Systematic analysis of the varied designs of 819 therapeutic antibodies and Fc fusion proteins assigned international nonproprietary names

Ian Wilkinson and Geoff Hale

mAbsolve Limited, Oxford, UK

ABSTRACT
By combining and cross-checking data from the World Health Organization’s lists of international nonproprietary names with three other databases, we assembled a dataset of amino acid sequences of 819 antibody-related therapeutics. This enabled the systematic tabulation of 57 different molecular formats and about 90 different constant-region variants (47 for silencing, 14 for enhancement, 8 for modifying binding to FcRn, 13 for heterodimerization, 4 for site-specific modification and 4 for stabilization), as well as the frequencies of different targets and immunoglobulin allotypes. The curated dataset provides a resource for researchers, giving insights into trends in antibody engineering and a guide to the most frequently tested designs.

Introduction
Since the first international nonproprietary name (INN) was assigned to a monoclonal antibody in 1988, there has been an explosion in the development of therapeutic antibodies and related biological drugs with a huge number of new formats and variants, as testified in the pages of this journal. The creativity of antibody engineers is such that systematic taxonomy lags behind and it is increasingly difficult for a student of the field to grasp its overall complexity. And although there are several valuable compilations of structural data, it is inevitable that some errors have occurred in the naming of amino acid sequences of these molecules. With this in mind, we have attempted to catalog and cross-check the primary structure of all of the antibodies and Fc fusion proteins published by the World Health Organization (WHO) up to April 2022.

An International nonproprietary Name (INN) is an official generic name given to a pharmaceutical drug or active ingredient, typically when it is in early-stage clinical trials. The INN system has been coordinated by WHO since 1953. Proposed names are first published in INN Proposed Lists (PL) for comment and when approved, they are published in INN Recommended Lists (RL). The nomenclature for monoclonal antibodies has undergone several revisions. Since the ad hoc naming of muromonab-CD3, the stem -mab was introduced, preceded by an infix to indicate the source: -o- (mouse), -axo- (rat/mouse), -u- (human), -xi- (chimeric), -zu- (humanized), -xizu- (chimeric/humanized). By 2014, new technologies were making it difficult to characterize antibodies simply by the way they were created and new criteria were proposed based on sequence homologies. However, this was severely criticized by the scientific community as being arbitrary and ambiguous. And so, in 2017 the source infixes were dropped. By 2021, 880 INNs with the stem -mab had been approved, and the system was becoming congested. A radical decision was made to discontinue the -mab stem (except for its use in -vetmab) and replace it with four new stems: -tug (monospecific, full-length, Fc unmodified immunoglobulins), -bart (monospecific, full-length immunoglobulins with engineered constant domains), -mig (bispecific or multispecific immunoglobulins of any format) and -ment (fragments derived from an immunoglobulin variable domain). Meanwhile, fusion proteins which contain an immunoglobulin Fc domain have been characterized by the stem -fusp (all components having a pharmacological activity) or the prefix ef- (where the Fc region has only a stabilizing function), though several also have the stem -cept (receptor molecules or ligands) or other stems relating to properties of the fusion partner.

Since about 2005, the WHO INN lists started to include protein sequences, but often not in a machine-readable format. The compilers of various databases, including Thera-SAbDab and the international ImMunoGeneTics information system (IMGT) 2Dstructure-DB have used such sources to create systematic sequence listings. However, it was difficult, if not impossible, to determine variations in Fc region sequences from these secondary sources. Therefore, we have re-extracted protein sequences from the WHO INN lists. We compared them with listings from the publicly available databases and manually curated our dataset to reconcile discrepancies. We believe this represents the most complete and accurate listing of sequences for antibodies and Fc fusion proteins that have been assigned INN names. However, it must be acknowledged that there could still be inaccuracies in the original WHO lists.
Pharmacologic properties of immunoglobulins depend very much on their Fc region. Interaction with C1q initiates activation of complement. Binding to various Fc receptors on leukocytes induces antibody-dependent cell-mediated cytotoxicity (ADCC) or antibody-dependent cell-mediated phagocytosis (ADCP). Binding to the FcRn receptor is responsible for the comparatively long half-life of IgG. For many years, scientists have investigated the structures involved in binding to these various ligands, with the aim of modifying the natural properties of antibodies either to enhance (e.g., to improve killing of tumor cells, or to extend half-life) or to reduce (e.g., to avoid unwanted side effects).\textsuperscript{11–14} As a result, many therapeutic antibodies and Fc fusion proteins have Fc regions which have been altered compared with wild-type immunoglobulins.

For example, binding to C1q and Fc-\gamma receptors is reduced by the mutations L234A/L235A ("LALA")\textsuperscript{15,16} or N297A (removing a carbohydrate attachment site).\textsuperscript{17} Conversely, binding can be enhanced by reducing the extent of fucosylation of the N-linked carbohydrate.\textsuperscript{18} Binding to FcRn at low pH and consequently half-life in vivo can be increased by various mutations such as M252Y/S254T/T256E ("YTE")\textsuperscript{19} or reduced by I253A/H310A/H435A.\textsuperscript{20} IgG4 antibodies can be stabilized and manufacturability improved by S228P\textsuperscript{21} and heterogeneity can be reduced by truncation of the C-terminal lysine K447A.\textsuperscript{22} Mutations such as S239C\textsuperscript{23} have been introduced to facilitate site-specific conjugation of drugs. Other mutations have been used to create heterodimeric bispecific antibodies for example, by replacing a small amino acid with a larger one on one chain and a larger with a smaller on the other to create "knobs into holes" modifications.\textsuperscript{24}

Using the new dataset of sequences, we can now catalog the modifications that have been incorporated into the Fc region of antibodies and fusion proteins, including those that reached late-stage clinical trials, revealing a remarkable diversity of variants and providing an overview of trends in antibody engineering over the past 25 years.

Results

Construction of sequence database

Protein sequences from a total of 772 monoclonal antibodies and Fc fusion proteins were extracted from the WHO INN lists. They were compared with corresponding sequences from the Thera-SAbDab, IMGT and Inxight databases. Some sequences that were not present in the INN lists were found in one or more of the alternate databases; presumably they had been obtained from other sources. The results are summarized in Table 1.

Sequences of 24 Fc fusion proteins and two antibodies were found only in INN lists. The Thera-SAbDab database contains only sequences of variable regions. The IMGT database contains full-length sequences, but in a format which was difficult to download and analyze because the sequences are separated into functional regions. Therefore, we only compared the IMGT sequences of variable regions. The Inxight database contained contiguous full-length protein sequences. It included sequences from 44 INNs that were not provided in the INN lists. They were compared with sequences published in the patent literature. No perfect matches were found for lemasolasmab (H + L chains), pexelizumab (scFv), tacatuzumab (H chain) or tucotuzumab (H chain) and several others matched only a few patent families. These are all included in our analysis of Fc variants but flagged as "unverified". A substantial number of INN sequences are reproduced in patent application US2020/0023076. However, it includes numerous alternate sequences with no indication which is correct, and so was not considered to be a reliable guide. We considered two other databases which contain sequences of protein drugs, ChEMBL\textsuperscript{25} and KEGG,\textsuperscript{26} but they did not add any significant amount of additional data.

