



The molecular basis of immunosuppression by soluble CD52 is defined by interactions of N-linked and O-linked glycans with HMGB1 box B

Received for publication, November 18, 2024, and in revised form, February 5, 2025. Published, Papers in Press, February 25, 2025.
<https://doi.org/10.1016/j.jbc.2025.108350>

Nicholas J. DeBono¹ , Silvia D'Andrea², Esther Bandala-Sanchez³, Ethan Goddard-Borger³,
Muhammad A. Zenaidee⁴, Edward S. X. Moh¹, Elisa Fadda^{2,5}, Leonard C. Harrison³, and Nicolle H. Packer^{1,*}

From the ¹ARC Centre of Excellence in Synthetic Biology, School of Natural Sciences, Macquarie University, Sydney, Australia; ²Department of Chemistry, Maynooth University, Maynooth, Ireland; ³The Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia; ⁴Australian Proteome Analysis Facility, Macquarie University, Sydney, New South Wales, Australia; ⁵School of Biological Sciences, University of Southampton, Southampton, United Kingdom

Reviewed by members of the JBC Editorial Board. Edited by Robert Haltiwanger

Human soluble CD52 is a short glycopeptide comprising 12 amino acids (GQNDSQTSSPS) which functions as an immune regulator by sequestering the pro-inflammatory high mobility group box protein 1 (HMGB1) and suppressing immune responses. Recombinant CD52 has been shown to act as a broad anti-inflammatory agent, dampening both adaptive and innate immune responses. This short glycopeptide is heavily glycosylated, with a complex sialylated N-linked glycan at N3 and reported O-linked glycosylation possible on several serine and threonine residues. Previously we demonstrated that specific glycosylation features of CD52 are essential for its immunosuppressive function, with terminal α -2,3-linked sialic acids required for binding to the inhibitory SIGLEC-10 receptor leading to T-cell suppression. Using high resolution mass spectrometry, we have further characterized the N- and O-linked glycosylation of Expi293 recombinantly produced CD52 at a glycopeptide and released glycan level, accurately determining glycan heterogeneity of both N- and O-linked glycosylation, and localizing the site of O-glycosylation to T8 with high confidence and direct spectral evidence. This detailed knowledge of CD52 glycosylation informed the construction of a model system, which we analyzed by molecular dynamics simulations to understand the mechanism of recognition and define interactions between bioactive CD52, HMGB1 and the SIGLEC-10 receptor. Our results confirm the essential role of glycosylation, more specifically hyper-sialylation, in the function of CD52, and identify at the atomistic level specific interactions between CD52 glycans and the Box B domain of HMGB1 that determine recognition, and the stability of the CD52/HMGB1 complex. These insights will inform the development of synthetic CD52 as an immunotherapeutic agent.

Human CD52 is a GPI-anchored glycopeptide composed of 12 amino acids (GQNDSQTSSPS) in its released soluble form (1). Despite its small size, this peptide has

previously been shown to carry both N- and O-linked glycosylation (2, 3). Our previous studies (4) identified soluble CD52 as having broad immunosuppressive properties. Activated T cells release CD52 which, after sequestering HMGB1, binds to the inhibitory sialic acid-binding immunoglobulin (Ig-like) lectin-10 (SIGLEC-10) on T cells (2, 5). HMGB1 is comprised of two structurally identical domains, Box A and Box B; CD52 interacts with Box B (2), which when uncomplexed is pro-inflammatory. The immunosuppressive properties of CD52 identify it as a potential immunotherapeutic agent.

Previous work on the characterization of CD52 has identified two glycans on the peptide, an N-linked glycan at N3, and an O-linked glycan at several possible O-glycosites, at T8, S9 and/or S10 (3). These glycosylations are highly complex in terms of both their heterogeneity and overall glycan structure, and act as the main feature for interaction between CD52, HMGB1-Box B and SIGLEC-10 (3). The structure of CD52 glycans has remained an ongoing analytical challenge, with its complexity and site of O-glycosylation being only partially solved (1, 3, 6). Glycans on CD52 contain a high abundance of α -2,3 sialyl glycoforms, which we confirmed are essential for its immunosuppressive activity (3). However, the relevance of sialylation has not been confirmed for any O-glycosylation present on bioactive CD52. Previously optimized methods for production of recombinant soluble CD52 (4) produce very low yields of the active glycopeptide (3). As glycosylation is required for activity of CD52, the glycoforms of active CD52 need to be fully characterized for development of CD52 as an immunotherapeutic agent. Here, we extend previous work characterizing the recombinant CD52 glycome by defining both the N- and O-linked active rCD52 glycoforms in greater detail, including N- and O-glycan structural characterization, and by determining both the site and occupancy of O-glycosylation. These experimental results then informed modelling to explain at the atomistic level how glycans mediate the interactions between CD52 and HMGB1-Box B, leading to recognition by SIGLEC-10. More specifically, we found through extensive model sampling that hyper-

* For correspondence: Nicolle H. Packer, nicki.packer@mq.edu.au.

Soluble CD52 activity modulated by N- and O-linked glycans

sialylation of both O- and N-glycans on CD52 determines its conformational propensity and allows the glycopeptide to engage with specific contact amino acid residues in HMGB1-Box B. These results explain the preference of CD52 for the HMGB1-Box B domain over the structurally identical Box A. Ultimately, the structure of the CD52/HMGB1-Box B complex we identified through Molecular Dynamics (MD) simulations favors the exposure of the CD52 N-glycans sialylated arms towards the solvent, in a conformation accessible for recognition by SIGLEC-10.

Results

Recombinant CD52 displays glycan heterogeneity and a low abundance of bioactive forms

When soluble CD52-Fc fusion protein, expressed in Expi293 cells, was fractionated by anion exchange chromatography, only a minority of individual fractions exhibited immunosuppressive activity (3). Activity increased with increasing negative charge, before decreasing rapidly (Fig. 1A). The most active fraction, #34, comprised only 0.75% of the total CD52-Fc protein produced.

N-glycosylation on asparagine 3 of CD52-Fc is highly heterogeneous and highly sialylated

After removal of the Fc tag on CD52 by cleavage with Factor Xa (FXa), N-glycans were released from anion exchange fractions of CD52, cleaned and analyzed *via* negative mode PGC ESI-LC tandem mass spectrometry. 192 unique glycan structures were completely or partially determined from 80 glycan compositions using previously reported diagnostic MS2 ions and reported retention time rules of PGC separation (Table S1) (7–12). Almost all N-glycans characterized on CD52-Fc were fucosylated, predominantly with a core fucose motif (Fig. 1, C and D, Table S1). When examining the comparative N-glycome of all fractions, an increase in degree of sialylation could be observed as fractions increased their elution from anion exchange (Fig. 1B), with 24.7% of the N-glycan population of the most active fraction 34 containing at least three sialic acids (Fig. 1C, Table S1). Interestingly, the N-glycome of the relatively inactive fraction 35, collected directly after the most active fraction 34 (Fig. 1A), only carried a 6.9% abundance of N-glycans with at least three sialic acids and contained 45.7% unsialylated N-glycans (Fig. 1C, Table S1). We have previously shown that sialylation, particularly 2,3 linked, is essential for the activity of CD52-Fc (3). Sulfated N-glycans were also observed in CD52 fractions, however the relative abundance of these fractions did not appear to contribute to increased activity of active CD52 (Table S1) as the proportion of sulfated glycans increased with increasing CD52 fraction number, and did not correlate with activity of recombinant CD52 (Fig. 1A). Of the 10 most abundant N-glycan compositions observed on the active rCD52 fraction, seven were sialylated, usually with two or more sialic acids (Fig. 1D). When comparing fractions, the tetra-sialylated composition Hex7HexNAc6Sial4Fuc1 was present in highest

relative abundance in the active fraction 34 (Figs. 1D, S1, Table S1).

Previously, we observed polyLacNAcs on native soluble CD52 N-glycans from spleen (3). Analysis of the tandem mass spectra of released N-glycans from CD52-Fc also showed some polyLacNAc N-glycan extensions (see Fig. S1 for example).

