

Non-human Animals that Make Single Domain Binding Proteins

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application Serial No. 61/968,986, filed 21 March 2014, and U.S. Provisional Application Serial No. 61/968,905, filed 21 March 2014, both of which applications are hereby incorporated by reference.

FIELD OF INVENTION

[0002] Non-human animals are provided that exhibit high diversity in the immunoglobulin heavy chain locus, and preferably very low diversity in the immunoglobulin light chain locus, which allows for selection of single domain antigen binding proteins, including V_H -single domain binding proteins and V_L single domain binding proteins, that bind antigen.

BACKGROUND

[0003] In most animals, normal immunoglobulin heavy chains are only well-expressed when coupled with their cognate light chains. In humans, lone heavy chains are found in heavy chain disease that is manifested by dysfunctional heavy chains that lack sequences of the variable heavy, the C_H1 , or the variable heavy and C_H1 domains. Heavy chains devoid of light chains are encountered in certain species of fish and in camels. Such heavy chains lack a functional C_H1 domain and have non-human features in their heavy chain variable domains. Attempts have been made to make camelized antibodies by modifying mice to express camelized genes that mimic V_{HH} domains found in camels or certain species of fish, in part by removal of IgM and IgG C_H1 domains and conforming the heavy chain variable regions to resemble those of camels and/or certain species of fish. Unfortunately, camelized antibodies would be expected to induce immune responses in non-camelid animals. Another challenge with previous versions of non-human animals genetically modified to comprise an inactivated C_H1 domain is the reduced expression levels of antigen-specific single domain antigen binding proteins, compared to traditional antibodies. Such reduction may be due to a lack of mechanisms available to non-camelid heavy chain variable regions that allow the heavy chain variable regions to compensate for the absence of a V_L domain. For example, camelid V_{HH} domains found in heavy chain-only binding proteins comprise a CDRH3 that is, on average, longer than those found in non-camelid antibodies, considered to be a major influence on overall

antigen affinity and specificity, and thought to compensate for the absence of a V_L domain in the camelid heavy chain-only antibody.

[0004] Thus there is a need in the art for genetically modified non-human animals that make diverse single domain binding proteins that have non-camelid V_H domains.

SUMMARY

[0005] As disclosed herein, immunoglobulin polypeptide chains comprising a light chain variable region and a heavy chain constant region may be expressed by non-human animals and form V_L -single domain antigen binding proteins, e.g., single domain antigen binding proteins comprising light chain variable domains operably linked to heavy chain constant domains, wherein the heavy chain constant domain(s) lack a functional C_H1 domain, e.g., of an immunoglobulin heavy chain constant region selected from IgG, IgA, IgE, IgD, or a combination thereof. The V_L -single domain antigen binding proteins may exhibit increased stability compared to V_H -single domain antigen binding proteins comprising heavy chain variable domains operably linked to heavy chain constant domains lacking a functional C_H1 domain. Accordingly, provided herein are non-human animals capable of expressing a V_L -single domain antigen binding protein comprising a light chain variable domain and a heavy chain constant region, wherein the heavy chain constant region lacks a functional C_H1 domain, and may also optionally lack a functional hinge region and methods of making and using non-human animals expressing V_L -single domain antigen binding proteins. Also provided are cells, proteins and nucleic acids derived from non-human animals expressing V_L -single domain antigen binding proteins, and use of the isolated cells, proteins and nucleic acids.

[0006] Also disclosed herein, the expression of a single rearranged light chain by non-human animals capable of producing single domain antigen binding proteins, e.g., V_L - or V_H -single domain antigen binding proteins, increases the titer of antigen specific single domain antigen binding proteins in response to antigen challenge compared to a similar non-human animal capable of producing single domain antigen binding proteins that do not express the single rearranged light chain. **FIG. 5.** This data suggests that the presence of the single rearranged light chain by non-human animals increases the likelihood of generating an antigen-specific single domain antigen binding protein. Accordingly, provided herein are non-human animals capable of expressing single domain antigen binding proteins (e.g., V_H - and/or V_L -single domain antigen binding proteins) and a single rearranged light chain and methods of making and using non-human animals expressing V_L -single domain

antigen binding proteins. Also provided are cells, proteins and nucleic acids derived from non-human animals expressing single domain antigen binding proteins and a single rearranged light chain, and use of the isolated cells, proteins and nucleic acids.

[0007] Also provided herein are non-human animals comprising V_L -single domain antigen binding proteins and a single rearranged light chain, methods of making non-human animals capable of producing a high titer of single domain antigen binding proteins and/or increasing the production of single domain binding proteins by non-human animals capable of producing such binding proteins, methods of using the non-human animals to make antigen-specific single domain antigen binding proteins, and single domain antigen binding proteins so made.

[0008] Genetically modified cells, non-human embryos, non-human animals and methods and compositions for making and using them are provided, wherein the animals are genetically modified to produce single domain antigen binding proteins, e.g., binding proteins comprising a heavy chain constant region that lacks a functional C_H1 sequence, and also optionally lack a functional hinge region sequence, and wherein the animals are further genetically modified to express the single domain antigen binding protein as a V_L -single domain antigen binding protein (e.g., encoded from a rearranged light chain variable region nucleotide sequence operably linked to a heavy chain constant region nucleic acid sequence modified to inactivate or delete a C_H1 domain encoding sequence) and/or a single rearranged light chain (e.g., encoded from a single rearranged $V_L:J_L$ sequence operably linked to a light chain constant region in the animal's germline).

[0009] The animals as disclosed herein may produce single domain binding proteins, which, in one aspect, comprise an IgG isotype, such as, e.g., the IgG1 isotype. In some embodiments, the single domain antigen binding protein is a V_H -single domain antigen binding protein, e.g., comprises a heavy chain variable region operably linked to a heavy chain constant region lacking a functional C_H1 . In other embodiments, the single domain antigen binding protein is a V_L -single domain antigen binding protein, e.g., comprises a light chain variable domain operably linked to a heavy chain constant region lacking a functional C_H1 , e.g., a heavy chain constant region comprising hinge, C_H2 , C_H3 , C_H4 , or a combination thereof.

[0010] Accordingly, in some aspects, a single domain antigen binding protein as described herein is encoded by a nucleic acid sequence derived from one or more unrearranged immunoglobulin light chain V segments and one or more unrearranged immunoglobulin light chain J segments operably linked to a heavy chain constant region (e.g., a heavy chain constant region domain selected from the group

consisting of C_H1, hinge, C_H2, C_H3, C_H4, and combination thereof), wherein the heavy chain constant region comprises a deletion or an inactivating mutation in a C_H1 region sequence. In one embodiment, the unrearranged light chain V and J segments replace one or more, substantially all, or all functional endogenous non-human immunoglobulin heavy chain variable region gene segments at the endogenous non-human immunoglobulin heavy chain locus. In some embodiments, a heavy chain locus modified to comprise light chain variable region gene segments operably linked to a heavy chain constant region comprising a deletion or inactivating mutation in a C_H1 region sequence as disclosed herein may be found in the germline of the non-human animal. Such a modified locus can be at the endogenous heavy chain locus, or present in a transgene at a locus other than the endogenous heavy chain locus (e.g., inserted at a random position in the genome).

[0011] In one aspect, animals disclosed herein comprising a nucleic acid sequence encoding a V_L-single domain antigen binding protein (e.g., a nucleic acid sequence derived from one or more unrearranged immunoglobulin light chain V segments and one or more unrearranged immunoglobulin light chain J segments operably linked to a heavy chain constant region comprising a deletion or an inactivating mutation in a C_H1 region sequence) may further comprise a second immunoglobulin polypeptide chain comprising a light chain variable region and a light chain constant region, which may be encoded by a second nucleic acid sequence comprising a human light chain V segment and a human light chain J segment operably linked to a light chain constant region. Such second nucleic acid sequence may also be found in the germline of the non-human animal. Such a nucleic acid sequence can be present at the endogenous light chain locus, or present in a transgene at a locus other than the endogenous light chain locus (e.g., inserted at a random position in the genome).

[0012] In another aspect, animals modified to comprise a nucleic acid sequence derived from one or more unrearranged immunoglobulin light chain V segments and one or more unrearranged immunoglobulin light chain J segments operably linked to a heavy chain constant region comprising a deletion or an inactivating mutation in a C_H1 region sequence may be further modified to express a single rearranged light chain, e.g., a common light chain (ULC).

[0013] In some embodiments, the single domain antigen binding protein such as, but not limited to, a V_L-single domain antigen binding protein, and/or single rearranged light chain comprises human idiotypes. For example, a single domain antigen binding protein and/or a genetically engineered single rearranged light chain as disclosed herein may comprise a human variable domain and, in one

embodiment, a non-human constant domain. In one embodiment, the non-human constant domain is an endogenous non-human constant domain. In one embodiment, the non-human constant domain is a rodent constant domain, e.g., a murine constant domain, e.g., a mouse constant domain. In another embodiment, the constant domain is a human constant domain. In one aspect, the single domain antigen binding protein is a V_L -single domain antigen binding protein comprising a human light chain variable domain and a non-human heavy chain constant domain. In one embodiment, the unrearranged light chain V and/or J segments encoding a V_L -single domain antigen binding protein as disclosed herein are human segments.

[0014] Animals genetically modified to produce the single domain antigen binding proteins as disclosed herein may comprise a heavy chain locus having a replacement of one or more, or all, endogenous immunoglobulin heavy chain variable region gene segments with one or more unrearranged human immunoglobulin heavy chain variable region gene segments, or one or more unrearranged human immunoglobulin light chain V segments and one or more unrearranged human immunoglobulin light chain J segments. In some aspects, all endogenous V_H , D_H , and J_H gene segments are replaced with one or more unrearranged human V_H , one or more unrearranged human D_H , and one or more unrearranged human J_H gene segments. In other aspects, all endogenous V_H , D_H , and J_H gene segments are replaced with one or more unrearranged human immunoglobulin light chain V_L gene segments and one or more unrearranged human immunoglobulin light chain J_L gene segments, e.g., human kappa (κ) V_κ and/or J_κ gene segments and/or human lambda (λ) V_λ and/or J_λ gene segments.

[0015] Animals genetically modified to produce the single domain binding protein that comprises a heavy chain variable region or a light chain variable region in the context of a heavy chain constant region comprising a deletion of a C_H1 region and/or hinge region may bear the modification (and/or other modifications of a constant gene locus described herein) at an endogenous heavy chain locus, or may bear the modification on a transgene, wherein the transgene is positioned anywhere in the genome, e.g., introduced into the genome by random insertion. In some embodiments, the modified heavy chain locus as described herein may be found in the germline of the animal. In animals also modified to express a single rearranged light chain, the single rearranged light chain variable region can be operably linked to a light chain constant region at the endogenous light chain locus, or can be present in a transgene comprising the single rearranged light chain variable region operably linked with a syngeneic (e.g., autologous; with respect to the non-human animal) or

heterologous light chain constant region and present at a locus other than the endogenous light chain locus, e.g., randomly inserted into the genome.

[0016] In further embodiments, the heavy chain loci of the animals disclosed herein may comprise a deletion or inactivating mutation in the hinge region(s).

[0017] Further, animals disclosed herein may be modified to comprise and/or express a single rearranged light chain variable gene sequence operably linked to a light chain constant region, also referred to as common or universal light chain (ULC), which may be encoded by a light chain locus comprising a single rearranged $V_L:J_L$ gene sequence. In some embodiments, the light chain locus comprises a single rearranged $V_L:J_L$ gene sequence in which the V_L sequence is a V_κ gene sequence. In some aspects, the V_κ sequence is selected from $V_{\kappa 1-39}$ or $V_{\kappa 3-20}$. In some aspects, the J_L sequence is a J_κ gene sequence, e.g., a $J_{\kappa 1}$ sequence, a $J_{\kappa 2}$ sequence, a $J_{\kappa 3}$ sequence, a $J_{\kappa 4}$ sequence, or a $J_{\kappa 5}$ sequence, etc. In some embodiments, the light chain locus comprises a single rearranged $V_\kappa:J_\kappa$ sequence selected from the group consisting of $V_{\kappa 1-39}J_{\kappa 5}$ and $V_{\kappa 3-20}J_{\kappa 1}$. In one embodiment, the light chain locus comprises a single rearranged $V_\kappa:J_\kappa$ sequence of $V_{\kappa 1-39}J_{\kappa 5}$. In another embodiment, the light chain locus comprises a single rearranged $V_\kappa:J_\kappa$ sequence of $V_{\kappa 3-20}J_{\kappa 1}$. In some embodiments, the single rearranged variable gene sequence is operably linked to a non-human light chain constant region gene, e.g., endogenous non-human light constant region gene. In another embodiment, the single rearranged variable gene sequence is operably linked to a human light chain constant region gene. In some aspects, the single rearranged variable gene sequence is a human $V:J$ sequence inserted to the endogenous immunoglobulin light chain locus such that the resulting non-human animal does not comprise functional unrearranged V and/or J gene segments in one or more light chain loci.

[0018] Accordingly, provided herein are non-human animals bearing a heavy chain constant region comprising a deletion or inactivating mutation in a C_H1 encoding region and either or both (a) a light chain variable region in the context of the heavy chain constant region comprising a deletion of or inactivating mutation in a C_H1 region and (b) the single rearranged light chain. For example, a non-human animal as disclosed herein may comprise a nucleic acid sequence derived from one or more unrearranged immunoglobulin light chain V segments and one or more unrearranged immunoglobulin light chain J segments operably linked to a heavy chain constant region comprising a deletion or an inactivating mutation in a C_H1 region sequence as described herein. In one aspect, a non-human animal as

disclosed herein comprises a deletion or an inactivating mutation in a nucleic acid sequence encoding an immunoglobulin C_H1 domain and a single rearranged light chain variable gene sequence operably linked to a light chain constant region as disclosed herein. In another aspect, a non-human animal as disclosed herein comprises a nucleic acid sequence derived from one or more unrearranged immunoglobulin light chain V segments and one or more unrearranged immunoglobulin light chain J segments operably linked to a heavy chain constant region comprising a deletion or an inactivating mutation in a C_H1 region sequence and a single rearranged light chain variable gene sequence operably linked to a light chain constant region. In some embodiments, the heavy chain constant region is a non-human constant region, e.g., an endogenous non-human constant region. In other embodiments, the heavy chain constant region is a human constant region.

[0019] In some aspects, a non-human animal comprises the modified heavy chain loci and/or genetically engineered rearranged light chain loci as disclosed herein in its germline. The non-human animal may also comprise a deletion or inactivating mutation in one or more of the following immunoglobulin genes: IgD, IgG3, IgG2a, IgG2b, IgG2c, IgE, IgA, and a combination thereof. In one embodiment, the non-human animal comprises a deletion or inactivating mutation in the IgG2a and IgG2b immunoglobulin genes. In another embodiment, the non-human animal comprises a deletion or inactivating mutation in the IgG3, IgD, IgA, and IgE immunoglobulin genes. In another embodiment, the non-human animal comprises a deletion or inactivating mutation in the IgG3, IgD, IgG2a, IgG2b, IgA, and IgE immunoglobulin genes.

[0020] In some aspects, a non-human animal as disclosed herein may further comprise an Adam6a gene (or fragment thereof) and/or an Adam6b gene (or fragment thereof) capable of retaining fertility of a male non-human animal. The Adam6a gene, Adam6b gene, or both may be placed ectopically, or may be at a position that approximates the position of the Adam6 gene(s) in the non-human animal. The Adam6a gene, Adam6b gene, or both are functional in a male non-human animal. For example, the non-human animal is a rodent (e.g., a mouse or a rat) and the Adam6a gene, Adam6b gene, or both are mouse or rat genes, respectively. In various embodiments, maintenance or insertion of the Adam6 gene(s) maintains or confers fertility on the male non-human animal (e.g., on the male mouse or rat).

[0021] In one aspect, the non-human animals disclosed herein comprise an IgM immunoglobulin encoded by an IgM gene sequence comprising a functional C_H1 domain encoding sequence, which may be associated with a cognate light chain,

e.g., a genetically engineered single rearranged light chain. In another embodiment, the non-human animal produces only heavy chains having an IgM and IgG1 isotype, wherein the IgM heavy chains comprise a functional C_H1 domain while the IgG1 heavy chains lack a functional C_H1 domain. In one aspect, the cognate light chain associated with the IgM heavy chain is encoded by or derived from a single rearranged light chain variable gene sequence operably linked to a light chain constant region.

[0022] Expression of the genetically engineered single rearranged light chain as disclosed herein results in the production of a high titer of antigen specific single domain antigen binding proteins after antigen challenged by the non-human animals. A titer, e.g., an antibody or binding protein concentration, e.g., as measured by ELISA, of at least 1×10^2 µg/mL, at least 1×10^3 µg/mL, at least 1×10^4 µg/mL, or at least 1×10^5 µg/mL may be considered a high titer. Alternatively, a non-human animal produces a high titer of binding protein if the binding protein concentration is at least 2-fold, at least 5-fold, at least 10-fold, or at least 100-fold the concentration of a corresponding control animal not comprising the genetically engineered rearranged light chain.

[0023] Methods of producing a non-human animal as disclosed herein are also provided. Such methods comprise modifying the non-human heavy chain constant region of the non-human animal such that the heavy chain constant region comprises a deletion or an inactivating mutation of a nucleotide sequence encoding a C_H1 domain, e.g., an IgG1 C_H1 domain.

[0024] Methods of producing a non-human animal as disclosed herein may further comprise replacing at an endogenous immunoglobulin heavy chain locus, one or more, all, or substantially all endogenous non-human heavy chain variable region gene segments with one or more unrearranged light chain V and/or one or more unrearranged light chain J gene segments such that the light chain V and J gene segments are operably linked to a heavy chain constant region comprising a deletion or an inactivating mutation of a nucleotide sequence encoding a C_H1 domain. In one embodiment, the unrearranged light chain V and J gene segments are capable of undergoing productive rearrangement, e.g., comprise recombination signal sequences (RSS) that allow the unrearranged light chain V and J gene segments to recombine such that the modified non-human animal comprises a rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence operably linked to a heavy chain constant region nucleic acid sequence, wherein the heavy chain constant region nucleic acid sequence comprises an inactivating mutation or deletion in a sequence encoding a C_H1 domain. In one embodiment, the

unrearranged light chain V and J gene segments recombine such that the modified non-human animal comprises a rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions and is operably linked to a heavy chain constant region nucleic acid sequence, and wherein the heavy chain constant region nucleic acid sequence comprises an inactivating mutation or deletion in a sequence encoding a C_H1 domain. In one embodiment, the unrearranged light chain V and J gene segments are human V and J segments. The method may also comprise causing the animal to express a V_L -single domain binding protein derived from the unrearranged light chain V gene segment, the unrearranged light chain J gene segment and the heavy chain constant region having an inactivated or deleted C_H1 domain.

