TARGET ANTIGEN DISCOVERY, PHENOTYPIC SCREENS AND USE THEREOF FOR IDENTIFICATION OF TARGET CELL SPECIFIC TARGET EPITOPES

RELATED APPLICATIONS

This application is claims priority to related to U.S. Provisional Application No. 61/840,583, filed June 28, 2013, entitled "Target Antigen Discovery and Phenotypic Screens", the contents of which is incorporated herein in their entirety.

BACKGROUND

Binding polypeptides, such as antibodies and fragments thereof, are commercially important as therapeutic and diagnostic agents. Traditional methods of screening for binding polypeptide generally employ soluble antigens. However, for certain cell-surface antigens, conformational epitopes on these antigens are altered when the antigens are solubilized from the plasma membrane, resulting in a failure to generate binding polypeptides that can recognize the native antigen. Accordingly, there is a need in the art for novel methods of screening for binding polypeptides that can specifically bind to cell-surface antigens in their native conformation.

SUMMARY

The invention provides methods and compositions for identifying binding polypeptides (e.g., antibodies or antigen binding fragments thereof) that specifically binds to a cell-surface antigen. The methods of the invention generally comprise contacting a variegated nucleic acid-display library of binding polypeptides with a cell-surface antigen displayed on the exterior surface of a cell; and isolating from the library at least one library member that specifically binds to the cell-surface antigen on the exterior surface of the cell. The methods and compositions of the invention are particularly advantageous in that they allow for the rapid identification of binding polypeptides that bind to native forms of the target cell surface antigen. These methods and compositions also allow for identification of novel, therapeutically useful cell-type specific antigens or epitopes.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic of exemplary DNA display compositions and screening methods of the invention.

Figure 2 is a schematic of exemplary DNA display compositions and screening methods of the invention.

Figure 3 is a schematic of exemplary target cell screening strategies employed in the methods of the invention.

Figure 4 is a schematic of exemplary parallel screening and deep sequencing strategies employed in the methods of the invention.

Figure 5 is a schematic of exemplary parallel screening and deep sequencing strategies employed in the methods of the invention.

Figure 6 is a schematic of exemplary parallel screening and deep sequencing strategies employed in the methods of the invention.

Figure 7 is a schematic of the results of exemplary parallel screening strategies employed in the methods of the invention.

Figure 8 depicts a graph showing the results of a FACS based binding assay of high affinity VH molecules selected using the methods of the invention.

Figure 9 depicts graphs showing the differential binding of VH molecules selected using the methods of the invention.

Figure 10 depicts graphs showing the differential binding of VH molecules selected using the methods of the invention.

DETAILED DESCRIPTION

I. Definitions

As used herein, the term "nucleic acid display library" refers to any art recognized *in vitro* cell-free phenotype-genotype linked display, including, without limitation those set forth in, for example, U.S. Patent Nos. 7,195,880; 6,951,725; 7,078,197; 7,022,479; 6,518,018; 7,125,669; 6,846,655; 6,281,344; 6,207,446; 6,214,553; 6,258,558; 6,261,804; 6,429,300; 6,489,116; 6,436,665; 6,537,749; 6,602,685; 6,623,926; 6,416,950; 6,660,473; 6,312,927; 5,922,545; and 6,348,315, and in WO2010/011944, which are all hereby incorporated by reference in their entirety.

As used herein, the term "antigen" refers to the molecule recognized by a binding polypeptide.

As used herein, the term "specifically binds to" refers to the ability of a binding molecule (e.g., a VH or VL domain) to bind to an antigen with an affinity of at least about 1 x 10^{-6} M, 1 x 10^{-7} M, 1 x 10^{-8} M, 1 x 10^{-9} M, 1 x 10^{-10} M, 1 x 10^{-11} M, 1 x 10^{-12} M, or more,

and/or bind to a target with an affinity that is at least two-fold greater than its affinity for a nonspecific antigen.

