Antibody Therapeutics, 2023, Vol. 6, No. 1 ***13–29***
https://doi.org/10.1093/abt/tbac029
Advance Access Publication on 11 November 2022

Review Article

Developability assessment at early-stage discovery to enable development of antibody-derived therapeutics

Weijie Zhang[1], Hao Wang[1], Nan Feng[1], Yifeng Li[2], Jijie Gu[1] and Zhuozhi Wang[1,*]

[1]Biologicals Innovation and Discovery, WuXi Biologicals, 1951 Huifeng West Road, Fengxian District, Shanghai 201400, China , and [2]Technology and Process Development, WuXi Biologicals, 288 Fute Zhong Road, Waigaoqiao Free Trade Zone, Shanghai 200131, China

Received: July 29, 2022; Revised: November 1, 2022; Accepted: November 2, 2022

ABSTRACT

Developability refers to the likelihood that an antibody candidate will become a manufacturable, safe and efficacious drug. Although the safety and efficacy of a drug candidate will be well considered by sponsors and regulatory agencies, developability in the narrow sense can be defined as the likelihood that an antibody candidate will go smoothly through the chemistry, manufacturing and control (CMC) process at a reasonable cost and within a reasonable timeline. Developability in this sense is the focus of this review. To lower the risk that an antibody candidate with poor developability will move to the CMC stage, the candidate's developability-related properties should be screened, assessed and optimized as early as possible. Assessment of developability at the early discovery stage should be performed in a rapid and high-throughput manner while consuming small amounts of testing materials. In addition to monoclonal antibodies, bispecific antibodies, multispecific antibodies and antibody-drug conjugates, as the derivatives of monoclonal antibodies, should also be assessed for developability. Moreover, we propose that the criterion of developability is relative: expected clinical indication, and the dosage and administration route of the antibody could affect this criterion. We also recommend a general screening process during the early discovery stage of antibody-derived therapeutics. With the advance of artificial intelligence-aided prediction of protein structures and features, computational tools can be used to predict, screen and optimize the developability of antibody candidates and greatly reduce the risk of moving a suboptimal candidate to the development stage.

Statement of Significance: This article reviews the assessment of developability at early-stage discovery of antibody-derived therapeutics, including monoclonal antibodies, bispecific antibodies and antibody-drug conjugates, and suggests general considerations and practices to mitigate the risk of moving a suboptimal candidate to the development stage. The article also comprehensively reviews the physicochemical properties of an antibody candidate that could affect its key attributes of homogeneity, stability, solubility and specificity. In addition, the article summarizes experimental and in silico methods for the assessment of developability at the discovery stage.

KEYWORDS: developability; antibody; bispecific; antibody-drug conjugate; discovery

INTRODUCTION

Antibody-related biological drugs, including monoclonal antibodies (mAbs), bispecific antibodies (bsAbs) and antibody-drug conjugations (ADCs), have become an important category of innovative drugs. At the discovery stage of research and development of such drugs, although the initial focus usually is on the unique mechanism of action and biological functions of drug candidates, developability of these candidates has also become critical. In consideration of the current competitive landscape, the

*To whom correspondence should be addressed. Biologics Innovation and Discovery, WuXi Biologicals, 1951 Huifeng West Road, Fengxian District, Shanghai 201400, China, Phone number: +86-21-50518899 Email: wang_zhuozhi@wuxibiologics.com

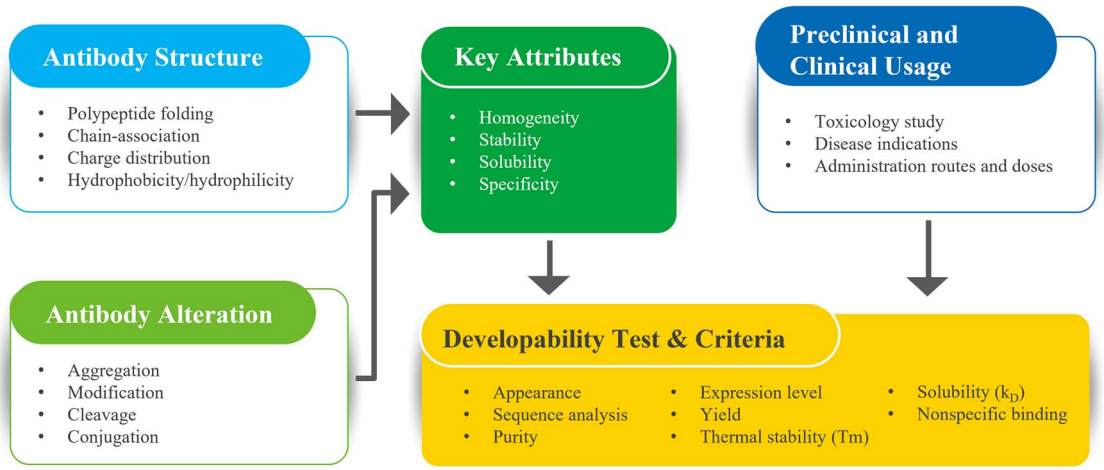

Figure 1. Key factors affecting developability assessment. Four key attributes of developability (homogeneity, stability, solubility and specificity) are determined by antibody structure and alteration. These attributes and the usage of antibody candidates are used to determine the developability criteria.

speed of advancing a drug candidate from discovery to preclinical development to clinical trial, and eventually to regulatory approval and commercialization, is critical for the success of a biological drug. Developability of mAbs at the discovery stage was recently reviewed [1, 2]. Developability refers to the attributes of a drug candidate that is readily manufactured, safe and efficacious [3]. This review focuses on the physicochemical properties of mAbs, bsAbs and ADCs that greatly contribute to their developability. The structure of an antibody and its alteration determine some key attributes, including homogeneity, stability, solubility and specificity. These attributes determine the observed developability-related features, such as appearance, expression level and yield, purity and others (Fig. 1). A simplified developability-related assessment is proposed for early-stage discovery, and the criteria of developability should be related to clinical usage of the drug candidates, including treated disease indication, dosage and route of administration of the drug candidates (Fig. 1). In general, a potent drug with low dosage and low concentration used in patients can have less strict developability criteria than a drug with high dosage and high concentration used in clinical settings.

GENERAL PROCESS OF ANTIBODY-DRUG DISCOVERY AND DEVELOPMENT

The process of antibody discovery and development includes several steps (Fig. 2). Once a target or a disease-related pathway is identified, a panel of antibodies can be generated using in vivo immunization, in vitro display technology, B-cell selection or a combination of techniques. Starting from hundreds to thousands of target-binding antibodies, several screening steps can be used to narrow down promising antibodies in a process referred to as a screening funnel. In this process, antibodies with appropriate affinity, specificity and functionality are identified. These antibodies can be further optimized or engineered by affinity maturation, humanization, Fc-engineering and other processes. The physicochemical properties of these antibodies can be altered in the process. If an antibody can be internalized upon binding to the targets on cells, the antibody may be conjugated with a payload to make ADCs. BsAbs candidates can be produced from original screening of appropriate pairs of mAbs or they can be engineered after a pair of mAbs is identified (Fig. 3). BsAbs candidates are screened and engineered by a similar process as mAbs, although complicated protein engineering and biological assays may be applied to make and characterize such bsAbs candidates.

Once an antibody is identified as a preclinical candidate, a cell line will be developed for the stable expression of the antibody, and a series of purification steps will be developed. During preclinical development, an antibody will likely be tested in monkeys or other relevant species for toxicologic studies. The International Conference on Harmonization (ICH) guidance (S6R1) suggests that 'dosage level should be selected to provide information on a dose–response relationship, including a toxic dose and a no observed adverse effect level (NOAEL)'. It is common and sometimes necessary to choose a high dose level in a toxicology study to guide clinical usage.

DEVELOPABILITY ANALYSES ARE DIFFERENT AT THE DISCOVERY AND DEVELOPMENT STAGES

Developability assessment at the discovery stage is different from developability study at the chemistry, manufacturing and control (CMC) stage, as shown in Fig. 4. At the discovery stage, it is important to identify a therapeutic molecule with target product profile, which is usually based on clinical usage of this potential therapeutic agent. Critical quality attributes (CQAs) of the lead molecule will be further assessed to setup quality target product profile at the CMC stage. Developability assessment at the discovery stage is an initial step towards establishing potential CQA. Assessment at the discovery stage is done to identify potential developability risk using rapid and high-throughput methods. At the early discovery stage, several antibody candidates with similar biological functions can be compared by

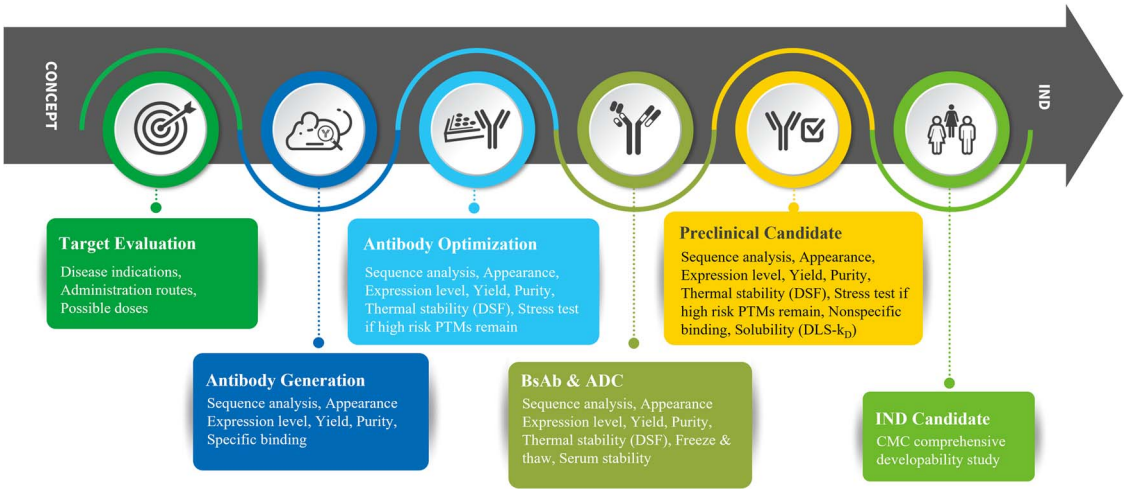

Figure 2. Considerations of developability-related measurements at the different stages from target evaluation to IND filing.

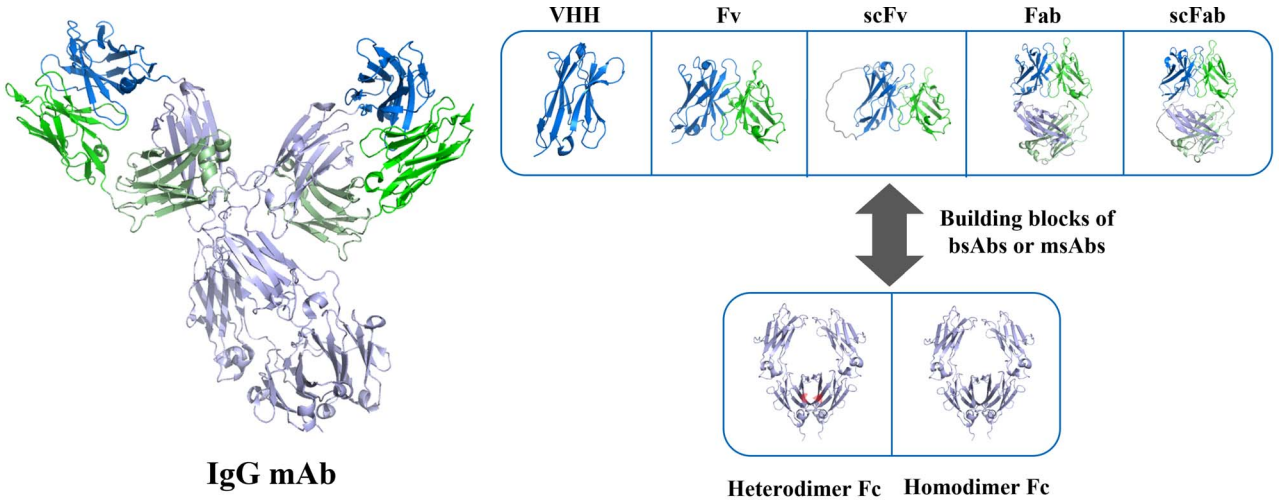

Figure 3. Structures of an IgG mAb and basic building blocks for bsAbs or msAbs. Structure of the IgG mAb (PDB code 5DK3) was created with PyMol. BsAbs or msAbs can be assembled from the five antigen-binding domains: VHH, Fv, scFv, Fab and scFab, and these five building blocks can also be assembled with heterodimeric or homodimeric Fc to form Fc-containing bsAbs or msAbs. The red mark in the Fc domain represents the mutations facilitating heterodimerization.

developability-related assays, and usually one or two with acceptable developability are selected to move to the CMC stage. Even if only one functionally optimal antibody is identified at the discovery stage, its developability still needs to be assessed and possibly optimized. The cost of engineering an antibody at the discovery stage is significantly lower than the cost of optimizing process and formulation at the development stage. In addition, poor physicochemical properties of an antibody may affect its biological functions in vitro and in vivo. For example, post-translational modifications (PTMs) in complementarity-determining regions (CDRs) of an antibody can significantly reduce its potency. It was reported that an asparagine (Asn) in CDR1 of an antibody light chain (LC) reached 92.6% deamidation after 40°C for 3 months, which correlated with loss of potency of this antibody [4]. The pharmacokinetics study of this antibody in the monkey also showed a loss of target binding over time. Asn deamidation and Asp isomerization located in CDRs correlated with reduced antibody potency in other reports [5].