[0025] In another embodiment, the methods of producing a non-human animal as disclosed herein may further comprise introducing a genetically engineered single rearranged light chain locus comprising a nucleic acid encoding a single rearranged light chain, e.g., a universal light chain (ULC), and optionally, causing the animal to express the heavy chain immunoglobulin locus having an inactivated C_H1 domain and the single rearranged light chain locus.

[0026] In one aspect, the methods of producing a non-human animal as disclosed herein comprises (a) modifying the non-human heavy chain constant region of the non-human animal such that the heavy chain constant region comprises a deletion or an inactivating mutation of a nucleotide sequence encoding a C_H1 domain, and either or both (b) replacing at an endogenous immunoglobulin heavy chain locus, one or more, all, or substantially all endogenous non-human heavy chain variable region gene segments with one or more unrearranged light chain V and/or one or more unrearranged light chain J gene segments such that the light chain V and J gene segments are operably linked to the non-human heavy chain constant region comprising a deletion or an inactivating mutation of a nucleotide sequence encoding a C_H1 domain, and/or (c) introducing a nucleic acid encoding a genetically engineered single rearranged light chain locus. The steps of the methods disclosed herein may be performed in any order, sequentially or simultaneously.

[0027] For example, a method as disclosed herein may comprise (a) modifying the non-human heavy chain constant region of the non-human animal such that the heavy chain constant region comprises a deletion or an inactivating mutation of a nucleotide sequence encoding a C_H1 domain, and (b) replacing at an endogenous immunoglobulin heavy chain locus, one or more, all, or substantially all endogenous non-human heavy chain variable region gene segments with one or more unrearranged light chain V and/or one or more unrearranged light chain J gene

segments such that the light chain V and J gene segments are operably linked to the non-human heavy chain constant region comprising a deletion or an inactivating mutation of a nucleotide sequence encoding a C_H1 domain. In one aspect, a method as disclosed herein comprises (a) modifying the non-human heavy chain constant region of the non-human animal such that the heavy chain constant region comprises a deletion or an inactivating mutation of a nucleotide sequence encoding a C_H1 domain, and (b) introducing a nucleic acid encoding a genetically engineered single rearranged light chain locus. In another aspect, the method of producing a non-human animal as disclosed herein comprises (a) modifying the non-human heavy chain constant region of the non-human animal such that the heavy chain constant region comprises a deletion or an inactivating mutation of a nucleotide sequence encoding a C_H1 domain, (b) replacing at an endogenous immunoglobulin heavy chain locus, one or more, all, or substantially all endogenous non-human heavy chain variable region gene segments with one or more unrearranged light chain V and/or one or more unrearranged light chain J gene segments such that the light chain V and J gene segments are operably linked to the non-human heavy chain constant region comprising a deletion or an inactivating mutation of a nucleotide sequence encoding a C_H1 domain, and (c) introducing a nucleic acid encoding a genetically engineered single rearranged light chain locus. The methods of making a non-human animal as disclosed herein may further comprise introducing an Adam6a gene, an Adam6b gene, or both into the genome of the non-human animal, e.g., into the germline of the animal. In one aspect, the method further comprises causing the animal to express the heavy chain immunoglobulin locus having an inactivated C_H1 domain and/or the genetically rearranged light chain locus (single rearranged light chain locus), e.g., by immunizing the animal.

[0028] In one aspect, the step of inactivating the C_H1 domain(s), and optionally the hinge region(s), of a heavy chain immunoglobulin locus enables the non-human animal to produce a single domain antigen binding protein as disclosed herein. In one embodiment, inactivating the C_H1 domain(s), and optionally the hinge region(s) of a heavy chain immunoglobulin locus comprises targeting an endogenous heavy chain immunoglobulin locus with a targeting vector that deletes or introduces an inactivating mutation into the C_H1 domain(s), and optionally the hinge region, of a heavy chain locus, e.g., of an IgG heavy chain locus. In one embodiment, inactivating the C_H1 domain(s), and optionally, the hinge region(s) of the heavy chain immunoglobulin locus comprises homologous recombination. In another embodiment, the inactivating step may further comprise replacing of one or more, substantially all, or all of the endogenous variable region gene segments of the heavy

chain gene locus with any one or more of the following: one or more heavy chain variable region gene segments, one or more light chain variable region gene segments, one or more human variable region gene segments, and one or more unrearranged variable region gene segments. The step of inactivating the C_H1 domain(s), and optionally hinge region(s) of the heavy chain gene locus may occur in the germline of the non-human animal.

[0029] In one aspect the method comprises introducing a nucleic acid encoding a genetically engineered rearranged light chain locus, e.g., a genetically engineered universal light chain as described herein. In one aspect, the step of introducing the nucleic acid encoding a genetically engineered rearranged light chain locus further comprises replacing all or substantially all endogenous immunoglobulin light chain loci of the non-human animal. In one aspect, the step of introducing the nucleic acid encoding a genetically engineered rearranged light chain locus further comprises functionally inactivating all or substantially all endogenous immunoglobulin light chain loci of the non-human animal. In another aspect the nucleic acid encoding a genetically engineered rearranged light chain is introduced into the germline of the animal.

[0030] In one embodiment, the method of making a non-human animal as disclosed herein comprises (a) obtaining a first non-human animal comprising a heavy chain locus having a deleted or inactivated C_H1 domain (and optionally a light chain variable region nucleotide sequence operably linked to heavy chain constant region sequence having the deleted or inactivated C_H1 domain), and optionally a deleted or inactivated hinge region such that the non-human animal produces single domain antigen binding proteins (such as a V_L single domain binding protein) as disclosed herein, and (b) breeding the first non-human animal of (a) with a second non-human animal, which in one aspect may be a different strain as the first non-human animal, wherein the second non-human animal expresses a universal light chain, and wherein the breeding results in offspring that produce, e.g., comprise, a single domain antigen binding protein and a genetically engineered rearranged light chain (single rearranged light chain; ULC).

[0031] In some aspects, the methods of producing a non-human animal as disclosed herein further comprises inactivating or deleting one or more immunoglobulin genes selected from the group consisting of IgD, IgG3, IgG2a, IgG2b, IgG2c, IgE, and IgA. In one embodiment, the IgG2b and IgG2a/IgG2c immunoglobulin genes are deleted. In another embodiment, the IgD, IgG3, IgG2b, IgG2a/IgG2c, IgE, and IgA genes are deleted such that the non-human animal produces immunoglobulin heavy chains having an IgM or IgG1 isotype, wherein the

IgG1 isotype has a deletion or inactivating mutation in C_H1 domain and optionally a hinge region.

[0032] In one aspect, a non-human animal is already capable of producing single domain antigen binding proteins and provided herein is a method of increasing the production of single domain antigen binding proteins by the non-human animal. Such method comprises causing B cells of the non-human animal to express a nucleic acid encoding a genetically engineered single rearranged light chain, e.g., a genetically engineered universal light chain as described herein. Causing the B cells to express such a nucleic acid may also comprise the step of inactivating or preventing expression of endogenous light chain genes by the B cells.

[0033] In one aspect a non-human animal as disclosed herein is a rat or a mouse. In another embodiment, a non-human animal as disclosed herein is a mouse. Accordingly, provided herein is a genetically modified mouse, comprising (a) a replacement at a mouse heavy chain locus of all or substantially all endogenous immunoglobulin heavy chain V, D, and J gene segments with one or more human heavy chain V, D, and J gene segments, wherein the one or more human heavy chain V, D, and J gene segments are operably linked to a mouse heavy chain constant region (e.g., endogenous mouse heavy chain constant region), wherein the mouse heavy chain constant region comprises a full-length IgM gene; and an IgG gene comprising a deletion or an inactivating mutation in a nucleotide sequence encoding a C_H1 region in an IgG gene selected from the group consisting of an IgG1, IgG2a, IgG2c, IgG2b, and a combination thereof, wherein the mouse expresses a B cell receptor that comprises an IgM with a C_H1 region, wherein the IgM comprises a heavy chain associated with a cognate light chain; and (b) a replacement of all or substantially all endogenous immunoglobulin light chain V and J gene segments with a single rearranged variable V_κ:J_κ gene sequence. In some embodiments, the cognate light chain is derived from the single rearranged variable V_κ:J_κ gene sequence. In some embodiments, the single rearranged variable V_κ:J_κ gene sequence is operably linked to a mouse light chain constant sequence, e.g., endogenous mouse light chain constant sequence.

[0034] In another aspect, provided herein is a non-human animal, e.g., a rat or a mouse, comprising (a) a deletion or functional inactivation at a mouse heavy chain locus of all or substantially all endogenous immunoglobulin heavy chain V, D, and J gene segments and introduction of one or more human heavy chain V, D, and J gene segments, wherein the one or more human heavy chain V, D, and J gene segments are operably linked to a mouse heavy chain constant region (e.g., endogenous

mouse heavy chain constant region), wherein the mouse heavy chain constant region comprises a full-length IgM gene; and an IgG gene comprising a deletion or an inactivating mutation in a nucleotide sequence encoding a C_H1 region in an IgG gene selected from the group consisting of an IgG1, IgG2a, IgG2c, IgG2b, and a combination thereof, wherein the mouse expresses a B cell receptor that comprises an IgM with a C_H1 region, wherein the IgM comprises a heavy chain associated with a cognate light chain; and/or (b) a deletion or functional inactivation of all or substantially all endogenous immunoglobulin light chain V and J gene segments and introduction of a single rearranged variable V_κ:J_κ gene sequence.

[0035] Also provided herein is a genetically modified mouse, comprising (a) a replacement at a mouse heavy chain locus of all or substantially all endogenous immunoglobulin heavy chain V, D, and J gene segments with one or more human light chain V and J gene segments, wherein the one or more human light chain V and J gene segments are operably linked to a mouse heavy chain constant region, wherein the mouse heavy chain constant region comprises a full-length IgM gene; and an IgG gene comprising a deletion or an inactivating mutation in a nucleotide sequence encoding a C_H1 region in an IgG gene selected from the group consisting of an IgG1, IgG2a, IgG2c, IgG2b, IgG3, and a combination thereof, wherein the mouse expresses a B cell receptor that comprises an IgM with a C_H1 region, wherein the IgM comprises a heavy chain associated with a cognate light chain; and (b) a replacement of all or substantially all endogenous immunoglobulin light chain V and J gene segments with a single variable V_κ:J_κ gene sequence. In some embodiments, the cognate light chain is derived from the single rearranged variable V_κ:J_κ gene sequence.

[0036] In one aspect, a method for making an binding protein that lacks a C_H1 domain is provided, comprising: (a) isolating from a non-human animal as described herein the binding protein, a cell that makes the binding protein, or a nucleotide sequence that encodes a sequence of the binding protein. In one aspect, the isolating step may comprise one or more of the following steps: (a) immunizing a non-human animal as described herein with an antigen; (b) maintaining the non-human animal under conditions sufficient for the non-human animal to make a binding protein and/or (c) identifying an binding protein made by the non-human animal that lacks a functional C_H1 domain and/or that lacks a functional hinge region. In some aspects the binding protein so isolated is a single domain antigen binding protein. In one aspects a single domain antigen binding protein is monomeric.

[0037] In one aspect, a method for making an antigen-binding protein is provided, comprising (a) immunizing a non-human animal as described herein with an antigen; (b) maintaining the non-human animal under conditions sufficient to make an binding protein; (c) identifying an binding protein made by the non-human animal, wherein the binding protein lacks a functional C_H1 domain or lacks a functional C_H1 domain and lacks a hinge region; (d) identifying a variable region sequence encoding a variable domain on an immunoglobulin polypeptide that lacks a C_H1 domain, or lacks a hinge region and a C_H1 domain, wherein the variable domain specifically binds the antigen; (e) expressing a protein encoded by a sequence identical to or substantially identical to the variable region sequence of (d) in a suitable expression system wherein the variable region sequence of (d) is linked with a nucleic acid sequence of a heavy chain variable sequence that lacks a C_H1 region or lacks a C_H1 region and a hinge; and/or (f) isolating the expressed protein of (e). In some embodiments, the steps of expressing a protein encoded by the variable region sequence and/or (f) isolating the expressed protein comprises culturing a cell, e.g., a cell transfected with the variable region sequence, a hybridoma formed from a cell isolated from an animal disclosed herein and/or collecting supernatant from a cultured cell.

[0038] In one aspect, a method for making an antigen-binding protein is provided, comprising immunizing a non-human animal as described herein with an antigen, identifying a variable region nucleic acid sequence encoding a variable domain that specifically binds the antigen, and employing the variable region nucleic acid sequence in a suitable expression system, wherein the variable region nucleic acid sequence is linked with a heavy chain constant gene that lacks a C_H1 or lacks a C_H1 and a hinge; wherein the expression system expresses an antigen-binding protein that specifically binds the antigen.

[0039] Accordingly, also provided herein are such isolated binding proteins, cells, and nucleic acid sequences.

[0040] Other embodiments are described and will become apparent to those skilled in the art from a review of the ensuing detailed description.

BRIEF DESCRIPTION OF THE FIGURES

[0041] **FIG. 1A** illustrates targeting a mouse IgG1 gene, IgG2b and IgG2a genes (not to scale) to make a genetically modified mouse immunoglobulin heavy chain locus that expresses an IgG1 lacking a C_H1 domain; human immunoglobulin heavy chain V, D and J segments, represented by empty triangles, are inserted to a mouse

constant locus wherein the IgG1 C_H1 exon* and IgG2a/ 2b** are deleted, ovals represent enhancers.

[0042] FIG. 1B illustrates a mouse immunoglobulin light chain locus (not to scale) comprising a single rearranged human V_L/J_L gene sequence***.

[0043] FIG. 1C illustrates an IgM expressed by a mouse having the Ig loci of FIGs. 1B and 1C, wherein the IgM comprises an intact C_H1 domain. FIG. 1C also illustrates that upon class switching, the IgG1 expressed by the mouse having the Ig loci of FIGs. 1A and 1C is a single domain heavy chain antigen binding protein lacking a C_H1 domain.

[0044] FIG. 2 illustrates the genomic structure (not to scale) of two universal light chains, one of which comprises a single rearranged human variable region comprising V_κ1-39J_κ5 and the other of which comprises a single rearranged human variable region comprising V_κ3-20J_κ1.

[0045] FIG. 3 illustrates a wild-type IgG1 locus in a mouse (IgG1, top), showing the J_H region gene segment fused to a C_H1 gene segment, followed by a hinge region, a C_H2 gene segment, and a C_H3 gene segment; an IgG1 locus targeted with a construct that deletes a C_H1 domain (IgG1ΔC_H1; I); an IgG1 locus targeted with a construct that deletes both a C_H1 domain and a hinge region (IgG1ΔC_H1-Δhinge; II); a constant region locus targeted with a construct that deletes an IgG1 C_H1 domain, IgG2b and IgG2a (IgG1ΔC_H1ΔIgG2b/2a; III); a constant region locus targeted with a construct that deletes an IgG1 C_H1 domain, a hinge region, IgG2b and IgG2a (IgG1ΔC_H1&Δhinge ΔIgG2b/2a; IV); or constant region locus targeted with a construct that deletes an IgG1 C_H1 domain, IgG2b, IgG2a, IgG3, IgD, IgA, and IgE, and optionally a hinge region (IgG1ΔC_H1ΔIgG2b/2a ΔIgG3 ΔIgD/A/E (optionally Δhinge); V). The schematic illustrations of the loci are not presented to scale. IgG2a/c designates either an IgG2a or IgG2c locus, as a mouse may have either an IgG2a allele or an IgG2c allele depending on its strain.

[0046] FIG. 4A illustrates targeting a mouse heavy chain sequence (not to scale) to make a genetically modified locus that contains human heavy chain variable gene segments (empty triangles) and lacks a functional IgG1 C_H1 domain as well as lacks IgG2a and IgG2b loci (in some embodiments referred to as 1673).

[0047] FIG. 4B illustrates targeting a mouse IgG1 gene (all variable gene segments are mouse and indicated by the filled triangles) to make a genetically modified locus (not to scale) that expresses an IgG1 lacking a C_H1 domain and a hinge (in some embodiments referred to as 1576).

[0048] FIG. 4C illustrates targeting a mouse heavy chain constant region (not to scale) to make a genetically modified locus lacking the IgM, IgD, IgG3, IgG1, IgG2b, IgG2a, IgE and IgA gene segments (first part of cloning IgG1 Δ C_H1 Δ IgG2b/IgG2a, Δ IgG3, Δ IgD/A/E continued in FIG. 4D). Human variable gene segments are indicated by empty triangles.

[0049] FIG. 4D illustrates targeting the mouse constant region of FIG. 4C to make a genetically modified locus (not to scale) that comprises human heavy chain variable gene segments, a complete and functional murine IgM gene region, and an IgG1 gene region lacking a functional C_H1 domain, and optionally, lacking a functional hinge region (the mouse also lacks IgG2b/IgG2a, IgG3, and IgD/A/E; in some embodiments referred to as 6180). Human variable gene segments are indicated by empty triangles.

[0050] FIG. 5A shows comparative total serum IgG1 titers between various single domain IgG1 mice, before and after immunization with β -galactosidase (β gal): mice homozygous for mV_HIgG1 Δ C_H1&hinge with mouse kappa chain (1576); mice homozygous for mV_HIgG1 Δ C_H1&hinge with a kappa chain that is a single rearranged light chain V κ 3-20J κ 1 ULC (1576/1635); or wild-type (WT) mouse with same genetic background. HO is homozygous for genetic modification.

[0051] FIG. 5B illustrates antigen specific IgG1 serum titers in immunized WT mice compared to mV_HIgG1 Δ C_H1&hinge homozygous mice (1576HO) or mV_HIgG1 Δ C_H1-hinge x V κ 3-20J κ 1 ULC1 homozygous mice (1576HO 1635HO). Mice were immunized with β -galactosidase (β gal) as a model antigen, antigen specific IgG1 titers were measured by ELISA.

