Molecular toolkits for engineering of self-navigating drug delivery vehicles

James W. Gillespie* and Valery A. Petrenko**

Department of Pathobiology, 264 Greene Hall, Auburn University, AL, USA
*gillejw@auburn.edu and **petreva@auburn.edu

ABSTRACT

Analysis of multi-thousand clone populations of the cancer cell-binding landscape phages and their multi-motif proteins using massively parallel sequencing techniques allowed the discovery of a variety of short motifs serving as elementary binding units during phage selection – a 'molecular LEGO' for engineering of proteins with predicted cell binding, penetrating and migrating performance. This finding inspired us to propose a novel "addressed drug navigation" concept, which relies on the use of "self-navigating ligands", selected from polyvalent landscape phage display libraries and accumulating EBUs responsible for binding to different tissue cells. Applied to the targeted drug delivery problem, this novel approach promises to replace the existing 'point to point' targeting concept for the novel 'self-navigating' drug delivery paradigm that can be used as a theoretical basis in development of a novel generation of molecular imaging probes and medications for precise and personal medicine.

Keywords: phage display, targeted drug delivery systems, nanomedicines, self-navigating drug delivery systems

1 INTRODUCTION

One of the current strategies for improving cancer treatments, while simultaneously reducing non-specific drug side effects, is the development of actively targeted drug delivery systems. At a minimum, these targeted systems are composed of three common features: a therapeutic drug, a carrier system to increase circulation time of the drug, and a targeting ligand that allows specific accumulation within the desired cell type [1]. Targeting ligands are commonly selected based upon the prior knowledge of known receptors that are over expressed on the surface of cancer cells or blood vessels. When these actively targeted drug delivery systems are administered intravenously, they are hypothesized to accumulate based on a combination of vascular defects (i.e., the enhanced permeability and retention or EPR effect) and specific interactions of a targeting ligand with the over expressed receptor [2]. These actively targeted nanomedicines have demonstrated sufficient activity in preclinical models. However, translation to spontaneous tumors in human patients has demonstrated limited improvement with poor accumulation of administered drug into the tumor [3, 4].

The currently accepted 'point-to-point' drug delivery concept, which involves using ligands based on individual points along the intended delivery path. For example, targeting ligands for tumor vasculature and cancer cells are used in these first-generation targeted drug delivery systems and accumulate drug at the desired cell type. However, the tumor microenvironment has been revealed to be more complex, leading us to propose the use of 'self-navigating' ligands to address the limitations observed with targeted drug delivery systems. This concept suggests the use of ligands that can navigate drug delivery systems through a complex microenvironment to the intended target [5].

Phage display is a commonly used technique to identify protein interactions with different types of targets including mammalian cell lines. Two of the predominant proteins of fd-tet phage vector that are used for displaying randomized fusions are the pIII and pVIII proteins, which are found in 5 or ~4,000 copies per phage particle respectively, as illustrated in Figure 1 [6]. We have extensively used polyvalent peptide display on the pVIII protein to identify ligands with specificity and selectivity towards different types of cancer cell types including: prostate [7], breast [8], lung [9], and pancreatic [10, 11] cancer cell lines. These 8 to 9-mer peptide fusions were used within the context of the full-length pVIII protein to create a panel of actively targeted drug delivery systems [12].

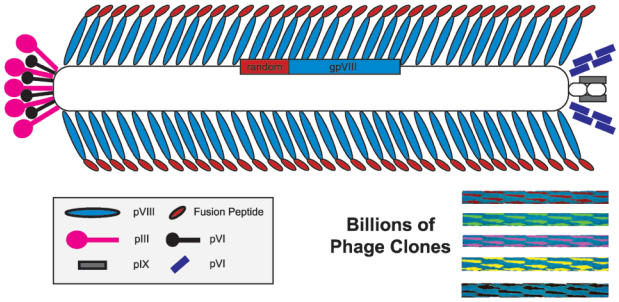

Figure 1: Schematic of f8-type landscape phage display libraries. Derivatives of filamentous phage fd-tet consist of phage with a randomized peptide (red) fused to every copy of the N-terminus of the pVIII major coat protein (blue).

