Glycopeptidomics Analysis of a Cell Line Model Revealing Pathogenesis and Potential Marker Molecules for the Early Diagnosis of Gastric MALT Lymphoma

Di Xiao[1]*[†], Qinghua Zou[2][†], Le Meng[3], Yanli Xu[4], Huifang Zhang[1], Fanliang Meng[1], Lihua He[1] and Jianzhong Zhang[1]*

[1] State Key Laboratory of Infectious Disease Prevention and Control, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China, [2] Department of Microbiology, School of Basic Medical Sciences, Peking University Health Science Center, Beijing, China, [3] Center for Biomedical Mass Spectrometry, Department of Biochemistry, Boston University, Boston, MA, United States, [4] Hebei University of Engineering, Affiliated Hospital, College of Medicine, Handan, China

Background & Aims: Gastric mucosa-associated lymphoma (GML) is a mature B cell tumor related to *Helicobacter pylori* (*H.pylori*) infection. The clinical manifestations of GML are not specific, so GML is often misdiagnosed, leading to excessive treatment. The pathogenesis of *H.pylori*-induced GML is not well understood and there are no molecular markers for early GML diagnosis.

Methods: Glycopeptidomics analyses of host cell lines (a BCG823 cell line, C823) and C823 cells infected by *H. pylori* isolated from patients with GML (GMALT823), gastritis (GAT823), gastric ulcer (GAU823) and gastric cancer (GAC823) were carried out to clarify the host reaction mechanism against GML and to identify potential molecular criteria for the early diagnosis of GML.

Results: Thirty-three samples were analyzed and approximately 2000 proteins, 200 glycoproteins and 500 glycopeptides were detected in each sample. O-glycans were the dominant glycoforms in GMALT823 cells only. Four specific glycoforms in GMALT823 cells and 2 specific glycoforms in C823 and GMALT823 cells were identified. Eight specific glycopeptides from 7 glycoproteins were found in GMALT823 cells; of these glycopeptides, 6 and 3 specific glycopeptides had high affinity for T cell epitopes and have conformational B cell epitopes, respectively.

Conclusion: The predominant glycoforms of host cells infected by MALT *H. pylori* isolates differ from others, and the glycoproteins, glycosylation sites and glycoforms might be closely related to the formation of GML, which provides new insights into the pathogenic mechanisms of *H. pylori* infection and suggests molecular indicators for the early diagnosis of GML.

Keywords: gastric MALT lymphoma, *Helicobacter pylori*, cell line model, glycopeptidomics, diagnosis

INTRODUCTION

Helicobacter pylori (*H. pylori*) is a gram-negative, microaerophilic bacterium adapted for survival in the human stomach, where it can cause chronic gastritis, peptic ulcer disease, gastric mucosa-associated lymphoma (GML) and gastric adenocarcinoma. Fifty percent of the world's total population is estimated to be infected with *H. pylori* (Deng et al., 2013). Mucosa-associated lymphoid tissue (MALT) lymphoma is a low-grade malignant lymphoma. It usually occurs in the stomach, lungs and lacrimal glands, which have no lymphoid tissue. MALT lymphoma is derived from GML (Dogan et al., 1997), a mature B cell tumor related to *H. pylori* infection that accounts for 30%-50% of all extranodal lymphomas and 2-8% of all gastric cancers (Zullo et al., 2014; Moleiro et al., 2016). *H. pylori* infection is the main factor leading to GML. When GML cells were isolated and cultured by standard methods, all the cells died within 5 days, but following the addition of heat-inactivated whole-cell *H. pylori*, the lymphocytes proliferated (Raderer et al., 2016). The phenotype of *H. pylori* is conservative, but its gene mutation rate is high. The genetic polymorphism of *H. pylori* is associated with the clinical outcome of *H. pylori* infection (Saribasak et al., 2004). As a form of chronic antigen stimulation, persistent *H. pylori* colonization recruits immune lymphocytes that migrate to and infiltrate the site of *H. pylori* infection in the stomach, leading to the loss of regulation of B-lymphocyte proliferation and differentiation (de Jong et al., 1999). The clinical manifestations of GML are not specific. It is easy for MALT lymphoma to escape diagnosis or be misdiagnosed, and excessive treatment is common (Moleiro et al., 2016). The pathogenesis of *H. pylori*-induced GML is not well understood, although some immunologic mechanisms are thought to be involved (de Jong et al., 1999; Du and Isaccson, 2002). Studies to identify virulence factors or genetic markers of *H. pylori* strains associated with GML have also been performed. Whole-genome comparisons showed that *H. pylori* strains isolated from patients with GML are different from those isolated from patients with gastritis and ulcers (Wang et al., 2015). In our previous studies, we also found there were great difference in the transcriptome between the MALT lymphoma isolates and gastritis isolate 26695. The differentially transcribed genes were mainly with functions in amino sugar and nucleotide sugar metabolism, porphyrin and chlorophyll metabolism, and glyoxylate and dicarboxylate metabolism (Wang et al., 2020). At present, there is no molecular marker for GML diagnosis. The diagnosis of GML mainly depends on pathological diagnosis. Therefore, it is important and of theoretical, clinical, and practical value to clarify the host reaction mechanism and establish molecular criteria for the early diagnosis of GML. Because *H. pylori* strains are difficult to colonize in the stomach of mice, research on *H. pylori* cannot be completed in animal models. Therefore, with the help of a host cell model, the molecular mechanism of *H. pylori* alone can be examined. More than 90% of known clinical diagnostic markers are glycoproteins. Protein glycosylation is one of the most common and functionally important forms of posttranslational modifications and plays a central role in many biological processes and pathways (Dube and Bertozzi, 2005; Napoletano et al., 2007; Hoseki et al., 2010;

Liu et al., 2011; Pinho and Reis, 2015). In this study, glycopeptidomics analyses of host cell models infected with *H. pylori* from different sources were carried out, and this study aimed to reveal the biological effects of GML-related isolates on host cells and identify potential early diagnostic markers among these glycopeptides.

METHODS

Bacterial Strains and Phenotypic Identification

A total of 32 *H. pylori* strains (2 isolated from GML patients, 10 each isolated from gastritis patients, gastric ulcer patients and gastric cancer patients) from the *H. pylori* strain library of the China CDC were used in this study. We performed both transcriptomic and proteomic analyses on gastric cancer cells infected by the two *H. pylori* isolates from GML patients and the reference strain 26695 in our previous studies, and found specific differentially expressed genes and differentially expressed proteins induced by the GML isolates (Zou et al., 2020). The *H. pylori* strains were identified by microscopic examination and urease, catalase, and oxidase activity tests. All strains were grown on 10% (vol/vol) Columbia blood agar (CM0331, OXOID) plates and inoculated at 37°C for 24 to 48 h under microaerophilic conditions (5% O_2, 5% CO_2, 2% H_2, and 88% N_2).

Sample Preparation

In this study, we used a human gastric carcinoma undifferentiated cell line BCG823 which have been used in previous studies to reveal the proteome variation infected by MALT isolates and other isolates (Zou et al., 2020). The cell line was cultured in RPMI 1640 supplemented with 10% FBS at 37°C in an atmosphere of 5% CO_2. BCG823 cells were infected with *H. pylori* at a bacteria/BCG823 cell ration of 100:1 for 4 h. For convenience, the BCG823 cell line and BCG823 cell lines infected with *H. pylori* isolates from the gastric mucosa of patients with gastritis, gastric ulcer, gastric cancer and GML were named C823, GAT823, GAU823, GAC823 and GMALT823 cells, respectively.

Protein extraction, concentration determination and trypsin digestion and peptide purification of BCG823 cell line and BCG823 cell lines infected with *H. pylori* were performed using TMT Kit and BCA Protein Assay Kit (Thermo Fisher Scientific).

