New Biotechnology 33 (2016) 790–796

Contents lists available at ScienceDirect

New Biotechnology

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

Full length article

Next generation sequencing of all variable loops of synthetic single framework scFv—Application in anti-HDL antibody selections

Janita Lövgren[a,*], Juha-Pekka Pursiheimo[b,1], Mikko Pyykkö[a], Jussi Salmi[b], Urpo Lamminmäki[a]

[a] Department of Biochemistry, Division of Biotechnology, University of Turku, Turku, Finland
[b] Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, Turku, Finland

ARTICLE INFO

Article history:
Received 4 March 2016
Received in revised form 1 July 2016
Accepted 17 July 2016
Available online 19 July 2016

Keywords:
scFv
Antibody library
Phage display
Next generation sequencing
Deep sequencing
High density lipoprotein

ABSTRACT

Next generation sequencing (NGS) can be applied to monitoring antibody phage display library selection processes to follow the enrichment of each individual antibody clone. Utilising the recent development of the Illumina sequencing platform enabling sequencing up to 2 × 300 bp, we have developed a method to deep sequence all complementarity determining regions (CDRs) in the clones obtained from a synthetic single framework antibody library. This was complemented by an in-house bioinformatics pipeline for efficient analysis of the sequencing results. The method was utilised to study antibody selections against high density lipoprotein (HDL) particles. Sequencing of the output from each selection round enabled extraction of useful information on both the total copy numbers as well as the relative enrichment rates of the clones. Ten antibody clones showing different ranking in terms of frequency were reproduced from synthetic DNA constructs and their capacity to bind HDL was verified by an immunoassay. The method thus facilitates the isolation of clones of interest, and in particular can assist retrieval of less efficiently enriched, yet interesting clones, which are unlikely to be identified by conventional, random colony picking based, screening.

1. Introduction

Phage display of recombinant antibodies is a well established method for developing specific affinity reagents for various applications [1]. In combination with universal antibody libraries, the method makes it possible to produce antibodies without the use of animals and enables rapid generation of binders against virtually any target molecule, as well as generating antibodies which are extremely difficult to obtain by animal immunisation. A two-stage process is used to isolate specific binders from an antibody phage library. The library is first subjected to phage display selections to enrich antibodies against the target of interest, followed by a screening step which conventionally involves immunoassay based activity analysis of antibody clones produced from individually picked bacterial colonies. The enrichment rate of an individual antibody phage clone can be affected by several factors [2], and in some situations some less efficiently enriched clones, which are rarely picked in the screening, may nevertheless also be of high interest. The risk of inefficient enrichment of some potentially useful clones is particularly prominent when the selections are performed against a complex target such as intact cells [3] comprising multiple antigens of different chemical natures and abundance. Recently, next generation sequencing (NGS) has been employed to analyse the entire output of antibody phage library selections [4–7], and the possibilities and benefits of such approaches have been recently reviewed [8]. The reported approaches were based on monitoring the sequences of the third complementary determining region of the heavy chains (CDRH3) and allowed the rapid assessment of the identity and abundance of virtually all enriched clones. In some cases, the sequence information obtained was also used for the isolation of individual interesting antibody clones from the pool of enriched library clones without the random colony picking based screening process [4,5,9,10]. However, as the sequence information was limited to the CDRH3, both hybridisation- and PCR-based cloning techniques were required to rescue the entire sequences of interesting antibody clones.

Abbreviations: CAD, coronary artery disease; CDR, complementarity determining region; HDL, high density lipoprotein; LDL, low density lipoprotein; NGS, next generation sequencing; scFv, single-chain variable fragment.

