





Position-specific *N*- and *O*-glycosylation of the reactive center loop impacts neutrophil elastase-mediated proteolysis of corticosteroid-binding globulin

Received for publication, March 21, 2023, and in revised form, November 20, 2023. Published, Papers in Press, November 30, 2023.

<https://doi.org/10.1016/j.jbc.2023.105519>

Anastasia Chernykh¹ , Jodie L. Abrahams^{1,2}, Oliver C. Grant³, Lucas Kambanis^{4,5}, Zeynep Sumer-Bayraktar^{1,6}, Julian Ugonotti¹ , Rebeca Kawahara^{1,7}, Leo Corcilius^{4,5}, Richard J. Payne^{4,5}, Robert J. Woods³, and Morten Thaysen-Andersen^{1,7,*}

From the ¹School of Natural Sciences, Macquarie University, Sydney, New South Wales, Australia; ²Glycosciences Laboratory, Department of Metabolism, Digestion and Reproduction, Imperial College London, London, United Kingdom; ³Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia, USA; ⁴School of Chemistry, The University of Sydney, Sydney, New South Wales, Australia; ⁵Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Sydney, Sydney, New South Wales, Australia; ⁶Glycometabolic Biochemistry Team, Cluster of Pioneering Research, RIKEN, Wako, Saitama, Japan; ⁷Institute for Glyco-core Research (iGCORE), Nagoya University, Nagoya, Japan

Reviewed by members of the JBC Editorial Board. Edited by Chris Whitfield

Corticosteroid-binding globulin (CBG) delivers anti-inflammatory cortisol to inflamed tissues through proteolysis of an exposed reactive center loop (RCL) by neutrophil elastase (NE). We previously demonstrated that RCL-localized Asn347-linked *N*-glycans impact NE proteolysis, but a comprehensive structure–function characterization of the RCL glycosylation is still required to better understand CBG glycobiochemistry. Herein, we first performed RCL-centric glycoproteomics of serum-derived CBG to elucidate the Asn347-glycans and then used molecular dynamics simulations to study their impact on NE proteolysis. Importantly, we also identified *O*-glycosylation (di/sialyl T) across four RCL sites (Thr338/Thr342/Thr345/Ser350) of serum CBG close to the NE-targeted Val344–Thr345 cleavage site. A restricted *N*- and *O*-glycan co-occurrence pattern on the RCL involving exclusively Asn347 and Thr338 glycosylation was experimentally observed and supported *in silico* by modeling of a CBG–GalNAc-transferase (GalNAc-T) complex with various RCL glycans. GalNAc-T2 and GalNAc-T3 abundantly expressed by liver and gall bladder, respectively, showed *in vitro* a capacity to transfer GalNAc (Tn) to multiple RCL sites suggesting their involvement in RCL *O*-glycosylation. Recombinant CBG was then used to determine roles of RCL *O*-glycosylation through longitudinal NE-centric proteolysis experiments, which demonstrated that both sialoglycans (disialyl T) and asialoglycans (T) decorating Thr345 inhibit NE proteolysis. Synthetic RCL *O*-glycopeptides expanded on these findings by showing that Thr345-Tn and Thr342-Tn confer strong and moderate protection against NE cleavage, respectively. Molecular dynamics substantiated that short Thr345-linked *O*-glycans abrogate NE interactions. In conclusion, we report on biologically relevant CBG RCL glycosylation events, which improve our understanding of mechanisms governing cortisol delivery to inflamed tissues.

Human corticosteroid-binding globulin (CBG) is a heavily *N*-glycosylated glucocorticoid-binding protein that binds and transports anti-inflammatory cortisol in blood within a dedicated hydrophobic steroid-binding pocket (1, 2). CBG, which is primarily produced by hepatocytes and less by other tissues, for example, gall bladder and kidney (3, 4), is responsible for the timely and tissue-specific delivery of cortisol to sites of inflammation by regulating the cortisol bioavailability to target cells under normal and stressed conditions (5–8).

Being a member of the serine protease inhibitor (serpin) superfamily, CBG adopts the common three-dimensional structure of serpins, but, importantly, does not confer protease inhibition (9, 10). The exposed reactive center loop (RCL), a serpin hallmark, spans Glu333–Ile354 of CBG. The RCL is a target for several proteases including the endogenous neutrophil elastase (NE) and chymotrypsin as well as several known exogenous virulence factors (11–13). These host-derived and microbial proteases target specific (but different) cleavage sites within the RCL. The potent NE, which is released *via* degranulation from activated neutrophils at sites of inflammation (14), rapidly and specifically cleaves the RCL at Val344–Thr345 (15). Importantly, NE-mediated RCL cleavage leads to profound stressed-to-relaxed conformational changes of CBG involving irreversible destruction of the cortisol-binding site, reduced cortisol binding affinity and, consequently, cortisol liberation from CBG (11, 15). Temperature and pH are other factors reported to affect the CBG:cortisol binding affinity independently of NE cleavage (16, 17).

While the exact biochemical changes in terms of altered site-specific glycosylation patterns and site occupancies are yet to be determined for CBG across various (patho)physiological conditions, the *N*-glycan modifications decorating CBG have been reported to vary with different physiological states including in pregnancy and with hormone fluctuations (18–20). Notably, *N*-glycosylation plays recognized roles in CBG

* For correspondence: Morten Thaysen-Andersen, morten.andersen@mq.edu.au.

RCL glycosylation modulates the proteolysis of CBG

secretion (21), circulation half-life (22), and steroid binding affinity and capacity (21, 23, 24). CBG harbors six *N*-glycosylation sites, including the Asn347 site located within the exposed RCL region. We previously performed site-specific structural characterization of serum CBG, which revealed the monosaccharide compositions of the *N*-glycans decorating the six *N*-glycosylation sites of the protein (25). We demonstrated that CBG carries site-specific patterns of highly sialofucosylated and branched complex-type *N*-glycans, but the glycan fine structural elements from each of the six *N*-glycosylation sites were not mapped in that study.

More recently, we published data indicating that several structural features of the Asn347-glycans located proximal to the Val344–Thr345 cleavage site affect the NE cleavage rate and thereby impact the cortisol release process (26). Specifically, we provided data suggesting that high branching, core fucosylation, and site occupancy of the Asn347-glycans reduce the NE-mediated cleavage efficiency, and, in turn, impede the cortisol release rate by conferring steric hindrance of the RCL cleavage site.

Despite providing new mechanistic insight into the cortisol delivery process, these previous studies have left some unaddressed questions regarding the fascinating relationship between RCL glycosylation and cortisol release. Not only are the fine structural details of the Asn347-glycans still missing, but a recent large-scale *O*-glycoproteomics study of human sera indicated (indirectly) the presence of *O*-glycosylation at Thr342 on the RCL of CBG amongst several thousand other serum glycopeptides not validated or investigated further in that study (27).

Inspired by these observations and building on our previous work in CBG glycobiology (25, 26), we here characterize the structure–function relationships of the glycosylated RCL region using comprehensive glycan- and glycopeptide-based profiling of serum-derived (native) CBG in combination with *in silico* molecular modeling, *in vitro* glycosylation assays, and longitudinal cleavage assays of recombinant CBG (rCBG) and synthetic RCL *O*-glycopeptides as well as nonglycosylated RCL peptide variants. We report on novel RCL *N*- and *O*-glycosylation events of CBG and document their functional importance in modulating the NE cleavage susceptibility. Collectively, these findings improve our understanding of the intriguingly complex CBG glycobiology governing cortisol delivery to inflamed tissues.

Results

Comprehensive Asn347-centric RCL *N*-glycoprofiling of serum CBG

We have previously reported on the site-specific *N*-glycosylation of serum CBG (25, 26), but the fine structures of the *N*-glycans decorating the highly occupied (>80%) Asn347 site located close to the NE cleavage site (Val344–Thr345) on the RCL remain uncharacterized (Fig. 1A). To this end, we first performed detailed Asn347-centric *N*-glycoprofiling using a strategy involving (i) SDS-PAGE-based isolation of the Asn347-containing C-terminal RCL peptide fragment after

exhaustive NE cleavage of serum CBG (Fig. S1), (ii) peptide:*N*-glycosidase F (PNGase F)-mediated in-gel release of *N*-glycans specifically from the Asn347 site on the C-terminal RCL peptide fragment, and (iii) porous graphitized carbon (PGC) liquid chromatography–tandem mass spectrometry (LC–MS/MS)-based glycomics performed with and without prior exoglycosidase treatment. The resulting *N*-glycomics data unveiled a diverse set of complex-type Asn347-localized *N*-glycans (Fig. 1B). By leveraging the information obtained from the PGC chromatographic elution profile, the MS/MS fragmentation patterns, and the sensitivity of the glycans to exoglycosidase treatment, we confidently mapped the isomeric Asn347-glycan species, established their relative abundances, and, in conjunction with glycopeptide analysis, determined the Asn347 site occupancy of serum CBG (Table S1A).

The abundant Asn347-glycans (~80% site occupancy) of serum CBG were found to span a heterogeneous population varying in their type and degree of sialylation, fucosylation, and antennary branching in line with (but expanding on) our previous reports of CBG glycosylation (25, 26) and recapitulating *N*-glycan features commonly observed in human serum (28, 29).

Several key Asn347-glycan structural features were identified including (1) prevalent 3'-arm antennary branching as compared with the less abundant 6'-arm branching for the triantennary *N*-glycans as identified after sialidase and galactosidase treatment (Fig. 1C, i), (2) abundant core fucosylation and a minor presence of antenna fucosylation as determined after sialidase treatment (Fig. 1C, ii and Fig. S2A), and (3) a dominance of α 2,6-sialylation over α 2,3-sialylation (Figs. 1C, iii and S2B).

Guided by the now elucidated Asn347-glycans, we then performed a series of molecular dynamics (MD) simulations of a CBG–NE complex to explore how key volume-enhancing features (core fucosylation and triantennary branching) of the Asn347-glycans impact the NE cleavage process. Specifically, MD simulations of four abundant and structurally related Asn347-glycoforms were performed, and the ability of NE to sterically access the RCL cleavage site in the presence of these glycans was assessed (Fig. 1D).

The two core fucosylated Asn347-glycans (FA2G2S2-a and FA3G3S3-a) displayed a significantly higher degree of steric contact with NE (both 100%) relative to their nonfucosylated counterparts (A2G2S2-a and A3G3S3-a) (93.3% and 97.6%, both $p < 0.05$, Student's *t* tests). Similarly, the triantennary Asn347-glycans (A3G3S3-a and FA3G3S3-a) showed a higher degree of steric contact with NE compared with their smaller biantennary counterparts (A2G2S2-a, FA2G2S2-a, both $p < 0.05$, Student's *t* tests). Collectively, these *in silico* analyses provide new insights into the spatial positioning and steric impact of the Asn347-glycans of the RCL of serum CBG and support experimental observations made in our previous work (Fig. 1E, replotted for clarity and support from Ref. (26)). Taken together, our data confirm that core fucosylation and triantennary branching of the Asn347-glycans are structural features that fine tune the NE cleavage rate of the RCL of serum CBG.

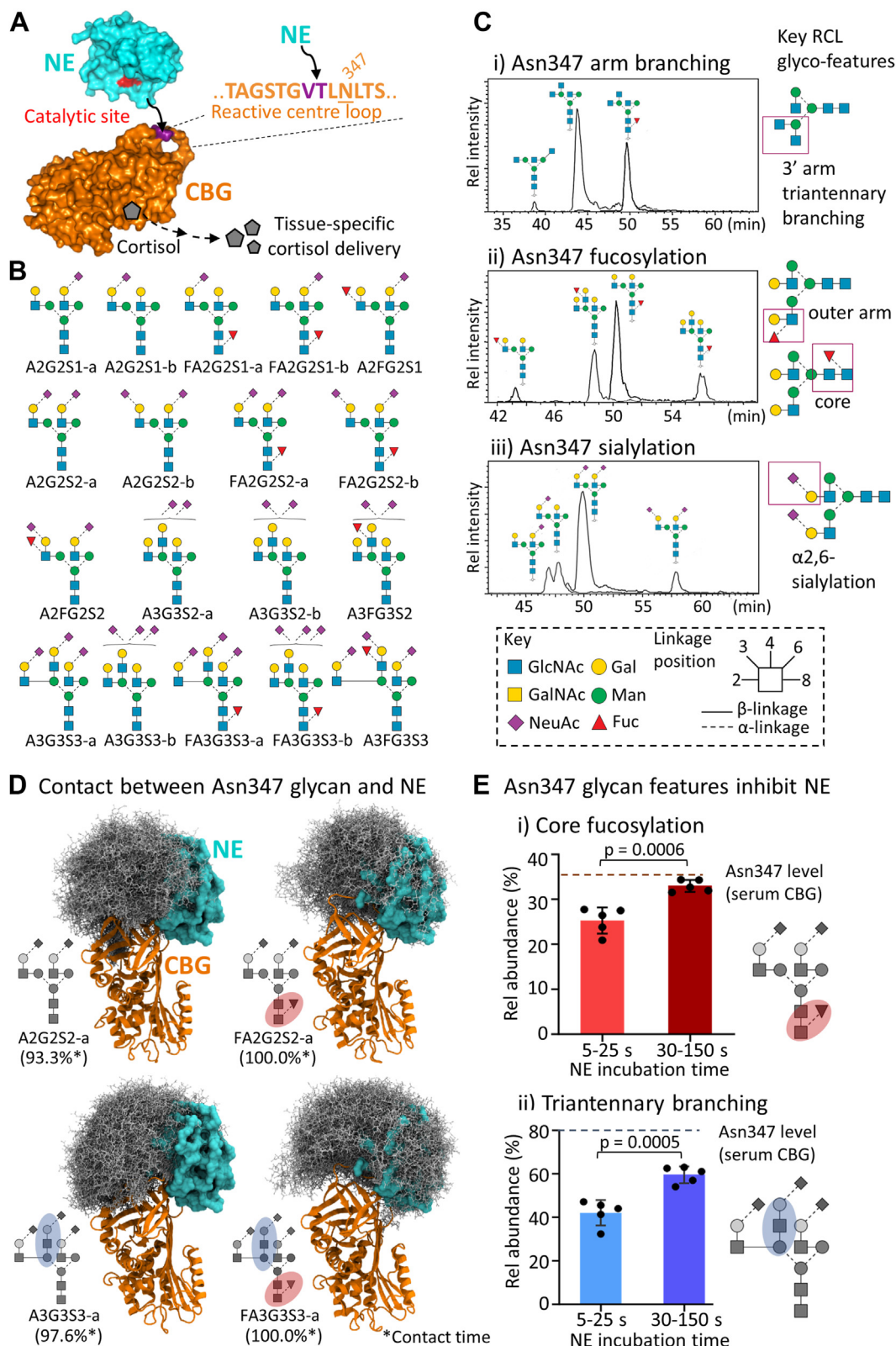


Figure 1. RCL Asn347-centric glycoprofiling of serum CBG. **A**, mechanism of cortisol delivery upon NE-mediated cleavage of the RCL and the subsequent structural rearrangement of the cortisol-carrying CBG illustrated using 3D structures of human NE (PDB code: 3Q77, cyan) and uncleaved (intact RCL) human CBG (homology model based on uncleaved human TBG, PDB code: 2CEO, orange). The NE cleavage site (VT) is in purple. **B**, RCL Asn347-glycans of serum CBG (and their short-hand nomenclature) identified by PGC-LC-MS/MS, see Table S1A for Asn347 site occupancy information. Trace glycoforms have been omitted for simplicity. **C**, identification of the fine structural features of the Asn347-glycans including (i) the type of antennary branching as determined after sialidase and galactosidase treatment, (ii) the fucose position as determined after sialidase treatment and (iii) the sialyl linkage type. Representative extracted ion chromatograms of key glycans are shown, and prominent Asn347-glycan features are highlighted (right). **D**, NE contact (steric clash) conferred by four prominent Asn347-glycans (gray, ball/stick representations) as determined *in silico* by MD simulations of NE (cyan, solid surface) complexed with

RCL glycosylation modulates the proteolysis of CBG

Newly discovered RCL O-glycosylation of serum CBG

We then investigated the intriguing CBG RCL O-glycosylation event recently reported (in passing) in a global O-glycoproteomics screen of human sera using the OpeRATOR enzyme within an “EXoO” method (27). Aided by the Bionic glycopeptide search engine (30), we reinterrogated the publicly available LC–MS/MS data from that serum O-glycoproteomics study (PRIDE: PXD009476) and surprisingly found robust spectral evidence for CBG O-glycopeptides not only from the Thr342 O-glycosylation site (as reported in the associated article) but also the nearby Thr345 site (Fig. S3). We also found evidence for previously unreported RCL Thr342 O-glycopeptides in both normal and tumorous kidney tissues from the high-quality data generated by the same study (Fig. S4) suggesting that O-glycosylation may span multiple RCL sites and is neither tissue-specific nor a rare event for CBG.

