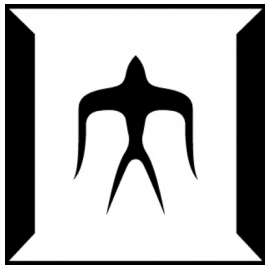
# 論文 / 著書情報
# Article / Book Information

| | |
|---|---|
| 題目(和文) | |
| Title(English) | Development of an Epitope Binning Platform for Parallel Antibody Classification with Mammalian Cell Surface Display |
| 著者(和文) | LinNing |
| Author(English) | Ning Lin |
| 出典(和文) | 学位:博士(工学),<br>学位授与機関:東京工業大学,<br>報告番号:甲第12901号,<br>授与年月日:2024年9月20日,<br>学位の種別:課程博士,<br>審査員:門之園 哲哉,大河内 美奈,越川 直彦,北口 哲也,石田 忠 |
| Citation(English) | Degree:Doctor (Engineering),<br>Conferring organization: Tokyo Institute of Technology,<br>Report number:甲第12901号,<br>Conferred date:2024/9/20,<br>Degree Type:Course doctor,<br>Examiner:,,,, |
| 学位種別(和文) | 博士論文 |
| Type(English) | Doctoral Thesis |



# Dissertation

## Development of an Epitope Binning Platform for Parallel Antibody Classification with Mammalian Cell Surface Display

Name: Ning Lin

Supervisor: Assoc. Prof. Tetsuya Kadonosono

Tokyo Institute of Technology

School of Life Science and Technology

Department of Life Science and Technology

# Table of Contents

## List of abbreviations

| | |
|---|---|
| ADCC | Antibody-dependent cell-mediated cytotoxicity |
| ADCP | Antibody-dependent cellular phagocytosis |
| AF488 | Alexa Fluor 488 |
| AF647 | Alexa Fluor 647 |
| AKT | Protein kinase B |
| BFP | Blue fluorescent protein |
| BLI | Biolayer interferometry |
| CDC | Complement-dependent cytotoxicity |
| CDR | Complementarity-determining region |
| CH | Heavy chain constant domain |
| CL | Light chain constant domain |
| CLL | Chronic lymphocytic leukemia |
| Dac | scFv of daclizumab |
| ECD | Extracellular domain |
| ELISA | Enzyme-linked immunosorbent assay |
| Fab | Antigen binding fragment |
| FACS | Fluorescence activated cell sorting |
| Fc | Crystallizable fragment |
| FCM | Flow cytometry |
| FcγR | Fcγ receptor |
| FRs | Framework regions |
| FS | Forward scatter |
| gAb | Guide antibody |
| gDNA | Genomic DNA |
| H6 | 6×His tag |
| HC | Heavy chain |
| HER | Human epidermal growth factor receptor |
| HER2 | Human epidermal growth factor receptor 2 |
| ICD | Intracellular domain |
| Ig | Immunoglobulins |

| | |
|---|---|
| $K_D$ | Dissociation constant |
| L | Linker peptide |
| LC | Light chain |
| MAGPIE | Monoclonal antibody-guided peptide identification and engineering |
| MAPK | Mitogen-activated protein kinase |
| MD | Molecular dynamics |
| MFI | Mean fluorescence intensity |
| MOA | Mechanism of action |
| MOI | Multiplicity of infection |
| NGS | Next generation sequencing |
| Niv | scFv of nivolumab |
| PCR | Polymerase chain reaction |
| Per | scFv of pertuzumab |
| Pert-AF647 | AF647-labeled pertuzumab |
| PI3K | Phosphatidylinositol 3'-kinase |
| qAb | Query antibody |
| rAb | Reference antibody |
| rAb(-) | rAb-non-binding |
| rAb(+) | rAb-binding |
| scFc | Single chain Fc |
| scFv | Single-chain variable fragment |
| scFv(Dac) | Daclizumab scFv |
| scFv(Niv) | Nivolumab scfv |
| scFv(Per) | Pertuzumab scfv |
| scFv(Tra) | Trastuzumab scFv |
| SDS-PAGE | Sodium dodecyl sulfate-polyacrylamide gel electrophoresis |
| sfGFP | Superfolder green fluorescent protein |
| SPR | Surface plasmon resonance |
| ss | Secretion signal peptide |
| TM | Transmembrane domain |
| Tra | scFv of trastuzumab |

| | |
|---|---|
| Tras-AF488 | AF488-labeled trastuzumab |
| Tras-AF647 | AF647-labeled trastuzumab |
| VH | Heavy chain variable domain |
| VL | Light chain variable domain |
| WES | Whole exome sequencing |
| WGS | Whole genome sequencing |

# CHAPTER 1

# General introduction

## 1.1 Antibodies

An antibody, also referred to as immunoglobulin (Ig), is a type of protective macromolecule naturally produced by immune cells to defend against infections caused by pathogens. Antibodies, renowned for their exceptional specificity and high affinity for particular antigens, have emerged as predominant therapeutic agents in the treatment of various diseases[1], such as cancers, autoimmune diseases, and infectious diseases. Human antibodies are classified into five distinct isotypes: IgA, IgD, IgE, IgG, and IgM. Among these, IgG is the most abundant isotype in human serum, and the majority of approved therapeutic antibodies belong to the IgG isotype. The following descriptions of antibody features are specifically focused on the IgG isotype.

### 1.1.1 Structure of IgG antibodies and antibody fragments

An IgG antibody has a molecular weight of approximately 150 kDa, comprising four polypeptides: two identical heavy chains (HCs) and two identical light chains (LCs) (Fig. 1-1-1-1a). Each HC pairs with an LC through disulfide bonds to form a heterodimer, and two identical heterodimers are interconnected through disulfide bonds linking two HCs, resulting in the formation of an intact antibody with a symmetric Y-shaped structure. The variable domain of HC (VH) and variable domain of LC (VL) are located at the distal end of the arm of the 'Y' structure, forming the antigen-binding site commonly called paratope. Each variable domain contains three complementarity-determining regions (CDRs), namely CDR-H1, CDR-H2, and CDR-H3 for HCs, and CDR-L1, CDR-L2, and CDR-L3 for LCs. CDRs are characterized by hypervariable diversity in sequences across different antibodies, thereby dominating the antigen-antibody binding interaction. The CDRs of VH and VL are separated by four conserved framework regions (FRs), which function as a scaffold to support the binding of CDRs to the antigen. The antigen-binding fragment (Fab) comprises the VL and the constant domain of the LC (CL), as well as the VH and a constant domain of the HC (CH, CH1). The crystallizable fragment (Fc) region, formed by the 'Y' stem consisting of CH2 and CH3, serves to stabilize full-size antibodies and interact with Fcγ receptors (FcγR) and complement proteins, facilitating the execution of antibody functions[2].

The antigen specificity of antibodies is determined by the antigen-binding sites which are located at CDRs, and antibody fragments that retain antigen-binding sites can maintain the binding specificity of their parental antibodies. The relatively smaller size and simpler structure of antibody fragments than those of full-length antibodies significantly reduce the production complexity and

manufacturing cost, providing a higher flexibility for engineering and modification to develop novel molecules with improved performance. Single-chain variable fragment (scFv) is an engineered antibody format, composed of a VH and a VL connected by a flexible linker peptide, typically $(G4S)_3$ /$(SG4)_3$[3]. Owing to the structural simplicity and ease of production, antibody fragment formats such as Fab and scFv (Fig. 1-1-1-1b) have found widespread applications in diverse research endeavors. For instance, Fab and scFv are crystallized in complexes with antigenic proteins to investigate the three-dimensional structure of antigen-antibody interactions. Additionally, these formats are displayed on the surface of phage, yeast, and mammalian cells for screening purposes. They also act as antigen recognition domains in chimeric antigen receptor T cell therapy and function as modules in the development of innovative multi-specific antibody formats. The utilization of such antibody fragments broadens the space for the exploration of antibody-based drugs and therapies.

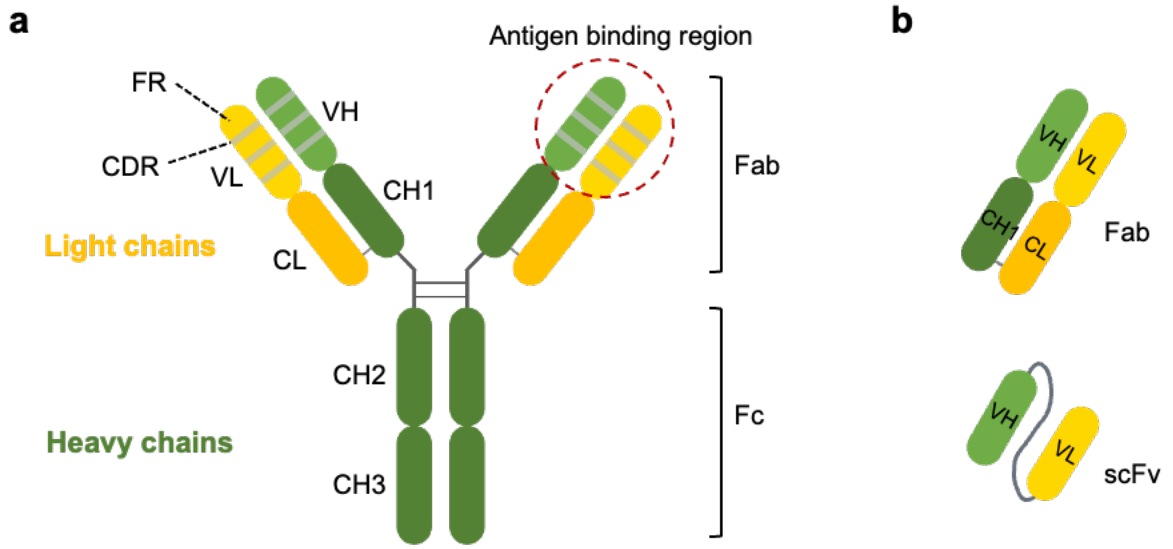



**Figure 1-1-1-1.** Diagram illustration of IgG antibody structure. (**a**) An IgG antibody comprises two identical heavy chains (HCs) (green) and two identical light chains (LCs) (yellow). The Fab region consists of the VH and CH1 of the HCs, as well as the VL and CL of the LCs. Each VH and VL domain is composed of four FRs and three CDRs (grey) that form the antigen binding region. The Fc region consists of the CH2 and CH3 from the HCs. (**b**) Representative antibody fragments. The Fab (top) is composed of the VL, CL, VH, and CH1 domains, and the scFv (bottom) is comprised of a VH and a VL connected by a flexible linker. VH, heavy chain variable domain; VL, light chain variable domain; CH, heavy chain constant domain; CL, light chain constant domain; Fab, antigen-binding fragment; CDRs, complementarity-determining regions; FRs, framework regions; Fc, crystallizable fragment;

scFv, single-chain variable fragment.

### 1.1.2 Mechanism of action (MOA) of antibody drugs

Therapeutic antibodies exert their efficacy through various MOAs, including neutralization, antibody-dependent cell-mediated cytotoxicity (ADCC), antibody-dependent cellular phagocytosis (ADCP), and complement-dependent cytotoxicity (CDC)[4]. Neutralization involves antibodies binding to the target antigen on cell surfaces, thereby blocking its pathological functions. For example, the antibody cetuximab binds to the human epidermal growth factor receptor 1 (HER1) and prevents the ligand epidermal growth factor from binding to the receptor, inhibiting HER1 activation and downstream events such as cell proliferation and migration[5]. The other three MOAs, ADCC, ADCP, and CDC, involve interactions between the Fc domain of antibodies and FcγRs or other proteins. In ADCC, antibodies bound to extracellular membrane antigens trigger the recruitment of immune-effector cells by binding their Fc domain to FcγR on the surface of immune cells. The binding to the FcγRIIIA receptor on natural killer cells promotes the release of lytic factors, leading to target cell destruction[6]. When the Fc domain binds to FcγRI receptors expressed in phagocytic cells (such as macrophages, neutrophils, and eosinophils), it initiates a phagocytosis process called ADCP, resulting in clearance of target cells. Additionally, CDC involves the activation of complement cascade by the antibody-antigen complex, leading to the formation of membrane attack complex and lysis of targeted cells. Antibodies can achieve their therapeutic efficacy through either a single MOA or synergistically involving several MOAs. For instance, daratumumab, an anti-CD38 monoclonal antibody approved for the treatment of multiple myeloma, utilizes multiple MOAs, including ADCC, ADCP, CDC, and modulating CD38 enzyme activities[4].

### 1.1.3 Antibodies exert their functions through binding to specific epitopes on antigens

Antibodies typically recognize and bind to specific and localized regions on the surface of their antigens. These regions on antigens, which interact with the paratope of antibodies, are termed epitopes. Epitopes represent one of the most critical intrinsic features of antibodies, as the therapeutic efficacy of an antibody is intimately correlated with the epitope it targets on the target antigen[7, 8]. Epitopes are broadly categorized into two types: linear epitopes and conformational epitopes. Linear epitopes consist of a continuous sequence of amino acid residues on antigens

(usually 8-15 residues), also called continuous epitopes. Conformational epitopes are composed of discontinuous amino acid residues that are brought together through the protein folding of antigens, also known as discontinuous epitopes. It has been reported that conformational epitopes are the predominant B-cell epitopes[9], and the majority of approved therapeutic antibodies target conformational epitopes.

The vast majority of eukaryotic antigenic proteins are typically composed of hundreds of amino acids, forming complex three-dimensional structures containing multiple domains that serve as the structural and functional units of proteins. Due to the chemical and structure complexity, it is common to find multiple epitopes within a single antigen. Therefore, it is a standard practice to develop different antibodies directly against distinct or partially overlapping epitopes on the same target [10]. Moreover, owing to the diverse functionalities of protein domains in physiological and pathological processes, antibodies targeting distinct epitopes on the same antigen may exert their functions through diverse MOAs, eliciting different antibody responses and leading to varied therapeutic outcomes. For example, therapeutic antibodies such as rituximab and ofatumumab target distinct epitopes on CD20, resulting in varying degrees of efficacy and approval for different indications[10]. Rituximab is approved for the treatment of non-Hodgkin's lymphoma, chronic lymphocytic leukemia (CLL), rheumatoid arthritis, as well as Wegener's granulomatosis and microscopic polyangiitis, whereas ofatumumab is specifically for CLL. Additionally, while rituximab and obinutuzumab both recognize a partially overlapping epitope on CD20, they differ in their interaction with CD20, resulting in different MOAs[11]. Numerous instances demonstrate the close relationship between the targeted epitope and antibody efficacy. Given the potential for cancer cells to develop resistance to specific antibodies, the strategy of simultaneously targeting different functional epitopes has proven effective in overcoming the resistance to single-drug therapy and achieving synergistic therapeutic effects[12]. Therefore, an in-depth understanding of epitopes and underlying molecular mechanisms of antibody efficacy holds the potential to facilitate the development of rational strategies for combination therapy. Furthermore, it paves the way for the development of novel antibodies with complementary mechanisms and functions to existing antibodies against established target antigens.

### 1.1.4 Epitope characterization—epitope mapping

The decisive role of epitopes in determining drug efficacy makes them essential considerations

during drug discovery, vaccine design, and immunotherapy. With the advancement in technologies, such as genetic and protein engineering, developing antibodies specific to the target antigen through screening antibody library has become a manageable task. However, despite the fact that the resulting lead candidates usually exhibit robust binding performance to the target, there remains a potential risk of lacking desired therapeutic effects if their binding is directed against functionally irrelevant epitopes. Therefore, comprehensive evaluations across various aspects are necessary to assess the efficacy and overall suitability of the obtained target antigen-specific binders for therapeutic use. Epitope characterization is a key aspect for comprehending the MOA, elucidating functionality, and evaluating the therapeutic potential of antibody candidates. It facilitates the study of sequence and structural properties of epitopes and reveals epitope coverage among various antibodies targeting the same antigen, thus providing valuable insights into the specificity and potential therapeutic benefits of antibodies.

Epitope mapping identifies antibody binding residues on antigens through various techniques. For linear epitopes, peptide-based approaches (Fig. 1-1-3-1) are commonly employed, involving the utilization of short overlapping peptide fragments that cover the entire antigen proteins. These peptides are presented either as synthetic peptide microarrays[13, 14], or displayed on the surfaces of yeast[15] or phage[16], followed by testing for their binding to the antibody of interest to identify those with strong reactivity. The identified peptides are then aligned with the original protein sequence, and the amino acid sequence within the overlapping regions with strong binding are considered as potential epitopes. Conformational epitope determination poses a greater challenge due to their reliance on the complete tertiary structure of antigens. While peptide-based methodologies can be used for conformational epitope identification, the linear stretches of peptides are inadequate for presenting the native three-dimensional arrangement of amino acids within a conformational epitope. This limitation may potentially result in the loss of critical information about binding residues and lead to false positive results. Techniques such as X-ray crystallography[17], nuclear magnetic resonance[18], hydrogen/deuterium exchange[19], and cryo-electron microscopy[20] are better suited for determining conformational epitopes. However, their application is restricted by the demanding requirements of sample preparation, high costs, and limited throughput capacity. These resource-intensive technologies are usually not prioritized in the early stages of drug discovery, where a large number of candidates need to be evaluated. Instead, they are primarily utilized in the later stages of antibody development when the most

promising candidates have emerged. What's more, although site-directed or comprehensive mutagenesis scanning combined with surface-display technologies and next-generation sequencing (NGS) provides a high-throughput approach for epitope mapping[21-23], constructing a comprehensive mutant library covering the entire antigen is challenging, especially for large and complex proteins. Collectively, although epitope mapping is able to provide residue-level resolution epitope information, this advantage comes at the expense of time consumption and labor intensity.

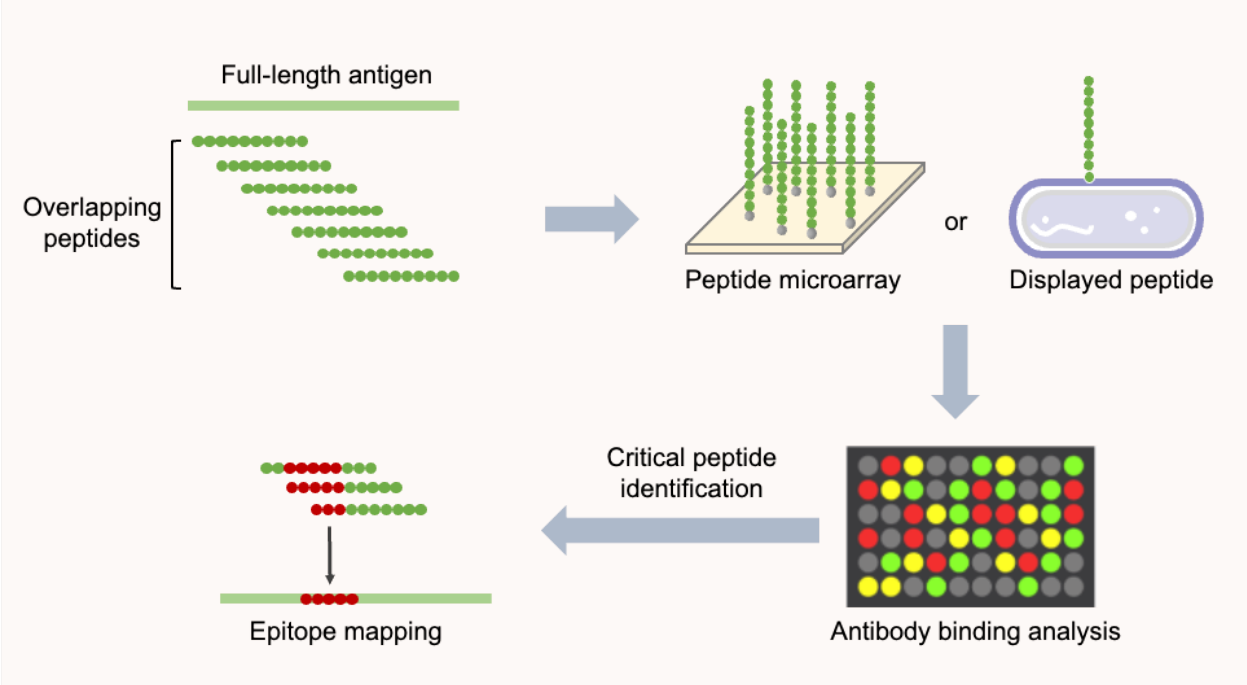



**Figure 1-1-3-1.** Peptide-based epitope mapping. Overlapping peptides covering the entire antigen sequence are presented on a peptide microarray or the surface of organisms, followed by antibody binding analysis to identify peptides responsible for antibody-antigen binding. Overlapping residues on these identified peptides are viewed as potential epitopes, which are then mapped onto the target antigen.

### 1.1.5 Epitope characterization—epitope binning

Given the ever-expanding number of antibodies under development for diverse targets, it becomes imperative to develop a cost-effective and efficient technique for epitope characterization of massive candidate antibodies, particularly in the initial stages of antibody discovery. Epitope binning emerges as a favored alternative to epitope mapping, involving profiling a collection of antibodies and categorizing them into distinct groups or bins based on epitope similarities[24].

Antibodies targeting similar epitopes often exhibit similar functional characteristics due to the corresponding relation between epitopes and functionality[7, 8]. Even though the exact antibody binding residues are not fully revealed, epitope binning can guide the rational selection of candidates for comprehensive characterization, especially assisting in the identification of those with overlapping functional epitopes alongside validated antibodies and exhibiting superior properties. This significantly improves productivity and reduces the occurrence of 'dead-end' candidates during therapeutic antibody discovery.

Various computational approaches have been devised to predict antibodies with similar epitopes, and these methods are applicable to antibodies with both similar and dissimilar sequences[25-28]. Nevertheless, their practical application is greatly restricted by limited accuracy and the lack of consideration for epitope variability under physical conditions. Experimental epitope binning (Fig. 1-1-4-1) utilizes pairwise competitive binding assays to assess whether the paired query antibody (qAb) and reference antibody (rAb) target overlapping or different epitopes on the same target antigen[24, 29]. This is accomplished through the detection of competitive inhibition, wherein the binding of the qAbs to antigens is blocked by the presence of rAbs. There are various competitive immunoassay formats, including classical sandwich, premix, or in-tandem assays[24]. In a sandwich assay, the binding of a qAb to the antigen is tested after the antigen has been initially captured by an rAb. In a premix assay, the antigen is mixed beforehand with the qAb at a saturating concentration, and then the binding of the immobilized rAb to the qAb-antigen complex is tested. In an in-tandem assay, the rAb and qAb are sequentially applied to an immobilized antigen to test whether the rAb inhibits the binding of qAb to the antigen. These competitive binding formats can be used in combination with the enzyme-linked immunosorbent assay (ELISA)[30], biolayer interferometry (BLI), or surface plasmon resonance (SPR)[31-34]. The continuous progress in biosensor technology has contributed to cost reduction, enabling high-throughput epitope binning using BLI or SPR as fundamental methods for epitope characterization of early antibody candidates. High-throughput epitope binning can also be achieved by integrating the competition strategy with the Luminex multiplex technique[35] or cellular barcoding technique[36].

Although various methodologies are available and additional strategies are being proposed to facilitate the effectiveness and simplicity of epitope binning, to the best of my knowledge, all current approaches require the individual production and even purification of antibodies, thereby

limiting their utility for large-scale evaluations. Therefore, to evaluate the epitope similarities for multiple antibodies more efficiently, it is necessary to develop an epitope binning approach that eliminates the requirement for individually produced antibodies.

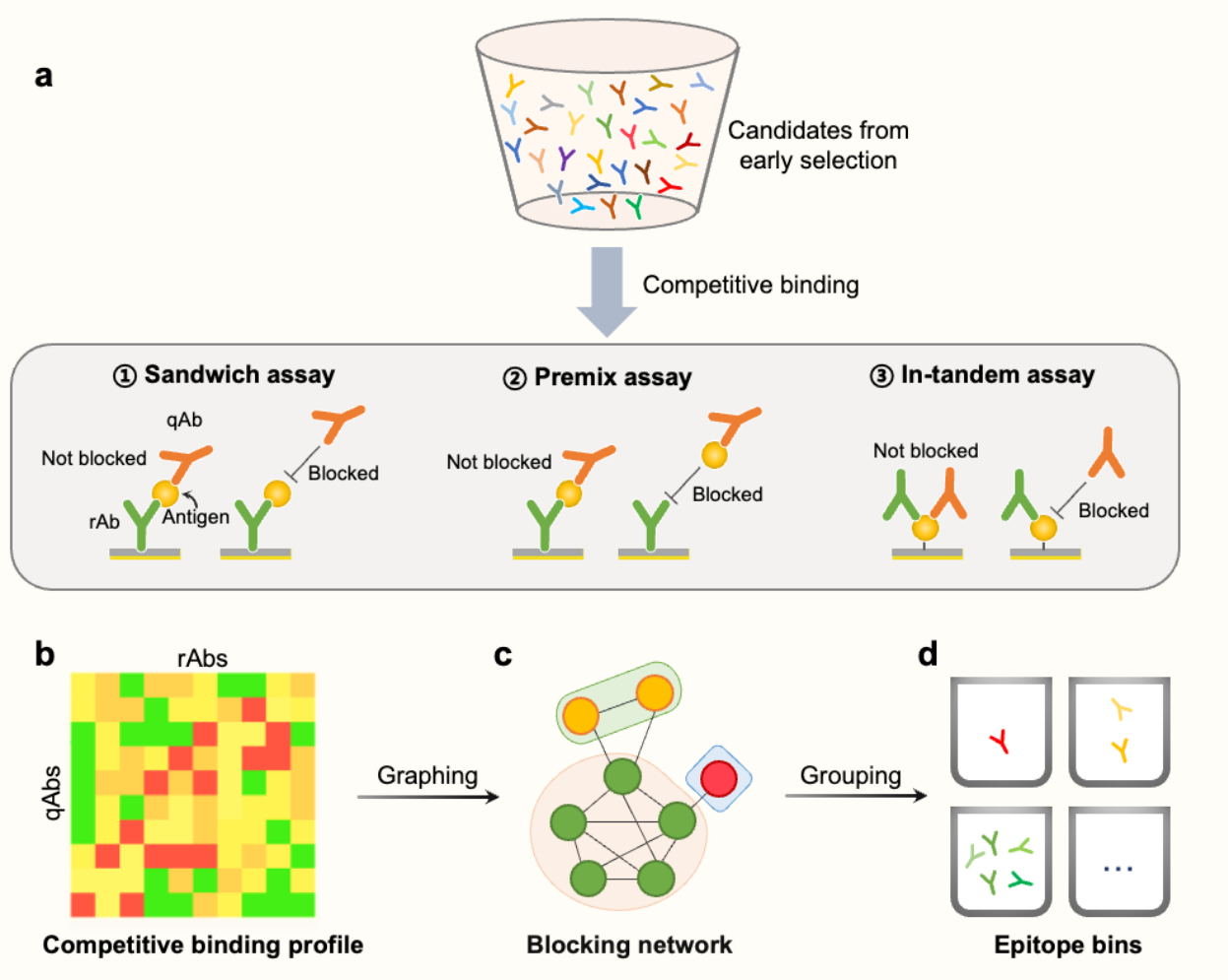



**Figure 1-1-4-1.** Schematic representation of epitope binning. (**a**) Epitope binning analysis of antibody candidates from early-stage screening using pairwise competitive binding assays. The carton in the rounded rectangle shows three formats of competitive binding assays, including sandwich assay where the qAb is tested for the binding to the antigen that is first captured by an rAb, premix assay where the premixed solution of qAb and antigen is tested for the binding to the mobilized rAb, and in-tandem assay where rAb and qAb are sequentially applied to the mobilized antigen and the binding signal of qAb is tested. (**b**) Diagram of heatmap summarizing the epitope binning results. (**c**) Diagram of blocking network plot where chords represent the blocking relationships, and the colored envelopes represent the bins. (**d**) Epitope bins representing antibody groups targeting distinct epitopes of an antigen.

