OXFORD

Site-specific glycosylation analysis of epidermal growth factor receptor 2 (ErbB2): exploring structure and function toward therapeutic targeting

Naoki Fujitani[1,*] , Yasuaki Uehara[1,2], Shigeru Ariki[1,3], Ukichiro Hashimoto[1], Jo Mukai[1], Yoshihiro Hasegawa[1,2], Motoko Takahashi[1,*]

[1]Department of Biochemistry, Sapporo Medical University School of Medicine, S1W17, Chuo-ku, Sapporo 060-8556, Japan, [2]Department of Respiratory Medicine and Allergology, Sapporo Medical University School of Medicine, S1W17, Chuo-ku, Sapporo 060-8556, Japan, [3]Department of Chemistry, Sapporo Medical University Center for Medical Education, S1W17, Chuo-ku, Sapporo 060-8556, Japan

*Corresponding author: Department of Biochemistry, Sapporo Medical University School of Medicine, S1W17, Chuo-ku, Sapporo 060-8556, Japan.
Email: fujitani@sapmed.ac.jp and takam@sapmed.ac.jp

Glycans found on receptor tyrosine kinases (RTKs) have emerged as promising targets for cancer chemotherapy, aiming to address issues such as drug resistance. However, to effectively select the target glycans, it is crucial to define the structure and function of candidate glycans in advance. Through mass spectrometric analysis, this study presents a "glycoform atlas" of epidermal growth factor receptor 2 (ErbB2), an RTK targeted for the treatment of ErbB2-positive cancers. Our analysis provides an in-depth and site-specific glycosylation profile, including both asparagine- and serine/threonine-linked glycosylation. Molecular dynamics simulations of N-glycosylated ErbB2 incorporating the identified glycan structures suggested that the N-glycan at N124 on the long flexible loop in the N-terminal region plays a role in stabilizing the ErbB2 structure. Based on the model structures obtained from the simulations, analysis employing an ErbB2 mutant deficient in N-glycosylation at N124 exhibited a significantly shorter intracellular half-life and suppressed autophosphorylation compared to wild-type ErbB2. Moreover, a structural comparison between the N-glycosylated forms of ErbB2 and its structurally homologous receptor, epidermal growth factor receptor (EGFR), demonstrated distinct variations in the distribution and density of N-glycans across these two molecules. These findings provide valuable insights into the structural and functional implications of ErbB2 glycosylation and will contribute to facilitating the establishment of glycan-targeted therapeutic strategies for ErbB2-positive cancers.

Key words: breast cancer; ErbB2; glycosylation; HER2; mass spectrometry.

Introduction

The ErbB family members, including epidermal growth factor receptor (EGFR), ErbB2 (frequently referred to as HER2), ErbB3, and ErbB4, are representative receptor tyrosine kinases (RTKs) involved in cell growth, survival, migration, and differentiation (Citri and Yarden 2006; Lemmon and Schlessinger 2010). Aberrant gene amplification, hyperactivation, and mutation of these proteins are closely associated with tumorigenesis and cancer progression (Yarden and Sliwkowski 2001). Consequently, various antitumor agents such as tyrosine kinase inhibitors, monoclonal antibodies, and antibody-drug conjugates, have been extensively developed and utilized in the clinical setting, particularly for targeting EGFR and ErbB2 (Gschwind et al. 2004; Hynes and Lane 2005; Yarden and Pines 2012). However, despite the high tumor response rates with first-line treatment, cancer cells almost inevitably acquire resistance to ErbB-targeted therapies, eventually leading to therapeutic limitations (Kumagai et al. 2021). Furthermore, aberrant ErbB signaling has been found to contribute to immune evasion, compromising the efficacy of cancer immunotherapies involving immune checkpoint inhibitors (Kumagai et al. 2021).

To overcome drug resistance and advance cancer treatment strategies, there is growing interest in exploring glycans on proteins as promising therapeutic targets. Glycans on proteins play essential roles in protein quality control and regulation of protein function, and are also closely related to diseases, including cancers (Moremen et al. 2012; Takahashi et al. 2016; Varki 2017; Reily et al. 2019). Targeting glycans holds the potential to address drug resistance resulting from genetic mutations, as glycans are synthesized in a non-template-driven manner, independent of genetic information. Consequently, research targeting the glycans of ErbB family is gaining steam with the aim of developing more effective cancer therapies (Duarte et al. 2022). Inhibitors that target asparagine-linked glycosylation (N-glycosylation), such as tunicamycin and NGI-1, have also shown promise in impairing ErbB signaling and suppressing cancer cell growth, highlighting the potential of ErbB family glycans as therapeutic targets (Contessa et al. 2008; Lopez Sambrooks et al. 2018).

Site-specific glycan analysis has become a practical approach for elucidating the structure and function of glycans. Glycoproteomics of the ErbB family, particularly EGFR and ErbB2, have unveiled their site-specific N-glycan structures. For EGFR, representative site-specific N-glycan structures have been shown for EGFR from lung cancer cell lines, secreted EGFR from skin cancer cell lines, and recombinant EGFR (Wu et al. 2006; Hanash et al. 2008; Liu et al. 2011;

Hasegawa et al. 2015). These studies have shed light on the regulation of EGFR dimer formation, activation, and proposed cancer-specific glycan markers from a N-glycan perspective. In the case of ErbB2, recent site-specific N-glycan structural analysis of ErbB2 from ErbB2-positive gastric cancer cells has revealed the involvement of $\alpha2,6$-sialic acids on ErbB2 N-glycans in regulating the sensitivity of gastric cancer cells to trastuzumab, an anti-ErbB2 antibody (Duarte et al. 2021). Site-specific N-glycan analysis of RTKs other than the ErbB family have also revealed the relationship between N-glycosylation and ligand-binding capacity, as observed in vascular endothelial growth factor receptor 2 (VEGFR2) (Chandler et al. 2017, 2019). More recently, our site-specific N-glycan analysis of the hepatocyte growth factor receptor (MET) revealed that N-glycosylation of MET is involved in intracellular processing (Saitou et al. 2022). Site-specific glycoform analysis of proteins has also contributed to the elucidation of glycan functions and drug discovery beyond RTKs, including the SARS-CoV-2 spike protein and its receptor, angiotensin converting enzyme 2, as well as integrin $\alpha2$ (Casalino et al. 2020; Shajahan et al. 2020, 2021; Watanabe et al. 2020; Huang et al. 2021).

ErbB2 is a key target RTK in the treatment for ErbB2-positive cancers, which are frequently found in breast, stomach, colon, bladder, and biliary cancers (Meric-Bernstam et al. 2019). The emergence of drug resistance remains a pressing challenge in the treatment of ErbB2-positive cancers (Meric-Bernstam et al. 2019). Among the members of ErbB family, ErbB2 stands out as an orphan receptor capable of ligand-independent activation, distinguishing it from other family members that require ligands for activation (Klapper et al. 1999; Yarden and Sliwkowski 2001; Citri and Yarden 2006). The extracellular region of the ErbB family is divided into four conformational subdomains (I to IV) that contain potential sites for N-glycosylation, consisting of an -Asn-X-Ser/Thr-sequence (sequon). The absence of sequons in subdomain III of ErbB2 is quite distinctive (Fig. S1) and may be responsible for the unique activation mechanism of ErbB2.

Unveiling the structure and function of ErbB2 glycans in a site-specific manner and cataloging potential ErbB2 glycan structures would provide valuable insights, not only for assessing whether glycans are specific to cancer in the glycoproteomics of ErbB2 derived from cancer cells but also for leading the evaluation of their potential as therapeutic or diagnostic targets. Currently, however, the identification of glycan structures of RTKs through glycoproteomics remains to the representation of predominant glycans in the majority of analyzed cases, with limited insights into site-specific functions. In this study, we therefore performed an in-depth and site-specific analysis of N-glycans and serine/threonine-linked glycans (O-glycans) on recombinant ErbB2 to produce a "glycoform atlas" that serves as a valuable reference for potential ErbB2 glycan structures. We also attempted to analyze glycan analysis of ErbB2 in ErbB2-positive breast cancer cells, comparing with recombinant ErbB2 and extracting characteristic glycans in cancers. Molecular dynamics simulation of glycosylated ErbB2 incorporating the identified glycans and analysis employing glycan-deficient mutant of ErbB2 revealed a N-glycosylation site responsible for stabilizing the conformation of ErbB2. Additionally, we compare the glycosylated structures of ErbB2 and EGFR, shedding light on the functional differences between these receptors from the perspective of N-glycosylation. We believe that this study will contribute
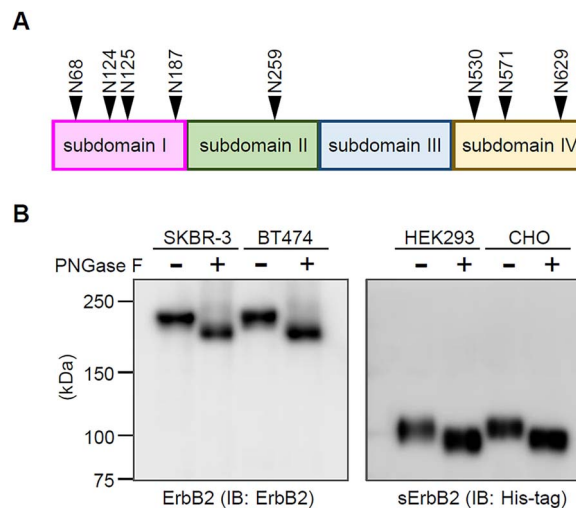

Fig. 1. N-glycosylation of ErbB2 from breast cancer cells and recombinant proteins of the ErbB2 extracellular region (sErbB2). A) Schematic diagram showing the location of sequons in the ErbB2 extracellular region. B) Western blot analysis showing shift in mobility of ErbB2 from cancer cells (left) and of recombinant sErbB2 (right) following PNGase F treatment.

to the acceleration of glycan-targeted drug development by providing detailed insights into the site-specific structure and function of the ErbB2 glycans.

Results

N-glycosylation status of ErbB2

The extracellular region of ErbB2 possesses eight sequons (Fig. 1A), suggesting that ErbB2 is highly N-glycosylated. To confirm the N-glycosylation status of ErbB2, lysates from ErbB2-positive breast cancer cell lines SKBR-3 and BT474, in which ErbB2 is highly expressed, were treated with peptide N-glycosidase F (PNGase F). The mobility shifts of the ErbB2 band in PNGase F-treated samples were clearly displayed in Western blotting, indicating that ErbB2 was N-glycosylated in both breast cancer cells (Fig. 1B). Since N-glycosylation is an event in the extracellular region of the protein, the His-tagged extracellular region of ErbB2 (soluble form of ErbB2; sErbB2) was expressed as a recombinant protein in HEK293 and CHO cells to assess whether these proteins were also N-glycosylated. Upon PNGase F treatment, sErbB2 from both host cells also displayed downshifts similar to those of ErbB2 from cancer cells, indicating that sErbB2 is also N-glycosylated (Fig. 1B).