Guided by these system-wide observations, we then performed site-specific glycopeptide profiling of serum CBG, which, excitingly, unveiled four O-glycosylation sites within the RCL, that is, Thr338, Thr342, Thr345, and Ser350 (Fig. 2A). The four sites were decorated with sialyl T (NeuAc₁Gal₁GalNAc₁) and/or disialyl T (NeuAc₂Gal₁GalNAc₁) O-glycans (Fig. 2B).

Importantly, all four O-glycosylation sites were unambiguously elucidated to amino acid resolution with electron transfer higher-energy collision dissociation (EThcD)–MS/MS (Figs. 2, C and D and S5–S7). While the occupancy of the individual O-glycosylation sites could not be accurately determined because of incomplete LC separation of the positional RCL O-glycopeptide isomers, the global level of RCL O-glycosylation was estimated to represent a relatively minor proportion compared with the dominant RCL Asn347-glycosylation as determined from PNGase F-treated (de-N-glycosylated) serum CBG (Table S1A). While experimental data are still lacking, it may be anticipated that Thr338, Thr342, and Thr345 exhibit higher site occupancy as compared with Ser350 because of the observed preference of polypeptide GalNAc transferases (GalNAc-Ts) for Thr over Ser acceptor residues (31).

We found no evidence of co-occurrence of multiple O-glycans across any of the four identified O-glycosylation sites within the RCL of serum CBG. However, one of the O-glycosylation sites, Thr338, demonstrated an interesting O-glycan co-occurrence with an Asn347-linked N-glycan (Fig. 3A and Table S1A), whereas no O- and N-glycan co-occurrence of Asn347 and the proximal Thr342, Thr345, and Ser350 O-glycosylation sites was detected (Fig. 3B, i).

Seeking to understand the molecular basis for the restricted co-occurrence pattern, we hypothesized that the initiation of the O-glycosylation process, which is mediated by a family of

cis-Golgi-resident GalNAc-Ts, is disfavored for O-glycosylation sites located proximal to the Asn347 site undergoing N-glycosylation initiated earlier in the endoplasmic reticulum (Fig. 3B, ii). To support this hypothesis, we explored through *in silico* modeling of a complex between CBG and GalNAc-T2 (see *in vitro* O-glycosylation assay for rationale for choosing GalNAc-T2) if the Asn347 site could accommodate an unprocessed oligomannosidic N-glycan (M5) when separately aligning the active site of the transferase across the four discrete RCL O-glycosylation sites (Fig. 3C). In line with the observed co-occurrence pattern, the bulky M5 could exclusively be accommodated when aligning GalNAc-T2 to Thr338, not the Thr342, Thr345, and Ser350 sites that displayed molecular crowding locally around Asn347.

Attempting to validate the lack of Asn347-glycan co-occurrence with the strategically located Thr345, we then performed MD simulations of a complex between GalNAc-T2 (aligned to Thr345) and CBG carrying an unprocessed Asn347-M5 glycan (Fig. S8). In agreement with the experimental data, only 0.4% of the shape conformers were compatible with O-glycan transfer to the Thr345 site when the RCL was installed with an Asn347-M5 N-glycan. Collectively, these *in silico* observations provide molecular-level clues to the restricted O- and N-glycan co-occurrence patterns experimentally observed for serum CBG.

GalNAc-T2 and GalNAc-T3 facilitate RCL O-glycosylation *in vitro*

To investigate the GalNAc-T isoenzyme(s) responsible for the RCL O-glycosylation observed for serum CBG, we used *in vitro* assays to explore if and to which sites human GalNAc-T2 and GalNAc-T3 transfer O-glycans to a synthetic RCL peptide (Fig. 4A). GalNAc-T2 and GalNAc-T3 were chosen for these experiments as these GalNAc-T isoenzymes exhibit high expression in liver and gall bladder, respectively, two of the principal tissue origins of CBG (4, 32) (Fig. 4B). As confirmed by EThcD–MS/MS (Figs. S9–S17), both GalNAc-T2 and GalNAc-T3 demonstrated the ability to transfer GalNAc residues (Tn) to multiple sites on the RCL (Fig. 4C). While both transferases readily transferred Tn to Thr349 and Ser350, only GalNAc-T2 modified Thr342 and GalNAc-T3 exclusively modified Thr345 with Tn residues. Taken together, these *in vitro* observations indicate that both GalNAc-T2 and GalNAc-T3 may facilitate the RCL O-glycosylation events observed for serum CBG.

Thr345 O-glycosylation inhibits NE-mediated RCL proteolysis

Next, we investigated if the newly identified RCL O-glycosylation influences the NE cleavage reaction. While only displaying relatively low occupancy for serum CBG (Table S1A),

glycosylated CBG (orange, ribbon structure). In total, 2500 snapshots were extracted from 2.5 μ s of cumulative MD simulation time for each of the four CBG glycoforms. Multiple MD-derived snapshots of the Asn347-glycans are shown with a single snapshot of the CBG–NE complex. *The proportion of NE contact time (in %) is indicated. E, longitudinal NE-based cleavage of serum CBG demonstrating the impact of (i) core fucosylation and (ii) antennary branching of the Asn347-glycan on the NE cleavage rate replotted, for clarity and support, for six dominating Asn347-glycoforms from our previous study (26). Dashed lines indicate the native modification level for serum CBG. Data plotted as mean \pm SD, n = 5, Student's *t* test. CBG, corticosteroid-binding globulin; MD, molecular dynamics; NE, neutrophil elastase; PDB, Protein Data Bank; RCL, reactive center loop.

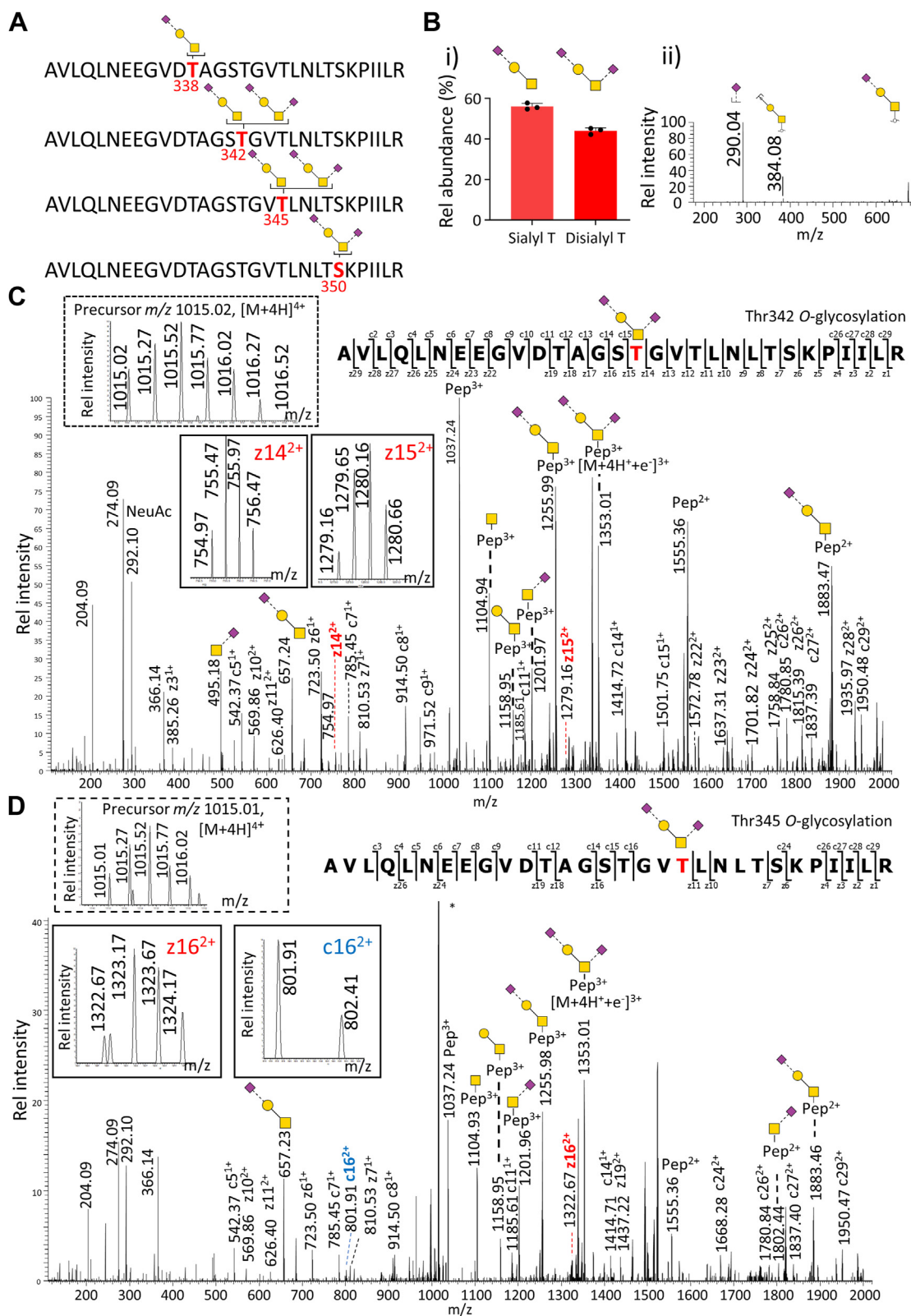
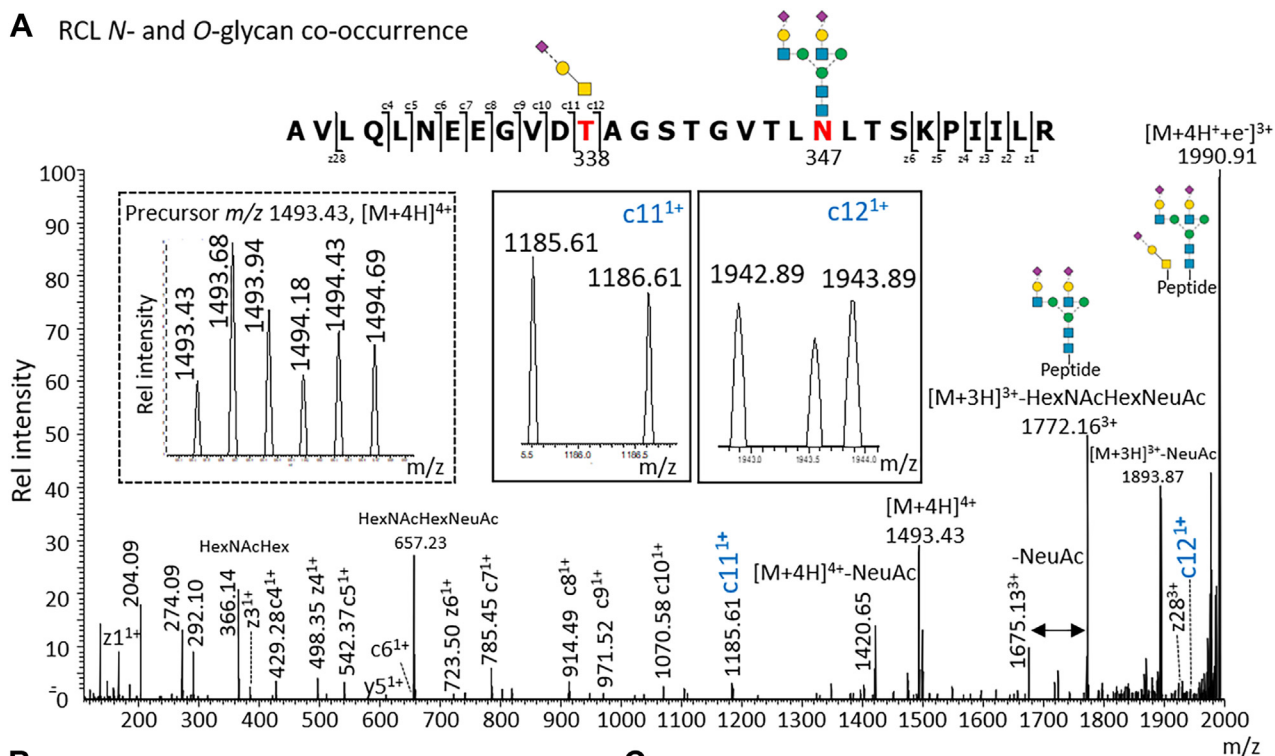


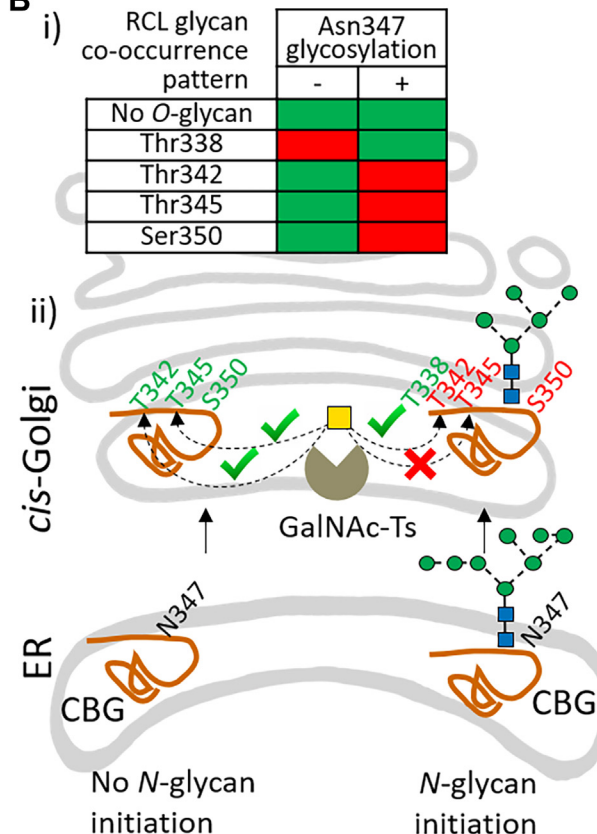
Figure 2. Newly discovered RCL O-glycosylation of serum CBG. A, four O-glycosylation sites each carrying one or more O-glycan structures were identified on the RCL of serum CBG at relatively low abundance, see Table S1 for site occupancy. B, (i) Distribution of O-glycans decorating the CBG RCL based on glycopeptide data (mean + SD, n = 3). (ii) CID-MS/MS spectrum of an O-glycan (sialyl T) released from serum CBG (glycomics data). C and D, annotated ETHcD-MS/MS spectra of two RCL O-glycopeptides providing evidence of Thr342 and Thr345 RCL O-glycosylation, respectively. These two RCL O-glycopeptides, which were not treated with any exoglycosidases or endoglycosidases, were not found to carry any Asn347-linked N-glycosylation. Insets, precursor (broken box) and key fragment ions (full boxes) documenting the site-specific O-glycosylation events. *Interfering fragment of a coisolated precursor ion unrelated to the targeted CBG glycopeptides. See Figure 1 for key to glycan symbols. CBG, corticosteroid-binding globulin; CID, collision-induced dissociation; ETHcD, electron transfer higher-energy collision dissociation; RCL, reactive center loop.