## 1.2 Mammalian cell display

Mammalian cell display is an *in vitro* technique used to present protein molecules, such as antibodies and peptides, on the surface of mammalian cells. This enables direct interactions between the surface-displayed proteins and other exogenous molecules, making it a valuable tool for studying protein-protein interactions. The surface display of the protein of interest has been achieved by fusing the gene encoding the target to a transmembrane domain(TM) of a transmembrane protein[37]. Various proteins, such as human platelet-derived growth factor receptor[38], FasL[39], CD8 and CD28[40], and PD-L1[41], have been reported as suitable sources of TM for display. Mammalian cell display systems can be divided into transient expression systems, in which the introduced expression vector is transcribed as mRNA for rapid and short-term expression, and stable expression systems, where the genes encoding the target protein are integrated into the host cell genome, resulting in stable and long-lasting expression[42].

In comparison to other display techniques such as phage display, yeast display, and bacteria display, the most attractive advantage of mammalian cell display is its ability to enable proper post-translational modification and protein folding[43, 44], allowing the surface-displayed protein to be presented in its native conformation. This favorable property makes mammalian cell display a preferred expression tool for investigating transmembrane proteins with complex structures and dynamic conformational changes, such as G protein-coupled receptor[45] and ion channel proteins. It ensures that these proteins undergo proper folding and modification, providing a natural environment conducive to exhibiting their physiological activity.

### 1.2.1 Mammalian cell display facilitates antibody development

Mammalian cell display has proven to be a versatile tool in antibody development, applicable to both antibody discovery and antibody characterization. In the whole-cell panning approach, the target antigen is displayed on mammalian cell surfaces and interacts with the phage- or yeast-displayed antibody libraries to analyze antigen-antibody interactions and identify antigen-specific antibodies[45, 46]. This approach avoids the need for purified proteins as targets, thereby enabling the development of antibodies against complex membrane proteins with difficulty in recombinant expression. Mammalian cells displaying the target antigen also serve as an antigen supplier for determining antibody affinity, where the binding of soluble antibodies to the cell surface antigen is quantified[47], promoting comprehensive antibody characterization. Additionally, mammalian

cell display has found widespread applications in presenting antibody libraries for screening against a variety of targets[48-50]. Moreover, because somatic hypermutation, the natural mechanism of antibody maturation, can be introduced and reproduced within the mammalian cells used for display, mammalian cell display can be applied for the selection of affinity-matured antibodies[51].

The process of screening antibodies or antibody fragments (hereafter collectively referred to as antibodies) using mammalian cell display (Fig. 1-2-1-1) generally begins with the generation of an antibody DNA library that may derived from antigen-immunized animals, convalescent patients, or computational design. The pooled gene fragments are then incorporated into a display vector that contains a linker and a TM, with a reporter gene often incorporated for expression visualization. In the construction of an antibody cell library, both transient and stable expression systems can be adopted. The transient system is suitable for a single round of selection before recovering antibody genes, whereas the stable system allows for multiple rounds of selection to enrich cells presenting the antibodies of interest[37], facilitating the identification of low-abundance species. For constructing cells that stably express the antibody library, a lentivirus library produced from the plasmid DNA library is used to infect mammalian cells, leading to the expression of the antibody library on cell surfaces. The cell library is then incubated with a labeled target protein, and when combined with the fluorescence-activated cell sorting (FACS), cells displaying antibodies with desired properties are isolated. After several rounds of selection, the isolated cells are subjected to NGS analysis to reveal antibody sequences. Following recombinant expression and further evaluation of selected candidates, the target-binding candidates can be identified.

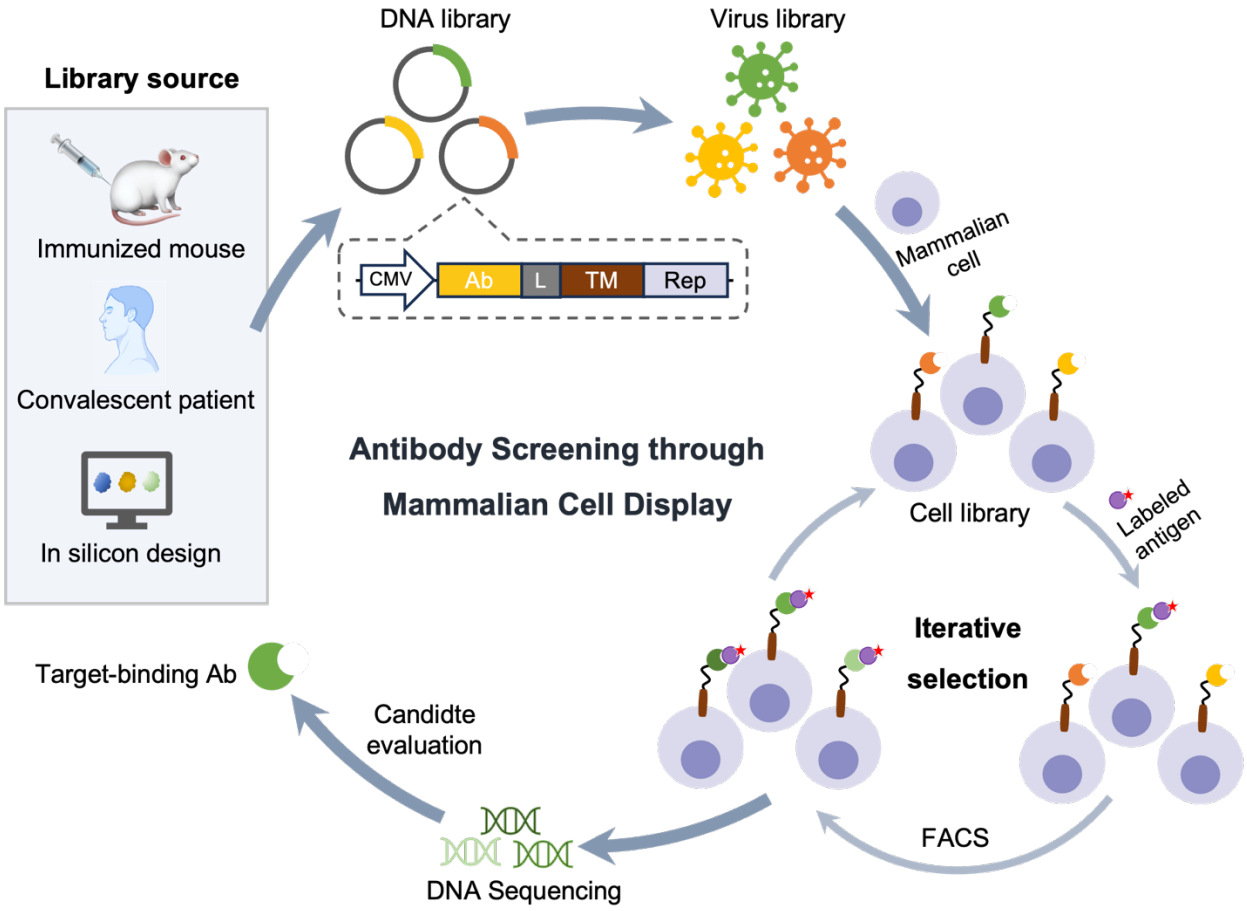



**Figure 1-2-1-1.** A general workflow for antibody screening via mammalian cell display. Antibody genes sourced from immunized mice, convalescent patients, or computational design are cloned into a display vector with a linker, a TM, and a reporter gene, generating a DNA library. This library is then utilized to produce a virus library, which is subsequently transduced into mammalian cells for the surface display. Then, several rounds of selection against the cell library using a labeled antigen are performed, followed by DNA sequencing analysis of sorted cells. Finally, the evaluation of top candidates leads to the identification of high-performance antibodies. Ab, antibody; L, linker peptide; FACS, fluorescence-activated cell sorting.

### 1.2.2 Mammalian cell co-display of antigen and antibody

Taking advantage of mammalian cell display, co-expressing both the antigen and antibody on the same mammalian cell allows the proper presentation of both molecules in their native conformation and physiological activity. This approach mimics the conditions under which antibodies naturally interact with antigens, which is beneficial for studying antibody-antigen

interactions and identifying promising antibody candidates. Using mammalian cells that co-express surface-displayed antigen and a secreted antibody library, Robertson, N. et al. implemented a FACS-based Forster resonance energy transfer assay to isolate cells self-labeled by secreted scFv and identify antigen-specific binders[52]. In addition, an affinity maturation screening system involving the mammalian cells displaying an antigen/death-signaling receptor chimera and a competing scFv and secreting an scFv library was designed, wherein the binding of the competing scFv to the antigen triggered cell-death signaling and induced cell death[53]. A secreted affinity-matured scFv mutant sharing the same epitope inhibits the binding of the competing scFv to the antigen, thereby preventing cell death and allowing the selection of affinity-matured epitope-specific antibodies. This example suggests the potential application of the competitive binding context provided by the co-expression system in epitope-related studies. However, these systems have a potential concern of 'cross-labeling', wherein cells may be labeled by antibodies secreted from nearby cells, resulting in false positive clones. Although integrating the antigen display/antibody secretion system with a droplet-based microfluidic technique ensures the autocrine binding of antibodies[54, 55], the complexity and high cost of microfluidic systems may pose barriers for many researchers.

See, K. et al. previously developed a screening system, monoclonal antibody-guided peptide identification and engineering (MAGPIE) (Fig. 1-2-2-1), for identifying antigen-binding peptides[41]. MAGPIE involves displaying both the target antigen and the peptide library on the surface of mammalian cells, thus eliminating the concern of 'cross-labeling'. In this system, both the library design and screening for antigen-binding peptides are guided by a validated antigen-binding antibody, called the guide antibody (gAb). The cell library is then incubated with the labeled gAb, followed by isolating low signal fraction cells for NGS analysis to identify peptides with high read count. Because the identification of candidate peptides binding to the antigen on the cell surface relies on competitive binding with the gAb, it is expected that the epitope of the identified antigen-binding peptides should be similar to that of the gAb, showing the potential of MAGPIE for evaluating epitope similarity between cell surface-displayed peptide and the gAb.

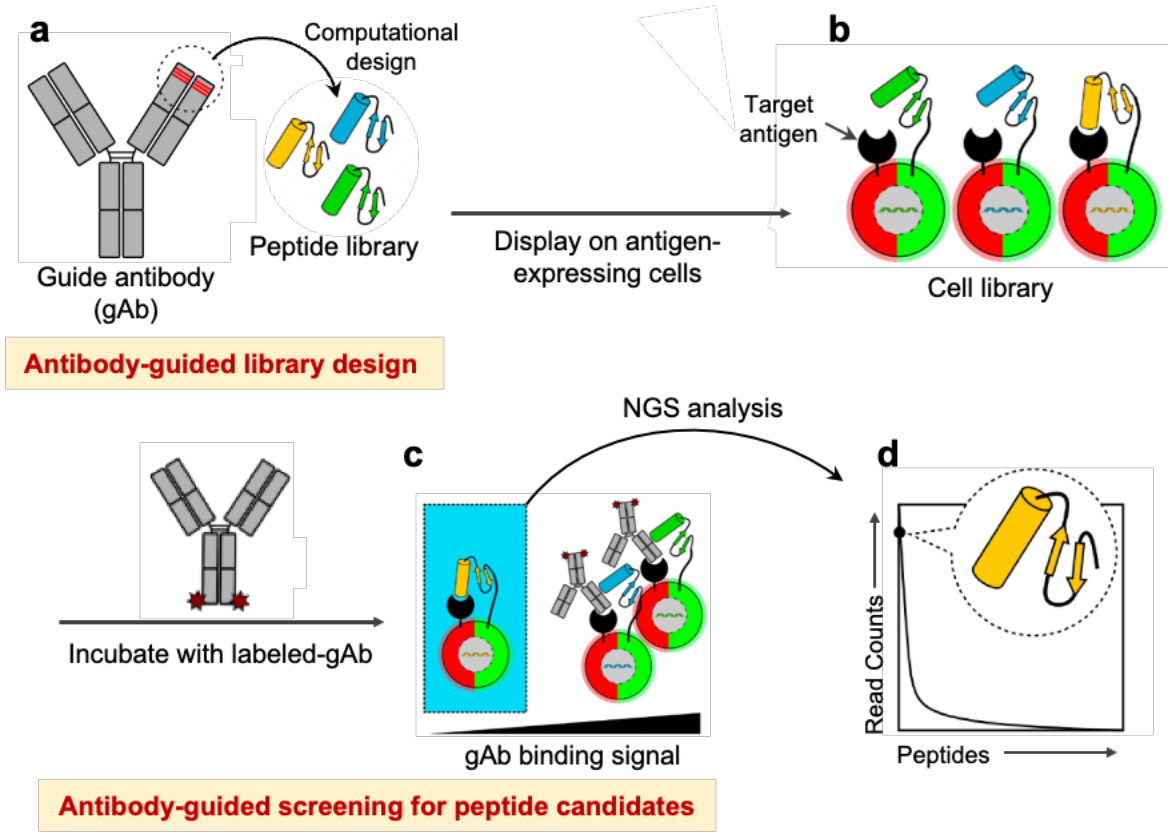



**Figure 1-2-2-1.** Overview of MAGPIE system. (**a**) Computational design of peptide library by grafting binding residues from a validated antibody onto a scaffold protein to construct a peptide library. (**b**) Generation of cells displaying both the target and the peptide library on their surface. (**c**) Isolation of cells with low gAb binding signal after incubation with labeled gAb. (**d**) Identification of peptide candidates through NGS analysis. Cited from See, Kyra et al.[41] and modified in part.

## 1.3 Research purposes of this study

The considerable success of antibody drugs has attracted increasing attentions from researchers and pharmaceutical companies dedicated to developing novel antibody drugs for addressing various refractory diseases. However, the antibody drug discovery process is resource-intensive and time-consuming, necessitating cost-effective, reliable, and efficient strategies to improve productivity and reduce the number of dead-end candidates in the early stage. Epitope binning provides investigators with general epitope information and preliminary insights into the functionality information of antibody candidates in the early development. This process effectively facilitates the elimination of ineffective hits targeting non-functional epitopes, mitigating the allocation of resources towards antibodies with limited functional potential. However, as

mentioned earlier, existing epitope binning methodologies require individually produced antibodies for evaluation, often necessitating purification to ensure accuracy. Besides, each rAb-qAb pair is typically tested separately. These limitations place a significant burden on antibody preparation and analysis time, especially during large-scale evaluations.

To address these shortages, I aim to develop an epitope binning platform that eliminates the need for individual antibody production and enables simultaneous evaluation of multiple antibodies in a single-tube assay. MAGPIE identified target-binding peptides based on the competitive binding between the displayed peptides and exogenous gAb, which is a consistent concept for experimental epitope binning. However, epitope similarities between the identified peptides and gAb have not been verified. This has inspired me to investigate whether the antigen/antibody co-displaying system can be used to assess epitope similarity and developed as a tool for parallel epitope binning for multiple antibody pairs. Through this study, I hope to propose a facile and efficient methodology for epitope characterization, which can be hopefully applied to the practical identification of therapeutic antibody drugs.

# CHAPTER 2

# Establishment of an antigen/antibody co-displaying system for epitope similarity evaluation

## 2.1 Overview

In this chapter, I attempted to extend MAGPIE as a platform for epitope similarity evaluation and investigated the feasibility of this system for this new utility. The peptide library in MAGPIE was replaced with qAbs, and the gAb was repurposed to serve as an rAb. As a proof of concept, the human epidermal growth factor receptor 2 (HER2) was utilized as the model antigen, and two well-studied anti-HER2 monoclonal antibodies, pertuzumab and trastuzumab, were selected as rAbs. To comprehensively assess the ability of this system for epitope similarity evaluation, the included qAbs either targeted the same or different epitopes as the rAb or did not bind to HER2 at all. Mammalian cells displaying both HER2 and qAbs were constructed through lentiviral transduction and flow cytometry (FCM) analysis using a fluorescently labeled rAb revealed the specific binding inhibition of the rAb by the surface-displayed qAb with the same epitope rather than those with different epitope or non-binding qAb. This inhibition was evidenced by the presence of a population of cells that did not bind with the rAb. Furthermore, when dual different rAbs were used simultaneously, epitope similarity evaluation could also be achieved, even in situations where several rAb-qAb pairs co-existed. Overall, the findings in this chapter demonstrate the applicability of this system for evaluating epitope similarity between qAbs and rAbs.

## 2.2 Introduction

### 2.2.1 HER2 as a biomarker for cancers

HER2, also known as ErBb2, is a transmembrane receptor tyrosine kinase and a member of the human epidermal growth factor receptor (HER) family, alongside HER1, HER3, and HER4. HER family members share a similar structure (Fig. 2-2-1-1a), all consisting of an extracellular domain (ECD), a single transmembrane segment, and a cytoplasmic domain containing tyrosine kinase region and carboxy-terminal tail. The ECD is subdivided into four segments: ECD I, ECD II, ECD III, and ECD IV, with the dimerization arm located at the ECD II and the ligand binding domains situated at ECD I and ECD III. In contrast to other HER members, HER2 exhibited some unique features. HER receptors are usually activated through ligand engagement with the ligand-binding domain, followed by the conformation change in the ECD and subsequent homo- or heterodimerization among HER family members. However, the activation of HER2 is ligand-

independent, and there is no known ligand for HER2. In addition, unlike other HER members, whose dimerization arms are hidden through the intramolecular interaction between ECD II and ECD IV in an inactivated state[56, 57], HER2 maintains a constitutively activated status, with its dimerization arms remaining exposed[58]. These unique properties of HER2 enable it to be a preferred dimerization partner for other members. The receptor dimerization brings the intracellular domain of each member into proximity and induces the phosphorylation of the cytoplasmic domain, leading to the recruitment of downstream effector proteins and triggering the relevant signaling pathways, such as the phosphatidylinositol 3′-kinase (PI3K)-protein kinase B (AKT) pathway and extracellular signal-regulated kinase/mitogen-activated protein kinase (ERK/MAPK) pathway (Fig. 2-2-1-1b). As a result, the transcription of genes involved in cell survival, cell-cycle progression, and cell proliferation is activated, which drives tumorigenesis.

It has been reported that heterodimers containing HER2 elicit stronger signal potency, with the HER2-HER3 heterodimer exhibiting the most potent activation strength of the signaling pathway, particularly the PI3K/AKT pathway[59]. As a potent driver of signal pathways that result in the activation of a variety of genes promoting cell proliferation and differentiation, HER2 plays a positive role in tumorigenesis. As reported, HER2 gene amplification and HER2 overexpression are frequently observed in various cancers, such as breast cancer[60], gastric cancer[61], esophageal cancer[62], ovarian cancer[63], and endometrial cancer[64], and are associated with poor diagnosis. Therefore, HER2 serves as a biomarker and target molecule for cancer diagnosis and treatment.

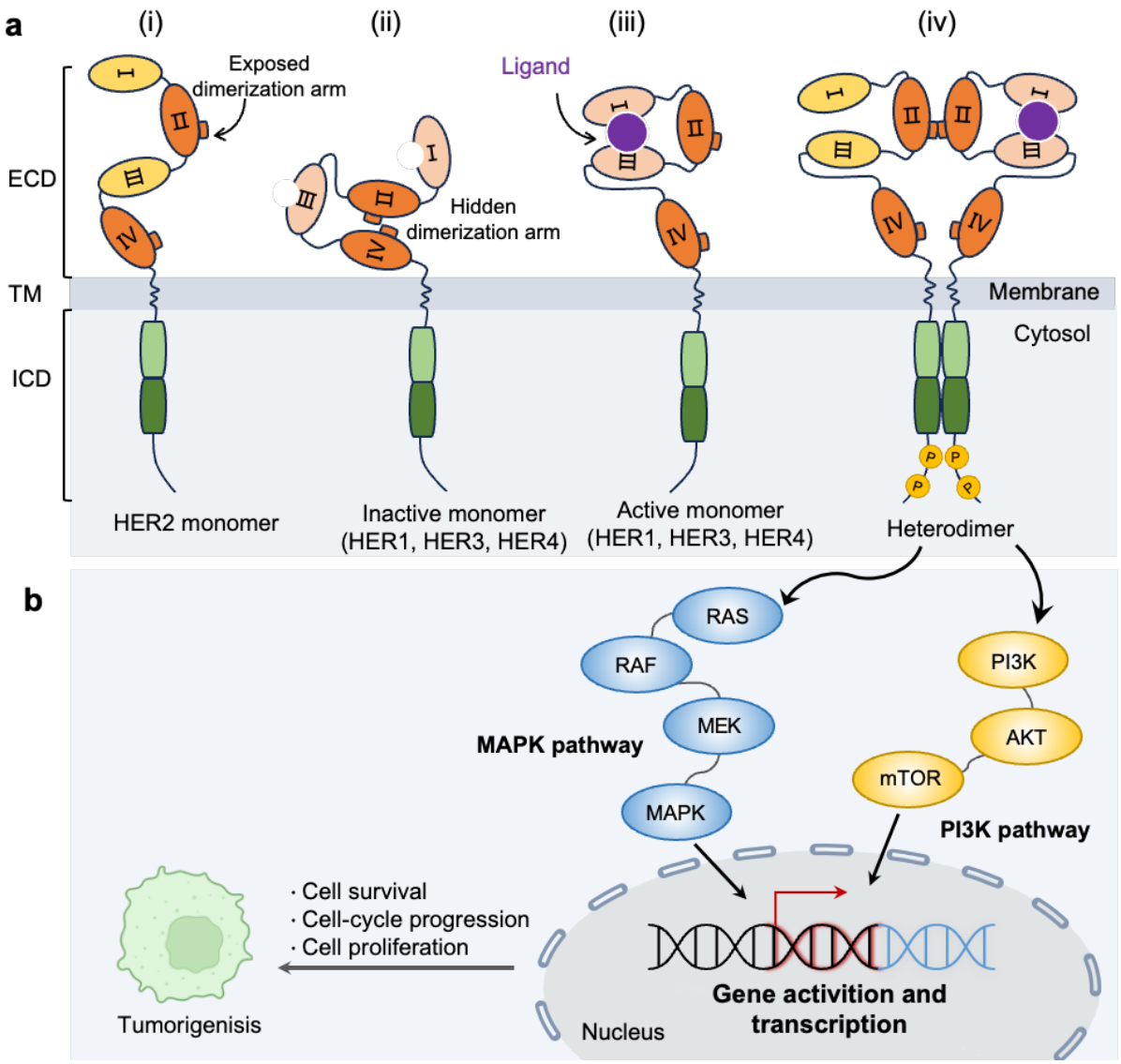



**Figure 2-2-1-1.** Schematic diagram of HER family structure and signaling pathways. (**a**) Overall structure of HER family members. A HER molecule comprises an extracellular domain (ECD), a transmembrane domain (TM), and an intracellular domain (ICD). The ECD consists of four subdomains numbered ECD I, II, III, and IV, with the dimerization arm located at ECD II and the ligand binding domains located at ECD I and III. HER2 (i) exists in a constitutively activated state, featuring the exposed dimerization arm, while other members (ii) remain in an inactive state, with the hidden dimerization arm formed by the interaction of ECD II and IV. Ligand binding triggers receptor activation, leading to the conformation rearrangement of ECDs and the exposure of the dimerization arm within ECD II (iii). The interaction between the dimerization arms within two active receptors brings the ICDs close and results in the phosphorylation of the C-terminal tail (iv). Scheme adapted from Wieduwilt, M. J. et al.,2008.[65]. (**b**) Major signaling pathways triggered by

HER receptors heterodimerization. The dimerization of HER receptors initiates various signaling pathways. The PI3K and MAPK are the two major pathways that play a pivotal role in activating genes that promote cell survival, cell-cycle progression, and cell proliferation, consequently contributing to tumorigenesis. PI3K, phosphatidyl inositol 3-kinase; MAPK, mitogen-activated protein kinase.

### 2.2.2 Anti-HER2 antibodies

HER2-targeted therapy offers a promising approach for the treatment of HER2-positive cancers. Anti-HER2 therapeutic reagents can be broadly classified into small molecule tyrosine kinase inhibitors (such as laptinib, neratinib, and afatinib) and monoclonal antibody drugs (such as trastuzumab, pertuzumab, and antibody-drug conjugate ado-trastuzumab emtansine)[59]. Generally, monoclonal antibodies exhibit higher specificity to the target compared to small molecule inhibitors.

Trastuzumab is a humanized IgG1 antibody against the ECD IV of HER2[66, 67] (Fig. 2-2-2-1) with a moderate affinity with an equilibrium dissociation constant ($K_D$) of approximately 5 nM[10]. It is approved by the US Food and Drug Administration for the treatment of HER2-positive breast, gastroesophageal, and gastric cancers. Although numerous studies have been carried out to investigate how trastuzumab exerts its anti-tumor effects, the exact and detailed MOA remains controversial and requires further elucidation. There are multiple proposed molecular mechanisms for trastuzumab[68, 69], including inhibition of ligand-independent HER2-HER3 heterodimerization[70], suppression of HER3 phosphorylation, induction of the HER2 endocytosis and degradation[71], inhibition of the PI3K/AKT signaling pathway[72], activation of ADCC process[73], and so on.

Pertuzumab is a recombinant humanized IgG1 antibody targeting the ECD II of HER2[67, 74] (Fig. 2-2-2-1) with a $K_D$ of approximately 1.8 nM[75]. It is approved for HER2-positive metastatic breast cancer in combination with other antitumor agents. Pertuzumab is a dimerization inhibitor that targets the dimerization arm of HER2[76], preventing the ligand-dependent heterodimerization of HER2 with other HER receptors and blocking ligand-stimulated signaling via the MAPK and PI3K pathways[77]. Additionally, pertuzumab mediates the ADCC to exert its therapeutic effect[78].

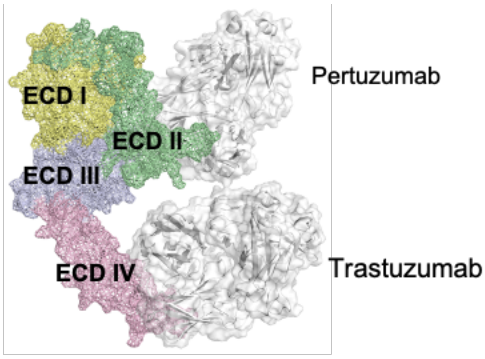



**Figure 2-2-2-1.** Binding structure model of pertuzumab and trastuzumab to the ECD of HER2. Pertuzumab binds to the extracellular domain (ECD) II, while trastuzumab binds to ECD IV of HER2. The model was created using data from PDB entries 1N8Z and 1S78.

### 2.2.3 Research purpose of this chapter

In this chapter, the objective was to investigate the feasibility of the antigen/antibody co-display system for epitope similarity evaluation by using HER2 and two anti-HER2 monoclonal antibodies as model antigen and rAbs, respectively. I aimed to explore the feasibility of evaluating epitope similarity between the surface-displayed qAb and rAb under conditions involving either a single rAb or dual rAbs.

## 2.3 Material and methods

### 2.3.1 Molecular dynamics (MD) simulations

All MD simulations were conducted utilizing the AMBER 16 program package[79] on TSUBAME (Global Scientific Information and Computing Center at Tokyo Institute of Technology) as previously described[41]. The AMBER ff14SB force field was applied for proteins, while the TIP3P model was employed for water molecules. The initial coordinates of the pertuzumab Fab-HER2 and trastuzumab Fab-HER2 complexes were retrieved from PDB accession codes 1S78 and 1N8Z, respectively. To calculate binding free energy, the systems were fully solvated with explicit solvent, and two Na+ counterions were introduced to maintain electrostatic neutrality. The systems were optimized through energy minimization and equilibrated with backbone restraints. A 100 ns production run was then performed, of which the final 10 ns was utilized by the molecular mechanics/generalized Born surface area (MM/GBSA) module to calculate the binding free energy.