[0052] FIG. 6 provides an image of a Western blot, prepared under non-reducing conditions and visualized with anti-mouse IgG, of mouse sera from two hV_HIgG1 Δ C_H1&hinge Δ IgG2b/2a x V κ 1-39J κ 5 ULC homozygous mice (1859HO 1633HO), three hV_HIgG1 Δ C_H1 Δ IgG2b/2a x V κ 3-20J κ 1 ULC homozygous mice (1673HO 1635 HO) and three VELOCIMMUNE® control mice (VI3 IgG1) having human variable regions (V_H and V κ) with mouse constant regions of dimeric single domain antigen binding proteins (37-37 homodimer) or monomeric Δ C_H1 single domain binding proteins (C_H1 hinge delete single chain).

[0053] FIG. 7 shows IgG1 titers (y-axis) found in the plasma of animals at different timepoints (x-axis) after intraperitoneal immunization with Antigen X, a cell surface protein, of hV_HIgG1 Δ C_H1 Δ IgG2b/2a x V κ 1-39J κ 5 homozygous mice (1673X1633) or hV_HIgG1 Δ C_H1 Δ IgG2b/2a x V κ 3-20J κ 1 homozygous mice (1673X1635).

[0054] FIG. 8 provides an image of a Western blot, prepared under non-reducing conditions and visualized with anti-mouse IgG, of mouse sera from three hV_HIgG1ΔC_H1ΔIgG2b/2a ΔIgG3 ΔIgD/A/E x V_κ1-39Jκ5 ULC homozygous mice (6180HO 1634 HO), two hV_HIgG1ΔC_H1ΔIgG2b/2a ΔIgG3 ΔIgD/A/E homozygous mice (6180 HO) and three VI3 control mice having human variable regions (V_H and V_κ) with mouse constant regions of dimeric single domain antigen binding proteins (37-37 homodimer) or monomeric ΔC_H1 single domain binding proteins (C_H1 delete single chain).

[0055] FIG. 9 shows the concentration of steady state IgM and IgG in the plasma of hV_HIgG1ΔC_H1ΔIgG2a/2bΔIgG3ΔIgD/A/E x V_κ1-39Jκ5 (6180 HO x1634 HO) mice and hV_HIgG1ΔC_H1ΔIgG2a/2bΔIgG3ΔIgD/A/E (6180 HO) and VI3 control animals.

[0056] FIG. 10 shows contour plots of splenocytes gated on singlets stained for CD19 and CD3 from a representative homozygous control VI3 and hV_HIgG1ΔC_H1ΔIgG2a/2bΔIgG3ΔIgD/A/E (6180 HO) x V_κ1-39Jκ5 (6180 HO 1634 HO) mouse (A). Also shown are contour plots of splenocytes gated on CD19+ B cells stained for immunoglobulin D (IgD) and immunoglobulin M (IgM) from a representative control VI3 mouse and a representative homozygous hV_HIgG1ΔC_H1ΔIgG2a/2bΔIgG3ΔIgD/A/E x V_κ1-39Jκ5 (6180 HO x1634 HO) mouse (C). Notably, the B cells are IgD⁻ as the IgD constant domain was deleted. Therefore, the term “mature” in this plot is merely an indication of the absence of IgD and not other non-IgM immunoglobulins. Also provided is a graph showing the total number of CD19⁺ B cells (y-axis; cells/spleen X10⁷) of three representative mice from each group (B). The animal for which the contour plots are included is encircled.

[0057] FIG. 11 shows contour plots of (A) CD19+ gated B cells isolated from the spleen stained for CD93 and B220, (B) immature or mature gated B cells stained for IgM and CD23, or (C) immature or mature gated B cells stained for CD21/35 and IgM from a representative control VI3 mouse and a representative homozygous hV_HIgG1ΔC_H1ΔIgG2a/2bΔIgG3ΔIgD/A/E x V_κ1-39Jκ5 (6180 HO x1634 HO) mouse.

[0058] FIG. 12 shows contour plots of bone marrow isolated from the femurs of a representative control VI3 mouse and a representative homozygous hV_HIgG1ΔC_H1ΔIgG2a/2bΔIgG3ΔIgD/A/E x V_κ1-39Jκ5 (6180 HO x1634 HO) mouse, stained with CD19 and CD3. Also shown are the total number of cells or CD19⁺ B cells per femur (y-axis; Cells/femur x 10⁷) of three representative mice from each group. The animal for which the contour plots are included is encircled.

[0059] FIG. 13 shows contour plots of bone marrow isolated from femurs of a representative control VI3 mouse and a representative homozygous

hV_HIgG1ΔC_H1ΔIgG2a/2bΔIgG3ΔIgD/A/E x V_κ1-39J_κ5 (6180 HO x1634 HO) mouse stained with IgM and B220. Also shown is the total number of mature (IgM⁺ B220^{hi}) and immature (IgM⁺ B220^{int}) B cells per femur (y-axis; Cells/femur x 10⁷) of three representative mice from each group. The animal for which the contour plots are included is encircled.

[0060] FIG. 14A illustrates a schematic of the mouse heavy chain locus (not to scale). The mouse heavy chain locus is about 3 Mb in length and contains approximately 200 heavy chain variable (V_H) gene segments, 13 heavy chain diversity (D_H) gene segments and 4 heavy chain joining (J_H) gene segments as well as enhancers (Enh) and heavy chain constant (CH) regions.

[0061] FIG. 14B illustrates a schematic of the human κ light chain locus (not to scale). The human κ light chain locus is duplicated into distal and proximal contigs of opposite polarity spanning about 440 kb and 600 kb, respectively. Between the two contigs is about 800 kb of DNA that is believed to be free of V_κ gene segments. The human κ light chain locus contains about 76 V_κ gene segments, 5 J_κ gene segments, an intronic enhancer (Enh) and a single constant region (C_κ).

[0062] FIG. 15 shows a targeting strategy (not to scale) for progressive insertion of 40 human V_κ and 5 human J_κ gene segments into a mouse heavy chain locus in which endogenous heavy chain variable region gene segments have been deleted. Hygromycin (HYG) and Neomycin (NEO) selection cassettes are shown with recombinase recognition sites (FRT). Also shown is a targeting strategy (not to scale) for the insertion of the Adam6a, Adam6b and IGCR1 genes. Human variable gene segments are indicated by empty triangles.

[0063] FIG. 16 shows the modified mouse heavy chain locus of FIG. 15 (hJ_κ, 40hV_κ, Adam6; top); a targeting strategy that results in the deletion and/or inactivation of the C_H1 domain of the IgG1 gene sequence, the IgG2b gene sequence, and the IgG2a gene sequence from the modified mouse heavy chain locus of FIG. 15 (hJ_κ, 40hV_κ, Adam6, ΔC_H1, ΔIgG2b, ΔIgG2a; middle), and a targeting strategy that results in the deletion of the selection cassette from a modified mouse heavy chain locus comprising unrearranged human J_κ gene segments and 40 human V_κ gene segments, wherein the modified mouse heavy chain locus lacks a functional C_H1 domain in the IgG1 gene and also lacks functional IgG2b and IgG2a genes (hJ_κ, 40hV_κ, Adam6, ΔC_H1, ΔIgG2b, ΔIgG2a selection cassette deleted; also termed 6082, bottom). Human variable gene segments are indicated by empty triangles.

[0064] FIG. 17A shows the relative mRNA expression (normalized to HPRT1 mRNA; y-axis) by B cells isolated from the spleen and bone marrow of three different groups of animals (x-axis): wild type (WT) mice, mice homozygous for the modified heavy chain locus of FIG. 16 (hJ κ , 40hV κ , Adam6, Δ C $_H$ 1 Δ IgG2b Δ IgG2a selection cassette deleted; KoH C $_H$ 1 del), and mice homozygous for both a modified mouse heavy chain locus that expresses human heavy chain V, D and J segments, lacks a functional C $_H$ 1 domain in the IgG1 genes, and also lacks functional IgG2b and IgG2a genes and comprises a single rearranged light chain locus (C $_H$ 1 del x ULC). Probes were designed to detect productive rearrangement, and consisted of a probe that detected recombination between a human J κ segment and murine IgG1 hinge (hJ κ /mIgG1 hinge probe; left panels) or a probe that detected recombination between a human J $_H$ segment and murine IgG1 hinge (hJ $_H$ /mIgG1 hinge probe; right panels). ND: not detected (Ct $\geq$ 35). n= 2 for WT, 2 for KoH C $_H$ 1 del, and 3 for C $_H$ 1 del x ULC.

[0065] FIG. 17B shows the relative mRNA expression (normalized to mKappaC mRNA; y-axis) by B cells isolated from the spleen and bone marrow of three different groups of animals (x-axis): wild type (WT) mice, mice homozygous for the modified heavy chain locus of FIG. 16 (hJ κ , 40hV κ , Adam6, Δ C $_H$ 1 Δ IgG2b Δ IgG2a selection cassette deleted; KoH C $_H$ 1 del), and mice homozygous for both a modified mouse heavy chain locus that expresses human heavy chain V, D and J segments, lacks a functional C $_H$ 1 domain in the IgG1 gene sequence, and also lacks functional IgG2b and IgG2a gene sequences and comprises a single rearranged light chain locus (C $_H$ 1 del x ULC). Probes were designed to detect productive rearrangement, and consisted of a probe that detected recombination between a human J κ segment and murine IgG1 hinge (hJ κ /mIgG1 Hinge probe; left panels) or a probe that detected human J $_H$ segment and murine IgG1 hinge (hJ $_H$ /mIgG1 hinge probe; right panels). ND: not detected (Ct $\geq$ 35). n= 2 for WT, 2 for KoH C $_H$ 1 del, and 3 for C $_H$ 1 del x ULC.

[0066] FIG. 18 provides an image of a Western blot, prepared under non-reducing conditions and visualized with anti-mouse IgG, of mouse sera from three hV κ IgG1 Δ C $_H$ 1 Δ IgG2a/2b x V κ 3-20J κ 1 ULC homozygous mice (6082HO 1635 HO), three VELOCIMMUNE® mice (VI3 IgG1) having human variable region (V $_H$ and V κ) with mouse constant regions (WT), and two hV $_H$ IgG1 Δ C $_H$ 1 Δ IgG2a/2b Δ IgG3 Δ IgD/A/E mice (6180 HO) demonstrating presence or absence of dimeric single domain antigen binding proteins (37-37 homodimer) or monomeric Δ C $_H$ 1 single domain binding proteins (C $_H$ 1 delete single chain).

DETAILED DESCRIPTION

[0067] The invention is not limited to particular methods, and experimental conditions described, as such methods and conditions may vary. The terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only by the claims.

[0068] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, particular methods and materials are now described. All publications and patent documents mentioned herein are incorporated herein by reference in their entirety.

[0069] The present invention provides genetically modified non-human animals (e.g., mice, rats, rabbits, hamsters, etc.) that comprise in their genome, e.g., in their germline, nucleotide sequence(s) encoding single domain antigen binding proteins, including V_H -single domain antigen binding proteins and V_L -single domain antigen binding proteins, and/or a single rearranged light chain; methods of making the same; as well as methods of using the same. Unless defined otherwise, all terms and phrases used herein include the meanings that the terms and phrases have attained in the art, unless the contrary is clearly indicated or clearly apparent from the context in which the term or phrase is used.

[0070] The term "antibody" includes typical immunoglobulin molecules comprising four polypeptide chains, two heavy (H) chains and two light (L) chains inter-connected by disulfide bonds. The term also includes an immunoglobulin that is reactive to an antigen or fragment thereof. Suitable antibodies include, but are not limited to, human antibodies, primatized antibodies, chimeric antibodies, monoclonal antibodies, monospecific antibodies, polyclonal antibodies, polyspecific antibodies, nonspecific antibodies, bispecific antibodies, multispecific antibodies, humanized antibodies, synthetic antibodies, recombinant antibodies, hybrid antibodies, mutated antibodies, grafted conjugated antibodies (i.e., antibodies conjugated or fused to other proteins, radiolabels, cytotoxins), and in vitro-generated antibodies. A skilled artisan will readily recognize common antibody isotypes, e.g., antibodies having a heavy chain constant region selected from the group consisting of IgG, IgA, IgM, IgD, and IgE, and any subclass thereof (e.g., IgG1, IgG2, IgG3, and IgG4).

[0071] The phrase "heavy chain," or "immunoglobulin heavy chain" includes an immunoglobulin heavy chain sequence, including immunoglobulin heavy chain constant region sequence, from any organism. Heavy chain variable domains include three heavy chain CDRs and four FR regions, unless otherwise specified. Fragments of heavy chains include CDRs, CDRs and FRs, and combinations thereof. A typical heavy chain has, following the variable domain (from N-terminal to C-terminal), a C_H1 domain, a hinge, a C_H2 domain, a C_H3 domain, and a C_H4 domain (in the context of IgM or IgE). A functional fragment of a heavy chain includes a fragment that is capable of specifically recognizing an epitope (e.g., recognizing the epitope with a KD in the micromolar, nanomolar, or picomolar range), that is capable of expressing and secreting from a cell, and that comprises at least one CDR. A heavy chain variable domain is encoded by a variable region gene sequence, which generally comprises V_H, D_H, and J_H segments derived from a repertoire of V_H, D_H, and J_H segments present in the germline. Sequences, locations and nomenclature for V, D, and J heavy chain segments for various organisms can be viewed at the website of the International Immunogenetics Information System (IMGT) found at www.imgt.org.

[0072] The phrase "light chain" includes an immunoglobulin light chain sequence from any organism, and unless otherwise specified includes human kappa and lambda light chains and a VpreB, as well as surrogate light chains. Light chain variable domains typically include three light chain CDRs and four framework (FR) regions, unless otherwise specified. Generally, a full-length light chain includes, from amino terminus to carboxyl terminus, a variable domain that includes FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4, and a light chain constant region. A light chain variable domain is encoded by a light chain variable region gene sequence, which generally comprises V_L and J_L gene segments, derived from a repertoire of V_L and J_L gene segments present in the germline. Sequences, locations and nomenclature for V and J light chain segments for various organisms can be viewed at the website of the International Immunogenetics Information System (IMGT) found at www.imgt.org. Light chains include those, e.g., that do not selectively bind either a first or a second epitope selectively bound by the epitope-binding protein in which they appear. Light chains also include those that bind and recognize, or assist the heavy chain with binding and recognizing, one or more epitopes selectively bound by the epitope-binding protein in which they appear. The phrase light chain includes a "common light chain," also referred to as a "universal light chain" (ULC).

[0073] Common or universal light chains (ULCs) include those derived from an immunoglobulin light chain locus comprising a single rearranged immunoglobulin light chain variable region encoding sequence operably linked with a light chain constant region, wherein expression of the immunoglobulin light chain locus produces only a light chain derived from the single rearranged immunoglobulin light chain variable region operably linked to the light chain constant region regardless of the inclusion of other nucleic acid sequences, e.g., other light chain gene segments, in the immunoglobulin light chain locus. Universal light chains include human V κ 1-39J κ gene (e.g., V κ 1-39J κ 5 gene) or a human V κ 3-20J κ gene (e.g., V κ 3-20J κ 1 gene), and include somatically mutated (e.g., affinity matured) versions of the same.

[0074] The phrase “gene segment,” or “segment” includes reference to a V (light or heavy) or D or J (light or heavy) immunoglobulin gene segment, which includes unrearranged sequences at immunoglobulin loci (in e.g., humans and mice) that can participate in a rearrangement (mediated by, e.g., endogenous recombinases) to form a rearranged V/J (light) or V/D/J (heavy) sequence. Unless indicated otherwise, the V, D, and J segments comprise recombination signal sequences (RSS) that allow for V/J recombination or V/D/J recombination according to the 12/23 rule. Unless indicated otherwise, the segments further comprise sequences with which they are associated in nature or functional equivalents thereof (e.g., for V segments, promoter(s) and leader(s)).

[0075] The term “unrearranged,” with reference to a nucleic acid sequence, includes nucleic acid sequences that exist in the germline of an animal cell, preferably a cell derived from an animal that has not been genetically modified, e.g., comprises a wild-type genome. Generally, in native germline configuration, the heavy chain variable region comprises unrearranged V_H gene segments, unrearranged D_H gene segments and unrearranged J_H gene segments while the light chain variable region comprises unrearranged V_L gene segments and unrearranged J_L gene segments. During the B cell maturation process, these gene segments rearrange to produce a rearranged variable region gene.

[0076] The term “germline” in reference to an immunoglobulin nucleic acid sequence includes a nucleic acid sequence that can be passed to progeny.

[0077] The phrase “complementarity determining region,” or the term “CDR,” includes an amino acid sequence encoded by a nucleic acid sequence of an organism’s immunoglobulin genes that normally (i.e., in a wild-type animal) appears between two framework regions in a variable region of a light or a heavy chain of an

immunoglobulin molecule (e.g., an antibody or a T cell receptor). A CDR can be encoded by, for example, a germline sequence or a rearranged or unrearranged sequence, and, for example, by a naive or a mature B cell or a T cell. A CDR can be somatically mutated (e.g., vary from a sequence encoded in an animal's germline), humanized, and/or modified with amino acid substitutions, additions, or deletions. In some circumstances (e.g., for a CDR3), CDRs can be encoded by two or more sequences (e.g., germline sequences) that are not contiguous (e.g., in an unrearranged nucleic acid sequence) but are contiguous in a B cell nucleic acid sequence, e.g., as the result of splicing or connecting the sequences (e.g., V-D-J recombination to form a heavy chain CDR3).

[0078] The phrase “somatically mutated” includes reference to a nucleic acid sequence from a B cell that has undergone class-switching, wherein the nucleic acid sequence of an immunoglobulin variable region (e.g., nucleotide sequence encoding a heavy chain variable domain or including a heavy chain CDR or FR sequence) in the class-switched B cell is not identical to the nucleic acid sequence in the B cell prior to class-switching, such as, for example, a difference in a CDR or framework nucleic acid sequence between a B cell that has not undergone class-switching and a B cell that has undergone class-switching. “Somatically mutated” includes reference to nucleic acid sequences from affinity-matured B cells that are not identical to corresponding immunoglobulin variable region sequences in B cells that are not affinity-matured (i.e., sequences in the genome of germline cells). The phrase “somatically mutated” also includes reference to an immunoglobulin variable region nucleic acid sequence from a B cell after exposure of the B cell to an epitope of interest, wherein the nucleic acid sequence differs from the corresponding nucleic acid sequence prior to exposure of the B cell to the epitope of interest. The phrase “somatically mutated” refers to sequences from binding proteins that have been generated in an animal, e.g., a mouse having human immunoglobulin variable region nucleic acid sequences, in response to an immunogen challenge, and that result from the selection processes inherently operative in such an animal.

