Predictions of Antibody Biophysical Properties for Improving Drug Development

by

Yulei Zhang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Engineering)
in the University of Michigan
2021

Doctoral Committee:

Professor Peter M. Tessier, Chair
Assistant Professor Bryan Goldsmith
Associate Professor Anna Schwendeman
Associate Professor Greg Thurber

Yulei Zhang

yuleizh@umich.edu

ORCID iD: 0000-0002-8151-0867

Dedication

This work is dedicated to my family at home and at work.

Acknowledgements

I would like to thank Dr. Tessier, my dissertation committee members, and members of the Tessier lab for their help with the work. I would also like to thank Scholar Space, UoM Library and Rackham OARD for help in preparing this document. I also want to thank Peng-Kai Kao, Yixuan Chen, Pengji Zhou and Jianxin Liu to help me with my PhD life.

Table of Contents

List of Figures

List of Tables

Abstract

Antibodies are natural proteins that are central to the function of the immune system. With high affinity and specificity, antibodies serve as adaptor molecules between recognition of foreign invaders and recruitment of immune cells. Monoclonal antibodies (mAbs) are attractive as therapeutic because of their drug-like properties (high solubility, stability, and specificity as well as excellent pharmacokinetics). However, the process of discovering and developing antibody drugs is time-consuming and expensive. Majority of the lead molecules fail in the clinical trial and only a few of them can survive from the clinical trial and finally be considered as drug. As a result, identifying antibodies with drug-like properties at early stages of the discovery process can reduce the risk of their rejection at late stages of development as well as reduce the time and cost of drug development. Our goal is to develop physicochemical descriptors and machine learning models for identifying and improving antibodies with drug-like properties during early stages of clinical development based on sequence based and predicted structure-based features. We argued that antibody specificity is the most important property to identify drug-like molecules since the measurements for self-interaction and non-specific interaction displayed a strong statistical significance in segregating approved drugs and those in clinical trial Phase II and III. The identified descriptors were used to guide the antibody library design to further improve antibody specificity while maintaining or increasing the humanness. Chapter 1 describes the reason why antibodies are considered as potential class of therapeutics and the background and motivation of molecular assessment for drug development; Chapter 2 shows how we developed the chemical rules for

identifying monoclonal antibodies with high specificity defined by previously reported experimental measurements; Chapter 3 shows a discovery that antibody variants with mutations in complementarity-determining regions (CDRs) that reduce non-specific binding also display drug-like specificity using a large set of human library antibodies; Chapter 5 describes the summary of the findings and discusses future directions in this field.

Chapter 1 **Introduction**

Antibodies are natural Y-shaped protein that are central to the function of the immune system specially to recognize various of foreign invaders such as bacteria and virus and recruit potent effector functions after target recognition (1). Natural antibodies are composed by two identical paired heavy and light polypeptide chains linked by a disulfide bond. Both light and heavy chains contain constant domains and variable domains. There are one variable region (V_H) and three constant regions (C_H1, C_H2 and C_H3) in heavy chain. Similarly, there are one variable region (V_L) and one constant region (C_L) in light chain (2). The fragment crystallizable region (Fc), which is composed by two constant regions in heavy chain (C_H2 and C_H3), determine the mechanism for destroying invaders. Based on Fc structure and immune function, antibodies are divided into five major classes, Ig (Immunoglobulin)M, IgG, IgA, IgD and IgE. Among these isotypes, nearly 3/4 of the human serum antibodies are IgGs which is considered as the most common type of antibody in the human circulation. The antigen-binding fragment (Fab) are composed by two variable region (V_H and V_L) and two constant regions from both heavy and light chain (C_H1 and C_L). The tip of the Fab region is known as the complementarity-determining regions (CDRs), consisted of six most diverse peptide loops (three from heavy chain and three from light chain) and folded in a Greek key barrel shape (Figure 1-1B) (2). Antibody diversity is due to the tremendous number of V(variable)-D(diversity)-J(joining) genes combination, where two types of light chains (kappa and lambda) are governed by *V* and *J* gene segments and heavy chain is governed by combination of *V*, *D*, and *J* gene segments from different chromosomes (3).

Nowadays, antibodies are considered as one of the most promising class of therapeutic based on their many attractive properties, for example, high binding affinity, specificity for target molecules, their ability to recruit potent effect functions after target recognition. Encouragingly, antibodies aimed to bind different target antigen share similar expression and purification approach which make them more attractive in the therapeutics (1,4). Up to date, over 200 monoclonal antibodies are investigated in clinical trials and half of them have be considered as approved drug by the Food and Drug Administration (FDA) (5), including the best-selling drug, Adalimumab (6).

However, the process of drug development and evaluation in the clinic is an extremely time consuming and expensive process, especially for antibody drug. After the panel of lead antibody molecules identified in the wet lab, they will be further engineered, formulated, and then go through the clinical trials to test safety and therapeutic efficacy (Figure 1-2) (4). The approved

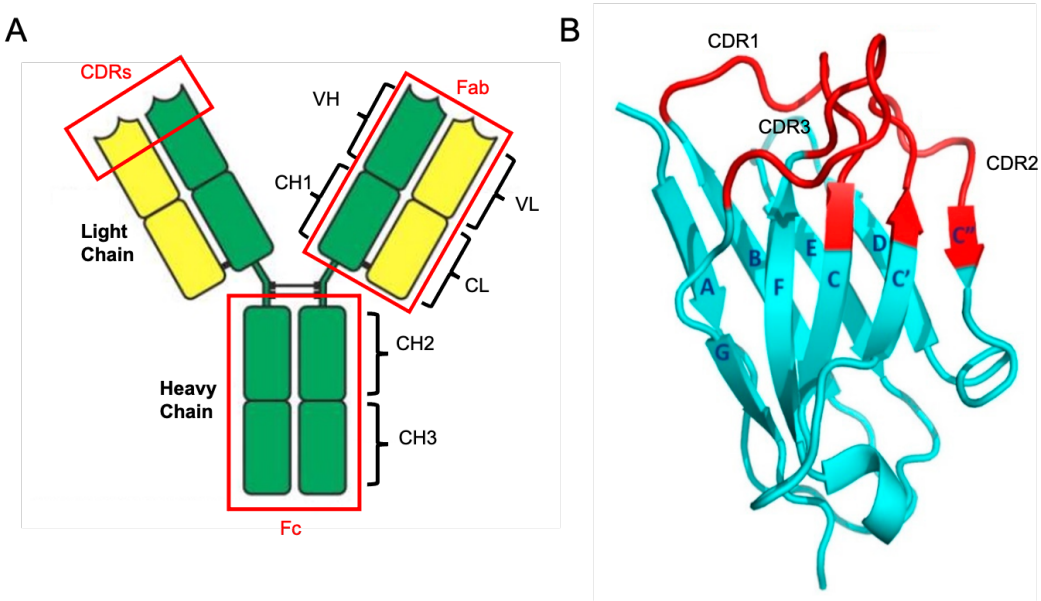


Figure 1-1: Structure of monoclonal antibodies and CDRs.
(A) Natural Ig (immunoglobulin) G1s are composed by two paired of heavy chains (green) and light chains (yellow). Heavy chain is composed by one variable region (V_H) and three constant regions (C_H1, C_H2 and C_H3). Light chain is composed by one variable (V_L) and one constant region (C_L). Antibody fragment Fc involves two heavy chain constant regions (CH2 and CH3) and Fab involves two variable regions (V_H and V_L) and two constant regions (C_H1 and C_L). CDR regions are located at the tip of the Y-shaped molecules. (B) CDRs are arranged in the Greek key barrel shape(2).

drug should contain both activity and developability properties (7). As a result, Optimizing the antibody development and identifying antibodies with drug-like properties at early stages of the discovery process can increase the likelihood of success at late clinical stages as well as reduce the time and cost. Back in 1990s, Lipinski brought up his idea about describing what kinds of small molecules have greatest chance to succeed in the clinic. It is the first time that people came up with descriptions about what meant to be drug-like with setting limits on both chemical and physical properties. It said "The orally acrive drugs can have no more than one violation of the following criteria: (1) less than or equal to 5 hydrogen bond donors; (2) less than or equal to 10 hydrogen bond acceptors; (3) less than 500 Daltons molecular mass; (4) less than 5 of log of octanol-water partition coefficient. The success of the Lipinski's rule inspired the idea of developing descriptors for identifying drug-like antibodies (8)". Although the details of the Lipinski's rules were argued time by time, the framework or the big picture approach that he drought up has been influence the field. Later, people developed more advanced methods to identify drug-like molecules, though they all followed the similar spirit. It also inspired we and others to develop rules based on chemical attributions, such as amino acid compositions, or physical properties, such as non-specific binding measurements that could guide us to identify drug-like antibodies.

In 2017, Adimab published a paper in *PNAS*, where the authors evaluated 137 monoclonal antibodies (mAbs) that are either approved drugs or in clinical trials Phase II and III (4). They measured various mAbs properties using twelve biophysical assays, including non-specific binding [poly-specific reagent assay (PSR) (9) using soluble membrane and cytosolic proteins and two ELISAs (10) using either baculovirus particles (BVP) (11) or a mixture of proteins, lipid and nucleic acid reagents], colloidal interactions [affinity-capture self-interaction nanoparticle spectrometry (AC-SINS) (12-14)and clone self-interaction by biolayer interferometry (CIS)(15)],

and thermal stability (Tm), etc (4,16). For each property, they established the upper limits based on the worst 10% of approved mAbs. The statistical significance of differentiating approved drug and those in Phase II and III for each assay was also investigated and reported in this work by logistic regression analysis. Out of the 12 assays, only the assays that measured non-specific binding and self-interactions displayed statistically significant manner in discriminating approved drug and those in clinical trials. This resulted in an argument that specificity is one of the most important properties to identify drug-like molecules. Antibodies with high specificity have increased likelihood to succeed in clinical trials. The statistical analysis in this work not only provided a possible approach to select or filter lead antibodies with drug-like properties but also inspired the research direction for generating physicochemical descriptors to reflect experimental measurements.

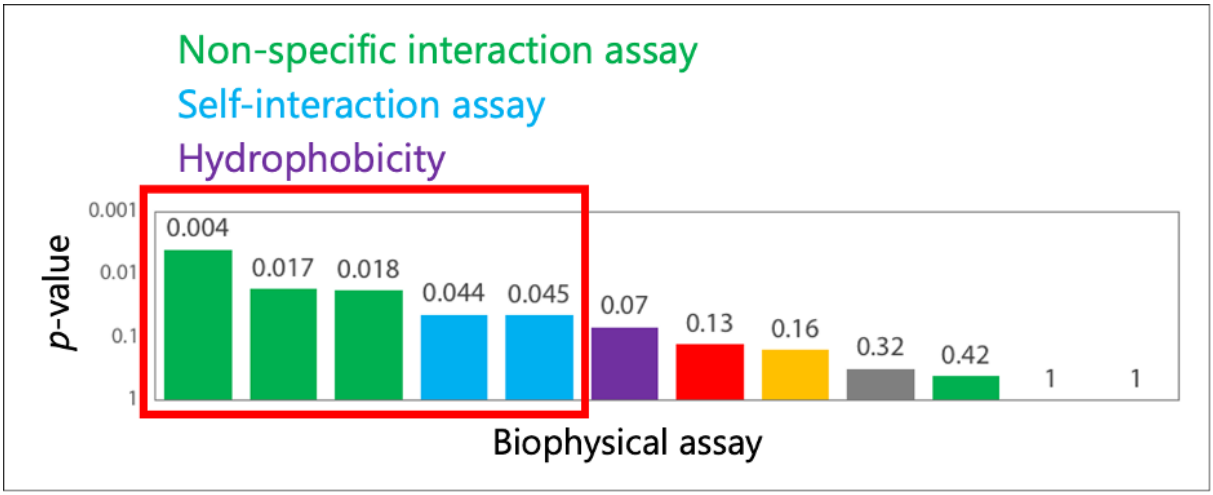


Figure 1-2: The statistical significance of the twelve reported biophysical measurement for differentiating approved drugs and antibodies in clinical trials Phase II and III.
Three assays that measure nonspecific binding (green) and two assays that evaluated antibody self-interactions (blue) displayed the most significant using logistic regression analysis. The lower the p-value represents the higher the statistical significance.

Antibody non-specific interactions and binding are the key reasons for high clearance rate and poor formulation (17-19). The high clearance is contributed by antibody non-specific interactions with cell surface instead of the target molecule. This resulted in highly uptake rate

therefore increase the chance for antibody get cleared in blood. Some other studies for antibody non-specific binding also declared that those antibodies are hardly released from the neonatal Fc receptor for IgG (FcRn) so that they are unable to be recycled back to the circulation and get degraded instead. Furthermore, antibody self-interactions lead to high viscosity when it is formulated at high concentration which is required for subcutaneous injection (20).

Given that specificity is one of the key properties and has been argued to be important in discriminating between approved drug and antibody in clinical trial Phase II and III, it is reasonable to develop physicochemical descriptors to identify antibodies with high non-specific interactions. According to the literature, electrostatic interactions was strongly correlated with antibody specificity. The evidence that high positively charged CDRs are enriched in early immune B cells suggest that non-specific interactions are strongly correlated to charged residues (21,22). And studies also showed that enrichment of negatively charged residues, especially in CDRs, not only contributed to increase solubility but also prevent aggregation (23). Other studies showed that amino acid motifs that composed by Trp and any of the four amino acids, Gly, Val, Trp and Arg, also show strong effects on increasing level of non-specific binding and interactions (24). These experimental results indicated that the composition of amino acids could potentially serve as physicochemical descriptors.

One of another key issues in the process of antibody development is immunogenicity reduction (25). To inhibit or weaken the immune respond, in other words, treating the antibody drug as foreign invaders after injection, various of antibody humanization approaches have been investigated, for example, the CDR-grafting techniques where CDRs from animal antibodies are grafted onto human FRs (human variable region acceptor frameworks) (26). The assumption for using CDR-grafting is to increase the global sequence similarity to human antibodies. Strategically,

the evaluation of humanness level based on global sequence identity and similarity to human serum antibodies is also important since CDRs are the most diverse region (27,28).

In this work, we have successfully generated chemical rules that can specifically flag monoclonal antibodies with drug-like specificity. The physiochemical descriptors in the chemical rules were involving information of composition of amino acids weighted by their predicted relative solvent accessible surface area. We then mutagenized one of the clinical-stage antibodies (emibetuzumab) that has been identified as a low specific mAb using experimental measurements. The library design for improving emibetuzumab specificity was guided by the chemical rules, aiming to reduce the number of problematic amino acids that enhance the non-specificity. The identified mutations that correlated with reducing non-specific interactions also display increasing level of humanness based on two humanization scores. We also developed a more general and robust machine learning method based on a larger dataset (>200,000) from a human library that encouragingly show strong prediction power to identify antibodies with low specificity.

References

1. Tiller, K. E., and Tessier, P. M. (2015) Advances in antibody design. *Annu Rev Biomed Eng* **17**, 191-216
2. Chiu, M. L., Goulet, D. R., Teplyakov, A., and Gilliland, G. L. (2019) Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies (Basel)* **8**
3. Sinkora, M., Sun, J., and Butler, J. E. (2000) Antibody repertoire development in fetal and neonatal piglets. V. VDJ gene chimeras resembling gene conversion products are generated at high frequency by PCR in vitro. *Mol Immunol* **37**, 1025-1034

4. Jain, T., Sun, T. W., Durand, S., Hall, A., Houston, N. R., Nett, J. H., Sharkey, B., Bobrowicz, B., Caffry, I., Yu, Y., Cao, Y., Lynaugh, H., Brown, M., Baruah, H., Gray, L. T., Krauland, E. M., Xu, Y. D., Vasquez, M., and Wittrup, K. D. (2017) Biophysical properties of the clinical-stage antibody landscape. *P Natl Acad Sci USA* **114**, 944-949

5. Umscheid, C. A., Margolis, D. J., and Grossman, C. E. (2011) Key concepts of clinical trials: a narrative review. *Postgrad Med* **123**, 194-204

6. Menter, A., Tyring, S. K., Gordon, K., Kimball, A. B., Leonardi, C. L., Langley, R. G., Strober, B. E., Kaul, M., Gu, Y., Okun, M., and Papp, K. (2008) Adalimumab therapy for moderate to severe psoriasis: A randomized, controlled phase III trial. *J Am Acad Dermatol* **58**, 106-115

7. Bailly, M., Mieczkowski, C., Juan, V., Metwally, E., Tomazela, D., Baker, J., Uchida, M., Kofman, E., Raoufi, F., Motlagh, S., Yu, Y., Park, J., Raghava, S., Welsh, J., Rauscher, M., Raghunathan, G., Hsieh, M., Chen, Y. L., Nguyen, H. T., Nguyen, N., Cipriano, D., and Fayadat-Dilman, L. (2020) Predicting Antibody Developability Profiles Through Early Stage Discovery Screening. *MAbs* **12**, 1743053

8. Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* **46**, 3-26

9. Xu, Y., Roach, W., Sun, T., Jain, T., Prinz, B., Yu, T. Y., Torrey, J., Thomas, J., Bobrowicz, P., Vasquez, M., Wittrup, K. D., and Krauland, E. (2013) 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* **26**, 663-670

10. Mouquet, H., Scheid, J. F., Zoller, M. J., Krogsgaard, M., Ott, R. G., Shukair, S., Artyomov, M. N., Pietzsch, J., Connors, M., Pereyra, F., Walker, B. D., Ho, D. D., Wilson, P. C., Seaman, M. S., Eisen, H. N., Chakraborty, A. K., Hope, T. J., Ravetch, J. V., Wardemann, H., and Nussenzweig, M. C. (2010) Polyreactivity increases the apparent affinity of anti-HIV antibodies by heteroligation. *Nature* **467**, 591-U117

11. Hotzel, I., Theil, F. P., Bernstein, L. J., Prabhu, S., Deng, R., Quintana, L., Lutman, J., Sibia, R., Chan, P., Bumbaca, D., Fielder, P., Carter, P. J., and Kelley, R. F. (2012) A strategy for risk mitigation of antibodies with fast clearance. *MAbs* **4**, 753-760

12. Sule, S. V., Sukumar, M., Weiss, W. F., Marcelino-Cruz, A. M., Sample, T., and Tessier, P. M. (2011) High-Throughput Analysis of Concentration-Dependent Antibody Self-Association. *Biophys J* **101**, 1749-1757

13. Geng, S. B., Cheung, J. K., Narasimhan, C., Shameem, M., and Tessier, P. M. (2014) Improving Monoclonal Antibody Selection and Engineering using Measurements of Colloidal Protein Interactions. *J Pharm Sci-Us* **103**, 3356-3363

14. Liu, Y. Q., Caffry, I., Wu, J. M., Geng, S. B., Jain, T., Sun, T. W., Reid, F., Cao, Y., Estep, P., Yu, Y., Vasquez, M., Tessier, P. M., and Xu, Y. D. (2014) High-throughput screening for developability during early-stage antibody discovery using self-interaction nanoparticle spectroscopy. *Mabs* **6**, 483-492

15. Sun, T., Reid, F., Liu, Y., Cao, Y., Estep, P., Nauman, C., and Xu, Y. (2013) High throughput detection of antibody self-interaction by bio-layer interferometry. *MAbs* **5**, 838-841

16. Estep, P., Caffry, I., Yu, Y., Sun, T. W., Cao, Y., Lynaugh, H., Jain, T., Vasquez, M., Tessier, P. M., and Xu, Y. D. (2015) An alternative assay to hydrophobic interaction

chromatography for high-throughput characterization of monoclonal antibodies. *Mabs* **7**, 553-561

17. Datta-Mannan, A., Lu, J., Witcher, D. R., Leung, D., Tang, Y., and Wroblewski, V. J. (2015) The interplay of non-specific binding, target-mediated clearance and FcRn interactions on the pharmacokinetics of humanized antibodies. *MAbs* **7**, 1084-1093

18. Piche-Nicholas, N. M., Avery, L. B., King, A. C., Kavosi, M., Wang, M., O'Hara, D. M., Tchistiakova, L., and Katragadda, M. (2018) Changes in complementarity-determining regions significantly alter IgG binding to the neonatal Fc receptor (FcRn) and pharmacokinetics. *Mabs* **10**, 81-94

19. Kelly, R. L., Yu, Y., Sun, T. W., Caffry, I., Lynaugh, H., Brown, M., Jain, T., Xu, Y. D., and Wittrup, K. D. (2016) Target-independent variable region mediated effects on antibody clearance can be FcRn independent. *Mabs* **8**, 1269-1275

20. Tomar, D. S., Kumar, S., Singh, S. K., Goswami, S., and Li, L. (2016) Molecular basis of high viscosity in concentrated antibody solutions: Strategies for high concentration drug product development. *Mabs* **8**, 216-228

21. Wardemann, H., Yurasov, S., Schaefer, A., Young, J. W., Meffre, E., and Nussenzweig, M. C. (2003) Predominant autoantibody production by early human B cell precursors. *Science* **301**, 1374-1377

22. Boughter, C. T., Borowska, M. T., Guthmiller, J. J., Bendelac, A., Wilson, P. C., Roux, B., and Adams, E. J. (2020) Biochemical patterns of antibody polyreactivity revealed through a bioinformatics-based analysis of CDR loops. *Elife* **9**

23. Datta-Mannan, A., Thangaraju, A., Leung, D., Tang, Y., Witcher, D. R., Lu, J., and Wroblewski, V. J. (2015) Balancing charge in the complementarity-determining regions of

humanized mAbs without affecting pI reduces non-specific binding and improves the pharmacokinetics. *MAbs* **7**, 483-493

24. Kelly, R. L., Le, D., Zhao, J., and Wittrup, K. D. (2018) Reduction of Nonspecificity Motifs in Synthetic Antibody Libraries. *J Mol Biol* **430**, 119-130

25. Harding, F. A., Stickler, M. M., Razo, J., and DuBridge, R. B. (2010) The immunogenicity of humanized and fully human antibodies: residual immunogenicity resides in the CDR regions. *MAbs* **2**, 256-265

26. Kettleborough, C. A., Saldanha, J., Heath, V. J., Morrison, C. J., and Bendig, M. M. (1991) Humanization of a mouse monoclonal antibody by CDR-grafting: the importance of framework residues on loop conformation. *Protein Eng* **4**, 773-783

27. Lazar, G. A., Desjarlais, J. R., Jacinto, J., Karki, S., and Hammond, P. W. (2007) A molecular immunology approach to antibody humanization and functional optimization. *Mol Immunol* **44**, 1986-1998

28. Gao, S. H., Huang, K. X., Tu, H., and Adler, A. S. (2013) Monoclonal antibody humanness score and its applications. *Bmc Biotechnol* **13**

Chapter 2 **Physicochemical Rules for Identifying Monoclonal Antibodies with Drug-like Specificity**

Abstract

The ability of antibodies to recognize their target antigens with high specificity is fundamental to their natural function. Nevertheless, therapeutic antibodies display variable and difficult-to-predict levels of non-specific and self-interactions that can lead to various drug development challenges, including antibody aggregation, abnormally high viscosity, and rapid antibody clearance. Here we report a method for predicting the overall specificity of antibodies in terms of their relative risk for displaying high levels of non-specific and/or self-interactions. We find that individual and combined sets of rules that limit the numbers of certain solvent-exposed residues in antibody variable regions, including both maximum limits (especially on the number of arginine residues) and minimum limits (especially on the number of aspartic acid residues), are strong predictors of specificity for large panels of preclinical and clinical-stage antibodies. These findings can be readily used to improve the selection and engineering of antibodies with drug-like specificity.

Introduction

Monoclonal antibodies (mAbs) are one of the most promising classes of therapeutics because of their many attractive properties, including their high affinity and specificity for target

molecules and their ability to recruit potent effector functions after target recognition (1,29,30). The generation and affinity maturation of mAbs involve introducing significant sequence variation in their six binding loops (complementarity-determining regions, CDRs) and, to a lesser extent, in their framework regions. Although maximal antibody sequence diversity is unimaginably large, the fraction of antibody sequences that give rise to mAbs with drug-like properties is expected to be dramatically lower. Natural filtering mechanisms used by the immune system eliminate many undesirable mAb sequences during antibody generation (21). However, antibodies generated by the immune system (as well as those discovered using *in vitro* display methods) are not optimized for the extreme requirements of many therapeutic applications (31,32). Indeed, several examples of poor physicochemical properties of mAbs have been reported that are linked to specific antibody sequences(23,32-38). Recent work suggests that high specificity is a key indicator of drug-like antibodies (4). Out of twelve biophysical assays of non-specific interactions, self-association, hydrophobicity and aggregation that were used to profile 137 clinical-stage antibodies, only assays that measured antibody non-specific interactions (three assays) and self-interactions (two assays) were able to identify approved antibody drugs as having superior biophysical properties relative to antibodies in phase 2 and 3 clinical trials. Nevertheless, it remains extremely challenging to identify the molecular determinants of antibody specificity for multiple reasons. First, antibody specificity is a relative concept that is dependent on the type of methods used to measure non-specific and self-interactions. Therefore, analysis of the molecular determinants of antibody specificity based on data obtained using a single type of assay may lead to conclusions that are not generally applicable to other types of antibody specificity measurements. Second, there has not been sufficient data available until recently (4,39) for detailed statistical analysis of the molecular determinants of antibody specificity. This new comprehensive data set provides several different

types of specificity measurements for a diverse panel of antibodies, which could enable a holistic analysis of the molecular determinants of antibody specificity for the first time.

In this work, we have sought to develop chemical (amino acid composition) rules that are able to identify mAbs with high, drug-like specificity and reduced risk of displaying high levels of non-specific and self-interactions. Our approach is to first segregate clinical-stage mAbs into two groups – namely those with low and high specificity – based on several different types of specificity measurements. Next, we have sought to develop chemical rules based on physicochemical properties of different regions in antibody variable fragments that are able to selectively identify mAbs with low specificity. This approach seeks to identify the most important chemical properties of the variable regions of antibodies linked to specificity in order to improve the identification and engineering of drug-like antibodies. Here we report individual and combined sets of chemical rules that are able to selectively identify mAbs with high specificity based only on antibody sequences and predicted site-specific solvent accessibilities.

Experimental Section

Antibody sequence and biophysical data. The amino acid sequences of the variable (V_H and V_L) regions of 137 clinical-stage antibodies and their corresponding measurements of non-specific and self-interactions were obtained from a previous publication (4). The amino acid sequences of a panel of preclinical antibodies were provided by Adimab (39). The relative solvent accessibilities of the clinical and preclinical antibodies were calculated using a Random Forest Regression method that was trained on over 900 antibodies in the Protein Data Bank (40). The maximum and minimum counts for each type of amino acid present in the CDRs and different regions of antibody variable regions weighted by their solvent accessible surface area for the clinical-stage antibodies are reported in Table 2-1. The similarities of mAbs relative to those in the

training sets are reported as the difference between 100% and the percentage of the 286 specified amino acid limits that are violated (Table 2-1). The preclinical mAbs have greater than 99% similarity relative to the clinical stage mAbs. The CDRs were defined using a combination of Chothia and Kabat numbering, and heavy chain CDR3 was defined to also include two additional N-terminal residues. The theoretical net charges of various antibody regions were calculated at pH 7.4 by assigning charges of +1 for Lys and Arg, +0.1 for His, and -1 for Asp and Glu.

Table 2-1: Maximum and minimum values for the observed counts of amino acids (weighted by their solvent accessibilities) and net charges (pH 7.4) of different regions within the variable domains of clinical-stage and preclinical mAbs in the training sets. Glycine is assumed to be fully exposed (SASA value of one).

The net charges were calculated by assigning values of +1 for R and K, +0.1 for H, and -1 for D and E. The CDRs were defined using a combination of Chothia and Kabat numbering, and heavy chain CDR3 was defined to also include two additional N-terminal residues. It is expected that the performance of the rules generated in this study will be highest for mAbs that do not violate any of these limits.

Max limits	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y	# of residues	charge
H1	0.81	0	1.12	0.64	0.78	6	0.62	1.11	0.73	0.28	0.10	1.28	0.61	0.02	0.68	2.01	1.54	0.63	0.41	2.02	12	1.1
H2	0.98	0	2.01	1.79	0.81	5	0.80	1.48	1.79	0.32	0.64	2.33	0.81	2.10	1.47	2.82	1.50	0.67	0.63	1.59	19	3
H3	1.15	0.05	1.88	0.70	0.93	4	0.76	1.07	0.92	1.67	0.90	0.81	0.79	0.60	1.13	1.67	1.20	1.83	1.06	2.07	23	3.1
H123	1.37	0.05	2.61	2.41	1.47	9	1.49	1.69	2.16	1.67	0.91	2.88	1.09	2.10	1.87	4.60	2.54	1.83	1.18	3.59	50	4.1
L1	0.82	0.02	2.43	0.80	0.68	4	0.92	1.30	1.74	0.61	0.01	1.51	0.61	1.31	1.73	3.28	1.37	0.84	0.47	1.07	17	3.2
L2	0.45	0	1.95	0.94	0.60	2	0.36	0.75	0.62	0.54	0.24	1.41	0.77	0.59	1.40	2.39	1.44	0.33	0.64	0.86	7	2.1
L3	0.68	0.05	0.78	0.85	0.49	4	0.67	0.69	0.53	0.92	0.50	0.97	0.72	0.57	0.69	1.30	1.03	0.59	0.47	0.95	12	2.1
L123	1.02	0.05	2.71	1.56	0.73	7	0.99	1.30	1.74	1.25	0.50	2.94	0.94	1.31	2.70	5.36	2.06	1.16	0.71	1.74	33	5.3
CDR	1.95	0.05	4.22	3.04	1.47	12	1.80	1.79	2.89	1.78	0.91	3.84	1.68	3.12	2.74	8.22	3.58	1.96	1.34	4.05	79	7.1
VH	3.39	0.05	3.63	4.34	1.47	18	1.49	2.18	4.73	2.65	1.02	4.11	3.47	4.68	3.84	11.50	7.26	2.82	1.47	3.80	130	7.1
VL	2.98	0.06	4.54	4.31	1.22	17	1.26	1.59	4.05	2.67	0.50	3.28	3.67	3.32	4.16	12.61	5.22	2.56	0.71	2.09	113	5.2
Fv	5.94	0.06	6.70	6.99	1.67	30	1.80	2.80	7.41	4.29	1.02	6.86	5.36	7.26	5.57	20.64	10.93	4.29	1.69	4.70	238	9.1
framework	4.66	0.03	4.43	5.51	0.99	19	0.88	1.59	6.37	3.37	0.64	4.02	4.75	6.70	4.52	15.17	9.43	3.39	0.54	1.52	161	5.1

Min limits	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y	# of residues	charge
H1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	-2
H2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	-3
H3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	-4
H123	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.19	32	-4.9
L1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.25	0	0	0	0	10	-4
L2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	-2
L3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	-2
L123	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.32	0	0	0	0	24	-5
CDR	0	0	0	0	0	1	0	0	0	0	0	0	0	9E-04	0	0.82	0.21	0	0	0.31	60	-7
VH	0.35	6E-05	0.29	0.30	0	7	0	0	1.17	0.02	0	0	0.53	1.17	0	5.26	1.25	0.27	0.01	0.29	112	-4.9
VL	0.13	3E-05	0.03	0.60	0.01	7	0	0.01	0.76	0.22	0	0	0.88	0.74	0.52	6.19	1.99	0.10	0.00	0.22	104	-5.9
Fv	0.75	3E-04	1.15	1.54	0.04	17	0	0.08	2.79	0.31	0	0	1.74	2.26	0.70	12.88	4.36	0.95	0.02	1.02	220	-5
framework	0.63	3E-04	0.10	1.54	0.01	13	0	0.01	1.85	0.31	0	0	1.55	1.90	0.70	8.78	2.92	0.74	0.01	0.13	158	-4

Chemical rules for identifying antibodies with low specificity. Rules for describing antibody specificity were calculated using the procedure described below and in the Supplemental Methods. First, the specificity of clinical-stage mAbs was experimentally evaluated using five specificity tests, which are maximum limits on the levels of non-specific interactions [>4.3 signal/background for baculovirus particle (BVP) binding, >0.27 for polyspecificity reagent (PSR) binding and >1.9 signal/background for ELISA] and self-interactions [>11.8 nm for affinity-

capture self-interaction nanoparticle spectroscopy (AC-SINS) and >0.01 response units for clone self-interaction by biolayer interferometry (CSI)](4). Second, each antibody was assigned to one of two groups based on its number of biophysical flags, as defined by the number of times an antibody exceeds the five maximum limits for non-specific and self-interactions. One group with <2 biophysical flags is defined as the high specificity group and a second group with ≥2 biophysical flags is defined as the low specificity group. The clinical-stage mAbs were assigned to the two groups based on their value of each individual biophysical measurement (BVP, PSR, ELISA, AC-SINS, and CSI) relative to the limits described above. The preclinical mAbs(39) were also assigned into two groups with high specificity (≤0.27 for PSR) or low specificity (>0.27 for PSR). Third, rules for the counts of amino acids in the CDRs and variable regions, weighted by their relative solvent exposure, were evaluated using the summed values (increments of 0.1) for various combinations of specific residues that spanned the values observed in the clinical-stage antibodies (as described in the Results section). An amino acid was considered solvent accessible if its relative solvent exposure was ≥10% (otherwise it was excluded from analysis), and glycine was assumed to be fully exposed.

Individual chemical rules for identifying antibodies with low specificity. The chemical rules were generated using threefold cross validation methods and were required to meet a number of constraints. First, the clinical-stage mAbs (137) were split into training (80%) and test (20%) sets in ten different ways using stratified sampling. The training sets were further divided into three partitions (folds), and two partitions were used for training and one for validation. Individual rules were required to satisfy the following constraints: i) adjusted accuracy (herein simply referred to as accuracy) of preclinical antibodies (training set) >55%; ii) % mAbs flagged with high specificity (preclinical, training set) < % mAbs flagged with low specificity (preclinical, training set); iii)

accuracy (clinical mAbs, training set) >55% in each fold (three constraints for each of the ten 80/20 splits); iv) difference between the accuracy of training (two folds) and validation (one fold) < 5% (three constraints for each of the ten 80/20 splits, clinical mAbs); v) % mAbs flagged with high specificity (defined by each individual assay) < % mAbs flagged with low specificity (defined by each individual assay; five constraints that were evaluated using the entire 80% of the clinical antibody training data for each of the ten 80/20 splits); vi) average validation accuracy for each of the ten 80/20 splits > 60% where the average validation accuracy is the average of the validation accuracy for all observed flag values within each 80/20 split; vii) average test accuracy for each of the ten 80/20 splits > 50% where the average test accuracy is the average of the test accuracies for all observed flag values within each 80/20 split. Finally, the rules were required to be observed in each of the ten 80/20 splits, although different values for the rules were allowed.

The significance of each rule for selectively flagging antibodies with low specificity was assessed using adjusted accuracies and 2x2 contingency tables (Fisher's exact test) for evaluating *p*-values. Adjusted accuracy was calculated by equally weighting the true and false positives with the true and false negatives. Because many simultaneous hypothesis tests were evaluated with a common procedure, the False Discovery Rate (FDR) method was used for individual chemical rules (for each antibody region) to balance the competing demands of sensitivity and specificity of the statistical analysis.

Combined rules for enhancing the identification of mAbs with low specificity. Sets of rules were generated by combining single rules together (up to six single rules per combined set were evaluated), as explained in the Results section. Each mAb was considered to have low specificity if flagged by four or five rules (as specified). Sets of rules in the first round of analysis were only accepted if they met several requirements: i) accuracy >60% (preclinical mAbs, training set); ii) %

mAbs flagged with high specificity <10% (preclinical mAbs, training set); iii) accuracy (clinical mAbs, training seet) >60% in each fold (three constraints for each of the ten 80/20 splits); iv) difference between the accuracy of training (two folds) and validation (one fold) <10% (three constraints for each of the ten 80/20 splits, clinical mAbs); v) flag <5% clinical-stage mAbs with high specificity in training sets (as defined combination of five assays); vi) % accuracy (clinical mAbs, training set) >60% (defined by each individual assay; five constraints that were evaluated using the entire 80% of the clinical antibody training data for each of the ten 80/20 splits) and vii) % accuracy (clinical mAbs, training set) >60% (defined by combination of five assays; one constraint was evaluated using the entire 80% of the clinical antibody training data for each of the ten 80/20 splits). Finally, the reuls were required to be observed in each of the ten 80/20 splits, although different values for the rules were allowed.

The significance of each rule for selectively flagging antibodies with low specificity was assessed using adjusted accuracies and 2x2 contingency tables (Fisher's exact test) for evaluating *p*-values. Adjusted accuracy was calculated by equally weighting the true and false positives with the true and false negatives. Because many simultaneous hypothesis tests were evaluated with a common procedure, the False Discovery Rate (FDR) method was used for individual chemical rules (for each antibody region) to balance the competing demands of sensitivity and specificity of the statistical analysis.

Combined rules for enhancing the identification of mAbs with low specificity. Sets of rules were generated by combining single rules together (up to six single rules per combined set were evaluated), as explained in the Results section. Each mAb was considered to have low specificity if flagged by four or five rules (as specified). Sets of rules in the first round of analysis were only accepted if they met several requirements: i) accuracy >60% (preclinical mAbs, training

set); ii) % mAbs flagged with high specificity 60% in each fold (three constraints for each of the ten 80/20 splits); iv) 2 difference between the accuracy of training (two folds) and validation (one fold) 60% (defined by each individual assay; five constraints that were evaluated using the entire 80% of the clinical antibody training data for each of the ten 80/20 splits); and vii) % accuracy (clinical mAbs, training set) >60% (defined by combination of five assays; one constraint was evaluated using the entire 80% of the clinical antibody training data for each of the ten 80/20 splits). These constraints are summarized in Table 2-2. Finally, the combined rules (with the same values for each rule) were required to be observed in each of the ten 80/20 splits. The best sets of combined rules in the first round of analysis were identified as those with the lowest coefficients of variation for the average validation accuracy (ten 80/20 splits). Next, mAbs that were not flagged as polyspecific in the first round of analysis ii) accuracy (clinical mAbs, training set) >75% in each fold (three constraints for each of the ten 80/20 splits); iii) difference between the accuracy of training (two folds) and validation (one fold) 70% (defined by each individual assay; five constraints that were evaluated using the entire 80% of the clinical antibody training data for each of the ten 80/20 splits); and vi) % accuracy (clinical mAbs, training set) >75% (defined by combination of five assays; one constraint that was evaluated using the entire 80% of the clinical antibody training data for each of the ten 80/20 splits). These constraints are summarized in Table 2-2. Finally, the combined rules (with the same values for each rule) were required to be observed in each of the ten 80/20 splits. The best five sets of combined rules in concert with Set A (Table 2-3) were identified as those with the lowest coefficients of variation for the average validation accuracy (ten 80/20 splits).

Table 2-2 Summary of the constraints used to generate the single and combined sets of rules for flagging mAbs with low specificity.

	Constraints for single rules	1st round	2nd round
1	Accuracy for preclinical mAbs (PSR assay)	>55%	>55%
2	% mAbs (preclinical, training set) flagged with low specificity - high specificity (PSR assay)	>0	>0
3	Accuracy of training set (clinical mAbs) in each fold	>55%	>55%
	* Three constraints for each of the ten 80/20 splits		
4	Accuracy of training set (clinical mAbs, two folds) - accuracy of validation set (one fold)	<5%	<5%
	* Three constraints for each of the ten 80/20 splits		
5	% mAbs (clinical, training set) flagged with low specificity - high specificity for each assay	>0	>0
	* Five constraints for each individual assay (PSR, ELISA, BVP, AC-SINS, CSI)		
6	Average validation accuracy for clinical mAbs	>60%	>60%
	* Average validation accuracy is evaluated based on the results for each flag value observed in each of the ten 80/20 sp		
7	Average test accuracy for clinical mAbs	>50%	>50%
	* Average test accuracy is evaluated based on the results for each flag value observed in each of the ten 80/20 splits		

	Constraints for combined rules	1st round	2nd round
1	Accuracy for preclinical mAbs (PSR assay)	>60%	>70%
2	% mAbs (preclinical, training set) flagged with high specificity (PSR assay)	<10%	
3	Accuracy of training set (clinical mAbs) in each fold	>60%	>75%
	* Three constraints for each of the ten 80/20 splits		
4	Accuracy of training set (clinical mAbs, two folds) - accuracy of validation set (one fold)	<10%	<10%
	* Three constraints for each of the ten 80/20 splits		
5	% mAbs (clinical, training set) flagged with high specificity	<5%	<10%
6	Accuracy of training set (clinical mAbs) for each assay (PSR, ELISA, BVP, AC-SINS, CSI)	>60%	>70%
	* Five constraints for each individual assay (PSR, ELISA, BVP, AC-SINS, CSI)		
7	Accuracy for training set (clinical mAbs)	>60%	>75%

Measurements of antibody non-specific and colloidal interactions. mAb variants (39 IgGs) with sequence differences in their frameworks and CDRs were expressed as IgG1 antibodies in the CHO-3E7 cell line (L-11992, National Research Council Canada) and purified via Protein A chromatography. Preparative size-exclusion chromatography was also performed (when necessary) to reduce the aggregate content below 5%.

Table 2-3: Summary of the best combined sets of rules generated in the first round of analysis that flag mAbs with low specificity. Five sets (A, B, C, D and E) of rules are reported.