Site-specific N-glycosylation analysis of sErbB2

To determine the N-glycoforms and of sErbB2 in a site-specific manner, we performed glycoproteomic analysis for purified sErbB2s using electrospray ionization mass spectrometry coupled with nanoflow-based liquid chromatography (nLC-ESI MS). Trypsin digestion covered all potential glycopeptides. The glycoforms in each glycopeptide were identified through a combination of manual analysis and Glyco-Decipher software (Fang et al. 2022), as detailed in the Materials and Methods section. A total of 96 N-glycoforms were identified in both HEK293- and CHO-expressed sErbB2 (Table S1). Figure 2A represents a representative high-energy collision-induced dissociation (HCD) MS/MS spectrum, demonstrating the identification of N-glycan structure at N530 of sErbB2 expressed

Table 1. Summary of the site-specific profiles of N-glycans on sErbB2.

sites	Most abundant N-glycoform		unmodified ratio (%)		fucosylation ratio (%)		sialylation ratio (%)	
	HEK293	CHO	HEK293	CHO	HEK293	CHO	HEK293	CHO
N68			0.0	0.0	96.7	96.0	2.9	41.8
N124/125			0.0	0.0	99.1	96	15.9	71.2
N187			74.5	84.5	100.0	96.7	30.2	49
N259			0.8	3.0	0.0	0.0	0.0	0.0
N530			17.5	6.5	55.7	21.4	10.6	43.3
N571			0.0	0.3	100.0	98.5	36.7	76.9
N629			8.7	24.7	100.0	99.9	30.1	67.7

Monosaccharide symbols follow the SNFG system (Varki et al. 2015).

in HEK293 cells. The MS/MS data for all identified glycopeptides form sErbB2 are summarized in Table S2. N-glycan profiles based on precursor ion intensity were summarized in Fig. 2B and Fig. S2. The glycoforms shown here are structures estimated from biosynthetic pathways, and differences in structural topology have not been experimentally determined. Table 1 provides an overview of the primary N-glycoforms. HCD MS/MS spectra of glycopeptides for identifying the most abundant N-glycan at each site of HEK293-expressed sErbB2 are displayed in Fig. S3. For the N124 and N125 sequons, no unmodified peptides were detected, suggesting that at least one of the asparagine residues was N-glycosylated. Determining the individual N-glycosylation ratio for N124 and N125 solely from HCD MS/MS measurements presented challenges. To assess the N-glycosylation of these asparagines, we performed deglycosylation of sErbB2s using PNGase F and observed the ratio of asparagines converted to aspartic acid through deglycosylation. HCD MS/MS analysis of PNGase F-treated sErbB2 revealed that N124 in both HEK293- and CHO-expressed sErbB2 were converted to aspartate through de-N-glycosylation at rates of 99.9% and 92.4%, respectively, whereas N-glycosylation at N125 was very limited (Fig. S4). This indicates that N124 is almost exclusively glycosylated in the consecutive sequons, N124 and N125.

At first glance, the distribution of N-glycan structural classes (high mannose- or hybrid/complex-type) in sErbB2 was similar between host cells (Fig. 2B and Fig. S2). Except for N259 in subdomain II, which was occupied by high mannose-type glycans, complex-type glycans were predominant at all other sites. To assess the similarity or dissimilarity of sErbB2 N-glycosylation across different host cells, comparative and correlation analyses based on structural classes of sErbB2 from HEK293 and CHO cells were performed. Classifying N-glycosylation into unmodified, high mannose-type, and complex/hybrid-type, their ratios at each site were similar between sErbB2 expressed with HEK293- and CHO cells (Fig. 3A). The correlation analysis, which was based on the ratio of these categories by site for sErbB2 expressed in HEK293 and CHO cells, exhibited remarkably high correlation coefficients (r^2) > 0.944 for all sites (Fig. 3B). The similarity of N187 is mainly due to the fact that a markedly higher proportion of unmodified peptides containing N187 (74.5% and 84.5% for HEK293- and CHO-expressed sErbB2s, respectively), suggesting a lower site occupancy of N187, is shared by both sErbB2s (Fig. 2B,

Fig. S2, and Fig. 3A). Glycans at N259 were characteristically limited to high mannose-type, and the most abundant structure in both HEK293- and CHO-expressed sErbB2 was $(Man)_3 + (Man)_3(GlcNAc)_2$, accounting for 72.6% and 85.5% respectively (Fig. 2B, Fig. S2 and Fig. 3A). The other sequons were predominantly occupied by complex-type glycans as the major glycans (Fig. 3A). A wide variety of complex-type N-glycans were identified in both sErbB2, showcasing their structural diversity (Fig. 2B and Fig. S2). Among the identified complex-type N-glycans, the site-specific fucosylation rates were interestingly comparable between HEK293- and CHO-expressed sErbB2 (Fig. 3C). The characteristics of low fucosylation rate for the complex-type glycans at N530 were also consistent in both sErbB2 (55.8% for HEK293- and 21.5% for CHO-expressed sErbB2), while the fucosylation rate for complex-type glycans at other sites was >95% (Fig. 3C). Most fucosylated N-glycans displayed fragment ions $[peptide+(HexNAc)_1(Fuc)_1]^+$ in HCD MS/MS, indicating the presence of core fucose. In Fig. 2B, only glycans that exhibited this fragment ion are marked with the structure with core fucose, while those without this fragment ion are labeled without core fucose. In contrast to fucosylation, the site-specific sialylation profiles of complex-type glycans differed markedly between host cells (Fig. 3D). Sialylation rates at each site were generally lower compared to fucosylation, especially in HEK293-expressed sErbB2, where the sialylation rate of complex-type glycans at all sites was less than half that in CHO-expressed sErbB2 (Fig. 3D). The similarity of the backbone structure of the complex-type glycans between both sErbB2 was clearly site-specific. HEK293-expressed sErbB2 tended to exhibit a higher abundance of glycans with bisecting GlcNAc, showing fragment ions $[peptide+(Hex)_1(HexNAc)_3]^+$ in HCD MS/MS, compared to CHO-expressed sErbB2 (Fig. S3). Classifying the estimated complex-type N-glycans into four structural categories (biantennary, triantennary or bisecting, triantennary, and tetraantennary), the site-specific population of each category and the correlation coefficients between both sErbB2s are presented in Fig. 3E and F, respectively. At N68, complex-type glycans in HEK293-expressed sErbB2 were predominantly triantennary structures (65.9%), while in CHO-expressed sErbB2, 89.7% of them converged to biantennary structures (Fig. 3E), revealing a notably low correlation coefficient (r^2 = 0.022) (Fig. 3F). The correlation coefficients for complex-type glycans at the other sites were
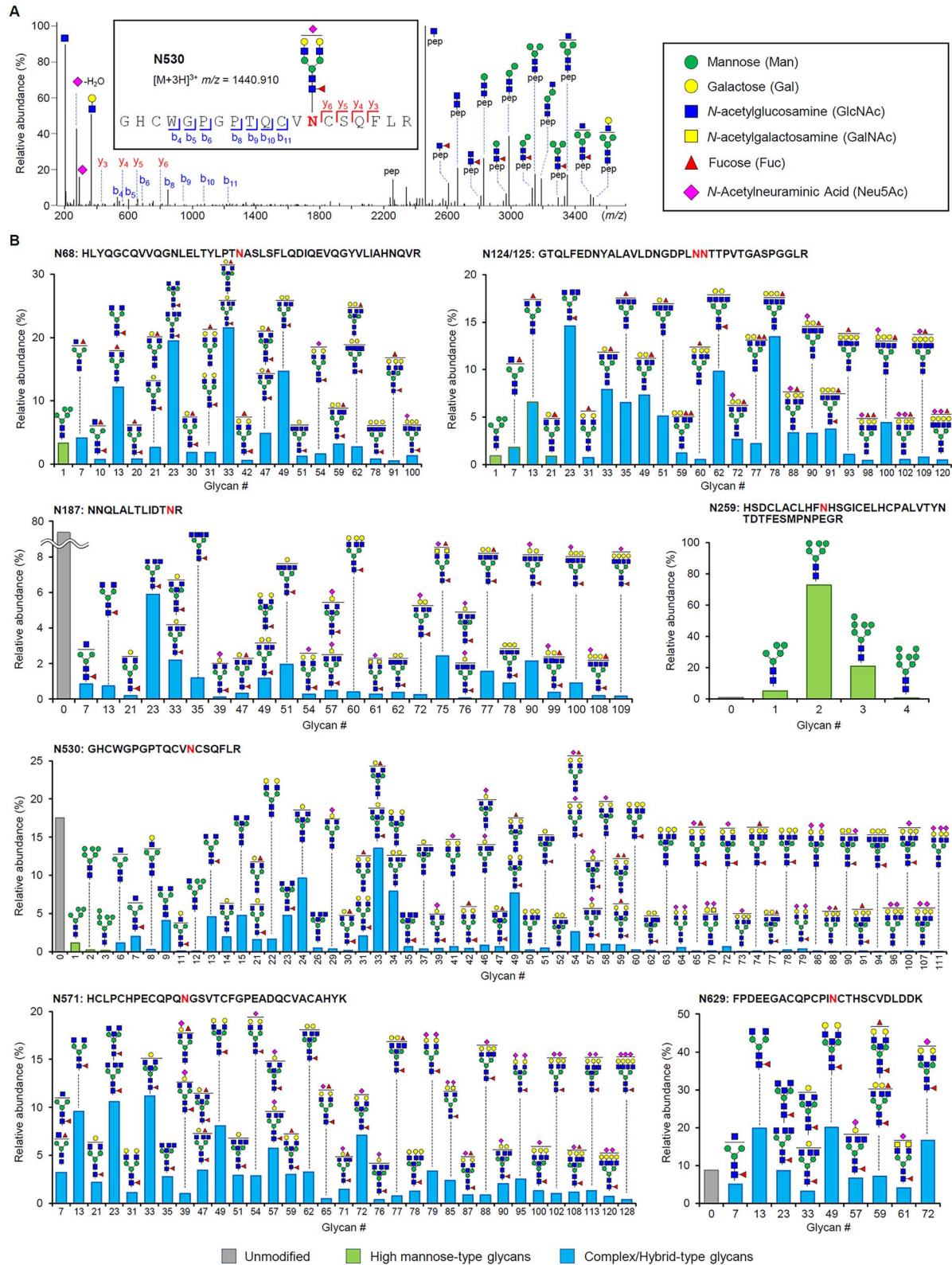
Fig. 2. Site-specific N-glycosylation profile of HEK293-expressed sErbB2. A) A representative HCD MS/MS spectrum for the identification of the N-glycoform at N530 of HEK293-expressed sErbB2. Identified fragment ions (b-ions and y-ions) were labeled. Monosaccharide symbols follow the SNFG (symbol nomenclature for glycans) system (Varki et al. 2015). B) Precursor ion intensity-based N-glycosylation profile on eight sequons. The colors of the bars correspond to unmodified, high mannose-type glycans, and complex-type glycans. Refer to Table S1 for glycan number. Structures exhibiting fragment ions [pep+(Hex)$_1$+(HexNAc)$_3$]$^+$ are labeled with structures containing bisecting GlcNAc.

generally high, except for N187, which had a relatively low correlation coefficient of 0.354. Nevertheless, the correlation coefficients at the other sites were higher than 0.657, indicating a high overall structural similarity between the two sErbB2s (Fig. 3E). This N-glycosylation profiling of sErbB2 suggests that the sialylation rate and glycan branching at N68 exhibit cell-dependent diversity, while there is high similarity in terms of modification rates and glycan structural classes.

Site-specific N-glycosylation analysis of ErbB2 from ErbB2-positive breast cancer cells

In addition to recombinant sErbB2, site-specific N-glycosylation analysis of endogenous ErbB2 derived from ErbB2-positive breast cancer cells (SKBR-3 and BT474) was also performed. Identification of the N-glycoform of ErbB2 from cancer cells was performed as for sErbB2s, and relative abundance and MS/MS data are summarized in Tables S1 and S2, respectively. Of the eight sequons in ErbB2, a total of 27 N-glycoforms were identified for all sequons in ErbB2 derived from SKBR-3 cells and three sequons (N68, N259 and N530) in ErbB2 from BT474 cells. The N-glycan profiles at sites commonly observed in both SKBR-3 and BT474 cells (N68, N259, and N530) are shown in Fig. 4A, and Fig. S5 presents the N-glycan profiles at sites where N-glycans were exclusively identified in SKBR-3 cells.

