



Full length article

Mass spectrometric determination of disulfide bonds and free cysteine in grass carp IgM isoforms

Yiling Su¹, Bing Wang¹, Ying Zhang¹, Zilun Ruan¹, Hao Bai, Jian Wan, Chen Xu, Guoqi Li, Shengqiang Wang, Hui Ai, Li Xiong, Hui Geng*

Hubei Key Laboratory of Genetic Regulation and Integrative Biology, School of Life Sciences, Central China Normal University, Wuhan, 430079, China

ARTICLE INFO

Keywords:

IgM
Teleost
Grass carp
Liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS)
Disulfide bonds

ABSTRACT

Disulfide bonds are fundamental in establishing Ig structure and maintaining Ig biological function. Here, we analysed disulfide bonds and free cysteine in three grass carp IgM isoforms (monomeric, dimeric/trimeric, and tetrameric IgM) by liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS). The results revealed that Cys⁵⁷⁴ residue status at the C-terminal tail differed substantially in monomeric IgM in comparison with polymeric IgM, Cys⁵⁷⁴ was found as free thiol in monomeric IgM, while it formed disulfide linkages in dimeric/trimeric and tetrameric IgM. Five intra-chain disulfide bonds in the CH1~CH4 and CL1 domains, as well as one H-H and one H-L inter-chain disulfide linkages, were also observed and shown identical connectivity in monomeric, dimeric/trimeric, and tetrameric IgM. These findings represent the first experimental assignments of disulfide linkages of grass carp IgM and reveal that grass carp IgM isoform formation is due to alternative disulfide bonds connecting the Cys⁵⁷⁴ residue at the C-terminal tail.

1. Introduction

Disulfide bonds are one of the most essential protein post-translational modifications in stabilizing structure and regulating the biological function of proteins [1–4]. Numerous studies have shown that disulfide bonds change as disulfide reduction, disulfide scrambling and free cysteine residue oxidation can lead to structural perturbations in proteins in the form of lower thermo stability, spontaneous unfolding, and biological activity change [5–7]. The determination of disulfide connectivity is an important aspect in understanding protein structures and defining the functional domains of proteins.

Abs or immunoglobulins (Ig) are high disulfide bond glycoproteins with the basic four-chain structure. Two identical heavy chains (IgH) and two identical light chains (IgL) are joined by inter-chain disulfide bonds and intra-chain disulfide loops in all five mammalian Ig classes (IgM, IgG, IgA, IgE and IgD) [1,2,8,9]. The κ and λ light chains are identified following the same disulfide connectivity, while the five Ig heavy chains show substantial differences in the number of cysteine residues and disulfide bond arrangement, specifically in the constant

region. For example, four human IgG subclasses (IgG1, IgG2, IgG3, and IgG4) were found to exhibit characteristically different inter-chain disulfide bonding in the constant region, with 2 inter-chain disulfide bonds in IgG1 and IgG4, 4 inter-chain disulfide bonds in IgG2, and 5 inter-chain disulfide bonds in IgG3 [10–12]. Besides the classical Ig disulfide connection, alternative (or non-classical) disulfide bonding has been found in the IgG2 and IgG4 subclasses, leading to the disulfide bond-mediated structural isoform and Ab effector function changes [5,9,13]. The constant region of IgM is composed of four classical Ig domains (C μ 1~C μ 4) and an additional C-terminal tail piece. It revealed that pentameric human IgM is assembled by J chain, inter-monomer disulfide linkages with the cysteine residues Cys414 and Cys575 located at the μ 3 domain and C-terminal tail piece, respectively [8,14–16].

Of the tree teleost Ig classes IgM, IgD and IgT/Z (for Teleost/Zebrafish), IgM is the most abundant Ab and the first one discovered decades ago, which serves in a key role in humoral immune response. Regarding teleost IgM disulfide connectivity patterns, our understanding is largely based on homology sequence comparison with mammalian IgM [17–20]. Compared to mammalian IgM, teleost IgM

Abbreviations: Ig, immunoglobulin; CH, heavy chain constant domain; CL, light chain constant domain; VH, heavy chain variable domain; VL, light chain variable domain; Fc, fragment of constant region; MS, mass spectrometry; LC-ESI-MS/MS, liquid chromatography-electrospray ionization tandem mass spectrometry; XIC, extracted ion chromatography

* Corresponding author. Hubei Key Laboratory of Genetic Regulation and Integrative Biology, School of Life Sciences, Central China Normal University, No152, Luoyu Road, Hubei province, Wuhan, 430079, China.

E-mail address: genghui@mail.ccnucnu.edu.cn (H. Geng).

¹ These authors contributed equally to this paper.

<https://doi.org/10.1016/j.fsi.2019.10.051>

Received 1 May 2019; Received in revised form 21 September 2019; Accepted 24 October 2019

Available online 25 October 2019

1050-4648/ © 2019 Elsevier Ltd. All rights reserved.

lacks the J chain, which regulates the polymeric structure of IgM. A universal feature of teleost IgM is the production of variable disulfide structural heterogeneity with different polymerization states. It has been reported that salmon fish IgM produces isoforms under denaturation non-reductive conditions, which indicated salmon IgM isoforms related to disulfide connectivity [21]. Much work also revealed similar findings, as channel catfish IgM, catfish IgM, sheepshead IgM, and carp IgM were found to exhibit structural diversity involved in disulfide bond-mediated isoforms [22–26]. However, the disulfide bond structures of teleost IgM and disulfide bond-mediated IgM isoforms are limited to homology, and no direct evidence has been presented for the complete structures.

Traditional methods for the analysis of disulfide bonds include Edman degradation, diagonal paper electrophoresis, X-ray crystallography and NMR spectroscopy [27–30]. However, none of these methods are ideal, due to requiring significant amounts of highly pure samples and due to the low throughput of solving structures. In recent years, mass spectrometry-based approaches have been employed in the analysis of disulfide bonds in a promising high throughput manner due to its high sensitivity [31,32].

In this study, the disulfide connectivity and free cysteine in grass carp covalent IgM isoforms are investigated using LC-MS/MS. In addition to complete mapping of the inter-chain and intra-chain disulfide bonds, the identification of Cys⁵⁷⁴ at the C-terminal tail piece, exhibited as a free thiol in monomeric IgM while forming disulfide linkages in dimeric/trimeric and tetrameric IgM, was achieved. The data represent the first experimental assignment of disulfide linkages and also reveal alternative disulfide bond arrangements at the C-terminal tail piece dominating the formation of grass carp IgM isoforms.

2. Materials and methods

2.1. Reagents

All chemicals and solvents used in this work were of HPLC grade. N-Ethylmaleimide (NEM), dithiothreitol (DTT), iodoacetamide (IAM), ammonium sulphate, ammonium bicarbonate, ethylenediaminetetraacetic acid (EDTA), sodium dodecyl sulphate (SDS), and urea were obtained from Sigma-Aldrich (Shanghai, China). Acetonitrile and 100% (w/v) formic acid were purchased from Thermo Scientific (San Jose, CA). PNGase F was obtained from New England Biolabs (Massachusetts, USA). Trypsin and endoproteinase Asp-N were purchased from Promega (Beijing, China) and Roche (Penzberg, Germany). ZipTip C18 was purchased from Millipore (Boston, USA). All buffers and solutions were prepared with deionized water purified by a Millipore purification system (Billerica, USA).

2.2. Animals

Grass carp (*Ctenopharyngodonidella idella*) about 300–500 g were obtained from Tangsun Lake aquatic culture company (Hubei, China). They were bred and kept at $25 \pm 2^\circ\text{C}$ in 1000-L tanks with re-circulating and UV-sterilized water for acclimating to the aquarium tanks. Grass carps were anaesthetized with 0.03% tricaine methane sulphate (Sigma, Shanghai, China) in water for blood collection. The Institutional Animal Care and Use Committee of Central China Normal University approved all of the experiments.