Set of flags	# of flags for mAs with low specificity	Individual rule numbers for each set of flags						Flag values for corresponding individual rule of each set of flags					
A	≥4	1	11	14	21	24	37	5.0	4.0	4.6	4.7	12.2	4.9
B	≥4	7	11	15	19	21	33	3.6	4.0	4.2	3.0	4.8	17.8
C	≥5	9	13	19	21	23	33	3.6	7.5	3.0	4.8	2.1	17.8
D	≥4	7	11	15	19	21	33	3.6	4.0	4.2	3.0	4.8	16.7
E	≥5	3	7	11	21	23	33	4.7	3.6	4.0	4.8	2.1	17.8

	Clinical mAbs (non-specific and self-interactions, five assays)								
Set of flags	% mAbs flagged (high specificity group)	% mAbs flagged (low specificity group)	Average training accuracy (%)	Std. dev. of training accuracy (%)	COV for training accuracy (%)	Average test accuracy (%)	Test accuracy Stdev (%)	COV for test accuracy (%)	*p*-value for entire panel of clinical mAbs
A	2.06	35.00	66.17	0.87	1.3	67.43	3.68	0.05	3.78E-07
B	2.06	40.00	69.07	0.93	1.3	68.42	3.74	0.05	2.15E-08
C	2.06	40.00	68.48	0.96	1.4	70.92	3.91	0.06	2.15E-08
D	0.00	37.50	68.71	1.06	1.5	68.75	4.17	0.06	1.04E-09
E	3.09	47.50	72.18	1.22	1.7	72.01	4.98	0.07	1.31E-09

The levels of antibody non-specific binding were evaluated using an ELISA method reported previously with minor modifications. Immulon 2HB plates (3655TS, Thermo Fisher Scientific) were coated separately with six non-antigens as reported previously except that insulin was immobilized at 0.2 mg/mL (1 h at 37 °C). The plates were washed three times (0.2 mL/well) using PBST (PBS with 0.05% Tween 20) and were not blocked. Next, each mAb (1 μM in PBST) was added to the wells for 1 h. After washing the wells three times (0.2 mL/well of PBST per wash), the secondary antibody (HRP conjugated goat anti-human IgG antibody, 10 ng/mL; 109-

035-008, Jackson Immuno Research) was added (1 h). Finally, after removal of unbound secondary antibody, TMB substrate (TMBS-1000-01, Surmodics) was added, and the plates were developed (5 min) before quenching with 2 M sulfuric acid. The volume of the solutions added to the wells was 50 μL/well unless otherwise specified. The absorbance values were evaluated at 450 nm using a Biotek Synergy 2 plate reader, and the signal over background was calculated using background values evaluated without mAb and with all other reagents.

The levels of antibody self-interactions were measured for the mAbs using affinity-capture self-interaction nanoparticle spectroscopy (AC-SINS), as reported previously. Briefly, the nanoparticle conjugates were prepared by first adsorbing goat anti-human Fc polyclonal antibody and then co-adsorbing human mAbs and human polyclonal antibodies at different ratios (fixed total concentration of 20 μg/mL of human antibody). The reported plasmon shifts are averages of those evaluated at three different percentages of human mAbs (5, 15 and 25%). The control values used to calculate the plasmon shifts were those for 0% human mAb (100% human polyclonal antibody). The absorbance spectra used to evaluate the plasmon shifts were measured using a Biotek Synergy 2 plate reader.

Antibody sub-library design and sorting. Sites for mutation in the variable heavy (V_H) region of emibetuzumab were identified using the combined chemical rules. In particular, sites in the CDRs were targeted if they were i) flagged by the maximum limits (rules 1-6 in the combined rules), ii) hydrophobic or positively charged, iii) solvent exposed (>10% SASA) and iv) relatively uncommon at a given antibody site (<50%) in tens of thousands of human antibodies. The last requirement that the wild-type residue must not be highly conserved aims to avoid mutations at sites that are critical to antibody folding and/or stability. The resulting antibody library was generated with mutations at eight sites in V_H (Y33, R50, R54, R55, G56, A95, W97 and Y102)

Table 2-4: Design of a mutant VH library for single-chain Fab fragments of emibetuzumab aimed at reducing the number of chemical flags. Eight sites in the CDRs flagged by the maximum chemical rules in Fig. 4 were mutated using degenerate codons that are designed to sample the wide-type residue and five mutations that reduce the number of chemical rules.

	H1	H2				H3		
	33	50	54	55	56	95	97	102
Degenerate codon	KHY	RVR	RVR	RVR	RVY	KHY	KBG	KHY
WT	Y	R	R	R	G	A	W	Y
1	F	K	K	K	A	F	V	F
2	V	A	A	A	D	Y	L	V
3	A	G	G	G	N	V	A	A
4	D	E	E	E	S	D	G	D
5	S	T	T	T	T	S	S	S

that were aimed to reduce the number of chemical flags in the variable regions of emibetuzumab. For each mutation site, degenerate codons were designed to sample the wild-type residue as well as at least one negatively charged residue and one polar residue, as well as up to three additional residues with similar properties relative to the wild-type. For example, Y33 in V_H was mutated to Tyr (wild type), Phe (aromatic and hydrophobic), Val and Ala (hydrophobic), Ser (polar) and Asp (negatively charged). The total library size was 10^6 variants, and the library design is summarized in Table 2-4.

The initial rounds of sorting were conducted by incubating 10^9 (round 1) or 10^7 (round 2) surface-displaying yeast with 10^7 Dynabeads (Protein A, 10002D; Thermo Fisher Scientific) saturated with antigen (hepatocyte growth factor receptor as an Fc fusion protein, HGFR-Fc; MET-H5256, Acro Biosciences) in PBSB (PBS with 1 mg/mL BSA) and 1% milk for 3 h at room temperature. The final round of sorting (round 3) was completed via FACS (MoFlo Astrios,

Beckman-Coulter) using 10^7 cells following incubation with soluble antigen or polyspecificity reagents. Ovalbumin (Sigma, A5503) and soluble membrane proteins isolated from CHO cells (polyspecificity reagent or PSR) were biotinylated using Sulfo-NHS-LC-Biotin (Pierce, P121335; Thermo Fisher Scientific) and HGFR-Fc was used as purchased. Cells were incubated with ovalbumin (260 μg/mL), PSR (130 μg/mL), or HGFR-Fc (1 nM with 1% milk) for 3 h at room temperature (ovalbumin, HGFR-Fc) or 20 min on ice (PSR) in PBSB with an anti-myc tag mouse mAb (1:1000 dilution; 2276S, Cell Signaling Technologies). After one wash with PBSB, cells were incubated with secondary reagents to detect scFab display (1:100 goat anti-mouse Alexa Fluor 488; A11001, Life Technologies) and binding (1:1000 streptavidin Alexa Fluor 647, S32357, Life Technologies for ovalbumin and PSR; 1:300 goat anti-human Fc Alexa Fluor 647, 109605098, Jackson Immuno Research Labs for HGFR-Fc). Finally, the cells were washed with PBSB and sorted for positive display and non-binding to ovalbumin and PSR or binding to HGFR-Fc.

Deep sequencing and data analysis. The sorted antibody library samples were evaluated using deep sequencing by extracting the single-chain Fab (scFab) plasmids from yeast using the Zymoprep Yeast Plasmid Miniprep II Kit (D2004; Zymo Research). The V_H region of the scFab gene was amplified via two-step PCR using Q5 polymerase (M0491; New England Biolabs). The first reaction was performed using primers that were complementary to the V_H domain in addition to Illumina adapter sequences and barcodes (see Supplemental Methods for more detail). The PCR product was gel purified (1% agarose) and isolated using a QIAquick Gel Extraction Kit (28704; Qiagen, Germantown, MD). The second reaction used 2 μL of the purified PCR product with primers identical to the Illumina adapter sequences and was also gel purified following the manufacturer's recommendations. Concentrations of each sample were determined using a Qubit 4 Fluorometer (Q33240; Waltham, MA) and pooled together at an equimolar ratio. The pooled

samples were evaluated using deep sequencing (Illumina MiSeq in a 300 bp paired-end sequencing reaction). The raw sequencing files from Illumina MiSeq (300 bp paired-end sequencing reaction_ were merged together using BBMerge with ".fastq" file was converted to a ".fasta" file and analyzed without any 'N' base calls. The frequency of each set of mutations (herein referred to as a mutational string) was counted if the first residue of was incorrect ot there was a stop codon, the translation of the reverse complement was checked. The frequency of each mutational string was determined and exported into a ".csv" file for calculation of the enrichment ratios.

Next, the deep sequencing data were analyzed to identify antibody variants observed in four library samples in two different biological repeats of the third round of sorting, namely the i) input library and the samples sorted for ii) negative ovalbumin binding (OVA-), iii) negative PSR binding (PSR-), and iv) positive HGFR-Fc binding. From these two repeats, 3465 unique scFabs were identified that were present in all of the eight analyzed samples. To identify the mutations that are most strongly linked to high specificity, sets of one to four mutations were evaluated in the 3465 scFabs that were most strongly correlated with enrichment in the samples sorted for low non-specific binding. First, all possible combinations of mutations for the eight mutated sites were evaluated. Because the statistical significance of the sets with four mutations was found to be highest, we focused on these 43,750 mutational sets. Each mutational set [e.g., Y33F, R54T, G56D, and Y102A in V_H] was evaluated by first identifying the clones that contain such mutations (regardless of whether they have wild-type or mutant residues at other sites), which are referred to as the four mutant (4MT) group. Similarly, the clones with wild-type residues at the same four sites (regardless of whether they have wild-type or mutant residues at the other sites) were identified, which are referred to as the four wild-type residue (4WT) group. Only the 4MT/4WT

sets that contain more than ten clones in each group were further evaluated to maximize statistical

significance. Next, a Spearman's rank correlation coefficient was evaluated for each set of clones

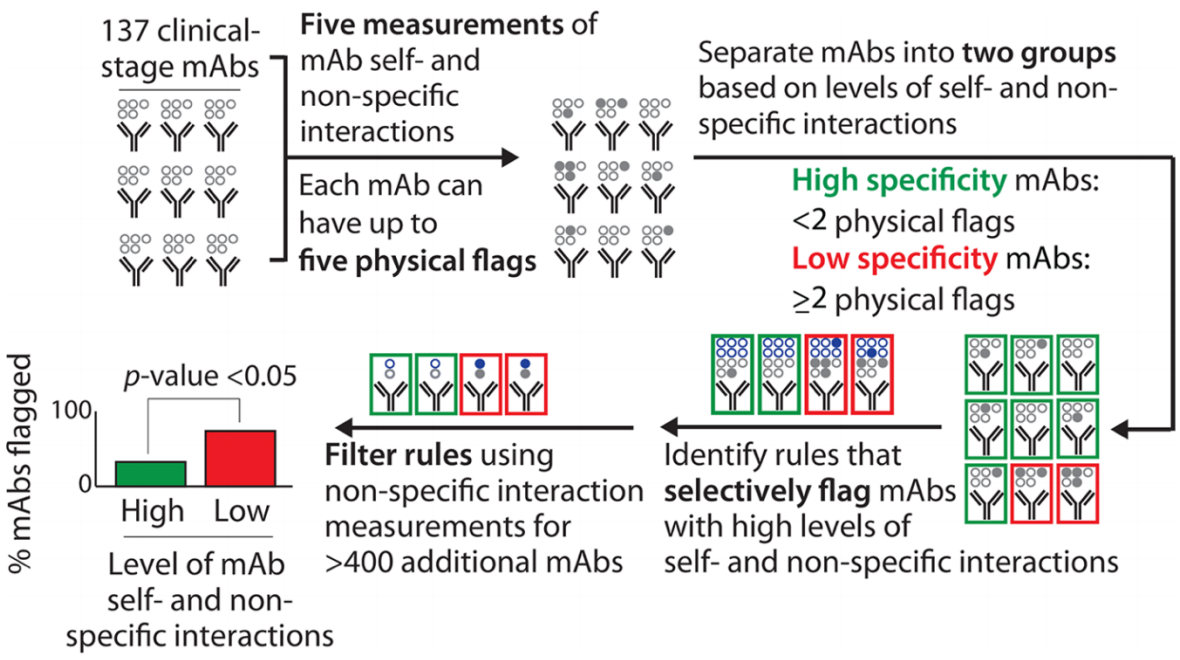


Figure 2-1: Overview of the methodology used to evaluate the molecular determinants of antibody specificity for monoclonal antibodies (mAbs).

Each mAb received up to five physical flags based on exceeding limits for two self-interaction tests (AC-SINS >11.8 nm and CSI >0.01 response units) and three nonspecific interaction tests (PSR >0.27, ELISA >1.9 signal/noise, and BVP >4.3 signal/noise). The experimental data and limits were reported in a previous publication. The statistical significance was evaluated for the ability of the chemical rues to selectively flag mAbs with low specificity (>2 physical flags) relative to mAbs with high specificity (<2 physical flags). The chemical rules were filtered using nonspecific interaction measurements for an additional set of 424 preclinical mAbs to identify the most robust and general chemical rules.

in the 4WT/4MT sets of antibodies based on whether they have the mutations (0 or 100%) relative to their enrichment ratios for PSR- and OVA- samples. Mutational sets were identified as significant if they have Spearman correlation coefficients ≥ 0.6 and *p*-values <0.05.

Results

Chemical rules for identifying antibodies with high specificity. Our approach to identify the molecular determinants of antibody specificity is outlined in Figure 2-1. We applied five tests of antibody specificity to 137 clinical-stage mAbs that are either approved drugs or are (or were) in

phase 2 and 3 clinical trials using previously reported specificity measurements (4). The specificity measurements were obtained using different variable (V_H and V_L) regions for each clinical-stage mAb and the same constant regions (IgG1) regardless of the actual isotype. The five assays included three non-specific binding assays that evaluate antibody interactions with various types of proteins, DNA and virus particles [ELISA (4), BVP (11) and PSR (9)] and two assays that evaluate antibody self-association [AC-SINS (41) and CSI (15)]. We assigned each mAb up to five biophysical flags if they exceeded previously reported upper limits for non-specific and self-interactions that segregate the top 90% of approved antibody drugs from the bottom 10% (4). We define antibodies with high specificity as those with few (<2) biophysical flags. Therefore, we segregated the clinical stage mAbs into two groups, namely those with high specificity (<2 biophysical flags, 97 mAbs) and low specificity (≥2 flags, 40 mAbs), and evaluated chemical rules that selectively identify mAbs with low specificity.

Our approach to identify such chemical rules involved first evaluating maximum limits on the combined numbers of specified residues (weighted by their solvent exposure) for all possible combinations of 19 amino acids (excluding cysteine due to its rarity) for rules composed of as few as one and as many as 10 residues. This process was performed for the entire antibody variable fragment and 66 subregions of Fv, including the Fv framework (without the CDRs), V_H, V_L, individual CDRs (heavy chain CDRs 1, 2 and 3 and light chain CDRs 1, 2 and 3), and all possible combinations of CDRs that include as few as two and as many as six CDRs (e.g., heavy chain CDRs 1 and 3 and light chain CDR2). These rules sampled values that spanned the minimum and maximum values observed in the clinical-stage antibodies in increments of 0.1. In total, we evaluated $>10^7$ maximum rules based on the antibody Fv.

We required that the rules meet a number of constraints (Table 2-2), including that they selectively flag clinical stage mAbs with low specificity relative to mAbs with high specificity. We also required that each rule flag mAbs with low specificity (as judged by the PSR assay) in an even more selective manner for a second training set of 424 human (preclinical) mAbs than for the clinical stage mAbs. The amino acid composition limits that define that clinical-stage and preclinical mAbs are given in Table 2-1.

Our findings for chemical rules that identify mAbs with poor specificity based on maximum limits on the number of solvent-exposed amino acids in antibody variable regions are summarized in Figure 2-2. Despite evaluating $>10^7$ different chemical rules, only 16 ultimately met our constraints and passed our statistical analyses, including our False Discovery Rate (FDR) analysis to account for performing many simultaneous comparisons. Our most significant rule flags mAbs with a sum of >5.0 solvent-exposed Gln, Arg, His, Pro, Met, Leu, Tyr and Trp residues in heavy chain CDRs 1, 2 and 3 (H123) and light chain CDRs 2 and 3 (L23; Figure 2-2A). This single rule flagged more than half (55%) of mAbs with low specificity while only flagging relatively few (14%) mAbs with high specificity (*p*-value of 2.6×10^{-6}). While eight residues contributed to the rule, we evaluated how each residue contributed to the differences in the observed values for low specific mAbs relative to high specific mAbs. Notably, the most important residues were Gln (accounts for >20% of the difference observed between low and high specific mAbs), and Arg and His (each of which contribute 10-20%). Conversely, Pro, Met, Leu and Tyr contributed modestly (0-10%) and Trp contributed negatively (<0%). The latter finding suggests the role of Trp is context-dependent and not simply due to increased Trp content in antibodies with low specificity. The distribution of values for this chemical rule reveals that most mAbs with values >5.0 possess low specificity and those with values <5.0 have high specificity. Finally, the

accuracy of the training (71%) and test (69%) sets of mAbs are similar, suggesting that our cross-validation procedures prevent overfitting of the training data.

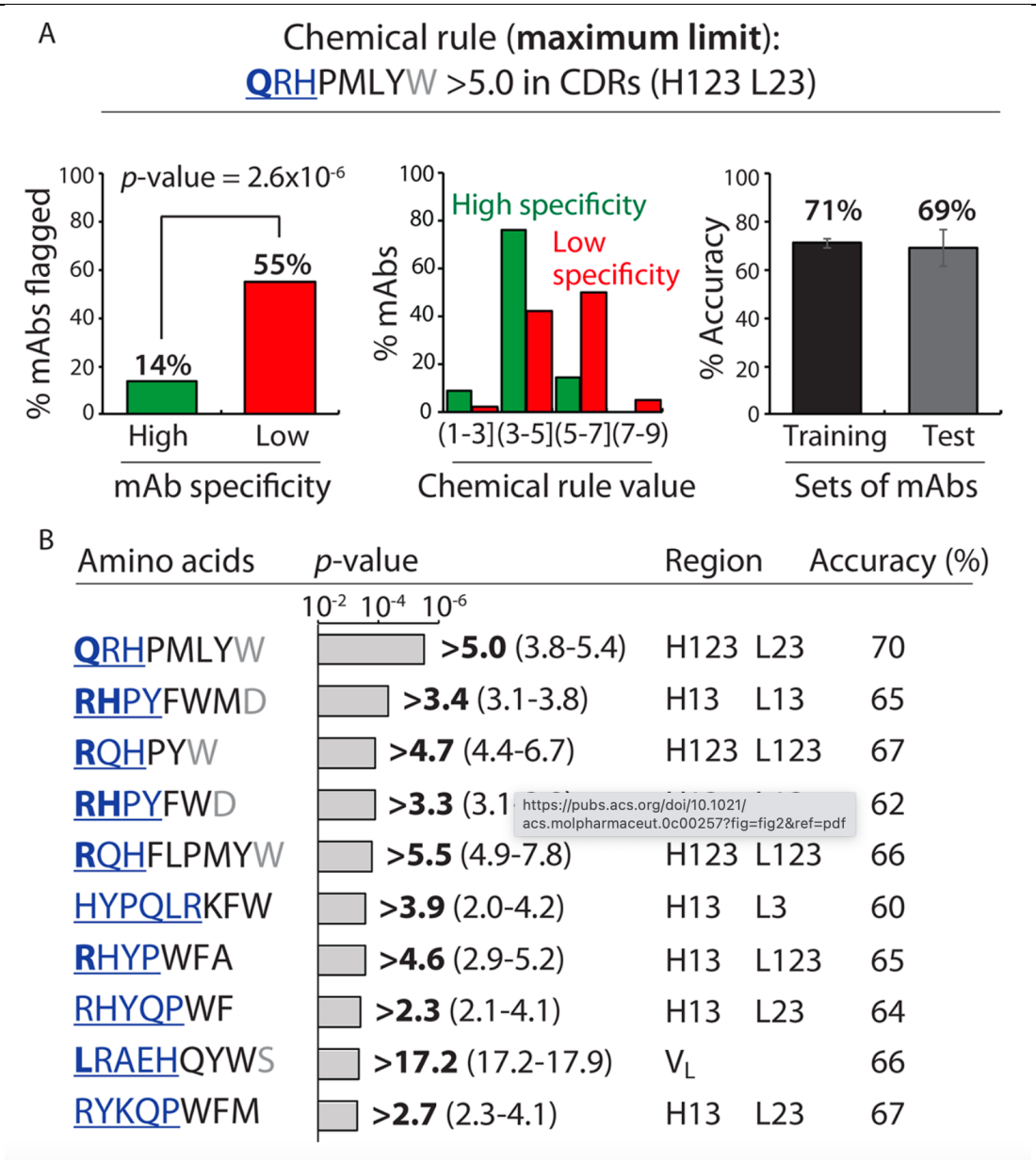


Figure 2-2: Chemical rules for selectively flagging mAbs with low specificity that limit the maximum allowable number of solvent-accessible residues in antibody variable regions.

Each chemical rule is a maximum limit on the summed counts of different types of amino acids in the CDRs weighted by their relative solvent accessibilities. (A) Most selective maximum chemical rule for identifying mAbs with low specificity. (B) Summary of the ten most selective chemical rules that limit the maximum sum of particular types of residues. The bolded value of each rule is the most statistically significant one when evaluated for the entire panel of clinical-stage mAbs, while the range of values reflect those that met the constraints used during cross validation. In (A) and (B), the contributions of the residues to each rule are reported in terms of their contributions to the differences in the observed rule values for mAbs with low specificity relative to those with high specificity. The relative contributions of each amino acid are represented as bold and underlined blue font (most important, >30%), regular and underlined blue font (important, 10-30%), black font (minor importance, 0-10%) and grey font (least important, <0%). mAbs with low and high specificity are defined as described in Fig. 1. The *p*-values were calculated using a 2x2 contingency table (Fisher's exact test), and the reported accuracies are adjusted to account for the different numbers of mAbs with high (97) and low (40) specificity. In (A), the average adjusted accuracies are calculated based on the training (80%) and test (20%) sets for each of the ten splits of the training and test sets. In (B), the adjusted accuracies are calculated for the entire set of 137 clinical-stage mAbs.

It is notable that Arg is the most important contributor to the maximum chemical rules (Figure 2-2B). Of the top ten maximum rules, Arg is one of the most significant contributors (>20% contribution) in half of the rules and a significant contributor (10-20%) in all the other rules. Moreover, His and Gln are also key contributors to the maximum rules (e.g., both contribute >20% in at least one of the rules), suggesting that certain positively charged and polar residues may be particularly important in mediating polyspecificity. Finally, although we considered many different subregions in the antibody variable regions, the chemical rules are dominated by various combinations of heavy and light CDRs.

A key hypothesis in our preceding analysis is that over-enrichment of specific types of solvent-accessible residues in antibody variable regions is linked to poor specificity. We also sought to test the converse hypothesis by evaluating if underrepresentation of other specific types of residues may also be predictive of antibody specificity. Therefore, we evaluated minimum limits on the number of the residues weighted by their solvent exposure in antibody variable regions for all possible combinations of as many as ten residues (19 amino acids excluding cysteine; a total of $>10^7$ rules).

We identified a small subset of minimum rules (24) that met our constraints (Figure 2-3). For example, the most significant minimum rule was a sum <11.6 Asn, Asp, Leu, Ala, Pro, Met, His, Glu and Gln residues in the variable heavy domain (V_H; Figure 2-3A). This single rule flagged half of the mAbs with low specificity while flagging few (13%) of mAbs with high specificity (*p*-value of 1.5×10^{-5}). The most significant contributors were negatively charged (Asp) and polar (Asn) residues. Of the top ten minimum chemical rules, it is notable that Asp is the single most important contributor, and Asn and Glu are also key contributors. These findings suggest that the presence

of negatively-charged and certain polar residues in antibody variable regions are linked to high specificity, which is consistent with previous work (32,34,35,42-46).

Combinations of rules are highly selective for identifying mAbs with high specificity. The selectivity of these rules led us to evaluate whether greater discrimination between antibodies with high and low specificity could be achieved using combinations of these rules (Figure 2-4). Therefore, we tested all possible combinations of 40 individual rules that passed our constraints to generate the best sets of rules that selectively identify mAbs with low specificity. We evaluated sets of rules with as few as four and as many as six (a total of $>10^6$ sets of rules) and identified antibodies with low specificity as those with ≥4-6 chemical flags (as defined on a case-by-case basis). We eliminated the vast majority of the sets of rules by requiring that they satisfy a number of constraints (Table 2-2), resulting in only 16 sets of rules meeting these constraints.

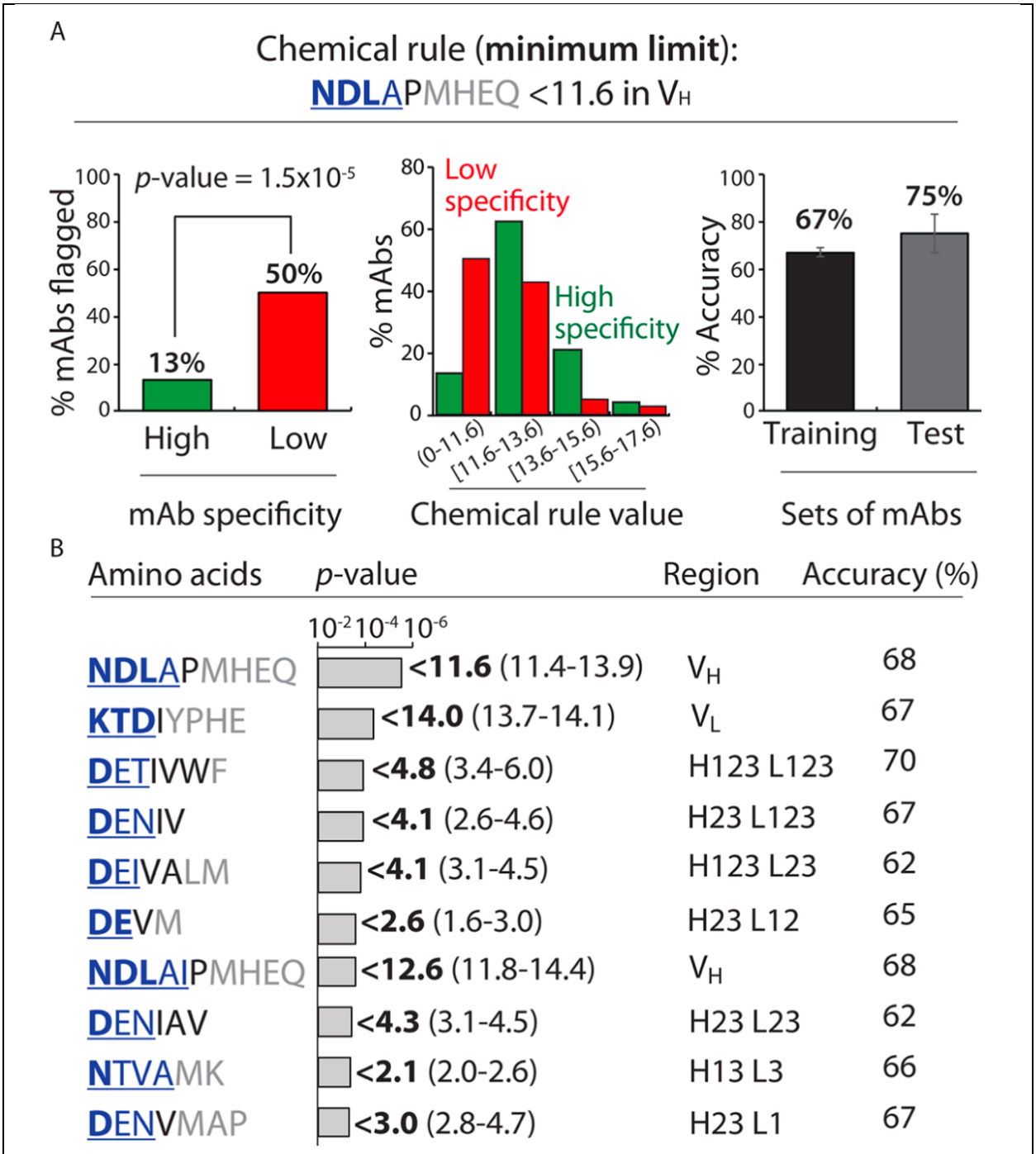


Figure 2-3: Chemical rules for selectively flagging mAbs with low specificity that limit the minimum allowable number of solvent accessible residues in antibody variable regions.

Each chemical rule is a minimum limit on the summed counts of different types of amino acids in the CDRs weighted by their relative solvent accessibilities. (A) Most selective minimum chemical rule for identifying mAbs with low specificity. (B) Summary of the ten most selective chemical rules that limit the minimum sum of particular types of residues. In (A) and (B), the contributions of the residues to each rule are reported are described in Fig. 2 except that the differences in the observed rule values are calculated for high specific mAbs relative to low specific mAbs. mAbs with low and high specificity are defined as described in Fig. 1. The *p*-values and accuracies were calculated as described in Figure 2-2.

The best set of chemical rules we identified contained six chemical rules and displayed a significant improvement in performance relative to the individual rules. This set of rules includes three maximum limits and three minimum limits, five of which are CDR-specific, and the other is V_H-specific. This set of rules was able to flag 35% of clinical-stage mAbs with low specificity while flagging only 2% of mAbs with high specificity (≥4 chemical flags correspond to low antibody specificity). Similar to the individual rules, this set of rules displays similar average validation (66%) and test (67%) accuracies.

We reasoned that the specificity predictions could be further improved by eliminating mAbs flagged by the first set of six rules (Table 2-3) and generating additional rules for selectively flagging liabilities that were not identified in the first round of analysis (Figure 2-4). Therefore, we eliminated mAbs from our training sets that were flagged by the first set of rules and identified individual maximum and minimum chemical rules that were best at selectively identifying the remaining mAbs with low specificity in our training sets. We applied similar constraints and statistical methods in generating the individual rules for the second specificity test as we used for the first test.

Interestingly, we identified several (45) chemical rules that were markedly different than those generated in the first round of analysis (Figure 2-2 and 2-3). For example, Lys was the most significant contributor (>20% contribution) in most (77%) of the maximum rules in the second round of analysis, while Arg was rarely observed as one the most significant contributors (19% of the maximum rules). Moreover, most (73%) of the maximum and minimum rules in the second round of analysis were specific for one of the variable regions (V_H or V_L), the entire Fv or the variable framework (Fv without the CDRs), which was markedly different than the findings in the first round of analysis (15%).

We next evaluated whether the best set of rules in the first round of analysis could be combined with rules generated in the second round of analysis to further improve the selectivity of identifying mAbs with low specificity (Figure 2-4). Therefore, we tested all possible combinations of the 45 individual rules with the six rules generated in the first round of analysis for a total of 8 to 14 chemical rules per set. As for the first round of analysis, we required that the sets of rules meet a number of constraints and statistical measures and we identified mAbs with low antibody specificity as those with ≥4-12 chemical flags (as defined on a case-by-case basis). Our best combined set of chemical rules is reported in Figure 2-4A. The expanded set of 12 rules significantly improves the overall identification of clinical-stage mAbs with high specificity, as defined as those with <8 of 12 chemical flags. This set of rules flags most (78%) of clinical-stage mAbs with low specificity while flagging few (8%) mAbs with high specificity (*p*-value of 1.5×10^{-15} and area under curve of 0.85; Figure 2-4B). Importantly, the average accuracy for our training (83%) and test (90%) sets of antibodies are similar, confirming that our cross-validation methods prevented overfitting.

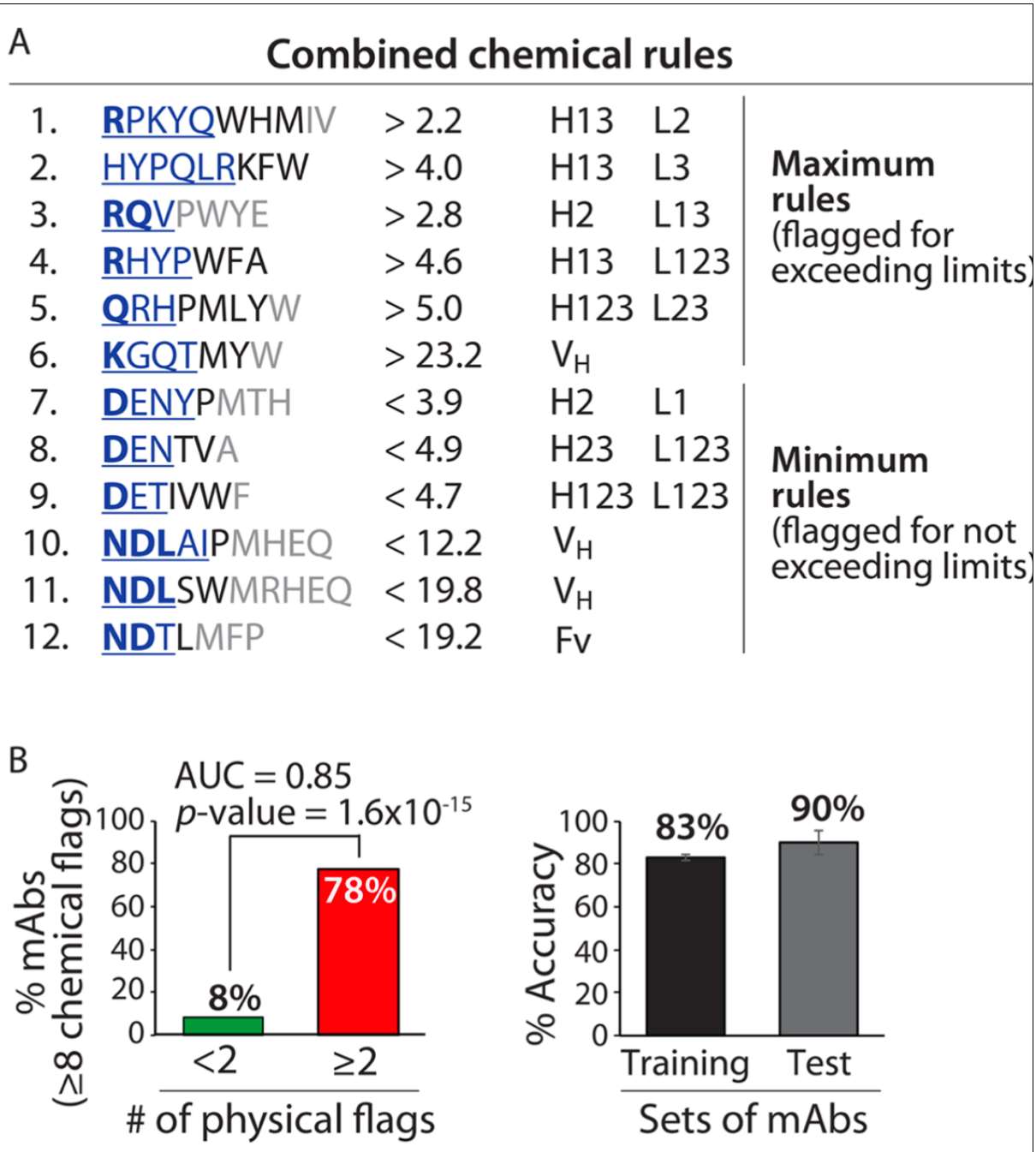


Figure 2-4: Combined chemical rules display high selectivity for identifying clinical-stage mAbs with low specificity.

(A) Antibodies with predicted high specificity are required to be flagged by <8 of 12 rules. The contributions of the residues to each rule are reported as described in Figure 2-2 and 2-3. (B) The combined rules selectively flag mAbs with low specificity (≥2 physical flags) and display similar average adjusted accuracies for the training and test sets. The experimentally determined antibody specificities – as judged by five measurements of non-specific and self-interactions – are defined as described in Fig. 1. The *p*-values and adjusted accuracies were calculated as described in Figure 2-2, and the area under the curve (AUC) is also reported.

The distribution of the number of chemical flags for the mAbs with high and low specificity reveals that most mAbs with <8 chemical flags have high specificity, while those with ≥8 chemical flags have low specificity (Figure 2-5). It is also notable that further refinement of the predictions of antibody specificity is possible. Antibodies with <4 chemical flags are all predicted correctly to have high specificity (accuracy of 100%). Likewise, antibodies with ≥8 flags are mostly predicted correctly to have low specificity (accuracy of 90%). Antibodies with 4-7 flags – which were

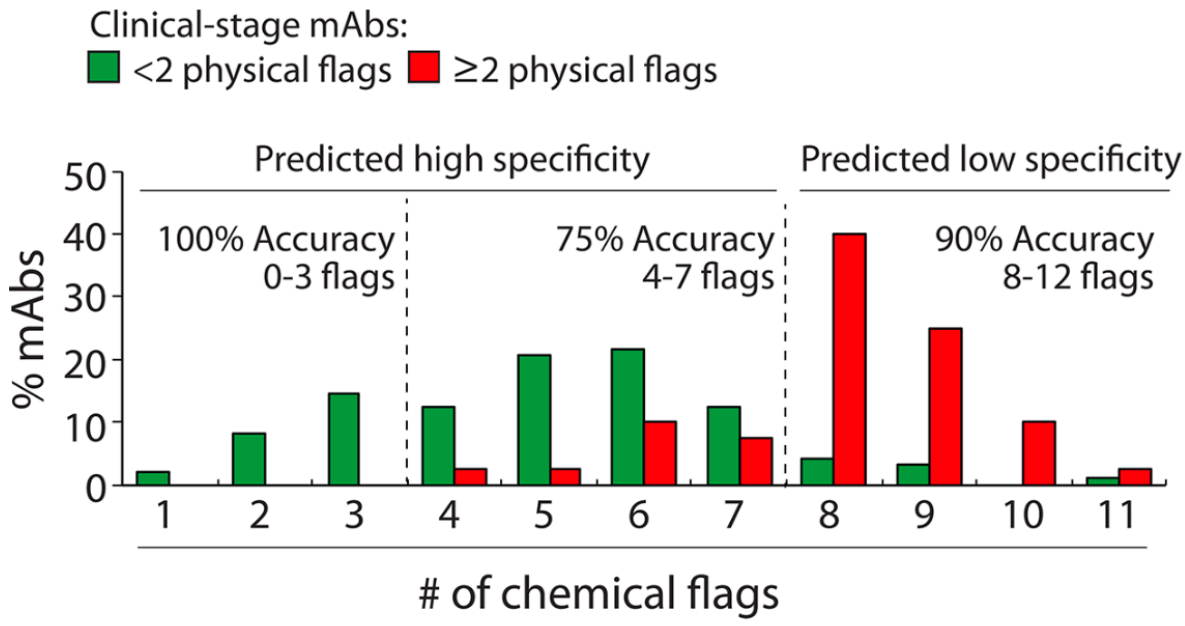


Figure 2-5: Distributions of the number of chemical flags for clinical-stage mAbs with high and low specificity.

The chemical flags are defined in Figure 2-4A. The experimentally determined antibody specificities – as judged by five measurements of non-specific and self-interactions – are defined as described in Figure 2-1. mAbs with high specificity are those with <2 physical flags and mAbs with low specificity are those with ≥2 physical flags. The adjusted accuracies are calculated as described in Figure 2-2.

considered in our original analysis as those with high specificity (<8 of 12 chemical flags) – are predicted correctly with more modest accuracy (75%). This suggests that a useful application of our chemical rules is to define three regions of specificity predictions, two with higher confidence (0-3 chemical flags for high specificity and 8-12 flags for low specificity), and a third with modest confidence in predicting high antibody specificity (4-7 flags).

We also sought to test the performance of our chemical rules if we eliminated the use of specific experimental limits to define antibody specificity (e.g., mAbs with PSR values >0.27 have low specificity) and instead simply ranked the antibodies from the most specific to the least specific based on experimental measurements (Figure 2-6). To do this, we ranked the 137 clinical-stage mAbs from best (lowest levels of non-specific or self-interactions) to worst (highest levels of non-specific interactions or self-interactions) for each of the five biophysical assays and used the average rank percentile of the five assays to define the most specific mAbs (lowest rank). Given that there are 97 of 137 mAbs with <2 physical flags in our original definition of high specificity, we would expect that these antibodies would be ranked mostly in the top 71% (97 of 137 mAbs). Indeed, we find that most (94%, 91 of 97 mAbs) of the antibodies ranked in the top 71% of the mAbs have <2 physical flags (Figure 2-6A), suggesting that our original definition of antibody specificity is weakly influenced by the use of experimental limits to determine poor specificity. Moreover, we find that our chemical rules segregate the best and worst antibodies in a similar manner as the physical rules (Figure 2-6B). We also observe that mAbs with 0-3 chemical flags are mostly (79%, 19 of 24 mAbs) ranked in the top half of the antibodies, while mAbs with 8-12 chemical flags are all (100%) ranked in the bottom half of the antibodies and those with 4-7 chemical flags show intermediate average ranks (Figure 2-6C). These findings further suggest that our chemical rules provide the strongest predictions of high specificity for mAbs with 0-3 chemical flags and low specificity for mAbs with 8-12 chemical flags.

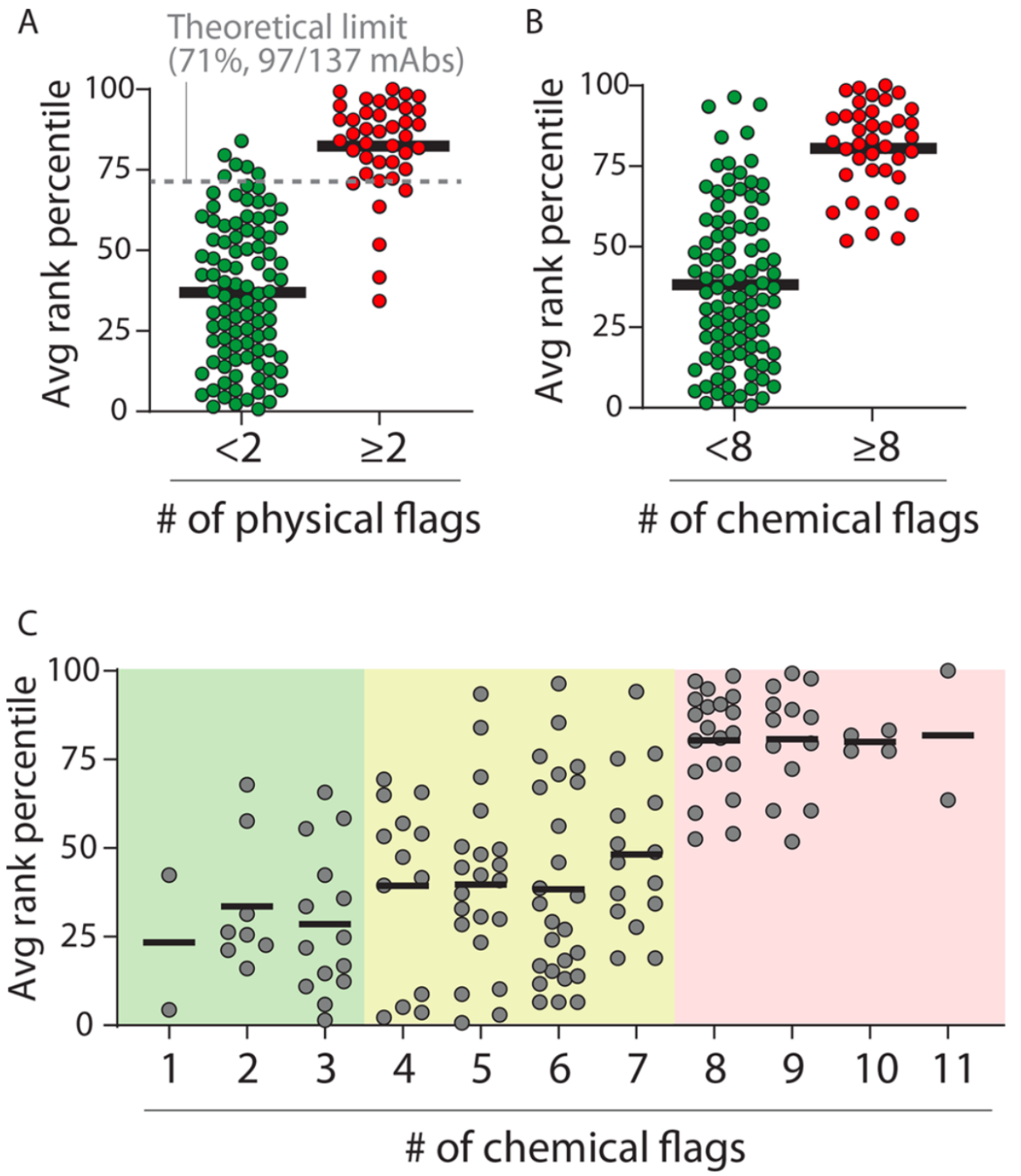


Figure 2-6：Comparison of the average rank for clinical stage mAbs based on five measures of non-specific and self-interactions and the corresponding number of physical and chemical flags.