To characterize cancer cell-specific N-glycosylation status, site-specific correlation analysis based on N-glycosylation profiles between endogenous ErbB2 from breast cancer cells and recombinant sErbB2 was performed (Fig. 4B). In the subdomain I sequons, the glycans at N68 observed in both SKBR-3 and BT474 (Fig. 4A), as well as N124/N125 observed exclusively in SKBR-3 (Fig. S5), were limited to high mannose-type glycans. This differs remarkably from sErbB2, where complex-type glycans were predominantly found in the subdomain I sequons (Figs 2B and 3A). The weak correlations between N-glycoforms at N68 were shown between ErbB2 from breast cancer cells and sErbB2s (Fig. 4B). However, the observation that N-glycans in subdomain I of ErbB2 from breast cancer cells exclusively consist of the high mannose-type structures aligns with the report that N124 is occupied by high mannose-type glycans, although no glycans at N68 have been identified, in ErbB2 from gastric cancer cells, NCI-N87 (Duarte et al. 2021). The N-glycoforms at N259 in subdomain II were also limited to the high mannose-type (Fig. 4A), which was common to recombinant sErbB2 (Figs 2B and 3A), leading to relatively high r^2 values between ErbB2 and sErbB2 from breast cancer cells (Fig. 4B). Although not identified in ErbB2 from wild-type NCI-N87 cells, N-glycans at N259 were limited to a high mannose structure $(Man)_3 + (Man)_3(GlcNAc)_2$ in the β-galactoside $\alpha 2,6$-sialyltransferase 1 knockout NCI-N87 (Duarte et al. 2021).

The correlations of N-glycans at N530 between sErbB2 and ErbB2 from breast cancer cells were weak, although breast cancer cells exhibited relatively similar characteristics to each other (Fig. 4B). The proportion of unmodified peptides in N530-containing peptides was below 20% in sErbB2 (Figs 2B and 3A), whereas it was comparatively higher in SKBR-3 and BT474 cells, measuring at 32.1% and 50.4%, respectively (Fig. 4A). The N-glycoforms at N530 were a mixture of high mannose- and complex-type glycans (Fig. 4A), which is consistent with the analysis of ErbB2 in gastric cancer cells, NCI-N87 (Duarte et al. 2021). The high mannose-type

glycans accounted for 31.8% and 30.1% of total N-glycans at N530 in SKBR-3 and BT474 cells, respectively, while in sErbB2, they were only a small fraction, comprising approximately 1.4% in HEK293-expressed sErbB2 and 1.7% in CHO-expressed sErbB2. The fucosylated glycans ratio to total complex-type glycans at N530 was 54.5% in SKBR-3 cells and 93.5% in BT474 cells (Fig. 4C), compared to 55.8% in HEK293-expressed sErbB2 and 21.4% in CHO-expressed sErbB2. Characteristically, the ratio of di- or tri-fucosylated glycans was 25.2% and 88.1% to the total fucosylated glycans in ErbB2 from SKBR-3 and BT474 cells, respectively, while it was less than 3% in sErbB2s (Fig. 4C). These fucosylation profiles suggested that a higher abundance of N-glycans displaying the Lewis antigens in ErbB2 form both breast cancer cell lines compared to sErbB2. Indeed, this was corroborated by the detection of a diagnostic fragment ion $[(Hex)_1(HexNAc)_1(Fuc)_1]^+$ at m/z 512.19 in the HCD MS/MS analysis of the glycopeptides modified with the difucosylated N-glycans (Fig. 4D). In particular, the presence of the fragment ion $[peptide+(Hex)_3(HexNAc)_3(Fuc)_2]^+$ in the example provided in Fig. 4D indicated that another fucose other than core fucose was bound to GlcNAc, the fundamental structure of Lewis^a or Lewis^x antigens, instead of galactose as in the H-type structure. The ratio of sialylation to total complex-type glycans was considerably enhanced in breast cancer cells compared to sEbB2, being less than 50% for sErbB2 compared to 100% for SKBR-3 cells and 94.2% for BT474 cells (Fig. 4C). ErbB2 from breast cancer cells was enriched in multiple sialylated structures, especially in ErbB2 of SKBR-3, where the ratio of di- and tri-sialylated structures to total complex-type N-glycans reached 70.7% (Fig. 4C).

Site-specific analysis of sErbB2 O-glycans

In addition to N-glycans, O-glycans on sErbB2 were also explored. To date, large-scale O-glycosylation analysis of cells and tissues has found that T127 of ErbB2 is O-glycosylated with $(Hex)_1(HexNAc)_1$, which is probably the core-1 type structure (Yang et al. 2018). To accurately analyze MS/MS spectra of glycopeptides when both N- and O-glycans coexist, O-glycans were surveyed using PNGase F-treated sErbB2 to remove interfering N-glycans. An O-glycosylated peptide, ^{104}GTQLFEDNYALAVLDNGDPLNNTTPVTGASPGGLR^{138}, containing the N124/N125 sequons, was identified. The identified O-glycoform profiles are summarized in Table S3. Representative HCD MS/MS spectra of glycopeptide displaying m/z corresponding to peptide+$(Hex)_2(HexNAc)_2(NeuAc)_2$ in HEK293-expressed sErbB2 are shown in Fig. 5A. This peptide contains four potential sites for O-glycosylation: T105, T126, T127 and S133. Definitive determination of the O-glycosylation site was challenging due to the absence of clear fragment ions derived from the O-glycosylated amino acids. However, in HCD MS/MS spectra, the sequential detection of the unmodified b-ion series downstream from b18 and the unmodified y-ion series downstream from y26 inferred that T105 and T126 were the most plausible O-glycosylated sites (Fig. 5A and Fig. S6). The HCD MS/MS spectra in Fig. 5A show that this glycopeptide is either T105 or T126 modified with the disialylated core-2 type, $(Hex)_2(HexNAc)_2(NeuAc)_2$, or both sites with sialylated core-1, $(Hex)_1(HexNAc)_1(NeuAc)_1$ (Fig. S6). Figure 5B shows the host cell-dependent structural profiles of O-glycans on sErbB2. The HEK293-expressed sErbB2 possessed either branched core-2 structure on one site or core-1 type structures
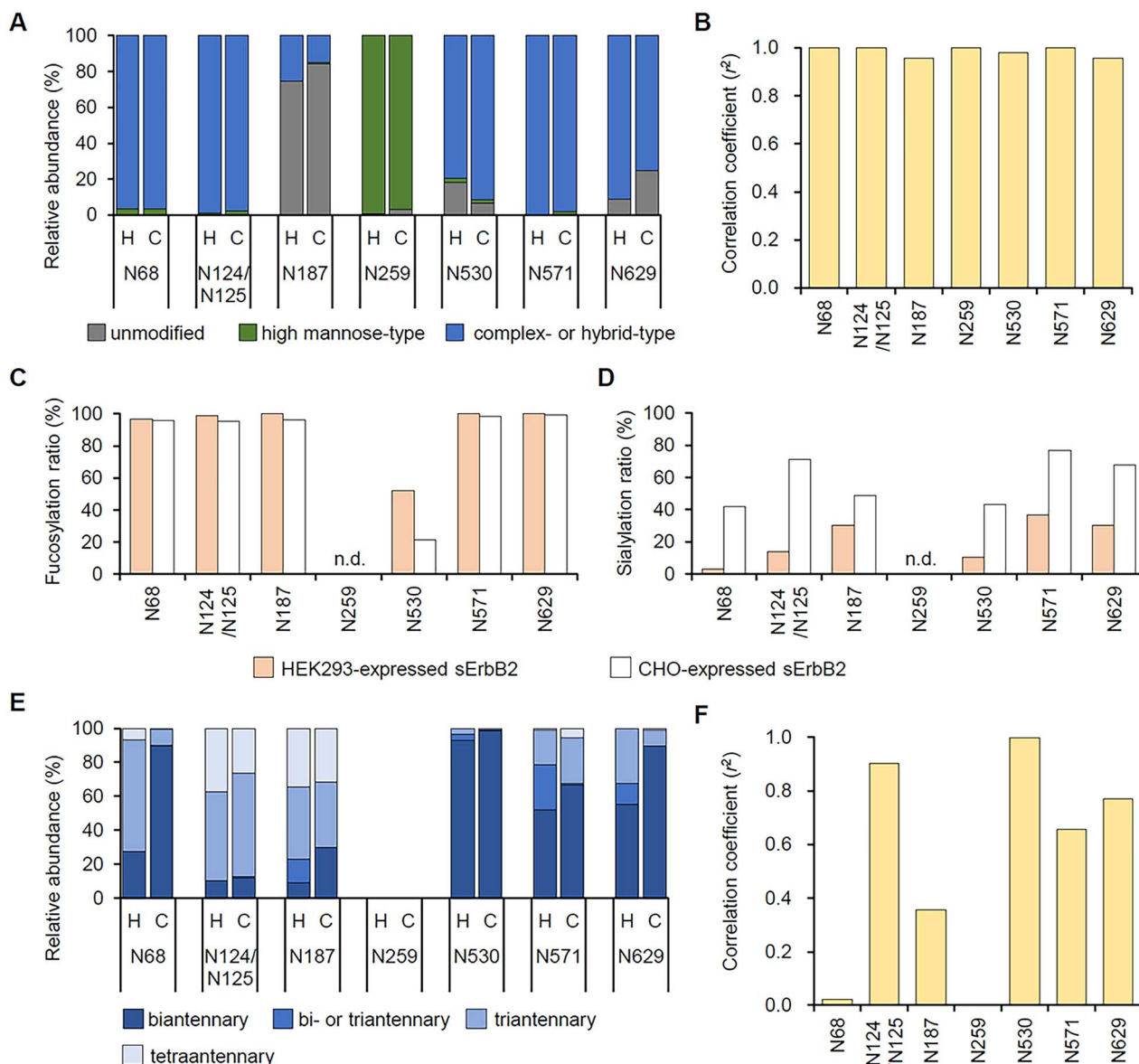
Fig. 3. Comparison of N-glycosylation profiles of HEK293- and CHO-expressed sErbB2. A) Ratio of unmodified peptides, high mannose-type N-glycans, and complex/hybrid-type N-glycans in each sequon. The labels for H and C indicate the expression hosts, HEK293 and CHO cells, respectively. B) Correlation coefficient (r^2) for each sequon based on structural class (A). C) Ratio of fucosylated glycans to total complex-type N-glycans for each sequon. D) Ratio of sialylated glycans to total complex-type N-glycans for each sequon. E) Ratio of the branching structure of complex-type N-glycans in each sequon. H and C labels indicate the expression hosts, HEK293 and CHO cells, respectively. The labels for H and C are the same as in (A). F) Correlation coefficient (r^2) for each sequon based on the ratio of branching structure of complex-type N-glycans (E). The value of N259 in (C), (D), (E) and (F) cannot be defined because N259 is occupied with only high mannose-type glycans.

on two sites, and CHO-expressed sErbB2 contains core-1 type structures on one site (Table S3). Although O-glycosylation rates were low in HEK293- and CHO-expressed sErbB2 (5.0% and 11.3%, respectively), these findings indicated the potential for O-glycosylation of ErbB2.

