

RESEARCH ARTICLE

QTY code designed antibodies for aggregation prevention: A structural bioinformatic and computational study

Mengke Li^{1,2}  | Yanze Wang³  | Fei Tao²  | Ping Xu²  | Shuguang Zhang¹ 

¹Laboratory of Molecular Architecture, Media Lab, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

³State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China

²Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Correspondence

Shuguang Zhang, Laboratory of Molecular Architecture, Media Lab, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA.
Email: shuguang@mit.edu

Funding information

Shanghai Jiao Tong University

Abstract

Therapeutic monoclonal antibodies are the most rapidly growing class of molecular medicine, and they are beneficial to the treatment of a broad spectrum of human diseases. However, the aggregation of antibodies during the process of manufacture, distribution, and storage poses significant challenges, potentially compromising efficacy and inducing adverse immune responses. We previously conceived a QTY (glutamine, threonine, tyrosine) code, a simple tool for enhancing protein water-solubility by systematically pairwise replacing hydrophobic residues L (leucine), V (valine)/I (isoleucine), and F (phenylalanine). The QTY code offers a promising alternative to traditional methods of controlling aggregation in integral transmembrane proteins. In this study, we designed variants of four antibodies applying the QTY code, changing only the β -sheets. Through the structure-based aggregation analysis, we found that these QTY antibody variants demonstrated significantly decreased aggregation propensity compared to their wild-type counter parts. Our results of molecular dynamics simulations showed that the design by QTY code is capable of maintaining the antigen-binding affinity and structural stability. Our structural informatic and computational study suggests that the QTY code offers a significant potential in mitigating antibody aggregation.

KEYWORDS

AlphaFold2 protein structural predictions, hydrophobic to hydrophilic conversion, hydrophobic/hydrophilic beta-sheet, molecular dynamics simulation, protein design

1 | INTRODUCTION

More than 100 antibody-based therapeutics have been approved and counting for the treatment of a diverse array of human diseases up to 2022, and this number is still growing rapidly in recent years.¹ Monoclonal antibodies (mAbs) have been widely employed to treat autoimmune diseases,^{2,3} chronic inflammatory diseases,⁴ various types of cancers,^{5,6} ophthalmic diseases,⁷ and more. Representative examples include: Infliximab, a mAb against Tumor Necrosis Factor-alpha (TNF- α), extensively employed in the treatment of rheumatoid arthritis⁸;

Trastuzumab, targeting the human epidermal growth factor receptor 2 (HER2), widely utilized in the management of early-stage breast cancer⁶; Ranibizumab, targeting vascular endothelial growth factor (VEGF-A), widely employed in the treatment of neovascular age-related macular degeneration.⁷ One of the most challenging issues in the manufacture, distribution, and storage of mAbs lies in the potential of mAb aggregation,⁹ which significantly limits the shelf-life of the mAbs in the point of care. Aggregation can not only compromise their efficacy, but also induce adverse immune response, even cause severe hypersensitivity responses,¹⁰ such as anaphylaxis.¹¹

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Proteins: Structure, Function, and Bioinformatics* published by Wiley Periodicals LLC.

A conventional solution to this problem is to optimize the environmental factors that affect the aggregation rate in the process of manufacture, formulation and storage of mAbs,¹⁰ such as temperature, protein concentration, pH, oxygen, shear forces, and the ionic strength.^{12–14} However, these methods tend to be costly, increasing complications for medical delivery and time-consuming.¹⁵ An alternative approach is to redesign the aggregation-prone hydrophobic patches to reduce the aggregation propensity of the antibodies.¹⁶ Numerous studies have demonstrated the effectiveness of this approach in reducing aggregation propensity while maintaining the bioactivity.^{15,17–20} Nonetheless, such methodology is case-by-case and dependent on the available structure inputs, such as crystal and CryoEM structures. There is no simple and universal method that can address the aggregation of diverse antibodies.

We previously reported a simple and general tool for enhancing protein water-solubility, QTY (glutamine, threonine, tyrosine) code for directly converting hydrophobic α -helices to hydrophilic α -helices in several integral membrane protein chemokine receptors.²¹ QTY code is based on two key molecular facts: (1) Several amino acids share strikingly similar structures: Leu (L) versus Gln (Q)/Asn (N), Ile (I)/Val (V) versus Thr (T) and Phe (F) versus Tyr (Y), despite the stark contrast between the hydrophobicity of L, I, V, and F and the hydrophilicity of Q, T, and Y. (2) The secondary structure propensity of Q, T, and Y is close to that of L, I/V, and F, respectively. Previous studies have experimentally demonstrated the robustness of QTY code on water-solubilization of proteins with transmembrane α -helices.^{21–25} The water-solubilization design by QTY code produced detergent-free transmembrane receptors that retained ligand-binding affinity^{21–23,25} and high thermostability. Structural informatic studies using the highly accurate machine learning-based structure prediction approach AlphaFold²⁶ (AF2) showed the significant superpositions between the native proteins and their QTY variants.^{27–30} This general QTY code could have far-reaching implications for designing protein water-solubility, opening the door to designing proteins for a wide range of applications for pharmaceutically therapeutic mAbs, protein, and peptide-based biologics that can treat a wide range of diseases.

The molecular structures of antibodies predominantly comprise of β -sheets, typically more than 50% of the entire protein sequence.³¹ The inherent inclination of β -strands or sheets to form unfavorable inter-molecular interactions can often result in the destabilization of soluble structures and the formation of aggregates without a precisely controlled register.³²

We ask if the QTY code could also be applied to β -sheets since the β -sheet propensity of Q, T, and Y is also close to that of L, I/V, and F, respectively.³³ Therefore, it is promising that we can design antibody variants using the QTY code to replace hydrophobic residues in β -sheets of the antibody structure with hydrophilic ones, thus reducing their propensity for aggregation.

Here, we report study of the designed QTY variants changing only the β -sheets of four antibodies with available full-length crystal structures: (i) IgG1 b12,³⁴ (ii) Pembrolizumab,³⁵ (iii) anti-NPRA IgG4,³⁶ and (iv) Mab 61.1.3.³⁷ We generated the full-length AF2 structure models of QTY variants of the four antibodies. Their structures were

all superposed well with the wild-type antibodies, for both antigen binding fragment (Fab) regions and crystallizable fragment (Fc) regions. We then evaluated the aggregation propensity of the wild-type and QTY antibodies by analyzing the three-dimensional structures and found that the QTY variants showed significantly decreased aggregation propensity, compared to the wild-type antibodies. Using molecular dynamics simulations, we showed that the QTY variant IgG1 b12^{QTY} interacted with the antigen with a close binding affinity in a similar manner, compared to wild-type IgG1 b12. We observed that the interactions of the polar residues introduced by QTY code mimicked the packing of the original hydrophobic residues in the β -sheets. Our study suggests the great potential of QTY code in mitigating aggregation of therapeutic antibodies.

2 | MATERIALS AND METHODS

2.1 | Design of QTY variants and characteristics calculation

The sequences of the selected antibodies were from PDB <https://www.rcsb.org>. The PDB entry numbers are listed in Table 1 and Table S1. The secondary structure information was extracted from the PDB files by PyMOL (version 2.0 Schrödinger, LLC). All the residues comprising β -sheets were regarded as the replacements of QTY design. The molecular weight and pI values of the proteins (combining the heavy and light chains) were calculated by the web server Expasy https://web.expasy.org/compute_pi/.

2.2 | Structure prediction by AF2 and structural analysis

The structure prediction of all proteins was performed using AF2 via the ColabFold³⁸ pipeline applying mostly default parameters (num_relax = 5, template mode: none, msa_mode: MMSeq2, max_recycles = 12). The resulting predicted structures with the highest model rank were used for subsequent analysis. PyMOL was used for structural visualization and analysis. The structures of the wild-type antibodies and their QTY variants were superposed by PyMOL “align” command, of which the cycle number was chosen when the objected atom number was the lowest. The monomeric Fab and Fc regions were aligned, separately, due to the flexibility of the hinge region.