Intact glycopeptide enrichment was performed *via* hydrophilic interaction liquid chromatography (HILIC) (The Nest Group, Inc.). Briefly, the tryptic and desalted peptides were resuspended in 80% ACN. The appropriate amounts of HILIC particle in 80% ACN were placed in Pierce spin columns (Thermo Scientific) and equilibrated three times using 80% ACN, which was followed by sample loading three times and washing two times with 80% ACN. Then, glycopeptides bound to the HILIC column were eluted three times with 100 μL of 0.1% TFA. The samples were dried by a SpeedVac and stored at -80°C until analysis.

Nano-HPLC-MS/MS Analysis

The dried and desalted glycopeptides were reconstituted in 0.1% formic acid (FA) and separated on a nanoAcquity ultraperformance

liquid chromatography (UPLC) system and an EASY-nLC 1000 system (Thermo Scientific) fitted with a nanoAcquity Symmetry C18 trap column (100 μm×2 cm, nanoViper C18, 5 μm, 100 Å) and an analytical column (75 μm×15 cm, nanoViper C18, 3 μm, 100 Å). Mobile phase A was 100:0.1 HPLC-grade water/FA, and mobile phase B was 80:20:0.1 ACN/HPLC-grade water/FA. Each sample was loaded on the trapping column at a 2.0 μL/min flow rate and then separated on the analytical column using a 100-minute 3-35% mobile phase B linear gradient at a 0.8 μL/min flow rate. A retention time calibration mixture (Thermo Scientific) was used to optimize liquid chromatography (LC) and mass spectrometry (MS) parameters and monitor the stability of the system.

The analytical column was coupled to a high-resolution Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific, San Jose, CA) with a nanoelectrospray ion source operated in positive ion mode. The source was operated at 2.0 kV with the transfer capillary temperature maintained at 250°C and the S-lens RF level set at 60. MS spectra were obtained by scanning over an m/z range of 350–2000. Mass spectra in both MS and MS/MS were acquired in an Orbitrap mass analyzer with 1 microscan per spectrum. The resolving power for MS and MS/MS was set at 70,000 and 17,500, respectively. Tandem MS data on the top 20 most abundant multiply charged precursors were acquired in parallel with MS data, with higher energy collisional dissociation (HCD) at a normalized collision energy of 30 V. Precursors were isolated using a 2.0 m/z window, and dynamic exclusion of 60 s was enabled during precursor selection.

Data Analysis
Protein and Glycopeptide Identification
Database searches were performed using Byonic software (v2.13.17, Protein Metrics, Inc.). The following parameters were set for the search: cleavage sites, RK; cleavage side, C-terminal; digestion specificity, fully specific; missed cleavages, 2; precursor mass tolerance, 10 ppm; fragmentation type, QTOF/HCD; fragment mass tolerance, 0.02 Da; modifications, carbamidomethyl and oxidation; glycans, 182 human N-glycans with no multiple fucose moieties in the N-glycan database and 70 human O-glycans in the O-glycan database; and protein false discovery rate (FDR), 1% FDR (or 20 reverse counts). All the other settings were set at their default values.

Glycopeptide, Glycoform and Glycosylation Analysis
Byonic scores reflect the absolute quality of the peptide-spectrum match and not the relative quality compared to other candidate peptides. The Byonic score ranges from 0 to approximately 1000, with 300 being a good score, 400 a very good score, and peptide-spectrum matches with scores over 500 almost certainly correct. The DeltaMod value indicates whether modifications are confidently localized; DeltaMod values over 10 indicate a high likelihood that all modification placements are correct. Therefore, a score over 300, a DeltaMod value over 10, a q-value<0.05, and an FDR <0.1% were set as thresholds in this study. Systematic and comprehensive analyses of specific glycopeptides, glycoforms and glycosylation sites related to GMALT823 from all the proteins identified by Byonic were carried out.

Ingenuity Pathway Analysis (IPA)
The commercial software IPA was used to analyze proteins related to the specific glycopeptides. The diseases and functions along with the upstream regulatory factors and downstream regulatory factors of these proteins were analyzed by the BUILD function of IPA with a species selection of human. The OVERLAY tool was used to analyze the canonical pathways and biomarker information for the proteins.

T Cell Epitope and Conformational B Cell Epitope Prediction
T cell epitopes were identified using prediction tools at the Immune Epitope Database and Analysis Resource (IEDB-AR), a database of experimentally characterized immune epitopes (http://tools.immuneepitope.org). T cell epitopes were classified based on their binding affinity for human major histocompatibility complex (MHC) alleles using the half-maximal inhibitory concentration of a biological substance (IC_{50}) as the unit of measure. The threshold was set as follows: peptides with IC_{50} values <50 nM were considered high-affinity peptides, those with IC_{50} values <500 nM were considered intermediate-affinity peptides and those with IC_{50} values <5000 nM were considered low-affinity peptides. MHC class I T cell epitope prediction was performed using the artificial neural network (ANN) method (Tenzer et al., 2005), and MHC class II T cell epitope prediction was performed using the consensus method, a combination of the average relative binding matrix method and the stabilization matrix alignment method (SMM-align) (Wang et al., 2008). For conformational B cell epitope prediction, SWISS-MODEL Workspace and ElliPro were used for three-dimensional (3D) structure template modeling for selected proteins (Arnold et al., 2006; Ponomarenko et al., 2008). A template model of each protein was obtained by submitting protein sequences in FASTA format and modeling those sequences. ElliPro was used to predict conformational B cell epitopes from selected proteins using a modeled 3D structure template for each protein. The default values (a minimum score of 0.5 and a maximum distance of 6 angstroms) were selected for 3D structure prediction. Each predicted epitope formed by a group of amino acid residues was viewed with a Java viewer for chemical structures in 3D to illustrate the epitope 3D structure and relative orientation to the protein molecule.

Other Bioinformatic Methods
Signal peptides and transmembrane helices were predicted using the online tool Phobius (http://phobius.sbc.su.se/index.html). Cartoon representations of proteins and protein features, which were mapped with all the peptides identified from all MS/MS experiments, were generated using the online software tool Protter v.1.0 (Haserick et al., 2017). The N- and O-glycosylation sites of specific proteins were predicted using the online software NetNGlyc 1.0 Server and NetOGlyc 4.0 Server (Steentoft et al., 2013), respectively.

RESULTS

Protein and Glycopeptide Identification
The number of total proteins, the number of glycoproteins, the kinds of glycoforms, and the numbers of total peptides and

glycopeptides in each sample were determined by Byonic analysis and independent comprehensive analysis (**Table 1**).

The enrichment efficiency of glycoproteins and glycopeptides in the C823 cell line infected by *H. pylori* was approximately 12% and 7%, respectively, and the distribution trend of glycoproteins and glycopeptides was consistent (**Supplementary Figure 1**). For C823, GMALT823, GAT823, GAU823 and GAC823 cells, O-glycans were the dominant glycoprotein, and there were more O-glycans than N-glycans. However, the distributions of O- and N-glycans in the samples were different. There were more kinds of glycoforms with O-glycans than those with N-glycans in C823 and GMALT823 cells, and the number of O-glycoforms was 1.9 times and 2.1 times (on average) higher than that of N-glycoforms in C823 and GMALT823 cells, respectively. In contrast, there were more kinds of glycoforms with N-glycans than O-glycans in GAT823, GAU823, and GAC823 cells, and the number of N-glycoforms was 4.3 times (on average) higher than that of O-glycoforms, with the number of total glycoforms significantly lower than that in C823 cells (**Table 1**, **Figure 1**).