If an antibody is found with a developability issue in the discovery stage, it can be fixed by protein engineering. For example, deletion of two hydrophobic residues in the CDR domains of a mAb significantly reduced its tendency to precipitation [6]. In another example, developability of bsAb against CD3 and GUCY2C was optimized by structure-guided mutagenesis and phage-displayed antibody panning and screening: removal of deamidation site in the CDR H2 of anti-GUCY2C and of proteolytic cleavage site in the CDR H2 of anti-CD3 reduced its polyreactivity and self-association potential. After multidimensional optimization campaigns, a well-behaved bsAb suitable for manufacturing was obtained [7].

In contrast, a developability study in the development stage is performed to obtain sufficient materials with reasonable process and formulation for preclinical and

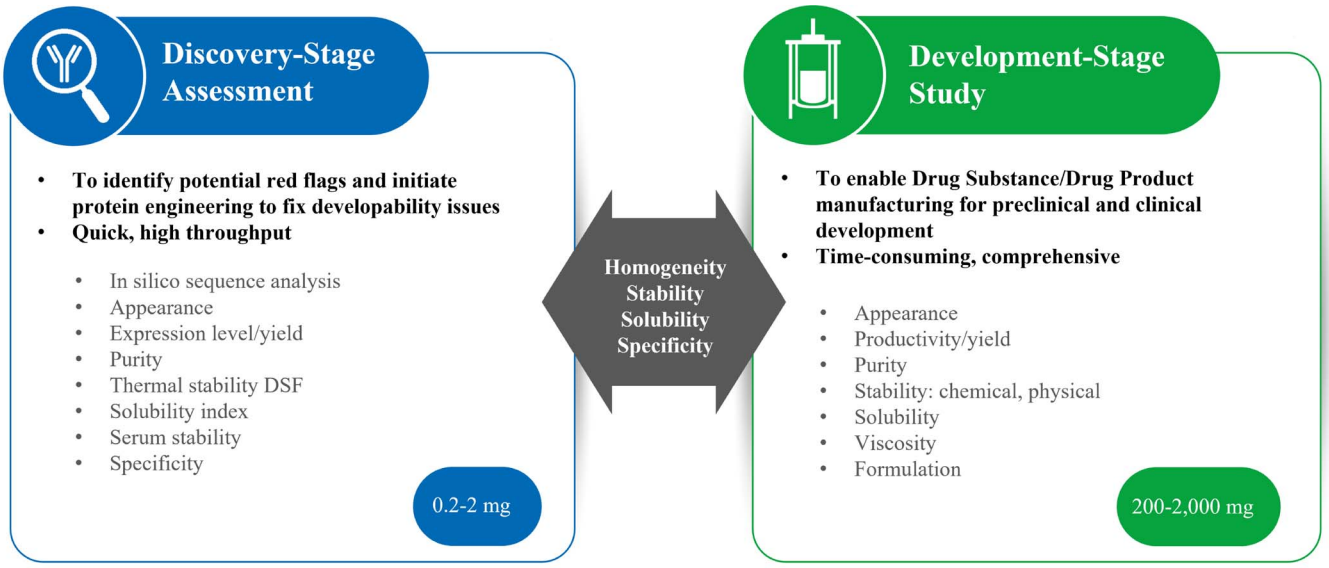

Figure 4. Different purposes and tests at the discovery and development stages.

clinical use. The study is usually comprehensive and time-consuming, and requires large amounts (hundreds of milligrams to a few grams) of testing materials that are usually generated from a stable cell line or stable pool. Solubility studies can be started by screening different buffers (e.g. acetate, citrate and phosphate buffers) at various pH values by polyethylene glycol precipitation to identify the one with the highest solubility. Solubility can be further assessed by screening buffers containing suitable excipients (e.g. sucrose and polysorbate 80). With each formulation, the antibody is dissolved to the highest possible concentration and stored at 2°C–8°C. Samples are taken at different time intervals, their appearance, concentration and size-exclusion chromatography (SEC) profiles are determined and the results are compared with those at T0. This part of the study can serve as a preformulation screening. Then developability study can then be performed using samples in the selected buffer with excipient at the target concentration under stress conditions. A panel of tests is typically used to evaluate a candidate's developability (summarized in Fig. 5), including general test methods and stability studies under stressed conditions (e.g. exposure to elevated temperature, low or high pH, light, oxidative reagents and freeze–thaw).

The developability of biological drug candidates greatly affects the speed, cost and success rate of these candidates from discovery to clinical development and commercialization. With the advance of modern biotechnology and accumulated experience, developability should be assessed at the early discovery stage to reduce the risk of development of such suboptimal drugs.

DEVELOPABILITY-RELATED CONSIDERATIONS AND ASSESSMENT AT THE DISCOVERY STAGE

Analysis of approved antibody drugs can suggest developability-related considerations at the discovery stage, beginning with the end in mind. Figure 6 summarizes the concentrations and administration routes of 129 antibody drugs approved by the agencies of the United States, Europe, Japan and China. Among 54 approved antibodies that are intravenously injected for cancer therapy, the median concentration is 11.25 mg/mL, which is much lower than the median concentration (100 mg/mL) of antibodies subcutaneously injected for autoimmune diseases. Antibodies injected intravitreally for ophthalmologic therapy are formulated in small volumes with high concentrations. For example, brolucizumab against vascular endothelial growth factor is formulated in a volume of 0.05 mL with concentration of 120 mg/mL.

Among several approved bsAbs [8, 9], three are T-cell engagers: blinatumomab against CD3 and CD19 formulated in lyophilized powder and injected at 0.009–0.028 mg/-day; tebentafusp against CD3 and gp100 stored in solution at a concentration of 0.2 mg/mL and dosed at 0.02–0.068 mg/day; and mosunetuzumab against CD3 and CD20 formulated in solution at a concentration of 1 mg/mL. Four other bsAbs are amivantamab, cadonilimab, emicizumab and faricimab, which are formulated in solution at a concentration of 50 mg/mL solution for IV injection, 12.5 mg/mL for IV injection, 150 mg/mL for subcutaneous (SC) injection and 120 mg/mL for intravitreal injection, respectively.

Twelve of 13 approved ADCs are formulated in lyophilized powder, likely due to instability of the linker and payload in the liquid state.

From analysis of these approved antibody drugs, we can conclude that the criteria of developability for different projects can be different, based on consideration of disease indications treated by the antibody drugs and possible dose and administration routes of the antibody drugs. Here we outline the developability-related considerations at different discovery stages and suggest some assays to assess developability-related attributes (Fig. 2). At the target evaluation stage, we suggest considering the possible clinical uses of the therapeutic antibodies. For example, therapeutic antibodies for ophthalmologic indications are

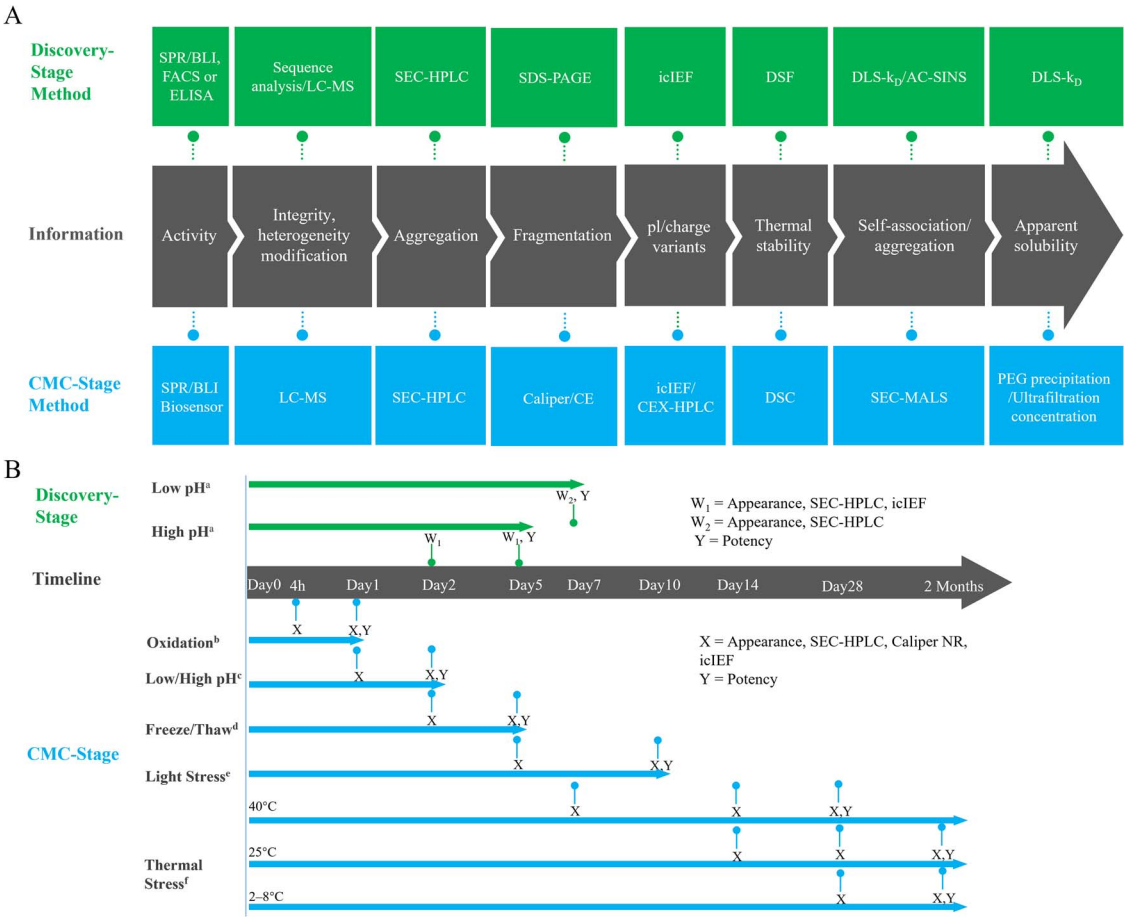

Figure 5. A panel of tests for developability assessment at the discovery and CMC stages. (A) Analytical methods generally used for developability assessment at the discovery and CMC stages. (B) Typical forced stress study plans at the discovery and CMC stages. [a.] In the discovery stage, low pH is for Asp isomerization susceptibility testing, whereas high pH is for deamidation susceptibility testing; [b.] 0.05% and 1% H_2O_2 treatment at 25°C; [c.] Low pH: pH 3.5 incubation; High pH: pH 9.0 incubation; [d.] 3 or 5 freezing (at −70°C) and thawing (at 25°C) cycles; [e.] An antibody is exposed to bright light of 5 000 lx for 4 h and 24 h; [f.] Thermal stress under 2°C–8°C, 25°C and 40°C.

usually intravitreally injected in small volumes, and therefore the antibody candidates need to have high solubility and low viscosity.