Following FXa cleavage and N-glycan release, O-glycan alditols were chemically removed from CD52 fractions using reductive β -elimination and analyzed using well established PGC-ESI tandem mass spectrometry (7, 13, 14). Eight O-glycan structures were observed from six compositions, with the (Hex2HexNAc2NeuAc2, GlyToucan ID G52120NK, Fig. S2) occupying 55.5% of the O-glycome of the most active fraction 34. Both observed doubly sialylated abundant O-glycans were only present in one isomeric form indicating one type of sialic acid linkage, most likely α -2,3 (Fig. 2B, Table S2). Double sialylation of the O-glycans on CD52 appeared to be conserved across less active, later eluting fractions (Fractions 23–35), with the second most abundant O-glycan as the doubly sialylated mucin-type core 1 structure (HexHexNAc-NeuAc2, GlyToucan ID G63110FE) present at an average 17.6% relative abundance (Fig. 2A, Table S2).

O-glycosylation of rCD52 is localized to threonine 8 with low site occupancy

The site of O-glycosylation was then determined on CD52, after treatment with FXa to remove the Fc fusion and treatment with PNGase F to remove N-glycans, using a glycopeptide diagnostic ion triggered ETHcD LC-MSMS workflow. Treatment with FXa of the CD52-Fc results in a 29-mer glycopeptide (Table S3), which was used in subsequent searches. All fractions were analyzed using the glyco-O-Pair default settings within Fragpipe in MSFragger software (15, 16), with the identified six compositions of the O-glycome as the glycan search space. As previous attempts to confidently identify the O-glycosite of CD52 have resulted in ambiguity (3), we chose a very strict cut off of reported confidence level (level one only) and an O-Pair score (>20) to ensure accurate determination of the O-glycosite. Using these filters within O-Pair, 47 glycoPSMs were identified (Table S4), which were then manually confirmed for the presence of peptide ions fragments on either side of the predicted site of glycosylation. All filtered glycoPSMs localized the site of O-glycosylation to T8 (Fig. 3), with the most highly confident spectral matches occurring with non- or partially-sialylated O-glycopeptides (Table S4). Comparing the proportion of O-glycosylated to non-O-glycosylated CD52, site occupancy at the T8 O-glycosite was calculated to be just 7.5% of analyzed CD52 peptides across all fractions (Table S5). This O-glycosite determination aligns with glycosite predictions by the O-glycosylation predictive software IsoGlyP (17), which predicts T8 as the most likely site for O-glycosylation by all modelled transferases (Table S6). By contrast, the N3 N-glycosite of CD52 was found to be occupied, on average, 99% of the time across all fractions on CD52 when comparing non-N-glycosylated CD52 to the

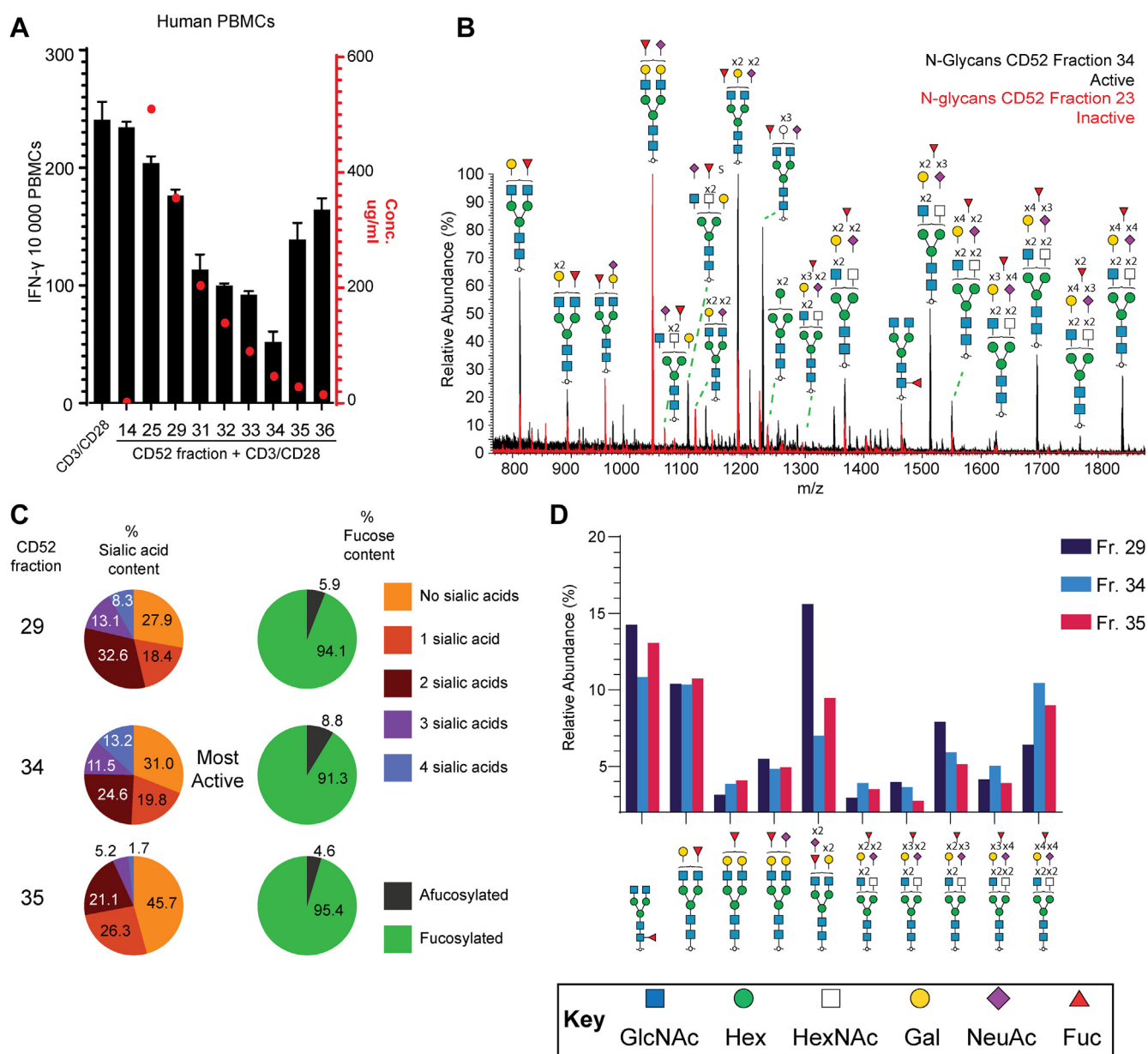


Figure 1. Characterization of the immunosuppressive activity and N-glycan compositions of recombinant CD52-Fc fractions. A, Bars: IFN- γ ELISpot assay of CD52-Fc activity after fractionation by anion exchange chromatography (representative fractions shown). 10,000 peripheral blood mononuclear cells (PBMCs) were incubated in 200 μl per well either with or without anti-CD3/CD28 Dynabeads (1 bead/cell) in the presence of CD52-Fc fractions (F14–F36) at 5 ng/ml. The left Y-axis shows the number of IFN- γ -positive ELISpots, a decrease indicating immunosuppressive activity of CD52-Fc. The red dots correspond to the protein concentration ($\mu\text{g/ml}$) of CD52-Fc in each fraction, measured by Qubit absorbance, with the right Y-axis referencing this data. B, summed MS1 spectrum of the active fraction 34 (black) is overlaid with the spectrum of the most abundant fraction by protein weight, #23 (red), which is inactive. Abundant N-glycan compositions have been annotated, with doubly charged forms of each glycan appearing most commonly. A greater proportion of highly sialylated N-linked compositions can be seen in the active fraction 34. C, % Sialic acid (left) and % fucosylation (right) content of N-glycans in selected anion exchange CD52-Fc fractions. The highly active fraction 34 contains a high abundance of highly sialylated N-glycans, while all fractions are heavily fucosylated. D, the top 10 N-glycan compositions with more than 2.5% relative abundance in active fraction 34, compared to two other anion exchange fractions and organized by increasing m/z (M-H). All compositions have been shown according to SNFG cartoon nomenclature guidelines (48).

deamidated CD52 peptides pre- and post-N-glycan removal by PNGase F digestion (Table S5).