RCL glycosylation modulates the proteolysis of CBG

A RCL N- and O-glycan co-occurrence



B



C

Asn347-M5 is accommodated when GalNAc-T is aligned to Thr338

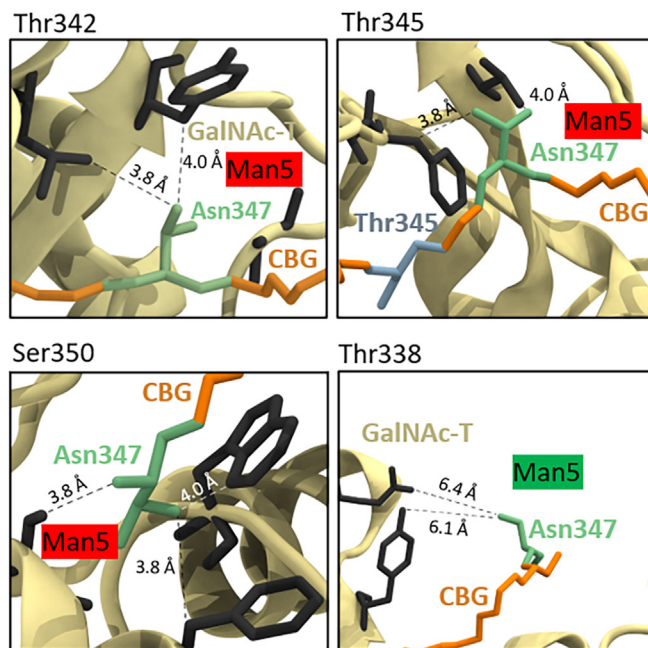


Figure 3. Co-occurrence of RCL N- and O-glycosylation on serum CBG. A, EthCD-MS/MS spectrum of an RCL glycopeptide co-occupied with an O- and N-glycan (Thr338 and Asn347). Inset, precursor (broken) and key fragment glycosylation events (full) documenting the site-specific glycosylation events. B, (i) Observed (green) and nondetected (red) RCL O-glycoforms in the absence (–) and presence (+) of Asn347-glycosylation. (ii) Proposed spatiotemporal interplay between RCL N- and O-glycosylation providing a possible explanation for the absence of specific co-occupied RCL glycopeptides (e.g., Thr345 > Asn347). C, four *in silico* models of a complex between nonglycosylated CBG and GalNAc-T2 (see below for rationale for choosing GalNAc-T2) produced by aligning the transferase across the four discrete RCL O-glycosylation sites of serum CBG. The local environment around the Asn347 residue (green) was visually assessed for the ability to accommodate a bulky oligomannosidic-type N-glycan (Man5). See Figure 1 for key to glycan symbols. CBG, corticosteroid-binding globulin; EthCD, electron transfer higher-energy collision dissociation; GalNAc-T, GalNAc transferase; RCL, reactive center loop.

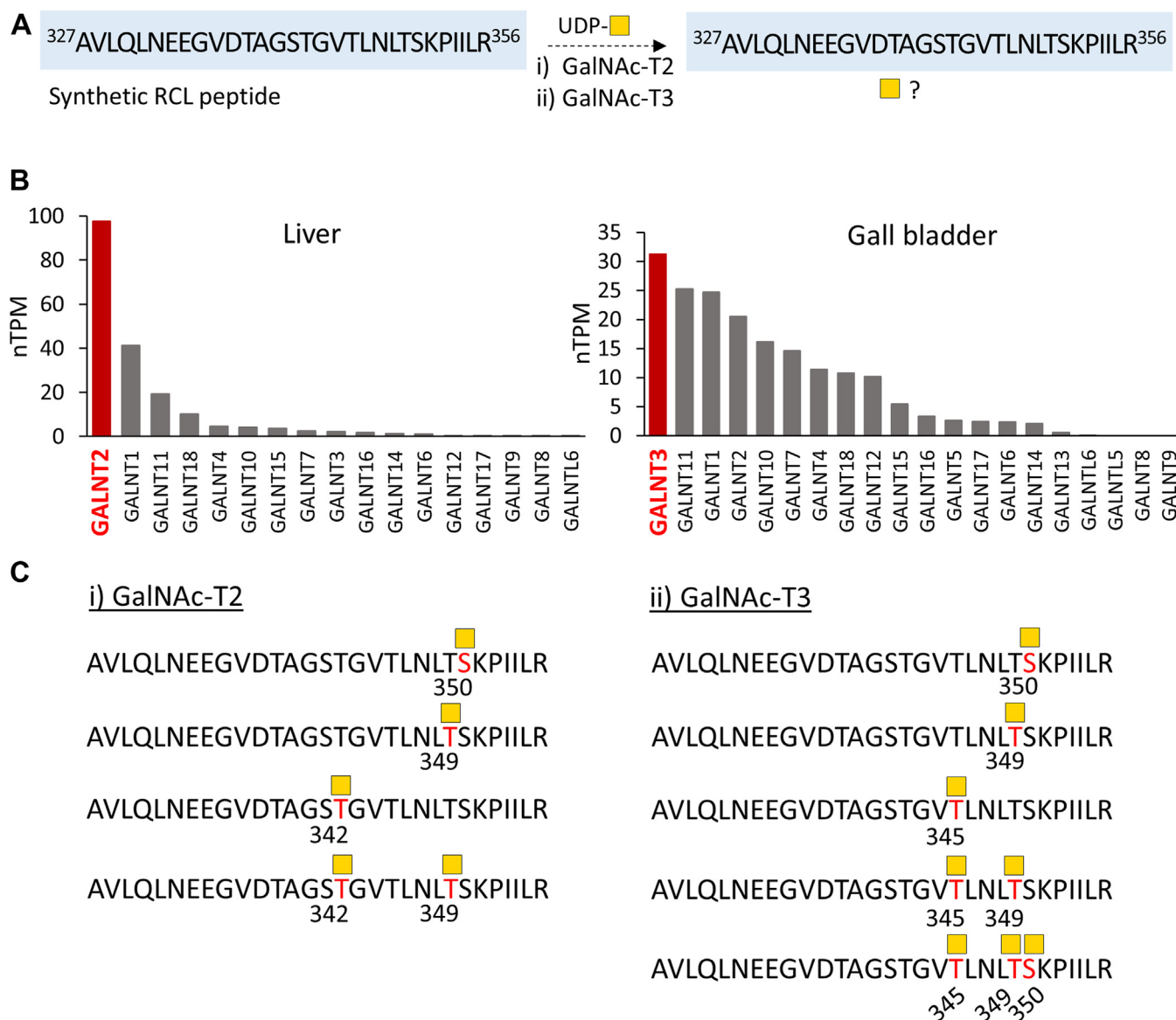


Figure 4. GalNAc-T2 and GalNAc-T3 transfer O-glycans to a synthetic RCL peptide *in vitro*. *A*, schematic representation of the *in vitro* glycosylation assay exploring how and where (i) GalNAc-T2 and (ii) GalNAc-T3 transfer GalNAc (Tn) moieties to a synthetic RCL peptide. *B*, transcript (mRNA) expression profile of *GALNTs* in liver and gall bladder (the principal tissue origins of CBG) based on RNA consensus tissue gene data from the Human Protein Atlas (4). *C*, glycopeptide products of the *in vitro* O-glycosylation reactions by (i) GalNAc-T2 and (ii) GalNAc-T3. The RCL O-glycans sites were identified using ETHcD-MS/MS (Figs. S9–S17). See Figure 1 for key to glycan symbols. ETHcD, electron transfer higher-energy collision dissociation; nTPM, normalized transcript per million; RCL, reactive center loop.

we decided to focus on the Thr345 O-glycosylation site strategically located at the Val344–Thr345 cleavage site using a human embryonic kidney 293 (HEK293)–derived rCBG model system. An initial comparative global (site unspecific) glycan profiling analysis demonstrated that rCBG and serum CBG carry qualitatively similar but quantitatively different N-glycans (Fig. S18 and Table S2). Proving ideal for these structure–function experiments, the RCL-specific glycosylation analysis of rCBG revealed a very low N-glycan occupancy of Asn347 (<2%) as compared with serum CBG (>80%) and, importantly, relative high levels of RCL O-glycosylation (18.8%) decorating exclusively the Thr345 O-glycosylation site (Figs. 5A and S19, and Table S3). Three O-glycans were identified at the Thr345 O-glycosylation site of rCBG including a disialyl T O-glycan

(NeuAc₂Gal₁GalNAc₁) representing by far the most abundant glycoform (93.5% relative abundance) (Fig. 5B). Given these characteristics, rCBG was deemed suitable to explore any protective roles of the Thr345 O-glycosylation against the NE-mediated proteolysis of RCL.

The NE-mediated cleavage of the RCL occurring at Val344–Thr345 results in the formation of two cleavage products that, in turn, give rise to two MS-friendly peptides spanning the RCL region upon further trypsin digestion (Fig. 5C). To simultaneously explore the importance of the terminal sialylation decorating the Thr345-glycans of rCBG, we also generated a sialic acid–deficient glycoform (asialo-rCBG) through exhaustive sialidase treatment under native conditions of rCBG. We then longitudinally followed the proteolysis

RCL glycosylation modulates the proteolysis of CBG

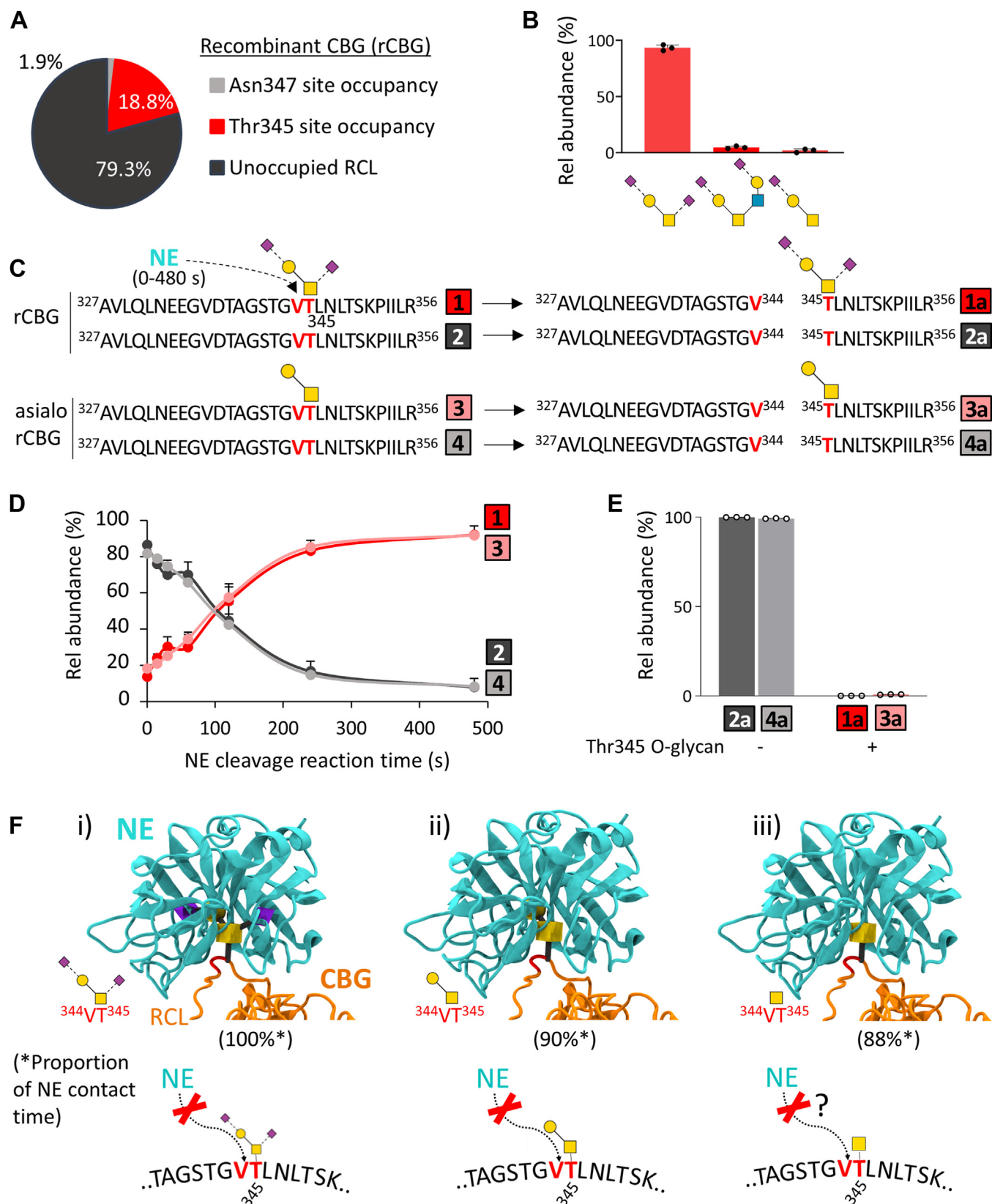


Figure 5. Protective role of Thr345 O-glycosylation of rCBG against NE-mediated proteolysis. A, RCL glycan occupancies of HEK293-derived rCBG. B, distribution of Thr345 O-glycans of rCBG determined based on glycopeptide data. Data points are plotted as mean + SD, n = 3. C, overview of the RCL peptides before and after NE cleavage with downstream trypsin digestion. D, NE cleavage of rCBG and asialylated rCBG (asialo-rCBG) over time. Distribution of unoccupied uncleaved (dark/light gray) and O-glycosylated (red/pink) uncleaved RCL tryptic peptides of rCBG and asialo-rCBG identified based on glycopeptide data. Data points are plotted as mean + SD, n = 3. E, distribution of nonglycosylated and O-glycosylated C-terminal peptide fragments detected after 480 s of NE incubation. Data points are plotted as mean + SD, n = 3. F, alignment of NE (cyan, PDB code: 7CBK) onto MD-generated shape

reaction of natively folded rCBG and asialo-rCBG by monitoring the unoccupied and *O*-glycosylated (sialo- and asialo-) forms of intact (uncleaved) RCL peptides and the generated cleavage products (after downstream trypsin digestion) over a 480 s incubation period with NE (Fig. 5, D and E and Table S4). Interestingly, the cleavage data revealed that both the sialylated and asialylated *O*-glycans at Thr345 strongly inhibit the NE proteolysis reaction, whereas the non-glycosylated RCL is rapidly and selectively cleaved by NE.

