

DISS. ETH NO. 28505

***Synthetic selection and engineering of
antibodies***

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES

(Dr. sc. ETH Zurich)

presented by

Roy Alexander Ehling

MSc ETH Biotechnology

born on 16.01.1992

accepted on the recommendation of

Prof. Dr. Sai Reddy

Prof. Dr. Ben Murrell

Prof. Dr. Yariv Wine

2022

Zusammenfassung

Antikörper, die von B-Zellen produziert werden, sind vielfältige Moleküle, die fremde Antigene mit hoher Spezifität binden und neutralisieren. Dadurch bilden sie eine entscheidende Komponente der adaptiven Immunität. Die Vielfalt der Antikörper ist das Ergebnis einer fehleranfälligen, quasi-zufälligen Neuordnung von DNS-Segmenten während der B-Zell-Entwicklung sowie einer weiteren Diversifizierung in Folge der antigenspezifischen B-Zell-Aktivierung. Das daraus resultierende Repertoire enthält schätzungsweise über 10^{13} einzigartige Antikörpersequenzen und maximiert damit die Wahrscheinlichkeit, dass für jeden beliebigen Erreger mindestens eine antigenreaktive oder neutralisierende B-Zelle vorhanden ist. Ende 2019 tauchte ein neuartiger viraler Erreger - SARS-CoV-2 - auf und verbreitete sich rasch, was zur COVID-19-Pandemie führte. Die Struktur und damit die Sequenz, durch die sich SARS-CoV-2 ausbreitet und der Immunantwort entgehen kann, wurde unter anderem durch Tiefensequenzierung und pseudotypisierte Virusneutralisationstests untersucht. Insbesondere die Tiefensequenzierung hat sich als leistungsfähiges und wichtiges Instrument erwiesen, nicht nur zur Bestimmung der Dynamik des pathogenbedingten Antikörperrepertoires, sondern insbesondere zur Überwachung der Entwicklung der laufenden Pandemie.

In Kapitel 1 dieser Arbeit werden zunächst die Prinzipien der natürlichen Antikörperentwicklung vorgestellt, gefolgt von den Instrumenten zur Analyse des Immunrepertoires und ihrer derzeitigen Verwendung in der Antikörperforschung. Die entschlüsselten Antikörperrepertoires werden für den Aufbau von Antikörperbibliotheken verwendet. Das Screening solcher Bibliotheken wird mit einer Reihe von Screening- und Display-Plattformen mit mittlerem bis hohem Durchsatz durchgeführt, die auf ihre Stärken und Schwächen hin bewertet werden, wobei ihr Nutzen für die Entdeckung und Optimierung von Antikörpern besonders im Vordergrund steht. Schließlich endet Kapitel 1 mit einer ausführlichen Beschreibung von SARS-CoV-2, den zugelassenen Antikörpertherapien und dem eskalierenden Wettstreit zwischen SARS-CoV-2 und dem Immunsystem.

Trotz der Fortschritte bei der raschen Identifizierung von Antigen-bindenden B-Zellen aus Blut mit Hilfe von Antigen-Baiting-Techniken ist die Identität der Antikörper-sezernierenden Zellen (Plasmablasten und Plasmazellen), die direkt für den Aufbau von Antikörperreaktionen gegen Krankheitserreger verantwortlich sind, noch weitgehend unerforscht. Auch die Charakterisierung von COVID-19-Antikörpern hat sich weitgehend auf Gedächtnis-B-Zellen konzentriert, während es die Antikörper-sezernierenden Zellen sind, die eine entscheidende Rolle bei der Abheilung der SARS-CoV-2-Infektion spielen.

In Kapitel 2 beschreiben wir eine Technologie-Pipeline, die die Sequenzierung des Einzelzell-Antikörper-Repertoires und das Säugetierzell-Display integriert, um die Spezifität der transitorischen Plasmazellen von 16 rekonvaleszenten Patienten zu untersuchen. Die Einzelzellsequenzierung ermöglicht es uns, ein Profil des Antikörperrepertoires zu erstellen und erweiterte klonale Abstammungslinien zu identifizieren. Das Screening mit Hilfe von Säugetierzell-Display zeigt, dass 43 Antikörper (von 132 Kandidaten), die aus erweiterten Plasmazelllinien stammen, spezifisch für SARS-CoV-2-Antigene sind, darunter Antikörper mit hoher Affinität zur SARS-CoV-2-Rezeptorbindungsdomäne (RBD), die eine starke Neutralisierung und eine breite Bindung an die RBD von SARS-CoV-2-Varianten aufweisen.