At the antibody generation stage, antibody candidates are usually screened by antigen-binding activity and biological function. Small amounts (usually <1 mg) of each antibody are expressed and purified in this stage. The appearance, expression level, yield, purity and binding to the homolog of the target should be examined. The appearance of the antibody should be observed during the entire process of manufacturing and using the antibody, including cell culture, harvest of cultural supernatant, purification, concentration, storage, freeze/thaw and dilution to the assay buffer. An antibody with visible particles or precipitates should be noticed and may need optimization. The expression level can be measured by analytic protein A chromatography or estimated from protein-A affinity-purification. Usually, 100 mg/L mAbs from Expi293 transient expression can be considered acceptable. Normally, greater than 90% purity is acceptable for most in vitro assays, although some assays require antibodies with higher purity to avoid the effect of impurities. Findings of particles, aggregates, low expression, low purity or binding to homolog can be warning signs of a developability problem. When sequences of the antibodies are available, in silico analysis should be performed to mark PTM hotspots, unusual residues at particular positions, deletion or addition of sequences, etc.

mAbs are usually further engineered to meet the requirements of the project, such as humanization, affinity maturation, Fc-engineering, etc. In this stage, in addition to monitoring for appearance, expression level and purity, PTM hotspots with high risk should be removed and thermal stability should be performed. High-risk motifs, including NG/NS/DG motifs, extra Cys residue and N-glycosylation motif on variable regions, should be removed to mitigate potential risks to on efficacy and safety. Once a few antibody candidates are identified for preclinical studies, we recommend testing solubility-related features and extensive nonspecific binding.

When a mAb is conjugated with payload via a linker to form ADCs, the payload, linker and conjugation may alter the conformation of the mAb. Hydrophobicity of payloads and uneven charge distribution may trigger the aggregation of ADCs. Therefore, upon conjugation, the purity, yield and drug-antibody ratio (DAR) of ADCs need

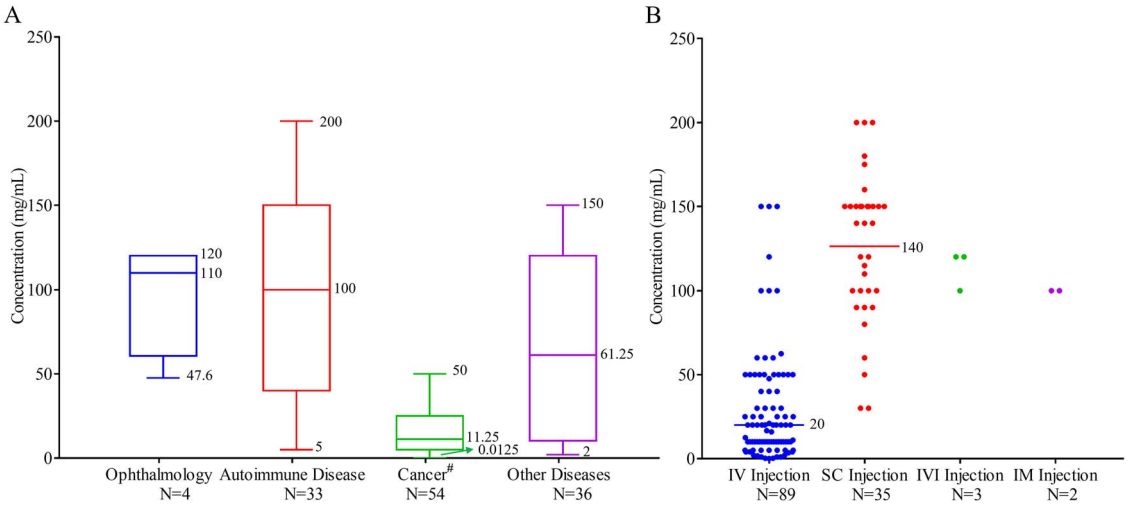

Figure 6. Concentration of the approved antibodies based on disease indications (A) and administration routes (B): 129 antibodies approved by the agencies of the United States, Europe, China and Japan are included in the data analysis. Biosimilar or withdrawn products are excluded. (A) Concentration of approved antibodies grouped based on disease indications. In the boxplots, the concentration for each indication with median (the centre line) and interquartile (25% and 75%) ranges are shown. # Two outliers found in the upper and lower whiskers (outside 1.5 interquartile ranges) are excluded: anti-PD-L1 antibodies Envafolimab (200 mg/mL) and Atezolizumab (60 mg/mL). (B) Concentration of approved antibodies based on administration routes. The concentration for IV and SC injection with median (the centre line) are shown. IV, intravenous; SC, subcutaneous; IVI, intravitreal; IM, intramuscular.

to be measured, with special attention to the percentage of high-molecular-weight fraction. The stability of ADCs after several cycles of freezing and thawing should also be measured. The serum stability of ADCs may be examined to measure the percentage of payload cleaved from the ADCs.

A bsAb is usually assembled from a pair of parental antibodies against different targets. Since the physicochemical properties of the variable regions of the parental antibodies will be carried over to the bsAb, it is highly recommended to assess and if necessary, optimize, parental antibodies prior to construction and optimization of the bsAb. BsAbs should also be examined for appearance, expression, purity, thermal stability, etc. In silico analysis should also be performed to avoid added PTM hotspots during the engineering of bsAbs.

For antibody candidates with potential developability risk, several strategies can be applied. Antibody molecules with slightly visible particles or precipitates can be tested in different formulation buffers. For antibodies with severe developability risk, such as molecules with continuous aggregation, poor stability, unexpected low solubility or significant nonspecific binding, it is necessary to engineer antibody sequences. After sequence optimization, all the developability properties should be reassessed, and if a red flag is found, further optimization will be needed. Finally, if PTM hotspots cannot be removed without altering antigen binding, a forced stress study should be performed to evaluate the potential risk (Fig. 5B).

INTRISIC PROPERTIES OF ANTIBODIES DETERMINING THEIR DEVELOPABILITY

For the analysis and optimization of developability-related attributes, such as the homogeneity, stability and solubility of antibody-derived candidates, it is necessary to understand the intrinsic properties of an antibody that affect its developability, including its multipolypeptide folding, charge distribution and hydrophobicity (Fig. 1).

Antibody aggregation, a common developability problem, can be triggered by partial unfolding of antibody domains, leading to monomer–monomer association followed by nucleation and growth [10]. Antibody aggregation/colloidal stability can also be affected by surface charge distribution of CDR loops. When asymmetric or varied surface charge distribution occurs, the localized surface charge patches may generate attractive protein–protein interactions via electrostatic interaction, eventually causing colloidal instability [11]. The electrostatic surface charge distribution of proteins plays a dominant role in governing intermolecular interactions and subsequently self-association and viscosity behaviour, especially at high antibody concentrations [11].

Hydrophobicity and hydrophilicity are determined by the side chains of 20 amino acids and their post-translationally or cotranslationally modified derivatives [12]. The strength of hydrophobic surface patches can be modulated by neighbouring charged residues [13]. Unfolding or misfolding of hydrophobic patches in CDRs can confer a high degree of hydrophobicity on antibodies and potentially lead to intermolecular attraction, low solubility, high viscosity and even aggregation [1].

BsAbs or multispecific antibodies (msAbs) are usually constructed by assembling several building blocks (Fig. 3), such as VHH (single-domain antibody), variable region (Fv), single-chain variable fragment (scFv), Fab and single-chain Fab (scFab), against different antigens or epitopes [14]. These antibody fragments may have lower folding stability than a full-length IgG. In addition, different domains may have polar charges, and multiple domains on one

molecule may increase its tendencies to aggregation or instability. The issues affecting chemical stability, such as deamination and isomerization, are similar to those of mAbs. However, the cumulative effects of the chemical instability of bsAbs can be more severe than those of mAbs.

HOMOGENEITY

Because antibodies are made from living cells, antibody-derived drugs cannot be as homogeneous as chemically synthesized small-molecule drugs. The ICH guideline (Q6B) suggests that manufacturing process-related impurities and some product-related impurities should be reduced to an acceptable level. Manufacturing process-related impurities include cell substrates (e.g. host cell proteins, host cell DNA), cell culture components (e.g. inducers, antibiotics or media components) and downstream processing substrates. Product-related impurities (e.g. precursors, certain degradation products) are molecular variants arising during manufacture and/or storage that do not have properties comparable to those of the desired product regarding activity, efficacy and safety. At the discovery stage, product-related impurities are the main concern in the assessment and optimization of developability.

Unpaired, missed or mismatched chains

Unpaired or missing LC is a common heterogeneity found in mAb and bsAb preparations. Generally, the size, charge or hydrophobicity/hydrophilicity of an antibody with missing light LCs are different from those of an intact antibody; the difference can be used to purify intact antibody products. Because two kinds of HCs and two kinds of LCs may be present in single cells, an IgG-like bsAb can compose a mismatched HC–LC or HC–HC. It is a particular challenge to prevent mispairing in a bsAb product, so that different bsAb formats, such as knobs-into-holes, dual-variable domains Ig, WuXiBody, $\kappa\lambda$-bodies, etc. have been developed in an attempt to obtain homogeneous bsAb products [15]. When asymmetric IgG-like bsAbs are expressed and assembled, there is a possibility of forming a homodimer instead of a heterodimer. The percentage of correctly assembled bsAb and mispaired by-products needs to be evaluated. Other undesired bsAb-specific by-products, such as aggregates and fragments, have been reported [16]. General downstream purification processes, such as affinity-, charge-, size-, hydrophobicity- and mixed mode-based purification, have been employed to remove undesired by-products [17].

Cleavage

Cleavage of antibodies may occur at the N-terminus, hinge region, constant region or C-terminus. Specific cleavage sites recognized by signal peptidase need to fulfil the A-X-B rule, where residues at positions A and B are usually small neutral amino acids, such as Ala, Gly and Ser [18]. Signal peptide cleavage can also be influenced by the composition of the N-terminal amino acids of the mature protein [18], which may lead to nonspecific cleavage of signal peptides, generating either elongation [19] or truncation [20] of the N-terminus of the heavy and LCs. Although the remaining signal peptide in antibodies has no effect on antigen-binding [21], truncated antibody variants may affect antibody efficacy and bioactivity [20]. Accurate and specific cleavage of signal peptides is critical for manufacturing intact and homogeneous products. Engineering signal peptide sequences or the N-terminal amino acids of mature proteins can be used for accurate cleavage of signal peptides [18].

The hinge region is highly flexible and exposed to solvent compared with other parts of the antibody. Peptide bond cleavage, either enzymatic or chemical cleavage, mostly on the hinge region and the CH2–CH3 interface, results in fragments of different sizes [22]. Cleavage susceptibility is linked to the length and flexibility of the hinge region. IgG1 antibodies were found to be more susceptible to cleavage processes than IgG2 and IgG4 antibodies [23]. Hinge cleavage may have negative effects on bivalent binding, Fc effector function and neonatal Fc receptor (FcRn) function, leading to decreased efficacy and faster clearance [22]. Therefore, the effect of hinge cleavage on the potency of an antibody should be evaluated according to the mechanism of action of the antibody. Effective removal of antibody fragments is necessary in CMC to ensure the stability and efficacy of the products.

The cleavage sites in the constant region are usually localized at the loop or domains–domain interfaces [22, 24]. Elevated temperatures or acidic or basic pH can accelerate the cleavage rate [24]. Cleavage of the constant region may have an impact on Fc-mediated effector function or FcRn-related pharmacokinetic (PK) profile [22].

IgG HCs terminate with Lys, which is susceptible to clipping by carboxypeptidases during cellular processing. The variation in C-terminal Lys processing generates antibodies with 0, 1 and 2 Lys residues, leading to different charge distributions owing to the positive charge of Lys [25, 26]. Given that C-terminal Lys is distant from the functional parts of the antibody, C-terminal heterogeneity does not affect the structure, thermal stability, antigen binding, potency or pharmacokinetics of IgG mAbs [27]. However, it has been reported that the IgG without C-terminal lysine has a lower titre in stable cell line culture [28]. Maximal complement activation may need removal of C-terminal Lys in some cases [29]. In some bsAb formats, an antigen-binding domain is fused to the C-terminus of Fc, where it is important to remove the C-terminal Lys to avoid a heterogeneous product with various sizes and functions.