We supported these observations by modeling the interactions between NE and the RCL of rCBG decorated with the abundant disialyl T *O*-glycan and with shorter asialo- (T) and agalacto- (Tn) glycoforms (Fig. 5F). Recapitulating the cleavage data, the MD simulations indicated that disialyl T and T glycoforms at the Thr345 site strongly interfere with the accessibility of NE to the RCL substrate (100% and 90% contact time, respectively). While not included in the rCBG cleavage experiments, the MD data suggested that even the ultrashort Thr345-Tn glycoform may obstruct the access by NE (88% contact time) and consequently potentially impede the cleavage process. The obstruction of the NE-mediated proteolysis reaction conferred by even ultrashort *O*-glycans decorating Thr345 may not be entirely surprising since the Thr side chain proposedly is involved in the catalytic action of NE (33).

Prompted by these observations, we then explored the relationship between RCL *O*-glycosylation and the NE cleavage reaction under controlled conditions using synthetic RCL peptides. Specifically, we generated synthetic RCL peptides with and without *O*-glycosylation (Tn) installed site-specifically at Thr342 or Thr345 and performed NE digestion experiments that were longitudinally monitored by MALDI-MS (Fig. 6A). As expected, the unoccupied RCL peptide was rapidly cleaved by NE as demonstrated by the near absence of the RCL peptide substrate (and concomitant presence of the expected peptide products) after only 120 s digestion (Fig. 6, B and C, i). Interestingly, the Tn *O*-glycans at Thr342 and Thr345 were, in contrast, found to confer moderate and strong (near-complete) protection against NE cleavage over the cleavage time course, respectively (Fig. 6, B and C, ii and iii).

Taken together, the longitudinal cleavage experiments of both intact rCBG and synthetic RCL glycopeptides have revealed previously unknown relationships between the site-specific *O*-glycosylation events of the RCL and their protection against NE-mediated cleavage. Our data demonstrate that both full-length and short *O*-glycans confer strong protection against NE proteolysis when installed at strategic RCL sites.

Discussion

CBG plays an essential role in regulating circulating plasma cortisol levels and is instrumental for the timely delivery of

cortisol to inflamed tissues (34). Cortisol release from CBG is primarily driven by the proteolysis of the RCL by NE present at sites of inflammation. We have previously shown that the prevalent RCL *N*-glycosylation at Asn347 affects the NE-mediated cleavage of the Val344–Thr345 site (26). In the present work, we expand on this previous finding, by first detailing the fine structural elements of the Asn347-glycans, their spatial orientation on the RCL and, in turn, their effect on the NE proteolysis process. Our detailed site-specific profiling of the RCL *N*-glycans enabled by PGC-LC-MS/MS-based glycomics of an isolated C-terminal peptide fragment containing only the Asn347 *N*-glycosylation site revealed a heterogeneous collection of complex-type *N*-glycans carrying a range of features commonly observed for liver-derived glycoproteins circulating in plasma including prominent α 2,6-sialylation, triantennary branching, and core fucosylation (28, 29, 35).

In line with our previous experimental observations (26), molecular modeling of the CBG-NE complex confirmed that volume-enhancing Asn347-glycan features (*i.e.*, core fucosylation and triantennary branching) reduce the accessibility of NE to the RCL cleavage site and therefore impede Val344–Thr345 cleavage and, consequently, inhibit the cortisol release rate. While the molecular modeling provides important support for (and expand our insight into) this intuitive relationship, it should be noted that the NE accessibility analysis is limited by considering only a single conformation of the protease, thus ignoring any potential induced fit of NE, as well as any alignment of the RCL peptide substrate in the active site and potential initial interactions between NE and the CBG glycans that may enhance the surface accessibility. Despite these limitations inherent to the modeling approaches used herein, our analysis has provided new structural and functional insights into the heterogeneous RCL *N*-glycosylation, which confirms that the NE-mediated cleavage process is modulated sterically by voluminous Asn347-glycans located close to the RCL cleavage site.

Human CBG is a well-studied glycoprotein featuring six utilized *N*-glycosylation sites. Adding another layer of complexity, a recent large-scale *O*-glycoproteomics study reported on CBG *O*-glycosylation at Thr342 amongst several thousand other serum *O*-glycopeptides neither validated nor investigated further in that study (27). Since data in that study were acquired from samples treated with sialidase, the *O*-glycans were observed without any terminating NeuAc residues as HexNAc₁ and HexNAc₁Hex₁ at Thr342. Another limitation of that study was the site-specific assignment of the reported *O*-glycosylation sites, which was based on indirect evidence through higher-energy collisional dissociation (HCD)-MS/MS data of peptides generated by OpeRATOR, a commercial *O*-glycoprotease. OpeRATOR is a bacterial enzyme that reportedly cleaves N-terminal to *O*-glycosylated Ser/Thr, thus generating peptides that generally carry N-terminal *O*-

conformers of uncleaved CBG (orange, PDB code: 4BB2/2CEO) carrying relevant Thr345 *O*-glycans, that is, (i) disialyl T, (ii) T, and (iii) Tn (3D-SNFG, top) and schematics of their demonstrated or suggested protective roles against NE-mediated RCL cleavage (bottom). The proportion of NE contact time (in %) of the respective CBG *O*-glycans is provided. See Figure 1 for key to glycan symbols. HEK293, human embryonic kidney 293 cell line; MD, molecular dynamics; NE, neutrophil elastase; rCBG, recombinant corticosteroid-binding globulin; RCL, reactive center loop.

RCL glycosylation modulates the proteolysis of CBG

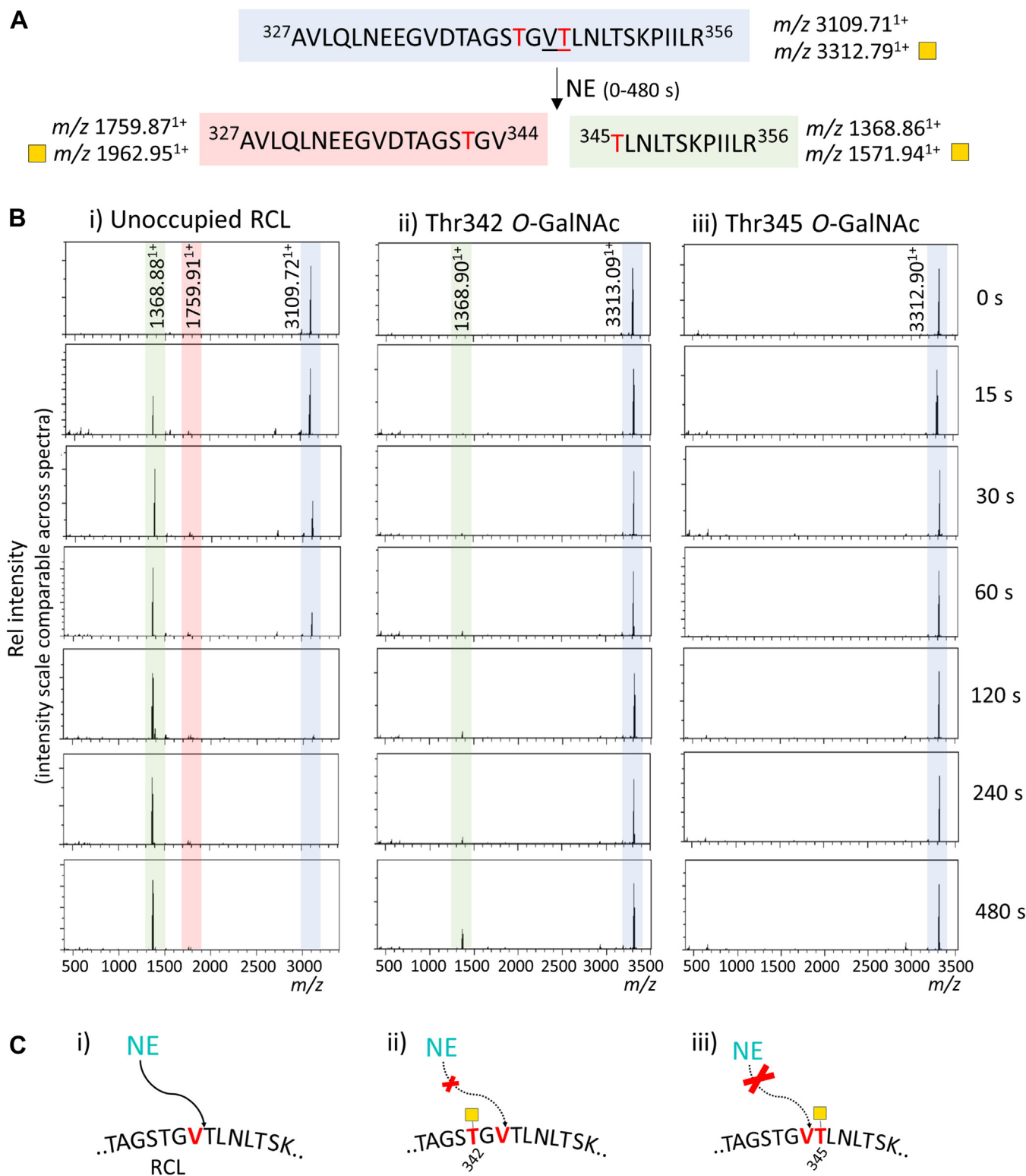


Figure 6. Thr342- and Thr345-GalNAc (Tn) O-glycans impede NE proteolysis of synthetic RCL peptides. *A*, schematic representation of the synthetic CBG RCL (glyco)peptides and their expected cleavage products upon NE proteolysis. m/z values for both the nonglycosylated and Tn peptides are provided. *B*, NE-based cleavage of (i) nonglycosylated, (ii) Thr342, and (iii) Thr345 O-glycosylated RCL peptides over time as monitored using MALDI-TOF MS. Intact (uncleaved) nonglycosylated and O-glycosylated RCL peptides are in blue, RCL N-terminal cleavage products in red, and RCL C-terminal cleavage products in green. *C*, schematics of the observed NE accessibility to (i) unoccupied, (ii) Thr342, and (iii) Thr345 O-glycosylated RCL. See Figure 1 for key to glycan symbols. CBG, corticosteroid-binding globulin; NE, neutrophil elastase; RCL, reactive center loop.

glycosylation (36, 37). As recently discussed (38), HCD-MS/MS data of OpeRATOR-generated peptides are less than ideal for confident *O*-glycosylation site localization, whereas EThcD-MS/MS in favorable cases provides direct evidence for unambiguous site assignment. Upon revisiting and reinterrogating the publicly available LC-MS/MS data from that previous study, we identified an additional potential *O*-glycosylation site (Thr345) within the RCL region of CBG from human sera.

Prompted by these preliminary observations, we performed comprehensive glycopeptide profiling of serum CBG and excitingly identified four utilized *O*-glycosylation sites on the RCL, that is, Thr338, Thr342, Thr345, and Ser350. Evidence from intact *O*-glycopeptides revealed that all four sites of serum CBG carry sialyl T (NeuAc₁Gal₁GalNAc₁) and/or disialyl T (NeuAc₂Gal₁GalNAc₁) *O*-glycans albeit at a relatively low occupancy level. Surprisingly, we did not find any *O*-sites carrying NeuAc₁GalNAc₁ (sialyl Tn), which, after desialylation, would give rise to the HexNAc₁ glycan of serum CBG previously reported by Yang *et al.* (27).

No co-occurrence of multiple *O*-glycans on the RCL was detected of serum CBG, but we identified an interesting co-occurrence relationship between the *N*- and *O*-glycans on the RCL. While the Thr342, Thr345, and Ser350 *O*-glycosylation sites located in -5, -2, and +3 positions relative to Asn347, respectively, showed no co-occupancy with the *N*-glycosylation of Asn347, *O*-glycosylation of the distal Thr338 (-9 position) was found to coexist with *N*-glycosylated Asn347. The restricted co-occurrence pattern was supported through *in silico* modeling of various complexes between GalNAc-T2 and CBG displaying different glycoforms, which supported that only the distal Thr338 can engage uninterrupted with the GalNAc-T2 when Asn347 is modified by *N*-glycosylation.

Densely glycosylated regions featuring multiple closely spaced *O*-glycosylation sites is a known characteristic of mucins and glycoproteins containing mucin-like domains (39, 40); however, reports on co-occurrence and “crosstalk” of *N*- and *O*-glycosylation are less common. Amongst the few examples we could find in the literature, several *O*-glycosylation sites of recombinant severe acute respiratory syndrome coronavirus 2 spike glycoprotein expressed in HEK293F cells were identified within or next to *N*-glycosylation sequons (NxS/T, x ≠ P) and found to be occupied primarily in the absence of *N*-glycans suggesting an antagonistic relationship between the two types of glycosylation (41). Opposing that report, another study on the severe acute respiratory syndrome coronavirus 2 spike glycoprotein found many utilized *O*-glycosites in the +2 position of *N*-glycosylated sequons (42). Separately, the analysis of human plasma protease C1-inhibitor also indicated the possibility of co-occupancy of closely spaced *N*- and *O*-glycosylation sites (43). Our study adds to these examples of *N*- and *O*-glycosylation co-occurrence, which remains under-reported and mechanistically incompletely understood, in part, since methods for *N*- and *O*-glycoproteomics differ and often are applied separately to address bespoke glycobiological questions (44, 45).