We noted 14 INNs that have had their names replaced, as shown in Table 2. Occasionally, the obsolete names had been used in other databases. We notified the respective curators and believe that these have now been updated. Our final consolidated list contained 894 INNs with sequence data available for 819 of them. The complete dataset is provided in the Supplementary Information (Supplementary Tables 1 and 2).

Applications and targets

Of the 819 immunoglobulin-related INNs with full sequences, 13 are for veterinary use (-vetmab), 2 are for imaging and 804 are for therapy. Our analysis will focus on the therapeutics. Most (89%) of these are antibodies with the -mab stem (719) or the newer -bart stem (5). There are 55 Fc fusions (28 -cept and 27 ef-) containing no targeting domain and 20 fusion proteins (-fusp) containing a targeting element. Five INNs fall outside these classifications: asfotase; blisibimod; dulaglutide; trebana-nib and melredableukin. Although we analyzed each INN independently, it should be noted that some are used in combination, especially in the field of anti-infectives. Table 3 shows the list of combinations of which we are aware.

A total of about 371 distinct human targets, 2 feline and 5 canine were identified in the complete listing of 894 INNs. The most common primary targets are listed in Table 4.

\begin{table}[h]
\centering
\caption{INNs with sequence data available from different databases.}
\begin{tabular}{ |l|c|c| }
\hline
Database & Total Sequences & INN lists Sequences not in INN lists \\
\hline
INN lists & 772 & n/a & n/a \\
Thera-SAbDab & 740 & 704 & 36 \\
IMGT & 676 & 661 & 15 \\
Inxight & 663 & 618 & 45 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\caption{List of obsolete INNs that have been replaced.}
\begin{tabular}{ |l|c|c| }
\hline
Old name & PL & New name & PL & RL \\
\hline
afutzumab & PL99 & obinutzumab & PL109 & RL69 \\
duligotumab & PL107 & duligotuzumab & PL110 & RL72 \\
fibatumumab & PL113 & ifabotuzumab & PL115 & RL77 \\
ratolimab & PL122 & ticemalimab & PL124 & RL86 \\
lambrolizumab & PL109 & pembrolizumab & PL110 & RL72 \\
lelandizumab & PL114 & olendalizumab & PL116 & RL78 \\
movezelinlb & PL121 & favezelinlb & PL123 & RL85 \\
nidainilimab & PL120 & nadunolimab & PL122 & RL84 \\
pavunalimab & PL123 & bavunalimab & PL125 & RL87 \\
pogalimub & PL114 & vonlerolizumab & PL116 & RL78 \\
sapeluzumab & PL114 & satralizumab & PL116 & RL78 \\
tavolxizumab & PL118 & tavolimab & PL118 & RL80 \\
vanalimb & PL118 & mitazalimab & PL119 & RL81 \\
zatuximab & PL107 & modotuximab & PL110 & RL72 \\
\hline
\end{tabular}
\end{table}
A total of 104 drugs had two binding targets (bispecific antibodies or antibody fusion proteins), two had three targets and one (emfizatamab) had four targets. CD3 was by far the most common secondary binding target for bispecific antibodies, with 41 examples.

### Classification of antibody-related INNs

Therapeutic biologics have undergone an explosion in design complexity with the generation of antibody fragments, Fc fusion proteins, and an ever-growing myriad of multispecifics. From personal experience we estimate that well over 200 antibody-related designs in preclinical use have been reported. However, there is no consistent nomenclature for these designs. Some are given names and potentially trademarked, such as BiTE 

Table 3. INNs used in combinations of two or three antibody-related drug substances. The list is arranged alphabetically. It may not be comprehensive as the status of clinical trials and regulatory approvals is constantly changing.

| INN 1 | INN 2 | INN 3 | Note |
|-------|-------|-------|------|
| actoxumab | bezlotoxumab | | Trialed for treatment of *C. difficile* infection, but only bezlotoxumab approved by FDA |
| amubarvimab | romlusevimab | | Approved in China for treatment of Covid-19 |
| altitivimab | mafitivimab | odesivimab | Approved by FDA for treatment of Ebola virus infection |
| bamlanivimab | etsevimab | | EUA granted for treatment of Covid-19 |
| atidortoxumab | berlativimab | | In development for treatment of S. aureus infection |
| casirivimab | imdevimab | | EUA granted for treatment of Covid-19 |
| cligavimab | tixagevimab | | EUA granted for treatment of Covid-19 |
| cosfoxivimab | lacraviXimab | porgaviximab | Approved by FDA for treatment of Ebola virus infection |
| crexavimat | ogalvimat | | In development for treatment of Covid-19 |
| dafsolimab | frivimab | | In development for treatment of hepatitis B virus infection |
| dacarivimab | miromavimab | | Approved in India for treatment of rabies |
| exkvivimab | librivimab | | In development for treatment of influenza A virus infection |
| futuximab | modotuximab | | In development for treatment of colorectal cancer |
| mazorelivimab | zamerovimab | | In development for treatment of rabies |

Table 4. Top targets for antibody drugs.

| Target | Number of INNs |
|--------|----------------|
| PD-1 | 39 |
| EGFR | 22 |
| SARS-CoV-2 | 22 |
| PD-L1 | 21 |
| CD20 | 19 |
| TNF alpha | 14 |
| HER2 | 13 |
| CD40 | 12 |
| CD19 | 11 |
| PCSK9 | 10 |
| VEGF | 10 |

| Name | Number of INNs |
|-------|----------------|
| Fc | 1 |
| IgG | 1 |
| Colored domain | 1 |
| Antigen | 1 |
| Antibody targeting | 1 |
| Antibody cross-over | 1 |

| Name | Number of INNs |
|-------|----------------|
| Young | 1 |
| Old | 1 |
| Focused | 1 |
| Non-focused | 1 |

| Name | Number of INNs |
|-------|----------------|
| Monospecific antibodies | 1 |
| Multispecific antibodies | 1 |

**Monospecific antibodies**

Monospecific antibodies are defined as molecules containing an immunoglobulin binding domain to only one epitope. There are 14 different formats in this class, as shown in Figure 1. The majority of therapeutic INNs can be categorized as IgGs, representing 77% (607/788). Twenty-three of the monospecific antibodies are fragments lacking an intact Fc domain, presumably to avoid effector function, reduce half-life, avoid receptor clustering and/or aid tumor penetration.