Modification

Modifications can occur at any stage of the antibody–drug manufacturing process and even after administration. Commonly investigated modifications include Cys-related modifications, deamination, isomerization, oxidation and glycosylation. These modifications may cause heterogeneity in the size or charge of the antibody and may also affect the stability, safety and efficacy of the antibody [30, 31].

Cysteine-related modifications

Cys residuals are typically involved in the formation of intra- or interchain disulphide bonds to support the

structural integrity of the antibody. Each IgG contains 12 intrachain disulphide bonds, two interchain disulphide bonds linking the LC and the HC and various numbers of interchain disulphide bonds in the hinge region of the HC: two for IgG1 and IgG4 and four for IgG2 [32]. Alkaline environments can promote disulphide bond scrambling and may result in the reconfiguration of interchain disulphide bonds in IgG antibodies, particularly in the IgG2 and IgG4 subclasses [33]. Formation of intrachain disulphides and subsequent dissociation of the half-antibody for the IgG4 subclass due to disulphide bond scrambling has been reported [34, 35].

Although most paired Cys residues can form disulphide bonds, a low percentage of free sulfhydryl is commonly present in the constant domains on mAbs [36] and bsAbs [37]. Under basic conditions, due to beta-elimination of the disulphide bond between the HC and the LC [38], peptide bond hydrolysis may lead to antibody fragmentation in the hinge region and the formation of a nonreducible thioether linkage [39].

Incomplete disulphide bond formation has been reported in many IgG molecules and has been shown to reduce antigen-binding affinity and potency [40]. Unformed disulphide bonds, especially due to additional free cysteine, can cause antibody heterogeneity, low thermal stability [41], aggregation, low solubility [42] and low potency [40].

It has been documented that antibodies with an extra cysteine, especially in CDRs, can cysteinylate [43]. Cysteinylation may disturb the structural stability of the antibodies, increase aggregation and eventually adversely affect the biofunction of the molecules [44]. It is recommended to identify extra cysteine by in silico analysis and mutate the cysteine to other residues to prevent the risk of heterogeneity and loss of potency or stability.

Trisulphide bonds, formed by insertion of a sulphur atom into a disulphide bond, can occur in all isotypes of IgGs. The trisulphide modification occurs mainly in HC–LC and HC–HC interchain linkages [45]. Antibodies with trisulphide bonds increase their molecular weight by 32 Da and generate relatively acidic variants [46]. The level of trisulphides can be affected by cell culture conditions and hydrogen sulphide concentration. Previous studies showed that trisulphide bonds did not affect the antigen-binding affinity [47] or thermal stability [46] of antibodies.

N-terminal modifications

Cyclisation of Gln/Glu is a major N-terminal modification that has no significant effect on the structure, stability and function of an antibody [25, 31].

Deamination of Asn

Asn deamination occurs widely in either CDRs or Fc regions [48]. Each Asn deamination increases the mass of the antibody by 1 Da and adds a negative charge to the antibody [30]. Asn deamination in Fc regions has no effect on antigen binding [31] and is not a concern at the discovery stage. However, Asn deamination in CDRs may decrease antigen-binding affinity [5, 49, 50], resulting in loss of potency and efficacy. In addition, Asn deamination may lead to immunogenicity [51, 52]. The deamination rate depends on the pH, the size and flexibility of the residue following Asn, temperature and the tertiary structure of the antibody [26]. The deamination rates at pH 8.0 are almost 40-fold faster than at pH 5.5 [53]. Common motifs for deamination are Asn-Gly (NG), Asn-Ser (NS), Asn-Asn (NN), Asn-Thr (NT) and Asn-His (NH) [54]. Previous studies showed that NG and NS are more susceptible to deamination than other motifs [48, 54]. Therefore, NG/NS motifs on CDRs should be identified and possibly removed by protein engineering. In addition, Asn in flexible structures may be prone to deamination even if the Asn is not in the NG/NS motif. Heterogeneity of charge caused by deamination may also affect the stability and viscosity of an antibody.

Asp isomerization

Asp isomerization is another common modification of antibodies. Asp isomerization introduces a methyl group to the peptide backbone [26], which may change the structure of the antibody and, when isomerization occurs in CDRs, impair its antigen-binding affinity [30]. In addition, Asp isomerization may alter susceptibility to proteolysis and potentially trigger immunogenicity. The isomerization rate is closely related to the size and flexibility of the residue following Asp, pH and temperature [55, 56]. Previously identified hotspots include Asp-Gly (DG), Asp-Ser (DS), Asp-Asp (DD), Asp-Thr (DT) and Asp-His (DH) [54]. Among these canonical motifs, DG and DS are more liable to isomerization than are other motifs [54]. The isomerization rate may accelerate as the pH decreases [57]. In the discovery stage, it is necessary to identify DG and DS motifs in the CDRs of antibody candidates and attempt to remove the motif without altering antigen binding.

Glycosylation

Oligosaccharides can be introduced to Asn to generate N-linked glycoforms or to the hydroxyl groups of Ser, Thr or Tyr to generate O-linked glycoforms [58]. Glycosylation can be heterogeneous, altering the size, functionality and half-life of antibodies [59]. In the discovery stage, attention is usually paid to N-linked glycosylation motifs, whereas there is no identified motif for O-linked glycosylation. Human IgGs have a conserved N-glycosylation site at Asn 297 of the Fc region [60], which is involved in effector functions, such as antibody-dependent cellular cytotoxicity, complement-dependent cytotoxicity (CDC) and antibody-dependent cellular phagocytosis and may affect the half-life of an antibody by binding to FcRn [61, 62].

Approximately 20% of human IgGs contain the N-glycosylation consensus NXS/T motif (X refers to any amino acid residue except proline) in the variable region [63], which may affect the antibody's binding affinity to antigens. Each glycosylation increases the molecular weight by 1.5–4 kDa [64] and results in size heterogeneity. Some glycans, such as sialic acid, may cause immunogenicity [65, 66]. Among approved antibody drugs, only cetuximab has an N-glycosylation site in the variable region [67]. At the

antibody discovery stage, it is a general practice to remove the N-glycosylation site in the variable region.

Oxidation

Oxidation mainly occurs on Met and Trp residues when they are exposed to light, incubated with oxidizing reagents or stored for a long time [26]. Oxidation of two conserved Met residues (M252 and M428) in the Fc regions has no effect on antigen binding but may decrease thermal stability [68], induce aggregation [69], decrease CDC activity [70], decrease binding to FcRn [71] and shorten in vivo half-life [72]. In contrast, Trp residues are commonly present in antibody CDRs and are usually exposed to solvent [73]. Oxidation of Trp in CDRs may lead to reduced antigen binding and potency of the antibody [74, 75]. The susceptibility of Trp to oxidation in CDRs should be evaluated, and appropriate formulation needs to be developed.

Aggregation

Aggregation is one of the CQAs for the production of therapeutic antibodies. The major mechanisms of protein aggregation include association of native monomers, aggregation of conformationally altered monomers, aggregation of chemically modified monomers, nucleation-controlled aggregation and surface-induced aggregation [76].

Aggregation of antibodies may result in increased immunogenicity potential and decreased biological function of the antibody candidates. The aggregation propensity of an antibody is mainly associated with folding stability and the hydrophobic properties of the surface [77]. Poor folding stability is associated with low thermal stability, which may increase the probability of exposure of the hydrophobic residue, leading to aggregation. At the same time, the hydrophobic patches on the antibody surface may lead to antibody self-interaction. In particular, an antibody aggregates mainly due to intermolecular interactions of hydrophobic regions in the CDRs or framework regions resulting from partial or transient unfolding of the proteins [26]. As mentioned above, oxidation of Trp in CDRs and modification of extra Cys may also lead to an increased propensity for aggregation [78]. Polar charged patches on mAbs or bsAbs are related to aggregation due to inter-molecule attraction.

Some extrinsic factors may induce or reduce antibody aggregation. Aggregation of antibody candidates can be accelerated at low pH values, at ultrafiltration/concentration steps, by physical stress, after freeze–thaw cycles, by interaction with containers or after long-term storage [79].

Excipients can act as general protein stabilizers. In the presence of sucrose, the protein state with the least surface area will be thermodynamically favoured; Sucrose can stabilize the proteins by driving them towards a compact native state [80]. Salts and buffers can interact with proteins through three mechanisms: changing the enthalpy of ionization of various side chains, a cosolvent exclusion mechanism and a Debye screening of charge fluctuations [80]. The stabilizing salts, through the preferential exclusion mechanism, can minimize the amount of protein surface area exposed to solvent by 'salting-out' the protein [81]. Specific interactions between some buffer molecules (phosphate and citrate buffers) and the Fc domain of IgG are involved in the aggregation propensity of heat-denatured IgG [82].

Aggregation can be a major issue for symmetric bsAbs with elongated chains. Chain elongation increases the flexibility and uneven charge distribution of bsAbs and enhances intra- and intermolecular interactions, resulting in aggregation.

STABILITY

Stability is a critical factor for the discovery and development of therapeutic antibody candidates, including thermal stability, light-sensitive stability, pH-sensitive stability, physical stress-related stability and freeze and thaw stability, etc.

Protein thermal stability is positively correlated with the number of hydrogen bonds and the polar surface area fraction [83]. High thermal stability indicates a low propensity for unfolding and aggregation. Adding hydrogen bonds and disulphide bonds is commonly used to improve antibody stability in the discovery stage.

Antibodies are inevitably subjected to light exposure during the manufacturing process of biologics. Light exposure can induce oxidation of multiple amino acid residues, such as Trp, Tyr, Phe, Cys and Met [84]. Installation of safe lights in manufacturing and storage areas can provide a balance between the safety of human operators and the conservation of product quality [85]. A high-concentration liquid formulation of human IgG1 mAb showed colour changes of the solution, oxidation, aggregation, fragmentation and a loss in bioactivity [86].

As mentioned above, antibodies are exposed to low pH conditions during protein A elution and viral inactivation. Low pH can induce antibody cleavage and modification and increase the propensity to aggregate [87]. Soluble aggregates of an IgG4 mAb were observed during low-pH viral inactivation due to poor mixing and exposure to low pH [88].

During manufacture, transport and final administration, antibodies are subjected to physical stress, such as shear, agitation and stirring, potentially resulting in physical and chemical instability [89]. Exposure to air–liquid interfaces may cause antibodies to unfold [90]. Foaming resulting from protein unfolding can affect the properties of air–liquid interfaces and protein aggregation [91]. pH and ionic strength are two primary factors for formulation screening to control unfolding and aggregation of antibody products [92]. Various surfactants have been used to minimize the agitation- and surface-induced aggregation [93]. The selection of surfactant requires a balance between prevention of aggregation and perturbation of structure [94].

Freezing and thawing can put antibodies under stress, for example, by generation of the ice–liquid interface, phase transition, pH changes and changes in the distribution and concentration of solutes [95]. When an antibody solution is frozen, the water molecules are removed from the protein surface, leading to the disruption of the hydration shell around the protein surface, and the damage of

three-dimensional structure of antibodies. Meanwhile, the concentration of solutes (salt, excipient, antibodies, etc.) may increase several folds due to liquid–liquid phase separation and partitioning of unfrozen solutes, generating crystallization or precipitation, or pH change and finally impair the protein stability [95]. Some proteins are prone to unfolding as a consequence of weakening electrostatic and hydrophobic interaction at cold temperature [96]. In addition, as ice forms, the protein can directly interact with the ice surface, which disturbs the native structure of the antibody and causes protein denaturation [97]. The composition of the formulation buffer, the freezing and thawing rate, the number of freeze–thaw cycles and the antibody concentration can affect the stability of antibodies during the freezing and thawing process. For example, surfactants such as polysorbate 80 can serve as cryoprotectants to stabilize antibodies by preventing or reducing unfolding of the protein at the ice–liquid interface [98].