As used herein, the term "antibody" refers to immunoglobulin molecules comprising four polypeptide chains, two heavy (H) chains and two light (L) chains inter-connected by disulfide bonds, as well as multimers thereof (e.g., IgM). Each heavy chain comprises a heavy chain variable region (abbreviated VH) and a heavy chain constant region. The heavy chain constant region comprises three domains, CH1, CH2 and CH3. Each light chain comprises a light chain variable region (abbreviated VL) and a light chain constant region. The light chain constant region comprises one domain (CL1). The VH and VL regions can be further subdivided into regions of hypervariability, termed complementarity determining regions (CDRs), interspersed with regions that are more conserved, termed framework regions (FR).

As used herein, the term "antigen-binding portion" of an antibody includes any naturally occurring, enzymatically obtainable, synthetic, or genetically engineered polypeptide or glycoprotein that specifically binds an antigen to form a complex. Antigen-binding fragments of an antibody may be derived, e.g., from full antibody molecules using any suitable standard techniques such as proteolytic digestion or recombinant genetic engineering techniques involving the manipulation and expression of DNA encoding antibody variable and optionally constant domains. Non-limiting examples of antigen-binding portions include: (i) Fab fragments; (ii) $F(ab')_2$ fragments; (iii) Fd fragments; (iv) Fv fragments; (v) single-chain Fv (scFv) molecules; (vi) dAb fragments; and (vii) minimal recognition units consisting of the amino acid residues that mimic the hypervariable region of an antibody (e.g., an isolated complementarity determining region (CDR)). Other engineered molecules, such as diabodies, triabodies, tetrabodies and minibodies, are also encompassed within the expression "antigen-binding portion."

As used herein, the terms "VH domain" and "VL domain" refer to single antibody variable heavy and light domains, respectively, comprising FR (Framework Regions) 1, 2, 3 and 4 and CDR (Complementary Determinant Regions) 1, 2 and 3 (see Kabat et al. (1991) Sequences of Proteins of Immunological Interest. (NIH Publication No. 91-3242, Bethesda).

II. Cell Surface Antigens

In certain aspects, the invention provides methods of identifying a binding polypeptide that specifically binds to a cell-surface antigen.

Any antigen that is capable of being displayed on the surface of a cell can be employed in the methods of the invention, including without limitation, protein, glycan, and/or lipid antigens. In certain embodiments, the antigen is a naturally occurring molecule. Suitable, non-limiting examples of naturally occurring antigens include transmembrane proteins (e.g., G-protein coupled receptors) and GPI-anchored proteins. In certain embodiments, the antigen is a non- naturally occurring recombinant or synthetic antigen. Suitable, non-limiting examples of naturally occurring antigens include chimeric antigens comprising portions from different antigen molecules. In certain embodiments, the identity of the antigen is known prior to preforming the methods of the invention. In certain embodiments, the identity of the antigen is unknown prior to preforming the methods of the invention.

The cell surface antigens employed in the methods of the invention can be displayed on any cell or cell-like particle (e.g., lipid vesicle). In certain embodiments, the cell is a cell type that naturally expresses the cell-surface antigen. In certain embodiments, the cell is a recombinant cell that is engineered to heterologously express the cell-surface antigen. In certain embodiments, the cell is a disease-associated variant of a normal cell (e,g, a tumor cell).

III. Binding Polypeptides

In certain aspects, the invention provides methods of identifying a binding polypeptide that specifically binds to a cell-surface antigen.

Any type of binding polypeptide can be employed in the methods of the invention including, without limitation antibodies, or fragments thereof, and immunoglobulin-like domains. Suitable immunoglobulin-like domains include, without limitation, fibronectin domains (see, for example, Koide et al. (2007), *Methods Mol. Biol.* 352: 95–109, which is incorporated by reference herein in its entirety), DARPin (see, for example, Stumpp et al. (2008) *Drug Discov. Today* 13 (15–16): 695–701, which is incorporated by reference herein in its entirety), Z domains of protein A (see, Nygren et al. (2008) *FEBS J.* 275 (11): 2668–76, which is incorporated by reference herein in its entirety), Lipocalins (see, for example, Skerra et al. (2008) *FEBS J.* 275 (11): 2677–83, which is incorporated by reference herein in its entirety), Affilins (see, for example, Ebersbach et al. (2007) *J. Mol. Biol.* 372 (1): 172–85, which is incorporated by reference herein in its entirety), Affitins (see, for example, Krehenbrink et al. (2008). *J. Mol. Biol.* 383 (5): 1058–68 , which is incorporated by reference herein in its entirety), Avimers (see, for example, Silverman et al. (2005) *Nat.*