2.3 | Aggregation propensity evaluation and visualization

The structure-based prediction of protein aggregation web server, Aggrescan3D (A3D) version 2.0³⁹ (<http://biocomp.chem.uw.edu.pl/A3D2/>), was used to evaluate the aggregation propensity and identify aggregation-prone regions of the antibodies. The monomeric Fab

regions and the dimeric Fc regions (without glycans) were input and analyzed, separately. For the settings, stability calculations: yes; dynamic mode: yes; mutate residues: no; distance of aggregation analysis: 10 Å. For calculation of the average A3D score of each protein, the scores of all 13 models (the dynamic models and the static model) were used. For Fc regions, the sites interacting with glycans were excluded for the average score calculation. P values were calculated by the unpaired two-tailed Student's t-test in the software Graphpad Prism (<https://www.graphpad.com>). For visualization of the aggregation-prone patches, the representative structure snapshots were directly exported from the A3D server.

2.4 | MD simulations and analysis

The simulations were performed using GROMACS 2022.3 (<https://manual.gromacs.org/2022.3/>). For wild-type IgG1 b12-gp120 complex, the available complex crystal structure (PDB: 2NY7) was used as the input structure. For IgG1 b12^{QTY}-gp120 complex, we first aligned the AF2 structure of IgG1 b12^{QTY} (Fab region) with IgG1 b12 of 2NY7, then manually removed some residue clashes by adjusting the side chain orientation in PyMOL and exported the IgG1 b12^{QTY} complexed with gp120 as the input structure. The topology files were established through Leap program of AmberTools22.⁴⁰ For the settings, ff14SB Amber force field⁴¹ and TIP3P water model were selected. To neutralize the system and enable the salt concentration of 150 mM, 141 Cl⁻ ions and 128 Na⁺ ions were added. The distance between the protein and the edge of the box was set to 12 Å. The system energy was minimized using the steepest descent method and converged to 1000 kJ/mol/nm. Electrostatics were treated with Particle Mesh Ewald, and the cutoff for both Coulomb and van der Waals interactions was 1.2 nm; 2 fs time step were used during the equilibration stage. The modified Berendsen thermostat was used, with system coupled to a 300 K bath. The Parrinello-Rahman barostat was used with isotropic coupling. Hydrogen bonds were constrained using LINCS algorithm. Finally, the production run of 50 ns was performed. Simulation frames were clustered (GROMACS built-in command “gmx cluster”) using linkage method (cutoff: 0.13 nm for wild-type, 0.125 nm for QTY) and the medoid structures were exported. To calculate the binding free energy between wild-type/QTY IgG1 b12 and gp120 during the simulation, gmx_mmpbsa tool⁴² was used. For settings: sys_name = “Prot-Prot,” forcefields = “leaprc.protein.ff14SB,” saltcon = 0.150, idecomp = 2, dec_verbose = 3, print_res = “within

5.” For the analysis of binding-contributing residues, gmx_mmpbsa_ana tool was used.

3 | RESULTS AND DISCUSSION

3.1 | Design of QTY variants of antibodies

We selected four antibodies as the subjects of design (Table 1), whose intact crystal structures are available, including both Fab regions and Fc regions. In this study, we only focus on applying the QTY code for the design of β -sheets since there are high content of β -sheets in all antibodies. For the residues comprising β -sheets in each antibody, because the striking structural similarity of the pairwise amino acids, namely, L:Q, I/V:T, F:Y, we replace all leucine (L) with glutamine (Q); isoleucine (I) and valine (V) with threonine (T); phenylalanine (F) with tyrosine (Y) (Figure 1, Figures S2–S5). We aligned the sequences of QTY variants with the wild-type antibodies. On average, about one-third residues in β -sheets were substituted and 14% ~ 17% in total were substituted, similarly for heavy chains (HCs) and light chains (LCs). We preliminarily analyzed the basic characteristics of the designed QTY variants. Despite the high variation rate introduced by QTY code, each QTY variant shows close molecular weight (MW) and isoelectric point (pI) value with those of the wild-type antibody, which is determined by the rationale of QTY code. QTY code adopts the residues of the comparable size for substitution and does not introduce any additional charge, so as to preserve the intrinsic properties of the proteins.

3.2 | Structure prediction of QTY variants and superposition with wild-type antibodies

In order to study the molecular structural properties of the QTY variants, we generated the structure models of full-length heterotetramers (~1300 amino acids) using AF2. By analyzing the predicted Local Distance Difference Test (pLDDT) data of all residues of the predicted models, Fab and Fc regions both showed a good prediction quality (Figure S1). The hinge regions showed poor quality, which was expected due to the intrinsic high flexibility.⁴³ The AF2 models of the four QTY variants all showed the similar manner of assembly of heterotetramers to native antibodies (Figure S1). Then, we superposed the structures of the QTY variants with the crystal structures of the

TABLE 1 The basic information of the selected antibodies.

Antibody	Target	Organism	PDB entry numbers
IgG1 b12	HIV-1 gp120 envelope glycoprotein	<i>Mus musculus</i>	1. 1HZH (2.70 Å) 2. 2NY7 (2.30 Å, complex)
Pembrolizumab	Programmed cell death protein 1	<i>Homo sapiens</i>	5DK3 (2.28 Å)
Anti-NPRA IgG4	Natriuretic peptide receptor A	<i>Homo sapiens</i>	6GFE (1.80 Å)
Mab 61.1.3	Phenobarbital	<i>Mus musculus</i>	1IGY (3.20 Å)

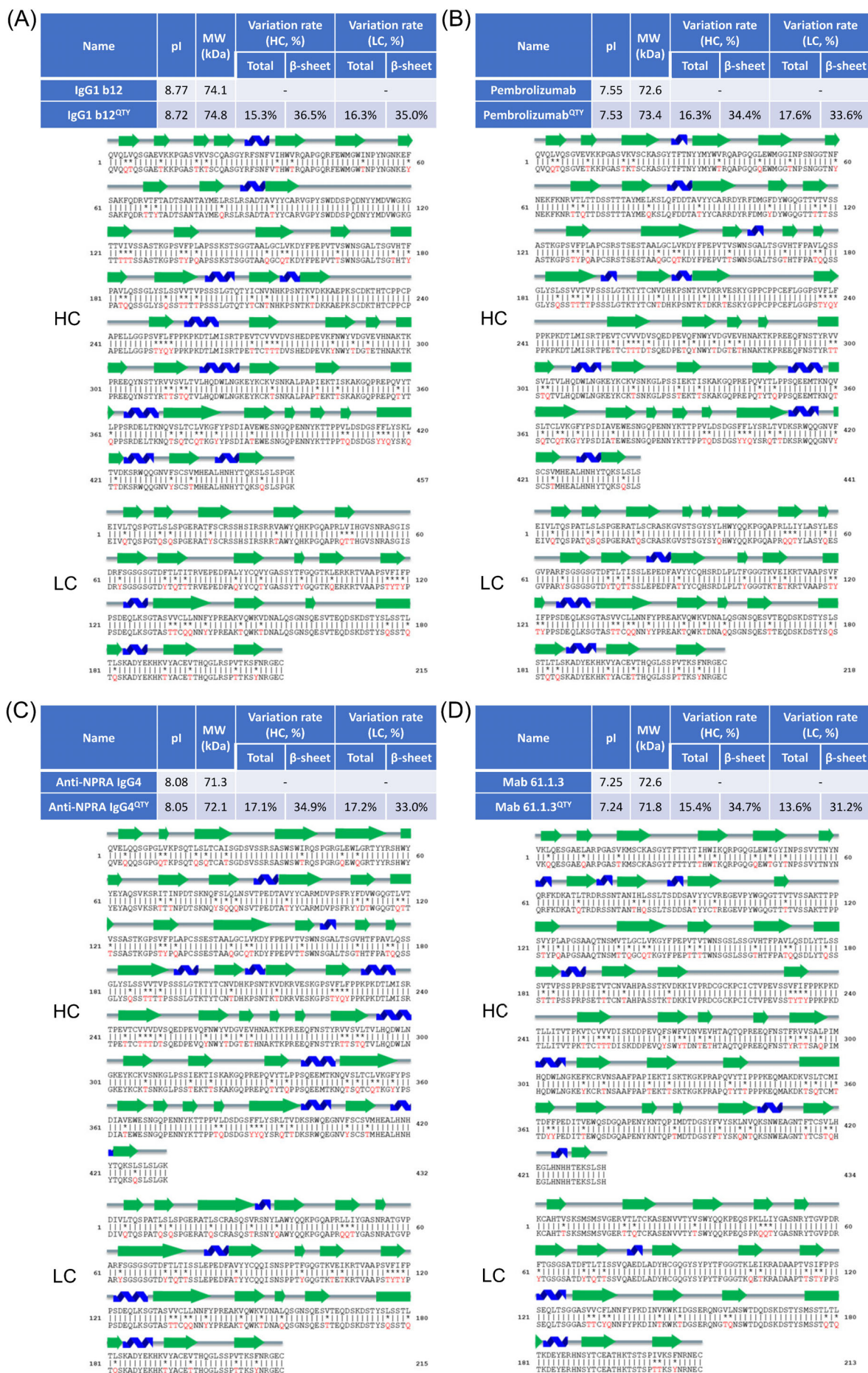


FIGURE 1 Legend on next page.

