD often do not allow confident glycosylation site assignment, which is still most ideally performed using ion-ion type fragmentation of glycopeptides, e.g. electron-transfer dissociation (ETD) that produces peptide fragments (c-/z-ions) which crucially still retain the conjugated glycan. Variants of the ETD method, including the hybrid-type electron-transfer/higher-energy collision dissociation (EThcD) are gaining popularity due to a higher fragmentation efficiency (77, 80, 101, 104, 105).

1.6 Aims of Research

From the above, it is beyond doubt that CBG is a key player in cortisol transport and delivery to hormone-sensitive tissues under normal organism homeostasis and in different pathophysiological conditions. *N*-glycosylation is known to directly impact CBG biology by influencing the secretion efficiency, serum half-life, steroid-binding capacity and affinity of the protein, and, importantly, has also been shown to impact the cortisol delivery mechanism. Nevertheless, biochemical knowledge is still not available to detail exactly how the heterogeneous *N*-glycosylation impacts the function of CBG. Moreover, solid evidence of the structure-function of the putative *O*-glycosylation of CBG is missing.

Adding to this, CBG is increasingly being recognised as a potential therapeutic agent that may be harnessed to deliver anti-inflammatory cortisol to tissues/organs undergoing severe acute or chronic inflammation or be used as a carrier molecule for the targeted delivery of drugs upon tailored engineering of the ligand-binding pocket or manipulation of the reactive centre loop (RCL) crucial for delivery (59). Thus, improving our understanding of the CBG structure-function, including the intriguing glycobiology underpinning the function of CBG is required to facilitate and further stimulate such important developments.

To this end, the overarching goal of this Masters of Research thesis was to structurally and functionally characterise the glycosylation of two related variants of recombinant human CBG (rhCBG) engineered to carry attractive features including a histidine tag (hereafter His-CBG) and a biotin group (hereafter biotin-CBG) beneficial for the *in vivo* visualisation and monitoring upon the future use of rhCBG as an anti-inflammatory therapeutic agent. Three specific aims were designed in order to achieve this overarching goal:

1. To perform detailed *N*- and *O*-glycan profiling of rhCBG
2. To map the site-specific glycosylation of rhCBG with a special focus on the RCL glycosylation
3. To investigate the role(s) of the RCL glycosylation of rhCBG

Achieving these aims is a crucial step to advance our understanding of the structure-function relationship of human CBG, knowledge that promises to be of immediate value for fundamental glycobiology and glycoimmunology as well as for applied areas of research within the biomedical and clinical sciences.

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Chemical and reagents

All chemicals used for the described experiments were sourced from Sigma Aldrich/Thermo Fisher Scientific (Sydney, Australia) unless otherwise specified. Ultra-high-quality water was from a Millipore Academic Milli-Q system (Melbourne, Australia). Peptide-*N*-glycosidase F (PNGase F) from *Elizabethkingia miricola,* and modified porcine trypsin were from Promega (Sydney, Australia).

Human CBG (UniProtKB, P08185) was isolated from pooled sera of donors (hereafter referred to as "native-CBG") to >90% purity as assessed by SDS-PAGE (Affiland, Belgium).

Two polyhistidine-tagged recombinant human CBG variants expressed in HEK293 cells with (hereafter "biotin-CBG", 409 amino acid residues) and without (hereafter "His-CBG", 394 amino acid residues) an additional AVI-tag and biotin group at the protein C-terminus were purchased from Sino Biological Inc. (Cat no. 10998-H08H and 10998-H27H-B, respectively, Beijing, China). Human neutrophil elastase (NE) (UniProtKB, P08246) from blood-derived resting neutrophils isolated to >96% purity was obtained from Lee Biosolutions (product code ELA-2 342-40, Maryland Heights, MO). *P. aeruginosa* elastase (PAE) (UniProtKB, P14756) was from the culture medium of a pathogenic laboratory wound strain (PAO1) of *P. aeruginosa* (see section 2.2.3).

2.2 Methods

2.2.1 *N*- and *O*-glycan profiling

2.2.1.1 *N*- and *O*-glycan release and handling

The cysteine residues of native-, His- and biotin-CBG were reduced using 10 mM aqueous dithiothreitol (DTT), 45 min, 56°C and carbamidomethylated using 40 mM aqueous iodoacetamide, 30 min, in the dark, 20°C (all stated as final concentrations). Native-, His- and biotin-CBG were immobilised in discrete spots (5 µg/spot) on a primed 0.45 µm polyvinylidene fluoride (PVDF) membrane (Merck-Millipore) and processed essentially as described previously (78). In brief, the dried protein spots were stained with Direct Blue, excised, transferred to separate wells in a flat bottom polypropylene 96-well plate (Corning Life Sciences), blocked with 1% (w/v) polyvinylpyrrolidone in 50% (v/v) aqueous methanol and washed with water. The *N*-glycans were exhaustively released from the immobilised proteins using 10 U PNGase F in 10 µL water/well, 16 h, 37°C. The released *N*-glycans were transferred into fresh tubes and deaminated (to prepare the glycans for reduction) by the addition of 10 µL 100 mM aqueous ammonium acetate, pH 5, 1 h, 20°C. The *N*-glycans were then dried and subsequently reduced to alditols using 20 µL 1 M sodium borohydride in 50 mM aqueous potassium hydroxide, 3 h, 50°C. The de-*N*-glycosylated proteins on the PVDF membrane were then treated with 0.5 M $NaBH_4$ in 50 mM potassium hydroxide 6 h, 50°C to release *O*-linked glycans by reductive β-elimination.

The glycan reduction reactions were quenched by the addition of 2 µL glacial acetic acid to each sample. Dual desalting of the reduced *N*- and *O*-glycans was performed using first strong cation exchange (SCX) micro-columns using AG 50W X8 resin (Bio-Rad, Hercules, CA) and then porous graphitised carbon (PGC) resin packed on top of C18 discs in P10 solid-phase extraction (SPE) formats. The *N*- and *O*-glycans were found in the flowthrough in the SCX micro-columns and eluted from the PGC-SPE micro-columns using 0.05% trifluoroacetic acid: 40% acetonitrile (ACN): 59.95% water (v/v/v), dried, and finally redissolved in 10 µL water for *N*- and *O*-glycan profiling.

2.2.1.2 PGC-LC-MS/MS-based *N*- and *O*-glycan profiling

The liberated *N*-glycans were profiled using LC-MS/MS performed on an LTQ Velos Pro ion trap mass spectrometer (Thermo Scientific, San Jose, USA) coupled to a Dionex Ultimate-3000 HPLC (Dionex, Sunnyvale, USA). The *N*-glycans were separated on a PGC-LC capillary column (Hypercarb KAPPA, 5 µm particle size, 250 Å pore size, 0.18 x 100 mm, Thermo Scientific) using a linearly increasing concentration of solvent B (2.6-64% (v/v) 10 mM NH_4HCO_3 in ACN over 86 min) in solvent A (10 mM aqueous NH_4HCO_3) at a constant flow rate of 3 µL/min. The setup employed a

post-column make-up flow configuration by supplying pure isopropyl alcohol at a constant 4 µL/min flow rate throughout the LC-MS/MS run for enhanced sensitivity (106). The MS1 acquisition scan range was set to *m/z* 570-2,000 with a zoom scan resolution of *m/z* 0.25 at full width at half maximum (FWHM) and a source voltage of +2.7 kV. Detection was performed in negative ion polarity mode with data-dependent MS/MS acquisition. The automatic gain control (AGC) for the MS1 scans was set to 5 x 10^4 with a maximum accumulation time of 50 ms. For the MS/MS events, the resolution was set to *m/z* 0.35 FWHM, the AGC was 2 x 10^4, and the maximum accumulation time was 300 ms. The five most abundant precursors in each MS1 scan were selected for MS/MS using resonance-activation (ion trap)-based collision-induced dissociation (CID) at a fixed 35% normalised collision energy (NCE). Dynamic exclusion was disabled.

2.2.1.3 Data analysis of the *N*- and *O*-glycans

The LC-MS/MS raw data was browsed and interrogated using Xcalibur v2.2 (Thermo Scientific). The fine structure of each glycan structure was characterised using the relative and absolute PGC-LC retention time, the monoisotopic precursor mass and the CID-MS/MS fragmentation pattern as previously described (79, 107). GlycoMod (108) and GlycoWorkBench (109) were used to aid the structural characterisation. No glycan quantitation was performed since the glycans were later site-specifically profiled (and quantified) using glycopeptide analysis (see below).

2.2.2 *N*- and *O*-glycopeptide analysis

2.2.2.1 Protein reduction, alkylation and trypsin digestion

The cysteine residues of native-, His- and biotin-CBG (5 µg/sample) were reduced using 2.5 mM aqueous DTT, 45 min, 56°C and carbamidomethylated using 10 mM aqueous iodoacetamide, 30 min, in the dark, 20°C (all stated as final concentrations). In solution digestion was performed using 1:20 (w/w, enzyme-to-protein ratio) overnight in 100 mM aqueous ammonium bicarbonate, pH 6.8, 37°C. The resulting CBG peptide mixtures were desalted using C18 stage-tips (Thermo Scientific) according to the manufacturer's instructions, dried and redissolved in 0.1% (v/v) aqueous formic acid prior to LC-MS/MS analysis.

2.2.2.2 Reversed phased LC-MS/MS-based (glyco)peptide profiling

For LC-MS/MS profiling of intact glycopeptides, the CBG (glyco)peptides were loaded on a trap column (20 mm length x 0.1 mm inner diameter) custom packed with ReproSil-Pur C18 AQ 5 µm resin (Dr. Maisch, Ammerbuch-Entringen, Germany). Approximately 1 µg total peptide was injected per LC-MS/MS run. The peptides were separated at a constant flow rate of 250 nL/min on an

analytical column (Reprosil-Pur C18-Aq, 250 mm length x 0.075 mm ID, 3 µm resin size, Dr. Maisch, Ammerbuch-Entringen, Germany) operated in reversed-phase mode using an UltiMate™ 3000 RSLCnano System. The mobile phases were 99.9% ACN in 0.1% (both v/v) aqueous formic acid (FA) (solvent B) and aqueous 0.1% (v/v) FA (solvent A). The linear gradient of solvent A increased from 2% to 30% (v/v) over 60 min, 30-50% (v/v) over 5 min, 50-95% (v/v) over 1 min, and, finally, 7 min at 95% B (v/v). The nanoLC was connected to a Q Exactive HF-X Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) operating in positive ion mode. The MS/MS data was collected using data-dependent acquisition. The full MS1 scan data was acquired in the Orbitrap with an AGC target value of 3×10^6 ions and a maximum fill time of 50 ms. MS1 scans were acquired at a high resolution of 60,000 FWHM (measured at *m/z* 200) with a range of *m/z* 350-2,000. The 20 most abundant precursor ions were selected from each MS full scan utilising higher-energy collision-induced dissociation (HCD) fragmentation with an NCE of 28%. Precursors selected for fragmentations were at least doubly charged ($z \geq +2$). Fragmentation was performed at high resolution (45,000 FWHM) with an AGC target of 1×10^5 product ions and a maximum injection time of 90 ms using a precursor isolation window of *m/z* 1.0 and a dynamic exclusion of 30 s after a single isolation and fragmentation of a given precursor.

In addition, (glyco)peptides from His- and biotin-CBG were also analysed on an Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer (Thermo Fisher Scientific, San Jose, CA) coupled to a Dionex 3500RS UPLC system. Peptides were separated on a 150 mm length x 0.075 mm ID LC-column packed in-house (1.9 µm particle size, C18AQ material, Dr. Maisch, Germany) using a gradient of 5-45% (v/v) solvent B (80% ACN containing 0.1% (both v/v) aqueous formic acid) over 61 min in solvent A consisting of 0.1% (v/v) aqueous FA at a constant flow rate of 350 nL/min at 55°C. The mass spectrometer was operated in positive mode using data-dependent acquisition of MS/MS data. Full MS1 scans were acquired with a scan range of *m/z* 350-2,000 using 120,000 resolution, 5×10^5 AGC and 100 ms maximum injection time. HCD-MS/MS was performed at top speed with an NCE of 30% in the Orbitrap using an AGC of 2×10^5, 120 ms injection time and a quadrupole isolation window of *m/z* 2.0. Following HCD, electron transfer higher collision dissociation (EThcD) at an AGC target of 2×10^5, and with a supplemental activation energy of 15% NCE in the Orbitrap and ion trap CID with an AGC target of 1×10^4 and a NCE of 35% were triggered by the presence of highly specific diagnostic sugar oxonium ions *m/z* 204.0867 (HexNAc), *m/z* 138.0545 (HexNAc) and *m/z* 366.1396 (HexHexNAc) among the 20 most intense fragment ions in each HCD-MS/MS spectrum (30 ppm mass tolerance). All MS and MS/MS data except the low resolution ion trap CID-MS/MS were acquired in the profile mode.

2.2.2.3 Data analysis of the *N*- and *O*-glycopeptides

The intact (glyco)peptide LC-MS/MS raw data was browsed using Xcalibur v2.2 (Thermo). The HCD-MS/MS data from each sample was searched against a defined protein and glycan search space with a precursor ion mass tolerance of 5 ppm and a product ion mass tolerance of 10 ppm using Byonic v2.6.46 (Protein Metrics) (110). Searches were performed with carbamidomethylation of Cys (+57.021 Da) as a fixed modification, and strict trypsin specific cleavage (C-terminal to R/K except when followed by P) with a maximum of two missed tryptic cleavages per peptide. Oxidation of Met (+15.994 Da), *N*-glycosylation of sequon-localised Asn with the predefined *N*-glycan database of 310 mammalian *N*-glycans without sodium and/or *O*-glycosylation of Thr/Ser with the predefined *O*-glycan database containing 78 common mammalian *O*-glycans without sodium available within Byonic were used as variable modifications. A maximum of two common and two rare variable modifications were allowed per peptide. Two protein search spaces were used: 1) the reviewed UniProtKB human protein database (released July, 2017 containing 20,201 entries) and 2) the canonical human CBG sequence (UniProtKB, P08185). All searches were filtered to <1% false discovery rate (FDR) at the protein level and 0% at the peptide level by using a decoy database of the reversed sequences of the list of proteins used for the original (forward) protein search space (only applicable for the searches using the entire human proteome). The Byonic-identified glycopeptides were manually checked with support from GPMAW v10.0 (Lighthouse, Odense, Denmark) (111). The relative abundance of observed glycopeptides was quantified using EIC-based area-under-the-curve measurements of all observed charge states of the monoisotopic precursor with the assistance of the Skyline software (112).

2.2.3 Bacterial cultures and crude isolation and identification of *P. aeruginosa* elastase (PAE)

The *P. aeruginosa* PAO1, a pathogenic lab strain, was plated on Luria-Bertani agar plates and incubated overnight, 37°C as described (113). A cell colony was transferred to 100 mL LB medium, incubated overnight, 37°C, and the culture medium containing the secreted PAE was collected. The culture medium was centrifuged at 5,000 x g and the supernatant containing PAE was mixed with a slurry of strong anion exchange (SAX) resin (Q-Sepharose, GE Healthcare) pre-equilibrated in 0.02 M sodium phosphate, pH 7.7 to crudely separate PAE from other more acidic molecular components. End-over-end mixing was performed for 30 min, room temperature. The unbound fraction containing the PAE was collected after 5 min centrifugation at 3,000 rpm. Small fractions of both the crude (unfiltered) and the semi-isolated (filtered) PAE were applied to SDS-PAGE to confirm the presence of PAE, a highly abundant secreted virulence factor with a molecular mass of ~35 kDa (85).

2.2.4 NE- and PAE-based digestion of CBG

His-CBG was incubated with NE in 50 mM sodium acetate, 0.6 M sodium chloride, pH 5.5 at an enzyme/substrate ratio of 1:25 (w/w) for 15 s, 30 s, 60 s, 120 s, 240 s, 480 s, and 960 s, 20°C. In addition, native-CBG and His-CBG were incubated with purified PAE in 50 mM Tris-HCl, 10 mM $CaCl_2$, pH 7.5 at an estimated enzyme/substrate ratio of ~1:5 (w/w) for 1 h, 4 h, and 16 h, 37°C. The digestion reactions were quenched by instantly boiling the samples in NuPAGE LDS sample buffer (Life Technologies) containing aqueous 10 mM dithiothreitol in glycerol, 1 M Tris-HCl (pH 6.8), 2-mercaptoethanol, and SDS for 5 min. Intact (uncleaved) CBG and the CBG fragments generated from the protease treatment were separated on 4-12% Nu-PAGE gels (Life Technologies) using a 2-(N-morpholino)ethanesulfonic acid (1x) running buffer (Life Technologies) for 40 min at 180 V. Gels were stained overnight using Coomassie brilliant blue, and destained in water.

2.2.5 In-gel trypsin digestion and extraction of CBG (glyco)peptides

Gel bands corresponding to intact (uncleaved) native- and His-CBG and the large N-terminal fragments (called "CBG-Nt") both located in the upper protein band (here termed "Band 1", 55-65 kDa) and the small C-terminal fragments (called "CBG-Ct") located in "Band 2" (~5 kDa) were excised from the gels, cut into small pieces and washed thoroughly with 50% (v/v) aqueous ACN, 100 mM aqueous ammonium bicarbonate and ACN and dried as previously described (114). The cysteine residues of the proteinaceous material were reduced using 10 mM aqueous DTT, 45 min, 56°C and carbamidomethylated using 55 mM aqueous iodoacetamide, 30 min in the dark at room temperature (all stated as final concentrations). The gel bands were then washed as described above.

In-gel overnight digestion was performed using 100 ng trypsin/gel band in 100 mM aqueous ammonium bicarbonate, pH 6.8, 37°C. The CBG (glyco)peptides were extracted from the gel pieces using repeated washing cycles of ACN and 5% (v/v) aqueous FA. The resulting (glyco)peptide mixtures were desalted using C18 stage-tips (Thermo Scientific) according to the manufacturer's instructions, dried and redissolved in 0.1% (v/v) aqueous FA. The (glyco)peptide mixtures from Band 1 and 2 were analysed separately using LC-MS/MS as previously described (see section 2.2.2.2).

2.2.6 Structural representation of CBG

The three dimensional structure of human CBG was visualised using a previously reported homology model of uncleaved (intact, "stressed") CBG created by using two available x-ray structures i.e. the RCL-cleaved ("relaxed") human CBG (PDB code 4BB2, 2.48 Å) and the uncleaved ("stressed") human TBG (PDB code 2CEO, 2.80 Å), a close structural homolog of human CBG (85).

N-glycans were added to each of the six glycosylation sites of human CBG using Glycam (58). The visualisation of this homology model of glycosylated human CBG was performed using PyMOL (115).

Chapter 3: Results

In this thesis, the *N*- and *O*-glycans decorating two biologically-active forms of HEK293-expressed recombinant human corticosteroid-binding globulin (rhCBG), i.e. His- and biotin-CBG were site-specifically profiled using mass spectrometry-based glycomics and glycopeptide analysis employing multiple dissociation methods. Further, digestion assays using neutrophil elastase (NE) and *P. aeruginosa* elastase (PAE) were employed to investigate if the CBG glycans strategically positioned on the reactive centre loop (RCL) of the protein were involved in protecting CBG from (or alternatively exposing CBG to) proteolysis and thus impacting the cortisol delivery mechanism.

3.1 Glycan profiling of rhCBG

PGC-LC-MS/MS-based glycomics was firstly utilised to profile the *N*- and *O*-glycans of His- and biotin-CBG. The glycan profiling data showed an extensive *N*-glycan heterogeneity by the characterisation of 71 monosaccharide compositions corresponding to chitobiose core-, paucimannosidic-, oligomannosidic- and complex-type *N*-glycans. The glycan fine structures were characterised using three key analytical features, i.e. the monoisotopic molecular mass, the glycan fragmentation pattern and the LC elution time, a common strategy in PGC-LC-MS/MS-based glycomics (79). Multiple *N*-glycan isomers, including perfectly isobaric glycan structure isomers often only differing in a single glycosidic linkage (e.g. α2,3- vs α2,6-sialyl) were chromatographically resolved by the PGC-LC and manually validated by the presence of diagnostic ions in the resonance-excitation (ion trap) CID-MS/MS spectral data recorded in negative polarity mode.

The His- and biotin-CBG *N*-glycan profiles were qualitatively very similar (see **Appendix 1**, **Figure S1** for summed PGC-LC-MS *N*-glycome profiling). Only relatively few *N*-glycans were observed uniquely for each of the two recombinant glycoproteins (**Figure 11A**). The complex-type *N*-glycans displayed a high degree of α1,6-linked (core) fucosylation as supported by a diagnostic fragment ion pair (*m/z* $350/368^{1-}$) in the CID-MS/MS (-) data. In addition, fragment ions corresponding to antenna fucosylation i.e. sialyl $Lewis^{x/a}$ (*m/z* 801/819) and $Lewis^{x/a}$ (*m/z* 510/528) were identified. The complex-type *N*-glycans displayed a high degree of branching (bi- to tetra-antennary) and most structures were found to terminate with either α2,3-/α2,6-sialylation (e.g. Structure 57-62, see Figure 11A for glycan identifiers) or β1,4-galactosylation (e.g. Structure 49-51).

Fewer structures were β-GlcNAc-capped (e.g. Structure 12-14). The less common *N,N*-diacetyllactosamine (LacdiNAc) and sialyl-LacdiNAc moieties (Structure 34-45) were also identified as supported by the diagnostic fragment ions *m/z* 465^{1-} ($^{1,3}A_{Man}$-ion, LacdiNAc), *m/z* 696^{1-} (B-ion, sialyl-LacdiNAc), *m/z* 858^{1-}/878^{1-} (B-/C-ion, ManGlcNAcHexNAcNeuAc) (**Figure 11D**) (for nomenclature see Introduction 1.5 and (95). Further, several *N*-glycans containing bisecting β1,4-GlcNAc were identified as indicated by the presence of the characteristic D-221 diagnostic ions (*m/z* 670.20^{1-}, 816.24^{1-}, 961.28^{1-} and 711.20^{1-}) and their early PGC-LC elution time (Structure 23-29 and 63-71) as reported on previously (79).

Several oligomannosidic-type *N*-glycans were observed for both His- and biotin-CBG (Structure 6-11). Interestingly, some oligomannosidic-type *N*-glycans ($Man_{5-6}GlcNAc_2Phos_1$, Structure 10-11) were modified with mannose-6-phosphate (M6P) as evidenced by two characteristic fragment ions (*m/z* 403^{1-}, BY-ion, Man_2Phos_1 and *m/z* 421^{1-}, CY-ion, Man_2Phos_1) (**Figure 11C**). Only few structures were of the chitobiose core- (Structure 1-2) and paucimannosidic-type (Structure 3-5). The latter were identified only for biotin-CBG. The observed *N*-glycans showed as expected biosynthetic relatedness (discussed below), which served to further validate the identified structures.

Unexpectedly, the subsequent *O*-glycan profiling indicated the presence of *O*-glycans carried by both His- and biotin-CBG, a previously unknown feature of human CBG (76, 85). The *O*-glycan profile revealed three monosaccharide compositions representing mucin-type core 1- and core 2-type *O*-glycan structures (**Figure 11B; Appendix 1**, **Figure S2)**.