2.3. Isolation of grass carp IgM

Serum was used as a source of grass carp IgM preparation. Blood was collected from the caudal veins from 10 to 15 naïve grass carp, allowed to clot for 2 h at room temperature, collected by centrifugation at 4000 rpm for 10 min, and pooled and stored at -80°C until use. To prevent disulfide bond shuffling during sample purification, all solution pH conditions were strictly maintained under 6.5 with the addition of

4 mM N-ethylmaleimide (NEM) to the reagents as described previously [33]. Grass carp IgM samples were purified by ammonium sulphate precipitation with sequential ion-exchange chromatography following the protocol described previously [34,35]. Briefly, approximately 30 ml of grass carp serum was precipitated by saturated ammonium sulphate and then 50% (w/v) ammonium sulphate. The collected proteins were re-suspended in 10 mM sodium phosphate buffer (pH 6.5) containing 5 mM EDTA and passed over a Source Q anion column (5 ml bed volume of Amersham Biosciences Source Q anion resin). The IgM retained in the column were eluted with a 0–1 M linear NaCl gradient at a flow rate of 1 ml/min. To further purify the IgM, the fraction containing grass carp IgM was collected and dialyzed against 10 mM sodium phosphate (pH 5.9), loaded onto a Source S cation exchange column (2 ml bed volume), and the retained IgM in the column were eluted with NaCl gradient (0–0.5 M). Fractions containing grass carp IgM were pooled, concentrated using Amicon ultra-centrifugal filters (50 kDa cut-off), checked by SDS-PAGE and stored at -80°C until use.

2.4. Blocking of free cysteine residues in grass carp IgM

Prior to SDS-PAGE to separate grass carp IgM, free cysteine residues in grass carp IgM were treated with 4 mM N-ethylmaleimide (NEM) in 0.1 M Tris-HCl, pH 6.5, for 1 h to ensure that free cysteine residues were completely blocked. NEM was chosen as the blocking agent because it can readily mask cysteine residues under acidic conditions and can also be identified during ESI-MS/MS by the mass increase of 125.13 Da, thus allowing confident determination of free cysteine residues [32].

2.5. SDS-polyacrylamide gel electrophoresis to separate IgM isoforms

To verify different IgM disulfide polymerization, the collected IgM samples were mixed with loading buffer without any reducing agent (β -mercaptoethanol or DTT in the loading buffer) and characterized by 6% SDS-polyacrylamide gel electrophoresis at 30 mA for 2.5 h. Protein bands were stained with Coomassie Brilliant Blue. The bands of interest were excised from the gel and then subjected to deglycosylation and proteolytic digestion as described below.

2.6. Deglycosylation and proteolytic digestion of IgM

The protein bands were excised and destained by alternating use of acetonitrile and 100 mM ammonium bicarbonate (pH 6.5) until no visible colour was present. The destained gel pieces were completely dehydrated with acetonitrile and then dried in a SpeedVac centrifuge. One unit of PNGase F (100 units/ml) was added to cover the dried gel pieces, after a 30-min incubation at 4°C , reaction buffer (pH 6.5) was added to cover the gel pieces overnight (~ 12 h) at 37°C . To ensure complete deglycosylation, a second aliquot of PNGase F was added to the samples for additional overnight digestion. The deglycosylated IgM samples were digested overnight with trypsin (1:20 w/w, in 0.1 M Tris, pH 6.5) at 37°C , and trypsin was added at an enzyme-to-protein ratio of 1:50 (w/w) and incubated for 4 h at 37°C . For endoproteinase Asp-N digestion, a portion of the sample was subjected to Asp-N digestion (1:50 w/w, in 0.1 M Tris, pH 6.5) for 6 h at 37°C after trypsin treatment. When reducing the disulfide bond-peptides, the protein gel piece was incubated with 64 mM DTT at 56°C for 30 min, and alkylation was achieved with 130 mM IAM at R.T. for 30 min in the dark prior to deglycosylation and proteolytic digestion. Proteolytic digestion was stopped by adding 10 μl of 10% formic acid. Peptides were extracted from the gel with 50% acetonitrile in 5% formic acid, lyophilized by SpeedVac and stored at -20°C until desalting treatment.

2.7. Peptide desalting

The lyophilized peptide samples were resuspended with 1% formic acid in water and desalted using ZipTip C18 (Millipore). Briefly, ZipTip

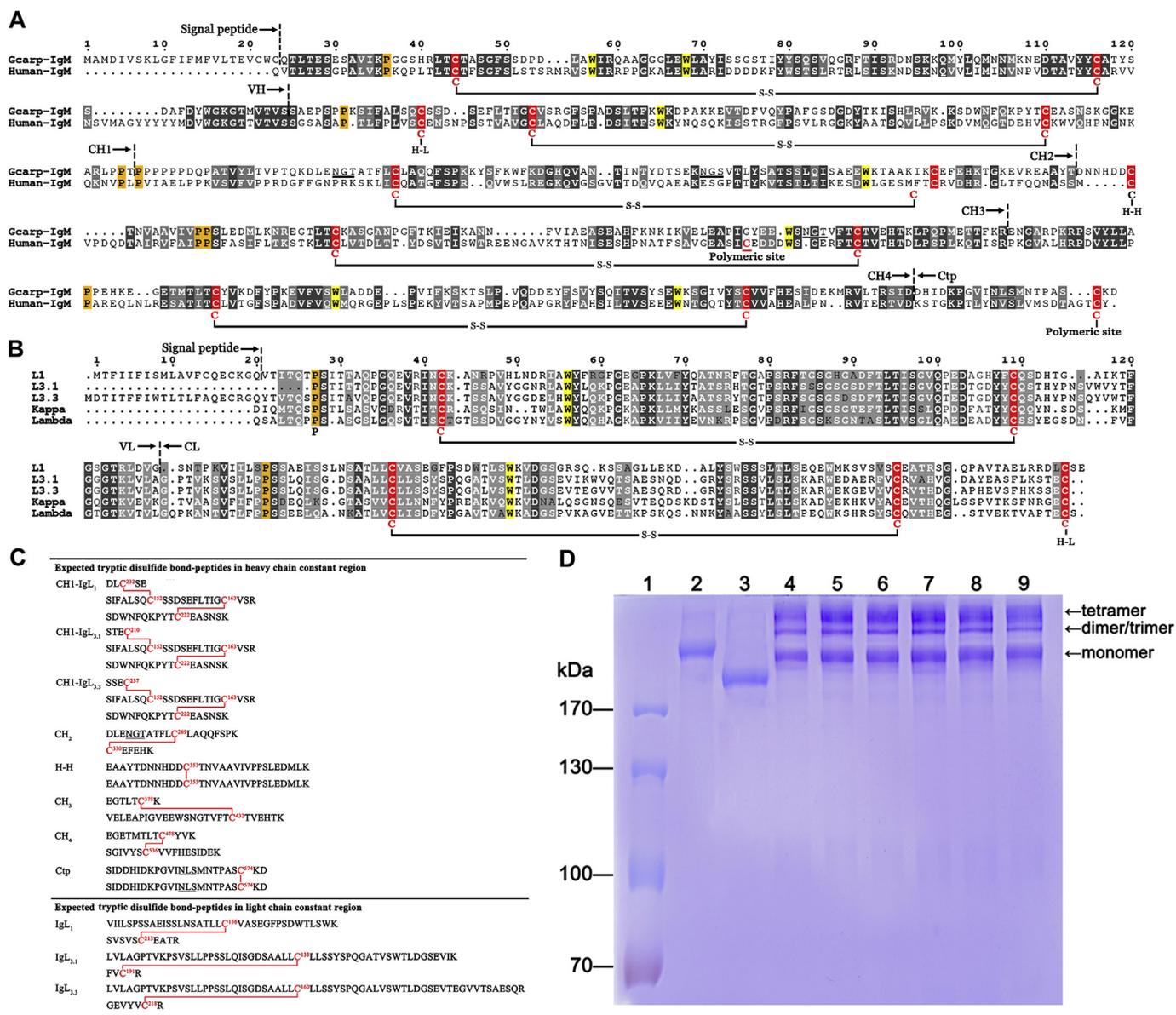


Fig. 1. Predicted disulfide bonds in grass carp IgM based on homology model. Sequence alignment of grass carp Ig heavy chain (A) and Ig light chain (B). Dark and light shadows indicate identical and similar residues, red shadow indicates the conserved cysteine involved in the disulfide bond, while yellow and brown shadows indicate crucial tryptophan and proline residues forming the fold of each Ig domain. The N-linked glycosylation sites in grass carp IgM heavy chain are underlined. (C) Expected tryptic disulfide bond-peptides of grass carp IgM heavy and light chain constant regions. (D) Analysis of purified grass carp IgM by SDS-PAGE under denaturation non-reducing conditions. Lane 1, Marker; lane 2, thyroglobulin (303 kD, Sigma); lane 3, mouse IgG1 mAb (anti-COMP, 15A11, 150 kD); lane 4 to lane 9, grass carp IgM. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

C18 was activated with 75 µl of acetonitrile containing 1% formic acid, and then balanced three times with deionized water containing 1% formic acid. The peptide samples were loaded onto the C18 tip and desalted with 1% formic acid in water. Peptides were desalted and eluted stepwise with 1% formic acid in deionized water and 1% formic acid in acetonitrile. The peptide samples were then lyophilized by SpeedVac and stored at -20 °C until analysis.