(A) The average rank of mAbs with <2 physical flags (97 of 137 mAbs, 71% of mAbs) and ≥2 physical flags (40 of 137 mAbs, 29% of mAbs) were calculated based on their ranks in five assays of self- and non-specific interactions. (B, C) The average experimental rank of mAbs compared to (B) <8 versus ≥8 chemical flags and (C) the number of chemical flags. In (C), three regions are shown, one with predicted high specificity (0-3 chemical flags), a second one with intermediate specificity (4-7 chemical flags) and a third one with low specificity (8-12 chemical flags).

Our goal in developing the chemical rules in this work was to broadly describe antibody specificity and not rely on any single type of experimental specificity measurement. Nevertheless, we next evaluated how the rules that emerged from our analysis would perform in the context of each individual specificity assay (three non-specific binding assays and two self-interaction assays). For each biophysical assay, the antibodies were segregated into two groups based on previously established limits for high levels of non-specific and self-interactions. Next, the specificity test in Figure 2-4 was applied to each group of antibodies defined by single non-specific and self-interaction measurements (Figures 2-7 to 2-11). Encouragingly, the performance of the chemical rules was both strong and relatively similar for the five individual assays (p-values $<10^{-5}$ and accuracy of ≥75% for each assay). Moreover, we observed that the accuracies for the PSR and AC-SINS assays were particularly strong for predicting high specificity of mAbs with <4 flags (100% accuracy). More generally, we also observed strong performance for the ELISA, BVP, PSR and AC-SINS assays for predicting mAbs with high specificity (>80% accuracy, <4 flags) and low specificity (>80% accuracy, 8-12 flags).

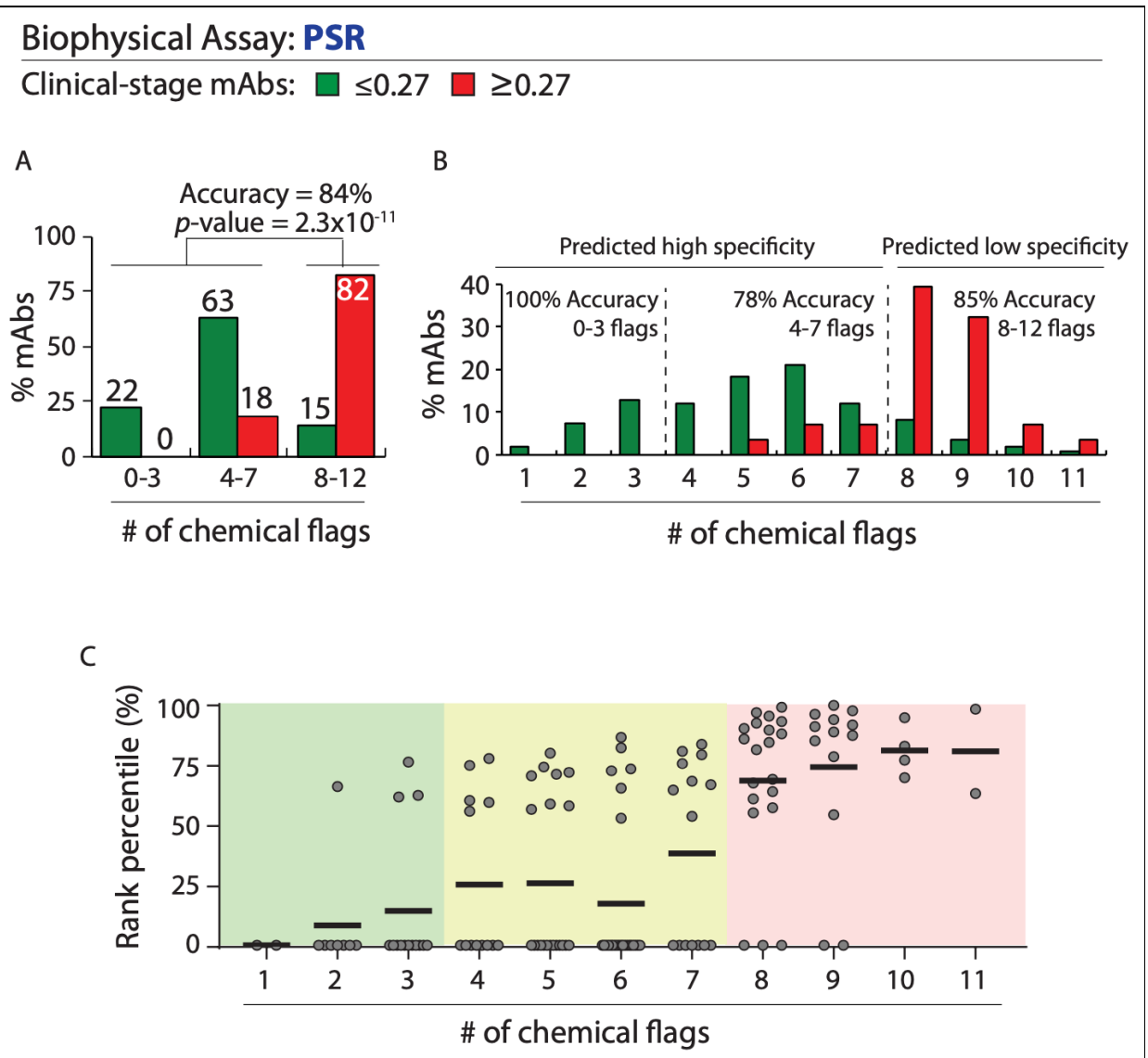


Figure 2-7：Performance of combined chemical rules for identifying clinical stage mAbs with high levels of non-specific interactions detected using the PSR assay.

(A, B) Evaluation of the percentage of mAbs with high and low specificity flagged by the combined set of chemical rules for (A) grouped numbers of chemical flags and (B) individual numbers of chemical flags. (C) Comparison of the rank for clinical stage mAbs based on PSR measurements and the corresponding number of chemical flags. In (A) and (B), the p-values were calculated using a 2×2 contingency table (Fisher's exact test). In (C), three regions are shown, one with predicted high specificity (0-3 chemical flags), a second one with intermediate specificity (4-7 chemical flags), and a third one with low specificity (8-12 chemical flags).

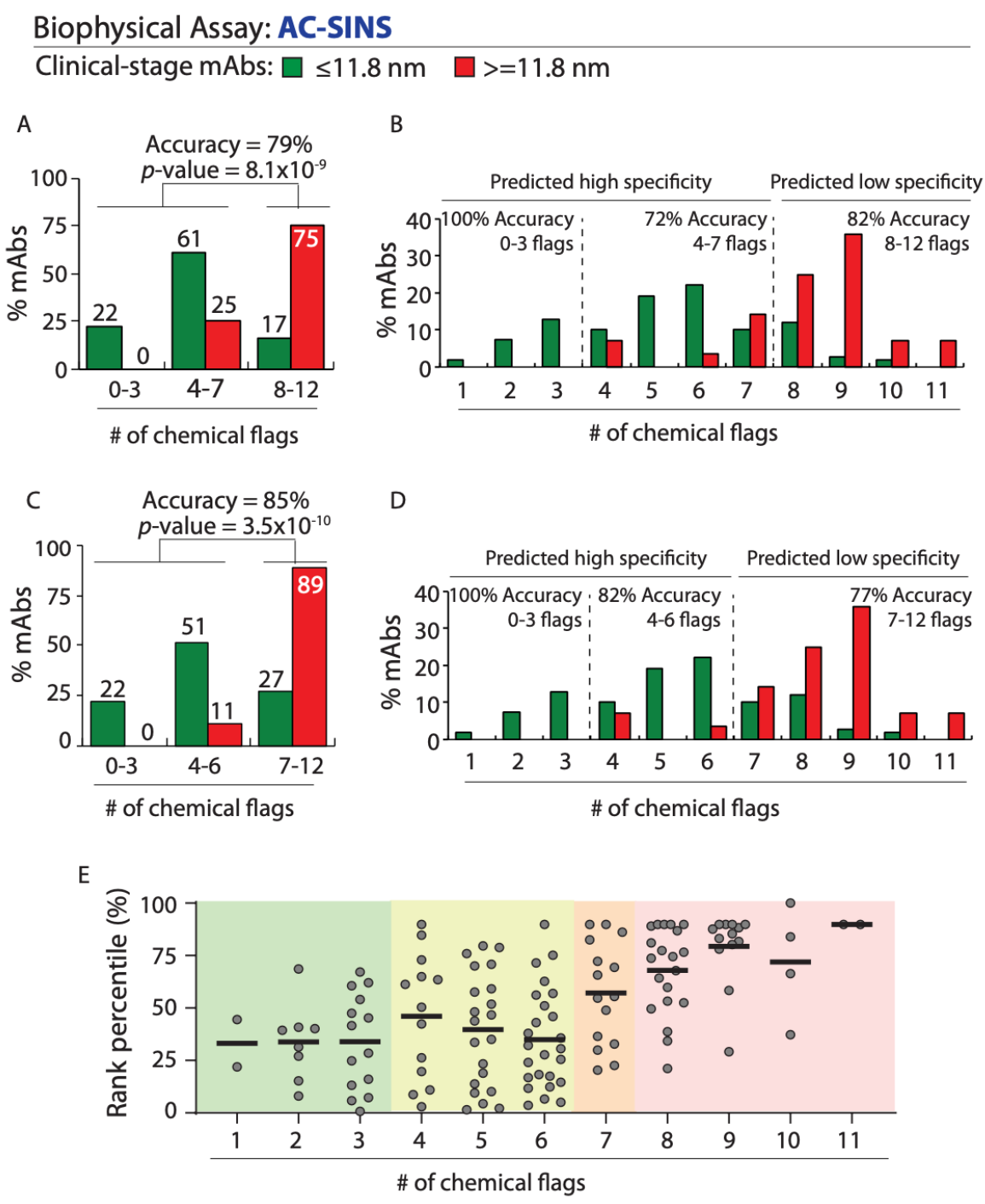


Figure 2-8: Performance of combined chemical rules for identifying clinical-stage mAbs with high levels of self-interactions detected using the AC-SINS assay.

(A-D) Evaluation of the percentage of mAbs with high and low specificity flagged chemical flags. (E) Comparison of the rank for clinical-stage mAbs based on AC-SINS measurements and the corresponding number of chemical rules. In (A-D), the p-values were calculated using a 2×2 contingency table (Fisher's exact test). In €, three regions are shown, one with predicted high specificity (0-3 chemical flags), a second one with intermediate specificity (4-6 or 4-7 chemical flags), and a third one with low specificity (7-12 or 8-12 chemical flags).

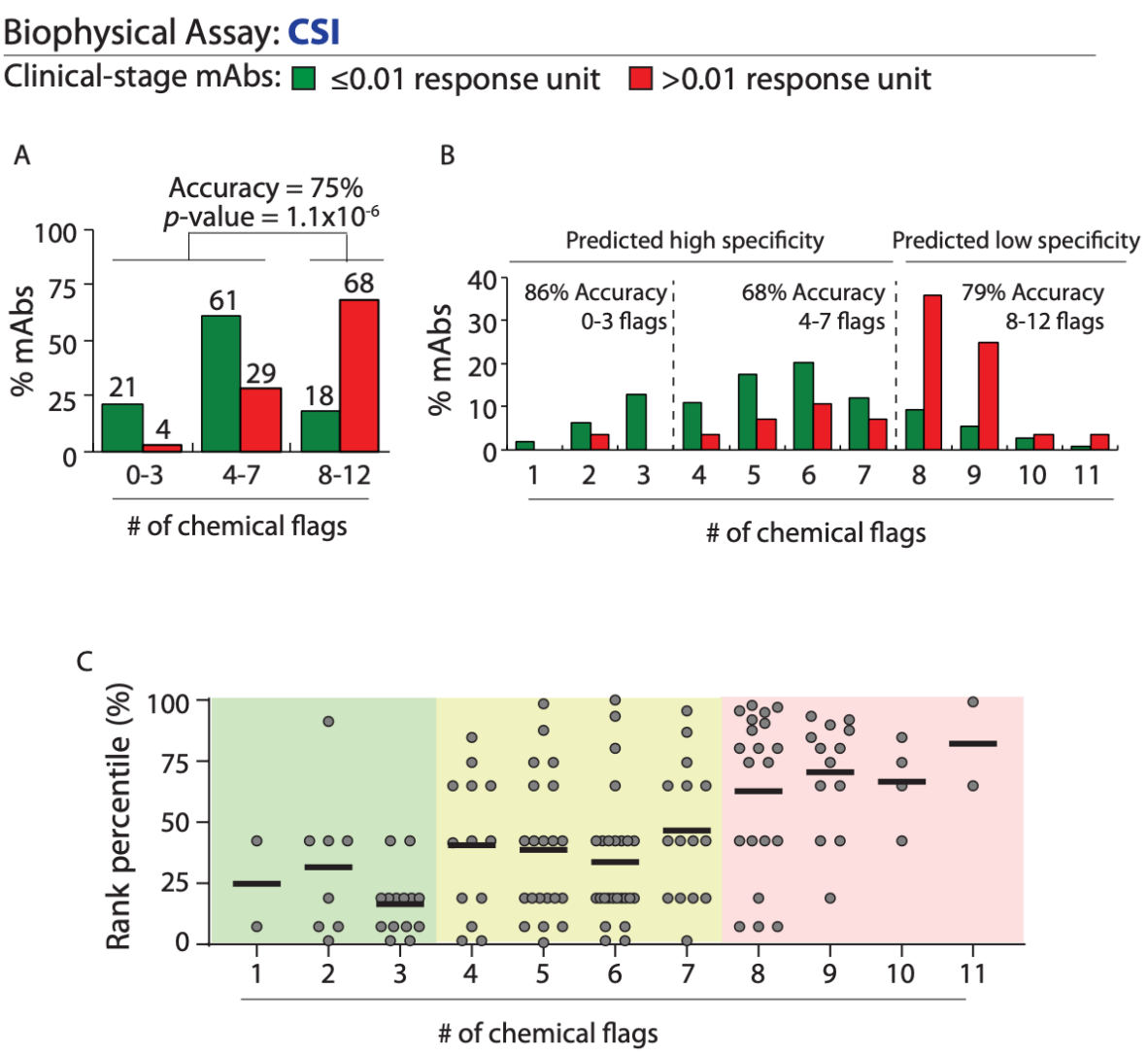


Figure 2-9: Performance of combined chemical rules for identifying clinical stage mAbs with high levels of self-interactions detected using the CSI assay.

(A, B) Evaluation of the percentage of mAbs with high and low specificity flagged by the combined set of chemical rules for (A) grouped numbers of chemical flags and (B) individual numbers of chemical flags. (C) Comparison of the rank for clinical stage mAbs based on CSI measurements and the corresponding number of chemical flags. In (A) and (B), the p-values were calculated using a 2×2 contingency table (Fisher's exact test). In (C), three regions are shown, one with predicted high specificity (0-3) chemical flags), a second one with intermediate specificity (4-7 chemical flags), and a third one with low specificity (8-12 chemical flags).

Evaluation of combinations of chemical rules using independent sets of antibodies. We also evaluated our rules using independent sets of antibodies not included in our training or test sets. We first evaluated an independent set of non-specific interaction (PSR) measurements for an additional 359 preclinical mAbs that largely fall within the amino acid composition space that we evaluated previously in this work (≥98% and <99% similarity based on the maximum and minimum limits). Importantly, the combined specificity rules correctly identified more than half (55%) of these preclinical mAbs with low specificity while flagging few mAbs (16%) with high specificity (*p*-value of 1.6×10^{-4} and accuracy of 69%; Figure 2-12). It is also notable that the accuracies for the training (70%) and test (69%) sets of preclinical antibodies were similar, suggesting that the combined rules do not overfit the training data.

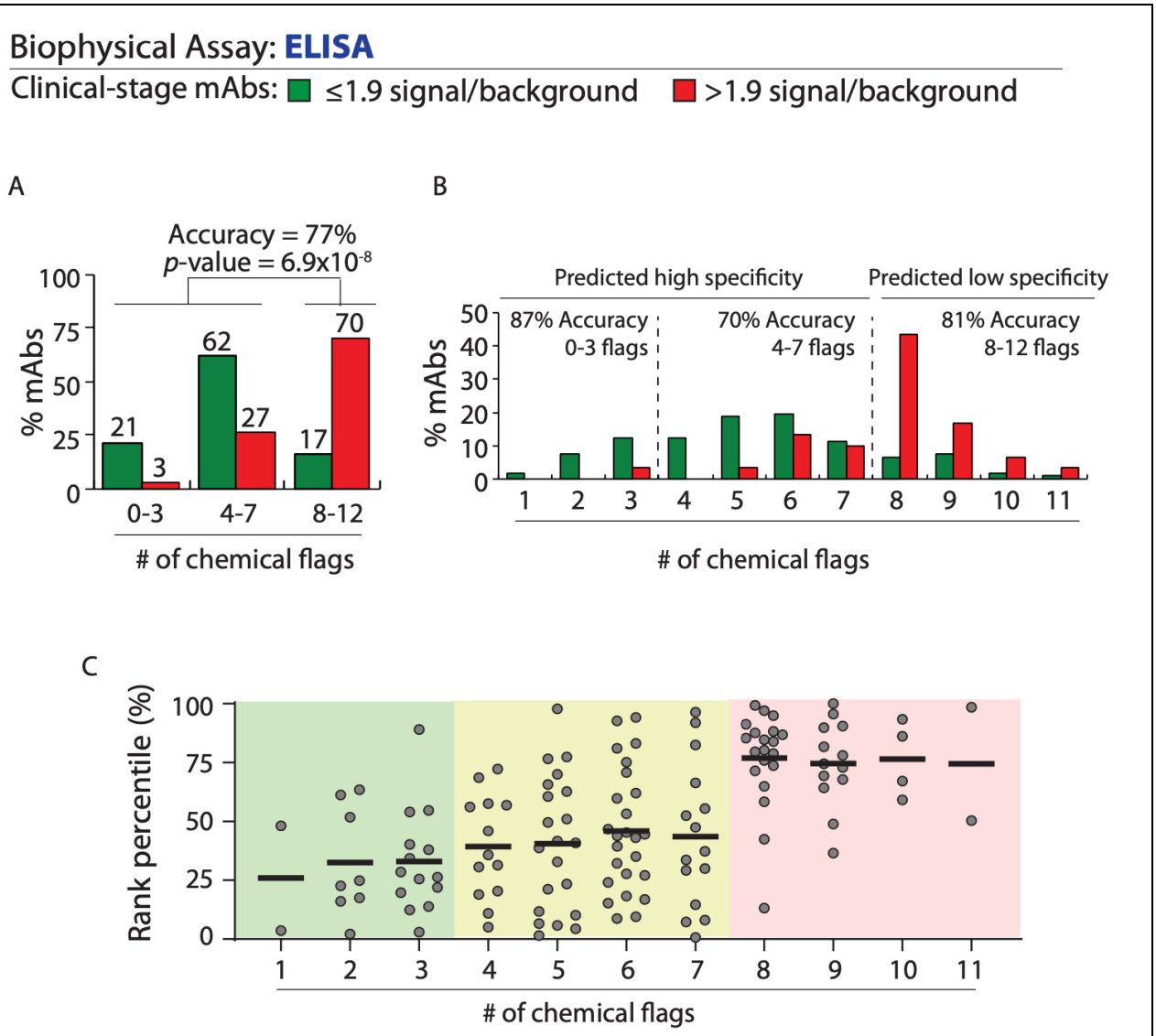


Figure 2-10: Performance of combined chemical rules for identifying clinical stage mAbs with high levels of non-specific interactions detected using the ELISA non-specific binding assay.

(A, B) Evaluation of the percentage of mAbs with high and low specificity flagged by the combined set of chemical rues for (A) grouped numbers of chemical flags and (B) individual numbers of chemical flags. (C) Comparison of the rank for clinical stage mAbs based on ELISA measurements and the corresponding number of chemical flags. In (A) and (B), the *p*-value was calculated using a 2×2 contingency table (Fisher's exact test). In (C), three regions are shown, one with predicted high specificity (0-3 chemical flags), a second one with intermediate specificity (4-7 chemical flags), and a third one with specificity (8-12 chemical flags).

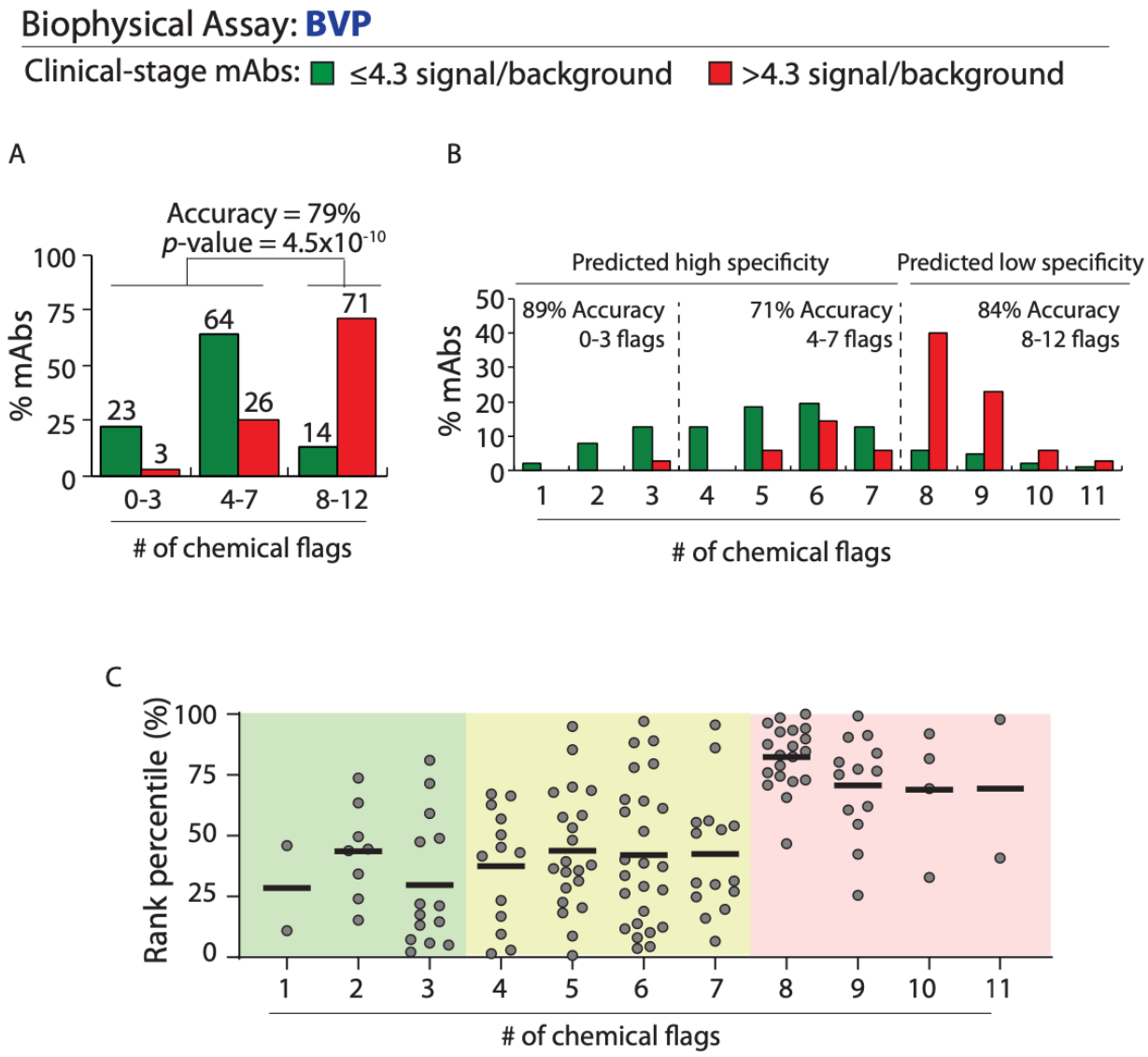


Figure 2-11: Combined chemical rules selectively flag preclinical mAbs with high levels of non-specific interactions.

Antibodies with predicted high specificity (as described in Figure 2-4) display reduced levels of non-specific binding to a poly-specificity reagent (PSR). (A) The selectivity of the combined chemical rules is similar for the training and test sets of preclinical antibodies. (B) Distributions of the number of flags for the training and test sets of antibodies with low and high specificity. The *p*-value was calculated using a 2×2 contingency table (Fisher's exact test).

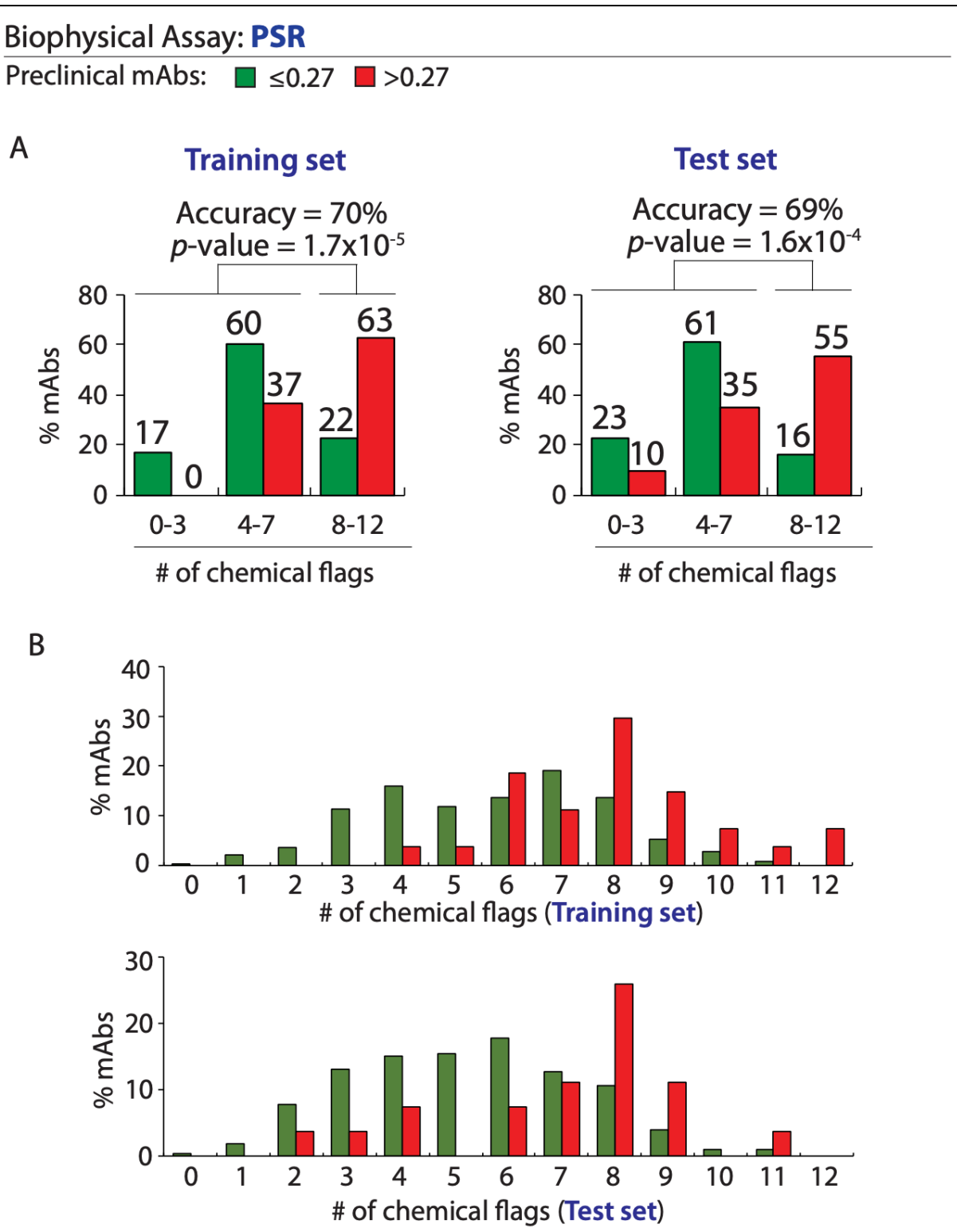


Figure 2-12: Combined chemical rules selectively flag preclinical mAbs with high levels of non-specific interactions.
Antibodies with predicted high specificity (as described in Figure 2-4) display reduced levels of non-specific binding to a poly-specificity reagent (PSR). (A) The selectivity of the combined chemical rules is similar for the training and test sets of preclinical antibodies. (B) Distributions of the number of flags for the training and test sets of antibodies with low and high specificity. The *p*-value was calculated using a 2×2 contingency table (Fisher's exact test).

We also tested our predictions using a second independent set of mAbs (39 IgGs) that we generated and characterized using non-specific binding (ELISA) and self-association (AC-SINS) assays (Figure 2-7). Importantly, the mAbs predicted by our specificity rules to have poor specificity (≥8 chemical flags) displayed significantly higher levels of non-specific binding than those that pass our test (<8 flags, *p*-value of 6.3×10^{-4}; Figure 2-7A). Moreover, we find that the same specificity test is also able to identify mAbs with high levels of self-association (Figure 2-7B). The mAbs identified by our specificity test (≥8 flags) displayed markedly higher levels of self-association relative to mAbs that were not flagged (<8 flags, *p*-value of 1.2×10^{-9}). Moreover, we observed that mAbs with <4 chemical flags generally had lower levels of non-specific binding and self-association than those with 4-7 flags, although the difference for non-specific binding is not significant. These results demonstrate the generality of our methodology for identifying antibodies with drug-like specificity.

The strong performance of our chemical rules for identifying antibodies with high specificity using multiple independent sets of data (Figure 2-12 and 13) suggest that our rules identify some of the most important determinants of antibody specificity. However, we sought to evaluate our predictions using much larger data sets to better evaluate their utility in identifying antibodies with high specificity. Therefore, we sought to mutagenize a clinical-stage antibody with poor specificity (emibetuzumab) that is flagged by 8 of 12 chemical rules. Our strategy was to identify sites in the variable regions that were flagged by our maximum rules and mutate them to residues that reduced the number of chemical flags, including those that are most important in the minimum rules (e.g., D and T). We identified eight sites in the heavy chain CDRs to mutagenize and sampled five mutations per site in addition to the wild-type residue using degenerate codons, which resulted in $>10^6$ variants. This library was then displayed on yeast as single-chain Fab

fragments, sorted for low non-specific binding against two reagents (PSR and ovalbumin) or high (specific) binding to the target antigen, and the selected antibody variants with high specificity were identified using deep sequencing.

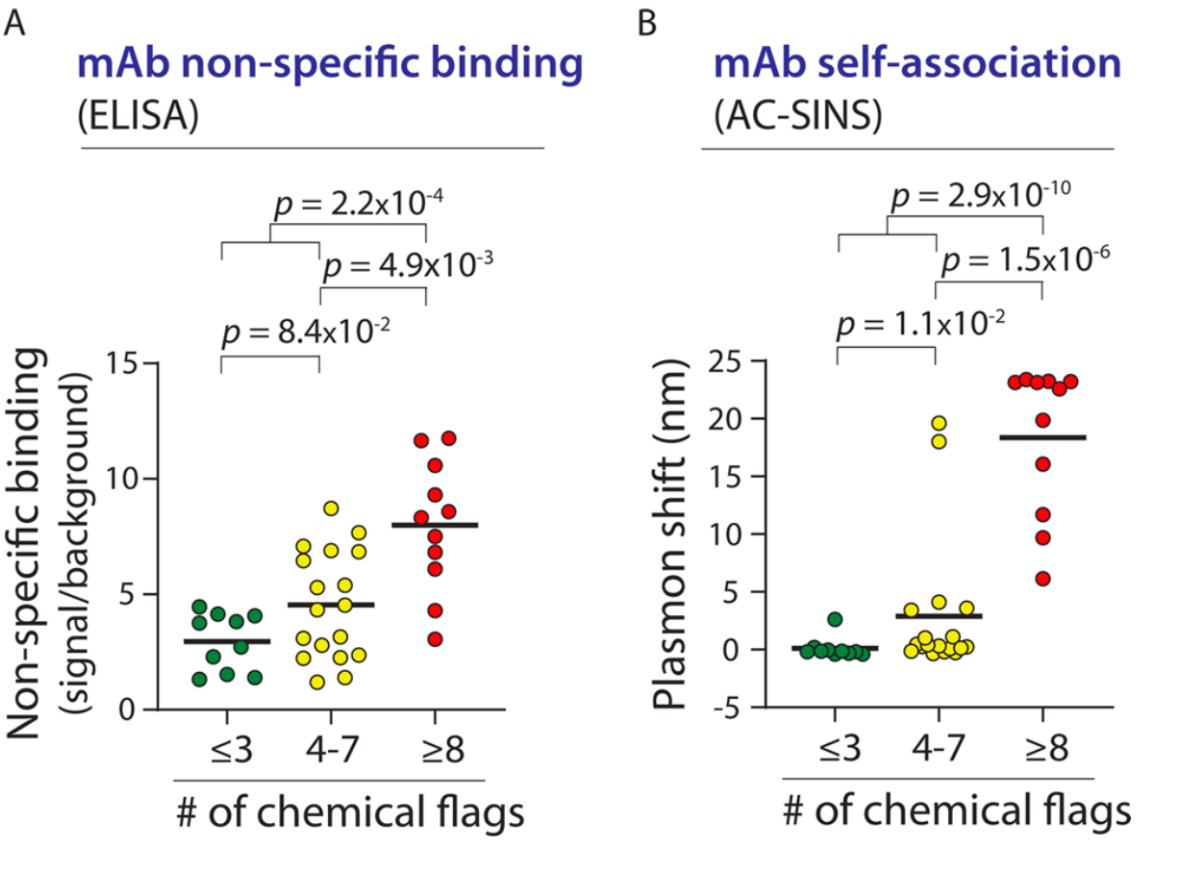


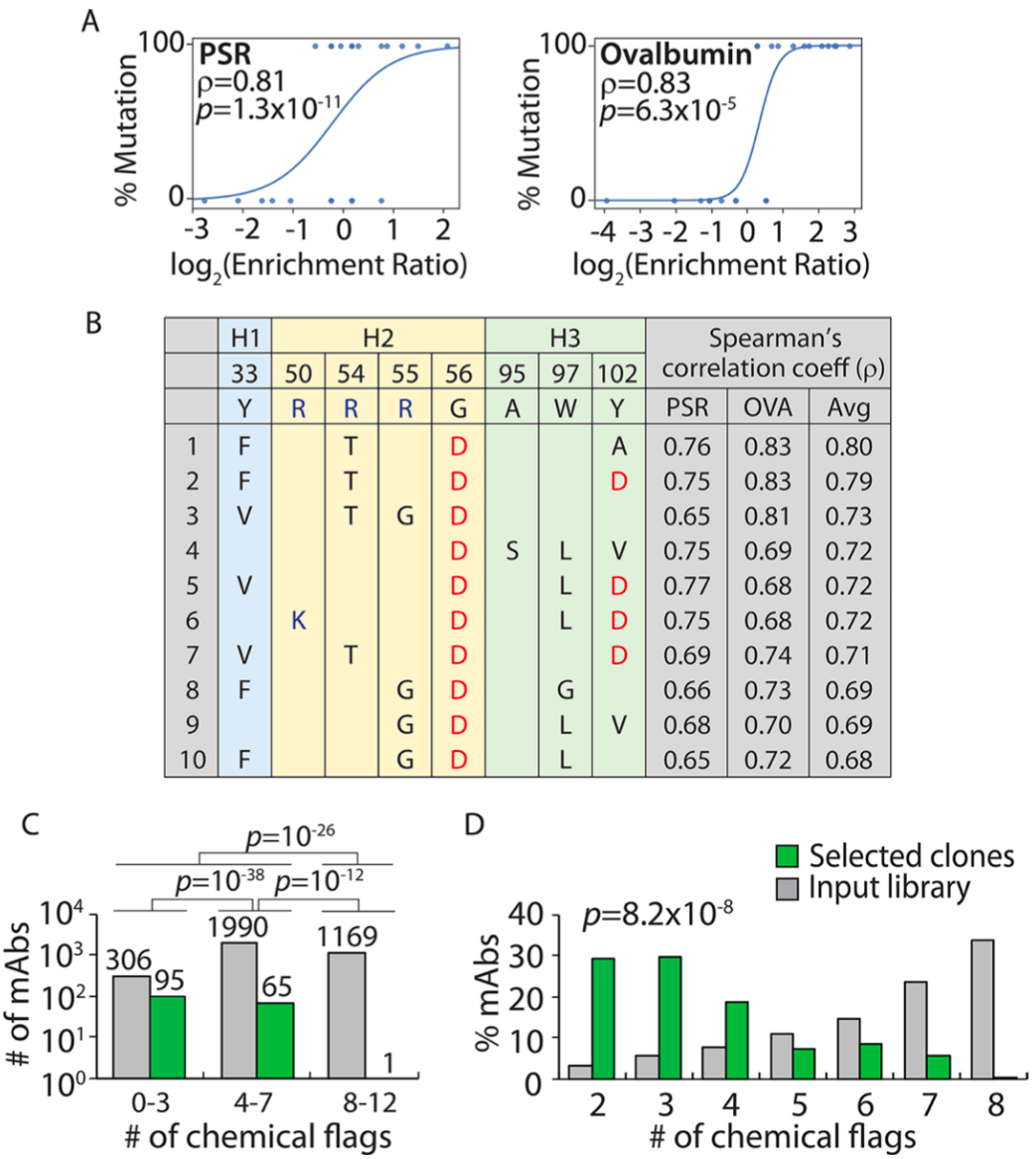
Figure 2-13：Combined chemical rules strongly differentiate between mAbs with different levels of non-specific and self-interactions for an independent set of antibodies.

(A) Non-specific interactions (ELISA) and (B) self-interactions (AC-SINS) were measured for 39 mAbs that were not included in the training and test sets used to generate the combined chemical rules. The *p*-values were calculated using a two sample Anderson-Darling test. In (A), the difference between ≤3 chemical flags and 4-7 chemical flags is not significant.

Our findings are summarized in Figure 2-8. Of the 3465 antibodies that we identified in two independent experiments, we first sought to identify sets of mutations that were most strongly correlated with significant enrichment for high specificity during selection. Our initial analysis led us to focus on sets of four mutations to maximize statistical significance. For example, we identified a set of four mutations (Y33F, R54T, G56D and Y102A in V_H) that shows strong

correlation between antibody variants that have such mutations and their enrichment ratios during selection for low polyspecificity (Spearman's correlation coefficients of 0.81 for PSR and 0.83 for ovalbumin; Figure 2-8A). Of the top ten sets of four mutations we identified, all of them included a mutation that introduces negative charge, further suggesting that negative charge is linked to increased antibody specificity (Figure 2-8B). It is also notable that most (80%) of top sets of mutations involved eliminating at least one Arg in the CDRs. In total, we identified 13 sets of four mutations and 161 antibodies with such mutations that are expected to possess high specificity.

We next sought to test the ability of our chemical rules to correctly predict antibodies with high specificity (Figure 2-14C). Strikingly, despite observing thousands of antibodies with 0-7 chemical flags (2,370 antibodies) and 8-12 flags (1,095 antibodies with 8 chemical flags) in our input library, our rules correctly identify almost all selected antibodies with high specificity (160 of 161 antibodies with <8 chemical flags, p-value of 10^{-26}). Moreover, we find that antibodies with 0-3 flags were much more strongly enriched than even those with 4-7 flags (p-value $<10^{-38}$), which further suggests that our predictions of specificity are strongest for antibodies with 0-3 chemical flags. More generally, we find that the distributions of the number chemical flags for the selected mAbs with high specificity relative to the input library are significantly different (p-value of 10^{-8}; Figure 2-14D). These findings demonstrate how our chemical rules can be used to guide the design of antibody libraries that target sites involved in polyspecificity and facilitate the identification of antibodies with significant improvements in specificity.



	H1	H2				H3			Spearman's correlation coeff (ρ)		
	33	50	54	55	56	95	97	102			
	Y	R	R	R	G	A	W	Y	PSR	OVA	Avg
1	F		T		D			A	0.76	0.83	0.80
2	F		T		D			D	0.75	0.83	0.79
3	V		T	G	D				0.65	0.81	0.73
4					D	S	L	V	0.75	0.69	0.72
5	V				D		L	D	0.77	0.68	0.72
6		K			D		L	D	0.75	0.68	0.72
7	V		T		D			D	0.69	0.74	0.71
8	F			G	D		G		0.66	0.73	0.69
9				G	D		L	V	0.68	0.70	0.69
10	F			G	D		L		0.65	0.72	0.68

Figure 2-14：Design of Fab sub-libraries of emibetuzumab guided by the combined chemical rules and evaluation of selected mutants with improved antibody specificity.