3.2 Site occupancy and site-specific glycoprofiling of rhCBG

Next, the *N*- and *O*-glycans attached to rhCBG were site-specifically profiled using high-resolution LC-MS/MS-based intact glycopeptide analysis. The intact glycopeptides were identified using Byonic database search along with manual annotation of the resulted LC-MS/MS spectra using accurate monoisotopic molecular mass, HCD-based fragment pattern, and reversed-phase LC elution time (see **Appendix 2, Table S1** for tabulated Byonic search data). The relative abundances of the observed glycoforms and the site occupancies were established using EIC-based quantitation as is commonly performed in glycoproteomics (77). The relative abundance of each tryptic glycopeptide was calculated as the percentage of the summed glycopeptide peak intensities of observed charge states in the relation to total glycopeptide intensity for each glycosylation site (116). The site occupancies were calculated as the percentage of glycopeptide peak intensities of observed charge states in the relation to the sum of glycosylated and non-glycosylated peptide peak intensities with

the assumption that different glycopeptides as well as non-glycosylated peptides have similar ionisation efficiency which is an approximation (117).

The site occupancies of the glycosylation sites of CBG were similar for His- and biotin-CBG. Five of the six *N*-glycosylation sites i.e. Asn9, Asn74, Asn154, Asn238, and Asn308 were found to be highly occupied (i.e. 24.4%-100.0%) of both His- and biotin-CBG (**Table 1**). The occupancy of the Asn154 site carried by the tryptic dipeptide (Asn154-Lys155) could only be estimated indirectly using a missed cleavage peptide (Gln147-Lys155).

The interesting Asn347 glycosylation site, strategically located in the RCL region of CBG (peptide Ala327-Arg356), was surprisingly very lowly occupied by *N*-glycans (0.4% and 1.9% for His- and biotin-CBG, respectively). Excitingly, the Ala327-Arg356 peptide was instead found to be decorated with a significant amount of *O*-glycosylation (5.9% and 5.0% for His- and biotin-CBG, respectively). Of note, no co-occurrence of *N*- and *O*-glycosylation was found on any Ala327-Arg356 peptides.

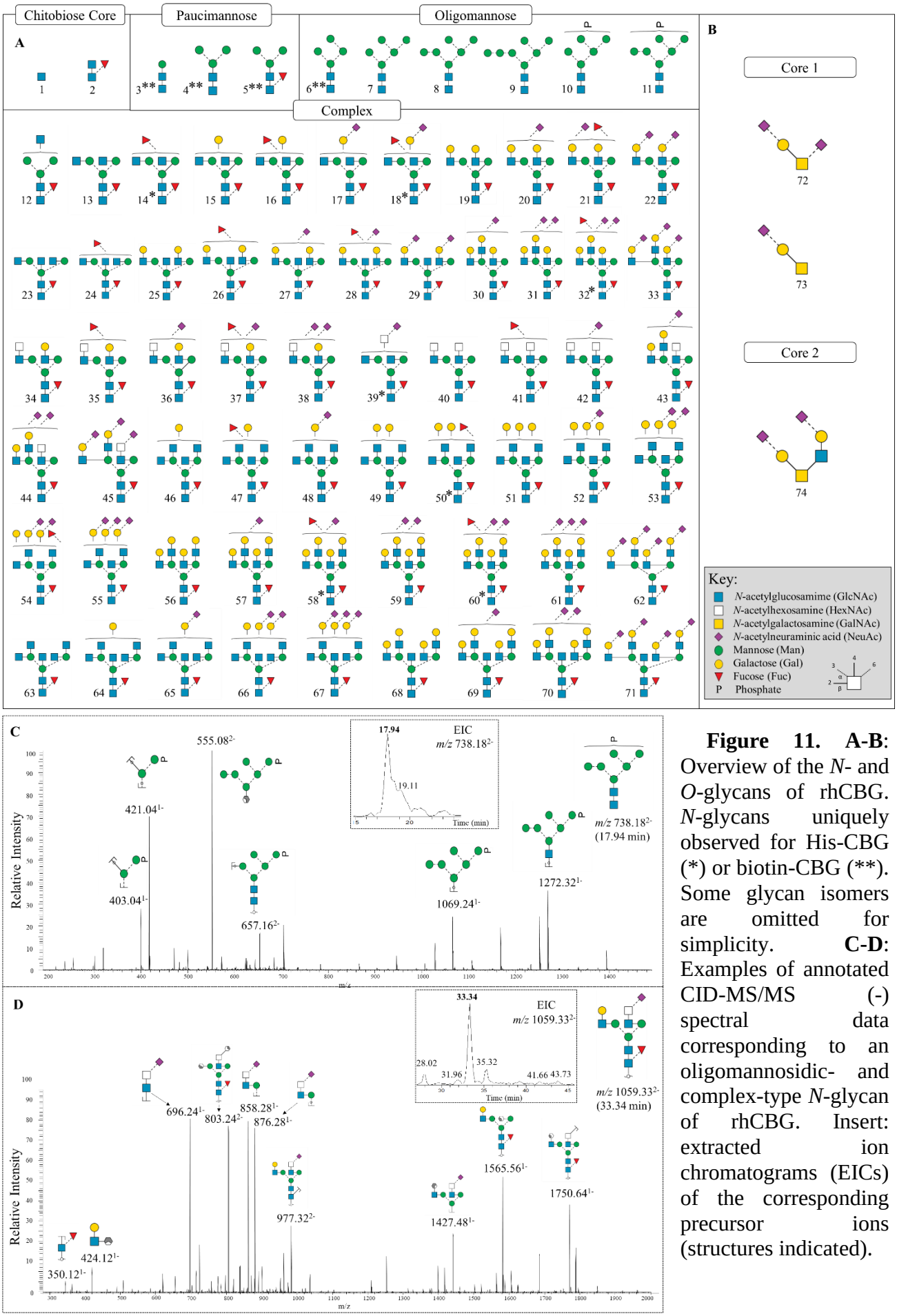


Figure 11. A-B: Overview of the *N*- and *O*-glycans of rhCBG. *N*-glycans uniquely observed for His-CBG (*) or biotin-CBG (**). Some glycan isomers are omitted for simplicity. **C-D**: Examples of annotated CID-MS/MS (-) spectral data corresponding to an oligomannosidic- and complex-type *N*-glycan of rhCBG. Insert: extracted ion chromatograms (EICs) of the corresponding precursor ions (structures indicated).

Table 1. Overview of the glycosylation sites of rhCBG (His- and biotin-CBG), the tryptic (glyco)peptide sequences and masses used for their identification, and their glycan occupancies. *The glycan occupancy of the dipeptide Asn154-Lys155 was not determined directly; instead Asn154-glycosylation was estimated from the Gln147-Lys155 tryptic peptide containing a missed cleavage.

Glycosylation site (residue in mature CBG)	Glycopeptide (trypsin)	Non-glycosylated peptide mass	Site occupancy His-CBG/biotin-CBG
1 (Asn9)	Met1-Arg15	1755.76 Da	31.5 % / 24.4 %
2 (Asn74)	Ala65-Arg78	1558.85 Da	52.6 % / 64.6 %
3 (Asn154)	Asn154-Lys155	260.15 Da	100.0 % / 98.3 %*
4 (Asn238)	Val209-Lys248	4571.22 Da	100.0 % / 100.0 %
5 (Asn308)	Val282-Arg311	3273.59 Da	88.7 % / 93.4 %
6 (Asn347)	Ala327-Arg356	3108.70 Da	*N*-glycans 0.4 % / 1.9 % *O*-glycans 5.9 % / 5.0 %

The intact glycopeptide profiling displayed dramatic site-specific *N*-glycosylation of CBG (**Figure 12**), a feature that also has been observed for native-CBG (discussed below) (76). Adding confidence to the glycoprofile, the identified peptides were found to carry *N*-glycans that accurately matched the already defined "*N*-glycome" of rhCBG (**Figure 11**).

The His-CBG *N*-glycoprofile comprised predominantly complex-type *N*-glycan structures. Asn154, Asn238 and Asn347 displayed markedly less micro-heterogeneity than Asn9, Asn74, and Asn308. Some glycoforms were observed across several sites, e.g. the biantennary mono-sialylated complex-type *N*-glycan (Structure 9, see **Figure 12** for glycan identifiers) was found at all sites except at Asn238. Further, glycans containing *N*,*N*-diacetyllactosamine (LacdiNAc) and sialyl-LacdiNAc moieties (e.g. Structure 20-21) were also observed across all six sites of His-CBG, while the closely related *N*-glycans (Structure 18, 19, 21, 22, and 23) were found with high relative abundance only at Asn74, Asn154, Asn238 and Asn308. The oligomannosidic-type *N*-glycans (i.e. $Man_5GlcNAc_2$, Structure 3) and the M6P-modified oligomannosidic-type *N*-glycans (i.e. $Man_{5-6}GlcNAc_2Phos_1$, Structure 4-5) were generally of low abundance for His-CBG but were still found to be relatively prominent features at Asn9, Asn74 and Asn308.

Biotin-CBG shared several *N*-glycosylation features with the His-CBG at each site but displayed clear qualitative and quantitative differences. As a distinct difference, biotin-CBG was highly glycosylated with oligomannosidic-type *N*-glycans comprising $Man_5GlcNAc_2$ (Structure 3) as the most abundant glycoform identified at Asn9, Asn74, Asn238 and Asn308. In fact, Asn308 (and to some extent also Asn347) displayed a high abundance of the M6P-modified oligomannosidic-type *N*-glycans (i.e. $Man_{5-6}GlcNAc_2Phos_1$, Structure 4-5).

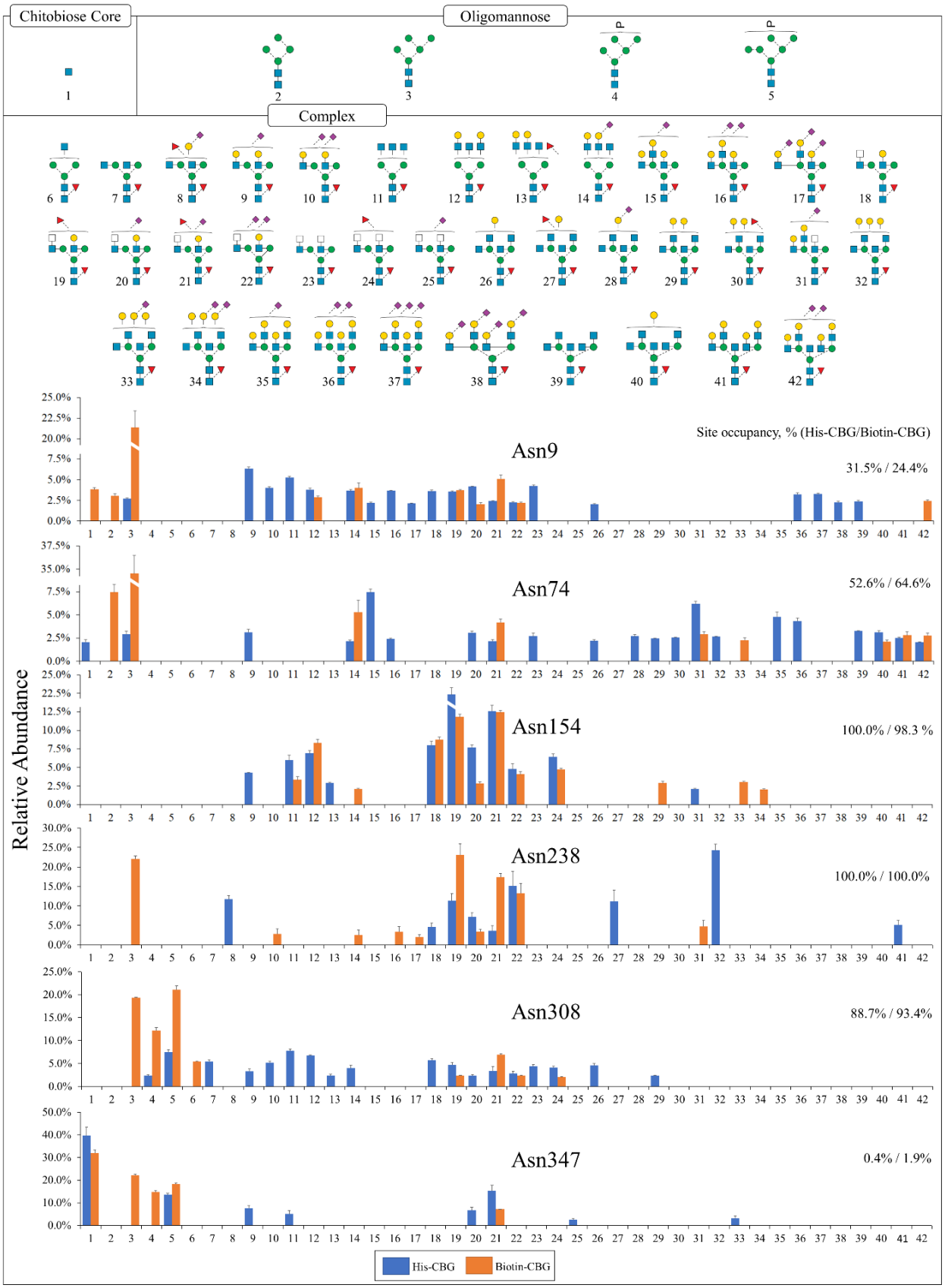


Figure 12. Site-specific *N*-glycoprofiling of His- and biotin-CBG using high-resolution LC-MS/MS-based glycopeptide analysis. The glycan structures, their relative abundances and site occupancies for each of the six *N*-glycosylation sites (Asn9, Asn74, Asn154, Asn238, Asn308, and Asn347) are shown. Low abundance *N*-glycans (<2%) have been omitted in this representation for simplicity.

To my great excitement, the glycopeptide analysis also resulted in the identification of RCL glycopeptides carrying compositions corresponding to *O*-glycans matching structures already identified in the *O*-glycan profiling (**Figure 11B; Appendix 2, Figure S3**). The putative RCL *O*-glycosylation of the His- and biotin-CBG, at that moment still non-localised on the RCL peptide, appeared qualitative and quantitative similar (**Figure 13**). The disialylated core 1-type *O*-glycan ($GalNAc_1Gal_1NeuAc_2$, Structure 1) was observed as the most abundant *O*-glycan carried by His- and biotin-CBG representing 95.8% and 96.5%, respectively, of all observed *O*-glycoforms on the RCL peptides. In contrast, the $GalNAc_1GlcNAc_1Gal_2NeuAc_2$ (Structure 2) and $GalNAc_1Gal_1NeuAc_1$ (Structure 3) were low abundance *O*-glycans of His- and biotin-CBG.

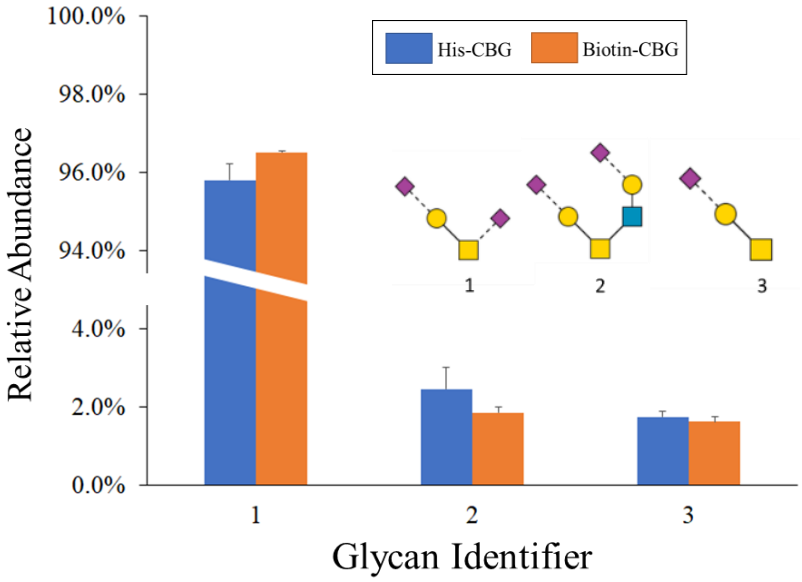


Figure 13. Quantitative RCL *O*-glycopeptide profiling of His- and biotin-CBG using reversed-phase LC-MS/MS (+).

3.3 Identification of a novel *O*-glycosylation site on the RCL of human CBG

The *O*-glycosylation site of the tryptic RCL peptide ^{327}AVLQLNEEGVDTAGSTGVTLNLTSKPIILR^{357} of His- and biotin-CBG was determined using a high-resolution LC-MS/MS-based approach employing orthogonal fragmentation types including HCD, EThcD and resonance-activation (ion trap) CID. The *O*-glycan was confirmed to be conjugated to Thr345 for both His- and biotin-CBG as evidenced by the presence of key glycopeptide fragment ions i.e. z_{10}^{2+} (*m/z* 569.86), z_{11}^{2+} (*m/z* 626.40), z_{12}^{2+} (*m/z* 1150.59), z_{13}^{2+} (*m/z* 1200.12), c_{21}^{2+} (*m/z* 1517.70) and c_{16}^{1+} (*m/z* 1602.80) in the EThcD-MS/MS data generated from an *O*-glycosylated RCL peptide decorated with a monosaccharide composition corresponding to the abundant CBG *O*-glycan

i.e. $GalNAc_1Gal_1NeuAc_2$, **Figure 14**. Importantly, none of the other five putative sites (Thr338, Ser341, Thr342, Thr349, and Ser350) of this (or other) RCL peptide were found to be *O*-glycosylated.

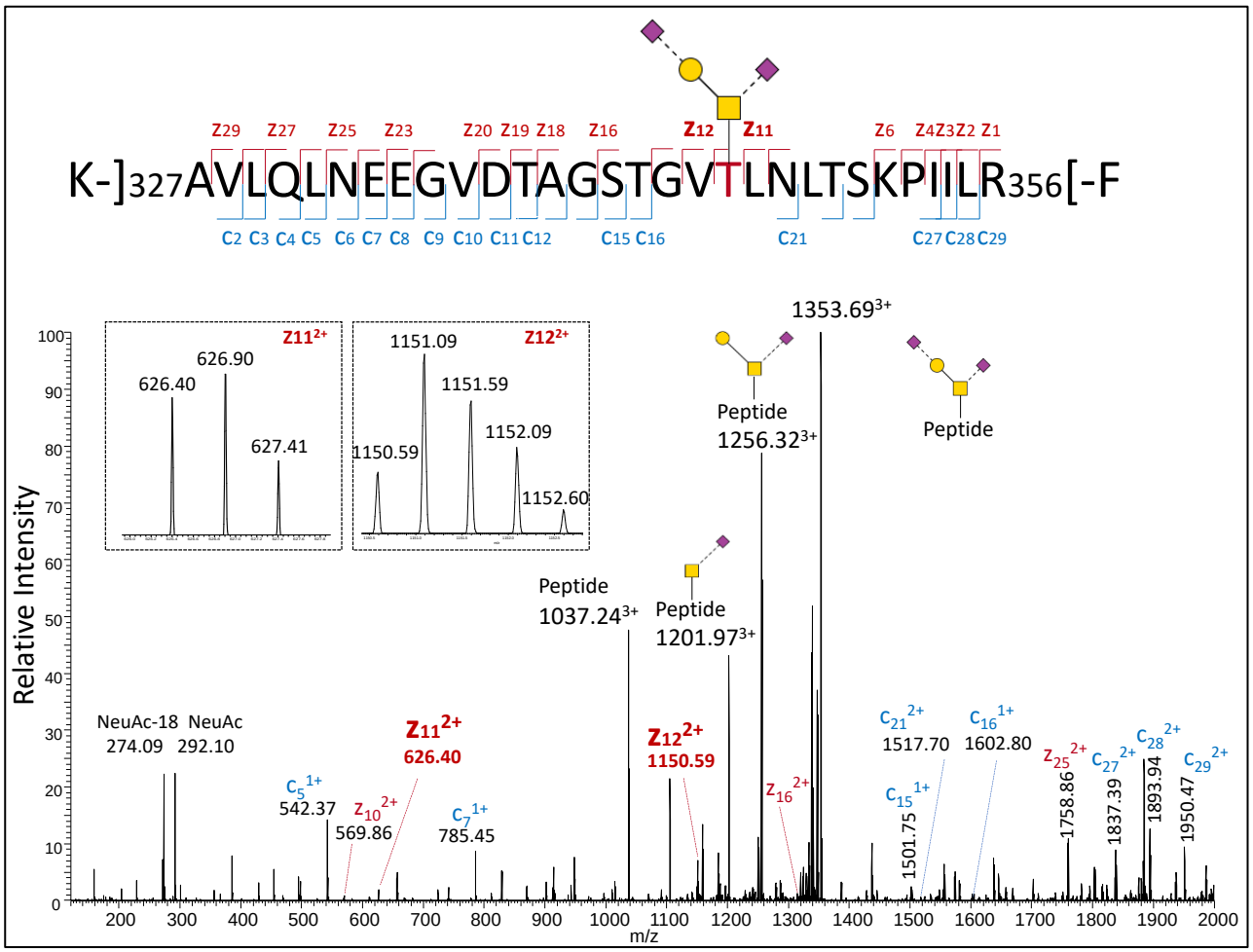


Figure 14. Site localisation of the RCL *O*-glycosylation of His-CBG using a high-resolution Fusion Lumos Orbitrap mass spectrometer capable of performing EThcD-MS/MS. Insert: Zoom of key peptide fragment ions still conjugated with *O*-glycans evidencing Thr345 as the *O*-glycan conjugation site with amino acid resolution. The EThcD-MS/MS of biotin-CBG contained a near-identical fragment pattern and the same key fragment ions also confirming Thr345 for this rhCBG variant (data not shown). The identification was supported by the matching HCD- and ion trap CID-MS/MS, albeit this data was less informative for the validation of the glycopeptide site localisation (data not shown).

No co-occurrence of *N*- and *O*-glycans of ^{327}AVLQLNEEGVDTAGSTGVTLNLTSKPIILR^{357} was identified from the LC-MS/MS-based profiling of the RCL glycopeptides.

3.4 The Thr345 *O*-glycan of CBG protects the RCL from proteolysis

The potential role of the unexpected Thr345 *O*-glycosylation of rhCBG in protecting the RCL from proteolysis (and thereby indirectly impacting the cortisol release mechanism) was firstly assessed using a neutrophil elastase (NE)-based digestion assay. The potent NE is known to rapidly and specifically cleave the RCL at Val344-Thr345 (**Figure 15A**) forming two CBG N- and C-

terminal fragments (CBG-Nt and CBG-Ct, respectively) (85). The working hypothesis was that the newly discovered *O*-glycosylation of Thr345 was likely to play a role in masking (as opposed to promoting) the NE-mediated cleavage event.

Short (15 s - 240 s) incubation of His-CBG with NE generated significant RCL cleavage products forming a large glycosylated CBG-Nt fragment (55-65 kDa) and a much smaller CBG-Ct fragment (~5 kDa) (**Figure 15B**). In-gel trypsin digestion of the CBG-Ct fragment (Band 2) resulted in the identification of the non-glycosylated semi-tryptic peptide ^{345}TLNLTSKPIILR^{357} (*m/z* 684.93^{2+}) using HCD-MS/MS (**Figure 15D**). All the non-glycosylated CBG molecules were cleaved by the 240 s time point as indicated by the absence of the non-glycosylated tryptic RCL peptide ^{327}AVLQLNEEGVDTAGSTGVTLNLTSKPIILR^{357} in Band 1 at this long digestion time point. This non-glycosylated tryptic RCL peptide made up 5% of the uncleaved molecules at 60 s and a staggering 94% before the NE treatment.

