Analytical Biochemistry 417 (2011) 25–35

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Mapping protein interactions by combining antibody affinity maturation and mass spectrometry

Michael R. Dyson [a], Yong Zheng [b], Cunjie Zhang [b], Karen Colwill [b], Kritika Pershad [c], Brian K. Kay [c], Tony Pawson [b,d], John McCafferty [a,*]

[a] Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK
[b] Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada M5G 1X5
[c] Department of Biological Sciences, University of Illinois at Chicago, Chicago, IL 60607, USA
[d] Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, Canada M5S 1A8

ARTICLE INFO

Article history:
Received 11 January 2011
Received in revised form 27 March 2011
Accepted 3 May 2011
Available online 31 May 2011

Keywords:
Antibody phage display
Affinity maturation
Immunoprecipitation
Mass spectrometry
Protein–protein interaction networks

ABSTRACT

Mapping protein interactions by immunoprecipitation is limited by the availability of antibodies recognizing available native epitopes within protein complexes with sufficient affinity. Here we demonstrate a scalable approach for generation of such antibodies using phage display and affinity maturation. We combined antibody variable heavy (V_H) genes from target-specific clones (recognizing Src homology 2 (SH2) domains of LYN, VAV1, NCK1, ZAP70, PTPN11, CRK, LCK, and SHC1) with a repertoire of 10^8 to 10^9 new variable light (V_L) genes. Improved binders were isolated by stringent selections from these new "chain-shuffled" libraries. We also developed a predictive 96-well immunocapture screen and found that only 12% of antibodies had sufficient affinity/epitope availability to capture endogenous target from lysates. Using antibodies of different affinities to the same epitope, we show that affinity improvement was a key determinant for success and identified a clear affinity threshold value (60 nM for SHC1) that must be breached for success in immunoprecipitation. By combining affinity capture using matured antibodies to SHC1 with mass spectrometry, we identified seven known binding partners and two known SHC1 phosphorylation sites in epidermal growth factor (EGF)-stimulated human breast cancer epithelial cells. These results demonstrate that antibodies capable of immunoprecipitation can be generated by chain shuffling, providing a scalable approach to mapping protein–protein interaction networks.

The creation of a resource of binders to all proteins in the proteome together with their splice variants and posttranslational modifications would be of great value to the research community [1,2]. Application of such antibodies to affinity capture of protein complexes (immunoprecipitation) would enable the characterization of cellular signaling networks in response to various stimuli. Current methods require the artificial "tagging" of proteins that may alter their activity and cellular location, whereas the use of binders to native proteins allows one to directly observe unmodified proteins. There are several initiatives and proposals relating to genome-wide antibody generation [1,3–5]. The ability to generate large sets of antibodies validated for use in immunoprecipitation represents an ambitious goal that would nonetheless be of enormous value to the research community in understanding protein interaction networks.

This goal can best be achieved by screening large recombinant phage [6,7], yeast [8,9], or ribosome [10] display libraries, which is a faster and more scalable approach than animal immunization. Antibody phage display, first described two decades ago [11], is a powerful method for selecting novel binders from large antibody libraries by linking the binding properties of an antibody displayed on the surface of filamentous bacteriophage to the encoding DNA within the bacteriophage. As an alternative to antibodies, other scaffolds have been used with phage display to generate binding molecules, including combinatorial peptides [12], the Z domain of protein A [13], the fibronectin type III domain [14], and designed ankyrin repeat proteins [15].

In a recent international collaboration, we and others demonstrated the feasibility of generating research reagents by phage display by targeting a group of 20 Src homology 2 (SH2)[1]

* Corresponding author. Fax: +44 1223 333345.
E-mail address: jm635@cam.ac.uk (J. McCafferty).