Biotechnol. 23 (12): 1556–61 , which is incorporated by reference herein in its entirety), Fynomers, (see, for example, Grabulovski et al. (2007) *J Biol Chem* 282 (5): 3196–3204 , which is incorporated by reference herein in its entirety), and Kunitz domain peptides (see, for example, Nixon et al. (2006) *Curr Opin Drug Discov Devel* 9 (2): 261–8, which is incorporated by reference herein in its entirety). In certain embodiments, the binding polypeptide is antibody VH or VL domain.

IV. Cell Surface Display Methods

In certain aspects, the invention provides a method of identifying a binding polypeptide that specifically binds to a cell-surface antigen, the method comprising: (a) contacting a variegated nucleic acid-display library of binding polypeptides with a cell-surface antigen displayed on the exterior surface of a first cell type; and (b) isolating from the library at least one library member that specifically binds to the cell-surface antigen on the exterior surface of the first cell type, thereby identifying a binding polypeptide that specifically binds to the cell surface antigen. In certain embodiments, prior to step (a), the variegated nucleic acid-display library of binding polypeptides is contacted with a second cell type that does not display the antigen displayed on the exterior surface, in order to pre-clear the library of binding polypeptides that do not specifically bind to the antigen.

In certain aspects, the invention provides a method of identifying a binding polypeptide that specifically binds to a cell-surface antigen, the method comprising: (a) contacting a variegated nucleic acid-display library of binding polypeptides with a first cell type expressing a cell-surface antigen, and isolating from the library at least one library member that specifically binds to the first cell type; (b) contacting the variegated nucleic acid-display library of binding polypeptides with a second cell type that does not express the cell surface antigen, and isolating from the library at least one library member that specifically binds to the second cell type; and (c) selecting library members that specifically bind to the first cell type but not to the second cell type, thereby identifying a binding polypeptide that specifically binds to the cell surface antigen.

Suitable nucleic acid-display libraries for use in the methods of the invention are set forth in, for example, U.S. Patent Nos. 7,195,880; 6,951,725; 7,078,197; 7,022,479; 6,518,018; 7,125,669; 6,846,655; 6,281,344; 6,207,446; 6,214,553; 6,258,558; 6,261,804; 6,429,300; 6,489,116; 6,436,665; 6,537,749; 6,602,685; 6,623,926; 6,416,950; 6,660,473; 6,312,927; 5,922,545; and 6,348,315, and in WO2010/011944, which are all hereby incorporated by reference in their entirety. In certain embodiments, the variegated nucleic

acid-display library is a DNA display library. In one embodiment, the nucleic acid-display library is a DNA display library described herein or in WO2010/011944, which is hereby incorporated by reference in its entirety.

In certain embodiments, each member of the DNA-display library comprises a binding polypeptide linked through an intervening DNA linker to a DNA coding sequence encoding the binding polypeptide, wherein the DNA linker comprising a restriction endonuclease site (see e.g., Figure 1). Any restriction endonuclease site can be employed. In one particular embodiment, the restriction endonuclease site is not present in the DNA coding sequence of members of the DNA-display library, thus avoiding cleavage of the DNA coding sequence upon restriction endonuclease digestion of the library members. In one particular embodiment, the restriction endonuclease site is a Not1 site.

In certain embodiments, it is desirable to physically separate the DNA coding sequence of the isolated library members from the linked binding polypeptide. Any methods of physical separation can be employed. Where the isolated library members comprise a DNA linker comprising a restriction endonuclease site (see e.g., Figure 1), the physical separation can be achieved by restriction endonuclease digestion of the isolated library members. The resultant liberated DNA coding sequences can be further separated from the cell/binding polypeptide complexes by any art recognized method, e.g., centrifugation .