### 2.3.2 Plasmid construction

All recombinant DNA experiments were conducted with the approval of the Recombinant DNA Experimental Safety Management Committee of Tokyo Institute of Technology (no. I2021017).

To express HER2, a cDNA encoding human HER2, a T2A self-cleaving peptide (EGRGSLLTCGDVEENPGP), and mCherry was inserted into the multiple cloning site of the CSII-CMV-MCS plasmid (RIKEN Bio-Resource Center, Ibaraki, Japan), resulting in the plasmid pCSII/HER2-mCherry.

To display scFv on cell surfaces, the DNA fragment of scFv(Dac) in the previously constructed plasmid pCSII/scFv-sfGFP[41] was replaced with scFv(Per) or scFv(Tra). The resultant plasmids expressed the superfolder green fluorescent protein (sfGFP) together with scFv-fused proteins. To express a blue fluorescent protein (BFP) together with scFv-fused proteins, the DNA fragment of sfGFP in the pCSII/scFv-sfGFP was replaced with that of BFP through Gibson Assembly using a Gibson Assembly Master Mix kit (New England Biolab, Ipswich, Massachusetts, USA), resulting in the construction of pCSII/scFv-BFP plasmids.

### 2.3.3 Cell culture

The human chronic myelogenous leukemia cell line K562 and a subclone of the human embryonic kidney HEK293 cell line Lenti-X 293T were purchased from JCRB Cell Bank (Osaka, Japan), and Clontech (Mountain View, CA, USA), respectively. The K562/HER2, K562/HER2/scFv, and K562/HER2/scFv$^{B}$ cell lines were established through lentiviral transduction of K562 and K562/HER2 cells with the plasmids pCSII/HER2-mCherry, pCSII/scFv-sfGFP, and pCSII/scFv-BFP, respectively.

K562, K562/HER2, K562/HER2/scFv, and K562/HER2/scFv$^{B}$ cells, as well as Lenti-X 293T cells, were cultured in 10% FBS-RPMI-1640 (Invitrogen, Waltham, MA, USA) and 10% FBS-DMEM (Nacalai Tesque, Kyoto, Japan), respectively. All media were supplemented with penicillin (100 U/mL) and streptomycin (100 μg/mL) (Nacalai Tesque). Cells were maintained in a 37°C incubator with 5% $CO_2$ and subject to regular checks for mycoplasma contamination using a mycoplasma test kit (Lonza, Basel, Switzerland).

### 2.3.4 Lentivirus production and transduction

For lentivirus production, plasmid pCSII/HER2-mCherry, pCSII/scFv-sfGFP, or pCSII/scFv-BFP was mixed with Lentiviral High Titer Packaging Mix (Takara Bio, Shiga, Japan) and

transfected into Lenti-X 293T cells using Lipofectamine LTX (Invitrogen). After 24 h of incubation, the medium was replaced with 10% FBS-DMEM containing 10 μM forskolin (Fujifilm Wako Pure Chemical, Osaka, Japan). At 48 h post-transfection, the lentiviruses in the medium were collected and concentrated using a Lenti-X Concentrator (Clontech) in accordance with the manufacturer's instructions.

To express HER2 and scFv on the cell surfaces, K562 cells were transduced with lentiviruses carrying genes from pCSII/HER2-mCherry, resulting in K562/HER2 cells. Subsequently, K562/HER2 cells were transduced with lentiviruses carrying genes from pCSII/scFv-sfGFP, or pCSII/scFv-BFP, resulting in K562/HER2/scFv, or K562/HER2/scFv$^{B}$ cells. Lentiviral transduction was performed using a modified spinoculation protocol[80, 81] at a multiplicity of infection (MOI) of 0.3 to maximize the cell population with a single copy integration of the target gene into the genome. Specifically, the cells were mixed with the lentivirus and 8 μg/mL polybrene (Nacalai Tesque) in 1 mL of culture medium, followed by centrifugation at 800 ×*g* for 90 min at room temperature. The resulting cell pellet was resuspended in fresh medium and transferred to a culture dish. After incubation for 48 h, the transduction efficiency was confirmed by observing the cells under a confocal microscope Zeiss LSM780 (Zeiss, Jena, Germany) with the appropriate filters (Ex/Em = 545 ± 25 nm/605 ± 70 nm for mCherry, Ex/Em = 470 ± 40 nm/520 ± 50 nm for sfGFP, and Ex/Em = 365 nm/445 ± 50 nm for BFP).

### 2.3.5 rAb labeling

To prepare rAbs conjugated with fluorescent dyes, pertuzumab (MedChemExpress, Monmouth Junction, NJ, USA), and trastuzumab (Chugai pharmaceutical, Tokyo, Japan) were labeled using the Alexa Fluor 647(AF647) Kit or Alexa Fluor 488 (AF488) Antibody Labeling Kit (Invitrogen) according to the manufacturer's instruction. Specifically, 100 μg of each antibody (1 mg/mL in PBS) was mixed with a one-tenth volume of freshly dissolved 1 M sodium bicarbonate buffer, followed by incubating the antibody solution with the provided reactive dye for 1 h at room temperature with gentle inversion on a rotator. The reaction mixture was then purified using purification resins and a column to remove free dye. The absorbance of the labeled antibody was measured using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) in the UV-Vis mode, and the concentration was calculated using the following formula:

$$\text{Protein concentration (M)} = [A280 - (A650 \times 0.03)] \times \text{dilution factor}/203{,}000$$

where 203,000 is the molar extinction coefficient (ε) in $cm^{-1}M^{-1}$ of a typical IgG at 280 nm, and 0.03 is a correction factor for the fluorophore's contribution to the absorbance at 280 nm.

### 2.3.6. FCM analyses

To confirm the cell surface localization of HER2 and scFv-fused proteins, $2\times10^5$ K562, K562/HER2, and K562/HER2/scFv cells were incubated with AF488-conjugated anti-human HER2 antibody (clone 24D2, 1:100; BioLegend, San Diego, CA, USA) and AF647-conjugated anti-His tag antibody (0.5 μg/mL; MBL Life Science, Tokyo, Japan), respectively, for 1 h at 4°C prior to FCM analysis.

To evaluate epitope similarity, $2\times10^5$ K562/HER2/scFv cells were incubated with 1 nM of AF647-labeled pertuzumab (Pert-AF647) and AF647-labeled trastuzumab (Tras-AF647) for 1 h at 4°C prior to FCM analysis using an iCyt ec800 flow cytometer (Sony Biotechnology, Tokyo, Japan). To investigate the feasibility of combinational use of two different rAbs, $2\times10^5$ K562/HER2, K562/HER2/$scFv^B$, and cell mixtures consisting of indicated K562/HER2/$scFv^B$ cells were incubated with rAb solution containing 1 nM Pert-AF647 and AF488-labeled trastuzumab (Tras-AF488) for 1 h at 4°C prior to FCM analysis using a FACSCanto II (BD Biosciences, Franklin Lakes, NJ, USA).

In all FCM experiments, antibody dilution as well as cell washing were performed using FACS buffer (PBS containing 3% FBS). Data acquisition was carried out using an iCyt ec800 flow cytometer (Sony Biotechnology) unless otherwise stated, and the flow cytometric data were analyzed using FlowJo software ver.10.8.1 (FlowJo, LLC, Ashland, OR, USA).

### 2.3.7 Statistics analysis

All experimental data were obtained from three independent experiments and are expressed as the mean ± SEM. Statistical analysis was performed using GraphPad Prism 10 (GraphPad Software, Boston, MA, USA). *P*-values were determined using Tukey's multiple comparisons test or two-side unpaired Student's *t*-tests, as specified in the figure legends. $P < 0.05$ was considered statistically significant.

## 2.4 Results

### 2.4.1 Strategy for epitope similarity evaluation

To develop an effective epitope binning methodology independent of antibody production and

purification, I repurposed MAGPIE as a system for epitope similarity evaluation. In this approach, a qAb is displayed in a scFv format on antigen-expressing cells through gene transduction (Fig. 2-4-1-1a), which enables direct interaction between the qAb and the antigen on the cell surface. A fluorescently labeled and antigen-specific antibody was employed as an rAb to specify a target epitope on the antigen and guide the determination of epitope similarity with qAbs during FCM analysis. When the rAb is incubated with the cells displaying a qAb sharing a similar epitope with the rAb, the binding of the rAb to the antigen is blocked, as the qAb has already preoccupied and masked the shared epitope. Consequently, the cells would manifest as an rAb-non-binding population [rAb(-)] (Fig. 2-4-1-1b). On the contrary, if the qAb binds to a different epitope than that of the rAb, it allows the epitope to remain accessible for rAb binding, resulting in an rAb-binding cell population [rAb(+)] (Fig. 2-4-1-1c).

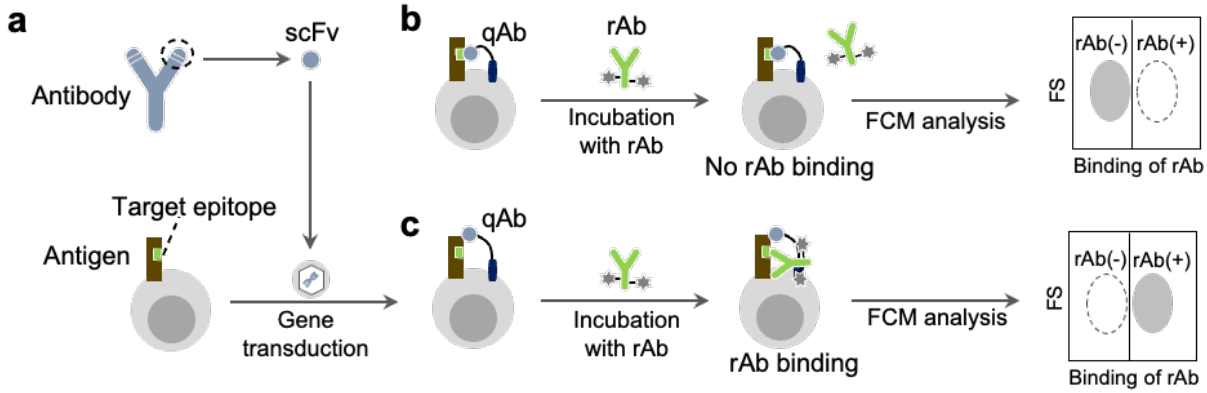



**Figure 2-4-1-1.** Schematic overview of the epitope similarity evaluation system. **(a)** Cell surface display of the query antibody (qAb) through gene transduction. The cDNA encoding a single chain variable fragment (scFv) of the antibody to be tested is transduced into the antigen-expressing cells and presented on cell surface as qAb. **(b)** Evaluation of qAb sharing the similar epitope with the fluorescently labeled reference antibody (rAb). Because the target epitope on the antigen is masked by the qAb, the rAb is unable to bind to the antigen, and an rAb-non-binding [rAb(-)] cell population is detected in FCM analysis. Forward scatter, FS. **(c)** Evaluation of qAbs with an epitope distinct from that of the rAb. The rAb binds to the target epitope, leading to the detection of an rAb-binding [rAb(+)] cell population in FCM analysis.

### 2.4.2 Construction of a system for epitope similarity evaluation

To demonstrate the feasibility of the antigen/antibody co-display system for epitope similarity evaluation, a model evaluation system was established using HER2 as the model antigen and

HER2-specific monoclonal antibodies pertuzumab and trastuzumab as rAbs. The binding free energy calculation through MD simulation revealed residues that comprised the pertuzumab- and trastuzumab-binding epitopes on HER2 (Fig. 2-4-2-1a). Residues with a predicted binding free energy value of < -1.5 kcal/mol were considered as binding residues and visualized on the surface of HER2 (Fig. 2-4-2-1b). These residues displayed a non-linear distribution, suggesting that pertuzumab and trastuzumab bind to distinct conformational epitopes on HER2.

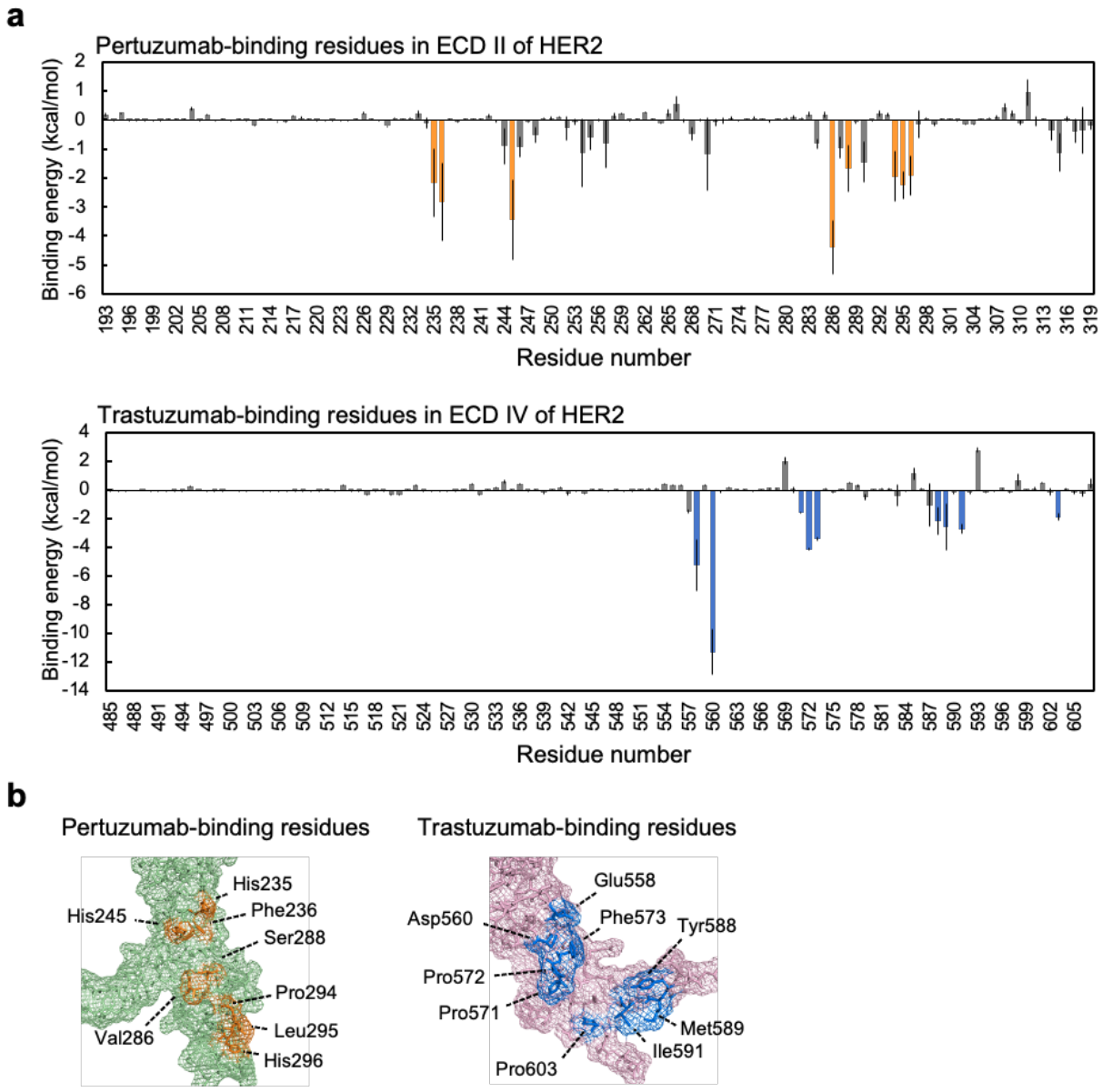



**Figure 2-4-2-1.** Identification of antibody-binding residues on HER2. **(a)** Determination of residues binding to pertuzumab and trastuzumab on HER2. The calculated binding free energy of each residue on HER2 to pertuzumab and trastuzumab is presented as the mean ± SEM. n = 3. Residues

with binding free energy values of < -1.5 kcal/mol, presented by orange and blue bars, are considered as epitope-constituting residues for pertuzumab and trastuzumab, respectively. **(b)** Visualization of amino acids comprising the conformational epitopes on HER2 for pertuzumab and trastuzumab. Antibody-binding residues are shown.

The system construction started with the establishment of antigen-expressing K562 cells (K562/HER2). A cDNA encoding HER2-T2A-mCherry (Fig. 2-4-2-2a, upper panel) was introduced into K562 cells through lentiviral transduction to allow for the cell surface display of the target antigen HER2 and the intracellular expression of a red fluorescent protein mCherry, which is achieved through the function of a self-cleaving T2A peptide. FCM analysis of the resultant K562/HER2 cells using an anti-HER2 antibody demonstrated a proportional correlation between the cell-surface expression of HER2 and the intracellular expression of mCherry (Fig. 2-4-2-2b, left). The scFvs derived from pertuzumab, trastuzumab, and a CD25-targeting monoclonal antibody daclizumab, designated scFv(Per), scFv(Tra), and scFv(Dac), respectively, were designed as qAbs (Fig. 2-4-2-2c). The cDNAs allowing the presentation of each scFv on the cell surface as a His-tagged qAb, along with an intracellular sfGFP[41], were constructed (Fig. 2-4-2-2a, lower panel). Following lentiviral transduction of the individual cDNAs into the K562/HER2 cell line, cells displaying each qAb, namely K562/HER2/scFv(Per), K562/HER2/scFv(Tra), and K562/HER2/scFv(Dac) were successfully established, with the expression visualized through fluorescence observation (Fig. 2-4-2-2d). The surface expression of qAbs was further confirmed using AF647-conjugated anti-His tag antibody, and the amount of cell-surface qAb was found to be proportionally correlated with the expression of sfGFP in the established cell lines (Fig. 2-4-2-2b, right).

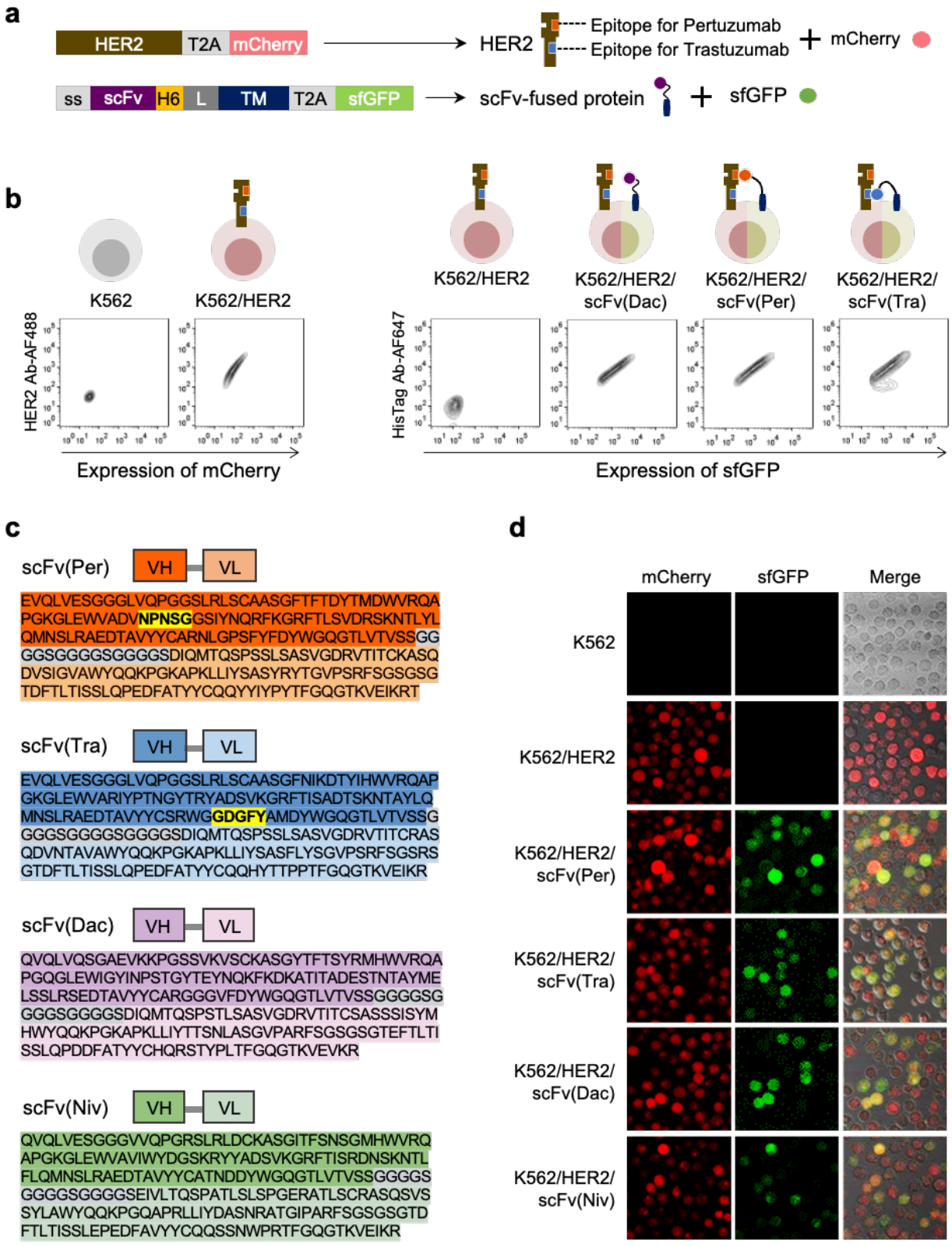



**Figure 2-4-2-2.** Validation of the protein expression of the evaluation system. **(a)** Schematic

representation of cDNA constructs in plasmids pCSII/HER2-mCherry (Upper panel) and pCSII/scFv-sfGFP (lower panel) for HER2 expression and surface display of scFv, respectively. ss, secretion signal peptide; H6, His6 tag peptide; L, linker peptide; TM, transmembrane domain. **(b)** Establishment of cell lines displaying qAbs on HER2-expressing cells. (Left) Expression of HER2 on K562/HER2 cells, validated by incubation with AF488-conjugated anti-HER2 antibody (HER2 Ab-AF488) and FCM analysis. (Right) Surface expression of qAbs on K562/HER2/scFv(Dac), K562/HER2/scFv(Per), and K562/scFv(Tra) cells, validated by incubation with AF647-conjugated anti-His tag antibody (HisTag Ab-AF647) and FCM analysis. (**c)** Structural diagram and the sequences of pertuzumab scFv [scFv(Per)], trastuzumab scFv [scFv(Tra)], daclizumab scFv [scFv(Dac)], and nivolumab scFv [scFv(Niv)]. The sequence of the heavy chain variable domain (VH) and light chain variable domain (VL) are highlighted in thick and thin colors, respectively. The linker sequence is highlighted in grey. Residues highlighted in yellow within scFv(Per) and scFv(Tra) denote positions for alanine substitution to generate mutant scFv(Per) and mutant scFv(Tra), respectively. **(d)** Expression of mCherry and sfGFP in K562/HER2/scFv(Per), K562/HER2/scFv(Tra), K562/HER2/scFv(Dac), and K562/HER2/scFv(Niv) cells. Scale bar = 50 μm.

### 2.4.3 Proof of concept validation for established system on epitope evaluation

To verify the capability of the system constructed in this study to evaluate epitope similarity, the three established cell lines were incubated with each of the two rAbs, AF647-labeled pertuzumab (Pert-AF647) and AF647-labeled trastuzumab (Tras-AF647). Upon treatment with the rAb Pert-AF647, K562/HER2/scFv(Per) showed an evident rAb(-) population, in contrast to K562/HER2/scFv(Dac) and K562/HER2/scFv(Tra) (Fig. 2-4-3-1a). This distinction arises because the cell surface expression of scFv(Per) effectively masked the epitope for Pert-AF647 on HER2, whereas the latter cells respectively express a qAb that either does not bind or binds to a distinct epitope as that of the rAb, Similarly, when treated these three cell lines with Tras-AF647 as rAb, the noticeable rAb(-) cell population was exclusively detected in K562/HER2/scFv(Tra) cells (Fig. 2-4-3-1b). These findings unequivocally validate that the system worked as intended and offer robust evidence of its effectiveness in evaluating epitope similarity.

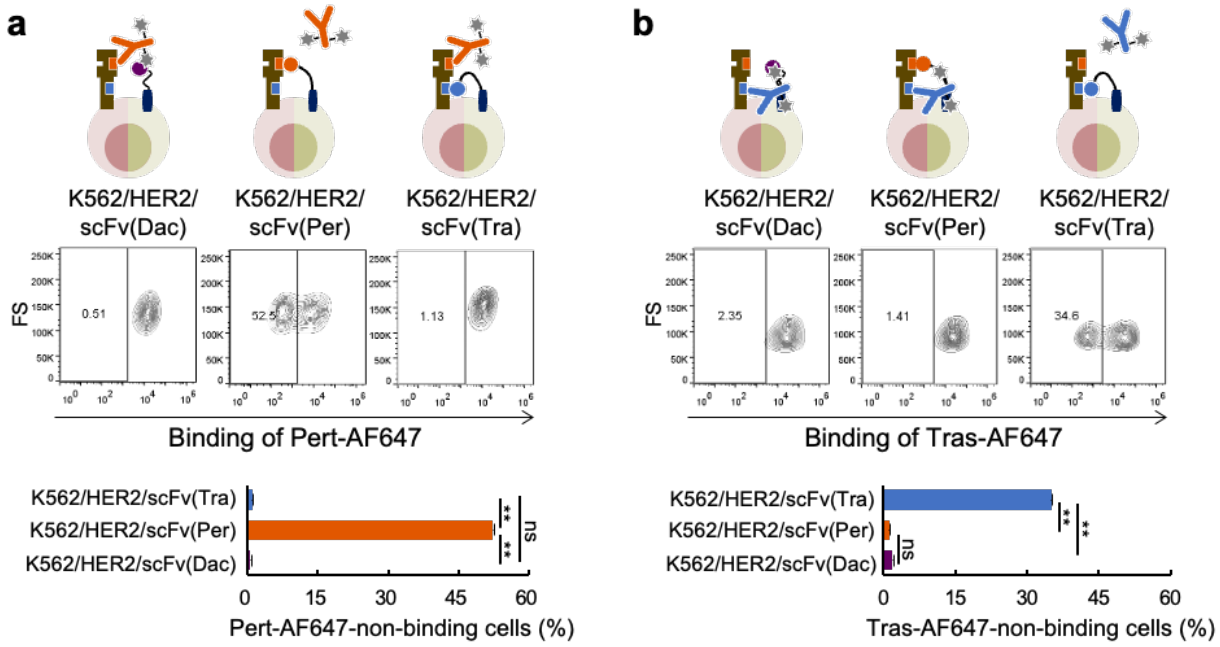



**Figure 2-4-3-1.** Validation of the feasibility for epitope similarity evaluation. **(a)** Assessment of epitope similarity between qAb and Pert-AF647. (Upper panel) K562/HER2/scFv(Dac), K562/HER2/scFv(Per), and K562/scFv(Tra) cells were incubated with 1 nM Pert-AF647 before FCM analysis. Representative contour plots are shown. (Lower panel) Percentage of Pert-AF647-non-binding cells calculated using FCM plots from three independent experiments is shown as the mean ± SEM. $^{**}p < 0.0001$. ns, not significant (Tukey's multiple comparisons test). n = 3. **(b)** Evaluation of epitope similarity between qAb and Tras-AF647. (Upper panel) K562/HER2/scFv(Dac), K562/HER2/scFv(Per), and K562/scFv(Tra) cells were incubated with 1 nM Tras-AF647 prior to FCM analysis. Representative contour plots are shown. (Lower panel) Percentage of Tras-AF647-non-binding cells calculated using FCM plots from three independent experiments is shown as the mean ± SEM. $^{**}p < 0.0001$. ns, not significant (Tukey's multiple comparisons test). n = 3.