Excitingly, the CBG molecules carrying Thr345 *O*-glycosylation were very inefficiently cleaved by NE as evidenced by the over-representation of the RCL tryptic peptide conjugated with the dominant CBG *O*-glycan i.e. $GalNAc_1Gal_1NeuAc_2$ (*m/z* 1353.01 $[M+3H]^{3+}$ and *m/z* 1015.01 $[M+4H]^{4+}$) at the 240 s digestion time point (Band 1, **Figure 15B, C**). The relative quantification of the CBG (glyco)peptides was performed using EIC-based area-under-the-curve measurements of all observed charge states of the monoisotopic precursor ions, a common approach in quantitative glycopeptide profiling studies (102). The inability of NE to cleave Thr345 glycosylated CBG was also visually supported by the generation of a homogenous CBG-Ct fragment migrating as a single protein band (Band 2); Indeed, in the event that NE was able to digest Thr345-glycosylated CBG, multiple bands carrying heterogenous CBG-Ct fragments would have been expected. As has been observed in other studies (85), it was found that an extended incubation with NE (beyond 240 s) resulted in unspecific proteolysis of the intact CBG and CBG-Nt fragments in Band 1 (marked with * in Figure 13B). Hence, the very extended time points were not considered further in this thesis.

Secondly, His-CBG was incubated with a zinc-metalloprotease *P. aeruginosa* elastase (PAE) from the culture media of the virulent *P. aeruginosa* strain PAO1. It is known that the exogenous PAE is less potent than the endogenous NE and consequently requires much longer incubation and more favourable enzyme-substrate ratios to cleave the RCL of human CBG (85).

Ahead of the digestion assay, the presence of PAE in the *P. aeruginosa* PAO1 culture media was firstly assessed using SDS-PAGE (**Figure 16**). Being a major virulence factor secreted from *P. aeruginosa*, PAE was, as expected, identified as an abundant protein band (~35 kDa) in both the crude

and filtered culture media as previously reported (85). The SDS-PAGE revealed a similar concentration of PAE in the two preparations.

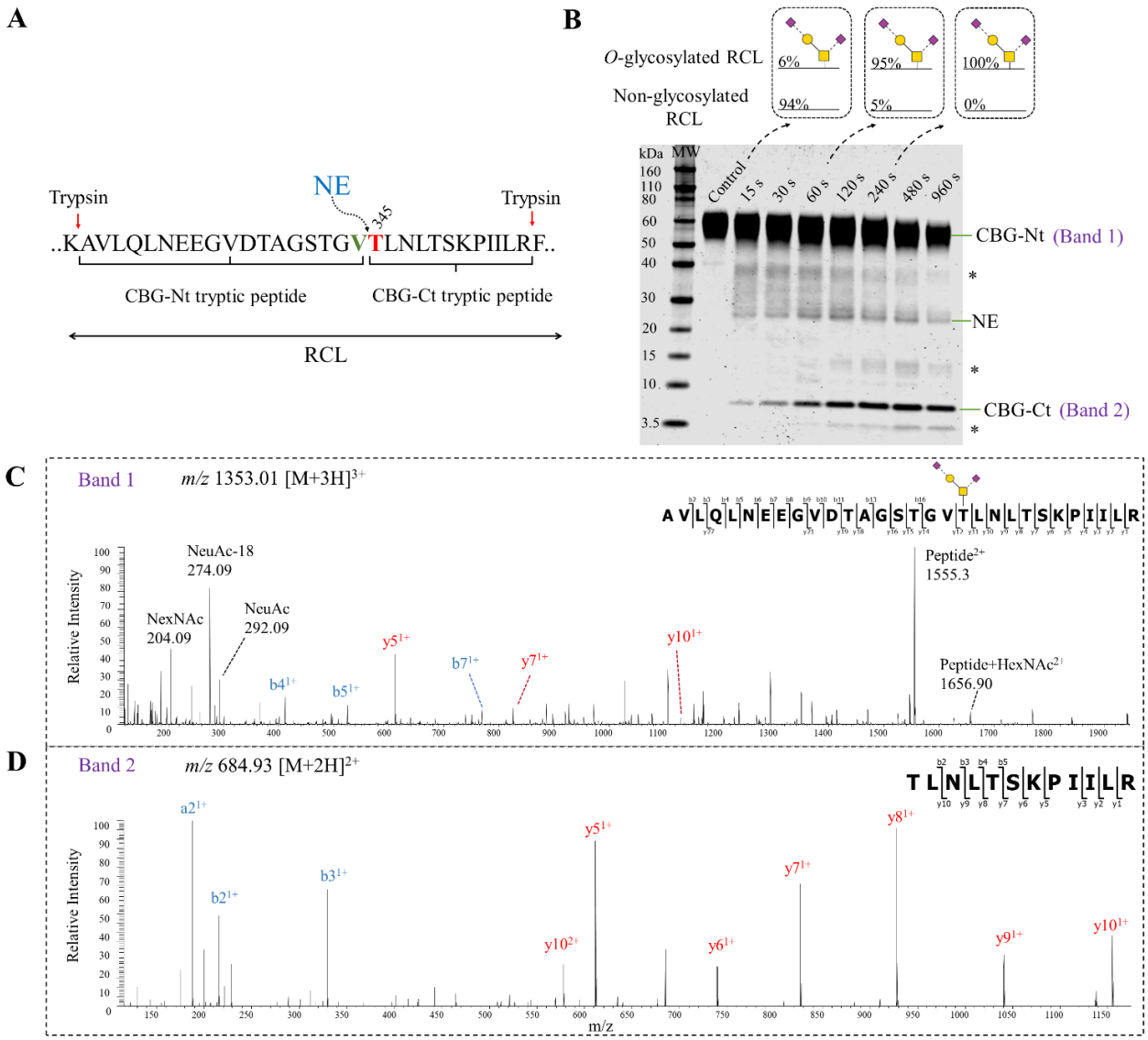


Figure 15. Thr345 *O*-glycosylation strongly retards neutrophil elastase (NE)-mediated RCL cleavage of His-CBG. **A.** Schematics of the CBG reactive center loop (RCL) with an indication of the CBG-Nt and CBG-Ct fragments and the corresponding tryptic peptides generated upon NE cleavage at Val344-Thr345. **B.** SDS-PAGE illustrating the two significant CBG products formed (Band 1 and Band 2) at different time points (0-960 s) resulting from incubation with NE. CBG-Nt-, CBG N-terminal fragment; CBG-Ct-, CBG C-terminal fragment, NE, neutrophil elastase, * Unspecific CBG-Nt fragments (non-Val344-Thr345 cleavage products) occurring after extended NE digestion as identified with LC-MS/MS (data not shown). **C.** Annotated HCD-MS/MS of an *O*-glycosylated RCL peptide (*m/z* 1353.01^{3+}) identified form Band 1. **D.** Annotated HCD-MS/MS of an NE-cleaved semi-tryptic non-glycosylated RCL peptide (*m/z* 684.93^{2+}) identified from Band 2.

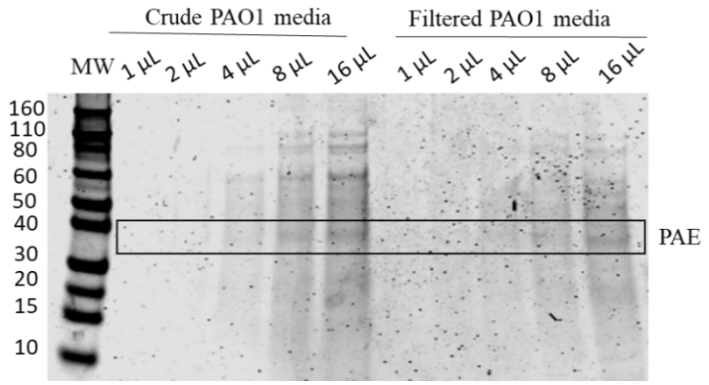


Figure 16. SDS-PAGE of proteins in the culture media of *P. aeruginosa* (PAO1) before (crude) and after (filtered) purification. Different volumes of media were separated to assess the secretion level of the virulent *P. aeruginosa* elastase (PAE) (~ 35 kDa, boxed) of interest.

The filtered PAO1 medium containing PAE was incubated with native- and His-CBG for 1 h, 4 h and 16 h (**Figure 17B**). Despite using identical conditions as previously reported (85), the SDS-PAGE separation of the resulting CBG cleavage products was unfortunately not optimal. Due to the strict time constraint of this thesis, this important experiment could unfortunately, not be repeated. Nevertheless, the incompletely separated CBG products arising from the PAE digestion assay (including an approximately 3.5 kDa CBG-Ct fragment, boxed) clearly indicated that PAE was able to cleave the RCL of both native- and His-CBG under extended incubation conditions.

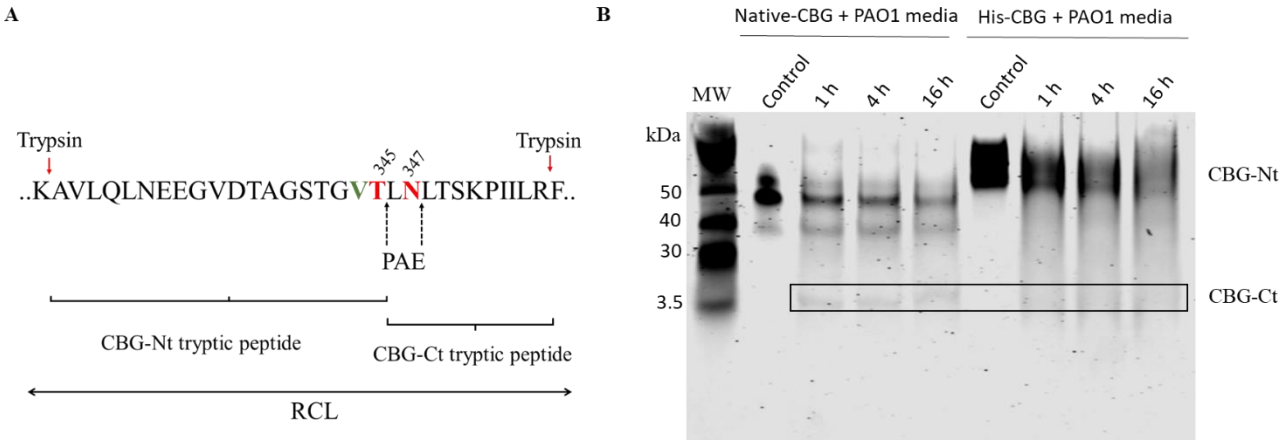


Figure 17. A. Schematics of the CBG RCL with indication of the CBG-Nt and CBG-Ct fragments and the corresponding tryptic peptides generated upon PAE cleavage at Val345-Leu346 and Asn347-Leu348. **B**. SDS-PAGE illustrating the CBG C-terminal fragment (CBG-Ct) resulted from the incubation of native- and His-CBG with filtered culture media of *P. aeruginosa* (PAO1) containing the virulent PAE for 1 h, 4 h, and 16 h.

As was previously reported, PAE is only capable of cleaving 15% of native-CBG representing the subpopulation of CBG molecules lacking Asn347-glycosylation (85). PAE cleaves the RCL at Val345-Leu346 and Asn347-Leu348 (**Figure 17A**). The observation of the single CBG-Ct protein band at 3.5 kDa is in agreement with the previous observation. Further LC-MS/MS-based analysis of the resulting CBG-Ct and CBG-Nt fragments are required to fully establish the PAE-mediated cleavage pattern and efficiency of His-CBG in order to explore the role of RCL Thr345 *O*-glycosylation in the PAE-based RCL cleavage (60, 85).

Taken together, this section has provided evidence for the glycan fine structures and the site-specific distribution of the *N*- and *O*-glycans carried by His- and biotin-CBG. This allowed for the discovery of a novel *O*-glycosylation site (Thr345) strategically located on the RCL of rhCBG. Excitingly, the Thr345-glycan was shown to strongly inhibit the NE-mediated RCL cleavage at Val344-Thr345 implying its biological involvement in regulating the cortisol delivery process.

Chapter 4: Discussion

Protein *N*- and *O*-linked glycosylation are complex and energy-demanding post-translational modifications of proteins trafficking the endoplasmic reticulum and Golgi apparatus of the secretory pathway. Diverse mammalian recombinant expression systems including CHO, BHK, and HEK are available for the production of recombinant glycoproteins that may hold academic value or be useful for biotechnological or clinical applications. The HEK293 cell line, a transformed type of human embryonic kidney cells, is a commonly used expression system capable of the large-scale manufacturing of recombinant glycoproteins (118, 119). The mammalian expression systems are all capable of producing complex-type "human-like" *N*-glycosylation. Integration of functional tags (His-, biotin, FLAG-, etc.) for *in vivo* visualisation or tissue targeting or alternatively for the biochemical purification and characterisation of the recombinant glycoproteins is a common strategy. However, being non-natural expression systems, the actual glycosylation features of recombinant human glycoproteins often differ from the glycosylation of the naturally occurring counterpart (120). This discrepancy from the native glycosylation pattern may not necessarily constitute a concern; it may, in fact, prove advantageous if used in an appropriate context (discussed below).

Detailed biochemical characterisation is often required to document the expressed glyco-epitopes and ensure production consistency prior to application of recombinant glycoproteins, in particular, if these are to be used for human applications.

In this thesis, two HEK293-produced recombinant human CBG (rhCBG) variants (His- and biotin-CBG) were structurally and functionally analysed and compared to already generated knowledge of the natural glycosylation of human serum CBG (native-CBG). As the main protein carrier of the anti-inflammatory cortisol, rhCBG has a significant yet still untapped therapeutic potential to alleviate acute and chronic inflammation as experienced in sepsis, bacterial infections and autoimmune disorders.

4.1 Similarities yet quantitative differences between His- and biotin- CBG

The *N*- and *O*-glycan profiling revealed significant structural overlap of the His- and biotin-CBG glycan structures. This was expected taking into consideration their common expression system and hence, their similar secretory pathways used for their expression. The identified *N*-glycans spanned chitobiose core-, paucimannosidic-, oligomannosidic- and complex type *N*-glycan structures (see Figure 11 for overview). With a few exceptions (e.g. GlcNAcβAsn, Structure 1, which appeared with a very high abundance at Asn347, see Figure 12), the structures were biosynthetically related, and most intermediates in the *N*-glycosylation biosynthetic pathway were observed. The identified *N*- and *O*-glycans of rhCBG resembled previously reported structures for other HEK293-produced recombinant proteins (121–123). The key *N*-glycosylation features included an extensive *N*-glycan heterogeneity, a high degree of core and antenna fucosylation and bisecting β1,4-GlcNAcylation, and a relatively low degree of terminal sialylation, which are all consistent with the glyco-analytical literature reporting on HEK293-expressed *N*-linked glycoproteins (121).

Despite similar *N*- and *O*-glycans observed for His- and biotin-CBG, the glycopeptide profiling revealed unexpected differences in the site-specific distribution of *N*-glycans between the two rhCBG forms. The less processed oligomannosidic-type *N*-glycans quantitatively dominated the biotin-CBG profile while the highly processed complex-type *N*-glycans were predominantly carried by His-CBG, see **Figure 18** for site-specific comparison of the observed *N*-glycan types. In particular, Asn308 and Asn347, but also Asn9 and Asn74 of biotin-CBG were rich in the immature oligomannosidic-type *N*-glycans including, most prominently, $Man_5GlcNAc_2$.

Limited information from the manufacturer was provided about the two recombinant products, leaving me speculating on the observed differences between His- and biotin-CBG. Firstly, the biotin-CBG is known to be co-expressed with a biotin-protein ligase derived from *E.coli* (BirA; EC 6.3.4.15) that recognises a specific 15 amino acid peptide (AviTag) as a biotin acceptor peptide (124). The resulting biotinylation reaction is known to occur in different cellular compartments with different efficiencies. For CBG, it can be assumed that the co-expressed BirA was retained in the endoplasmic reticulum (ER), the potential organelle of the biotinylation of CBG. Following this line of thought, the enzymatic transfer of a biotin group to only weakly processed biotin-CBG trafficking the ER may have created significant occlusion and sterical hindrance of some of the proximal *N*-glycosylation sites of these glycoprotein intermediates, and, as a consequence, retarded their further processing in ER-Golgi pathway. Indirectly supporting the hypothesis of differential processing of the two rhCBG forms in or after the ER, relatively similar site occupancies (a feature introduced pre-protein folding) were observed between His- and biotin-CBG, **Figure 12**. However, other possible reasons for the

differential glycan processing cannot be ruled out including the potential impact on the secretory pathway and residence of protein in each compartment by various factors including, but not limited to, the culture medium, nutrients support, culturing conditions, cell growth rate, cell density and viability and metabolism (125, 126).

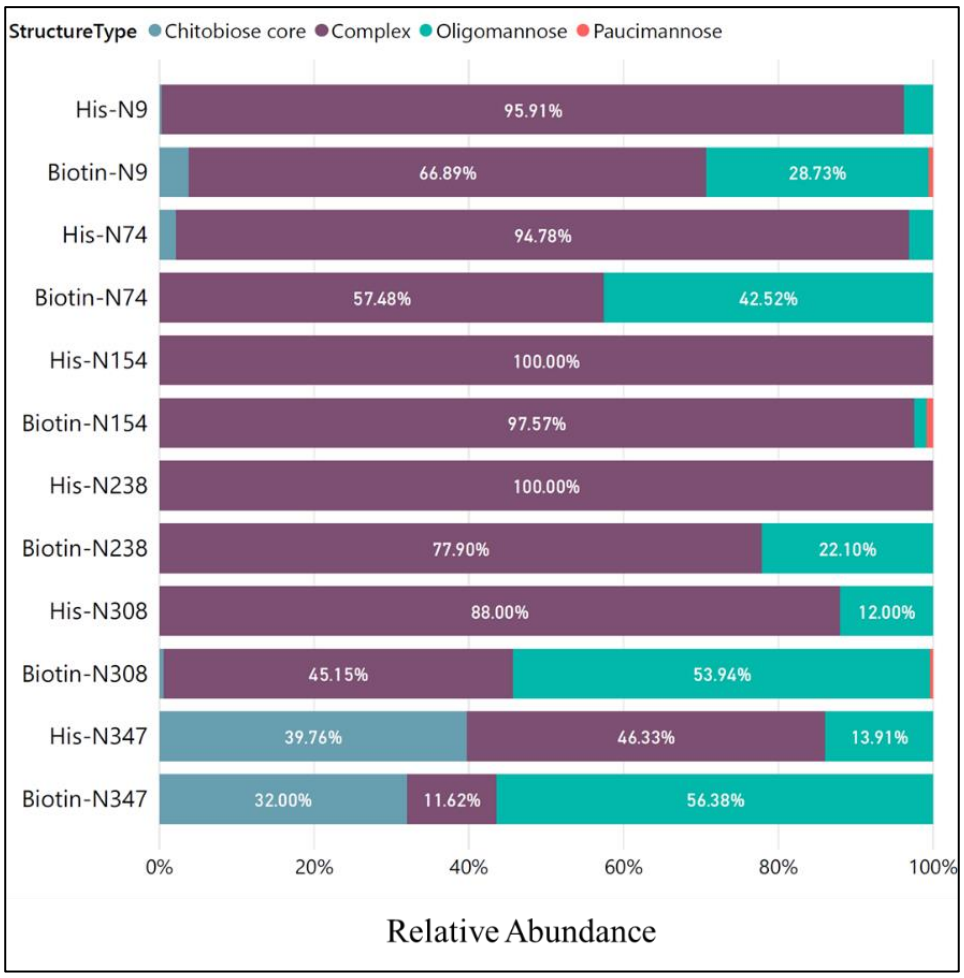


Figure 18. Site-specific distribution of the *N*-glycan types (chitobiose core-, complex-, oligo-mannosidic- and paucimannosidic) of His- and biotin-CBG across the six *N*-glycosylation sites. Strong differences were observed at N9, N74, N308 and N347 between the two rhCBG forms.

Given the therapeutic potential of rhCBG, it becomes relevant to compare the glycosylation of His- and biotin-CBG to the already established glycosylation pattern of naturally occurring CBG (native-CBG) (76). The quantitative comparison of rhCBG and native-CBG revealed significant differences in the site-specific glycosylation and site occupancies. Only approximately 10% of the *N*-glycans observed for His- and biotin-CBG were also found in native-CBG comprising the expected bi-, tri- and tetra-antennary, core-fucosylated and sialylated structures (Structure 17, 22, 31, 33, 59, 61, 62, see glycan identifiers in **Figure 11**). The level of fucosylation was another important difference; while most *N*-glycans of rhCBG were either core or antenna fucosylated (11.62–100.0% across sites), human CBG carried a comparably lower degree of only core fucosylation (0-34.9% across sites) (**Table 2**) (76). As expected, the degree of sialylation of the *N*-glycans was significantly lower for rhCBG than the blood-derived native-CBG. The sialylation of rhCBG was identified as NeuAc(α2,3/6)Gal(β1,4)GlcNAc and NeuAc(α2,3/6)GalNAc(β1,4)GlcNAc (sialyl-LacdiNAc),

whereas native-CBG carried no sialyl-LacdiNAc. Further, native-CBG is known to have a high site occupancies for all sites (70.5-99.5% across sites) (76), whereas His- and biotin-CBG showed a high degree of macro-heterogeneity with some sites being almost unoccupied and other sites being almost fully occupied. The RCL-localised Asn347 was very lowly occupied in rhCBG (0.1-1.9%) as compared to the highly occupied Asn347 in native-CBG (84.7%), a significant difference discussed further below.

Unexpectedly, the glycan profiling data revealed a few prominent mucin-type *O*-glycans for the two rhCBG variants, structures not previously reported for native-CBG (76, 85). Core 1- and core 2-type *O*-glycans are known to be efficiently synthesised by HEK293 (122). The unexpected *O*-glycosylation of rhCBG will be further discussed below in a site-specific context.

Table 2. Quantitative comparison of the site-specific *N*-glycoprofiles generated of rhCBG (His- and biotin-CBG) to native-CBG (76). The *N*-glycan types and total level of fucosylation and sialylation (regardless of linkage type and topology) were compared across the six *N*-glycosylation sites (N9, N74, N154, N238, N308, N347 in unique colour/shading).