The V_H domain of emibetuzumab was mutated at eight solvent-exposed sites (Y33, R50, R54, R55, G56, A95, W97 and Y102) in the three heavy chain CDRs that were flagged by the maximum chemical rules. The mutations sampled the wild-type residue as well as five mutations that are predicted to reduce the number of chemical flags. The libraries were constructed as single-chain Fab fragments (scFabs) on yeast, sorted for non-binding to two polyspecificity reagents [PSR and ovalbumin (OVA)], and evaluated via deep sequencing. (A) Enrichment ratios for antibody variants with a set of four mutations (F33, T54, D56 and A102 in V_H) relative to antibody variants with wild-type residues at the same positions (Y33, R54, G56 and Y102 in V_H) for two different polyspecificity reagents. (B) Top ten sets of four mutation combinations that are most strongly correlated with reduced binding to polyspecificity reagents and increased specificity. (C, D) The (C) number and (D) percentage of mAb variants selected with high specificity as a function of the number of chemical flags relative to the corresponding values for the input library. In (A), the mAbs included in the wild-type or mutant groups are only required to have wild-type or mutant residues at the four sites that are evaluated and can have either wild-type or mutant residues at the other four sites. Moreover, logistic regression curves are shown as guides to the eye, and the *p*-values are for the Spearman's correlation coefficients (rho). In (C), the *p*-values for the comparisons of the number of mAbs were calculated using a 2x2 contingency table (Fisher's exact test). In (D), the *p*-value for comparing the distributions of mAbs was calculated using paired sample *t*-test (two tailed).

Discussion

Our findings provide a relatively simple yet powerful description of the overall specificity of monoclonal antibodies. To the best of our knowledge, there are no prior methods for predicting the overall propensity of antibodies to interact non-specifically with various types of molecules or themselves based simply on their primary structures and their corresponding sequence-based solvent accessibilities. Importantly, the conceptual framework that we have developed is able to explain many previous disparate findings and observations. Our finding that antibodies with high specificity have low levels of positively charged CDR residues is consistent with previous findings that increased levels of positive charge in the CDRs and variable (V_H and V_L) regions are linked to poor antibody specificity (17,21,23,46-50) and pharmacokinetics (17,23,50-53) as well as high viscosity in concentrated antibody solutions (54). Conversely, our finding that antibodies with high specificity have increased levels of negative charge in the CDRs is consistent with previous findings that increased levels of negative charge in antibody CDRs and variable regions are linked to high solubility and low aggregation propensity (32,34,35,55-59), low self-association (60) (although exceptions have been noted for mAbs with abnormal levels of negative charge (56,57)) and favorable pharmacokinetics (52,53,61). Moreover, our finding that antibodies with high specificity also have increased levels of polar CDR residues (e.g., Asn) is consistent with the fact that increased levels of polar residues in the CDRs are linked to increased antibody specificity (47,48) and improved pharmacokinetics (36). Our specificity rules capture multiple factors that govern the physicochemical properties of antibodies and simplify them into powerful guidelines for rapidly identifying mAbs that are expected to display favorable specificity, solubility and biodistribution.

There are also multiple factors to consider when interpreting and applying our findings. First, we defined the limits of sequence space for our analysis in Table 2-1 based on the maximum and minimum numbers of each type of amino acid (weighted by their relative solvent accessibilities) in the CDRs and variable regions of preclinical and clinical-stage antibodies in our training sets. It is expected that the performance of our rules will be reduced for antibodies whose amino acid compositions and site-specific solvent accessibilities are outside the range of chemical and structural diversity that we explored in this analysis. Encouragingly, we find that the accuracy of our specificity predictions for non-specific binding (PSR assay) is weakly impacted by reducing the similarity of antibodies in our test set relative to the those in our training sets (Figure 2-12). However, the accuracy of such predictions is expected to decrease as the antibodies to be analyzed become more dissimilar relative to those in our training sets. Second, the antibody specificity measurements were obtained using common antibody constant regions (IgG1 isotype with corresponding kappa and lambda light chains) (4,39). Therefore, the effects of different antibody isotypes on the physicochemical properties of mAbs have not been addressed in our specificity analysis and will need to be addressed in future work.

Although our findings generally suggest that increased negative charge in antibody variable regions is linked to reduced non-specific and self-interactions, it is well established that over-enrichment in negative charge in antibody variable regions is also linked to increased self-association and viscous solution behavior at high antibody concentrations (20,56,62-65). One obvious difference between this work and previous studies related to viscous antibody solutions are the solution conditions, as we analyzed antibody self-interaction measurements at pH 7.4 (PBS) and previous studies have evaluated antibody viscosity measurements in typical formulation conditions (e.g., pH 5-6 without salt or with low salt concentrations). Another notable difference

is that the antibodies with high specificity in our analysis have modest amounts of negative charge – based on their theoretical net charge calculated at pH 7.4 – in their V_H (0.8±1.9), V_L (0.8±2.0), Fv (1.5±2.5), heavy chain CDRs (-1.5±1.7), light chain CDRs (0.6±1.9) and overall CDRs (-0.9±2.4; Table S16). Moreover, the isoelectric points of the variable regions of the antibodies with high specificity in our analysis are relatively typical of antibodies in general, as evidenced by the values for V_H (7.4±1.4), V_L (7.5±1.4) and Fv (7.7±1.3).

Caution should be used when interpreting our predictions of specificity for antibodies with variable region charges and isoelectric points that fall beyond the range of values represented in our study. For example, omalizumab is viscous at high antibody concentrations at typical mAb formulation conditions (e.g., ~40 cP at ~120 mg/mL in histidine buffer of pH 6.0)(56,57,66) and has abnormal charge properties (pH 7.4) relative to those for high specific antibodies in this study, including a more negatively charged V_L (-3.9 relative to 0.8±2.0) and light chain CDRs (-4.9 relative to 0.6±1.9) as well as a lower V_L isoelectric point (pI of 4.7 relative to 7.5±1.4). Moreover, mutations in the light chain CDRs of omalizumab (D28A, D30N, H92Y, E93T, D94T and Y96P in V_L) that reduce the amount of negative charge (-1.0 relative to -4.9 for wild type at pH 7.4) and increase the isoelectric point of V_L (pI 6.3 relative to 4.7 for wild type) to be within the range of the high specific antibodies significantly reduce the viscosity (<10 cP at ~120 mg/mL) (56,57,67). Future work is needed to better define limits on the amount of negative charge in antibody variable regions that is favorable for specificity without promoting attractive electrostatic interactions that are unfavorable for high concentration viscoelastic behavior.

It is also important to consider the impact of the methods we used to calculate solvent-accessible surface areas on our findings. We employed a published machine learning method that only requires antibody variable region sequences (40), which is particularly convenient for

antibodies of unknown structure. Nevertheless, it is important to note that this machine learning method yields slightly different models each time it is trained, which leads to small differences in the predicted solvent accessibility values and the resulting chemical rule values. For example, we compared two different models generated from this machine learning method and observed minor differences in the number of chemical flags for the clinical-stage antibodies. This results in minor changes in the accuracy of the predictions for the specificity of the clinical-stage antibodies (e.g., 82-85% accuracy). We recommend not only using the same machine learning method, but also the same compiled model that we used in this work to calculate solvent accessibilities for evaluating the number of chemical flags for additional antibodies. It is also notable that the machine learning method is mostly trained on antibodies with kappa light chains, and caution should be used when applying this method to antibodies with lambda light chains. More generally, our training sets of non-specific and self-interaction measurements primarily contained antibodies with kappa light chains given that most clinical-stage antibodies have kappa light chains (91% of the clinical-stage antibodies in this study), and our chemical rules are expected to be most useful for predicting the specificity of antibodies with kappa light chains.

We expect that our findings will immediately impact therapeutic antibody development in multiple ways. First, our specificity rules will serve as valuable design guidelines for generating antibody libraries with drug-like specificity. This is particularly important for both *in vitro* antibody discovery and affinity maturation given that it is only possible to sample an extremely small fraction of maximum CDR chemical diversity, and it is critical to focus the CDR diversity on combinations of residues that give rise to drug-like properties. We also expect that our specificity rules will provide powerful guidelines for both identifying antibody candidates with high specificity during early antibody discovery and re-engineering existing antibodies with drug-

like properties later in the optimization and development process. More generally, we expect that our novel conceptual framework – which can be readily expanded in the future to include additional structural information and incorporate additional biophysical data sets – will accelerate the generation of potent antibody therapeutics with drug-like properties.

Acknowledgements

We thank Laura Walker for providing antibody sequences and PSR measurements of non-specific antibody binding. We thank Tushar Jain for providing support and guidance in using his machine learning methods for calculating antibody solvent assessible surface areas. We also thank Laura Walker, Tushar Jain, Mark Julian and Charles Starr for providing valuable feedback on multiple versions of this manuscript. This work was supported by the National Institutes of Health (R01GM104130, R01AG050598 and RF1AG059723 to P.M.T., T32 fellowship to L.W.), National Science Foundation (CBET 1813963, 1605266 and 1804313 to P.M.T., Graduate Research Fellowships to L.A.R. and M.D.S), and the Albert M. Mattocks Chair (to P.M.T).

Note: This chapter is adapted from manuscript titled 'Physicochemical rules for identifying monoclonal antibodies with drug-like specificity.

Citation: Zhang, Y., Wu, L., Gupta, P., Desai, A. A., Smith, M. D., Rabia, L. A., ... & Tessier, P. M. (2020). Physicochemical rules for identifying monoclonal antibodies with drug-like specificity. *Molecular Pharmaceutics*, *17*(7), 2555-2569

Chapter 3 **Mutations in Antibody Complementarity-determining Regions that Strongly Reduce Non-specific Binding also Significantly Increase Humanness**

Abstract

The development of safe and effective therapeutic antibodies necessitates robust methods for isolating human or humanized antibodies with drug-like properties, including high affinity, stability and solubility as well as low levels of off-target binding. In particular, antibodies with high levels of off-target binding have an increased risk for purification challenges *in vitro* and fast antibody clearance *in vivo*. Here we report a facile approach for first identifying sites in antibody complementarity-determining regions (CDRs) that are linked to high levels of off-target binding and then designing libraries of mutations at such sites aimed at reducing non-specific binding while increasing humanness. This approach generally seeks to modestly reduce the density of positively charged and hydrophobic CDR residues while modestly increasing the density of negatively charged and hydrophilic CDR residues based on the predicted impacts of specific CDR sequences on antibody non-specific binding. Using emibetuzumab as an example of a clinical-stage antibody with high levels of off-target binding, we designed a CDR library of variants, displayed the Fab library on yeast, sorted for variants with reduced non-specific binding while maintaining high affinity, and deep sequenced the enriched libraries. We identified a panel of antibody variants that retain high affinity while significantly reducing non-specific binding. Surprisingly, sets of mutations most strongly correlated with reduced non-specific binding resulted in significantly increased antibody humanness, while similar sets of mutations that were poorly correlated with

reduced non-specific binding significantly reduced humanness. Modeling analysis revealed that the reduced non-specific binding of the emibetuzumab variants was mainly due to disruption of hydrophobic and positively charged patches in the heavy chain CDRs. We expect that this systematic approach of co-optimizing antibody affinity, non-specific binding and humanness will simplify the engineering of humanized antibody therapeutics with drug-liike properties.

Introduction

Antibody therapeutics are being widely used to treat diverse human diseases, including cancer, autoimmune disorders, viral infections and neurological disorders (29,30). There are a staggering number of requirements that antibody candidates must meet to advance to the clinic. On the molecular level, antibody therapeutics must meet a number of functional (affinity, potency), developability (solubility, physical and chemical stability, viscosity, non-specific binding, self-association,) and safety (humanization, immunogenicity) requirements in the development pipeline (1,68-70). Of the many developability properties, a previous study of clinical-stage antibodies revealed that one of their most distinguishing qualities is that they display lower levels of non-specific and self-interactions relative to antibody candidates either in clinical trials or that failed in clinical trials (4). A key challenge in co-optimizing these developability properties and functional properties (e.g., affinity) is that both types of properties are highly interconnected because they are mediated primarily by solvent-exposed residues in complementarity-determining regions (CDRs). For example, several previous reports have demonstrated that CDR mutations that reduce antibody non-specific or self-interactions also reduce affinity,(38,71) which highlights the overlapping nature of these properties. Moreover, given that antibody candidates need to be as human as possible to reduce immunogenicity concerns, it would be valuable if methods could be

developed that enable co-optimization of antibody function (e.g., affinity), developability (e.g., non-specific binding) and safety (e.g., humanness).

We have developed a bioinformatics approach for identifying antibodies with increased risk of high levels of non-specific and self-interactions in physiological conditions (PBS, pH 7.4).(72) Our approach evaluates the number of specific amino acids in variable fragment (Fv) of antibodies, including the individual variable regions (V_H or V_L) or one or more of the CDRs, weighted by their predicted solvent exposure. This approach generally flags the Fv regions of antibodies with either i) over enrichment of positively charge and hydrophobic residues or ii) under representation of negatively charged and hydrophilic residues. Using twelve molecular flags that limit either the maximum (six flags) or minimum (six flags) numbers of solvent-exposed Fv residues, this approach flags most (78%) of the clinical-stage antibodies with high levels of non-specific and self-interactions while flagging few of the corresponding antibodies with low levels of non-specific and self-interactions (8%). Notably, most (7 out of 12) of the molecular flags are based only on the CDR sequences, which suggests that sites that mediate non-specific and self-interactions will strongly overlap with sites that mediate antibody affinity. This raises important questions about the feasibility of re-engineering antibodies to display reduced non-specific and self-interactions without compromising their function.

Therefore, we have sought to develop a methodology for co-optimizing antibody affinity and non-specific binding that addresses key challenges associated with maintaining affinity while introducing CDR mutations needed to reduce non-specific binding. Our approach, which is summarized in Figure 3-1, is to first predict sites in antibody CDRs that mediate non-specific binding, and then design combinatorial libraries of mutations at such CDR sites that are predicted to reduce non-specific binding. These sets of mutations are also chosen based on the most frequent

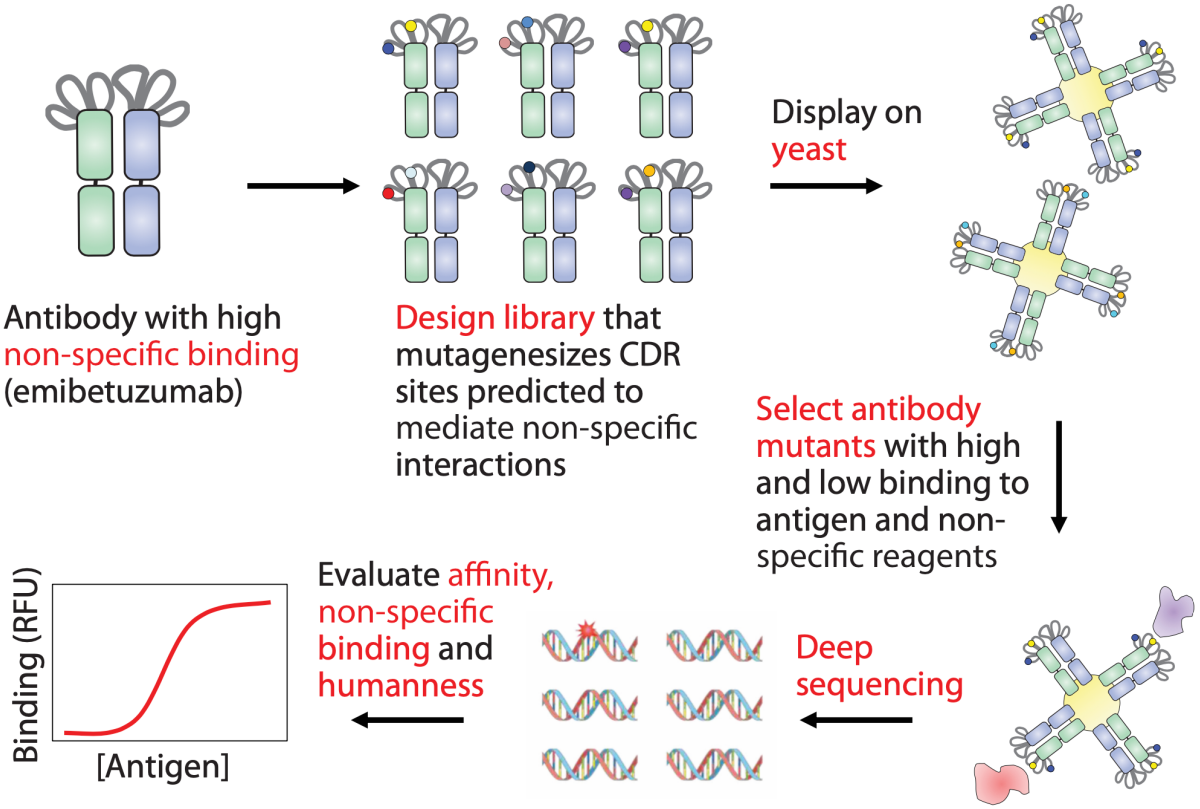


Figure 3-1： Overview of molecular engineering approach used to reduce the non-specific binding and improve the humanness of a therapeutic antibody (emibetuzumab) while maintaining high affinity.

Eight sites in the heavy chain complementarity-determining regions (CDRs) were first identified that were predicted to mediate non-specific binding, and focused libraries were designed that were predicted to reduce non-specific binding while maximizing humanness. Next, the libraries were displayed on the surface of yeast as single-chain Fabs, and the libraries were sorted for high affinity and high and low levels of non-specific interactions. Finally, the enriched libraries were deep sequenced and IgGs were identified that have desirable combinations of properties, including high affinity, low off-target binding and high humanness.

residues observed at each CDR site in human antibodies to maximize the humanness of the antibody libraries. Next, we generate these libraries as Fab fragments on yeast, and sort them for high affinity and low non-specific binding using fluorescence-activated cell sorting (FACS). Finally, we deep sequence the libraries and identify clones with favorable combinations of affinity and non-specific binding. Herein, we report our integrated approach for co-optimizing antibody and non-specific binding for a clinical-stage antibody (emibetuzumab) with high levels of non-specific binding and demonstrate that mutations strongly linked to reduce antibody non-specific binding are strongly correlated with increased humanness.

Experimental Section

Antibody sequence datasets. The synthesis and initial stages of sorting (sorts 1-3) of the emibetuzumab single-chain Fab (scFab) library on yeast were described previously.(72) Briefly, the antibody library was sorted against antigen (HGFR-Fc) for two rounds using MACS. Next, in round 3, the library was sorted for low and high levels of non-specific binding against two reagents (PSR and ovalbumin) or for high affinity against antigen, and the enriched antibody variants with high or low specificity and high affinity were evaluated using deep sequencing. The library was also further sorted for high affinity (positive antigen selections in rounds 4 and 5) or low levels of non-specific binding (negative selections against PSR or ovalbumin in round 4), as summarized in Figure 3-7. For antigen selections by FACS), the HGFR-Fc concentration was 1 nM [PBS with 1 mg/mL BSA (PBSB) and 1% milk]. The MACS selections were performed using HGFR-Fc immobilized on Protein A Dynabeads (Invitrogen, 10001D). The antigen was immobilized on the beads as per manufacturer's instruction. For the non-specific binding selections (FACS), the concentration of ovalbumin was 0.26 mg/mL (PBSB) and the concentration of PSR was 0.13 mg/mL (PBSB).

The emibetuzumab deep sequencing data from round 3 of sorting was analyzed to identified antibody variants in four library samples in two independent replicates: (i) input library, 5.3×10^5 variants in the first replicate and 2.8×10^5 variants in the second replicate; (ii) enriched library sorted for negative ovalbumin binding (OVA-), 3.4×10^5 variants in the first replicate and 1.9×10^5 variants in the second replicate; (iii) enriched library sorted for negative PSR binding (PSR-), 1.7×10^5 variants in the first replicate and 1.9×10^5 in the second replicate; and (iv) enriched library sorted for positive antigen (HGFR-Fc) binding, 6.9×10^4 variants in the first replicate and

4.9×10^4 variants in the second replicate. This resulted in a total of 3,465 unique scFabs (antibody set #1) that were present in all the eight samples.

Two additional sets of Fabs was identified by analyzing the deep sequencing data from round 3 in six library samples in two independent replicates: (i) input library, 5.3×10^5 variants in the first replicate and 2.8×10^5 variants in the second replicate; (ii) enriched library sorted for positive ovalbumin binding (OVA+), 3.8×10^5 variants in the first replicate and 2.2×10^5 variants in the second replicate; (iii) enriched library sorted for positive PSR binding (PSR+), 5.3×10^5 variants in the first replicate and 1.8×10^5 variants in the second replicate; (iv) enriched library sorted for negative ovalbumin binding (OVA-), 3.4×10^5 variants in the first replicate and 1.9×10^5 variants in the second replicate; and (v) enriched library sorted for negative PSR binding (PSR-), 1.7×10^5 variants in the first replicate and 1.9×10^5 in the second replicate. This resulted in a total of 3418 unique scFabs (antibody set #2) that were present in enriched library samples (i)-(iii) for the low specificity group and 4814 unique scFabs (antibody set #3) that were present in the enriched library samples (i) and (iv-v) for the high specificity group with enrichment ratio greater than 1.

The human dataset (antibody set #4) for evaluating level of humanness was composed of three subsets of human antibodies with paired V_H/V_L sequences: (i) 1.4×10^4 human antibodies from a previous publication;(73-77) (ii) 3.7×10^4 human antibodies from the Observed Antibody Space (OAS) database;(74-77) (iii) 6.5×10^4 naïve and antigen-experienced human antibodies from a previous study.(78) The variable heavy (V_H) and variable light (V_L) sequences were processed using the Antigen receptor Numbering And Classification (ANARCI) method (79) to assign Kabat numbering. Antibodies were eliminated if the number of amino acids in their variable regions or CDRs exceeded preset maximum limits: 134 residues for V_H; 7 for HCDR1; 19 residues for

HCDR2; 21 residues for HCDR3; 122 residues for V_L; 17 residues for LCDR1; 15 residues for LCDR2 and 12 residues for LCDR3. This resulted in 9.4×10^4 human antibodies in antibody set #4.

Mutational bioinformatics analysis. For antibody set #1, antibody subsets were identified for each possible two mutation combination observed in the library. For a given set of two mutations (2MT; e.g., T54 and D56 in V_H), all antibodies with wild-type residues at the same two sites (2WT; e.g., R54 and G56 in V_H) were also isolated regardless of whether they possess mutant or wild-type residues at the other six sites that were mutated in the library. Next, the enrichment ratios were calculated for each antibody based on the frequency of the clone in the low non-specific FACS gates (PSR- and ovalbumin-) relative to the input frequencies for two independent replicates. Moreover, we required at least ten antibodies with a given set of mutations and at least ten antibodies with wild-type residues for further analysis. To identify pairs of mutations that were most strongly correlated with enrichment for low non-specific binding, Spearman's rho values were evaluated for the relationship between % mutant antibodies with a given unique enrichment ratio and the corresponding enrichment ratio value. Sets of mutations with the strongest correlations for reduced non-specific binding were those with Spearman's rho values >0.5 and corresponding *p*-values <0.01. This same analysis was performed for sets of three mutations except that the requirement for Spearman's rho values were >0.55.

Humanness analysis. Two methods were used to evaluate the level of humanness for a given antibody variant based on the variable (V_H and V_L) domains. The first approach, which is referred to as the Human String Content (HSC) method(27), involves first defining a string (scanning window) length for comparing each peptide within the antibody of interest to a reference set of human antibodies. In this work, a string length of nine residues was used for calculating the sequence identity between each possible 9mer peptide and all possible 9mer peptides at the same

aligned positions in a reference set of human antibodies (antibody set #4 in this work). The formula that was used is:

$$HSC = 100 \times \frac{1}{(L-9+1)\times 9} \sum_{i=1}^{L-9+1} \left(\sum_{j=i}^{i+9-1} \delta_{aa_j^s, aa_j^h} \right)$$

where L is the length of the target sequence, i is the first position in the string of the 9mer peptide, aa_j^s is the amino acid at position j of the target antibody s, and aa_j^h is the amino acid at position j of human antibody h. In this work, all possible 9mer peptides in variable heavy and light domains were used to calculate the HSC score. This analysis only excludes the first eight positions in V_L and last eight positions in V_H.

The second method that was used to evaluate the level of antibody humanness, which is referred to as the T20 method,(28) uses global sequence similarity instead of local identity between small peptide regions for the HSC method. A normalized sequence similarity score was used to evaluate the difference between the antibody of interest and each human antibody in the reference set (antibody set #4 in this work). The sequence similarity score was determined by sum of blocks substitution matrix (BLOSUM) score for each amino acid pair of the aligned sequences (80). The gap extension and gap opening for comparing two aligned sequences were set as -1 and -5, respectively. The sequence similarity score was then normalized by the average of self-sequence similarity scores for each pair of antibodies. Finally, the T20 humanness score for the target antibody was determined by the average of the top twenty normalized similarity scores (28).

Molecular feature analysis for identifying antibody variants with high and low specificity. Molecular features were evaluated that differentiate between antibodies in sets #2 and #3. The molecular features that were considered fall into three categories, namely (i) sequence-based (SB) features, (ii) sequence-based features weighted by the predicted solvent exposure of each amino

acid (SB-SE), and (iii) sequence-based features weighted by predicted solvent exposures of each amino acid and their relative side chain surface areas (SB-SE-SA). SB features included several different types of features, including the number of: (a) hydrophobic amino acids (Phe, Ile, Leu, Pro, Val, Trp, Met and Tyr), (b) positively charged amino acids (Arg, Lys, and His), (c) negatively charged amino acids (Asp and Glu), (d) polar residues (Gln, Ser, Thr, Asn and Tyr), (e) combination of (a) and (b), (f) combination of (c) and (d), and (g) net charge at pH 7.4 (+1 for Arg and Lys, +0.1 for His and -1 for Asp and Glu). The SB features were evaluated for each heavy chain CDR (H1, H2 and H3), the combined three heavy chain CDRs (H123) and the combined overall CDRs for both the heavy and light chains (H123 L123). The SB-SE features were the same as the 12 chemical rules reported previously (72), and the predicted solvent-accessible surface area (SASA) values were calculated using a previously reported method.(40) The SB-SE features were evaluated for the same antibody regions as the SB features as well as for V_H, V_L, Fv and combinations of the heavy and light chain CDRs. The SB-SE-SA features were calculated using the SB features (a-f) weighted by the predicted SASA of each residue (as done for the SB-SE features) and multiplied by the side chain surface area of each residue (based on A-X-A peptide calculations).(81) The SB-SE-SA features were evaluated for the same antibody regions as the SB features. Finally, the ability of these three sets of features to differentiate between antibody sets #2 and #3 in Fig. 6 were evaluated by logistic regression analysis and ranked in terms of their area under the curve (AUC) values.

Molecular feature analysis of improved emibetuzumab variants. Molecular features in Figure 3-11C were evaluated for their ability to selectively differentiate between a subset of antibodies in set #1 (identified in Figure 3-3) with either sets of two (2MT) or three (3MT) mutations most strongly correlated with improved antibody specificity (referred to as set #1A)

relative to their corresponding wild-type variants (2WT and 3WT; referred to as set #1B). This resulted in a total number of 1520 variants (973 variants in set #1A and 547 variants in set #1B). Molecular features were evaluated via logistic regression to identify features with the highest AUC values that were best at differentiating between sets #1A and #1B. The molecular features that were considered fall into two categories described above, namely (i) SB-SE and (ii) SB-SE-SA features. The SB-SE were evaluated as described above. The SB-SE-SA features in this analysis were more extensive, and included eight groups of features based on their properties in V_H ,V_L and Fv regions: (a) hydrophobic amino acid group 1 (Ala, Phe, Ile, Leu, Pro, Val, Trp, Met and Tyr); (b) hydrophobic amino acid group 2 (Phe, Ile, Leu, Pro, Val and Trp); (c) hydrophobic and positively charged amino acids (Phe, Ile, Leu, Pro, Val, Trp, Arg, Lys and His); (d) hydrophilic amino acid group 1 (Gln, Ser, Thr and Asn); (e) hydrophilic amino acid group 2 (Gln, Ser, Thr, Asn and Tyr); (f) hydrophilic and negatively charge amino acids (Gln, Ser, Thr, Asn, Asp and Glu); and (g) aromatic amino acids (Phe, Trp and Tyr); (h) aromatic amino acids and proline (Phe, Trp, Tyr and Pro). The SB-SE-SA features were otherwise calculated as described above. Finally, the features that most strongly differentiated between antibodies in sets #1A and #1B were selected if their AUC values from logistic regression analysis were >0.7 and the average of normalized feature values for four emibetuzumab variants (43-01, EM-02 and 27-10) were reduced by >15%.

Antibody homology models and visualization. The antibody homology models were generated using Molecular Operating Environment (MOE) software with the Amber10:EHT forcefield and with a dielectric constant value of 4. The MOE antibody modeler was used to build an initial homology model based on searching templates from a Fab/antibody structure database, including antibody structures in Protein Data Bank (PDB), for V_H, V_L and individual CDR loops. A model for wild-type emibetuzumab and each variant were generated first by grafting each CDR

loop template onto the framework templates for V_H and V_L. The V_H and V_L chains of PDB:4LIQ were selected to be the framework templates for emibetuzumab. The templates were chosen from PDB:3W2D for light chain CDR loops. The templates were chosen from PBD:4YHL for heavy chain CDR 1 and 2 were and from PDB:5NHW for heavy chain CDR 3. Finally, the initial antibody model was energy minimized with a minimum gradient setting of 0.00001 RMS kcal/mol/A^2. The final homology models for the emibetuzumab variants were then exported to Pymol for visualization.

Antibody cloning and production. DNA fragments encoding V_L and V_H sequences were ordered as geneblocks (IDT) for selected antibodies and cloned into pTT5 mammalian expression plasmids containing a common IgG1 heavy and light chain (kappa) frameworks, as described previously.(82) Briefly, the PCR-amplified fragments and expression vectors were digested with desired restriction enzymes (*EcoRI-HF* and *NheI-HF* for V_H; *EcoRI-HF* and *BsiWI-HF* for V_L; New England Biolabs). Finally, digested DNA fragments and vectors were purified (Qiagen, 28104), ligated with T4 (New England Biolabs, M0202L) and transformed into competent DH5α cells. Antibody sequences were confirmed by Sanger sequencing.

Antibodies were expressed with the HEK293-6E cell line (L-11565, National Research Council of Canada). Cells were cultured in F17 media (Thermo Fisher, 50591354) with Kolliphor (Thermo Fisher, NC0917244), Glutamine (Invitrogen, 25030081) and G418 (Thermo Fisher, 10131035) until the density reached approximately 2-3 million viable cells per ml. Cells (25 mL) were then transfected with 15 μg DNA (7.5 μg of light chain plasmid and 7.5 μg of heavy chain plasmid) and 45 μg polyethylenimine (PEI MAX, Polysciences Inc., 247651). 20% w/v yeastolate (Gibco, 292804) was added after 24 to 48 h post transfection. On day 5, culture supernatant was collected by centrifuging at 3500xg for 40 min. mAbs were then purified with Protein A agarose

resin (Pierce, 20034). Protein A beads (1 mL) were added to each tube and the mixture was gentle rocked overnight at 4 °C. After incubation, Protein A beads were collected with vacuum filter columns (Thermo Fisher Scientific, 89898), and washed with 50-100 mL of PBS. Glycine buffer (0.1 M, pH 3.0) was then used to elute antibodies followed by adding 1 M Tris to neutralize the pH to 7.4. Next, protein concentrations were measured via absorbance measurements at 280 nm and purity was evaluated by SDS-PAGE (Thermo Fisher Scientific, WG1203BOX).

All antibodies analyzed in this study were second-step purified using size-exclusion chromatography (SEC) with a Shimadzu Prominence HPLC System outfitted with a LC-20AT pump, SIL-20AC autosampler and FRC-10A fraction collector. After purification with Protein A, antibodies were loaded onto a SEC column (Superdex 200 Increase 10/300 GL column; GE, 28990944) and analyzed at 0.75 mL/min using a PBS running buffer supplemented with 200 mM arginine (pH 7.4). Absorbance was monitored at 220 and 280 nm, and the 280 nm signal was primarily used for analysis. Fractions were collected during times correlating to monomeric protein elution based on a mAb standard. The fractions were collected, buffer exchanged into PBS (pH 7.4), filtered, aliquoted, and stored at -80 °C.

Affinity analysis. Individual emibetuzumab clones (wild type, EM-02, 43-06, 43-01 and 27-10) were evaluated as single-chain Fabs on the surface of yeast. Antigen affinity was evaluated via flow cytometry by measuring the binding of 10^6 yeast cells to antigen (HGFR-Fc; Acro Biosystems, MET-H5256) at a range of antigen concentrations in PBSB. The incubation volume was adjusted to ensure that the molar concentration of HGFR-Fc exceeded that of the surface-displayed scFab (assuming 50,000 Fabs per cell) by at least tenfold. Detection of antigen binding (1:1000; Jackson ImmunoResearch, 109-606-008) and antibody display (1:500; Southern Biotech, 2062-30) was performed after incubation for 15 min on ice in PBSB. Average EC_{50} values were

calculated by fitting the median antigen binding signal of the population of antibody-expressing yeast cells to the three-parameter Hill equation.

Polyspecificity analysis. The polyspecificity reagent (SMP) was prepared as previously described.(9) CHO cells (10^9, Gibco, A29133) were pelleted, the cell pellets were washed separately with PBSB and Buffer B (50 mM HEPES, 0.15 M NaCl, 2 mM $CaCl_2$, 5 mM KCl, 5 mM $MgCl_2$, 10% Glycerol, pH 7.2), and then pelleted again. The pellets were resuspended in 5 mL of Buffer B supplemented with a protease inhibitor (Sigma Aldrich, 4693159001). Next, the resuspended cells were homogenized for 90 s (three cycles of 30 s). The cell suspension was then spun down at 40000xg for 1 h and the supernatant was discarded. The pellet, containing the membrane proteins, was resuspended in Buffer B with a Dounce homogenizer. The protein concentration was determined using a detergent-compatible protein assay kit (BioRad, 5000116). The suspension was diluted to a theoretical concentration of 1 mg/mL in solubilization buffer (pH 7.2), which contained 50 mM HEPES, 0.15 M NaCl, 2 mM $CaCl_2$, 5 mM KCl, 5 mM $MgCl_2$, 1% n-dodecyl-□-D-maltopyranoside (Sigma Aldrich, D4641), and a protease inhibitor (Sigma Aldrich, 11873580001). The solution was then mixed overnight at 4 °C. The soluble membrane protein (SMP) fraction was centrifuged at 40000xg for 1 h and the supernatant was collected. The final concentration of the supernatant was diluted to 1.0 mg/mL.

Sulfo-NHS-LC-biotin (Thermo Fisher, PI21335) was dissolved in distilled water at ~11.5 mg/mL. Stock solution of Sulfo-NHS-LC-biotin (150 μL) and the SMP reagent (4.5 mL at 0.8-0.9 mg/mL) were mixed via end-over-end mixing at room temperature (45 min). The reaction was quenched (10 μL of 1.5 M hydroxylamine at pH 7.2), and biotinylated SMP was aliquoted and stored at -80 °C. Ovalbumin (Sigma, A5503) was dissolved at 5 mg/mL in PBS. Sulfo-NHS-LC-biotin (Thermo Fisher, PI21335) was dissolved in distilled water at ~11.5 mg/mL. For

biotinylation, the stock solution of Sulfo-NHS-LC-biotin (590 μL) was added to ovalbumin (5 mL) and incubated for 30 min at room temperature. The reactions were quenched with hydroxylamine (10 μL of 1.5 M at pH 7.2). The reagent was aliquoted and stored at -80 °C for up to six months.

The specificity assay was performed as previously described. Protein A magnetic beads (Invitrogen, 88846) were washed with PBSB and incubated with antibodies overnight at 4 °C. Next, the protein-coated beads were washed and resuspended with a 10x diluted solution of biotinylated SMP or biotinylated ovalbumin and incubated on ice for 20 min (SMP) or 3 h at room temperature (ovalbumin). Beads were again washed with PBSB and incubated with streptavidin AF-647 (Invitrogen, S32357) and 1000x diluted solution of goat anti-human Fc F(ab')2 AF-488 (Invitrogen, H10120) on ice (4 min). Bead were washed, resuspended in PBSB, and analyzed via flow cytometry. Three independent repeats were performed with all results normalized between the parental wild-type IgG (emibetuzumab) and a high specificity control IgG (elotuzumab).

Results

Identification of mutations in antibody CDRs that strongly reduce non-specific binding. Toward our goal of developing a systematic approach for co-optimizing antibody affinity, non-specific binding and humanness, we first generated a sub-library of emibetuzumab V_H mutants in which the heavy chain CDRs were mutated at eight sites (Figure 3-2A). The sites for mutagenesis were selected based on four requirements for the heavy chain CDR residues: 1) hydrophobic or basic; 2) flagged by at least one of the six maximum chemical rules for identifying mAbs with poor specificity that we reported previously(72); 3) solvent exposed [solvent-exposed surface area (SASA) >0.1 except for glycine (no SASA requirement)]; and 4) not highly conserved [average frequency in human antibodies <50% (as defined by the abYsis database(83))]. Only one residue (K62) that meets the four requirements was not mutated because it was distant from other eight

residues (Figure 3-2B). The residues that were sampled at each site in the library are shown in Figure 3-2C, which were chosen based on three factors: 1) one or more mutations at each site were predicted to reduce non-specific binding; 2) mutations sample one or more residues common to human antibodies at the same CDR site; and 3) ability to encode the mutations using degenerate codons. The library was library was generated as Fab fragments displayed on yeast and sorted for high antibody binding and low and high non-specific binding to two non-specific reagents [ovalbumin and soluble membrane proteins (referred to as PSR for polyspecificity reagent(9)], and the enriched library samples were deep sequenced.(72)

To identify sets of mutations that are most strongly linked to reduced non-specific binding, we processed the deep sequencing data to isolate a set of 3,465 unique Fabs (antibody set #1) that were observed in two independent replicates of the input libraries and the enriched libraries sorted for positive antigen binding and negative binding to both ovalbumin and PSR (8 library samples in total; see Methods for more detail). Next, we evaluated sets of two mutations that were most strongly correlated with reduced non-specific binding using a five step process: 1) identify a given set of two CDR mutations observed in antibody set #1; 2) isolate all antibodies that possess the two CDR mutations regardless of whether they possess wild-type or mutant residues at the other six sites; 3) isolate all antibodies that possess the corresponding two wild-type residues regardless of whether they possess wild-type or mutant residues at the other six positions; 4) evaluate the enrichment ratios for each antibody based on their frequency in the libraries enriched for low non-specific binding relative to the corresponding input libraries; and 5) evaluate the Spearman's ρ values for the relationship between the % antibodies with a given set of two mutations and the corresponding enrichment ratios. We defined sets of two mutations that are strongly linked to

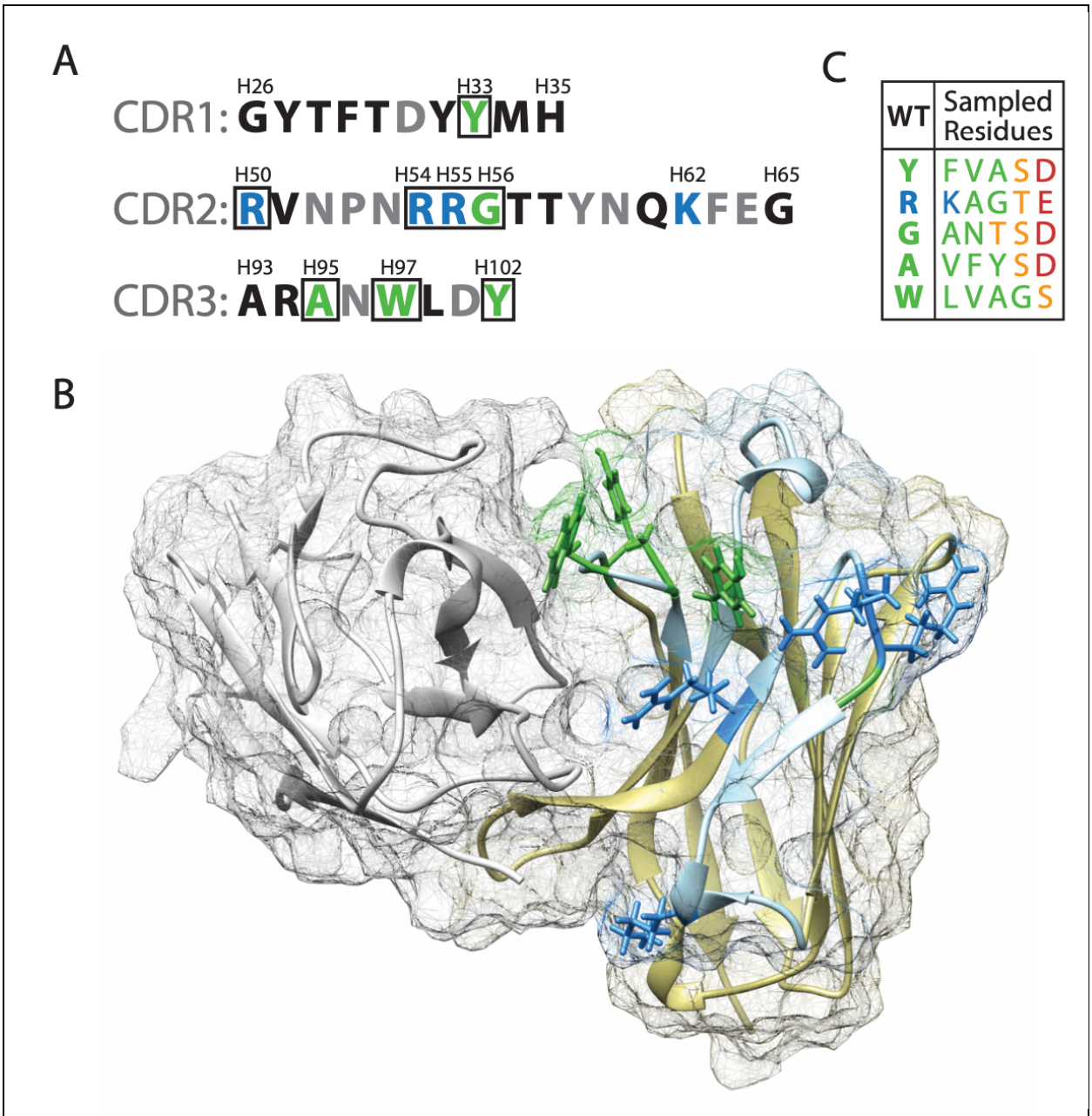


Figure 3-2: Design of emibetuzumab V_H sublibrary aimed at reducing non-specific binding and increasing humanness.