Exploring the function of N-glycans for ErbB2 molecules

To address the function of N-glycans on ErbB2 molecule, we performed in silico structural modeling of N-glycosylated ErbB2 was performed by incorporating the most abundant N-glycans at each site identified in sErbB2 expressed in HEK293 (Table 1) into the structure of the extracellular region of ErbB2 extracted from the trastuzumab-ErbB2 complex

structure (Cho et al. 2003) (Fig. S7). No glycans were added to N125 and N187 because proportions of unmodified peptides containing these sites in sErbB2 expressed in HEK293 were 100% and 74.5%, respectively (Figs 2, S4 and S7), suggesting that the occupancy of these sites was expected to be low. Comparison of the constructed structures of N-glycosylated and non-N-glycosylated sErbB2 (Movies S1 and S2, respectively) showed that the fluctuations in the loop structure of [382]DGDPASNTA[390] in subdomain III of N-glycosylated ErbB2 were suppressed. The root mean square of fluctuation (RMSF) for each residue is shown in Fig. 6A. The average RMSFs for each residue of the loop, [382]DGDPASNTA[390], of N-glycosylated and unmodified ErbB2 were found to be 1.48 and 2.59 Å, respectively. The structure of N-glycosylated ErbB2 showed that the antenna
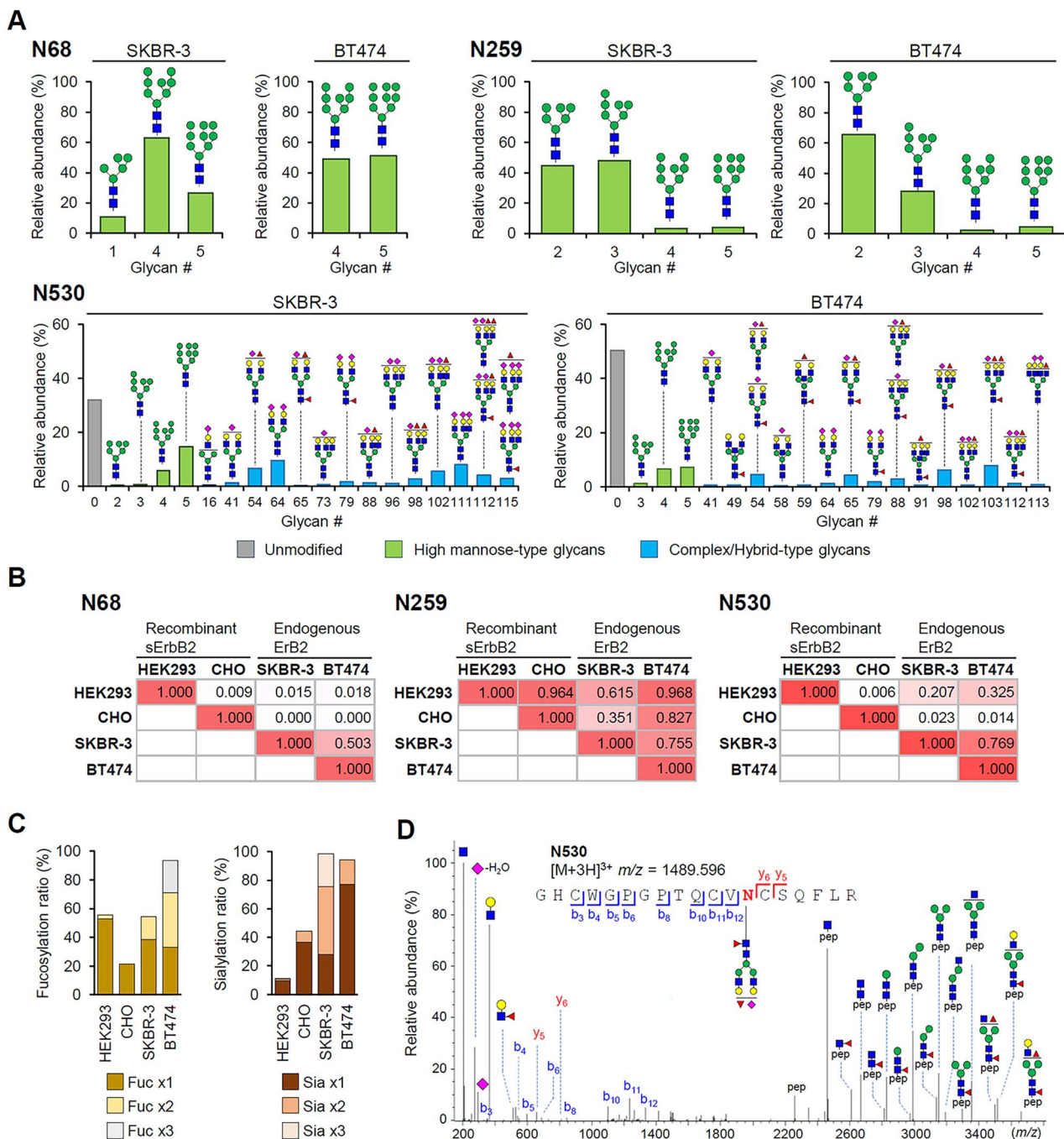
Fig. 4. N-glycan profiles of ErbB2 derived from ErbB2-positive breast cancer
cells, SKBR-3 and BT474. A) Bars colored gray, green and blue represent unmodified peptides, high mannose-type glycans and hybrid/complex-type glycans, respectively. As in Fig. 2, the colors of the bars correspond to unmodified, high mannose-type glycans, and complex-type glycans.
Refer to Table S1 for glycan number. B) Correlation coefficient (r^2) matrix between endogenous ErbB2 from cancer cells and recombinant sErbB2
calculated from the ratio of individual N-glycoforms at N68, N259 and N530. C) Fucosylation and sialylation rates of sErbB2s and cancer cell ErbB2s to
total complex-type N-glycans at N530. D) A representative HCD MS/MS spectrum to identify N-glycoform displaying the Lewis antigen at N530 of ErbB2
in SKBR-3. Identified fragment ions (b-ions and y-ions) were labeled. Monosaccharide symbols in (A) and (D) follow the SNFG system (Varki et al. 2015).

GlcNAc (oxygen at position 4) of the N-glycan at N124 was able to be close as 4.9 Å to P385 (proton at position γ) on the short loop of subdomain III (Fig. 6B). The model structure of N-glycosylated ErbB2 suggested that the N-glycan at N124 may function structurally to stabilize the conformation of ErbB2.

Based on the structure of N-glycosylated ErbB2, we investigated the functional impact of N-glycans at N124 in cellular contexts using ErbB2-KO-293 cells in which endogenous ErbB2 was knocked out by the CRISPR/Cas9 system. Two stable cell lines were generated: one expressing wild-type ErbB2 and the other expressing a mutant ErbB2 lacking the N-glycan at N124. To investigate the compensatory effect of replacing N124 with glutamine on N-glycosylation at N125, considering their consecutive sequons presence in ^{121}DPTLNNTTPVTGASPGG^{136}, mutant sErbB2_N124Q was expressed, purified, and examined for N125 glycosylation. A small amount of glycosylation was detected at N125
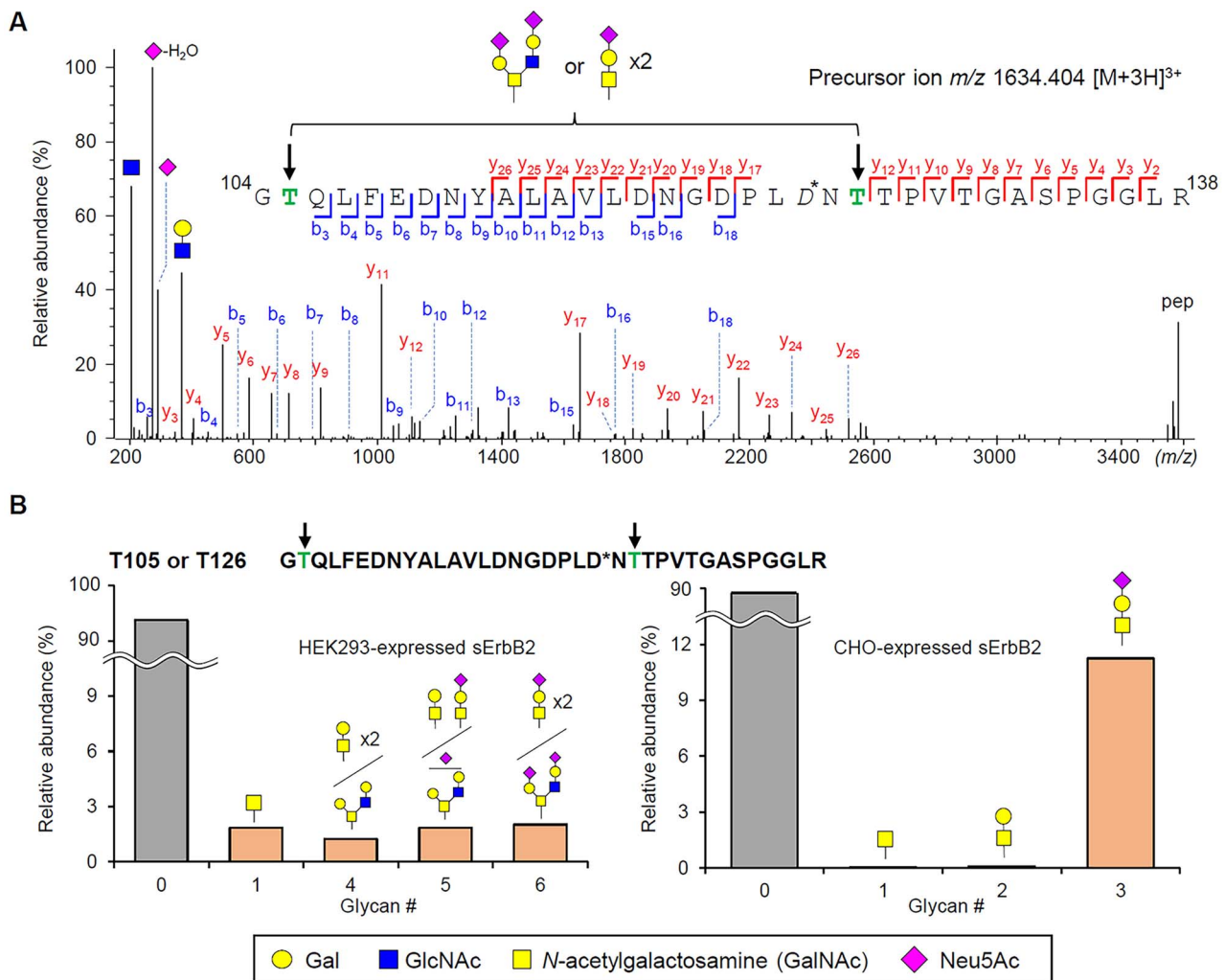
Fig. 5. Site-specific O-glycosylation analysis of sErbB2. A) A representative HCD MS/MS spectrum for the identification of the O-glycosylated peptide, [104]GTQLFEDNYALAVLDNGDPLNNTTPVTGASPGGLR[138]. The HCD MS/MS spectrum from the trivalent precursor ion, m/z 1634.4044 [M+3H][3+], are considered to be derived from the following two types of O-glycosylated peptides: The peptide modified in one site with core-2 type glycan "peptide+(Hex)$_2$(HexNAc)$_2$(NeuAc)$_2$", and the peptide modified in two sites with core-1 type glycans "peptide+2x(Hex)$_1$(HexNAc)$_1$(NeuAc)$_1$". Asterisked D (D∗) indicates N-glycosylated asparagine converted to aspartate by PNGase F treatment, and O-glycosylated sites were indicated by arrows. Note that the observed fragment ion series (b- and y-ions) are derived from unmodified amino acids. Monosaccharide symbols follow the SNFG system (Varki et al. 2015). B) Precursor ion intensity-based O-glycan profiles. The bar with glycan number 0 represents the ratio of unmodified peptides. Refer to Table S3 for glycan number.

in sErbB2_N124Q, (Fig. S8), whereas no glycosylation was observed at N125 in HEK293-expressed sErbB2 (Fig. S4). Therefore, to eliminate the influence of N125 glycans in deleting the N124 glycans, the expression vector was constructed where both N124 and N125 were replaced with glutamine (ErbB2_N124Q/N125Q).

Both wild-type and N124Q/N125Q mutant localized to the plasma membrane surface without accumulating in the endoplasmic reticulum, indicating that N-glycans at N124 do not impact on the intracellular transport of ErbB2 (Fig. S9). To compare the intracellular stability, their half-lives were estimated using the protein synthesis inhibitor cycloheximide (CHX) chase experiments. The wild-type and N124Q/N125Q ErbB2s had estimated intracellular half-lives of 18.9 ± 4.5 and 10.8 ± 2.2 h, respectively, with 100 μg/mL CHX chases (Fig. 6C). The mutant ErbB2 deficient in N-glycans at N124 exhibited a shorter intracellular half-life compared to the wild-type ErbB2, indicating that N-glycans at N124 contribute to the intracellular stability of ErbB2. Additionally, the

phosphorylation of ErbB2 was compared between the wild-type and mutant deficient in N-glycans at N124. ErbB2 is activated in a ligand-independent manner, leading to phosphorylation of the ErbB2-ErbB2 homodimer even at steady state in the absence of exogenous stimulation. The mutant deficient in N-glycans at N124 showed significantly reduced phosphorylation rates compared to the wild-type ErbB2 (Fig. 6D), suggesting that N-glycans at N124 not only contribute to the ErbB2 molecular stability but also affect its phosphorylation.