**Multispecific antibodies**

Multispecific antibodies are defined as molecules containing immunoglobulin binding domains to more than one antigen or epitope. Numerous reviews over the last five years have noted the growing number of designs for multispecific antibodies, covering IgG-like, fragments, and appended molecules in either symmetric or asymmetric design. For many years these were the plaything of protein engineers, but increasingly they are now entering the clinic with a total of 79 multispecific antibodies targeting the first, second, third or fourth antigen are colored in yellow, orange, red and purple, respectively. Constant domains are colored in dark blue, unless they are part of an asymmetric design utilizing heterodimerization mutations or domain cross-over in which case they are colored cyan and green. Additionally designs incorporating domain cross-over have a circular arrow symbol next to the green/cyan domains. Non-immunoglobulin proteins are shown as teardrop shapes colored in silver and bronze for the first and second types, respectively, of protein. T-cell receptor (TCR) is shown as a spring image with cyan and green coloring to represent heterodimerization of the α and β chains. Short peptides, including the J chain, are shown as green lines, with all other linker and hinge regions shown as black lines. Small chemicals are shown as black cartoon image of carbon rings. To minimize the complexity of the system heavy and light chains have not been colored differently. A more detailed description of each format is provided in Supplementary Table 3.
covering 25 different formats (Figure 2). Analyzing the first INN listing of each multispecific antibody shows a lag period after the first, blinatumomab, in 2008 followed by an upward trend from 2015 onwards (Figure 3). The number of new multispecific formats has also steadily risen over this period.

**Fc fusion proteins**

Fc fusion proteins are defined as non-immunoglobulin proteins fused to intact immunoglobulin Fc domains without the presence of an immunoglobulin targeting domain. There are 7 different design formats within our dataset, as shown in Figure 4, with the fusion of a single protein at the N-terminus of the Fc domain being by far the most common option (35/55). The first reported Fc fusion protein, lenercept, was followed by a lag period until the early 2000s since which there has been a steady increase in the number of INNs within this category (Figure 5).

**Antibody fusion proteins**

Antibody fusion proteins are defined as non-immunoglobulin proteins fused to immunoglobulin targeting domains, whole antibodies or fragments thereof. There are a total of 26 antibody fusion proteins covering 8 formats (Figure 6). In the last 5 years there has been a steady rise in the number of new INNs filed within this category (Figure 7).

**Miscellaneous formats**

There are six INNs that fall outside of our categorization, representing three different formats (Figure 8). Efgartigimod is an isolated human IgG1 Fc domain used for FcRn blockade. Efgartigimod is a small molecule chemically conjugated to an Fc domain. Efnopegudtide, efapagartim, eflaperogudtide and epsegomab are composed of peptides or proteins chemically conjugated to an Fc domain. Chemical conjugates to the Fc domain have not been included in the Fc fusion protein category, as this is intended for genetic fusions.

**Therapeutic antibody-related design complexity**

The listing of molecules that are not categorized as monospecific antibodies has increased dramatically since the early 2000s (Figure 9, Figure 10). In the early to mid-2000s (2001–2005), an average of 7% of listings were for non-monospecific antibodies, and these were exclusively within the Fc fusions category. In contrast, from 2018 to April 2022, non-monospecific antibodies are on average 18% of listings. This is largely due to the rise of multispecific antibodies, which in 2022 accounted for 10% of all antibody-related INN listings. More specifically, a comparison of traditional IgG molecules with all other formats demonstrates that in recent years approximately one-third of all listings are for non-IgG molecules (30% in 2020, 40% in 2021 and 33% up till April 2022). This large increase in non-IgG molecules is also reflected in the number of new formats, which has continued to accelerate from the early 2000s to date (Figure 11).
Although insightful, analyzing biological drug design at the macromolecular level masks the additional complexity that can be found at the protein sequence level. In particular, here, we focus on immunoglobulin constant domains covering isotypes, subtypes, allotypes and non-natural modifications of these.

**Figure 2.** Caption: Multispecific antibody formats. Below each image the number of occurrences of this format within the dataset is shown. Notes: Twenty-five cartoons showing different arrangements of protein domains represented as colored ellipses.
### Subtype and allotype

Analysis of allotypes in particular becomes complex in formats that do not contain all immunoglobulin constant domains. For this reason, we have restricted our analysis to therapeutic IgGs. There are 606 therapeutic IgGs for which full sequences are available: 409 human IgG1, 131 human IgG4, 48 human IgG2, 6 mouse IgG1, 4 human IgG2/4, 4 mouse IgG2b, 1 human IgG2/1, 1 human IgG1/2/4, 1 mouse IgG2a and 1 mouse IgG3. Of these, 6 are non-natural chimeras of human IgG constant domains to create a molecule containing preferable properties of human IgG1, 2 and/or 4. With respect to light chains, there are 516 human kappa, 79 human lambda, 10 mouse kappa and 1 mouse lambda.

Human sequences have been further analyzed for allotypic variants, as shown in Table 5. Human IgG3 has been omitted, as only one IgG3 is present within our dataset. For human IgG2 and 4 antibodies there is a clear preference for allotypes IGH2*01 and IGHG4*01, respectively. For human IgG1s there is more variation, with three allotypic variants making up 96% of the listings. The choice of allotypic variants for human light chains is less complex. Lambda light chains contain no natural allotopes, and for kappa all but one light chain are of the Km3 allotype.

Although the choice of allotype is generally believed to have minimal impact on quality, function or toxicity of the molecule, the choice of IgG subtype is a critical consideration with respect to interaction with Fcγ receptors and/or components of the complement pathway. In basic terms human IgG1 and IgG3 have higher ADCC and CDC activity, whereas IgG2 and IgG4 have substantially lower. Nevertheless, this has often been the basis for selection of IgG subtype there are also many Fc variants that have been reported with amino acid substitutions to further enhance or reduce this activity.

### Fc silencing

Antibody-related molecules can be silenced to limit antibody-mediated effector function by the use of formats lacking an immunoglobulin Fc domain or through the use of natural or modified Fc domains with reduced binding to Fcγ receptors and/or C1q. Of the 804 therapeutic antibody-related listings, 48 lack an intact immunoglobulin Fc domain, but from sequence analysis alone it is unclear whether this is for Fc silencing, half-life reduction, enhancement of tumor penetration, the avoidance of receptor clustering or some other
Table: Antibody fusion protein formats

| Format Description          | Number |
|-----------------------------|--------|
| IgG-protein fusion (type 1) | 12     |
| IgG-protein fusion (type 2) | 3      |
| IgG-protein fusion (type 3) | 1      |
| Fab-protein fusion (type 1) | 3      |
| Fab-protein fusion (type 2) | 1      |
| scFv-protein fusion         | 4      |
| Fv-protein fusion           | 1      |

**Figure 6.** Caption: Antibody fusion protein formats. Below each image the number of occurrences of this format within the dataset is shown. Notes: Eight cartoons showing different arrangements of protein domains represented as colored ellipses.