Molecular dynamics simulation reveals an intricate interaction between glycosylated CD52 and HMGB1 box B

We used an approach based on deterministic sampling by molecular dynamics (MD) simulations to understand the role of sialylation on the recognition and binding of the 12 residue

CD52 glycopeptide (aa 25–36 region from the Uniprot P31358 CAMPATH-1 antigen with sequence GQNDTSQTSSPS) to the HMGB1 box B. We first analyzed the effects of sialylation on the conformational equilibrium of the isolated CD52 glycopeptide by running two sets of five uncorrelated MD simulations, in one set the peptide was linked to the fully sialylated N- and O-glycans (GlyToucan ID's G80552MJ and G42089IU respectively) based on the analytical glycoanalysis described above and in the other, the peptide was linked to the

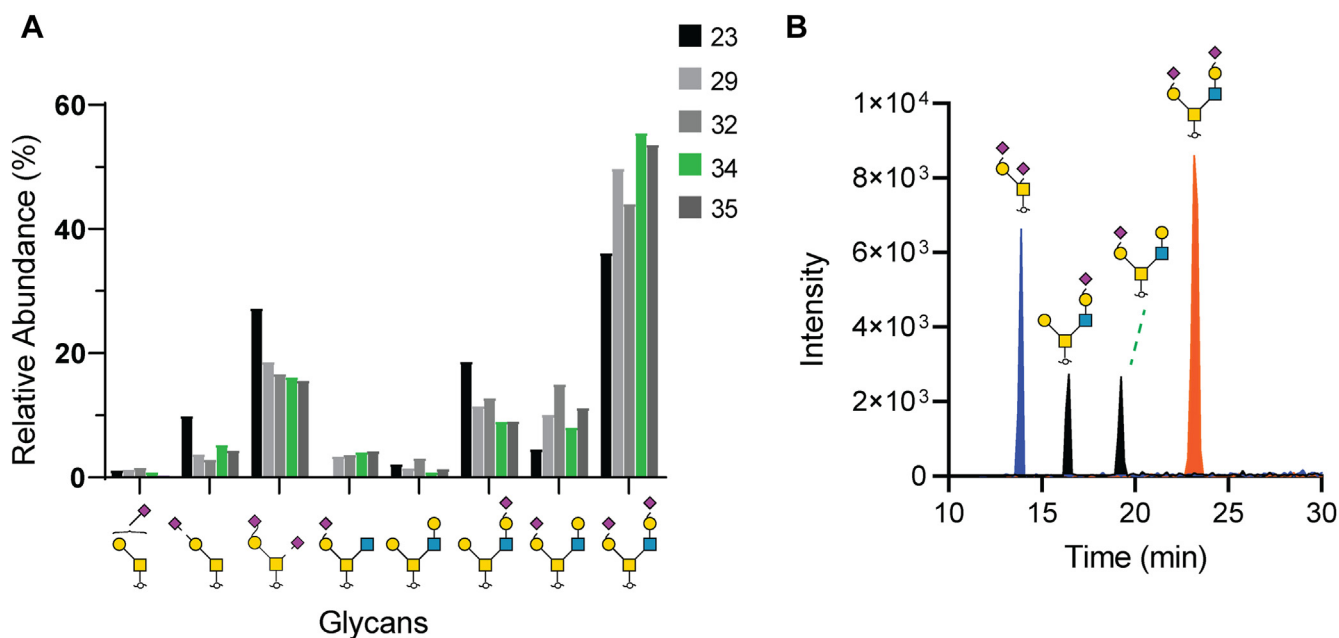
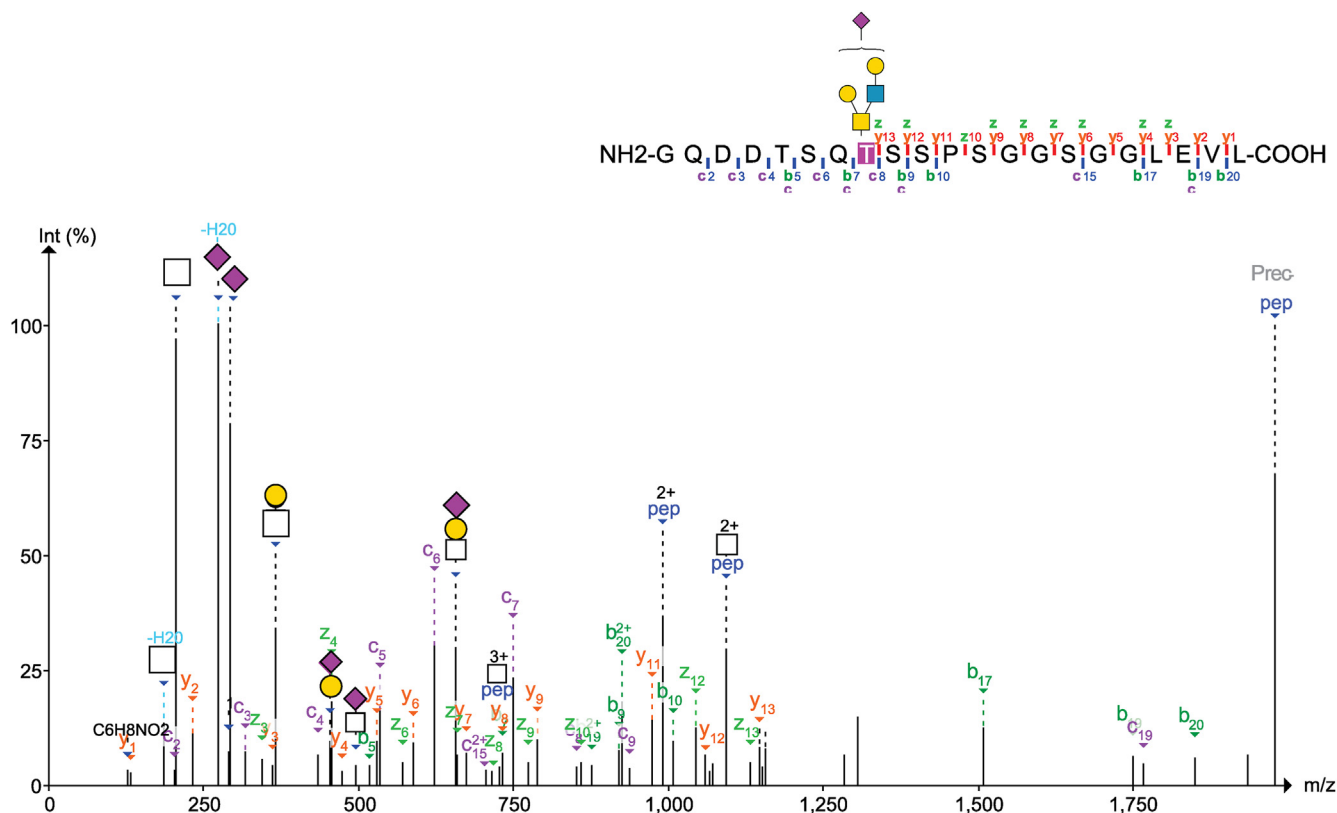


Figure 2. Characterization of O-glycan structures and sialic acid linkage variants in active CD52 fraction 34. A, O-glycan structures identified in specific CD52 fractions, with the most active fraction 34 highlighted in green. Fraction 34 exhibited the highest relative abundance of di-sialylated core 2 O-glycan (GlyToucan ID [G52120NK](#)). B, extracted ion chromatograms displaying the three most intense O-glycan mass compositions in fraction 34. Both doubly sialylated O-glycans appeared as a single isomer, suggesting only one type of sialic acid linkage, while glycans containing one sialic acid appeared as two unique isomers, due to alternative arm sialic acid linkage addition.