SOLUBILITY AND VISCOSITY

Solubility is inversely associated with self-interaction of antibodies. The self-interaction is caused by hydrophobic effects, van der Waals interactions, hydration forces and electrostatic interactions [99]. These interactions are determined both by the three-dimensional structure of the antibody and by the properties of the solution [6]. Previous studies indicated that Fab–Fab interactions can speed up the irreversible aggregation [99]. The impact of asymmetric charge patches on solubility may be more significant for bsAb due to their complicated structures. In addition, the properties of the solution, such as buffer composition, pH, temperature, salt concentration and excipients, have different effects on protein–protein interactions and finally on solubility [100]. Low solubility usually leads to high viscosity, off-target binding and rapid clearance [100].

Solubility is an important parameter for the assessment of developability. Aggregation-prone regions, asymmetric charged patches, conformational changes and chemical modification of antibodies can significantly influence the solubility of antibodies. Solubility can also be related to buffer formulation, viscosity of the antibody [11] and phase separation [100] and can affect the in vivo clearance rate of the antibody [100].

Because of the limited volume (usually <2 mL) used for SC injection, high-concentration products are required. In addition to good solubility properties, low viscosity is required for drug production and patient use [101].

SPECIFICITY

The antigen-binding sites of an antibody comprise three (for VHH) or six (for Fab) CDR loops, which can bind to a specific antigen epitope through noncovalent interactions, mainly by hydrogen bonds, hydrophobic interactions, van der Waals forces and electrostatic interactions [102]. These multiple noncovalent interactions can provide relatively strong binding between antibody and antigen. Occasionally, an antibody may bind to an undesired target. Recent studies showed that antibodies derived from phage-displayed libraries bound to the plastic nonspecifically via aromatic stacking [103]. In addition, antibodies can react with different targets if the targets share a similar conformational epitope [104].

A well-characterized therapeutic antibody should be exquisitely specific to the single target. However, not all antibodies are truly specific. Nonspecific binding can be due to imbalanced positive charge distribution or excess hydrophobicity in variable regions of the antibodies [105, 106]. Imbalanced positive charge on an antibody may result in its interactions with negatively charged polymers such as extracellular matrix or FcRn [105]. Excess hydrophobicity may cause nonspecific membrane interactions [105]. Nonspecific binding of an antibody may lead to poor PK [107] and decreased bioactivity. For example, bococizumab, an anti-PCSK9 antibody with excess positive charge, showed poor PK, poor biodistribution and high immunogenicity after SC treatment in patients, and finally failed in late-stage clinical development [108, 109]. In addition, a high level of structural redundancy related to the targets may also cause nonspecific binding [110]. Nonspecific binding can cause serious adverse effects during preclinical and clinical studies. For example, camrelizumab (an anti-PD1 antibody) treatment can cause capillary haemangioma in patients [111], a specific toxicity compared with other approved anti-PD1 antibodies. Recent studies showed that camrelizumab can mediate nonspecific binding to human receptors VEGFR2, FZD5 and ULBP2, which can stimulate vascular neogenesis and eventually lead to haemangioma toxicity [110]. In the discovery stage, nonspecific binding of antibody candidates needs to be evaluated, especially for the antibodies with potent cell-killing activity, such as ADC and bispecific T-cell engagers [110].

EXPERIMENTAL ASSESSMENT AND IMPROVEMENT OF DEVELOPABILITY AT THE DISCOVERY STAGE

In general, developability assessment at the discovery stage identifies the potential risks of the selected candidate, such as high propensity for aggregation or fragmentation, poor solubility, low stability, etc. Developability assessment at the discovery stage is usually rapid and high throughput, consuming small amounts (micrograms to a few milligrams) of testing materials. Developability assessment should include evaluation of expression level, yield, purity, homogeneity, stability, solubility and specificity, etc.

Expression level and yield

Low expression level is usually a sign of poor folding and assembly and low stability of an antibody. When the expression level of an antibody is normal, a low yield of the antibody may be due to removal of side products during multiple steps of purification or to aggregation and precipitation during the process. Generally, low expression levels and yields indicate poor developability of the antibody. In addition, low expression levels and yields increase the

overall cost and time required to obtain enough material to characterize the antibody in vitro and in vivo. Protein modelling and engineering tools can be used to improve the expression level and yield of an antibody [112].

Purity and homogeneity

During purification and storage of an antibody, aggregation and degradation may occur and generate size or charge heterogeneity. Size heterogeneity can be determined by size-exclusion chromatography high-performance liquid chromatography (SEC-HPLC) and reducing/nonreducing SDS-PAGE or capillary electrophoresis sodium dodecyl sulphate (CE-SDS), whereas charge heterogeneity can be detected by isoelectric focusing (IEF), capillary isoelectric focusing, imaged capillary isoelectric focusing (icIEF) and ion exchange chromatography (IEX) and sometimes can also be observed on hydrophobic interaction chromatography-high-performance liquid chromatography (HIC-HPLC) and reversed phase-high-performance liquid chromatography (RP-HPLC) [2, 25, 30].

Thermal stability

Classically, differential scanning calorimetry (DSC) and differential scanning fluorimetry (DSF) are used to evaluate the thermal stability of antibodies. The thermal transition midpoint (T_m) obtained from DSC is based on the basis of the heat capacity as a function of temperature [63], whereas DSF determines T_m by recording the change in fluorescence intensity of hydrophobic fluorescent dye added to the solution. The melting profiles measured by DSF and DSC are generally consistent [87]. However, DSF is more suitable than DSC in the discovery stage due to high throughput of DSF (with a 96-well or 384-well plate) with minimal protein consumption. In addition, DSF can also be used for preformulation screening in the discovery stage. Nano-DSF is another option for measurement of T_m with moderate throughput. Unlike DSF, which utilizes the interaction between the antibodies and extrinsic fluorescent dye, nano-DSF obtains T_m by monitoring changes in the fluorescence intensity of intrinsic Trp or Tyr at the emission wavelengths of 330 and 350 nm [2].

Alternatively, aggregation onset temperature (T_{agg}) measured by dynamic light scattering (DLS) can also indicate thermal stability. With elevated temperatures, proteins will change conformation, unfold and eventually aggregate. Unfolding and aggregation of the antibody lead to increase in radius that can be measured by DLS.

A previous report showed that most marketed or preclinical antibodies have a melting temperature of Fab greater than 59°C (ranging from 59°C to 90°C) [113] and a T_{agg} greater than 55°C [63]. mAbs with T_{m1} <55°C may be at risk for thermal stability, and T_{onset} should also be considered.

Chemical stability

The major factor affecting chemical stability is PTM. Although PTMs hotspots with high risk should be engineered, sometimes these motifs are critical for antigen binding and cannot be removed. In this situation, it is necessary to perform stress testing to evaluate the risk of PTM hotspots (deamination, isomerization or oxidation) [63]. Antibody candidates incubated under stress conditions may degrade or generate size or charge variants. Size variants can be detected by SEC-HPLC, and charge variants can be detected by icIEF, IEX-HPLC or HIC-HPLC to evaluate the chemical stability of antibody candidates. In addition, antigen-binding or functional assays can be performed to evaluate the antibody variants. Different antibody candidates can be ranked based on potency and stability.

Solubility

Therapeutic antibodies, especially those administered by SC, intravitreal or intramuscular injection, are required to have high solubility. Concentration by ultrafiltration is a conventional method of measuring solubility [114]. This method requires hundreds of milligrams of antibodies and has low throughput, which is not practical for screening and evaluating dozens to hundreds of antibody candidates at the discovery stage. Addition of polyethylene glycol to antibody solutions causes precipitation, even at low concentrations, and thus can be utilized to determine the solubility index of antibodies in a relatively high-throughput manner with 1 mg each of protein samples [114, 115]. However, this approach evaluates solubility through extrapolation and may not measure true solubility [116]. Notably, lower solubility is related to strong self-association of antibodies, which is mainly determined by hydrophobic and charged moieties on the surface of antibodies [99]. Several high-throughput methods have been reported to evaluate self-interaction of antibody candidates, such as DLS [117], affinity-capture self-interaction nanoparticle spectroscopy (AC-SINS) [118] and bio-layer interferometry (BLI) [119]. AC-SINS and BLI have been reported recently for characterizing self-interaction of unpurified antibodies and can be used for candidate selection at the early discovery stage. They both are able to distinguish between poorly and highly soluble antibodies. DLS can be used to calculate the diffusion interaction parameter (k_D): positive and negative k_D values indicate repulsive and attractive forces, respectively [120]. This method measures the diffusion coefficient within a protein concentration range of 1–20 mg/mL, which consumes ~1–2 mg of protein. In addition, k_D was found to be strongly correlated with viscosity [121].

Specificity

Nonspecific and off-target binding can lead to rapid clearance in vivo, unexpected toxicity or risk of immunogenicity [122]. Several methods have been reported to assess antibody nonspecific interactions, such as BLI [123], surface plasmon resonance [124], cross-interaction chromatography [125], enzyme-linked immunosorbent assay (ELISA) and fluorescence-activated cell sorting (FACS) [126].

A common, classical method is ELISA: antibodies are incubated with immobilized biomolecules such as proteins, lipopolysaccharides and DNA to evaluate the antibodies' nonspecific interactions [127]. Because the types of immobilized proteins used in this assay are limited, baculovirus

particles are used to present diverse types of lipids and proteins in nonspecific binding ELISA. Binding of baculovirus particles was reported to correlate well with faster serum clearance [128].

Protein biochips, in which a diverse set of proteins is printed on a chip, can also be used for detection of nonspecific binding. A chip manufactured by Protagen AG® with 384 different recombinant human proteins has been successfully used to validate off-target activities of adalimumab and infliximab [129].

FACS-based methods have also been developed for nonspecific binding: the polyspecificity reagent (PSR) assay [113] and the polyspecificity particle (PSP) assay [130]. The PSR assay uses complex yeast display technology to present antibodies and mixtures of soluble membrane proteins from Chinese hamster ovary cells as nontarget molecules to evaluate nonspecific binding [131], a complex and expensive method. The PSP assay uses protein A magnetic beads to capture antibodies and ovalbumin as nontarget molecules to incubate with antibodies, a more convenient and cost-saving method than PSR [130]. PSR and PSP yield similar results, and the quality of the data is higher than that of ELISA. In very early stages, if these methods are too expensive or time-consuming to test on dozens to hundreds of antibodies, we suggest testing the binding of antibodies to a molecule that is homologous in sequence or structure with the target of interest. For example, anti-CTLA4 antibodies could be tested on binding to CD28.

IN SILICO PREDICTION

Computational tools have been used to evaluate developability-related attributes, including chemical stability, folding energy and surface colloidal properties [132]. Good chemical stability of antibodies is associated with high homogeneity [54]. Low folding energy of an antibody is associated with high thermal stability, high expression level and robustness in different physicochemical conditions (such as cell culture, formulation buffer, storage, different pHs, light exposure, etc.) [133]. Colloidal properties of the antibody surface, such as hydrophobicity and asymmetric electrical distribution, may affect solubility, viscosity and aggregation [77]. A considerable number of in silico methods is available for predicting the intrinsic and phenomenological properties of antibodies. The inputs of these computational tools can be sequences, structures and combined sequences and structures. In the context of the theories used, these methods can be divided into energy-based, statistics-based and machine learning-based [132]. The machine learning-based method retrieves sequence and structural features from large amounts of data to train models, and then uses these models to predict the related properties of new inputs [133]. The quantity and quality of training data are key factors for machine learning models. Some recently released or commonly used in silico methods are described in the Supplementary Table 1.

PTM and heterogeneity

High-risk motifs of PTM include NS, NG (deamidation), DG (isomerization) and NXT/S (N-glycosylation) in the variable region and unpaired Cys (aggregation, oxidation) at any position on an antibody [54]. These motifs can be easily identified. Recently, machine learning-based models have been used to predict PTMs. Delmar *et al*. reported a machine learning model trained by an experimentally determined dataset of 766 peptides from a mAb [134]. Sankar *et al*. reported a quantitative model trained by a dataset of methionine oxidation in 122 distinct mAbs to predict methionine oxidation [135]. Both methods employed structure and sequence features (such as net charge and solvent-accessible surface area) for training, whereas Sankar *et al*. also took dynamic features into consideration by using coarse-grained elastic network models.