Analysis of Specific Glycopeptides, Glycoforms, and Glycosylation Sites

Two new glycosylation sites in GMALT823 cells were observed (these proteins and peptide sequences were both found in C823, GAT823, GAU823, and GAC823 cells but were not glycosylated); these glycosylation sites were the 497T site of protein disulfide isomerase A4 (PDIA4), in which the glycoform was HexNAc(2), and the 71S site of heterogeneous nuclear ribonucleoproteins A2/B1 (HNRNPA2B1), in which the glycoform was HexNAc(1). Two glycoforms of two glycosylation sites were absent in GMALT823 cells (these sites had these glycoforms in C823, GAT823, GAU823, and GAC823 cells, but the glycoforms of these sites were different in GMALT823 cells.); these were the 221N site of cathepsin L1 (CTSL), in which the glycoform was HexNAc(2)Hex(6), and the 69N site of galectin-3-binding protein (LGALS3), in which the glycoform was HexNAc(2)Hex(9) (**Table 2**, **Figures 2**, **3**).

Compared with the glycosylation sites in C823 cells, four glycosylation sites (O-glycan) in four proteins were consistently

TABLE 1 | Summary of glycosylation modifications.

Description	Name of cell line or isolate	Number		Number of glycosylated peptides* N-glycan	Kind of glycoform		
		Protein/ Peptide	Glycoprotein (O/N**)/ Glycopeptide		O-glycan	Total	
Cell line	BCG 823	2323/13899	229 (155/74)/539	134	14	26	40
MALT lymphoma isolates	MALT1	1825/10347	191 (131/60)/470	126	12	29	41
	H879	2571/16322	313 (185/128)/767	297	30	54	84
Gastritis-related *H. pylori* isolates	liuwanhua	915/3259	125 (64/61)/345	78	15	4	19
	zhanghaiying	835/3045	80 (46/34)/227	37	9	1	10
	HLJ208	1266/4831	176 (93/83)/436	86	17	4	21
	HZ00020	1297/5164	151 (86/65)/312	63	15	4	19
	YN3-77	1137/4346	164 (99/65)/403	98	20	4	24
	HZ00073	1169/4324	144 (80/64)/314	31	9	2	11
	YN3-35	1229/4536	151 (84/67)/367	126	22	12	34
	1-192	1440/8608	57 (42/15)/115	42	13	8	21
	26695	1868/11204	131 (81/50)/322	73	12	8	20
	XA91	576/1777	91 (49/42)/228	32	9	2	11
Gastric ulcer-related *H. pylori* isolates	XA203	929/3469	92 (48/44)/215	45	9	4	13
	XA137	855/2198	72 (41/31)/205	39	8	2	10
	HZ00124	1119/4589	142 (78/64)/372	89	16	8	24
	YN4-102	1156/4621	136 (81/55)/292	48	10	3	13
	HZ00058	1287/5206	135 (75/60)/381	89	23	5	28
	HLJ253	1387/5791	148 (81/67)/389	107	22	11	33
	wumingqun	1098/3893	137 (69/68)/330	57	14	1	15
	gaosonggui	1138/4355	160 (86/74)/388	74	13	4	17
	YN1-122	1253/4413	173 (96/77)/348	64	12	3	15
	4	1470/6888	168 (88/80)/339	92	16	5	21
Gastric cancer-related *H. pylori* isolates	HLJ011	931/3517	102 (52/50)/260	45	11	5	16
	HLJ021	525/1585	62 (34/28)/169	33	9	3	12
	Zhou183	1239/4923	165 (85/80)/440	88	21	5	26
	YTC-6	931/3256	134 (69/65)/377	78	13	4	17
	YTC-3	632/3121	71 (40/31)/225	40	11	2	13
	YTC-9	853/2769	120 (67/53)/270	49	11	3	14
	BJWJ-18	884/3174	143 (76/67)/315	61	15	2	17
	BJWJ-11	998/3404	114(58/56)/272	62	11	5	16
	HZ00066	895/2616	108 (62/46)/247	41	9	4	13
	HLJ005	1180/6551	44 (29/15)/90	34	13	1	14

Data that did not satisfy the cutoff values in this study (score over 300, DeltaMod value over 10, q-value<0.05, FDR <0.1%) were removed. Some peptides and glycosylation sites appeared more than two times due to the collection of different scans. These peptides were incorporated into one entry.
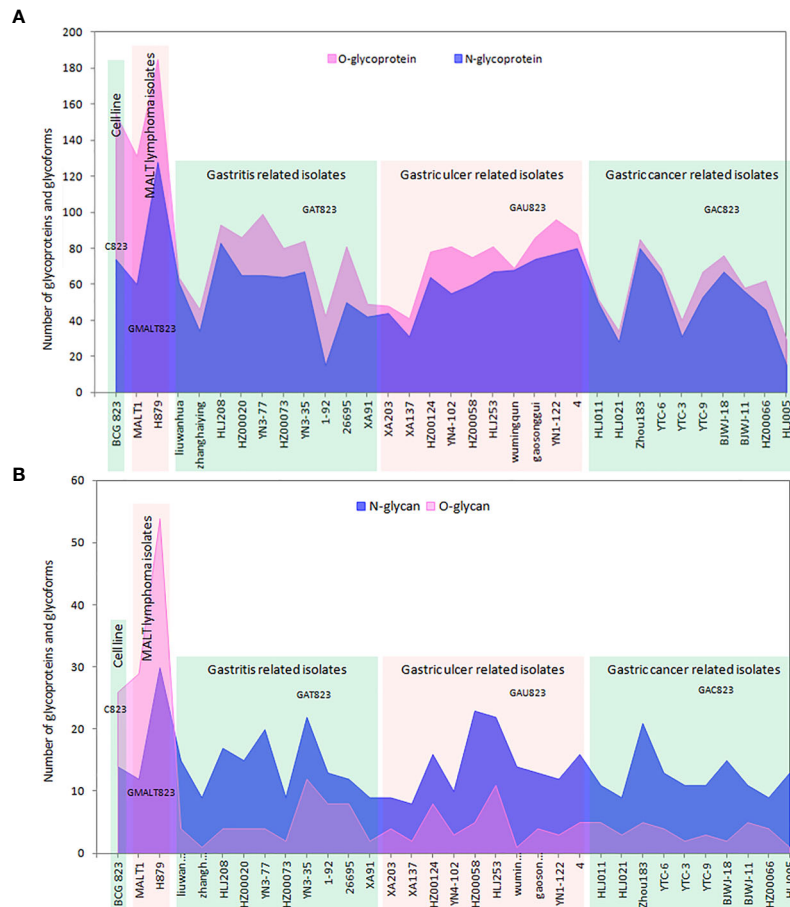***The number of glycoproteins, O-glycoproteins and N-glycoproteins.*

FIGURE 1 | Summary of the number of glycoproteins and glycoforms of all the samples used in this study. **(A)** The distribution of the numbers of glycoproteins; there are more O-glycans than N-glycans, and O-glycans are the dominant type of glycoprotein. **(B)** The distribution of the numbers of glycoforms; there are more O-glycans than N-glycans in C823 and GMALT823 cells, but in GAT823, GAU823 and GAC823 cells, the opposite is true.

found in GMALT823 cells. Protein microtubule-associated tumor suppressor 1 (MTUS1) was not expressed in GAU823 and GAC823 cells, but in C823 and GMALT823 cells, two glycoforms HexNAc(2)Hex(1)NeuAc(2) and HexNAc(1)Hex(1)NeuAc(1) at the 734S site were stable. The 259S site of the HNRNPA2B1 protein, which was deglycosylated in GAT823, GAU823 and GAC823 cells, was the only glycosylation site with the glycoform HexNAc(2)Hex(2)Fuc(2)NeuAc(1) in C823 and GMALT823 cells. The glycoform of the 45S site of protein calnexin (CANX) was different or deglycosylated in GAT823, GAU823 and GAC823 cells, but the glycoform HexNAc(1)Hex(1)NeuAc(2) at this site was stable in C823 and GMALT823 cells. The glycoform of the 271T site of nuclease-sensitive element-binding protein 1 (YBX1) was different in GAT823, GAU823 and GAC823 cells, but the glycoform HexNAc(1)Hex(2)Fuc(1) was stable in C823 and GMALT823 cells (**Table 2**, **Figures 2**, **3**). The glycosylation sites 497T of PDIA4 and the 71S and 259S sites of HNRNPA2B1 were previously unreported glycosylation sites found in this study (**Table 2**).