In addition to CBG, the aforementioned *O*-glycoproteomics study by Yang *et al.* (27) identified *O*-glycosylation of the RCL region of a few other clade A members of the serpin family closely related to CBG including alpha-1-antichymotrypsin (UniProtKB: P01011) in positions P3 and P11 (numbering according to Schechter and Berger (46)) and thyroxine-binding globulin (P05543) in position P7 (27). Importantly, these *O*-glycosylation sites were located in proximity to proteolytic sites on their respective RCL peptide regions suggesting the potential involvement of *O*-glycans in modulating the proteolysis process for those proteins as also demonstrated herein for CBG. RCL *O*-glycosylation was also previously reported for serpins from other clades, including alpha-2-antiplasmin (P08697, position P7), antithrombin-III (P01008, position P8), and heparin cofactor 2 (P05546, positions P7, P10, and P11) (27). Taken together, these observations demonstrate that *O*-glycosylation of the RCL region is a widespread phenomenon of serpins not unique to CBG. Furthermore, a number of serpin proteins are known to be *O*-glycosylated in non-RCL regions of the protein, that is, plasma protease C1-inhibitor (P05155) (27, 43, 47–50), plasminogen activator inhibitor 1 (P05121) (47), plasma serine protease inhibitor (P05154) (47–50), antithrombin-III (50), and alpha-2-antiplasmin and alpha-1-antitrypsin (P01009) (27). We did not identify any *O*-glycosylation outside the RCL region of neither serum CBG nor rCBG.

Our *in vitro* *O*-glycosylation experiments pointed to both GalNAc-T2 and GalNAc-T3 commonly expressed in liver and gall bladder, respectively, as likely candidates to facilitate the RCL *O*-glycosylation of serum CBG. The gall bladder and in particular the liver are principal tissue origins of CBG in circulation (4). However, the fact that neither GalNAc-T2 nor GalNAc-T3 glycosylated the Thr338 site in our *in vitro* assay, suggests that multiple (some still-to-be-defined) GalNAc-T isoenzymes perhaps from a variety of tissue origins contribute to the heterogenous *O*-glycosylation of the RCL of serum CBG, which therefore requires further investigation.

Given the known biological roles of GalNAc-type (mucin-type) *O*-glycosylation to impact and regulate proteolysis (51–53), we decided to investigate how the newly discovered CBG *O*-glycosylation affects the RCL cleavage. Indeed, the longitudinal cleavage experiments of rCBG and synthetic RCL *O*-glycopeptides by NE demonstrated a moderate and strong (near-complete) protective role of Thr342 (P3) and Thr345 (P1') *O*-glycosylation, respectively, against NE cleavage. A previous study, which performed an *in silico* analysis of known *O*-glycosylation sites and NE cleavage sites within the human proteome, revealed an enrichment of *O*-glycosylated amino acid residues in vicinity of the NE cleavage sites (spanning the region from P7 to P7') suggesting that beneficial relationships exist between *O*-glycosylation and NE cleavage events (54). In support, the same study also found that the loss of GalNAc-T2 in HepG2 SimpleCells resulted in less cleavage of several NE substrates including proteins with NE cleavage sites located within known *O*-glycosylated regions. In line with the observations reported in our present study, other work by the same authors, however, revealed that an *O*-glycan in the vicinity of a

RCL glycosylation modulates the proteolysis of CBG

cleavage site inhibited NE proteolysis as shown *in vitro* using synthetic peptides (47). As highlighted by these disparate reports, the contribution of the individual *O*-glycosylation sites, *O*-glycan structures, and local peptide environments to the overall NE-mediated “degradome” of cells and tissues remains poorly understood.

While our CBG data indicated that the Thr345 *O*-glycan confers near-complete protection against NE cleavage, the RCL may still be susceptible to proteolysis at other accessible cleavage sites by enzymes other than NE present at sites of inflammation and infection. For example, the *Pseudomonas aeruginosa* virulence factor lasB (also known as *P. aeruginosa* elastase, that is, PAE, Q02RJ6) is known to target several cleavage sites within the RCL including Thr345–Leu346 and Asn347–Leu348 (12, 26). However, it remains to be explored if (and how) the elaborate RCL *N*- and *O*-glycosylations of serum CBG reported herein are affecting the PAE-mediated cleavage of the RCL.

Our data point to relatively low levels of RCL *O*-glycosylation in serum CBG from healthy individuals. Given the high levels of RCL *O*-glycosylation observed in HEK293-derived rCBG, it is, however, tempting to speculate that CBG expressed by other tissues (*e.g.*, kidney) and in specific conditions (*e.g.*, during rapid growth/proliferation including in cancer and stem cell development) may exhibit considerably higher levels of RCL *O*-glycosylation that potentially plays regulatory roles in cortisol transport and delivery. Investigations of such putative tissue-specific CBG glycoforms potentially featuring high RCL *O*-glycosylation are now enabled by recent advances in *O*-glycoproteomics and mucin glycobiology (55–57).

In summary, this study reports on the structure–function relationship of complex RCL *N*- and *O*-glycosylation patterns of human CBG in the context of RCL cleavage and cortisol release (see Fig. 7 for summary of findings). We have unveiled previously unknown *N*- and *O*-glycosylation events with site-specific resolution and demonstrated their functional importance in NE-mediated proteolysis. Our study contributes to an improved understanding of mechanisms governing cortisol delivery to sites of inflammation and highlights the complex nature of how glycosylation contributes to CBG-guided delivery of cortisol to inflamed tissues.

Experimental procedures

Key materials

Human CBG (UniProtKB: P08185) was purified from pooled sera of healthy donors to >90% purity as assessed by SDS-PAGE. Polyhistidine-tagged recombinant human CBG expressed in HEK293 cells (rCBG) was from Sino Biological, Inc (catalog number: 10998-H08H). Human NE (P08246) from blood-derived resting neutrophils isolated to >96% purity was from Lee Biosolutions (catalog number: ELA-2342-40). *Arthrobacter ureafaciens* α ,2,3/6/8/9-linkage–sensitive neuraminidase A (broad substrate specificity) recombinantly expressed in *Escherichia coli* was from New England Biolabs (catalog number: P0722). Bovine β 1,3/4-linkage–sensitive

galactosidase was from Prozyme (catalog number: GKX-5013). *Elizabethkingia miricola* PNGase F recombinantly expressed in *E. coli* was from Promega (catalog number: 9PIV483). Modified porcine trypsin was from Promega (catalog number: V5111). Ultra–high-quality water was used for all experiments. All chemicals and LC solvents were from Sigma–Aldrich/Thermo Fisher Scientific unless otherwise specified.

Synthetic RCL peptides with and without *O*-glycosylation

Three RCL peptide variants were synthesized, two of which carried Tn (GalNAc) *O*-glycans at either Thr342 or Thr345, whereas the remaining peptide was nonglycosylated (see later for sequences). To enable the peptide synthesis, 2-chlorotriethylamine resin (1 g, 0.8–1.5 mmol/g, catalog number: 48101-025) was loaded by treating with a solution of Fmoc-Arg(Pbf)-OH (550 μ mol), and a 5 M excess of *i*-Pr₂NEt (2.75 mmol) in 5 ml CH₂Cl₂. The resin was agitated gently on an orbital shaker (18 h), then drained, and washed five times with 5 ml CH₂Cl₂. The derivatized resin was subsequently capped with a mixture of CH₂Cl₂/MeOH/*i*-Pr₂NEt (17:2:1 v/v/v, 4 ml, 1 h) before washing five times sequentially with 3 ml dimethylformamide (DMF), 3 ml dichloromethane, and 3 ml DMF. The amino acid loading was determined to be 0.4 mmol/g. The synthesis of the three RCL peptides was initiated on a 25 μ mol scale. Peptide coupling cycles for standard amino acids involved Fmoc-protected amino acid (4 eq), ethyl cyanohydroxyiminoacetate (Oxyma) (4.4 eq), and *N,N'*-diisopropylcarbodiimide (4 eq) in DMF (0.17 M final concentration for all components after mixing). Coupling steps were performed for 40 min at 40 °C. The resin was then washed five times sequentially with 3 ml DMF, 3 ml CH₂Cl₂, and 3 ml DMF. A capping step was performed after each coupling step, which involved treatment of the resin with a 0.3 M Ac₂O/0.3 M *i*-Pr₂NEt solution in DMF for 5 min. Fmoc-deprotection steps were performed by treatment of the resin twice with a 20% (v/v) piperidine solution in DMF at room temperature for 5 min. Following each coupling, capping, or deprotection step, the resin was washed four times with DMF for 30 s. The underlined residues of the three synthesized RCL (glyco)peptides were double coupled:

RCL peptide 1: H₂N-AVLQLNEEGVDTAGSTGVTLNLTSKPIILR-OH.

RCL peptide 2: H₂N-AVLQLNEEGVDTAGSTGV[T³⁴⁵(GalNAc)]LNLTSKPIILR-OH.

RCL peptide 3: H₂N-AVLQLNEEGVDTAGS[T³⁴²(GalNAc)]GVTLNLTSKPIILR-OH.

The resin-bound peptides synthesized previously were treated with a solution of TFA:*i*-Pr₃SiH:H₂O (90:5:5, v/v/v, 2 ml), and the mixture was agitated at room temperature for 2 h. The resins were filtered and washed with 2 ml TFA. The combined filtrates were concentrated under a stream of nitrogen, and the residue was suspended in diethyl ether and cooled to 0 °C. The precipitate was collected *via* centrifugation. The diethyl ether was decanted leaving behind the crude peptides.

The crude RCL peptide 1 was split, and a third of the material was directly purified by preparative reversed-phase

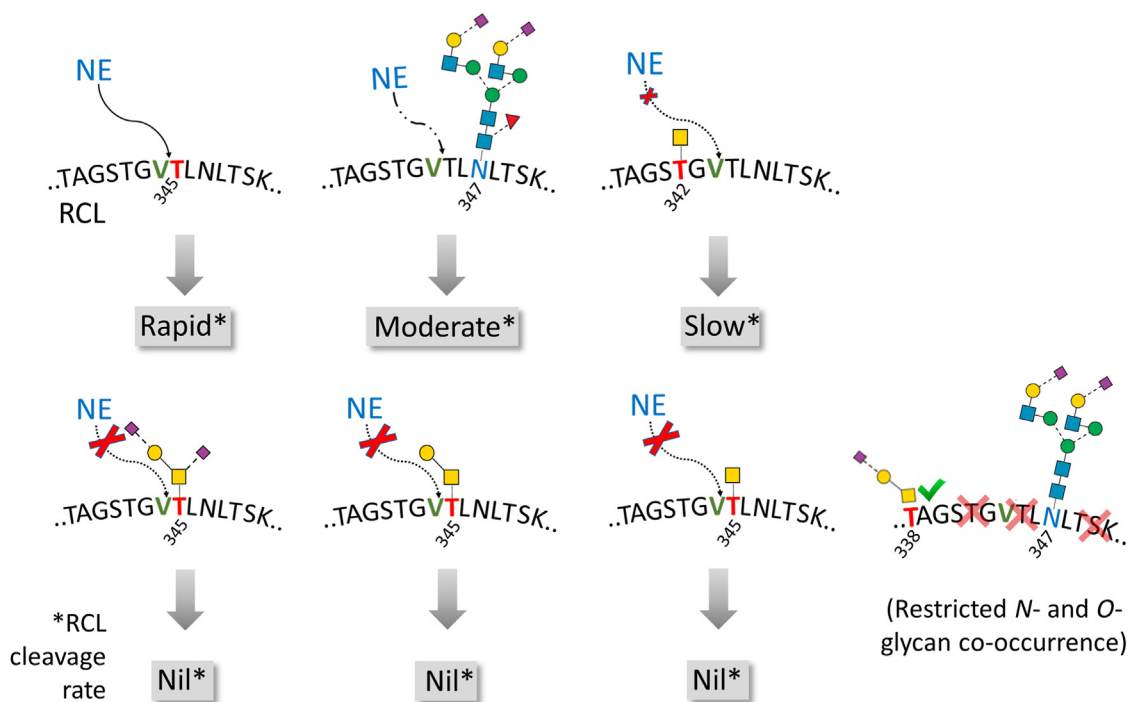


Figure 7. Summary of findings. In this study, we report on newly identified *N*- and *O*-glycosylation events modifying the RCL of human CBG, and we have determined how they impact the NE-mediated RCL cleavage process (*left side*), in turn, leading to an altered cortisol release rate from CBG, and we have established their crosstalk on the RCL (*right side*). See [Figure 1](#) for key. CBG, corticosteroid-binding globulin; NE, neutrophil elastase; RCL, reactive center loop.

HPLC with a gradient of 0 to 60% solvent B (99.9% acetonitrile [ACN] in 0.1% [both v/v] aqueous TFA) in solvent A (aqueous 0.1% [v/v] TFA) over 65 min, on a Waters Symmetry C8 column (5 μ m particle size, 30 mm inner diameter \times 150 mm column length; flow rate 30 ml/min) to afford a highly pure form of RCL peptide 1 as a fluffy white solid after lyophilization (2.6 mg, 0.84 μ mol, 10% overall yield, see [Fig. S20](#) for LC-MS/MS data of RCL peptide 1–3). The crude acetyl-protected RCL peptides 2 to 3 were split, and half of the material had the acetyl groups removed from the GalNAc units by incubating the crude glycopeptide in a solution containing 5% (v/v) hydrazine monohydrate and 25% (v/v) ACN in H₂O (30 mg crude RCL peptide in 4 ml solution) at 37 °C. Deprotection of the acetyl groups was completed after \sim 2 h as demonstrated by UPLC-MS analysis of an aliquot of the reaction mixture. The solution was filtered and directly purified by preparative reversed-phase HPLC as described previously to afford highly pure target glycopeptides as fluffy white solids after lyophilization, that is, RCL peptide 2 (7.7 mg, 2.3 μ mol, 19% overall yield) and RCL peptide 3 (2.5 mg, 0.75 μ mol, 6% overall yield).

Isolation of the Asn347-containing RCL peptide fragment of serum CBG

Serum CBG was incubated with NE in 50 mM sodium acetate, 0.6 M sodium chloride, pH 5.5 at a substrate (CBG)/protease (NE) ratio of 50:1 (w/w) for 10 min at room temperature to ensure exhaustive cleavage of intact CBG and the formation of two well-defined CBG peptide fragments including a low molecular weight (\sim 10 kDa) C-terminal peptide fragment spanning the RCL region and containing Asn347 as the only *N*-glycosylation site (26). The

digestion reaction was quenched by boiling the samples in NuPAGE LDS sample buffer containing 10 mM aqueous DTT for 5 min. The low molecular weight C-terminal peptide fragment of interest was separated from the large N-terminal CBG peptide fragment (and any remaining intact CBG molecules) on a 4–12% precast NuPAGE gel using a 2-(*N*-morpholino)ethanesulfonic acid (1 \times) running buffer for 35 min at 180 V. The gel was stained overnight using Coomassie brilliant blue.