(A) Heavy chain CDR sequences of emibetuzumab identified for mutagenesis, as indicated by boxed sites and colored based on their properties (blue for basic residues and green for aliphatic and aromatic residues). The identified sites were selected based on four requirements for the heavy chain CDR residues: 1) hydrophobic or basic; 2) flagged by at least one of the six maximum chemical rules for identifying mAbs with poor specificity; 3) solvent exposed [SASA >0.1 except for Gly (no SASA requirement)]; and 4) not highly conserved [average frequency in human antibodies <0.5 (as defined by the abYsis database)]. The sites that were flagged by at least one maximum chemical rule – but do not meet all of the other requirements – are indicated in bold black text. Only one residue (K62) that meets the four requirements was not mutated because it was distant from other eight residues. (B) Residues sampled at each selected site in the antibody sublibrary are shown, which include the wild-type residue and five mutations encoded by degenerate codons that are predicted to reduce non-specific binding while increasing humanness. (C) Homology model of emibetuzumab Fv regions with nine sites that meet the four design requirements highlighted in blue (basic residues) and green (aliphatic and aromatic residues). The highlighted blue (lysine) residue at position 62 at the bottom of the V_H structure was not mutated.

reduce non-specific binding as those with Spearman's ρ values were >0.5 and corresponding *p*-values <0.01.

We find only two sets of two mutation combinations that are strongly linked to reduced non-specific binding (Figure 3-3A). Interestingly, both sets of mutations involve replacement of a non-charged residue in HCDR2 (glycine at position 55) with a negatively charged residue (aspartic acid). Moreover, the other mutation in each case replaces a positively charged residue in heavy chain CDR2 (arginine at positions 54 or 55) with a non-charged residue (threonine at position 54 or glycine at position 55). This change in net charge of -2 in HCDR2 appears to be critical to strongly reducing non-specific binding to both ovalbumin and soluble membrane proteins (PSR). For example, there are 120 antibodies with wild-type residues at positions 54 (Arg) and 56 (Gly) in HCDR2, and 203 antibodies with mutant residues at the corresponding positions (Thr54 and Asp56) regardless of the mutant or wild-type residues at the other six positions. The distributions of these antibodies in terms of their enrichment ratios for low non-specific binding (i.e., enrichment into negative FACS gates for low non-specific binding) are shown in Figure 3-2B (ovalbumin) and 3-2C (PSR). These distributions reveal that antibodies with such mutants are significantly enriched for low non-specific binding relative to antibodies with wild-type residues at the same CDR positions. This finding is notable because the identities of the residues at the other six mutated sites were ignored in this analysis, suggesting that the effects of the set of two mutations (e.g., Thr54 and Asp56 in HCDR2) are dominant relative to the other mutations sampled in the library. We also repeated this analysis for sets of three mutations and found four sets of mutations most strongly linked to reduced non-specific binding (Figure 3-3D-F). Interestingly, each of them also possessed the same Gly-to-Asp mutation at position 56 in HCDR2 observed for the two mutation sets, while three of the four of them possessed mutations in HCDR3 as well.

A

	H1	H2				H3		
	33	50	54	55	56	95	97	102
WT	Y	R	R	R	G	A	W	Y
1			T		D			
2				G	D			

D

	H1	H2				H3		
	33	50	54	55	56	95	97	102
WT	Y	R	R	R	G	A	W	Y
1			T	E	D			
2				G	D		L	
3					D		L	S
4					D		S	A

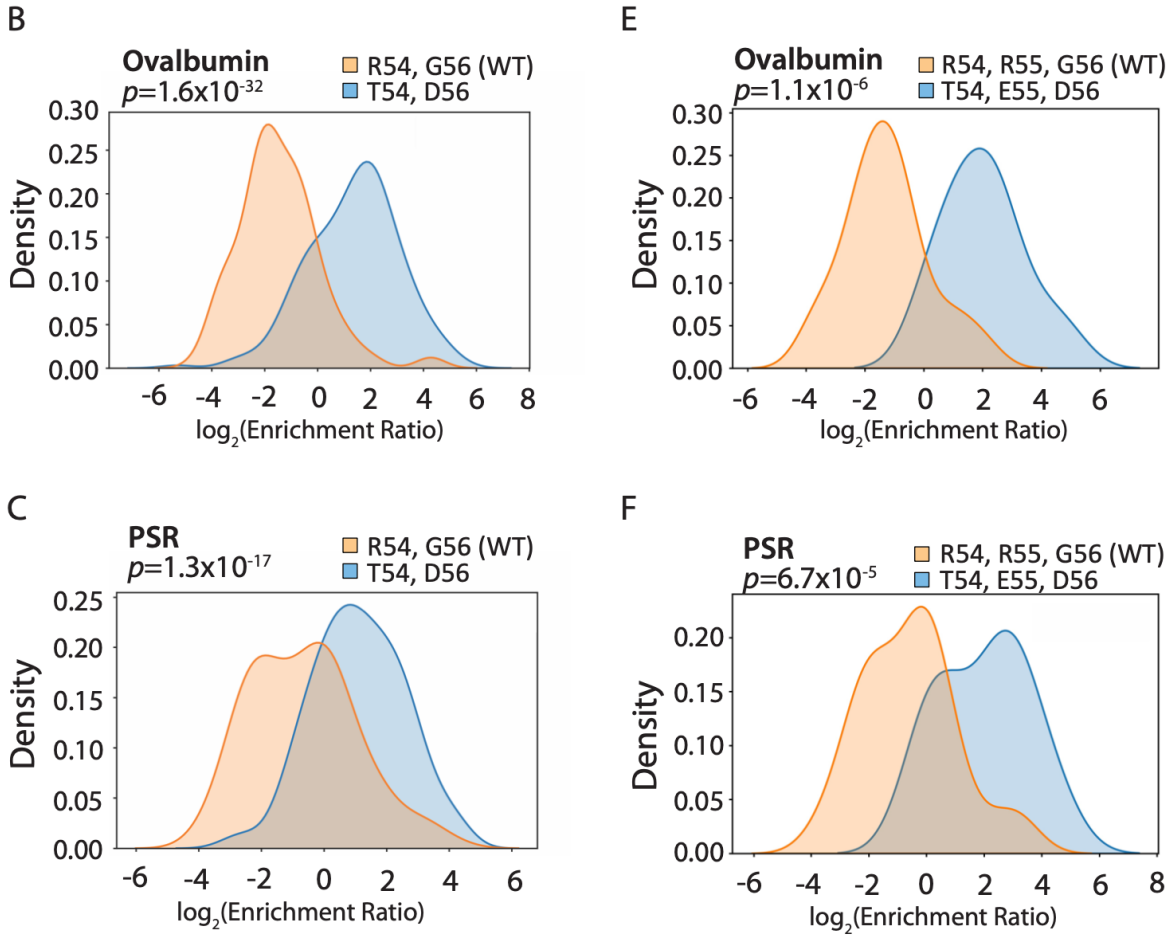


Figure 3-3：Sets of two and three mutations in the heavy chain CDRs of emibetuzumab that are most strongly correlated with reduced non-specific binding.

(A) Identified sets of two mutations that are strongly correlated with reducing non-specific binding. (B, C) Enrichment ratio distributions for antibody variants with either a pair of two mutations (T54 and D56 in V_H; blue) or the corresponding pair of wild-type residues (R54 and G56 in V_H; orange) for variants enriched for low levels of non-specific binding to (B) ovalbumin and (C) soluble membrane proteins (referred to as polyspecificity reagent or PSR). (D) Identified sets of three mutations strongly correlated with reduced non-specific binding. (E, F) Enrichment ratio distributions for antibody variants with either a set of three mutations (T54, E55 and D56 in V_H; blue) or a set of the corresponding wild-type residues (R54, R55 and G56 in V_H; orange) for antibodies sorted for low levels of non-specific binding to (E) ovalbumin and (F) PSR. In (A-F), the antibody variants that were analyzed were only required to have either the mutant or wild-type residues at the identified positions regardless of the residues at the other five or six sites that were mutated. In (B), (C), (E) and (F), the *p*-values were evaluated using the Mann-Whitney *U* test to evaluate significance for segregating the two distributions.

We next evaluated if the experimentally identified sets of mutations that strongly reduce non-specific binding were consistent with the predictions used to design the libraries (Figure 3-4). Strikingly, we observed that the sets of mutations experimentally identified to reduce non-specific binding led to excellent segregation of antibody mutants with low (e.g., 1-4 flags out of a maximum of 12 flags; wild type emibetuzumab has 8 chemical flags) and intermediate (e.g., 6-8 flags) numbers of chemical flags. For example, the set of antibodies with Thr-54 and Asp-56 in HCDR2 that strongly reduced non-specific binding also strongly reduced the number of chemical flags (2-5 flags) relative to the corresponding set of antibodies with wild-type residues at these two positions (5-8 flags; Figure 3-4A). These results are even more striking for the sets of antibodies with three mutations (Figure 3-4B).

We also observed similar results for a second set of two and three CDR mutations (Figure 3-5). Importantly, we also performed similar analysis for sets of two and three mutations that are poorly correlated with reduced non-specific binding by performing the same analysis as done for Figure 3-3 but instead requiring that the correlations are not significant (p-value <0.05) with Spearman's ρ values of $< |0.1|$. Notably, we observe little difference between the distributions in the number of chemical flags for such panels of antibodies with two (Figure 3-4C) and three (Figure 3-4D) mutations, demonstrating the significance of the predictions of antibody CDR mutations that reduce non-specific binding.

CDR mutations that strongly reduce antibody non-specific binding also significantly increase humanness. We next evaluated if the sets of CDR mutations most strongly linked to reduce non-specific binding influenced the level of humanness (Figure 3-6). Emibetuzumab is a humanized antibody generated in mouse and has a *VH1-02* human germline. Given that CDR mutations were only in the heavy chain CDRs, we evaluated the humanness of the emibetuzumab

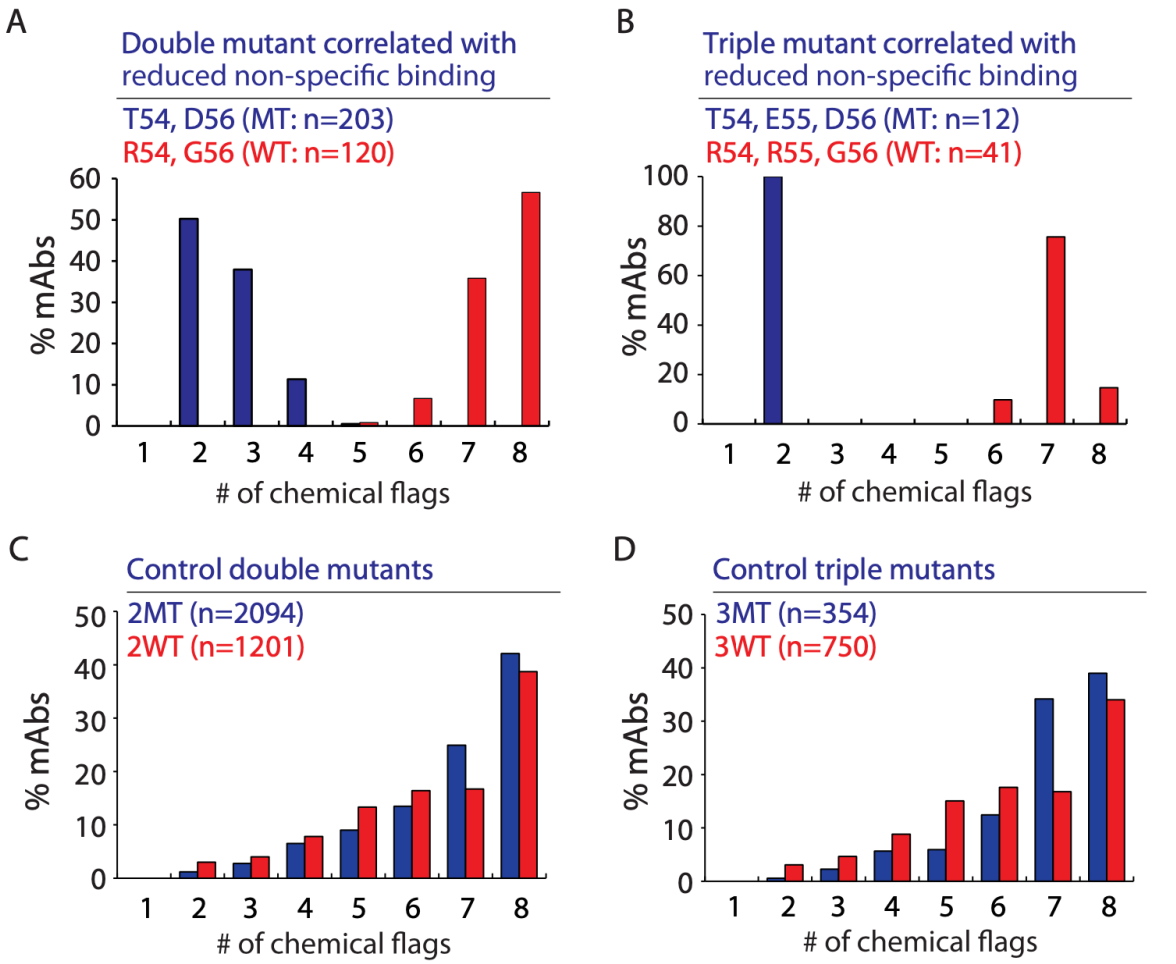


Figure 3-4: Emibetuzumab variants with sets of mutations most correlated with reduced non-specific binding are flagged significantly less by chemical rules that predict antibody specificity.

(A-B) Distributions of the number of chemical flags for antibody variants with (A) two or (B) three mutations strongly linked to reduced non-specific binding relative to the antibodies with wild-type residues at the corresponding CDR sites. (C-D) Distributions of the number of chemical flags for control antibody variants with (C) two or (D) three mutations that are not statistically correlated with reductions in non-specific binding relative to the antibodies with wild-type residues at the corresponding CDR sites. The number of antibody variants in this analysis are specified in the figure, which includes only antibodies with either specific mutations or wild-type residues regardless of the identities of the residues at the other five or six mutated sites in the heavy chain CDRs.

V_H region for the antibody mutants using a panel of 61,467 antibodies with unique V_H regions from antibody set #4, which included 12,358 *VH1* antibodies and 3,498 *VH1-02* antibodies. We first evaluated the human string score (HSC) based on the local sequence identity between each consecutive 9mer peptide of a given target antibody and all possible 9mer peptides at the same aligned positions in a reference set of human antibodies. For the HSC score, we first considered the 3,498 *VH1-02* antibodies in the reference set. Interestingly, we observed that sets of two CDR

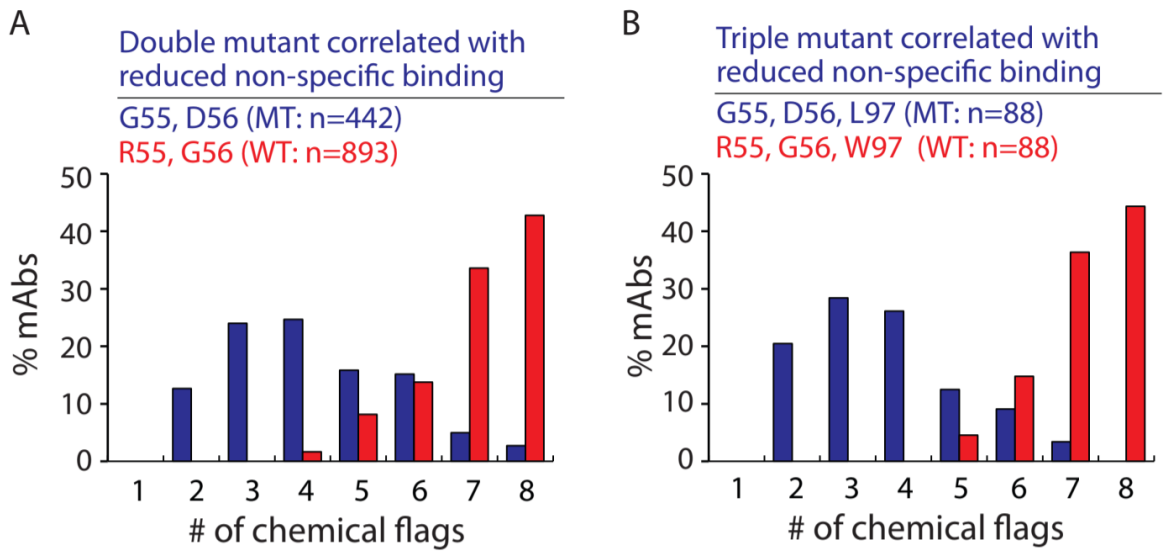


Figure 3-5：Emibetuzumab variants with additional sets of mutations most correlated with reduced non-specific binding are flagged significantly less by chemical rules that predict antibody specificity.

(A-B) Distributions of the number of chemical flags for antibody variants with (A) two or (B) three mutations strongly linked to reduced non-specific binding relative to the antibodies with wild-type residues at the corresponding CDR sites. The analysis was performed as described in Fig. 4 for unique sets of two and three mutations.

mutations most strongly linked to increased HSC scores were also strongly correlated to increased humanness, as shown for the distributions of HSC scores and relationship between HSC scores and the % of mAbs with sets of two mutations strongly linked to reduced non-specific binding (Figure 3-6A). We obtained similar results if we calculated the HSC scores using the 12,358 *VH1* antibodies (Figure 373). These results are even more notable when considering that sets of two mutations in the same antibody library that are poorly correlated with reduced non-specific binding (as described for Figure 3-4C and 4D) show the opposite trend, as such mutations are correlated with reduced humanness (Figure 3-6B and 3-7).

There are many different methods for evaluating antibody humanness, and we sought to test if our conclusions for HSC scores (Figure 3-4A-B and 3-7), which are based on local sequence identity, are general and observed for humanness scores based on global sequence similarity.

Therefore, we also evaluated humanness scores by calculating the average sequence similarities of the 20 most similar antibody sequences in a given set of human antibodies relative to each to

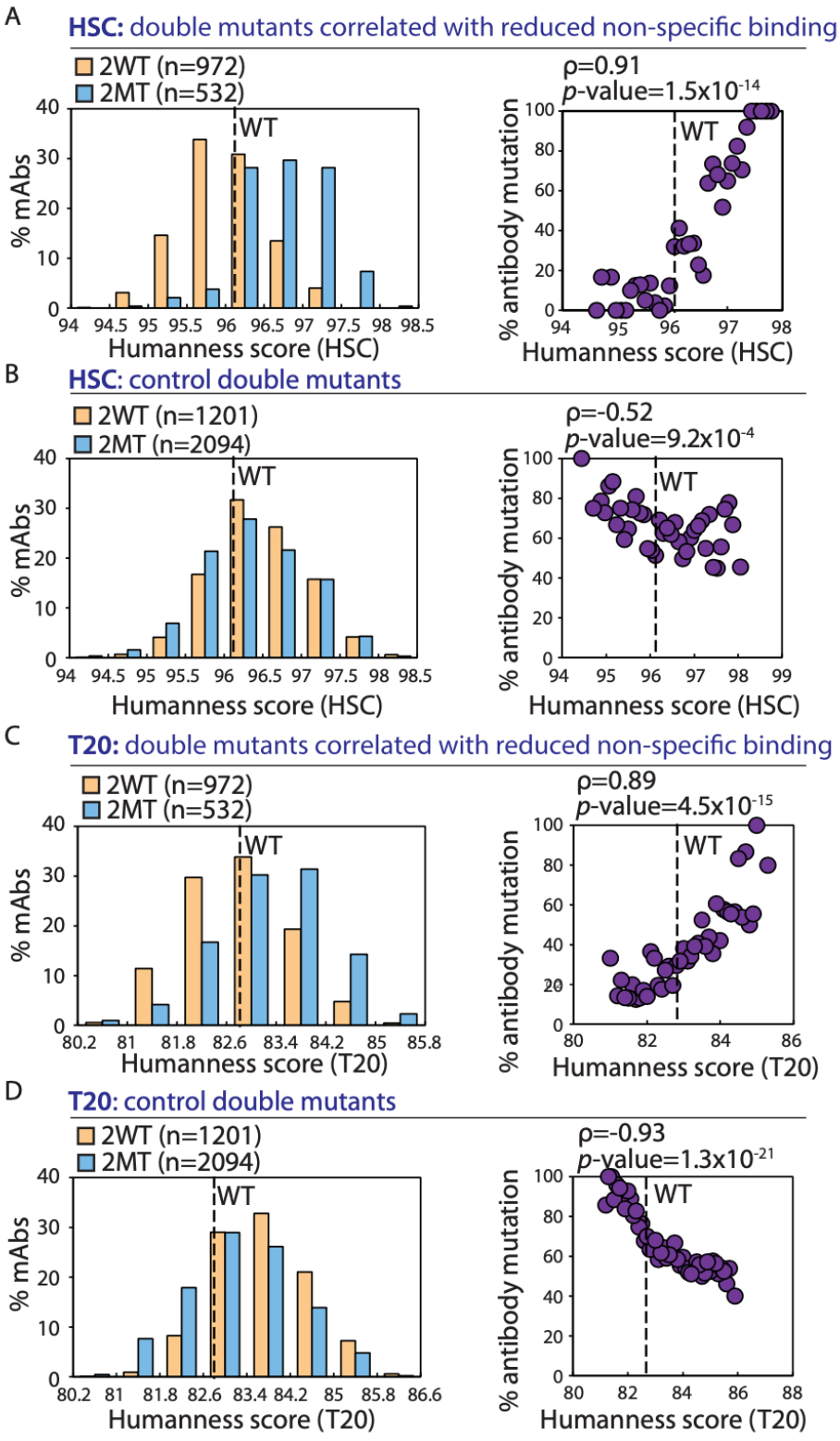


Figure 3-6: Emibetuzumab variants with mutations most correlated with reduced non-specific binding display increased levels of humanness.

(A, left) Distribution of human string content (HSC) scores for antibodies with sets of two mutations most strongly correlated with reduced non-specific binding (2MT) or the corresponding sets of antibodies with wild-type residues (2WT). (A, right) Correlation between HSC scores and percentage of antibody variants with two mutation sets that are most strongly correlated with reduced non-specific binding. (B, left) Distribution of HSC scores for control antibodies with sets of two mutations that are not statistically correlated with reduced non-specific binding (2MT) or the corresponding sets of antibodies with wild-type residues (2WT). (B, right) Correlation between HSC scores and percentage of antibody variants with two mutation sets that are not statistically correlated with reduced non-specific binding. (C, left) Distribution of T20 scores for antibodies with sets of two mutations most strongly correlated with reduced non-specific binding (2MT) or the corresponding sets of antibodies with wild-type residues (2WT). (C, right) Correlation between T20 scores and percentage of antibody variants with two mutation sets that are most strongly correlated with reduced non-specific binding. (D, left) Distribution of T20 scores for control antibodies with sets of two mutations that are not statistically correlated with reduced non-specific binding (2MT) or the corresponding sets of antibodies with wild-type residues (2WT). (D, right) Correlation between T20 scores and percentage of antibody variants with two mutation sets that are not statistically correlated with reduced non-specific binding. In (A-D), the dotted lines are the emibetuzumab (WT) humanness scores, and the *p*-values are reported for the Spearman's ρ values. In (A-B), the HSC scores are calculated using a panel of 3498 *VH1-02* human antibodies. In (C-D), the T20 scores are calculated using a panel of 12358 *VH1* human antibodies.

the emibetuzumab variants; these humanness scores are referred to T20 scores. As expected, we find that the top 20 most similar human V_H sequences relative to emibetuzumab V_H are in the *VH1* family. However, we find that not all of the 20 most similar sequences are in the *VH1-02* subfamily (4 of 20 are in the *VH1-46* subfamily). Therefore, we evaluated the T20 scores based on the 12,358 *VH1* antibodies, and observe similar trends for the T20 scores (Figure 3-5C and 5D) as for the HSC scores (Figure 3-5A and 3B), demonstrating the generality of these findings. We also obtained similar findings if we calculated the T20 scores using only the 3,498 *VH1-02* antibodies (Figure 3-8). Moreover, we repeated this analysis for sets of three emibetuzumab mutations and also find similar observations (Figure 3-9). Overall, these findings demonstrate that mutations strongly linked to reduce non-specific binding are also selectively linked to increased humanness.

Molecular features that strongly distinguish between antibody mutants with low and high levels of non-specific binding. We next sought to evaluate the physiochemical features of the emibetuzumab variants that were most strongly linked to non-specific interactions for the CDR mutants (Figure 3-10). To perform this analysis, we used two sets of emibetuzumab mutants, namely those in antibody sets #2 and #3, as described in the Methods section. Briefly, antibody set #2 (3,418 variants) contained antibody mutants that were linked to high non-specific binding and

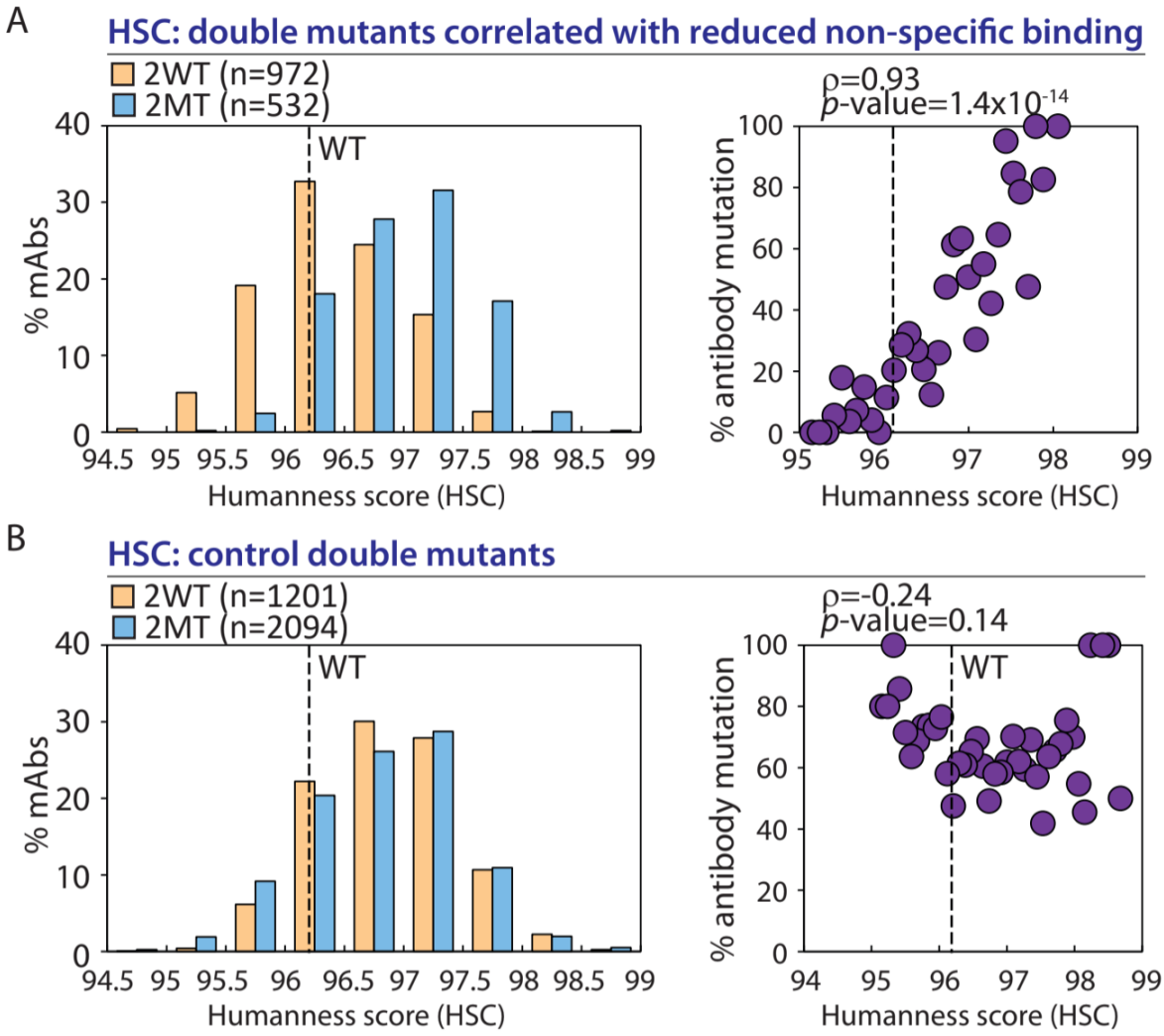


Figure 3-7: Emibetuzumab variants with mutations most correlated with reduced non-specific binding display increased humanness (HSC) scores calculated using *VH1* human antibodies.

(A, left) Distribution of human string content (HSC) scores for antibodies with sets of two mutations most strongly correlated with reduced non-specific binding (2MT) or the corresponding sets of antibodies with wild-type residues (2WT). (A, right) Correlation between HSC scores and percentage of antibody variants with two mutation sets that are most strongly correlated with reduced non-specific binding. (B, left) Distribution of HSC scores for control antibodies with sets of two mutations that are not statistically correlated with reduced non-specific binding (2MT) or the corresponding sets of antibodies with wild-type residues (2WT). (B, right) Correlation between HSC scores and percentage of antibody variants with two mutation sets that are not statistically correlated with reduced non-specific binding. In (A-B), the dotted lines are the emibetuzumab (WT) humanness scores, and the *p*-values are reported for the Spearman's ρ values, and the HSC scores are calculated using a panel of 12358 *VH1* human antibodies.

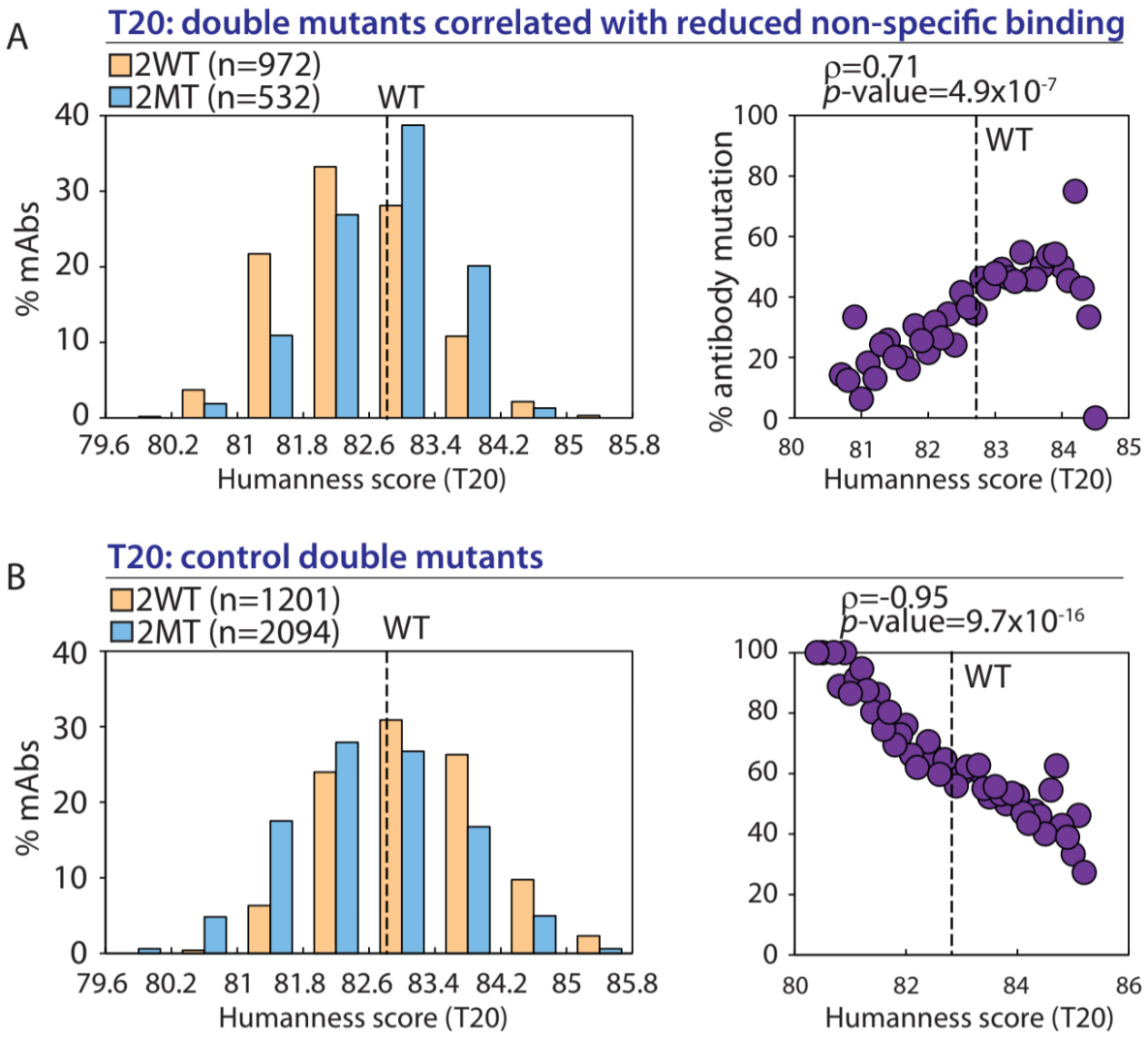


Figure 3-8：Emibetuzumab variants with mutations most correlated with reduced non-specific binding display increased humanness (T20) scores calculated using *VH1-02* human antibodies.

(A, left) Distribution of T20 scores for antibodies with sets of two mutations most strongly correlated with reduced non-specific binding (2MT) or the corresponding sets of antibodies with wild-type residues (2WT). (A, right) Correlation between T20 scores and percentage of antibody variants with two mutation sets that are most strongly correlated with reduced non-specific binding. (B, left) Distribution of T20 scores for control antibodies with sets of two mutations that are not statistically correlated with reduced non-specific binding (2MT) or the corresponding sets of antibodies with wild-type residues (2WT). (B, right) Correlation between T20 scores and percentage of antibody variants with two mutation sets that are not statistically correlated with reduced non-specific binding. In (A-B), the dotted lines are the emibetuzumab (WT) humanness scores, and the *p*-values are reported for the Spearman's ρ values, and the T20 scores are calculated using a panel of 3498 *VH1-02* human antibodies.

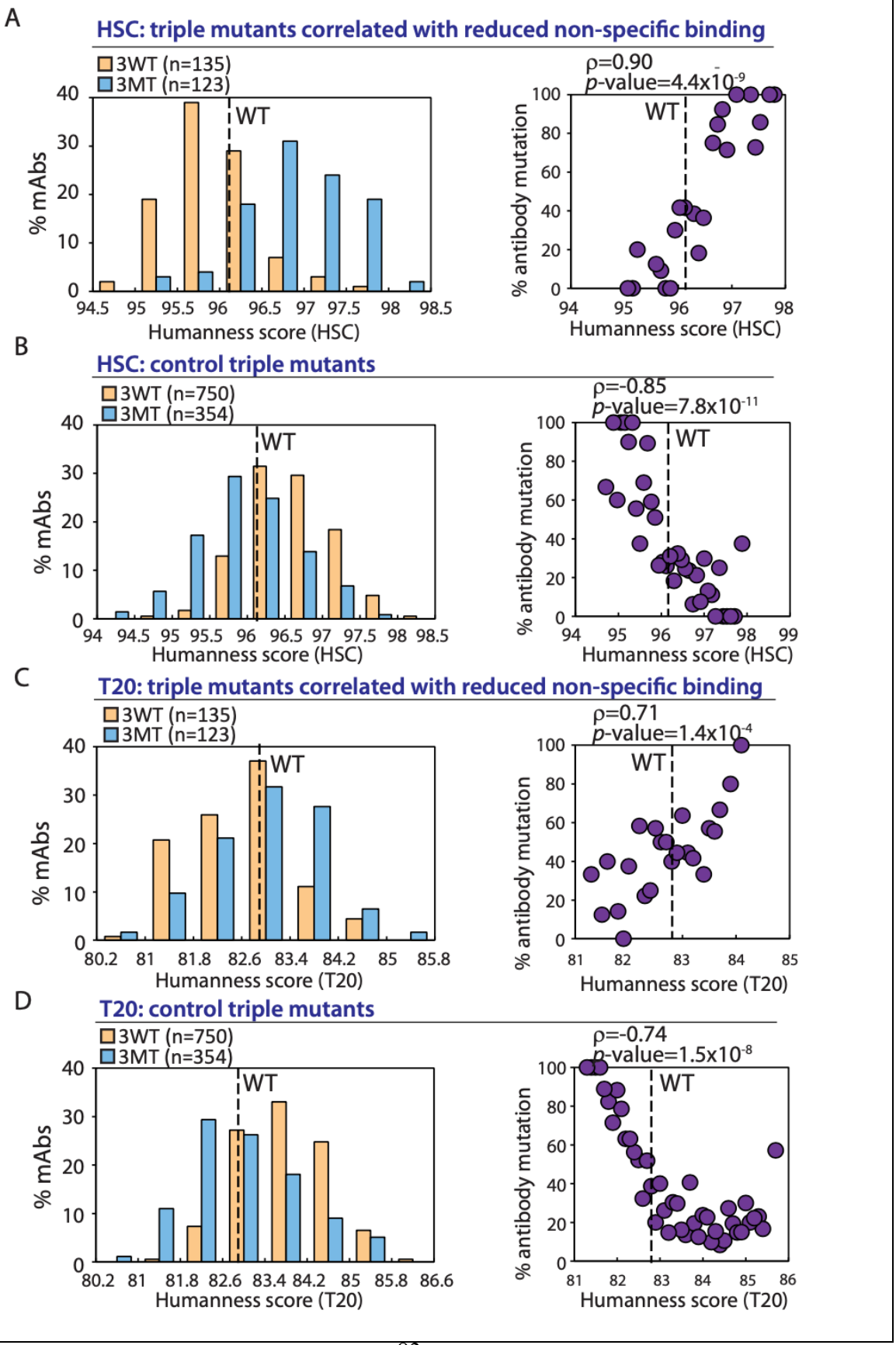

A
HSC: triple mutants correlated with reduced non-specific binding
3WT (n=135)
3MT (n=123)
WT
% mAbs
Humanness score (HSC)
ρ=0.90
p-value=4.4x10-9
% antibody mutation
B
HSC: control triple mutants
3WT (n=750)
3MT (n=354)
ρ=-0.85
p-value=7.8x10-11
C
T20: triple mutants correlated with reduced non-specific binding
Humanness score (T20)
ρ=0.71
p-value=1.4x10-4
D
T20: control triple mutants
ρ=-0.74
p-value=1.5x10-8

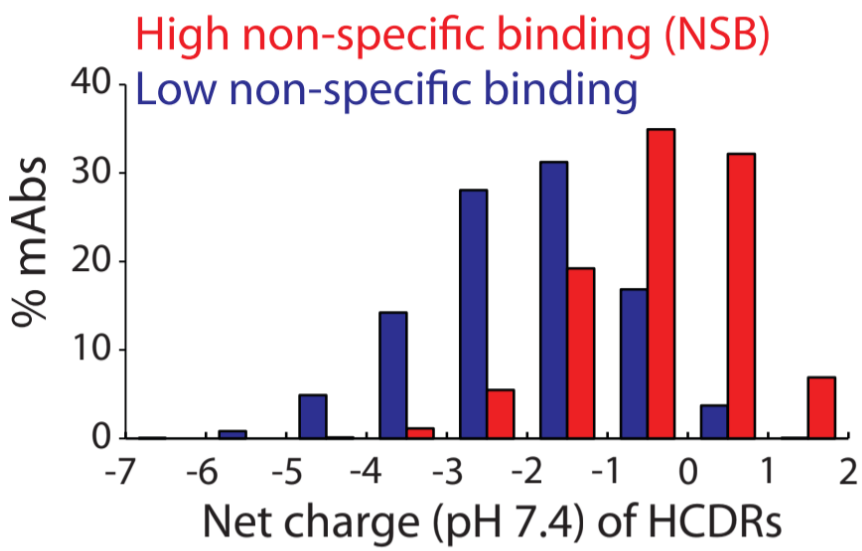
Figure 3-9: Emibetuzumab variants with three mutations most correlated with reduced non-specific binding display increased levels of humanness.

(A, left) Distribution of human string content (HSC) scores for antibodies with sets of three mutations most strongly correlated with reduced non-specific binding (3MT) or the corresponding sets of antibodies with wild-type residues (3WT). (A, right) Correlation between HSC scores and percentage of antibody variants with three mutation sets that are most strongly correlated with reduced non-specific binding. (B, left) Distribution of HSC scores for control antibodies with sets of three mutations that are not statistically correlated with reduced non-specific binding (3MT) or the corresponding sets of antibodies with wild-type residues (3WT). (B, right) Correlation between HSC scores and percentage of antibody variants with three mutation sets that are not statistically correlated with reduced non-specific binding. (C, left) Distribution of T20 scores for antibodies with sets of three mutations most strongly correlated with reduced non-specific binding (3MT) or the corresponding sets of antibodies with wild-type residues (3WT). (C, right) Correlation between T20 scores and percentage of antibody variants with three mutation sets that are most strongly correlated with reduced non-specific binding. (D, left) Distribution of T20 scores for control antibodies with sets of three mutations that are not statistically correlated with reduced non-specific binding (3MT) or the corresponding sets of antibodies with wild-type residues (3WT). (D, right) Correlation between T20 scores and percentage of antibody variants with three mutation sets that are not statistically correlated with reduced non-specific binding. In (A-D), the dotted lines are the emibetuzumab (WT) humanness scores, and the *p*-values are reported for the Spearman's ρ values. In (A-B), the HSC scores are calculated using a panel of 3498 *VH1-02* human antibodies. In (C-D), the T20 scores are calculated using a panel of 12358 *VH1* human antibodies.

enriched in positive selections for binding to two different polyspecificity reagents [ovalbumin and soluble membrane proteins (PSR)]. Conversely, antibody set #3 (4,814 variants) contained antibody mutants that were linked to low non-specific binding and enriched in positive selections for a lack of binding to the two different polyspecificity reagents. We evaluated a diverse set of molecular features (e.g., charge, hydrophobicity, hydrophilicity and combinations thereof), including: i) Fv sequence features (SB); ii) Fv sequence features weighted by the predicted solvent exposure of each amino acid (SB-SE); and (iii) sequence-based features weighted by predicted solvent exposures of each amino acid and their relative side chain surface areas (SB-SE-SA; see Methods for more detail). For example, SB features included the number of hydrophobic residues (Phe, Ile, Leu, Pro, Val, Trp, Met and Tyr) in the heavy chain CDRs, SB-SE features included the number of hydrophobic residues in the heavy chain CDRs weighted by the relative solvent exposure of each amino acid, and the SB-SE-SA features included the number of hydrophobic residues in the heavy chain CDRs weighted by the relative solvent exposure of each amino acid and multiplied by the relative side chain surface area of each hydrophobic residue. Logistic regression analysis was used to evaluate the significance of each molecular feature for differentiating between antibody mutants with high and low levels of non-specific binding, and

the features were ranked in terms of their area under the curve (AUC) values.