**Figure 7.** Caption: Analysis of the first reporting date for each INN and molecular format within the antibody fusion proteins category. The cumulative number of INNs is shown in Orange and formats in blue. Notes: Graph showing the increase in numbers of INNs and formats between 2008 and 2022. It shows a plateau from 2009 but the number of INNs increases rapidly from 2017.

**Figure 8.** Caption: Miscellaneous formats. Below each image the number of occurrences of this format within the dataset is shown. Notes: Three cartoons showing different arrangements of protein domains represented as colored ellipses.

reason. We therefore focused our attention on the 756 listings that contain an intact immunoglobulin Fc domain. Of these, 339 (44.8%) have reduced effector function, through the use of a human IgG2 or IgG4 Fc domain or a modified IgG1 Fc domain. Analyzed by their INN stems there are 304 -mab, 2 -bart, 9 -cept, 8 -fusp, 14 ef- and 2 others. This means that 44.9% (306/681) of antibodies (-mab and -bart), 46.0% (23/50) of Fc fusion proteins (-cept and -ef) and 53.3% (8/15) of fusion proteins (-fusp) are silenced. We also note that INN listings within the multispecific antibodies category are disproportionately more likely to be silenced, with 51/82 (62.2%) of all listings in this category silenced, rising to 51/67 (76.1%) when only considering those containing an immunoglobulin Fc domain.

There are a remarkable 49 different Fc variants that have been used to reduce effector function. Almost half of all silenced Fcs contain wild-type sequences with or without mutations that do not affect FcγR binding (36.2% IgG4 and 12.4% IgG2), with the remainder (51.4%) having additional mutations to reduce FcγR binding. To analyze trends in the use of Fc silencing, variants have been grouped into nine
categories as shown in Table 6. Trends in their usage are shown in Figure 12, where it is notable that wild-type or stabilized IgG4 continues to be widely used despite the catastrophic side effects attributable to unwanted FcyR binding which were seen with the superagonistic CD28 antibody TGN1412 in 2006.35

**Fc enhancement**

In some situations, it may be beneficial to enhance effector function over and above what can typically be achieved with a wild-type human IgG1. Twenty-eight of the therapeutic INN listings, 1 Fc fusion protein (conbercept), 1 bispecific antibody (ivicentamab) and 26 IgGs, contain IgG1-based Fc domains that have been modified to enhance Fcγ receptor and/or C1q binding. These can be broadly categorized as: 1) glycoengineered variants (9/28), through the use of cell lines where expression of enzymes in the glycosylation pathway has been altered leading to the production of antibodies with reduced levels of fucose that have enhanced binding to FcγRIIIα,36 or 2) Fc-modified variants (19/28) containing a variety of mutations to enhance binding to one or more receptors. Table 7 summarizes the variety of modifications that have been used.

**Half-life modifications**

IgG antibodies bind in a pH-dependent manner to FcRn and are efficiently rescued from catabolism, thus gaining an unusually long serum half-life. This is often deemed an advantage of Fc-containing biologics, but there have been a number of efforts to further modulate half-life through Fc engineering. Although we focus on Fc sequence variants here, we note that five of the INNs (alacizumab, certolizumab, daizolizumab, lulizumab, rivabazumab) are PEGylated fragments, a further five INNs (gefurulimab, isecarosmab, ozoralizumab, sonelokimab and vobarilizumab) contain albumin binding domains and one INN (imvotamab) is fused to human serum albumin, all of which are alternative means of half-life extension that do not use an immunoglobulin Fc domain.

Thirty-eight of the therapeutic INNs have Fc mutations to further extend half-life over and above what is achieved by a natural Fc domain. These cover six different Fc modifications, as shown in Table 8. Thirty-three of these use a human IgG1-based Fc domain with the other 5 using a human IgG4-based Fc.

One INN in our dataset (faricimab) includes the Fc mutations I235A/H310A/H435A to abolish FcRn binding, and thus reduce half-life.37 There is a further INN (efgartigimod) that contains the mutations M252/S254Y/T256E/H433K/N434F. These greatly enhance FcRn binding but lose pH dependency,38 thus these mutations are used for FcRn blockade.

**Heterodimerization**

The human IgG Fc domain is highly useful for cell killing through immune effector functions, half-life extension, ease of purification by Protein A or G and dimerization of immunoglobulin binding domains or other proteins, but the native sequence restricts construct design to symmetrical molecules. For multispecific antibodies or fusion proteins it can be advantageous to use asymmetrical designs to generate, for instance,
Table 5. Analysis of human heavy and light chain allotypes. Amino acids for IgG1 allotypes are shown in light blue, IgG2 in tan, IgG4 in light green and kappa in light yellow. For each allotype the total number of INN listings is reported.

| Domain | CH1 | CH2 | CH3 | CL | INN |
|--------|-----|-----|-----|----|-----|
| IgG1   | P   | S   | L   | K   | V   | L   | D   | L   | A   | K   | V   | A   | 173 |
| IGGF*01 | P   | S   | L   | R   | V   | L   | E   | M   | A   | K   | V   | A   | 134 |
| Other  | P   | S   | L   | K   | V   | L   | E   | M   | A   | K   | V   | A   | 85  |
| IGGF*03 | P   | S   | L   | R   | V   | L   | D   | L   | A   | K   | V   | A   | 11  |
| IGGF*07 | P   | S   | L   | K   | V   | L   | D   | L   | A   | K   | V   | G   | 5   |
| Other  | P   | S   | L   | R   | V   | L   | D   | L   | A   | K   | V   | G   | 1   |
| IGGF*04 | P   | S   | L   | K   | V   | L   | D   | L   | A   | K   | I   | A   | 0   |
| IgG2   |     |     |     |     |     |     |     |     |     |     |     |     |     |
| IGGH2*01 | P   | N   | F   | T   | V   | V   | E   | M   | A   | K   | V   | A   | 45  |
| Other  | T   | N   | F   | T   | M   | V   | E   | M   | A   | K   | V   | A   | 2   |
| Other  | T   | N   | F   | T   | V   | V   | E   | M   | A   | K   | V   | A   | 1   |
| IGGH2*04 | P   | S   | L   | T   | V   | V   | E   | M   | A   | K   | V   | A   | 0   |
| IGGH2*06 | P   | N   | F   | T   | V   | V   | E   | M   | S   | K   | V   | A   | 0   |
| IgG4   |     |     |     |     |     |     |     |     |     |     |     |     |     |
| IGHG4*01 | P   | S   | L   | R   | V   | L   | E   | M   | A   | R   | V   | A   | 129 |
| IGHG4*03 | P   | S   | L   | R   | V   | L   | E   | M   | A   | K   | V   | A   | 1   |
| Other  | P   | S   | L   | R   | V   | V   | E   | M   | A   | K   | V   | A   | 1   |
| IGHG4*02 | P   | S   | L   | R   | V   | V   | E   | M   | A   | R   | V   | A   | 0   |
| Kappa  |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Km3    |     | A   | V   | 515 |
| Km1,2 |     | A   | L   | 1   |
| Km1   |     | V   | L   | 0   |