Release and handling of Asn347-glycans of serum CBG

The Asn347-linked *N*-glycans were released directly from the C-terminal peptide fragment of NE-cleaved CBG isolated in a broad SDS-PAGE gel band ([Fig. S1](#)) as described (58). For that purpose, the gel band of interest was excised, cut into 1 mm³ pieces, and washed sequentially three times with 500 μ l ACN and twice with 100 mM aqueous ammonium bicarbonate buffer while shaking (15 min each step). After washing, 30 μ l PNGase F (3 U) in a buffer containing 20 mM aqueous sodium bicarbonate, pH 7.0 was added to the gel pieces. The gel pieces were left for 5 min, covered with \sim 100 μ l additional digestion buffer, and incubated at 37 °C overnight. The released *N*-glycans were extracted from the gel pieces by sonication for 15 min and by repeat water and ACN washes (200 μ l each step). All extractions were combined and dried in a vacuum centrifuge. The released *N*-glycans were resuspended in water and hydroxylated by the addition of 100 mM aqueous ammonium acetate solution, pH 5 for 1 h at 20 °C (59). The *N*-glycans were then dried and subsequently reduced using 1 M sodium borohydride in 50 mM aqueous potassium hydroxide for 3 h at 50 °C. The reduction reaction was quenched using glacial acetic acid. Dual desalting

RCL glycosylation modulates the proteolysis of CBG

of the reduced detached *N*-glycans was performed using, first, SCX resin (AG 50W-X8 Resin; Bio-Rad) and then, second, PGC resin custom-packed as microcolumns on top of C18 discs (catalog number: 66883-U; Supelco) in P10 pipette tips in a solid-phase extraction (SPE) format. The *N*-glycans were eluted from the PGC-SPE microcolumns using 0.05% TFA:40% ACN:59.95% water (v/v/v), dried, and reconstituted in 10 μ l water. Samples were then centrifuged at 14,000g for 10 min at 4 °C and transferred into high recovery glass vials (Waters) for LC-MS/MS analysis. Bovine fetuin was included as a control to ensure efficient sample handling and LC-MS/MS performance.

Exoglycosidase treatment of Asn347-glycans of serum CBG

Aliquots of the released Asn347-linked *N*-glycans of serum CBG were dried in a vacuum centrifuge before digestion with combinations of various exoglycosidases. Digestion reactions were carried out in a total volume of 10 μ l with 2 μ l 250 mM sodium acetate buffer, pH 5.5 (final concentration of 50 mM). Enzymes were used at the following concentrations: 2 μ l α 2,3/6/8/9-linkage-sensitive neuraminidase A (0.5 mU/ μ l) and 2 μ l β 1,3/4-linkage-sensitive galactosidase (0.5 mU/ μ l) at 37 °C for 16 h as described (60). The exoglycosidase-treated *N*-glycan mixtures were desalted using PGC-SPE in preparation for LC-MS/MS analysis as per aforementioned.

Global *N*- and *O*-glycan release from serum CBG and rCBG

The cysteine residues of serum CBG and rCBG were reduced using 10 mM aqueous DTT for 45 min at 56 °C and carbamidomethylated using 40 mM aqueous iodoacetamide (IAA; final concentrations stated) for 30 min in the dark at 20 °C. Serum CBG and rCBG were deposited in discrete spots (5 μ g/spot) on a primed 0.45 μ m polyvinylidene fluoride membrane (Merck-Millipore) and processed as described (59). 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 immobilized proteins using 10 U PNGase F in 10 μ l water/well, 16 h, 37 °C. The released *N*-glycans were transferred into fresh tubes and deaminated by the addition of 10 μ l 100 mM aqueous ammonium acetate, pH 5 for 1 h at 20 °C. The *N*-glycans were then dried and subsequently reduced to alditols using 20 μ l 1 M NaBH₄ in 50 mM aqueous potassium hydroxide for 3 h at 50 °C. The de-*N*-glycosylated proteins on the polyvinylidene fluoride membrane were then treated with 0.5 M NaBH₄ in 50 mM potassium hydroxide for 16 h at 50 °C to release the *O*-linked glycans by reductive β -elimination. The glycan reduction reactions were quenched by the addition of 2 μ l glacial acetic acid to each sample. The *N*- and *O*-glycan samples were desalted using PGC-SPE in preparation for LC-MS/MS analysis as described previously.

Glycan profiling of serum CBG and rCBG using PGC-LC-MS/MS

An established PGC-LC-MS/MS glycomics method (59) was used to separately profile (i) the RCL *N*-glycans of serum

CBG and (ii) *N*- and *O*-glycans from all sites of serum CBG and rCBG.

For (i), the RCL (Asn347-linked) *N*-glycans were analyzed using a HyperCarb KAPPA PGC LC column (3 μ m particle size, 180 μ m inner diameter \times 100 mm column length; Thermo Fisher Scientific) as described (61). Glycans were separated over an 85 min gradient of 0 to 45% (v/v) ACN in 10 mM aqueous ammonium bicarbonate at a flow rate of 2 μ l/min using a 1260 series HPLC system (Agilent) interfaced with an LCD/MSD Trap XCT Ultra (Agilent) mass spectrometer. MS1 spectra were acquired in negative ion polarity mode in the range *m/z* 440 to 2200. The following MS settings were used: drying gas temperature: 300 °C; drying gas flow rate: 6 l/min, nebulizer gas: 12 psi; skimmer, trap drive, and capillary exit: -40 V, -99.1 V, and -166 V, respectively. Ions were detected with ion charge control enabled (targeting 30,000 ions per scan) with a maximum accumulation time of 200 ms. Using data-dependent acquisition (DDA) of the three most intense ions in each MS1 scan, resonance-activation (ion trap)-based collision-induced dissociation (CID) was performed at 35% normalized collision energy (NCE) and a precursor isolation window of 4 *m/z*.

For (ii), *N*- and *O*-glycans liberated from serum CBG and rCBG were globally profiled using LC-MS/MS performed on an LTQ Velos Pro ion trap mass spectrometer (Thermo Fisher Scientific) coupled to a Dionex Ultimate-3000 HPLC system (Thermo Fisher Scientific). All glycans were separated on a PGC-LC capillary column (HyperCarb KAPPA, 5 μ m particle size, 250 Å pore size, 0.18 mm inner diameter \times 100 mm column length, Thermo Fisher Scientific) heated to 50 °C, but slightly different gradients were applied. For the *N*-glycans, the gradient was 8 min at 2.6% solvent B (10 mM ammonium bicarbonate in ACN) in solvent A (10 mM aqueous ammonium bicarbonate), 2 min from 2.6 to 13.5% B, 55 min from 13.5 to 37.3% B, 10 min from 37.3 to 64% B, 1 min from 64 to 98% B, 5 min at 98% B, 1 min from 98 to 2.6% B, and 4 min at 2.6% B. For the *O*-glycans, the gradient was 5 min at 2.8% B, 33 min from 2.8 to 30% B, 5 min from 30 to 98% B, 5 min at 98% B, 2 min from 98 to 2.8% B, and 5 min at 2.8% B.

Both setups used a constant LC flow rate of 3 μ l/min and employed a postcolumn make-up flow configuration by supplying pure isopropyl alcohol at a constant 4 μ l/min flow rate throughout the LC-MS/MS run for enhanced sensitivity (62). The MS1 acquisition scan range was *m/z* 570 to 2000 for *N*-glycans and *m/z* 375 to 1800 for *O*-glycans with a zoom scan resolution of *m/z* 0.25 at full width at half maximum (FWHM). The following MS settings were used for all acquisitions. The source potential was 2.7 kV. Detection was performed in negative ion polarity mode. The automatic gain control (AGC) for the MS1 scans was 5×10^4 with 50 ms maximum accumulation time. For MS/MS, the resolution was *m/z* 0.35 FWHM, the AGC was 2×10^4 , and the maximum accumulation time was 300 ms. Employing DDA, the five most abundant precursors in each MS1 scan were selected for MS/MS using resonance-activation CID at a fixed 33% NCE with an activation Q of 0.250 and an activation time of 10 ms. Dynamic exclusion was enabled with repeat counts set at a maximum of

5, repeat duration at 15 s, and exclusion duration at 30 s. All spectra were acquired in profile mode.

Data analysis of N- and O-glycans of serum CBG and rCBG

DataAnalysis software, version 4.0 (Bruker Daltonics) and Xcalibur, version 2.2 (Thermo Fisher Scientific) were used to inspect, browse, and annotate raw glycan LC–MS/MS data. The glycan identification was aided by RawMeat, version 2.1 (Vast Scientific, www.vastscientific.com), GlycoMod (63), and GlycoWorkBench (64) but otherwise performed by manual *de novo* characterization using relative and absolute PGC-LC retention time, accurate monoisotopic precursor mass, and the CID–MS/MS fragmentation pattern as described (60, 65, 66). The elucidation of glycan fine structural elements was also guided by data obtained from the parallel experiments performed with exoglycosidase treatment. For the relative glycan quantitation, the monoisotopic glycan precursor ions of all identified glycan species in all observed charge states were extracted using DataAnalysis, version 4.0 to enable area under the curve (AUC)–based quantitation. The MIRAGE guidelines were followed for the reporting of the glycomics experiments and resulting data (67).

In vitro O-glycosylation of synthetic RCL peptides

The synthetic nonglycosylated RCL peptide 1 (1 μ g, see aforementioned for synthesis) was mixed with 0.5 mM uridine 5'-diphospho-GalNAc disodium salt (catalog number: U5252; Sigma), 5 mM manganese chloride, 50 mM Tris–HCl buffer, pH 7.4 (all final concentrations), and 1 μ g of either recombinant human polypeptide *N*-acetylgalactosaminyltransferase 2 (GalNAc-T2, catalog number: 7507-GT; R&D Systems) or recombinant human polypeptide *N*-acetylgalactosaminyltransferase 3 (GalNAc-T3, catalog number: 7174-GT; R&D Systems). The solutions were incubated overnight at 37 °C. The next day, the peptides were desalted *via* C18-SPE clean-up and dried prior to LC–MS/MS analysis.

The RCL peptides from the *in vitro* O-glycosylation experiments were analyzed on an Orbitrap Eclipse Tribrid Mass Spectrometer (Thermo Fisher Scientific) coupled to a Dionex Ultimate 3000 RSLCnano HPLC system (Thermo Fisher Scientific). Peptides were separated on a 75 μ m inner diameter \times 150 mm length LC column (3.0 μ m particle size, C18, Nikkyo Technologies). For this purpose, a gradient of 3 to 35% (v/v) solvent B (99.9% ACN containing 0.1% (both v/v) formic acid (FA)) over 30 min in solvent A consisting of 0.1% (v/v) aqueous FA was used at a constant flow rate of 300 nl/min. Full MS1 scans were acquired in the range *m/z* 350 to 2000 using 120,000 resolution, AGC of 5×10^5 ions, and 100 ms maximum injection time. The mass spectrometer was operated in positive ion polarity mode. DDA was used to collect EThcD–MS/MS data within a 3 s cycle time. EThcD–MS/MS of precursors isolated with a quadrupole isolation window of *m/z* 2 was performed with a supplemental activation energy of NCE 25%. Fragment ions were detected in the Orbitrap using 60,000 resolution, AGC of 2×10^5 ions, and a 120 ms maximum injection time.

Longitudinal NE digestion of rCBG and asialo-rCBG

Longitudinal cleavage experiments of rCBG and asialylated rCBG pretreated with 2 μ l neuraminidase A (40 U) were performed overnight at 37 °C. In a set of parallel cleavage reactions, natively folded rCBG and asialo-rCBG were incubated with NE at a substrate (rCBG)/protease (NE) ratio of 25:1 (w/w) for 15, 30, 60, 120, 240, and 480 s on ice. For all reactions, the NE digestion was quenched by incubating the samples in a freshly prepared 1 mM (final concentration) PMSF solution (catalog number: P7626; Sigma–Aldrich) for 8 min on ice. The samples were then heated in 5 mM DTT (final concentration) in 100 mM aqueous ammonium bicarbonate for 5 min at 95 °C and then for 45 min at 56 °C. Samples were carbamidomethylated using 20 mM IAA (final concentration) in 100 mM aqueous ammonium bicarbonate for 30 min at room temperature in the dark, and reactions were quenched by the addition of another 5 mM DTT solution (final concentration) for 15 min in the dark. In solution trypsin digestion was performed using 20:1 (w/w) substrate (CBG)/protease (trypsin) ratio overnight in 100 mM aqueous ammonium bicarbonate at 37 °C. The resulting CBG peptide mixtures were acidified by the addition of 0.1% (v/v) aqueous TFA (final concentration), desalted using home-made Oligo R3 reversed-phase SPE microcolumns (Oligo R3 resin particle size, 30 μ m, catalog number: 1133903; Thermo Fisher Scientific), dried, and redissolved in 0.1% (v/v) aqueous FA prior to LC–MS/MS analysis.

Longitudinal NE digestion of synthetic RCL peptides

The synthetic RCL peptides, that is, the nonglycosylated variant (RCL peptide 1) and the two O-glycosylated variants (Thr345-Tn, RCL peptide 2 and Thr342-Tn, RCL peptide 3, see aforementioned) were dissolved in 100 mM aqueous ammonium bicarbonate to a concentration of 1 μ g/ μ l in preparation for the longitudinal cleavage experiments. In a set of parallel cleavage reactions, each peptide variant was incubated with NE at a substrate (RCL peptide)/protease (NE) ratio of 200:1 (mol/mol) for 15, 30, 60, 120, 240, and 480 s on ice. For all cleavage reactions, the NE digestion was quenched by incubating the samples in 1 mM PMSF solution (final concentration) for 8 min on ice. The samples were then heated in 10 mM DTT (final concentration) for 5 min at 95 °C and left to cool.

The synthetic RCL peptides 1 to 3 were longitudinally profiled using a 4800 Plus MALDI-TOF/TOF mass spectrometer (Sciex). For each cleavage time point, 2 μ l peptides (\sim 180 pmol) were mixed with a 2 μ l matrix solution containing 4 mg/ml α -cyano-4-hydroxycinnamic acid dissolved in 0.1% aqueous TFA, 70% ACN (both v/v) in water, and deposited on a MALDI steel target. MS1 spectra (*m/z* 400–4000) were recorded in positive ion polarity using the reflector mode. In total, 500 individual laser shots were manually acquired per spectrum using a fixed laser intensity set at 4000. Spectra were visualized with the Mass++ tool (68).

Preparation of glycopeptides and de-N-glycosylated peptides of serum CBG

Serum CBG was reduced using 5 mM aqueous DTT for 45 min at 56 °C and then carbamidomethylated using 20 mM

RCL glycosylation modulates the proteolysis of CBG

IAA (final concentrations stated) for 30 min in dark. The alkylation reaction was quenched by the addition of excess DTT. In solution trypsin digestion was performed using a substrate (CBG)/protease (trypsin) ratio of 20:1 (w/w) with overnight incubation in 100 mM aqueous ammonium bicarbonate at 37 °C. Samples were acidified by the addition of 0.1% (v/v) aqueous TFA (final concentration), desalted using home-made reversed-phase C18 SPE microcolumns, and dried. An aliquot was resuspended in 18 μ l 50 mM ammonium bicarbonate and exhaustively de-*N*-glycosylated using 2 μ l PNGase F (20 U) with overnight incubation at 37 °C. The de-*N*-glycosylated peptides were desalted using C18 SPE microcolumns and dried for LC-MS/MS analysis.