A



B

	Feature	Feature type	Region	Correlation with NSB	AUC
1	net charge	SB	H123	+	0.833
2	net charge	SB	H2	+	0.792
3	DENTYMPH	SB-SE	H2 L1	-	0.782
4	DENQTY	SB-SE-SA	H123	-	0.775
5	RKHFILPVW	SB	H2	+	0.766
6	RKH	SB	H123	+	0.765
7	RKH	SB	H2	+	0.765
8	RKH	SB-SE-SA	H123	+	0.755
9	RKH	SB-SE-SA	H2	+	0.752
10	RKHFILPVW	SB-SE-SA	H2	+	0.750

Figure 3-10: Evaluation of molecular features that most strongly differentiate emibetuzumab variants with low and high levels of non-specific interactions.

(A) The theoretical net charge of the heavy chain CDRs (pH 7.4) best differentiates between antibody variants that display low and high levels of binding to two non-specific reagents (ovalbumin and PSR). (B) Summary of molecular features that best discriminate between emibetuzumab variants with low and high non-specific binding (NSB). The feature types are SB (sequence based), SB-SE (sequence based and weighted by relative solvent exposures) and SB-SE-SA (sequenced based and weighted both by relative solvent exposures and side chain surface areas). In (A) and (B), the analysis was performed using antibody sets #2 (3418 unique variants) and #3 (4814 unique variants), and only features with AUC values >0.75 (logistic regression) are shown.

Interestingly, the net charge of the heavy chain CDRs (pH 7.4), a sequence-based (SB) feature, was best at differentiating between emibetuzumab mutants with low and high levels of nonspecific binding (AUC of 0.833; Figure 3-10A). Antibody mutants with increased positive charge in the heavy chain CDRs were strongly correlated with increased non-specific binding. The net charge of heavy chain CDR2 (also a sequence-based feature) was also strongly linked to non-specific binding, suggesting that reducing the charge of this CDR is key to reduced non-specific binding. This is interesting because three of the four mutated residues in heavy chain CDR2 are positively charged (arginine) residues (Figure 3-2A).

There were several other molecular features identified that were strongly linked to non-specific binding (AUC >0.75; Figure 3-10B). Six of these features display higher values for antibodies with high non-specific binding relative to those with low non-specific binding, as observed for the two charge features for heavy chain CDRs and H2 in particular. These six features are either linked to the number of positively charged residues alone or in combination with hydrophobic residues, including three SB features and three SB-SA-SE features. Two features display higher values for antibodies with low non-specific binding, and both are linked to increased numbers of negatively charged (Asp, Glu), hydrophilic (Thr, Asn) and polar aromatic (Tyr) residues. One of the features (SB-SE) is specific to two CDRs (heavy chain CDR2 and light chain CDR1) and is one of the 12 features we reported previously that are strongly linked to non-specific binding for clinical-stage antibodies (72). The second feature (SB-SE-SA) is specific for the heavy chain CDRs and is linked to residues that are either negatively charged or polar (uncharged). Collectively, these results suggest that CDR mutations that reduce positive charge and hydrophobicity and/or increase negative charge and hydrophilicity are most strongly linked to reduce non-specific binding.

Co-optimized antibody mutants with high affinity, low non-specific binding and high humanness. These results suggested that the heavy chain CDR mutations in the designed library could significantly reduce non-specific binding while significantly increasing humanness. However, as expected, we found that many of the mutants with reduced non-specific binding also displayed reduced antigen affinity (data not shown). Therefore, we further sorted the libraries to enrich for antigen binding, as described in Figure 3-11, and isolated clones from different stages of sorting. This led to the isolation of 17 mutants that were characterized in more detail.

We first evaluated the humanness of the selected clones and found that most (16 of 17) of them had either higher HSC or T20 scores than wild type emibetuzumab, and more than half (11 of 17) had both HSC and T20 scores that were similar (>99% of wild type) or higher than wild type (Figure 3-12. For these selected 12 mutants, we evaluated their binding to antigen and non-specific reagents (ovalbumin and PSR) and found that all of them displayed reduced non-specific binding relative to wild type (Figure 3-13. Many of them (9 of 12) displayed similar (>75%) or higher antigen binding than wild type at 1 nM antigen. We further evaluated the monovalent affinities of three of the most promising clones (EM-02, 43-06 and 27-10), and found that they displayed higher affinity than wild type (Figure 3-14).

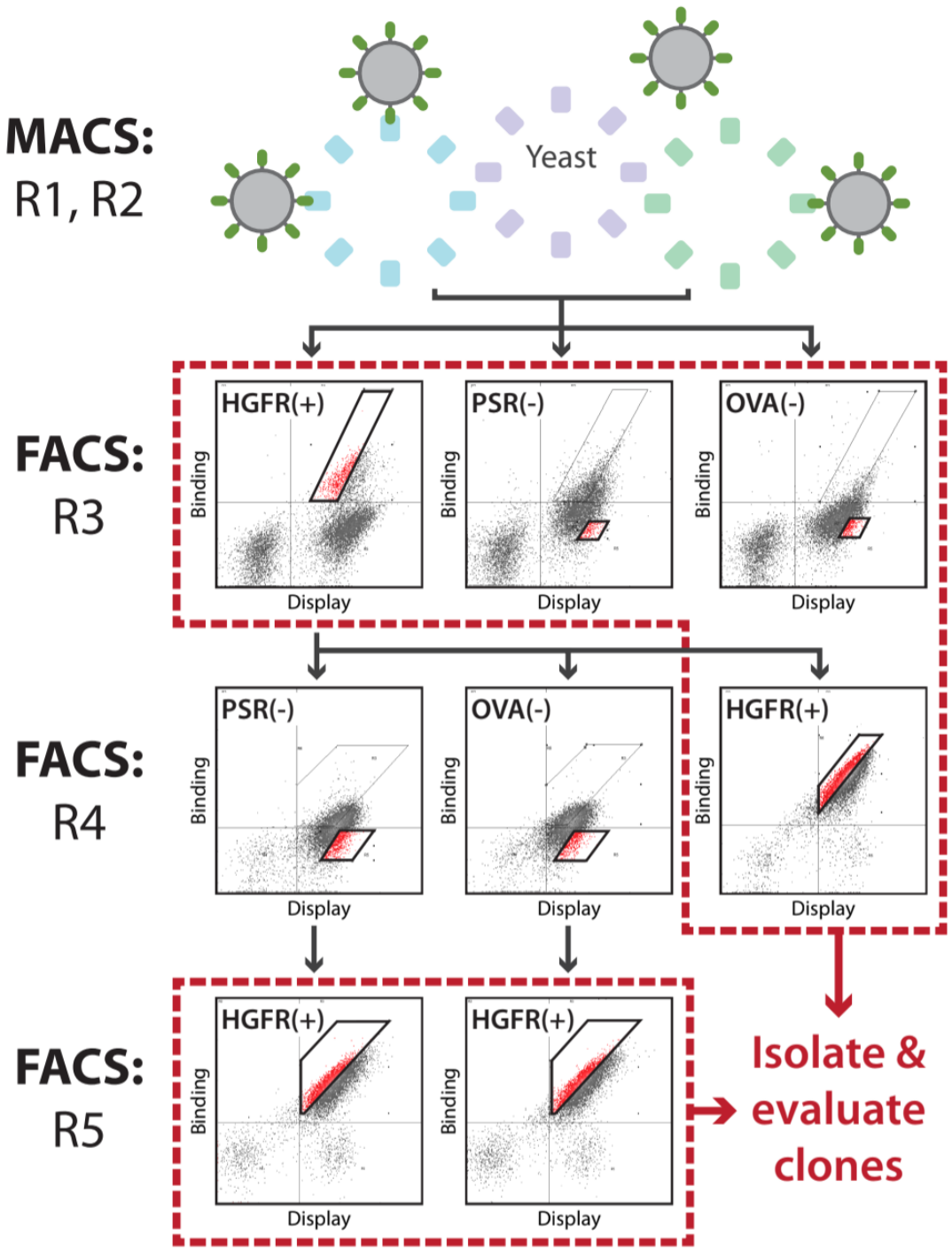


Figure 3-11: Sorting strategy used for isolating emibetuzumab variants with reduced non-specific binding while maintaining high affinity.

Magnetic-activated cell sorting (MACS) was performed in rounds 1 (R1) and 2 (R2) against antigen (hepatocyte growth factor receptor-Fc fusion protein, HGFR-Fc), and then fluorescence-activated cell sorting (FACS) was performed in rounds 3 (R3), 4 (R4) and 5 (R5) against antigen and non-specific reagents (ovalbumin and PSR). Antibody variants at different stages of sorting were isolated, sequenced and evaluated for their affinity and non-specific binding as single-chain Fabs (scFabs) on yeast and as soluble antibodies (IgGs).

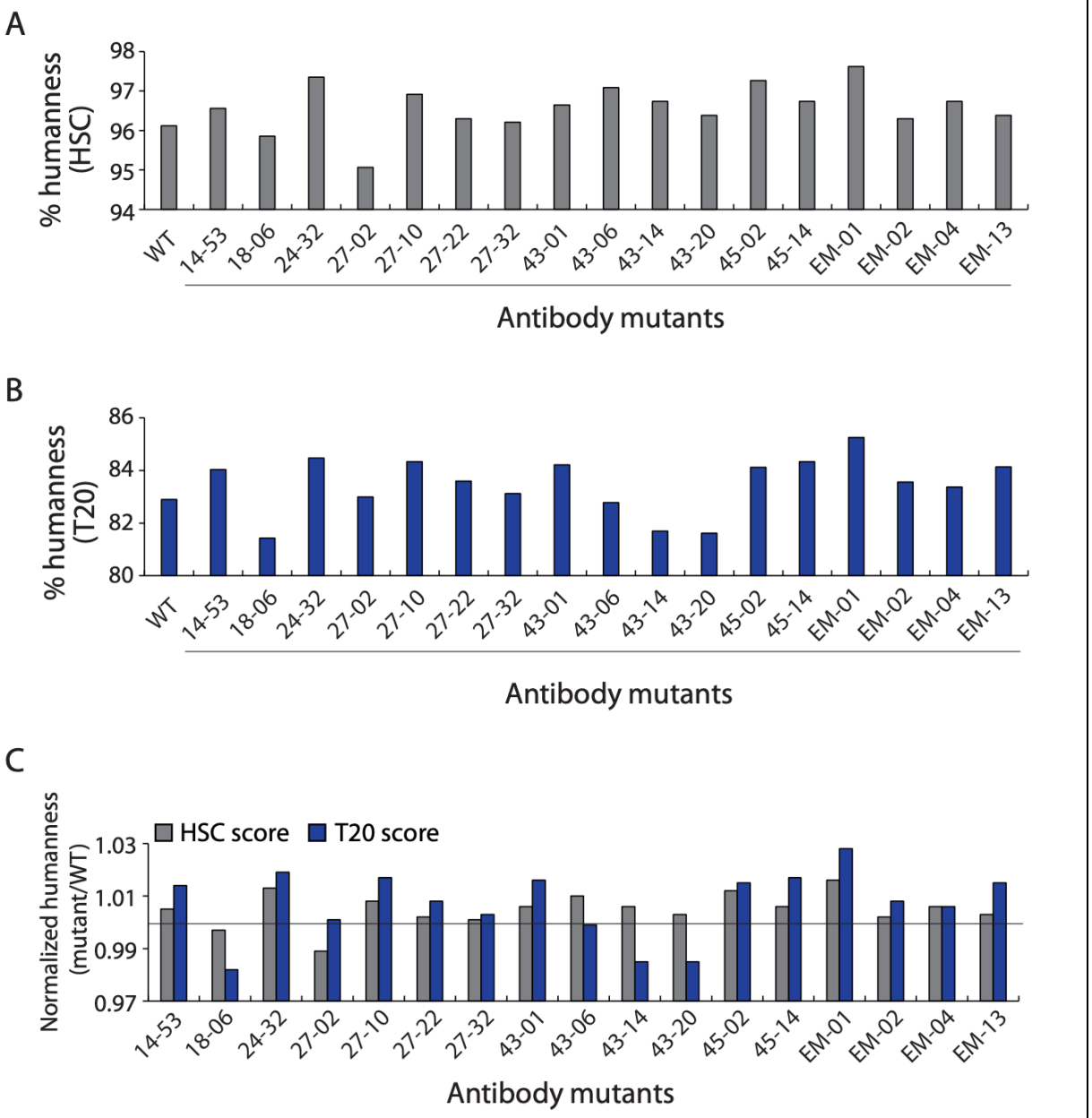


Figure 3-12: Evaluation of humanness scores of emibetuzumab variants selected using yeast surface display.

(A, B) Humanness scores calculated using the (A) HSC (*VH1-02*) and (B) T20 (*VH1*) methods. (C) Normalized humanness scores of selected antibody variants relative to wild type.

Finally, we evaluated the molecular origins of the reduced non-specific binding for the selected promising clones (Figure 3-15). Of the eight mutated sites in the three heavy chain CDRs, the three selected variants displayed similar patterns of mutations (Figure 3-15A). For example, in each of the three selected variants, arginine in heavy chain CDR2 (site 54) was mutated to a small hydrophobic residue (Gly or Ala), while two or three hydrophobic residues in heavy chain CDR3 were mutated to less hydrophobic or negatively charged residues. This pattern of reduced positive charge and hydrophobicity in the CDRs appears to explain the reduced level of non-specific binding of these variants. On the other hand, the three variants also displayed conservation of residues in heavy chain CDR1 (Tyr-33) and CDR2 (Arg-55 and Gly-56), suggesting that these residues may be important for antigen binding. In particular, it appears that Gly-56 is particularly important for antigen binding because it is commonly mutated in the emibetuzumab variants with low non-specific binding (Figure 3-3) that also display low antigen binding (data not shown).

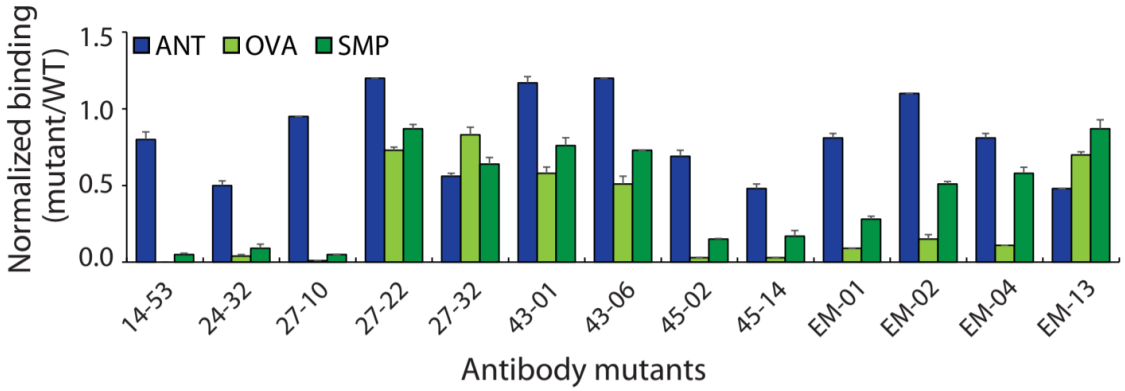


Figure 3-13: Yeast surface display evaluation of the normalized binding of emibetuzumab variants to antigen and non-specific reagents.

Single-chain Fab fragments were displayed on yeast and their binding to antigen (HGFR-Fc, 1 nM) and non-specific reagents [ovalbumin (0.1 mg/mL) and PSR (0.1 mg/mL)] was evaluated using flow cytometry. The measurements are averages of three independent experiments, and the error bars are standard deviations.

To further interrogate the origins of the important antibody properties, we generated homology models of the three emibetuzumab variants and visualized their surface charge and hydrophobic patches (Figure 3-15B). The wild-type antibody has a patch of positively charged residues (blue box) that overlaps with heavy chain CDR2, and this patch was disrupted in each of the three emibetuzumab mutants, especially for the mutant 27-10. Moreover, the hydrophobic patch in heavy chain CDR3 in the wild-type antibody is disrupted in two (27-10 and EM-02) of the mutants.

To quantify the molecular differences that differentiate the wild-type and mutant antibodies, we first identified molecular features that best differentiate between sets of antibodies with two or three mutations most strongly linked to reduced non-specific binding (n=973 antibodies) relative to sets of antibodies with the corresponding two or three wild-type residues (n=547 antibodies; see Methods for more detail). The molecular features considered were SB-SE features (number of

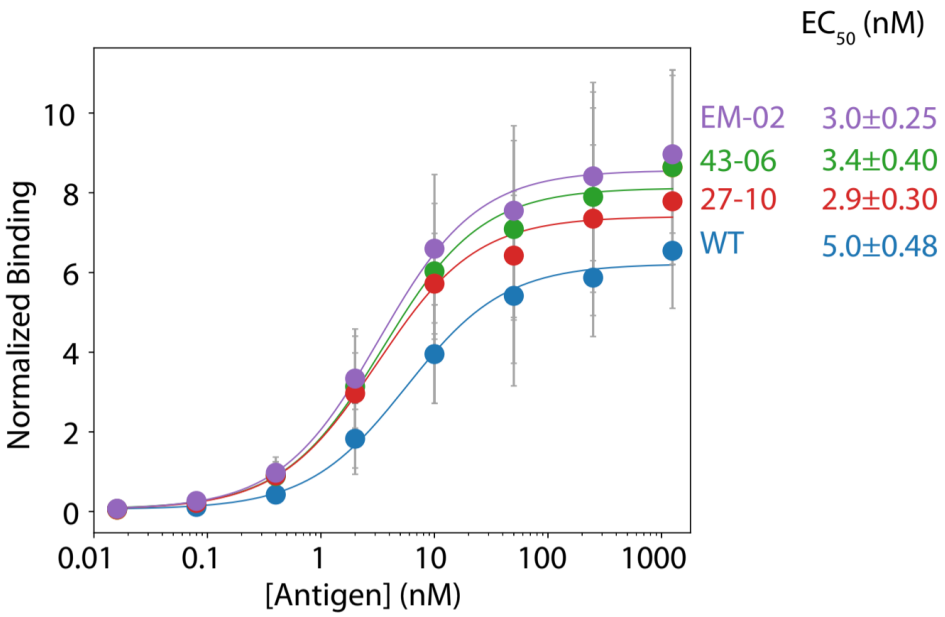


Figure 3-14: Emibetuzumab variants with reduced non-specific binding display high affinity.

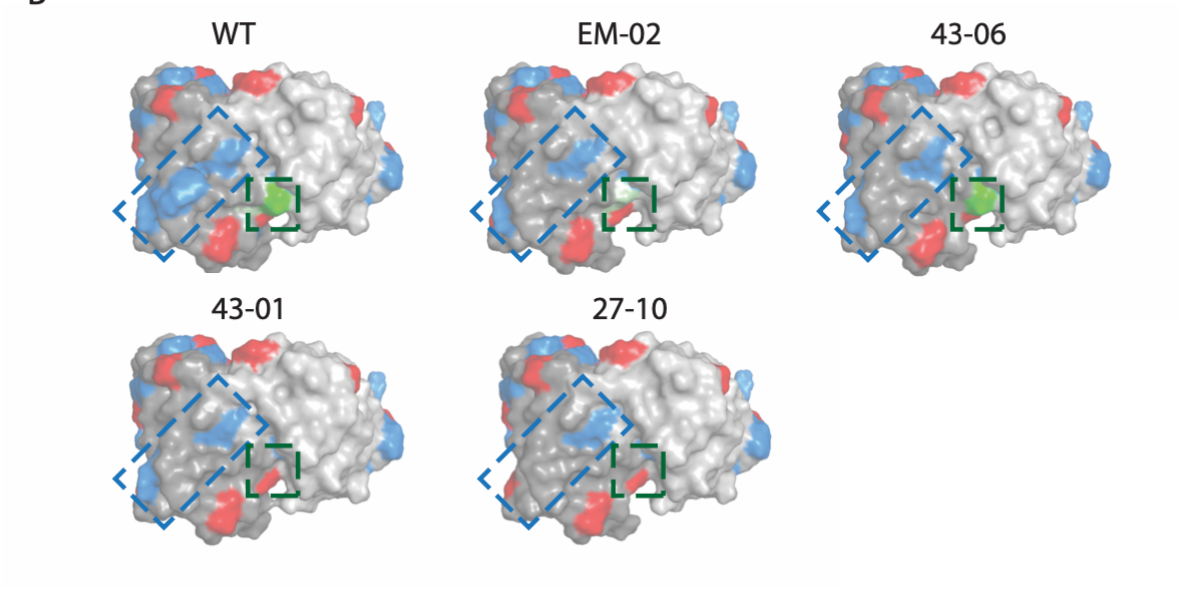
Single-chain Fab fragments were displayed on yeast, incubated with different concentrations of antigen, and evaluated via flow cytometry. The normalized binding signals were fit using a three-parameter Hill equation to obtain the EC_{50} values. The data shown are mean values of three independent repeats, the error bars are standard deviations, and the EC_{50} values are reported as mean ± standard deviation.

residues weighted by their relative solvent exposures) and SB-SE-SA features (number of residues weighted by their solvent exposure and multiple by their side chain surface area). Notably, the three most significant features were primarily combinations of positively charged and hydrophobic residues (Figure 3-15C). Each of the three features display >10% reduction relative to wild type, and the magnitude of change of the features was generally correlated with the level of reduction of non-specific binding. For example, mutants with the largest (27-10) and smallest (43-06) reductions in non-specific binding (Figure 3-15A) also displayed the largest (27-10) and smallest (43-06) reductions in each feature (Figure 3-15C). These results further highlight the importance of reducing positive charge and hydrophobicity to reduce non-specific binding.

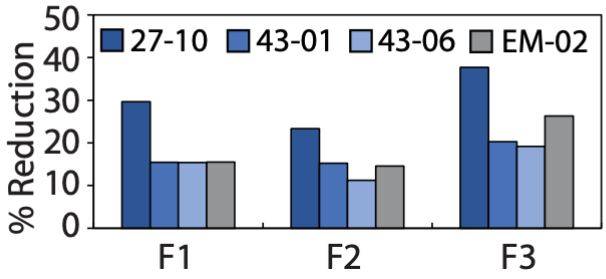
A

	H1	H2				H3			Antibody properties (normalized)		
	33	50	54	55	56	95	97	102	Antigen	Ovalbumin	SMP
WT	Y	R	R	R	G	A	W	Y	1.00	1.00	1.00
EM-02	Y	R	G	R	G	A	A	A	1.10	0.19	0.51
43-06	Y	R	G	R	G	S	L	D	1.20	0.42	0.73
43-01	Y	R	G	R	G	S	G	Y	1.15	0.53	0.76
27-10	Y	R	A	G	G	S	G	D	0.95	0.01	0.05

B



C



	Feature residues	Region
F1	RPVWYQE	H2 L13
F2	RKHFILPVW*	V_H
F3	RHLMPWYQ	H123 L23

Figure 3-15: Sequence and structural analysis of emibetuzumab variants with improved specificity and humanness.

(A) Summary of the mutations and normalized levels of antigen (1 nM) and non-specific binding for the selected antibody variants. (B) Fv homology models of the antibody variants with selected residues highlighted, including positively charged residues in Fv (blue), negatively charged residues in Fv (red) and hydrophobic residues (Ala, Trp and Leu) in HCDR3 (green). The V_H domain is indicated by dark gray and the V_L domain is indicated by light gray. The blue and green boxes indicate key regions in the heavy chain CDRs that are altered due to the selected mutations. (C) Molecular features most strongly correlated with reduced non-specific binding for the three emibetuzumab variants. The feature type (SB-SE and SB-SE-SA) are defined as described in Figure 3-10 and the Methods section.

Chapter 4 **Machine Learning Predictions of Antibody Polyreactivity**

Abstract

The development of monoclonal antibodies and antibody fragments for therapeutic applications requires the identification of immunoglobulins with drug-like biophysical properties, including high physical and chemical stability, high solubility, high specificity, low off-target binding, low viscosity, and long half-life. These many requirements have motivated the development of models for predicting antibody biophysical properties to reduce the amount of experimentation required during antibody candidate selection and optimization. However, the development of such models is typically limited by small datasets (<100-1000 antibodies)(4) that prevent the development of robust models, including machine-learning models that require much larger datasets (>10,000 antibodies). Here we report the generation of large deep sequencing datasets (>100,000 antibodies) for single-chain variable fragment (scFv) antibodies with low and high levels of non-specific binding. Two human scFv libraries displayed on yeast were sorted successively against four different polyspecificity reagents via fluorescence-activated cell sorting to enrich for variants that were either broadly polyreactive or broadly non-polyreactive. Analysis of the deep sequencing data reveals that the single molecular features that best classified the levels of antibody non-specific binding were the net charge of the variable fragments (Fv) and variable heavy domains (V_H), as highly positively charged Fv and V_H regions were correlated with high levels of non-specific binding. Moreover, high combined numbers of hydrophobic and positively charged residues and high numbers of aromatic residues in the Fv and V_H regions were also

correlated with high levels of non-specific binding. Random forest classification models were trained on the deep sequencing data and displayed strong performance for predicting the levels of scFv and mAb non-specific binding for hold-out human scFv sequencing data (area under the curve of 0.90) and clinical-stage scFv and mAb experimental data (areas under the curve of 0.83-0.99). These models require only sequence-based Fv features and are simple to integrate into computational screens for identifying antibody variants with drug-like biophysical properties.

Introduction

Monoclonal antibodies are being used to treat a wide range of human disorders, including those ranging from traditional therapeutic areas such as cancer and autoimmune diseases to emerging areas such as neurological and neurodegenerative diseases. The success of antibody therapeutics is governed by several factors, including their activity, safety and developability. Activity is the single most important property, which encompasses several key molecular actions of antibody therapeutics, including high-affinity antigen binding, modulation of antigen activity (e.g., reduced activity for agonists and increased activity for agonists) and recruitment of effector functions. Safety is also a key property that encompasses adverse biological events related to the specific mode of action of the antibody drug and more general safety issues such as immunogenicity. Developability is a third key property that encompasses the physical and chemical attributes of antibody drugs, including stability, solubility, aggregation, viscosity, non-specific binding, and pharmacokinetics.

A longstanding goal in the field is to be able to reliably and predictably generate antibody therapeutic candidates that are co-optimized for high activity, safety and developability. Despite important progress toward this goal, there is still much need for predictive models that can reduce the need for extensive experimentation and improve the drug development process. The most

limiting factor in developing predictive models of antibody properties is the availability of large datasets that are required for training and testing of robust models, which is especially true for the development of machine learning models. In the area of antibody developability, there is great potential to develop machine learning models for predicting antibody non-specific interactions because of the feasibility of generating large datasets using high-throughput screening methods. Multiple studies to date have reported panels of thousands of single-chain variable fragments (scFvs) and IgGs, which have revealed a number of molecular features that appear linked to non-specific binding. One study identified specific hydrophobic peptide motifs enriched in polyreactive scFvs (22,73), while a second study found that polyreactive IgGs displayed increased heavy chain inter-loop crosstalk and a propensity for neutral binding surfaces both in terms of charge and hydrophobicity (24). The latter findings were integrated into a machine learning model for classifying antibody polyreactivity with an accuracy at least 75% using a set of ~1000 IgG and IgA antibodies.

Despite this strong progress toward understanding and predicting antibody polyreactivity, the single issue that is most limiting further progress is the lack of large (>100,000 antibodies), high quality and self-consistent datasets defined for polyreactive antibody binding (or a lack thereof) to diverse panels of non-antigen (polyspecificity) reagents. The lack of such datasets has resulted in an incomplete analysis of two key outstanding questions. First, what are the physicochemical features of antibodies that are most strongly linked to antibody non-specific binding? Second, to what extent can be antibody non-specific binding be predicted? We have sought to address aspects of both questions by generating large (>100,000 variants) and self-consistent datasets by sorting and deep sequencing two yeast-displayed human scFv libraries for either binding or a lack of binding to four diverse polyspecificity reagents. Herein we report the

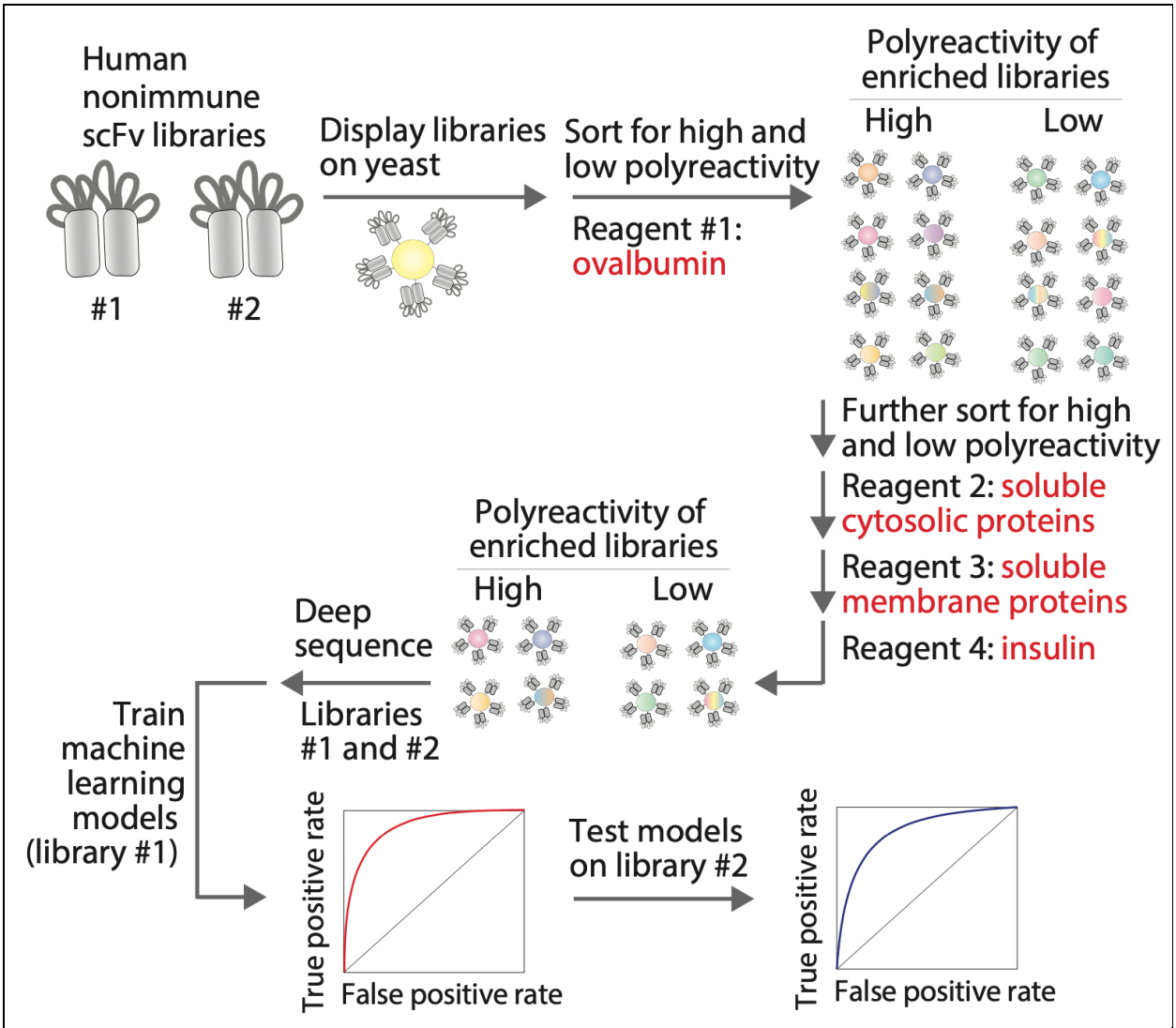


Figure 4-1: Overview of the experimental and computational methods used to generate machine learning models for predicting antibody polyspecificity.

Two human single-chain variable fragment (scFv) libraries displayed on the surface of yeast were sorted for positive or negative binding to multiple polyspecificity reagents. The enriched libraries were deep sequenced to generate large datasets of antibody sequences and corresponding classifications for ether high or low polyspecificity. One of large human scFv datasets was used along with a smaller dataset for clinical-stage antibodies to train machine learning models to predi t antibody polyspecificity. Finally, the models were tested on the second large human scFv dataset (not used for training) and additional independent datasets for preclinical and clinical-stage antibodies.

generation and analysis of such large datasets and our findings of both previously reported and novel molecular features strongly linked to antibody polyreactivity as well as the development of machine learning models that accurately predict antibody polyreactivity.

Experimental section

Human scFv library sorting and clone evaluation. To generate large datasets of scFvs with low and high levels of non-specific binding, two human scFv libraries (library #1 (84)and #2 (85)) were displayed on yeast (Aga2-V_H-linker-V_L) and sorted first for full-length scFv display by MACS and then progressively against four different polyspecificity reagents by FACS. The initial sort using MACS (sort 1) was conducted with cMyc antibody (Cell Signaling Technology, 2276S; library #1) or V5 antibody (Abcam, ab27671; library #2) to select full-length scFvs. To perform the MACS selections, $2x10^9$ cells were incubated with 80 nM cMyc or V5 antibody in PBSB (1 g/L BSA, PBS) with 1% milk at room temperature (3 h). Afterward, the cells were then washed once with PBSB and incubated with Protein G beads and gently rocked for 30 min at 4 °C. The cells suspension was then passed through a MACS column for bound cells. After washing, the cell/bead complexes were transferred into SDCAA media (20 g dextrose, 5 g casamino acids, 6.7 g yeast nitrogen base without amino acids, 16.75 g sodium citrate, and 4 g anhydrous citric acid per liter) and grown at 30 °C for 2 d.

Next, the scFv libraries were progressively sorted for positive and negative binding to four biotinylated polyspecificity reagents (0.13 mg/mL), namely ovalbumin (Sigma Aldrich, A5503; sort 2), soluble cytosolic proteins from CHO cells (SCP; sort 3), soluble membrane proteins from CHO cells (SMP; sort 4) and insulin (Sigma Aldrich, I9278; sort 5). The SCP and SMP reagents were prepared as described previously.(86) The libraries were sorted via a FACS MoFlo Astrios (Beckman) using standard protocols (87). Yeast cells were incubated the polyspecificity reagents for 3 h at room temperature (ovalbumin, SCP and insulin) or 20 min on ice (SMP). The polyspecificity reagents were biotinylated and detected via streptavidin AlexaFluor 647 secondary (1:1000; Invitrogen, S21374). The top 10% and bottom 10% of the scFv binding population of

yeast cells that expressed human scFvs were collected separately and cultured for the next round of positive or negative selection. The enriched library samples for high non-specific binding to ovalbumin in sort 2 were next selected for high non-specific binding to SCP in sort 3, while the enriched libraries for low non-specific binding to ovalbumin in sort 2 were next selected for low non-specific binding to SCP in sort 3. This process was repeated in successive sorts to generate enriched scFv library samples that were either broadly polyreactive or non-polyreactive.

Yeast plasmids were extracted from the positive and negative populations from sort 5 via yeast minipreps (Zymo, D2004), and transformed into DH5-α component cells. Single colonies were picked, and scFv coding fragments was confirmed via Sanger sequencing. The sequenced scFv plasmids were transformed into yeast cells (EBY100) using a standard media solution [40% (w/v) PEG 3350, 100 mM Lithium acetate, 10 mM Tris (pH 7.5), 0.4 mM EDTA]. Yeast harboring the plasmids were then recovered in SDCAA medium with shaking at 30 °C. Yeast surface display of the scFvs was induced by transferring mid-exponential phase cells (OD600 at 2 to 4) into SD-GCAA medium with shaking at 30 °C for 20 h before flow analysis. Non-specific binding of the polyspecificity reagents to the scFvs was evaluated using a 1:10 dilution of each biotinylated polyspecificity reagent (1.3 mg/ml) and streptavidin AlexaFluor 647 conjugate (1:1000; Invitrogen). Human scFv expression was detected using mouse cMyc (1:1000; Cell Signaling Technology) and goat anti-mouse AlexaFluor 488 secondary (1:200; Invitrogen). To compare the non-specific binding levels of different scFvs, the median binding signals were normalized for each scFv between two control scFvs. The median value for each scFv was normalized between a positive control (FP3) with a score of 1 and a negative scFv (FN4) with a score of 0.

Deep sequencing of human libraries. Plasmids from the sorted libraries were isolated using yeast mini preps (Zymo, D2004). The variable light and heavy domains were amplified from the

plasmids and purified with a 1% agarose gel. The products were then gel purified using a QIAquick Gel Extraction Kit (Qiagen, 28704). The concentrations of the purified amplicons were then determined using a Qubit 4 Fluorometer. Each sample (50 μL at 10 ng/μL) was sent to the University of Wisconsin Sequencing Core for PacBio sample preparation and sequencing. The output bam files from PacBio sequencing were first converted to fastq files via BedTools and then converted into a fasta format for ease of processing.(88) The fasta files were evaluated and full-length sequences without stop codons were collected into a dictionary containing the number of occurrences of each sequence.

Human library #1 scFv dataset. The deep sequencing data from library #1 was analyzed to identified scFvs in human library samples sorted either for binding or a lack of binding to ovalbumin, insulin, CHO soluble cytosolic proteins (SCP) and CHO soluble membrane proteins (SMP). The resulting library samples, which are referred to as either positive or negative samples against each reagent, were collected in three independent experiments. The antibody sequences were pooled for the twelve antibody library samples (four reagents with three biological replicates) sorted positively for non-specific binding, while the corresponding twelve antibody library samples sorted for a lack of non-specific binding were also pooled. Next, the two sets of pooled library sequences were processed to eliminate duplicates with each pooled sequence sets and any duplicates between the two different sequence sets. Afterward, the remaining variable heavy (V_H) and variable light (V_L) domains in each pooled sequence set were aligned in ANARCI using a database of Hidden Markov Models,(79) and scFvs were eliminated from further consideration if V_H and V_L germlines could not be assigned. Finally, scFvs were eliminated if the number of amino acids in specific antibody regions exceeded maximum limits: 134 residues for V_H; 7 residues for HCDR1; 19 residues for HCDR2; 21 residues for HCDR3; 122 residues for V_L; 17 residues for

LCDR1; 15 residues for LCDR2; and 12 residues for LCDR3. The antibody CDRs were defined using a combination of Chothia and Kabat numbering. This resulted in a final set of 131,754 scFv sequences selected for low non-specific binding and 115,233 scFv sequences selected for high non-specific binding, which together (246,987 scFvs) is referred to as antibody set #1 (Dataset S1).

Human scFv library #2 dataset. The deep sequencing data for library #2 was processed in the same manner as the corresponding data for library #1. The library samples were collected by sorting against the same four polyspecificity reagents as used for library #1, and two biological replicates were performed. After processing the sequencing data, the final sets of antibodies included 38,500 scFv sequences selected for low non-specific binding and 99,120 scFv sequences selected for high non-specific binding, which together (137,620 scFvs) is referred to as antibody set #2 (Dataset S2).

Additional antibody datasets with experimental measurements. The models developed in this work were trained or evaluated using three additional sets of antibodies. Antibody set #3 contained 20 clinical-stage antibodies that were selected based on displaying either generally high or generally low levels of non-specific binding (4). These antibodies were selected based on the previously reported non-specific binding values from three assays, namely the polyspecificity reagent (PSR) assay using soluble membrane and cytosolic proteins (9) and two ELISAs using either baculovirus particles(4) or a mixture of proteins, lipid and nucleic acid reagents.(4) Each of the selected antibodies displayed either higher or lower values for all three assays than the published threshold values for high levels of non-specific binding. This resulted in 10 mAbs with experimentally confirmed high levels of non-specific binding and 10 mAbs with low levels of non-specific binding. Two additional mAbs were initially identified in this analysis (crenezumab and natalizumab), but their V_H and V_L germlines were not observed in antibody set #1 and were

excluded from further analysis. A second set of clinical-stage antibodies was also used during the training process (antibody set #4). This set included the variable regions of 103 mAbs and their corresponding values of non-specific binding to the PSR reagent. Finally, a set of 27 scFvs provided by Boehringer Ingelheim (antibody set #5) was used for model evaluation.

Identification of molecular features linked to antibody specificity. A diverse panel of molecular features were evaluated for their ability to discriminate between antibodies with high and low levels of non-specific binding in antibody set #1. The features were evaluated for the antibody V_H, V_L and Fv regions, and were based on three categories of features, namely: (i) sequence-based (SB) features; (ii) sequence-based features weighted by site-specific solvent accessibilities (SB-SE);(40) and (iii) sequence-based features weighted by site-specific solvent accessibilities and multiplied by side chain surface areas (SB-SE-SA). For evaluating SB features, three sub-categories were evaluated: (i) number of amino acids (e.g., number of Asp and Glu residues in V_H); (ii) overall sequence hydrophobicity (e.g., total number of each of the 20 residues in V_L multiplied by their residue-specific hydrophobicity values determined by the Black and Mould scale(89)); (iii) theoretical net charge at pH 7.4 (+1 for Lys and Arg, +0.1 for His, and -1 for Asp and Glu); (iv) hydrophobicity asymmetry (e.g., the absolute value of the difference between overall sequence hydrophobicities for V_H and V_L); and (v) charge asymmetry (e.g., the product of the net charge of the V_H and V_L regions at pH 7.4).