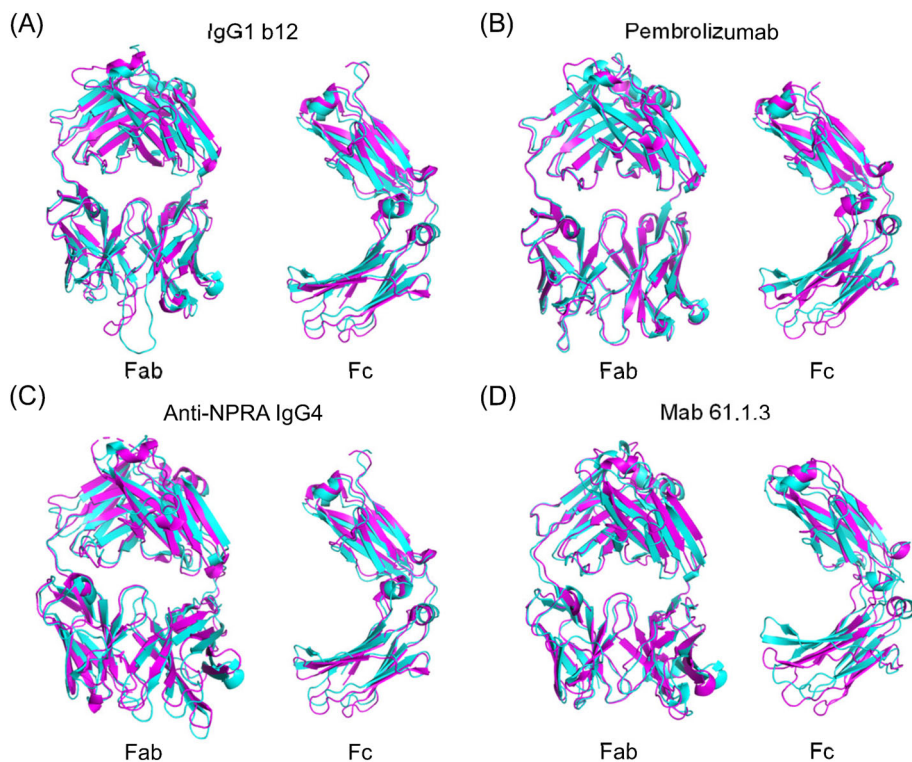


FIGURE 2 Structural superposition between the crystal structures of the four antibodies (magenta) and the structures of their QTY variants (cyan) predicted by AF2. For each antibody, Fab fragment superposition is on the left, Fc on the right. (A) IgG1 b12 versus IgG1 b12^{QTY} (RMSD: 2.107 Å, 1.670 Å). (B) Pembrolizumab versus Pembrolizumab^{QTY} (RMSD: 1.015 Å, 1.968 Å). (C) Anti-NPRA IgG4 versus Anti-NPRA IgG4^{QTY} (RMSD: 2.533 Å, 2.307 Å). (D) Mab 61.1.3 versus Mab 61.1.3^{QTY} (RMSD: 1.931 Å, 2.475 Å), respectively.

TABLE 2 RMSD between the crystal structures of the antibodies and the AF2 structures of their QTY variants.

Antibody	RMSD (Fab, Å)	RMSD (Fc, Å)
IgG1 b12	2.107	1.670
Pembrolizumab	1.015	1.968
Anti-NPRA IgG4	2.533	2.307
Mab 61.1.3	1.931	2.475

corresponding wild-type antibodies. Due to the flexibility of hinge regions, we performed the superposition of Fab regions and Fc regions separately (Figure 2). All the four QTY variants showed good structural similarity with the wild-type antibodies, for both Fab and Fc regions. The root-mean-square deviation (RMSD) values of all superposition are below 2.6 Å (Table 2). These results imply that despite the significant substitutions of hydrophobic residues with hydrophilic residues, QTY code preserved the overall structures of the antibodies.

3.3 | Aggregation propensity evaluation of wild-type/QTY antibodies

To assess if and how our design might ameliorate the aggregation of the antibodies, we compared the aggregation propensity of the

wild-type antibodies with their QTY variants by analyzing their three-dimensional structures using the tool A3D 2.0. According to the specific conformational context of the individual residue, A3D calculates a structurally corrected aggregation value (A3D score). By averaging the A3D scores of all residues in the input structure, we obtained an “average A3D score” to indicate the aggregation propensity of this structure. Moreover, A3D can generate 12 dynamic models by its dynamic mode, which can conduct fast simulations of protein flexibility, sampling multiple conformations instead of only the static model.

For each wild-type antibody and its QTY variant, we compared the average A3D scores, taking all 13 models (12 dynamic models and the static model) into account. For the Fab region, all the four antibodies showed significantly lower aggregation propensity after the water-solubilizing design by QTY code (QTY design) (Figure 3A–D), especially IgG1 b12 (b12) and Mab 61.1.3 (with p values < .01). We then analyzed the aggregation-prone regions of the structural models by A3D (Figure 4). Generally, by comparing the QTY variants with the wild-type antibodies, we observed the lower A3D scores for most residues, indicating the decreased overall aggregation propensity.

From these structural models, we observed a conserved aggregation-prone patch (in contrast to the individual aggregation-prone residues) in the Fab regions shared by all the four antibodies, which was previously identified as a hydrophobic region,¹⁷ or in sequence aspect as an aggregation-prone motif.⁴⁴ This motif,

FIGURE 1 Alignments of wild-type antibodies with QTY variants. (A–D) Wild-type (upper) IgG1 b12, Pembrolizumab, Anti-NPRA IgG4, Mab 61.1.3 versus their QTY variants (lower), respectively. The Q, T, and Y amino acid substitutions are in red. The secondary structure of each residue is shown above the protein sequences. Green arrows indicate β -sheets, blue helices indicate α -helices, and gray lines indicate loops. The symbols “|” and “*” indicate similar and different amino acids, respectively. Characteristics of wild-type antibodies and QTY variants with isoelectric point (pI), molecular weight (MW), variation rate in heavy/light chain (HC/LC) are shown in the tables.