### 2.4.4 Simultaneous evaluation of epitope similarity using two distinct rAbs

Having confirmed the effectiveness of the constructed system in evaluating epitope similarity of a single rAb-qAb pair using both rAbs Pert-AF647 and Tras-AF647, I explored its functionality to simultaneously evaluate four rAb-qAb pairs using dual rAbs with different epitopes. To verify the feasibility of simultaneous evaluation, pertuzumab and trastuzumab were labeled with AF647 and AF488, respectively, resulting in two rAbs, Pert-AF647 and Tras-AF488. To avoid the

excitation/emission spectral overlap between AF488 and sfGFP, the qAb cDNA was reconstructed to co-express with BFP (Fig. 2-4-4-1a). Following lentiviral transduction of the cDNAs into K562/HER2 cells, the resulting K562/HER2/scFv(Per)$^{B}$, K562/HER2/scFv(Tra)$^{B}$, and K562/HER2/scFv(Dac)$^{B}$ cells showed the co-expression of mCherry and BFP (Fig. 2-4-4-1b).

The established K562/HER2/scFv$^{B}$ cells were then simultaneously incubated with two different rAbs, Pert-AF647 and Tras-AF488, and subjected to FCM analysis. Both Pert-AF647 and Tras-AF488 bound to K562/HER2 and K562/HER2/scFv(Dac)$^{B}$ cells, with a linear correlation between the binding amount of each rAb (Fig. 2-4-4-1c). This observation confirmed that each rAb bound independently to HER2 without mutual interference when used simultaneously. Additionally, a Pert-AF647-non-binding/Tras-AF488-binding cell population was observed in K562/HER2/scFv(Per)$^{B}$ cells, while a Pert-AF647-binding/Tras-AF488-non-binding cell population was detected in K562/HER2/scFv(Tra)$^{B}$ cells (Fig. 2-4-4-1c). Furthermore, to validate whether the utilization of two rAb allow to simultaneously evaluate multiple rAb-qAb pairs with distinct epitopes, cells displaying different qAbs were combined pairwise in a 1:1 ratio. Upon incubation with both rAbs, the presence of cells expressing scFv(Per) or scFv(Tra) with the same epitope as either Pert-AF647 or Tras-AF488 led to the observation of the corresponding populations of cells that showed no binding to Pert-AF647 or Tras-AF488 (Fig. 2-4-4-1d). These results suggested that using two distinct rAbs enables the simultaneous evaluation of multiple rAb-qAb pairs with distinct epitopes in a single experiment.

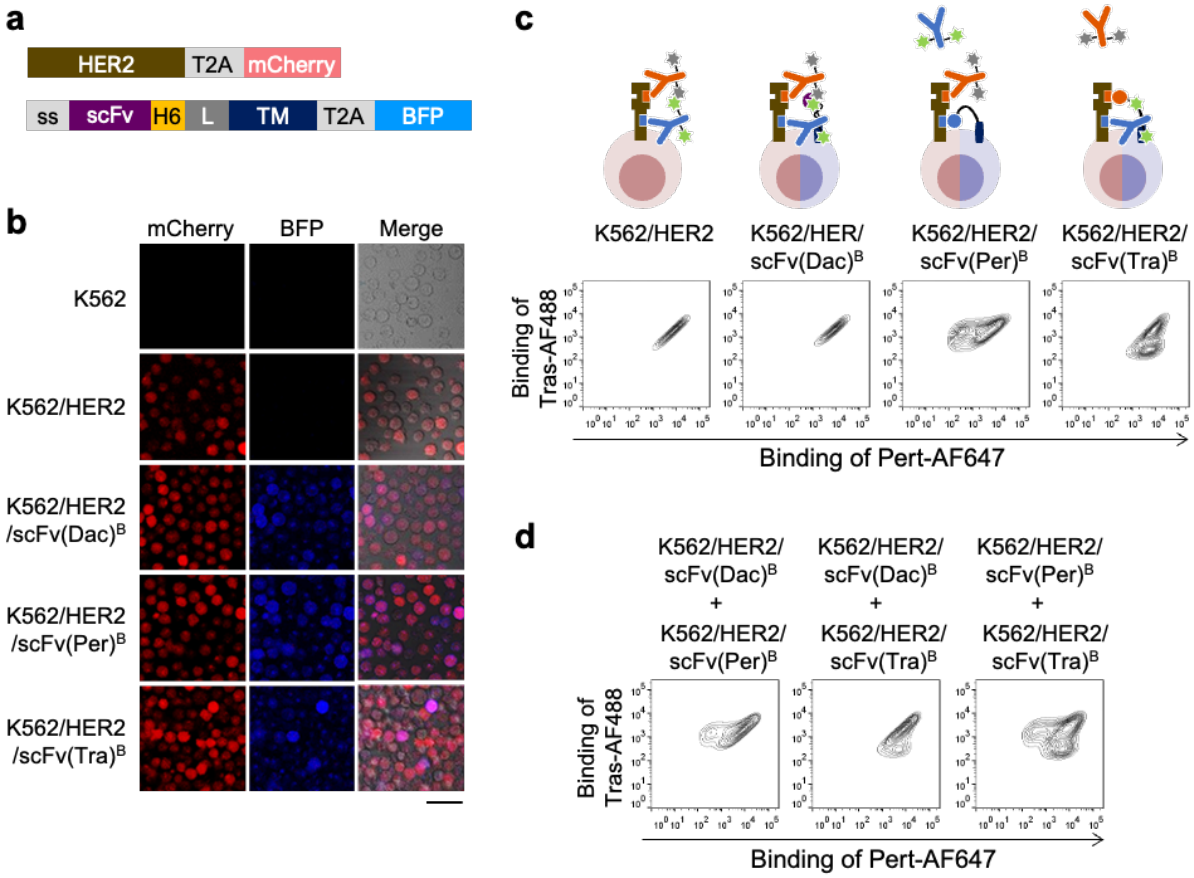



**Figure 2-4-4-1.** Simultaneous evaluation of multiple rAb-qAb pairs. **(a)** Schematic representation of the cDNA constructs in plasmids pCSII/HER2-mCherry (upper panel) and pCSII/scFv-BFP (lower panel) for HER2 expression and surface display of scFv, respectively. ss, secretion signal peptide; H6, His6 tag peptide; L, linker peptide; TM, transmembrane domain. **(b)** Expression of mCherry and BFP in K562/HER2/scFv(Dac)$^{B}$, K562/HER2/scFv(Per)$^{B}$, and K562/HER2/scFv(Tra)$^{B}$ cells. Scale bar = 50 μm. **(c)** Evaluation of epitope similarity between qAb and two rAbs. K562/HER2/scFv(Dac)$^{B}$, K562/HER2/scFv(Per)$^{B}$, and K562/HER2/scFv(Tra)$^{B}$ cells were treated with both 1 nM Pert-AF647 and 1 nM Tras-AF488 prior to FCM analysis. Representative contour plots are shown. **(d)** Evaluation of epitope similarity between two qAbs and two rAbs. K562/HER2/scFv(Dac)$^{B}$, K562/HER2/scFv(Per)$^{B}$, and K562/HER2/scFv(Tra)$^{B}$ cells were mixed pairwise in a 1:1 ratio and treated with both 1 nM Pert-AF647 and 1 nM Tras-AF488 prior to FCM analysis. Representative contour plots are shown.

## 2.5 Discussion

Competitive binding assays are typical approaches used to determine whether a pair of

antibodies binds to overlapping or distinct epitopes on the same target. By combining the co-presentation of both the qAb and the target antigen on the surface of mammalian cells with the use of exogenous rAb, a context for cell surface-based competitive binding detection was successfully established. The surface-displayed qAb specifically blocks the binding of the rAb sharing the same epitope, enabling the determination of the epitope similarity between qAb and rAb. The established system exhibits high specificity in indicating the epitope similarity between the qAb and rAb, as evidenced by the presence of the rAb(-) cell population only when the displayed qAb shares the same epitope with the rAb (Fig. 2-4-3-1). Thus, the rAb(-) population serves as an indication that qAb and rAb target a similar epitope. The findings in this chapter demonstrate the novel utility of the antigen/antibody co-display system in epitope characterization, extending beyond its conventional application in screening target-binding antibodies/peptides.

Furthermore, through the simultaneous utilization of two different rAbs, this platform could assess epitope similarities for at least four different rAb-qAb pairs, highlighting its capacity to simultaneously evaluate epitopes for a larger number of qAbs using various rAbs. Integration of the described approach in this study with multicolor FCM, wherein rAbs are distinguished by distinct fluorophores, could facilitate the analysis of multiple epitope groups in a single experiment. Such a multiplexed epitope binning approach would significantly expand the pool of antibody candidates targeting distinct epitopes on the same antigen and provide a greater variety of epitope combinations for the development of engineered antibody therapeutics, such as bispecific biologics with two antigen-binding arms for targeting distinct epitopes[82, 83].

# CHAPTER 3

# Characterization of the epitope evaluation system to unveil the factors influencing sensitivity

## 3.1 Overview

The interactions between antigens and antibodies are influenced by numerous external factors. In this chapter, I delved into several critical factors that impact the sensitivity of epitope similarity evaluation, including the concentrations of rAb, the relative abundance of antigen and qAb, as well as the affinity of qAb. To investigate the effect of rAb concentrations, cells co-expressing antigen and qAb were incubated with varying concentrations of rAb, revealing that the sensitivity decreased with increasing rAb concentration beyond a certain concentration threshold. Next, the effect of the relative levels between qAb and antigen was investigated from two aspects. Firstly, by estimating qAb and antigen expression in rAb(-) and rAb(+) populations within the same qAb/antigen-expressing cell line, I found higher qAb but lower antigen expression in rAb(-) compared with rAb(+) populations. Secondly, using cell lines with different levels of antigen expression, I analyzed the percentage of rAb(-) population and found that higher sensitivity could be achieved using cells with a higher qAb/antigen ratio. Furthermore, for investigating the effect of affinity of qAbs, qAbs with varying affinities to the target HER2 but targeting the similar epitope were generated and displayed on HER2-expressing cells. I found that the established system was able to evaluate epitope similarities for qAbs with varying affinities, and the sensitivity for evaluating high-affinity qAbs was less susceptible to the elevation of rAb concentration compared to low-affinity qAbs.

## 3.2 Introduction

### 3.2.1 Factors governing the antigen-antibody interactions

Antigen-antibody interactions are reversible processes characterized by dynamic association and dissociation, expressed as: antigen + antibody ⇌ antigen-antibody complex. In the initial stages, the reaction predominantly proceeds in the forward direction to form the antigen-antibody complex, while the reverse rate progressively increases over time until the reaction reaches equilibrium, where the forward and revere rates are equal[84]. The process is influenced by various factors, such as temperature, pH, and ionic strength, which can impact the equilibrium constant. Besides, according to the law of mass action: [antigen-antibody]/([antigen] ×[antibody]) = $k_a/k_d$ = $K_{eq}$, where the $k_a$ represents the association rate constant, $k_d$ represents the dissociation rate constant, $K_{eq}$ represents the equilibrium constant. The concentration of antigen and antibody, along

with their binding affinity, critically determines the rate and extent of the reaction. A higher concentration of antigen or antibody generally leads to a faster and more pronounced formation of antigen-antibody complex. In terms of affinity, the parameter indicating binding strength in antibody-antigen interactions, higher-affinity antibodies bind to antigens more efficiently and form more stable complexes compared to antibodies with lower affinity. As reported, low-affinity antibodies exhibited greater sensitivity to alterations in antibody concentration[84].

In the context of competitive binding experiments, where two distinct antibody molecules compete for the same binding site in a mutually exclusive manner, the aforementioned factors also affect the reaction dynamics, ultimately determining the predominant formation of a certain antibody-antigen complex.

### 3.2.2 Antibodies with similar sequences behave similarly

The sequence-structure-function paradigm posits that the primary structure of a protein, represented by the amino acid sequence, governs its three-dimensional structure, thereby determining its functional properties[85, 86]. This notion suggests that proteins with similar sequences tend to have similar structures and functions of proteins. It is widely acknowledged that antibodies sharing significant sequence similarity, particularly in regions responsible for antigen binding, are likely to target similar epitopes and exhibit functional similarity[26, 27]. This concept facilitates many immunological studies and biomedical research. For example, understanding the relationship between antibody sequence similarity and epitope specificity contributes to the development of vaccines with broad protective immune responses, because sequence-similar antibodies facilitate the identification of conserved epitopes on pathogens owing to their binding epitope similarities. Incorporating identified conserved epitopes into subunit vaccines can enhance vaccine efficacy against pathogen variants[87]. Additionally, antibodies sharing similar sequences are promising candidates to be engineered as novel therapeutic antibodies with enhanced properties, for instance, substituting critical residues of these antibodies to adjust the binding affinity without altering the binding epitopes of the engineered antibodies.

### 3.2.3 Research purpose of this chapter

To maximize the performance of the established epitope similarity evaluation system in future practical applications, the objective of this chapter is to explore factors that impact its evaluation

sensitivity. I aimed to investigate the impact of various factors influencing the antibody-antigen interactions and competitive binding assays on this evaluation system, including the concentration of rAb, the relative abundance of qAb and antigen, as well as the affinity of qAbs.

## 3.3 Material and methods

### 3.3.1 MD simulation

All MD simulations were conducted utilizing the AMBER 16 program package[79] on TSUBAME (Global Scientific Information and Computing Center at Tokyo Institute of Technology) as previously described[41]. The AMBER ff14SB force field was applied for proteins, while the TIP3P model was employed for water molecules. The initial coordinates of the pertuzumab Fab-HER2 and trastuzumab Fab-HER2 complexes were retrieved from PDB accession codes 1S78 and 1N8Z, respectively. To calculate binding free energy, the systems were fully solvated with explicit solvent, and two Na+ counterions were introduced to maintain electrostatic neutrality. The systems were optimized through energy minimization and equilibrated with backbone restraints. A 100 ns production run was then performed, of which the final 10 ns was utilized by the molecular mechanics/generalized Born surface area (MM/GBSA) module to calculate the binding free energy.

### 3.3.2 Plasmid construction

All recombinant DNA experiments were conducted with the approval of the Recombinant DNA Experimental Safety Management Committee of Tokyo Institute of Technology (no. I2021017).

To enable the extracellular secretion of scFvs, a cDNA encoding a fusion protein was constructed. The fusion protein consists of the PD-L1 secretion signal (amino acids 1–18, UniProt Q9NZQ7), the scFv, a single chain fragment of two human IgG1 Fc domains (scFc) (amino acids 104–330, UniProt P01857), a His6 tag peptide, a T2A peptide, and sfGFP. Subsequently, this construct was inserted into the multiple cloning site of the CSII-CMV-MCS plasmid. The resulting plasmid was designated pCSII/scFv-scFc-H6.

To display scFv variants on cell surfaces, the DNA fragment of scFv(Dac) in the previously constructed plasmid pCSII/scFv-sfGFP[41] was substituted with that of other scFv(Per) mutants carrying N52A, P53A, G54A, S55A, G56A mutation, and scFv(Tra) mutants carrying G101A,

D102A, G103A, F104A, Y105A mutation, and scFv from the anti-PD1 nivolumab. This resulted in the construction of pCSII/scFv-sfGFP plasmids.

### 3.3.3 Cell culture

The subclones of the human embryonic kidney cell line HEK293, namely HEK293T and Lenti-X 293T, were purchased from RIKEN Bio-Resource Center and Clontech, respectively. The K562/HER2/scFv, K562/HER2$^{Lo}$/scFv, and K562/HER2$^{Hi}$/scFv cell lines were established through lentiviral transduction of K562/HER2 cells with the plasmids pCSII/scFv-sfGFP.

K562/HER2/scFv, K562/HER2$^{Lo}$/scFv, and K562/HER2$^{Hi}$/scFv cells, as well as Lenti-X 293T and HEK293T cells, were cultured in 10% FBS-RPMI-1640 (Invitrogen) and 10% FBS-DMEM (Nacalai Tesque), respectively. All media were supplemented with penicillin (100 U/mL) and streptomycin (100 μg/mL) (Nacalai Tesque). Cells were maintained in a 37°C incubator with 5% $CO_2$ and subject to regular checks for mycoplasma contamination using a mycoplasma test kit (Lonza).

### 3.3.4 Lentivirus production and transduction

Plasmid pCSII/scFv-scFc-H6 or pCSII/scFv-sfGFP was mixed with Lentiviral High Titer Packaging Mix (Takara Bio) and transfected into Lenti-X 293T cells using Lipofectamine LTX (Invitrogen). After 24 h of incubation, the medium was replaced with 10% FBS-DMEM containing 10 μM forskolin (Fujifilm Wako Pure Chemical). At 48 h post-transfection, the lentiviruses in the medium were collected and concentrated using a Lenti-X Concentrator (Clontech) in accordance with the manufacturer's instructions.

To establish K562/HER2/scFv cells with different expression levels of HER2, K562/HER2 cells with low HER2 expression (K562/HER2$^{Lo}$) and high HER2 expression (K562/HER2$^{Hi}$) were separately transduced with lentiviruses carrying gene from pCSII/scFv-sfGFP using a modified spinoculation protocol[80, 81] at a MOI of 0.3, resulting in the construction of K562/HER2$^{Lo}$/scFv and K562/HER2$^{Hi}$/scFv cells.

To establish HEK293T cells stably secreting scFv-scFc-H6 fusion proteins, the lentivirus carrying genes from pCSII/scFv-scFc-H6 was resuspended in the medium with 8 μg/mL polybrene (Nacalai Tesque). The mixture was then added to approximately 80%–90% confluent HEK293T cells. After incubation for 48 h, the transduction efficiency was verified under a Biorevo BZ-710 fluorescence microscope (Keyence, Osaka, Japan) with the appropriate filters for sfGFP (Ex/Em = 470 ± 40 nm/520 ± 50 nm).

To establish cells co-expressing HER2 and sfGFP along with scFv variant, K562/HER2 cells were transduced with lentiviruses carrying gene from pCSII/scFv-sfGFP using a modified spinoculation protocol[80, 81] at a MOI of 0.3, resulting in the construction of K562/HER2/scFv cells presenting individual scFv variants.

### 3.3.5 Protein purification

To avoid the influence of FBS on the purity of purified product, HEK293T cells secreting the scFv-scFc-H6 proteins were cultured in 1% FBS-DMEM medium for 48 h, after which the culture supernatant was harvested. The proteins were purified from the supernatant using Ni-NTA agarose (QIAGEN, Hilden, Germany) following the manufacturer's manual. The purity of the isolated protein was validated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and silver staining using the Silver Stain KANTO III kit (KANTO CHEMICAL, Tokyo, Japan).

### 3.3.6 FCM analyses

To assess the effect of rAb concentrations on evaluation sensitivity, $2\times10^5$ K562/HER2/scFv cells were incubated with either Pert-AF647 or Tras-AF647 at varying concentrations of 0.1, 1, 10, 100 nM for 1 h at 4°C prior to FCM analysis.

To assess the effect of antigen density on evaluation sensitivity, $2\times10^5$ K562/HER2/scFv, K562/HER2$^{Lo}$/scFv, or K562/HER2$^{Hi}$/scFv cells were incubated with either Pert-AF647 or Tras-AF647 at the concentration of 1 nM for 1 h at 4°C prior to FCM analysis using a FACSAria III (BD Biosciences).

To evaluate the HER2-binding affinity of scFvs, $2\times10^5$ K562/HER2 cells were incubated with varying concentrations (0.5, 1, 5, 10, 25, 50, 100, 200, 400, 800, and 1600 nM) of scFv-scFc-H6 proteins for 90 min at 4°C. The cells were then stained with AF488-conjugated anti-His tag antibody (clone J099B12, 1:400; BioLegend) for 40 min at 4°C and analyzed by FCM analysis.

To assess epitope similarity for qAbs with different affinities, $2\times10^5$ K562/HER2/scFv cells displaying each scFv variant were stained with their respective rAb, Pert-AF647 or Tras-AF647 at the concentrations of 0.1 and 10 nM, for 1 h at 4°C prior to FCM analysis.

In all FCM experiments, protein and antibody dilution as well as cell washing were performed using 3% FBS-PBS. Data acquisition was carried out using an iCyt ec800 flow cytometer (Sony Biotechnology) unless otherwise stated, the flow cytometric data were analyzed using FlowJo software ver.10.8.1 (FlowJo).

### 3.3.7 Statistics analysis

All experimental data were obtained from three independent experiments and are expressed as the mean ± SEM. Statistical analysis was performed using GraphPad Prism 10 (GraphPad Software). *P*-values were determined using Tukey's multiple comparisons test or two-side unpaired Student's *t* tests, as specified in the figure legends. $P < 0.05$ was considered statistically significant.

## 3.4 Results

### 3.4.1 Effect of rAb concentration on evaluation sensitivity

The basis for judging the similar epitopes between rAb and qAb lies in the ability of qAb to inhibit rAb binding to the antigen. Given that, the percentage of rAb(-) cells was utilized to indicate the sensitivity of epitope similarity evaluation, reflecting the proportion of epitope-similar qAbs that can be effectively identified when a single rAb-qAb pair exists. As this evaluation system lies in the competitive binding between qAb and rAb with the antigen, the proportions of qAb and rAb become crucial in determining which antibody gains a competitive advantage. In addition, considering the impact of antibody concentration on antigen-antibody interactions, an investigation into the influence of rAb concentration on assessment sensitivity was conducted using the established model system. K562/HER2/scFv(Per) and K562/HER2/scFv(Dac) cells were incubated with varying concentrations of Pert-AF647, followed by FCM analysis to determine the percentage of rAb(-) cell population (Fig. 3-4-1-1a). When reacting with rAb concentrations ranging from 0.1 to 10 nM, the differences in the percentage of the rAb(-) cell population among K562/HER2/scFv(Per) cells were negligible. However, a significant reduction in the rAb(-) cell population percentage was observed when treated with a higher rAb concentration of 100 nM. K562/HER2/scFv(Dac) cells, the negative control without the expression of an scFv against HER2, exhibited no observable rAb(-) cell population across any concentration of Pert-AF647. When Tras-AF647 was used as the rAb in an experiment with K562/HER2/scFv(Tra) cells, no significant difference in the percentage of rAb(-) cells was observed at relatively lower rAb concentrations of 0.1 nM and 1 nM (Fig. 3-4-1-1b). However, unlike Pert-AF647, a reduction was observed upon increasing the concentration of Tras-AF647 to 10 nM (Fig. 3-4-1-1b). Although the concentration points at which the sensitivity reduction occurred slightly vary between the two different rAbs, these observations are still robust enough to demonstrate that the rAb concentration does affect the

sensitivity of the assessment, particularly when it exceeds a certain threshold.

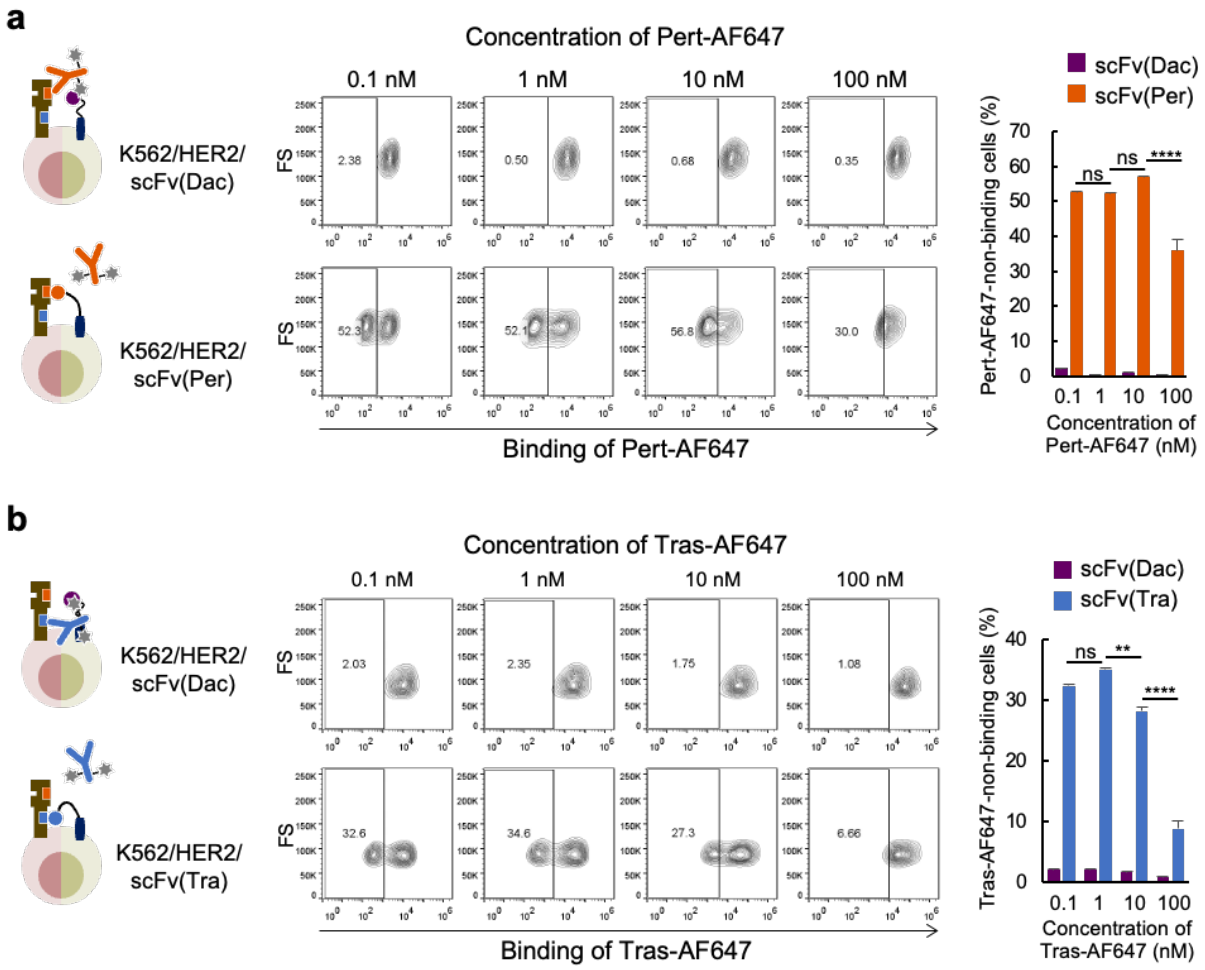



**Figure 3-4-1-1.** Effects of rAb concentrations on evaluation sensitivity. **(a)** Assessment of epitope similarity between qAb and rAb using varying concentrations of Pert-AF647. (Left) K562/HER2/scFv(Dac) and K562/HER2/scFv(Per) cells were incubated with 0.1–100 nM Pert-AF647 prior to FCM analysis. Representative contour plots are shown. (Right) Percentage of Pert-AF647-non-binding cells calculated using FCM plots from three independent experiments is presented as the mean ± SEM. n = 3. ns, not significant; ****$p < 0.0001$ (Tukey's multiple comparisons test). **(b)** Evaluation of epitope similarity between qAb and various concentrations of Tras-AF647. (Left) K562/HER2/scFv(Dac) and K562/HER2/scFv(Tra) cells were incubated with 0.1–100 nM Tras-AF647 prior to FCM analysis. Representative contour plots are shown. (Right) Percentage of Tras-AF647-non-binding cells calculated using FCM plots from three independent experiments is presented as the mean ± SEM. n = 3. ns, not significant; **$p < 0.01$, ****$p < 0.0001$ (Tukey's multiple comparisons test).