[0079] The term “cognate,” when used in the sense of “cognate with,” e.g., a first V_L domain that is “cognate with” a second V_L domain, is intended to include reference to the relation between two V_L domains from a same binding protein made by a mouse in accordance with the invention. For example, a mouse that is genetically modified in accordance with an embodiment of the invention, e.g., a mouse having a heavy chain locus in which V_H , D_H , and J_H regions are replaced with V_L and J_L regions, makes antibody-like binding proteins that have two identical polypeptide

chains made of the same mouse C_H region (e.g., an IgM isotype fused with a first human V_L domain, and two identical polypeptide chains made of the same mouse C_L region fused with a second human V_L domain. During clonal selection in the mouse, the first and the second human V_L domains were selected by the clonal selection process to appear together in the context of a single antibody-like binding protein. Thus, first and second V_L domains that appear together, as the result of the clonal selection process, in a single antibody-like molecule are referred to as being "cognate." In contrast, a V_L domain that appears in a first antibody-like molecule and a V_L domain that appears in a second antibody-like molecule are *not* cognate, *unless* the first and the second antibody-like molecules have identical heavy chains (i.e., unless the V_L domain fused to the first human heavy chain region and the V_L domain fused to the second human heavy chain region are identical).

[0080] Early in antibody development, antibody heavy chains undergo a selection process wherein nature chooses, through a variety of selection schemes, suitable heavy chains to undergo further selection to eventually form functional and affinity-matured antibodies. Diversity arising from heavy and light chain variable gene rearrangement occurs in the bone marrow and precedes class switching. Antibody heavy chains expressed from recombined heavy chain gene segments in progenitor B cells (or, pro-B cells) are normally paired with a surrogate light chain for presentation on the surface of the pro-B cell in an IgM isotype to form a structure (which includes other co-receptors) referred to as a pre-B cell receptor, or pre-BCR. Once the pre-BCR is presented on the cell surface, the pre-BCR is believed to signal its appropriate formation of the complex to the cell, effectively instructing the cell that the heavy chain has passed this early selection step. Thus the cell is informed that the heavy chain may undergo further selection. If the heavy chain contains a defect that is deleterious to the formation of a pre-BCR when presented in the context of an IgM and a surrogate light chain, the cell will undergo apoptosis. If the cell undergoes apoptosis, the usefulness, or contribution to diversity, of the heavy chain variable region of the heavy chain will be lost. Thus, a very early step in antibody selection requires presentation of the heavy chain together with a surrogate light chain in the context of an IgM isotype. Normal development of antibody-producing B cells generally requires the presence of a C_H1 domain. All heavy chain isotypes, including IgM, comprise a C_H1 domain. Both the surrogate light chain and a cognate light chain are believed to interact with a given heavy chain through the heavy chain's C_H1 domain in the context of an IgM.

[0081] After B-cells exit the bone marrow, engagement with antigen (which requires a low affinity interaction between the rearranged antibody expressed as a cell-surface IgM) stimulates concerted induction of somatic hypermutation and class switching. After class switching, differential antigen recognition by the surface B-cell receptor allows antibodies of increased affinity to be selected from a pool of hyper-mutated derivatives of the original IgM.

[0082] The term "heavy chain only antibody," "heavy chain only antigen binding protein," "single domain antigen binding protein," "single domain binding protein" or the like refers to a monomeric or homodimeric immunoglobulin molecule comprising an immunoglobulin-like chain comprising a variable domain operably linked to a heavy chain constant region, that is unable to associate with a light chain because the heavy chain constant region typically lacks a functional C_H1 domain.

Accordingly, the term "heavy chain only antibody," "heavy chain only antigen binding protein," "single domain antigen binding protein," "single domain binding protein" or the like encompasses a both (i) a monomeric single domain antigen binding protein comprising one of the immunoglobulin-like chain comprising a variable domain operably linked to a heavy chain constant region lacking a functional C_H1 domain, or (ii) a homodimeric single domain antigen binding protein comprising two immunoglobulin-like chains, each of which comprising a variable domain operably linked to a heavy chain constant region lacking a functional C_H1 domain. In various aspects, a homodimeric single domain antigen binding protein comprises two identical immunoglobulin-like chains, each of which comprising an identical variable domain operably linked to an identical heavy chain constant region lacking a functional C_H1 domain. Additionally, each immunoglobulin-like chain of a single domain antigen binding protein comprises a variable domain, which may be derived from heavy chain variable region gene segments (e.g., V_H, D_H, J_H), light chain gene segments (e.g., V_L, J_L), or a combination thereof, linked to a heavy chain constant region (C_H) gene sequence comprising a deletion or inactivating mutation in a C_H1 encoding sequence (and, optionally, a hinge region) of a heavy chain constant region gene, e.g., IgG, IgA, IgE, IgD, or a combination thereof. A single domain antigen binding protein comprising a variable domain derived from heavy chain gene segments may be referred to as a "V_H-single domain antibody" or "V_H-single domain antigen binding protein". A single domain antigen binding protein comprising a variable domain derived from light chain gene segments may be referred to as a "V_L-single domain antigen binding protein".

[0083] As disclosed above, the production of single domain antigen binding proteins by non-human animals engineered to do so results in the relatively low expression of antigen-specific single domain antigen binding proteins in response to antigen compared to traditional antibodies. The art suggests that a high titer is possible only when there is little to no expression of a rearranged light chain. Specifically, it has been asserted that animals that do not express a rearranged light chain are capable of producing higher levels of single domain antigen binding proteins that specifically bind antigen. Janssens et al. (2006) *PNAS* 103:15130-15135; Zou et al. (2004) *J. Exp. Med.* 204:3271-32. Contrary to the art, the data provided herein show that animals that express a genetically engineered single rearranged light chain will generate in high titers of antigen-specific single domain antigen binding proteins after challenge. Also shown herein is the ability of light chain variable region gene segments to rearrange and recombine with a heavy chain constant region gene comprising a deletion or inactivating mutation in a C_H1 sequence to encode for a V_L-single domain antigen binding protein capable of specifically binding antigen. It is possible that the light chain variable domain of such a V_L-single domain antigen binding protein compensates for the lack of a cognate light chain by the addition of, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions in a rearranged light chain variable region gene sequence, which is not typically seen in light chain variable region gene sequences rearranged from an endogenous and unmodified immunoglobulin light chain locus.

[0084] Accordingly, in one aspect, a non-human animal is provided, wherein the animal comprises a single domain antigen binding protein and a genetically engineered single rearranged light chain, e.g., a common light chain, wherein at least one heavy chain of the single domain antigen binding protein lacks a functional C_H1 domain. In another aspect, a non-human animal is provided, wherein the animal comprises V_L-single domain antigen binding protein that comprises a light chain variable region and a heavy chain constant region lacking a functional C_H1 domain. In another aspect, a non-human animal is provided, wherein the animal comprises a V_L-single domain antigen binding protein that comprises a light chain variable region and a heavy chain constant region lacking a functional C_H1 domain and a genetically engineered single rearranged light chain, e.g., a common light chain. Also provided are methods of making the genetically modified non-human animal, proteins (e.g., single domain antigen binding proteins) cells isolated from the genetically modified non-human animals, and methods of isolating proteins and cells from the genetically modified animals.

[0085] The term “high affinity” antibody refers to an antibody that has a K_D with respect to its target epitope about of 10^{-9} M or lower (e.g., about 1×10^{-9} M, 1×10^{-10} M, 1×10^{-11} M, or about 1×10^{-12} M). In one embodiment, K_D is measured by surface plasmon resonance, e.g., BIACORE™; in another embodiment, K_D is measured by ELISA.

[0086] The term “cell” includes any cell that is suitable for expressing a recombinant nucleic acid sequence. Cells include those of prokaryotes and eukaryotes (single-cell or multiple-cell), bacterial cells (e.g., strains of *E. coli*, *Bacillus* spp., *Streptomyces* spp., etc.), mycobacteria cells, fungal cells, yeast cells (e.g., *S. cerevisiae*, *S. pombe*, *P. pastoris*, *P. methanolica*, etc.), plant cells, insect cells (e.g., SF-9, SF-21, baculovirus-infected insect cells, *Trichoplusia ni*, etc.), non-human animal cells, human cells, or cell fusions such as, for example, hybridomas or quadromas. In some embodiments, the cell is a human, monkey, ape, hamster, rat, or mouse cell. In some embodiments, the cell is eukaryotic and is selected from the following cells: CHO (e.g., CHO K1, DXB-11 CHO, Veggie-CHO), COS (e.g., COS-7), retinal cell, Vero, CV1, kidney (e.g., HEK293, 293 EBNA, MSR 293, MDCK, HaK, BHK), HeLa, HepG2, WI38, MRC 5, Colo205, HB 8065, HL-60, (e.g., BHK21), Jurkat, Daudi, A431 (epidermal), CV-1, U937, 3T3, L cell, C127 cell, SP2/0, NS-0, MMT 060562, Sertoli cell, BRL 3A cell, HT1080 cell, myeloma cell, tumor cell, and a cell line derived from an aforementioned cell. In some embodiments, the cell comprises one or more viral genes, e.g. a retinal cell that expresses a viral gene (e.g., a PER.C6™ cell).

[0087] The term “conservative,” when used to describe a conservative amino acid substitution, includes substitution of an amino acid residue by another amino acid residue having a side chain R group with similar chemical properties (e.g., charge or hydrophobicity). In general, a conservative amino acid substitution will not substantially change the functional properties of interest of a protein, for example, the ability of a variable region to specifically bind a target epitope with a desired affinity. Examples of groups of amino acids that have side chains with similar chemical properties include aliphatic side chains such as glycine, alanine, valine, leucine, and isoleucine; aliphatic-hydroxyl side chains such as serine and threonine; amide-containing side chains such as asparagine and glutamine; aromatic side chains such as phenylalanine, tyrosine, and tryptophan; basic side chains such as lysine, arginine, and histidine; acidic side chains such as aspartic acid and glutamic acid; and, sulfur-containing side chains such as cysteine and methionine. Conservative amino acids substitution groups include, for example, valine/leucine/isoleucine,

phenylalanine/tyrosine, lysine/arginine, alanine/valine, glutamate/aspartate, and asparagine/glutamine. In some embodiments, a conservative amino acid substitution can be substitution of any native residue in a protein with alanine, as used in, for example, alanine scanning mutagenesis. In some embodiments, a conservative substitution is made that has a positive value in the PAM250 log-likelihood matrix disclosed in Gonnet et al. (1992) Exhaustive Matching of the Entire Protein Sequence Database, Science 256:1443-45, hereby incorporated by reference. In some embodiments, the substitution is a moderately conservative substitution wherein the substitution has a nonnegative value in the PAM250 log-likelihood matrix.

[0088] In some embodiments, residue positions in an immunoglobulin light chain or heavy chain differ by one or more conservative amino acid substitutions. In some embodiments, residue positions in an immunoglobulin light chain or functional fragment thereof (e.g., a fragment that allows expression and secretion from, e.g., a B cell) are not identical to a light chain whose amino acid sequence is listed herein, but differs by one or more conservative amino acid substitutions.

[0089] The phrase "epitope-binding protein" or "antigen binding protein" includes a protein having at least one CDR and that is capable of selectively recognizing an epitope, e.g., is capable of binding an epitope with a K_D that is at about one micromolar or lower (e.g., a K_D that is about 1×10^{-6} M, 1×10^{-7} M, 1×10^{-9} M, 1×10^{-9} M, 1×10^{-10} M, 1×10^{-11} M, or about 1×10^{-12} M). Therapeutic epitope-binding proteins (e.g., therapeutic binding proteins) frequently require a K_D that is in the nanomolar or the picomolar range.

[0090] The phrase "functional fragment" includes fragments of epitope-binding proteins that can be expressed, secreted, and specifically bind to an epitope with a K_D in the micromolar, nanomolar, or picomolar range. Specific recognition includes having a K_D that is at least in the micromolar range, the nanomolar range, or the picomolar range.

[0091] The term "identity" in connection with a comparison of sequences includes identity as determined by any of a number of different algorithms known in the art that can be used to measure nucleotide and/or amino acid sequence identity. In some embodiments, identities as described herein are determined using a ClustalW v. 1.83 (slow) alignment employing an open gap penalty of 10.0, an extend gap penalty of 0.1, and using a Gonnet similarity matrix (MACVECTOR™ 10.0.2, MacVector Inc., 2008). The term "identity" includes the overall relatedness between

polymeric molecules, e.g., between nucleic acid molecules (e.g., DNA molecules and/or RNA molecules) and/or between polypeptide molecules. In some embodiments, polymeric molecules are considered to be "substantially identical" to one another if their sequences are at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% identical. As will be understood by those skilled in the art, a variety of algorithms are available that permit comparison of sequences in order to determine their degree of homology, including by permitting gaps of designated length in one sequence relative to another when considering which residues "correspond" to one another in different sequences. Calculation of the percent identity between two nucleic acid sequences, for example, can be performed by aligning the two sequences for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second nucleic acid sequences for optimal alignment and non-corresponding sequences can be disregarded for comparison purposes). In certain embodiments, the length of a sequence aligned for comparison purposes is at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or substantially 100% of the length of the reference sequence. The nucleotides at corresponding nucleotide positions are then compared. When a position in the first sequence is occupied by the same nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which needs to be introduced for optimal alignment of the two sequences. Representative algorithms and computer programs useful in determining the percent identity between two nucleotide sequences include, for example, the algorithm of Meyers and Miller (CABIOS, 1989, 4: 11-17), which has been incorporated into the ALIGN program (version 2.0) using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. The percent identity between two nucleotide sequences can, alternatively, be determined for example using the GAP program in the GCG software package using an NWSgapdna.CMP matrix.

[0092] The phrase "micromolar range" is intended to mean 1-999 micromolar; the phrase "nanomolar range" is intended to mean 1-999 nanomolar; the phrase "picomolar range" is intended to mean 1-999 picomolar.

[0093] The term "operably linked" refers to a relationship wherein the components operably linked function in their intended manner. In one instance, a nucleic acid sequence encoding a protein may be operably linked to regulatory

sequences (e.g., promoter, enhancer, silencer sequence, *etc.*) so as to retain proper transcriptional regulation. In one instance, a nucleic acid sequence of an immunoglobulin variable region (or V(D)J segments) may be operably linked to a nucleic acid sequence of an immunoglobulin constant region so as to allow proper recombination between the sequences into a rearranged immunoglobulin heavy or light chain gene sequence.

[0094] The term "replacement" in reference to gene replacement refers to placing exogenous genetic material at an endogenous genetic locus, thereby replacing all or a portion of the endogenous gene with an orthologous or homologous nucleic acid sequence.

[0095] The term "non-human animals" is intended to include any vertebrate such as cyclostomes, bony fish, cartilaginous fish such as sharks and rays, amphibians, reptiles, mammals, and birds. Suitable non-human animals include mammals. Suitable mammals include non-human primates, goats, sheep, pigs, dogs, cows, and rodents.

[0096] In some aspects of the invention, the non-human animal includes a small mammal, e.g., of the superfamily Dipodoidea or Muroidea. In one embodiment, the genetically modified animal is a rodent. In one embodiment, the rodent is selected from a mouse, a rat, a squirrel, a porcupine, or a hamster. In one embodiment, the rodent is selected from the superfamily Muroidea. In one embodiment, the genetically modified animal is from a family selected from Calomyscidae (e.g., mouse-like hamsters), Cricetidae (e.g., hamster, New World rats and mice, voles), Muridae (true mice and rats, gerbils, spiny mice, crested rats), Nesomyidae (climbing mice, rock mice, white-tailed rats, Malagasy rats and mice), Platacanthomyidae (e.g., spiny dormice), and Spalacidae (e.g., mole rates, bamboo rats, and zokors). In a specific embodiment, the genetically modified rodent is selected from a true mouse or rat (family Muridae), a gerbil, a spiny mouse, and a crested rat. In one embodiment, the genetically modified mouse is from a member of the family Muridae. In one embodiment, the animal is a rodent. In a specific embodiment, the rodent is selected from a mouse and a rat. In one embodiment, the non-human animal is a mouse.

[0097] In a specific embodiment, the non-human animal is a rodent that is a mouse of a C57BL strain selected from C57BL/A, C57BL/An, C57BL/GrFa, C57BL/KaLwN, C57BL/6, C57BL/6J, C57BL/6ByJ, C57BL/6NJ, C57BL/10, C57BL/10ScSn, C57BL/10Cr, and C57BL/Ola. In another embodiment, the mouse is

a 129 strain selected from the group consisting of a strain that is 129P1, 129P2, 129P3, 129X1, 129S1 (*e.g.*, 129S1/SV, 129S1/SvIm), 129S2, 129S4, 129S5, 129S9/SvEvH, 129S6 (129/SvEvTac), 129S7, 129S8, 129T1, 129T2 (*see, e.g.*, Festing *et al.* (1999) Revised nomenclature for strain 129 mice, Mammalian Genome 10:836, *see also*, Auerbach *et al.* (2000) Establishment and Chimera Analysis of 129/SvEv- and C57BL/6-Derived Mouse Embryonic Stem Cell Lines). In a specific embodiment, the genetically modified mouse is a mix of an aforementioned 129 strain and an aforementioned C57BL/6 strain. In another specific embodiment, the mouse is a mix of aforementioned 129 strains, or a mix of aforementioned BL/6 strains. In a specific embodiment, the 129 strain of the mix is a 129S6 (129/SvEvTac) strain. In another embodiment, the mouse is a BALB strain, *e.g.*, BALB/c strain. In yet another embodiment, the mouse is a mix of a BALB strain and another aforementioned strain.

[0098] In one embodiment, the non-human animal is a rat. In one embodiment, the rat is selected from a Wistar rat, an LEA strain, a Sprague Dawley strain, a Fischer strain, F344, F6, and Dark Agouti. In one embodiment, the rat strain is a mix of two or more strains selected from the group consisting of Wistar, LEA, Sprague Dawley, Fischer, F344, F6, and Dark Agouti.

[0099] “Genetically engineered,” “genetically modified,” and the like, as used herein includes the artificial manipulation, modification and/or recombination of a nucleic acid sequence resulting in the production of a non-native polypeptide, *e.g.*, by an animal.