[1] *Abbreviations used:* SH2, Src homology 2; pTyr, phosphorylated tyrosine; scFv, single chain variable fragment; V_H, variable heavy; V_L, variable light; MS, mass spectrometry; EGF, epidermal growth factor; PCR, polymerase chain reaction; DTT, dithiothreitol; BSA, bovine serum albumin; ELISA, enzyme-linked immunosorbent assay; PBS, phosphate-buffered saline; PBSM, 2% milk powder in PBS; PBST, PBS and 0.1% Tween 20; IP-ELISA, immunoprecipitation ELISA; RU, response units; PMSF, phenylmethylsulfonyl fluoride; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; MRM, multiple reaction monitoring; SPR, surface plasmon resonance; EGFR, EGF receptor.

0003-2697/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2011.05.005

domain-containing proteins [5,16,17]. SH2 domains mediate many signal transduction processes through their ability to bind phosphorylated tyrosine (pTyr)-containing polypeptides [18,19]. In humans, there are 120 SH2 domains found within 110 distinct proteins that represent several different protein classes, including kinases, phosphatases, adapters, and transcription factors [20]. The SH2 domain consists of approximately 100 amino acids that fold into an antiparallel β-sheet flanked by two α-helices. Binding specificity of the SH2 domain is typically mediated by the 3–6 amino acids C terminal to the pTyr in the target sequence [21,22]. Given their widespread role in signal transduction and the availability of many characterized antibodies, in this study we focused on eight different SH2 domain proteins: LYN, VAV1, NCK1, ZAP70, PTPN11, CRK, LCK, and SHC1. For more detailed quantitative studies, we concentrated on the SH2 domain-containing protein SHC1. SHC1 is a prototypic scaffold protein that is essential for cancer progression [23]. It links activated receptor protein tyrosine kinases to the Ras-MAPK pathway and is a mediator of cell responsiveness to different external stimuli involving G protein-coupled receptors, immunoglobulin receptors, integrins, and non-receptor tyrosine kinases [24]. SHC1 has been shown to direct tissue morphogenesis during development [24] and encodes three splice variants of 46 kDa (p46SHC1), 52 kDa (p52SHC1), and 66 kDa (p66SHC1). Overexpression of p66SHC1 accelerates ES cell neural induction and modulates the Wnt/β-catenin pathway [25].

The availability of immunoprecipitation reagents to these and other signaling molecules would be of enormous benefit for studying signaling interactions and dynamics either individually or at a systems biology level. Despite the long history of the use of antibodies in immunoprecipitation, and despite the importance of the technique, there have been no studies examining the relationship between antibody affinity and performance in immunoprecipitation. Immunoprecipitation is a particularly challenging application for antibodies because it requires affinity capture and retention of native proteins and their complexes present at relatively low concentrations in cells or tissues. Given these requirements, we anticipated that high affinity would be a crucial determinant of success and sought to improve the affinity of antibodies emerging from phage display selections. Following the initial selection of antibodies recognizing SH2 domains [17], we employed "chain-shuffling" to create secondary gene-specific libraries.

In our antibody display library, antibodies are presented in the form of single chain variable fragments (scFvs), where the heavy chain variable region genes (V_H) and light chain variable region genes (V_L) are joined by a flexible linker peptide. Although the initial phage antibody selections [17] were performed with a very large antibody library consisting of more than 10^{10} clones [7], we reasoned that any V_H will have paired with a limited number of V_L partners (and vice versa) and that each might not have found its optimal partner from the available repertoires. Therefore, we employed chain shuffling as a simple method for creating secondary diversified libraries from individual antibodies from which we could select higher affinity variants. Because the greatest diversity resides within the V_H region, we retained selected V_H genes and shuffled these with a repertoire of V_L genes. Therefore, we created diversified libraries to eight targets in parallel by cloning selected mixes of V_H genes from primary selections [17]. These libraries were subjected to stringent selections using limiting concentrations of biotinylated antigen.

The availability of the target epitope is an important factor in determining success in immunoprecipitation. With this in mind, we developed a novel 96-well immunocapture screen to rapidly identify binders recognizing available epitopes with sufficient affinity to capture low levels of endogenous SHC1 in a breast cancer epithelial cell line. The affinities and off-rates of a panel of anti-SHC1, all sharing the same V_H gene, were measured, and these correlated with their ability to work in immunoprecipitation. We demonstrated the use of our anti-SHC1 antibodies in immunoprecipitation coupled to mass spectrometry (MS) to identify known binding partners of SHC1 during epidermal growth factor (EGF) signaling.