IPA Analysis

The 7 proteins related to specific glycopeptides were analyzed using IPA. The related diseases and functions, upstream regulatory factors and downstream regulatory factors, canonical pathways, biomarkers, roles in the cell, Entrez Gene summary and molecular function of these proteins are listed in **Supplementary Table 1**.

T Cell Epitope and Conformational B Cell Epitope Prediction

For T cell epitope prediction, the predicted output was given in units of IC_{50} nM for combinatorial library and SMM_align. Therefore, a lower number indicated a higher affinity. Here, peptides with IC_{50} values <50 nM (high affinity) and <500 nM (intermediate affinity) are listed in **Supplementary Table 2**. The glycopeptide MEPEEFDSDTLREFVTFKK from protein PDIA4 was predicted to have high-affinity and intermediate-affinity MHC class I epitopes and intermediate-affinity MHC class II epitopes. The glycopeptide FGFVTFSSMAEVDAAMAARPHSIDGRV

TABLE 2 | Summary of specific glycosylation modifications in the BCG823 cell line infected with MALT lymphoma-related *H. pylori* isolates.

Protein Name	Gene Name	Sequence	Glycoform	Description	Score	Delta Score
>sp|Q9ULD2| Microtubule-associated tumor suppressor 1	MTUS1	K.VGPPVS*[+1150.4024]C[+57.0214]LR.R	HexNAc(2)Hex(1)NeuAc(2)	This protein was not expressed in GAU823 and GAC823 cells. This site and glycoform were only found in C823 and GMALT823 cells.	434.5	35.6
		K.VGPPVS*[+656.2276]C[+57.0214]LR.R	HexNAc(1)Hex(1)NeuAc(1)		464	39.7
>sp|P22626| Heterogeneous nuclear ribonucleoproteins A2/B1	HNRNPA2B1	G.GNFGGS**[+1313.47563]PGYGGGR.G	HexNAc(2)Hex(2)Fuc(2)NeuAc(1)	This site was deglycosylated in GAT823, GAU823 and GAC823 cells.	466.4	61.2
>sp|P27824| Calnexin	CANX	V.IDIEDDLDDVIEEVEDS*[+947.3230]KPDTTAPPSSPK.V	HexNAc(1)Hex(1)NeuAc(2)	The glycoform of this site was changed or this site was deglycosylated in GAT823, GAU823 and GAC823 cells.	635.5	61.3
>sp|P67809| Nuclease-sensitive element-binding protein 1	YBX1	E.DKENQGDET*[+673.2429]QGQQPPQRR.Y	HexNAc(1)Hex(2)Fuc(1)	The glycoform of this site was different in GAT823, GAU823 and GAC823 cells.	612.2	474.5
>sp|P13667| Protein disulfide isomerase A4	PDIA4	M.EPEEFDSDT**[+406.15875]LREFVTAFK.K	HexNAc(2)	This protein and sequence were found in all samples, but glycosylation of this site was found in only GMALT823 cells.	309.3	195.9
>sp|P22626| Heterogeneous nuclear ribonucleoproteins A2/B1	HNRNPA2B1	F.GFVTFSS**[+203.07937]MAEVDAAMAARPHSIDGR.V	HexNAc(1)		556.1	389.0
>sp|P07711| Cathepsin L1	CTSL	K.YSVAN*[+1378.47569]DTGFVDIPKQEK.A	HexNAc(2)Hex(6)	All samples had this protein and sequence, but only GMALT823 cells lacked this glycoform on this glycosylation site.	543.8	355.0
>sp|Q08380|Galectin-3-binding protein	LGALS3	R.ALGFEN*[+1864.63416]ATQALGR.A	HexNAc(2)Hex(9)		524.5	257.6

*Indicates that this site is a glycosylation site by online prediction. ** Indicates that this site is not a glycosylation site by online prediction. The modifications with a gray background are stable glycosylation modifications in C823 and GMALT823 cells. The modifications with a white background are specific modifications in GMALT823 cells.

from the HNRNPA2B1 protein was predicted to have high-affinity and intermediate-affinity MHC class I and MHC class II epitopes. The glycopeptide GGNFGGSPGYGGGRG from the HNRNPA2B1 protein was predicted to have high-affinity and intermediate-affinity MHC class II epitopes. The specific glycopeptide sequences from the CTSL, LGALS3, and MTUS1 proteins were predicted to have intermediate-affinity MHC class II epitopes.
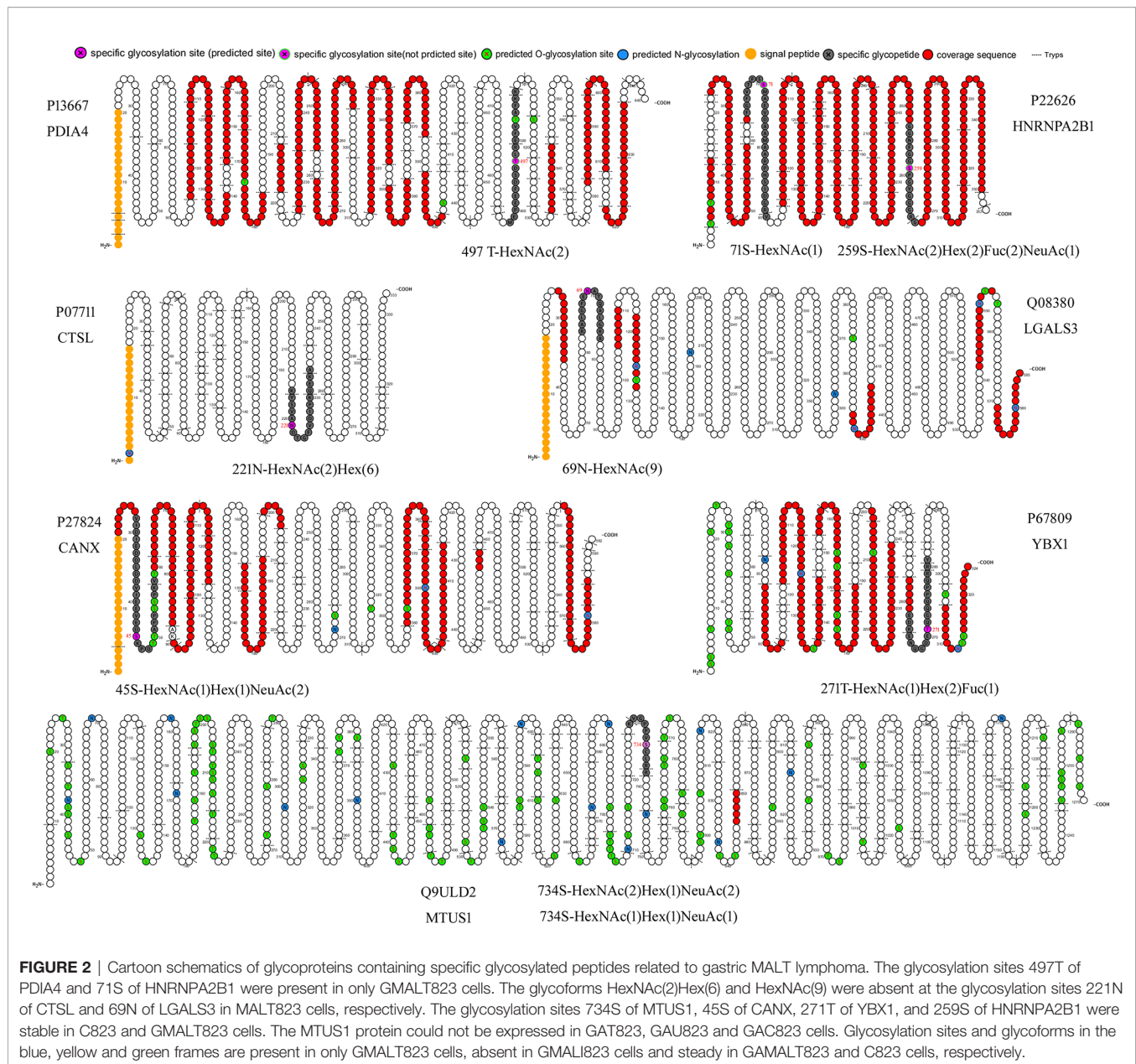
For conformational B cell epitope prediction, neither program found a suitable model for the specific glycopeptide sequences from the HNRNPA2B1, MTUS1, CANX and YBX1 proteins. The specific glycopeptide sequences MEPEEFDSDTLREFVTAFKK, KYSVANDTGFVDIPKQEKA and RALGFENATQALGRA from the PDIA4, CTSL, and LGALS3 proteins, respectively, were predicted to have conformational B cell epitopes (**Supplementary Table 3**). The 3D structures of B cell epitopes in specific glycopeptide sequences and model-template alignments of the PDIA4, CTSL, and LGALS3 proteins are shown in **Figure 4**.