### 3.4.2 Effect of relative expression levels between qAb and antigen on evaluation sensitivity

In the investigation utilizing K562/HER2/scFv(Per) and K562/HER2/scFv(Tra) cells along with their corresponding rAb, Pert-AF647 and Tras-AF647, I noticed the presence of rAb(+) cell populations (Fig. 2-4-3-1), which contradicts the expected result based on the assumption that the qAb had already occupied the epitope for the rAb. These observations led to a hypothesis that the unexpected rAb(+) populations may result from the heterogeneous expression levels of HER2 and qAbs within individual cells. In situations where both the antigen and qAb are presented on the same cell surface, the availability of the epitope for rAb is contingent upon the relative molecular ratio between the antigen and qAb. When qAb is expressed in relatively higher quantities compared to HER2, it occupies all available epitopes on HER2, effectively preventing rAb from binding to HER2; conversely, if qAb is expressed relatively less than HER2, rAb may be able to bind to spare epitope on HER2. To elucidate this hypothesis, I analyzed the expression levels of qAb and antigen in rAb(-) and rAb(+) cell populations using the expression of sfGFP and mCherry, respectively (Fig. 3-4-2-1a, b). As expected, rAb(-) cells exhibited higher qAb expression but lower HER2 expression levels relative to rAb(+) cells. The consistent results were observed in both the Pert-AF647 treated K562/HER2/scFv(Per) cells and Tras-AF647 treated K562/HER2/scFv(Tra) cells, suggesting the impact of the relative expression levels of qAb and antigen on evaluation sensitivity.

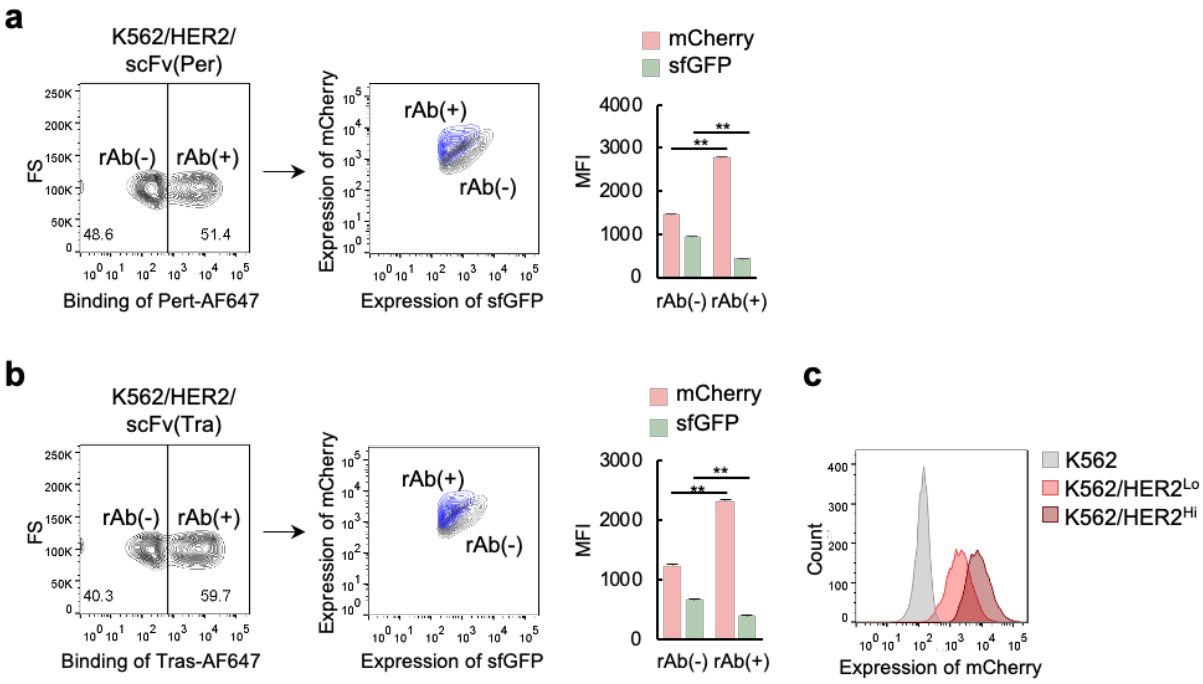



**Figure 3-4-2-1.** Difference in expression levels of qAb and HER2 within rAb(-) and rAb(+) cell populations. **(a)** Comparison of scFv(Per) and HER2 expression levels in Pert-AF647-non-binding [rAb(-)] and Pert-AF647-binding [rAb(+)] cell populations. K562/HER2/scFv(Per) cells were incubated with 1 nM Pert-AF647 prior to FCM analysis, with the expression levels of scFv(Per) and HER2 estimated by the mean fluorescence intensity (MFI) values of sfGFP and mCherry, respectively. **(b)** Comparison of scFv(Tra) and HER2 expression levels in Trast-AF647-non-binding [rAb(-)] and Tras-AF647-binding [rAb(+)] cell populations. K562/HER2/scFv(Tra) cells were incubated with 1 nM Tras-AF647 prior to FCM analysis, with the expression levels of scFv(Tra) and HER2 estimated by the MFI of sfGFP and mCherry, respectively. Representative contour plots are shown, and data from three independent experiments are presented as mean ± SEM. n = 3. ** $p$ <0.0001 ($t$-test). **(c)** Expression level of HER2 on K562/HER2$^{Lo}$ and K562/HER2$^{Hi}$ cells, as determined by the fluorescent intensity of mCherry. Representative histograms are shown.

To further validate the effect of the qAb/antigen balance, two K562/HER2 cell lines with different HER2 expression levels, specifically low HER2 expression (K562/HER2$^{Lo}$) and high HER2 expression (K562/HER2$^{Hi}$) (Fig. 3-4-2-1c), were selected for establishing K562/HER2$^{Lo}$/scFv(Per) and K562/HER2$^{Hi}$/scFv(Per) cells, respectively. As indicated by the sfGFP/mCherry ratio, the molecular ratio of HER2 to qAb in K562/HER2$^{Lo}$/scFv(Per) cells was lower than that in K562/HER2$^{Hi}$/scFv(Per) cells (Fig. 3-4-2-2a). Upon treatment with Pert-AF647,

a notably smaller rAb(-) cell population was observed in K562/HER2$^{Hi}$/scFv(Per) cells in comparison to K562/HER2$^{Lo}$/scFv(Per) cells (Fig. 3-4-2-2b). Similarly, K562/HER2$^{Lo}$/scFv(Tra) and K562/HER2$^{Hi}$/scFv(Tra) cells, which respectively exhibited relatively higher and lower qAb/HER2 ratios, were established (Fig. 3-4-2-2c). The treatment with Tras-AF647 yielded the expected result, as a significantly larger rAb(-) population was observed in K562/HER2$^{Lo}$/scFv(Tra) cells compared to K562/HER2$^{Hi}$/scFv(Tra) cells (Fig. 3-4-2-2d). Collectively, these findings demonstrate that a higher qAb/antigen ratio correlates with improved performance in the epitope similarity evaluation system.

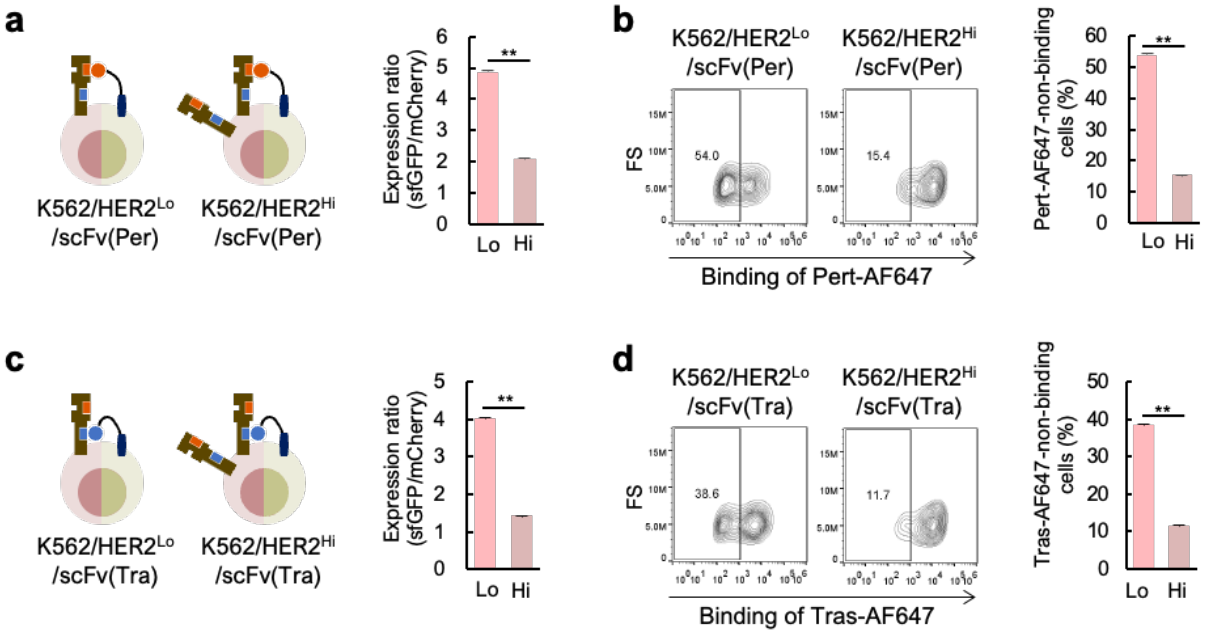
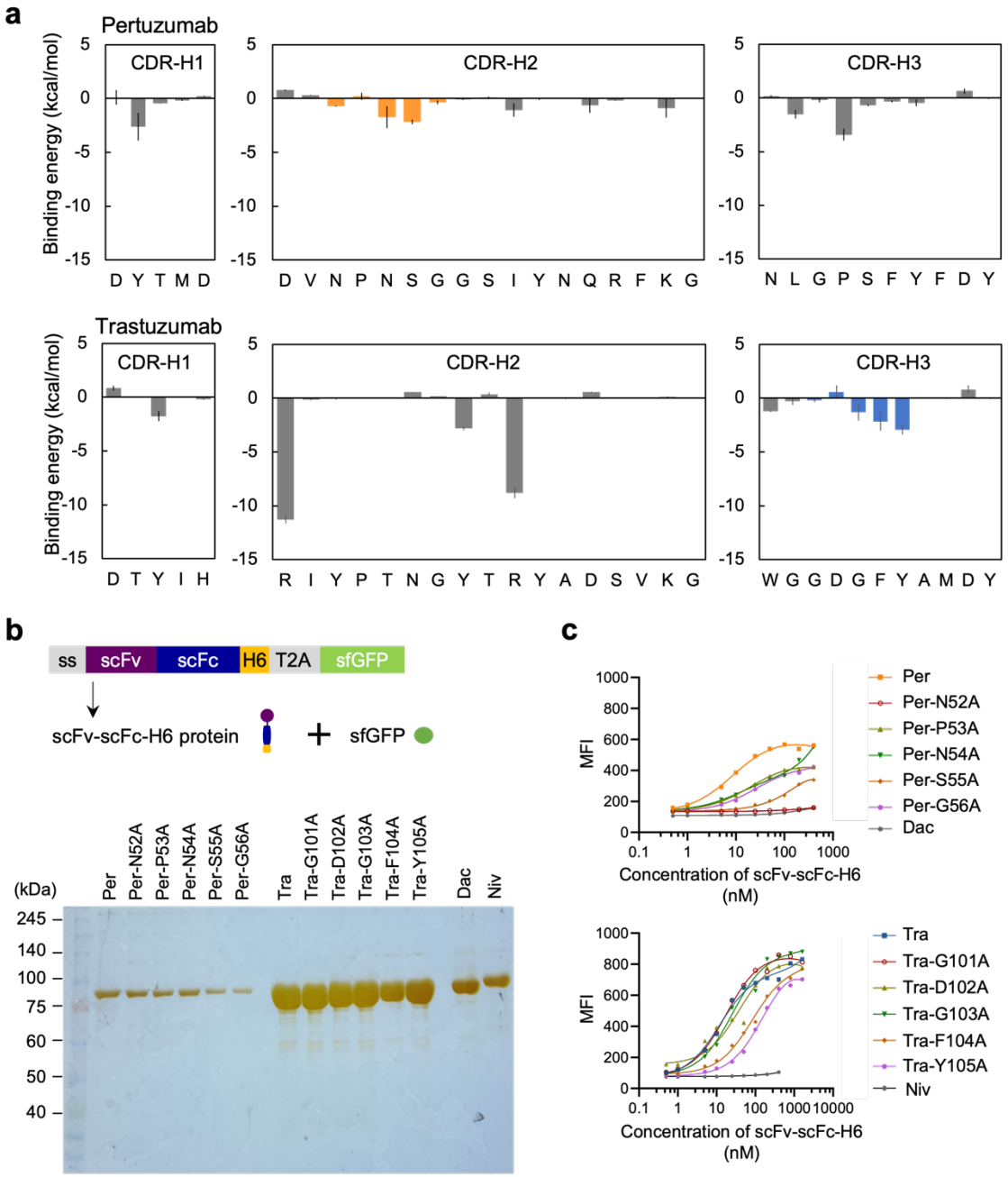



**Figure 3-4-2-2.** Effects of the ratio between qAb and HER2 expression levels on evaluation sensitivity. **(a)** Relative expression levels between scFv(Per) and HER2 in K562/HER2$^{Lo}$/scFv(Per) (Lo) and K562/HER2$^{Hi}$/scFv(Per) (Hi) cells. Expression levels of scFv(Per) and HER2 were estimated by the fluorescent intensity of sfGFP and mCherry, respectively. The relative expression ratio was calculated using the MFI values of sfGFP and mCherry. **(b)** Evaluation of epitope similarity between qAb and Pert-AF647 using cells with different HER2 expression levels. (Left) K562/HER2$^{Lo}$/scFv(Per) (Lo) and K562/HER2$^{Hi}$/scFv(Per) (Hi) cells were incubated with 1 nM Pert-AF647 prior to FCM analysis. (Right) Percentage of Pert-AF647-non-binding cells calculated using FCM plots. **(c)** Relative expression levels between scFv(Tra) and HER2 in K562/HER2$^{Lo}$/scFv(Tra) (Lo) and K562/HER2$^{Hi}$/scFv(Tra) (Hi) cells. Expression levels of scFv(Tra) and HER2 were estimated by the fluorescent intensity of sfGFP and mCherry, respectively. The relative expression ratio was calculated using the MFI values of sfGFP and

mCherry. **(d)** Evaluation of epitope similarity between qAb and Tras-AF647 using cells with different HER2 expression levels. (Left) K562/HER2$^{Lo}$/scFv(Tra) (Lo) and K562/HER2$^{Hi}$/scFv(Tra) (Hi) cells were incubated with 1 nM Tras-AF647 prior to FCM analysis. (Right) Percentage of Tras-AF647-non-binding cells calculated using FCM plots. Representative contour plots are shown, and data from three independent experiments are presented as mean ± SEM. n = 3. ** $p$ <0.0001 ($t$-test).

### 3.4.3 Effect of qAb affinity on evaluation sensitivity

Antibodies with higher affinity generally form more stable antigen-antibody complexes, exhibiting stronger competition in competitive binding assays compared to those with lower affinity. Consequently, antibodies with lower affinity to a functional epitope may be overlooked due to their weaker competitive ability. To assess the ability of this system to evaluate qAbs with weak HER2-binding affinities, as well as to unveil the impact of qAb affinity on the performance of this system, mutant scFvs were designed based on the results of binding free energy calculations (Fig. 3-4-3-1a). Five consecutive residues contributing to HER2 binding to varying degrees were individually replaced with alanine, resulting in the creation of scFv(Per) and scFv(Tra) mutants. The scFvs, including both wild-type and mutants, namely scFv(Per), N52A mutant scFv(Per), P53A mutant scFv(Per), N54A mutant scFv(Per), S55A mutant scFv(Per), G56A mutant scFv(Per), scFv(Tra), G101A mutant scFv(Tra), D102A mutant scFv(Tra), G103A mutant scFv(Tra), F104A mutant scFv(Tra), Y105A mutant scFv(Tra), scFv(Dac), and scFv of PD-1 targeting nivolumab (Fig. 2-4-2-2c), are hereafter referred to as Per, Per-N52A, Per-P53A, Per-N54A, Per-S55A, Per-G56A, Tra, Tra-G101A, Tra-D102A, Tra-G103A, Tra-F104A, Tra-Y105A, Dac, Niv, respectively. The cDNA plasmids encoding each of the aforementioned scFvs were constructed (Fig. 3-4-3-1b, upper panel), and all 14 scFvs were purified as fusion proteins with a C-terminal His-tagged single chain Fc (scFc) domain from the culture supernatant of transduced HEK293T cells (Fig. 3-4-3-1b, lower panel). To determine their binding affinity, the surface binding of purified scFvs to HER2-expressing cells was quantified by measuring the binding of an anti-His antibody to the His-tagged scFv (Fig. 3-4-3-1c). The parental scFvs, Per and Tra, along with their variants, exhibited a wide range of binding affinities to HER2, as evidenced by various $K_D$ values, while Dac and Niv exhibited no observable binding, as expected (Fig. 3-4-3-1d). These results demonstrated the successful construction of affinity qAb variants.




d

| scFvs | $K_D$ (nM)* |
|---|---|
| Per | 9.0 ± 2.2 |
| Per-N52A | N.D. |
| Per-P53A | 20 ± 2.6 |
| Per-N54A | N.D. |
| Per-S55A | 230 ± 12 |
| Per-G56A | 26 ± 4.5 |
| Dac | N.B. |

| scFvs | $K_D$ (nM)* |
|---|---|
| Tra | 12 ± 0.12 |
| Tra-G101A | 20 ± 1.7 |
| Tra-D102A | 16 ± 7.2 |
| Tra-G103A | 30 ± 2.2 |
| Tra-F104A | 131 ± 35 |
| Tra-Y105A | 170 ± 0.27 |
| Niv | N.B. |

* N.D., not determined; N.B., no binding.

**Figure 3-4-3-1.** Generation of mutant scFvs with various HER2-binding affinity. **(a)** Identification of HER2-binding residues on pertuzumab and trastuzumab. The binding free energy of individual residues within CDRs of pertuzumab and trastuzumab to HER2 was calculated and is presented as the mean ± SEM. Bars in orange and blue represent residues for alanine substitution to generate mutant scFv(Per) and mutant scFv(Tra), respectively. **b)** Production of scFc- and 6×His tag (H6)-fused scFv (scFv-scFc-H6) proteins. (Upper panel) Schematic representation of cDNA construct in the plasmids pCSII/scFv-scFc designed for the expression of scFv-scFc-H6 proteins which are secreted extracellularly. (Lower panel) Purity of purified scFv-scFc-H6 proteins. 14 scFvs, specifically Per, Per-N52A, Per-P53A, Per-N54A, Per-S55A, Per-G56A, Tra, Tra-G101A, Tra-D102A, Tra-G103A, Tra-F104A, Tra-Y105A, Dac, and Niv, were expressed as scFv-scFc-H6 proteins and purified using affinity chromatography resins. The purified protein solutions were separated via SDS-PAGE and visualized with silver staining. The approximate size of all fusion proteins was 80 kDa. **(c)** Binding of scFv-scFc-H6 proteins to HER2. K562/HER2 cells were incubated with various concentrations of scFv-scFc-H6 proteins prior to FCM analysis using AF488-conjugated anti-His tag antibody. Representative plots of MFI values against the concentrations of scFv-scFc-H6 proteins are shown. Plots were fitted using a one-site total nonlinear regression curve fitting model in GraphPad Prism 10 for $K_D$ calculation. **(d)** HER2-binding affinity of qAbs. $K_D$ ± SEM values are shown. n = 3.

To assess the performance of this system in evaluating qAbs with varying affinities, these qAb variants were presented on the K562/HER2 cell surface and detected for rAb binding by FCM analysis. The treatment of cells displaying Per and its mutants with 0.1 nM Pert-AF647 resulted in a consistent percentage (49.4-59.0%) of the rAb(-) cell population for most qAbs, except for a notable reduction observed for Per-N52A (29.3%), which exhibited the lowest binding affinity (Fig. 3-4-3-2a). Moreover, when increasing the concentration of Pert-AF647 to 10 nM, the percentage of the rAb(-) cell population decreased as the binding affinity of the qAbs reduced (Fig. 3-4-3-2b). In the analysis of cells displaying Tra and its mutants with Tras-AF647 as the rAb, it was still observed that the percentages of the rAb(-) cell population decreased with diminishing affinity at the higher concentration of 10 nM but remained consistent at the lower concentration of 0.1 nM (Fig. 3-4-3-2c, d). These results clearly indicate the capability of this system in assessing epitope similarity for qAbs with various affinities. Furthermore, they also reveal the influence of qAb affinity on the sensitivity of epitope evaluation, particularly under the condition of a relatively

high rAb concentration.

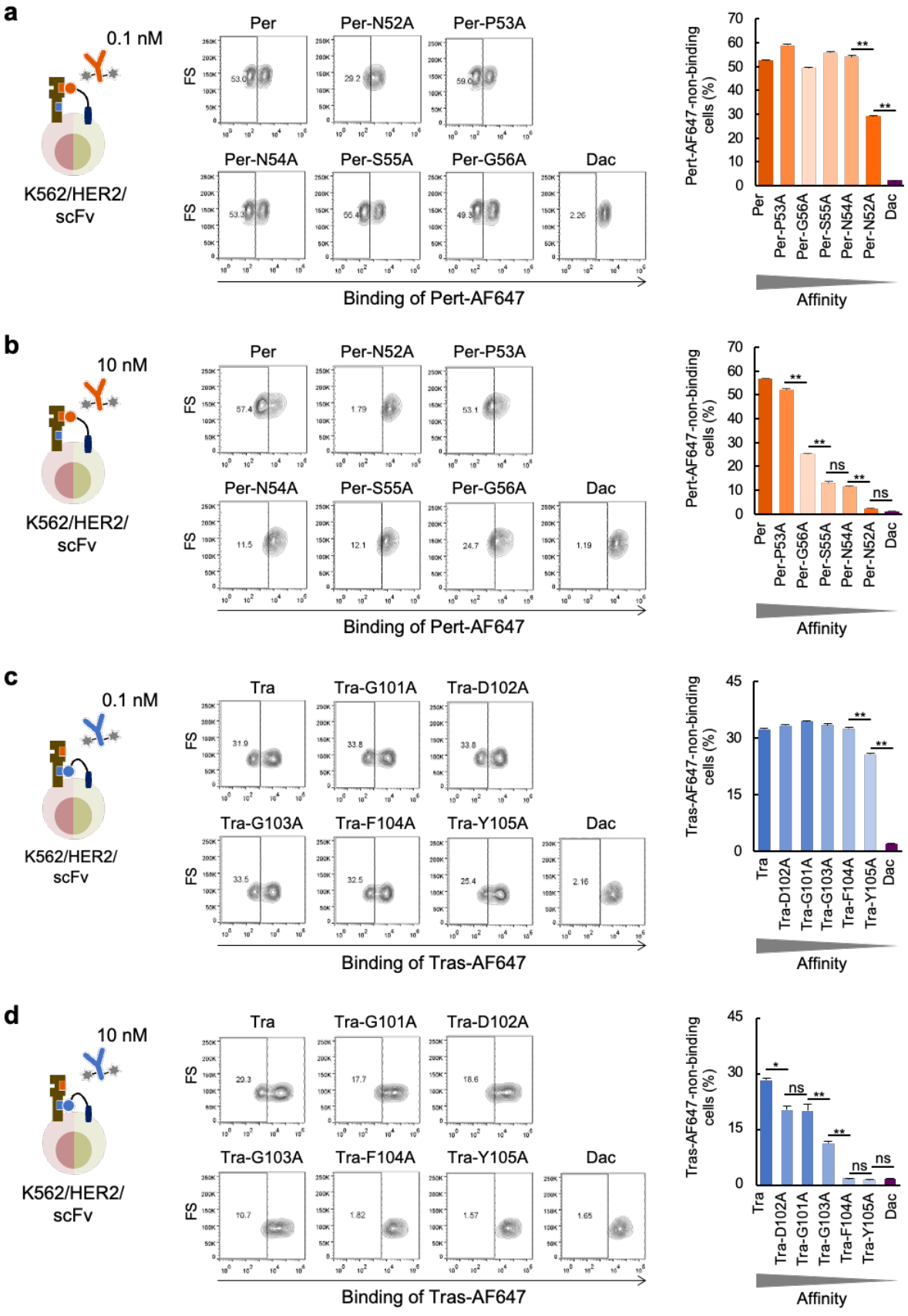

a
0.1 nM
K562/HER2/
scFv
Per
Per-N52A
Per-P53A
Per-N54A
Per-S55A
Per-G56A
Dac
FS
53.0
29.2
59.0
53.3
55.4
49.3
2.26
Binding of Pert-AF647
Pert-AF647-non-binding
cells (%)
70
60
50
40
30
20
10
0
**
**
Per
Per-P53A
Per-G56A
Per-S55A
Per-N54A
Per-N52A
Dac
Affinity
b
10 nM
K562/HER2/
scFv
Per
Per-N52A
Per-P53A
Per-N54A
Per-S55A
Per-G56A
Dac
FS
57.4
1.79
53.1
11.5
12.1
24.7
1.19
Binding of Pert-AF647
Pert-AF647-non-binding
cells (%)
70
60
50
40
30
20
10
0
**
**
ns
**
ns
Per
Per-P53A
Per-G56A
Per-S55A
Per-N54A
Per-N52A
Dac
Affinity
c
0.1 nM
K562/HER2/
scFv
Tra
Tra-G101A
Tra-D102A
Tra-G103A
Tra-F104A
Tra-Y105A
Dac
FS
31.9
33.8
33.8
33.5
32.5
25.4
2.16
Binding of Tras-AF647
Tras-AF647-non-binding
cells (%)
45
30
15
0
**
**
Tra
Tra-D102A
Tra-G101A
Tra-G103A
Tra-F104A
Tra-Y105A
Dac
Affinity
d
10 nM
K562/HER2/
scFv
Tra
Tra-G101A
Tra-D102A
Tra-G103A
Tra-F104A
Tra-Y105A
Dac
FS
29.3
17.7
18.6
10.7
1.82
1.57
1.65
Binding of Tras-AF647
Tras-AF647-non-binding
cells (%)
45
30
15
0
*
ns
**
**
ns
ns
Tra
Tra-D102A
Tra-G101A
Tra-G103A
Tra-F104A
Tra-Y105A
Dac
Affinity

**Figure 3-4-3-2.** Epitope evaluation for qAbs with various HER2-binding affinities. **(a-b)** Epitope similarity evaluation between qAb and Pert-AF647 at 0.1 nM or 10 nM. (Left) K562/HER2 cells displaying Per, Per mutants, and Dac were incubated with 0.1 nM (**a**) or 10 nM (**b**) Pert-AF647 prior to FCM analysis. (Right) Percentage of Pert-AF647-non-binding cells calculated using FCM plots. **(c-d)** Epitope similarity evaluation between qAb and Tras-AF647 at 0.1 nM or 10 nM. (Left) K562/HER2 cells displaying Tra, Tra mutants, and Dac were incubated with 0.1 nM (**c**) or 10 nM (**d**) Tras-AF647 prior to FCM analysis. (Right) The percentage of Tras-AF647-non-binding cells calculated using FCM plots. Representative contour plots are shown, and data from three independent experiments are presented as mean ± SEM. n = 3. $^{**}p < 0.0001$, $^{*}p < 0.001$. ns, not significant (Tukey's multiple comparisons test).