* Corresponding author at: Department of Biochemistry, Biotechnology, University of Turku, Tykistökatu 6A 6th floor, FIN-20520, Finland.
E-mail addresses: janita.lovgren@utu.fi, janita.lovgren@pp.inet.fi (J. Lövgren).
[1] Present address: Department of Medical Biochemistry and Genetics, Institute of Biomedicine, Turku, Finland

http://dx.doi.org/10.1016/j.nbt.2016.07.009
1871-6784/

The recent advances in the Illumina sequencing technology allow sequencing of about 600 base pair fragments (V3 2 × 300 bp chemistry). Based on this technology, we have developed an amplicon deep sequencing method, which allows one to reveal all the sequence regions determining the binding specificity of the antibodies in our synthetic single-chain Fv (scFv) antibody libraries [11,12]. As a result, any antibody clone of interest in the enriched library can easily be reproduced by designing a synthetic DNA to construct.

Here, we describe the use of this method both for monitoring the enrichment process and the isolation of the specific binders in a library selection campaign against a high density lipoprotein (HDL) antigen from coronary artery disease (CAD) patients. HDL is a supramolecular particle composed of both protein and lipids, with up to 85 associated proteins of varying abundance [13]. The heterogeneity of HDL is thought to play a role in CAD. Some HDL subspecies are believed to be atheroprotective and others atherogenic [14,15]. Due to these features, HDL is a very interesting but also challenging target antigen. In a previous study, we observed that recombinant antibodies selected through phage display and colony picking based screening were mostly against the most abundant apolipoproteins in HDL, apolipoprotein A-I and A-II (apoA-I and apoA-II) [16]. By using a deep sequencing approach a more extensive understanding of the selection process is obtained.

2. Materials and methods

2.1. CAD HDL pannings

The HDL isolated from serum of CAD patients and purified apoA-I was obtained from Matti Jauhiainen from the Genomics and Biomarkers unit, National Institute for Health and Welfare, Helsinki, Finland. HDL was isolated by sequential ultracentrifugation using KBr for density adjustment as described earlier [16]. Protein content of isolated HDL was measured by the Bradford method and equal amounts from 24 individuals were pooled and then biotinylated with EZ-Link NHS-PEG_4-Biotin (ThermoScientific, USA). ApoA-I was biotinylated using bioisothiocyanate. Synthetic single framework antibody libraries scFvP [11] and scFvM [12] were displayed in a multivalent form using Hyperphage (Progen Biotechnik, Germany). Dynabeads M280 streptavidin (Invitrogen, Thermo Fisher Scientific, USA) were saturated with the biotinylated apoA-I or CAD HDL and biotin. For the first panning cycle 1E13 scFv-phage were incubated with 50 μl apoA-I coated beads in 400 μl 50 mM Tris-HCl pH 7.5, 150 mM NaCl (TBS) to remove apoA-I binding and unspecific antibodies. The unbound phage were incubated with 50 μl of CAD HDL coated beads under mixing for 1 h at room temperature. The beads with the attached phage were washed five times with panning buffer and the bound phage were eluted with 50 μg/ml trypsin for 30 min at room temperature. Trypsin was inhibited by adding 5 x molar excess of Soy Bean Trypsin Inhibitor. After the first round, the scFv-DNA was recovered from the eluate and eluted beads with PCR. The primers used for annealing on the 5′ and 3′ ends of the scFv-DNA, were 5′-CGGCAGCCGCTGGATTGTTATTAC and 5′-ACCAGAACCGCCACGACCTTC, respectively. The PCR protocol was asymmetric starting with an initial amplification performed with the mere upstream primer using initial denaturation 2 min 95 °C, 10 cycles of denaturation 95 °C 30 s, annealing 60 °C 30 s and amplification 72 °C 2 min. This was followed by the conventional PCR using both the primers and initial denaturation 2 min 95 °C, 30 cycles of denaturation 95 °C 30 s, annealing 60 °C 30 s, amplification 72 °C 2 min and a final elongation step at 72 °C of 10 min. The PCR product was purified, cut with SfiI, and ligated back in the pEB32 x vector [12]. *E.coli* SS320 cells were transformed with the ligation mixture and plated on a large LA (10 μg/ml tetracyclin and 25 μg/ml chloramphenicol and 0.5% glucose) dish and grown overnight at 30 °C. In the morning the cells were suspended in SB and used to make a vector miniprep of the panning output for sequencing and for culture for production of scFv-phages using VCS M13 helper phage. For the second panning cycle 20 μl apoA-I beads were incubated with 1E11 scFv-phage in 400 μl TBS, 0.001% Tween 20 and 1% BSA. The unbound phage were incubated with 10 μl of CAD HDL coated beads under mixing for 1 h at room temperature. The beads with the attached phage were washed four times with panning buffer and once with TBS. The bound phage were eluted with trypsin as described above. Antibodies were recovered by infection of *E.coli* XL 1 Blue cells. The cells were plated on LA (10 μg/ml tetracyclin and 25 μg/ml chloramphenicol and 0.5% glucose) dish and grown overnight at 30 °C. After the cells had grown sufficiently they were suspended in SB and used to make a vector miniprep for sequencing and cultured for production of scFv-phage for the third panning round. The third panning round was done as the second one using 10 μl apoA-I beads and 5 μl CAD HDL beads.