Table 6. Types of Fc used to minimize binding to Fc gamma receptors. Most, if not all, of these variants also exhibit lower binding to C1q compared with wild-type IgG1.

| Category       | Description                                                                 | Number of INNs |
|----------------|-----------------------------------------------------------------------------|----------------|
| IgG4           | Wild-type IgG4 (26 INNs) or IgG4 with mutations (e.g., S228P) that do not affect binding to FcγR (97 INNs) | 123            |
| IgG4-FALA      | IgG4 with F234A/L235A mutations                                              | 17             |
| IgG4-other     | IgG4 with other mutations to reduce binding to FcγR                        | 16             |
| IgG2           | Wild-type IgG2 (32 INNs) or IgG2 with mutations that do not affect binding to FcγR (10 INNs) | 42             |
| IgG2-other     | IgG2 with other mutations to reduce binding to FcγR                        | 12             |
| IgG1-agly      | IgG1 with mutation to prevent glycosylation                                 | 32             |
| IgG1-LALA      | IgG1 with L234A/L235A mutations                                              | 28             |
| IgG1-other     | IgG1 with other mutations to reduce binding to FcγR                        | 64             |
| chimeric       | Hybrid molecules containing segments of IgG1, IgG2 and/or IgG4              | 6              |

Figure 12. Caption: Cumulative use over time of different groups of variants for Fc silencing. Notes: Graph showing the increase in use of different Fc variants between 1998 and 2022. It shows a steady increase in most variants since about 2007 but a steeper rise in IgG4 since 2017.
IgG-like structures with each Fab arm binding different antigens or molecules with different numbers of binding sites to each antigen. To minimize the occurrence of side-products, efficient generation of such heterodimeric molecules requires the introduction of mutations in the CH3 domain to drive correct heavy chain pairing and, where two different light chains are present, in the CH1 and CL domains to drive correct light chain pairing. (We found nine examples of IgG bispecific antibodies having two heavy chains but only a single light chain.)

The first reported method to achieve efficient heavy chain pairing was so-called knobs-into-holes mutations, formed by replacement of small amino acids with large on one CH3 domain, creating a knob, and large amino acids with small on the opposite side, creating a hole.24 Other similar mutations followed this principle of creating steric hindrance and thus preferential pairing. Alternatively, charged amino acids have been introduced or switched at the CH3:CH3 interface to drive heterodimerization through charge pairing, often referred to as electrostatic steering. The third mechanism of generating heterodimers is through controlled Fab arm exchange. Molecules are generated separately, but contain mutations that drive efficient heterodimerization when the molecules are mixed together under mildly reducing conditions.39,40 We also note that additional mutations are occasionally included to alter the charge or Protein A binding capability of one or both chains to enable more efficient separation of homo- and hetero-dimeric species. The variants used for heavy chain heterodimerization are shown in Table 9.

Antibody engineering solutions to the light chain pairing problem are far fewer in number (though the problem does not arise with controlled Fab arm exchange). Domain cross-over11 is the dominant solution, where CH1 and CL domains are swapped onto the opposite chains for one heavy-light pair. There are 7 INNs which use this approach; 6 are multispecific antibodies and one is an IgG-protein fusion. Of these, 3 (cibisamatab, faricimab and vanucizumab) just use domain cross-over to favor correct pairing, whereas 4 (alnuctamab, glothitamab, trontinemab, ensomafusp) have additional mutations in the CL (E123R/Q124K) and CH1 (K147E/213E) in the non-crossed domains to further enhance efficient pairing.42 There is one INN (reozalimab) that uses mutations without domain cross-over to achieve efficient light chain pairing. Specifically, these mutations are S131R/V133G/S176R in the CL and L128E/K147T/Q175E in the CH1.43

Tuvonralimab is an exception in that it contains mutations for heavy and light chain pairing (HC mutations K147D/
F170C/V173C/C220G/R255K/D399R/K409E; light chain mutations S131K/Q160C/S162C/C214S), but not for the generation of a bispecific antibody. The final drug substance is a mixture of two monospecific IgGs, iparolimab and tuvonralimab. The two antibodies are produced from the same cell line as a means of reducing manufacturing costs, with the mutations within tuvonralimab ensuring efficient pairing of the correct chains. Iparolimab contains no mutations.

**Site-specific conjugation**

Antibody-drug conjugates (ADCs) are a rapidly growing class of biologics, combining the specificity of an antibody with the potency of a toxic or radioactive payload. We identified 72 ADCs, but most (65/72; 90%) contain no modifications to the protein sequence to optimize conjugation, and thus are presumably conjugated via surface-exposed lysines, cysteines in the hinge region or the Fc domain glycans. Seven of the ADCs contain constant domain mutations for site-specific conjugation, as shown in Table 10. These mutations are all present in the heavy chain of human IgGs. Although site-specific conjugation of light chains has been reported, we have not identified any within our INN dataset. The most common approach to site-specific conjugation is the substitution of a surface-exposed amino acid with cysteine. Alternatively, two antibodies (ispектамаб and лувелтамаб) incorporate the non-natural amino acid para-azidomethyl-l-phenylalanine (pAMF) at position 180.

**Stabilization**

A substantial number of the antibody-related INNs contain mutations to enhance stability of the molecule. This is most prominent with human IgG4, with 84.6% (126/149) of the IgG4-based molecules containing the S228P mutation to prevent Fab arm exchange. Additionally, R409K and L445P mutations are used in a small number of INNs to further prevent Fab arm exchange and/or reduce acid-induced fragmentation and aggregation of IgG4s.

Aside from IgG4s, 10 IgG1-based INNs have been reported with mutations R292C/V302C, which introduce an additional intrachain disulfide into the CH2 domain. These mutations are always used in combination with the aglycosylation mutation N297G. Aglycosylation typically reduces the thermal stability of antibodies by as much as 10°C and introduction of the additional disulfide counteracts this.

One IgG (regdanvimab) contains mutations S152G/T163K in the lambda light chain, which are reported to improve stability of lambda constant domains.

**C-terminal truncation**

The C-terminal lysine, EU number 447, of IgG Fc domains is highly prone to cleavage by carboxypeptidases during manufacturing processes. This creates product heterogeneity, and thus some developers prefer to remove this lysine at the sequence level, especially since the C-terminal lysine is also quickly removed by carboxypeptidases in vivo. According to the INN lists, 111 of the 606 IgGs contain a C-terminal truncation. A total of 107 consist of the removal of K447 alone, while 4 have both G446 and K447 removed. However, we noted numerous discrepancies between the INN lists and the Inxight dataset regarding the presence or absence of a C-terminal lysine, suggesting that manufacturers may not always have reported this consistently.