Using polyvalent landscape phage display libraries, we previously identified a collection of ligands with promiscuous cell binding properties, in which a single peptide ligand can interact with two phenotypically different cell types [11]. This discovery allowed us to propose the concept of elementary binding units or "EBUs". One of the first commonly recognized elementary

binding units, RGD, was identified more than two decades ago [13] and has been used extensively for targeting integrins overexpressed on the blood vessels within the tumor microenvironment. Later it was recognized that the affinity and selectivity of the RGD motif could be modulated by changing the trailing C-terminal residue and also by constraining the conformation of the peptide [13]. We hypothesized that these promiscuous peptides had accumulated multiple EBUs into a single peptide motif during enrichment through multiple rounds of selection from the initial phage display library.

We suggest that rational design of motifs in the future will be accomplished by the combination of multiple EBUs as molecular LEGO® building blocks to construct a desired peptide motif. Here, we use landscape phage display libraries in combination with massively paralleled sequencing to identify potential EBUs for cell-specific targeting applications including personalized, precision imaging and drug delivery systems.

2 EXPERIMENTAL DESIGN

2.1 Phage Display

Landscape phage display libraries, f8/8 and f8/9, were constructed previously by introducing a randomized oligonucleotide sequence into the gpVIII of an fd-tet vector resulting in the randomized sequence being expressed as an 8- or 9- amino acid peptide fusion to the N-terminus of every copy of the mature pVIII major coat protein [14, 15]. Biopanning experiments to enrich cancer cell-specific population were enriched previously and all experimental details are described elsewhere [8-11]. Following each round of selection, recovered phages are amplified in *E. coli* and purified using standard procedures to produce phage populations enriched in cancer cell-specific binding phages. A schematic of the biopanning procedure is depicted in Figure 2.

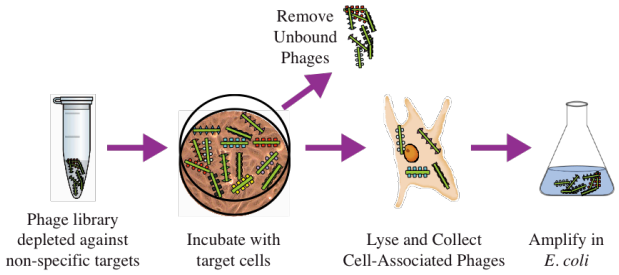

Figure 2: Conceptual *in vitro* biopanning procedure to enrich phage populations.

2.2 Next Generation Sequencing

Purified phages were diluted 1:100 in PCR grade water and gpVIII-specific amplicons were generated using a standard PCR with gpVIII-specific primers. Amplicons were purified using a QIAquick PCR purification kit (Qiagen) and eluted with ddH_2O. Amplicon size and integrity were verified by Bioanalyzer (Agilent) by the MGH NextGen sequencing core facility. Illumina multiplex barcodes and adapter sequences were ligated onto a representative portion of the amplicon library using the NEBNext DNA library prep master mix set and multiplex oligos. Final multiplexed library samples were purified and loaded with 15% PhiX control library onto a 150-cycle MiSeq flow cell for cluster generation. Amplicons were sequenced by the MGH NextGen sequencing core facility on an Illumina MiSeq for 150 cycles in single-read mode using standard Illumina primers.

2.3 Bioinformatics

Files were obtained as demultiplexed FASTQ files and analyzed in a bioinformatics workflow, adapted from [16], as depicted in Figure 3 to generate a library of peptide sequences interacting with the indicated target cell line.

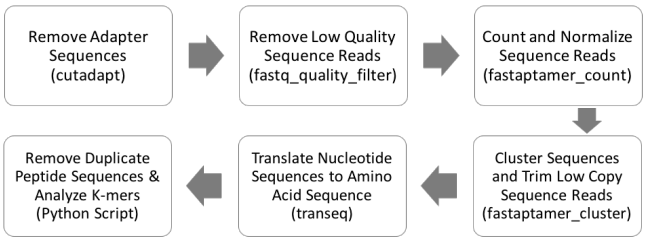

Figure 3: Bioinformatics workflow for analyzing NGS datasets following biopanning experiments.