Genetically Engineered Animals for Production of Single Domain Antigen Binding Proteins

[00100] Provided herein are genetically modified non-human animals that comprise (a) single domain antigen binding proteins, *e.g.*, V_H- or V_L- single domain binding proteins, which may be respectively encoded by heavy chain variable region or light chain variable region gene sequences in a modified immunoglobulin heavy chain locus that contains one or more non-IgM immunoglobulin constant regions in which a functional C_H1 domain has been inactivated and/or removed while retaining an intact IgM C_H1 constant region, and/or (b) a genetically engineered single rearranged light chain, which may be encoded by a single rearranged variable gene sequence in a light chain locus, *e.g.*, a single rearranged V_κ:J_κ gene sequence inserted into an immunoglobulin kappa locus.

[00101] Antibodies are useful as human therapeutics. Single domain binding proteins are also useful as human therapeutics. Because single domain binding proteins lack a light chain, they are smaller and thus expected to exhibit better tissue penetration than antibodies that contain light chains, yet have a similar or more favorable pharmacokinetic profile and yet retain similar effector function as compared to a conventional antibody. Because they are smaller, single domain binding proteins are also capable of administration at a higher dose in a given volume. A frequent method of administering binding proteins is by subcutaneous injection, and a reduction in administration volume for a given dosage of antibody can provide benefits to patients and avoid complications and pain due to subcutaneous injections of large volumes.

[00102] Another advantage of single domain binding proteins is the ability to make bispecific antibodies by heterodimerizing immunoglobulin chains with specificity for two different epitopes in a single therapeutics. Because single domain binding proteins lack a light chain, they are particularly suited for making bispecific antibodies since there is no light chain rearrangement that would create a light chain that would interfere with binding affinity or specificity of the other chain.

[00103] Observations in camelids, in certain fish, and in pathological conditions reveal that under some circumstances a binding protein that lacks a functional C_H1 domain in its heavy chain constant region can be expressed in the absence of a cognate light chain. Accordingly, in one embodiment, a binding protein may be referred to a "single domain binding protein," which may also be well-known in the art as an antibody devoid of light chains that comprise a light chain variable region and a light chain constant region, i.e., a "heavy chain only antibody" comprising only one or two immunoglobulin polypeptide chains each comprising heavy chain constant region, wherein at least one of the immunoglobulin polypeptide chains of the heavy chain only antibody lacks a functional C_H1 domain. Teachings on heavy chain only antibodies are found in the art, for example, see PCT publications WO02085944, WO02085945, WO2006008548, and WO2007096779. See also U.S. Pat. No. 5,840,526; U.S. Pat. No. 5,874,541; U.S. Pat. No. 6,005,079; U.S. Pat. No. 6,765,087; U.S. Pat. No. 5,800,988; EP 1589107; WO 9734103; and U.S. Pat. No. 6,015,695, incorporated herein by reference.

[00104] Non-human animals genetically modified to produce heavy chain only antigen binding proteins are well-known in the art. See, e.g., *Janssens et al.* (2006) *PNAS* 103:15130-15135; Zou et al. (2004) *J. Exp. Med.* 204:3271-32. For example, animals, and in particular, rodents (e.g., mice) that have been genetically modified to

lack a functional C_H1 sequence, e.g., in an immunoglobulin G (IgG) gene, subsequently expressed single domain antigen binding proteins.

[00105] Although observations in camelids, certain fish, and in pathological conditions reveal that under some circumstances a binding protein that lacks a C_H1 domain of its heavy chain constant region can be expressed in the absence of a cognate light chain, normal development of antibody-producing B cells generally requires the presence of a C_H1 domain. All heavy chain isotypes, including IgM, comprise a C_H1 domain. Both the surrogate light chain and a cognate light chain are believed to interact with a given heavy chain through the heavy chain's C_H1 domain in the context of an IgM. To the extent that development of single domain binding proteins depends upon structural integrity or functionality of an IgM isotype heavy chain, disruption of the IgM's structural integrity or function would be undesirable.

[00106] Normal development of antibodies requires that antibodies survive throughout a multiplicity of complex selection schemes that result in the survival and ultimate expression of functional and useful antibodies. Disruptions in antibody structure can prove deleterious to the survival and ultimate expression of an antibody to the extent that the structural disruption results in the inability of the antibody to effectively compete and evolve to the satisfaction of one or more of nature's antibody selection schemes.

[00107] Early in antibody development, antibody heavy chains undergo a selection process wherein nature chooses, through a variety of selection schemes, suitable heavy chains to undergo further selection to eventually form functional and affinity-matured antibodies. Antibody heavy chains expressed from recombined heavy chain gene segments in progenitor B cells (or, pro-B cells) are normally paired with a surrogate light chain for presentation on the surface of the pro-B cell in an IgM isotype to form a structure (which includes other co-receptors) referred to as a pre-B cell receptor, or pre-BCR. Once the pre-BCR is presented on the cell surface, the pre-BCR is believed to signal its appropriate formation of the complex to the cell, effectively instructing the cell that the heavy chain has passed this early selection step. Thus the cell is informed that the heavy chain may undergo further selection. If the heavy chain contains a defect that is deleterious to the formation of a pre-BCR when presented in the context of an IgM and a surrogate light chain, the cell will undergo apoptosis. If the cell undergoes apoptosis, the usefulness, or contribution to diversity, of the heavy chain variable region of the heavy chain will be lost. Thus, a very early step in antibody selection requires presentation of the heavy chain together with a surrogate light chain in the context of an IgM isotype. The surrogate light chain is believed to interact with IgM at least in part through IgM's C_H1 domain.

A failure or disruption in antibody structure at this early juncture (e.g., a nonfunctional C_H1 domain) can result in clonal selection failure, loss of the pro-B cell that expresses the heavy chain, and loss of the possibility of employing the particular heavy chain variable domain in a useful antibody.

[00108] Once the cell bearing the pre-BCR passes this selection step, the next selection step requires that the heavy chain be paired with a cognate light chain (e.g., either kappa or lambda in mice and humans). The paired heavy chain/cognate light chain structure is again presented on the surface of the cell, now a naive pre-B cell, in the context of an IgM isotype through the IgM's C_H1 domain. This complex on the surface results in a functional, membrane-bound, B cell receptor (BCR). This BCR is believed to signal to the cell that the heavy chain is suitable for further selection, and that the cell may now commit to expressing this particular light chain and proceed to further B cell maturation steps, including affinity maturation and class switching. If the heavy chain contains a defect that is deleterious to the formation of a BCR when presented in the context of an IgM and its cognate light chain, the cell will undergo apoptosis. If the cell undergoes apoptosis, the usefulness, or contribution to diversity, of the heavy chain variable region of the heavy chain will be lost. Thus, a very early step in antibody selection requires presentation of the heavy chain together with a cognate light chain in the context of an IgM isotype. Again, a failure or disruption in antibody structure (e.g., a non-functional C_H1 domain) at this early juncture can result in clonal selection failure and concomitant loss of the pre-B cell that expresses the heavy chain.

[00109] Having survived selection thus far, the pre-B cell that presents the heavy chain paired with its cognate light chain in the IgM context then undergoes a maturation process that ultimately results in class switching and further selection mechanisms in which the heavy chain and cognate light chain are presented on the B cell surface in the context of an IgG isotype. It would be at this step that any selection of IgG heavy chains that lack a C_H1 domain or that lack a C_H1 domain and a hinge region would occur. In animals according to the invention, it is believed that an increased repertoire of variable regions at a heavy chain locus would be available for selection based upon whether the variable domain would survive to be expressed in an IgG heavy chain that lacks a C_H1 domain or that lacks a C_H1 domain and a hinge region. In contrast, mice that have impaired IgM would likely not present a full repertoire of heavy chain variable regions, since only those variable regions capable of surviving selection in the context of an impaired IgM would be available for class switching.

[00110] Thus, an animal lacking a functional IgM may experience a marked reduction in the ability to make a B cell population following rearrangement of otherwise suitable heavy chain variable gene segments. In such a case, even where an ample supply of variable regions is available (i.e., the animal has a suitable number of variable region gene segments capable of rearranging and operably linking to a heavy chain constant region, e.g., in a heavy chain immunoglobulin locus), a satisfactory population of B cells that display a desirable degree of diversity may not form because of an IgM impairment that mitigates against survival of a heavy chain during the selection process.

[00111] A suitable number of rearranged variable regions in a heavy chain immunoglobulin locus that can effectively survive selection when presented during B cell development in the context of an IgM is desirable to be maintained in order to generate sufficient diversity to make antibodies by immunizing a non-human animal with an immunogen of interest. Thus, a genetically modified non-human animal that comprises a nonfunctional C_H1 domain or a nonfunctional C_H1 domain and, optionally, a nonfunctional hinge region, in an immunoglobulin heavy chain should not comprise a C_H1 deletion in either or both both IgM alleles. Such animals, disclosed in US 2011/0145937, which is incorporated herein by reference, exhibit class switching to a IgG constant gene region wherein the C_H1 domain has been deleted or inactivated and express a single-domain, surface IgG (B-cell receptor) that both 1) folds and expresses at the cell surface without a light-chain partner, and 2) still recognizes antigen in the absence of light chain, in order to be stimulated by antigen and selected.

[00112] In various embodiments of the present invention, genetically modified non-human animals are provided that make binding proteins that lack a C_H1 domain, including single domain antigen binding proteins, such as but not limited to V_H and V_L single domain binding proteins. The genetically modified non-human animals may comprise a genetic modification that includes a lack of a functional immunoglobulin heavy chain domain (a C_H1 domain), e.g., an IgG1 C_H1 domain, and in some embodiments a further modification comprising a deletion of a hinge region in the immunoglobulin heavy chain that lacks the functional C_H1 domain, wherein the non-human animal expresses a functional IgM. Other modifications include rendering isotypes other than IgG1 and IgM to be nonfunctional, e.g., making deletions in genes, or deletions of genes, or inactivating mutations in genes, for IgD, IgG3, IgG2a, IgG2c, IgG2b, IgA, and IgE, such as deletions or inactivating mutations of CH1 domains or hinge regions of IgD, IgG3, IgG2a, IgG2c, IgG2b, IgA, and IgE.

Genetically modified non-human embryos, cells, and targeting constructs for making the non-human animals, non-human embryos, and cells are also provided.

[00113] Compositions and methods are also provided for making an animal that makes a binding protein that lacks an immunoglobulin C_H1 domain (and optionally a hinge region), including single domain antigen binding proteins, which may comprise V_H domains (e.g., endogenous or human V_H domains) or V_L domains (e.g., human V_L domains). The methods include selectively rendering an endogenous non-IgM C_H1 region to be nonfunctional (e.g., by a deletion or inactivation of a sequence of a C_H1 domain), and employing either unrearranged endogenous heavy chain variable region (HCVR) gene segments, unrearranged human variable region (hHCVR) gene segments, or unrearranged human light chain variable region (hLCVR) at the endogenous variable region locus to make a chimeric human binding protein in a non-human. The deletion of the C_H1 domain is made in one or more immunoglobulin constant region genes (e.g., IgG1, IgD, IgG3, IgG2a, IgG2c, IgG2b, IgA, or IgE genes), but not in an IgM gene. In an embodiment wherein the deletion is in an IgG, this approach selectively renders one or more IgG C_H1 domains nonfunctional while retaining a functional IgM. In addition to a deletion of the one or more IgG C_H1 domains, a further embodiment provides for deleting or rendering nonfunctional the hinge region(s) of the IgG(s) lacking a functional C_H1 domain.

[00114] In this particular embodiment, the IgG C_H1 deletion approach employs a relatively conservative disruption in natural B cell development in the animal, because not all Ig isotypes of the genetically modified non-human animal will exhibit a nonfunctional C_H1 or a deletion of the C_H1 domain (and, optionally, hinge). Thus, the C_H1 modification does not occur in IgM molecules and thus does not affect those steps, as described above, in early B cell development that depend on an IgM having a functional C_H1. Because the IgM is not modified, animals bearing one or more deletions of the C_H1 domain of an IgG (and optionally a hinge region of the IgG), but not the C_H1 domain of an IgM, should be able to process a satisfactorily large repertoire of variable regions in clonal selection steps prior to presentation of the variable domain in the context of an IgG. Thus in various embodiments, any deleterious effect of the genetic modification(s) on the diversity of variable regions available for use in a single domain binding protein should not negatively impact the pool of variable regions available for selection in an IgG context. Further, where the C_H1 sequence that is rendered nonfunctional (e.g., deleted) in the germline is an IgG1, the animal will lack the ability to make any RNA that encodes a C_H1 domain.

[00115] Genetically modifying a non-human animal to render a C_H1 domain or a C_H1 domain and, optionally, a hinge region of one or more non-IgM immunoglobulin

isotype nonfunctional may result in an animal that is able to select, from a full or substantially full repertoire of V region gene segments, e.g., V_H or V_L regions, a suitable V region to express in a single domain binding protein. Selectively modifying IgG isotypes (but not IgM) avoids a potential reduction in the number of variable regions that survive selection due to a lack of a C_H1 domain or a lack of a C_H1 domain in IgM. Thus, a fuller repertoire of V regions is available for selection in the context of an IgG (that lacks a C_H1 domain or that lacks a C_H1 domain and that lacks a hinge region). Thus, selection of a V domain in a genetically modified animal in accordance with the invention does not depend, e.g., on which V domain might help overcome early IgM-dependent B cell developmental hurdles that are due to modified IgM structures. Instead, early IgM-dependent steps should occur as normal, resulting in a large repertoire of heavy chains available for selection as to their suitability to express in the context of an IgG that lacks a C_H1 domain or that lacks a C_H1 domain and lacks a hinge region.

[00116] Thus, in various embodiments, a genetically modified animal in accordance with the invention should maintain functional IgM expression, which should provide an opportunity for a more natural clonal selection process. For example, with a functional IgM (e.g., an IgM that comprises a functional C_H1 domain), both surrogate light chain and the cognate light chain will be able to associate through the IgM's C_H1 domain and participate in selection processes in early B cell development. In a genetically modified animal in accordance with the invention, it is believed that class switching to an IgG isotype is the first selection step where any selection of heavy chain variable domains that can be expressed in the context of a constant domain lacking a functional C_H1 domain or lacking a functional C_H1 domain and a functional hinge is encountered.

[00117] In various embodiments, the non-IgM heavy chain constant region in the non-human animal that comprises a deletion or an inactivating mutation in the $CH1$ domain is a non-human, e.g., endogenous non-human heavy chain constant region. In another embodiment, the non-IgM heavy chain constant region in the non-human animal that comprises a deletion or an inactivating mutation in the $CH1$ domain is a human heavy chain constant region. In yet other embodiments, wherein the animal additionally comprises a single rearranged light chain gene sequence operably linked to a light chain constant region, the light chain constant region is a non-human, e.g., endogenous non-human light chain constant region; or the light chain is a human light chain constant region.

V_L-Single Domain Binding Proteins, e.g., Single Domain Antigen Binding Proteins Having a Light Chain Variable Domain

[00118] Provided herein are genetically modified non-human animals comprising two types of single domain antigen binding proteins: (1) a V_H single domain binding protein, which is encoded by a rearranged heavy chain variable region gene sequence and (2) a V_L single domain binding protein, which is encoded by a rearranged light chain variable region gene sequence, each also encoded by a modified immunoglobulin heavy chain locus that contains one or more non-IgM immunoglobulin constant regions in which a functional C_H1 domain has been inactivated and/or removed while retaining an intact IgM C_H1 constant region.

[00119] Thus, in one embodiment provided herein is a heavy chain locus genetically engineered to comprise light chain variable region gene segments, e.g., V_κ, J_κ, V_λ, and/or J_λ gene segments, operably linked to a heavy chain constant region. The genetic engineering of a heavy chain locus to comprise a light chain variable region has been described. For example, generation of a non-human animal comprising an immunoglobulin heavy chain locus comprising a replacement of one or more, substantially all, or all immunoglobulin heavy chain variable region V_H, D_H, and/or J_H gene segments with one or more light chain variable region V_L and/or J_L gene segments is described in, e.g., U.S. Patent Application No. 20120096572, which is incorporated herein by reference in its entirety.

[00120] A skilled artisan will readily recognize that the replacement V_L and/or J_L gene segments may comprise unrearranged V_L and/or unrearranged J_L gene segments, which are capable of undergoing productive rearrangement. Additionally, the V_L and/or J_L gene segments may be one or more segments selected from V_κ, J_κ, V_λ, J_λ gene segments, and may be a combination thereof. In one embodiment, the one or more heavy chain variable region gene segments are replaced with one or more human light chain variable gene segments, which allows for the production of a variable domain having human idiotypes.

[00121] As provided herein, a heavy chain locus genetically engineered to comprise light chain variable region gene segments, e.g., V_κ, J_κ, V_λ, and/or J_λ gene segments, operably linked to a heavy chain constant region may undergo productive gene rearrangement to form an immunoglobulin chain, even when one or more domains or gene segments of the heavy chain constant region is inactivated or deleted. As shown herein, replacement of the heavy chain variable region gene segments with light chain variable region gene segments coupled with deletion of a

C_H1 domain, e.g., in the IgG1 gene, results in single domain antigen binding proteins having light chain variable regions. Specifically, replacement of endogenous heavy chain variable region gene segments of a heavy chain locus with kappa (κ) V and J gene segments (V κ and J κ) results in a kappa variable region operably linked to a heavy chain constant region (KoH). Further modification to the heavy chain locus to delete the C_H1 domain(s) (C_H1 del) results in an immunoglobulin locus (KoH C_H1 del) that encodes for an immunoglobulin polypeptide chain comprising a light chain variable region and a heavy chain constant region that lacks a functional C_H1 domain, wherein the immunoglobulin polypeptide chain may form a single domain antigen binding protein, e.g., a V_L-single domain antigen binding protein.

[00122] Accordingly, in some embodiments, provided are genetically modified non-human animals that comprise a V_L single domain binding protein comprising a light chain variable domain operably linked to a heavy chain constant region that lacks a functional C_H1 domain, wherein the immunoglobulin polypeptide chain may form a single domain antigen binding protein, which may be encoded by light chain variable regions gene sequences in a modified immunoglobulin heavy chain locus that contains one or more non-IgM immunoglobulin constant regions in which a functional C_H1 domain has been inactivated and/or removed while retaining an intact IgM constant region.

