doi:10.1016/j.jmb.2007.08.005 J. Mol. Biol. (2007) 373, 924–940

JMB

Available online at www.sciencedirect.com
ScienceDirect

ELSEVIER

High-throughput Generation of Synthetic Antibodies from Highly Functional Minimalist Phage-displayed Libraries

Frederic A. Fellouse[1]†, Kaori Esaki[2], Sara Birtalan[1], Demetrios Raptis[2] Vincenzo J. Cancasci[2], Akiko Koide[2], Parkash Jhurani[1], Mark Vasser[1] Christian Wiesmann[1], Anthony A. Kossiakoff[2] Shohei Koide[2]* and Sachdev S. Sidhu[1]*

[1]*Department of Protein Engineering, Genentech Inc 1 DNA Way, South San Francisco, CA 94080, USA*

[2]*Department of Biochemistry and Molecular Biology The University of Chicago 929 East 57th Street, Chicago IL 60637, USA*

Received 22 May 2007; received in revised form 1 August 2007; accepted 1 August 2007
Available online 19 August 2007

We have previously established a minimalist approach to antibody engineering by using a phage-displayed framework to support complementarity determining region (CDR) diversity restricted to a binary code of tyrosine and serine. Here, we systematically augmented the original binary library with additional levels of diversity and examined the effects. The diversity of the simplest library, in which only heavy chain CDR positions were randomized by the binary code, was expanded in a stepwise manner by adding diversity to the light chain, by diversifying non-paratope residues that may influence CDR conformations, and by adding additional chemical diversity to CDR-H3. The additional diversity incrementally improved the affinities of antibodies raised against human vascular endoethelial growth factor and the structure of an antibody–antigen complex showed that tyrosine side-chains are sufficient to mediate most of the interactions with antigen, but a glycine residue in CDR-H3 was critical for providing a conformation suitable for high-affinity binding. Using new high-throughput procedures and the most complex library, we produced multiple high-affinity antibodies with dissociation constants in the single-digit nanomolar range against a wide variety of protein antigens. Thus, this fully synthetic, minimalist library has essentially recapitulated the capacity of the natural immune system to generate high-affinity antibodies. Libraries of this type should be highly useful for proteomic applications, as they minimize inherent complexities of natural antibodies that have hindered the establishment of high-throughput procedures. Furthermore, analysis of a large number of antibodies derived from these well-defined and simplistic libraries allowed us to uncover statistically significant trends in CDR sequences, which provide valuable insights into antibody library design and into factors governing protein–protein interactions.

Edited by I. Wilson

Keywords: phage display; protein engineering; combinatorial mutagenesis; antibody library; vascular endothelial growth factor

**Corresponding authors.* E-mail addresses: skoide@uchicago.edu; sidhu@gene.com.

Present address: F. A. Fellouse, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario, Canada M5G 1X5.

Abbreviations used: BSA, bovine serum albumin; CDR, complementarity determining region; CDR-H*n*, (where $n = 1, 2,$ or 3), heavy chain CDR 1, 2, or 3; CDR-L3, light chain CDR3; ELISA, enzyme-linked immunosorbent assay; Fab, antigen-binding fragment; HTP, high-throughput pipeline; hVEGF, human vascular endothelial growth factor; mAb, monoclonal antibody; scFv, single-chain variable fragment.

Introduction

The natural immune system can produce antibodies that recognize essentially any antigen with high affinity and specificity, and thus, antibodies produced by animal immunization and hybridoma methods[1] are indispensable tools for biological research. As a powerful alternative to hybridoma technology, phage display allows for the display of libraries of antigen-binding fragments (Fabs) or single-chain variable fragments (scFvs) on the surfaces of bacteriophage particles that encapsulate the encoding DNA.[2–6] These libraries can be cycled through binding selections *in vitro* to select antigen-specific antibodies. Selected antibodies are inherently monoclonal, and their amino acid sequences can be readily decoded from the DNA. Furthermore, the use of *in vitro* methods allows for exquisite control over selection conditions, thus allowing for tailored selection strategies for obtaining antibodies with extremely precise specificities, and also, for dealing with difficult antigens, such as integral membrane proteins. Importantly, phage display selection and analysis protocols are well suited for adaptation to a high-throughput pipeline (HTP), and consequently, the technology is likely to play an important role in proteomics research.[7–9]