DISCUSSION

Histological examination, immunohistochemistry and clonal analysis of B cells can be used for the pathological diagnosis of GML (Hu et al., 2016). To diagnose small lymphomas, which cannot be detected by the naked eye under endoscopy, the "observe and wait" method is usually adopted, after which a diagnosis can be made when the lymphoma is observed (Doglioni et al., 2011). Unlike other gastric cancers, early GML

is a reversible tumor, and the eradication of *H. pylori* at the early stage of GML causes 60-80% of MALT lymphomas to subside (Nobre-Leitao et al., 1998; Moleiro et al., 2016). Early detection and treatment of GML can greatly reduce the burden of this disease. Therefore, identifying molecular markers for the early diagnosis of GML has important clinical and social value.

Glycoforms are related to specific biological properties. Certain glycoforms are mass-produced in a disease-specific manner (Lee et al., 2015). Aberrant protein glycosylation has been observed during cancer development and progression. Given the functional role of aberrant glycosylation in cancer, many researchers have attempted to utilize specific glycoforms as cancer biomarkers for various clinical purposes (Kim et al., 2008; Lee et al., 2015). This study showed that the dominant glycoforms in the BCG823 cell line both before and after *H. pylori* infection are O-glycans. However, the glycoform type was different before and after infection with *H. pylori* from patients with different diseases. Before and after infection with MALT lymphoma isolates, the BCG823 cell line expressed more O-glycan glycoforms than N-glycan glycoforms; however, after infection with isolates from patients with gastritis, gastric ulcer and gastric cancer, the dominant glycoforms in BCG823 cell lines were N-glycan glycoforms. So whether a predominance of N-glycans in host cells may lead to the development of gastritis, gastric ulcers and gastric cancer, while a predominance of O-glycans in host cells may lead to the development of MALT lymphoma deserve further investigation. In addition, the differentially and stably expressed glycoforms in GMALT823 cells suggest that the HexNAc(2)Hex(6) glycoform on the 221N site of the CTSL protein or/and the

FIGURE 2 | Cartoon schematics of glycoproteins containing specific glycosylated peptides related to gastric MALT lymphoma. The glycosylation sites 497T of PDIA4 and 71S of HNRNPA2B1 were present in only GMALT823 cells. The glycoforms HexNAc(2)Hex(6) and HexNAc(9) were absent at the glycosylation sites 221N of CTSL and 69N of LGALS3 in MALT823 cells, respectively. The glycosylation sites 734S of MTUS1, 45S of CANX, 271T of YBX1, and 259S of HNRNPA2B1 were stable in C823 and GMALT823 cells. The MTUS1 protein could not be expressed in GAT823, GAU823 and GAC823 cells. Glycosylation sites and glycoforms in the blue, yellow and green frames are present in only GMALT823 cells, absent in GMALI823 cells and steady in GAMALT823 and C823 cells, respectively.

HexNAc(2)Hex(9) glycoform on the 69N site of the LGALS3 protein are absent in host cells in MALT lymphoma, in contrast, the HexNAc(2) glycoform at the 497T site of the PDIA4 protein and/or the HexNAc(1) glycoform at the 71S site of HNRNPA2B1 are detected only in MALT lymphoma. Therefore, the predominant types of glycoforms in host cells identified in this study are potential molecular markers to predict the developmental course of disease from *H. pylori* infection.

The pathogenesis of GML is not yet clear (Hu et al., 2016). Previous studies have elucidated some immunological and virulence factors and gene marker-related mechanisms involved in GML (Lin et al., 2010; Munari et al., 2011; Saito et al., 2012; Zullo et al., 2014). *H. pylori*-induced T cell-dependent B-cell

activation and deficient cytotoxic control of B-cell growth may link H. pylori infection, local T-cell response, and genesis of low-grade gastric MALT lymphoma (D'Elios et al., 1999). Our results provided some clues for revealing the pathogenesis of MALT lymphoma. First, among the eight specific glycopeptides in seven proteins found in this study, six T cell receptor epitopes and three conformational B cell epitopes mediate cellular and humoral immunity, respectively. Changes in specific glycosylation sites and glycoforms caused by infection with MALT lymphoma-related *H. pylori* isolates may affect the binding of T/B cell epitopes, thus positively or negatively affecting cellular or humoral immunity. Second, seven specific glycoproteins (MTUS1, HNRNPA2B1, CANX, YBX1, PDLA4, CTSL and
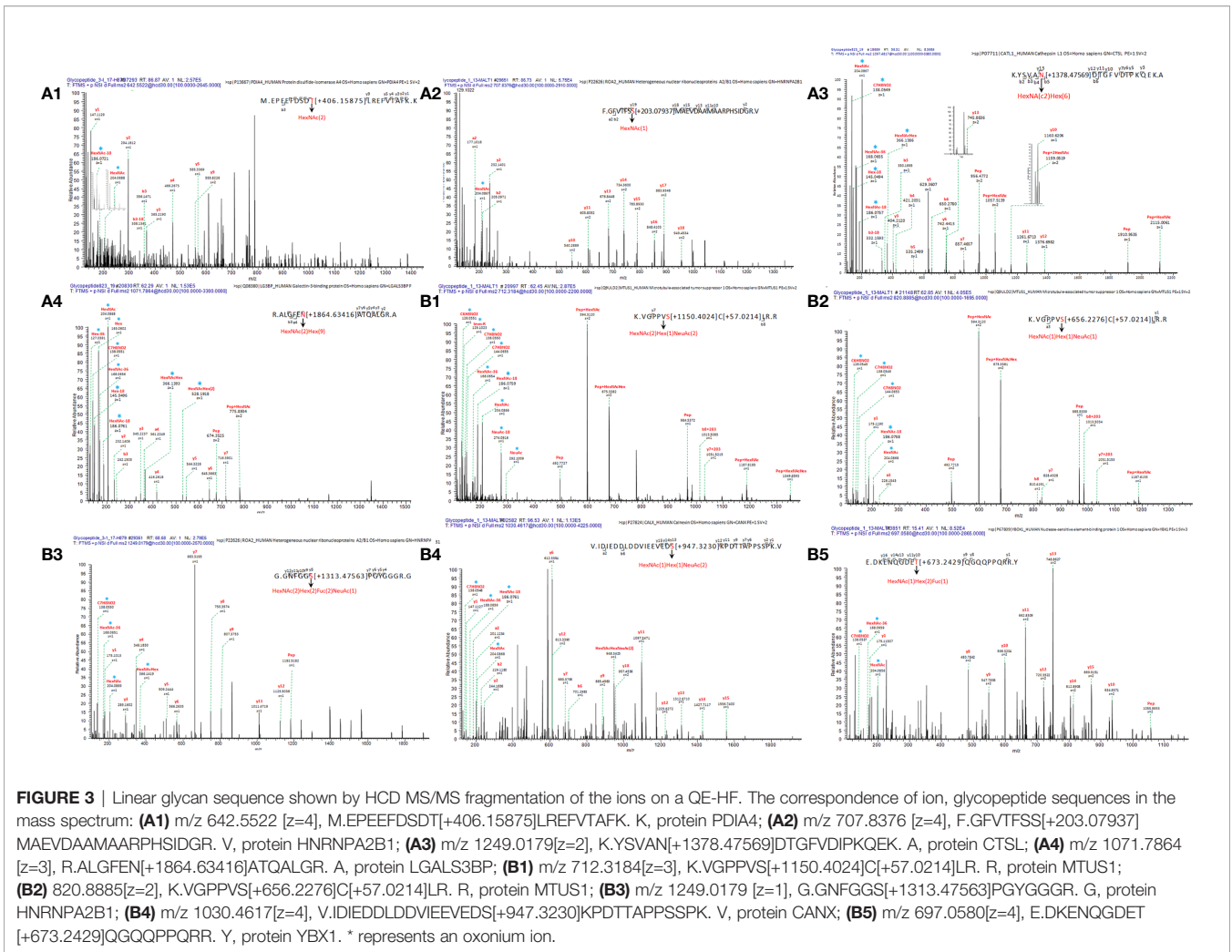
FIGURE 3 | Linear glycan sequence shown by HCD MS/MS fragmentation of the ions on a QE-HF. The correspondence of ion, glycopeptide sequences in the mass spectrum: **(A1)** m/z 642.5522 [z=4], M.EPEEFDSDT[+406.15875]LREFVTAFK. K, protein PDIA4; **(A2)** m/z 707.8376 [z=4], F.GFVTFSS[+203.07937] MAEVDAAMAARPHSIDGR. V, protein HNRNPA2B1; **(A3)** m/z 1249.0179[z=2], K.YSVAN[+1378.47569]DTGFVDIPKQEK. A, protein CTSL; **(A4)** m/z 1071.7864 [z=3], R.ALGFEN[+1864.63416]ATQALGR. A, protein LGALS3BP; **(B1)** m/z 712.3184[z=3], K.VGPPVS[+1150.4024]C[+57.0214]LR. R, protein MTUS1; **(B2)** 820.8885[z=2], K.VGPPVS[+656.2276]C[+57.0214]LR. R, protein MTUS1; **(B3)** m/z 1249.0179 [z=1], G.GNFGGS[+1313.47563]PGYGGGR. G, protein HNRNPA2B1; **(B4)** m/z 1030.4617[z=4], V.IDIEDDLDDVIEEVEDS[+947.3230]KPDTTAPPSSPK. V, protein CANX; **(B5)** m/z 697.0580[z=4], E.DKENQGDET [+673.2429]QGGQQPPQRR. Y, protein YBX1. * represents an oxonium ion.