Briefly, random nucleotide inserts were first trimmed of upstream and downstream gpVIII elements using the Cutadapt python script [17]. In-frame nucleotide inserts recovered from the trimmed sequence population were limited to 27 base pairs corresponding to the expected 9 amino acid insert size. Recovered insert sequences were filtered for high quality sequence reads with the FASTQ Quality Filter program found within the FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Sequences were filtered from the final library if any nucleotide position had a FASTQ quality score less than 30. Duplicate nucleotide sequences were combined into a single sequence using the FASTX Collapser program within the FASTX-toolkit. Finally, nucleotide sequences were translated to amino acid sequences using the EMBOSS transeq program using the standard codon table [18]. Duplicate peptide sequences were similarly combined using a custom Python script. Peptide k-mers were analyzed and compared between populations using custom Python scripts.

3 RESULTS

Phage PCR amplicons were generated from representative breast cancer cell-specific phage sublibraries. Following massively parallel sequencing on a MiSeq instrument and analysis by the bioinformatics workflow above, representative summary data is presented in Table 1.

In general, highly stringent parameters were applied throughout the analysis of phage amplicons to ensure high quality sequence data with low error rates. As such, a minimum Phred quality score of 30 was required to increase the base call accuracy to a minimum of 99.9% accurate. An average Phred quality score of 37.7 was obtained for each nucleotide position after removal of low quality sequence reads. These representative data demonstrate the high read coverage (>100 reads expected per sequence) that was obtained for each unique sequence within the overall population.

Table 1: Representative Summary of Next Generation Sequencing Data Analysis

Sample	Total Reads	Unique Sequences	Tri-mers
BC[†] (Eluate)	1,977,328	1,072	2,873
BC (Lysate)	1,775,077	2,942	4,245
Normal	2,040,792	8,116	5,022

We hypothesized that short k-mer domains containing the proposed EBUs would be enriched within a given sequence population for each sample. After generating a peptide sequence database for each sample analyzed, a list of k-mers was generated to identify potential EBUs. It was suggested previously that short linear motifs (SLiMs) are expected to be 3 to 4 amino acid residues in length [19]. Since the SLiMs are expected to be equivalent to EBUs, we analyzed the frequency of each tri-mer peptide domain found within the recovered sequence population for each sample. The number of tri-mers identified in each sample are listed in Table 1. When comparing the tri-mers identified from the breast cancer cell-specific phage population with those identified from normal breast epithelial cell-specific phage population, 32.6% of all tri-mers are shared between the two populations. However, there were 140 (3.27%) tri-mers enriched specifically in the breast cancer cell-specific phage population and 2,743 (64.12%) tri-mers enriched in the normal breast epithelial cell-specific population. These data suggest that there may be specific EBUs that interact only with overexpressed receptors on the breast cancer cells. Identification of these potential EBUs provides a theoretical basis for the rational design of cell-specific ligands. Confirmation of EBU interactions with the target cell population and the appropriate receptor is required to verify specificity and selectivity of a rationally designed peptide ligand.

4 PROSPECTS

Previously, it was suggested that SLiMs are expected to be 3-4 amino acid residues in length and interact transiently with their cognate receptor with low-to-medium affinity [19]. Landscape phage display vectors were demonstrated to identify these moderate affinity, transient interactions better than using traditional pIII display systems and may be better suited for identifying EBUs.

Using massively parallel sequencing of landscape phage populations following enrichment during selection experiments, we demonstrate here the feasibility of identifying potential EBUs that can be combined, like molecular LEGO® building blocks, into larger peptide motifs for the rational design of targeting ligands. The concept to combine EBUs into a single targeting ligand, as illustrated in Figure 4, was discovered when promiscuous ligands were identified with specificity towards both lung and pancreatic cancer cells *in vitro*.

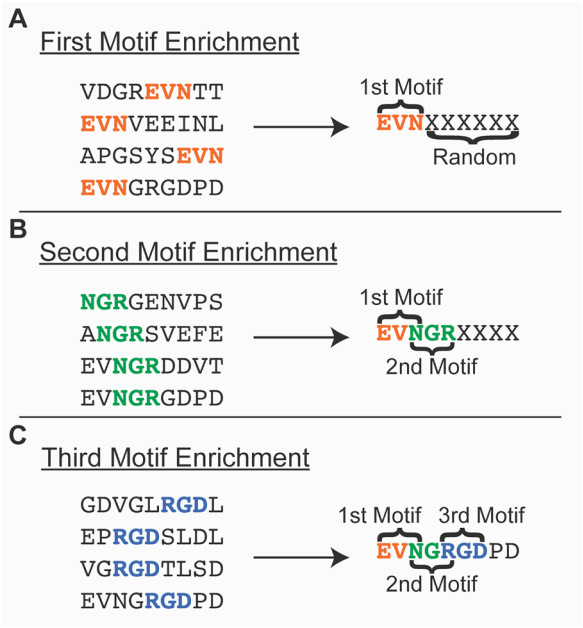

Figure 4: Theoretical accumulation of proposed EBUs during biopanning experiments to enrich promiscuous binding peptide ligands [11].