Glycosylation Site	Chitobiose Core	Paucimannose	Oligomannose	Complex	Fucosylation	Sialylation	Site Occupancy
His-N9	0.33%	0.00%	3.76%	95.91%	95.91%	56.82%	31.50%
Biotin-N9	3.81%	0.57%	28.73%	66.89%	67.35%	37.29%	22.40%
Native-N9	0.00%	0.00%	0.00%	100.00%	17.20%	100.00%	96.70%
His-N74	2.15%	0.00%	3.07%	94.78%	94.12%	59.60%	52.60%
Biotin-N74	0.00%	0.00%	42.52%	57.48%	57.07%	32.64%	64.60%
Native-N74	0.00%	0.00%	0.00%	100.00%	13.70%	100.00%	99.50%
His-N154	0.00%	0.00%	0.00%	100.00%	100.00%	41.41%	100.00%
Biotin-N154	0.00%	0.85%	1.58%	97.57%	98.42%	39.56%	98.30%
Native-N154	0.00%	0.00%	0.00%	100.00%	14.10%	100.00%	89.00%
His-N238	0.00%	0.00%	0.00%	100.00%	100.00%	40.31%	100.00%
Biotin-N238	0.00%	0.00%	22.10%	77.90%	77.90%	54.79%	100.00%
Native-N238	0.00%	0.00%	0.00%	100.00%	0.00%	100.00%	70.50%
His-N308	0.47%	0.00%	12.00%	87.53%	87.53%	31.49%	88.70%
Biotin-N308	0.56%	0.34%	53.94%	45.15%	44.12%	22.29%	93.40%
Native-N308	0.00%	0.00%	0.00%	100.00%	4.70%	100.00%	96.30%
His-N347	39.76%	0.00%	13.91%	46.33%	46.33%	39.14%	0.40%
Biotin-N347	32.00%	0.00%	56.38%	11.62%	11.62%	9.47%	1.90%
Native-N347	0.00%	0.00%	0.00%	100.00%	34.90%	100.00%	84.70%

4.2 Unexpected RCL *O*-glycosylation of rhCBG

High-resolution LC-MS/MS employing orthogonal fragmentation types including HCD-, and subsequent production-dependent EThcD- and CID-M/MS was used to provide solid biochemical evidence for the presence of RCL Thr345 *O*-glycosylation of the two investigated rhCBG variants, and, further, demonstrated an intriguing role of the Thr345-glycosylation in blocking NE-mediated cleavage, thus, implying a role in impacting the cortisol delivery mechanism. Given the potential

biological importance of this novel CBG modification, several immediate questions arise from these findings, including 1) *Is Thr345-glycosylated CBG expressed and naturally present in human?* 2) *Which tissue(s) produce Thr345-glycosylated CBG and under which physiological conditions?* and finally, 3) *Which glycosylation enzyme(s) are responsible for the Thr345-glycosylation of human CBG?* This section will attempt to briefly address these and related questions.

4.2.1 Does human CBG naturally carry *O*-glycosylation?

A commercial source of purified human CBG, reportedly arising from pooled sera of donors but otherwise isolated using a non-disclosed protocol (Affiland, Belgium), is, to date, the type of natural CBG (herein called native-CBG) that has been characterised in the greatest details (76, 85). Although native-CBG in those papers was not reported to carry any *O*-glycosylation, it can still be challenged if *O*-glycosylated variants of CBG were overlooked in those studies, e.g. via a glycoform-biased isolation strategy of CBG, due to low abundance of the *O*-glycosylation or, more simply, perhaps due to a lack of attention to these unexpected modifications.

Recently, two large-scale LC-MS/MS-based *O*-glycoproteomics studies indicated *O*-glycosylation of serum CBG by reporting (amongst an extensive list of *O*-glycopeptides from other serum glycoproteins) intact *O*-glycopeptides spanning the RCL of CBG (93, 94).

In one of these studies, the bacterial enzyme OpeRATOR, which recognises *O*-linked glycans presented on polypeptide scaffolds and generates glycopeptides by cleaving N-terminal of the *O*-glycosylated Ser/Thr (94), was employed in combination with linkage-unspecific neuraminidase and trypsin to create a large population of relatively easy-to-identify *O*-glycopeptides all carrying N-terminally-located asialo-*O*-glycans. Of relevance to this thesis, the study reported the *O*-glycopeptide ^{342}**T**GVTLNLTSKPIILR^{356} spanning the RCL of human CBG. Two *O*-glycan compositions, i.e. HexNAc(1) and HexNAc(1)Hex(1) were reportedly conjugated to Thr342 corresponding to the precursor ions *m/z* 624.37, $[M+3H]^{3+}$ and *m/z* 678.38, $[M+3H]^{3+}$, respectively. Lys351 was guanidinylated in the identified peptides due to the utilised sample processing procedure.

To scrutinise the report of Thr342 *O*-glycosylation of the CBG RCL peptide, which differed from my own findings of Thr345-glycosylation of rhCBG, I obtained the raw LC-MS/MS data from the OpeRATOR study (publicly available via https://www.ebi.ac.uk/pride/archive/projects/PXD009476) and manually re-interrogated the relevant fragment spectra corresponding to the reported RCL *O*-glycopeptides. Of note in this context, the LC-MS/MS data was surprisingly obtained using

exclusively HCD-MS/MS (NCE 36%), which is known to be less-than-ideal to localise *O*-glycans on peptides due to the prominent loss of the glycan on the generated peptide (b-/y-ion) fragments (80).

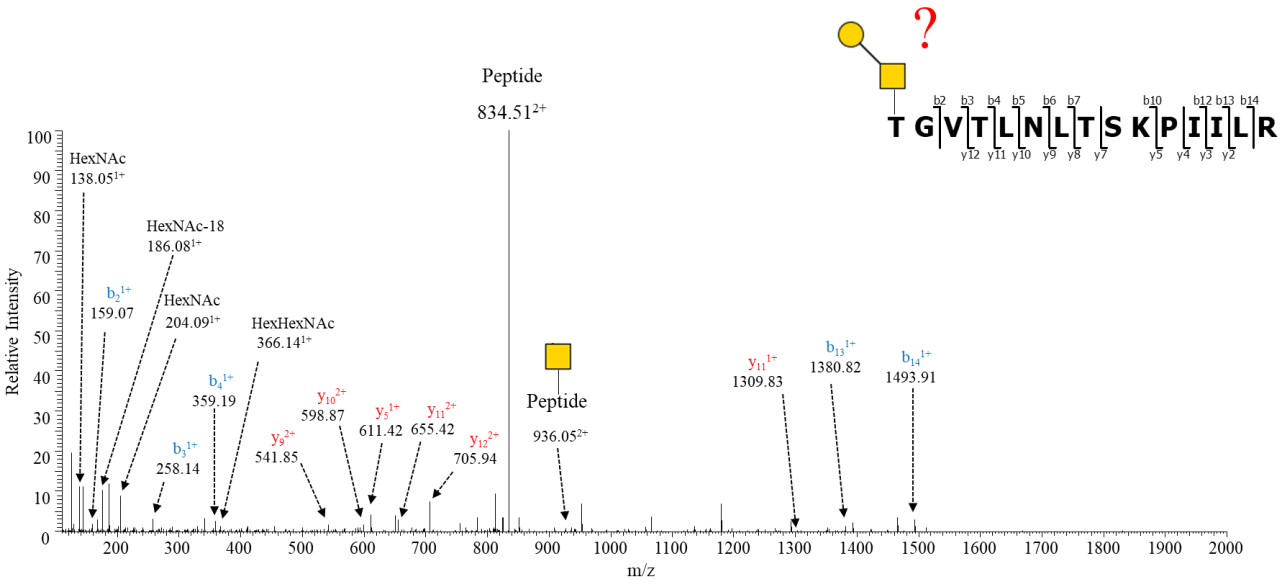


Figure 19. Manual re-annotation of an HCD-MS/MS spectrum previously reported to correspond to a Thr342-glycosylated RCL peptide (^{342}**T**GVTLNLTSKPIILR^{356}) from human serum CBG conjugated with HexNAc(1)Hex(1) (94). The *O*-glycopeptides were generated using 1) trypsin, 2) guanidination of Lys-containing peptides, 3) glycopeptide capture, 4) *O*-glycopeptide release using OpeRATOR, a protease which reportedly cleaves glycopeptides N-terminal of the *O*-glycosylated Ser/Thr, and 5) linkage-unspecific neuraminidase treatment. The re-interrogation clearly demonstrated characteristic oxonium ions and confirmed the peptide backbone sequence (identified b- and y-ions are indicated), but as expected, the peptide fragments appeared without the conjugated *O*-glycan thereby precluding confident assignment of the *O*-glycan site. Thus, no direct evidence for the reported *O*-glycan site (Thr342) could be obtained for this peptide.

Manual annotation of the spectrum corresponding to the CBG peptide ^{342}**T**GVTLNLTSKPIILR^{356} modified with HexNAc(1)Hex(1) confirmed *O*-glycosylation of reported CBG peptide, but did not afford confident site identification (**Figure 19**). Peptide b- and y-ion fragments retaining the sugar moiety (in part or in full) required to confirm the position of the *O*-glycan were not found for any of the putative glycosylation sites of this peptide (Thr342, Thr345, Thr349, and Ser350). Hence, the assigned Thr342 glycosylation site was seemingly deduced solely from the OpeRATOR cleavage pattern, which only provides indirect biochemical evidence for *O*-glycosylation on this site (94). In fact, the OpeRATOR specificity has been questioned (unpublished observation, personal communication with Prof Henrik Clausen, Univ Copenhagen, Denmark), stressing the need to obtain orthogonal LC-MS/MS data using ETD- and/or EThcD-MS/MS that arguably still is the most useful strategy to obtain direct evidence of *O*-glycan site localisation with amino acid resolution (80).

Interestingly, further Byonic-assisted and manual exploration of the OpeRATOR LC-MS/MS raw data also indicated the presence of a CBG *O*-glycopeptide ^{345}**T**LNLTSKPIILR^{356} conjugated with HexNAc(1)Hex(1) (*m/z* 592.68, $[M+3H]^{3+}$) that surprisingly was not reported in the study despite a

high sequence coverage (**Figure 20**). Although peptide-glycan b-/y-ion fragments were also absent in this specific HCD-MS/MS spectrum, this observation supports the natural presence of Thr345- (and possibly Thr342-) *O*-glycosylation of CBG in non-depleted human serum. Upon re-checking my own glycopeptide data, it was confirmed that Thr342 was not utilised for *O*-glycosylation in rhCBG.

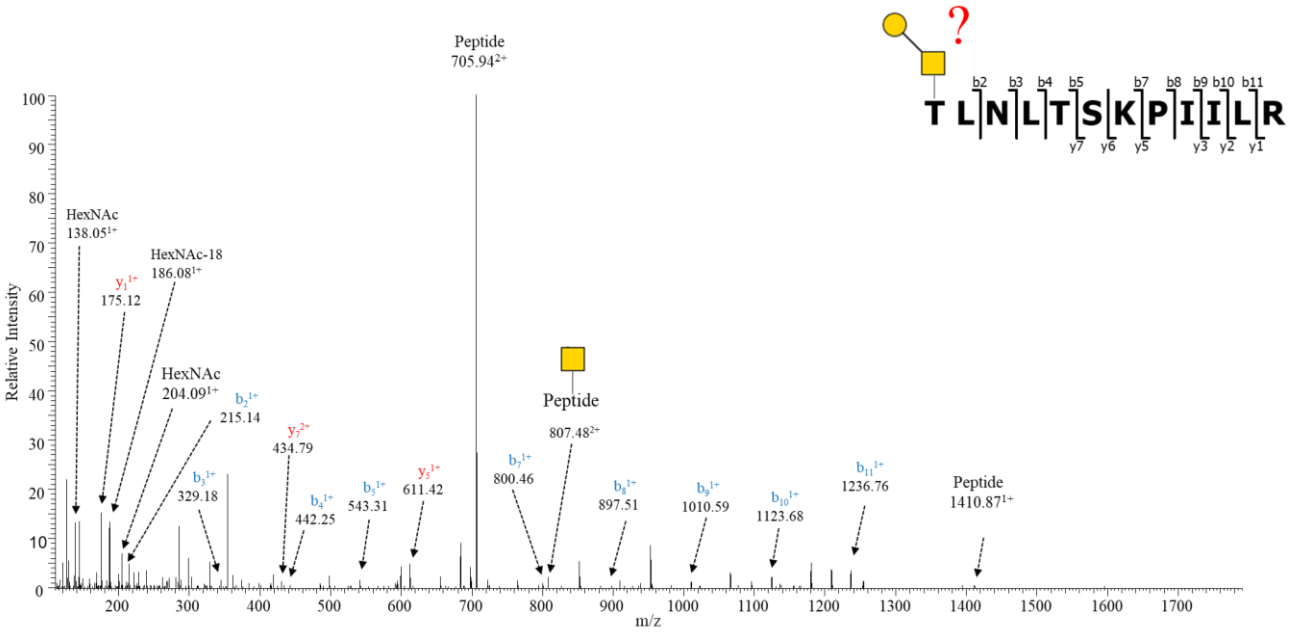


Figure 20. Manual and Byonic-assisted annotation of an HCD-MS/MS spectrum surprisingly not reported in the OpeRATOR study (94). The spectrum was found to correspond to an RCL-located *O*-glycopeptide ^{345}**T**LNLTSKPIILR^{356} modified with HexNAc(1)Hex(1) (*m/z* 592.68, $[M+3H]^{3+}$) from human CBG. The OpeRATOR cleavage pattern points in this case to Thr345 being the *O*-glycosylated site, albeit no direct biochemical evidence from peptide-glycan fragments allowed confirmation of this. Oxonium ions corresponding to the *O*-glycan HexNAc(1)Hex(1) were observed.

Adding further support for Thr345-glycosylation of human CBG, another recent *O*-glycoproteomics study indicated RCL *O*-glycosylation using the SimpleCell technology (127). In this and related studies (128), extensive ETD- and HCD-MS/MS spectral libraries of *O*-glycopeptides identified with confident site assignment were used in conjunction with a high-resolution mass spectrometry-based data-independent acquisition method (called "Glyco-DIA") to search for *O*-glycosylated peptides in serum. Excitingly the Thr345-glycosylation was reported for two peptides spanning the CBG RCL, i.e. ^{330}QLNEEGVDTAGSTGV**T**L^{346} and ^{334}EGVDTAGSTGV**T**L^{346} generated with trypsin and chymotrypsin (127). The conjugated *O*-glycans were reported to be HexNAc(1) moieties as expected for the engineered SimpleCells that were used. This specific information was obtained from the "*O*-glyco Proteome DB Human, *O*-GalNAc human glycoproteome" (http://www.glycoproteomics.somee.com/), an on-line resource build to support the findings presented in the impressive Glyco-DIA study (93). However, it was unfortunately not possible to obtain and scrutinise the underpinning LC-MS/MS data to obtain direct spectral evidence for the Thr345-glycosylation of human serum CBG from this study (93).

The above provides support for the natural presence of CBG modified with Thr345-glycosylation in blood circulation. While these studies did not quantitate the levels of Thr345-glycosylated CBG, the absence of reports of Thr345-glycosylation (or any other CBG *O*-glycosylation) from other studies dealing with human serum CBG (22, 76, 85) indicate that serum CBG, which is originating mostly from liver tissues (129), is only lowly modified with Thr345 *O*-glycosylation. This leads directly to a related question: *Are tissues other than the liver producing Thr345-glycosylated CBG?*

4.2.2 Which tissue(s) produce Thr345-glycosylated CBG and under which conditions?

Solid data on this topic is yet to be established, and the following is largely based on speculations.

The principal human tissue producing CBG is the liver. However, CBG mRNA expression was also found at significant levels in the gallbladder, kidney, colon and pancreas (129). These tissues can, therefore, potentially produce *O*-glycosylated variants of human CBG.

Taking into consideration the identification of Thr345 *O*-glycosylation of HEK293-derived rhCBG communicated herein, it may be proposed that human kidney tissue is a potential source of RCL *O*-glycosylation. Further, it can also be suggested that CBG RCL *O*-glycosylation possibly is a sex-specific modification since studies in mice have indicated that CBG expression in the kidney is only produced in the collecting ducts in females (130). The female kidney CBG was suggested to decrease the intracellular cortisol concentrations since cortisol may compete with another corticosteroid aldosterone for the aldosterone receptor binding. Aligning with this line of thoughts, HEK239 cell line was originally derived from the female embryonic kidney epithelial cells (118, 131, 132), which supports the hypothesis of a gender-specific source of *O*-glycosylated CBG from female kidney tissue. This hypothesis naturally requires further experimental support.

Based on previous reports (93, 94), it may be suggested that RCL *O*-glycosylated CBG can be produced under normal physiology. Further, the observation of significant *O*-glycosylated CBG from the HEK293- and SimpleCell expression system using immortalised (cancerogenic) cell lines with some embryonic (stem cell/differentiation) features may indicate that *O*-glycosylated CBG could be associated with cancer and/or early development as a potential oncofetal antigen. Nevertheless, the *N*- and *O*-glycosylation of CBG in pathophysiological conditions remain to be investigated.

4.2.3 Which glycosylation enzyme(s) are responsible for Thr345-glycosylation of human CBG?

Mucin-type *O*-glycosylation is initiated by a large family of the polypeptide *N*-acetylgalactosamine transferases (GalNAc-Ts) in the Golgi apparatus. In total, 20 GalNAc-Ts, which

display unique substrate (polypeptide) preferences, kinetic properties, and tissue-specific expression, have been identified (74). Polypeptides with dense *O*-glycosylation typically require the cooperation of multiple GalNAc-Ts whereas the site-specific *O*-glycosylation (as is the case for Thr345-glycosylation of CBG) is considered to be catalysed by a single or only few GalNAc-Ts (133).

Identification of the GalNAc-T(s) that may be responsible for the Thr345-glycosylation of human CBG is not trivial, however, the access to the glycoengineered cellular technology platforms (134) developed by the Copenhagen Centre for Glycomics, University of Copenhagen, Denmark may be valuable in this task. As the first step, the level of *O*-glycosylated CBG could be assessed by characterising rhCBG expressed in engineered HEK293 cells where the individual GalNAc-Ts have been systematically eliminated (134). This strategy may provide some insights into potential GalNAc-Ts that are important for CBG *O*-glycosylation. Further, by consulting the existing databases of tissue-specific expression of different GalNAc-T isoforms, it may be possible to identify the most likely tissue(s) producing *O*-glycosylated CBG. Female kidney, being suggested as a potential tissue source of *O*-glycosylated CBG (see above), is known to produce GalNAc-T 2/3 (74), making these candidate enzymes for the initiation of the Thr345 *O*-glycosylation, but concrete evidence is yet to be provided to support these claims and speculations.

4.3 Role of CBG *O*-glycosylation in RCL proteolysis

The proteolysis rate of the RCL of CBG that critically defines the cortisol release process has previously been shown to be impacted by the Asn347-glycosylation localised close to the RCL cleavage sites (85). In fact, the use of glycosylation to directly dictate or perhaps more often just fine-tune the susceptibility of the RCL region to undergo proteolysis is seemingly an often-used feature shared by members of the serpin family (135). In this context, site-specific *O*-glycosylation rather than *N*-glycosylation have more often been implicated in altering the sensitivity to proteolysis or directly change the protease specificity (133, 136). Apart from the steric hindrance elicited by the *O*-glycan, an effect that is likely to contribute to the prevention of proteolysis at a proximal proteolytic cleavage site, *O*-glycans may also, in some cases, change the protein structure in allosteric ways that may affect the susceptibility of that protein to proteolysis.

The role of the intriguing Thr345 *O*-glycosylation of rhCBG reported herein, was investigated using first a neutrophil elastase (NE)-based digestion assay of His-CBG. This assay, which explored the His-CBG cleavage products generated after very short exposure (second-minute time scale) to the potent NE, revealed that the Thr345-glycans strongly inhibited, in fact completely precluded, the NE-mediated RCL cleavage at Val344-Thr345 (**Figure 21A-B**). The highly dense yet flexible nature of

the core 1- and core-2 type *O*-glycans observed at Thr345 of rhCBG and their immediate proximity to the NE cleavage sites are likely to explain the attenuation of NE-mediated proteolysis.

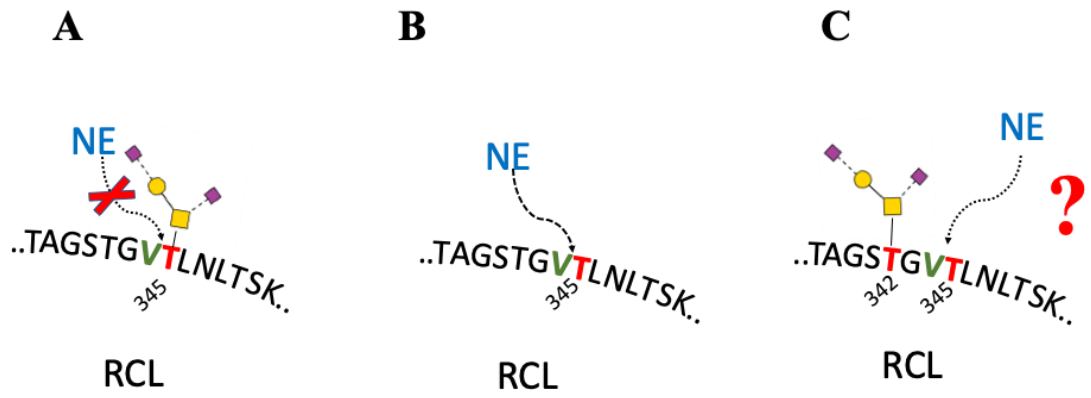


Figure 21. **A**. Thr345-glycosylation of the RCL prevents the NE-mediated proteolysis of rhCBG. **B.** Unoccupied RCL is readily digested by NE at Val344-Thr345. **C**. The effect of RCL Thr342 *O*-glycosylation (if existing at all) on the NE-based RCL cleavage is still unknown.

Thr342 *O*-glycosylation was previously suggested to be a feature of serum CBG (see above) (94). Although this site has still not been experimentally confirmed, it can be speculated that RCL *O*-glycosylation at Thr342 may not negatively impact the NE-mediated proteolysis rate as is the case for Thr345-glycosylation. Based on a similar observation for Asn347-glycosylated human CBG (85), it could be proposed that a position of the *O*-glycan relatively distal to the NE cleavage site and the potential benefit of the electrostatic interactions formed between the positively-charged arginine-rich surface of NE (137) and the negatively-charged *O*-glycan, in concert, may promote the NE-based cleavage efficiency, a hypothesis that deserves to be empirically tested (**Figure 21C**).

Interestingly, the detailed glycopeptide profiling carried out in this thesis did not reveal any co-occurrence of Thr345- and Asn347-glycosylation of the same CBG molecule. Doubly glycosylated peptides are notoriously challenging from an analytical viewpoint, but the use of both HCD- and EThcD-MS/MS in conjunction with deep manual and Byonic-assisted annotation of the resulting fragment spectra provided confidence that Thr345 and Asn347 are not simultaneously utilised for glycosylation in the two investigated rhCBG variants, although the possibility of poor ionisation of doubly glycosylated peptide exists. The “cross-talk” between proximal glycosylation sites has been reported before; it is known that the occupancy of *N*-glycosylation sites of serum proteins is often low when such sites are located near utilised *O*-glycosylation sites (135). Since *N*-glycosylation is initiated as a co-translational event in the ER before the subsequent *O*-glycosylation in the Golgi, it is likely that the occupancy of Asn347 is dictating the utilisation of the Thr345 in the recombinant expression of rhCBG (or Thr342 in other tissues); however, both the biosynthetical aspects of the

interplay or “cross-talk” between the CBG *N*- and *O*-glycosylation and any biological relevance require further exploration.

RCL cleavage may also be induced by other proteases including the *P. aeruginosa* elastase (PAE) secreted as a prominent virulence factor by this opportunistic pathogen commonly found in wounds and in the lungs of individuals suffering from pneumonia and cystic fibrosis (60). PAE is biologically relevant by inducing the release of the anti-inflammatory cortisol via RCL cleavage of CBG, which, in turn, serves to inhibit the host immune response, an outcome that is seen as being beneficial for *P. aeruginosa* (85). The exact role of Thr345 *O*-glycosylation of rhCBG on the PAE-mediated RCL cleavage is yet to be determined as pointed out in the Results section where preliminary data using a PAE digestion assay were presented. Based on the previous studies showing that Asn347-glycosylation of human CBG inhibits the PAE cleavage at Val345-Leu346 and Asn347-Leu348, it may be speculated that Thr345 *O*-glycosylation will also protect, in part or completely, for RCL cleavage.