## 3.5 Discussion

### 3.5.1 Mechanisms underlying the various factors influencing evaluation sensitivity

Dynamic and reversible antigen-antibody interactions are influenced by a variety of factors. In the established epitope similarity evaluation system, which involves competitive binding inhibition conducted on the cell surface, two critical factors must be considered: the relative molecular ratios of qAb to the antigen, and the ratio of qAb to rAb. The qAb/antigen ratio governs the availability of unoccupied epitopes for rAb, and the qAb/rAb ratio determines whether the predominant complex formed is the qAb-antigen or the rAb-antigen. Essentially, the competitive binding and binding inhibition are closely associated with the shift in antigen-antibody reaction toward a specific direction: forming more rAb-antigen or more qAb-antigen complexes (Fig. 3-5-1-1). The elevated concentration of rAb promotes the replacement of qAb that binds with the antigen on the cell surface, favoring the formation of antigen-rAb complexes. Consequently, more cells are labeled as rAb(+), leading to diminished sensitivity in epitope evaluation. Besides, when the expression level of qAb is kept constant, an increase in the amount of antigen provides additional binding sites for rAb. In this case, even if cells expressing a qAb sharing the same epitope as the rAb can be labeled as rAb(+), leading to decreased sensitivity in the evaluation. In addition, the relative difference in affinity between the rAb and qAb also affects the direction of the reaction. A higher-affinity antibody tends to form more stable complexes with the antigen compared to one with a lower affinity. If the qAb exhibits a lower affinity for the antigen than the rAb, the latter can outcompete the former, making the antigen-qAb complex easier to dissociate.

As a result, this drives the formation of more antigen-rAb complexes and a higher proportion of cells being labeled as rAb(+), reducing the sensitivity. A comprehensive understanding of these parameters is crucial for optimizing and harnessing the full potential of this epitope evaluation system, which will enable more sensitive and accurate assessments of epitope similarity.

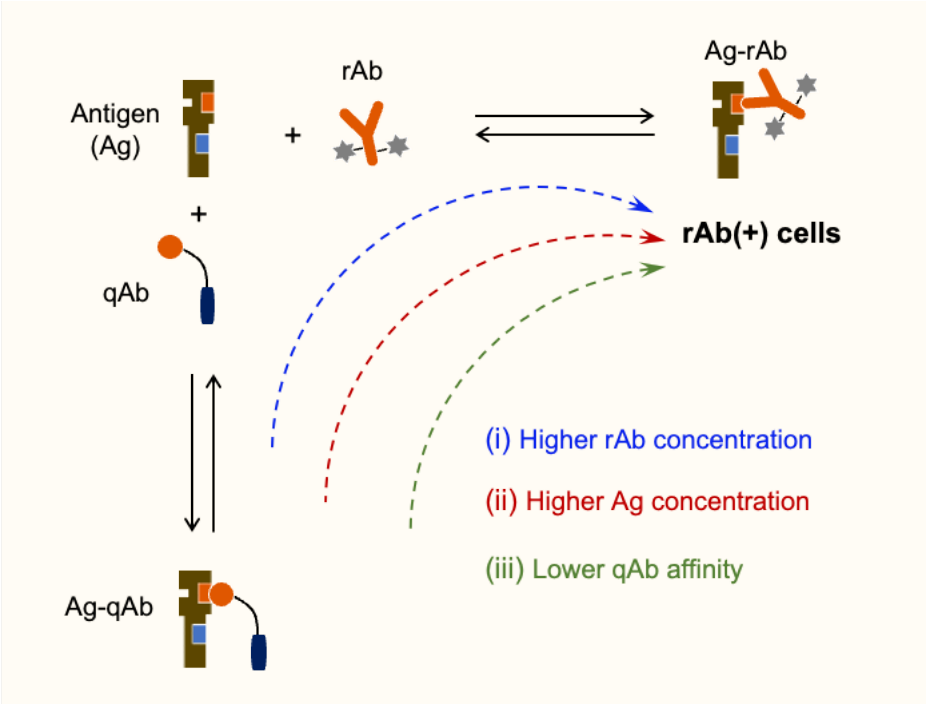



**Figure 3-5-1-1.** Illustration of competitive binding inhibition between qAb and rAb**.** On the cell surface, qAb binds with antigen to form the antigen-qAb complex. In the presence of the rAb, it binds to the empty epitope or competes for binding sites on antigen with the qAb, forming the antigen-rAb complex. In the context of a fixed amount of qAb: (i) an increase in the concentration of rAb shifts the reaction toward the formation of Ag-rAb; (ii) an increase in the concentration of antigen shifts the reaction toward the formation of Ag-rAb; (iii) a lower-affinity qAb is unable to outcompete the rAb, leading to the reaction shift towards the formation of Ag-rAb. The presence of the Ag-rAb complex on the cell surfaces allows the cells to be classified as the rAb(+) population. Ag, antigen.

### 3.5.2 Considerations for future optimization

In the developed system, when antigen expression is low, a high qAb/antigen molecular ratio enables the specific epitope on the antigen to be efficiently masked, preferring the collection of rAb(-) cell populations (Fig. 3-4-2-1). However, it is also important to ensure that the antigen expression level surpasses the minimal threshold to guarantee the sensitivity of this system.

Extremely low antigen expression in cells renders the binding of rAb undetectable, regardless of the epitope of qAb. These cells may be misclassified as rAb(-) population, resulting in false positive results. To eliminate this potential issue, it is recommended to establish an antigen-expressing cell line with a detectable antigen level before introducing the gene encoding qAbs.

Another potential factor affecting the performance of this system is the valence of the antibody, which refers to the number of antigen-binding sites on an antibody molecule. Higher valence allows an antibody to bind more antigen molecules simultaneously, thereby enhancing the strength of the antigen-antibody interaction and forming more stable complexes less susceptible to displacement by other antibodies. In this study, the qAbs utilized are monovalent scFvs, while the rAbs pertuzumab and trastuzumab are bivalent full IgGs that usually exhibit superior binding efficiency than scFvs. The difference in the valence between the qAb and rAb may partially account for the presence of rAb(+) population in cells expressing qAb with the same epitope as the rAb (Fig. 2-4-3-1). Using rAb and qAb with the same format to ensure consistent valence might further improve the sensitivity of this system. For instance, displaying qAb in the form of a full-length antibody with two binding sites to match the valence with that of the rAb. The feasibility of presenting full-length antibodies is supported by mammalian cell display techniques capable of presenting intact antibody structures[88].

The optimal concentration for achieving a desired sensitivity varies among different rAbs. When Pers-AF647 was used at various concentrations as the rAb, the highest percentage of rAb(-) cells exceeded 50%, while with Tras-AF647, the highest percentage of rAb(-) cells was approximately 35%. Besides, for Pert-AF647, a concentration of 10 nM maintained a comparable rAb(-) percentage to those at 0.1 nM and 1 nM; while for Tras-AF647 at 10 nM, the percentage of rAb(-) decreased. These disparities can be attributed to differing binding behaviors of rAbs, such as their binding kinetics, which influence their competition ability in the competitive binding assay. Thus, when using multiple rAbs, comprehensive optimization is necessary to determine the optimal working concentration for each rAb, which is crucial for enhancing the sensitivity of epitope evaluation.

# CHAPTER 4

# Epitope Binning-seq for simultaneous evaluation of multiple qAbs

## 4.1 Overview

In this chapter, a platform called Epitope Binning-seq, which integrates the epitope similarity evaluation system with DNA sequencing, was described. To validate the feasibility of this platform for parallel epitope binning, a model cell library displaying 14 different qAbs with known binding information, was established. Then, using either Pers-AF647 or Tras-AF647 as rAb to treat the qAb cell library, rAb(-) cells were isolated to analyze the DNA sequence of qAbs. By observing the specific enrichment of qAbs within particular rAb(-) populations, qAbs sharing the same epitope as the rAb were identified and categorized into respective epitope bins. The successful classification demonstrates the effectiveness of Epitope Binning-seq in assessing the epitope similarity and epitope binning of a large number of qAbs.

## 4.2 Introduction

### 4.2.1 Targeted NGS

NGS is a powerful technology that reads and determines the nucleotide sequences of DNA or RNA fragments in a massively parallel manner. This technique has revolutionized genomic studies and provided researchers with vast genetic information, significantly facilitating biological and medical research. For instance, whole-genome sequence (WGS) analysis via NGS facilitates the elucidation of the gene landscape and the identification of disease-relevant genes or genetic mutations, making NGS a useful tool for rare disease diagnosis[89]. Besides, genomic information on diseases obtained from NGS can also promote the development of rational treatment regimens, targeted therapy, and novel therapeutic biologics.

NGS can be generally classified into three categories based on the object being sequenced: WGS, whole exome sequencing (WES), and targeted sequencing[90]. Compared with WGS and WES, which respectively investigate the whole genome and whole exome, targeted NGS specifically focuses on particular regions of the genome containing genes of interest. This targeted approach provides higher sequencing depths for the genes of interest while outputting smaller data sets that are more manageable for interpretation. Importantly, the high target specificity of targeted NGS enhances its accuracy and sensitivity in identifying variants at low frequencies[91, 92]. Therefore, the targeted NGS is extensively employed in various cancer biology research, such as head and neck squamous cell carcinoma[93], and so on.

Targeted NGS (Fig. 4-2-1-1) usually starts with sample preparation, which involves the extraction and isolation of DNA or RNA from tissues, blood, or cells. An enrichment process for nucleic acids is required to gather the target DNA fragments for subsequent library preparation. The primary techniques for the target enrichment include hybrid capture and polymerase chain reaction (PCR) amplification[94, 95]. In the capture-based method, the DNA undergoes fragmentation followed by enrichment using DNA or RNA probes that are complementary to the target region[96]; while in the PCR amplicon-based method, the genes of interest are amplified using the target-specific primers[94]. The amplicon-based approach is broadly favored due to its simple workflow and cost-effectiveness.

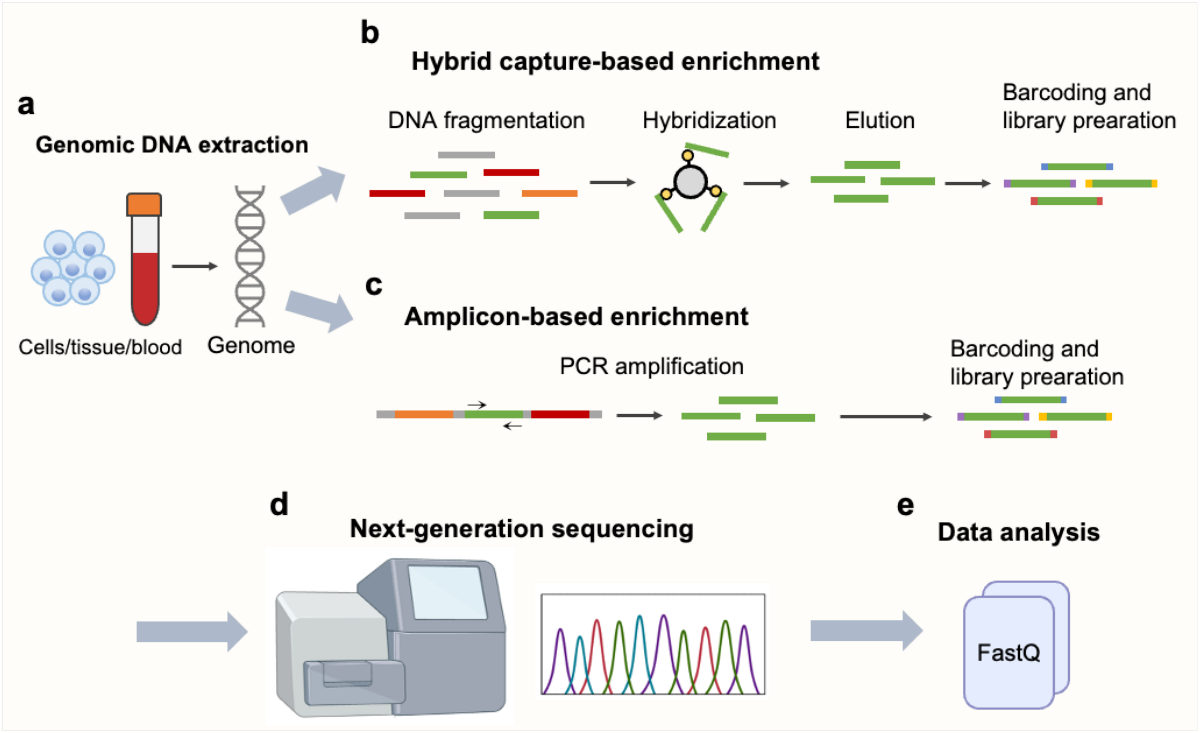



**Figure 4-2-1-1.** General workflow of targeted NGS. (a) Isolation of genomic DNA from collected cells, tissue, or blood. (b) Hybrid capture-based enrichment of specific genes of interest using beads coated with oligonucleotide probes. The captured DNA fragments are then eluted and barcoded for library preparation. (c) Amplicon-based enrichment of specific genes of interest. The target DNA fragments are amplified using primers specific to the target region, followed by barcoding and library preparation. (d) DNA sequencing via the NGS platform. (e) NGS data analysis to interpret the results.

In addition to its widespread application in genomics research, the targeted NGS is introduced

into the antibody screening procedure for identifying potential binders on the basis of sequence enrichment during selection cycles[97, 98]. It can be combined with various *in vitro* surface display technologies, such as phage[99, 100], yeast[101], and mammalian cell surface display[39, 102]. For the analysis of the DNA sequences for enriched clones using targeted NGS, it's important to prepare PCR amplicons with a length compatible with the maximum read length of the selected NGS platform. Depending on the library design, sequencing may focus on regions such as the CDR-H3[99], which are considered as the most contributing binding regions, the VH domain[103], the VL domain[104], or regions subjected to mutation or randomization for library generation. Compared to conventional ELISA-based binding screening, targeted NGS offers high-throughput capabilities, allowing for parallel screening of antibodies or peptide libraries containing thousands of candidates. This significantly accelerates the identification of target-specific binders in a more cost-effective manner. Furthermore, targeted NGS is also applied to epitope mapping to reveal the antibody binding sites within antigens[15, 105]. During the COVID-19 pandemic, detailed investigations of SARS-CoV-2 variant strains[106] and epitope identification[107] through NGS technology have rapidly advanced our understanding of COVID-19, facilitating rapid clinical diagnosis, treatment, and vaccine design.

### 4.2.2 Parallel epitope binning

In the combinatorial competitive binding assays for epitope binning, each rAb-qAb pair is typically assessed individually to determine their binding competition. While automation can streamline the process by utilizing SPR or BLI, evaluating each antibody pair separately still consumes considerable time and resources. A method capable of simultaneously evaluating multiple antibodies would be highly advantageous in expediting the process. Carterra LSA, a platform integrating flow microfluidics with high-throughput SPR, represents the cutting-edge technology for epitope binning. It notably accelerates the process of therapeutic antibody discovery, as demonstrated by the rapid identification of the approved therapeutic bamlanivimab against COVID-19 within a short time[108].

The strategy integrating the competition assay with the Luminex multiplex technique[35] has provided a solution for multiplexed epitope binning. In this approach, spectrally distinct beads were used to couple and differentiate each rAb, and the bead-rAb complexes were then reacted with the antigen and each qAb individually, allowing for binning of a panel of antibodies based on

their cross-binding competition to the antigen in a single multiplex assay. Similarly, using antigen-expressing cells with different fluorescent signatures in combination with FCM analysis enabled the determination of competitive binding profiles and the simultaneous binning of antibodies[36]. However, the capacity to evaluate qAbs using these methods heavily relies on the number of channels equipped within the flow cytometer used, which may pose scalability limitations. Besides, there are few methodologies other than those mentioned above available for simultaneous epitope binning, highlighting the urgent need for additional effective approaches in this area. In this chapter, a methodology called Epitope Binning-seq was described. It provides a simple and efficient strategy to achieve high-throughput epitope binning, offering higher scalability and simplicity compared with existing methods.

### 4.2.3 Research purpose of this chapter

The objective of this chapter is to assess the feasibility of the developed system, Epitope Binning-seq, for simultaneously binning multiple qAbs by integrating it with NGS analysis. I aimed to validate the effectiveness of this system by investigating whether the 14 qAbs used can be accurately classified into their respective bins.

## 4.3 Material and methods

### 4.3.1 Cell culture

The K562/HER2/scFv cells established in Chapter 3 were cultured in 10% FBS-RPMI-1640 (Invitrogen). The media was supplemented with penicillin (100 U/mL) and streptomycin (100 μg/mL) (Nacalai Tesque). Cells were maintained in a 37°C incubator with 5% $CO_2$ and subject to regular checks for mycoplasma contamination using a mycoplasma test kit (Lonza).

### 4.3.2 Cell sorting and genomic DNA (gDNA) extraction

To collectively evaluate epitope similarity for multiple qAbs, 14 different K562/HER2/scFv cells that were established in Chapter 3 were mixed to generate a library containing 14 qAbs with known affinities. Then, $1\times10^6$ library cells were incubated with 0.1 and 10 nM Pert-AF647 and Tras-AF647 for 1 h at 4°C. The cells that did not bind with Pert-AF647 and Tras-AF647 were isolated with a FACSAria III cell sorter (BD Biosciences). Antibody dilution was performed using FACS buffer (PBS containing 3% FBS).

The sorted cells were cultured until reaching approximately $5\times10^5$ cells. gDNA from both the library and the rAb(-) cells was then extracted using a GenElute Mammalian Genomic DNA Miniprep Kit (Sigma-Aldrich, St. Louis, MO, USA) according to the manufacturer's instructions.

### 4.3.3 Primer design and amplicon preparation

To amplify full-length scFv genes from gDNA, a primer pair (Forward-1 and Reverse-1) (Table. 4-4-2-1) was designed to specifically target the signal sequence (ss) located at the N-terminus of scFv and the linker located at the C-terminus of scFv.

To design primers for amplifying target DNA fragments containing the mutated region, the DNA sequence of scFv(Per), scFv(Tra), scFv(Dac), and scFv(Niv) were aligned to identify regions with high similarity among the four scFvs, which were then utilized as the primer binding regions for primer design. Different nucleotide bases at the same position of the primer binding regions as those of other scFvs were uniformed to the same nucleotide based on the principle of maximizing sequence similarity between each scFv and primer, resulting in a forward primer Forward-2 in Table. 4-4-2-2 and a reverse primer Reverse-2 in Table. 4-4-2-3. This primer pair enabled the amplification of partial scFv fragments ranging from 212 bp to 233 bp, with lengths of 230 bp for scFv(Per), 233 bp for scFv(Tra), 221 bp for scFv(Dac), and 212 bp for scFv(Niv).

For confirmation of the performance of designed primers, gDNA isolated from K562/HER2/scFv(Per), K562/HER2/scFv(Tra), K562/HER2/scFv(Dac), and K562/HER2/scFv(Niv) cells was used as first template to amplify four different full-length scFvs by PCR with the primer pair Forward-1 and Reverse-1. Subsequently, the PCR products were utilized as templates for amplifying the target region of each scFv using the primer pair Forward-2 and Reverse-2. Following primer confirmation, unique 6-bp barcodes were added at the 5'-end of the primer Forward-2 to generate barcoded primers through chemical synthesis (Eurofins Genomics, Tokyo, Japan).

Amplicons were prepared using the nested PCR method described above, utilizing gDNA from either the library or the sorted rAb(-) cells, along with the primer pair Forward-1 and Reverse-1, and barcoded primers. Amplicons with a size of around 200 bp were purified using AMPure XP magnetic beads (Beckman Coulter Life Sciences, Indianapolis, IN, USA) according to the manufacturer's instructions. Each purified amplicon was then mixed with an equal amount of nanograms for subsequent NGS analyses.

All PCRs were performed using a KOD FX kit (TOYOBO, Osaka, Japan), and the resulting

PCR products were analyzed by agarose electrophoresis.

### 4.3.4 NGS analyses

Targeted NGS was conducted on the NovaSeq 6000 platform by Macrogen Japan Corp. (Tokyo, Japan) with a read depth of 2,600,000 reads.

To eliminate low-quality reads from the raw data, Trimmomatic (version 0.39)[109] was utilized, with the MINLEN parameter set to 50 and SLIDINGWINDOW set to 20:20. For ensuring high sequencing accuracy, only reads that matched the primer sequences were retained for subsequent analysis. Then, reads were allocated to different samples based on unique barcodes. These reads were then translated into amino acid sequences using Biostrings (version 2.66.0-1)[110] and assigned to corresponding antibody species. The occupancy of each antibody in the samples was calculated by dividing the count of qAb by the total count of a given sample. Then, the enrichment ratio of the negative control Niv was set to 1, and the relative enrichment ratio of other qAbs compared to Niv was calculated. The qAb with a relative enrichment ratio greater than 3 was considered to be significantly enriched.

## 4.4 Results

### 4.4.1 Strategy for parallel epitope binning

For the simultaneous evaluation of epitope similarities for multiple antibodies, a qAb library is displayed on the surface of antigen-expressing cells (Fig. 4-4-1-1a). After the treatment of the cell library with individual rAb, the corresponding rAb(-) cell population is sorted for NGS analysis (Fig. 4-4-1-1b). The qAbs enriched within the sorted cells are then identified and grouped into respective rAb epitope bins (Fig. 4-4-1-1c). This platform is referred to as Epitope binning-seq. As the cDNA encoding the qAb is stably integrated into the genome of the cells, the identity of qAb can be determined through the analysis of the cellular DNA sequence. This advantageous feature of Epitope Binning-seq platform enables it to be a high-throughput methodology for simultaneous epitope binning of a large number of antibodies.

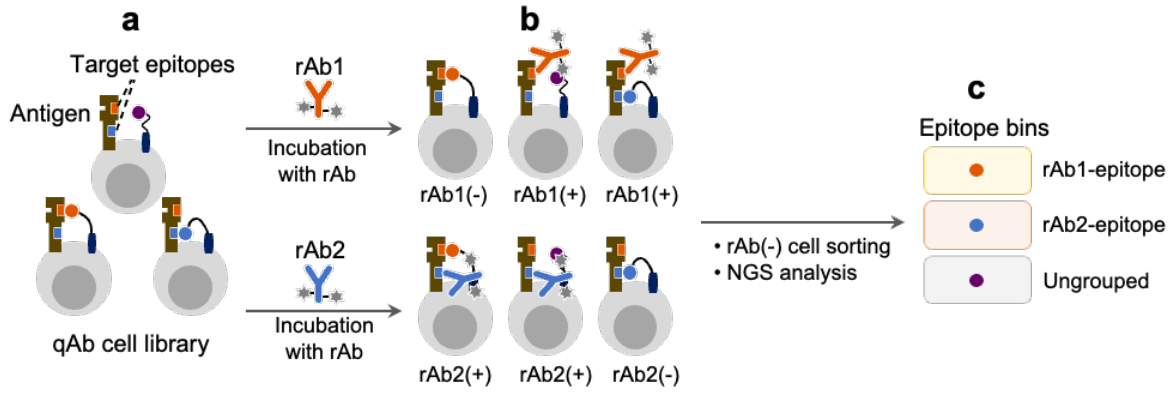



**Figure 4-4-1-1.** Schematic overview of the Epitope Binning-seq platform. **(a)** Presentation of the qAb library on the surface of antigen-expressing cells. (**b**) Treatment of the qAb cell library with individual rAbs, which bind to distinct epitopes on the antigen, followed by isolation of each rAb(-) cell population for NGS analysis to reveal qAb sequences. (**c**) Classification of qAbs enriched in the corresponding rAb(-) populations into different epitope bins, which comprise qAbs with epitopes similar to each rAb.

### 4.4.2 Parallel epitope binning for multiple qAbs through Epitope Binning-seq

For validating the utility of Epitope Binning-seq platform for simultaneous epitope binning of multiple antibodies, a cell library consisting of 14 distinct scFvs as qAbs was generated by mixing K562/HER2/scFv cells and employed as a model library (Fig. 4-4-2-1a). To comprehensively assess the performance of the platform, qAbs were presented in the library at varied abundances. Specifically, there were relatively higher presentations of Per, Per mutants, and Niv, while Tra, Tra mutants, and Dac were relatively less presented. The sfGFP-positive (sfGFP+) cells were sorted from the mixed cells, serving as the original library to determine the abundance of each variant (Fig. 4-4-2-1b, top). Meanwhile, rAb(-) cells were sorted out after incubating the mixed cells with specified rAb (Fig. 4-4-2-1b, bottom).

While scFv(Per), scFv(Tra), scFv(Dac), and scFv(Niv) share similarity in their amino acid sequences within the FRs (Fig. 2-4-2-2c), the degeneracy in the genetic code of these four different scFv scaffolds leads to diversity in the nucleotide bases (Fig.4-4-2-1c), which poses challenges for designing a pair of primer completely complementary to all scFvs. To address this issue, a strategy of nested PCR was adopted for amplicon preparation, which involved two consecutive rounds of PCR amplification using a set of outer primers to amplify full-length scFv, followed by using a pair of inner primers to specifically amplify the smaller and more specific region. Regions that

exhibit high sequence similarity among the four scFvs and enable the amplification of DNA fragments spanning the mutated region of scFv(Per) and scFv(Tra) were selected as the primer binding regions for designing inner primers (Fig. 4-4-2-1c). These resulting primers (Table. 4-4-2-1, 2 and 3) led to the amplification of the DNA fragments with a desired size of approximately 200 bp (Fig.4-4-2-1d, e), which were applied for amplicon preparation.

Table. 4-4-2-1 Primers for amplifying full-length scFvs

| Primer | Sequence |
|---|---|
| Forward-1 | 5'-CGGTCTCGAGATGAGGATATTTGCTGTCTTTATATTCATG-3' |
| Reverse-1 | 5'-CTTCCTCGCTAATCAGTTTCTGTTCTC-3' |

Table. 4-4-2-2 Sequence analysis of designed forward primer and scFv scaffolds*

| | Forward -2 | 5'-TGGGTAAGACAAGCTCCAGG-3' | Sequence similarity with primer |
|---|---|---|---|
| **Primer Binding regions** | scFv(Per) | TGGGTGAGACAGGCTCCAGG | 90.0% |
| | scFv(Tra) | TGGGTGAGACAGGCACCAGG | 85.0% |
| | scFv(Dac) | TGGGTACGCCAAGCTCCGGG | 85.0% |
| | scFv(Niv) | TGGGTAAGACAGGCACCCGG | 85.0% |

*Nucleotide bases highlighted in green indicate differences from the bases at the same positions within the primer sequence.

Table. 4-4-2-3 Sequence analysis of designed reverse primer and scFv scaffolds*

| | Reverse-2 | 5'-CCTTGTCCCCAATAGTC-3' | Sequence similarity with primer |
|---|---|---|---|
| **Primer Binding regions** | scFv(Per) | CCTTGACCCCAATAGTC | 94.1% |
| | scFv(Tra) | CCTTGTCCCCAGTAGTC | 94.1% |
| | scFv(Dac) | CCTTGACCCCAATAGTC | 94.1% |
| | scFv(Niv) | CCTTGGCCCCAATAGTC | 94.1% |

*Nucleotide bases highlighted in green indicate differences from the bases at the same positions within the primer sequence.