For the features in category (i) of the SB features, the number of amino acids were evaluated in terms of nine groups based on their properties: (a) hydrophobic amino acid group 1 (Ala, Phe, Ile, Leu, Pro, Val, Trp, Met and Tyr); (b) hydrophobic amino acid group 2 (Phe, Ile, Leu, Pro, Val and Trp); (c) hydrophobic and positively charged amino acid group (Phe, Ile, Leu, Pro, Val, Trp, Arg, Lys and His); (d) hydrophilic amino acid group 1 (Gln, Ser, Thr and Asn); (e)

hydrophilic amino acid group 2 (Gln, Ser, Thr, Asn and Tyr); (f) hydrophilic and negatively charge amino acid group (Gln, Ser, Thr, Asn, Asp and Glu); (g) aromatic amino acid group (Phe, Trp and Tyr); (h) aromatic amino acid and proline group (Phe, Trp, Tyr and Pro); and (i) the twenty amino acids (tested individually). This resulted in 84 SB features in this sub-category.

The SB-SE features require evaluation of the site-specific solvent accessibilities of residues in the V_H and V_L regions. This was performed using a published Random Forest machine learning model that was trained on over 900 antibodies in the Protein Data Bank.(40) The SB-SE-SA features required both the solvent accessibilities and the side chain surface areas of each residue, the latter of which was obtain from a reported previously.(81) The SB-SE-SA features were otherwise evaluated in a generally similar manner as the SB features for category (i) and sub-categories (a-h). This resulted in 24 SB-SE-SA features in this sub-category.

The germlines of the VH and VL regions were also evaluated as molecular features for distinguishing between and low specific antibodies. The germlines were identified via ANARCI predictions using a database of Hidden Markov Models.(79) The germlines were one-hot encoded into a one-dimensional vector for analysis. There were seven V_H (*VH1*, *VH2, VH3, VH4, VH5, VH6, VH7*), six Vκ (*Vκ1, Vκ2*, *Vκ3*, *Vκ4*, *Vκ5*, *Vκ6*) and nine Vλ (*Vλ1*, *Vλ2*, *Vλ3*, *Vλ4*, *Vλ5*, *Vλ6*, *Vλ7*, *Vλ8*, *Vλ10*) germlines used for analysis.

The most significant features that discriminated between antibodies with low and high levels of non-specific binding (antibody set #1) were identified by calculating the area under of the ROC curve (AUC) for logistic regression analysis. The top twenty features with AUC values >0.75 are reported in Fig. 5. Moreover, the best threshold (rule) value for each feature for differentiating between antibodies with high and low non-specific binding (antibody set #1) was evaluated by calculating the adjusted accuracy for the confusion (2x2) matrix for a range of

threshold values. In particular, the possible threshold values tested for each feature ranged from the minimum to maximum observed values in increments of 0.1 for SB-SE and charge-based SB features and increments of 1 for everything else. Both maximum and minimum threshold values were evaluated.

Random forest model generation and analysis. The molecular features were preprocessed by normalizing feature values from 0 to 1 using MinMaxScaler function in the scikit-learn preprocessing package in Python. The scFvs with high and low specificity in antibody set #1 (Dataset S1) were split into training (80%) and test (20%) sets using stratified sampling. The training set was further divided into ten random partitions (folds) for tenfold cross-validation, nine of which were used for training and the other for validation (and this was repeated ten times). The first-generation model was trained using the random forest classifier package in the scikit-learn ensemble package in Python. The model was optimized by tuning the hyperparameters and selecting the best model based on two criteria: (i) minimum coefficient of variation for the ten validation accuracies calculated for antibody set #1 (80% training set); and (ii) area under the ROC curve >0.85 (logistic regression) for predicting the relative specificity (as judged by PSR experimental measurements) for the 103 clinical-stage antibodies (antibody set #5). The parameters of the best first-generation model n_estimator of 5, max_depth of 15, min_sample_split of 30 and min_samples_leaf of 1.

The first-generation model was used to identify the ten most important molecular features for predicting antibody non-specific binding (Table 4-2). These ten features, which are five SB features and five SB-SE-SA features, were combined with the twelve previously reported SB-SE features,(72) for training the second-generation random forest model using a total of 22 features. The model was trained as described for the first-generation model except the best model was

chosen based on the following three criteria: (i) >95% for the average of the ten validation accuracies calculated for antibody set #1 (80% training set); (ii) minimum coefficient of variation for the corresponding calculations in (i); and (iii) area under the ROC curve >0.85 (logistic regression) for predicting the relative specificity for antibody set #5. The parameters of the best first-generation model were n_estimator of 14, max_depth of 20, min_sample_split of 10 and min_samples_leaf of 1.

Random forest model with adding CDR features generation and analysis. The prior random forest model was generated using features in VH, VL and Fv domain. In this model, features in HCDR, LCDR and all six CDR were added to the feature pool. The model was optimized by tuning the hyperparameters and selecting the best model based on the same two criteria: (i) minimum coefficient of variation for the ten validation accuracies calculated for antibody set #1 (80% training set); and (ii) area under the ROC curve >0.75 (logistic regression) for predicting the relative specificity (as judged by PSR experimental measurements(4)) for the 123 clinical-stage antibodies (antibody set #5). The parameters of the best first-generation model n_estimator of 10, max_depth of 50, min_sample_split of 100 and min_samples_leaf of 1. The first-generation model was used to identify the ten most important molecular features for predicting antibody non-specific binding. Within the top ten features, three features were CDR related features (net charge at LCDR, net charge at CDR and number of aromatic amino acid in CDR weighted by solvent accessible surface area. These ten features were combined with the twelve previously reported SB-SE features,(72) for training the second-generation random forest model using a total of 22 features. The model was trained as described for the first-generation model except the best model was chosen based on the following two criteria: (i) >95% for the average of the ten validation accuracies calculated for antibody set #1 (80% training set); (ii)

maximum area under the ROC curve (logistic regression) for predicting the relative specificity for antibody set #5. The parameters of the best first-generation model were n_estimator of 20, max_depth of 20, min_sample_split of 40 and min_samples_leaf of 1.

Additional model generation and analysis. For the linear discriminant analysis (LDA) model, the V_H and V_L sequences from antibody set #1 were converted into One-Hot Encoded (OHE) matrices. The selected residues at each position in V_H and V_L that were encoded were those that occurred at least 1% of the time among all the scFv sequences. For example, there were five amino acids that occurred at the first position in V_H (Ala, Gln, Glu, Phe, and Asp), but only Glu and Gln occurred at least 1% of the time and were encoded in the OHE matrix. This resulted in the scFv sequences being encoded in a matrix with 1187 columns. The same approach was applied for the other antibody sets (#2-6) to evaluate the performance of the LDA model. For the support vector classifier (SVC) model, two position-specific scoring matrices (PSSMs) were generated for scFvs with low and high polyreactivity in antibody set #1 (PSSM1 for low polyreactivity scFvs and PSSM2 for high polyreactivity scFvs). Each scFv sequence was given two scores, one from PSSM1 and the other from PSSM2. Next, for the SVC model, an optimized radial basis function (RBF) was trained using antibody set #1 to differentiate between antibodies with low and high polyreactivity using a C parameter of 10.

Results

Generation of large datasets of human scFvs with high and low levels of polyreactivity. To develop robust models for predicting human antibody polyreactivity, we first sought to generate large datasets of human scFvs enriched for high and low levels of polyreactivity. We used two different previously reported libraries, namely library #1(84) and #2(90), displayed on the surface

of yeast for our analysis. To ensure the enriched libraries would be generally polyreactive or non-polyreactive, we sorted the libraries against four different polyspecificity reagents, namely ovalbumin, soluble cytosolic proteins (SCP; from CHO cells), soluble membrane proteins (SMP; from CHO cells) and insulin. We first selected the top 10% and the bottom 10% of each library against the first reagent (ovalbumin) and then subsequently sorted each enriched population (ovalbumin+ and ovalbumin-) against soluble cytosolic proteins (top 10% of ovalbumin+ library and bottom 10% of ovalbumin- sample) to generate libraries doubly enriched for non-specific binding (ovalbumin+/SMP+) or a lack of non-specific binding (ovalbumin-/SMP-). This process was repeated two more times against SMP and insulin, which resulted in libraries generally enriched for polyreactivity (ovalbumin+/SCP+/SMP+/insulin+) or a lack thereof (ovalbumin-/SCP-/SMP-/insulin-). After four rounds of sorting, the final enriched libraries were sorted one final time against each of the four reagents to collect positive and negative non-specific binding samples for deep sequencing analysis, as shown for library #1 in Figure 4-2.

To evaluate the quality of the sorted library samples, we randomly isolated 28 scFvs that were in the library sample enriched for high polyreactivity and 34 clones in the library sample enriched for low polyreactivity from library #1. The levels of non-specific binding for the clones enriched for high non-specific binding were much higher than those clones enriched for low non-specific binding for each of the four reagents (Figure 4-3). These results demonstrate that enriched libraries display the expected large differences in their levels of non-specific binding, and are well suited for further analysis of the determinants of antibody polyreactivity.

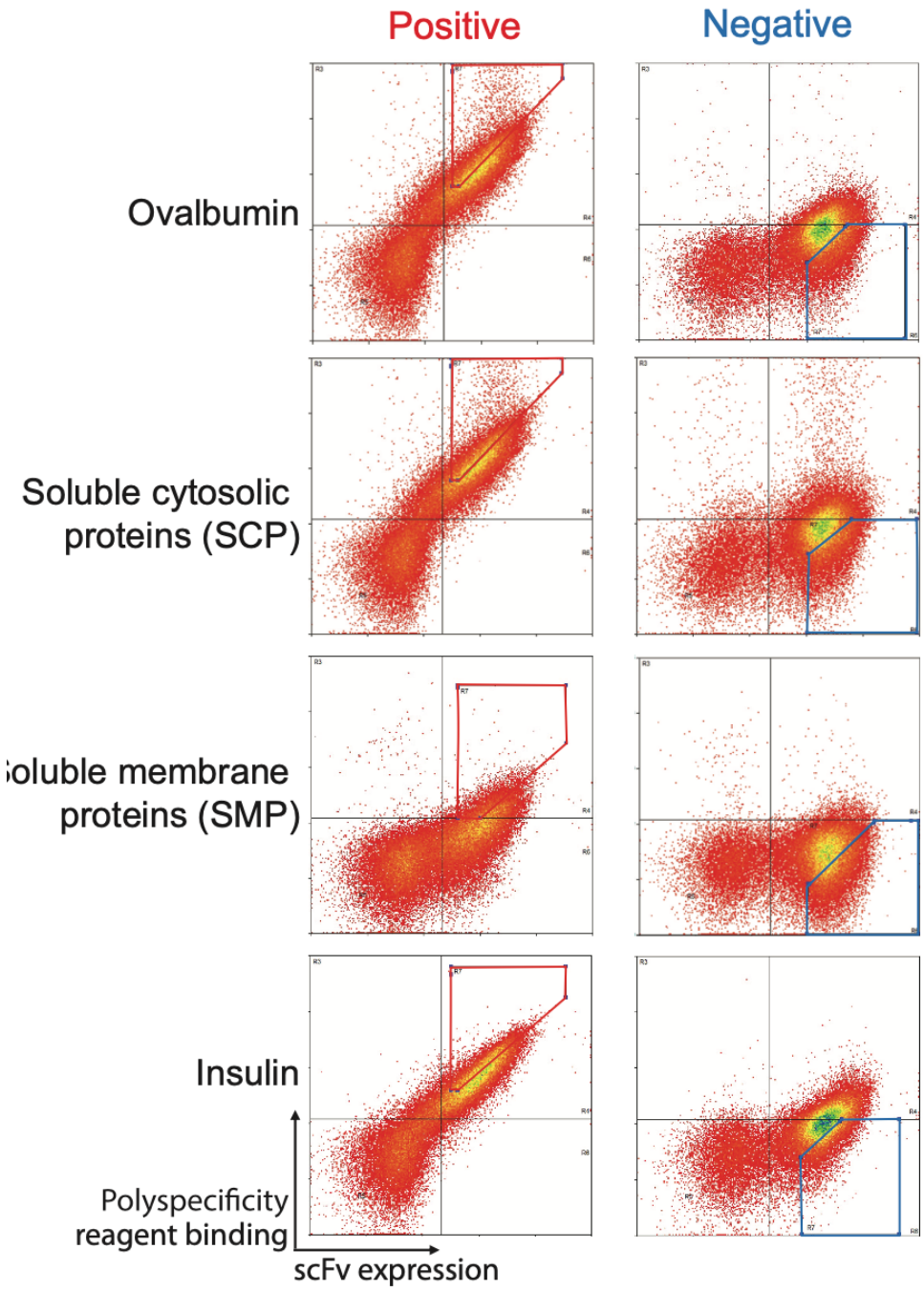


Figure 4-2: Summary of FACS sorting of a human scFv library to entich for high and low levels of antibody non-specific binding to multiple polyspecificity reagents.

The human scFv library reported previously was displayed on the surface of yeast and sorted (in order) against ovalbumin (0.13 mg/mL; FACS sort #1), soluble cytosolic proteins (SCP) from CHO cells (0.13 mg/mL; FACS sort #2), soluble membrane proteins from VHO cells 0.13 mg/mL; FACS sort #3) and insulin (0.13 mg/mL; FACS sort #4). FACS cytograms are shown for FACS sort #5, which was performed using the output libraries from FACS sort #4 to collect samples for deep sequencing analysis. The FACS cytograms report the antibody (scFv) expression onn the x-axis and the non-specific binding on the y-axis. The positive and negative non-specific binding populations that were selected for deep sequencing analysis are shown in red (high non-specific binding) and blue (low non-specific binding) gates.

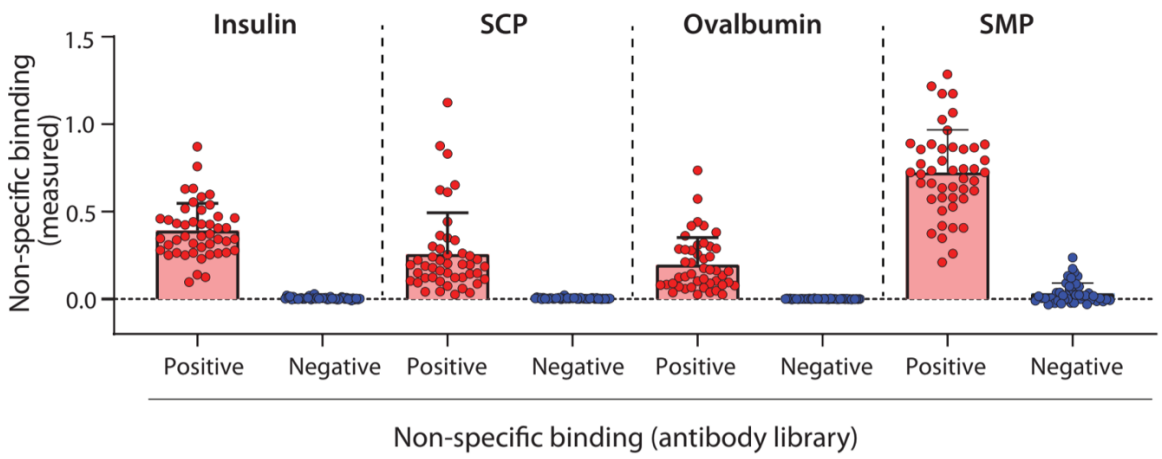


Figure 4-3: Evaluation of non-specific binding of human scFvs enriched during library sorting for high and low levels of polyspecificity.

Individual scFv variants were isolated after FACS sort #4 (as defined in Fig. 2) via Sanger sequencing and their levels of non-specific binding were evaluated using yeast surface display and flow cytometry. The four polyspecificity reagents were used as described in Fig. 2 except that 0.026 mg/mL of each reagent was used. The reported levels of non-specific binding were first normalized to their scFv expression levels and then normalized between two scFv standards (FN4 for the negative control and FP3 for positive control; Dataset S3). Data shown are averages of two independent repeats, and the error bars are standard errors.

Molecular features strongly linked to antibody polyreactivity. The enriched library samples were deep sequenced using PacBio, which resulted in the generation of two large datasets, namely Dataset S1 for library #1 with 115,233 scFvs with low polyreactivity and 131,754 scFvs with high polyreactivity, and Dataset S2 for library #2 with 99,120 scFvs with low polyreactivity and 38,500 scFvs with high polyreactivity. Herein we use Dataset S1 for analyzing molecular features linked to antibody polyreactivity and developing models of predictions of antibody polyreactivity.

We first evaluated the average theoretical net charges and hydrophobicities for each position in VH and VL for the scFvs identified with high and low polyreactivity in antibody set #1 (n=246,987 scFvs; Figs. 4 and S2). The net charge was evaluated at pH 7.4 assuming charges of +1 for lysine and arginine, +0.1 for histidine, and -1 for glutamate and aspartate. The hydrophobicity was evaluated for each amino acid using the scale of Black and Mould.(89) The profiles for the charged and hydrophobicity values for each site and the changes in each property at each site in VH and VL are shown in Figure 4-4. It is notable that the differences in charge for

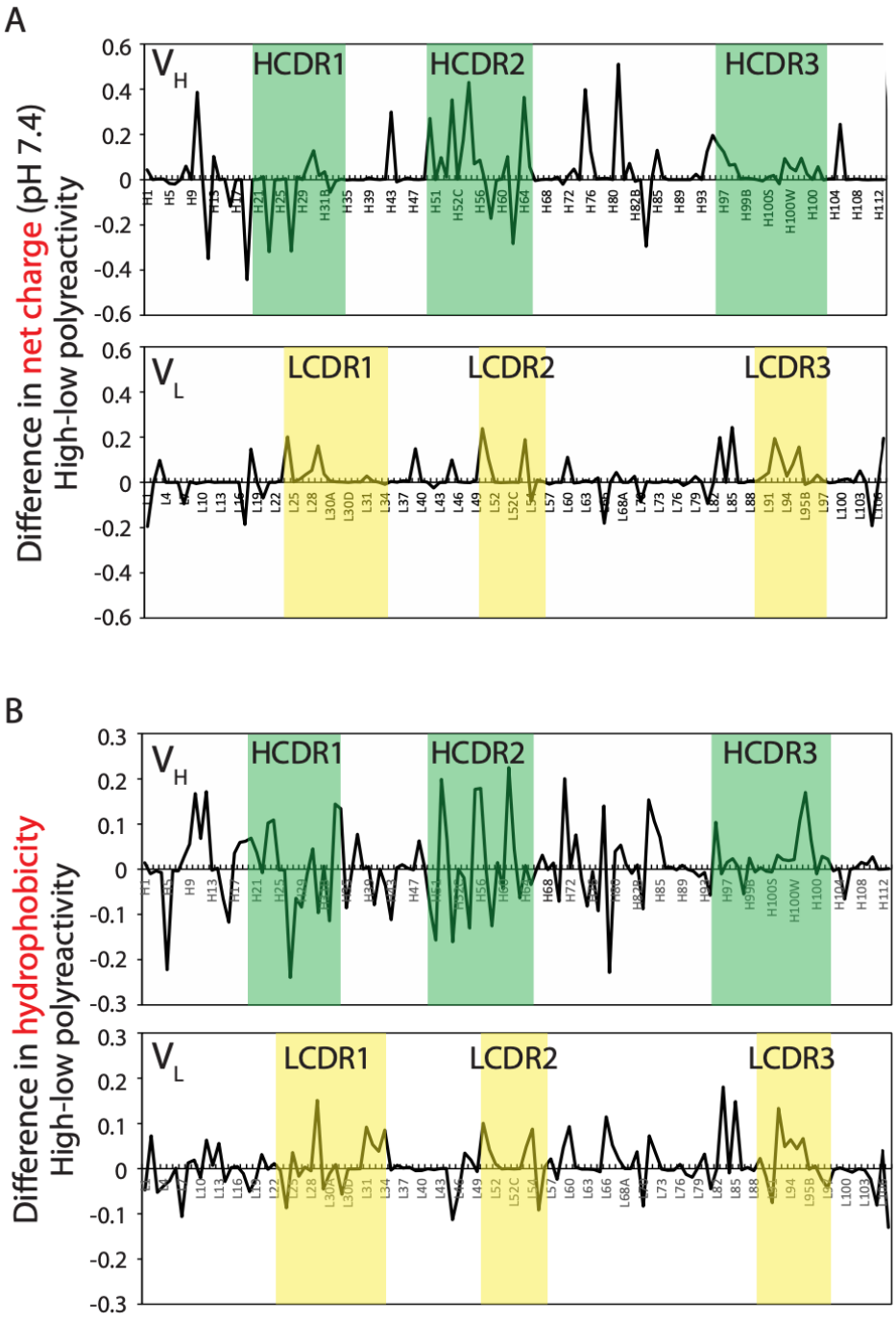


Figure 4-4: Differences in site-specific charge and hydrophobicity properties in the V_H and V_L regions of human scFvs with high polyreactivity relative to those with low polyreactivity.

The average theoretical net charges and hydrophobicities for each position in V_H and V_L for the scFvs (antibody set #1, n=246,987) were evaluated, and the differences are reported for scFvs with high polyreactivity relative to those with low polyreactivity. The net charge was evaluated at pH 7.4 assuming charges of +1 for lysine and arginine, +0.1 for histidine, and -1 for glutamate and aspartate. The hydrophobicity was evaluated for each amino acid using the scale of Black and Mould.(89) The antibody numbering is described in a previous report.(40)

scFvs with high polyreactivity relative to those with low polyreactivity occur both in the CDRs and frameworks (Figure 4-4A). Moreover, the differences in charge are more pronounced for VH than for VL, and heavy chain CDR2 displays the most obvious increases in positive charge of the six CDRs. Likewise, the differences in hydrophobicity for scFvs with high polyreactivity relative to those with low polyreactivity occur both in the CDRs and frameworks (Figure 4-4B). Moreover, the differences in hydrophobicity, like charge, are more pronounced for VH than for VL, and heavy chain CDR2 displays the most obvious changes in hydrophobicity of the six CDRs, including relatively large alternating increases and decreases in hydrophobicity. Overall, these findings suggest that scFvs with increased polyreactivity display changes in both charge and hydrophobicity in the CDRs and framework regions, and these changes are most pronounced in the VH region.

We next evaluated three main types of molecular features for differentiating between scFvs with high and low levels of polyreactivity: i) sequence-based (SB) features; ii) SB features weighted by the predicted site-specific solvent exposure of each amino acid (SB-SE)(72); and iii) SB-SE features multiplied by the relative surface area of each amino acid side chain (SB-SE-SA; see Methods for more detail). To maximize generality of our findings and reduce reliance on assumptions associated with structural modeling, we mainly focused on molecular features that are based on the VH, VL and Fv domains, although the SB-SE features are based primarily on the CDRs. Interestingly, the most significant molecular feature linked to polyreactivity was Fv net charge (pH 7.4), as antibodies with high non-specific binding generally have positively charged Fv regions (4.5±2.7) while antibodies with low non-specific binding generally have negative or weakly positively charged Fv regions (-1.1±3.3; Figure 4-5A). A second feature that is strongly linked to non-specific binding is the combined number of positively charged (Arg, Lys, His) and

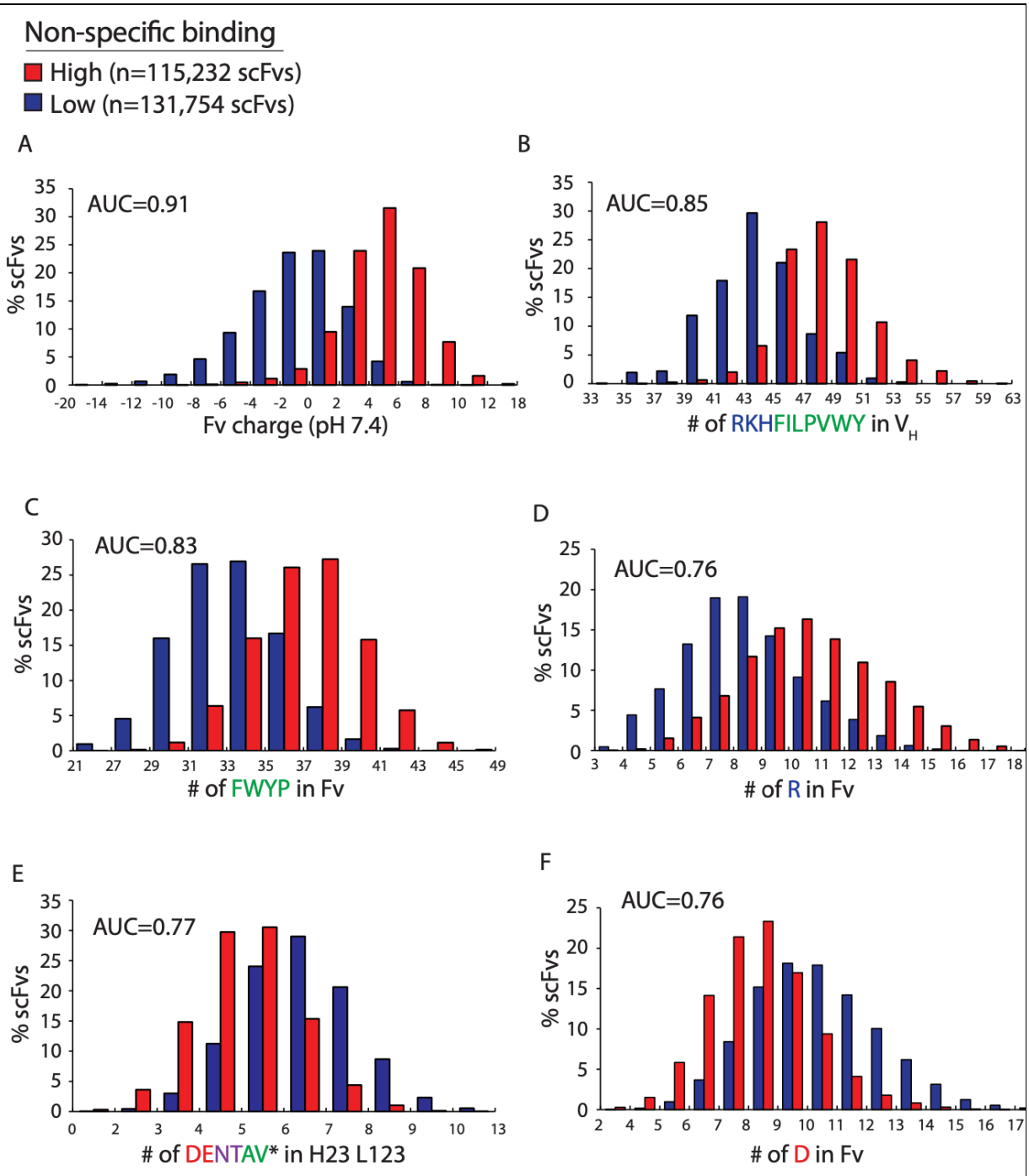


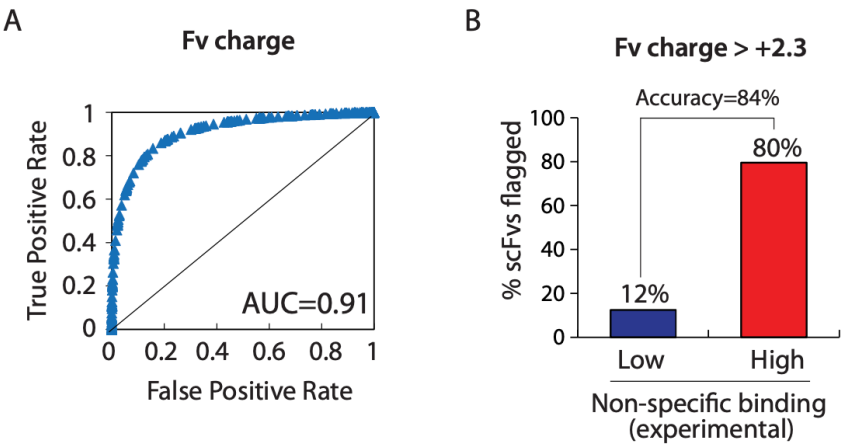
Figure 4-5: Molecular features that strongly differentiate between human scFvs with high and low levels of non-specific binding.

Distributions of key molecular features linked to non-specific binding of human scFvs (antibody set #1) were identified based on the area under the ROC curve (AUC) calculated from logistic regression analysis: (A) Fv charge calculated at pH 7.4 (Arg and Lys charge of +1, His charge of +0.1, and Asp and Glu charge of -1); (B) number of hydrophobic (Phe, Ile, Leu, Val, Trp, Tyr) and positively charged (Arg, Lys, His) amino acids in V_H; (C) number of aromatic amino acids (Phe, Tyr and Trp) and Proline residues; (D) number of arginines in Fv; (E) number of negatively charged (Asp, Glu), polar (Asn, Thr) and small hydrophobic (Ala, Val) residues in five CDRs (heavy chain CDRs 2 and 3 and light chain CDRs 1, 2 and 3) weighted by their site-specific solvent accessibilities; and (F) number of aspartic acid residues in Fv.

hydrophobic (Phe, Ile, Leu, Pro, Val, Trp, Tyr) residues in the VH domain, as polyreactive antibodies generally have higher levels of such residues (48.9±3.0) than non-polyreactive antibodies (44.6±3.1; Figure 4-5B). Other key features are the combined number of aromatic and proline residues (37.5±2.8 for polyreactive scFvs and 33.7±2.7 for non-polyreactive scFvs; Fig. 5C) and arginine residues (11.3±2.5 for polyreactive scFvs and 8.9±2.2 for non-polyreactive scFvs; Fig. 5D) in Fv.

While over-enrichment of positively charged and hydrophobic residues in antibody Fv and VH regions is linked to polyreactivity, we also find that underrepresentation of negatively charged, polar and small hydrophobic residues is also linked to polyreactivity (Figure 4-5E-F). For example, we find that the number of negatively charged (Asp and Glu), polar (Asn and Thr) and small hydrophobic (Ala and Val) residues, weighted by their site-specific solvent accessibilities, in five CDRs (heavy chain CDRs 2 and 3, light chain CDRs 1, 2 and 3) are linked to non-specific binding (Figure 4-5E). Antibodies with low non-specific binding have more of these residues (6.4±1.3) than those with low non-specific binding (5.1±1.2). Likewise, antibodies with low non-specific binding have more aspartic acid residues (10.8±2.1) than antibodies with high non-specific binding (8.9±1.8; Figure 4-5F).

The top 20 features (out of the 151 features evaluated) linked to non-specific binding with area under the curve (AUC) values >0.75 are summarized in Figure 4-6 along with the best single limits of each feature that segregate human scFvs with high and low levels of non-specific binding. For example, the receiver operating characteristic (ROC) curve of Fv charge (pH 7.4) reveals an AUC value of 0.91 (Fig. 6A) and the best limit of Fv charge (pH 7.4) that flags antibodies with high non-specific binding is >+2.3 (Figure 4-6B). The other top six molecular features that are enriched in antibodies with high non-specific binding are positive charge in VH (pH 7.4; AUC of



C

Region	Feature	Type	AUC	Rule	Accuracy (%)
Fv	Charge (pH7.4)	SB	0.91	>2.3	84
V_H	Charge (pH7.4)	SB	0.91	>2.1	83
V_H	FILPVWRKH	SB	0.85	>47	78
Fv	FILPVWRKH	SB	0.84	>86	76
Fv	FWYP	SB	0.83	>36	76
V_H	FWYP	SB	0.82	>20	75
Fv	FWY	SB	0.82	>25	74
Fv	FILPVWRKH	SB-SE-SA	0.82	>3024 Å^2	73
Fv	FWYP	SB-SE-SA	0.81	>1233 Å^2	73
V_L	FILPVWRKH	SB-SE-SA	0.81	>1788 Å^2	72
V_H	FILPVW	SB	0.80	>36	73
V_H	FWY	SB	0.79	>15	71
V_L	FWYP	SB-SE-SA	0.79	>746 Å^2	71
Fv	FILPVW	SB	0.78	>66	71
H23 L123	DENTAV	SB-SE	0.77	<5.8	71
V_H	W	SB	0.77	>5	70
Fv	R	SB	0.76	>10	70
Fv	D	SB	0.76	<9	69
Fv	FWY	SB-SE-SA	0.75	>717 Å^2	68
V_H	R	SB	0.75	>7	70

Figure 4-6: Summary of the top 20 molecular features and their optimal limits that best differentiate between human scFvs with high and low levels of non-specific binding.

The molecular features were identified for scFvs in antibody set #1 based on the highest 20 AUC values, as described in Fig. 4. (A) Receiver operating characteristic (ROC) curve calculated from logistic regression analysis of Fv charge (pH 7.4) versus the experimental classification (high or low non-specific binding) of the human scFvs. (B) Maximum limit of Fv charge (pH 7.4) that most selectively flags antibodies with high non-specific binding. Most (80%) of the scFvs that were enriched for non-specific binding are those with Fv charge >+2.3, while few of the scFvs enriched for low non-specific binding have such positively charged Fv regions. (C) Summary of the 20 most selective molecule features for differentiating between scFvs with high and low levels of non-specific binding. The features were selected based on the AUC values, as described in (A), and the optimal limits of each feature are reported, as described in (B). Three types of features are reported: sequence based (SB); SB features weighted by site-specific solvent exposure (SB-SE); and SB-SE features multiplied by the side chain surface area of each amino acid (SB-SE-SA). Most of the features have maximum limits, which are defined as those that flag scFvs as having high non-specific binding when the limit value is exceeded, while two of the features have minimum limits and are defined in the opposite manner. The accuracies for each molecular feature and the corresponding limit value were calculated using a confusion matrix and adjusting for the different number of scFvs in antibody set #1 with high (n=115232) and low (n=131754) non-specific binding.

0.91), the combined number of hydrophobic (Phe, Ile, Leu, Pro, Val and Trp) and positively (Arg, Lys and His) charged residues in VH and Fv, and the combined number of aromatic (Phe, Trp and Tyr) and proline residues in Fv and VH (Figure 4-6C). Interesting, the four of the top 20 features are individual amino acids, as over-enrichment in tryptophan in VH and arginine in Fv and VH and underrepresentation of aspartic acid in Fv is linked to high non-specific binding.

Random forest models for classifying antibody polyreactivity. We next sought to evaluate if the molecular features linked to antibody polyreactivity could be combined into a random forest model to accurately predict antibody polyreactivity. This process was conducted in two steps. For the first-generation model, we generated in a random forest model using all of the 151 molecular features to identify the most important features for the second-generation model. Our approach used standard tenfold cross-validation (80% training and 20% test for antibody set #1), and our primary performance objective was to minimize the coefficient of variation for the ten calculated validation accuracies (see Methods for more detail). To make the model more general, we also required the model predict the level of non-specific binding for a set of 103 clinical-stage mAbs (antibody sets #5)(4) with an area under the ROC curve of >0.75. This resulted in model #1 that predicted scFv non-specific binding with a training accuracy of 97.3% and test accuracy of 96.2% (antibody set #1), and an area under the ROC curve of 0.75 for clinical-stage mAbs (antibody sets #4 and #5).

Therefore, we next sought to develop a second-generation random forest model (model #2) with improved performance. Our approach involved three key modifications to those used to train model #1. First, we limited the molecular features used by model #2 to the top 10 features that most strongly impacted the performance of model #1. Second, we also used twelve SB-SE features reported previously for predicting antibody non-specific binding.(72) Third, we added an

additional performance requirement that the accuracy of model #2 is >95% for the average of the ten validation accuracies (antibody set #1).

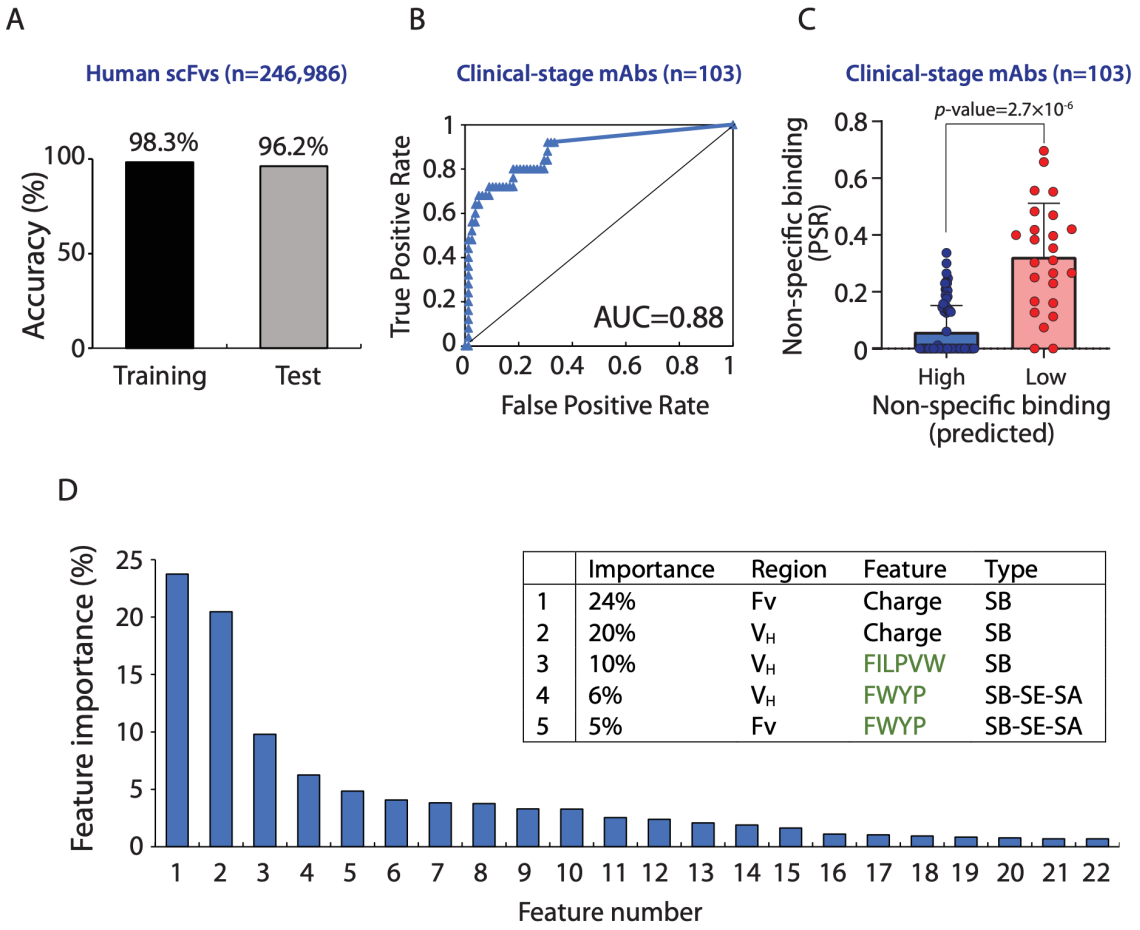


Figure 4-7: Training performance of random forest model for predicting human scFvs with low and high levels of non-specific binding.

(A) Validation (training) and test performance of the model trained on 80% of antibody set #1 (tenfold cross validation) and tested on the other 20%. (B) ROC curve for logistic regression analysis of previously reported measurements of non-specific (PSR) binding for clinical-stage antibodies (antibody set #5)(4) relative to the model predictions. (C) Comparison of non-specific binding (PSR) values for clinical-stage antibodies and model predictions. (D) The relative importance of each of the 22 molecular features used during model training. In (B) and (C), the 103 clinical-stage antibodies (antibody set #5) and the non-specific binding (PSR) measurements were used during model training. In (D), the relative importance of the features were evaluated using the training set of antibodies (80% of antibody set #1), and the net charge feature was evaluated at pH 7.4.

The performance of our second-generation model is summarized in Figure 4-8. The training and test accuracies were >96% for antibody set #1 (Figure 4-8A). Moreover, the area under the ROC curve was 0.88 for the 103 clinical-stage mAbs (antibody set #5; Figure 4-8B), which corresponded to a highly significant p-value (2.7x10-6) for flagging mAbs with high levels of non-specific binding (Figure 4-7C). The relative feature importance of model #2 is summarized

in Figure 4-7D, which reveals that charge of Fv and VH and combined numbers of aromatic and hydrophobic residues are the most important features in predicting antibody non-specific binding.

We next evaluated model #2 using several sets of independent data. First, we evaluated the model predictions for the 62 human scFvs identified via Sanger sequencing after sorting library #1, namely 45 scFvs in the library sample enriched for high non-specific binding and 43 scFvs in the library sample enriched for low non-specific binding (antibody set #3). Importantly, none of these 62 scFvs were in the sets of antibodies used for training the model. Encouragingly, we find an overall (adjusted) accuracy of 98.8% for predicting the level of non-specific binding for the 62 scFvs for a total of four polyspecificity reagents, including area under the ROC curve of 1 for three reagents (insulin, ovalbumin and soluble cytosolic proteins) and 0.99 for the fourth reagent (soluble membrane proteins).

Next, we generated an independent panel of 47 human scFvs that were based on 27 preclinical antibodies (antibody set #6) and 20 clinical-stage antibodies (antibody set #4), and evaluated their levels of non-specific binding for four polyspecificity reagents (Figure 4-8). Notably, the model results in area under the ROC curve values of 0.83-0.99 (Figure 4-8), and the predictions for flagging antibodies with high non-specific binding are significant for each set of measurements (*p*-values <10-3; Figure 4-9B). Similar strong results were also observed for the predictions of mAb non-specific binding for a panel of 20 clinical-stage antibodies (antibody set #4; Figure 4-9), as the area under the ROC curve was 0.96 and the *p*-value for flagging mAbs with high non-specific binding was 9.8x10-7. Finally, to complement the predictions of experimental measurements of non-specific binding for individual human scFvs, we also evaluated the model performance for predicting the classification of >100,000 scFvs in library #2 (Figure 4-10). Importantly, none of the scFvs in antibody set #2 were observed in antibody set #1. Encouragingly,

we observed that a high area under the ROC curve value (0.91; Figure 4-10A) and accuracy (83%; Figure 4-10B). These results collectively demonstrate the strong and general predictive power of the second-generation model.