Taken together, this thesis have provided data evidencing that the RCL Thr345-glycosylation, similar to the Asn347-glycosylation (85), modulates the NE-mediated RCL proteolysis of CBG, and hence impacts the cortisol delivery mechanism. These exciting findings open new avenues for the tailored glycoengineering of rhCBG with unique features for a range of therapeutic applications.

4.4 Therapeutic potential of rhCBG

The chance of survival for individuals undergoing septic shock (multiorgan failure) is strongly associated with the tissue levels of cortisol and the serum concentration of CBG (138). Interestingly, septic shock non-survivors display significantly lower serum CBG concentrations than survivors. Based on this observation, exogenously produced rhCBG or rhCBG loaded with cortisol administered either locally (subcutaneously) or systemically (intravenously) have been proposed as a new therapeutic avenue that may benefit sepsis patients. Currently, individuals displaying sepsis symptoms are often treated using antibiotics, intravenous fluid treatment, as well as vasoactive drugs (139). It is recognised that the effectiveness of the current treatment regime for sepsis and septic shock is low; hence, highly efficacious new treatment strategies are urgently required (139).

Several of the *N*-glycosylation features identified for His- and biotin-CBG can be predicted to affect the pharmacokinetics properties if rhCBG is to be used as a therapeutic agent in the future. For example, rhCBG carrying *N*-glycans capped with galactose, a feature that is relatively common across all *N*-glycosylation sites except for Asn347 (12.3%-58.8%) may be quickly recognised by the hepatic

asialoglycoprotein receptors leading to a rapid clearance of rhCBG from circulation (see **Appendix 3, Figure S4A**) (140). Although the total level of sialylation (predominantly α2,6-, data not shown in graph) was relatively high at each site (9.5%-59.6%), a feature known to increase the blood circulation time (141), the fact that the average rhCBG molecule would still contain multiple asialylated glycans would predictably translate into a relatively low circulation half-life of rhCBG (see **Appendix 3, Figure S4B**). The rhCBG *N*-glycans of most sites (except Asn347) also displayed high levels of total fucosylation of which most were found to be core fucosylation (see **Appendix 3**, **Figure S4C**) while still significant proportions of the fucosylated structures in particular at Asn154 and Asn238 were found to be antennae fucosylated (e.g. sialyl-Lewis^x and Lewis^x epitopes) (see **Appendix 3**, **Figure S4D**). Given binding of the sialyl-Lewis^x epitopes to matching selectin receptors expressed at elevated levels by the endothelial cells lining the blood vasculature around inflamed tissues, rhCBG carrying these features may be an avenue to target cortisol delivery selectively to areas of inflammation (142). In support, the *N*-glycans from serum proteins in septic patients and in acute inflammation after heart surgery show increased levels of sialyl-Lewis^x and Lewis^x epitopes (143, 144).

Mannose-6-phosphorylation (M6P) was also observed as a prevalent modification of the *N*-glycans decorating several sites of His- and biotin-CBG (see **Appendix 3, Figure S4E**). In particular, the Asn308 and Asn347 of biotin-CBG were M6P-rich sites (33.1%-33.2%). M6P-modified glycoproteins are typically regarded as lysosomal since the M6P is a recognised trafficking signal to the lysosome (145). However, serum (non-lysosomal) glycoproteins including several serpins (e.g. TBG (145), are also known to carry M6P-glycosylation. In fact, it was suggested that the SERPIN domain may be structurally similar to the structural determinant of lysosomal proteins important for the M6P-modification of these proteins. The M6P-receptor (MPR-300) is expressed on the surface of T-cells residing at the sites of inflammation (146). Thus, the high level of M6P-glycosylation of biotin-CBG may be a feature that can be used for targeting rhCBG to the inflamed tissues.

Finally, another key feature of rhCBG that makes this artificial form different from (and perhaps beneficial over) native-CBG is the very low Asn347-glycan occupancy (0.4%-1.9%). The under-utilisation of the Asn347 glycosylation site strategically located on the RCL in close proximity to the Val344-Thr345 cleavage site may be beneficial to allow for fast RCL proteolytic cleavage by NE leading, in turn, to a quick cortisol release, a highly desirable feature in sepsis and septic shock. In contrast, the strong inhibitory effect of the Thr345 *O*-glycosylation of rhCBG on the RCL proteolysis rate may be a feature that can be advantageous for some applications of rhCBG in which slow release of cortisol is required, e.g. chronic inflammation. Alternatively, glycoengineering is a powerful tool

that can be used to eliminate the Thr345-glycosylation of rhCBG (e.g. by silencing the relevant GalNAc-T or via site-directed mutagenesis of Thr345) to create a rhCBG variant with completely non-glycosylated RCL for quick NE-mediated proteolysis and release of anti-inflammatory cortisol as it would be required in sepsis and septic shock.

Chapter 5: Conclusions and Future Directions

Detailing the glycosylation carried by CBG with site-specific resolution is essential to improve our understanding of the structure and function of this central immunomodulatory protein. Unravelling the intricate details of the fascinating cortisol release mechanism has been a focus of CBG biologists for many years. CBG is recognised as a potential therapeutic agent for efficient, targeted and timely delivery of cortisol (and potentially other drugs) to sites of inflammation. An impressive body of literature generated over the last decades has beyond doubt demonstrated the significance of glycosylation with respect to CBG biology, but has come short of fully detailing the underpinning glycobiological mechanisms.

To this end, this thesis set out to investigate the structure-function of the glycosylation features carried by two biologically-relevant types of HEK293-derived recombinant human CBG (rhCBG) engineered to harbour attractive functional groups for research and future clinical application purposes including polyhistidine (His-CBG) and biotin (biotin-CBG) tags.

Advanced mass spectrometry-based glycan and glycopeptide profiling were used in concert to obtain both glycan fine structure and site-specific information of rhCBG. His- and biotin-CBG were in general found to carry the same set of *N*-glycans, some mirroring the glycosylation of the natural form of CBG (native-CBG). The *N*-glycans, however, displayed different distribution profiles at each site and between the His- and biotin-CBG variants despite the fact that these two rhCBG variants originated from a common expression system. Biotin-CBG abundantly carried oligomannosidic-type *N*-glycans, while complex-type *N*-glycans were prominent His-CBG features. Of potential utility for future applications of rhCBG, mannose-6-phosphorylation, antenna-fucosylation, bisecting β1,4-GlcNAcylation and LacdiNAcylation, all with matching endogenous or exogenous receptors (lectins) for potential *in vivo* tissue targeting or visualisation or to achieve desired pharmacokinetic properties of rhCBG including prolonged circulation half-life (or rapid clearance) and anti- and pro-immunogenic features were identified with site-specific resolution of the two rhCBG variants.

Relevant to the focus of this thesis, Asn347, which strategically is located on the functionally important RCL region, was found to be largely unoccupied by *N*-glycans for the two rhCBG variants.

Instead, the RCL was excitingly found to be *O*-glycosylated in position Thr345, a feature not previously acknowledged in the large body of CBG literature. Recently, however, a large-scale *O*-glycoproteomics study of human serum indicated RCL *O*-glycosylation (94), but as discussed above, the site assignments, which were based on HCD-MS/MS data and Operator cleavage specificity, were ambiguous. Consequently, it remains unclear if human CBG naturally carries RCL *O*-glycosylation.

Importantly, by using protease digestion assays, it was determined that the Thr345 *O*-glycans of rhCBG strongly inhibit neutrophil elastase (NE)-mediated cleavage of the RCL at position Val344-Thr345. Being able to block the cleavage of NE, a potent serine protease released by activated neutrophils, is a potentially important function of the RCL *O*-glycan that may be used naturally by certain human tissues to express CBG molecules with slow cortisol release properties. Further, this relationship can be used when designing and engineering CBG variants for therapeutic applications. The intriguing role of Thr345 *O*-glycosylation was further investigated using another protease assay employing PAE, a virulence factor secreted by *P. aeruginosa*. The incubation of rhCBG with PAE resulted in RCL cleavage, but due to severe time constraints in this thesis period, it remains to be understood if (and exactly how) the Thr345 *O*-glycosylation impacts the RCL proteolysis by PAE.

This thesis has generated novel data that advance our knowledge of CBG glycobiology, but, at the same time, raised many important questions that are left unanswered. Most importantly, and as already discussed above, future efforts should explore if natural CBG indeed carries RCL *O*-glycosylation, and if so, specify which tissue(s), enzymes and conditions are associated with CBG *O*-glycosylation. Important details regarding the CBG *O*-glycosylation may be studied using the SimpleCell technology, cells with simple *O*-glycosylation (127, 128, 134) or in glycoengineered CHO, HEK239 and/or HepG2 cell systems that have been genetically modified to express (or alternatively silenced for) the ensemble of GalNAc-Ts that putative are responsible for catalysing the CBG Thr345-glycosylation. These *in vitro*-based experiments may then subsequently lead to glycobiological studies involving CBG obtained directly from donors of different age/gender and from patients affected by different conditions relevant to CBG biology, including sepsis and pregnancy. These are indeed examples of research questions and approaches that can be explored in a potential PhD project building on findings presented herein. Ultimately, improving our knowledge of the structure-function of CBG may aid the development of CBG-based therapeutical products.

During this short nine-month project period, I have been introduced to the fascinating world of analytical glycoimmunology, and have, in this process, been getting a taste of cutting-edge research in an exciting and rapidly progressing area of science. While becoming proficient in the arts of advanced mass spectrometry and in the immensely powerful glycomics and glycopeptide profiling

technologies have been highly rewarding, stimulating and educational, I am delighted that this thesis has contributed, albeit only in an incremental manner (see **Appendix 5** for manuscript submitted to *Endocrinology* and **Appendix 4** for poster presented at 18th HUPO World Congress, Sep 2019, based on data presented in this thesis), to a knowledge base that will be of immediate and future use to glycoimmunologists and biochemists alike.

References

1. Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., and Seeberger, P. H. (2015) *Essentials of Glycobiology* (Cold Spring Harbor Laboratory Press, New York) 3rd edition

2. Underhill, D. A., and Hammond, G. L. (1989) Organization of the Human Corticosteroid Binding Globulin Gene and Analysis of Its 5′-Flanking Region. *Mol. Endocrinol.* 3, 1448–1454

3. Seralini, G. E., Bérubé, D., Gagné, R., and Hammond, G. L. (1990) The human corticosteroid binding globulin gene is located on chromosome 14q31–q32.1 near two other serine protease inhibitor genes. *Hum. Genet.* 86, 73–75

4. Zhao, H., Friedman, R. D., and Fournier, R. E. K. (2007) The Locus Control Region Activates Serpin Gene Expression through Recruitment of Liver-Specific Transcription Factors and RNA Polymerase II. *Mol. Cell. Biol.* 27, 5286–5295

5. Gettins, P. G. W. (2002) Serpin Structure, Mechanism, and Function. *Chem. Rev.* 102, 4751–4804

6. Breuner, C. W., and Orchinik, M. (2002) Plasma binding proteins as mediators of corticosteroid action in vertebrates. *J. Endocrinol.* 175, 99–112

7. Schlechte, J. A., and Hamilton, D. (1987) The Effect of Glucocorticoids on Corticosteroid Binding Globulin. *Clin. Endocrinol. (Oxf).* 27, 197–203

8. Cruz-Topete, D., and Cidlowski, J. A. (2014) One hormone, two actions: Anti- and pro-inflammatory effects of glucocorticoids. *Neuroimmunomodulation* 22, 20–32

9. Underhill, D. A., and Hammond, G. L. (1995) cis-Regulatory elements within the proximal promoter of the rat gene encoding corticosteroid-binding globulin. *Gene* 162, 205–211

10. Verhoog, N., Allie-Reid, F., Berghe, W. Vanden, Smith, C., Haegeman, G., Hapgood, J., and Louw, A. (2014) Inhibition of corticosteroid-binding globulin gene expression by glucocorticoids involves C/EBPβ. *PLoS One* 9, 1–14

11. Beishuizen, A., Thijs, L. G., and Vermes, I. (2001) Patterns of corticosteroid-binding globulin and the free cortisol index during septic shock and multitrauma. *Intensive Care Med.* 27, 1584–1591

12. Bernier, J., Jobin, N., Emptoz-Bonneton, A., Pugeat, M. M., and Garrel, D. R. (1998) Decreased corticosteroid-binding globulin in burn patients: Relationship with interleukin-6 and fat in nutritional support. *Crit. Care Med.* 26, 452–460

13. Bartalena, L., Hammond, G. L., Farsetti, A., Flink, I. L., and Robbins, J. (1993) Interleukin-6 inhibits corticosteroid-binding globulin synthesis by human hepatoblastoma-derived (Hep G2) cells. *Endocrinology* 133, 291–296

14. Emptoz-Bonneton, A., Crave, J. C., Lejeune, H., Brébant, C., and Pugeat, M. (1997) Corticosteroid-Binding Globulin Synthesis Regulation by Cytokines and Glucocorticoids in

Human Hepatoblastoma-Derived (HepG2) Cells. *J. Clin. Endocrinol. Metab.* 82, 3758–3762

15. Mihrshahi, R., Lewis, J. G., and Ali, S. O. (2006) Hormonal effects on the secretion and glycoform profile of corticosteroid-binding globulin. *J. Steroid Biochem. Mol. Biol.* 101, 275–285

16. Crave, J. C., Lejeune, H., Brébant, C., Baret, C., and Pugeat, M. (1995) Differential effects of insulin and insulin-like growth factor I on the production of plasma steroid-binding globulins by human hepatoblastoma-derived (Hep G2) cells. *J. Clin. Endocrinol. Metab.* 80, 1283–1289

17. White, T., Özel, B., Jain, J. K., and Stanczyk, F. Z. (2006) Effects of transdermal and oral contraceptives on estrogen-sensitive hepatic proteins. *Contraception* 74, 293–296

18. Sitruk-Ware, R., and Nath, A. (2011) Metabolic effects of contraceptive steroids. *Rev. Endocr. Metab. Disord.* 12, 63–75

19. Lei, J. H., Yang, X., Peng, S., Li, Y., Underhill, C., Zhu, C., Lin, H. Y., Wang, H., and Hammond, G. L. (2015) Impact of corticosteroid-binding globulin deficiency on pregnancy and neonatal sex. *J. Clin. Endocrinol. Metab.* 100, 1819–1827

20. Hammond, G. L. (2016) Plasma steroid-binding proteins: Primary gatekeepers of steroid hormone action. *J. Endocrinol.* 230, R13–R25

21. Shen, M., Cao, J., and Shi, H. (2018) Effects of Estrogen and Estrogen Receptors on Transcriptomes of HepG2 Cells: A Preliminary Study Using RNA Sequencing. *Int. J. Endocrinol.* 2018, 1–16

22. Hammond, G. L., Smith, C. L., and Underhill, D. A. (1991) Molecular studies of corticosteroid binding globulin structure, biosynthesis and function. *J. Steroid Biochem. Mol. Biol.* 40, 755–762

23. Hammond, G. L., Smith, C. L., Goping, I. S., Underhill, D. A., Harley, M. J., Reventos, J., Musto, N. A., Gunsalus, G. L., and Bardin, C. W. (1987) Primary structure of human corticosteroid binding globulin, deduced from hepatic and pulmonary cDNAs, exhibits homology with serine protease inhibitors. *Proc. Natl. Acad. Sci. United States Am.* 84, 5153–5157

24. Schäfer, H. H., Gebhart, V. M., Hertel, K., and Jirikowski, G. F. (2014) Expression of Corticosteroid-Binding Globulin CBG in the Human Heart. *Horm. Metab. Res.* 47, 596–599

25. Miska, W., Peña, P., Villegas, J., and Sánchez, R. (2004) Detection of a CBG-like protein in human Fallopian tube tissue. *Andrologia* 36, 41–46

26. Misao, R., Itoh, N., Mori, H., Fujimoto, J., and Tamaya, T. (1994) Sex hormone-binding globulin mRNA levels in human uterine endometrium. *Eur. J. Endocrinol.* 131, 623–629

27. Misao, R., Iwagaki, S., Sun, W. S., Fujimoto, J., Saio, M., Takami, T., and Tamaya, T. (1999) Evidence for the synthesis of corticosteroid-binding globulin in human placenta. *Horm. Res.* 51, 162–167

28. Pugeat, M., Grasa, M. A. R., Vendrell, J., Brun, J., and Ricart, W. (2002) Serum Corticosteroid-Binding Globulin Concentration and Insulin Resistance Syndrome: A Population Study. *J.*

Clin. Endocrinol. Metab. 87, 4686–4690

29. Nenke, M. A., Zeng, A., Meyer, E. J., Lewis, J. G., Rankin, W., Johnston, J., Kireta, S., Jesudason, S., and Torpy, D. J. (2017) Differential Effects of Estrogen on Corticosteroid-Binding Globulin Forms Suggests Reduced Cleavage in Pregnancy. *J. Endocr. Soc.* 1, 202–210

30. Ho, J. T., Al-Musalhi, H., Chapman, M. J., Quach, T., Thomas, P. D., Bagley, C. J., Lewis, J. G., and Torpy, D. J. (2006) Septic shock and sepsis: A comparison of total and free plasma cortisol levels. *J. Clin. Endocrinol. Metab.* 91, 105–114

31. Garrel, D. R. (1996) Corticosteroid-Binding Globulin during Inflammation and Burn Injury: Nutritional Modulation and Clinical Implications. *Horm. Res.* 45, 245–251

32. Lewis, J. G., Bagley, C. J., Elder, P. A., Bachmann, A. W., and Torpy, D. J. (2005) Plasma free cortisol fraction reflects levels of functioning corticosteroid-binding globulin. *Clin. Chim. Acta* 359, 189–194

33. Mickelson, K. E., Forsthoefel, J., and Westphal, U. (1981) Steroid-protein interactions. Human corticosteroid binding globulin: some physicochemical properties and binding specificity. *Biochemistry* 20, 6211–6218

34. Henley, D., Lightman, S., and Carrell, R. (2016) Cortisol and CBG — Getting cortisol to the right place at the right time. *Pharmacol. Ther.* 166, 128–135

35. Simard, M., Hill, L. A., Lewis, J. G., and Hammond, G. L. (2015) Naturally Occurring Mutations of Human Corticosteroid-Binding Globulin. *J. Clin. Endocrinol. Metab.* 100, 129–139

36. Emptoz-Bonneton, A., Cousin, P., Pugeat, M., Seguchi, K., Bully, C., Avvakumov, G. V., and Hammond, G. L. (2000) Novel Human Corticosteroid-Binding Globulin Variant with Low Cortisol-Binding Affinity. *J. Clin. Endocrinol. Metab.* 85, 361–367

37. Van Baelen, H., Brepoels, R., and De Moor, P. (1982) Transcortin Leuven: A variant of human corticosteroid-binding globulin with decreased cortisol-binding affinity. *J. Biol. Chem.* 257, 3397–3400

38. Perogamvros, I., Underhill, C., Henley, D. E., Hadfield, K. D., Newman, W. G., Ray, D. W., Lightman, S. L., Hammond, G. L., and Trainer, P. J. (2010) Novel corticosteroid-binding globulin variant that lacks steroid binding activity. *J. Clin. Endocrinol. Metab.* 95, E142–E150

39. Hill, L. A., Vassiliadi, D. A., Simard, M., Pavlaki, A., Perogamvros, I., Hadjidakis, D., and Hammond, G. L. (2012) Two Different Corticosteroid-Binding Globulin Variants that Lack Cortisol-Binding Activity in a Greek Woman. *J. Clin. Endocrinol. Metab.* 97, 4260–4267

40. Lin, H. Y., Underhill, C., Lei, J. H., Helander-Claesson, A., Lee, H. Y., Gardill, B. R., Muller, Y. A., Wang, H., and Hammond, G. L. (2012) High frequency of SERPINA6 polymorphisms that reduce plasma corticosteroid-binding globulin activity in Chinese subjects. *J. Clin. Endocrinol. Metab.* 97, 678–686

41. Torpy, D. J., Bachmann, A. W., Grice, J. E., Fitzgerald, S. P., Phillips, P. J., Whitworth, J. A., and Jackson, R. V (2001) Familial Corticosteroid-Binding Globulin Deficiency Due to a Novel

Null Mutation : Association with Fatigue and Relative Hypotension. *J. Clin. Endocrinol. Metab.* 86, 3692–3700

42. Oakley, R. H., and Cidlowski, J. A. (2013) The biology of the glucocorticoid receptor: New signaling mechanisms in health and disease. *J. Allergy Clin. Immunol.* 132, 1033–1044

43. Zänkert, S., Bellingrath, S., Wüst, S., and Kudielka, B. M. (2018) HPA axis responses to psychological challenge linking stress and disease: What do we know on sources of intra- and interindividual variability? *Psychoneuroendocrinology*, 1–12

44. Liyanarachchi, K., Ross, R., and Debono, M. (2017) Human studies on hypothalamo-pituitary-adrenal (HPA) axis. *Best Pract. Res. Clin. Endocrinol. Metab.* 31, 459–473

45. Oakley, R. H., and Cidlowski, J. A. (2013) The biology of the glucocorticoid receptor: New signaling mechanisms in health and disease. *J. Allergy Clin. Immunol.* 132, 1033–1044

46. Perogamvros, I., Ray, D. W., and Trainer, P. J. (2012) Regulation of cortisol bioavailability—effects on hormone measurement and action. *Nat. Rev. Endocrinol.* 8, 717–727

47. Moisan, M. P., and Castanon, N. (2016) Emerging Role of Corticosteroid-binding globulin in Glucocorticoid-Driven metabolic disorders. *Front. Endocrinol. (Lausanne).* 7, 1–5

48. Kumsta, R., Entringer, S., Hellhammer, D. H., and Wüst, S. (2007) Cortisol and ACTH responses to psychosocial stress are modulated by corticosteroid binding globulin levels. *Psychoneuroendocrinology* 32, 1153–1157

49. Perogamvros, I., Aarons, L., Miller, A. G., Trainer, P. J., and Ray, D. W. (2011) Corticosteroid-binding globulin regulates cortisol pharmacokinetics. *Clin. Endocrinol. (Oxf).* 74, 30–36

50. Mendel, C. M. (1989) The Free Hormone Hypothesis: A Physiologically Based Mathematical Model. *Endocr. Rev.* 10, 232–274

51. Meyer, E. J., Nenke, M. A., Lewis, J. G., and Torpy, D. J. (2017) Corticosteroid-binding globulin: acute and chronic inflammation. *Expert Rev. Endocrinol. Metab.* 12, 241–251

52. De Bosscher, K. (2000) Mechanisms of anti-inflammatory action and of immunosuppression by glucocorticoids: negative interference of activated glucocorticoid receptor with transcription factors. *J. Neuroimmunol.* 109, 16–22

53. Busillo, J. M., and Cidlowski, J. A. (2013) The five Rs of glucocorticoid action during inflammation: Ready, reinforce, repress, resolve, and restore. *Trends Endocrinol. Metab.* 24, 109–119

54. Defaye, G., Basset, M., Monnier, N., and Chambaz, E. M. (1980) Electron spin resonance study of human transcortin thiol groups and binding site topography. *Biochim. Biophys. Acta - Protein Struct.* 623, 280–294