same glycan structures but without terminal sialic acids (GlyToucan ID's [G56655CC](#) and [G96017QA](#) respectively) (Fig. 4, A and B). The analysis of the CD52 structures obtained from 10 μ s of cumulative sampling, 5 μ s for each set, showed that hyper-sialylation stabilizes an extended structure of the peptide, using conformational changes in the torsion angles values of the backbone at, or in the immediate vicinity of, the glycosylation sites. More specifically, the Ramachandran analysis of the conformation of the CD52 glycopeptide shows

significant changes in the conformational propensity at bonds 2 and 8, corresponding to the peptide bonds centered at N3 and S9, respectively (Fig. 4, C and D). In the case of N3 (bond 2), the conformational preference changes from an extended β strand conformation in the presence of a core-fucosylated tetra-antennary and fully sialylated N-glycan (GlyToucan ID [G80552MJ](#)) to a 3 to 10 helical turn when the same N-glycan structure does not have terminal sialic acids (GlyToucan ID [G56655CC](#)). The absence of sialylation on the extended core 2

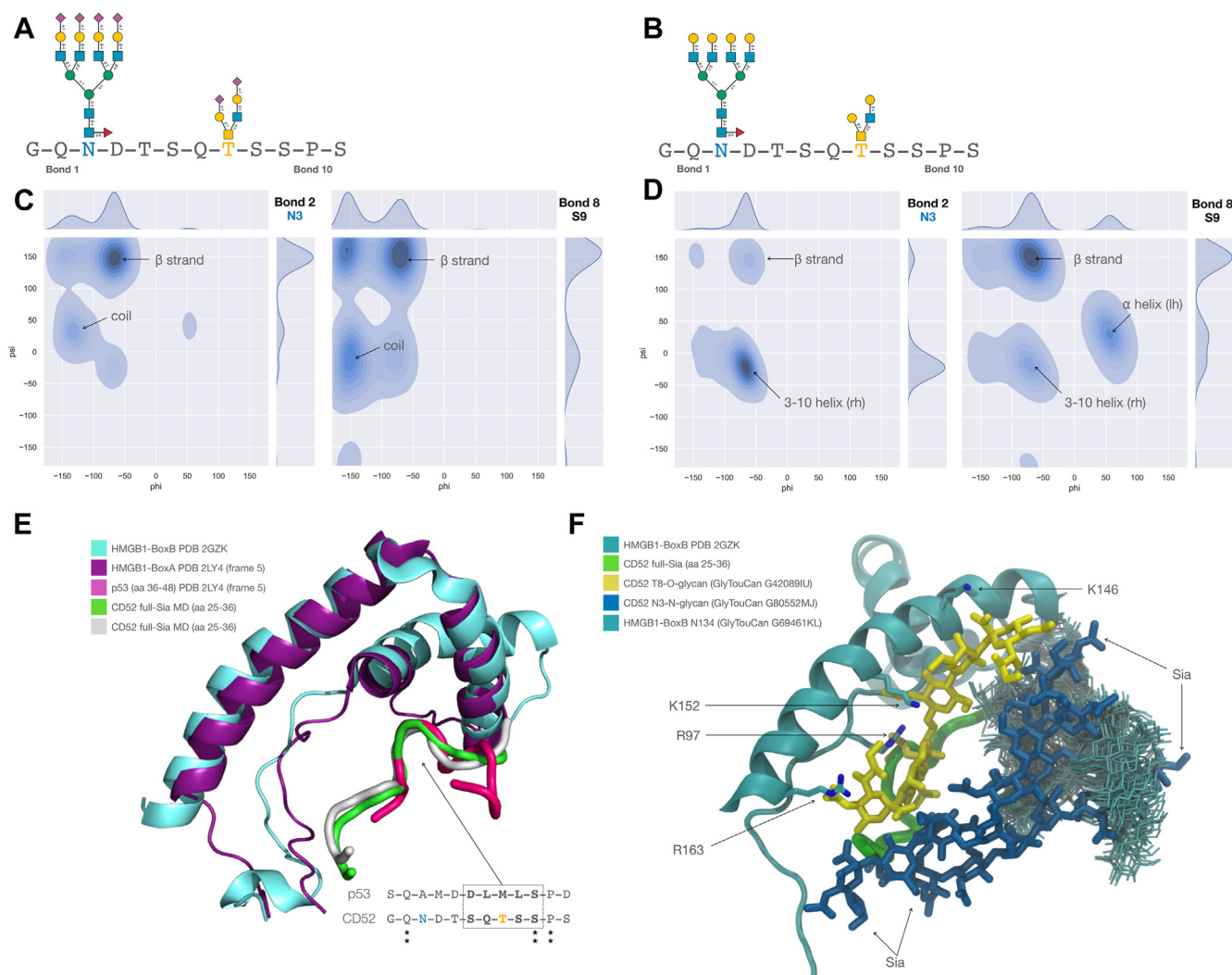


Figure 4. Structural and dynamic analysis of CD52 glycopeptide variants and their interaction with HMGB1 domains. A, sequence of the 12 residue CD52 active peptide (aa 25–36) with the hyper-sialylated glycan structures at N3 (GlyToucan ID [G80552MJ](#)) and T8 (GlyToucan [G42089IU](#)) depicted using SNFG nomenclature. B, sequence of the 12-residue soluble CD52 peptide (aa 25–36) with the non-sialylated glycan structures (GlyToucan ID [G56655CC](#) and [G96017QA](#)) at N3 and T8, respectively. C, Ramachandran plot of peptide bonds 2 and 8, centered at N3 and S9, respectively, obtained from 5 μ s of cumulative MD sampling of the hyper-sialylated CD52 glycopeptide. The secondary structure types associated with torsion angles are indicated. Kernel Density Estimates (KDE) are shown on the top and right of the plot. D, Ramachandran plot of the peptide bonds 2 and 8, centered at N3 and S9, respectively, obtained from the 5 μ s of cumulative sampling of the non-sialylated CD52 glycopeptide, with corresponding KDE values on above and on the right-hand side of the plot. All plots were made with seaborn (<https://seaborn.pydata.org/>). E, structural alignment of the HMGB1 box A (purple cartoon representation, PDB 2LY4) bound to p53 (pink cartoon representation, PDB 2LY4) to the HMGB1 box B (cyan cartoon representation, PDB 2GZK) with two selected representative snapshots (green and white cartoon representation) from the cumulative 5 μ s MD simulation of the hyper-sialylated CD52 glycopeptide. Sequence alignment of the HMGB1 box A-bound p53 peptide to the active CD52 glycopeptide where the sequence corresponding to the helical turn is highlighted in the dotted box. The T8 of CD52 is highlighted in yellow. Graphical rendering and sequence alignment with pymol (<https://pymol.org/2/>). F, snapshot (frame 200 ns) from the 1 μ s MD trajectory of the reconstructed 3D model of HMGB1 box B (cyan) in complex with the hyper-sialylated CD52 glycopeptide (green). The N- and O-glycans are shown with sticks in blue and yellow, respectively. The key protein residues identified to engage in interactions with the O-glycan at T8 are indicated with labels and shown with sticks. The position of the terminal sialic acids of the CD52 N-glycan are also highlighted with labels. The conformation of the HMGB1 box B N-glycan is shown with multiple snapshots collected every 20 frames (1 ns each) through the duration of the trajectory. Graphical rendering with VMD (<https://www.ks.uiuc.edu/Research/vmd/>). All cartoon representations of glycans shown according to SNFG symbol nomenclature guidelines, with key in Figure 1.

Soluble CD52 activity modulated by N- and O-linked glycans

O-glycan (GlyTouCan ID [G96017QA](#)) at T8 stabilizes the localized propensity for a left-handed α helical conformation, which extends to S9. The increased helicity of the CD52 glycopeptide in the absence of sialylation ultimately leads to a more compact structure, as shown also by the analysis of the radii of gyration (R_g) (Fig. S2). The Ramachandran plots for all the peptide bonds in the presence and absence of sialylation are shown in Figs. S4 and S5, respectively.