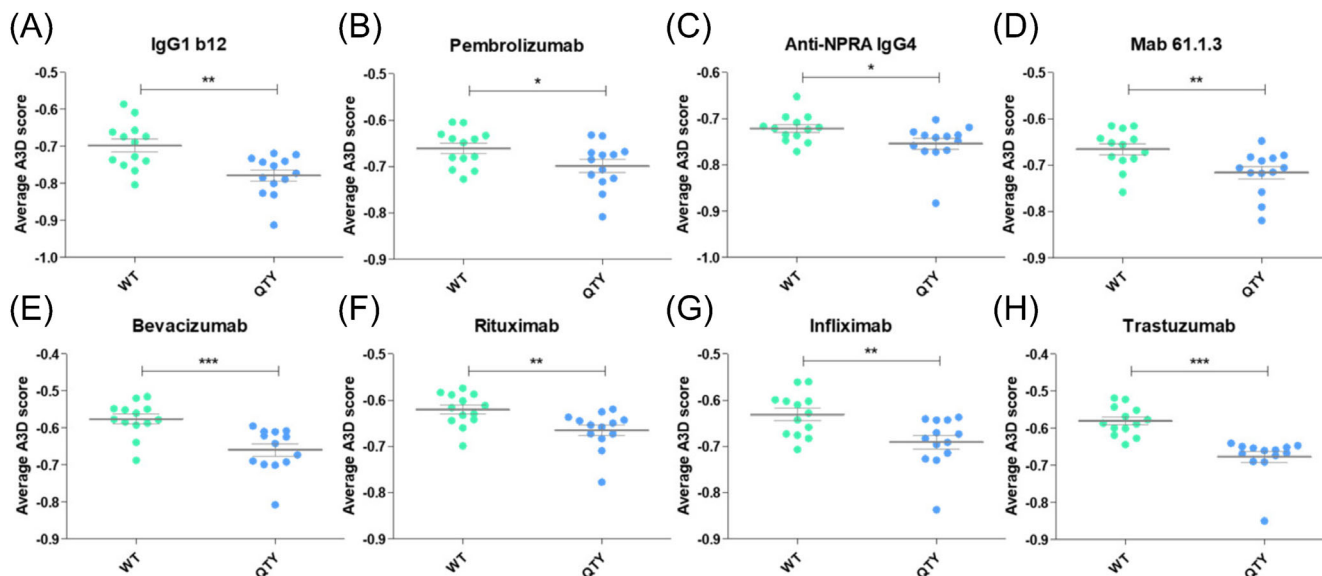


FIGURE 3 Aggregation propensity of the Fab regions of the antibodies was lowered down after QTY design. For each protein, the average A3D scores of 13 structure models (the dynamic models and the static model) were calculated by A3D 2.0, indicating the aggregation propensity. Low A3D scores indicate low aggregation propensity. For example, the average A3D score of IgG1 b12 is around -0.7 , which indicates a higher aggregation propensity compared to IgG1 b12^{QTY}, with the average A3D score around -0.8 . (A–H) Wild-type (WT, in green) IgG1 b12, Pembrolizumab, anti-NPRA IgG4, Mab 61.1.3, Bevacizumab, Rituximab, Infliximab, Trastuzumab versus their QTY variants (in blue), respectively. Data are presented as the mean \pm standard error of the mean (SEM). p values were calculated by the Student's t -test. * $p < .05$, ** $p < 0.01$, *** $p < .001$.

“AVLQSSGLY” (or “AVLQSDLY” in Mab 61.1.3), caused considerable local aggregation propensity, which, however, was greatly ameliorated by QTY design, for all the four cases, although only two hydrophobic residues were changed into polar residues (“ATQQSSGLY”). This could be interpreted by the high hydrophobicity resulted from these two successive highly-exposed hydrophobic residues (“VL”) and the effects of the surrounding water-solubilization tendency resulted from QTY design.

For the Fc region, the residues interacting with glycans were excluded for the average score calculation, because AF2 cannot predict the glycan structure. All the QTY variants showed significantly lower aggregation propensity compared with wild-type antibodies (three with p value $< .001$ and one $< .01$; Figure 5). For the analysis of aggregation-prone regions, similarly, we identified a conserved aggregation-prone patch, also previously identified as a hydrophobic region,¹⁷ or in sequence aspect as an aggregation-prone motif,⁴⁴ “VLDSG~~S~~FF” (or “IMDTDGSYF” in Mab 61.1.3), which was significantly eliminated by QTY design (Figure 6). In this motif, four hydrophobic residues were changed into polar residues (one in Mab 61.1.3). It is worth noting that these two conserved aggregation-prone patches are not reported to be engaged in formation of common complexes with other proteins.¹⁷ Thus, the design of these regions should have negligible effects on the functionality of the antibodies.

Some therapeutic mAbs are explicitly reported to be aggregation-prone. In order to show the potential of QTY code on ameliorating the aggregation of these antibodies, we selected another four therapeutic mAbs with available structures of the Fab region,

Bevacizumab,¹⁵ Rituximab,⁴⁵ Infliximab,⁸ and Trastuzumab⁴⁶ (Table S1). Using the same QTY code protocol, we designed the QTY variants of these antibodies and generated their AF2 structure models. The aggregation analysis using A3D showed that QTY design remarkably mitigated the aggregation propensity (two with p value $< .001$ and the other two $< .01$; Figure 3E–H). It is observed that the Fab region average A3D scores of these four antibodies (~ -0.6) were overall lower than those of IgG1 b12, Pembrolizumab, anti-NPRA IgG4, and Mab 61.1.3 (~ -0.7), supporting the aggregation-prone nature of these antibodies and the QTY design to reduce the aggregation.

Our results suggest that QTY design is capable of ameliorating the aggregation propensity of the antibodies.

3.4 | Molecular dynamics simulations of the wild-type/QTY antibody–antigen complex

In order to examine the effects of QTY design on the functionality and structural stability of the antibody, we conducted 50-ns all-atom molecular dynamic (MD) simulations of the Fab region of b12/b12^{QTY} complexed with the HIV CD4-binding domain of gp120 (~ 730 amino acids in total), respectively. The structure of b12 Fab region complexed with the antigen, HIV-1 envelope glycoprotein gp120, is available (PDB: 2NY7) and has been well studied.^{47–49} During the 50-ns simulation, both b12 and b12^{QTY} complex showed good structural stability (Figure 7A), with comparable backbone RMSD values, 1.72

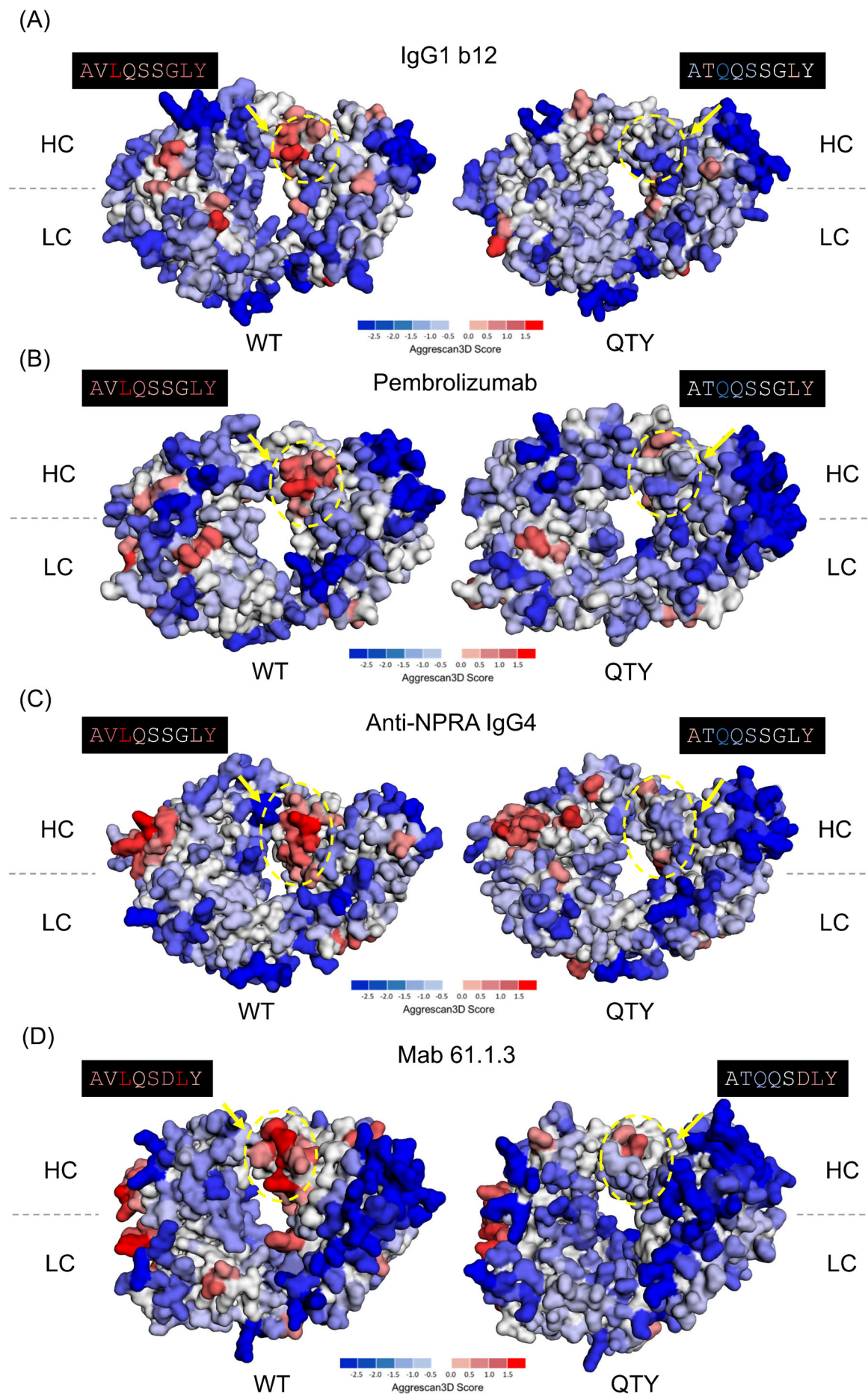


FIGURE 4 Legend on next page.