55. Hammond, G. L. (1990) Molecular properties of corticosteroid binding globulin and the sex-steroid binding proteins. *Endocr. Rev.* 11, 65–79

56. Klieber, M. A., Underhill, C., Hammond, G. L., and Muller, Y. A. (2007) Corticosteroid-binding Globulin, a Structural Basis for Steroid Transport and Proteinase-triggered Release. *J.*

Biol. Chem. 282, 29594–29603

57. Gardill, B. R., Vogl, M. R., Lin, H. Y., Hammond, G. L., and Muller, Y. A. (2012) Corticosteroid-Binding Globulin: Structure-Function Implications from Species Differences. *PLoS One* 7, 1–11

58. Woods Group GLYCAM Web. *Complex Carbohydr. Res. Center, Univ. Georg. Athens, GA.(http//glycam.org)*

59. Chan, W. L., Zhou, A., and Read, R. J. (2014) Towards engineering hormone-binding globulins as drug delivery agents. *PLoS One* 9, 1–21

60. Simard, M., Hill, L. A., Underhill, C. M., Keller, B. O., Villanueva, I., Hancock, R. E. W., and Hammond, G. L. (2014) Pseudomonas aeruginosa elastase disrupts the cortisol-binding activity of corticosteroid-binding globulin. *Endocrinology* 155, 2900–2908

61. Lewis, J. G., and Elder, P. A. (2014) The reactive centre loop of corticosteroid-binding globulin (CBG) is a protease target for cortisol release. *Mol. Cell. Endocrinol.* 384, 96–101

62. Pemberton, P. A., Stein, P. E., Pepys, M. B., Potter, J. M., and Carrell, R. W. (1988) Hormone binding globulins undergo serpin conformational change in inflammation. *Nature* 336, 257–258

63. Hammond, G. L., Smith, C. L., Paterson, N. A. M., and Sibbald, W. J. (1990) A role for corticosteroid-binding globulin in delivery of cortisol to activated neutrophils. *J. Clin. Endocrinol. Metab.* 71, 34–39

64. Lin, H. Y., Muller, Y. A., and Hammond, G. L. (2010) Molecular and structural basis of steroid hormone binding and release from corticosteroid-binding globulin. *Mol. Cell. Endocrinol.* 316, 3–12

65. Zhou, A., Wei, Z., Stanley, P. L. D., Read, R. J., Stein, P. E., and Carrell, R. W. (2008) The S-to-R Transition of Corticosteroid-Binding Globulin and the Mechanism of Hormone Release. *J. Mol. Biol.* 380, 244–251

66. Milroy, L.-G., Zhou, A., Ley, S. V., Wei, Z., Myers, R. M., Qi, X., Read, R. J., Yan, Y., Chan, W. L., Loiseau, F., and Carrell, R. W. (2011) Allosteric Modulation of Hormone Release from Thyroxine and Corticosteroid-binding Globulins. *J. Biol. Chem.* 286, 16163–16173

67. Moremen, K. W., Tiemeyer, M., and Nairn, A. V. (2012) Vertebrate protein glycosylation: Diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* 13, 448–462

68. Li, H., and d'Anjou, M. (2009) Pharmacological significance of glycosylation in therapeutic proteins. *Curr. Opin. Biotechnol.* 20, 678–684

69. Kelleher, D. J., and Gilmore, R. (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* 16, 47–62

70. Larkin, A., and Imperiali, B. (2011) The Expanding Horizons of Asparagine-Linked Glycosylation. *Biochemistry* 50, 4411–4426

71. Loke, I., Kolarich, D., Packer, N. H., and Thaysen-Andersen, M. (2016) Emerging roles of

protein mannosylation in inflammation and infection. *Mol. Aspects Med.* 51, 31–55

72. Loke, I., Østergaard, O., Heegaard, N. H. H., Packer, N. H., and Thaysen-Andersen, M. (2017) Paucimannose-Rich N-glycosylation of Spatiotemporally Regulated Human Neutrophil Elastase Modulates Its Immune Functions. *Mol. Cell. Proteomics* 16, 1507–1527

73. Thaysen-Andersen, M., Venkatakrishnan, V., Loke, I., Laurini, C., Diestel, S., Parker, B. L., and Packer, N. H. (2015) Human Neutrophils Secrete Bioactive Paucimannosidic Proteins from Azurophilic Granules into Pathogen-Infected Sputum. *J. Biol. Chem.* 290, 8789–8802

74. Bennett, E. P., Mandel, U., Clausen, H., Gerken, T. A., Fritz, T. A., and Tabak, L. A. (2012) Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* 22, 736–756

75. Jensen, P. H., Kolarich, D., and Packer, N. H. (2010) Mucin-type O-glycosylation - putting the pieces together. *FEBS J.* 277, 81–94

76. Sumer-Bayraktar, Z., Kolarich, D., Campbell, M. P., Ali, S., Packer, N. H., and Thaysen-Andersen, M. (2011) N-glycans Modulate the Function of Human Corticosteroid-Binding Globulin. *Mol. Cell. Proteomics* 10, 1–14

77. Thaysen-andersen, M., Packer, N. H., and Schulz, B. L. (2016) Maturing Glycoproteomics Technologies Provide Unique Structural Insights into the N -glycoproteome and Its Regulation in Health and Disease *. *Mol. Cell. Proteomics* 15.6, 1773–1790

78. Jensen, P. H., Karlsson, N. G., Kolarich, D., and Packer, N. H. (2012) Structural analysis of N- and O-glycans released from glycoproteins. *Nat. Protoc.* 7, 1299–1310

79. Everest-Dass, A. V., Abrahams, J. L., Kolarich, D., Packer, N. H., and Campbell, M. P. (2013) Structural Feature Ions for Distinguishing N- and O-Linked Glycan Isomers by LC-ESI-IT MS/MS. *J. Am. Soc. Mass Spectrom.* 24, 895–906

80. Darula, Z., and Medzihradszky, K. F. (2018) Analysis of Mammalian O-Glycopeptides—We Have Made a Good Start, but There is a Long Way to Go. *Mol. Cell. Proteomics* 17, 2–17

81. Hossnerf, K. I. M. L., and Billiar, R. B. (1981) Plasma Clearance and Organ Distribution of Native and Desialylated Rat and Human Transcortin: Species Specificity. *Endocrinology* 108, 1780–1786

82. Avvakumov, G. V., Warmels-Rodenhiser, S., and Hammond, G. L. (1993) Glycosylation of human corticosteroid-binding globulin at aspargine 238 is necessary for steroid binding. *J. Biol. Chem.* 268, 862–866

83. Underhill, C., and Hammond, G. L. (2018) Functional implications of corticosteroid-binding globulin N-glycosylation. *J. Mol. Endocrinol.* 60, 71–84

84. Chan, W. L., Carrell, R. W., Zhou, A., and Read, R. J. (2013) How Changes in Affinity of Corticosteroid-binding Globulin Modulate Free Cortisol Concentration. *J. Clin. Endocrinol. Metab.* 98, 3315–3322

85. Sumer-Bayraktar, Z., Grant, O. C., Venkatakrishnan, V., Woods, R. J., Packer, N. H., and Thaysen-Andersen, M. (2016) Asn347 Glycosylation of Corticosteroid-binding Globulin Fine-

tunes the Host Immune Response by Modulating Proteolysis by *Pseudomonas aeruginosa* and Neutrophil Elstase. *J. Biol. Chem.* 291, 17727–17742

86. Ho, J. T., Lewis, J. G., O'Loughlin, P., Bagley, C. J., Romero, R., Dekker, G. A., and Torpy, D. J. (2007) Reduced maternal corticosteroid-binding globulin and cortisol levels in pre-eclampsia and gamete recipient pregnancies. *Clin. Endocrinol. (Oxf).* 66, 869–877

87. Avvakumov, G. V., and Strel'chyonok, O. A. (1987) Properties and serum levels of pregnancy-associated variant of human transcortin. *Biochem. Biophys. Acta* 925, 11–16

88. Torpy, D. J., Mitchell, E., and Bagley, C. J. (2004) Pregnancy-associated Corticosteroid-binding Globulin: High Resolution Separation of Glycan Isoforms. *Horm. Metab. Res.*, 357–359

89. Avvakumov, G. V., and Strel'chyonok, O. A. (1988) Evidence for the involvement of the transcortin carbohydrate moiety in the glycoprotein interaction with the plasma membrane of human placental syncytiotrophoblast. *Biochim. Biophys. Acta - Biomembr.* 938, 1–6

90. Hoffmann, M., Marx, K., Reichl, U., Wuhrer, M., and Rapp, E. (2016) Site-specific O-Glycosylation Analysis of Human Blood Plasma Proteins. *Mol. Cell. Proteomics* 15, 624–641

91. Zhang, Y., Xie, X., Zhao, X., Tian, F., Lv, J., Ying, W., and Qian, X. (2018) Systems analysis of singly and multiply O -glycosylated peptides in the human serum glycoproteome via EThcD and HCD mass spectrometry. *J. Proteomics* 170, 14–27

92. Darula, Z., Sarnyai, F., and Medzihradszky, K. F. (2016) O-glycosylation sites identified from mucin core-1 type glycopeptides from human serum. *Glycoconj. J.* 33, 435–445

93. Ye, Z., Mao, Y., Clausen, H., and Vakhrushev, S. Y. (2019) Glyco-DIA: a method for quantitative O-glycoproteomics with in silico-boosted glycopeptide libraries. *Nat. Methods* 16, 902–910

94. Yang, W. (2018) Mapping the O-glycoproteome using site-specific extraction of O-linked glycopeptides (EXoO). *Mol. Syst. Biol.* 14, 1–12

95. Domon, B., and Costello, C. E. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J.* 5, 397–409

96. Han, L., and Costello, C. E. (2013) Mass spectrometry of glycans. *Biochem.* 78, 710–720

97. Roepstorff, P., and Fohlman, J. (1984) Letter to the editors. *Biol. Mass Spectrom.* 11, 601–601

98. Hoffmann, M., Pioch, M., Pralow, A., Hennig, R., Kottler, R., Reichl, U., and Rapp, E. (2018) The Fine Art of Destruction: A Guide to In-Depth Glycoproteomic Analyses-Exploiting the Diagnostic Potential of Fragment Ions. *Proteomics* 18, 1800282

99. Nishikaze, T., Kawabata, S., and Tanaka, K. (2014) In-Depth Structural Characterization of N -Linked Glycopeptides Using Complete Derivatization for Carboxyl Groups Followed by Positive- and Negative-Ion Tandem Mass Spectrometry. *Anal. Chem.* 86, 5360–5369

100. Nishikaze, T. (2017) Sensitive and Structure-Informative N-Glycosylation Analysis by MALDI-MS; Ionization, Fragmentation, and Derivatization. *Mass Spectrom.* 6, A0060–A0060

101. Reiding, K. R., Bondt, A., Franc, V., and Heck, A. J. R. (2018) The benefits of hybrid fragmentation methods for glycoproteomics. *TrAC Trends Anal. Chem.* 108, 260–268

102. Thaysen-Andersen, M., Packer, N. H., and Schulz, B. L. (2016) Maturing Glycoproteomics Technologies Provide Unique Structural Insights into the *N* -glycoproteome and Its Regulation in Health and Disease. *Mol. Cell. Proteomics* 15, 1773–1790

103. Hu, H., Khatri, K., and Zaia, J. (2017) Algorithms and design strategies towards automated glycoproteomics analysis. *Mass Spectrom. Rev.* 36, 475–498

104. Yu, Q., Wang, B., Chen, Z., Urabe, G., Glover, M. S., Shi, X., Guo, L.-W., Kent, K. C., and Li, L. (2017) Electron-Transfer/Higher-Energy Collision Dissociation (EThcD)-Enabled Intact Glycopeptide/Glycoproteome Characterization. *J. Am. Soc. Mass Spectrom.* 28, 1751–1764

105. Pap, A., Klement, E., Hunyadi-Gulyas, E., Darula, Z., and Medzihradszky, K. F. (2018) Status Report on the High-Throughput Characterization of Complex Intact O-Glycopeptide Mixtures. *J. Am. Soc. Mass Spectrom.* 29, 1210–1220

106. Hinneburg, H., Chatterjee, S., Schirmeister, F., Nguyen-Khuong, T., Packer, N. H., Rapp, E., and Thaysen-Andersen, M. (2019) Post-Column Make-Up Flow (PCMF) Enhances the Performance of Capillary-Flow PGC-LC-MS/MS-Based Glycomics. *Anal. Chem.* 91, 4559–4567

107. Ashwood, C., Lin, C.-H., Thaysen-Andersen, M., and Packer, N. H. (2018) Discrimination of Isomers of Released N- and O-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS. *J. Am. Soc. Mass Spectrom.* 29, 1194–1209

108. Cooper, C. A., Gasteiger, E., and Packer, N. H. (2001) GlycoMod - A software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics* 1, 340–349

109. Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., and Haslam, S. M. (2008) GlycoWorkbench: A Tool for the Computer-Assisted Annotation of Mass Spectra of Glycans. *J. Proteome Res.* 7, 1650–1659

110. Bern, M., Kil, Y. J., and Becker, C. (2012) Byonic: Advanced peptide and protein identification software. *Curr. Protoc. Bioinforma.*, 1–14

111. Peri, S., Steen, H., and Pandey, A. (2001) GPMAW – a software tool for analyzing proteins and peptides. *Trends Biochem. Sci.* 26, 687–689

112. Schilling, B., Rardin, M. J., MacLean, B. X., Zawadzka, A. M., Frewen, B. E., Cusack, M. P., Sorensen, D. J., Bereman, M. S., Jing, E., Wu, C. C., Verdin, E., Kahn, C. R., MacCoss, M. J., and Gibson, B. W. (2012) Platform-independent and Label-free Quantitation of Proteomic Data Using MS1 Extracted Ion Chromatograms in Skyline. *Mol. Cell. Proteomics* 11, 202–214

113. Luria, S. E., and Burrous, J. W. (1957) Hybridization between Escherichia coli and Shigella. *J. Bacteriol.* 74, 461–76

114. Kolarich, D., Jensen, P. H., Altmann, F., and Packer, N. H. (2012) Determination of site-specific glycan heterogeneity on glycoproteins. *Nat. Protoc.* 7, 1285–1298

115. Schrödinger, L. (2015) The PyMol Molecular Graphics System, Versión 1.8. *Thomas Hold.*

116. Rebecchi, K. R., Wenke, J. L., Go, E. P., and Desaire, H. (2009) Label-free quantitation: A new glycoproteomics approach. *J. Am. Soc. Mass Spectrom.* 20, 1048–1059

117. Stavenhagen, K., Hinneburg, H., Thaysen-Andersen, M., Hartmann, L., Varón Silva, D., Fuchser, J., Kaspar, S., Rapp, E., Seeberger, P. H., and Kolarich, D. (2013) Quantitative mapping of glycoprotein micro-heterogeneity and macro-heterogeneity: an evaluation of mass spectrometry signal strengths using synthetic peptides and glycopeptides. *J. Mass Spectrom.* 48, 627–639

118. Russell, W. C., Graham, F. L., Smiley, J., and Nairn, R. (1977) Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *J. Gen. Virol.* 36, 59–72

119. Dumont, J., Euwart, D., Mei, B., Estes, S., and Kshirsagar, R. (2016) Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Crit. Rev. Biotechnol.* 36, 1110–1122

120. Goh, J. B., and Ng, S. K. (2018) Impact of host cell line choice on glycan profile. *Crit. Rev. Biotechnol.* 38, 851–867

121. Böhm, E., Seyfried, B. K., Dockal, M., Graninger, M., Hasslacher, M., Neurath, M., Konetschny, C., Matthiessen, P., Mitterer, A., and Scheiflinger, F. (2015) Differences in N-glycosylation of recombinant human coagulation factor VII derived from BHK, CHO, and HEK293 cells. *BMC Biotechnol.* 15, 87

122. Yang, X., Tao, S., Orlando, R., Brockhausen, I., and Kan, F. W. K. (2012) Structures and biosynthesis of the N- and O-glycans of recombinant human oviduct-specific glycoprotein expressed in human embryonic kidney cells. *Carbohydr. Res.* 358, 47–55

123. Thaysen-Andersen, M., Chertova, E., Bergamaschi, C., Moh, E. S. X., Chertov, O., Roser, J., Sowder, R., Bear, J., Lifson, J., Packer, N. H., Felber, B. K., and Pavlakis, G. N. (2016) Recombinant human heterodimeric IL-15 complex displays extensive and reproducible N- and O-linked glycosylation. *Glycoconj. J.* 33, 417–433

124. Tykvart, J., Šácha, P., Bařinka, C., Knedlík, T., Starková, J., Lubkowski, J., and Konvalinka, J. (2012) Efficient and versatile one-step affinity purification of in vivo biotinylated proteins: Expression, characterization and structure analysis of recombinant human glutamate carboxypeptidase II. *Protein Expr. Purif.* 82, 106–115

125. Hunter, M., Yuan, P., Vavilala, D., and Fox, M. (2019) Optimization of Protein Expression in Mammalian Cells. *Curr. Protoc. Protein Sci.* 95, e77

126. Costa, A. R., Rodrigues, M. E., Henriques, M., Oliveira, R., and Azeredo, J. (2014) Glycosylation: impact, control and improvement during therapeutic protein production. *Crit. Rev. Biotechnol.* 34, 281–299

127. Vakhrushev, S. Y., Steentoft, C., Vester-Christensen, M. B., Bennett, E. P., Clausen, H., and Levery, S. B. (2013) Enhanced Mass Spectrometric Mapping of the Human GalNAc-type O -Glycoproteome with SimpleCells. *Mol. Cell. Proteomics* 12, 932–944

128. Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B.,

Schjoldager, K. T.-B. G., Lavrsen, K., Dabelsteen, S., Pedersen, N. B., Marcos-Silva, L., Gupta, R., Paul Bennett, E., Mandel, U., Brunak, S., Wandall, H. H., Levery, S. B., and Clausen, H. (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* 32, 1478–1488

129. Fagerberg, L., Hallström, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., Tahmasebpoor, S., Danielsson, A., Edlund, K., Asplund, A., Sjöstedt, E., Lundberg, E., Szigyarto, C. A.-K., Skogs, M., Takanen, J. O., Berling, H., Tegel, H., Mulder, J., Nilsson, P., Schwenk, J. M., Lindskog, C., Danielsson, F., Mardinoglu, A., Sivertsson, Å., von Feilitzen, K., Forsberg, M., Zwahlen, M., Olsson, I., Navani, S., Huss, M., Nielsen, J., Ponten, F., and Uhlén, M. (2014) Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. *Mol. Cell. Proteomics* 13, 397–406

130. Rinn, J. L., Rozowsky, J. S., Laurenzi, I. J., Petersen, P. H., Zou, K., Zhong, W., Gerstein, M., and Snyder, M. (2004) Major Molecular Differences between Mammalian Sexes Are Involved in Drug Metabolism and Renal Function. *Dev. Cell* 6, 791–800

131. Shah, K., McCormack, C. E., and Bradbury, N. A. (2014) Do you know the sex of your cells? *Am. J. Physiol. Physiol.* 306, C3–C18

132. Lin, Y.-C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., Chen, J., Speleman, F., Lambrechts, D., Van de Peer, Y., Tavernier, J., and Callewaert, N. (2014) Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nat. Commun.* 5, 4767

133. Goth, C. K., Vakhrushev, S. Y., Joshi, H. J., Clausen, H., and Schjoldager, K. T. (2018) Fine-Tuning Limited Proteolysis: A Major Role for Regulated Site-Specific O -Glycosylation. *Trends Biochem. Sci.* 43, 269–284

134. Narimatsu, Y., Joshi, H. J., Schjoldager, K. T., Hintze, J., Halim, A., Steentoft, C., Nason, R., Mandel, U., Bennett, E. P., Clausen, H., and Vakhrushev, S. Y. (2019) Exploring Regulation of Protein O-Glycosylation in Isogenic Human HEK293 Cells by Differential O-Glycoproteomics. *Mol. Cell. Proteomics* 18, 1396–1409

135. King, S. L., Joshi, H. J., Schjoldager, K. T., Halim, A., Madsen, T. D., Dziegiel, M. H., Woetmann, A., Vakhrushev, S. Y., and Wandall, H. H. (2017) Characterizing the O-glycosylation landscape of human plasma, platelets, and endothelial cells. *Blood Adv.* 1, 429–442

136. King, S. L., Goth, C. K., Eckhard, U., Joshi, H. J., Haue, A. D., Vakhrushev, S. Y., Schjoldager, K. T., Overall, C. M., and Wandall, H. H. (2018) TAILS N-terminomics and proteomics reveal complex regulation of proteolytic cleavage by O-glycosylation. *J. Biol. Chem.* 293, 7629–7644

137. Hansen, G., Gielen-Haertwig, H., Reinemer, P., Schomburg, D., Harrenga, A., and Niefind, K. (2011) Unexpected Active-Site Flexibility in the Structure of Human Neutrophil Elastase in Complex with a New Dihydropyrimidone Inhibitor. *J. Mol. Biol.* 409, 681–691

138. Meyer, E. J., Nenke, M. A., Rankin, W., Lewis, J. G., Konings, E., Slager, M., Jansen, T. C., Bakker, J., Hofland, J., Feelders, R. A., and Torpy, D. J. (2019) Total and high-affinity corticosteroid-binding globulin depletion in septic shock is associated with mortality. *Clin. Endocrinol. (Oxf).* 90, 232–240

139. Polat, G., Ugan, R. A., Cadirci, E., and Halici, Z. (2017) Sepsis and Septic Shock: Current Treatment Strategies and New Approaches. *Eurasian J. Med.* 49, 53–58

140. Ashwell, G., and Harford, J. (1982) Carbohydrate-Specific Receptors of the Liver. *Annu. Rev. Biochem.* 51, 531–554

141. Lalonde, M.-E., and Durocher, Y. (2017) Therapeutic glycoprotein production in mammalian cells. *J. Biotechnol.* 251, 128–140

142. Mulligan, M. S., Warner, R. L., Till, G. O., Ward, P. A., Crouch, L. D., Rittershaus, C. W., Thomas, L. J., Ryan, U. S., and Foreman, K. E. (1999) Endothelial targeting and enhanced antiinflammatory effects of complement inhibitors possessing sialyl Lewis(x) moieties. *J. Immunol.* 162, 4952–4959

143. Novokmet, M., Lukić, E., Vučković, F., –Durić, Ž., Keser, T., Rajšl, K., Remondini, D., Castellani, G., Gašparović, H., Gornik, O., and Lauc, G. (2015) Changes in IgG and total plasma protein glycomes in acute systemic inflammation. *Sci. Rep.* 4, 4347

144. Gornik, O., Royle, L., Harvey, D. J., Radcliffe, C. M., Saldova, R., Dwek, R. A., Rudd, P., and Lauc, G. (2007) Changes of Serum Glycans During Sepsis and Acute Pancreatitis. *Glycobiology* 17, 1321–1332

145. Sleat, D. E., Wang, Y., Sohar, I., Lackland, H., Li, Y., Li, H., Zheng, H., and Lobel, P. (2006) Identification and Validation of Mannose 6-Phosphate Glycoproteins in Human Plasma Reveal a Wide Range of Lysosomal and Non-lysosomal Proteins. *Mol. Cell. Proteomics* 5, 1942–1956

146. Hindmarsh, E. J., Staykova, M. A., Willenborg, D. O., and Parish, C. R. (2001) Cell surface expression of the 300 kDa mannose-6-phosphate receptor by activated T lymphocytes. *Immunol. Cell Biol.* 79, 436–443

Appendices

Appendix 1. PGC-LC-MS/MS *N*- and *O*-glycome profiling of His- and biotin-CBG.