We also sought to compare the performance of our best random forest model to a second random forest model of similar complexity that also included CDR (SB and SB-SE-SA) features as well as different types of models or simple rules based on single molecular properties First, we compared the performance of our random forest model (#1) based primarily on molecular features from the Fv, VH and VL regions to that of a similar model that also includes SB and SB-SE-SA features for the three heavy chain CDRs, three light chain CDRs and all six CDRs. While this type of model was able to perform relatively well for the deep sequencing datasets (accuracy >0.8 for sets #1 and #2) and for the measurements of scFv non-specific binding for clones from library #1 (AUC of 1), it displayed reduced performance for an additional set of pre-clinical and clinical-stage scFvs (AUC of 0.772 for antibody sets #4 and #6 relative to 0.902 for random forest model #1) and clinical-stage mAbs (AUC of 0.733 for antibody set #5 relative to 0.880 for random forest model #2).

The other models, namely the LDA model based on One-Hot Encoded sequence features and the SVC model based on PSSM features, displayed reasonable performance for the primary training set (accuracies of 0.87-0.98 for antibody set #1), but displayed reduced performance for other datasets (e.g., accuracies of 0.501-0.766 for antibody set #2; Table 1). In the case of the SVC model, it strongly predicted the opposite specificity of clinical-stage mAbs (AUC of 0.138 for antibody set #5). We also tested the performance of three of the molecular features that were most strongly correlated with antibody specificity, namely Fv and VH charge (pH 7.4) and the total number of specific hydrophobic and positively-charged residues (Phe, Ile, Leu, Val, Pro, Trp, Arg,

Lys and His) in VH, in the form of rules for predicting polyreadctivity. These simple rules, while not performing as well as our best random forest model, displayed reasonable performance (e.g., accuracy of 0.793 for antibody set #2 for the Fv charge rule relative to 0.830 for random forest model #1), which further demonstrates the significance of these molecular features in mediating antibody polyreactivity. Overall, these findings demonstrate the challenge of generally predicting antibody polyreactivity and the superior performance of the machine model based primary on VH, VL and Fv features.

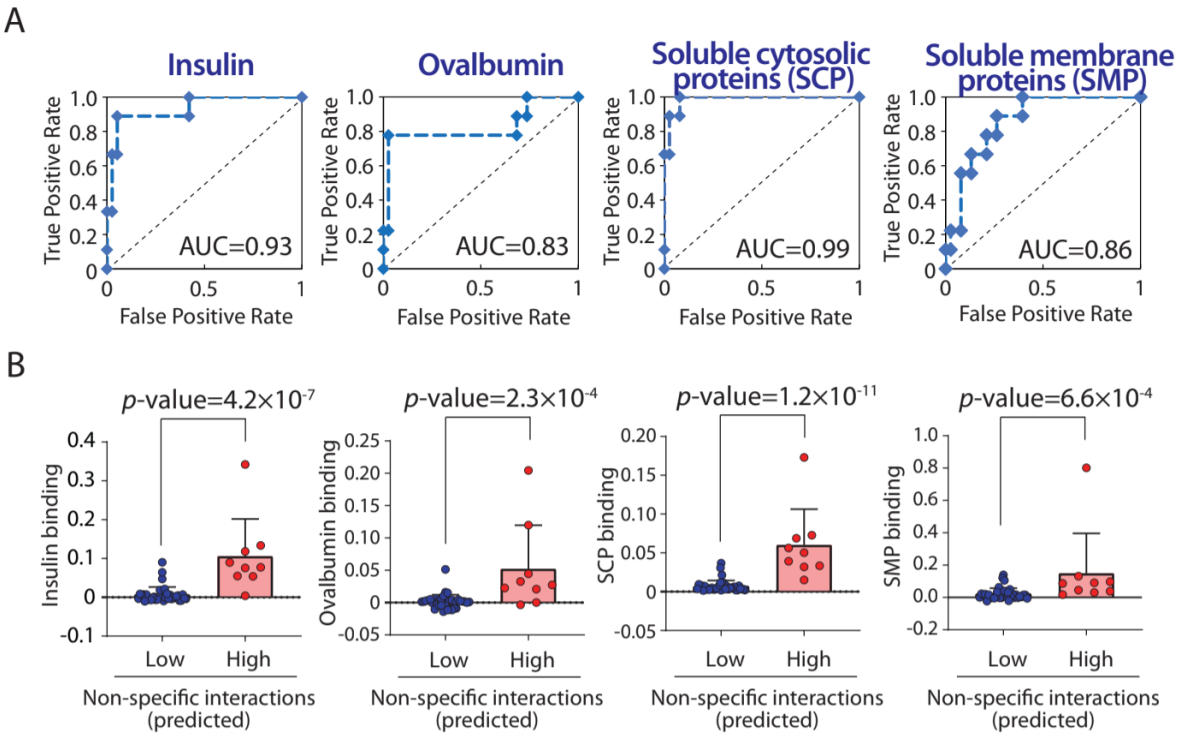


Figure 4-8: Evaluation of random forest model predictions using an independent set of human scFvs generated from pre-clinical and clinical-stage antibodies.

Human scFvs were generated for 27 pre-clinical antibodies (antibody set #6) and 20 clinical-stage antibodies (antibody set #4), and their level of non-specific binding was evaluated for four polyspecificity reagents, namely insulin, ovalbumin, soluble cytosolic proteins (SCP) and soluble membrane proteins (SCP). (A) ROC curves for logistic regression analysis of the non-specific binding measurements relative to the model predictions. (B) Comparison of scFv non-specific binding measurements and model predictions. In (A) and (B), the non-specific binding measurements were evaluated using yeast surface display and flow cytometry. In (B), the data are averages of four independent experiments, and the *p*-values were calculated using the Anderson-Darling test.

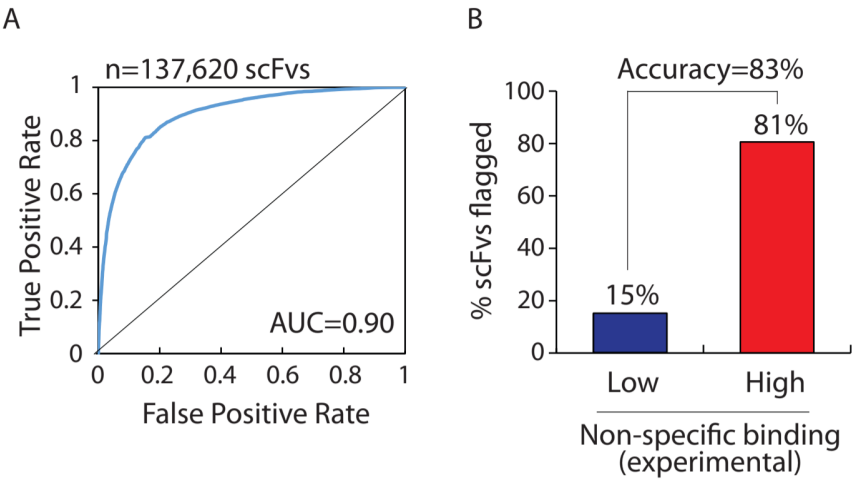


Figure 4-9: Evaluation of random forest model using deep sequencing data obtained for a second human scFv library that was sorted for high and low levels of non-specific binding.

Human scFvs in antibody set #2 (n=137,620) were evaluated using the random forest model reported in Figure 4-4. (A) ROC curve calculated from logistic regression analysis of the model prediction of the probability of high non-specific binding versus the experimental classification (high or low non-specific binding) of the human scFvs. (B) Experimental classification of antibody non-specific binding based on deep sequencing data versus the percentage of mAbs predicted by the model to display high non-specific binding. The accuracy was calculated using a confusion matrix and adjusting for the different number of scFvs in antibody set #2 with high (n=99,210) and low (n=38,500) non-specific binding.

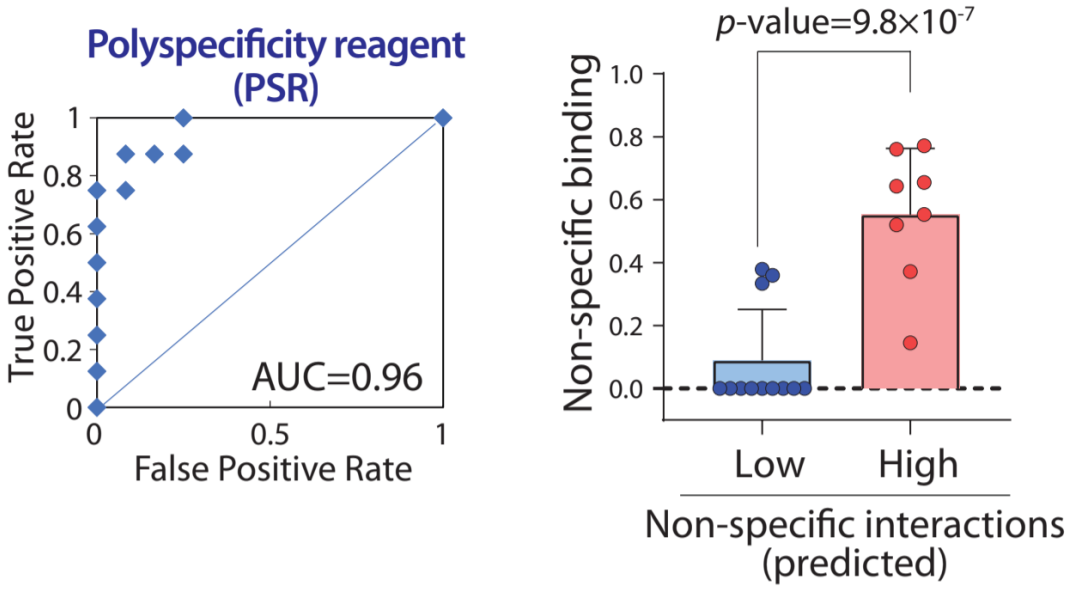


Figure 4-10: Evaluation of the random forest model for predicting non-specific binding for a panel of clinical stage mAbs.

(A) ROC curve for logistic regression analysis of previously reported measurements of non-specific (PSR) binding for clinical-stage mAbs in the IgG format (antibody set #4) relative to the model predictions. (B) Comparison of non-specific binding (PSR) values for clinical stage mAbs and model predictions. In (A) and (B), the 20 clinical-stage antibodies (antibody set #4) and the non-specific binding (PSR) measurements were not used during model training.

Chapter 5 **Conclusion**

Antibodies are considered as one of the most promising drugs in therapeutics, widely used in cancer, Alzheimer and Parkinson treatments, based on their attractive properties, including high binding affinity and specificity for target molecules, high solubility, and the ability to initiate immune respond after target recognition. However, the process of discovering and developing antibody drugs is time-consuming and expensive. Nowadays, only approximately 200 antibodies are investigating in the clinical trials. It is critical to identify drug like antibodies in the early discovery stage to reduce the risk of failure in the late clinical stage, aiming to save time and resources. Back in 1990s, Lipinski's rule of five, which summarized and developed using properties from thousands small molecule drugs, can successfully identify drug-like small molecules based on their number of hydrogen bond doner and acceptors, molecular mass and water-octanal partition. Inspired from that, it is also valuable to develop Lipinski-like rules for identifying antibodies with drug-like properties. Our goal is to develop physicochemical descriptors or machine learning models for identifying antibodies with drug-like properties during early stages of clinical development.

In 2017, 137 clinical stage monoclonal antibodies were evaluated using twelve biophysical assays for different antibody properties including binding specificity, stability, and aggregation etc. The authors also established the upper limit for each assay based on the worst 10% of the approved drugs. According to their results, five of the biophysical assays (PSR, AC-SINS, CSI, ELISA and BVP) that either evaluate level of antibody non-specific binding or self-interactions displayed

statistical significance in discriminating approved drugs and those in phase II and III using logistic regression *p*-value (the lower the *p*-value, the higher the significance). As a result, we argue that antibody specificity is one of the most important properties to identify drug-like candidates. Practically, high level of antibody non-specific binding or interaction can lead to high clearance rate and poor formulation for injection which lower the effect of dosage.

In the first project, we developed a set of twelve chemical rules using clinical and pre-clinical stage monoclonal antibodies for identifying antibodies with high specificity. Each chemical rule was composed by amino acid composition weighted by their predicted solvent accessible surface area in all combination of CDR and Fv regions. Antibodies with more than eight chemical flags were identified as antibodies with low specificity with 84% accuracy. Positively charged amino acid (e.g., Arg) and negatively charged and polar amino acid (e.g., Asp, Glu, and Asn) were most contributed to chemical rules with maximum and minimum allowable limits, indicating that there were enriched number of exposed of these amino acids in the group of antibodies with low and high specificity, respectively. This finding also matched with previous experimental measurements. The developed chemical rules displayed a strong identification power evaluated using two panels of pre-clinical stage mAbs with non-specific measurements.

We further mutagenesis one of the clinical stage mAb (emibetuzumab) that displayed high level of non-specific binding based on biophysical measurements to improve the specificity using the developed chemical rules. The target amino acids that we identified as problematic sites were those exposed hydrophobic and positively charged amino acids (>10%) that were involved in the maximum chemical rules in the specified CDR domains in heavy chain. The mutations included at least one polar or negatively charged amino acid using degenerate codon. The final designed library of emibetuzemab included eight mutation sites with library diversity of 10^7 and then this

library was sorted using two poly-specific reagents, PSR and ovalbumin, and collected samples from next generating sequencing. Surprisingly, the clones with identified mutations that linked to reducing non-specific interactions had higher enrichment ratio in samples negatively binding with poly-specific reagents and these clones also showed decreased number of chemical flags compared with those clones in the input library sample. To summarize this project, our findings provide a relatively simple but powerful descriptions for overall specificity of monoclonal antibodies. The identified chemical rules can not only predict antibody specificity but also can improve non-specific interactions by guiding the library design.

In the second project, we investigated sets of two or three mutations in antibody CDR regions that strongly reduce non-specific binding in the same emibezutumab library samples. The clones with identified mutations had higher enrichment ratio in both PSR and ovalbumin negative gates, compared with clones with wide type residues at corresponding sites. We then evaluated the level of humanness for these two groups of clones using two approaches, human string content and T20 score. Human string content was focusing on calculating sequence identity for all pieces of nine amino acid peptides between target sequence and a set of human antibodies. T20 score was focusing on global sequence similarity, including amino acid evolution information (Bolsum62). We also identified sets of two or three mutations that were not linked to non-specific binding as control. Interestingly, our findings showed that clones with mutations that reduce non-specific binding not only displayed fewer number of chemical flags but also showed higher level of humanness compared with those with wide type residues at corresponding sites using spearman's correlation. However, in the control set, clones with mutations that were not linked to non-specific binding displayed a similar number of chemical flag distribution to those with wide type residues. Clones with mutations displayed lower level of humanness. The results indicated that mutations

in CDR regions that strongly reduce non-specific binding also significantly increase humanness. This conclusion needs to be further evaluated using a different antibody library.

In the third project, we developed a robust Random Forest classification model using a panel of >200,000 human antibodies with high and low specificity. Within 150 sequence-based features we evaluated in this work, theoretical net charge at pH 7.4 in Fv region is the most significant feature to differentiate scFvs with high and low specificity which also matched with the previous results. Surprisingly, the selected random forest model not only displayed an excellent performance on training and test set for the large human antibodies, but also had a strong ability to identify antibodies with high level of non-specific on an additional human antibody library (>130,000 scFvs) and the panel of clinical stage mAbs. Notably, all features were trained in this model were all sequence-based features. No antibody structures were generated which made this model more impressive.

In this thesis we investigated chemical rules and machine learning models for identifying antibodies with low specificity and discovered the correlation between level of humanness and non-specific interaction, however, there are much to be further studied in this topic. First, in terms of library designed method guided by the identified chemical rules, it is valuable to evaluate humanization score for each possible mutation and select the ones that can not only reduce non-specific and self-interaction but also can increase humanness. The improved library design approach can maximize the likelihood of success in the clinic. Second, in Chapter 3, we investigated the correlation between the mutation that reduce non-specific binding and level of humanness from one humanized mouse antibody library. The conclusion should be further validated using another antibody library. It is also interesting to investigate the correlation between the mutation that increase antigen binding affinity and level of humanness. This can be done by

identifying mutations that show increased enrichment ratio in the antigen positive gate. Third, the random forest model developed in Chapter 4 can be improved by adding other features, for example, position-specific features PSSM (position specific scoring matrix) and structural based features (area of hydrophobic patches, dipole moment). Protein structure prediction for large set of antibodies is the time limiting step which make the generation of structural based features challenging. Furthermore, it will be valuable to check the liability of the model predictions for the target antibodies by comparing the feature values with the training set, so that we can make assessment if the model is reliable when applying to a new dataset. Finally, people have been proved that maximize antibody specificity can lead to decreased antigen binding affinity. Although we argue that specificity is the most important property to identify drug-like molecules, it is vital important to maintain or even increase binding affinity. As a result, co-optimization of two or multiple properties can be considered as a significant breakthrough in antibody development.

References

1. Tiller, K. E., and Tessier, P. M. (2015) Advances in antibody design. *Annu Rev Biomed Eng* **17**, 191-216
2. Chiu, M. L., Goulet, D. R., Teplyakov, A., and Gilliland, G. L. (2019) Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies (Basel)* **8**
3. Sinkora, M., Sun, J., and Butler, J. E. (2000) Antibody repertoire development in fetal and neonatal piglets. V. VDJ gene chimeras resembling gene conversion products are generated at high frequency by PCR in vitro. *Mol Immunol* **37**, 1025-1034
4. Jain, T., Sun, T. W., Durand, S., Hall, A., Houston, N. R., Nett, J. H., Sharkey, B., Bobrowicz, B., Caffry, I., Yu, Y., Cao, Y., Lynaugh, H., Brown, M., Baruah, H., Gray, L. T., Krauland, E. M., Xu, Y. D., Vasquez, M., and Wittrup, K. D. (2017) Biophysical properties of the clinical-stage antibody landscape. *P Natl Acad Sci USA* **114**, 944-949
5. Umscheid, C. A., Margolis, D. J., and Grossman, C. E. (2011) Key concepts of clinical trials: a narrative review. *Postgrad Med* **123**, 194-204
6. Menter, A., Tyring, S. K., Gordon, K., Kimball, A. B., Leonardi, C. L., Langley, R. G., Strober, B. E., Kaul, M., Gu, Y., Okun, M., and Papp, K. (2008) Adalimumab therapy for moderate to severe psoriasis: A randomized, controlled phase III trial. *J Am Acad Dermatol* **58**, 106-115
7. Bailly, M., Mieczkowski, C., Juan, V., Metwally, E., Tomazela, D., Baker, J., Uchida, M., Kofman, E., Raoufi, F., Motlagh, S., Yu, Y., Park, J., Raghava, S., Welsh, J., Rauscher, M., Raghunathan, G., Hsieh, M., Chen, Y. L., Nguyen, H. T., Nguyen, N., Cipriano, D., and Fayadat-Dilman, L. (2020) Predicting Antibody Developability Profiles Through Early Stage Discovery Screening. *MAbs* **12**, 1743053
8. Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* **46**, 3-26
9. Xu, Y., Roach, W., Sun, T., Jain, T., Prinz, B., Yu, T. Y., Torrey, J., Thomas, J., Bobrowicz, P., Vasquez, M., Wittrup, K. D., and Krauland, E. (2013) 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* **26**, 663-670
10. Mouquet, H., Scheid, J. F., Zoller, M. J., Krogsgaard, M., Ott, R. G., Shukair, S., Artyomov, M. N., Pietzsch, J., Connors, M., Pereyra, F., Walker, B. D., Ho, D. D., Wilson, P. C., Seaman, M. S., Eisen, H. N., Chakraborty, A. K., Hope, T. J., Ravetch, J. V., Wardemann, H., and Nussenzweig, M. C. (2010) Polyreactivity increases the apparent affinity of anti-HIV antibodies by heteroligation. *Nature* **467**, 591-U117
11. Hotzel, I., Theil, F. P., Bernstein, L. J., Prabhu, S., Deng, R., Quintana, L., Lutman, J., Sibia, R., Chan, P., Bumbaca, D., Fielder, P., Carter, P. J., and Kelley, R. F. (2012) A strategy for risk mitigation of antibodies with fast clearance. *MAbs* **4**, 753-760
12. Sule, S. V., Sukumar, M., Weiss, W. F., Marcelino-Cruz, A. M., Sample, T., and Tessier, P. M. (2011) High-Throughput Analysis of Concentration-Dependent Antibody Self-Association. *Biophys J* **101**, 1749-1757
13. Geng, S. B., Cheung, J. K., Narasimhan, C., Shameem, M., and Tessier, P. M. (2014) Improving Monoclonal Antibody Selection and Engineering using Measurements of Colloidal Protein Interactions. *J Pharm Sci-Us* **103**, 3356-3363

14. Liu, Y. Q., Caffry, I., Wu, J. M., Geng, S. B., Jain, T., Sun, T. W., Reid, F., Cao, Y., Estep, P., Yu, Y., Vasquez, M., Tessier, P. M., and Xu, Y. D. (2014) High-throughput screening for developability during early-stage antibody discovery using self-interaction nanoparticle spectroscopy. *Mabs* **6**, 483-492
15. Sun, T., Reid, F., Liu, Y., Cao, Y., Estep, P., Nauman, C., and Xu, Y. (2013) High throughput detection of antibody self-interaction by bio-layer interferometry. *MAbs* **5**, 838-841
16. Estep, P., Caffry, I., Yu, Y., Sun, T. W., Cao, Y., Lynaugh, H., Jain, T., Vasquez, M., Tessier, P. M., and Xu, Y. D. (2015) An alternative assay to hydrophobic interaction chromatography for high-throughput characterization of monoclonal antibodies. *Mabs* **7**, 553-561
17. Datta-Mannan, A., Lu, J., Witcher, D. R., Leung, D., Tang, Y., and Wroblewski, V. J. (2015) The interplay of non-specific binding, target-mediated clearance and FcRn interactions on the pharmacokinetics of humanized antibodies. *MAbs* **7**, 1084-1093
18. Piche-Nicholas, N. M., Avery, L. B., King, A. C., Kavosi, M., Wang, M., O'Hara, D. M., Tchistiakova, L., and Katragadda, M. (2018) Changes in complementarity-determining regions significantly alter IgG binding to the neonatal Fc receptor (FcRn) and pharmacokinetics. *Mabs* **10**, 81-94
19. Kelly, R. L., Yu, Y., Sun, T. W., Caffry, I., Lynaugh, H., Brown, M., Jain, T., Xu, Y. D., and Wittrup, K. D. (2016) Target-independent variable region mediated effects on antibody clearance can be FcRn independent. *Mabs* **8**, 1269-1275
20. Tomar, D. S., Kumar, S., Singh, S. K., Goswami, S., and Li, L. (2016) Molecular basis of high viscosity in concentrated antibody solutions: Strategies for high concentration drug product development. *Mabs* **8**, 216-228
21. Wardemann, H., Yurasov, S., Schaefer, A., Young, J. W., Meffre, E., and Nussenzweig, M. C. (2003) Predominant autoantibody production by early human B cell precursors. *Science* **301**, 1374-1377
22. Boughter, C. T., Borowska, M. T., Guthmiller, J. J., Bendelac, A., Wilson, P. C., Roux, B., and Adams, E. J. (2020) Biochemical patterns of antibody polyreactivity revealed through a bioinformatics-based analysis of CDR loops. *Elife* **9**
23. Datta-Mannan, A., Thangaraju, A., Leung, D., Tang, Y., Witcher, D. R., Lu, J., and Wroblewski, V. J. (2015) Balancing charge in the complementarity-determining regions of humanized mAbs without affecting pI reduces non-specific binding and improves the pharmacokinetics. *MAbs* **7**, 483-493
24. Kelly, R. L., Le, D., Zhao, J., and Wittrup, K. D. (2018) Reduction of Nonspecificity Motifs in Synthetic Antibody Libraries. *J Mol Biol* **430**, 119-130
25. Harding, F. A., Stickler, M. M., Razo, J., and DuBridge, R. B. (2010) The immunogenicity of humanized and fully human antibodies: residual immunogenicity resides in the CDR regions. *MAbs* **2**, 256-265
26. Kettleborough, C. A., Saldanha, J., Heath, V. J., Morrison, C. J., and Bendig, M. M. (1991) Humanization of a mouse monoclonal antibody by CDR-grafting: the importance of framework residues on loop conformation. *Protein Eng* **4**, 773-783
27. Lazar, G. A., Desjarlais, J. R., Jacinto, J., Karki, S., and Hammond, P. W. (2007) A molecular immunology approach to antibody humanization and functional optimization. *Mol Immunol* **44**, 1986-1998

28. Gao, S. H., Huang, K. X., Tu, H., and Adler, A. S. (2013) Monoclonal antibody humanness score and its applications. *Bmc Biotechnol* **13**
29. Carter, P. J., and Lazar, G. A. (2018) Next generation antibody drugs: pursuit of the 'high-hanging fruit'. *Nat Rev Drug Discov* **17**, 197-223
30. Beck, A., Wurch, T., Bailly, C., and Corvaia, N. (2010) Strategies and challenges for the next generation of therapeutic antibodies. *Nat Rev Immunol* **10**, 345-352
31. Maynard, J., and Georgiou, G. (2000) Antibody engineering. *Annu Rev Biomed Eng* **2**, 339-376
32. Perchiacca, J. M., Ladiwala, A. R., Bhattacharya, M., and Tessier, P. M. (2012) Aggregation-resistant domain antibodies engineered with charged mutations near the edges of the complementarity-determining regions. *Protein Eng Des Sel* **25**, 591-601
33. Wu, S. J., Luo, J., O'Neil, K. T., Kang, J., Lacy, E. R., Canziani, G., Baker, A., Huang, M., Tang, Q. M., Raju, T. S., Jacobs, S. A., Teplyakov, A., Gilliland, G. L., and Feng, Y. (2010) Structure-based engineering of a monoclonal antibody for improved solubility. *Protein Eng Des Sel* **23**, 643-651
34. Perchiacca, J. M., Bhattacharya, M., and Tessier, P. M. (2011) Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions. *Proteins* **79**, 2637-2647
35. Dudgeon, K., Rouet, R., Kokmeijer, I., Schofield, P., Stolp, J., Langley, D., Stock, D., and Christ, D. (2012) General strategy for the generation of human antibody variable domains with increased aggregation resistance. *Proc Natl Acad Sci U S A* **109**, 10879-10884
36. Dobson, C. L., Devine, P. W., Phillips, J. J., Higazi, D. R., Lloyd, C., Popovic, B., Arnold, J., Buchanan, A., Lewis, A., Goodman, J., van der Walle, C. F., Thornton, P., Vinall, L., Lowne, D., Aagaard, A., Olsson, L. L., Ridderstad Wollberg, A., Welsh, F., Karamanos, T. K., Pashley, C. L., Iadanza, M. G., Ranson, N. A., Ashcroft, A. E., Kippen, A. D., Vaughan, T. J., Radford, S. E., and Lowe, D. C. (2016) Engineering the surface properties of a human monoclonal antibody prevents self-association and rapid clearance in vivo. *Sci Rep* **6**, 38644
37. Kelly, R. L., Zhao, J., Le, D., and Wittrup, K. D. (2017) Nonspecificity in a nonimmune human scFv repertoire. *MAbs* **9**, 1029-1035
38. Starr, C. G., and Tessier, P. M. (2019) Selecting and engineering monoclonal antibodies with drug-like specificity. *Curr Opin Biotech* **60**, 119-127
39. Shehata, L., Maurer, D. P., Wec, A. Z., Lilov, A., Champney, E., Sun, T., Archambault, K., Burnina, I., Lynaugh, H., Zhi, X., Xu, Y., and Walker, L. M. (2019) Affinity Maturation Enhances Antibody Specificity but Compromises Conformational Stability. *Cell Rep* **28**, 3300-3308 e3304
40. Jain, T., Boland, T., Lilov, A., Burnina, I., Brown, M., Xu, Y. D., and Vasquez, M. (2017) Prediction of delayed retention of antibodies in hydrophobic interaction chromatography from sequence using machine learning. *Bioinformatics* **33**, 3758-3766
41. Sule, S. V., Dickinson, C. D., Lu, J., Chow, C. K., and Tessier, P. M. (2013) Rapid analysis of antibody self-association in complex mixtures using immunogold conjugates. *Mol Pharm* **10**, 1322-1331
42. Rabia, L. A., Zhang, Y., Ludwig, S. D., Julian, M. C., and Tessier, P. M. (2018) Net charge of antibody complementarity-determining regions is a key predictor of specificity. *Protein Eng Des Sel* **31**, 409-418

43. Narula, J., Petrov, A., Bianchi, C., Ditlow, C. C., Lister, B. C., Dilley, J., Pieslak, I., Chen, F. W., Torchilin, V. P., and Khaw, B. A. (1995) Noninvasive localization of experimental atherosclerotic lesions with mouse/human chimeric Z2D3 F(ab')2 specific for the proliferating smooth muscle cells of human atheroma. Imaging with conventional and negative charge-modified antibody fragments. *Circulation* **92**, 474-484
44. McCarthy, B. J., and Hill, A. S. (2001) Altering the fine specificity of an anti-Legionella single chain antibody by a single amino acid insertion. *J Immunol Methods* **251**, 137-149
45. Sakhnini, L. I., Greisen, P. J., Wiberg, C., Bozoky, Z., Lund, S., Wolf Perez, A. M., Karkov, H. S., Huus, K., Hansen, J. J., Bulow, L., Lorenzen, N., Dainiak, M. B., and Pedersen, A. K. (2019) Improving the Developability of an Antigen Binding Fragment by Aspartate Substitutions. *Biochemistry* **58**, 2750-2759
46. Julian, M. C., Rabia, L. A., Desai, A. A., Arsiwala, A., Gerson, J. E., Paulson, H. L., Kane, R. S., and Tessier, P. M. (2019) Nature-inspired design and evolution of anti-amyloid antibodies. *J Biol Chem* **294**, 8438-8451
47. Birtalan, S., Zhang, Y., Fellouse, F. A., Shao, L., Schaefer, G., and Sidhu, S. S. (2008) The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies. *J Mol Biol* **377**, 1518-1528
48. Birtalan, S., Fisher, R. D., and Sidhu, S. S. (2010) The functional capacity of the natural amino acids for molecular recognition. *Mol Biosyst* **6**, 1186-1194
49. Tiller, K. E., Li, L., Kumar, S., Julian, M. C., Garde, S., and Tessier, P. M. (2017) Arginine mutations in antibody complementarity-determining regions display context-dependent affinity/specificity trade-offs. *J Biol Chem* **292**, 16638-16652
50. Hong, G., Chappey, O., Niel, E., and Scherrmann, J. M. (2000) Enhanced cellular uptake and transport of polyclonal immunoglobulin G and fab after their cationization. *J Drug Target* **8**, 67-77
51. Sharma, V. K., Patapoff, T. W., Kabakoff, B., Pai, S., Hilario, E., Zhang, B., Li, C., Borisov, O., Kelley, R. F., Chorny, I., Zhou, J. Z., Dill, K. A., and Swartz, T. E. (2014) In silico selection of therapeutic antibodies for development: viscosity, clearance, and chemical stability. *Proc Natl Acad Sci U S A* **111**, 18601-18606
52. Igawa, T., Tsunoda, H., Tachibana, T., Maeda, A., Mimoto, F., Moriyama, C., Nanami, M., Sekimori, Y., Nabuchi, Y., Aso, Y., and Hattori, K. (2010) Reduced elimination of IgG antibodies by engineering the variable region. *Protein Eng Des Sel* **23**, 385-392
53. Boswell, C. A., Tesar, D. B., Mukhyala, K., Theil, F. P., Fielder, P. J., and Khawli, L. A. (2010) Effects of charge on antibody tissue distribution and pharmacokinetics. *Bioconjug Chem* **21**, 2153-2163
54. Arora, J., Hu, Y., Esfandiary, R., Sathish, H. A., Bishop, S. M., Joshi, S. B., Middaugh, C. R., Volkin, D. B., and Weis, D. D. (2016) Charge-mediated Fab-Fc interactions in an IgG1 antibody induce reversible self-association, cluster formation, and elevated viscosity. *MAbs* **8**, 1561-1574
55. Perchiacca, J. M., Lee, C. C., and Tessier, P. M. (2014) Optimal charged mutations in the complementarity-determining regions that prevent domain antibody aggregation are dependent on the antibody scaffold. *Protein Eng Des Sel* **27**, 29-39
56. Yadav, S., Laue, T. M., Kalonia, D. S., Singh, S. N., and Shire, S. J. (2012) The influence of charge distribution on self-association and viscosity behavior of monoclonal antibody solutions. *Mol Pharm* **9**, 791-802

57. Yadav, S., Sreedhara, A., Kanai, S., Liu, J., Lien, S., Lowman, H., Kalonia, D. S., and Shire, S. J. (2011) Establishing a link between amino acid sequences and self-associating and viscoelastic behavior of two closely related monoclonal antibodies. *Pharm Res* **28**, 1750-1764
58. Jespers, L., Schon, O., Famm, K., and Winter, G. (2004) Aggregation-resistant domain antibodies selected on phage by heat denaturation. *Nat Biotechnol* **22**, 1161-1165
59. Arbabi-Ghahroudi, M., To, R., Gaudette, N., Hirama, T., Ding, W., MacKenzie, R., and Tanha, J. (2009) Aggregation-resistant VHs selected by in vitro evolution tend to have disulfide-bonded loops and acidic isoelectric points. *Protein Eng Des Sel* **22**, 59-66
60. Alam, M. E., Geng, S. B., Bender, C., Ludwig, S. D., Linden, L., Hoet, R., and Tessier, P. M. (2018) Biophysical and Sequence-Based Methods for Identifying Monovalent and Bivalent Antibodies with High Colloidal Stability. *Mol Pharm* **15**, 150-163
61. Li, B., Tesar, D., Boswell, C. A., Cahaya, H. S., Wong, A., Zhang, J., Meng, Y. G., Eigenbrot, C., Pantua, H., Diao, J., Kapadia, S. B., Deng, R., and Kelley, R. F. (2014) Framework selection can influence pharmacokinetics of a humanized therapeutic antibody through differences in molecule charge. *MAbs* **6**, 1255-1264
62. Nichols, P., Li, L., Kumar, S., Buck, P. M., Singh, S. K., Goswami, S., Balthazor, B., Conley, T. R., Sek, D., and Allen, M. J. (2015) Rational design of viscosity reducing mutants of a monoclonal antibody: hydrophobic versus electrostatic inter-molecular interactions. *MAbs* **7**, 212-230
63. Li, L., Kumar, S., Buck, P. M., Burns, C., Lavoie, J., Singh, S. K., Warne, N. W., Nichols, P., Luksha, N., and Boardman, D. (2014) Concentration dependent viscosity of monoclonal antibody solutions: explaining experimental behavior in terms of molecular properties. *Pharm Res* **31**, 3161-3178
64. Yadav, S., Shire, S. J., and Kalonia, D. S. (2012) Viscosity behavior of high-concentration monoclonal antibody solutions: correlation with interaction parameter and electroviscous effects. *J Pharm Sci* **101**, 998-1011
65. Buck, P. M., Chaudhri, A., Kumar, S., and Singh, S. K. (2015) Highly viscous antibody solutions are a consequence of network formation caused by domain-domain electrostatic complementarities: insights from coarse-grained simulations. *Mol Pharm* **12**, 127-139
66. Yadav, S., Liu, J., Shire, S. J., and Kalonia, D. S. (2010) Specific interactions in high concentration antibody solutions resulting in high viscosity. *J Pharm Sci* **99**, 1152-1168
67. Yadav, D. B., Sharma, V. K., Boswell, C. A., Hotzel, I., Tesar, D., Shang, Y. L., Ying, Y., Fischer, S. K., Grogan, J. L., Chiang, E. Y., Urban, K., Ulufatu, S., Khawli, L. A., Prabhu, S., Joseph, S., and Kelley, R. F. (2015) Evaluating the Use of Antibody Variable Region (Fv) Charge as a Risk Assessment Tool for Predicting Typical Cynomolgus Monkey Pharmacokinetics. *Journal of Biological Chemistry* **290**, 29732-29741
68. Goulet, D. R., and Atkins, W. M. (2020) Considerations for the Design of Antibody-Based Therapeutics. *J Pharm Sci* **109**, 74-103
69. Jarasch, A., Koll, H., Regula, J. T., Bader, M., Papadimitriou, A., and Kettenberger, H. (2015) Developability assessment during the selection of novel therapeutic antibodies. *J Pharm Sci* **104**, 1885-1898
70. Makowski, E. K., Wu, L., Gupta, P., and Tessier, P. M. (2021) Discovery-stage identification of drug-like antibodies using emerging experimental and computational methods. *MAbs* **13**, 1895540

71. Rabia, L. A., Desai, A. A., Jhajj, H. S., and Tessier, P. M. (2018) Understanding and overcoming trade-offs between antibody affinity, specificity, stability and solubility. *Biochem Eng J* **137**, 365-374
72. Zhang, Y. L., Wu, L. N., Gupta, P., Desai, A. A., Smith, M. D., Rabia, L. A., Ludwig, S. D., and Tessier, P. M. (2020) Physicochemical Rules for Identifying Monoclonal Antibodies with Drug-like Specificity. *Mol Pharmaceut* **17**, 2555-2569
73. Raybould, M. I. J., Marks, C., Krawczyk, K., Taddese, B., Nowak, J., Lewis, A. P., Bujotzek, A., Shi, J. Y., and Deane, C. M. (2019) Five computational developability guidelines for therapeutic antibody profiling. *P Natl Acad Sci USA* **116**, 4025-4030
74. Kovaltsuk, A., Leem, J., Kelm, S., Snowden, J., Deane, C. M., and Krawczyk, K. (2018) Observed Antibody Space: A Resource for Data Mining Next-Generation Sequencing of Antibody Repertoires. *J Immunol* **201**, 2502-2509
75. Setliff, I., Shiakolas, A. R., Pilewski, K. A., Murji, A. A., Mapengo, R. E., Janowska, K., Richardson, S., Oosthuysen, C., Raju, N., Ronsard, L., Kanekiyo, M., Qin, J. S., Kramer, K. J., Greenplate, A. R., McDonnell, W. J., Graham, B. S., Connors, M., Lingwood, D., Acharya, P., Morris, L., and Georgiev, I. S. (2019) High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. *Cell* **179**, 1636-+
76. King, H. W., Orban, N., Riches, J. C., Clear, A. J., Warnes, G., Teichmann, S. A., and James, L. K. (2021) Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. *Sci Immunol* **6**
77. Eccles, J. D., Turner, R. B., Kirk, N. A., Muehling, L. M., Borish, L., Steinke, J. W., Payne, S. C., Wright, P. W., Thacker, D., Lahtinen, S. J., Lehtinen, M. J., Heymann, P. W., and Woodfolk, J. A. (2020) T-bet plus Memory B Cells Link to Local Cross-Reactive IgG upon Human Rhinovirus Infection. *Cell Rep* **30**, 351-+
78. DeKosky, B. J., Lungu, O. I., Park, D., Johnson, E. L., Charab, W., Chrysostomou, C., Kuroda, D., Ellington, A. D., Ippolito, G. C., Gray, J. J., and Georgiou, G. (2016) Large-scale sequence and structural comparisons of human naive and antigen-experienced antibody repertoires. *P Natl Acad Sci USA* **113**, E2636-E2645
79. Dunbar, J., and Deane, C. M. (2016) ANARCI: antigen receptor numbering and receptor classification. *Bioinformatics* **32**, 298-300
80. Henikoff, S., and Henikoff, J. G. (1992) Amino-Acid Substitution Matrices from Protein Blocks. *P Natl Acad Sci USA* **89**, 10915-10919
81. Chennamsetty, N., Voynov, V., Kayser, V., Helk, B., and Trout, B. L. (2010) Prediction of Aggregation Prone Regions of Therapeutic Proteins. *J Phys Chem B* **114**, 6614-6624
82. Desai, A. A., Smith, M. D., Zhang, Y., Makowski, E. K., Gerson, J. E., Ionescu, E., Starr, C. G., Zupancic, J. M., Moore, S. J., Sutter, A. B., Ivanova, M. I., Murphy, G. G., Paulson, H. L., and Tessier, P. M. (2021) Rational affinity maturation of anti-amyloid antibodies with high conformational and sequence specificity. *J Biol Chem* **296**, 100508
83. Swindells, M. B., Porter, C. T., Couch, M., Hurst, J., Abhinandan, K. R., Nielsen, J. H., Macindoe, G., Hetherington, J., and Martin, A. C. (2017) abYsis: Integrated Antibody Sequence and Structure-Management, Analysis, and Prediction. *J Mol Biol* **429**, 356-364
84. Feldhaus, M. J., Siegel, R. W., Opresko, L. K., Coleman, J. R., Feldhaus, J. M., Yeung, Y. A., Cochran, J. R., Heinzelman, P., Colby, D., Swers, J., Graff, C., Wiley, H. S., and Wittrup, K. D. (2003) Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library. *Nat Biotechnol* **21**, 163-170

85. Sheets, M. D., Amersdorfer, P., Finnern, R., Sargent, P., Lindquist, E., Schier, R., Hemingsen, G., Wong, C., Gerhart, J. C., and Marks, J. D. (1998) Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens. *Proc Natl Acad Sci U S A* **95**, 6157-6162
86. Makowski, E. K., Wu, L., Desai, A. A., and Tessier, P. M. (2021) Highly sensitive detection of antibody nonspecific interactions using flow cytometry. *MAbs* **13**, 1951426
87. Van Deventer, J. A., and Wittrup, K. D. (2014) Yeast surface display for antibody isolation: library construction, library screening, and affinity maturation. *Methods Mol Biol* **1131**, 151-181
88. Quinlan, A. R., and Hall, I. M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842
89. Black, S. D., and Mould, D. R. (1991) Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. *Anal Biochem* **193**, 72-82
90. Sheets, M. D., Amersdorfer, P., Finnem, R., Sargent, P., Lindqvist, E., Schier, R., Hemingsen, G., Wong, C., Gerhart, J. C., and Marks, J. D. (1999) Efficient construction of a large nonimmune phage antibody library: The production of high-affinity human single-chain antibodies to protein antigens (vol 95, pg 6157, 1998). *P Natl Acad Sci USA* **96**, 795-795