**Discussion**

The design complexity of biological drugs has greatly increased over the 40 years since the first recombinant protein, insulin, was approved by the Food and Drug Administration (FDA) in 1982. There have been numerous reviews on the designs of preclinical molecules or of the relatively small number of molecules that have reached approval or late-stage trials, but we are unaware of an analysis on complexity across the whole clinical pipeline for antibody-related drugs. Although our dataset only accounts for antibody-related molecules that have been filed with the WHO for an INN, we believe this captures a large proportion of all clinical stage candidates. However, there are doubtless many others which are early in development and have not yet received an INN or have been abandoned and will never be assigned an INN.

Our dataset includes at least 378 different targets, 57 molecular formats and about 90 different variants (47 for silencing, 14 for enhancement, 8 for modifying binding to FcRn, 13 for heterodimerization, 4 for site-specific modification and 4 for stabilization). The ingenuity of antibody engineers does not seem to be exhausted and the range of clinical applications for antibody-related drugs continues to grow. However, in the absence of systematic studies to directly compare the effectiveness of formats and variants, it is hard to know which of the many alternatives to select. In a future study, we will begin to address this by synthesizing and comparing the biological activity of a matched set of antibodies with identical variable regions and different constant region variants.

While our analysis was in progress, a recent publication described a new notation language for antibody-based drug formats. It provides a systematic way to describe many possible formats and is well suited to provide the kind of definitions required for international regulatory purposes, including INNs. Software is also provided which can convert the rather complex-looking notation into visual representations similar to those we have proposed. In view of this parallel and rather more comprehensive development, we do not propose to provide further updates to our own classification system.

Accurate sequence listings are a vital prerequisite for researchers who wish to recreate and study the properties of antibody-related drugs. Patent literature can be a rich source of information, but is extremely difficult to use because of the vast number of
different sequences and inconsistent naming conventions. In contrast, the WHO INN lists appear to provide definitive protein sequences of drugs which actually are used in the clinic. For many years, however, these documents have been difficult to use because the sequences were not presented in a consistent machine-readable format and the descriptive information cannot be easily tabulated and analyzed. Secondary databases such as Thera-SAbDab, IMGT/2Dstructure-DB, Inxight, and The Antibody Society’s therapeutics database are therefore extremely valuable sources of data which have been curated and are freely accessible. However, none of them contained all the antibody-related sequences included in the INN lists and there were also a number of discrepancies in the sequence information. Some might be caused by errors or inconsistencies in the original deposition of sequence data to different authorities. For example, we noted numerous discrepancies between the WHO lists and the Inxight database (which uses US FDA sources) with regard to the presence or absence of a C-terminal lysine residue. We had no way of knowing which was the correct version. But in other cases, discrepancies could be resolved by careful inspection of the original INN lists to identify database anomalies that were probably caused by errors in transcription or optical character recognition. Such anomalies were notified to the respective database curators and have since been corrected where appropriate. We believe that our dataset is now the most complete and accurate, but it is unlikely to be entirely free from sequence errors and therefore users are well advised to make their own checks before embarking on a substantial project.

The focus of this work is on the structure and function of immunoglobulin constant domains, to the extent that it is determined by their amino acid sequence. This study represents a snapshot in time of the diversity of antibody-related therapeutics. We do not plan to provide regular updates and therefore have not included dynamic metadata such as disease indications, stage of clinical trials or regulatory status which is already available elsewhere. But we hope that this catalog of antibody formats and variants will be useful to designers of the future generation by exemplifying the range of possibilities.

**Methods**

**Nomenclature**

The EU numbering system is used throughout this article to describe variants in immunoglobulin constant domains. Amino acid alterations are described thus: XnnnY, where X is the single letter code for the residue in the native amino acid sequence, nnn is the EU index position and Y is the single letter code for the replacement amino acid residue. The symbol Δ refers to a residue which is deleted.

**Data sources**

The primary source for protein sequences was the lists of proposed and recommended INNs, published by WHO and downloaded as .pdf files from https://www.who.int/teams/health-product-and-policy-standards/inn/inn-lists. In addition, some INN sequences were obtained from individual .doc files at the School of INN, https://extranet.who.int/soinn/. Secondary sources were: 1) the Therapeutic Structural Antibody Database (Thera-SabDab), IMGT/2Dstructure-DB and IMGT/2Dstructure-DB Query page (Domains and sequence alignment) of IMGT downloaded as a .mhtml file and extracted into Microsoft Excel on 28 Mar 2022; 2) the IMGT/3Dstructure-DB and IMGT/2Dstructure-DB Query page (Domains and sequence alignment) of IMGT downloaded as a .mhtml file and extracted into Microsoft Excel on 28 Mar 2022; 3) the EMD-DB, Inxight and the Antibody Society’s database of antibody therapeutics in late-stage clinical studies at www.antisodyssey.org/antibodies-in-late-stage-clinical-studies/ accessed on 1 Jun 2022. Tertiary sources were the protein sequences found using PatSeq Finder at the Lens database, available at https://www.lens.org/lens/bio/patseqfinder (accessed 6 May 2022) and the sequence listing of patent application US2020/0023076, available at https://seqdata.uspto.gov/?pageRequest=docDetail&DocID=US20200023076A1 (accessed on 6 May 2022).

**Data processing**

We attempted to identify all monoclonal antibodies and Fc fusion proteins with sequence listings in the WHO INN lists (i.e., molecules containing some component of an immunoglobulin). These included INNs with the suffixes -mab, -cept, -fus, -cog, -bart, the prefix ef- and a few others. Since we focused on protein sequences, additional terms indicating glycoforms, drug conjugates and the like were removed. Where possible, sequences were extracted directly from the plain text of the INN .pdf or individual .doc files, unwanted spaces and numbers were stripped out and the contiguous sequences were assembled in an Excel spreadsheet. However, many of the original sequences were present as embedded images and could not be read directly. To extract these sequences, the pdf files were converted to text using the optical character recognition (OCR) function of Adobe Acrobat Pro DC (64 bit). This process was by no means free from error. Some of the recognized text appeared visually to be correct, but contained non-alphabetical characters. Sometimes characters were misread (e.g., F for VV or M for AA) and sometimes the sequences were misaligned due to erratic insertion of text boxes. To identify and correct errors, the sequences were aligned and compared with corresponding sequences from Thera-SabDab, IMGT and Inxight databases. Discrepancies were checked by manually inspecting the original documents and errors in our transcribed INN sequences were corrected. Discrepancies between the INN lists and published databases were flagged to the respective database curators. Some entries in the IMGT database had been assigned to an obsolete or unofficial INN; these have now been corrected. The Inxight database contained a sequence listing for edrecolomab which was a duplicate of nimotuzumab; it has now been removed. We have been informed by the curators that other sequence discrepancies have been checked and corrected where necessary. We understand that the Inxight database is being superseded by the GSRS (Global Substance Registration System) software.
and database produced by NIH/NCATS (see https://gscs.ncats.nih.gov/#!). When a complete sequence was only available in one list, it was used as a probe to search for identical sequences in the Lens patent database using the default BLAST options with a maximum of 500 hits. The hits were filtered to count the number of exact matches (100% coverage, 100% identity) and if there were few, then close matches (100% coverage, 98% identity) (Supplementary Table 4). A substantial number of exact matches was taken to indicate a higher probability that the probe sequence was correct. However, the lack of a complete match in patents did not necessarily show that the sequence was incorrect since it is quite common for only the sequence of variable regions to be reported. The final data were assembled and checked in an Excel spreadsheet using Microsoft Excel 2016 MSO version 2.204 32-bit.