[00123] Aspects described herein include V_L binding proteins that comprise a hybrid chain encoded by a hybrid immunoglobulin gene comprising or derived from a, preferably unrearranged and more preferably human, V_L gene segment (or portion thereof) rearranged with a, preferably unrearranged and more preferably human, J_L gene segment (or portion thereof) operably linked to nucleotide sequences that encode one or more heavy chain constant region genes, e.g., IgM, IgD, IgG, IgA or IgE, wherein the IgD, IgG, IgA or IgE gene comprises a deletion or inactivating mutation in a C_H1 encoding sequence. V_L binding protein, antigen binding V_L protein, or the like, includes an antigen binding protein comprising an antigen binding site comprising two light chain variable domains. In one embodiment, at least two light chain variable domains of the V_L binding proteins are cognate. In some embodiments, each of the two light chain variable domains are encoded by or derived from a light chain variable region (V_L) gene segment and/or a light chain joining region (J_L) gene segment. In preferred embodiments, one of the two light chain variable domains may be part of a hybrid immunoglobulin chain, and the other of the two light chain variable domains may be part of an immunoglobulin light chain (L).

[00124] The phrase “immunoglobulin hybrid chain,” “hybrid chain,” “hybrid immunoglobulin chain,” or the like as used herein refers to an immunoglobulin protein that includes, from amino terminus to carboxyl, a light chain variable domain (which may or may not be somatically mutated) and a heavy chain constant region. Generally, a hybrid chain is encoded by a rearranged light chain variable region gene sequence operably linked to a heavy chain constant region gene sequence. As disclosed herein, a V_L -single domain binding protein comprises a hybrid chain, wherein the hybrid chain is encoded by a rearranged light chain variable region gene sequence operably linked to a heavy chain constant region gene sequence having a deletion or inactivating mutation in a C_H1 encoding sequence.

[00125] The light chain variable region gene sequence of a hybrid immunoglobulin chain may generally comprise sequences from light chain variable (V_L) gene segment (or portion thereof) and a light chain joining (J_L) gene segment (or portion thereof). In preferred embodiments, the light chain variable region gene sequence, e.g., the rearranged V_L - J_L gene sequence, encoding the hybrid chain variable domain is derived from a repertoire of unrearranged V_L and J_L gene segments, preferably germline unrearranged V_L - and J_L gene segments, which are (a) capable of undergoing productive gene rearrangement, e.g., capable of rearranging to form an in-frame light chain variable region gene sequence and (b) operably linked to one or more heavy chain constant region gene segments, e.g., an unrearranged cluster of constant region gene segments or one constant region gene segment.

[00126] Upon rearrangement of the light chain gene segments, a rearranged nucleotide sequence is obtained that comprises a sequence encoding a light chain variable region fused with a sequence encoding a heavy chain constant region. This sequence encodes a hybrid immunoglobulin chain that has a light chain variable domain fused with a heavy chain constant domain. Thus, in one embodiment, a hybrid immunoglobulin as disclosed herein consists essentially of, from N-terminal to C-terminal, a V_L domain and a C_H domain. In one embodiment, the C_H domain comprises a functional C_H1 region (in the context of IgM), a hinge, a C_H2 region, a C_H3 region, and optionally a C_H4 region. In another embodiment, the C_H domain lacks a functional C_H1 region, e.g., lacks a C_H1 region in whole or in part, and may additionally lack a hinge region, e.g., in the context of IgG, IgA, IgE and/or IgD. In another embodiment, the C_H domain lacks a functional C_H1 region, e.g., lacks a C_H1 region in whole or in part, and may additionally lack other non-IgM isotype constant regions.

[00127] The modified non-human animals described herein may generate V_L binding proteins having an IgM isotype that also comprise a cognate light chain

paired with a hybrid chain to make a V_L binding protein that is antibody-like, e.g., may be tetrameric, but wherein instead of a heavy chain (or pair of heavy chains) the V_L binding protein comprises a hybrid chain (or pair of hybrid chains) that comprises V_L domain—not a V_H domain—fused to a IgM C_H domain.

[00128] Since the non-human animals disclosed herein preferably comprise an IgM constant region gene having a functional C_H1 domain, the non-human animals disclosed herein also encompasses the humanization of immunoglobulin loci resulting in expression of V_L binding proteins that resemble some conventional antibodies' tetrameric structure yet differ in binding characteristics, and resulting in expression of said V_L binding proteins on the membrane surface of cells of the non-human animal. In some embodiments, non-human animals of the present invention are capable of generating human V_L domains, on either or both the hybrid and light chains of the V_L binding protein, that bind to antigen; in some embodiments, such non-human mammals develop and/or have a B cell population that express binding proteins comprising variable domains that are not encoded by or derived from any V_H , D_H and/or J_H gene segment sequences. In some embodiments, V_L binding proteins expressed by such non-human animals are characterized in that the antigen-binding portion is comprised exclusively of human V_L domains. In some embodiments, non-human animals of the present invention comprise at an endogenous immunoglobulin heavy chain locus genetic material from the non-human animal and a heterologous species (e.g., a human) and comprise at an endogenous immunoglobulin light chain locus genetic material from the non-human animal and a heterologous species (e.g., human).

[00129] In various embodiments, the modified non-human animals make V_L single domain binding proteins, wherein the V_L domain of a hybrid chain exhibits an enhanced degree of somatic hypermutation over a V_L domain of a light chain. In some embodiments, a V_L region of a hybrid chain exhibits about 1.5-fold, 2-fold, 2.5-fold, 3-fold, 3.5-fold, 4-fold, 4.5-fold, or 5-fold or more somatic hypermutations than a V_L region fused with a C_L region. In some embodiments, the modified non-human animal, e.g., mouse, in response to an antigen exhibits a population of V_L single domain binding proteins that comprise a V_L domain of a hybrid chain, wherein the population of V_L single domain binding proteins exhibits an average of about 1.5-fold, 2-fold, 2.5-fold, 3-fold, 3.5-fold, 4-fold, 4.5-fold, 5-fold or more somatic hypermutations in the V_L domain of the hybrid chain than is observed in a population of light chains, e.g., a V_L domain of a light chain, exhibited by a wild-type mouse in response to the same antigen.

[00130] In one embodiment, the somatic hypermutations in the V_L domain of the hybrid chain comprises one or more or two or more N additions in a CDR3. In various embodiments, the V_L binding proteins, e.g., V_L single domain binding proteins, comprise hybrid chains comprising variable domains encoded by immunoglobulin light chain sequences that comprise a larger number of N additions than observed in nature for light chains rearranged from an endogenous light chain locus, e.g., the V_L and human J_L gene segments rearrange to form a rearranged variable region gene operably linked with a heavy chain constant region gene, wherein the rearranged light chain variable region comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions.

[00131] In various embodiments, V_L binding proteins, e.g., V_L single domain binding proteins, as disclosed herein, e.g., those produced by the genetically modified non-human animals, e.g., mice, disclosed herein, may be on average smaller than conventional antibodies or heavy chain only antibodies, respectively and possess advantages associated with smaller size. Smaller size is realized at least in part through the absence of an amino acid sequence encoded by a D_H region, normally present in a V_H domain. Smaller size can also be realized in the formation of a CDR3 that is derived, e.g., from a V_K region and a J_K region.

[00132] In one aspect, a non-human animal, e.g., a mouse, is provided, comprising an immunoglobulin hybrid chain locus. In one embodiment, the hybrid chain locus is created within an endogenous heavy chain locus, wherein one or more immunoglobulin heavy chain variable region (V_H) gene segments, heavy chain diversity (D_H) gene segments, and heavy chain joining (J_H) gene segments at an endogenous mouse immunoglobulin heavy chain locus are replaced with one or more light chain variable region (V_L) gene segments and one or more light chain joining region (J_L) gene segments that rearrange to form a rearranged variable region V_L/J_L gene sequence recombining with an endogenous mouse C_H gene to form a rearranged gene that is derived from a V_L gene segment, a J_L gene segment, and an endogenous mouse C_H gene, wherein the C_H gene is IgM, IgD, IgG, IgA, IgE, and wherein the IgD, IgG, IgA, or IgE lack a functional C_H1 domain. In one aspect, a non-human animal is provided, comprising a hybrid chain locus that replaces the endogenous immunoglobulin heavy chain locus, e.g., all or substantially all endogenous V_H , D_H , and J_H gene segments of one or both heavy chain loci are replaced with one or more V_L gene segments and one or more J_L gene segments that form a rearranged variable region V_L/J_L gene sequence recombining with an endogenous mouse C_H gene to form a rearranged gene that is derived from a V_L gene segment, a J_L gene segment, and an endogenous mouse C_H gene, wherein the

C_H gene is IgM, IgD, IgG, IgA, IgE, and wherein the IgD, IgG, IgA, or IgE lack a functional C_H1 domain.

[00133] In some embodiments, non-human animals of the present invention comprise an immunoglobulin hybrid chain locus that includes unrearranged human V_L gene segments and/or human J_L gene segments and an immunoglobulin light chain locus that includes unrearranged human V_L gene segments and/or human J_L gene segments. In some embodiments, non-human animals of the present invention comprise an immunoglobulin hybrid chain locus that includes unrearranged human V_L gene segments and/or human J_L gene segments and, preferably, an immunoglobulin light chain locus that includes a single rearranged human V_L/J_L variable region gene sequence operably linked to a light chain constant region gene sequence, e.g., and that encodes a common light chain.

Genetically Engineered Non-human Animals Expressing Single Domain Binding Proteins and a Rearranged Light Chain

[00134] In additional embodiments, provided herein are non-human animals comprising (a) a deletion or inactivating mutation in a nucleotide sequence encoding a C_H1 domain of at least one endogenous immunoglobulin heavy chain constant region gene at an endogenous immunoglobulin heavy chain locus, wherein the at least one endogenous immunoglobulin heavy chain constant region gene is IgG, IgA, IgE, IgD, or a combination thereof, (b) an immunoglobulin light chain locus that comprises a single rearranged immunoglobulin light chain variable region V_L/J_L gene sequence comprising V_L and J_L gene segment sequences, wherein the single rearranged immunoglobulin light chain variable region gene sequence is operably linked to an immunoglobulin light chain constant region gene sequence, and, e.g., encodes a single light chain, and optionally also (c) a replacement of endogenous V_H, D_H, J_H gene segments at the endogenous immunoglobulin heavy chain locus with a nucleic acid sequence comprising at least one unrearranged immunoglobulin light chain variable region (V_L) gene segment and at least one unrearranged immunoglobulin light chain joining (J_L) gene segment, wherein each of the unrearranged V_L and J_L gene segments are capable of recombining to form a rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence operably linked to the immunoglobulin heavy chain constant region gene comprising the deletion or inactivating mutation in the nucleotide sequence encoding the C_H1 domain.

[00135] The genetic engineering of a single rearranged light chain, e.g., a light chain comprising a rearranged light chain variable region has been described. For

example, generation of a universal light chain mouse (ULC) comprising a single rearranged variable gene sequence $V_L:J_L$ and generation of antigen-specific antibodies in those mice is described in, e.g., U.S. Patent Application Nos. 13/022,759, 13/093,156, 13/412,936, 13/488,628, 13/798,310, and 13/948,818 (Publication Nos. 2011/0195454, 2012/0021409, 2012/0192300, 2013/0045492, US20130185821, and US20130302836 respectively), each of which is incorporated herein by reference in its entirety. The expression of a genetically engineered single rearranged light chain, e.g., a universal light chain, causes expansion of antibodies at the early IgM stage, where the bulk of the diversity and thus antigen recognition occurs on the heavy chain. Without limitation as to the invention, it is proposed that expansion at the early IgM stage with a genetically engineered single rearranged light chain will result in more cells that bear the heavy or light chain variable regions capable of surviving to undergo class-switching to an IgG isotype and selection in the context of an IgG that lacks a functional C_H1 domain or that lacks a functional C_H1 domain and lacks a functional hinge region.

[00136] Accordingly, a genetically modified non-human animal is provided, along with methods and compositions for making the animal, wherein the genetic modification results in lack of a functional C_H1 domain (in a further embodiment lack of a functional hinge region) in an Ig domain that is not an IgM domain, and wherein the animal further expresses a genetically engineered single rearranged light chain, e.g., an engineered common light chain (ULC), which may associate with an intact IgM.

[00137] The engineered common light chain mouse described in U.S. Application Publication Nos. 2011/0195454, 2012/0021409, 2012/0192300 and 2013/0045492 comprised nucleic acid sequence encoding a limited repertoire of light chain options, e.g., common or universal light chain "ULC" that comprised no more than two V_L gene segments or a single rearranged human immunoglobulin light chain variable region sequence. To achieve such limited repertoire, a mouse was engineered to render nonfunctional or substantially nonfunctional its ability to make, or rearrange, a native mouse light chain variable domain. In one aspect, this was achieved, e.g., by deleting the mouse's light chain variable region gene segments. As previously described, the endogenous mouse locus can then be modified by exogenous suitable light chain variable region gene segments of choice, preferably human light chain variable region gene segments, operably linked to the endogenous mouse light chain constant domain, in a manner such that the exogenous variable region gene segments can combine with the endogenous mouse light chain constant region gene and form a rearranged reverse chimeric light chain gene (human variable, mouse

constant). In various embodiments, the light chain variable region is capable of being somatically mutated. In various embodiments, to maximize ability of the light chain variable region to acquire somatic mutations, the appropriate enhancer(s) is retained in the mouse. In one aspect, in modifying a mouse κ light chain locus to replace endogenous mouse κ light chain gene segments with human κ light chain gene segments, the mouse κ intronic enhancer and mouse κ 3' enhancer are functionally maintained, or undisrupted.

[00138] Thus, provided was a genetically engineered mouse that expresses a limited repertoire of reverse chimeric (human variable, mouse constant) light chains associated with a diversity of reverse chimeric (human variable, mouse constant) heavy chains. In various embodiments, the endogenous mouse κ light chain gene segments are deleted and replaced with a single (or two) rearranged human light chain region, operably linked to the endogenous mouse C κ gene. In embodiments for maximizing somatic hypermutation of the rearranged human light chain region, the mouse κ intronic enhancer and the mouse κ 3' enhancer are maintained. In various embodiments, the mouse also comprises a nonfunctional λ light chain locus, or a deletion thereof or a deletion that renders the locus unable to make a λ light chain.

[00139] Thus, in one embodiment, provided herein is a non-human animal (e.g., a rodent, e.g., a mouse or a rat) that comprises in its genome, e.g., in its germline, a limited repertoire of preferably human light chain variable regions, or a single rearranged human light chain variable region, from a limited repertoire of preferably human light chain variable gene segments, wherein the non-human animal also comprises in its genome, e.g., in its germline, a deletion or inactivating mutation in a nucleotide sequence encoding a C_H1 domain.

[00140] Genetically engineered animals are provided that express a limited repertoire of human light chain variable domains, or a single human light chain variable domain, from a limited repertoire of human light chain variable region gene sequences. In one embodiment, the single rearranged V/J human light chain sequence is selected from V κ 1-39J κ and V κ 3-20J κ , e.g., V κ 1-39J κ 5 and V κ 3-20J κ 1. In some embodiments, a non-human animal as disclosed herein comprises a modified light chain locus comprising a replacement all endogenous V_L and all endogenous J_L gene segments with the single rearranged V/J light chain sequence, wherein the single rearranged V/J light chain sequence is operably linked to an endogenous light chain constant region gene. In some embodiments, the modified

light chain locus is in the germline genome of the non-human animal. In one embodiment, the non-human animal comprises in its germline genome a single rearranged light chain variable gene sequence operably linked to a light chain constant region gene sequence, wherein the single rearranged light chain variable region gene sequence comprises human germline V_L and human germline J_L gene segments, e.g., human germline $V_{\kappa 1-39}$ and human germline $J_{\kappa 5}$ or human germline $V_{\kappa 3-20}$ and $J_{\kappa 1}$. In some embodiments, a non-human animal as disclosed herein comprises a B cell, e.g., a B cell that has not undergone class switching, comprising in its genome a single rearranged V/J light chain sequence operably linked to an endogenous light chain constant region gene, wherein the single rearranged V/J light chain does not comprise somatic mutations compared to a single rearranged V/J light chain sequence operably linked to an endogenous light chain constant region gene found in the germline genome of the non-human animal. In other embodiments, a non-human animal as disclosed herein comprises a B cell, e.g., a B cell that has undergone class switching, comprising in its genome a single rearranged V/J light chain sequence operably linked to an endogenous light chain constant region gene, wherein the single rearranged V/J light chain comprises somatic mutations compared to a single rearranged V/J light chain sequence operably linked to an endogenous light chain constant region gene found in the germline genome of the non-human animal.

Making Genetically Modified Animals

[00141] Methods of producing a non-human animal as disclosed herein are also provided. Such methods comprise (a) inactivating or deleting the C_H1 domain, and optionally the hinge region(s), of a heavy chain immunoglobulin locus of the non-human animal, such as the IgG1 heavy chain locus, introducing a nucleic acid encoding a genetically engineered rearranged light chain locus, and causing the animal to express the heavy chain immunoglobulin locus having an inactivated C_H1 domain and the genetically rearranged light chain locus (ULC).

[00142] Genetic modifications for making an animal that expresses a single domain binding protein are conveniently described herein by using the mouse as an illustration, although such modifications may be easily adapted and applied to other animals. A genetically modified animal according to the invention can be made in a variety of ways, particular embodiments of which are discussed below.

[00143] An exemplary schematic illustration (not to scale) of an IgG1 locus is provided in **FIG. 1** (top) to show C_H domain arrangement at the IgG1 locus. As

illustrated, domains C_H1, C_H2, and C_H3 and the hinge region are present in readily identifiable spans of nucleotide downstream of a switch region.

[00144] A genetically modified non-human animal, e.g., mouse, lacking a functional nucleotide sequence encoding a C_H1 domain of an IgG1 but containing a hinge region can be made by any method known in the art. For example, a targeting vector can be made that replaces the IgG1 gene with a truncated IgG1 lacking a C_H1 domain but containing the hinge. In one example, a mouse genome is targeted by a targeting construct having a 5' (with respect to the direction of transcription of the genomic IgG1 gene) homology arm containing sequence upstream of the endogenous C_H1 domain, followed by nucleotide sequences that encode an IgG1 hinge, an IgG1 C_H2 domain, an IgG1 C_H3 domain, a drug selection cassette (e.g., a loxed resistance gene), and an IgG1 transmembrane domain, and a 3' homology arm containing sequences downstream with respect to the transmembrane domain. Upon homologous recombination at the locus and removal of the drug selection cassette (e.g., by Cre treatment), the endogenous IgG1 is replaced by an IgG1 that lacks a C_H1 domain (**FIG. 3**) (IgG1 Δ C_H1; I). In some embodiments, the structure of the resulting locus, which will express an IgG1 has a J region sequence fused to the hinge sequence.