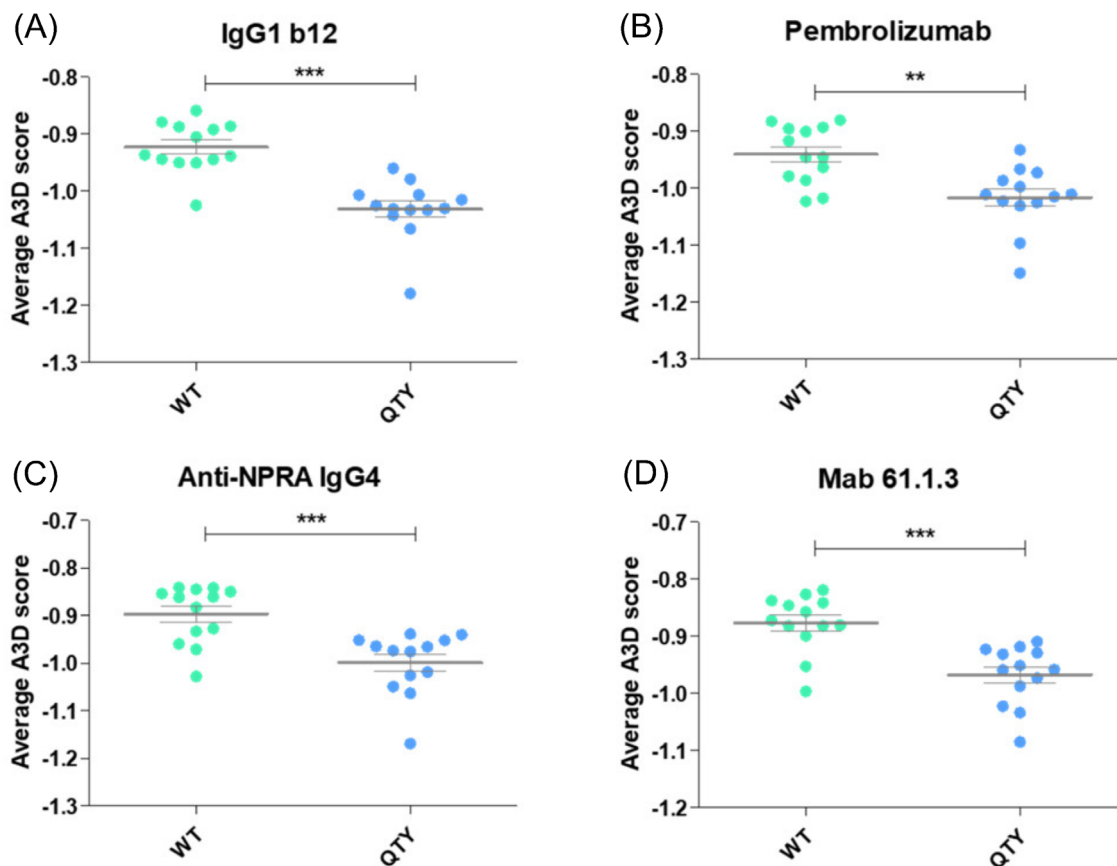


FIGURE 5 Aggregation propensity of the Fc regions of the antibodies was lowered down after QTY design. For each protein, the average A3D scores of 13 structure models (the dynamic models and the static model) were calculated by A3D 2.0, indicating the aggregation propensity (low A3D scores mean low aggregation propensity). For example, the average A3D score of IgG1 b12 is around -0.9 , which indicates a higher aggregation propensity compared to IgG1 b12^{QTY}, with the average A3D score around -1.0 . The residues interacting with glycans were excluded for the average score calculation. (A–D) Wild-type (WT, in green) IgG1 b12, Pembrolizumab, anti-NPRA IgG4, Mab 61.1.3 versus their QTY variants (in blue), respectively. Data are presented as the mean \pm SEM. *p* values were calculated by the Student's *t*-test. **p* < .05, ***p* < .01, ****p* < .001.

± 0.36 Å of b12, and 1.95 ± 0.29 Å of b12^{QTY}. The low standard deviation of RMSD in b12^{QTY} simulation also shows little significant structural deviation along the simulation.

To compare the binding affinity of b12 and b12^{QTY} with gp120, we used Molecular Mechanics Poisson-Boltzmann Surface Area (MMPBSA) algorithm⁴² to estimate the binding free energy. During the 50-ns simulation, the binding free energy of b12^{QTY}-gp120 complex (-64.7 ± 15.2 kcal/mol) was very close to wild-type b12 (-64.3 ± 7.2 kcal/mol; Figure 7B). These values are in good agreement with the value in previous MD simulations of b12-gp120 complex.⁴⁹ Next, we also analyzed the contributions of the individual residues to the binding on the wild-type/QTY b12-gp120 interface. In b12-gp120 complex simulation, we found a region that

contributed the most to the binding (Figure 7C), which was mainly composed of three bulky residues, Y98 and W100 in complementarity-determining region (CDR) H3 of b12, and R419 in gp120. These three residues were previous reported to be the main contributors to the binding.⁴⁹ In b12^{QTY}-gp120 complex simulation, likewise, this region was also the predominant contributor to the binding (Figure 7D).

Previous experimental data of the antigen-binding activity test in a study of engineering antibodies against aggregation showed that the activity was lost in the variants with mutations in the CDR regions, whereas the activity was retained in the variants with mutations elsewhere.¹⁷ Consistently, our QTY design changing only β -sheets circumvented these regions and minimized the effects on the functionality.

FIGURE 4 Visualization and comparison of the aggregation-prone regions of the Fab regions of native antibodies and their QTY variants. The structure snapshots are from the output structure models by A3D 2.0 dynamic modes <http://biocomp.chem.uw.edu.pl/A3D2/>. (A–D) Wild-type IgG1 b12, Pembrolizumab, anti-NPRA IgG4, Mab 61.1.3 versus their QTY variants, respectively. In the color scale, deep blue indicates the most soluble residues and deep red indicates the most aggregation-prone residues. As indicated by dash lines, the upper portion is the heavy chain (HC), lower the light chain (LC). The yellow dash circles indicate the conserved aggregation-prone patches, in the sequences of which are shown above the structures with black backgrounds.