To our knowledge, at this time there is no available structural information on the complex between CD52 and HMGB1 box B, or on the basis for the preference of the CD52 glycopeptide for box B relative to the structurally identical domain (2), box A, to which CD52 does not bind. To gain insight on the molecular basis for the preferential recognition of hyper-sialylated CD52 glycopeptide for HMGB1 box B, we analyzed the structure of the complex *in silico* between the HMGB1 box A and a section of the p53 transactivation domain (18) (aa 36–48). Although the functions of these two complexes are entirely unrelated, the structural features of the HMGB1 box A/p53 complex provide important clues that ultimately led us to build a working model for our complex. The rationale behind choosing the HMGB1-box A/p53 as a template is based on two main considerations. First, HMGB1 box A and box B share a high structural similarity, with a backbone RMSD value of 2.27 Å over 54 atoms, and a sequence identity of 25.33% for the human HMGB1 (Uniprot P09429) (19) (Fig. 4E). Second, the structural alignment of representative snapshots from the MD simulations of the CD52 hyper-sialylated glycopeptide to the box A-bound p53 conformation indicates a keen similarity between the two peptides around the O-glycosylation site on CD52 at T8 (Fig. 4E). Guided by these considerations, we generated a 3D model of the HMGB1 box B in complex with the CD52 glycopeptide, in which glycosylation on the peptide and the HMGB1 box B at N134 was restored with GlycoShape (20) to complement the tight steric requirements.

The most striking and unique feature of this 3D model is that the main contacts occur between the HMGB1 and the sialylated core 2 O-glycan, which directly faces the box B surface, and not between HMGB1 box B and the CD52 peptide (Fig. 4F). Also, in this model the large tetra-antennary N-glycan at N3 is orientated with the arms directed towards the solvent, and thus accessible for binding (potentially to SIGLEC10 (21)). We assessed the stability of this complex through a single trajectory of MD simulation with a 1 μ s production phase, ran all degrees of freedom unrestrained. The simulation allowed us to identify a set of key residues of HMGB1 box B that make specific contacts with the fully sialylated core 2 at T8, namely R97, K146, K152, R163, see Figure 4F. Interestingly, all of these residues, except K152, are not conserved in HMGB1 box A, where R97 is G, K146 is D and R163 is I, providing a rationale for the preference of the hyper-sialylated peptide for box B and not box A.

Ultimately, aside from other considerations and based on this 3D model of the complex alone, lack of sialylation of the core 2 O-glycan at T8 would prevent binding of CD52 to HMGB1 box B as most contacts with the protein involve

directly the terminal sialic acids (Fig. 4F). The MD simulation also shows that the terminal sialic acids on the CD52 N-glycan are mostly oriented toward the solvent and that the N-glycan on HMGB1 box B wraps around the CD53 N3 N-glycan. As a note of caution, additive force fields tend to enhance glycan-glycan interactions (22), so the extent of these contacts may be overestimated.

Discussion

In this study, we have extended previous work to determine the structure of active immunosuppressive CD52, produced recombinantly (CD52-Fc) (3), further characterizing both the complex, heavily sialylated N- and O-glycome. Glycosylation is a critical requirement for both CD52 activity (3) and SIGLEC-10 interaction (21), in both cases 2,3-sialylation conferring greater bioactivity than 2,6 sialylation. N-glycans identified in active fractions of CD52-Fc were overwhelmingly core fucosylated and heavily sialylated, with N-glycans containing at least two sialic acids comprising 49.3% of the relative abundance of all observed N-glycans on active CD52-Fc. Poly-LacNAcs were also present in recombinant CD52 (Table S1), as in native soluble spleen CD52 (3). Elongated linear and branched polyLacNAc chains are ubiquitously expressed in normal human adults (23), and have been broadly implicated in biological processes (24–28); however, their direct influence on glycan binding is relatively poorly understood due to inability to confidently resolve individual polyLacNAc isomers (24). In the context of soluble CD52, we propose that the presence of linear polyLacNAc structures ending in a terminal sialic acid could effectively enable deeper interaction between CD52 N-glycan sialic acids and the immunoinhibitory SIGLEC-10.

Compared to the N-glycome, the O-glycome observed on CD52 had low heterogeneity, with at least half of all observed O-glycan abundance being the extended, doubly sialylated core 2 O-glycan structure. The low variability of this O-glycome was further tested in the molecular dynamic simulations. Previously, O-glycosylation on CD52-Fc was putatively localized to either T8 or its neighbor, S9 (3). Specific O-glycosylation sites have previously been identified as critical to protein function and activity in other well studied glycoproteins (14, 29–31), and the molecular dynamic modelling informed by experimental data infers this for CD52 binding to HMGB1 box B. High resolution mass spectrometry with paired EThcD fragmentation techniques was able to localize the site of O-glycosylation of soluble CD52-Fc (produced in Expi293 host cells) to T8. Informed dynamic modelling of the soluble CD52 interaction with HMGB1 box B showed that a lack of sialylation of the O-glycan on T8 would prevent binding of the CD52 peptide to HMGB1 box B since contacts with the protein directly involve the anionic terminal sialic acids binding to the cationic amino acids. From this model, the measured low site O-glycan site occupancy of an average of 7.5% O-glycosylated rCD52 then appears to be the limiting factor in the low level of immunosuppressive activity in Expi293 cells. HMGB1-CD52

binding has been reported to be essential for CD52 activity (2) and our modelling of the doubly sialylated extended core 2, α -2,3 sialylated O-glycan on T8 of the soluble CD52 peptide confirmed that these sialic acids are essential for peptide binding to HMGB1 box B, specifically interacting with positively charged amino acids unique to HMGB1 box B, R91, K140, K146, R157 (PDB identification). Of note, we have shown (2) that CD52-Fc does not bind to HMGB1 box A, which is seen by the modelling to assume the same structure as the HMGB1 box B binding cleft, but without having three positively charged amino acids available to interact with the anionic sialic acids in the binding cleft.

Terminal sialic acids of the N-glycans on CD52 were orientated towards the solvent front with the N-glycan on HMGB1 wrapping around the N3 CD52 glycan in the MD simulations. This further implies that the N-glycan of CD52 plays a different role to the O-glycan, with the primary role of the N-glycan being to interact with the downstream immunoregulator SIGLEC-10. While additive force fields like those used here can enhance glycan-glycan interactions (22) overestimating glycan-glycan contacts, these predictions appear to be in accord with what is currently experimentally known of the soluble CD52 interaction.

Here we propose that the HMGB1-CD52 complex formed by the interaction of the doubly α -2,3 sialylated core 2 O-glycan on T8 of the CD52 peptide with the HMGB1 box B binding cleft, forces the highly sialylated CD52 N-glycans at site N3 of CD52 to extend outwards, allowing them to interact with SIGLEC10, thus initiating immunosuppressive activity.

In conclusion, we were able to characterize both the N-glycome and O-glycome of recombinant soluble CD52 using glycopeptide and released glycomics mass spectrometry and have identified the probable most active glycoforms of soluble CD52. The O-glycosylation site of CD52 has been confidently localized to T8 for the first time. This deep characterization of CD52 extends the current knowledge we have on the glycosylation of this therapeutically relevant glycopeptide. Experimental data obtained in this study then informed a molecular dynamics simulation which was able to infer a bi-modal mechanism of action involving both the O-glycans and N-glycans of CD52. The N- and O-glycans are shown to contribute to different interactions, with extended core 2 sialylated O-glycosylation needed to bind to HMGB1 box B, while heavily sialylated N-glycosylation is needed to activate SIGLEC receptors (21), that mediate immune suppression. Our findings further define the structural basis for the bioactivity of CD52 and will assist in the glycoengineered production of CD52 as a therapeutic agent.

Method

Recombinant CD52-Fc production

The CD52-Fc construct was engineered and expressed as previously described DNA encoding the signal peptide and CD52 extracellular domain was joined to that for human IgG1 Fc by PCR, then ligated into the pCAGGS vector for transfection into Expi293 cells. The CD52-Fc fusion protein is not

membrane-bound and is secreted into the extracellular medium. The construct included a flexible GGSGG linker, a cleavage site for Factor Xa protease between the signal peptide and Fc, and a C-terminal Strep-tag II sequence for purification (Table S3). Soluble CD52-Fc was purified from Expi293 cell medium by affinity chromatography on Streptactin resin and eluted with 2.5mM desthiobiotin (5).