LC-MS/MS analysis of glycopeptides from serum CBG and rCBG

Multiple peptide-centric LC-MS/MS experiments were performed to (i) profile RCL glycopeptides and assign RCL *O*-glycosylation sites of serum CBG, (ii) profile rCBG RCL glycopeptides, and (iii) assign RCL *O*-glycosylation sites of rCBG.

- i. For the RCL glycoproteomics and *O*-glycosylation site assignment of serum CBG, mixtures of (glyco)peptides and de-*N*-glycopeptides from serum CBG were analyzed on an Orbitrap Eclipse Tribrid Mass Spectrometer coupled to a Dionex 3500RS UPLC HPLC system (Thermo Fisher Scientific). Peptides were separated on a reversed-phase LC column (75 μ m inner diameter \times 400 mm column length) packed in-house with ReproSil-Pur C18-AQ resin (3 μ m particle size, Dr Maisch). For this, a gradient of 5 to 45% (v/v) solvent B (80% ACN containing 0.1% [both v/v] aqueous FA) over 61 min in solvent A consisting of 0.1% (v/v) aqueous FA was employed at 350 nl/min, 60 °C. The mass spectrometer was operated in positive ion polarity mode. Full MS1 scans were acquired with a scan range of *m/z* 550 to 1750 at a resolution of 120,000 (FWHM at *m/z* 200), standard AGC, and 100 ms maximum injection time. DDA was used to generate MS/MS data following precursor ion selection and isolation of the ten most abundant precursor ions in each MS1 scan using a quadrupole isolation window of *m/z* 2.0. HCD-MS/MS was performed at top speed with 30% NCE with fragment ions detected in the Orbitrap using standard AGC, and 200 ms maximum injection time. Re-isolation and two types of orthogonal fragmentation of select precursor ions were triggered by the presence of at least one of three specific diagnostic sugar oxonium ions, that is, *m/z* 204.0867 (HexNAc), *m/z* 138.0545 (HexNAc), and *m/z* 366.1396 (HexHexNAc) among the ten most intense fragment ions in each HCD-MS/MS spectrum (10 ppm mass tolerance) including (1) EThcD-MS/MS with Orbitrap detection of fragment ions using standard AGC target and a supplemental activation energy of 15% NCE and (2) ion trap CID-MS/MS with standard AGC target and 35% NCE. All MS and MS/MS data except the ion trap CID-MS/MS data were acquired in profile mode.
- ii. For LC-MS/MS profiling of rCBG, (glyco)peptides were loaded on a trap column (0.1 mm inner diameter \times 20 mm

column length) custom packed with ReproSil-Pur C18-AQ 5 μ m resin (Dr Maisch). Approximately 1 μ g of total peptide material was injected per LC-MS/MS run. The peptides were separated at a constant flow rate of 600 nl/min at 45 °C on an analytical column (Reprosil-Pur C18-AQ, 3 μ m particle size, 75 μ m inner diameter \times 250 mm column length, Dr Maisch) operated in reversed-phase mode using an UltiMate 3000 RSLCnano HPLC system (Thermo Fisher Scientific). The mobile phases were 99.9% ACN in 0.1% (both v/v) FA (solvent B) and aqueous 0.1% (v/v) FA (solvent A). The linear gradient of solvent B increased from 2 to 30% (v/v) over 60 min, 30 to 50% (v/v) over 5 min, 50 to 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 polarity mode. Full MS1 scans were acquired in the Orbitrap with an AGC of 3×10^6 and a maximum fill time of 50 ms. MS1 scans were acquired at high resolution of 60,000 (FWHM at *m/z* 200) with a range of *m/z* 350 to 2000. Using DDA, MS/MS data were collected of the 20 most abundant precursor ions in each MS1 full scan utilizing HCD fragmentation with an NCE of 28%. Precursors selected for fragmentations were at least doubly charged ($z \geq +2$). Fragment ions were measured at a resolution of 15,000 (FWHM at *m/z* 200) with an AGC target of 2×10^5 and a maximum injection time of 28 ms using a precursor isolation window of *m/z* 1.2 and a dynamic exclusion of 30 s after a single isolation and fragmentation of a given precursor.

- iii. To assign *O*-glycosylation sites of rCBG, (glyco)peptides from rCBG were analyzed on an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific) coupled to a Dionex 3500RS UPLC system. Peptides were separated on a 75 μ m inner diameter \times 150 mm length LC-column packed in-house (1.9 μ m particle size, ReproSil-Pur C18-AQ resin; Dr Maisch). For this purpose, a gradient of 5 to 45% (v/v) solvent B (80% ACN containing 0.1% (both v/v) aqueous FA) over 61 min in solvent A consisting of 0.1% (v/v) aqueous FA was used running at a constant flow rate of 350 nl/min at 55 °C. Full MS1 scans were acquired in the range *m/z* 350 to 2000 using 120,000 resolution, 5×10^5 AGC, and 100 ms maximum injection time. The mass spectrometer was operated in positive ion polarity mode. DDA was used to collect MS/MS data of the ten most abundant precursors in each MS1 scan. HCD-MS/MS of precursors isolated with a quadrupole isolation window of *m/z* 2 was performed at top speed with an NCE of 30%. Fragment ions were detected in the Orbitrap using an AGC of 2×10^5 and a 120 ms maximum injection time. Re-isolation and two types of orthogonal fragmentation of select precursor ions were triggered by the presence of at least one of three diagnostic oxonium ions, that is, *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 (30 ppm mass tolerance) including (1) EThcD-MS/MS with Orbitrap detection of fragment ions using an AGC target of 2×10^5 and a supplemental

activation energy of 15% NCE and (2) ion trap CID–MS/MS with an AGC target of 1×10^4 and 35% NCE. All MS and MS/MS data except for the ion trap CID–MS/MS data were acquired in profile mode.

Data analysis of glycopeptides from serum CBG and rCBG

The (glyco)peptide LC–MS/MS raw data were browsed using Xcalibur, version 2.2. The HCD– and EThcD–MS/MS data from each sample were searched against the canonical sequence of human CBG (UniProtKB: P08185) or the human proteome (reviewed UniProtKB database, released November, 2021 containing 20,360 entries) using Byonic, version 3.9.4 (Protein Metrics) (30, 69). Decoy and contaminant databases were enabled. The precursor and product ion mass tolerance thresholds were set to 10 and 20 ppm, respectively. Searches were performed with carbamidomethylation of Cys (+57.021 Da) as a fixed modification, and fully specific trypsin cleavage (C-terminal to R/K) with a maximum of two missed tryptic cleavages per peptide was allowed. Searches also included oxidation of Met (+15.994 Da) and Asn/Gln deamidation (+0.9840 Da) that were both considered “common” variable modifications. Searches against *N*-glycosylation of sequon-localized Asn were performed with a predefined *N*-glycan database of 309 mammalian *N*-glycans or using a glycomics-informed *N*-glycan database (70). Searches against *O*-glycosylated Thr/Ser residues were performed with a predefined *O*-glycan database containing nine common mammalian *O*-glycans available within Byonic. Both *N*- and *O*-glycans were considered as “rare” variable modifications. A maximum of two common and up to six rare variable modifications were allowed per peptide. All searches were filtered to <1% false discovery rate at the protein level and 0% at the peptide level. PEP-2D <0.001 was used as a general confidence threshold for the reported glycopeptides. Furthermore, the Byonic-identified glycopeptides were manually inspected with support from *in silico* fragmentation data produced by GPMW, version 10.0 (Lighthouse data, Odense, Denmark) (71). The relative abundance of the observed (glyco)peptides was quantified using AUC measurements based on extracted ion chromatograms generated from the monoisotopic precursor ions of all identified glycopeptide species in all detected charge states using the Skyline software (72). The relative abundance of each glycoform was determined as the proportion of the AUC of the individual glycopeptide forms relative to the total AUC of all glycopeptides spanning each glycosylation site, an approach commonly employed in quantitative glycopeptide analysis (73–75).

The publicly available HCD–MS/MS *O*-glycoproteomics raw datasets of human sera as well as normal and tumorous kidney tissues (PRIDE: PXD009476) (27) were searched against the human proteome (reviewed UniProtKB database, released June, 2022 containing 20,371 entries) using Byonic, version 3.9.4 (30, 69). Decoy database was enabled. The precursor and product ion mass tolerance were 10 and 20 ppm, respectively. Searches were performed with carbamidomethylation of Cys (+57.021 Da) as a fixed modification and fully specific cleavages C-terminal to R/K

(tryptic) and N-terminal to S/T (OpeRATOR) with a maximum of five missed cleavages per peptide. Searches also included oxidation of Met (+15.994 Da) and guanidinylation of Lys (+42.022 Da) that were both considered “common” variable modifications. Searches against *O*-glycosylation of Thr/Ser were performed with an *O*-glycan database containing six *O*-glycans as “rare” variable modifications. A maximum of one common and one rare variable modification were allowed per peptide.

MD simulations of CBG glycoforms

To generate a 3D structure of CBG that could serve as an initial model for advanced MD and glycoprotein structure prediction, a homology model of CBG was generated as described (26). In brief, uncleaved (intact, “stressed”) CBG was created by using two experimental X-ray structures, that is, the RCL-cleaved (“relaxed”) human CBG (Protein Data Bank [PDB] code: 4BB2, 2.48 Å) and the uncleaved (“stressed”) human thyroxine-binding globulin (PDB code: 2CEO, 2.80 Å), a close structural homolog of human CBG, which was used to predict the RCL region of uncleaved CBG. An initial 3D structure of nine realistic glycoforms of serum CBG (Table S5) was generated using the Glycoprotein Builder program in the GLYCAM Molecular Modeling Library (<https://github.com/GLYCAM-Web/gmml/tree/stable/internalPrograms/GlycoproteinBuilder>). All simulations were performed using the Amber16 software suite (76). Using tleap, counterions were added as necessary to neutralize the system, following which the 3D structures were placed in a cubic box of TIP5P water with a 10 Å water buffer. The FF14SB force field was employed with a nonbonded cutoff of 10.0 Å for vdW interactions. Long-range electrostatic interactions were modeled with the PMEMD approximation, using an 8.0 Å cutoff for electrostatics. Initial energy minimization (10,000 steps of steepest descent followed by 10,000 steps of conjugate gradient) was performed with Cartesian restraints (5 kcal/mol) on all solute heavy atoms to optimize the water molecule positions and orientations. The same restraints were employed during heating to 300 K using a Langevin thermostat and during a subsequent 400 ps nPT equilibration. This was followed by a 1 ns structural equilibration phase with Cartesian restraints (5 kcal/mol) only on the protein C α atoms. The atomic positions from the last step of the equilibration were used to start ten replicates (with random initial velocities) of 100 ns production runs, where no restraints were employed.

In silico modeling of the NE–CBG glycoform interactions

Using tleap, 2500 snapshots of each CBG glycoform were extracted at 1 ns intervals from the MD trajectories. Using UCSF Chimera, version 1.14 (build 42094) run from an in-house bash script (Supporting File S1), the 3D structure of human NE in complex with a peptide inhibitor (PDB code: 7CBK) was aligned with the CBG polypeptide backbone by superimposing the C α atoms of the bound inhibitor onto the corresponding atoms in the V344–T345 sequence. Any overlaps between the aligned NE and the glycosylated CBG were then computed. Shapes where the NE alignment caused backbone overlaps with the CBG protein were discarded. The

RCL glycosylation modulates the proteolysis of CBG

percentage of glycosylated CBG snapshots aligned with NE that gave rise to no overlaps was then calculated using an in-house bash script (Supporting File S1). Statistical significance was determined by performing Student's *t* tests for the average number of close contacts for each glycoform.

In silico modeling of the GalNAcT2–CBG glycoform interactions

Using tleap, 1000 snapshots of each CBG glycoform were extracted at 1 ns intervals from the MD trajectories. Using UCSF Chimera, version 1.14 (build 42094) run from an in-house bash script (Supporting File S1), the 3D structure of human GalNAc-T2 in complex with a glycopeptide (PDB code: 5AJP) was aligned with the CBG polypeptide chain by superimposing the N, C α , and C β atoms of residue 5 of the glycopeptide onto the corresponding atoms in Thr338, Thr342, T345, and Ser350 of CBG, respectively. Any overlaps between the aligned GalNAc-T2 and the glycosylated CBG were then computed. The percentage of glycosylated CBG snapshots aligned with GalNAc-T2 that gave rise to no overlaps was then calculated using an in-house bash script (Supporting File S1).

Data visualization and statistics

The number of technical replicates and statistical tests applied (if any) have been stated for each experiment. Data were plotted as the mean, whereas error bars represent their SD. The individual data points have also been shown for all bar graphs.

Data availability

All LC–MS/MS raw data generated for this study have been made available through ProteomeXchange Consortium *via* the PRIDE repository (PXD046405) (77).

Supporting information—This article contains supporting information (27, 78).

Acknowledgments—We thank Prof David Torpy, Dr Emily Meyer, and Dr Ellen Swan for valuable discussions related to this work. Some of the research described herein was facilitated by access to the Australian Proteome Analysis Facility funded under the Australian Government's National Collaborative Research Infrastructure Strategy/Education Investment Fund.

Author contributions—A. C. and M. T.-A. conceptualization; A. C., J. L. A., L. K., J. U., Z. S.-B., R. K., L. C., R. J. P., and M. T.-A. methodology; A. C., J. L. A., O. C. G., L. K., R. J. W., and M. T.-A. formal analysis; A. C., J. L. A., J. U., O. C. G., L. K., and Z. S.-B. investigation; L. C., R. J. P., R. J. W., and M. T.-A. resources; A. C., O. C. G., L. K., and M. T.-A. writing—original draft; A. C., J. L. A., O. C. G., L. K., Z. S.-B., J. U., R. K., L. C., R. J. P., R. J. W., and M. T.-A. writing—review & editing; M. T.-A. supervision; M. T.-A. funding acquisition.