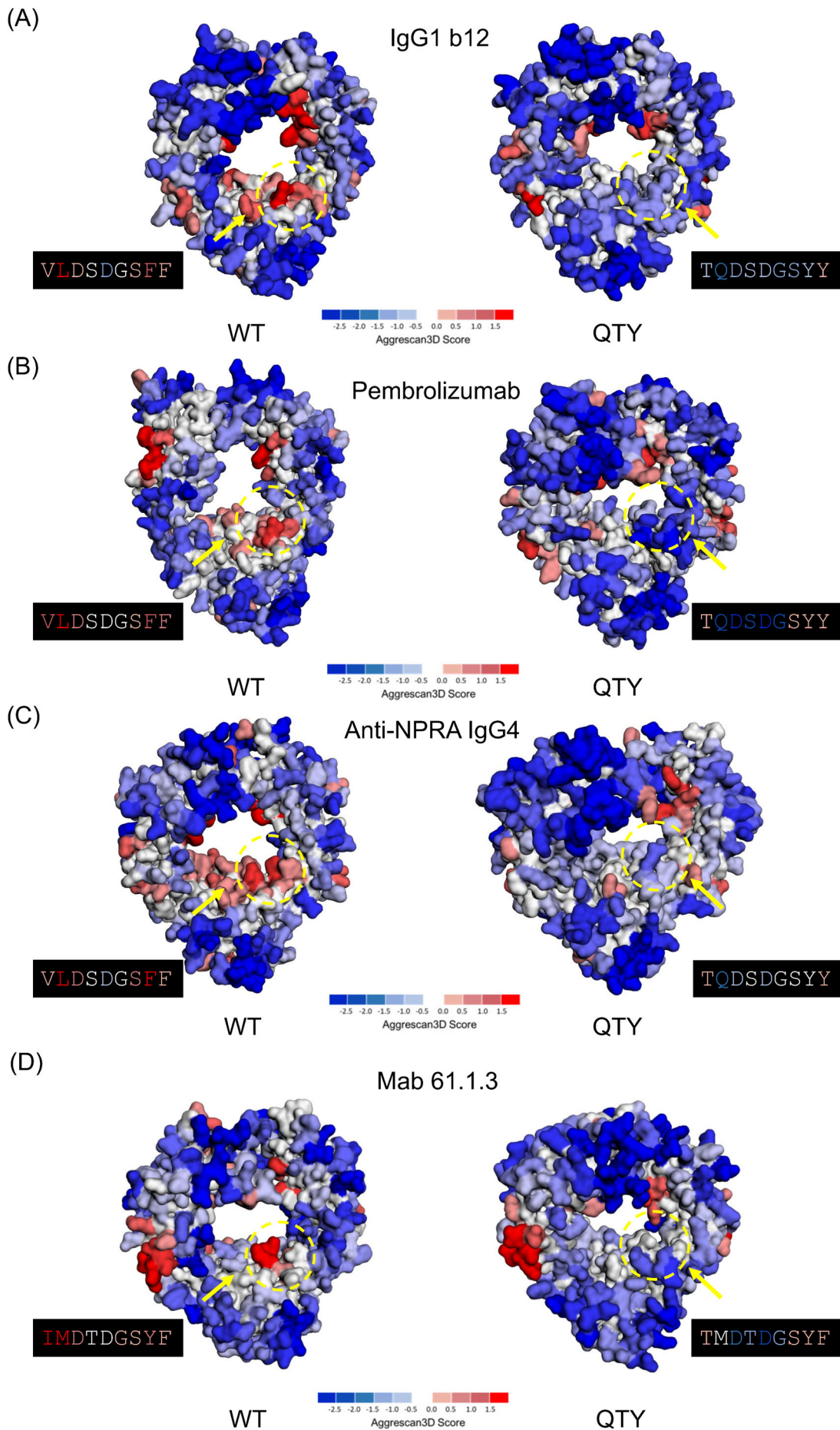


FIGURE 6 Legend on next page.

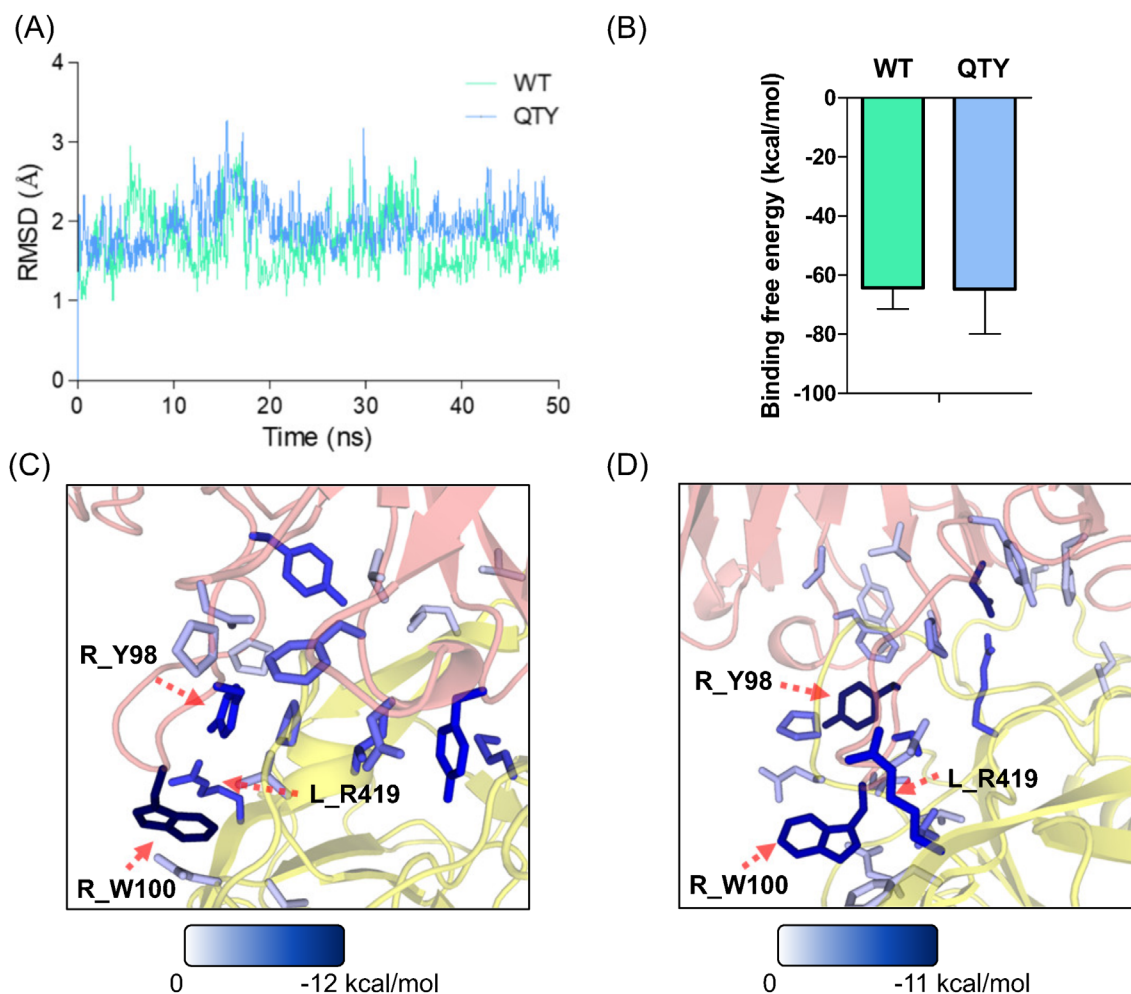


FIGURE 7 MD simulations of wild-type/QTY b12-gp120 complex. (A) RMSD curve of wild-type (in green)/QTY (in blue) b12-gp120 complex. (B) Binding free energy of wild-type/QTY b12-gp120 complex. Data are presented as the mean \pm standard deviation. (C,D) The major residues (binding energy < -2 kcal/mol) contributing to binding in the interface of wild-type (C)/QTY (D) b12-gp120 complex. The binding energy contribution of each residue is denoted by the color map below the figure. The antibody part (regarded as the receptor, R) is in red and gp120 (regarded as the ligand, L) is in yellow. For example, R_Y98 indicates Tyr at position 98 in the receptor, and L_R419 indicates Arg at position 419 in the ligand. The residue numbering is based on the sequence of PDB: 2NY7. The structure snapshots are from gmx_mmpbsa_ana tool.

We previously observed the formation of significant numbers of hydrogen-bonds (H-bonds) in the interior of the helical bundle of a QTY variant of α -helical membrane protein by MD simulations.²² These H-bonds would not exist in the native hydrophobic core areas with L, I/V, and F hydrophobic contacts. In this study in the context of β -sheet design, we asked how the polar residues introduced by QTY design (simplified as “QTY residues”) were engaged in the stabilization. We observed the interactions of these residues in b12^{QTY} mimicked the packing of the original hydrophobic residues in b12 (Figure 8A,B,E,F). In the cases shown in Figure 6, between the layers of β -sheets, the QTY residues tightly interacted with each other by forming inter-residue H-bonds, in similarity with

the Van der Waals contact of the original hydrophobic residues (Figure 8C,D,G,H).

In summary, these results indicate that QTY design targeting β -sheets could preserve the antigen-binding activity and maintain the structural stability.