MonoQ anion exchange column fractionation and protein quantitation

After recombinant CD52 production, samples were separated as previously described (3). In short, rCD52-Fc was diluted into 5 ml 50mM Tris-HCl, pH 8.3, and applied to a Mono Q column (Mono Q 5/50 GL, 10 μ m particle size, 5 $\times$ 50 mm column dimensions, 1 ml column volume, 2 ml/min flow rate (GE Lifesciences). The column was washed with 10 column volumes of 50mM Tris-HCl, pH 8.3, and then eluted with 50 column volumes of 50mM Tris-HCl, 500mM NaCl, pH 8.3 in 0.5ml fractions. Protein content of eluted fractions was estimated using a Qubit protein assay (ThermoFisher Cat: Q33212).

ELISpot assay of CD52-Fc activity

CD52-Fc MonoQ fractions were tested by interferon gamma (IFN- γ) ELISpot assay to measure immune suppression in peripheral blood mononuclear cells (PBMCs), as previously described (3). After approval by Melbourne Health Human Ethics Committee, blood samples were obtained with informed consent from healthy males. PBMCs were isolated from fresh blood on Ficoll/Hypaque (Cytiva), washed in phosphate-buffered saline (PBS) and re-suspended in Iscove's Modified Dulbecco's medium (IMDM) containing 5% (v/v) pooled, heat-inactivated human serum (PHS; Australian Red Cross, Melbourne, Australia), 100mM non-essential amino acids, 2 mM glutamine, and 50 μ M 2-mercaptoethanol (IP5 medium). PBMCs (10^4 /well) were incubated in 200 μ l IP5 medium in replicates of three in 96-well ELISpot plates (MultiScreen HTS, Millipore) for 18 to 24 h at 37 °C in 5% CO₂ air. Wells had been conditioned by washing with 35% ethanol before being coated with anti-human IFN- γ mAb (10 μ g/ml) in PBS overnight at 4 °C. PBMCs were incubated with anti-CD3/28 Dynabeads (1 bead/cell) $\pm$ CD52-Fc fractions (5 ng/ml). After 24 h, cells were lysed with water and discarded. Wells were washed with PBS between sequential incubations with biotinylated anti-human IFN- γ (1 μ g/ml), streptavidin-alkaline phosphatase (Mabtech) and 5-bromo-4-chloro-3-indolyl-phosphate/nitro blue tetrazolium substrate solution (Mabtech). The color reaction was stopped by the addition of water and IFN- γ spots counted with an AID ELISpot Reader (Autoimmun Diagnostika GmbH).

Removal of Fc fragment from CD52-Fc

CD52-Fc fusion protein (20 μ g) was cleaved with 2 μ l of FXa protease (New England Biolabs) in a volume of 350 μ l of water containing 2 mM of CaCl₂. Following cleavage, samples were applied to a C18 spin tip made from a frit of Empore C18 SPE

Soluble CD52 activity modulated by N- and O-linked glycans

disc (cat #66883U) and CD52 glycopeptide without Fc was eluted with 50 μ l of 85% (v/v) acetonitrile (ACN) and the sample dried in a vacuum concentrator.

Bottom-up MS/MS analysis of purified rCD52 peptide

Dried, purified FXa-cleaved CD52 peptide was resuspended to an approximate concentration of 3 μ g/ μ l equivalent of initial measured concentration and analyzed on an Orbitrap Eclipse Tribrid Mass Spectrometer (Thermo Fisher Scientific) coupled to a Dionex Ultimate 3000 RSLC nano System (Thermo Fisher Scientific). Peptides were separated on an in-house packed C18 column (75 μ m $\times$ 15 cm, 2 μ m particle size, ReproSil-Pur 120 C18-AQ, Dr Maisch). For this purpose, a gradient of (0–2 min—0%, 3 min—2%, 5 min—20%, 25 min—30%, 30–40 min—98%, 43–50 min—2%) solvent B (99.9% ACN containing 0.1% (both v/v) formic acid with solvent A consisting of 0.1% (v/v) aqueous formic acid) was used at a constant flow rate of 300 nL/min. Full MS scans were acquired in the range m/z 350–2000 using 120,000 resolution, Automatic Gain Control (AGC) of 5×10^5 ions and 100 ms maximum injection time. The mass spectrometer was operated in positive ion polarity mode. Data-dependent acquisition (DDA) was used to collect electron transfer higher-energy collision dissociation (ET_hCD)—MS/MS data within a 3 s cycle time. ET_hCD-MS/MS of precursors isolated with a quadrupole isolation window of m/z 2 was performed with a supplemental activation energy of normalized collision energy (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.

PNGase F release and O-glycan release

N-glycans were released from MonoQ fractionated FXa-cleaved CD52 using PNGase F (Promega). Briefly 10 U of PNGase F in MilliQ water was added to dried samples, then incubated at 37 degrees overnight. Following PNGase F release, samples were added to a self-packed C18 spin column made from a frit of Empore C18 SPE disc (cat #66883U) and pre-equilibrated with 0.1% formic acid. The flowthrough containing N-glycans was collected for reduction and analysis of the eluted N-glycans after carbon SPE cleaning, described below. The de-N glycosylated peptides retained on the C18 spin column (de-N-CD52 peptides) were eluted with 50 μ l of 85% ACN, dried, then resuspended in deionized water to an equivalent concentration of 1 μ g/ μ l of rCD52. O-glycans were released from de-N-CD52 peptides using β elimination with 0.5 M NaBH₄ in 50 mM KOH for 16 h. Following release, O-glycans were cleaned as described below, then resuspended for negative mode PGC (porous graphitized carbon) LC-MS/MS analysis, also described below.

PGC cleaning of released N- and O-glycan alditols

Released N- and O-glycans were cleaned as previously described with modifications (32). Briefly, N-glycans were resuspended in 20 μ l of 1 M NaBH₄ in 50 mM KOH and incubated at 50 °C for 3 h. After this incubation, all samples

were diluted to 120 μ l total volume with deionized water, then acidified with 2 μ l of glacial acetic acid. Following acidification, N-glycan and O-glycan samples were both taken for cleaning of carbon solid phase extraction (SPE) as described (32). After cleaning, samples were dried, then resuspended in 9 μ l of MilliQ for negative mode PGC LC-MS/MS analysis.

Negative mode PGC-LC-ESI-MS/MS of released glycan alditols

Eight microlitre of purified released N- and O- glycans was injected onto a Thermo HyperCarb PGC column (3 μ m particle, 1 mm $\times$ 30 mm) using an Agilent 1260 HPLC coupled with a Thermo LTQ Velos Pro linear ion trap mass spectrometer. Glycans were chromatographically resolved over 60 min at a flow rate of 15 μ l/min at 50 °C with the following gradient: Buffer A: 10 mM ammonium bicarbonate, Buffer B: 70% (v/v) acetonitrile in 10 mM ammonium bicarbonate. The gradient parameters were: 0 to 3 min—0% B, 4 min—14% B, 40 min—40% B, 48 min—56% B, 50 to 54 min—100% B, 56 to 60 min—0% B. The mass spectrometer was operated in negative ion mode and configured to perform one full zoom scan MS experiment (HESI source temperature 55 °C, spray voltage 2.75 kV, sheath gas flow 13, auxiliary gas flow 7, capillary temperature 275 °C, scan range 500–2000 m/z , AGC of $3e^4$, 3 microscans, and maximum IT of 100 msec) with the top 5 precursors (dynamic exclusion window of 15 s) selected for MS/MS (scan range 150–2000 m/z , AGC of $1e^4$, maximum IT of 100 msec, isolation window 1.4 m/z , and normalized collision energy set as 33).

Data analysis and search settings

Released glycan analysis was performed manually using the following criteria: m/z signals corresponding to biosynthetically possible glycan compositions according to GlycoMod (33) were selected for area-under-curve (AUC) quantitation using Skyline (34). Each peak area was then expressed as a percentage of the total area of all glycans in the sample. Structural characterization was performed for glycans with MS/MS fragmentation using diagnostic ions (12, 35) as well as known PGC elution patterns of specific glycan features (36, 37).