In Kapitel 3 etablieren wir eine Methode zur synthetischen Koevolution von neutralisierenden Antikörpern und der SARS-CoV-2 RBD. Unsere Studie beginnt mit der Auswahl eines Panels von neutralisierenden Antikörpern, die während der frühen Phasen der COVID-19-Pandemie entdeckt wurden. Diese Antikörper wurden ursprünglich aufgrund ihrer Bindung an die ursprüngliche RBD von SARS-CoV-2 (Wu-Hu-1) ausgewählt, aber viele dieser Antikörper haben seither die Bindung und Neutralisierung an nachfolgende Varianten von SARS-CoV-2 (z. B. Beta, Delta, BA.1) verloren. Als Nächstes führen wir eine synthetische somatische Hypermutation (sSHM) durch, um den natürlichen Prozess der Antikörperrevolution zu imitieren, das in einer gezielten Mutagenese der komplementärbestimmenden Regionen (CDRs) der Antikörper besteht. Die sSHM-Bibliotheken werden dann mit Hilfe des Säugetierzell-Displays gescreent, um Antikörper-Klonotyp-Linien zu isolieren, die die Bindung an SARS-CoV-2 RBD-Varianten wiederhergestellt haben. Anschließend wird das Hefe-Oberflächen-Display von RBD-Mutagenese-Bibliotheken in Kombination mit Tiefensequenzierung eingesetzt, um Escape-Varianten für die neu ausgewählten sSHM-Antikörper zu identifizieren. Dieser integrierte Ansatz für das Library-on-Library-Screening offenbart koevolutionäre Mutationslandschaften von Antikörperklontypen und SARS-CoV-2 RBD-Varianten und kann die Identifizierung von zukünftigen Linien von SARS-CoV-2 ermöglichen. Schließlich beschreiben wir mehrere Antikörper aus sSHM-Bibliotheken, die die Bindung und Neutralisierung an SARS-CoV-2-Varianten wiederherstellen, darunter zu Omikron (BA.1).

In Kapitel 4 versuchen wir, die derzeitigen Plattformen für die Diversifizierung von Antikörpern zu verbessern. Um die Schwächen des Hefe-Displays für die Antikörperentwicklung zu bekämpfen, beschreiben und validieren wir Vektoren für ein schnelles und effizientes Shuttling von auf Hefe exprimierten Antigen-bindenden Fragmenten (Fabs) zu kompletten Antikörpern in Säugetierzellen. Trotz erheblicher Anstrengungen können nicht alle Antikörper durch leichte Diversifizierung gerettet werden. Bestimmte

Mutationen in der RBD von SARS-CoV-2, insbesondere in den VOCs Beta, Gamma und neuerdings auch Omikron BA.1, stören das Epitop einer bestimmten Klasse von Antikörpern. Darüber hinaus untersuchen wir, wie eine groß angelegte Umkodierung von Antikörper-CDR-Sequenzen unter Verwendung homologer Sequenzen zur Rettung solcher Antikörper eingesetzt werden könnte. Dies geschieht mit einem Verfahren, das als "in silico Deep Mutational Scanning" (isDMS) bezeichnet wird. Anstatt eine oder einige wenige Positionen auf einmal zu mutieren, werden Antikörperdatenbanken nach ähnlichen Sequenzen innerhalb einer bestimmten Distanz durchsucht. Aus diesen ähnlichen Sequenzen werden degenerierte Oligoschemata erstellt, um Stellen zu erhalten, die unveränderlich sind, während andere variiert werden, was im Wesentlichen die Sättigungsmutagenese überspringt. Dennoch kommen wir zu dem Schluss, dass isDMS eine ungeeignete Methode für die gestellte Aufgabe ist und sich viel besser für die Entwicklung breiter Bibliotheken zur Entdeckung von Antikörpern eignet als für die Antikörperoptimierung.