Structural comparison to evaluate functional differences between ErbB2 and EGFR from the perspective of N-glycosylation

To gain insight into the involvement of N-glycosylation in the distinctive activation mechanism of ErbB2, which differs from other members of the ErbB family by not requiring a ligand for activation, we compared the site-specifically mapped N-glycoforms in the structure of the extracellular
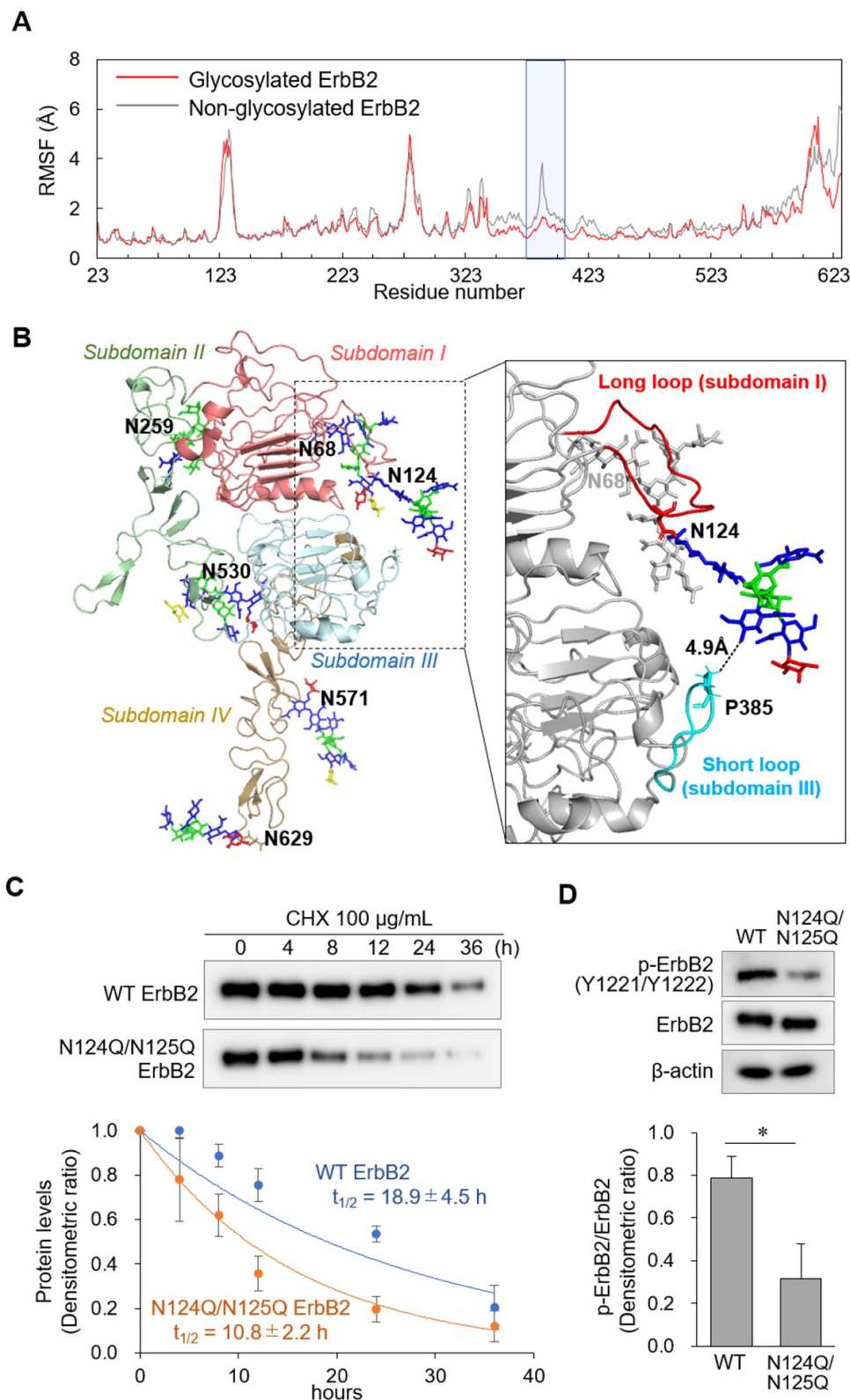
Fig. 6. Potential contribution of N-glycans at N124 in subdomain I to the stability of ErbB2. A) RMSF values (Å) for each residue in the extracellular region of glycosylated and unmodified ErbB2 calculated via molecular dynamics simulations. The perturbation of the short loop [382]DGDPASNTA[390] in subdomain III was reduced by N-glycosylation at N124. B) Structure of the extracellular region of N-glycosylated ErbB2 where the N-glycan at N124 on the loop of subdomain I is closest to the short loop of subdomain III. C) CHX chases to estimate the intracellular half-lives of wild-type and N-glycan-deficient mutant N124Q/N125Q ErbB2. Quantitative values are the mean ± S.D. of three trials using cells treated independently with CHX and the densitometric intensities at 0 h are normalized to 1.0. D) Change in phosphorylation (Y1221/Y1222) at steady state between wild-type and N-glycan-deficient mutant N124Q/N125Q ErbB2. The p-ErbB2/ErbB2 ratios are the mean ± S.D. of three independent trials (Student's t-test, ∗$P < 0.05$).

region of EGFR (sEGFR) and sErbB2. In addition to glycan analysis of sErbB2, we also performed a site-specific N-glycosylation analysis of sEGFR expressed in CHO cells, identifying a total of 97 N-glycan structures. These structures were generally consistent with the major N-glycan structures previously reported for EGFR (Tables S1 and S2 and Fig. S10) (Wu et al. 2006; Hanash et al. 2008; Liu et al. 2011; Hasegawa et al. 2015). Due to the high proportion of unmodified peptides containing N128, N196 and N603 of sEGFR (77.5%, 58.2% and 65.5%, respectively), the model structures of N-glycosylated sEGFR without N-glycans at these sites were generated as for sErbB2. Focusing on the ligand-binding site, a structural comparison between N-glycosylated sErbB2 and sEGFR revealed that the glycans at N68 and N124 of sErbB2 protrude outward in bundles and mask the open space between subdomains I and III, which corresponds to the ligand binding site of EGFR (Fig. 7). In contrast, in sEGFR, the N-glycans at N175 in subdomain I and N352 in subdomain III appear to form glycan clusters that help maintain the cavity between subdomains I and III. The corresponding glycosylation sites for these are not present in ErbB2. Apart from the ligand-binding site, there are other notable differences between the two molecules. The structural comparison revealed that the extracellular region of EGFR contains more glycan-dense regions compared to ErbB2. Specifically, in EGFR, the high mannose-type glycans at N361, the bundles of glycans at N413 and N444 in subdomain III, and the glycans at N504 and N544 in subdomain IV were clustered and prominently exposed, effectively masked subdomain IV (Fig. 7). These densely-packed glycans provide a protective shield against antibody recognition of peptide epitopes, thereby reducing immunogenicity. This seems reasonable from the glycosylation perspective, given that antibody therapies against targeting ErbB2, such as trastuzumab, are more responsive than those targeting EGFR and have been put into practical use as first-line agents.

Discussion

This study showed the potential glycosylation state of ErbB2, including N- and O-glycosylation, through the site-specific glycan analysis of recombinant sErbB2s expressed with HEK293 and CHO cells (Figs 2 and 5). The structural classes of N-glycans were distinctly exclusive of certain sites: N259 in subdomain II was dominated by high mannose-type glycans in both HEK293- and CHO-expressed sErbB2, while other sites were dominated by complex-type glycans. Comparison of N-glycans based on structural class revealed that sErBb2 from both cell types correlated extremely strongly, while complex-type N-glycans showed a wide cell-dependent structural diversity (Fig. 3).

The occupancy of N259 with high mannose-type glycans was consistently observed in ErbB2 of breast cancer cells (Fig. 4A). On the other hand, complex-type N-glycans were predominant in the recombinant sErbB2 subdomain I sequons (N68, N124/N125, N187), whereas high mannose-type N-glycans were predominant in ErbB2 from cancer cells (Fig. 4A), which is consistent with the pattern of glycan addition observed in ErbB2-positive gastric cancer cells (Duarte et al. 2021). The presence of high mannose-type glycans on mature proteins implies that remodeling to complex-type glycans in the Golgi is evaded. This suggests that

the regions modified with high mannose-type N-glycans are either buried within the molecule or associated with proteins such as chaperones during the folding process, making them inaccessible to mannosidases and glycosyltransferases. Thus, the regions where high mannose-type N-glycans are bound could be targets for disabling ErbB2 function by interfering with the folding process of ErbB2. The complex-type N-glycans showed wide structural diversity across expressing host cells and cancer cells. The complex-type N-glycans of N530 of ErbB2 derived from cancer cells were characterized by a higher ratio of multifucosylated to multisialylated glycans compared to sErbB2s (Fig. 4C). Sialylated and fucosylated glycans have been implicated in regulating the function of RTKs. For instance, in lung cancer cells, sialylation and fucosylation of N-glycans in EGFR have been shown to the downregulate receptor dimerization and activation (Liu et al. 2011). Similarly, sialylation of N-glycans in VEGFR2 has been demonstrated to modulate ligand binding capacity and signal transduction (Chandler et al. 2019). Among the fucosylated N-glycans, monofucosylated structures were predominant in sErbB2s, while the ratio of difucosylated structures was increased in both SKBR-3 and BT474 breast cancer cells (Fig. 4C). Of particular interest, these glycans were enriched in the Lewis antigens (Fig. 4D). Fucosylation has been implicated in the regulation of cancer cell metastasis (Schneider et al. 2017), and targeting increased fucosylation of ErbB2 may hold potential for future drug discovery. Sialylation was greatly enhanced in both breast cancer cells compared to sErbB2 (Fig. 4C). Targeting sialylation represents an attractive strategy for cancer therapy due to the involvement of sialylated glycans on the cell surface in evading immune responses through binding to sialic acid-binding Ig-like lectins, Siglecs (Crocker et al. 2007; Hudak et al. 2014). Additionally, sensitivity to trastuzumab for chemotherapy against ErbB2-positive cancers has been reported to be increased in α2,6-sialylation-deficient cells (Duarte et al. 2021). Thus, the site-specific N-glycosylation analysis of ErbB2 in this study enables to describe the promotion of ErbB2 sialylation in cancer cells and supports the efficacy of sialidase-bound trastuzumab at the molecular level (Xiao et al. 2016; Stanczak et al. 2022). However, sialylation is prone to fluctuations depending on the cell type and condition. In fact, our analysis revealed considerable differences in sialylation rates between sErbB2 expressed in HEK293 and CHO cells (Table 1 and Fig. 4C), and we have recently reported that the sialylation status is altered depending on the cellular conditions (Fujitani et al. 2021). Furthermore, the higher ratio of unmodified peptides containing N530 in subdomain IV compared to sErbB2 suggests a low site occupancy at this site. This may be advantageous for the effective binding properties of antibodies targeting subdomain IV, such as trastuzumab, to ErbB2.