Funding and additional information—A. C. was supported by an International Macquarie Research Excellence Scholarship (iMQRES

2017152). R. K. was supported by an Early Career Fellowship ECF 181259 (Cancer Institute NSW). This research was funded in part from the Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science (grant no.: CE200100012) to R. J. P. R. J. W. gratefully acknowledges support from the National Institutes of Health (grant nos.: 1R24GM136984 and 5R01GM135473) and the National Science Foundation (grant no.: CHE2002628). R. J. W. was supported in part by GlycoMIP, a National Science Foundation Materials Innovation Platform funded through Cooperative Agreement DMR-1933525. M. T. A. is supported by an Australian Research Council Future Fellowship (grant no.: FT210100455). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: ACN, acetonitrile; AGC, automatic gain control; AUC, area under the curve; CBG, corticosteroid-binding globulin; CID, collision-induced dissociation; DDA, data-dependent acquisition; DMF, dimethylformamide; EThcD, electron transfer higher-energy collision dissociation; FA, formic acid; FWHM, full width at half maximum; GalNAc-T, GalNAc transferase; HCD, higher-energy collisional dissociation; HEK293, human embryonic kidney 293 cell line; IAA, iodoacetamide; LC–MS/MS, liquid chromatography–tandem mass spectrometry; MD, molecular dynamics; NCE, normalized collision energy; NE, neutrophil elastase; PDB, Protein Data Bank; PGC, porous graphitized carbon; PNGase F, peptide:N-glycosidase F; rCBG, recombinant CBG; RCL, reactive center loop; SPE, solid-phase extraction.

References

1. 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
2. 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
3. Hammond, G. L., Smith, C. L., Goping, I. S., Underhill, D. A., Harley, M. J., Reventos, J., *et al.* (1987) Primary structure of human corticosteroid binding globulin, deduced from hepatic and pulmonary cDNAs, exhibits homology with serine protease inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 5153–5157
4. Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., *et al.* (2015) Tissue-based map of the human proteome. *Science* **347**, 1260419
5. 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
6. 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
7. Meyer, E. J., Nenke, M. A., Davies, M. L., Chapman, M., Rankin, W., Rushworth, R. L., *et al.* (2022) Corticosteroid-binding globulin deficiency independently predicts mortality in septic shock. *J. Clin. Endocrinol. Metab.* **107**, 1636–1646
8. Meyer, E. J., Nenke, M. A., Rankin, W., Lewis, J. G., Konings, E., Slager, M., *et al.* (2019) Total and high-affinity corticosteroid-binding globulin depletion in septic shock is associated with mortality. *Clin. Endocrinol.* **90**, 232–240
9. Klieber, M. A., Underhill, C., Hammond, G. L., and Muller, Y. A. (2007) Corticosteroid-binding globulin, a structural basis for steroid transport and proteinase-trigged release. *J. Biol. Chem.* **282**, 29594–29603

10. 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**, e52759
11. 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
12. Simard, M., Hill, L. A., Underhill, C. M., Keller, B. O., Villanueva, I., Hancock, R. E. W., et al. (2014) *Pseudomonas aeruginosa* elastase disrupts the cortisol-binding activity of corticosteroid-binding globulin. *Endocrinology* **155**, 2900–2908
13. 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
14. Loke, I., Ostergaard, 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
15. 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
16. Cameron, A., Henley, D., Carrell, R., Zhou, A., Clarke, A., and Lightman, S. (2010) Temperature-responsive release of cortisol from its binding globulin: a protein thermocouple. *J. Clin. Endocrinol. Metab.* **95**, 4689–4695
17. Meyer, E. J., Torpy, D. J., Chernykh, A., Thaysen-Andersen, M., Nenke, M. A., Lewis, J. G., et al. (2020) Pyrexia and acidosis act independently of neutrophil elastase reactive center loop cleavage to effect cortisol release from corticosteroid-binding globulin. *Protein Sci.* **29**, 2495–2509
18. 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
19. Mitchell, E., Torpy, D. J., and Bagley, C. J. (2004) Pregnancy-associated corticosteroid-binding globulin: high resolution separation of glycan isoforms. *Horm. Metab. Res.* **36**, 357–359
20. 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
21. Simard, M., Underhill, C., and Hammond, G. L. (2018) Functional implications of corticosteroid-binding globulin N-glycosylation. *J. Mol. Endocrinol.* **60**, 71–84
22. Hossner, K. 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
23. Avvakumov, G. V., Warmels-Rodenhiser, S., and Hammond, G. L. (1993) Glycosylation of human corticosteroid-binding globulin at asparagine 238 is necessary for steroid binding. *J. Biol. Chem.* **268**, 862–866
24. 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
25. 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**, M111.009100
26. 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 elastase. *J. Biol. Chem.* **291**, 17727–17742
27. Yang, W., Ao, M., Hu, Y., Li, Q. K., and Zhang, H. (2018) Mapping the O-glycoproteome using site-specific extraction of O-linked glycopeptides (EXoO). *Mol. Syst. Biol.* **14**, e8486
28. Chatterjee, S., Kawahara, R., Tjondro, H. C., Shaw, D. R., Nenke, M. A., Torpy, D. J., et al. (2021) Serum N-glycomics stratifies bacteremic patients infected with different pathogens. *J. Clin. Med.* **10**, 516
29. Clerc, F., Reiding, K. R., Jansen, B. C., Kammeijer, G. S. M., Bondt, A., and Wuhler, M. (2016) Human plasma protein N-glycosylation. *Glycoconj. J.* **33**, 309–343
30. Bern, M., Kil, Y. J., and Becker, C. (2012) Byonic: advanced peptide and protein identification software. *Curr. Protoc. Bioinformatics* **Chapter 13**, 13.20.1–13.20.14
31. Daniel, E. J. P., Las Rivas, M., Lira-Navarrete, E., García-García, A., Hurtado-Guerrero, R., Clausen, H., et al. (2020) Ser and Thr acceptor preferences of the GalNAc-Ts vary among isoenzymes to modulate mucin-type O-glycosylation. *Glycobiology* **30**, 910–922
32. 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
33. Hedstrom, L. (2002) Serine protease mechanism and specificity. *Chem. Rev.* **102**, 4501–4524
34. Hammond, G. L. (2016) Plasma steroid-binding proteins: primary gatekeepers of steroid hormone action. *J. Endocrinol.* **230**, R13–R25
35. Sumer-Bayraktar, Z., Nguyen-Khuong, T., Jayo, R., Chen, D. D., Ali, S., Packer, N. H., et al. (2012) Micro- and macroheterogeneity of N-glycosylation yields size and charge isoforms of human sex hormone binding globulin circulating in serum. *Proteomics* **12**, 3315–3327
36. Yang, W., Song, A., Ao, M., Xu, Y., and Zhang, H. (2020) Large-scale site-specific mapping of the O-GalNAc glycoproteome. *Nat. Protoc.* **15**, 2589–2610
37. Yang, S., Onigman, P., Wu, W. W., Sjogren, J., Nyhlen, H., Shen, R. F., et al. (2018) Deciphering protein O-glycosylation: solid-phase chemoenzymatic cleavage and enrichment. *Anal. Chem.* **90**, 8261–8269
38. Riley, N. M., Malaker, S. A., and Bertozzi, C. R. (2020) Electron-based dissociation is needed for O-glycopeptides derived from OPERATOR proteolysis. *Anal. Chem.* **92**, 14878–14884
39. Malaker, S. A., Riley, N. M., Shon, D. J., Pedram, K., Krishnan, V., Dorigo, O., et al. (2022) Revealing the human mucinome. *Nat. Commun.* **13**, 3542
40. Nason, R., Bull, C., Konstantinidi, A., Sun, L., Ye, Z., Halim, A., et al. (2021) Display of the human mucinome with defined O-glycans by gene engineered cells. *Nat. Commun.* **12**, 4070
41. Bagdonaite, I., Thompson, A. J., Wang, X., Søgaard, M., Fougeroux, C., Frank, M., et al. (2021) Site-specific O-glycosylation analysis of SARS-CoV-2 spike protein produced in insect and human cells. *Viruses* **13**, 551
42. Tian, W., Li, D., Zhang, N., Bai, G., Yuan, K., Xiao, H., et al. (2021) O-glycosylation pattern of the SARS-CoV-2 spike protein reveals an “O-Follow-N” rule. *Cell Res.* **31**, 1123–1125
43. Stavenhagen, K., Kayili, H. M., Holst, S., Koeleman, C. A. M., Engel, R., Wouters, D., et al. (2018) N- and O-glycosylation analysis of human C1-inhibitor reveals extensive mucin-type O-glycosylation. *Mol. Cell. Proteomics* **17**, 1225–1238
44. Chau, T. H., Chernykh, A., Kawahara, R., and Thaysen-Andersen, M. (2023) Critical considerations in N-glycoproteomics. *Curr. Opin. Chem. Biol.* **73**, 102272
45. Caval, T., de Haan, N., Konstantinidi, A., and Vakhruhev, S. Y. (2021) Quantitative characterization of O-GalNAc glycosylation. *Curr. Opin. Struct. Biol.* **68**, 135–141
46. Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162
47. King, S. L., Joshi, H. J., Schjoldager, K. T., Halim, A., Madsen, T. D., Dziegiel, M. H., et al. (2017) Characterizing the O-glycosylation landscape of human plasma, platelets, and endothelial cells. *Blood Adv.* **1**, 429–442
48. Darula, Z., Sarnyai, F., and Medzihradsky, K. F. (2016) O-glycosylation sites identified from mucin core-1 type glycopeptides from human serum. *Glycoconj. J.* **33**, 435–445
49. Zhang, Y., Xie, X., Zhao, X., Tian, F., Lv, J., Ying, W., et al. (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
50. Ye, Z., Mao, Y., Clausen, H., and Vakhruhev, S. Y. (2019) Glyco-DIA: a method for quantitative O-glycoproteomics with in silico-boosted glycopeptide libraries. *Nat. Methods* **16**, 902–910
51. Goth, C. K., Vakhruhev, 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
52. Wandall, H. H., Nielsen, M. A. I., King-Smith, S., de Haan, N., and Bagdonaite, I. (2021) Global functions of O-glycosylation: promises and challenges in O-glycobiology. *FEBS J.* **288**, 7183–7212

RCL glycosylation modulates the proteolysis of CBG

53. Wang, S., Foster, S. R., Sanchez, J., Corcilius, L., Larance, M., Canals, M., *et al.* (2021) Glycosylation regulates N-terminal proteolysis and activity of the Chemokine CCL14. *ACS Chem. Biol.* **16**, 973–981
54. King, S. L., Goth, C. K., Eckhard, U., Joshi, H. J., Haue, A. D., Vakhrushev, S. Y., *et al.* (2018) TAILS N-terminomics and proteomics reveal complex regulation of proteolytic cleavage by O-glycosylation. *J. Biol. Chem.* **293**, 7629–7644
55. Narimatsu, Y., Joshi, H. J., Schjoldager, K. T., Hintze, J., Halim, A., Steentoft, C., *et al.* (2019) Exploring regulation of protein O-glycosylation in isogenic human HEK293 cells by differential O-glycoproteomics. *Mol. Cell. Proteomics* **18**, 1396–1409
56. Schjoldager, K. T., Vakhrushev, S. Y., Kong, Y., Steentoft, C., Nudelman, A. S., Pedersen, N. B., *et al.* (2012) Probing isoform-specific functions of polypeptide GalNAc-transferases using zinc finger nuclease glyco-engineered simplecells. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 9893–9898
57. Nielsen, M. L., de Haan, N., Kightlinger, W., Ye, Z., Dabelsteen, S., Li, M., *et al.* (2022) Global mapping of GalNAc-T isoform-specificities and O-glycosylation site-occupancy in a tissue-forming human cell line. *Nat. Commun.* **13**, 6257
58. Royle, L., Radcliffe, C. M., Dwek, R. A., and Rudd, P. M. (2006) Detailed structural analysis of N-glycans released from glycoproteins in SDS-PAGE gel bands using HPLC combined with exoglycosidase array digestions. *Methods Mol. Biol.* **347**, 125–143
59. 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
60. Abrahams, J. L., Packer, N. H., and Campbell, M. P. (2015) Relative quantitation of multi-antennary N-glycan classes: combining PGC-LC-ESI-MS with exoglycosidase digestion. *Analyst* **140**, 5444–5449
61. Abrahams, J. L., Campbell, M. P., and Packer, N. H. (2018) Building a PGC-LC-MS N-glycan retention library and elution mapping resource. *Glycoconj. J.* **35**, 15–29
62. Hinneburg, H., Chatterjee, S., Schirmeister, F., Nguyen-Khuong, T., Packer, N. H., Rapp, E., *et al.* (2019) Post-column make-up flow (PCMF) enhances the performance of capillary-flow PGC-LC-MS/MS-based glycomics. *Anal. Chem.* **91**, 4559–4567
63. 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
64. 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
65. 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
66. 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
67. Kolarich, D., Rapp, E., Struwe, W. B., Haslam, S. M., Zaia, J., McBride, R., *et al.* (2013) The minimum information required for a glycomics experiment (MIRAGE) project: improving the standards for reporting mass-spectrometry-based glycoanalytic data. *Mol. Cell. Proteomics* **12**, 991–995
68. Tanaka, S., Fujita, Y., Parry, H. E., Yoshizawa, A. C., Morimoto, K., Murase, M., *et al.* (2014) Mass++: a visualization and analysis tool for mass spectrometry. *J. Proteome Res.* **13**, 3846–3853
69. Roushan, A., Wilson, G. M., Kletter, D., Sen, K. I., Tang, W., Kil, Y. J., *et al.* (2020) Peak filtering, peak annotation, and wildcard search for glycoproteomics. *Mol. Cell. Proteomics* **20**, 100011
70. Chau, T. H., Chernykh, A., Ugonotti, J., Parker, B. L., Kawahara, R., and Thaysen-Andersen, M. (2023) Glycomics-assisted glycoproteomics enables deep and unbiased N-glycoproteome profiling of complex biological specimens. *Methods Mol. Biol.* **2628**, 235–263
71. Peri, S., Steen, H., and Pandey, A. (2001) Gpmaw – a software tool for analyzing proteins and peptides. *Trends Biochem. Sci.* **26**, 687–689
72. Schilling, B., Rardin, M. J., MacLean, B. X., Zawadzka, A. M., Frewen, B. E., Cusack, M. P., *et al.* (2012) Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation. *Mol. Cell. Proteomics* **11**, 202–214
73. 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
74. Thaysen-Andersen, M., Venkatakrisnan, V., Loke, I., Laurini, C., Diestel, S., Parker, B. L., *et al.* (2015) Human neutrophils secrete bioactive paucimannosidic proteins from azurophilic granules into pathogen-infected sputum. *J. Biol. Chem.* **290**, 8789–8802
75. Tjondro, H. C., Ugonotti, J., Kawahara, R., Chatterjee, S., Loke, I., Chen, S., *et al.* (2020) Hyper-truncated Asn355- and Asn391-glycans modulate the activity of neutrophil granule myeloperoxidase. *J. Biol. Chem.* **296**, 100144
76. Case, D. A., Betz, R. M., Cerutti, D. S., Cheatham, T. E., III, Darden, T. A., Duke, R. E., *et al.* (2016) *AMBER 2016*, University of California, San Francisco
77. Perez-Riverol, Y., Bai, J., Bandla, C., Garcia-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., *et al.* (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552
78. Neelamegham, S., Aoki-Kinoshita, K., Bolton, E., Frank, M., Lisacek, F., Lutteke, T., *et al.* (2019) Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology* **29**, 620–624