LGALS3) were found in GMALT823 cells. MTUS1 is a tumor suppressor in lung cancer that promotes cell proliferation and migration (Gu et al., 2017). MTUS1 has a significant impact on the proliferation and metastatic potential of gastric cancer cell lines, as it has shown a potential anticancer effect in gastric cancer cell lines (Li et al., 2014). In this study, stable glycopeptides of the MTUS1 protein were found in host cells and host cells infected by MALT lymphoma isolates, but these glycopeptides did not express or exhibit specific glycopeptides in host cells infected by other *H. pylori* isolates. The specific glycopeptides and glycosylation sites of MTUS1 may play an important role in the development of MALT lymphoma. HNRNPA2B1 plays a direct role in cancer development, cancer progression, gene expression, and signal transduction. Knockdown of HNRNPA2B1 reduced breast cancer cell proliferation, induced apoptosis, and prolonged the S phase of the cell cycle *in vitro*. In addition, HNRNPA2B1 knockdown suppressed subcutaneous tumorigenicity *in vivo* (Hu et al., 2017). HNRNPA2B1 can specifically bind to the COX-2 core promoter and regulate the growth of non-small-cell tumors (Xuan et al., 2016). In this study, a new specific 71S glycosylation site with the glycoform HexNAc(1) was observed in the HNRNPA2B1

protein from host cells infected by MALT lymphoma isolates, and the 259S site with the glycoform HexNAc(2)Hex(2)Fuc(2) NeuAc(1) was expressed steadily in only host cells and host cells infected by MALT lymphoma isolates, suggesting that these two specific glycosylation sites play an important role in the pathogenesis of MALT lymphoma caused by *H. pylori* infection. CANX is a prognostic marker and potential therapeutic target in colorectal cancer (Ryan et al., 2016). In melanoma models, CANX knockout enhanced the infiltration and effector functions of T cells in the tumor microenvironment and inhibited tumor growth (Chen et al., 2019). YBX1 is involved in the antigen presentation pathway and phage maturation, which is related to gastric cancer and gastric epithelial cancer (**Supplementary Table 1**). YBX1 binds to HOXC-AS3 to mediate the tumorigenesis of gastric cancer (Zhang et al., 2018). The YBX1 protein is associated with cancer proliferation in numerous tissues, and its gene may be a prognostic marker for poor outcome and drug resistance in certain cancers (**Supplementary Table 1**). In this study, the glycoforms of the 45S site of CANX and the 271T site of YBX1 were stable in host cells infected with *H. pylori* isolates from patients with MALT lymphoma; however, glycoforms in these glycosylation sites

FIGURE 4 | Conformational B cell epitopes of specific glycosylated peptides. In the ball-and-stick models **(A)**, yellow balls are the residues of the predicted epitopes, and white sticks are the structures of nonepitopes and core residues. In the model-template alignment **(B)**, each epitope is shown with predicted residues and residue positions: P13667, PEEFDSDTLRE (490-500); P07711, QYVQDNGGLDSEESYPYEATEESCKYNPKYSVANDTG (188-224); and Q08380, LGFENA (65-70). The amino acid in the green frame is the glycosylation site.

changed or were deglycosylated in host cells infected with *H. pylori* isolates from patients with gastritis, gastric ulcers and gastric cancer. Changes in the glycosylation sites of these two glycoproteins may be related to the development of MALT lymphoma in cells. PDIA4 facilitates tumor cell growth *via* inhibiting the degradation and activation of procaspases 3 and 7 *via* their mutual interaction in a CGHC-dependent manner (Kuo et al., 2017). PDIA4 regulates MTTP (**Supplementary Table 1**), a major and an essential lipid transfer protein that transfers phospholipids and triacylglycerols to nascent ApoB for the assembly of lipoproteins (Sirwi and Hussain, 2018). Clinical studies have shown that the level of ApoB/ApoA1 in gastric cancer patients is negatively correlated with the survival rate, while PDIA4 is positively correlated with the survival rate of gastric cancer patients (Ma et al., 2018). In this study, a new specific 497T glycosylation site with the glycoform HexNAc(2) was observed in host cells infected by MALT lymphoma isolates, indicating that this glycosylation site may be an important factor in the development of MALT lymphoma in host cells. CTSL is

involved in autophagy and phage maturation-related pathways and associated with many cancers (**Supplementary Table 1**). CTSL expression may be linked to cancer grade, invasion and stage (Sudhan and Siemann, 2015). LGALS3 binds Galectin-3 (Gal3), which regulates its function. Gal3 plays an important role in innate immunity against infection and the colonization of *H. pylori*. Large lymphoid clusters consisting of mostly B cells have frequently been observed in the gastric submucosa of Gal3-deficient mice (Park et al., 2016), which is consistent with our results. Changes in the glycoforms on the glycosylation site of LGALS3 may lead to changes in its biological function, which may decrease the expression of Gal3 or cause a loss of biological function, resulting in B lymphocyte aggregation. This result suggests that deletion of the 221N glycosylation site of CTSL and the 69N glycosylation site of LGALS3 may be important factors in the development of MALT lymphoma. These seven glycoproteins can all be detected in bodily fluids (**Supplementary Table 1**).