In certain embodiments, it is desirable to physically separate the intact isolated library members from the from the first and/or second cell type. Any methods of physical separation can be employed. In certain embodiments, the isolated library members are separated from the first or second cell type by enzymatic cleavage of the cell-surface antigen. Any methods of enzymatic cleavage of the antigen can be employed, e.g., protease, lipid, and/or glycosidase enzymatic cleavage. In certain embodiments, where the cell-surface antigen is attached to the cell surface by a glycolipid anchor, the isolated library members are separated from the first or second cell type by phospholipase cleavage of the glycolipid anchor. The resultant liberated isolated library members can be further separated from the first or second cell type by any art recognized method, e.g., centrifugation.

Once the library members that specifically bind to the first and/or second cell type have been isolated, the DNA coding sequence of these molecules can be determined. Accordingly, in certain embodiments, the methods of the invention further comprise the step of determining the DNA coding sequence of at least a portion of the isolated library members. Any art recognized means for DNA sequence determination can be employed. In one particular embodiment, the DNA coding sequence is determined by single molecule,

deep sequencing techniques (e.g., pyrosequencing). Single molecule, deep sequencing techniques are well known in the art (see e.g., those described in US6,210,891, which is hereby incorporated by reference in its entirety). In certain embodiments, where the binding polypeptides are antibodies, or antigen binding fragments thereof, the DNA coding sequence of the CDR3 region is determined. In certain embodiments, the DNA coding sequences of the library member that bind to the first and second cell types are determined. Library members that specifically bind to the first cell type but not to the second cell type are considered to comprise binding polypeptides that specifically bind to an antigen specific for the first cell type.

Once a binding polypeptide that specifically binds to the cell-surface antigen has been identified, it can be heterologously expressed *in vitro* (e.g., in cells or in a cell-free expression system) or *in vivo* (e.g., in a transgenic animal). Accordingly, in certain embodiments, the methods of the invention further comprise the step of heterologously expressing *in vitro* (e.g., in cells or in a cell-free expression system) or *in vivo* (e.g., in a transgenic animal), the identified binding polypeptide.

In certain embodiments, the identity of the antigen is known prior to preforming the methods of the invention. However, it is not necessary to know the identity of the antigen. Indeed, in certain embodiments, the identity of the antigen is unknown prior to preforming the methods of the invention. Thus, in this latter case, the methods of the invention allow for the identification of novel antigens and epitopes present on the surface of a cell type of interest (e.g., a tumor cell).

In certain embodiments, the methods disclosed herein comprise the selection of binding polypeptides that are capable of functional internalization upon binding to the cell surface antigen. Such binding polypeptides are particularly useful in the production of drug conjugates because they allow for the delivery of a cytotoxic drug to the interior of a target cell. Any methodology for screening for functional internalization can be employed. For example, the variegated nucleic acid-display library of binding polypeptides can be contacted with target cells under conditions that to allow for binding polypeptide internalization (e.g., for about for 1-2 hours at 37°C). The cells can then washed and lysed with cell lysis buffer in the presence of protease inhibitors. The internalized library members can then be ethanol precipitated and the DNA coding sequences enriched by PCR amplification.

The methods disclosed herein can be applied to any target epitope discovery process. For example, target epitopes can include: homing domains for inflammation; tumor specific target epitopes from primary tumors with or without resistance to treatment, tumor cell lines,

and tumors that harbor any mutations that may result in neoepitopes; and other disease specific epitopes that mediate disease specific malfunction and be targeted for biologic therapy

The methods disclosed herein can also be applied for biomarker discovery to monitor the presence or absence of particular cell surface epitopes over a course of a drug treatment to patients. The antibodies derived from the biomarker discovery can also be used as tools for biomarker detection.

Additionally or alternatively, the methods disclosed herein can be applied to target or epitope discovery in other species, such as in transgenic animals and animal disease models.

V. Exemplification

A. Summary

A fully human antibody VH library obtained from bone marrow, peripheral blood and splenocytes of human donors was constructed and numerous high affinity and specific VH binders to multiple targets were identified using dsDNA display technology. Target cell specific epitopes were identified by live cell selection and deep sequencing analysis using human VH library and dsDNA display technology. Strategies of parallel, differential selections and target cell selective selections were applied in these methods (see Figures 4, 5, 6). The tabulation of the CDR3 frequency from deep sequencing of all pools across all rounds predicted the selectivity of the VH clones. High affinity VHs were identified that selectively bind to target cells and related cell types.