This study could be feasible in the future as a way to compare the site-specific N-glycan profiles of ErbB2 in cells and propose N-glycoforms that are expected to be characteristic of cancer cells. However, there are several limitations of this study that need to be addressed in terms of the practical application of therapeutic targeting of ErbB2 glycans. There is concern that the ErbB2 N-glycan profiles shown here were generated from cell line systems and may not represent the real picture of a complex environment such as tissue. To confirm whether the increased ratio of fucosylation, sialylation, and unmodification of ErbB2 N-glycans is truly cancer-specific, follow-up studies using clinical specimen samples
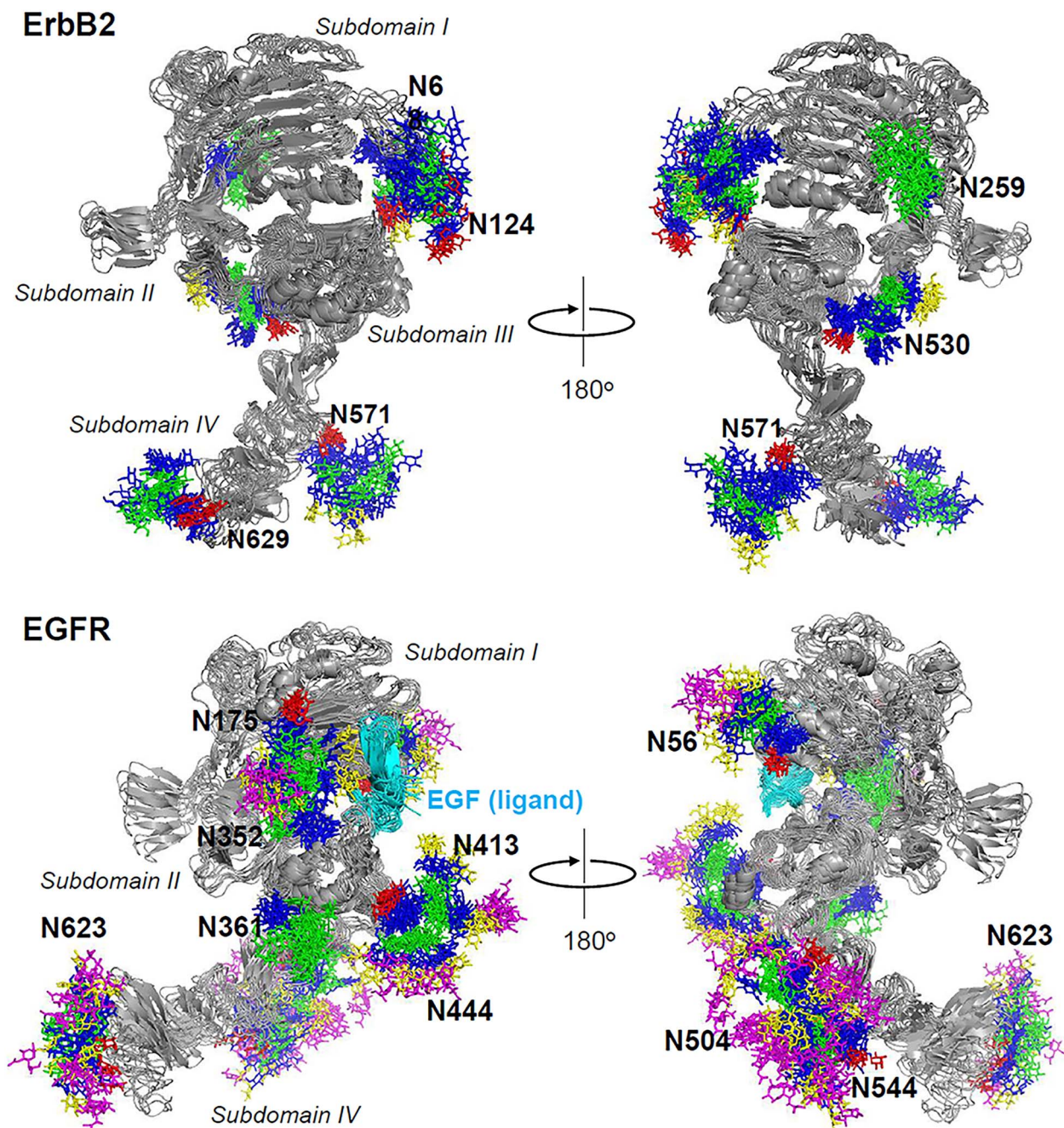
Fig. 7. Structural comparison of N-glycosylated extracellular regions of ErbB2 and EGFR. Superimposed 10 structures for every 100 ps obtained by N-glycosylated structure modeling. Protein backbone structure is colored gray. The monosaccharide structure colors follow the SNFG system (Varki et al. 2015).

representing actual pathological conditions, such as cancer tissue, are essential. To our knowledge, there is no report so far of achieving site-specific glycan analysis of ErbB2 from tissue, although there is report of visualization of N-glycans throughout ErbB2-positive breast cancer tissue (Scott et al. 2019). Complete site-specific glycan structure analysis of tissue-derived proteins is currently expected to be challenging. In fact, it is still difficult to completely determine the glycan structures of all glycosylation sites of endogenous RTKs even in cell-based experiments, including the example of ErbB2 derived from cancer cells in this study (Fig. 4 and Fig. S5). For practical applications, it is essential to develop a more robust

system capable of extracting site-specific glycan information from ErbB2 in complex tissues, extending beyond cell-derived endogenous ErbB2. In addition to N-glycans, this study proposed a specific structure for the O-glycans that ErbB2 may possess, although the exact site was not definitively determined. Future experiments integrating HCD and electron transfer dissociation, which minimizes disruption of glycan structures (Mechref 2012), will allow more detailed profiling of O-glycosylation and contribute to the characterization of ErbB2 in terms of O-glycosylation.

The structure modeling of N-glycosylated ErbB2 brought attention to the long flexible loop around [121]DPLNNTTPVT

GASPGG[136], which contains sequons N124/N125. To the best of our knowledge, this loop has previously received little structural or functional attention. To date, structural information on 20 extracellular regions of human ErbB2 is available from the Protein Data Bank (PDB). Except for the complex structure of the ErbB2 extracellular region with the antibody HuA21 (Wang et al. 2019), all other structures determined with X-ray crystallography or cryo-electron microscopy lack coordinate data to define the conformation of this loop. This study revealed that this loop was highly glycosylated, including N- and O-glycans (Figs 2B and 5). The inability to determine the conformation of this loop, even after N-glycan removal, may be due to its highly flexible nature as well as structural heterogeneity resulting from O-glycosylation. Regarding the structural role of this loop, our structure model of N-glycosylated ErbB2 indicated that the N-glycan at N124 contributes to molecular stability by attenuating fluctuations of the short loop [382]DGDPASNTA[390] in subdomain III (Fig. 6A, Movies S1 and S2). This observation aligns with previous study demonstrating that N-glycosylation has a dampening effect on molecular dynamics, leading to increased conformational stability of proteins (Lee et al. 2015). In vitro experiments based on the structural insights obtained showed that N124 N-glycan-deficient mutant ErbB2 exhibited a reduced intracellular half-life and autophosphorylation compared to wild-type ErbB2 (Fig. 6C and D). These findings suggest that targeting the N-glycans at N124 is a promising approach aiming to destabilize ErbB2 and inhibit its activation. This study shows that N-glycosylation at N124 of ErbB2 plays a role in the structural stability and autophosphorylation of ErbB2, suggesting that this glycan is a potential therapeutic target. However, characterizing the specific glycan structures essential for these functions has proven challenging. This complexity arises from the fact that the glycan structure at N124/N125 of recombinant sErbB2 is predominantly complex-type (Fig. 2B), while that of endogenous ErbB2 in breast cancer cells SKBR-3 is primarily high mannose-type (Fig. S5). The identification of glycan structures crucial for these functions in future studies will enable more precise proposals for therapeutic targets.

The structural comparison of N-glycosylated ErbB2 and EGFR incorporating N-glycans identified in this study could extract the distinctions between them from the viewpoint of N-glycosylations. A notable finding was that the N-glycan at N124 of ErbB2 obscured the region between subdomains I and III, which corresponds to the ligand binding site of EGFR (Fig. 7). We believe that elucidating the structural features of glycosylated ErbB2, as demonstrated in this study, will provide valuable insights into the ligand-independent activation mechanism of ErbB2, which sets it apart from other members of the ErbB family. On the other hand, it is well-known that the extracellular region of the ErbB family adopts an inactive conformation with subdomains II and IV in close proximity in the absence of ligand, and undergoes a drastic conformational change with subdomains I and III brought into proximity upon activation (Cho and Leahy 2002; Ferguson et al. 2003; Bouyain et al. 2005; Kaplan et al. 2016). The inability of the present structural modeling to reveal the effects of glycosylation on these dramatic conformational changes upon activation is one of the limitations of the current study. Future studies employing more extensive MD simulations shed light on the contribution of glycans to the drastic conformational changes observed in the ErbB family.

This study highlights the utility of employing site-specific structural analysis of glycans. Thorough comparative analysis of the glycans of recombinant sErbB2 and endogenous ErbB2 derived from cancer cell lines allowed extraction of N-glycoforms that are expected to be distinctive to cancer cells. Particularly, structural modeling of glycosylated ErbB2 using the identified glycan structures suggests that the N-glycans at N124/N125 may be involved in ErbB2 stability and activation, making them promising target candidates. Furthermore, this study showed that ErbB2 and EGFR, members of the ErbB family with high structural homology, exhibited different glycan distribution and density profiles. These findings offer insights into the differences in their activation mechanisms and may lead to the development of new approaches to regulate each molecule from a glycan perspective. We believe that the site-specific glycan profiles of ErbB2 obtained in this study will serve as a valuable reference for understanding ErbB2 glycosylation and that this approach will accelerate the identification of potential therapeutic targets.

Materials and methods

Cell cultures

Human breast cancer cell lines, SKBR-3 and BT474, were obtained from ATCC. To establish cells stably expressing sErbB2, sEGFR, wild-type ErbB2 and N-glycan deficient mutant ErbB2, Flp-In HEK293 and Flp-In CHO-K1 cells (Thermo Fisher Scientific, Waltham, MA) containing a single integrated Flp recombination target were used as an expression host. SKBR-3 and BT474 cells were cultured in the medium McCoy's 5A (Cytiva, Marlborough, MA) and RPMI 1640 (Nacalai tesque, Kyoto, Japan), respectively. Flp-In HEK293 and Flp-In CHO-K1 were maintained in Dulbecco's modified Eagle's medium (Nacalai tesque). All media were supplemented with 10% (v/v) of fetal bovine serum (Equitech-Bio, Kerrville, TX) and antibiotics (10 U/mL of penicillin and 10 μg/mL of streptomycin, Nacalai tesque). In experiments to determine the intracellular half-lives of ErbB2, the cells were incubated for 36 h in medium containing 100 μg/mL cycloheximide (FUJIFILM Wako Pure Chemical, Osaka, Japan). All cell lines were cultured at 37 °C under 5% CO_2 atmosphere.

Expression and purification of sErbB2 and sEGFR

The constructs for expression of sErbB2 and sEGFR, the open reading frame encoding extracellular region of ErbB2 and EGFR (from residue 1 to 644 for ErbB2, and from 1 to 642 for EGFR) were subcloned into the expression vector pcDNA5/FRT (Thermo Fisher Scientific) as previously reported (Takahashi et al. 2013; Hasegawa et al. 2015). These vectors were co-transfected into Flp-In HEK293 or Flp-In CHO K1 cells using Lipofectamine 2000 (Thermo Fisher Scientific) with the vector pOG44 (Thermo Fisher Scientific) encoding recombinase genes according to manufacturer's instruction. Transfected cells stably expressing sErbB2 or sEGFR were selected by continuous culture ($\sim$ 2 weeks) in medium containing 200 μg/mL for Flp-In HEK293 cells and 600 μg/mL hygromycin B (Merck, Darmstadt, Germany) for Flp-In CHO K1 cells. The culture media containing secreted

sErbB2 or sEGFR were collected with centrifugation (6,000 *g* for 1 h at 4 °C) and filtration (0.22 *μ*m pore, Thermo Fisher Scientific). sErbB2 and sEGFR were purified from the collected media by the liquid chromatography AKTA GO system (Cytiva) with HisTrap HP for affinity chromatogram, MonoQ 5/50 GL for anion exchange chromatogram, and HiLoad 16/600 Superdex200pg for gel filtration. Protein quantification was performed by the BCA assay (Thermo Fisher Scientific).

Protein extraction

All cell lysates used in this study were prepared by sonication in the cell lysis buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1% (w/v) Nonidet P-40, 1 mM EDTA) supplemented with protease inhibitor cocktail (Thermo Fisher Scientific) and phosphatase inhibitor cocktail (Roche, Basel, Switzerland) on ice and centrifugation (15,000 *g* for 15 min at 4 °C). The resulting supernatant was subjected to the following BCA protein assay, PNGase F treatment, Western blotting and immunoprecipitation.