Glycopeptide samples were searched using both MSFragger (38) (v20.0) with the following search settings: a custom FASTA protein database was used including common decoy contaminants provided from MSFragger and CD52 proteins identified in Supporting Information 4; glycan search list was defined by observed compositions from N-glycan and O-glycan analysis (Tables S1, S2) and a cutoff Pep2D filter of < 0.001 was applied. MS2 search parameters were set according to Thermo Scan headers, and all search tolerances were set to 20 ppm. MSFragger searching used the default Glyco-O-Hybrid workflow with a combined composition list of N-glycans and O-glycans analytically identified (Tables S1, S2). Site occupancy of N- and O- glycosites was calculated using a targeted mass list in Skyline by comparing the AUC of non-glycosylated CD52 fragments to deamidated CD52 fragments (Table S5).

Modelling of interactions between CD52 and HMGB1 box B

To understand the molecular determinants for recognition of the CD52 glycopeptide by HMGB1 box B, and the role of sialic acid in this interaction, a series of all-atoms, classical MD simulations based on deterministic sampling was run, first on the isolated glycopeptide, then on the CD52/HMGB1 box B complex which was built based on information gathered from the simulations of the isolated CD52 glycopeptide. To build the starting structure of the CD52 peptide the aa 25 to 36 region was selected from the Uniprot [P31358](#) CAMPATH-1 antigen (12 residue peptide with sequence GQNDS QTSSPS) from the AlphaFold Protein Structure Database (<https://alphafold.ebi.ac.uk/>). This sequence is recognized and bound by the HMGB1 box B and corresponds to the functional part of the peptide used in the experimental assays in this work. The glycan structures were chosen based on the results of the released glycome and glycoproteome analysis in this work. To represent the fully sialylated CD52 glycopeptide, a core-fucosylated tetra-antennary α -2,3 tetra-sialylated N-glycan (GlyTouCan ID [G80552MJ](#)) was linked at position N3, and a α -2,3 di-sialylated core 2 oligosaccharide (GlyTouCan ID [G42089IU](#)) at position T8. For the reconstruction of the peptide's glycosylation, the glycan 3D structures were sourced from the GlycoShape Glycan Database (<https://glycoshape.org>) and used the ReGlyco tool (20) to link the N- and O-glycans at positions N3 and T8, respectively. Two sets of MD simulations of the isolated CD52 were run, one with the peptide linked to fully sialylated N- and O-glycans and the other where neither oligosaccharide had terminal sialic acids (Fig. 4, A and B). Five uncorrelated MD simulations of 1 μ s production were run for each set, with a cumulative sampling of 10 μ s for the CD52 glycopeptide in the presence and absence of sialic acid. The starting structures described above were used for the first MD simulation in each set, while the remaining trajectories were started from uncorrelated snapshots taken from these MD trajectories.

The system set-up consisted of an initial energy minimization phase of 500,000 steps of steepest descent. The system was then brought at room temperature through a two-step heating scheme in the NVT ensemble of 500 ps each. An integration time step of 2 fs was used throughout. The temperature was raised from 0 K to 100 K during the first step and from 100 K to 300 K during the second step through Langevin dynamics with a friction coefficient (γ_{ln}) set to 1.0 ps⁻¹. The heating phase was followed by a 500 ps NPT equilibration to bring the system into the isothermal-isobaric (NPT) ensemble at 1 atm. Pressure was controlled with a Berendsen barostat in the simulations run with version 2018 of the AMBER package (39) and with the Parrinello-Rahman in the simulations run with version 2021.4-Ubuntu-2021.4-2 of GROMACS (40). Long range electrostatics were treated with Particle Mesh Ewald (PME) with an 11 Å cutoff. Long range dispersion interactions were truncated with an 11 Å cutoff. In all simulations, the GLYCAM_06j-1 force field (41) was used to represent the carbohydrate atoms, the TIP3P model (42)

was used to represent the solvent, and ff14SB parameters were used to represent the peptide and counterions (43). Simulations were run with both AMBER 2018 on GPU (44) and with 2021.4-Ubuntu-2021.4-2 GROMACS (45) on ORACLE Cloud Infrastructure (OCI).

The complex between the fully sialylated CD52 glycopeptide and HMGB1 was obtained by structural alignment of the dominant conformation at equilibrium of the CD52 backbone, obtained from Ramachandran analysis of the backbone, to the 12 aa sequence SQAMDDLMLSPD of the p53 transactivation domain in complex with HMGB1 box A (PDB-ID 2LY4) with pymol (www.pymol.org). This p53 transaction domain complex had been previously modelled and was used for starting structural alignment. After structural alignment of the peptides, the starting structure for the complex was obtained by replacing the structure of HMGB1 box A with the structure of box B from PDB 2GZK, from which we removed the previously modelled bound DNA and the box A domain. The MD simulation of this complex was set-up following the protocol described above and the production trajectory was run for 1 μ s.

Data availability

Raw files relating to released glycans for this work are accessible on GlycoPost (Accession ID GPST000491 with reviewer pin 1669) (46), with glycoproteomic data accessible via ProteomeXchange (PXD056562) (47).

Supporting information—This article contains supporting Information.

Author contributions—N. H. P., L. C. H., E. F., and E. S. X. M. writing—review & editing; N. H. P., L. C. H., E. F., and E. S. X. M. supervision; N. H. P., L. C. H., E. F., and E. D. G. resources; N. H. P. and L. C. H. project administration; N. H. P., N. J. D., E. F., E. S. X. M., E. B., S. D., M. A. Z., and E. D. G. methodology; N. H. P. and L. C. H. funding acquisition; N. H. P., L. C. H., E. F., and E. S. X. M. conceptualization; N. J. D. writing—original draft; N. J. D., L. C. H., E. F., E. B., and S. D. visualization; N. J. D., E. S. X. M., and E. B. validation; N. J. D., E. S. X. M., E. B., and M. A. Z. investigation; N. J. D., E. F., and S. D. formal analysis; N. J. D., L. C. H., E. F., and E. S. X. M. data curation; E. F. and S. D. software.

Funding and additional information—N.D. was supported by a Macquarie University Research Excellence Scholarship (MQRES) and a CSIRO Future Science Platform Top-Up Scholarship. E.S.X.M. was supported by Australian Research Council Centre of Excellence in Synthetic Biology (CE200100029). LCH was supported by a NHMRC Leadership Investigator Grant (APP1173945), a NHMRC Program Grant (APP1150425), Victorian State Government Operational Infrastructure Support and the NHMRC Research Institute Infrastructure Support Scheme. Some of the research described herein was facilitated by access to the Australian Proteome Analysis Facility (APAF) funded under the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS)/Education Investment Fund.

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

Abbreviations—The abbreviations used are: FXa, Factor Xa; MD, molecular dynamics; PGC, porous graphitized carbon; SIGLEC-10, sialic acid-binding immunoglobulin (Ig-like) lectin-10.