2.2. Sample preparation for NGS

The CDRL1-CDRH3 regions (from 597 to 636 nucleotides in length) of the scFv were amplified by PCR from the panning output DNA minipreps. Before PCR the scFv DNA was linearised by cutting the vector with SfiI. SfiI. The sequencing library was generated in a single PCR with primers containing the adapters and specific index sequences required for Illumina sequencing. The adapter-index-primers used were

(FRW1) 5′-AATGATACGGCGACCACCGAGATCTACAC-ATCGTACG-TATGGTAATT-GG-CCAGTCGGTGTCCAGCTC,

(FRW2) 5′-AATGATACGGCGACCACCGAGATCTACAC-ACTATCTG-TATGGTAATT-GG-CCAGTCGGTGTCCAGCTC,

(REV1) 5′-CAAGCAGAAGACGGCATACGAGAT-AACTCTCG-AGT-CAGTCAG-GC-GACTAGTGTACCCTGACC,

(REV2) 5′-CAAGCAGAAGACGGCATACGAGAT-ACTATGTC-AGT-CAGTCAG-GC-GACTAGTGTACCCTGACC,

(REV3) 5′-CAAGCAGAAGACGGCATACGAGAT-AGTAGCGT-AGT-CAGTCAG-GC-GACTAGTGTACCCTGACC.

The primers were used in designated combinations to allow identification of specific sample subsets. 20 ng of the Sfi I cut pEB32x-scFv plasmid DNA was used in a 50 μl reaction. KapaHiFi polymerase (Kapa Biosystems, Boston, USA) was used with the following PCR protocol: initial denaturation 3 min 95 °C, 15 cycles of denaturation 98 °C 20 s, annealing 62 °C 20 s, amplification 72 °C 30 s. The 693–728 bp PCR product was gel purified and extracted with MinElute kit (Qiagen).

2.3. Next generation sequencing (NGS)

Indexed libraries were pooled and sequenced with an Illumina MiSeq instrument (Illumina, USA) at the Finnish Microarray and Sequencing Centre. The molar concentration of the purified PCR products was determined using Bioanalyzer (Agilent) and QuBit v2 fluorometer (InVitrogen/LifeTechnologies/ThermoFisher). For the sequencing 12 pM sample was loaded together with 10% phiX viral DNA. The sequencing was performed using the V3 2 × 300 bp Illumina sequencing kit. Libraries were sequenced using custom sequencing primers for Read1 (5′-TATGGTAATT-GG-CCAGTCGGTGTCCAGCTC), Read2 (5′-AGTCAGTCAG-GC-GAC-TAGTGTACCCTGACC) and Index read (5′-GGTCAGGGTACAC-TAGTCGCCTGACTGACT).