PNGase F treatment

To remove N-glycans, 1 U of PNGase F (Roche) was added to aliquots containing 15 *μ*g of total protein from cancer cell lysates or 1 *μ*g of recombinant sErbB2 and incubated at 37 °C for 16 h. To assess the degree of deglycosylation by mobility shift, these aliquots were subjected to SDS-PAGE and western blotting.

Western blot

Protein-transferred PVDF membranes were blocked in 5% (w/v) skim milk in TBST (20 mM Tris, 150 mM NaCl, 0.1% Tween-20) for 1 h at room temperature, then the membranes were incubated with rabbit anti-ErbB2 (#2165, Cell Signaling Technology, Danvers, MA; 1:1000) and rabbit anti-His-tag (#PM032, Medical and Biological Laboratories, Tokyo, Japan; 1:1000) antibodies to detect endogenous ErbB2 from cancer cells and recombinant sErbB2, respectively in 5% (w/v) bovine serum albumin for overnight at 4 °C. To detect phosphorylation of ErbB2, rabbit anti-phosphorylated ErbB2 (Y1221/Y1222) (#2243, Cell Signaling Technology; 1:1000) was used as the primary antibody. After washing in TBST, the membranes were incubated with HRP-conjugated anti-rabbit IgG (Promega, Madison, WI; 1:1000) for 1 h at room temperature. After reaction with chemiluminescence reagent (Thermo Fisher Scientific), the immunoreactive bands were imaged on ImageQuant LAS500 system (Cytiva). The intracellular half-lives and changes in phosphorylation of ErbB2 were estimated by quantifying the densities of Western blot bands using ImageJ software.

Immunoprecipitation

Whole cell lysates of SKBR-3 and BT474 cells containing 300 *μ*g total proteins were pre-cleared with 20 *μ*L of Ab-Capcher MAG2 magnetic beads coupled with the modified protein A (ProteNova, Higashikagawa, Japan) for 2 h at 4 °C. To bind the antibody to the beads, 4 *μ*L of anti-ErbB2 antibody (#2165, Cell Signaling Technology) was mixed with 20 *μ*L of beads and incubated at 4 °C for 2 h. ErbB2 was immunoprecipitated by mixing pre-cleared cell lysates with antibody-conjugated beads followed by incubation at 4 °C for 4 h. After wash the beads with lysis buffer, the immune complexes were eluted with 100 mM Gly-HCl buffer (pH 3.0).

The eluents were immediately neutralized for pH with a small amount of 1 M Tris-HCl (pH 7.5).

Mass spectrometric analysis

Purified sErbB2, sEGFR and the immunoprecipitated ErbB2 from cancer cells were reduced with 10 mM dithiothreitol (FUJIFILM Wako Pure Chemical) for 40 min at 56 °C, followed by alkylation with 20 mM iodoacetamide (FUJIFILM Wako Pure Chemical) for 30 min at room temperature in the dark. The reductive alkylated samples were digested with 0.5 *μ*g of sequence grade trypsin/Lys-C (Promega) for 16 h at 37 °C. The resultant peptides were desalted with a styrene divinylbenzene polymer tip column (GL Science, Tokyo, Japan). Flow-through fractions that would contain more hydrophilic peptides, including glycopeptides, were further desalted on graphitized carbon tip columns (GL Science). The collected eluates were combined and evaporated on a centrifugal evaporator. Finally, the resulting desalted peptides were redissolved with 15 *μ*L of ultrapure water containing 0.1% formic acid.

All ErbB2-derived peptide samples were subjected to Orbitrap Q Exactive Plus Spectrometer (Thermo Fisher Scientific) through an EASY-nLC system equipped with an C18 column (0.075 mm × 125 mm, Nikkyo Technos, Tokyo, Japan). The peptides were eluted with acetonitrile gradient from 0% to 30% for 135 min at the flow rate of 300 nL/min. All data were acquired in data-dependent mode. Full MS scans were performed applying the following parameters: mass resolution of 70,000 at m/z 400, AGC target 3e6, maximum injection time 60 msec, mass range 400–2,000. HCD MS/MS scans were acquired by applying the following parameters: mass resolution of 17,500 at m/z 400, AGC target 5e5, maximum injection time 100 msec, loop count 10, isolation window 1.6 m/z, scan range 200–2,000 (fixed first mass m/z 200.0), 27% normalized collision energy, minimum AGC target 5.00e2, intensity threshold 5.0e3, dynamic exclusion duration 15 s.

Acquired data were processed with MaxQuant software 1.6.3.3 (Cox and Mann 2008). The peptide searches referred to amino acid sequence data obtained from UniProt registration numbers P04626 (ErbB2) and P00533 (EGFR). Up to two trypsin miscleavage were allowed. For fragment ion match, a mass tolerance was set to 20 ppm. The false discovery rate was set to 0.01 for the identification of peptides. Methionine oxidation and acetylation of the protein N-terminus were set as variable modifications and carbamidomethyl on cysteine as a fixed modification. For the analysis of PNGase F-treated samples, a substitution from asparagine to aspartate was applied as a variable modification. Glycopeptide identification was performed by a combination of manual analysis and Glyco-Decipher v1.0.4 software (Fang et al. 2022). For manual analysis, first, the molecular weight values deconvolved on MaxQuant from the *m/z* and charge number of the precursor ions, along with the amino acid sequence of ErbB2 or EGFR, are entered into the GlycoMod Tool (https://web.expasy.org/glycomod/), molecular weight-based glycopeptide candidates are obtained. Then, to determine the authenticity of the proposed glycopeptides by GlycoMod Tool, all candidate precursor ions were visually confirmed with MaxQuant, referring to the oxonium ions (*m/z* 204.087 for HexNAc, *m/z* 274.092 for NeuAc-H_2O, *m/z* 292.103 for NeuAc, and *m/z* 366.140 for Hex+HexNAc) and fragment ion for peptide+HexNAc in the MS/MS spectra. In addition,

ensure that all relevant glycopeptides elute within a narrow retention time window. For analysis by Glyco-Decipher, tolerance to MS1 and MS2 was set to 5 and 20 ppm, respectively. In addition, 28 oxonium ions (default setting), including the above oxonium ions, were set as diagnostic peaks for glycopeptides, and the false discovery rate was set to 0.01. We compared the results of manual analysis and analysis by Glyco-Decipher, and adopted the glycan structures identified by both methods. However, glycan structures that could not be identified by Glyco-Decipher, especially N68 of ErbB2 and N528 of EGFR, were identified on the condition that the diagnostic peak was clearly observed. The MS/MS data for all identified glycopeptides form sErbB2 and sEGFR are summarized in Table S2. Relative abundance of glycopeptides and the glycosylation rate were estimated from the intensity of precursor ions.

Structure modeling of N-glycosylated ErbB2 and EGFR

The template structures of the extracellular regions of ErbB2 and EGFR were obtained from PDB (PDB ID 1n8z for ErbB2 (Cho et al. 2003) and 3njp for EGFR (Lu et al. 2010)). All atoms except amino acids were removed from the structure files. The trastuzumab structure in 1n8z and one EGFR coordinate in 3jnp were deleted to prepare monomeric structure for simulations. Four regions with missing structural information in the ErbB2 structure, ^{124}NNTTPVTGA^{132}, ^{325}EDG^{327}, ^{383}GDPA^{386}, and ^{603}GVKPDLSYMP^{612}, were complemented using MODELLER software (Webb and Sali 2016). The most abundant N-glycan at each site identified in this study was attached to the prepared ErbB2 and EGFR structures using Glycan Reader and Modeler module on the CHARMM-GUI (Park et al. 2019). A water box containing 150 mM NaCl with a distance of 30 Å from the protein edge was set up and the system was equilibrated under constant temperature (303.15 K) in Glycan Reader and Modeler module on the CHARMM-GUI. The energy-minimized structures were applied to the GROMACS 2021.2 (Abraham et al. 2015). The obtained structures were visualized with the PyMOL Molecular Graphics System Version 2.0, Schrödinger, LLC.

Establishment of cell lines expressing wild-type and glycan-deficient mutant ErbB2

To eliminate the effects of endogenous ErbB2, we established Flp-In HEK293 cells in which endogenous ErbB2 was knocked out by the CRISPR/Cas9 system. The complex of synthetic gRNA containing the sequence 5′-UCAUCGCUCACAA CCAA-3′ (Thermo Fisher Scientific) and Cas9 protein (Thermo Fisher Scientific) was transfected into the cells by Lipofectamine CRISPRMAX (Thermo Fisher Scientific), then clonal cell lines were obtained by limiting dilution. The expression vector for the glycan-deficient mutant ErbB2 (N124Q/N125Q) was generated by introducing mutations into the plasmid pcDNA5/FRT encoding the wild-type ErbB2. The mutations were introduced in two steps using specific primer sets and PrimeSTAR MAX DNA polymerase (Takara Bio, Kusatsu, Japan). First, the N124Q mutation was introduced into the plasmid pcDNA5/FRT encoding wild-type ErbB2 using the following primer set:

Forward primer: 5′-CCGCTGCAAAATACCACCCCTGT CACA-3′

Reverse primer: 5′-GGGTATTTTGCAGCGGGTCTCCATT GTC-3′.

Then, the N125Q mutation was introduced into the plasmid carrying the N124Q mutation using the following primer set:

Forward primer: 5′-CTGCAACAAACCACCCCTGTCACA GGGG-3′

Reverse primer: 5′-GGGTGGTTTGTTGCAGCGGGTCT CCATT-3′.

To establish the cell lines stably expressing wild-type and glycan-deficient mutant (N124Q/N125Q) ErbB2, the expression vectors pcDNA5/FRT encoding wild-type ErbB2 and the mutant ErbB2 (N124Q/N125Q) were co-transfected with the pOG44 vector into ErbB2 knockout HEK293 cells. Stable expression cell lines were selected by treating the cells with hygromycin B, and clonal cell lines were obtained by performing a limiting dilution.

Immunofluorescent staining

ErbB2-KO 293 cells stably expressing wild-type and ErbB2_N 124Q/N125Q ErbB2s were on chamber slides (Thermo Fisher Scientific), fixed with ice-cold methanol, and blocked with 5% (w/v) bovine serum albumin in PBS containing 22.52 mg/mL of glycine and 0.1% (v/v) Tween-20 for 1 h at room temperature. Blocked cells were stained with rabbit anti-ErbB2 (#2165, Cell Signaling Technology; 1:500) and mouse anti-protein disulfide isomerases (#45596, Cell Signaling Technology; 1:500) antibodies for overnight at 4 °C. Following washing with PBS, goat anti-rabbit and anti-mouse secondary antibodies conjugated Alexa Fluor594 and 488 (Abcam #ab50080 and #ab50113; 1:1000), respectively, were added and incubated for 1 h at room temperature. After washing, cells were encapsulated in ProLong Gold antifade regent containing DAPI (Themo Fisher Scientific). Confocal cellular images were captured with Zeiss LSM 780 laser confocal microscope (Carl Zeiss, Oberkochen, Germany).

Author contributions

Naoki Fujitani (Conceptualization [equal], Data curation [lead], Formal analysis [lead], Funding acquisition [equal], Investigation [lead], Methodology [lead], Project administration [equal], Supervision [equal], Writing—original draft [lead]), Yasuaki Uehara (Data curation [equal], Formal analysis [supporting], Investigation [supporting], Supervision [equal], Writing—review & editing [equal]), Shigeru Ariki (Investigation [supporting], Project administration [supporting], Supervision [supporting], Writing—review & editing [equal]), Ukichiro Hashimoto (Data curation [supporting], Investigation [supporting]), Jo Mukai (Investigation [supporting]), Yoshihiro Hasegawa (Supervision [supporting]), and Motoko Takahashi (Conceptualization [equal], Funding acquisition [equal], Investigation [supporting], Project administration [lead], Supervision [lead], Writing—review & editing [supporting]).

Supplementary material

Supplementary material is available at *Glycobiology Journal* online.