Through the power of landscape phage display and the principles of combinatorial chemistry, we may be able to identify ligands with increased promiscuity based on the presence of previously identified EBUs. After identification of the EBUs, targeting ligands may be constructed using rational design to create personalized precision medicines for improved treatment or monitoring of complex diseases like cancer.

REFERENCES

[1] Bertrand N, Wu J, Xu X, Kamaly N, Farokhzad OC. Cancer nanotechnology: the impact of passive and active targeting in the era of modern cancer biology. Adv Drug Deliv Rev 2014;66:2-25.

[2] Danhier F, Feron O, Préat V. To exploit the tumor microenvironment: Passive and active tumor targeting of

[†] BC = breast cancer

nanocarriers for anti-cancer drug delivery. J Control Release 2010;148:135-46.
[3] Torrice M. Does Nanomedicine Have a Delivery Problem? ACS Cent Sci 2016;2:434-7.
[4] Danhier F. To exploit the tumor microenvironment: Since the EPR effect fails in the clinic, what is the future of nanomedicine? J Control Release 2016;244:108-21.
[5] Petrenko VA, Gillespie JW. Paradigm shift in bacteriophage-mediated delivery of anticancer drugs: from targeted 'magic bullets' to self-navigated 'magic missiles'. Expert Opin Drug Deliv 2017;14:373-84.
[6] Smith GP, Petrenko VA. Phage Display. Chem Rev 1997;97:391-410.
[7] Jayanna PK, Bedi D, Deinnocentes P, Bird RC, Petrenko VA. Landscape phage ligands for PC3 prostate carcinoma cells. Protein Eng Des Sel 2010;23:423-30.
[8] Fagbohun OA, Bedi D, Grabchenko NI, Deinnocentes PA, Bird RC, Petrenko VA. Landscape phages and their fusion proteins targeted to breast cancer cells. Protein Eng Des Sel 2012;25:271-83.
[9] Gillespie JW, Wei L, Petrenko VA. Selection of Lung Cancer-Specific Landscape Phage for Targeted Drug Delivery. Combinatorial Chemistry & High Throughput Screening 2016;19:412-22.
[10] Bedi D, Gillespie JW, Petrenko VA. Selection of pancreatic cancer cell-binding landscape phages and their use in development of anticancer nanomedicines. Protein Engineering Design & Selection 2014;27:235-43.
[11] Gross AL, Gillespie JW, Petrenko VA. Promiscuous tumor targeting phage proteins. Protein Engineering Design & Selection 2016;29:93-103.
[12] Petrenko VA, Jayanna PK. Phage protein-targeted cancer nanomedicines. FEBS Lett 2014;588:341-9.
[13] Ruoslahti E. RGD and other recognition sequences for integrins. Annu Rev Cell Dev Biol 1996;12:697-715.
[14] Kuzmicheva GA, Jayanna PK, Sorokulova IB, Petrenko VA. Diversity and censoring of landscape phage libraries. Protein Eng Des Sel 2009;22:9-18.
[15] Petrenko VA, Smith GP, Gong X, Quinn T. A library of organic landscapes on filamentous phage. Protein Eng 1996;9:797-801.
[16] Brinton LT, Bauknight DK, Dasa SS, Kelly KA. PHASTpep: Analysis Software for Discovery of Cell-Selective Peptides via Phage Display and Next-Generation Sequencing. PLoS One 2016;11:e0155244.
[17] Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 2011;17.
[18] Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 2000;16:276-7.
[19] Davey NE, Seo MH, Yadav VK, Jeon J, Nim S, Krystkowiak I, Blikstad C, Dong D, Markova N, Kim PM, Ivarsson Y. Discovery of short linear motif-mediated interactions through phage display of intrinsically disordered regions of the human proteome. FEBS J 2017;284:485-98.