Aggregation and solubility

Aggregation and solubility, which are closely related attributes, can be predicted by using a wide range of computational methods, such as SAP, CamSol and Aggrescan3D [136–138]. These methods have assisted in the rational design of soluble and stable therapeutic proteins [77]. SAP and Aggrescan3D share a common hypothesis: the exposed hydrophobic residues are likely to form aggregation-prone regions. These methods employ the amount of solvent-accessible surface of hydrophobic residues as a measurement of the aggregation-prone region. Aggrescan3D exploits an experimentally derived intrinsic aggregation propensity scale for each residue, whereas SAP uses a normalized hydrophobicity value form literature [12]. The static protein structure and dynamic trajectory from structural simulation can be used as inputs for SAP and Aggrescan3D, where dynamic inputs may have high accuracy and static inputs have high throughput [136, 137].

The CamSol method can be used to evaluate the protein solubility properties. Based on protein structure, CamSol exploits the physicochemical properties of amino acid residues, such as hydrophobicity, electrostatic charges and the spatial residue interaction, to calculate a protein solubility score [138]. The solubility score provided by CamSol can be used to rank libraries of proteins. These methods require a high resolution of protein structures to ensure the accuracy of the prediction. However, in the discovery stage, antibody structures are usually predicted by homology modelling with relatively lower precision due to CDR3 diversity, which hinders the application of these structure-based in silico methods. Recently released AlphaFold [139], with well-modelled protein structure, may improve the predictive accuracy of these structure-based methods.

A number of machine learning and deep learning-based methods have been reported for prediction of protein solubility. SOLart and PON-Sol2 [140, 141] are machine learning models that predict protein solubility by high-dimensional protein features. SOLart integrates structure-based features (backbone torsion angles, solvent accessibility) and solubility-dependent statistical potentials into a random forest model. SOLart is trained by a set of protein solubility data, with a Pearson correlation coefficient between experimental and predicted solubility values of ~0.7 in a validation dataset. PON-Sol2 is a gradient boosting model trained on a dataset of 6 328 variants [141]. PON-Sol2 started with up to 1 081 features in the

categories of amino acid propensities and characteristics, conservation, variation type, neighbourhood features and length. These two methods can be used to identify amino acid mutations that change protein solubility.

Thermal stability and folding energy

The stability of an antibody refers to its ability to exist in its folded form relative to its denatured state. The thermal stability and folding energy of antibodies are critical to their biological function and developability. Evaluating antibody folding energy can improve our fundamental understanding of antibody structure and suggest possible routes to improve antibody stability [133]. Numerous in silico methods have been developed to predict the effect of mutations on stability and folding energy. FoldX [142] and Rosetta-ΔΔG [143] are two structure-based methods validated by experimental data. FoldX and Rosetta-ΔΔG use the folding energy (ΔΔG) between the wild type and the variant as an assessment of protein stability. Statistical potentials and various structural-sampling techniques are used for calculation. The statistical potentials include a variety of features, such as the energies calculated by molecular force field, probabilities of specific skeletal conformations and probabilities of rotational isomers. However, these two methods are highly computing-intensive and may not be accurate in predicting variants with a number of mutations.

Some deep learning-based methods are being developed to predict changes in the folding energy of proteins after mutations [144, 145]. Harmalkar *et al*. proposed a supervised convolutional neural network model with Rosetta energetic features for prediction of antibody thermostability. This model uses sequence one hot encoding and the residue energy matrix as two features to feed the CNN network. The model is trained on a dataset of 2 700 antibody sequences, with the accuracy up to 0.93 in the test dataset [146]. SCONES and ProS-GNN are two graph neural networks models that treat proteins as a network graph, where a vertex is an atom and edges are chemical bonds [144, 145]. The total free energy of a protein is the sum of all atomic energy contributions. The contributions of a specific residue to ΔG can be calculated from its local environment by defining a map of neighbouring residues with a contact cut-off (such as Cβ atom distance). The local environment can then be treated as a graph with residues/atoms as nodes and interactions as edges for training. These deep learning-based methods offer a new viable solution to elucidating the structure–property relationship directly from protein structural data.

SUMMARY AND FUTURE PERSPECTIVE

mAbs, bsAbs and ADCs have become important therapeutic entities. In an effective antibody discovery programme, both the biology- and developability-related parameters of the leading candidates need to be carefully examined. At present, screening and optimization of the developability of the leading candidates have been advanced to the early discovery stage. These developability assessments are usually required to have high throughput to screening dozens to hundreds of antibody candidates while consuming relatively small amounts of testing materials. In addition, computational tools have been used to design, predict, screen and optimize the developability of antibody candidates. AlphaFold and other artificial intelligence tools show high accuracy in the prediction of protein structure. Structure- and force field-based methods can utilize structures predicted by artificial intelligence tools to improve their accuracy for the evaluation of protein developability. In addition, deep generative learning models are proposed for de novo design of proteins with desired properties, such as affinity, solubility and stability. It is important to note that, with improved quality and increased quantity of protein datasets, computational tools with high accuracy will be widely used in the evaluation of developability. Multiple properties, such as affinity, stability and solubility of antibody-derived molecules, can be predicted at the same time. These high-accuracy, multifaceted models will significantly reduce wet lab experiments and accelerate the development of biological drugs.

SUPPLEMENTARY DATA

Supplementary Data are available at ABT Online.

FUNDING STATEMENT

This work was supported by WuXi Biologics.

CONFLICT OF INTEREST STATEMENT

Weijie Zhang, Hao Wang, Nan Feng, Yifeng Li, Jijie Gu and Zhuozhi Wang are current employees of WuXi Biologics and may hold WuXi Biologics' stocks.

DECLARATIONS

Weijie Zhang, Hao Wang, Nan Feng, Yifeng Li, Jijie Gu and Zhuozhi Wang are current employees of WuXi Biologics and may hold WuXi Biologics' stocks.

DATA AVAILABILITY

The data that support this study are openly available.

ETHICS AND CONSENT

Consent was not required.

ANIMAL RESEARCH

This is not applicable.

ABBREVIATION

ADC, antibody drug conjugate; AI, artificial intelligence; mAb, monoclonal antibody; bsAb, bispecific antibody;

msAb, multispecific antibody; CMC, chemistry, manufacturing and control; HC, heavy chain; LC, light chain; Fab, antigen-binding fragment; Fc, fragment of crystallizable region; CDR, complementary-determining region; CDR L1, L2, L3, complementary-determining region 1, 2, 3 of light chain; CDR H1, H2, H3, complementary-determining region 1, 2, 3 of heavy chain; VH, variable domain of heavy chain; VL, variable domain of light chain; kD, diffusion interaction parameter; PPI, protein–protein interaction; HIC, hydrophobic chromatography; DSF, differential scanning fluorimetry; DLS, dynamic light scattering; SPR, surface plasmon resonance; BLI, biolayer Interferometry; SEC-HPLC, size-exclusion chromatography high-performance liquid chromatography; SDS-PAGE, sodium dodecyl sulphate–polyacrylamide gel electrophoresis; CE-SDS, capillary electrophoresis sodium dodecyl sulfate; IEF, isoelectric focusing; icIEF, imaged capillary isoelectric focusing; IEX, ion-exchange chromatography; RP-HPLC, reversed-phase high-performance liquid chromatography; AC-SINS, affinity-capture self-interaction nanoparticle spectrometry; PTM, post-translational modification.

REFERENCES

1. Xu, Y, Wang, D, Mason, B *et al.* Structure, heterogeneity and developability assessment of therapeutic antibodies. *MAbs* 2019; **11**: 239–64.
2. Bailly, M, Mieczkowski, C, Juan, V *et al.* Predicting antibody developability profiles through early stage discovery screening. *MAbs* 2020; **12**: 1743053.
3. Yang, Y, Velayudhan, A, Thornhill, NF *et al.* Multi-criteria manufacturability indices for ranking high-concentration monoclonal antibody formulations. *Biotechnol Bioeng* 2017; **114**: 2043–56.
4. Yang, X, Xu, W, Dukleska, S *et al.* Developability studies before initiation of process development: improving manufacturability of monoclonal antibodies. *MAbs* 2013; **5**: 787–94.
5. Vlasak, J, Bussat, MC, Wang, S *et al.* Identification and characterization of asparagine deamidation in the light chain CDR1 of a humanized IgG1 antibody. *Anal Biochem* 2009; **392**: 145–54.
6. Pindrus, M, Shire, SJ, Kelley, RF *et al.* Solubility challenges in high concentration monoclonal antibody formulations: relationship with amino acid sequence and intermolecular interactions. *Mol Pharm* 2015; **12**: 3896–907.
7. Root, AR, Guntas, G, Katragadda, M *et al.* Discovery and optimization of a novel anti-GUCY2c x CD3 bispecific antibody for the treatment of solid tumors. *MAbs* 2021; **13**: 1850395.
8. Suurs, FV, Lub-de Hooge, MN, de Vries, EGE *et al.* A review of bispecific antibodies and antibody constructs in oncology and clinical challenges. *Pharmacol Ther* 2019; **201**: 103–19.
9. Rader, C. Bispecific antibodies in cancer immunotherapy. *Curr Opin Biotechnol* 2020; **65**: 9–16.
10. Li, W, Prabakaran, P, Chen, W *et al.* Antibody aggregation: insights from sequence and structure. *Antibodies* 2016; **5**: 19.
11. Yadav, S, Laue, TM, Kalonia, DS *et al.* The influence of charge distribution on self-association and viscosity behavior of monoclonal antibody solutions. *Mol Pharm* 2012; **9**: 791–802.
12. Black, SD, Mould, DR. Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. *Anal Biochem* 1991; **193**: 72–82.
13. Perchiacca, JM, Ladiwala, ARA, Bhattacharya, M *et al.* Aggregation-resistant domain antibodies engineered with charged mutations near the edges of the complementarity-determining regions. *Protein Eng Des Sel* 2012; **25**: 591–601.
14. Nie, S, Wang, Z, Moscoso-castro, M *et al.* Biology drives the discovery of bispeci fi c antibodies as innovative therapeutics. *Antib Ther* 2020; **3**: 18–62.
15. Fan, G, Wang, Z, Hao, M *et al.* Bispecific antibodies and their applications. *J Hematol Oncol* 2015; **8**: 130.
16. Andrade, C, Arnold, L, Motabar, D *et al.* An integrated approach to aggregate control for therapeutic bispecific antibodies using an improved three column mab platform-like purification process. *Biotechnol Prog* 2019; **35**: e2720.
17. Chen, SW, Zhang, W. Current trends and challenges in the downstream purification of bispecific antibodies. *Antib Ther* 2021; **4**: 73–88.
18. Gibson, SJ, Bond, NJ, Milne, S *et al.* N-terminal or signal peptide sequence engineering prevents truncation of human monoclonal antibody light chains. *Biotechnol Bioeng* 2017; **114**: 1970–7.
19. Ambrogelly, A, Liu, YH, Li, H *et al.* Characterization of antibody variants during process development: the tale of incomplete processing of N-terminal secretion peptide. *MAbs* 2012; **4**: 701–9.
20. Shaw, SY, Margolies, MN. A spontaneous variant of an antidigoxin hybridoma antibody with increased affinity arises from a heavy chain signal peptide mutation. *Mol Immunol* 1992; **29**: 525–9.
21. Lyubarskaya, Y, Houde, D, Woodard, J *et al.* Analysis of recombinant monoclonal antibody isoforms by electrospray ionization mass spectrometry as a strategy for streamlining characterization of recombinant monoclonal antibody charge heterogeneity. *Anal Biochem* 2006; **348**: 24–39.
22. Vlasak, J, Ionescu, R. Fragmentation of monoclonal antibodies. *MAbs* 2011; **3**: 253–63.
23. Neergaard, MS, Nielsen, AD, Parshad, H *et al.* Stability of monoclonal antibodies at high-concentration: head-to-head comparison of the IgG1 and IgG4 subclass. *J Pharm Sci* 2014; **103**: 115–27.
24. Liu, H, Gaza-Bulseco, G, Lundell, E. Assessment of antibody fragmentation by reversed-phase liquid chromatography and mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 2008; **876**: 13–23.
25. Brorson, K, Jia, AY. Therapeutic monoclonal antibodies and consistent ends: terminal heterogeneity, detection, and impact on quality. *Curr Opin Biotechnol* 2014; **30**: 140–6.
26. Liu, H, Gaza-Bulseco, G, Faldu, D *et al.* Heterogeneity of monoclonal antibodies. *J Pharm Sci* 2008; **97**: 2426–47.
27. Jiang, G, Yu, C, Yadav, DB *et al.* Evaluation of heavy-chain C-terminal deletion on product quality and pharmacokinetics of monoclonal antibodies. *J Pharm Sci* 2016; **105**: 2066–72.
28. Hu, Z, Tang, D, Misaghi, S *et al.* Evaluation of heavy chain C-terminal deletions on productivity and product quality of monoclonal antibodies in Chinese hamster ovary (CHO) cells. *Biotechnol Prog* 2017; **33**: 786–94.
29. van den Bremer, ETJ, Beurskens, FJ, Voorhorst, M *et al.* Human IgG is produced in a pro-form that requires clipping of C-terminal lysines for maximal complement activation. *MAbs* 2015; **7**: 672–80.
30. Beck, A, Liu, H. Macro- and micro-heterogeneity of natural and recombinant IgG antibodies. *Antibodies* 2019; **8**: 18.
31. Liu, H, Ponniah, G, Zhang, HM *et al.* In vitro and in vivo modifications of recombinant and human IgG antibodies. *MAbs* 2014; **6**: 1145–54.
32. Liu, H, May, K. Disulfide bond structures of IgG molecules: structural variations, chemical modifications and possible impacts to stability and biological function. *MAbs* 2012; **4**: 17–23.
33. Liu-Shin, LPY, Fung, A, Malhotra, A *et al.* Evidence of disulfide bond scrambling during production of an antibody-drug conjugate. *MAbs* 2018; **10**: 1190–9.
34. Rispens, T, Ooijevaar-De Heer, P, Bende, O *et al.* Mechanism of immunoglobulin G4 Fab-arm exchange. *J Am Chem Soc* 2011; **133**: 10302–11.
35. Peters, SJ, Smales, CM, Henry, AJ *et al.* Engineering an improved IgG4 molecule with reduced disulfide bond heterogeneity and increased fab domain thermal stability. *J Biol Chem* 2012; **287**: 24525–33.
36. Chumsae, C, Gaza-Bulseco, G, Liu, H. Identification and localization of unpaired cysteine residues in monoclonal antibodies by fluorescence labeling and mass spectrometry. *Anal Chem* 2009; **81**: 6449–57.
37. Wei, B, Jia, W, Yang, Y *et al.* Development of a rapid reversed-phase liquid chromatographic method for total free thiol