2.8. Chromatography and mass spectrometry analysis

Peptide analyses were performed by using a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific, USA) coupled to an ultra-performance liquid chromatography (UPLC) system (EASY-nLC 1200 system, Thermo Fisher Scientific, USA) as described in previous research [34]. Peptides were separated using a C18 reversed-phase column (75 µm × 150 mm, particle size 3 µm, 100 Å; Thermo Fisher

Scientific, USA) with solvent A (95% water and 4.9% acetonitrile with 0.1% formic acid) and solvent B (90% acetonitrile and 9.9% water with 0.1% formic acid) at a flow rate of 500 nl/min. Lyophilized peptides were dissolved in 10 µl of solvent A. The gradient profile settings were as follows: 3–5% B for 6 min, 5–25% B for 54 min, 25–40% B for 25 min and holding for 5 min, 40–80% B for 5 min, and 80–100% B for 5 min, for a total of 100 min.

Mass spectrometric analysis was performed in data-dependent acquisition (or a pre-selected ion mode) and positive ion mode with dynamic exclusion duration of 40 s. Data-dependent acquisition was set up to acquire the most intense ion (top 20) events for every full high-resolution MS scan in the mass range from 350 to 2000 m/z. The width for precursor ion isolation was set to 2.0, with an activation time of 50 ms. The spray voltage was 2.2 kV, and the capillary temperature was 275 °C. In MS1, the precursor ion was set at an automatic gain control (AGC) target value of 3 × 10⁶ (3e⁶), the resolution was 70,000, and the

maximum injection time was 50 ms. In MS2, the AGC target was set at 5×10^4 ($5e^4$), while the resolution was 17,500, the maximum injection time was 50 ms, and normalized collision energy was 27% for HCD. A blank wash and BSA control were performed between every two samples to avoid sample carryover.

2.9. Peptide identification and disulfide bond assignment

Disulfide bond assignment was performed using plink SS (version 2.3.5) (<http://pfind.ict.ac.cn/software/plink/index.html#Downloads>) as described previously [36]. Raw MS/MS spectra data were searched with grass carp serum IgM heavy chain (accessions ABD76396.1) and 3 known Ig light chain type obtained from GenBank (accessions AEH76777.1, AEH76781.1, and AEH76779.1). Search criteria included trypsin (and Asp-N) and allowed a maximum of two missed cleavages; cysteine alkylation by NEM (+125.13 Da), cysteine alkylation by IAM (carbamidomethylation, +57.02 Da), oxidation of methionine (+15.9949 Da) and conversion of Asn to Asp (+0.98 Da) were considered as variable modifications. Mass tolerance of 10 ppm was used in all searches. The false-discovery rates of peptide and disulfide bonds were set at 0.01. The minimum and maximum peptide lengths were four and sixty amino acids, respectively. The reduced peptide spectra were also processed using Proteome Discoverer V2.2 (Thermo Fisher Scientific). Xcalibur 4.0 (Thermo Fisher Scientific) was used for generating extracted ion chromatography (XIC). Peptide identification, disulfide bond assignment and all spectra were confirmed manually.

3. Results

3.1. Characterization of grass carp IgM isoforms

There are 14 cysteine residues in each grass carp IgM heavy constant region, 3 cysteine residues in three known grass carp Ig light chain constant region. The 14 cysteines of heavy chain potentially involved in the disulfide linkage in CH1~CH4, heavy-heavy and light-heavy disulfide linkage (Fig. 1A). Notably, Cys⁴¹⁴ in the human IgM Fc region involved in inter-monomeric disulfide bonding has no equivalent in the grass carp IgM molecule, implying that Cys⁵⁷⁴ in the C-terminal tail piece is potentially associated with inter-monomeric disulfide linkage. Cysteine residues involved in CL and light-heavy disulfide linkages in 3 known grass carp Ig light chain type are shown in Fig. 1B. The predicted tryptic disulfide bond-peptides are shown in Fig. 1C.

Purified grass carp IgM migrated three apparent protein bands under denaturation non-reduced SDS-PAGE conditions, according to the marker and reference proteins (thyroglobulin, 303 kDa; mouse IgG1 mAb, 150 kDa), these bands were recognized as tetramer, dimer/trimer (mostly trimer), and monomer (Fig. 1D). Grass carp IgM samples used in this study were collected from serum, and the amino acid sequences of the variable regions are polyclonal, thus, only disulfide bonds and free cysteine in the constant region are mapped below.

3.2. Disulfide bond and free cysteine residue analysis approach

To map disulfide bonding and free cysteine residues in grass carp IgM isoforms, IgM samples were first blocked with NEM to mask free cysteine and eliminate thiol-mediated disulfide scrambling prior to SDS-PAGE and in-gel protease digestion. Grass carp IgM heavy chain has 4 conserved N-glycosylation sites, namely, Asn²⁶², Asn³⁰³, Asn⁴²⁶, and Asn⁵⁶⁵, and examination of disulfide bond-peptides containing Cys²⁶⁹ residue in the CH1 domain and Cys⁵⁷⁴ residue at the C-terminal tail piece may be affected by their surrounding Asn²⁶² and Asn⁵⁶⁵ glycosylations, thus, deglycosylation by PNGase F treatment prior to trypsin digestion was performed to reduce analysis complexity (Fig. 2). To limit disulfide shuffling during sample preparation and digestion, the experiment was strictly controlled under pH 6.5 as described in the Materials and methods section. The resultant tryptic peptides

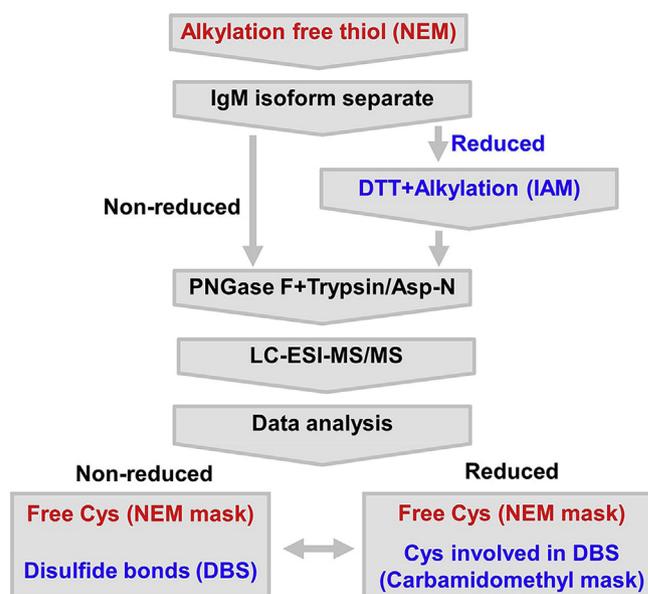


Fig. 2. Schematic representation of the disulfide bonds and free cysteine mapping approach.

containing disulfide linkage and NEM-modified free cysteine residue were subjected to reversed phase HPLC followed by ESI-MS and MS/MS analysis. In a parallel experiment, to support that cysteine residues are indeed involved in disulfide bond formation, samples were reduced with DTT and cysteine was labelled with alkylation IAM and thus could be distinguishable from free cysteine residues (Fig. 2). MS analysis indicated that complete deglycosylation was achieved, and trypsin digestion was sufficient with a low number of missed cleavages at pH 6.5 (see Figs. 3–6 and Table 1 listed below).