**Identification of targets**

There is no uniform system for naming of binding targets in the INN lists. Generally, a number of alternative names for the targets are provided and some or all of them may be used in the various secondary databases. Starting from the target designation used by The Antibody Society, we manually curated the data to ensure that nomenclature was as consistent as possible, indicating the target molecule without regard to fine specificity. For Fc fusion proteins, the target molecule was identified (where possible) as the binding partner of the chimeric protein. Five Fc fusion proteins were identified by their disease target: hemophilia (efanesoctocog, efumorococog, efotrenacog, efanesoctocog) and hypophosphatasia (asotase).

**Characterization of immunoglobulin structures**

Sequences were classified according to the composition and order of globular protein domains and/or peptide sequences, including linkers, and designated a particular “format”. The format indicates the macromolecular domain assembly, thus sequence variations due to specificity (i.e., complementarity-determining regions), allotype or other Fc mutations do not alter the format.

| Species | Isotype | Probe(s) |
|---------|---------|----------|
| Human   | IgG1    | ASTKGPSVFPLAPSSKTSGG and EPKSCDKTHT |
| Human   | IgG2    | ASTKGPSVFPLAPCSR and ERKCCVECPFC |
| Human   | IgG3    | ASTKGPSVFPLAPCSRSTSGG and ELXTPLGDTTH |
| Human   | IgG4    | ASTKGPSVFPLAPCSR and ESXYGPP |
| Mouse   | IgG1    | AKTPPSYPLAPSGS and VPRDCGCK |
| Mouse   | IgG2a   | AKTTAPSVPLAPVC and EPRGPTIK |
| Rat     | IgG1    | AETTAPSVPLAPGT and VPRNCGDCX |
| Rat     | IgG2b   | AGTTPPSYPLAPGCS and ERRNGGGHCK |
| Dog     | IgG1    | ASTTAPSVPLAPSC, PASNTRVKDPFNSEC and VFNECRCTD |
| Dog     | IgG2    | ASTTAPSVPLAPSC, PASKTRVKDPPKRE and VPKRENGR |
| Human   | Kappa   | RTVAAPSVF |
| Human   | Lambda  | OPKAAAPSVTL |
| Mouse   | Kappa   | RADAAPTVSIFPPSSEQTLS |
| Mouse   | Lambda  | OPKSSPSVTLPPSESEELET |
| Rat     | Kappa   | RADAAPTVSIFPPSTEQLATGG |
| Rat     | Lambda  | QPKSTPPTLTV |

**Identification of constant domain variants**

Sequences containing immunoglobulin constant domains were isotyped using short protein sequences as probes to enable rapid identification in Excel. Unique probes were designed in regions of the constant domains that are absent of allotypic variants and other common antibody engineering mutation sites (Table 11). Sequences that could not be isotyped in this way were analyzed using the NCBI protein BLAST tool available at https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins to find the closest germline match.

Further analysis of constant domain sequences was only performed for human IgGs, as these were deemed to be of most interest to therapeutic antibody development. Sequences were aligned to a germline sequence for the appropriate heavy or light chain subtype and any differences from germline were identified. Allelic variations were identified for the purpose of determining allotypes, but not classified as mutations. Alignment, allotype identification and mutation reporting was performed using a custom-built Excel tool.

Many molecules contain mutations for multiple purposes, such as for ablation of effector function and simultaneous extension of half-life. As far as possible we grouped mutations into their purpose and report them separately, creating nine categories: effector function enhancement; effector function reduction; half-life extension; half-life reduction; heterodimerization; site-specific conjugation; stabilization; C-terminal truncation; and other.

**Abbreviations**

ADC antibody-drug conjugate, ADCC antibody-dependent cell-mediated cytotoxicity, ADCP antibody-dependent cell-mediated phagocytosis, CDC complement-dependent cytotoxicity, dAb, domain antibody, EUA emergency use authorization, Fab fragment antigen binding, Fc fragment crystallizable, FcyR Fc gamma receptor, FcRn neonatal Fc receptor, FDA Food and Drug Administration, Fv fragment variable, IMGT International ImMunoGeneTics information system, INN International nonproprietary name, PEG polyethylene glycol, PL Proposed List, RL Recommended List, TCR T-cell receptor, scFv single chain Fv, WHO World Health Organization.
Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Geoff Hale https://orcid.org/0000-0001-9473-423X

References

1. Koch SSG, Thorpe R, Kawasaki N, Lefranc M-P, Malan S, Martin ACR, Mignot G, Pluckthun A, Rizzi M, Shubat S, et al. International nonproprietary names for monoclonal antibodies: an evolving nomenclature system. mAbs. 2022;14. doi:10.1080/19420862.2022.2075078.

2. World Health Organisation. Geneva, Switzerland. 2014. https://www.who.int/publications/i/item/who-emp-rht-ts-2019-1.

3. Jones TD, Carter PJ, Pluckthun A, Vasquez M, Holgate RGE, Hotzel I, Poppelwell AG, Parren PWHI, Enzelberger M, Rademaker HJ, et al. The INNs and outs of antibody nomenclature names. mAbs. 2016;8:1–9. doi:10.1080/19420862.2015.1114320.

4. Parren PWHI, Carter PJ, Pluckthun A. Changes to international nonproprietary names for antibody therapeutics 2017 and beyond: of mice, men and more. mAbs. 2017;9:898–906. doi:10.1080/19420862.2017.1341029.

5. World Health Organisation. Switzerland, Geneva. 2017. https://cdn.who.int/media/docs/default-source/international-nonproprietary-names-(inn)-64th-executive-summary.pdf?sfvrsn=985b862c_8&download=true.

6. Balocco R, Sdg K, Thorpe R, Weisser K, Malan S. New INN nomenclature for monoclonal antibodies. Lancet. 2022;399:24. doi:10.1016/S0140-6736(21)02732-X.