quantitation in protein therapeutics. *J Pharm Biomed Anal* 2020; **189**: 113434.
38. Nashef, AS, Osuga, DT, Lee, HS *et al.* Effects of alkali on proteins. Disulfides and their products. *J Agric Food Chem* 1977; **25**: 245–51.
39. Tous, GI, Wei, Z, Feng, J *et al.* Characterization of a novel modification to monoclonal antibodies: thioether cross-link of heavy and light chains. *Anal Chem* 2005; **77**: 2675–82.
40. Harris, RJ. Heterogeneity of recombinant antibodies: linking structure to function. *Dev Biol (Basel)* 2005; **122**: 117–27.
41. Lacy, ER, Baker, M, Brigham-Burke, M. Free sulfhydryl measurement as an indicator of antibody stability. *Anal Biochem* 2008; **382**: 66–8.
42. Li, X, Xiao, L, Kochert, B *et al.* Extended characterization of unpaired cysteines in an IgG1 monoclonal antibody by LC-MS analysis. *Anal Biochem* 2021; **622**: 114172.
43. Buchanan, A, Clementel, V, Woods, R *et al.* Engineering a therapeutic IgG molecule to address cysteinylation, aggregation and enhance thermal stability and expression. *MAbs* 2013; **5**: 255–62.
44. Banks, DD, Gadgil, HS, Pipes, GD *et al.* Removal of cysteinylation from an unpaired sulfhydryl in the variable region of a recombinant monoclonal IgG1 antibody improves homogeneity, stability, and biological activity. *J Pharm Sci* 2008; **97**: 775–90.
45. Gu, S, Wen, D, Weinreb, PH *et al.* Characterization of trisulfide modification in antibodies. *Anal Biochem* 2010; **400**: 89–98.
46. Pristatsky, P, Cohen, SL, Krantz, D *et al.* Evidence for trisulfide bonds in a recombinant variant of a human IgG2 monoclonal antibody. *Anal Chem* 2009; **81**: 6148–55.
47. Aono, H, Wen, D, Zang, L *et al.* Efficient on-column conversion of IgG1 trisulfide linkages to native disulfides in tandem with Protein A affinity chromatography. *J Chromatogr A* 2010; **1217**: 5225–32.
48. Sydow, JF, Lipsmeier, F, Larraillet, V *et al.* Structure-based prediction of asparagine and aspartate degradation sites in antibody variable regions. *PLoS One* 2014; **9**: 901–12.
49. Harris, RJ, Kabakoff, B, Macchi, FD *et al.* Identification of multiple sources of charge heterogeneity in a recombinant antibody. *J Chromatogr B Biomed Sci Appl* 2001; **752**: 233–45.
50. Huang, L, Lu, J, Wroblewski, VJ *et al.* In vivo deamidation characterization of monoclonal antibody by LC/MS/MS. *Anal Chem* 2005; **77**: 1432–9.
51. Mamula, MJ, Gee, RJ, Elliott, JI *et al.* Isoaspartyl post-translational modification triggers autoimmune responses to self-proteins. *J Biol Chem* 1999; **274**: 22321–7.
52. Hermeling, S, Crommelin, DJA, Schellekens, H *et al.* Structure-immunogenicity relationships of therapeutic proteins. *Pharm Res* 2004; **21**: 897–903.
53. Wakankar, AA, Borchardt, RT. Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization. *J Pharm Sci* 2006; **95**: 2321–36.
54. Lu, X, Nobrega, RP, Lynaugh, H *et al.* Deamidation and isomerization liability analysis of 131 clinical-stage antibodies. *MAbs* 2019; **11**: 45–57.
55. Sreedhara, A, Cordoba, A, Zhu, Q *et al.* Characterization of the isomerization products of aspartate residues at two different sites in a monoclonal antibody. *Pharm Res* 2012; **29**: 187–97.
56. Eakin, CM, Miller, A, Kerr, J *et al.* Assessing analytical methods to monitor isoAsp formation in monoclonal antibodies. *Front Pharmacol* 2014; **5**: 87.
57. Oliyai, C, Borchardt, RT. Chemical pathways of peptide degradation. IV. Pathways, kinetics, and mechanism of degradation of an aspartyl residue in a model hexapeptide. *Pharm Res An Off J Am Assoc Pharm Sci* 1993; **10**: 95–102.
58. Jefferis, R. Posttranslational modifications and the immunogenicity of biotherapeutics. *J Immunol Res* 2016; **2016**: 5358272.
59. Wang, Z, Zhu, J, Lu, H. Antibody glycosylation: impact on antibody drug characteristics and quality control. *Appl Microbiol Biotechnol* 2020; **104**: 1905–14.
60. Wright, A, Morrison, SL. Effect of glycosylation on antibody function: implications for genetic engineering. *Trends Biotechnol* 1997; **15**: 26–32.
61. Krapp, S, Mimura, Y, Jefferis, R *et al.* Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol* 2003; **325**: 979–89.
62. Arnold, JN, Wormald, MR, Sim, RB *et al.* The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu Rev Immunol* 2007; **25**: 21–50.
63. Jarasch, A, Koll, H, Regula, JT *et al.* Developability assessment during the selection of novel therapeutic antibodies. *J Pharm Sci* 2015; **104**: 1885–98.
64. Subedi, GP, Barb, AW. The structural role of antibody n-glycosylation in receptor interactions. *Structure* 2015; **23**: 1573–83.
65. Raju, TS, Briggs, JB, Borge, SM *et al.* Species-specific variation in glycosylation of Igc: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics. *Glycobiology* 2000; **10**: 477–86.
66. Kaneko, Y, Nimmerjahn, F, Ravetch, JV. Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* 2006; **313**: 670–3.
67. Qian, J, Liu, T, Yang, L *et al.* Structural characterization of N-linked oligosaccharides on monoclonal antibody cetuximab by the combination of orthogonal matrix-assisted laser desorption/ionization hybrid quadrupole-quadrupole time-of-flight tandem mass spectrometry and sequential enzy. *Anal Biochem* 2007; **364**: 8–18.
68. Liu, D, Ren, D, Huang, H *et al.* Structure and stability changes of human IgG1 Fc as a consequence of methionine oxidation. *Biochemistry* 2008; **47**: 5088–100.
69. Liu, H, Gaza-Bulseco, G, Xiang, T *et al.* Structural effect of deglycosylation and methionine oxidation on a recombinant monoclonal antibody. *Mol Immunol* 2008; **45**: 701–8.
70. Mo, J, Yan, Q, So, CK *et al.* Understanding the impact of methionine oxidation on the biological functions of IgG1 antibodies using hydrogen/deuterium exchange mass spectrometry. *Anal Chem* 2016; **88**: 9495–502.
71. Bertolotti-Ciarlet, A, Wang, W, Lownes, R *et al.* Impact of methionine oxidation on the binding of human IgG1 to FcRn and Fcγ receptors. *Mol Immunol* 2009; **46**: 1878–82.
72. Wang, W, Vlasak, J, Li, Y *et al.* Impact of methionine oxidation in human IgG1 Fc on serum half-life of monoclonal antibodies. *Mol Immunol* 2011; **48**: 860–6.
73. Sharma, VK, Patapoff, TW, Kabakoff, B *et al.* In silico selection of therapeutic antibodies for development: viscosity, clearance, and chemical stability. *Proc Natl Acad Sci U S A* 2014; **111**: 18601–6.
74. Dashivets, T, Stracke, J, Dengl, S *et al.* Oxidation in the complementarity-determining regions differentially influences the properties of therapeutic antibodies. *MAbs* 2016; **8**: 1525–35.
75. Hageman, T, Wei, H, Kuehne, P *et al.* Impact of tryptophan oxidation in complementarity-determining regions of two monoclonal antibodies on structure-function characterized by hydrogen-deuterium exchange mass spectrometry and surface plasmon resonance. *Pharm Res* 2019; **36**: 24.
76. Manning, MC, Chou, DK, Murphy, BM *et al.* Stability of protein pharmaceuticals: an update. *Pharm Res* 2010; **27**: 544–75.
77. Prabakaran, R, Rawat, P, Thangakani, AM *et al.* Protein aggregation: in silico algorithms and applications. *Biophys Rev* 2021; **13**: 71–89.
78. Stourac, J, Dubrava, J, Musil, M *et al.* FireProtDB: database of manually curated protein stability data. *Nucleic Acids Res* 2021; **49**: D319–24.
79. Kueltzo, LA, Wang, W, Randolph, TW *et al.* Effects of solution conditions, processing parameters, and container materials on aggregation of a monoclonal antibody during freeze-thawing. *J Pharm Sci* 2008; **97**: 1801–12.
80. Kendrick, BS, Li, T, Chang, BS. Physical stabilization of proteins in aqueous solution. *Pharm Biotechnol* 2002; **13**: 61–84.
81. Arakawa, T, Timasheff, SN. Preferential interactions of proteins with salts in concentrated solutions. *Biochemistry* 1982; **21**: 6545–52.
82. Kameoka, D, Masuzaki, E, Ueda, T *et al.* Effect of buffer species on the unfolding and the aggregation of humanized IgG. *J Biochem* 2007; **142**: 383–91.
83. Vogt, G, Argos, P. Protein thermal stability: hydrogen bonds or internal packing? *Fold Des* 1997; **2**: S40–6.
84. Kerwin, BA, Remmele, RL. Protect from light: photodegradation and protein biologics. *J Pharm Sci* 2007; **96**: 1468–79.