3.3. Identification of free cysteine in monomeric IgM

To demonstrate that disulfide bonding patterns may be different among grass carp IgM isoforms, we first investigated the potential free cysteine in monomeric IgM. Tryptic digests from deglycosylated monomeric IgM with NEM pretreatment were separated and analysed by on-line LC-ESI-MS/MS. For the determination of free cysteine residue, the mass increase of 125.13 Da associated with NEM alkylation can serve as a useful marker. A major peak with a retention time of between 46 and 48 min was observed, with monoisotopic m/z values of 678.82 in the 4^+ charge state and 904.75 in the 3^+ charge state (shown in Fig. 3A top panel inset), corresponding to NEM capped and deglycosylated Cys⁵⁷⁴ peptides. The HCD fragmentation ions y_3-y_4 and y_6-y_{10} compared to y_2 exhibit a mass increase of 125.13 Da, unambiguously locating the NEM capping to the Cys⁵⁷⁴ residue. In addition, b_{13} , b_{14} , and b_{15} fragment ions displayed an additional 0.98 Da mass increase, implying that deglycosylation converts the occupied asparagine residue to aspartic acid in the Asn⁵⁶⁵ glycosite (Fig. 3B top panel). In addition, another peak with a retention time of approx. 45.3 min was observed, with monoisotopic masses of 678.57 in the 4^+ charge state and 904.42 in the 3^+ charge state, which closely matched the expected masses of NEM-capped non-glycosylation tryptic Cys⁵⁷⁴ peptides. Its identity was further confirmed by MS-MS data, where numerous b and y ions were consistent with the NEM-capped non-glycosylation tryptic Cys⁵⁷⁴ peptides (Fig. 3B bottom panel). These results demonstrated that Cys⁵⁷⁴ residues in the C-terminal tail piece were present in free sulfhydryl form and were not involved in the disulfide linkage in monomeric IgM.

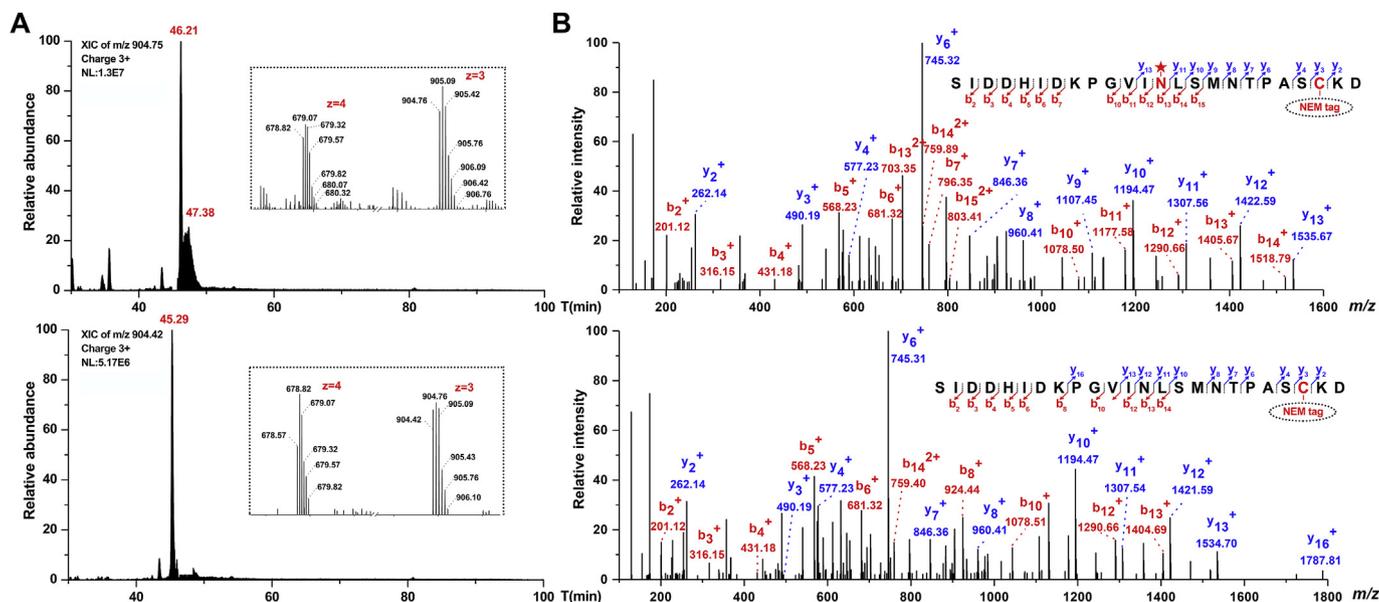


Fig. 3. NEM masked Cys⁵⁷⁴ peptides in monomeric grass carp IgM. (A) Extracted ion chromatograms, corresponding monoisotopic precursor mass spectra (inset) of NEM masked Cys⁵⁷⁴ peptide precursor with deglycosylation (top panel) and with non-glycosylation (bottom panel). (B) HCD-MS² spectra correspond to NEM masked Cys⁵⁷⁴ peptide with deglycosylation (top panel, m/z 904.76, charge 3⁺) and Cys⁵⁷⁴ peptide with non-glycosylation (bottom panel, m/z 904.42, charge 3⁺). The star denotes the deglycosylation site. XIC, extracted ion chromatography. NL, normalized level.

3.4. Determination of the status of Cys⁵⁷⁴ on oligomeric IgM

To compare the structural differences between grass carp IgM isoforms, the nonreduced tryptic peptides from dimeric/trimeric and tetrameric IgM were examined by similar on-line LC-ESI-MS/MS methods, and the analysis was still focused on the Cys⁵⁷⁴ residue in the C-terminal tail piece. Disulfide bond identification was based on extracted ion chromatograms (XIC), matching the observed precursor ion mass to the theoretical mass of two Cys-containing peptides (the sum masses of the dipeptide minus 2 Da). Low energy HCD preferentially cleaves the carbon backbone and leaves disulfide bonds intact, producing b, y ions with and without disulfide bonds. Assignments of disulfide bond-peptides could further be verified by corresponding MS2 fragmentation spectra. A peak with a retention time of approx. 55–57 min was observed, and three monoisotopic masses of 1289.85 in the 4⁺ charge state, 1031.89 in the 5⁺ charge state, and 860.07 in the 6⁺ charge state were detected in dimeric/trimeric and tetrameric IgM products that were not found in monomeric IgM. These observed m/z values are in good agreement with disulfide bonding in the tryptic Cys⁵⁷⁴ peptide (Fig. 4A). Further confirmation of the presence and location of the Cys⁵⁷⁴ disulfide bond-peptides in polymeric IgM was obtained from the MS/MS spectrum, as shown in Fig. 4B. This spectrum contained a series of b and y ions resulting from backbone cleavage of peptide α with methionine oxidation and peptide β without methionine oxidation, which indicated two Cys⁵⁷⁴ disulfide bond-peptides.

To provide additional support for the finding that Cys⁵⁷⁴ residues of polymeric IgM are indeed involved in disulfide linkages, the tryptic digestion samples were subjected to reduction with dithiothreitol (DTT) and IAM alkylation. The resulting reduced Cys⁵⁷⁴ peptide fragment ion spectrum exhibited an addition of IAM alkylation tag (carbamidomethylation, + 57.02 Da) to the Cys⁵⁷⁴ residue (shown in Supporting Figures S1). Remembering that free sulfhydryl groups have been blocked by NEM, only cysteine residues involved in disulfide bonding should be masked by carbamidomethylation. In all, these data indicated that Cys⁵⁷⁴ residues at the C-terminal tail piece are involved in disulfide linkages in polymeric IgM. Thus, the oligomeric form of grass carp IgM possesses a different disulfide bonding network at the C-terminal tail piece compared to monomeric IgM.