Missed diagnosis and misdiagnosis of GML are common in the clinic. The overtreatment of early GML, such as treatment with

systemic chemotherapy, occurs in China and many developing countries. Therefore, it is of great theoretical and clinical value to clarify the host response mechanism and identify and establish molecular markers for the early diagnosis of GML. In this study, the relationship between the predominant glycoform of host cells and the development of host disease was determined by glycopeptidomics analysis of a cell line model. Seven glycoproteins, eight glycosylation sites and 9 glycoforms might be closely related to the formation of GML. Further studies on these glycoforms may provide new insight into the pathogenic mechanisms of *H. pylori* infection and molecular indicators for the early diagnosis of GML. A limitation of this study is the low number of *H. pylori* isolates from GML that are used. Future studies should include a larger number of strains isolated from GML patients to clarify if the changes observed are seen with all strains isolated from cases of GML.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by National Institute for Communicable Disease Control and Prevention Chinese Center for Disease Control and Prevention Ethical Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

DX and JZ conceived the study. DX, QZ, LM, YX, HZ, FM, and LH completed the bacterial and cell cultures, sample preparation, and data acquisition. DX, QZ, and LM did data analysis and bioinformatics. DX wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2021. 715454/full#supplementary-material

REFERENCES

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL Workspace: A Web-Based Environment for Protein Structure Homology Modelling. *Bioinformatics* 22 (2), 195–201. doi: 10.1093/bioinformatics/bti770

Chen, Y., Ma, D., Wang, X., Fang, J., Liu, X., Song, J., et al. (2019). Calnexin Impairs the Antitumor Immunity of CD4(+) and CD8(+) T Cells. *Cancer Immunol. Res.* 7 (1), 123–135. doi: 10.1158/2326-6066.CIR-18-0124

D'Elios, M. M., Amedei, A., Manghetti, M., Costa, F., Baldari, C. T., Quazi, A. S., et al. (1999). Impaired T-Cell Regulation of B-Cell Growth in Helicobacter Pylori–Related Gastric Low-Grade MALT Lymphoma. *Gastroenterology* 117 (5), 1105–1112. doi: 10.1016/s0016-5085(99)70395-1

de Jong, D., Aleman, B. M., Taal, B. G., and Boot, H. (1999). Controversies and Consensus in the Diagnosis, Work-Up and Treatment of Gastric Lymphoma: An International Survey. *Ann. Oncol.* 10 (3), 275–280. doi: 10.1023/a:1008392022152

Deng, B., Li, Y., Zhang, Y., Bai, L., and Yang, P. (2013). Helicobacter Pylori Infection and Lung Cancer: A Review of an Emerging Hypothesis. *Carcinogenesis* 34 (6), 1189–1195. doi: 10.1093/carcin/bgt114

Dogan, A., Du, M., Koulis, A., Briskin, M. J., and Isaacson, P. G. (1997). Expression of Lymphocyte Homing Receptors and Vascular Addressins in Low-Grade Gastric B-Cell Lymphomas of Mucosa-Associated Lymphoid Tissue. *Am. J. Pathol.* 151 (5), 1361–1369. doi: 10.1007/s004280050125

Doglioni, C., Ponzoni, M., Ferreri, A. J., Savio, A., Gruppo Italiano Patologi Apparato, D.I.d. Societa Italiana di Anatomia Patologica e Citopatologia Diagnostica/International Academy of Pathology. (2011). Gastric Lymphoma: The Histology Report. *Dig. Liver Dis.* 43 (Suppl 4), S310–S318. doi: 10.1016/S1590-8658(11)60587-2

Dube, D. H., and Bertozzi, C. R. (2005). Glycans in Cancer and Inflammation–Potential for Therapeutics and Diagnostics. *Nat. Rev. Drug Discov.* 4 (6), 477–488. doi: 10.1038/nrd1751

Du, M. Q., and Isaccson, P. G. (2002). Gastric MALT Lymphoma: From Aetiology to Treatment. *Lancet Oncol.* 3 (2), 97–104. doi: 10.1016/s1470-2045(02)00651-4

Gu, Y., Liu, S., Zhang, X., Chen, G., Liang, H., Yu, M., et al. (2017). Oncogenic miR-19a and miR-19b Co-Regulate Tumor Suppressor MTUS1 to Promote Cell Proliferation and Migration in Lung Cancer. *Protein Cell* 8 (6), 455–466. doi: 10.1007/s13238-017-0393-7

Haserick, J. R., Klein, J. A., Costello, C. E., and Samuelson, J. (2017). Cryptosporidium Parvum Vaccine Candidates Are Incompletely Modified With O-Linked-N-Acetylgalactosamine or Contain N-Terminal N-Myristate and S-Palmitate. *PLoS One* 12 (8), e0182395. doi: 10.1371/journal.pone.0182395

Hoseki, J., Ushioda, R., and Nagata, K. (2010). Mechanism and Components of Endoplasmic Reticulum-Associated Degradation. *J. Biochem.* 147 (1), 19–25. doi: 10.1093/jb/mvp194

Hu, Y., Sun, Z., Deng, J., Hu, B., Yan, W., Wei, H., et al. (2017). Splicing Factor Hnrnpa2b1 Contributes to Tumorigenic Potential of Breast Cancer Cells Through STAT3 and ERK1/2 Signaling Pathway. *Tumour Biol.* 39 (3), 1010428317694318. doi: 10.1177/1010428317694318

Hu, Q., Zhang, Y., Zhang, X., and Fu, K. (2016). Gastric Mucosa-Associated Lymphoid Tissue Lymphoma and Helicobacter Pylori Infection: A Review of Current Diagnosis and Management. *Biomark Res.* 4, 15. doi: 10.1186/s40364-016-0068-1

Kim, Y. S., Hwang, S. Y., Kang, H. Y., Sohn, H., Oh, S., Kim, J. Y., et al. (2008). Functional Proteomics Study Reveals That N-Acetylglucosaminyltransferase V Reinforces the Invasive/Metastatic Potential of Colon Cancer Through Aberrant Glycosylation on Tissue Inhibitor of Metalloproteinase-1. *Mol. Cell Proteomics* 7 (1), 1–14. doi: 10.1074/mcp.M700084-MCP200

Kuo, T. F., Chen, T. Y., Jiang, S. T., Chen, K. W., Chiang, Y. M., Hsu, Y. J., et al. (2017). Protein Disulfide Isomerase A4 Acts as a Novel Regulator of Cancer Growth Through the Procaspase Pathway. *Oncogene* 36 (39), 5484–5496. doi: 10.1038/onc.2017.156

Lee, J. H., Cho, C. H., Kim, S. H., Kang, J. G., Yoo, J. S., Chang, C. L., et al. (2015). Semi-Quantitative Measurement of a Specific Glycoform Using a DNA-Tagged Antibody and Lectin Affinity Chromatography for Glyco-Biomarker Development. *Mol. Cell Proteomics* 14 (3), 782–795. doi: 10.1074/mcp.O114.043117

Li, X., Liu, H., Yu, T., Dong, Z., Tang, L., and Sun, X. (2014). Loss of MTUS1 in Gastric Cancer Promotes Tumor Growth and Metastasis. *Neoplasma* 61 (2), 128–135.