Abstract

Antibodies produced by B cells are diverse molecules that bind and neutralize foreign antigens with high specificity and serve as a critical component of adaptive immunity. Antibody diversity stems from an error-prone, quasi-random rearrangement of germline DNA segments during B-cell development, as well as further diversification during B cell maturation. The resulting repertoire is estimated to contain over 10^{13} unique antibody sequences, thus maximizing the probability to contain at least one antigen reactive or neutralizing B cell for any given pathogen. In late 2019, a newly emerged viral pathogen - SARS-CoV-2 - emerged and quickly spread, leading to the COVID-19 pandemic. The structure, and by extension the sequence, through which SARS-CoV-2 can spread and evade the immune response, has been studied by deep sequencing and pseudotyped viral neutralization assays. Deep sequencing in particular has emerged as a powerful and important tool, not only for determining pathogen-induced antibody repertoire dynamics but particularly for monitoring the evolution of the ongoing pandemic.

Chapter 1 of this thesis first introduces the principles of natural antibody development, followed by the tools to dissect and analyze the immune repertoire, as well as their current usage in antibody discovery. Decoded antibody repertoires are utilized to inform the construction of antibody engineering libraries. The screening of such libraries is performed using a variety of mid- to high-throughput screening and display platforms, which are evaluated for their strengths and weaknesses, with particular focus on their utility for antibody discovery and optimization. Finally, chapter 1 closes on an in-depth description of SARS-CoV-2, approved antibody therapies and the escalating arms race between SARS-CoV-2 and the immune system.

Despite the progress in the rapid identification of antigen-binding B cells from blood using antigen-baiting techniques, the identity of antibody-secreting cells that are directly responsible for mounting antibody responses against pathogens remains largely unexplored. Similarly, characterization of COVID-19 antibodies has largely focused on memory B cells; while it is the antibody-secreting cells that play a critical role in resolving SARS-CoV-2 infection.

In chapter 2, we describe a technology pipeline that integrates single-cell antibody repertoire sequencing and mammalian display to interrogate the specificity of plasma cells from 16 convalescent patients. Single-cell sequencing allows us to profile antibody repertoire features and identify expanded clonal lineages. Mammalian display screening is used to

reveal that 43 antibodies (of 132 candidates) derived from expanded plasma cell lineages are specific to SARS-CoV-2 antigens, including antibodies with high affinity to the SARS-CoV-2 receptor binding domain (RBD) that exhibit potent neutralization and broad binding to the RBD of SARS-CoV-2 variants (of concern/interest).

In chapter 3, we establish a method for synthetic co-evolution of neutralizing antibodies and the SARS-CoV-2 RBD. Our study starts by selecting a panel of neutralizing antibodies discovered during the early stages of the COVID-19 pandemic, which were initially selected based on binding to the original ancestral SARS-CoV-2 RBD (Wu-Hu-1), but many of these antibodies have since lost binding and neutralization to subsequent variants of SARS-CoV-2 (e.g., beta, delta, omicron). Next, we perform synthetic somatic hypermutation (sSHM) to mimic the natural process of antibody evolution, which consists of targeted mutagenesis of antibody complementary-determining regions (CDRs). The sSHM libraries are then screened using mammalian display in order to isolate antibody clonotype lineages that have recovered binding to SARS-CoV-2 RBD variants. Next, yeast surface display of RBD mutagenesis libraries combined with deep sequencing is used to identify escape variants for the newly selected sSHM antibodies. This integrated approach to library-on-library screening reveals co-evolutionary mutational landscapes of antibody clonotypes and SARS-CoV-2 RBD variants and may enable the identification of prospective lineages of SARS-CoV-2. Finally, we describe several antibodies from sSHM libraries that recover binding and neutralization to SARS-CoV-2 variants including omicron (BA.1).

In chapter 4, we attempt to improve upon current platforms for the diversification of antibodies. To combat the weaknesses of yeast display for antibody development, we describe and validate vectors for rapid and streamlined shuttling of yeast displayed antigen-binding fragments (Fabs) to full-length antibodies in mammalian cells. Despite significant effort, not all antibodies can be rescued by light diversification. Certain mutations in the RBD of SARS-CoV-2, particularly in the VOCs Beta, Gamma and, recently, Omicron, disrupt the epitope of a particular class of antibodies. Furthermore, we explore how large-scale recoding of antibody CDR sequences using homologous sequences could be employed to rescue such antibodies. This is done using a process called “in silico Deep Mutational Scanning” (isDMS). Instead of mutating one or a few positions at a time, antibody databases are mined for similar sequences within a defined edit distance. From these similar sequences, degenerate oligo schemes are created to preserve sites that are immutable while varying others that are, essentially skipping site-saturation mutagenesis. Nevertheless, we conclude that isDMS is an inappropriate method for the given task and may be much more suited to develop broad libraries for antibody discovery instead, rather than antibody optimization.