3.5. Mapping disulfide bonds in CH2~CH4 and between heavy-heavy chains in grass carp IgM

The disulfide bonds in CH2, CH3, CH4, and heavy-heavy chains in three grass carp IgM isoforms were identified by LC-ESI-MS/MS analysis in a similar manner. An example to determine disulfide bonding in the CH2 domain is shown in Fig. 5, these disulfide bond-peptides, which consist of a fully deglycosylated peptide DLENGTAT-FLC²⁶⁹LAQQFSPK and a non-glycosylated peptide C³³⁰EFEHK connected by a disulfide bond, have a theoretical m/z of 958.44 in the 3⁺ charge state, which was detected within 6 ppm in the mass spectrum (Fig. 5A). The identity of these ions was determined by MS2 data (Fig. 5B), where numerous b and y ions from the peptide backbone fragmentation of peptides α and β , as well as $\alpha\gamma$ - y_{16} , βb_2 , and βb_4 - b_5 ions, were consistent with the disulfide bond-peptides. Hence, the CH2 domain disulfide bond was assigned. A summary of the experimentally determined disulfide bonding in the Fc region and C-terminal tail piece of grass carp IgM isoforms is shown in Table 1. These results indicated that monomeric and oligomeric IgM isoforms contain identical disulfide bond patterns in CH2~CH4 and between heavy-heavy chains. MS/MS spectra showing the assignments of the CH3, CH4, and heavy-heavy chain disulfide bond-peptides, and their corresponding reduced peptides with carbamidomethylation masking, are shown in Supporting Figures S2-S4.

3.6. Mapping disulfide bonding in CH1, CL1 and heavy-light chains in grass carp IgM

Among 3 known Ig light chains type, only disulfide bonding of the IgL3.3 with CH1 was able to be detected (shown in Fig. 6A), while IgL1 and IgL3.1 connecting with CH1 were not covered in this study. This lack of coverage may be due to low-abundance disulfide bond-peptides that were not readily detected by MS analysis. Usage percentages of light chain was estimated by comparing the peptide spectral matches (PSMs) of the maximum observed frequency unique peptide for each isotype. No IgL1-specific peptides were detected on monomeric and polymeric IgM in non-reduced and reduced samples in this study, suggesting that IgL1 is not used by grass carp IgM (data not shown). The percentage of IgL3.3 isotype was estimated to be approximately

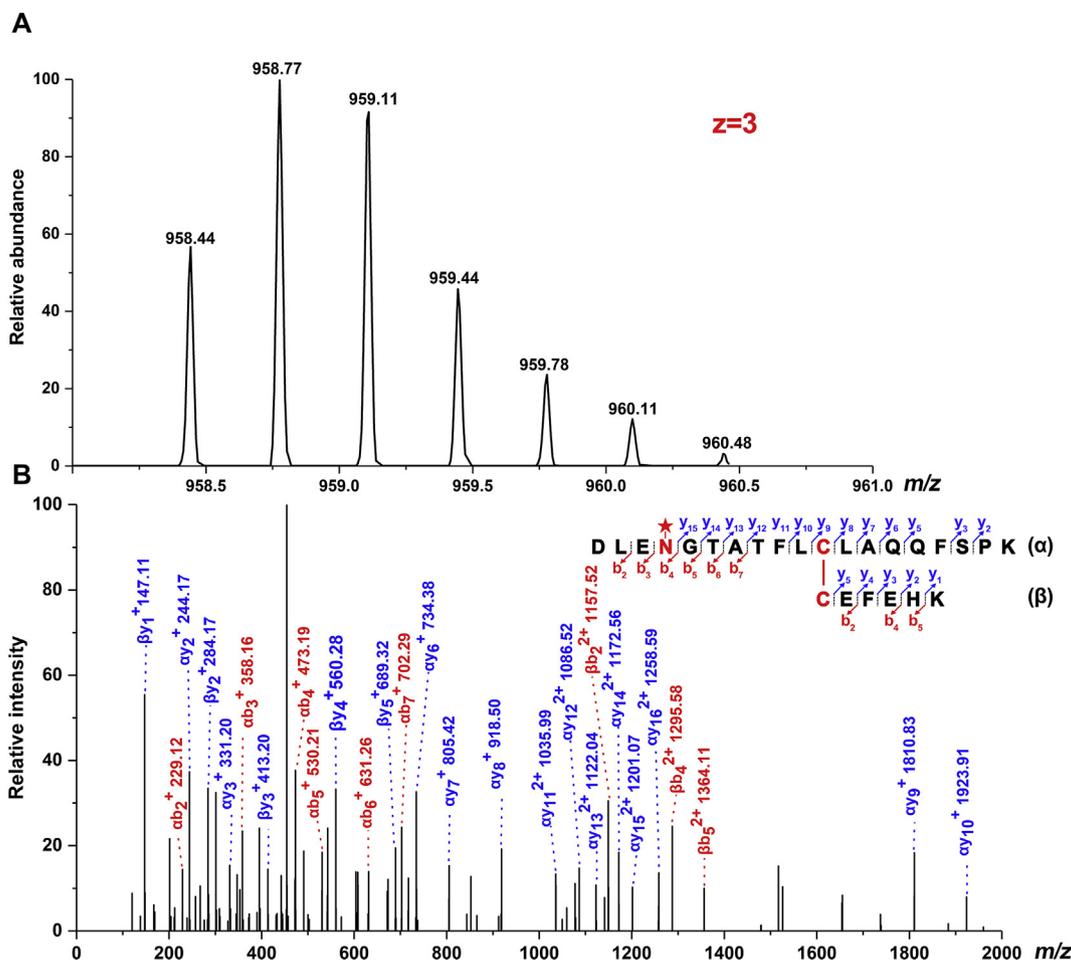


Fig. 5. Typical monoisotopic mass spectra (A) and HCD spectrum (B) that support the assignment of the disulfide bond in the CH2 domain. Sequence coverage maps of each peptide and annotation of the fragment ions are shown. The star denotes the deglycosylation site.

generation of disulfide bond-mediated isoforms. Disulfide isoforms of teleost IgM therefore have the potential to provide a significant level of *in vivo* control of antibody structure, thereby adding a new dimension to the diversity of the antibody family. Besides specific disulfide bonding at the C-terminal tail piece dominating grass carp IgM isoform formation, carbohydrates at the C-terminal tail piece may be another important motif for regulating IgM polymer assembly. In their absence, human IgM is inclined to form larger hexamers instead of pentamers, supporting the suspicion of oligosaccharide attached at the C-terminal tail piece exerting an influence on IgM oligomerization [38]. Additionally, crystal structure studies on Ig molecules reveal that carbohydrates influence Ig conformation, as it has been found that the conformation of the IgG4 CH2 domain to which the oligosaccharide is attached is altered in the absence of carbohydrate [39]. In case of grass carp IgM, our recent paper and the results in this study illustrated that the tetrameric IgM Asn⁵⁶⁵ glycan site lacked oligosaccharide, and carbohydrates linked to Asn⁵⁶⁵ sites were found only in monomers, with greatly reduced amounts on dimeric IgM [34]. Oligosaccharides are bulky and hydrophilic, and they might exist at the surface of the IgM, likely influencing cysteine redox-coupled processes. It has been reported that molecular chaperones ERp44/ERGIC53 bind to the conserved Asn⁵⁶³ glycan on human IgM molecules, controlling a thiol retention mechanism via Cys⁵⁷⁵ with the ERp44/ERGIC53 assembly platform in the endoplasmic reticulum [40–42].

To date, disulfide linkage mapping is a crucial component of mAb characterization for ensuring mAb drug safety and efficacy. It has been reported that some disulfide bonds broken in IgG1-based mAb drug as rituximab could change the binding affinity to FcγRIIIA receptor [43],

Additional disulfide bonds added into recombinant proteins could increase protein stability and prolonged protein half-life [44,45]. The observation that disulfide polymerization of trout IgM confers significantly longer serum half-life compared to lightly polymerized IgM is an interesting finding suggesting that disulfide polymerization could govern IgM half-life [46]. Disulfide polymerization could increase antigen binding affinity/avidity and regulate Ab effector functions, such as binding with FcR and complement activation, thereby leading to some level of physiological control. Evidence has been presented that human hexameric IgM activates the complement system 15- to 20-fold more efficiently than pentameric IgM [47]. The disulfide bond-mediated human IgG2 isoform has been revealed to possess different antigen binding activities [9,48]. Structural diversity is the universal feature of teleost IgM, and functional implications related to different teleost IgM isoforms are an enigmatic issue for investigation.