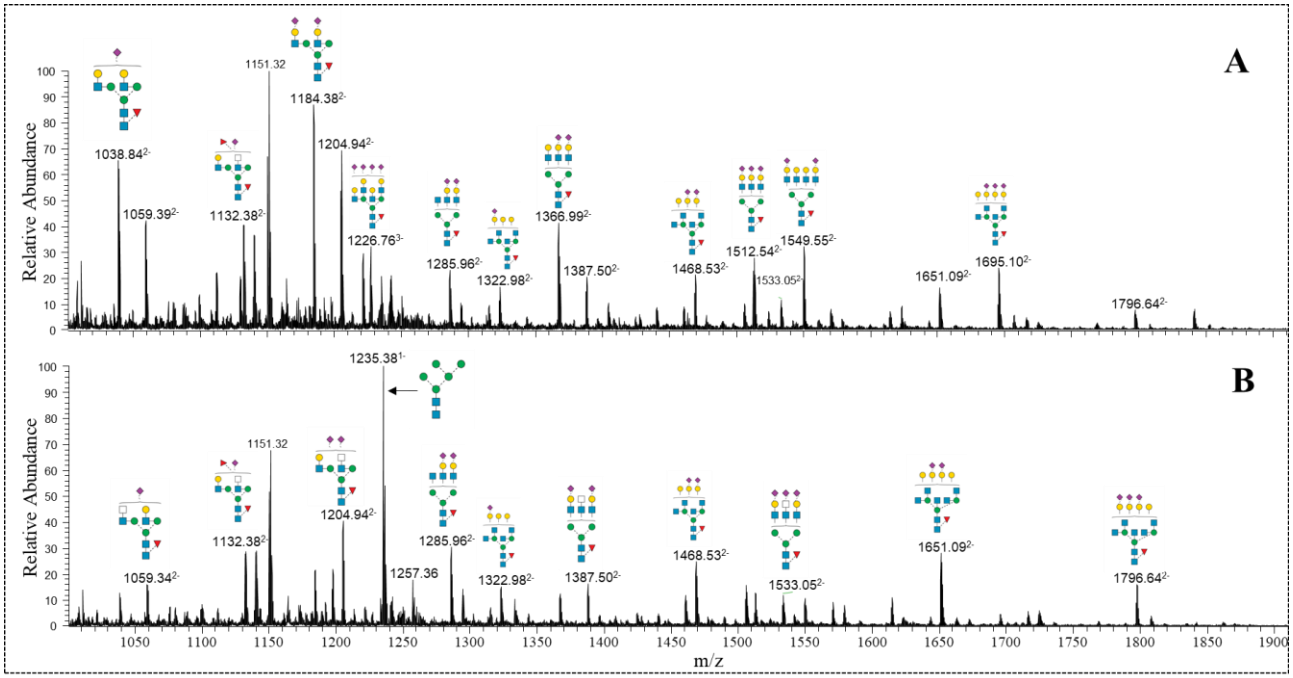


Figure S1. Summed MS profiles obtained using PGC-LC-MS/MS *N*-glycome profiling of His- (**A**) and biotin-CBG (**B**). Several *N*-glycan structures corresponding to the observed *m/z* values as supported by CID-MS/MS data and PGC-LC retention time patterns are shown.

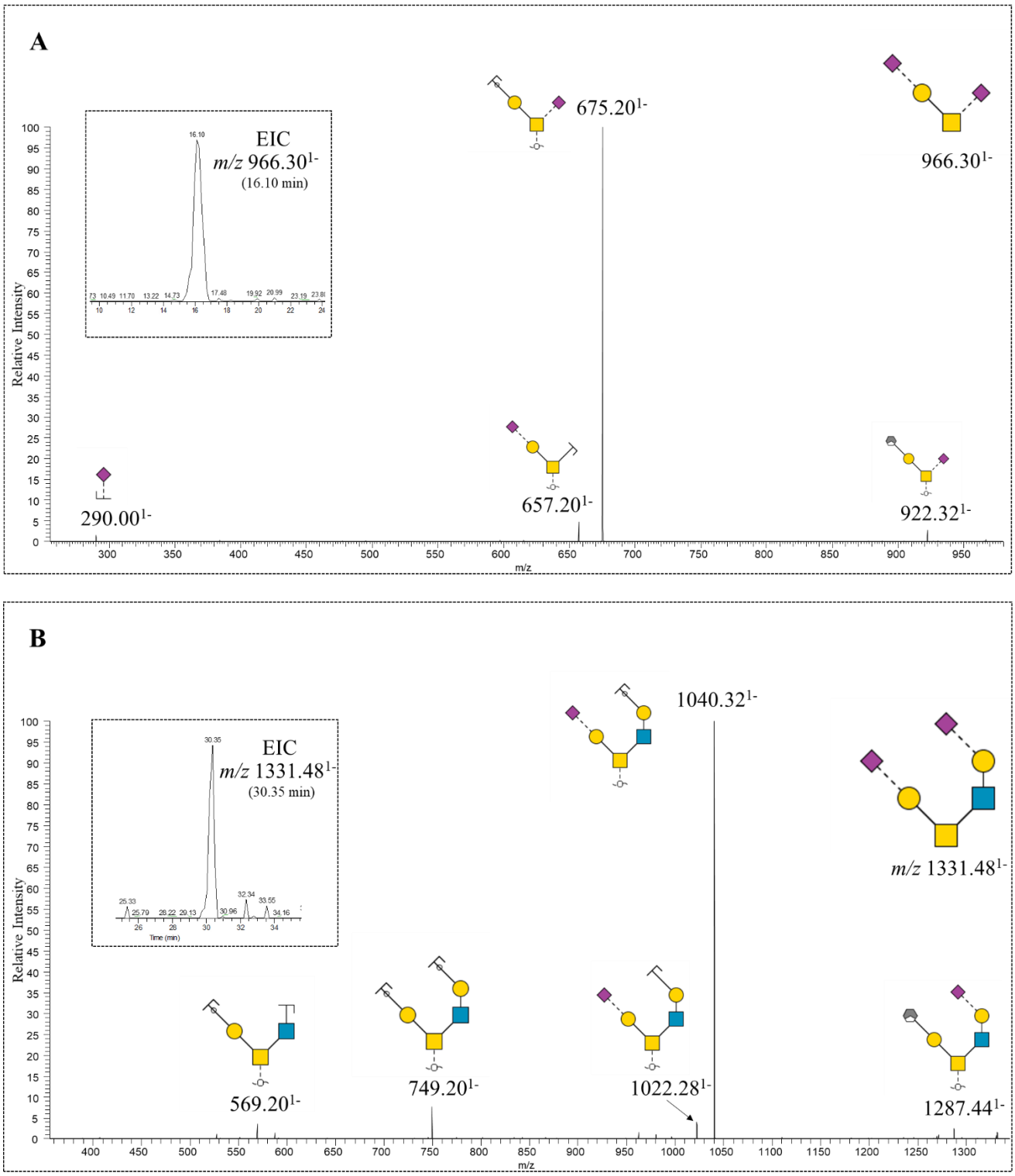


Figure S2. Examples of annotated CID-MS/MS (-) spectral data corresponding to core 1- (**A**) and core 2- (**B**) type *O*-glycans of His-CBG (data for biotin-CBG not shown). Insert: extracted ion chromatograms (EICs) of the corresponding precursor ions.

Appendix 2. Examples of Byonic search results for rhCBG.

Table S1. A partial list of intact His-CBG (glyco)peptides identified by LC-MS/MS followed by Byonic database search.

Pos.	Sequence	Mods (variable)	Glycans	PEP 2D	z	Obs. m/z	Calc. m/z
1	-.MDPNAAYVNMSNHHR.G			5.20E-09	2	878.8875	878.8857
1	-.MDPNAAYVNMSNHHR.G			3.10E-10	4	439.9485	439.9465
1	-.MDPNAAYVNMSNHHR.G			7.80E-12	3	586.2606	586.2596
1	-.MDPNAAYVN[+1216.42286]MSNHHR.G	N9(NGlycan / 1216.4229)	HexNAc(2)Hex(5)	4.20E-09	4	744.0518	744.0522
1	-.MDPNAAYVN[+1216.42286]MSNHHR.G	N9(NGlycan / 1216.4229)	HexNAc(2)Hex(5)	0.00016	3	991.7343	991.7339
1	-.MDPNAAYVN[+1241.45450]MSNHHR.G	N9(NGlycan / 1241.4545)	HexNAc(3)Hex(3)Fuc(1)	1.10E-06	4	750.3106	750.3101
1	-.MDPNAAYVN[+1241.45450]MSNHHR.G	N9(NGlycan / 1241.4545)	HexNAc(3)Hex(3)Fuc(1)	2.10E-06	3	1000.0746	1000.0778
1	-.MDPNAAYVN[+1606.58669]MSNHHR.G	N9(NGlycan / 1606.5867)	HexNAc(4)Hex(4)Fuc(1)	4.20E-09	3	1121.7897	1121.7885
1	-.MDPNAAYVN[+1647.61324]MSNHHR.G	N9(NGlycan / 1647.6132)	HexNAc(5)Hex(3)Fuc(1)	1.50E-09	4	851.8499	851.8498
1	-.MDPNAAYVN[+1850.69261]MSNHHR.G	N9(NGlycan / 1850.6926)	HexNAc(6)Hex(3)Fuc(1)	2.70E-07	4	902.6190	902.6197
1	-.MDPNAAYVN[+1897.68211]MSNHHR.G	N9(NGlycan / 1897.6821)	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	0.00034	4	914.3664	914.3670
1	-.MDPNAAYVN[+1938.70866]MSNHHR.G	N9(NGlycan / 1938.7087)	HexNAc(5)Hex(3)Fuc(1)NeuAc(1)	8.00E-05	4	924.6236	924.6237
1	-.MDPNAAYVN[+1955.72397]MSNHHR.G	N9(NGlycan / 1955.724)	HexNAc(5)Hex(4)Fuc(2)	2.40E-07	4	928.8761	928.8775
1	-.MDPNAAYVN[+1971.71889]MSNHHR.G	N9(NGlycan / 1971.7189)	HexNAc(5)Hex(5)Fuc(1)	2.10E-07	4	932.8752	932.8762
1	-.MDPNAAYVN[+1996.75052]MSNHHR.G	N9(NGlycan / 1996.7505)	HexNAc(6)Hex(3)Fuc(2)	5.30E-07	4	939.1339	939.1341
1	-.MDPNAAYVN[+2053.77199]MSNHHR.G	N9(NGlycan / 2053.772)	HexNAc(7)Hex(3)Fuc(1)	2.30E-09	4	953.3895	953.3895
1	-.MDPNAAYVN[+2059.73493]MSNHHR.G	N9(NGlycan / 2059.7349)	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	4.20E-08	4	954.8800	954.8802
1	-.MDPNAAYVN[+2100.76148]MSNHHR.G	N9(NGlycan / 2100.7615)	HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	2.30E-07	3	1286.5137	1286.5134
1	-.MDPNAAYVN[+2141.78803]MSNHHR.G	N9(NGlycan / 2141.788)	HexNAc(6)Hex(3)Fuc(1)NeuAc(1)	1.10E-06	4	975.3943	975.3935
1	-.MDPNAAYVN[+2350.83035]MSNHHR.G	N9(NGlycan / 2350.8304)	HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	7.10E-08	3	1369.8719	1369.8697
1	-.MDPNAAYVN[+2391.85690]MSNHHR.G	N9(NGlycan / 2391.8569)	HexNAc(5)Hex(4)Fuc(1)NeuAc(2)	4.90E-09	4	1037.9112	1037.9107
1	-.MDPNAAYVN[+2424.86713]MSNHHR.G	N9(NGlycan / 2424.8671)	HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	1.10E-08	4	1046.1626	1046.1633
1	-.MDPNAAYVN[+2465.89368]MSNHHR.G	N9(NGlycan / 2465.8937)	HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	5.60E-09	4	1056.4214	1056.4199
1	-.MDPNAAYVN[+2919.04192]MSNHHR.G	N9(NGlycan / 2919.0419)	HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	2.20E-08	4	1169.7104	1169.7070
1	-.MDPNAAYVN[+3007.05796]MSNHHR.G	N9(NGlycan / 3007.058)	HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	2.80E-09	4	1191.7126	1191.7110
1	-.MDPNAAYVN[+3048.08451]MSNHHR.G	N9(NGlycan / 3048.0845)	HexNAc(6)Hex(5)Fuc(1)NeuAc(3)	1.80E-06	4	1201.9687	1201.9676
1	-.MDPNAAYVN[+3081.09474]MSNHHR.G	N9(NGlycan / 3081.0947)	HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	1.60E-06	3	1613.2955	1613.2912
1	-.MDPNAAYVN[+3284.17411]MSNHHR.G	N9(NGlycan / 3284.1741)	HexNAc(7)Hex(7)Fuc(1)NeuAc(2)	9.50E-08	4	1260.9921	1260.9900
1	-.MDPNAAYVN[+3372.19016]MSNHHR.G	N9(NGlycan / 3372.1902)	HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	5.10E-08	4	1282.9941	1282.9940
1	-.MDPNAAYVN[+3663.28558]MSNHHR.G	N9(NGlycan / 3663.2856)	HexNAc(6)Hex(7)Fuc(1)NeuAc(4)	3.00E-09	4	1355.7678	1355.7679

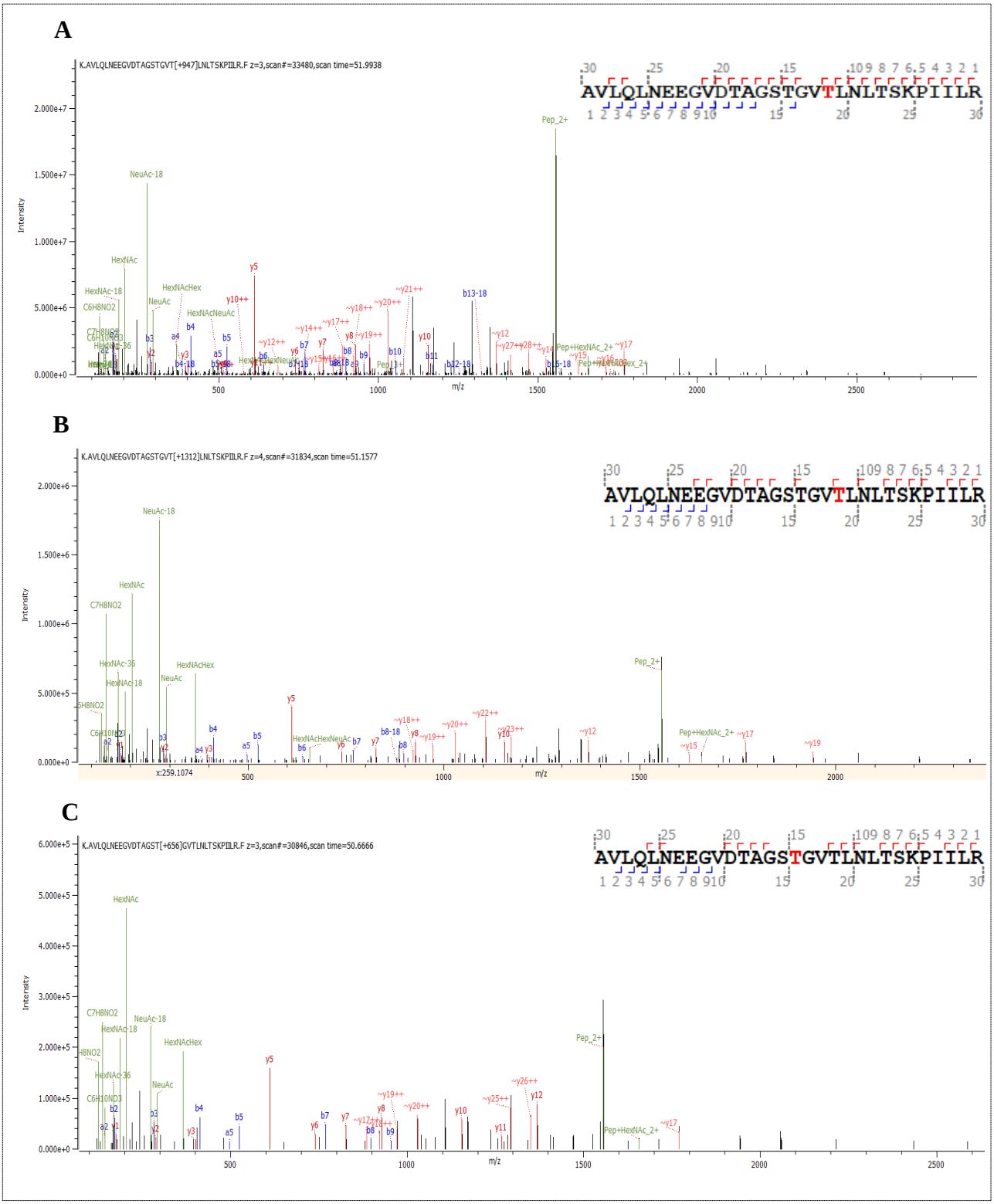
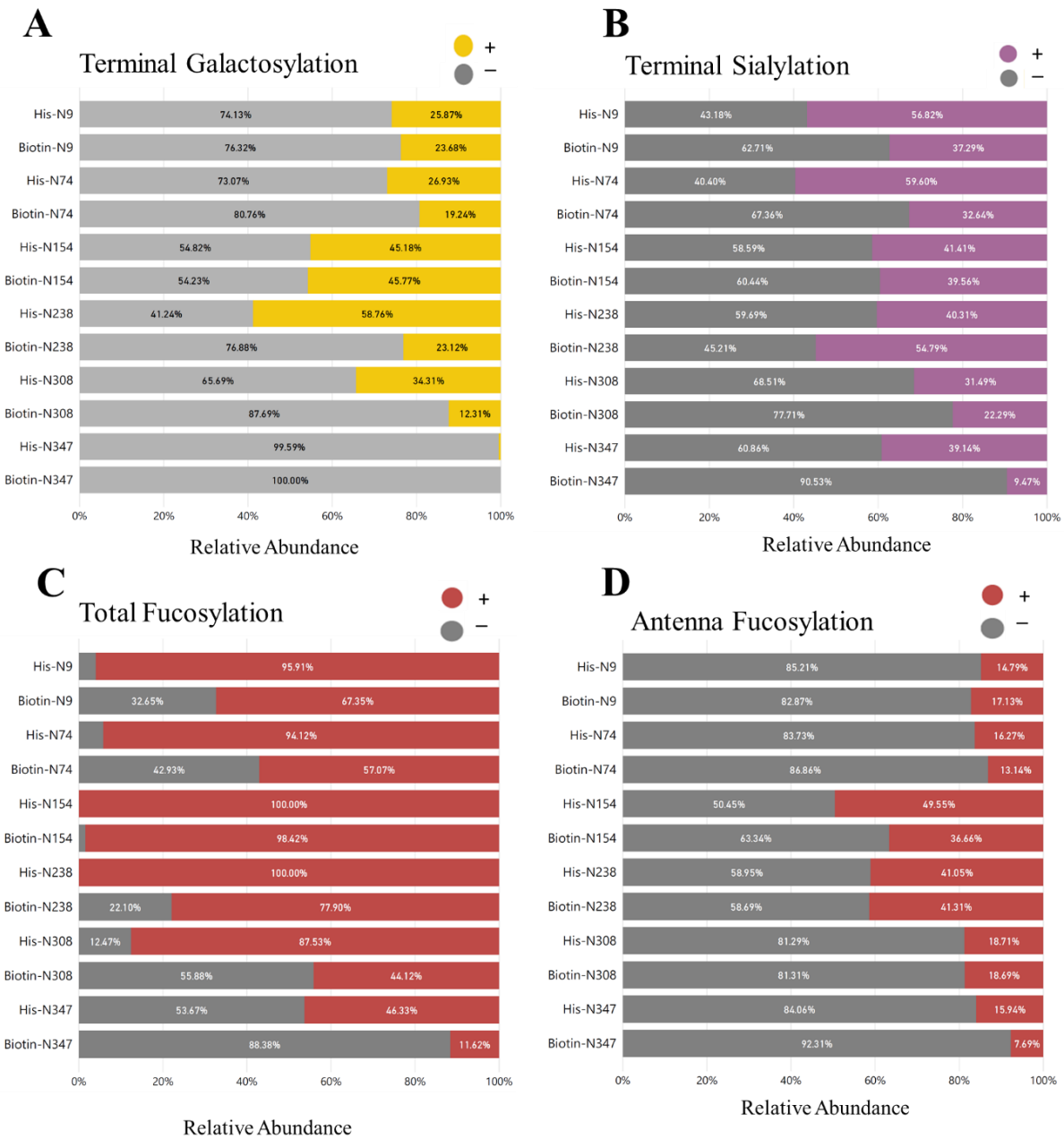


Figure S3. An example of Byonic-assisted annotation of HCD-MS/MS spectra of His-CBG RCL *O*-glycopeptides conjugated with (**A**) $HexNAc_1Hex_1NeuAc_2$ (*m/z* = 1353.01, charge 3+, PEP 2D score = 6.3E-24); (**B**) $HexNAc_2Hex_2NeuAc_2$ (*m/z* = 1106.29, charge 4+, PEP 2D score = 3.2E-17); (**C**) $HexNAc_1Hex_1NeuAc_1$ (*m/z* = 1255.98, charge 3+, PEP 2D score = 2.2E-14). Peptide backbone b- and y-ions are shown in blue and red, respectively. Oxonium ions and Y-ions are depicted in green.



Appendix 3.

Figure S4. Site-specific overview of key *N*-glycosylation features that may impact the pharmacokinetics of rhCBG (His- and biotin-CBG) including the level of terminal **A.** galactosylation **B.** sialylation (α2,3- and α2,6-) **C.** total fucosylation (core and antennae) **D.** antennae fucosylation and **E.** mannose-6-phosphorylation (M6P).

Appendix 4. Poster at 18th HUPO World Congress, Adelaide, SA

Unexpected RCL *O*-Glycosylation of Human Corticosteroid-Binding Globulin Impacts Proteolysis

Anastasia Chernykh[1], Krishnatej Nishtala[1], Sayantani Chatterjee[1], Harry C. Tjondro[1], Hannes Hinneburg[1], Zeynep Sumer-Bayraktar[2], Emily J. Meyer[3,4], David J. Torpy [3,4], Morten Thaysen-Andersen[1*]

[1]Department of Molecular Sciences, Macquarie University, Sydney, Australia, [2]Charles Perkins Centre, School of Life and Environmental Sciences, The University of Sydney, Sydney, Australia, [3]Endocrine and Metabolic Unit, Royal Adelaide Hospital, Adelaide, Australia, [4]Discipline of Medicine, University of Adelaide, Adelaide, Australia

Background and Aim

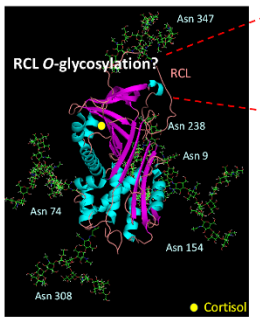
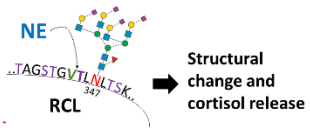


Figure 1. Homology model of human CBG based on human TBG (2CEO). *N*-glycans are mapped on the PDB.