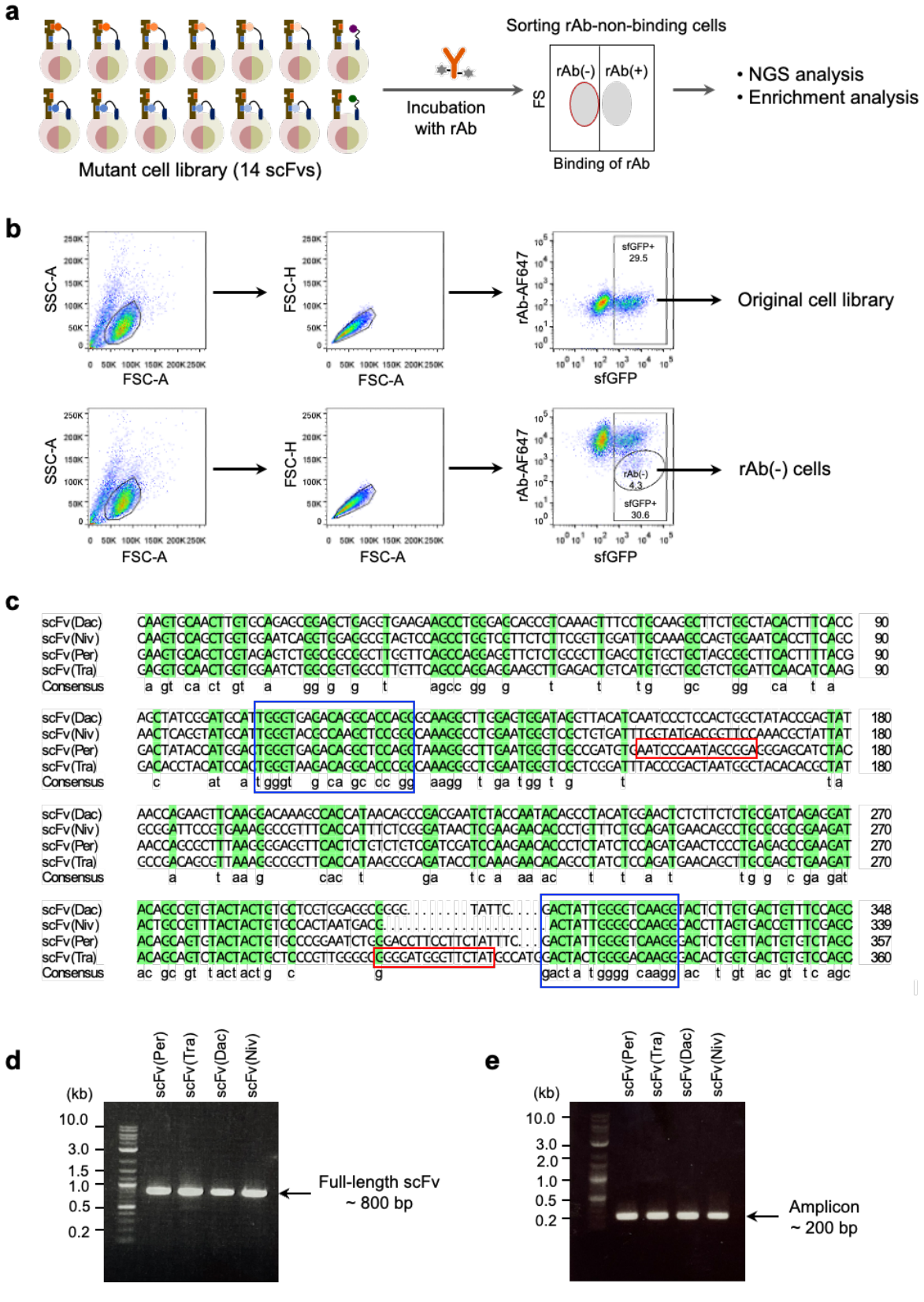

a
Mutant cell library (14 scFvs)
Incubation with rAb
Sorting rAb-non-binding cells
rAb(-)
rAb(+)
FS
Binding of rAb
• NGS analysis
• Enrichment analysis
b
SSC-A
FSC-A
FSC-H
rAb-AF647
sfGFP
sfGFP+ 29.5
Original cell library
rAb(-) 4.3
sfGFP+ 30.6
rAb(-) cells
c
scFv(Dac)
scFv(Niv)
scFv(Per)
scFv(Tra)
Consensus
d
(kb)
10.0
3.0
1.5
1.0
0.5
0.2
Full-length scFv ~ 800 bp
e
2.0
Amplicon ~ 200 bp

**Figure 4-4-2-1.** Amplicon preparation for NGS. **(a)** Schematic diagram of Epitope Binning-seq. A cell library is generated by mixing K562/HER2 cells expressing 14 different qAbs in the form of scFv, followed by incubation with an rAb and collection of the rAb-non-binding [rAb(-)] cells using a cell sorter. The genome is extracted from rAb(-) cells, and the occupancy of each scFv is analyzed by NGS. The relative enrichment ratio of each qAb is calculated based on the occupancy of each qAb sequence before and after sorting. (**b)** Representative gating schemes for sorting library cells and rAb(-) cells. (top) The sfGFP-positive (sfGFP+) cells were gated and sorted as the original library from the mixed cells. (bottom) Mixed cells were reacted with rAb, followed by gating sfGFP+ cells for sorting rAb(-) cells. (**c**) Sequence alignment among scFv(Per), scFv(Tra), scFv(Dac), and scFv(Niv). Consensus nucleotide bases among the four scFvs are highlighted in green. Highly similar regions marked by blue rectangles are used as primer binding regions for primer design. Regions marked by red rectangles represent mutation regions of scFv(Per) and scFv(Tra), which have to be included in the amplicon. (**d** and **e**) Confirmation of primer performance for nested PCR by agarose electrophoresis. (**d**) Full-length scFv amplified from gDNA of K562/HER2/scFv(Per), K562/HER2/scFv(Tra), K562/HER2/scFv(Dac), and K562/HER2/scFv(Niv) cells using Forward-1 and Reverse-1 primers. (**e**) Amplicons amplified from PCR products in **d** using Forward-2 and Reverse-2**.**

After NGS analysis of DNA from the sorted rAb(-) cells, the relative enrichment ratio, which reflects the enrichment degree of qAbs with epitopes similar to rAbs compared to the negative control Niv, was calculated using NGS read counts. Upon collecting the rAb(-) cell population treated with 0.1 nM Pert-AF647, there was a notable increase in the proportion of cells displaying Per and Per mutants (Fig. 4-4-2-2a), with all exhibiting a relative enrichment ratio greater than 3 (Fig. 4-4-2-2b), indicating the significant enrichment of desired qAbs through the Epitope Binning-seq platform. Conversely, the relative enrichment ratios of Tra and Tra mutants (which target different epitopes from pertuzumab), and Dac and Niv (which bind to different antigens other than HER2), were observed to be less than 3. Additionally, when increasing the concentration of Pert-AF647 to 10 nM, only Per and Per-P53A, which bound HER2 strongly, showed a relative enrichment ratio above 3 (Fig. 4-4-2-2c, d). In the treatment using Tras-AF647 as the rAb, Tra and its mutants could be distinguished from other qAbs using the relative enrichment ratio of 3 as the cutoff at the rAb concentration of 0.1 nM (Fig. 4-4-2-3a, b), while only slight enrichment was observed for qAbs with high binding affinity at the rAb concentration of 10 nM (Fig. 4-4-2-3c, d), as expected. In this study case, by identifying significantly enriched qAb clones upon the

utilization of each rAb at the concentration of 0.1 nM, the 14 qAbs were grouped into three distinct epitope bins: pertuzumab-, trastuzumab-, and ungrouped-epitope bins (Fig. 4-4-2-3e). Specifically, Per and Per mutants, as well as Tra and Tra mutants, were respectively classified into the pertuzumab- and trastuzumab-epitope groups, which was consistent with their origins. Meanwhile, qAbs Dac, Niv, which exhibited no binding to the HER2, did not belong to either of the two rAb-epitope bins. These results demonstrate the effectiveness of Epitope Binning-seq in classifying antibodies with similar epitopes as rAbs across a wide variety of antibodies. Moreover, under certain conditions, the varying responses to different rAb concentrations could also roughly reflect the relative binding affinity.

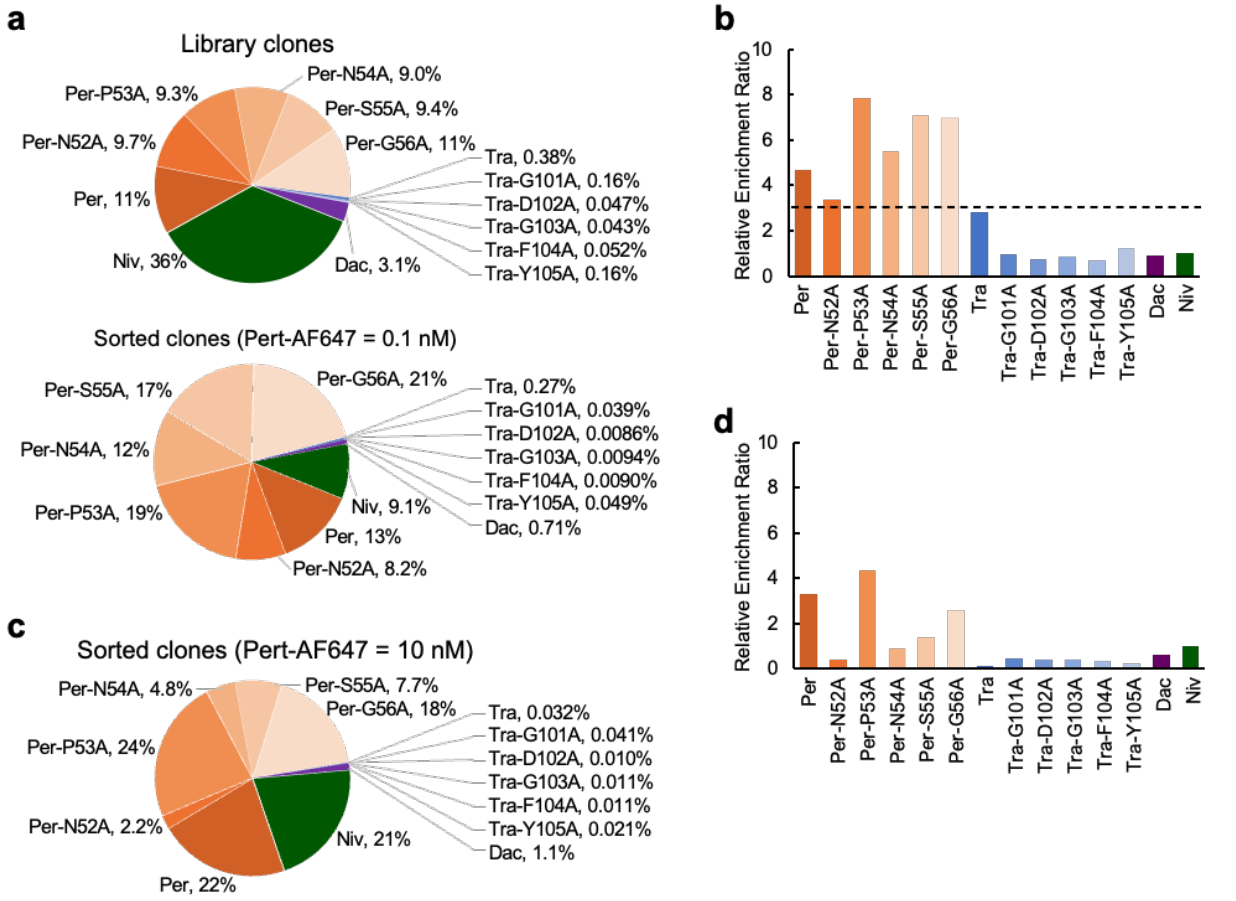



**Figure 4-4-2-2.** Epitope Binning-seq for evaluating multiple qAbs using Pert-AF647 as the rAb. (**a**) Occupancy of each qAb before and after sorting using Pert-AF647 at the concentration of 0.1 nM. The cell library was incubated with 0.1 nM Pert-AF647 and the rAb(-) cell population was isolated for NGS analysis. (**b**) Relative enrichment ratio of each qAb calculated from the occupancy in **a**. (**c**) Occupancy of each qAb after sorting using Pert-AF647 at the concentration of 10 nM. The same cell library in **a** was incubated with 10 nM Pert-AF647 and the rAb(-) cell population was isolated

for NGS analysis. (**d**) Relative enrichment ratio of each qAb calculated from the occupancy in **a** (library cells) and **c**.

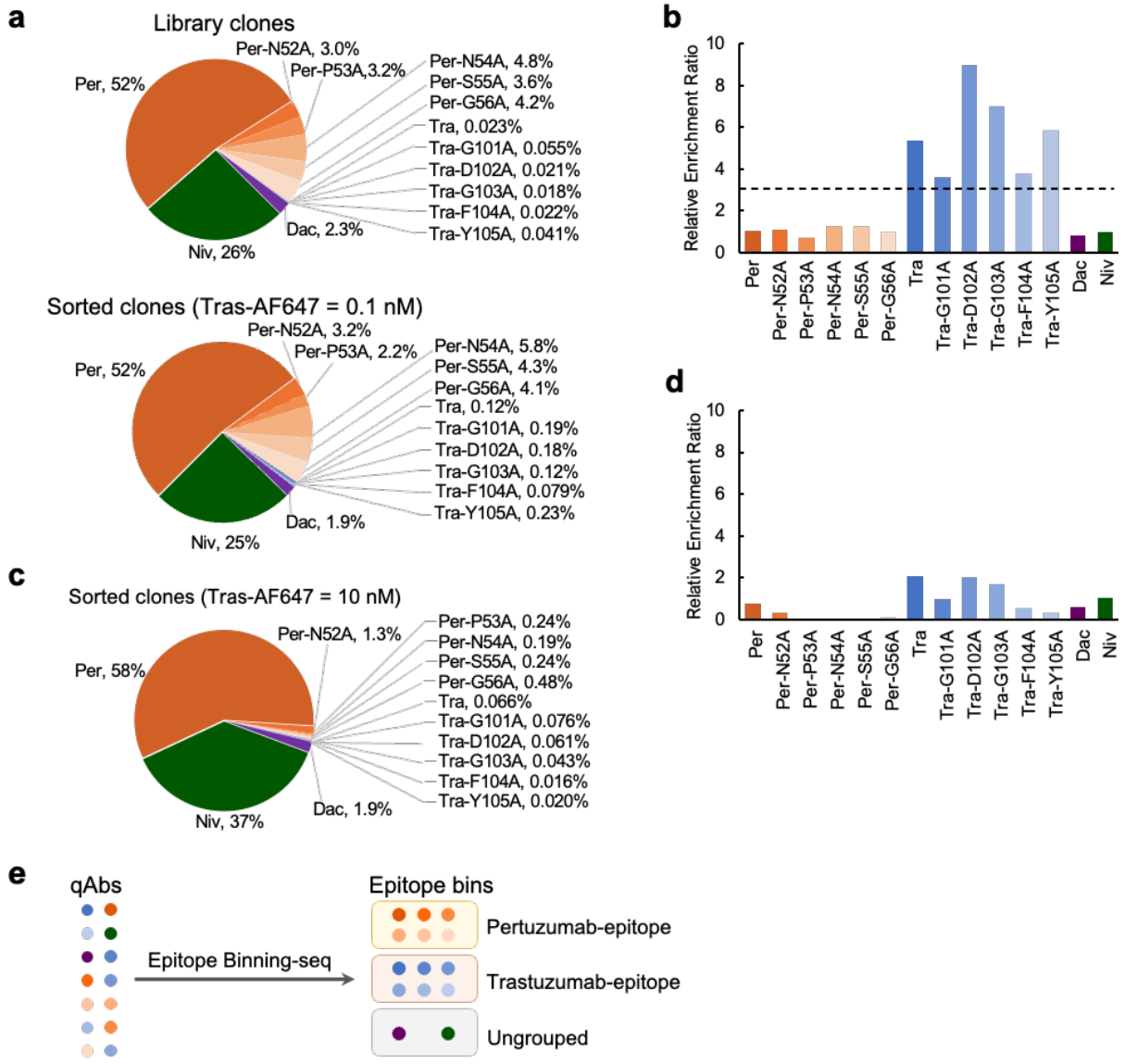



**Figure 4-4-2-3.** Epitope Binning-seq for evaluating various qAbs using Tras-AF647 as the rAb. **(a)** Occupancy of each qAb before and after sorting using Tras-AF647 at the concentration of 0.1 nM. The cell library was incubated with 0.1 nM Tras-AF647 and the rAb(-) cell population was isolated for NGS analysis. (**b**) Relative enrichment ratio of each qAb calculated from the occupancy in **a**. (**c**) Occupancy of each qAb after sorting using Tras-AF647 at the concentration of 10 nM. The same cell library in **a** was incubated with 10 nM Tras-AF647 and the rAb(-) cell population was isolated for NGS analysis. (**d**) Relative enrichment ratio of each qAb calculated from the occupancy in **a** (library cells) and **c**. (**e**) Parallel qAb classification via Epitope Binning-seq. The 14 qAbs were

classified into three bins, pertuzumab-epitope, trastuzumab-epitope, and ungrouped.

## 4.5 Discussion

Comparing the binding epitopes between approved and newly developed antibodies is crucial for developing antibody-based drugs with novel mechanisms of action. Meanwhile, epitope-directed antibody discovery utilizes available information to develop drugs with desired mechanisms and therapeutic effects, thereby notably reducing the time and risks involved in drug development. Classifying antibody candidates obtained from the preliminary screening based on their binding epitopes, a process known as epitope binning, enhances the possibility of selecting even more promising candidates for further investigation. However, existing approaches developed for epitope binning typically involve the production and purification of individual antibodies, which presents a bottleneck in simultaneously evaluating numerous antibodies. In this chapter, for the first time, I have developed an epitope binning platform, Epitope Binning-seq, an effective tool capable of simultaneously assessing numerous genetically encoded antibodies. The capacity to evaluate qAbs using Epitope Binning-seq platform depends on factors such as the size of the library, the number of cells to be sorted, and the volume of NGS reads. Nonetheless, it is feasible to simultaneously evaluate up to 10 million distinct qAbs using mammalian display in a laboratory-level experiment. Therefore, this platform can be applied for epitope binning of both natural and artificial antibody libraries, including polyclonal antibody pools resulting from animal immunization and antibody libraries created by random mutation.

Epitope Binning-seq effectively segregated epitope-similar qAbs with high reliability and sensitivity. In the analysis using Pert-AF647 as the rAb, the negative control qAb Niv, despite its initial high abundance (36%), showed no enrichment after sorting (Fig. 4-4-2-2a-d). In the analysis utilizing Tras-AF647, Tra-G103A, which initially comprised only 0.018%, was significantly enriched (Fig. 4-4-2-3a, b), demonstrating the capability to identify and evaluate a single clone from a pool containing at least 5500 species. With Epitope Binning-seq, 14 qAbs were accurately grouped into their corresponding epitope bins, demonstrating the accuracy of this platform in preliminary epitope analysis.

Epitope Binning-seq offers several distinct advantages comparable to Carterra LSA, a leading platform for epitope characterization. While the Carterra LSA can analyze combinations of 384 qAbs and 384 rAbs in a single run, Epitope Binning-seq holds huge potential to simultaneously

assess combinations of over 10 million qAbs and multiple rAbs. Furthermore, the Carterra LSA requires the production of individual qAbs for analysis, while Epitope Binning-seq eliminates this need, providing a significant advantage for large-scale binning analysis. These characteristic differences make Epitope Binning-seq particularly suitable for analyzing numerous qAbs in the early stages of drug discovery, while the Carterra LSA is more effective for examining epitope similarity among qAbs in later stages, especially after narrowing down the pool of candidate qAbs to some extent.

Theoretically, within Epitope Binning-seq, high-affinity qAbs targeting the similar epitope as the rAb should show a higher enrichment in the rAb(-) population, especially in a relatively high concentration of rAb. However, the relative enrichment ratios of some qAbs did not show a strong correlation with their affinity magnitude as expected. For instance, the qAb Per, which has the highest binding affinity to HER2 among six scFvs from pertuzumab, did not exhibit the anticipated highest relative enrichment ratio within the Pert-AF647 (-) cell population (Fig. 4-4-2-2b, d). Alongside the varied initial abundance of each qAb in the original library, potential reasons for these unexpected results include PCR errors and NGS errors (such as base-calling errors, sequencing artifacts, and biases arising from technical variability or experimental noise). To address these issues, optimizing PCR conditions and performing replicate sequencing that involves repeated sequencing of the same sample using the same sequencing platform and protocol, might be useful.

Moreover, the successful isolation of epitope-similar qAbs from the libraries (Fig. 4-4-2-2, Fig. 4-4-2-3a-d) demonstrates the utility of Epitope Binning-seq in epitope-specific antibody screening. While there are various epitope-directed selection strategies, such as animal immunizations with epitope-containing polypeptides[111, 112] or recombinant fusion proteins[113, 114], the removal of off-target binders from the library using epitope-masked antigen mutants[115, 116], and the incorporation of noncanonical photoreactive amino acids near the desired epitope to capture specific binders[117], these methods often require engineered antigens, which may compromise native conformations of target antigens. In contrast, the approach I described allows for epitope-directed antibody identification using surface-displayed antigens that retain their native conformation, without the need for introducing any mutations. This is particularly advantageous for identifying antibodies targeting conformational epitopes.

Intriguingly, Epitope Binning-seq showed that high-affinity qAbs could be specifically

isolated even under high rAb concentrations (Fig. 4-4-2-2d, Fig. 4-4-2-3d). In competitive binding assays, the balance between the quantities of qAb and rAb determines the predominant interaction, deciding whether qAb-antigen or rAb-antigen complexes are formed. Under conditions of low rAb concentrations, qAbs preferentially bind to antigens on the cell surface because of the limited presence of competing rAbs targeting the same epitope. Conversely, high concentrations of rAb exert intense competitive pressure on antigen-bound qAbs, leading to their displacement. The qAbs with high affinity are capable of withstanding this pressure and maintaining binding to the antigen, whereas those with low-affinity surrender to the pressure and are replaced by the rAb. As a result, when subjected to high concentrations of rAb, high-affinity qAbs were preferentially enriched, while a low concentration of rAb enabled the identification of low-affinity qAbs as well. This property, which depends on the concentration of rAb, suggests that this system performs epitope binning while also considering the antigen-binding affinity.

Epitope Binning-seq holds promise for application in antibody affinity maturation. In this scenario, parental and mutant antibodies would act as rAb and qAb, respectively, and Epitope Binning-seq would identify qAbs with superior binding affinity compared to the rAb. Previously, a competitive binding-based affinity maturation method has been studied, in which mutant antibodies were secreted from the engineered mammalian cells co-displaying the chimera antigen and parental antibody[53]. The secreted mutant antibodies bound to the antigen, competitively inhibiting the binding of the engineered parental antibody, thereby preventing the dimerization of chimera antigen and activation of cell death signaling, ultimately enabling the survival of cells that secreted affinity-matured antibodies. Although this method successfully identified affinity-matured anti-HER2 antibodies, it does have some potential weaknesses, such as a high false positive rate resulting from secreted mutant antibodies binding to antigens expressed by nearby cells, as well as limitations on the applicable antigens that cannot endure complex engineering. In contrast, Epitope Binning-seq platform allows for the presentation of mutant antibodies with minimal interference from neighboring cells and can be applied to a wide range of native antigens, including single-pass membrane proteins, G-protein coupled receptors, and ion channel proteins. In other words, Epitope Binning-seq holds the potential to serve as a more precise and versatile affinity maturation method than existing techniques.

# CHAPTER 5

# Conclusions and future prospects

Throughout the extensive process of therapeutic antibody discovery, the most primary and straightforward strategy for candidate selection relies on the strength of antigen-antibody binding. While this strategy yields numerous target-binding antibodies, mere binding to the antigen does not guarantee an antibody the ability to mediate the function of the target antigen, let alone disease treatment. Targeting a physiologically functional epitope is a prerequisite for the development of effective therapeutic drugs. Epitope characterization represents one of the most important aspects during antibody drug development, it not only aids in elucidating molecular MOA but also facilitates the optimization of therapeutic strategies. Epitope binning, which focuses on epitope similarity among pairs of antibodies rather than exact epitope residues, offers a cost-effective option for epitope characterization during the early stages of antibody discovery. The fundamental experimental approach for epitope binning involves competitive binding assays, wherein pairs of antibodies are assessed for their ability to inhibit each other's binding to the same antigen. The use of recombinantly expressed antibodies, whether in the form of culture supernatant or purified antibodies, poses a challenge when testing large-scale candidates due to the requirement for individual construction of recombinant plasmids, expression, and potential purification for each antibody. In addition, in most available methods, each rAb-qAb pair is tested separately, leading to labor-intensive processes and complicated data processing when evaluating a larger number of antibodies. The purpose of this study is to develop an epitope binning approach capable of assessing multiple antibodies simultaneously while circumventing the need for individual antibody production. To achieve this objective, I repurposed MAGPIE, originally designed as a mammalian display system co-presenting antigen and a peptide library for screening target-binding peptides through competitive bindings, for the epitope similarity evaluation and epitope binning. The platform, now designated as Epitope Binning-seq, successfully assessed epitope similarities for multiple antibodies simultaneously and grouped them into respective epitope bins without requiring separate antibody production.

**Chapter 2** investigates the feasibility of the antigen/antibody co-display system to assess epitope similarity, both for a single rAb-qAb pair and multiple rAb-qAb pairs. As a proof of concept, HER2 and HER2-specific antibodies pertuzumab and trastuzumab were selected as the model antigen and rAbs, respectively. The scFvs, derived from either rAbs or non-HER2 binding antibodies, were presented on the surface of HER2-expressing cells as qAbs, ensuring that qAbs targeting distinct epitopes could be investigated. FCM analysis of cells co-expressing HER2/qAb

treated with labeled-rAb revealed the presence of the rAb(-) population in cells expressing a qAb sharing the same epitope as the rAb, suggesting the feasibility of evaluating epitope similarity between qAb and rAb using the presence of rAb(-) population as an indicator. Furthermore, two different rAbs were simultaneously applied to treat cell mixtures containing various scFvs, the observed respective rAb(-) populations demonstrated the feasibility of epitope similarity assessment for multiple qAb-rAb pairs using dual rAbs with distinct epitopes simultaneously.

**Chapter 3** describes the investigation of factors influencing the sensitivity of epitope assessment, including rAb concentrations, antigen density, and qAb affinity. First, treatment of cells presenting antigen/qAb with rAbs at varying concentrations revealed that high rAb concentrations beyond a specific range are unfavorable for the efficient separation of rAb(-) cells. Subsequently, the investigation into the reason for the presence of the rAb(+) population in cells expressing a qAb targeting the same epitope as the rA showed that rAb(-) cells expressed a higher level of qAb but a lower level of HER2 than those of rAb(+) cells, which inspired me to explore the effect of relative qAb/HER2 ratio. By treating K562/scFv/HER2 cells with different HER2 expression levels with rAb and analyzing the proportion of rAb(-) cells, I found that cells with a higher qAb/HER2 ratio (HER2 low expression) were more advantageous for epitope similarity evaluation. Additionally, to investigate the effect of qAb affinity on the sensitivity of epitope similarity evaluation, I performed MD simulations to identify critical residues involved in the interactions of HER2 with pertuzumab and trastuzumab to varying degrees. I then conducted single alanine substitutions to generate qAb variants with diverse affinities to HER2 while remaining the same or similar epitope as the parental scFv. The resultant variants were individually displayed on antigen-expressing cells, followed by the rAb treatment and FCM analysis. It was observed that both low-affinity and high-affinity qAbs can be efficiently evaluated for their epitopes at relatively low rAb concentrations, while the evaluation sensitivity decreased as the qAb affinity decreased at a relatively high rAb concentration.

**Chapter 4** describes Epitope binning-seq, an efficient platform for epitope binning that enables simultaneous assessment of epitope similarity for multiple genetically encoded qAbs. To demonstrate the effectiveness of this platform, a library of 14 qAbs was displayed on the surface of antigen-expressing cells. The resultant cell libraries of qAbs were treated with rAbs, and the rAb(-) cells were sorted for NGS analysis to identify qAbs sharing epitopes similar to those of rAb. Using relative enrichment ratios calculated from NGS read counts, I successfully evaluated epitope

similarities between rAbs and these 14 qAbs at various abundances in libraries and classified these qAbs into respective epitope bins based on their enrichment within each rAb group.