[00145] A genetically modified non-human animal, e.g., mouse, lacking a nucleotide sequence encoding a C_H1 domain of an IgG1 and lacking a nucleotide sequence encoding a hinge region can be made by any method known in the art. For example, a targeting vector can be made that replaces the IgG1 gene with a truncated IgG1 lacking a sequence encoding a C_H1 domain and lacking a sequence encoding the hinge region. In another embodiment, a mouse genome is targeted by a targeting construct having a 5' (with respect to the direction of transcription of the genomic IgG1 gene) homology arm containing sequence upstream of the endogenous C_H1 domain, followed by nucleotide sequences that encode an IgG1 C_H2 domain, an IgG1 C_H3 domain, a drug selection cassette (e.g., a loxed resistance gene), and an IgG1 transmembrane domain, and a 3' homology arm containing sequences downstream with respect to the transmembrane domain. Upon homologous recombination at the locus and removal of the drug selection cassette (e.g., by Cre treatment), the endogenous IgG1 gene is replaced by an IgG1 gene that lacks a sequence encoding a C_H1 domain (**FIG. 3**) (IgG1 Δ C_H1&hinge; II). In some embodiments, the structure of the resulting locus will express an IgG1 having a J region sequence fused to the C_H2 domain.

[00146] A genetically modified non-human animal, e.g., mouse, lacking an IgG1 C_H1 sequence (IgG1ΔC_H1), or lacking an IgG1 C_H1 sequence and lacking a hinge (IgG1ΔC_H1&hinge), can be further modified to favor usage of the modified IgG1 isotype by deleting one or more other IgG isotypes, e.g., IgG2b and IgG2a/IgG2c, and/or one or more other Ig isotypes, e.g., IgD, IgA, and/or IgE, by deleting or functionally disabling sequences encoding these isotypes. For example, a targeting construct is made having a 5' homology arm containing sequence upstream of the endogenous hinge region sequence (or upstream of the endogenous C_H1 domain sequence), sequences that encode the IgG1 C_H2 and C_H3 domains, a drug selection cassette followed by a sequence encoding the IgG1 transmembrane domain, followed by another drug selection cassette if desired, and a 3' homology arm containing sequences downstream with respect to the IgG2a/c gene. Upon homologous recombination at the locus and removal of the drug selection cassette(s) (e.g., by Cre treatment), the endogenous heavy chain constant locus contains only two IgG genes: an endogenous IgG3 and the IgG1ΔC_H1 or IgG1ΔC_H1&hinge. (**FIG. 3**) (IgG1ΔC_H1ΔIgG2b/2a; III or IgG1ΔC_H1&hinge ΔIgG2b/2a; IV).

[00147] An animal engineered as described above may be further modified to comprise a deletion or inactivating mutation in the IgG2a, IgG2b, IgG2c, IgG3, IgD, IgA, and/or IgE gene segments of a heavy chain locus. For example, a targeting vector can be made that deletes the constant region gene sequence of the heavy chain locus. In one example, a mouse genome is targeted by a targeting construct having a 5' (with respect to the direction of transcription of the genomic constant region gene sequence) homology arm containing sequence upstream of the endogenous IgM domain, followed by nucleotide sequences that encode a drug selection cassette (e.g., a loxed resistance gene) and a 3' homology arm containing sequence downstream of the IgA gene segment. Upon homologous recombination at the locus and removal of the drug selection cassette (e.g., by Cre treatment), the endogenous constant region is deleted and/or replaced with a selectable marker. (**FIG. 4C**). The animal may be further modified with a targeting vector can be made that reintroduces an IgM gene segment and an IgG1 gene that lacks a functional C_H1 domain sequence and optionally lacks a functional hinge region. (**FIG. 4D**). In one example, a genome of an animal is targeted by a targeting construct having a 5' homology arm containing sequence upstream of the selectable marker gene, followed by nucleotide sequences that encode a complete IgM constant region and an IgG1 constant region lacking a functional C_H1 domain, and optionally lacking a functional hinge, a drug selection cassette (e.g., a loxed resistance gene), and a 3'

homology arm containing sequence downstream with respect to the selectable marker. Upon homologous recombination at the locus and removal of the drug selection cassette (e.g., by Cre treatment), an IgM gene segment and an IgG1 gene that lacks a functional C_H1 domain sequence and optionally lacks a functional hinge region is reintroduced. **(FIG. 3)** (IgG1 Δ C_H1 Δ IgG2b/2a Δ IgG3 Δ IgD/A/E (optionally Δ hinge); V). Other manipulations of endogenous immunoglobulin loci, e.g., deletions or inactivating mutations of CH1 region(s) of various non-IgM immunoglobulin isotypes are also provided.

[00148] In addition to genetic manipulation that introduces a deletion or inactivation into a C_H1 domain and, optionally a hinge, of a non-IgM immunoglobulin constant region by designing an appropriate constant region construct and introducing said construct into the locus by homologous recombination as described above, a deletion or an inactivation in a non-IgM C_H1 may be made by other methods known in the art, e.g., a conditional non-IgM C_H1 deletion that is induced in a mouse only upon antigen immunization, etc. Methods for conditional inactivation of loci are known in the art.

[00149] Genetic modification of the heavy chain locus as described above may further comprise replacement of one or more, substantially all, or all of the endogenous heavy chain variable gene segments, e.g., the V_H gene segments, D_H gene segments and/or J_H gene segments with (a) human V_H gene segments, D_H gene segments and/or J_H gene segments, which may be rearranged or capable of undergoing rearrangement to encode for binding proteins having human idiotypes, (b) light chain variable gene segments, e.g., light chain V gene segments and/or light chain J gene segments, which may be rearranged or capable of undergoing rearrangement to encode for immunoglobulin polypeptide chains having a light chain variable region linked to a heavy chain constant region lacking a functional C_H1 domain, e.g., V_L single domain binding proteins, e.g., a single domain antigen binding protein comprising a light chain variable region, or (c) human light chain variable gene segments, e.g., human light chain V gene segments and/or human light chain J gene segments, which may be rearranged or capable of undergoing rearrangement to encode for immunoglobulin polypeptide chains having a human light chain variable region linked to a heavy chain constant region lacking a functional C_H1 domain, e.g., V_L single domain binding proteins, e.g., a single domain antigen binding protein comprising a human light chain variable region having human idiotypes.

[00150] A schematic illustration (not to scale) of a mouse heavy chain and a human κ light chain loci is provided in **FIG. 14** to show the approximately 200 heavy

chain variable (V_H) gene segments, 13 heavy chain diversity (D_H) gene segments and 4 heavy chain joining (J_H) gene segments as well as enhancers (Enh) and heavy chain constant (C_H) regions of the mouse locus, and the about 76 V_k gene segments, 5 J_k gene segments, an intronic enhancer (Enh) and a single constant region (C_k) of the human κ locus.

[00151] Shown in **FIG. 15** is a schematic illustration (not to scale) for inserting human κ gene segments into a murine heavy chain locus, which was modified by homologous recombination to inactivate the endogenous mouse heavy chain locus through targeted deletion of mV_H , mD_H and mJ_H gene segments. As shown in **FIG. 15**, four separate targeting vectors may be used to progressively insert human V_k gene segments and human J_k gene segments into the inactivated mouse heavy chain locus using standard molecular techniques recognized in the art. The human κ gene segments used for engineering the four targeting constructs may be naturally found in proximal contig of the germline human κ light chain locus.

[00152] A genetically modified mouse comprising a genetically engineered rearranged light chain can be made by any method known in the art. For example, a targeting vector can be made that replaces either the endogenous unrearranged light chain variable V and J gene segments of an endogenous light chain locus with a single rearranged V:J gene, or the entire unrearranged light chain locus with a genetically engineered light chain locus comprising a single rearranged V:J gene operably linked to a light chain constant region.

[00153] In another aspect, a non-human animal as described herein is further engineered to comprise an ectopic nucleotide sequence encoding ADAM 6 (ADAM6a and/or ADAM6b), a functional fragment, homolog or ortholog thereof. In some embodiments, a heavy chain locus of a non-human animal described herein is further engineered to comprise an ectopic nucleotide sequence encoding a mouse ADAM6 (ADAM6a and/or ADAM6b), a functional fragment, homolog or ortholog thereof. In various embodiments, the ADAM6 protein is functional in a non-human male animal. Methods and compositions for engineering such non-human animals are described, e.g., in U.S. Patent No. 8,642,835, which is incorporated herein by reference.

[00154] In some embodiments, genetically modified animal as described above, and others, are made by introducing a suitable targeting construct into a suitable ES cell (in one or more independent targetings), and positive clones comprising a marker or selection cassette of the targeting construct are identified and grown. Clones are then employed as donor ES cells in a host embryo under conditions suitable for making a chimeric animal or a fully ES cell-derived animal. The marker or selection

cassette can be optionally removed, either at the ES cell stage or in the chimeric or ES cell-derived mouse, e.g., by employing a loxed cassette and breeding to a Cre-containing strain, or by electroporating the ES cell with a Cre expression vector. Accordingly, in some embodiments, the genetic modification occurs in the germline of the animal.

[00155] In some embodiments, the method of making an animal as disclosed herein comprises crossing a first animal capable of producing a single domain antigen binding protein, e.g. a first animal comprising in its germline an IgG heavy chain locus lacking a functional C_H1 domain, with a second animal capable of producing a genetically engineered rearranged light chain, e.g., a second animal comprising in its germline a genetically engineered light chain locus having a single rearranged V:J variable region operably linked to a light chain constant region, to produce an F1 genetically engineered animal, wherein the F1 animal comprises the IgG heavy chain locus of the first animal and the light chain locus of the second animal. The crossing may be done by animal breeding or by otherwise combining gametes, including in vitro manipulations.

[00156] For the non-human animals where suitable genetically modifiable ES cells are not readily available, methods distinct from those described herein are employed to make a non-human animal comprising the genetic modification. Such methods include, e.g., modifying a non-ES cell genome (e.g., a fibroblast or an induced pluripotent cell) and employing nuclear transfer to transfer the modified genome to a suitable cell, e.g., an oocyte, and gestating the modified cell (e.g., the modified oocyte) in a non-human animal under suitable conditions to form an embryo.

Making Single Domain Antigen Binding Proteins

[00157] Once a genetically engineered animal capable of producing single domain antigen binding proteins and/or a genetically engineered single rearranged light chain is obtained, immunoglobulins and binding protein preparations against an antigen can be readily obtained by immunizing the animal with the antigen. "Polyclonal antisera composition" as used herein includes affinity purified polyclonal binding protein preparations.

[00158] In one aspect, a method for making a binding protein that lacks a C_H1 domain is provided, comprising: (a) immunizing a non-human animal as described herein with an antigen; (b) maintaining the non-human animal under conditions sufficient for the non-human animal to make a binding protein; (c) identifying a binding protein made by the mouse that lacks a functional C_H1 domain and/or that lacks a functional hinge region; and, (d) isolating from the non-human animal the

binding protein, a cell that makes the binding protein, or a nucleotide sequence that encodes a sequence of the binding protein.

[00159] A variety of antigens can be used to immunize a transgenic animal. Such antigens include but are not limited to, cellular proteins, microorganisms, e.g. viruses and unicellular organisms (such as bacteria and fungi), alive, attenuated or dead, fragments of the microorganisms, or antigenic molecules isolated from the microorganisms.

[00160] The antigens can be administered to a transgenic animal in any convenient manner, with or without an adjuvant, and can be administered in accordance with a predetermined schedule.

[00161] For making a monoclonal binding protein, spleen cells are isolated from the immunized transgenic animal and used either in cell fusion with transformed cell lines for the production of hybridomas, or cDNAs encoding antibodies are cloned by standard molecular biology techniques and expressed in transfected cells. The procedures for making monoclonal antibodies are well established in the art. See, e.g., European Patent Application 0 583 980 A1 ("Method For Generating Monoclonal Antibodies From Rabbits"), U.S. Pat. No. 4,977,081 ("Stable Rabbit-Mouse Hybridomas And Secretion Products Thereof"), WO 97/16537 ("Stable Chicken B-cell Line And Method of Use Thereof"), and EP 0 491 057 B1 ("Hybridoma Which Produces Avian Specific Immunoglobulin G"), the disclosures of which are incorporated herein by reference. In vitro production of monoclonal antibodies from cloned cDNA molecules has been described by Andris-Widhopf et al., "Methods for the generation of chicken monoclonal antibody fragments by phage display", J Immunol Methods 242:159 (2000), and by Burton, D. R., "Phage display", Immunotechnology 1:87 (1995).

[00162] Once monoclonal single domain antigen binding proteins have been generated, such binding proteins can be easily converted into fully human binding proteins using standard molecular biology techniques, if desired. Fully human monoclonal binding proteins are not immunogenic in humans and are appropriate for use in the therapeutic treatment of human subjects.

[00163] Thus, in one embodiment, wherein the single domain antigen binding protein comprises a human V_H or a human V_L region and a mouse heavy chain constant region comprising a deletion or an inactivating mutation in a non-IgM C_H1 domain, the sequence of V_H or V_L domain of the single domain antigen binding protein can be cloned upstream of a human constant region, optionally lacking a C_H1 domain in a suitable expression vector resulting in an expression construct encoding

fully human single domain antigen binding protein that can be expressed in a suitable cell, e.g., cell typically for antibody expression, e.g., eukaryotic cell, e.g., a CHO cell.

[00164] Accordingly, also provided herein are monoclonal binding protein producing cells derived from animals genetically modified as disclosed herein, as well as nucleic acids derived therefrom. Also provided are hybridomas derived therefrom. Also provided are fully human single domain binding proteins, as well as encoding nucleic acids, derived therefrom.

[00165] Single domain antigen binding proteins described herein may also be used to make bispecific antibodies. An advantage of single domain antigen binding proteins described herein is the ability to make bispecific antibodies by heterodimerizing heavy chains with specificity for two different epitopes in a single therapeutics.

EXAMPLES

[00166] The following examples are provided so as to describe to those of ordinary skill in the art how to make and use methods and compositions of the invention, and are not intended to limit the scope of what the inventors regard as their invention. Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. The Examples do not include detailed descriptions of conventional methods that would be well known to those of ordinary skill in the art (molecular cloning techniques, etc.).

Example 1: Mice encoding a V_H single domain binding protein: Mice comprising immunoglobulin chain having a heavy chain variable region and a heavy chain constant region lacking a functional C_H1 domain, and comprising a single rearranged light chain (ULC).

Example 1.1: Generation of animals

[00167] Mice genetically modified to comprise a heavy chain locus comprising a complete and functional IgM gene sequence and an IgG1 gene sequence that lacks a functional C_H1 gene sequence, and optionally lacks a functional hinge region (**FIG. 1A**), were made according to the methods described in US2011/0145937, Macdonald et al., which is incorporated herein by reference. Several versions of mice that lacked various combinations of heavy chain constant gene sequences and contained deletion of C_H1 domain(s) but comprised a complete and functional IgM were made (**FIG. 3**). For example, mice homozygous for a heavy chain locus

comprising mouse heavy chain variable gene segments and a complete and functional IgM gene sequence and an IgG1 gene sequence that lacks a functional C_H1 gene sequence and hinge region (mV_HIgG1ΔC_H1&hinge; 1576 HO, “II” in **FIGs. 3 and 4B**) were made. Additionally, mice homozygous for a heavy chain locus comprising human heavy chain variable gene segments and a complete and functional IgM gene sequence, an IgG1 gene sequence that lacks a functional C_H1 gene sequence and hinge region, and lacking the IgG2b and IgG2a gene sequences (hV_HIgG1ΔC_H1&hingeΔIgG2b/2a; 1859 HO, see, e.g., “IV” in **FIG. 3**) were made. Additionally, mice homozygous for a heavy chain locus comprising human heavy chain variable gene segments and a complete and functional IgM gene sequence, an IgG1 gene sequence that lacks a functional C_H1 gene sequence, and lacking the IgG2b and IgG2a gene sequences (hV_HIgG1ΔC_H1ΔIgG2b/2a; 1673 HO, see, e.g., “III” in **FIGs. 3 and 4A**) were made. Additionally, mice homozygous for a heavy chain locus comprising human heavy chain variable gene segments and a complete and functional IgM gene sequence, an IgG1 gene sequence that lacks a functional C_H1, and lacking IgD, IgG2a, IgG2b, IgG3, IgE, and IgA gene sequences (hV_HIgG1ΔC_H1ΔIgG2b/2aΔIgG3ΔIgD/A/E; 6180 HO, see “V” in **FIGs. 3 and 4C-D**) were made. Additional exemplary versions of modification in the heavy chain constant region are presented in **FIG. 3**. Other variations of combinations of C_H1 deletions/inactivations and/or immunoglobulin constant gene deletions/inactivations are made, e.g., a mouse is made wherein both IgG1 and IgG2a comprise C_H1 domain deletions, and the mouse also comprises a deletion of IgD, IgE, IgG3, and IgG2b. As shown in **FIG. 4**, heavy chain loci may also be modified to comprise human variable regions, which may be human heavy chain variable regions [or human light chain variable regions (**FIG. 16**, see Examples 2 and 3 below)]. Heavy chain loci may also be modified to comprise an Adam 6a gene, an Adam 6b gene, or both, or a fragment of the gene, wherein the gene or the fragment thereof is functional in a male mouse (see, e.g., U.S. 2012/0322108, incorporated herein by reference).

[00168] Generation of a common light chain mouse (also referred to as universal light chain or ULC mice) comprising a single rearranged variable gene sequence V:J (e.g., Vκ1-39Jκ5 or Vκ3-20Jκ1 common light chain mouse) and generation of antigen-specific antibodies in those mice is described in, e.g., U.S. Patent Application Nos. 13/022,759, 13/093,156, 13/412,936, 13/488,628, 13/798,310, and 13/948,818 (Publication Nos. 2011/0195454, 2012/0021409, 2012/0192300, 2013/0045492, US20130185821, and US20130302836 respectively), each of which is incorporated herein by reference in its entirety. Specifically, mice that express the genetically

engineered V κ 1-39J κ 5 kappa light chain (1633 HO or 1634 HO) or the genetically engineered V κ 3-20J κ 1 kappa light chain (1635 HO or 1636 HO) in their germline were made.