In summary, this study provides experimental analysis of disulfide bonding and free cysteines of the native grass carp IgM isoforms. Our studies illustrate that disulfide bond heterogeneity at the C-terminal tail piece dominates grass carp IgM isoform formation. The cysteine residue at the C-terminal tail piece is highly conserved in many species of teleost IgM (trout, zebrafish, channel catfish) [49–51]. It is assumed that the finding that the cysteine at the C-terminal tail piece is a key residue responsible for inter-monomeric disulfide bonding, which dominates grass carp IgM isoforms, might also be true for many teleost IgM isoforms.

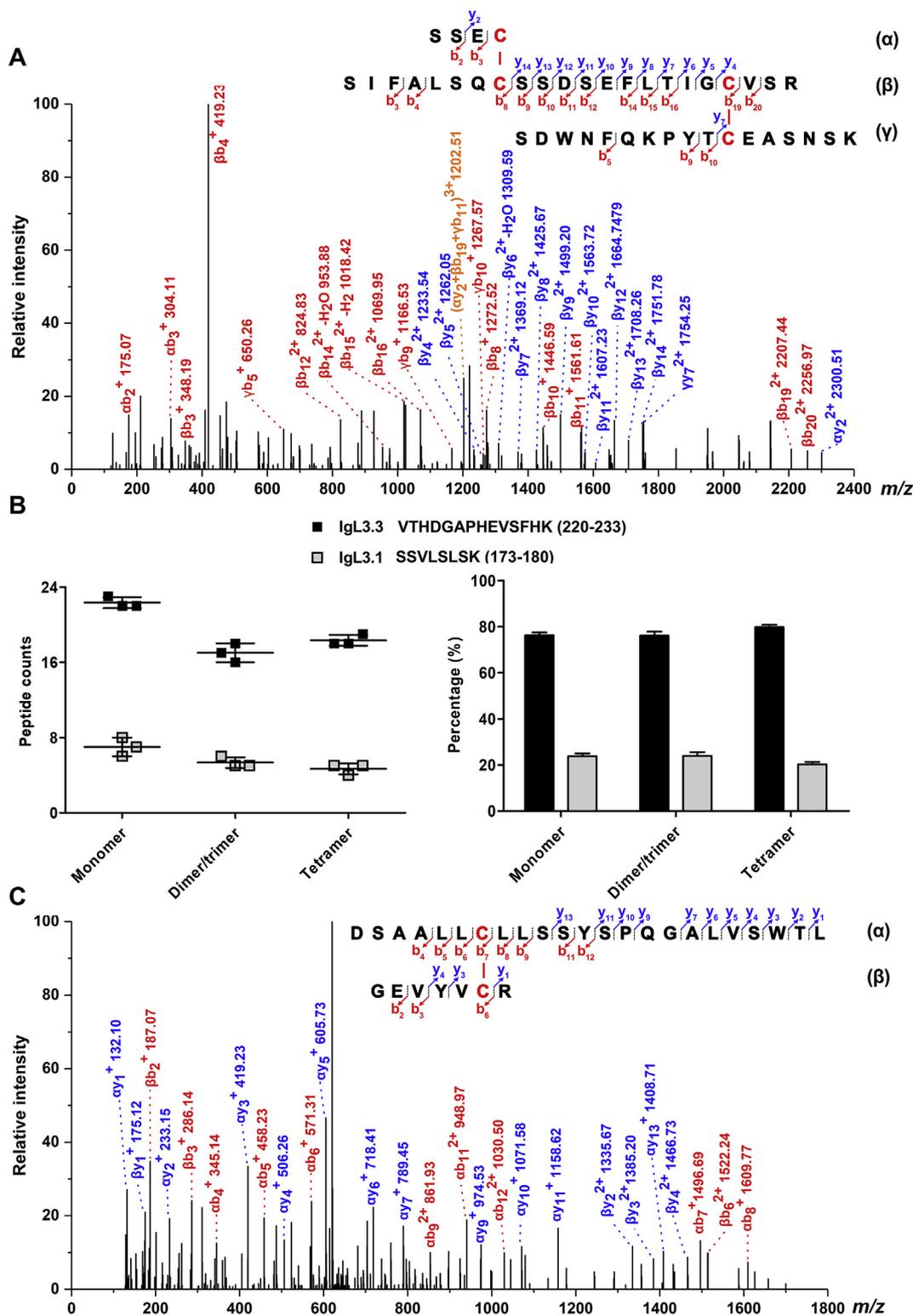


Fig. 6. Assignments of the CH1, CL1 and heavy-light chain disulfide bonds. (A) HCD spectrum of the two disulfide bond-peptides in the CH1 domain and between heavy-light chains. Fragment ions, y and b ions resulting from the cleavage of each peptide with and without disulfide bonds are designated and annotated. (B) Analyses of the IgL3.3 and IgL3.1% in grass carp IgM. The percentage was calculated by the sum of the total spectral counts of IgL3.3 and IgL3.1 unique peptides. Error bars indicate the standard deviation of the same IgM samples with three repeats. (C) Representative HCD spectrum that supports the assignment of disulfide bonding in CL1 in IgL3.3.

Table 1
Summary of disulfide bond assignment in Fc region and C-terminal tail piece of grass carp IgM isoform.

Domain	Disulfide-bond peptides	Charge	Theoretical <i>m/z</i>	Monomeric IgM		Dimeric/trimeric IgM		Tetrameric IgM	
				Observed <i>m/z</i>	Mass error (ppm)	Observed <i>m/z</i>	Mass error (ppm)	Observed <i>m/z</i>	Mass error (ppm)
CH2	DLEN*GTATFLC ²⁶⁹ LAQQFSPK	3 ⁺	958.442	958.447	4.904	958.451	9.390	958.439	-3.130
	C ³³⁰ EF ³⁷⁸ EHK	4 ⁺	719.083	719.089	8.205	719.086	3.059	719.091	5.285
CH3	EGTLTC ³⁷⁸ K	3 ⁺	1229.911	1229.918	3.792	N.D.	—	1229.918	5.204
	VELEAPIGVEEWSNGTVFTC ⁴³² TVEHTK	4 ⁺	922.685	922.692	7.727	922.693	8.454	922.6924	8.0201
		5 ⁺	738.349	738.353	5.204	738.346	-5.281	738.3525	4.7403
CH4	EGETMTLTC ⁴⁷⁸ YVK	3 ⁺	1100.5100	1100.5013	3.7922	1100.5124	2.1808	1100.5013	-7.9054
	SGIVYSC ⁵³⁶ VVFHESIDEK	4 ⁺	825.6344	825.6415	7.7274	825.6210	-4.1180	825.6415	8.5994
		5 ⁺	660.7090	660.7102	5.2036	660.7079	-1.6649	660.7102	1.8162
H-H	EAAYTDNNHDDC ³⁵³ TNVAIVPPSLEDMLK	5 ⁺	1258.3793	1258.3823	2.3841	1258.3786	4.6091	1258.3823	2.3840
	EAAYTDNNHDDC ³⁵³ TNVAIVPPSLEDMLK	6 ⁺	1048.8173	1048.8237	5.8445	1048.8228	9.1532	1048.8237	6.1021
		7 ⁺	899.1302	N.D.	—	899.1381	8.7640	899.1381	8.7418
Ctp	SIDDHIDKPGVIN ⁵ LSMNTPASC ⁵⁷⁴ KD ^a	3 ⁺	904.7546	904.7601	6.1011	N.D.	—	N.D.	—
		4 ⁺	678.8177	678.8229	7.1300	N.D.	—	N.D.	—
	SIDDHIDKPGVINLSMNTPASC ⁵⁷⁴ KD ^a	3 ⁺	904.4266	904.4185	3.7922	N.D.	—	N.D.	—
		4 ⁺	678.5697	678.5619	5.2036	N.D.	—	N.D.	—
	SIDDHIDKPGVINLSMNTPASC ⁵⁷⁴ KD	4 ⁺	1289.6017	N.D.	—	1289.6052	2.7140	1289.5996	-1.6517
	SIDDHIDKPGVINLSMNTPASC ⁵⁷⁴ KD	5 ⁺	1031.8828	N.D.	—	1031.8715	-5.1362	1031.8801	-2.6166
		6 ⁺	860.0702	N.D.	—	860.0724	2.5579	860.0705	0.2907

An asterisk denoted deglycosylation ($\Delta m = +0.98$ Da), ^acysteine masked with NEM ($\Delta m = +125.13$ Da), oxidation of methionine ($\Delta m = +125.13$ Da) shown in underline.