In conclusion, this study has provided a proof-of-concept for the DNA sequencing-based epitope binning platform. The successful implementation of simultaneous epitope similarity evaluation of multiple qAbs and effective classification of epitope bins without involving the individual production for each antibody demonstrates the effectiveness of Epitope Binning-seq as an approach for epitope characterization of numerous qAbs in both natural and artificial libraries. Epitope Binning-seq would be applicable for a broad range of antibodies and antigens, especially membrane proteins with complex structure and dynamic conformation changes. Last but not least, the versatility of the platform extends far beyond epitope binning. Its potential applications in antibody identification, such as epitope-directed selection and affinity maturation, hold great promise for significantly facilitating antibody drug discovery in the future.

Overall, the methodology described in this study demonstrates high reliability, scalability, and versatility, and would lay a strong foundation for further exploration in epitope analysis and antibody drug development and engineering. Continued efforts to enhance the performances of currently existing methodologies, including Epitope Binning-seq described here, and to develop additional novel strategies will facilitate more informed epitope analyses and candidate selection processes. This has the potential to substantially reduce drug development timelines and costs, ultimately benefiting a greater number of patients.

# Reference


1. Lu, R.M., et al., *Development of therapeutic antibodies for the treatment of diseases.* J Biomed Sci, 2020. **27**(1): p. 1.
2. Chiu, M.L., D.R. Goulet, A. Teplyakov, and G.L. Gilliland, *Antibody Structure and Function: The Basis for Engineering Therapeutics.* Antibodies (Basel), 2019. **8**(4) : p. 55.
3. Chen, X., J.L. Zaro, and W.C. Shen, *Fusion protein linkers: property, design and functionality.* Adv Drug Deliv Rev, 2013. **65**(10): p. 1357-1369.
4. Rodríguez-Nava, C., et al., *Mechanisms of Action and Limitations of Monoclonal Antibodies and Single Chain Fragment Variable (scFv) in the Treatment of Cancer.* Biomedicines, 2023. **11**(6): p. 1610.
5. Chidharla, A., M. Parsi, and A. Kasi, *Cetuximab [Updated 2023 May 1]*. 2023: In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan.
6. Sulica, A., R. Morel, D. Metes, and R.B. Herberman, *Ig-binding Receptors on Human NK Cells as Effector and Regulatory Surface Molecules.* International Reviews of Immunology, 2001. **20**(3-4): p. 371-414.
7. Topalian, S.L., J.M. Taube, R.A. Anders, and D.M. Pardoll, *Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy.* Nat Rev Cancer, 2016. **16**(5): p. 275-287.
8. Klein, C., et al., *Epitope interactions of monoclonal antibodies targeting CD20 and their relationship to functional properties.* MAbs, 2013. **5**(1): p. 22-33.
9. Hou, Q., et al., *SeRenDIP-CE: sequence-based interface prediction for conformational epitopes.* Bioinformatics, 2021. **37**(20): p. 3421-3427.
10. Shim, H., *One target, different effects: a comparison of distinct therapeutic antibodies against the same targets.* Experimental & Molecular Medicine, 2011. **43**(10): p. 539-549.
11. Klein, C., et al., *Epitope interactions of monoclonal antibodies targeting CD20 and their relationship to functional properties.* MAbs, 2013. **5**(1): p. 22-33.
12. Scheuer, W., et al., *Strongly enhanced antitumor activity of trastuzumab and pertuzumab combination treatment on HER2-positive human xenograft tumor models.* Cancer Res, 2009. **69**(24): p. 9330-9336.
13. Hansen, L.B., S. Buus, and C. Schafer-Nielsen, *Identification and Mapping of Linear Antibody Epitopes in Human Serum Albumin Using High-Density Peptide Arrays.* PLOS ONE, 2013. **8**(7): p. e68902.
14. Jin, Z., et al., *Identification of a potential neutralizing linear epitope of hemagglutinin-neuraminidase in Newcastle disease virus.* Virol J, 2021. **18**(1): p. 8.
15. Qi, H., et al., *Antibody Binding Epitope Mapping (AbMap) of Hundred Antibodies in a Single Run.* Mol Cell Proteomics, 2021. **20**: p. 100059.

16. Guo, J.Y., et al., *Identification of COVID-19 B-cell epitopes with phage-displayed peptide library.* J Biomed Sci, 2021. **28**(1): p. 43.
17. Padavattan, S., et al., *High-affinity IgE recognition of a conformational epitope of the major respiratory allergen Phl p 2 as revealed by X-ray crystallography.* J Immunol, 2009. **182**(4): p. 2141-2151.
18. Bardelli, M., et al., *Epitope mapping by solution NMR spectroscopy.* J Mol Recognit, 2015. **28**(6): p. 393-400.
19. Di Muzio, M., et al., *Hydrogen/deuterium exchange memory NMR reveals structural epitopes involved in IgE cross-reactivity of allergenic lipid transfer proteins.* J Biol Chem, 2020. **295**(51): p. 17398-17410.
20. Raghavan, S.S.R., et al., *Cryo-EM reveals the conformational epitope of human monoclonal antibody PAM1.4 broadly reacting with polymorphic malarial protein VAR2CSA.* PLoS Pathog, 2022. **18**(11): p. e1010924.
21. Najar, T.A., et al., *Mapping Protein Binding Sites and Conformational Epitopes Using Cysteine Labeling and Yeast Surface Display.* Structure, 2017. **25**(3): p. 395-406.
22. Datta, R., et al., *A facile method of mapping HIV-1 neutralizing epitopes using chemically masked cysteines and deep sequencing.* Proc Natl Acad Sci U S A, 2020. **117**(47): p. 29584-29594.
23. Kowalsky, C.A., et al., *Rapid fine conformational epitope mapping using comprehensive mutagenesis and deep sequencing.* J Biol Chem, 2015. **290**(44): p. 26457-26470.
24. Brooks, B.D., A.R. Miles, and Y.N. Abdiche, *High-throughput epitope binning of therapeutic monoclonal antibodies: why you need to bin the fridge.* Drug Discov Today, 2014. **19**(8): p. 1040-1044.
25. Raybould, M.I.J., A.R. Rees, and C.M. Deane, *Current strategies for detecting functional convergence across B-cell receptor repertoires.* MAbs, 2021. **13**(1): p. 1996732.
26. Wong, W.K., et al., *Ab-Ligity: identifying sequence-dissimilar antibodies that bind to the same epitope.* MAbs, 2021. **13**(1): p. 1873478.
27. Ghanbarpour, A., M. Jiang, D. Foster, and Q. Chai, *Structure-free antibody paratope similarity prediction for in silico epitope binning via protein language models.* iScience, 2023. **26**(2): p. 106036.
28. Mahita, J., et al., *Computational epitope binning reveals functional equivalence of sequence-divergent paratopes.* Comput Struct Biotechnol J, 2022. **20**: p. 2169-2180.
29. Brooks, B.D., *The importance of epitope binning for biological drug discovery.* Curr Drug Discov Technol, 2014. **11**(2): p. 109-112.
30. Kwak, J.-W. and C.-S. Yoon, *A convenient method for epitope competition analysis of two monoclonal antibodies for their antigen binding.* J Immunol Methods, 1996. **191**(1): p. 49-54.
31. Abdiche, Y., D. Malashock, A. Pinkerton, and J. Pons, *Determining kinetics and affinities of protein*

*interactions using a parallel real-time label-free biosensor, the Octet.* Anal Biochem, 2008. **377**(2): p. 209-217.

32. Abdiche, Y.N., et al., *Antibodies Targeting Closely Adjacent or Minimally Overlapping Epitopes Can Displace One Another.* PLoS One, 2017. **12**(1): p. e0169535.
33. Abdiche, Y.N., et al., *High-throughput epitope binning assays on label-free array-based biosensors can yield exquisite epitope discrimination that facilitates the selection of monoclonal antibodies with functional activity.* PLoS One, 2014. **9**(3): p. e92451.
34. Abdiche, Y.N., D.S. Malashock, A. Pinkerton, and J. Pons, *Exploring blocking assays using Octet, ProteOn, and Biacore biosensors.* Anal Biochem, 2009. **386**(2): p. 172-180.
35. Jia, X.C., et al., *A novel method of Multiplexed Competitive Antibody Binning for the characterization of monoclonal antibodies.* J Immunol Methods, 2004. **288**(1-2): p. 91-98.
36. Chan, B.M., et al., *Flow Cytometry-Based Epitope Binning Using Competitive Binding Profiles for the Characterization of Monoclonal Antibodies against Cellular and Soluble Protein Targets.* SLAS Discov, 2018. **23**(7): p. 613-623.
37. King, D.J., P.M. Bowers, M.R. Kehry, and R.A. Horlick, *Mammalian cell display and somatic hypermutation in vitro for human antibody discovery.* Curr Drug Discov Technol, 2014. **11**(1): p. 56-64.
38. Ho, M. and I. Pastan, *Mammalian cell display for antibody engineering.* Methods Mol Biol, 2009. **525**: p. 337-352.
39. Crook, Z.R., et al., *Mammalian display screening of diverse cystine-dense peptides for difficult to drug targets.* Nature Communications, 2017. **8**(1): p. 2244.
40. Taube, R., et al., *Lentivirus display: stable expression of human antibodies on the surface of human cells and virus particles.* PLoS One, 2008. **3**(9): p. e3181.
41. See, K., et al., *Antibody-guided design and identification of CD25-binding small antibody mimetics using mammalian cell surface display.* Scientific Reports, 2021. **11**(1): p. 22098.
42. Qin, C.-F. and G.-C. Li, *Mammalian cell display technology coupling with AID induced SHM in vitro: An ideal approach to the production of therapeutic antibodies.* International Immunopharmacology, 2014. **23**(2): p. 380-386.
43. Wang, Y., et al., *Construction and Applications of Mammalian Cell-Based DNA-Encoded Peptide/Protein Libraries.* ACS Synthetic Biology, 2023. **12**(7): p. 1874-1888.
44. Gray, D., *Overview of protein expression by mammalian cells.* Curr Protoc Protein Sci, 2001. **Chapter 5**(1): p. Unit5.9.
45. Krohl, P.J., et al., *Discovery of antibodies targeting multipass transmembrane proteins using a suspension cell-based evolutionary approach.* Cell Rep Methods, 2023. **3**(3): p. 100429.

46. Yoon, H., et al., *An efficient strategy for cell-based antibody library selection using an integrated vector system.* BMC Biotechnology, 2012. **12**(1): p. 62.
47. Moore, S.J. and J.R. Cochran, *Engineering knottins as novel binding agents.* Methods Enzymol, 2012. **503**: p. 223-251.
48. Jin, Y.-j., et al., *A novel and effective approach to generate germline-like monoclonal antibodies by integration of phage and mammalian cell display platforms.* Acta Pharmacologica Sinica, 2022. **43**(4): p. 954-962.
49. Ehling, R.A., et al., *SARS-CoV-2 reactive and neutralizing antibodies discovered by single-cell sequencing of plasma cells and mammalian display.* Cell Rep, 2022. **38**(3): p. 110242.
50. Beerli, R.R., et al., *Isolation of human monoclonal antibodies by mammalian cell display.* Proceedings of the National Academy of Sciences, 2008. **105**(38): p. 14336-14341.
51. Bowers, P.M., et al., *Coupling mammalian cell surface display with somatic hypermutation for the discovery and maturation of human antibodies.* Proc Natl Acad Sci U S A, 2011. **108**(51): p. 20455-20460.
52. Robertson, N., et al., *Development of a novel mammalian display system for selection of antibodies against membrane proteins.* J Biol Chem, 2020. **295**(52): p. 18436-18448.
53. Eguchi, A., et al., *An epitope-directed antibody affinity maturation system utilizing mammalian cell survival as readout.* Biotechnol Bioeng, 2019. **116**(7): p. 1742-1751.
54. Gaa, R., et al., *Mammalian display to secretion switchable libraries for antibody preselection and high throughput functional screening.* MAbs, 2023. **15**(1): p. 2251190.
55. Gaa, R., et al., *An integrated mammalian library approach for optimization and enhanced microfluidics-assisted antibody hit discovery.* Artif Cells Nanomed Biotechnol, 2023. **51**(1): p. 74-82.
56. Sergina, N.V. and M.M. Moasser, *The HER family and cancer: emerging molecular mechanisms and therapeutic targets.* Trends Mol Med, 2007. **13**(12): p. 527-534.
57. Burgess, A.W., et al., *An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors.* Mol Cell, 2003. **12**(3): p. 541-552.
58. Garrett, T.P., et al., *The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors.* Mol Cell, 2003. **11**(2): p. 495-505.
59. Iqbal, N. and N. Iqbal, *Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications.* Mol Biol Int, 2014. **2014**: p. 852748.
60. Mitri, Z., T. Constantine, and R. O'Regan, *The HER2 Receptor in Breast Cancer: Pathophysiology, Clinical Use, and New Advances in Therapy.* Chemother Res Pract, 2012. **2012**: p. 743193.
61. Ruschoff, J., et al., *HER2 testing in gastric cancer: a practical approach.* Mod Pathol, 2012. **25**(5): p. 637-650.

62. Smyth, E.C., et al., *Oesophageal cancer.* Nat Rev Dis Primers, 2017. **3**: p. 17048.
63. Hu, X., et al., *Advances in the Application of Radionuclide-Labeled HER2 Affibody for the Diagnosis and Treatment of Ovarian Cancer.* Front Oncol, 2022. **12**: p. 917439.
64. Buza, N., D.M. Roque, and A.D. Santin, *HER2/neu in Endometrial Cancer: A Promising Therapeutic Target With Diagnostic Challenges.* Arch Pathol Lab Med, 2014. **138**(3): p. 343-350.
65. Wieduwilt, M.J. and M.M. Moasser, *The epidermal growth factor receptor family: biology driving targeted therapeutics.* Cell Mol Life Sci, 2008. **65**(10): p. 1566-1584.
66. Cho, H.S., et al., *Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab.* Nature, 2003. **421**(6924): p. 756-760.
67. Hao, Y., et al., *Cryo-EM Structure of HER2-trastuzumab-pertuzumab complex.* PLoS One, 2019. **14**(5): p. e0216095.
68. Maadi, H., et al., *Trastuzumab Mechanism of Action; 20 Years of Research to Unravel a Dilemma.* Cancers (Basel), 2021. **13**(14): p. 3540.
69. Valabrega, G., F. Montemurro, and M. Aglietta, *Trastuzumab: mechanism of action, resistance and future perspectives in HER2-overexpressing breast cancer.* Ann Oncol, 2007. **18**(6): p. 977-984.
70. Junttila, T.T., et al., *Ligand-independent HER2/HER3/PI3K complex is disrupted by trastuzumab and is effectively inhibited by the PI3K inhibitor GDC-0941.* Cancer Cell, 2009. **15**(5): p. 429-440.
71. Austin, C.D., et al., *Endocytosis and sorting of ErbB2 and the site of action of cancer therapeutics trastuzumab and geldanamycin.* Mol Biol Cell, 2004. **15**(12): p. 5268-5282.
72. Nagata, Y., et al., *PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients.* Cancer Cell, 2004. **6**(2): p. 117-127.
73. Petricevic, B., et al., *Trastuzumab mediates antibody-dependent cell-mediated cytotoxicity and phagocytosis to the same extent in both adjuvant and metastatic HER2/neu breast cancer patients.* J Transl Med, 2013. **11**: p. 307.
74. Franklin, M.C., et al., *Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex.* Cancer Cell, 2004. **5**(4): p. 317-328.
75. Persson, M., et al., *[(177)Lu]pertuzumab: experimental studies on targeting of HER-2 positive tumour cells.* Eur J Nucl Med Mol Imaging, 2005. **32**(12): p. 1457-1462.
76. Adams, C.W., et al., *Humanization of a recombinant monoclonal antibody to produce a therapeutic HER dimerization inhibitor, pertuzumab.* Cancer Immunology, Immunotherapy, 2006. **55**(6): p. 717-727.
77. Agus, D.B., et al., *Targeting ligand-activated ErbB2 signaling inhibits breast and prostate tumor growth.* Cancer Cell, 2002. **2**(2): p. 127-137.
78. Mandó, P., et al., *Targeting ADCC: A different approach to HER2 breast cancer in the immunotherapy*

*era.* Breast, 2021. **60**: p. 15-25.

79. D. A.Case, D.S.C., T. E. I.Cheatham, T. A.Darden, R. E.Duke, T. J.Giese, H.Gohlke, A. W.Goetz, D.Greene, N.Homeyer, S.Izadi, A.Kovalenko, T. S.Lee, S.LeGrand, P.Li, C.Lin, J.Liu, T.Luchko, R.Luo, D.Mermelstein, K. M.Merz, G.Monard, H.Nguyen, I.Omelyan, A.Onufriev, F.Pan, R.Qi, D. R.Roe, A.Roitberg, C.Sagui, C. L.Simmerling, W. M.Botello-Smith, J.Swails, R. C.Walker, J.Wang, R. M.Wolf, X.Wu, L.Xiao, D. M.York and P. A.Kollman, *AMBER 2017*. 2017, San Francisc: University of California.
80. Obenaus, M., et al., *Identification of human T-cell receptors with optimal affinity to cancer antigens using antigen-negative humanized mice.* Nat Biotechnol, 2015. **33**(4): p. 402-407.
81. Sharma, S., et al., *Genome-scale identification of cellular pathways required for cell surface recognition.* Genome Res, 2018. **28**(9): p. 1372-1382.
82. Labrijn, A.F., M.L. Janmaat, J.M. Reichert, and P. Parren, *Bispecific antibodies: a mechanistic review of the pipeline.* Nat Rev Drug Discov, 2019. **18**(8): p. 585-608.
83. Nie, S., et al., *Biology drives the discovery of bispecific antibodies as innovative therapeutics.* Antib Ther, 2020. **3**(1): p. 18-62.
84. Reverberi, R. and L. Reverberi, *Factors affecting the antigen-antibody reaction.* Blood Transfus, 2007. **5**(4): p. 227-240.
85. Serçinoğlu, O. and P. Ozbek, *Sequence-structure-function relationships in class I MHC: A local frustration perspective.* PLoS One, 2020. **15**(5): p. e0232849.
86. Koehler Leman, J., et al., *Sequence-structure-function relationships in the microbial protein universe.* Nature Communications, 2023. **14**(1): p. 2351.
87. Jaiswal, V. and H.-J. Lee, *Conservation and Evolution of Antigenic Determinants of SARS-CoV-2: An Insight for Immune Escape and Vaccine Design.* Frontiers in Immunology, 2022. **13**: p.832106.
88. Zhou, C. and W.D. Shen, *Mammalian cell surface display of full length IgG.* Methods Mol Biol, 2012. **907**: p. 293-302.
89. Souche, E., et al., *Recommendations for whole genome sequencing in diagnostics for rare diseases.* European Journal of Human Genetics, 2022. **30**(9): p. 1017-1021.
90. Pei, X.M., et al., *Targeted Sequencing Approach and Its Clinical Applications for the Molecular Diagnosis of Human Diseases.* Cells, 2023. **12**(3): p.493.
91. Shin, H.-T., et al., *Prevalence and detection of low-allele-fraction variants in clinical cancer samples.* Nature Communications, 2017. **8**(1): p. 1377.
92. Yates, L.R., et al., *Subclonal diversification of primary breast cancer revealed by multiregion sequencing.* Nature Medicine, 2015. **21**(7): p. 751-759.
93. Dongre, H.N., et al., *Targeted Next-Generation Sequencing of Cancer-Related Genes in a Norwegian*

*Patient Cohort With Head and Neck Squamous Cell Carcinoma Reveals Novel Actionable Mutations and Correlations With Pathological Parameters.* Front Oncol, 2021. **11**: p. 734134.

94. Bewicke-Copley, F., et al., *Applications and analysis of targeted genomic sequencing in cancer studies.* Comput Struct Biotechnol J, 2019. **17**: p. 1348-1359.
95. Mertes, F., et al., *Targeted enrichment of genomic DNA regions for next-generation sequencing.* Brief Funct Genomics, 2011. **10**(6): p. 374-386.
96. Gaudin, M. and C. Desnues, *Hybrid Capture-Based Next Generation Sequencing and Its Application to Human Infectious Diseases.* Front Microbiol, 2018. **9**: p. 2924.
97. Zambrano, N., et al., *High-Throughput Monoclonal Antibody Discovery from Phage Libraries: Challenging the Current Preclinical Pipeline to Keep the Pace with the Increasing mAb Demand.* Cancers (Basel), 2022. **14**(5): p.1325.
98. Yang, W., et al., *Next-generation sequencing enables the discovery of more diverse positive clones from a phage-displayed antibody library.* Exp Mol Med, 2017. **49**(3): p. e308.
99. Ravn, U., et al., *Deep sequencing of phage display libraries to support antibody discovery.* Methods, 2013. **60**(1): p. 99-110.
100. Ferrara, F., et al., *Exploiting next-generation sequencing in antibody selections - a simple PCR method to recover binders.* MAbs, 2020. **12**(1): p. 1701792.
101. Huang, A., et al., *Strategies to Screen Anti-AQP4 Antibodies from Yeast Surface Display Libraries.* Antibodies (Basel), 2022. **11**(2): p.39.
102. Huhtinen, O., R. Salbo, U. Lamminmäki, and S. Prince, *Selection of biophysically favorable antibody variants using a modified Flp-In CHO mammalian display platform.* Front Bioeng Biotechnol, 2023. **11**: p. 1170081.
103. Sasso, E., et al., *One-Step Recovery of scFv Clones from High-Throughput Sequencing-Based Screening of Phage Display Libraries Challenged to Cells Expressing Native Claudin-1.* Biomed Res Int, 2015. **2015**: p. 703213.
104. Forsyth, C.M., et al., *Deep mutational scanning of an antibody against epidermal growth factor receptor using mammalian cell display and massively parallel pyrosequencing.* MAbs, 2013. **5**(4): p. 523-532.
105. Van Blarcom, T., et al., *Precise and efficient antibody epitope determination through library design, yeast display and next-generation sequencing.* J Mol Biol, 2015. **427**(6 Pt B): p. 1513-1534.
106. Nasereddin, A., et al., *Identification of SARS-CoV-2 Variants of Concern Using Amplicon Next-Generation Sequencing.* Microbiol Spectr, 2022. **10**(4): p. e0073622.
107. Haynes, W.A., et al., *High-resolution epitope mapping and characterization of SARS-CoV-2 antibodies in large cohorts of subjects with COVID-19.* Communications Biology, 2021. **4**(1): p. 1317.

108. Jones, B.E., et al., *The neutralizing antibody, LY-CoV555, protects against SARS-CoV-2 infection in nonhuman primates.* Sci Transl Med, 2021. **13**(593).
109. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-2120.
110. Pagès, H., et al., *Biostrings: Efficient manipulation of biological strings.* 2023.
111. Strang, K.H., et al., *Generation and characterization of new monoclonal antibodies targeting the PHF1 and AT8 epitopes on human tau.* Acta Neuropathol Commun, 2017. **5**(1): p. 58.
112. Arur, S. and T. Schedl, *Generation and purification of highly specific antibodies for detecting post-translationally modified proteins in vivo.* Nat Protoc, 2014. **9**(2): p. 375-395.
113. Hwang, L.A., et al., *Monoclonal Antibodies against Specific p53 Hotspot Mutants as Potential Tools for Precision Medicine.* Cell Rep, 2018. **22**(1): p. 299-312.
114. Liew, O.W., et al., *Epitope-directed monoclonal antibody production using a mixed antigen cocktail facilitates antibody characterization and validation.* Commun Biol, 2021. **4**(1): p. 441.
115. Puri, V., et al., *Highly efficient selection of epitope specific antibody through competitive yeast display library sorting.* MAbs, 2013. **5**(4): p. 533-539.
116. Ge, Q., et al., *An epitope-directed selection strategy facilitating the identification of Frizzled receptor selective antibodies.* Structure, 2023. **31**(1): p. 33-43.e5.
117. Chen, L., et al., *Epitope-directed antibody selection by site-specific photocrosslinking.* Sci Adv, 2020. **6**(14): p. eaaz7825.

# Achievements

## Publication

Ning Lin, Kotaro Miyamoto, Takumi Ogawara, Saki Sakurai, Shinae Kizaka-Kondoh, and Tetsuya Kadonosono. Epitope binning for multiple antibodies simultaneously using mammalian cell display and DNA sequencing. *Communications Biology*. 2024;7(1):652.

## Conferences

[1] Ning Lin, Kotaro Miyamoto, Takumi Ogawara, Tetsuya Kadonosono. A monoclonal antibody-guided screening platform for the efficient identification of epitope-specific antibodies. The 40th Medicinal Chemistry Symposium, Nov. 2023, Nagoya, Japan. (Poster presentation)

[2] Ning Lin, Tetsuya Kadonosono. Development of Evaluation System for Antibody Epitopes by Co-display of Antibodies and Antigen on Mammalian Cells, The 75th Annual Meeting of the Society for Biotechnology, Sep. 2023, Nagoya, Japan. (Oral presentation)

[3] Ning Lin, Saki Sakurai, Shinae Kizaka-Kondoh, Tetsuya Kadonosono. Establishment of an antibody-guided screening system for identifying epitope-directed anti-HER2 antibody mimetics, The 45th Annual Meeting of the Molecular Biology Society of Japan, Dec.2022, Chiba, Japan. (Poster presentation)

# Acknowledgment

This dissertation marks a significant milestone in my life, symbolizing the conclusion of over two decades of my studenthood. With this opportunity, I would like to sincerely thank all those who have supported and helped me during my doctoral period, as their contributions have facilitated my personal growth and academic accomplishments.

First and foremost, I would like to express my sincere appreciation to my supervisor, Associate Professor Tetsuya Kadonosono, for his patient guidance, unwavering support, and continuous encouragement throughout my doctoral journey. His expertise and insightful suggestions have played critical roles in the completion of this research. Moreover, I would like to express sincere gratitude to Professor Shinae Kizaka-Kondoh for her invaluable advice on my research. Her rigorous and serious attitude towards academic research serves as an exemplary model for me.

I would like to thank Mr. Kotaro Miyamoto (D2) for his kind help on NGS raw data processing, and to Mr. Takumi Ogawara (M2) for his dedicated efforts in MD simulation. In addition, I would extend my thanks to the other members of Kadonosono and Kondoh laboratories for generously sharing their time and experiences. It was a great pleasure to study in a harmonious environment, which provided support for my endeavors.

I would like to express my sincere acknowledgment to the Biomaterials Analysis Division of the Open Facility Center at Tokyo Institute of Technology for performing DNA sequencing analysis. Additionally, I would like to thank the FACS Core Laboratory at the Institute of Medical Science, University of Tokyo for providing technical assistance with operating flow cytometers. Furthermore, I am grateful to the Global Scientific Information and Computing Center of Tokyo Institute of Technology for granting permission to use TSUBAME.

I am indebted to my family, despite the physical distance that separated us a lot, they have always been supportive of my academic pursuits, giving me the freedom to chart my own path. Furthermore, I would like to thank China Scholarship Council for their financial support, which enabled me to complete my doctoral studies without any financial burden.

Last but not least, I would like to express appreciation for all the failed experiences I’ve encountered during my studies. Though initially daunting, they have proven to be invaluable lessons for me. The resilience I gained from overcoming these challenges and obstacles has empowered me with courage to face the subsequent setbacks with a more peaceful mindset, which supported me in completing my doctoral studies.

Ning Lin

February 2024