Lin, W. C., Tsai, H. F., Kuo, S. H., Wu, M. S., Lin, C. W., Hsu, P. I., et al. (2010). Translocation of Helicobacter Pylori CagA Into Human B Lymphocytes, the Origin of Mucosa-Associated Lymphoid Tissue Lymphoma. *Cancer Res.* 70 (14), 5740–5748. doi: 10.1158/0008-5472.CAN-09-4690

Liu, Y. C., Yen, H. Y., Chen, C. Y., Chen, C. H., Cheng, P. F., Juan, Y. H., et al. (2011). Sialylation and Fucosylation of Epidermal Growth Factor Receptor Suppress its Dimerization and Activation in Lung Cancer Cells. *Proc. Natl. Acad. Sci. U. S. A.* 108 (28), 11332–11337. doi: 10.1073/pnas.1107385108

Ma, M. Z., Yuan, S. Q., Chen, Y. M., and Zhou, Z. W. (2018). Preoperative Apolipoprotein B/apolipoprotein A1 Ratio: A Novel Prognostic Factor for Gastric Cancer. *Onco. Targets Ther.* 11, 2169–2176. doi: 10.2147/OTT.S156690

Moleiro, J., Ferreira, S., Lage, P., and Dias Pereira, A. (2016). Gastric Malt Lymphoma: Analysis of a Series of Consecutive Patients Over 20 Years. *United Eur. Gastroenterol. J.* 4 (3), 395–402. doi: 10.1177/2050640615612934

Munari, F., Lonardi, S., Cassatella, M. A., Doglioni, C., Cangi, M. G., Amedei, A., et al. (2011). Tumor-Associated Macrophages as Major Source of APRIL in Gastric MALT Lymphoma. *Blood* 117 (24), 6612–6616. doi: 10.1182/blood-2010-06-293266

Napoletano, C., Rughetti, A., Agervig Tarp, M. P., Coleman, J., Bennett, E. P., Picco, G., et al. (2007). Tumor-Associated Tn-MUC1 Glycoform Is Internalized Through the Macrophage Galactose-Type C-Type Lectin and Delivered to the HLA Class I and II Compartments in Dendritic Cells. *Cancer Res.* 67 (17), 8358–8367. doi: 10.1158/0008-5472.CAN-07-1035

Nobre-Leitao, C., Lage, P., Cravo, M., Cabecadas, J., Chaves, P., Alberto-Santos, A., et al. (1998). Treatment of Gastric MALT Lymphoma by Helicobacter Pylori Eradication: A Study Controlled by Endoscopic Ultrasonography. *Am. J. Gastroenterol.* 93 (5), 732–736. doi: 10.1111/j.1572-0241.1998.215_a.x

Park, A. M., Hagiwara, S., Hsu, D. K., Liu, F. T., and Yoshie, O. (2016). Galectin-3 Plays an Important Role in Innate Immunity to Gastric Infection by Helicobacter Pylori. *Infect. Immun.* 84 (4), 1184–1193. doi: 10.1128/IAI.01299-15

Pinho, S. S., and Reis, C. A. (2015). Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nat. Rev. Cancer* 15 (9), 540–555. doi: 10.1038/nrc3982

Ponomarenko, J., Bui, H. H., Li, W., Fusseder, N., Bourne, P. E., Sette, A., et al. (2008). ElliPro: A New Structure-Based Tool for the Prediction of Antibody Epitopes. *BMC Bioinf.* 9, 514. doi: 10.1186/1471-2105-9-514

Raderer, M., Kiesewetter, B., and Ferreri, A. J. (2016). Clinicopathologic Characteristics and Treatment of Marginal Zone Lymphoma of Mucosa-Associated Lymphoid Tissue (MALT Lymphoma). *CA Cancer J. Clin.* 66 (2), 153–171. doi: 10.3322/caac.21330

Ryan, D., Carberry, S., Murphy, A. C., Lindner, A. U., Fay, J., Hector, S., et al. (2016). Calnexin, an ER Stress-Induced Protein, Is a Prognostic Marker and Potential Therapeutic Target in Colorectal Cancer. *J. Transl. Med.* 14 (1), 196. doi: 10.1186/s12967-016-0948-z

Saito, Y., Suzuki, H., Tsugawa, H., Imaeda, H., Matsuzaki, J., Hirata, K., et al. (2012). Overexpression of miR-142-5p and miR-155 in Gastric Mucosa-Associated Lymphoid Tissue (MALT) Lymphoma Resistant to Helicobacter Pylori Eradication. *PLoS One* 7 (11), e47396. doi: 10.1371/journal.pone.0047396

Saribasak, H., Salih, B. A., Yamaoka, Y., and Sander, E. (2004). Analysis of Helicobacter Pylori Genotypes and Correlation With Clinical Outcome in Turkey. *J. Clin. Microbiol.* 42 (4), 1648–1651. doi: 10.1128/JCM.42.4.1648-1651.2004

Sirwi, A., and Hussain, M. M. (2018). Lipid Transfer Proteins in the Assembly of apoB-Containing Lipoproteins. *J. Lipid Res.* 59 (7), 1094–1102. doi: 10.1194/jlr.R083451

Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T., et al. (2013). Precision Mapping of the Human O-GalNAc Glycoproteome Through SimpleCell Technology. *EMBO J.* 32 (10), 1478–1488. doi: 10.1038/emboj.2013.79

Sudhan, D. R., and Siemann, D. W. (2015). Cathepsin L Targeting in Cancer Treatment. *Pharmacol. Ther.* 155, 105–116. doi: 10.1016/j.pharmthera.2015.08.007

Tenzer, S., Peters, B., Bulik, S., Schoor, O., Lemmel, C., Schatz, M. M., et al. (2005). Modeling the MHC Class I Pathway by Combining Predictions of Proteasomal Cleavage, TAP Transport and MHC Class I Binding. *Cell Mol. Life Sci.* 62 (9), 1025–1037. doi: 10.1007/s00018-005-4528-2

Wang, H. C., Cheng, F. C., Wu, M. S., Shu, H. Y., Sun, H. S., Wang, Y. C., et al. (2015). Genome Sequences of Three Helicobacter Pylori Strains From Patients With Gastric Mucosa-Associated Lymphoid Tissue Lymphoma. *Genome Announc.* 3 (2), e00229–15. doi: 10.1128/genomeA.00229-15

Wang, L., Zhu, Q., Zhang, H., and Xiao, D (2020). Comparison of Transcriptomes of Gastric Mucosa-Associated Lymphoid Tissue Lymphoma Helicobacter Pylori Isolates and Gastritis Isolates[J]. *Dis. Surveillance* 35 (4), 288–293. doi: 10.3784/j.issn.1003–9961.2020.04.005

Wang, P., Sidney, J., Dow, C., Mothe, B., Sette, A., and Peters, B. (2008). A Systematic Assessment of MHC Class II Peptide Binding Predictions and Evaluation of a Consensus Approach. *PLoS Comput. Biol.* 4 (4), e1000048. doi: 10.1371/journal.pcbi.1000048

Xuan, Y., Wang, J., Ban, L., Lu, J. J., Yi, C., Li, Z., et al. (2016). Hnrnpa2/B1 Activates Cyclooxygenase-2 and Promotes Tumor Growth in Human Lung Cancers. *Mol. Oncol.* 10 (4), 610–624. doi: 10.1016/j.molonc.2015.11.010

Zhang, E., He, X., Zhang, C., Su, J., Lu, X., Si, X., et al. (2018). A Novel Long Noncoding RNA HOXC-AS3 Mediates Tumorigenesis of Gastric Cancer by Binding to YBX1. *Genome Biol.* 19 (1), 154. doi: 10.1186/s13059-018-1523-0

Zou, Q., Zhang, H., Meng, F., He, L., Zhang, J., and Xiao, D (2020). Proteomic and Transcriptomic Studies of BGC823 Cells Stimulated With Helicobacter Pylori Isolates From Gastric MALT Lymphoma. *PLoS One* 15 (9), e0238379. doi: 10.1371/journal.pone.0238379

Zullo, A., Hassan, C., Ridola, L., Repici, A., Manta, R., and Andriani, A. (2014). Gastric MALT Lymphoma: Old and New Insights. *Ann. Gastroenterol.* 27 (1), 27–33.