A particularly promising branch of phage-displayed library technology is represented by the so-called "synthetic" antibody libraries, which contain antigen-binding sites constructed entirely from man-made diversity.[10,11] Libraries of this type are ideal for HTP applications, since frameworks can be chosen for optimal stability and performance, and the defined nature of the repertoires allows for rapid downstream sequence analysis and protein purification. In addition, the use of well-defined frameworks makes it possible to incorporate modular design features that enable facile affinity maturation or reformatting between different vector systems.

Over the last few years, significant advances in synthetic antibody engineering have been enabled by an enhanced understanding of the relationships between antibody structure and function. Analysis of antibody sequences and structures has revealed that antigen recognition is primarily mediated by the six complementarity determining regions (CDRs), or hypervariable loops, which form a functional antigen-binding site that is supported by more conserved framework regions.[12,13] Even within the CDRs, there are considerable biases in spatial and chemical diversity that is compatible with antigen recognition; antibody–antigen complex structures show that residues at certain CDR positions most often make contact with antigen and certain types of amino acids are most often involved in productive binding contacts.[14–17] Furthermore, while natural repertoires use combinations of dozens of different heavy and light chain frameworks to support CDR diversity, certain frameworks are much more stable than others, and consequently appear to be used more often in functional antibodies.[18] Highly functional synthetic antibody libraries have been constructed by using these natural clues to guide the choice of frameworks that are the best for supporting naïve diversity and to determine which types of chemical diversity should be introduced at which positions within the antigen-binding site.[10,11]

We have pursued a particularly simple approach to antibody engineering by using a single, highly stable Fab framework to support CDR diversity restricted to a binary code of only Tyr and Ser, two amino acids that are highly abundant in natural antigen-binding sites.[19–21] Surprisingly, these chemically minimalist libraries have proven to be highly effective in generating specific antibodies against a wide array of antigens, thus demonstrating a dominant role for Tyr in antigen recognition and providing a rationale for the high abundance of Tyr in the CDRs of the natural immune repertoire. Recently, we have shown that the effectiveness of the Tyr/Ser binary code in mediating molecular recognition is likely an intrinsic property not restricted to antibody scaffolds, as libraries of this type were also effective in providing specific molecular recognition in the context of a much smaller, single-domain scaffold of the fibronectin type III domain.[22]

These studies have established a benchmark for the minimal requirements for generating molecular recognition from naïve repertoires. However, while the Tyr/Ser binary libraries provided high affinity recognition in the nanomolar range for some antigens, in other cases affinities were in the micromolar range, suggesting that additional diversity is required for high affinity binding to a greater proportion of antigens. In the original Tyr/Ser binary Fab libraries, only positions that are expected to make a direct contact with the antigen ("paratope" residues) were diversified.[20] The natural antibody repertoire contains two additional types of CDR diversity.[12,13] First, amino acid diversity is present at CDR positions that do not make direct contact with the antigen. Such diversity at "non-paratope" residues contributes to generating conformational diversity of the paratope residues. Second, additional chemical diversity beyond the Tyr/Ser binary code is likely to augment affinity and fine-tune specificity.

Here, we have systematically augmented the diversity in the minimalist Tyr/Ser binary code to produce simple, yet highly functional synthetic antibodies. We introduced additional CDR diversity in a stepwise manner, first at non-paratope positions and then in all CDR-H3 positions, and evaluated the effects systematically using human vascular endoethelial growth factor (hVEGF) as the test antigen. The introduction of additional diversity incrementally improved the effectiveness of the Fab libraries. Our most complex Fab library was further evaluated with a diverse panel of antigens using newly established, HTP procedures for Fab generation. The library produced antibodies with tight affinities in the single-digit nanomolar range against all of the antigens tested, indicating that this fully