B. Library engineering and selection methods for live cell epitope discovery

Two methods were developed to effectively recover library members that bound to live cells.

The first method involved stripping off binders from live cells by restriction digestion of the DNA fused to bound antibodies. This method enables the full recovery of the VHs bound to all epitopes on cells. A C-terminal of VH DNA library was engineered to carry a NotI restriction site (see Figure 1). There are no NotI sites in naïve VH frameworks and therefore only full length VH binders are eluted from cells for subsequent amplification. The NotI restriction digestion buffer was tested on live cells, and with up to a 2 hour incubation at 37C the cells were viable. The efficiency of the NotI digestion was high. Following the binding of the library to cells for 1 hour at 4C, the cells were washed and digested with NotI

buffer at 37C for 1 hour, the cells were then spun down, supernatant (containing bound VH DNA) was collected for PCR amplification (see Figure 1).

The second method is to clip off VH binders from live cells using phospholipase C (PLC) digestion. This method enables the elution of the VHs that bound to the epitopes of any GPI anchored membrane protein (i.e., a subset of epitopes). The PLC clipping efficiency is high, as validated on FACS with control molecule. After incubation of library with cells for 1 hour at 4C, the cells were washed and incubated with PLC at 37C for 15 mins. The cells were subsequently spun down and the supernatant, containing fusion VH complexed with extracellular domain of the GPI anchored membrane protein, was PCR amplified (see Figure 2).

C. Parallel selections, differential selections and target cell selective selections on target and related/ undesired cell types

Master naïve fusion library was produced according to the protocol set forth in WO2010/011944 (which is hereby incorporated by reference in its entirety). For first round of selection, the purified fusion library was split equally into multiple libraries that carry the same diversity for all selection branches (see Figure 5).

Primary cells, obtained from normal donors or patients, were either thawed fresh or isolated from cell culture flasks, following standard cell biology protocols. The cells were then recovered for 1 hour in full media at 37C followed by blocking in selection buffer for 30 mins on ice. All the selections were carried out on ice to prevent antibody and target internalization.

For parallel selections, libraries were pre-cleared with 200 ul of pre-blocked streptavidin beads and 200 ul of pre-blocked hIgG-Epoxy beads for 30 mins at room temperature sequencially to remove any misfold and sticky library members. The pre-cleared libraries were then chilled on ice and subjected to pre-blocked cells and incubated on ice for 1 hour.

For target cell selective selections, pre-clearance was performed on undesired and closely related cell types for 1 hour on ice to remove any non-selective binders and then subjected on target cells.

At selection round 4, differential selection methods were applied to the branches of selection on target cells (with and without pre-clearance on cells). In this round, libraries were split into multiple tubes and bound to each cell type and patient's cells from different stage of the diseases in parallel. This strategy allowed for direct comparison of target cells

versus other cell types by deep sequencing analysis and identification of binders recognizing different epitopes that arose with disease progression (see Figure 6).

For all selection branches, after binding, cells were washed with 10 mL of binding buffer and subject to either NotI restriction digestion to recover all binders to membrane protein or PLC clipping to recover binders to GPI anchored membrane proteins as described above.

D. Deep Sequencing analysis to predict selective binders to target cells

After each round of selection, binding pools were PCR amplified. The HCDR3 of each binding pool was lifted up by PCR with oligos priming C-terminal of framework 3 and N-terminal of framework 4 of VH fragments. These HCDR3 fragments derived from individual selection rounds and branches were then tagged with specific DNA bar code used for Illumina sequencing by PCR. The tagged HCDR3 were pooled and sent for high throughput sequencing with Hi Seq technology. The round 4 binding pools from target cells were also tagged with DNA bar code and submitted for 454 sequencing to get full length VH sequences.

The sequences were deconvoluted based on the DNA bar code after sequencing. Millions of sequences derived from each selection round and selection branch were tabulated by comparing the frequency of a particular CDR3 sequence present at different rounds and selection branches. The criteria used for identification of selective binders were: 1) specific enrichment of a CDR3 sequence from earlier round to later round on target cells, not on control or close related cell types; 2) higher frequency on target specific cell type and low on control or closely related cell type at differential selection round (see Figure 7); and 3) sequences not present in other target or cell selections from other programs in database. The selective clones identified by Illumina sequencing were then synthesized based on the 454 full length sequence information.