Materials and methods

Generation of scFv libraries by chain shuffling

The V_H region of selected scFvs specific to the SH2 domains (LYN, VAV1, NCK1, ZAP70, PTPN11C, CRK, LCK, and SHC1) were polymerase chain reaction (PCR) amplified from pSANG10-TEV plasmid DNA [17] with the primers pSANG10-PelB (CGCTGCCC AGCCGGCCATGG) and HLINK (ACCGCCAGAGCCACCTCCGCC). The PCR reactions consisted of primers (0.5 μM), 2× Qiagen Hot Start *Taq* PCR mix (25 μl), and plasmid DNA (100 ng) in a total volume of 50 μl. The set of amplified V_H genes for each target antigen were pooled and purified by Qiagen spin column, and inserts (15 μg) were digested with *Nco*I (40 U) and *Xho*I (60 U) in 200 μl of NEB buffer 2 (50 mM NaCl, 10 mM Tris–HCl [pH 7.9], 10 mM $MgCl_2$, 1 mM dithiothreitol [DTT], and 0.1 mg/ml bovine serum albumin [BSA]) for 3 h at 37 °C. Digested inserts were purified by Qiagen spin column and ligated to the *Nco*I/*Xho*I cut pSANG4 vector harboring the naive V_L κ- and λ-chain libraries [7]. Ligation reactions contained V_H insert (1 μg), pSANG4 vector (2.5 μg), 50 mM Tris–HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT, and T4 DNA ligase (1000 U, New England Biolabs) in a total volume of 125 μl and were incubated at 16 °C for 16 h. Ligation reactions were phenol–chloroform extracted by the addition of 175 μl of 10 mM Tris–HCl (pH 8.0) to each ligation reaction, followed by extraction with 300 μl of phenol–chloroform–isoamylalcohol (25:24:1), and then vortexed and spun at 15,600*g* (2 min, 4 °C). The top aqueous layer was transferred to a fresh microfuge tube containing 280 μl of chloroform, vortexed, and spun at 15,600*g* (2 min, 4 °C), and the top aqueous layer was transferred to a fresh microfuge tube containing 1/10 volume (20 μl) of 3 M potassium acetate (pH 7.0). To this, 2.5 volumes (500 μl) of chilled 100% ethanol was added, incubated on ice for 10 min, spun at 15,600*g* (10 min, 4 °C), supernatant decanted, pellet washed with 600 μl of chilled 75% ethanol, spun at 15,600*g* (5 min, 4 °C), supernatant decanted, pellet air dried at 37 °C for 10 min, and dissolved in 4 μl of 10 mM Tris–HCl (pH 8.0). Here 2 μl of ligation reaction was used to electroporate 40 μl of electrocompetent TG1 cells in duplicate (Bio-Rad MicroPulser, EC1 channel), followed by the addition of 2 ml of SOC medium and incubation at 37 °C for 40 min. Cells were plated on 2× TY agar plates containing ampicillin (100 μg/ml) and glucose (2%). Dilutions of the transformation were also plated to determine library size, which ranged from 1.5×10^8 to 2×10^9 clones, with 77–93% of the transformants being positive for insertion scFv encoding insert, as determined by colony PCR.

Phage display selections, subcloning scFv populations into an expression vector, and primary ELISA

Rescue of phage particles from the chain-shuffled libraries and two rounds of selections against the eight SH2 domains were as described previously [7] except that selections employed biotinylated antigens (0.1–100 nM) that were captured with 25 μl of streptavidin-coated Dynabeads (Invitrogen, cat. no. M-280). Selected scFv populations were subcloned en masse from the phagemid vector, pSANG4, into the expression plasmid, pSANG-TEV-3F, which expresses scFvs with a six-histidine tag, a cleavage site for tobacco etch virus (TEV) protease, and three tandem copies of the FLAG