N.D. not determined. Ctp, C-terminal tail piece.

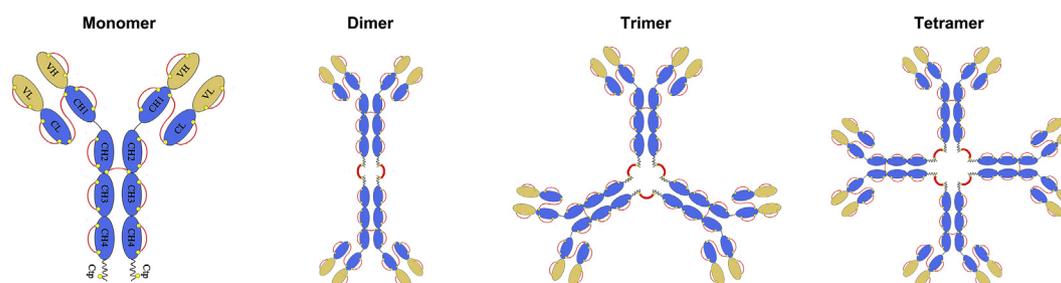


Fig. 7. Disulfide structural isoforms of grass carp IgM with varying disulfide bond involving the C terminal tail piece.

Funding

This work was supported by the Natural Science Foundation of China (21804048), the Fundamental Research Funds for the Central University of Central China Normal University (CCNU16A02044), and the Hubei Province Natural Science Key Foundation Project (2014CFA079).

Declaration of competing interest

The authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.fsi.2019.10.051>.

References

- [1] M.J. Feige, L.M. Hendershot, J. Buchner, How antibodies fold, *Trends Biochem. Sci.* 35 (4) (2010) 189–198.
- [2] T. Wan, R.L. Beavil, S.M. Fabiane, A.J. Beavil, M.K. Sohi, M. Keown, R.J. Young, A.J. Henry, R.J. Owens, H.J. Gould, B.J. Sutton, The crystal structure of IgE Fc reveals an asymmetrically bent conformation, *Nat. Immunol.* 3 (7) (2002) 681–686.
- [3] T.H. Truong, K.S. Carroll, Redox regulation of epidermal growth factor receptor signaling through cysteine oxidation, *Biochemistry* 51 (50) (2012) 9954–9965.
- [4] A. McAuley, J. Jacob, C.G. Kolvenbach, K. Westland, H.J. Lee, S.R. Brych, D. Rehder, G.R. Kleemann, D.N. Brems, M. Matsumura, Contributions of a disulfide bond to the structure, stability, and dimerization of human IgG1 antibody CH3 domain, *Protein Sci. : Publ. Protein Soc.* 17 (1) (2008) 95–106.
- [5] H. Liu, K. May, Disulfide bond structures of IgG molecules: structural variations, chemical modifications and possible impacts to stability and biological function, *mAbs* 4 (1) (2012) 17–23.
- [6] Y. Zuo, B. Xiang, J. Yang, X. Sun, Y. Wang, H. Cang, J. Yi, Oxidative modification of caspase-9 facilitates its activation via disulfide-mediated interaction with Apaf-1, *Cell Res.* 19 (4) (2009) 449–457.
- [7] N.M. Giles, G.I. Giles, C. Jacob, Multiple roles of cysteine in biocatalysis, *Biochem. Biophys. Res. Commun.* 300 (1) (2003) 1–4.
- [8] F.W. Putnam, G. Florent, C. Paul, T. Shinoda, A. Shimizu, Complete amino acid sequence of the μ heavy chain of a human IgM immunoglobulin, *Science (N.Y.)* 182 (4109) (1973) 287–291.
- [9] J. Wypych, M. Li, A. Guo, Z. Zhang, T. Martinez, M.J. Allen, S. Fodor, D.N. Kelner, G.C. Flynn, Y.D. Liu, P.V. Bondarenko, M.S. Ricci, T.M. Dillon, A. Balland, Human IgG2 antibodies display disulfide-mediated structural isoforms, *J. Biol. Chem.* 283 (23) (2008) 16194–16205.
- [10] B. Frangione, C. Milstein, Variations in the S-S bridges of immunoglobins G: interchain disulfide bridges of gamma G3 myeloma proteins, *J. Mol. Biol.* 33 (3) (1968) 893–906.
- [11] W. Zhang, L.A. Marzilli, J.C. Rouse, M.J. Czupryn, Complete disulfide bond assignment of a recombinant immunoglobulin G4 monoclonal antibody, *Anal. Biochem.* 311 (1) (2002) 1–9.
- [12] J.C. Lakub, D.F. Clark, I.S. Shah, Z. Zhu, E.P. Go, T.J. Tolbert, H. Desaire, Disulfide bond characterization of endogenous IgG3 monoclonal antibodies using LC-MS: an investigation of IgG3 disulfide-mediated isoforms, *Anal. Methods : Adv. Methods Appl.* 8 (31) (2016) 6046–6055.
- [13] B. Zhang, A.G. Harder, H.M. Connelly, L.L. Maheu, S.L. Cockrill, Determination of Fab-hinge disulfide connectivity in structural isoforms of a recombinant human immunoglobulin G2 antibody, *Anal. Chem.* 82 (3) (2010) 1090–1099.
- [14] A.C. Davis, K.H. Roux, J. Pursey, M.J. Shulman, Intermolecular disulfide bonding in IgM: effects of replacing cysteine residues in the μ heavy chain, *EMBO J.* 8 (9) (1989) 2519–2526.
- [15] R. Muller, M.A. Grawert, T. Kern, T. Madl, J. Peschek, M. Sattler, M. Groll,