4 | CONCLUSIONS

Protein aggregation is the predominant problem during the storage of mAbs for therapeutic use.¹⁷ QTY code, as a simple and robust tool of water-solubilizing design, has been demonstrated to be capable of

FIGURE 6 Visualization and comparison of the aggregation-prone regions of the Fc regions of native antibodies and their QTY variants. The structure snapshots are from the output structure models by A3D 2.0 dynamic modes <http://biocomp.chem.uw.edu.pl/A3D2/>. (A–D) Wild-type (WT, in green) IgG1 b12, Pembrolizumab, Anti-NPRA IgG4, Mab 61.1.3 versus their QTY variants, respectively. In the color scale, dark blue indicates the most soluble residues and dark red indicates the most aggregation-prone residues. The yellow dash circles indicate the conserved aggregation-prone patches, the sequences of which are shown above the structures with black backgrounds.

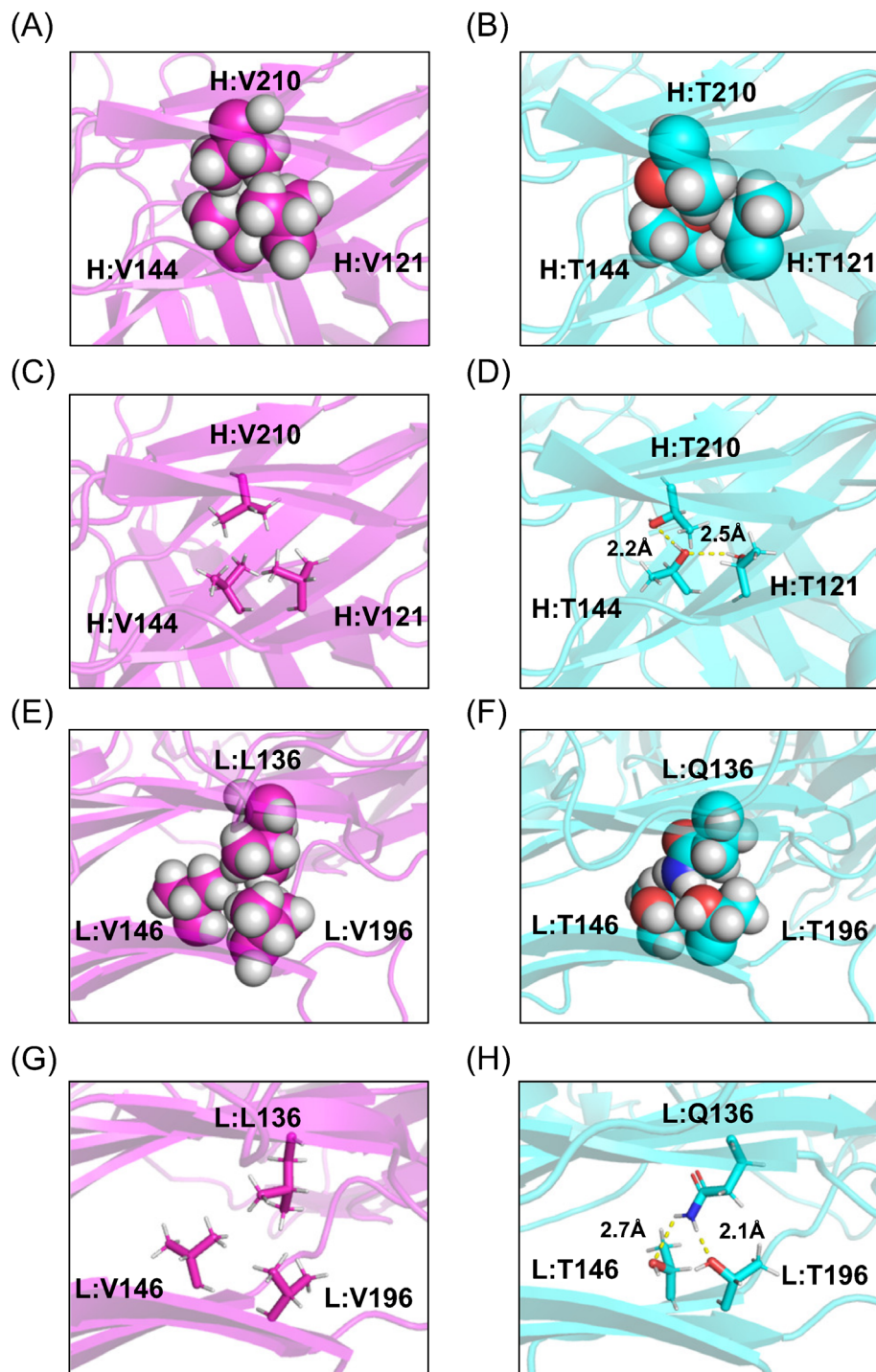


FIGURE 8 The interactions of polar residues in IgG1 b12^{QTY} mimicked the packing of the original hydrophobic residues in IgG1 b12. V121, V144, and V210 in heavy chain of native IgG1 b12 versus T121, T144, and T210 in heavy chain of IgG1 b12^{QTY}, shown as spheres, (A) versus (B), and sticks, (C) versus (D). L136, V146, and V196 in light chain of native IgG1 b12 versus Q136, T146 and T196 in light chain of IgG1 b12^{QTY}, shown as spheres, (E) versus (F), and sticks, (G) versus (H). Highlighted residues are shown as spheres. The hydrogen bonds are shown as yellow dashed line. The residue numbering is based on the sequence of PDB: 1HZH. H, heavy chain; L, light chain. The snapshots were from the simulation medoids.

designing the detergent-free, functional membrane proteins. In this study, our structural informatic and computational results showed that, compared to the wild-type antibodies, the antibody variants designed by QTY code exhibited lower aggregation propensity, with the comparable antigen-binding activity and good structural stability. We believe that QTY design targeting β -sheets performed as well as targeting α -helices previously. Our study suggests that QTY code is a promising and useful tool for mitigating the aggregation of therapeutic mAbs, and hopefully for other aggregation-prone proteins.

AUTHOR CONTRIBUTIONS

Mengke Li: Data curation; formal analysis; visualization; writing – original draft; methodology; investigation; validation; writing – review and editing. **Yanze Wang:** Formal analysis. **Fei Tao:** Writing – review and editing; funding acquisition. **Ping Xu:** Writing – review and editing; resources; funding acquisition. **Shuguang Zhang:** Conceptualization; writing – review and editing; investigation; supervision; project administration; writing – original draft.

ACKNOWLEDGMENTS

The authors acknowledge the MIT SuperCloud and Lincoln Laboratory Supercomputing Center for providing (HPC, database, consultation) resources that have contributed to the research results reported within this paper. Mengke Li acknowledges the student fellowship from Shanghai Jiao Tong University.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/prot.26603>.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Mengke Li  <https://orcid.org/0009-0006-3602-4129>