E. Production, purification and FACS binding assays

The binding pools and synthesized VHs were then subcloned into pET22b expression vectors. The VHs were produced in BL-21 E.coli cells and purified through C-terminal His tag using standard protocols. FACS assay was performed to assess the binding and selectivity of VHs to different cell types and the EC50 of the binders. High affinity and selective VH binders were identified through the live cell selection process (see Figures 8, 9 and 10).

We Claim:

1. A method of identifying a binding polypeptide that specifically binds to a cell-surface antigen, the method comprising:
    a. contacting a variegated nucleic acid-display library of binding polypeptides with a cell-surface antigen displayed on the exterior surface of a first cell type; and
    b. isolating from the library at least one library member that specifically binds to the cell-surface antigen on the exterior surface of the first cell type, thereby identifying a binding polypeptide that specifically binds to the cell surface antigen.

2. The method of claim 1, wherein prior to step (a), the variegated nucleic acid-display library of binding polypeptides is contacted with a second cell type that does not display the antigen displayed on the exterior surface.

3. A method of identifying a binding polypeptide that specifically binds to a cell-surface antigen, the method comprising:
    a. contacting a variegated nucleic acid-display library of binding polypeptides with a first cell type expressing a cell-surface antigen, and isolating from the library at least one library member that specifically binds to the first cell type;
    b. contacting the variegated nucleic acid-display library of binding polypeptides with a second cell type that does not express the cell surface antigen, and isolating from the library at least one library member that specifically binds to the second cell type; and
    c. selecting library members that specifically bind to the first cell type but not to the second cell type, thereby identifying a binding polypeptide that specifically binds to the cell surface antigen.

4. The method of claim 1, 2, or 3, wherein the variegated nucleic acid-display library is a DNA display library.

5. The method of claim 4, wherein each member of the DNA-display library comprises a binding polypeptide linked through an intervening DNA linker to a DNA coding

sequence encoding the binding polypeptide, wherein the DNA linker comprising a restriction endonuclease site.

6. The method of claim 5, wherein the restriction endonuclease site is not present in the coding sequence of members of the DNA-display library.

7. The method of claim 5 or 6, wherein the method further comprises physically separating the DNA coding sequence and the linked binding polypeptide of the isolated library members.

8. The method of any one of the preceding claims, wherein the method further comprises separating the intact isolated library members from the first or second cell type.

9. The method of claim 8, wherein the isolated library members are separated from the first or second cell type by enzymatic cleavage of the cell-surface antigen.

10. The method of claim 9, wherein the cell-surface antigen is attached to the cell surface by a glycolipid anchor and the isolated library members are separated from the first or second cell type by phospholipase cleavage of the glycolipid anchor.

11. The method of any one of the preceding claims, wherein the method further comprises determining the DNA coding sequence of at least a portion of the isolated library members.

12. The method of claim 11, wherein the DNA coding sequence is determined by pyrosequencing.

13. The method of any one of the preceding claims, where the antigen is a naturally occurring protein, glycan or lipid.

14. The method of any one of the preceding claims, where the antigen is a glycophosphatidylinositol (GPI) anchored protein.

15. The method of any one of the preceding claims, where the antigen is a recombinant antigen.

16. The method of any one of the preceding claims, where the antigen is a chimeric antigen.

17. The method of any one of the preceding claims, wherein the first cell type is a cell that naturally expresses the cell-surface antigen.

18. The method of any one of the preceding claims, wherein the first cell type is a recombinant cell that is engineered to heterologously express the cell-surface antigen.

19. The method of any one of the preceding claims, wherein the first cell type is a disease-associated variant of a normal cell.

20. The method of any one of the preceding claims, wherein the first cell type is a tumor cell.

21. The method of any one of the preceding claims, wherein the binding polypeptide is an antibody of fragment thereof.

22. The method of any one of the preceding claims, wherein the binding polypeptide is an antibody VH or VL domain.