Human corticosteroid-binding globulin (CBG) carries plasma cortisol, **Fig. 1**. We have shown a regulatory role of Asn347-glycosylation in the neutrophil elastase (NE) based cleavage of the reactive centre loop (RCL), which ultimately controls cortisol delivery [1,2]. Recent large-scale studies have indicated the presence of RCL *O*-glycosylation [3, 4].

Here, we set out to confirm the presence and investigate the function of RCL *O*-glycosylation of human CBG.

Methods

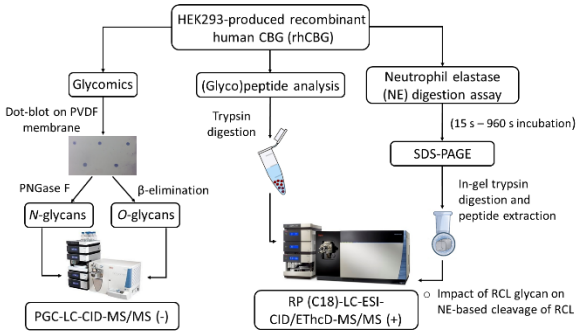


Figure 2. Workflow.

Results and Discussion

- Site-specific RCL *N*- (top) and *O*- (bottom) glycoprofiling of rhCBG, **Fig. 3**.

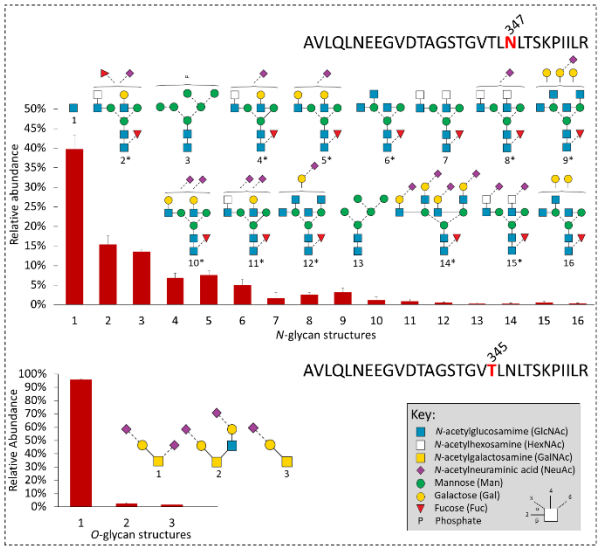


Figure 3. RCL *N*- and *O*-glycosylation of rhCBG. *Less abundant isomers were also detected.

- RCL *O*-glycosylation site identification (Thr345) of rhCBG, **Fig. 4**.

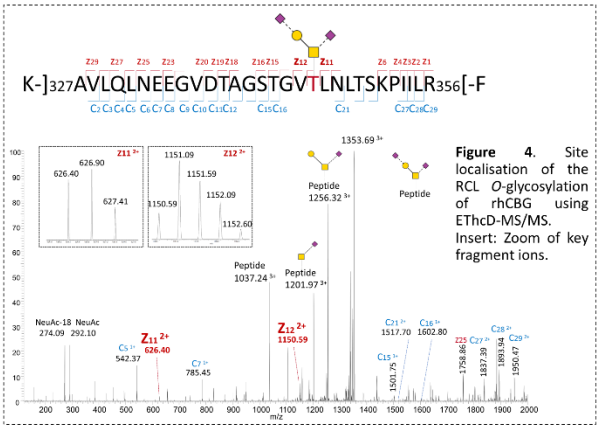


Figure 4. Site localisation of the RCL *O*-glycosylation of rhCBG using EThcD-MS/MS. Insert: Zoom of key fragment ions.

- Thr345 *O*-glycosylation strongly inhibits NE-mediated RCL cleavage, **Fig. 5**.

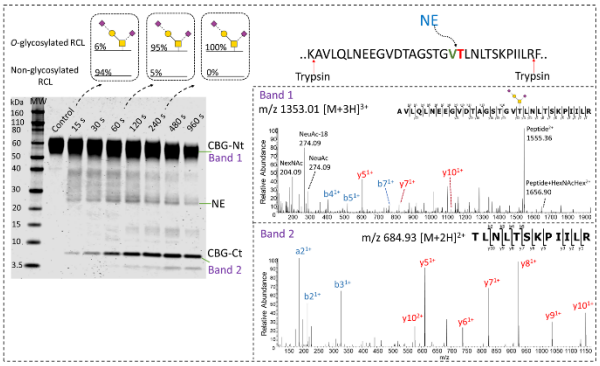
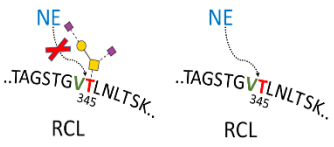


Figure 5. SDS-PAGE illustrating the CBG products formed from NE-based cleavage over time. The resulting CBG N- (CBG-Nt, band 1) and C-terminal (CBG-Ct, band 2) fragments were analysed using LC-MS/MS.

Conclusion and Future Perspectives

- First study to confirm RCL *O*-glycosylation of human CBG.
- Disialyl core 1 *O*-glycan occupies Thr345 next to the NE cleavage site.
- RCL *O*-glycosylation strongly inhibits NE-based RCL cleavage and may together with Asn347-glycans control cortisol release.
- Roles of RCL *O*-glycosylation and the interplay with Asn347-glycans of native human CBG is being studied.



MACQUARIE University

Our Group

Acknowledgements: AC and SC were supported by Macquarie University Scholarships. MTA was supported by a Macquarie University Research Seeding grant.

References: [1] Sumer-Bayraktar, Z., et al., *N*-glycans Modulate the Function of Human Corticosteroid-Binding Globulin. *Mol. Cell. Proteomics* 2011, 10, 1–14. [2] Sumer-Bayraktar, Z., et al., Asn347 Glycosylation of Corticosteroid-binding Globulin Fine-tunes the Host Immune Response by Modulating Proteolysis by *Pseudomonas aeruginosa* and Neutrophil Elstase. *J. Biol. Chem.* 2016, 291, 17727–17742. [3] Yang, W., et al., Mapping the *O*-glycoproteome using site-specific extraction of *O*-linked glycopeptides (EXoO). *Mol. Syst. Biol.* 2018, 14, 1–12. [4] Ye, Z., et al., Glyco-DIA: a method for quantitative *O*-glycoproteomics with in silico-boosted glycopeptide libraries. *Nat. Methods* 2019, 16, 902–910.

Appendix 5. Manuscript under review (my contributions in yellow)

Steroid binding to high- and low-affinity corticosteroid-binding globulin; effect of pyrexia and acidosis on affinity

Emily J Meyer [1 2], David J Torpy [1 2], **Anastasia Chernykh** [3], Morten Thaysen-Andersen [3], John Lewis[4], Wayne Rankin [1 2 5]*, Steven Polyak [6]*.

[1]Endocrine and Metabolic Unit, Royal Adelaide Hospital, Adelaide, SA, Australia
[2]Discipline of Medicine, University of Adelaide, Adelaide, South Australia, Australia
[3]Department of Molecular Sciences, Macquarie University, Sydney, New South Wales, Australia
[4]Steroid & Immunobiochemistry Laboratory, Canterbury Health Laboratories, Christchurch, New Zealand
[5]Chemical Pathology Directorate, SA Pathology, Adelaide, South Australia, Australia
[6]School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, South Australia, Australia
***** Equal senior co-authors

Short title: Steroid binding profile of high- and low-affinity corticosteroid-binding globulin.

Abstract

Context
Corticosteroid-binding globulin (CBG) is the principal transporter of cortisol and other steroid hormones. High-affinity CBG (haCBG) changes conformation to low-affinity CBG (laCBG) following proteolysis of the reactive centre loop (RCL) by neutrophil elastase (NE), which liberates anti-inflammatory cortisol at inflammatory sites. Pyrexia also reduces CBG:cortisol binding affinity.

Objective
To measure the equilibrium binding constants of a panel of steroids to haCBG and laCBG over temperature and pH ranges mimicking the pathophysiological conditions of septic shock.

Design
Surface plasmon resonance was used to determine the binding profiles of 19 steroid ligands to haCBG and laCBG at 25°C, 37°C and 39°C and pH 7.4 and 7.0. The RCL-recognising 9G12 antibody was used to assess the RCL integrity and availability across conditions.

Results
A 4-8-fold reduction in affinity for cortisol, cortisone, corticosterone, 11-deoxycortisol, progesterone, 17-hydroxyprogesterone and prednisolone occurred with haCBG-to-laCBG conversion. Binding affinity consistently decreased at higher temperatures and acidic pH for both haCBG and laCBG. 9G12 antibody RCL binding was preserved for haCBG across temperatures.

Conclusions
These studies reveal that steroid binding to CBG is selective and in all cases reduced by NE cleavage of CBG. Cortisol:progesterone CBG binding competition may be important in the developing foetus/placenta. Moreover, reduced CBG:cortisol binding affinity by temperature occurs with an intact RCL, suggesting a non-RCL mechanism for this phenomenon. Synergy of NE cleavage and pyrexia/acidosis may serve for local inflammatory site cortisol delivery and increase free cortisol. Use of RCL antibody binding affinity appears to be a valid measure of haCBG for clinical studies.

Tables

Table S2. Equilibrium binding constants (K_D, nM) of uncleaved (haCBG) and cleaved (laCBG) CBG determined at a constant condition of 25°C and pH 7.4. All data points are represented as the mean values and their variance indicated in brackets as SEM as determined from three technical replicates.

Steroid ligand	**haCBG** (K_D, nM)	**laCBG** (K_D, nM)	**ΔK_D** (fold change)
Hydrocortisone (cortisol)	42.0 (3.8)	217.5 (8.7)	5.2
Corticosterone	45.2 (6.4)	170.8 (6.6)	3.8
Prednisolone	77.9 (14.7)	458 (48.5)	5.9
11-deoxycortisol	85.3 (25.5)	348.7 (42.7)	4.1
17-OH progesterone	118.4 (10.2)	516.4 (53.9)	4.4
Progesterone	318.2 (67.6)	1570.3 (214.9)	4.9
Cortisone	400.5 (97.8)	2957.5 (243.5)	7.4
Testosterone	609.5 (22.7)	3102.3 (564.5)	5.1
Aldosterone	1008.8 (236.3)	4634.5 (846.4)	4.6
17-OH pregnenolone	>6000	>6000	
B-estradiol	>6000	>6000	
Dexamethaxone	>6000	>6000	
Dehydroepiandrosterone (DHEA)	>6000	>6000	
Estrone	>6000	>6000	
5a-Androstan-17B-ol-3-one	>6000	>6000	
Adrenosterone	>6000	>6000	
Dehydroisoandrosterone 3-sulphate	>6000	>6000	
Pregnenolone	>6000	>6000	
Allopregnanolone	>6000	>6000	

Table S3. Equilibrium binding constants (K_D, nM) of uncleaved (haCBG) and cleaved (laCBG) CBG at different temperatures and pH. All data points are represented as mean values and their variance indicated in brackets as SEM as determined from three technical replicates.

Steroid ligand	**haCBG** (K_D, nM)				**laCBG** (K_D, nM)			
	25°C, pH 7.4	37°C, pH 7.4	39°C, pH 7.4	39°C, pH 7.0	25°C, pH 7.4	37°C, pH 7.4	39°C, pH 7.4	39°C, pH 7.0
Hydrocortisone (cortisol)	42.0 (3.8)	195.2 (20.1)	330.2 (29.5)	418.3 (14.3)	217.5 (8.7)	1270 (183)	1232.4 (200.9)	1537.5 (125.5)
Corticosterone	45.2 (6.4)	181.3 (2.3)	450.3 (100.5)	552.2 (41.1)	170.8 (6.6)	897.95 (29.3)	1256 (116.7)	1812 (242)
11-deoxycortisol	85.3 (25.5)	249.8 (22.6)	485.9 (55.8)	557.2 (64.7)	348.7 (42.7)	1931 (201)	2238.7 (596.9)	2963 (292)
Prednisolone	77.9 (14.7)	329.8 (53.2)	647.7 (101.1)	759.4 (27.8)	458 (48.5)	2265.5 (345.5)	2688.7 (477.6)	4279.5 (940.5)
17-OH progesterone	118.4 (10.2	889.1 (373.9)	846.7 (110.2)	1567.5 (116.5)	516.4 (53.9)	2837.5 (257.5)	1595.5 (497.5)	5134
Cortisone	400.5 (97.8)	1338 (281)	1632.3 (290.5)	2978 (242)	2957.5 (243.5)	5242	>6000	>6000
Progesterone	318.2 (67.6)	1680.5 (101.5)	2838.5 (224.5)	3587 (510)	1570.3 (214.9)	6143 (1266)	>6000	>6000

Table S4. Antibody binding to uncleaved (haCBG) and cleaved (laCBG) CBG over different temperatures at physiological pH (pH 7.4). RU, response units. All data points are represented as mean values and their variance indicated in brackets asSEM as determined from three technical replicates of three technical replicates.

Temperature	**haCBG**		**laCBG**	
	Polyclonal (RU)	9G12 (RU)	Polyclonal (RU)	9G12 (RU)
25°C	51.7 (3.5)	34.2 (7.2)	85 (3.9)	1.1 (0.4)
37°C	63.9 (4.1)	32.2 (1.0)	87.6 (1.6)	-0.6 (0.3)
39°C	84.5	31.9	126.8	0.7

Figures

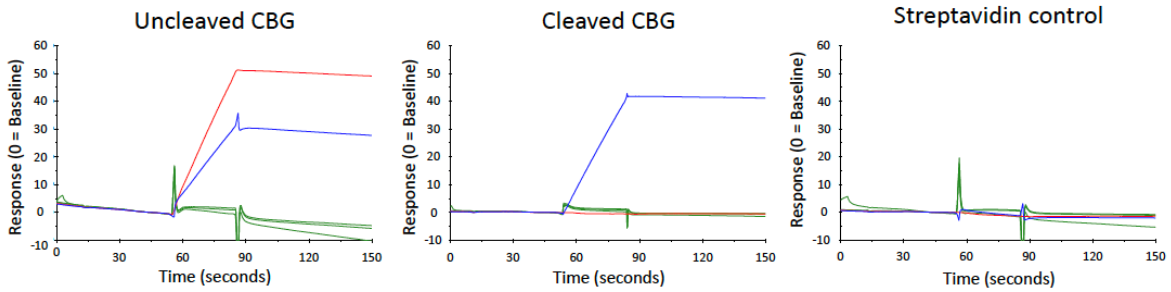


Figure S5. Validation of the immobilisation of haCBG and laCBG on separate SPR flow cells using the 9G12 and polyclonal antibodies recognising different human CBG epitopes. Polyclonal anti-CBG antibody measuring CBG regardless of confirmation (blue trace), 9G12 monoclonal antibody measuring only haCBG (red trace), buffer control (green trace). The streptavidin flow cell devoid of CBG was included as a negative control testing for unspecific antibody binding.

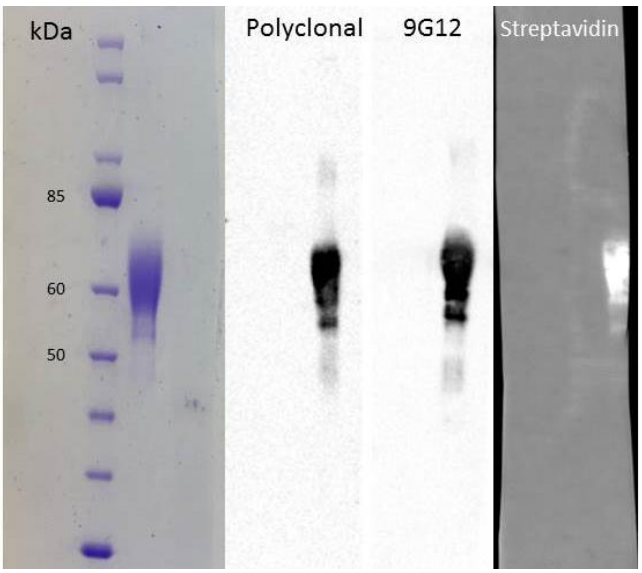


Figure S6. Coomassie stained SDS-PAGE (left lane) and the matching Western blots (three right lanes) of biotinylated recombinant human CBG probed with RCL antibodies 9G12 and CBG rabbit polyclonal antibody, as well as a streptavidin tag.

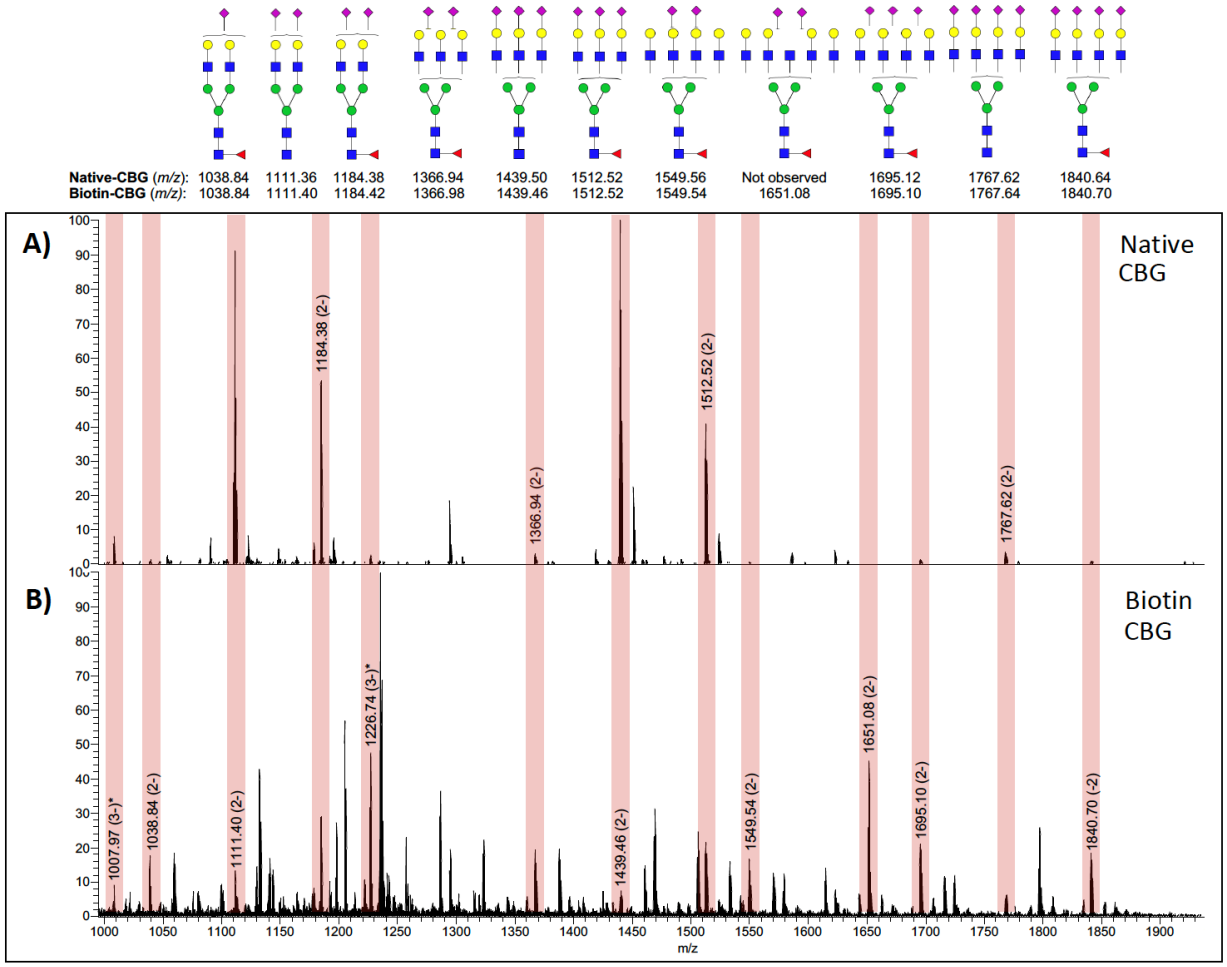


Figure S7. Biotinylated recombinant human CBG expressed in HEK293 cells (Biotin-CBG) and native human CBG isolated from donor blood (Native CBG) carry similar yet quantitatively different *N*-glycosylation features as documented using comparative PGC-LC-MS/MS *N*-glycome profiling. Vertical boxes in light red on the summed mass spectra highlight the *N*-glycans that are similar yet distinctly different between the two CBG forms. The *N*-glycan structures corresponding to the observed m/z values as supported by CID-MS/MS data are shown using the conventional symbol nomenclature for glycans. * Denotes an already depicted *N*-glycan structure also present in another charge state.

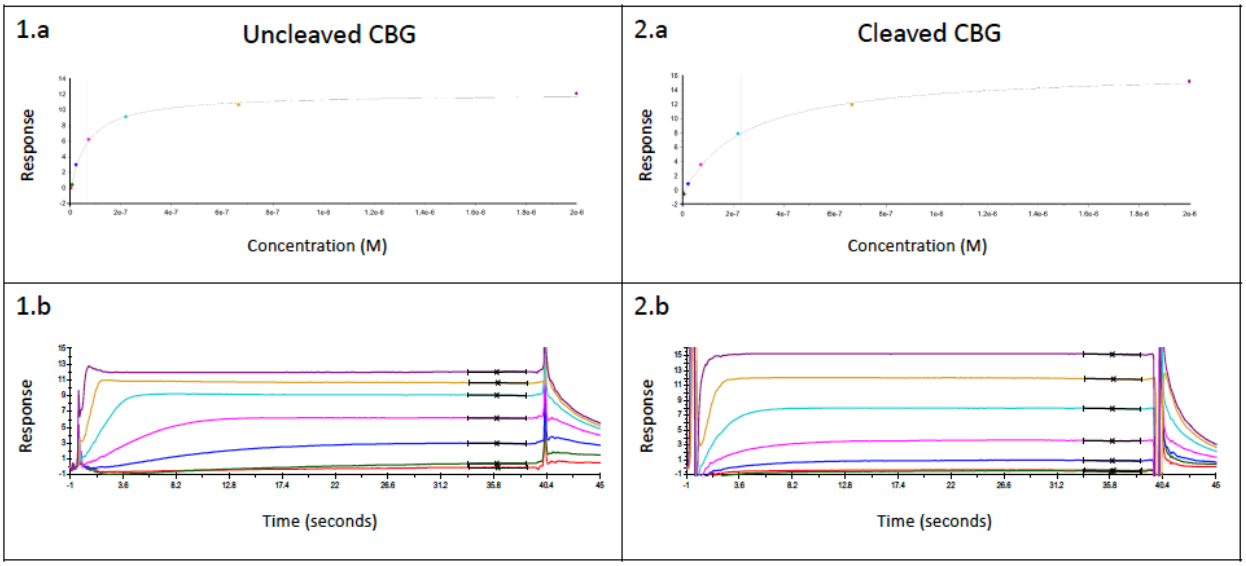


Figure S8. Binding affinity of hydrocortisone to haCBG at 25°C and pH 7.4. Representative SPR sensorgrams (a) and steady state analysis (b) for uncleaved (1) and cleaved CBG (2) and hydrocortisone.