- J. Buchner, High-resolution structures of the IgM Fc domains reveal principles of its hexamer formation, *Proc. Natl. Acad. Sci. U.S.A.* 110 (25) (2013) 10183–10188.
- [16] T.D. Randall, J.W. Brewer, R.B. Corley, Direct evidence that J chain regulates the polymeric structure of IgM in antibody-secreting B cells, *J. Biol. Chem.* 267 (25) (1992) 18002–18007.
- [17] Y.A. Zhang, I. Salinas, J. Oriol Sunyer, Recent findings on the structure and function of teleost IgT, *Fish Shellfish Immunol.* 31 (5) (2011) 627–634.
- [18] M.C. Piazzon, J. Galindo-Villegas, P. Pereiro, I. Estensoro, J.A. Calduch-Giner, E. Gomez-Casado, B. Novoa, V. Mulero, A. Sitja-Bobadilla, J. Perez-Sanchez, Differential modulation of IgT and IgM upon parasitic, bacterial, viral, and dietary challenges in a perciform fish, *Front. Immunol.* 7 (2016) 637.
- [19] S. Bilal, K.K. Lie, O.A. Karlsen, I. Hordvik, Characterization of IgM in Norwegian cleaner fish (lumpfish and wrasses), *Fish Shellfish Immunol.* 59 (2016) 9–17.
- [20] I. Hordvik, F.S. Berven, S.T. Solem, F. Hatten, C. Endresen, Analysis of two IgM isotypes in Atlantic salmon and brown trout, *Mol. Immunol.* 39 (5–6) (2002) 313–321.
- [21] A. Kamil, A. Raae, P.G. Fjelldal, E.O. Koppang, K.E. Fladmark, I. Hordvik, Comparative analysis of IgM sub-variants in salmonid fish and identification of a residue in mu3 which is essential for MAb4C10 reactivity, *Fish Shellfish Immunol.* 34 (2) (2013) 667–672.
- [22] S. Mashoof, M.F. Criscitiello, *Fish Immunoglobulins, Biology (Basel)* 5 (4) (2016).
- [23] J.O. Sunyer, Fishing for mammalian paradigms in the teleost immune system, *Nat. Immunol.* 14 (4) (2013) 320–326.
- [24] E.S. Bromage, J. Ye, S.L. Kaattari, Antibody structural variation in rainbow trout fluids, *Comparative biochemistry and physiology, Part B, Biochem. Mol. Biol.* 143 (1) (2006) 61–69.
- [25] C.J. Lobb, Covalent structure and affinity of channel catfish anti-dinitrophenyl antibodies, *Mol. Immunol.* 22 (8) (1985) 993–999.
- [26] S. Ryo, R.H. Wijdeven, A. Tyagi, T. Hermesen, T. Kono, I. Karunasagar, J.H. Rombout, M. Sakai, B.M. Verburg-van Kemenade, R. Savan, Common carp have two subclasses of bonyfish specific antibody IgZ showing differential expression in response to infection, *Dev. Comp. Immunol.* 34 (11) (2010) 1183–1190.
- [27] R.N. Perham, A diagonal paper-electrophoretic technique for studying amino acid sequences around the cysteine and cystine residues of proteins, *Biochem. J.* 105 (3) (1967) 1203–1207.
- [28] B. Zhang, S.L. Cockrill, Methodology for determining disulfide linkage patterns of closely spaced cysteine residues, *Anal. Chem.* 81 (17) (2009) 7314–7320.
- [29] A.A. McCarthy, P.W. Haebel, A. Torronen, V. Rybin, E.N. Baker, P. Metcalf, Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*, *Nat. Struct. Biol.* 7 (3) (2000) 196–199.
- [30] D. Sharma, K. Rajarathnam, ¹³C NMR chemical shifts can predict disulfide bond formation, *J. Biomol. NMR* 18 (2) (2000) 165–171.
- [31] J.C. Lakub, J.T. Shipman, H. Desaire, Recent mass spectrometry-based techniques and considerations for disulfide bond characterization in proteins, *Anal. Bioanal. Chem.* 410 (10) (2018) 2467–2484.
- [32] M.M. Quick, C.M. Crittenden, J.A. Rosenberg, J.S. Brodbelt, Characterization of disulfide linkages in proteins by 193 nm ultraviolet photodissociation (UVPD) mass spectrometry, *Anal. Chem.* 90 (14) (2018) 8523–8530.
- [33] S. Lu, Y. Cao, S.B. Fan, Z.L. Chen, R.Q. Fang, S.M. He, M.Q. Dong, Mapping disulfide bonds from sub-micrograms of purified proteins or micrograms of complex protein mixtures, *Biophys. Rep.* 4 (2) (2018) 68–81.
- [34] Y.L. Su, B. Wang, M.D. Hu, Z.W. Cui, J. Wan, H. Bai, Q. Yang, Y.F. Cui, C.H. Wan, L. Xiong, Y.A. Zhang, H. Geng, Site-specific N-glycan characterization of grass carp serum IgM, *Front. Immunol.* 9 (2018) 2645.
- [35] H. Boshra, A.E. Gelman, J.O. Sunyer, Structural and functional characterization of complement C4 and C1s-like molecules in teleost fish: insights into the evolution of classical and alternative pathways, *J. Immunol.* 173 (1) (2004) 349–359.
- [36] S. Lu, S.B. Fan, B. Yang, Y.X. Li, J.M. Meng, L. Wu, P. Li, K. Zhang, M.J. Zhang, Y. Fu, J. Luo, R.X. Sun, S.M. He, M.Q. Dong, Mapping native disulfide bonds at a proteome scale, *Nat. Methods* 12 (4) (2015) 329–331.
- [37] N. Zhang, X.J. Zhang, Y.L. Song, X.B. Lu, D.D. Chen, X.Q. Xia, J.O. Sunyer, Y.A. Zhang, Preferential combination between the light and heavy chain isotypes of fish immunoglobulins, *Dev. Comp. Immunol.* 61 (2016) 169–179.
- [38] D. Pasalic, B. Weber, C. Giannone, T. Anelli, R. Muller, C. Fagioli, M. Felkl, C. John, M.F. Mossuto, C.F.W. Becker, R. Sitia, J. Buchner, A peptide extension dictates IgM assembly, *Proc. Natl. Acad. Sci. U.S.A.* 114 (41) (2017) E8575–e8584.
- [39] A.M. Davies, R. Jefferis, B.J. Sutton, Crystal structure of deglycosylated human IgG4-Fc, *Mol. Immunol.* 62 (1) (2014) 46–53.
- [40] T. Anelli, S. Ceppi, L. Bergamelli, M. Cortini, S. Masciarelli, C. Valetti, R. Sitia, Sequential steps and checkpoints in the early exocytic compartment during secretory IgM biogenesis, *EMBO J.* 26 (19) (2007) 4177–4188.
- [41] S. Watanabe, M. Harayama, S. Kanemura, R. Sitia, K. Inaba, Structural basis of pH-dependent client binding by ERp44, a key regulator of protein secretion at the ER-Golgi interface, *Proc. Natl. Acad. Sci. U.S.A.* 114 (16) (2017) E3224–e3232.
- [42] L. Hampe, M. Radjainia, C. Xu, P.W. Harris, G. Bashiri, D.C. Goldstone, M.A. Brimble, Y. Wang, A.K. Mitra, Regulation and quality control of adiponectin assembly by endoplasmic reticulum chaperone ERp44, *J. Biol. Chem.* 290 (29) (2015) 18111–18123.
- [43] M. Suzuki, A. Yamanoi, Y. Machino, E. Kobayashi, K. Fukuchi, M. Tsukimoto, S. Kojima, J. Kohroki, K. Akimoto, Y. Masuho, Cleavage of the interchain disulfide bonds in rituximab increases its affinity for FcγRIIIA, *Biochem. Biophys. Res. Commun.* 436 (3) (2013) 519–524.
- [44] Y. Machino, H. Ohta, E. Suzuki, S. Higurashi, T. Tezuka, H. Nagashima, J. Kohroki, Y. Masuho, Effect of immunoglobulin G (IgG) interchain disulfide bond cleavage on efficacy of intravenous immunoglobulin for immune thrombocytopenic purpura (ITP), *Clin. Exp. Immunol.* 162 (3) (2010) 415–424.
- [45] L. Zhang, C.P. Chou, M. Moo-Young, Disulfide bond formation and its impact on the biological activity and stability of recombinant therapeutic proteins produced by *Escherichia coli* expression system, *Biotechnol. Adv.* 29 (6) (2011) 923–929.
- [46] J. Ye, E.S. Bromage, S.L. Kaattari, The strength of B cell interaction with antigen determines the degree of IgM polymerization, *J. Immunol.* 184 (2) (2010) 844–850.
- [47] T.D. Randall, L.B. King, R.B. Corley, The biological effects of IgM hexamer formation, *Eur. J. Immunol.* 20 (9) (1990) 1971–1979.
- [48] T.M. Dillon, M.S. Ricci, C. Vezina, G.C. Flynn, Y.D. Liu, D.S. Rehder, M. Plant, B. Henkle, Y. Li, S. Deechongkit, B. Varnum, J. Wypych, A. Balland, P.V. Bondarenko, Structural and functional characterization of disulfide isoforms of the human IgG2 subclass, *J. Biol. Chem.* 283 (23) (2008) 16206–16215.
- [49] J.D. Hansen, E.D. Landis, R.B. Phillips, Discovery of a unique Ig heavy-chain isotype (IgT) in rainbow trout: implications for a distinctive B cell developmental pathway in teleost fish, *Proc. Natl. Acad. Sci. U.S.A.* 102 (19) (2005) 6919–6924.
- [50] N. Danilova, J. Bussmann, K. Jekosch, L.A. Steiner, The immunoglobulin heavy-chain locus in zebrafish: identification and expression of a previously unknown isotype, immunoglobulin Z, *Nat. Immunol.* 6 (3) (2005) 295–302.
- [51] M.R. Wilson, A. Marcuz, F. van Ginkel, N.W. Miller, L.W. Clem, D. Middleton, G.W. Warr, The immunoglobulin M heavy chain constant region gene of the channel catfish, *Ictalurus punctatus*: an unusual mRNA splice pattern produces the membrane form of the molecule, *Nucleic Acids Res.* 18 (17) (1990) 5227–5233.