References

- Lifely, M. R., Treumann, A., Schneider, P., and Ferguson, M. A. J. (1995) Primary structure of CD52 (*). *J. Biol. Chem.* **270**, 6088–6099
- Bandala-Sanchez, E. G., Bediaga, N., Goddard-Borger, E. D., Ngui, K., Naselli, G., Stone, N. L., et al. (2018) CD52 glycan binds the proinflammatory B box of HMGB1 to engage the Siglec-10 receptor and suppress human T cell function. *Proc. Natl. Acad. Sci.* **115**, 7783–7788
- Shathili, A. M., Bandala-Sanchez, E., John, A., Goddard-Borger, E. D., Thaysen-Andersen, M., Everest-Dass, A. V., et al. (2019) Specific sialoforms required for the immune suppressive activity of human soluble CD52. *Front. Immunol.* **10**, 1967
- Rashidi, M., Bandala-Sanchez, E., Lawlor, K. E., Zhang, Y., Neale, A. M., Vijayaraj, S. L., et al. (2018) CD52 inhibits Toll-like receptor activation of NF- κ B and triggers apoptosis to suppress inflammation. *Cell Death Differ.* **25**, 392–405
- Bandala-Sanchez, E., Zhang, Y., Reinwald, S., Dromey, J. A., Lee, B.-H., Qian, J., et al. (2013) T cell regulation mediated by interaction of soluble CD52 with the inhibitory receptor Siglec-10. *Nat. Immunol.* **14**, 741–748
- Parry, S., Wong, N.-K., Easton, R. L., Panico, M., Haslam, S. M., Morris, H. R., et al. (2007) The sperm agglutination antigen-1 (SAGA-1) glycoforms of CD52 are O-glycosylated. *Glycobiology* **17**, 1120–1126
- 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
- Ashwood, C., Pratt, B., Maclean, B. X., Gundry, R. L., and Packer, N. H. (2019) Standardization of PGC-LC-MS-based glycomics for sample specific glycotyping. *The Analyst*. **144**, 3601–3612
- Abrahams, J. L., Campbell, M. P., and Packer, N. H. (2018) Building a PGC-LC-MS N-glycan retention library and elution mapping resource. *Glycoconjugate J.* **35**, 15–29
- 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
- Harvey, D. J. (2000) Electrospray mass spectrometry and fragmentation of N-linked carbohydrates derivatized at the reducing terminus. *J. Am. Soc. Mass Spectrom.* **11**, 900–915
- Harvey, D. J., Royle, L., Radcliffe, C. M., Rudd, P. M., and Dwek, R. A. (2008) Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption/ionization and negative ion nanospray mass spectrometry. *Anal. Biochem.* **376**, 44–60
- 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
- Nilsson, J., Nilsson, J., Larson, G., and Grahn, A. (2010) Characterization of site-specific O-glycan structures within the mucin-like domain of α -dystroglycan from human skeletal muscle. *Glycobiology* **20**, 1160–1169
- Polasky, D. A., Yu, F., Teo, G. C., and Nesvizhskii, A. I. (2020) Fast and comprehensive N- and O-glycoproteomics analysis with MSFragger-Glyco. *Nat. Methods* **17**, 1125–1132
- Polasky, D. A., Lu, L., Yu, F., Li, K., Shortreed, M. R., Smith, L. M., et al. (2024) Quantitative proteome-wide O-glycoproteomics analysis with FragPipe. *Anal. Bioanal. Chem.* **417**, 921–930
- Mohl, J. E., Gerken, T. A., and Leung, M.-Y. (2021) ISOGLyP: de novo prediction of isoform-specific mucin-type O-glycosylation. *Glycobiology* **31**, 168–172
- Rowell, P., John, S. L., Kathryn, Stott, K., Watson, M., and Thomas, J. O. (2012) HMGB1-Facilitated p53 DNA binding occurs via HMG-box/p53 transactivation domain interaction, regulated by the acidic tail. *Structure* **20**, 2014–2024
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* **7**, 539
- [preprint] Ives, C. M., Singh, O., D'Andrea, S., Fogarty, C. A., Harbison, A. M., Satheesan, A., et al. (2023) Restoring protein glycosylation with GlycoShape. *bioRxiv*. <https://doi.org/10.1101/2023.12.11.571101>
- Forgione, R. E., Di Carluccio, C., Guzmán-Caldentey, J., Gaglione, R., Battista, F., Chiodo, F., et al. (2020) Unveiling molecular recognition of sialoglycans by human siglec-10. *iScience* **23**, 101231
- Fadda, E. (2022) Molecular simulations of complex carbohydrates and glycoconjugates. *Curr. Opin. Chem. Biol.* **69**, 102175
- van den Eijnden, D. H., Koenderman, A. H., and Schiphorst, W. E. (1988) Biosynthesis of blood group i-active polylectosaminoglycans. *J. Biol. Chem.* **263**, 12461–12471
- Vos, G. M., Wu, Y., van der Woude, R., de Vries, R. P., and Boons, G.-J. (2024) Chemo-enzymatic synthesis of isomeric I-branched polylectosamines using traceless blocking groups. *Chem. – A Eur. J.* **30**, e202302877
- Ujita, M., McAuliffe, J., Hindsgaul, O., Sasaki, K., Fukuda, M. N., and Fukuda, M. (1999) Poly-N-acetylglucosamine synthesis in branched N-glycans is controlled by complementary branch specificity of i-extension enzyme and β 1,4-galactosyltransferase. *J. Biol. Chem.* **274**, 16717–16726
- Kadirvelraj, R., Yang, J. Y., Kim, H. W., Sanders, J. H., Moremen, K. W., and Wood, Z. A. (2021) Comparison of human poly-N-acetylglucosamine synthase structure with GT-A fold glycosyltransferases supports a modular assembly of catalytic subsites. *J. Biol. Chem.* **296**, 100110
- Lee, C.-G., Oh, M. J., Park, S.-Y., An, H. J., and Kim, J. H. (2018) Inhibition of poly-LacNAc biosynthesis with release of CMP-Neu5Ac feedback inhibition increases the sialylation of recombinant EPO produced in CHO cell. *Sci. Rep.* **8**, 7273
- Stanley, P., Wuhler, M., Lauc, G., Stowell, S. R., and Cummings, R. D. (2022) Structures common to different glycans in essentials of Glycobiology. In: Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., et al. eds. *Cold Spring Harbor Laboratory Press Copyright © 2022 the Consortium of Glycobiology Editors*, Cold Spring Harbor (NY), La Jolla, California: 165–184
- Chernykh, A., Abrahams, J. L., Grant, O. C., Kambanis, L., Sumer-Bayraktar, Z., Ugonotti, J., et al. (2023) Position-specific N- and O-glycosylation of the reactive centre loop impacts neutrophil elastase-mediated proteolysis of corticosteroid-binding globulin. *J. Biol. Chem.* **300**, 105519
- Stalnaker, S. H., Hashmi, S., Lim, J.-M., Aoki, K., Porterfield, M., Gutierrez-Sanchez, G., et al. (2010) Site mapping and characterization of O-glycan structures on B1-dystroglycan isolated from rabbit skeletal muscle. *J. Biol. Chem.* **285**, 24882–24891
- Liu, D., Wang, S., Zhang, J., Xiao, W., Miao, C. H., Konkole, B. A., et al. (2021) Site-specific N- and O-glycosylation analysis of human plasma fibronectin. *Front. Chem.* **9**, 691217
- Moh, E. S. X., Dalal, S., DeBono, N. J., Kautto, L., Wongtrakul-Kish, K., and Packer, N. H. (2024) SSSMuG: same sample sequential multi-glycomics. *Anal. Chem.* **96**, 3025–3033
- 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
- MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., et al. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968
- Ashwood, H. E., Ashwood, C., Schmidt, A. P., Gundry, R. L., Hoffmeister, K. M., and Anani, W. Q. (2021) Characterization and statistical modeling of glycosylation changes in sickle cell disease. *Blood Adv.* **5**, 1463–1473
- Anugraham, M., Jacob, F., Nixdorf, S., Everest-Dass, A. V., Heinzelmann-Schwarz, V., and Packer, N. H. (2014) Specific glycosylation of membrane proteins in epithelial ovarian cancer cell lines: glycan structures reflect gene expression and DNA methylation status. *Mol. Cell Proteomics* **13**, 2213–2232

37. Pabst, M., and Altmann, F. (2011) Glycan analysis by modern instrumental methods. *Proteomics* **11**, 631–643
38. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D., and Nesvizhskii, A. I. (2017) MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* **14**, 513–520
39. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., DeBolt, S., *et al.* (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules computer. *Phys. Commun.* **91**, 1–41
40. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., *et al.* (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25
41. Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., *et al.* (2008) GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. J. Comput. Chem.* **29**, 622–655
42. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935
43. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theor. Comput.* **11**, 3696–3713
44. Lee, T.-S., Cerutti, D. S., Mermelstein, D., Lin, C., LeGrand, S., Giese, T. J., *et al.* (2018) GPU-accelerated molecular dynamics and free energy methods in Amber18: performance enhancements and new features. *J. Chem. Inf. Model.* **58**, 2043–2050
45. Kutzner, C., Knip, C., Cherian, A., Nordstrom, L., Grubmüller, H., de Groot, B. L., *et al.* (2022) GROMACS in the cloud: a global supercomputer to speed up alchemical drug design. *J. Chem. Inf. Model.* **62**, 1691–1711
46. Watanabe, Y., Aoki-Kinoshita, K. F., Ishihama, Y., and Okuda, S. (2021) GlycoPOST realizes FAIR principles for glycomics mass spectrometry data. *Nucleic Acids Res.* **49**, D1523–D1528
47. Vizcaino, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Ríos, D., *et al.* (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32**, 223–226
48. 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