Yanze Wang  <https://orcid.org/0000-0002-0929-6544>

Fei Tao  <https://orcid.org/0000-0002-5997-8770>

Ping Xu  <https://orcid.org/0000-0002-4418-9680>

Shuguang Zhang  <https://orcid.org/0000-0002-3856-3752>

REFERENCES

- Kaplon H, Chenoweth A, Crescioli S, Reichert JM. Antibodies to watch in 2022. *MAbs*. 2022;14(1):2014296.
- Cassotta A, Mikol V, Bertrand T, et al. A single T cell epitope drives the neutralizing anti-drug antibody response to natalizumab in multiple sclerosis patients. *Nat Med*. 2019;25(9):1402-1407.
- Rajabally YA. Immunoglobulin and monoclonal antibody therapies in Guillain-Barré syndrome. *Neurotherapeutics*. 2022;19(3):885-896.
- Baert F, Noman M, Vermeire S, et al. Influence of immunogenicity on the long-term efficacy of infliximab in Crohn's disease. *N Engl J Med*. 2003;348(7):601-608.
- Salles G, Barrett M, Foà R, et al. Rituximab in B-cell hematologic malignancies: a review of 20 years of clinical experience. *Adv Ther*. 2017;34(10):2232-2273.
- Adams GP, Weiner LM. Monoclonal antibody therapy of cancer. *Nat Biotechnol*. 2005;23(9):1147-1157.
- Lee A, Shirley M. Ranibizumab: a review in retinopathy of prematurity. *Paediatr Drugs*. 2021;23(1):111-117.
- Lerch TF, Sharpe P, Mayclin SJ, et al. Infliximab crystal structures reveal insights into self-association. *MAbs*. 2017;9(5):874-883.
- Lundahl MLE, Fogli S, Colavita PE, Scanlan EM. Aggregation of protein therapeutics enhances their immunogenicity: causes and mitigation strategies. *RSC Chem Biol*. 2021;2(4):1004-1020.
- Vazquez-Rey M, Lang DA. Aggregates in monoclonal antibody manufacturing processes. *Biotechnol Bioeng*. 2011;108(7):1494-1508.
- Rosenberg AS. Effects of protein aggregates: an immunologic perspective. *AAPS J*. 2006;8(3):E501-E507.
- Svilenov HL, Winter G. Formulations that suppress aggregation during long-term storage of a bispecific antibody are characterized by high Refoldability and colloidal stability. *J Pharm Sci*. 2020;109(6):2048-2058.
- Svilenov HL, Kulakova A, Zalar M, Golovanov AP, Harris P, Winter G. Orthogonal techniques to study the effect of pH, sucrose, and arginine salts on monoclonal antibody physical stability and aggregation during long-term storage. *J Pharm Sci*. 2020;109(1):584-594.
- Svilenov H, Winter G. The RefOLD assay for protein formulation studies and prediction of protein aggregation during long-term storage. *Eur J Pharm Biopharm*. 2019;137:131-139.
- Courtois F, Agrawal NJ, Lauer TM, Trout BL. Rational design of therapeutic mAbs against aggregation through protein engineering and incorporation of glycosylation motifs applied to bevacizumab. *MAbs*. 2016;8(1):99-112.
- Kuroda D, Tsumoto K. Engineering stability, viscosity, and immunogenicity of antibodies by computational design. *J Pharm Sci*. 2020;109(5):1631-1651.
- Chennamsetty N, Voynov V, Kayser V, Helk B, Trout BL. Design of therapeutic proteins with enhanced stability. *Proc Natl Acad Sci U S A*. 2009;106(29):11937-11942.
- Clark RH, Latypov RF, De Imus C, et al. Remediating agitation-induced antibody aggregation by eradicating exposed hydrophobic motifs. *MAbs*. 2014;6(6):1540-1550.
- Courtois F, Schneider CP, Agrawal NJ, Trout BL. Rational design of biobetters with enhanced stability. *J Pharm Sci*. 2015;104(8):2433-2440.
- Reslan M, Sifniotis V, Cruz E, Sumer-Bayraktar Z, Cordwell SP, Kayser V. Enhancing the stability of adalimumab by engineering additional glycosylation motifs. *Int J Biol Macromol*. 2020;158:189-196.
- Zhang S, Tao F, Qing R, et al. QTY code enables design of detergent-free chemokine receptors that retain ligand-binding activities. *Proc Natl Acad Sci U S A*. 2018;115(37):E8652-E8659.
- Qing R, Han Q, Skuhersky M, et al. QTY code designed thermostable and water-soluble chimeric chemokine receptors with tunable ligand affinity. *Proc Natl Acad Sci U S A*. 2019;116(51):25668-25676.
- Hao S, Jin D, Zhang S, Qing R. QTY code-designed water-soluble fc-fusion cytokine receptors bind to their respective ligands. *QRB Discov*. 2020;1:e4.
- Tegler L, Corin K, Pick H, et al. The G protein coupled receptor CXCR4 designed by the QTY code becomes more hydrophilic and retains cell signaling activity. *Sci Rep*. 2020;10(1):21371.
- Qing R, Tao F, Chatterjee P, et al. Non-full-length water-soluble CXCR4(QTY) and CCR5(QTY) chemokine receptors: implication for overlooked truncated but functional membrane receptors. *iScience*. 2020;23(12):101670.
- Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589.
- Skuhersky MA, Tao F, Qing R, Smorodina E, Jin D, Zhang S. Comparing native crystal structures and AlphaFold2 predicted water-soluble G protein-coupled receptor QTY variants. *Life (Basel)*. 2021;11(12):1285.
- Tao F, Tang H, Zhang S, Li M, Xu P. Enabling QTY server for designing water-soluble alpha-helical transmembrane proteins. *mBio*. 2022;13(1):e0360421.
- Smorodina E, Tao F, Qing R, Jin D, Yang S, Zhang S. Comparing 2 crystal structures and 12 AlphaFold2-predicted human membrane glucose transporters and their water-soluble glutamine, threonine and tyrosine variants. *QRB Discovery*. 2022;3:e5.
- Smorodina E, Diankin I, Tao F, Qing R, Yang S, Zhang S. Structural informatic study of determined and AlphaFold2 predicted molecular structures of 13 human solute carrier transporters and their water-soluble QTY variants. *Sci Rep*. 2022;12(1):20103.
- Moro Pérez L, Rodríguez Taño AC, Martín Márquez LR, Gómez Pérez JA, Valle Garay A, Blanco Santana R. Conformational characterization of a novel anti-HER2 candidate antibody. *PLoS One*. 2019;14(5):e0215442.
- Qing R, Hao S, Smorodina E, Jin D, Zalevsky A, Zhang S. Protein design: from the aspect of water solubility and stability. *Chem Rev*. 2022;122(18):14085-14179.
- Creighton TE. *Proteins: Structures and Molecular Properties*. 2nd ed. W. H. Freeman; 1993.

34. Saphire EO, Parren PW, Pantophlet R, et al. Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. *Science*. 2001;293(5532):1155-1159.
35. Scapin G, Yang X, Prosis WW, et al. Structure of full-length human anti-PD1 therapeutic IgG4 antibody pembrolizumab. *Nat Struct Mol Biol*. 2015;22(12):953-958.
36. Blech M, Hörer S, Kuhn AB, et al. Structure of a therapeutic full-length anti-NPRA IgG4 antibody: dissecting conformational diversity. *Biophys J*. 2019;116(9):1637-1649.
37. Harris LJ, Skaletsky E, McPherson A. Crystallographic structure of an intact IgG1 monoclonal antibody. *J Mol Biol*. 1998;275(5):861-872.
38. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods*. 2022;19(6):679-682.
39. Kuriata A, Iglesias V, Pujols J, Kurcinski M, Kmiecik S, Ventura S. Aggrescan3D (A3D) 2.0: prediction and engineering of protein solubility. *Nucleic Acids Res*. 2019;47(W1):W300-W307.
40. Case HMA DA, Belfon K, Ben-Shalom IY, et al. AMBER 2022. University of California; 2022.
41. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput*. 2015; 11(8):3696-3713.
42. Valdés-Tresanco MS, Valdés-Tresanco ME, Valiente PA, Moreno E. gmx_MMPBSA: a new tool to perform end-state free energy calculations with GROMACS. *J Chem Theory Comput*. 2021;17(10):6281-6291.
43. Stanfield RL, Wilson IA. Antibody structure. *Microbiol Spectr*. 2014; 2(2):1-11.
44. Chennamsetty N, Helk B, Voynov V, Kayser V, Trout BL. Aggregation-prone motifs in human immunoglobulin G. *J Mol Biol*. 2009; 391(2):404-413.
45. Bansal R, Dash R, Rathore AS. Impact of mAb aggregation on its biological activity: rituximab as a Case study. *J Pharm Sci*. 2020;109(9): 2684-2698.
46. Reslan M, Ranganathan V, Macfarlane DR, Kayser V. Choline ionic liquid enhances the stability of Herceptin® (trastuzumab). *Chem Commun (Camb)*. 2018;54(75):10622-10625.
47. Zhou T, Xu L, Dey B, et al. Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature*. 2007;445(7129): 732-737.
48. Bhattacharyya S, Singh P, Rathore U, et al. Design of an *Escherichia coli* expressed HIV-1 gp120 fragment immunogen that binds to b12 and induces broad and potent neutralizing antibodies. *J Biol Chem*. 2013;288(14):9815-9825.
49. Martí D, Alemán C, Ainsley J, Ahumada O, Torras J. IgG1-b12-HIV-gp120 Interface in solution: a computational study. *J Chem Inf Model*. 2022;62(2):359-371.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Li M, Wang Y, Tao F, Xu P, Zhang S. QTY code designed antibodies for aggregation prevention: A structural bioinformatic and computational study. *Proteins*. 2024;92(2):206-218. doi:[10.1002/prot.26603](https://doi.org/10.1002/prot.26603)