Data availability

Mass spectrometric data have been deposited at the ProteomeXchange Consortium (https://www.proteomexchange.org/) via the PRIDE partner repository (Perez-Riverol et al. 2022) with the dataset identifiers PXD039834 and 10.6019/PXD039834 for HEK293 cells-expressed sErbB2, PXD039835 and 10.6019/PXD039835 for CHO cells-expressed sErbB2, PXD039836 and 10.6019/PXD039836 for CHO cells-expressed sEGFR, PXD039837 and 10.6019/PXD039837 for the immunoprecipitated ErbB2 from SKBR-3 cell lysate, and PXD039838 and 10.6019/PXD039838 for the immunoprecipitated ErbB2 from BT474 cell lysate.

References

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015:**1-2**:19–25.

Bouyain S, Longo PA, Li S, Ferguson KM, Leahy DJ. The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. *Proc Natl Acad Sci U S A*. 2005:**102**(42):15024–15029.

Casalino L, Gaieb Z, Goldsmith JA, Hjorth CK, Dommer AC, Harbison AM, Fogarty CA, Barros EP, Taylor BC, McLellan JS, et al. Beyond shielding: the roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent Sci*. 2020:**6**(10):1722–1734.

Chandler KB, Leon DR, Meyer RD, Rahimi N, Costello CE. Site-specific N-glycosylation of endothelial cell receptor tyrosine kinase VEGFR-2. *J Proteome Res*. 2017:**16**(2):677–688.

Chandler KB, Leon DR, Kuang J, Meyer RD, Rahimi N, Costello CE. N-glycosylation regulates ligand-dependent activation and signaling of vascular endothelial growth factor receptor 2 (VEGFR2). *J Biol Chem*. 2019:**294**(35):13117–13130.

Cho HS, Leahy DJ. Structure of the extracellular region of HER3 reveals an interdomain tether. *Science*. 2002:**297**(5585):1330–1333.

Cho HS, Mason K, Ramyar KX, Stanley AM, Gabelli SB, Denney DW Jr, Leahy DJ. Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature*. 2003:**421**(6924):756–760.

Citri A, Yarden Y. EGF-ERBB signalling: towards the systems level. *Nat Rev Mol Cell Biol*. 2006:**7**(7):505–516.

Contessa JN, Bhojani MS, Freeze HH, Rehemtulla A, Lawrence TS. Inhibition of N-linked glycosylation disrupts receptor tyrosine kinase signaling in tumor cells. *Cancer Res*. 2008:**68**(10):3803–3809.

Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.- range mass accuracies and proteome- wide protein quantification. *Nat Biotechnol*. 2008:**26**(12):1367–1372.

Crocker PR, Paulson JC, Varki A. Siglecs and their roles in the immune system. *Nat Rev Immunol*. 2007:**7**(4):255–266.

Duarte HO, Rodrigues JG, Gomes C, Hensbergen PJ, Ederveen ALH, de Ru AH, Mereiter S, Polónia A, Fernandes E, Ferreira JA, et al. ST6Gal1 targets the ectodomain of ErbB2 in a site-specific manner and regulates gastric cancer cell sensitivity to trastuzumab. *Oncogene*. 2021:**40**(21):3719–3733.

Duarte HO, Reis CA, Gomes J. Insights on ErbB glycosylation - contributions to precision oncology. *Trends Cancer*. 2022:**8**(6):448–455.

Fang Z, Qin H, Mao J, Wang Z, Zhang N, Wang Y, Liu L, Nie Y, Dong M, Ye M. Glyco-decipher enables glycan database-independent peptide matching and in-depth characterization of site-specific N-glycosylation. *Nat Commun*. 2022:**13**(1):1900.

Ferguson KM, Berger MB, Mendrola JM, Cho HS, Leahy DJ, Lemmon MA. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol Cell*. 2003:**11**(2):507–517.

Fujitani N, Ariki S, Hasegawa Y, Uehara Y, Saito A, Takahashi M. Integrated structural analysis of N-glycans and free oligosaccharides allows for a quantitative evaluation of ER stress. *Biochemistry*. 2021:**60**(21):1708–1721.

Gschwind A, Fischer A, Ullrich A. The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nat Rev Cancer*. 2004:**4**(5):361–370.

Hanash SM, Pitteri SJ, Faca VM. Mining the plasma proteome for cancer biomarkers. *Nature*. 2008:**452**(7187):571–579.

Hasegawa Y, Takahashi M, Ariki S, Asakawa D, Tajiri M, Wada Y, Yamaguchi Y, Nishitani C, Takamiya R, Saito A, et al. Surfactant protein D suppresses lung cancer progression by downregulation of epidermal growth factor signaling. *Oncogene*. 2015:**34**(7):838–845.

Huang YL, Liang CY, Labitzky V, Ritz D, Oliveira T, Cumin C, Estermann M, Lange T, Everest-Dass AV, Jacob F. Site-specific N-glycosylation of integrin $\alpha 2$ mediates collagen-dependent cell survival. *iScience*. 2021:**24**(10):103168.

Hudak JE, Canham SM, Bertozzi CR. Glycocalyx engineering reveals a Siglec-based mechanism for NK cell immunoevasion. *Nat Chem Biol*. 2014:**10**(1):69–75.

Hynes NE, Lane HA. ERBB receptors and cancer: the complexity of targeted inhibitors. *Nat Rev Cancer*. 2005:**5**(5):341–354.

Kaplan M, Narasimhan S, de Heus C, Mance D, van Doorn S, Houben K, Popov-Čeleketić D, Damman R, Katrukha EA, Jain P, et al. EGFR dynamics change during activation in native membranes as revealed by NMR. *Cell*. 2016:**167**(5):1241–1251.e11.

Klapper LN, Glathe S, Vaisman N, Hynes NE, Andrews GC, Sela M, Yarden Y. The ErbB-2/HER2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors. *Proc Natl Acad Sci U S A*. 1999:**96**(9):4995–5000.

Kumagai S, Koyama S, Nishikawa H. Antitumour immunity regulated by aberrant ERBB family signalling. *Nat Rev Cancer*. 2021:**21**(3):181–197.

Lee HS, Qi Y, Im W. Effects of N-glycosylation on protein conformation and dynamics: protein data bank analysis and molecular dynamics simulation study. *Sci Rep*. 2015:**5**(1):8926.

Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. *Cell*. 2010:**141**(7):1117–1134.

Liu YC, Yen HY, Chen CY, Chen CH, Cheng PF, Juan YH, Chen CH, Khoo KH, Yu CJ, Yang PC, et al. Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. *Proc Natl Acad Sci U S A*. 2011:**108**(28):11332–11337.

Lopez Sambrooks C, Baro M, Quijano A, Narayan A, Cui W, Greninger P, Egan R, Patel A, Benes CH, Saltzman WM, et al. Oligosaccharyltransferase inhibition overcomes therapeutic resistance to EGFR tyrosine kinase inhibitors. *Cancer Res*. 2018:**78**(17):5094–5106.

Lu C, Mi LZ, Grey MJ, Zhu J, Graef E, Yokoyama S, Springer TA. Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal growth factor receptor. *Mol Cell Biol*. 2010:**30**(22):5432–5443.

Mechref Y. Use of CID/ETD mass spectrometry to analyze glycopeptides. *Curr Protoc Protein Sci*. 2012:**68**(1):12.11.1–12.11.11.

Meric-Bernstam F, Johnson AM, Dumbrava EEI, Raghav K, Balaji K, Bhatt M, Murthy RK, Rodon J, Piha-Paul SA. Advances in HER2-targeted therapy: novel agents and opportunities beyond breast and gastric cancer. *Clin Cancer Res*. 2019:**25**(7):2033–2041.

Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol*. 2012:**13**(7):448–462.

Park SJ, Lee J, Qi Y, Kern NR, Lee HS, Jo S, Joung I, Joo K, Lee J, Im W. CHARMM-GUI glycan modeler for modeling and simulation

of carbohydrates and glycoconjugates. *Glycobiology*. 2019:**29**(4): 320–331.

Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S, Kamatchinathan S, Kundu DJ, Prakash A, Frericks-Zipper A, Eisenacher M, et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res*. 2022:**50**(D1):D543–D552.

Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. *Nat Rev Nephrol*. 2019:**15**(6):346–366.

Saitou A, Hasegawa Y, Fujitani N, Ariki S, Uehara Y, Hashimoto U, Saito A, Kuronuma K, Matsumoto K, Chiba H, et al. N-glycosylation regulates MET processing and signaling. *Cancer Sci*. 2022:**113**(4):1292–1304.

Schneider M, Al-Shareffi E, Haltiwanger RS. Biological functions of fucose in mammals. *Glycobiology*. 2017:**27**(7):601–618.

Scott DA, Casadonte R, Cardinali B, Spruill L, Mehta AS, Carli F, Simone N, Kriegsmann M, Del Mastro L, Kriegsmann J, et al. Increases in tumor N-glycan polylactosamines associated with advanced HER2-positive and triple-negative breast cancer tissues. *Proteomics Clin Appl*. 2019:**13**(1):e1800014.

Shajahan A, Supekar NT, Gleinich AS, Azadi P. Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology*. 2020:**30**(12):981–988.

Shajahan A, Archer-Hartmann S, Supekar NT, Gleinich AS, Heiss C, Azadi P. Comprehensive characterization of N- and O- glycosylation of SARS-CoV-2 human receptor angiotensin converting enzyme 2. *Glycobiology*. 2021:**31**(4):410–424.

Stanczak MA, Rodrigues Mantuano N, Kirchhammer N, Sanin DE, Jacob F, Coelho R, Everest-Dass AV, Wang J, Trefny MP, Monaco G, et al. Targeting cancer glycosylation repolarizes tumor-associated macrophages allowing effective immune checkpoint blockade. *Sci Transl Med*. 2022:**14**(669):eabj1270.

Takahashi M, Hasegawa Y, Ikeda Y, Wada Y, Tajiri M, Ariki S, Takamiya R, Nishitani C, Araki M, Yamaguchi Y, et al. Suppression of heregulin β signaling by the single N-glycan deletion mutant of soluble ErbB3 protein. *J Biol Chem*. 2013:**288**(46):32910–32921.

Takahashi M, Kizuka Y, Ohtsubo K, Gu J, Taniguchi N. Disease-associated glycans on cell surface proteins. *Mol Asp Med*. 2016:**51**: 56–70.

Varki A. Biological roles of glycans. *Glycobiology*. 2017:**27**(1):3–49.

Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, Stanley P, Hart G, Darvill A, Kinoshita T, et al. Symbol nomenclature for graphical representations of glycans. *Glycobiology*. 2015:**25**(12):323–324.

Wang Z, Cheng L, Guo G, Cheng B, Hu S, Zhang H, Zhu Z, Niu L. Structural insight into a matured humanized monoclonal antibody HuA21 against HER2-overexpressing cancer cells. *Acta Crystallogr D Struct Biol*. 2019:**75**(6):554–563.

Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science*. 2020:**369**(6501): 330–333.

Webb B, Sali A. Comparative protein structure modeling using MOD-ELLER. *Curr Protoc Bioinformatics*. 2016:**54**(1):5.6.1–5.6.37.

Wu SL, Kim J, Bandle RW, Liotta L, Petricoin E, Karger BL. Dynamic profiling of the post-translational modifications and interaction partners of epidermal growth factor receptor signaling after stimulation by epidermal growth factor using extended range proteomic analysis (ERPA). *Mol Cell Proteomics*. 2006:**5**(9):1610–1627.

Xiao H, Woods EC, Vukojicic P, Bertozzi CR. Precision glycocalyx editing as a strategy for cancer immunotherapy. *Proc Natl Acad Sci U S A*. 2016:**113**(37):10304–10309.

Yang W, Ao M, Hu Y, Li QK, Zhang H. Mapping the O-glycoproteome using site-specific extraction of O-linked glycopeptides (EXoO). *Mol Syst Biol*. 2018:**14**(11):e8486.

Yarden Y, Pines G. The ERBB network: at last, cancer therapy meets systems biology. *Nat Rev Cancer*. 2012:**12**(8):553–563.

Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol*. 2001:**2**(2):127–137.