7. World Health Organisation. Switzerland, Geneva. 2021. https://www.who.int/publications/i/item/inn-21-331.

8. World Health Organisation. Geneva, Switzerland. 2018. Notes from the fusion protein Working Group https://www.who.int/publications-m/en/inn-17-414-rev.

9. Raybould MJ, Marks C, Lewis AP, Shi J, Buijstrek A, Taddese B, Deane CM, Thera-SAbDab: the therapeutic structural antibody database. Nucleic Acids Res. 2020;48:D383–D388. doi:10.1093/nar/gkz827.

10. Lefranc M-P, Giudicelli V, Ginestoux C, Bodmer J, Muller W, Bontrop R, Lemaître M, Malik A, Barbie C, Chaume D. IMGT®, the international ImMunoGeneTics database. Nucleic Acids Res. 2009;37:D1006–D1012. doi:10.1093/nar/gkn750.

11. Stohl WR. Current progress in innovative engineered antibodies. Protein Cell. 2018;9:86–120. doi:10.1007/s13238-017-0457-8.

12. Wang X, Mathieu M, Brezski RJ. IgG fc engineering to modulate antibody effector functions. Protein Cell. 2018;9:63–73. doi:10.1007/s13238-017-0473-8.

13. Saunders KO. Conceptual approaches to modulating antibody effector functions and circulation half-life. Frontiers Immunol. 2019;10:1296. doi:10.3389/fimmu.2019.01296.

14. Liu R, Oldham RJ, Teal E, Beers SA, Cragg MA. Fc-engineering for modulated effector functions – improving antibodies for cancer treatment. Antibodies. 2020;9:64. doi:10.3390/antib9040064.

15. Lund J, Winter G, Jones PT, Pound JD, Tanaka T, Walker MR, Artyumiuk PJ, Arata Y, Burton DR, Jefferis R, et al. Human Fc gamma RI and Fc gamma RII interact with distinct but overlapping sites on human IgG. J Immunol. 1991;147:2657–62.

16. Xu D, Allegre M-L, Varga SS, Rothermel AL, Collins AM, Pulito VL, Hanna LS, Dolan KP, Parren PWHI, Bluestone JA, et al. In vitro characterization of five humanized OKT3 effector function variant antibodies. Cell Immunol. 2000;200:16–26. doi:10.1006/cimm.2000.1617.

17. Bolt S, Routledge E, Lloyd I, Chatenoud L, Pope H, Gorman SD, Clark M, Waldmann H. The generation of a humanised, non-mitogenic CD3 monoclonal antibody which retains in vitro immunosuppressive properties. Eur J Immunol. 1993;23:403–11. doi:10.1002/eji.1830230216.

18. Umana P, Jean-Mairet J, Moudry R, Amstutz H, Bailey JE. Engineered glycoforms of an anti-neuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. Nat Biotechnol. 1999;17:176–80. doi:10.1038/6179.

19. Dall’Acqua WF, Woods RM, Ward ES, Palaszynski SR, Patel NK, Brewah YA, Wu H, Kiener PA, Langermann S. Increasing the affinity of a human IgG1 for the neonatal Fc receptor: biological consequences. J Immunol. 2002;169:5171–80. doi:10.4049/jimmunol.169.9.5171.

20. Medesan C, Mateosi D, Radu C, Ghetie V, Ward ES. Delineation of the amino acid residues involved in transcytosis and catabolism of mouse IgG1. J Immunol. 1997;158:2211–7.

21. Angal S, King DJ, Bodmer MW, Turner A, Lawson ADG, Roberts G, Pedley B, Adair JR. A single amino acid substitution abolishes the heterogeneity of chimeric mouse/human (IgG4) antibody. Mol Immunol. 1993;30:105–8. doi:10.1016/0161-5890(93)90432-B.

22. Harris RJ. Processing of C-terminal lysine and arginine residues of proteins isolated from mammalian cell culture. J Chromatogr A. 1995;705:129–34. doi:10.1016/0021-9673(94)90125-5.

23. Sutherland MSK, Walter RB, Jeffrey SC, Burke PJ, Yu C, Kostner H, Stone I, Ryan MC, Sussman D, Lyon RP, et al. SGN-CD33A: a novel CD33-targeting antibody-drug conjugate using a pyrrolobenzodiazepine dimer is active in models of drug-resistant AML. Blood. 2013;122:1455–63. doi:10.1182/blood-2013-03-491506.

24. Ridgway JBB, Presta LG, Carter P. ‘Knobs into holes’ engineering of antibody CH3 domains for heavy chain heterodimerization. Protein Eng. 1996;9:617–71. doi:10.1093/protein/9.7.617.

25. Gaulton A, Hersey A, Nowotka M, Bento AP, Chambers J, Mendez D, Mutowo P, Atkinson F, Bells LJ, Cibrian-Uhalte E, et al. The ChEMBL database in 2017. Nucleic Acids Res. 2017;45: D945–D954. doi:10.1093/nar/gkw1074.

26. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2016;44:D457–D462. doi:10.1093/nar/gkv1070.

27. Brinkmann U, Kontermann RE. The making of bispecific antibodies. mAbs. 2017;9:182–212. doi:10.1080/19420862.2016.1268307.

28. Labrijn AF, Janmaat ML, Reichert JM, Parren PWHI. Bispecific antibodies: a mechanistic review of the pipeline. Nature Rev. 2019;18:585–608. doi:10.1038/s41573-019-0028-1.

29. Dickopf S, Georges GJ, Brinkmann U. Format and geometrics matter: structure-based design defines the functionality of bispecific antibodies. Comput Struct Biotechnol J. 2020;18:1221–27. doi:10.1186/s40588-020-00506.

30. Stiegmaier J, Benjamin J, Nagerson D. Utilizing the BiTE (bispecific T-cell engager) platform for immunotherapy of cancer. Expert Opin Biol Ther. 2015;15:1093–99. doi:10.1517/17425985.2015.1041373.

31. Moore PA, Zhang W, Rainer GI. Application of dual affinity retargeting molecules to achieve optimal redirected T-cell killing of C-cell lymphoma. Blood. 2011;117:4542–51. doi:10.1182/blood-2010-09-306449.

32. Ulrichs P, Guglietta A, Dreier T, van Bragt T, Hanssens V, Hofman E, Vankerkhoven B, Verheespen P, Ongenae N, Lykkopiy V, et al. Neonatal Fc receptor antagonist efgartigimod safely and sustainably reduces IgGs in humans. J Clin Invest. 2018;128:4372–86. doi:10.1172/JCI97911.

33. Gerstein HC, Sattar N, Rosenstock J, Ramsundarbethtic C, Pratley R, Lopes RD, Lam CSP, Khurmi NS, Heenan L, Del