85. Du, C, Barnett, G, Borwankar, A *et al.* Protection of therapeutic antibodies from visible light induced degradation: use safe light in manufacturing and storage. *Eur J Pharm Biopharm* 2018; **127**: 37–43.
86. Qi, P, Volkin, DB, Zhao, H *et al.* Characterization of the photodegradation of a human IgG1 monoclonal antibody formulated as a high-concentration liquid dosage form. *J Pharm Sci* 2009; **98**: 3117–30.
87. He, F, Hogan, S, Latypov, RF *et al.* High throughput thermostability screening of monoclonal antibody formulations. *J Pharm Sci* 2010; **99**: 1707–20.
88. Jin, W, Xing, Z, Song, Y *et al.* Protein aggregation and mitigation strategy in low pH viral inactivation for monoclonal antibody purification. *MAbs* 2019; **11**: 1479–91.
89. Hawe, A, Wiggenhorn, M, van de Weert, M *et al.* Forced degradation of therapeutic proteins. *J Pharm Sci* 2012; **101**: 895–913.
90. Bee, JS, Stevenson, JL, Mehta, B *et al.* Response of a concentrated monoclonal antibody formulation to high shear. *Biotechnol Bioeng* 2009; **103**: 936–43.
91. Clarkson, J, Cui, Z, Darton, R. Protein denaturation in foam. *J Colloid Interface Sci* 1999; **215**: 333–8.
92. Li, Y, Mach, H, Blue, JT. High throughput formulation screening for global aggregation behaviors of three monoclonal antibodies. *J Pharm Sci* 2011; **100**: 2120–35.
93. Wang, W, Wang, YJ, Wang, DQ. Dual effects of Tween 80 on protein stability. *Int J Pharm* 2008; **347**: 31–8.
94. Wang, S, Wu, G, Zhang, X *et al.* Stabilizing two IgG1 monoclonal antibodies by surfactants: balance between aggregation prevention and structure perturbation. *Eur J Pharm Biopharm* 2017; **114**: 263–77.
95. Jain, K, Salamat-Miller, N, Taylor, K. Freeze–thaw characterization process to minimize aggregation and enable drug product manufacturing of protein based therapeutics. *Sci Rep* 2021; **11**: 1133.
96. Franks, F. Protein destabilization at low temperatures. *Adv Protein Chem* 1995; **46**: 105–39.
97. Strambini, GB, Gabellieri, E. Proteins in frozen solutions: evidence of ice-induced partial unfolding. *Biophys J* 1996; **70**: 971–6.
98. Kolhe, P, Amend, E, Singh, SK. Impact of freezing on pH of buffered solutions and consequences for monoclonal antibody aggregation. *Biotechnol Prog* 2010; **26**: 727–33.
99. Bethea, D, Wu, SJ, Luo, J *et al.* Mechanisms of self-association of a human monoclonal antibody CNTO607. *Protein Eng Des Sel* 2012; **25**: 531–7.
100. Geng, SB, Cheung, JK, Narasimhan, C *et al.* Improving monoclonal antibody selection and engineering using measurements of colloidal protein interactions. *J Pharm Sci* 2014; **103**: 3356–63.
101. Narasimhan, C, MacH, H, Shameem, M. High-dose monoclonal antibodies via the subcutaneous route: challenges and technical solutions, an industry perspective. *Ther Deliv* 2012; **3**: 889–900.
102. Sela-Culang, I, Kunik, V, Ofran, Y. The structural basis of antibody-antigen recognition. *Front Immunol* 2013; **4**: 302.
103. Adey, NB, Mataragnon, AH, Rider, JE *et al.* Characterization of phage that bind plastic from phage-displayed random peptide libraries. *Gene* 1995; **156**: 27–31.
104. Ma, D, Baruch, D, Shu, Y *et al.* Using protein microarray technology to screen anti-ERCC1 monoclonal antibodies for specificity and applications in pathology. *BMC Biotechnol* 2012; **12**: 88.
105. Cunningham, O, Scott, M, Zhou, ZS *et al.* Polyreactivity and polyspecificity in therapeutic antibody development: risk factors for failure in preclinical and clinical development campaigns. *MAbs* 2021; **13**: 1999195.
106. Kohli, N, Jain, N, Geddie, ML *et al.* A novel screening method to assess developability of antibody-like molecules. *MAbs* 2015; **7**: 752–8.
107. Liu, L. Pharmacokinetics of monoclonal antibodies and Fc-fusion proteins. *Protein Cell* 2018; **9**: 15–32.
108. Ridker, PM, Revkin, J, Amarenco, P *et al.* Cardiovascular efficacy and safety of bococizumab in high-risk patients. *N Engl J Med* 2017; **376**: 1527–39.
109. Ridker, PM, Tardif, J-C, Amarenco, P *et al.* Lipid-reduction variability and antidrug-antibody formation with bococizumab. *N Engl J Med* 2017; **376**: 1517–26.
110. Finlay, WJJ, Coleman, JE, Edwards, JS *et al.* Anti-PD1 'SHR-1210′ aberrantly targets pro-angiogenic receptors and this polyspecificity can be ablated by paratope refinement. *MAbs* 2019; **11**: 26–44.
111. Chen, X, Ma, L, Wang, X *et al.* Reactive capillary hemangiomas: a novel dermatologic toxicity following anti-PD-1 treatment with SHR-1210. *Cancer Biol Med* 2019; **16**: 173–80.
112. Popovic, B, Gibson, S, Senussi, T *et al.* Engineering the expression of an anti-interleukin-13 antibody through rational design and mutagenesis. *Protein Eng Des Sel* 2017; **30**: 303–11.
113. Jain, T, Sun, T, Durand, S *et al.* Biophysical properties of the clinical-stage antibody landscape. *Proc Natl Acad Sci U S A* 2017; **114**: 944–9.
114. Chai, Q, Shih, J, Weldon, C *et al.* Development of a high-throughput solubility screening assay for use in antibody discovery. *MAbs* 2019; **11**: 747–56.
115. Gibson, TJ, Mccarty, K, Mcfadyen, IJ *et al.* Application of a high-throughput screening procedure with PEG-induced precipitation to compare relative protein solubility during formulation development with IgG1 monoclonal antibodies. *J Pharm Sci* 2011; **100**: 1009–21.
116. Hofmann, M, Winzer, M, Weber, C *et al.* Limitations of polyethylene glycol-induced precipitation as predictive tool for protein solubility during formulation development. *J Pharm Pharmacol* 2018; **70**: 648–54.
117. He, F, Becker, GW, Litowski, JR *et al.* High-throughput dynamic light scattering method for measuring viscosity of concentrated protein solutions. *Anal Biochem* 2010; **399**: 141–3.
118. Liu, Y, Caffry, I, Wu, J *et al.* High-throughput screening for developability during early-stage antibody discovery using self-interaction nanoparticle spectroscopy. *MAbs* 2014; **6**: 483–92.
119. Sun, T, Reid, F, Liu, Y *et al.* High throughput detection of antibody self-interaction by bio-layer interferometry. *MAbs* 2013; **5**: 838–41.
120. He, F, Woods, CE, Becker, GW *et al.* High-throughput assessment of thermal and colloidal stability parameters for monoclonal antibody formulations. *J Pharm Sci* 2011; **100**: 5126–41.
121. Neergaard, MS, Kalonia, DS, Parshad, H *et al.* Viscosity of high concentration protein formulations of monoclonal antibodies of the IgG1 and IgG4 subclass – prediction of viscosity through protein-protein interaction measurements. *Eur J Pharm Sci* 2013; **49**: 400–10.
122. Avery, LB, Wade, J, Wang, M *et al.* Establishing in vitro in vivo correlations to screen monoclonal antibodies for physicochemical properties related to favorable human pharmacokinetics. *MAbs* 2018; **10**: 244–55.
123. Kim, DM, Yao, X, Vanam, RP *et al.* Measuring the effects of macromolecular crowding on antibody function with biolayer interferometry. *MAbs* 2019; **11**: 1319–30.
124. García-Ojeda, PA, Hardy, S, Kozlowski, S *et al.* Surface plasmon resonance analysis of antipolysaccharide antibody specificity: responses to meningococcal group C conjugate vaccines and bacteria. *Infect Immun* 2004; **72**: 3451–60.
125. Jacobs, SA, Wu, SJ, Feng, Y *et al.* Cross-interaction chromatography: a rapid method to identify highly soluble monoclonal antibody candidates. *Pharm Res* 2010; **27**: 65–71.
126. Starr, CG, Tessier, PM. Selecting and engineering monoclonal antibodies with drug-like specificity. *Curr Opin Biotechnol* 2019; **60**: 119–27.
127. Wardemann, H, Yurasov, S, Schaefer, A *et al.* Predominant autoantibody production by early human B cell precursors. *Science* 2003; **301**: 1374–7.
128. Hötzel, I, Theil, FP, Bernstein, LJ *et al.* A strategy for risk mitigation of antibodies with fast clearance. *MAbs* 2012; **4**: 753–60.
129. Feyen, O, Lueking, A, Kowald, A *et al.* Off-target activity of TNF-α inhibitors characterized by protein biochips. *Anal Bioanal Chem* 2008; **391**: 1713–20.
130. Makowski, EK, Wu, L, Desai, AA *et al.* Highly sensitive detection of antibody nonspecific interactions using flow cytometry. *MAbs* 2021; **13**: 1951426.
131. Xu, Y, Roach, W, Sun, T *et al.* Addressing polyspecificity of antibodies selected from an in vitro yeast presentation system: a

FACS-based, high-throughput selection and analytical tool. *Protein Eng Des Sel* 2013; **26**: 663–70.
132. Khetan, R, Curtis, R, Deane, CM *et al.* Current advances in biopharmaceutical informatics: guidelines, impact and challenges in the computational developability assessment of antibody therapeutics. *MAbs* 2022; **14**: 2020082.
133. Narayanan, H, Dingfelder, F, Butté, A *et al.* Machine learning for biologics: opportunities for protein engineering, developability, and formulation. *Trends Pharmacol Sci* 2021; **42**: 151–65.
134. Delmar, JA, Wang, J, Choi, SW *et al.* Machine learning enables accurate prediction of asparagine deamidation probability and rate. *Mol Ther* 2019; **15**: 264–74.
135. Sankar, K, Hoi, KH, Yin, Y *et al.* Prediction of methionine oxidation risk in monoclonal antibodies using a machine learning method. *MAbs* 2018; **10**: 1281–90.
136. Chennamsetty, N, Voynov, V, Kayser, V *et al.* Prediction of aggregation prone regions of therapeutic proteins. *J Phys Chem B* 2010; **114**: 6614–24.
137. Zambrano, R, Jamroz, M, Szczasiuk, A *et al.* AGGRESCAN3D (A3D): server for prediction of aggregation properties of protein structures. *Nucleic Acids Res* 2015; **43**: W306–13.
138. Sormanni, P, Aprile, FA, Vendruscolo, M. The CamSol method of rational design of protein mutants with enhanced solubility. *J Mol Biol* 2015; **427**: 478–90.
139. Jumper, J, Evans, R, Pritzel, A *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 2021; **596**: 583–9.
140. Hou, Q, Kwasigroch, JM, Rooman, M *et al.* SOLart: a structure-based method to predict protein solubility and aggregation. *Bioinformatics* 2020; **36**: 1445–52.
141. Yang, Y, Zeng, L, Vihinen, M. PON-Sol2: prediction of effects of variants on protein solubility. *Int J Mol Sci* 2021; **22**: 8027.
142. Delgado, J, Radusky, LG, Cianferoni, D *et al.* FoldX 5.0: working with RNA, small molecules and a new graphical interface. *Bioinformatics* 2019; **35**: 4168–9.
143. Kellogg, EH, Leaver-Fay, A, Baker, D. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. *Proteins Struct Funct Bioinforma* 2011; **79**: 830–8.
144. Samaga, YBL, Raghunathan, S, Priyakumar, UD. SCONES: self-consistent neural network for protein stability prediction upon mutation. *J Phys Chem B* 2021; **125**: 10657–71.
145. Wang, S, Tang, H, Zuo, L. ProS-GNN: predicting effects of mutations on protein stability using graph neural networksbioRxiv. 2021. https://doi.org/10.1101/2021.10.25.465658.
146. Harmalkar, A, Rao, RM, Honer, J *et al.* Towards generalizable prediction of antibody thermostability using machine learning on sequence and structure features*bioRxiv*. 2022. https://doi.org/10.1101/2022.06.03.494724.