[00169] VELOCIMMUNE® mice containing a single rearranged human germline light chain region (ULC V κ 1-39J κ 5; 1633 or 1634) or (ULC V κ 3-20J κ 1; 1635 or 1636) are bred to mice carrying a modified IgG constant region. Specifically, such ULC mice were bred to mice having a murine heavy chain variable region operably linked to a murine heavy chain constant region wherein the IgG1 C_H1 and IgG1 hinge regions were deleted or inactivated (mV_HIgG1 Δ C_H1&hinge, 1576), mice having a human heavy chain variable region operably linked to a murine heavy chain constant region wherein the IgG1 C_H1 and IgG1 hinge regions, and the IgG2a and IgG2b genes were deleted or inactivated (hV_H IgG1 Δ C_H1&hinge Δ IgG2b/2a; 1859), mice having a human heavy chain variable region operably linked to a murine heavy chain constant region wherein the IgG1 C_H1 regions, and the IgG2a and IgG2b genes were deleted or inactivated (hV_H IgG1 Δ C_H1 Δ IgG2b/2a; 1673), or mice having a human heavy chain variable region operably linked to a murine heavy chain constant region wherein the IgG1 C_H1 regions, and the IgG2a and IgG2b, IgD, IgG3, IgA, and IgE genes were deleted or inactivated (hV_H IgG1 Δ C_H1 Δ IgG2b/2a Δ IgG3 Δ IgD/A/E; 6180), to obtain the following progeny mice:

[00170] mV_HIgG1 Δ C_H1&hinge x V κ 3-20J κ 1 ULC or mV_HIgG1 Δ C_H1&hinge x V κ 1-39J κ 5 ULC homozygous mice (1576HO 1635HO or 1576 HO 1633 HO),

[00171] hV_H IgG1 Δ C_H1&hinge Δ IgG2b/2a x V κ 3-20J κ 1 ULC or hV_H IgG1 Δ C_H1&hinge Δ IgG2b/2a x V κ 1-39J κ 5 ULC homozygous mice (1859 HO 1635 HO or 1859HO 1633HO),

[00172] hV_H IgG1 Δ C_H1 Δ IgG2b/2a x V κ 3-20J κ 1 ULC or hV_H IgG1 Δ C_H1 Δ IgG2b/2a x V κ 1-39J κ 5 ULC homozygous mice (1673HO 1635 HO or 1673 HO 1633 HO), and

[00173] hV_HIgG1 Δ C_H1&hinge Δ IgG2a/2b Δ IgG3 Δ IgD/A/E x V κ 3-20J κ 1 ULC or hV_HIgG1 Δ C_H1&hinge Δ IgG2a/2b Δ IgG3 Δ IgD/A/E x V κ 1-39J κ 5 ULC homozygous mice (6180 HO 1635 HO or 6180 HO 1634 HO).

[00174] Other versions of the mice comprising a deletion or inactivating mutation in a C_H1 domain and a deletion or inactivating mutation in an immunoglobulin constant region gene are bred to mice containing a single rearranged human germline light chain region as described above.

Example 1.2: Immunization of mice with antigen and expression of single domain binding proteins.

[00175] Mice homozygous for modifications were immunized with different antigens and boosted by various routes using a variety of adjuvants. Titers for IgG1 specific responses were evaluated by ELISA or western blot.

[00176] As shown in **FIG. 5A**, mice homozygous for both IgG1 ΔC_H1 /hinge and ULC modifications exhibit increased expression of total amount of IgG1 in the serum both before and after immunization, compared to a mouse having IgG1 ΔC_H1 alone. The higher titers in ΔC_H1 /hinge x ULC mice suggest that presence of universal light chain increases the likelihood of generating an antigen specific single domain antigen binding protein. **FIG. 5B**.

[00177] **FIG. 7** demonstrates that high titers may be obtained with various versions of mice genetically engineered to comprise ΔC_H1 and ULC modifications.

[00178] Additionally, **FIGs. 6** and **8** show that single domain antigen-binding proteins are present and may be isolated from mice comprising both human heavy chain variable region with a deletion in a C_H1 region and a single rearranged light chain (universal light chain).

Example 1.3: B cell development and maturation in mice expressing single domain binding proteins and a single rearranged light chain (ULC)

[00179] B cell contents of the spleen, blood and bone marrow compartments from mice homozygous for a modified C_H1 domain and a single rearranged light chain (ULC) (see **FIG. 2**) were analyzed for progression through B cell development and B cell maturation using flow cytometry of various cell surface markers as indicated herein.

[00180] Briefly, ULC mice and mice homozygous a modified C_H1 domain and a rearranged light chain were sacrificed and blood, spleens and bone marrow were harvested. Blood was collected into microtainer tubes with EDTA (BD Biosciences). Bone marrow was collected from femurs by flushing with complete RPMI medium (RPMI medium supplemented with fetal calf serum, sodium pyruvate, Hepes, 2-mercaptoethanol, non-essential amino acids, and gentamycin). RBCs from spleen and bone marrow preparations were lysed with ACK lysis buffer (Lonza Walkersville), followed by washing with complete RPMI medium.

[00181] Cells (1×10^6) were incubated with anti-mouse CD16/CD32 (2.4G2, BD) on ice for ten minutes, followed by labeling with the following antibodies for thirty minutes on ice: APC-H7 conjugated anti-mouse CD19 (clone 1D3, BD), Pacific Blue conjugated anti-mouse CD3 (clone 17A2, BIOLEGEND®), PeCy7-IgM (II/41,

EBIOSCIENCE®), PerCP-Cy5.5-IgD (11-26c.2a, BIOLEGEND®), APC-eFluor 780-B220 (RA3-6B2, EBIOSCIENCE®), APC-CD19 (MB19-1, EBIOSCIENCE®), PE-CD93 (AA4.1, BIOLEGEND®), FITC-CD23 (B3B4, BD), APC-CD21/CD35 (7G6, BD). Bone marrow: immature B cells (B220^{int}IgM⁺), mature B cells (B220^{hi}IgM⁺). Blood and spleen: B cells (CD19⁺), mature B cells (CD19⁺IgM^{int}IgD^{hi}), transitional/immature B cells (CD19⁺IgM^{hi}IgD^{int}).

[00182] Following staining, cells were washed and fixed in 2% formaldehyde. Data acquisition was performed on a LSRII flow cytometer and analyzed with FLOWJO™ software (Tree Star, Inc.). **FIG. 9** shows that these mice have normal serum steady state IgM and IgG levels. **FIGs. 10 and 11** show the results for the splenic compartment, demonstrating nearly normal B cell numbers and nearly normal B cell maturation in the spleen. **FIG. 12 and 13** show the results for the bone marrow compartment, demonstrating normal B cell numbers and nearly normal B cell development in the bone marrow.

Example 2: Mice encoding a V_L single domain binding protein: Mice comprising an immunoglobulin chain having a light chain variable region and a heavy chain constant region lacking a functional C_H1 domain.

Example 2.1: Generation of animals

[00183] A mouse having light chain gene segments introduced into a heavy chain locus was generated as described in U.S. Patent Publication No. 2012/0096572. Specifically, various targeting constructs were made using VELOCIGENE® genetic engineering technology to modify mouse genomic Bacterial Artificial Chromosome (BAC) libraries (see, e.g., US Pat. No. 6,586,251 and Valenzuela, D.M., Murphy, A.J., Frendewey, D., Gale, N.W., Economides, A.N., Auerbach, W., Poueymirou, W.T., Adams, N.C., Rojas, J., Yasenchak, J., Chernomorsky, R., Boucher, M., Elsasser, A.L., Esau, L., Zheng, J., Griffiths, J.A., Wang, X., Su, H., Xue, Y., Dominguez, M.G., Noguera, I., Torres, R., Macdonald, L.E., Stewart, A.F., DeChiara, T.M., Yancopoulos, G.D. (2003). High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. *Nat Biotechnol* 21, 652-659). Mouse BAC DNA was modified by homologous recombination to inactivate the endogenous mouse heavy chain locus through targeted deletion of V_H, D_H and J_H gene segments for the ensuing insertion of unrearranged human germline κ light chain gene sequences (top of **FIG. 15**).

[00184] Briefly, the mouse heavy chain locus was deleted in two successive targeting events using recombinase-mediated recombination. The first targeting event included a targeting at the 5' end of the mouse heavy chain locus using a

targeting vector comprising from 5' to 3' a 5' mouse homology arm, a recombinase recognition site, a neomycin cassette and a 3' homology arm. The 5' and 3' homology arms contained sequence 5' of the mouse heavy chain locus. The second targeting event included a targeting at the 3' end of the mouse heavy chain locus in the region of the J_H gene segments using a second targeting vector that contained from 5' to 3' a 5' mouse homology arm, a 5' recombinase recognition site, a second recombinase recognition site, a hygromycin cassette, a third recombinase recognition site, and a 3' mouse homology arm. The 5' and 3' homology arms contained sequence flanking the mouse J_H gene segments and 5' of the intronic enhancer and constant regions. Positive ES cells containing a modified heavy chain locus targeted with both targeting vectors (as described above) were confirmed by karyotyping. DNA was then isolated from the double-targeted ES cells and subjected to treatment with a recombinase thereby mediating the deletion of genomic DNA of the mouse heavy chain locus between the 5' recombinase recognition site in the first targeting vector and the 5' recombinase recognition site in the second targeting vector, leaving a single recombinase recognition site and the hygromycin cassette flanked by two recombinase recognition sites (see top of **FIG. 15**). Thus a modified mouse heavy chain locus containing intact C_H genes was created for progressively inserting human κ germline gene segments in a precise manner using targeting vectors as outlined in **FIG. 15**.

[00185] Four separate targeting vectors were engineered to progressively insert 40 human V κ gene segments and five human J κ gene segments into the inactivated mouse heavy chain locus (described above) (**FIG. 15**). The human κ gene segments used for engineering the four targeting constructs are naturally found in proximal contig of the germline human κ light chain locus (**FIG. 14B**).

[00186] Mice heterozygous for such modified heavy chain loci were bred to obtain a mouse homozygous for the heavy chain locus as described above. Embryonic stem cells comprising such modified heavy chain loci comprising light chain variable region gene segments were targeted according to the scheme provided in **FIG. 16** and using the methods described in U.S. Patent Publication No. US2011/0145937, Macdonald et al., which is incorporated herein by reference, to produce mice homozygous for a heavy chain locus comprising a light chain variable region and a heavy chain constant region that lacks a functional C_H1 domain in the IgG1 gene, and further lacks the IgG2b and IgG2a genes.

[00187] Thus, the germline of the modified heavy chain loci mice comprising light chain variable region gene segments described, e.g., in US 2012/0096572, were

further modified using targeting vectors as described in **FIG. 16** to engineer the heavy chain locus such that the IgG1 gene segment lacks a functional C_H1 domain and the IgG2a and IgG2b genes are deleted to obtain mice homozygous for a single domain antigen binding protein comprising a human kappa variable domain and a murine IgG1 constant region, wherein the IgG1 constant domain lacks a functional C_H1 region (hVκIgG1ΔC_H1ΔIgG2a ΔIgG2b; 6082 HO). Additional variations of combinations of CH1 deletions and/or immunoglobulin constant gene deletions are made, e.g., a mouse is made that comprises a heavy chain locus comprising a human light chain kappa variable region wherein both IgG1 and IgG2a comprise CH1 domain deletions, and the mouse also comprises a deletion of IgD, IgE, IgG3, and IgG2b.

[00188] Western blotting has confirmed that light chain only single domain binding proteins are present and may be isolated from mice genetically modified to comprise a human kappa variable domain and a murine IgG1 constant region, wherein the IgG1 constant domain lacks a functional C_H1 region (6082 HO, data not shown).

Example 2.2: Confirmation of productive rearrangement of gene sequences encoding VL single domain binding proteins

[00189] The mRNA of B cells was isolated from the spleen and bone marrow of (a) mice homozygous for a heavy chain locus comprising a light chain variable region and a heavy chain constant gene sequence that lacks a functional C_H1 domain in the IgG1 gene, and further lacks the IgG2b and IgG2a genes, (b) control wild type and (c) control C_H1 del x ULC mice homozygous for both a modified mouse heavy chain locus that expresses human heavy chain V, D and J segments, lacks a functional C_H1 domain in the IgG1 genes, and also lacks functional IgG2b and IgG2a genes, and comprises a single rearranged light chain locus also referred to as a common or universal light chain, see, e.g., U.S. Patent Publication No. 2011/0195454. The isolated mRNA was analyzed for productive rearrangement using the following probes and primers in a TAQMAN assay:

hJk/mlgG1 hinge-set 71 (ordered from Biosearch Technologies)

(sense) 5'- GGACCAAGCTGGAGATCAAAC -3' (SEQ ID NO:1),

(anti-sense) 5'- CTTCTGGGACTGTACATATGCAA -3' (SEQ ID NO:2),

(probe) 5'-FAM- CCCAGGGATTGTGGTTGTAAGCC –BHQ1-3' (SEQ ID NO:3);

hJH/mlgG1 hinge-set 72 (ordered from Applied Biosystems)

(sense) 5'- TGGTCACCGTCTCCTCAGTG -3' (SEQ ID NO:4),

(anti-sense) 5'- CACACGTGACCTTAGGAGTCAGAG -3' (SEQ ID NO:5),

(probe) 5'-FAM- TGGTTGTAAGCCTTGC -MGB-3' (SEQ ID NO:6);

mHPRT1-set 51 (ordered from Biosearch Technologies)

(sense) 5'- CGAGTCTGAAGCTCTCGATTTCCT -3' (SEQ ID NO:7),

(anti-sense) 5'- CAGCCAACACTGCTGAAACATG -3' (SEQ ID NO:8),

(probe) 5'-FAM- CAGCATCTAAGAGGTTTTGCTCAGTGGA -BHQ-3 (SEQ ID NO:9)';

[00190] As shown in **FIGs. 17A** and **17B**, unrearranged light chain variable region gene segments that replace endogenous heavy chain variable region gene segments are capable of undergoing productive rearrangement with the endogenous heavy chain constant IgG1 gene lacking a functional C_H1 domain.

Example 3: Mice encoding a VL single domain binding protein: Mice comprising an immunoglobulin chain having a light chain variable region and a heavy chain constant region lacking a functional C_H1 domain, and comprising a single rearranged light chain (ULC)

VELOCIMMUNE® humanized mice containing a single rearranged human germline light chain region (ULC V_κ3-20J_κ1; 1635, alternatively ULC V_κ1-39J_κ5; 1633 is also used) were bred to mice carrying a modified heavy chain locus comprising a human light chain kappa variable region operably linked to a murine constant region wherein the IgG1 C_H1 domain, and the IgG2a and IgG2b genes, were deleted or inactivated (hV_κlgG1ΔC_H1ΔIgG2a/2b; 6082) to obtain the following progeny mice: hV_κlgG1ΔC_H1ΔIgG2a/2b x V_κ3-20J_κ1 ULC homozygous mice (6082HO 1635 HO). These mice expressed V_L single domain binding proteins (**FIG. 18**).

We claim:

1. A genetically modified non-human animal comprising in its germline
 - (a) a deletion or inactivating mutation in a nucleotide sequence encoding a C_H1 domain of at least one endogenous immunoglobulin heavy chain constant region gene at an endogenous immunoglobulin heavy chain locus, wherein the at least one endogenous immunoglobulin heavy chain constant region gene is IgG, IgA, IgE, IgD, or a combination thereof, and
 - (b) either or both
 - (i) a nucleic acid sequence comprising at least one unrearranged immunoglobulin light chain variable region (V_L) gene segment and at least one unrearranged immunoglobulin light chain joining (J_L) gene segment, wherein the unrearranged V_L and J_L gene segments are capable of recombining to form a rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence operably linked to the immunoglobulin heavy chain constant region gene comprising the deletion or inactivating mutation in the nucleotide sequence encoding the C_H1 domain, and/or
 - (ii) an immunoglobulin light chain locus that comprises a single rearranged immunoglobulin light chain variable region V_L/J_L gene sequence comprising V_L and J_L gene segment sequences, wherein the single rearranged immunoglobulin light chain variable region gene sequence is operably linked to an immunoglobulin light chain constant region gene sequence.
2. The genetically modified non-human animal of claim 1, wherein the at least one unrearranged V_L gene segment, the at least one unrearranged J_L gene segment, and/or the single rearranged immunoglobulin light chain variable region V_L/J_L gene sequence is human.
3. The genetically modified non-human animal of claim 2, wherein the immunoglobulin light chain constant region and/or the immunoglobulin heavy chain constant region are non-human.
4. The genetically modified non-human animal of claim 2, wherein the at least one unrearranged V_L gene segment and the at least one unrearranged J_L gene segment is a human kappa segment.

5. The genetically modified non-human animal of claim 1, wherein the at least one unrearranged V_L gene segment and the at least one unrearranged J_L gene segment replaces one or more endogenous V_H , D_H , J_H gene segments at the endogenous immunoglobulin heavy chain.
6. The genetically modified non-human animal of claim 5, wherein the at least one unrearranged V_L gene segment and the at least one unrearranged J_L gene segment replace all or substantially all endogenous non-human immunoglobulin heavy chain variable region gene segments at the endogenous non-human immunoglobulin heavy chain locus.
7. The genetically modified non-human animal of claim 1, wherein all or substantially all endogenous heavy chain variable region gene segments and/or all or substantially all endogenous light chain variable region gene segments are deleted or functionally inactivated.
8. The genetically modified non-human animal of claim 2, wherein the single rearranged immunoglobulin light chain variable region V_L/J_L gene sequence is a human $V\kappa 1$ -39/J gene sequence, or a human $V\kappa 3$ -20/J gene sequence.
9. The genetically modified non-human animal of claim 8, wherein the human $V\kappa 1$ -39/J gene sequence comprises a human $V\kappa 1$ -39 gene segment rearranged with a human $J\kappa 5$ gene segment.
10. The genetically modified non-human animal of claim 8, wherein a human $V\kappa 3$ -20/J gene sequence comprises a human $V\kappa 3$ -20 gene segment rearranged with a human $J\kappa 1$ gene segment.
11. The genetically modified non-human animal of claim 1, wherein the animal does not comprise (i) a nucleic acid sequence comprising at least one unrearranged immunoglobulin light chain variable region (V_L) gene segment and at least one unrearranged immunoglobulin light chain joining (J_L) gene segment, wherein the unrearranged V_L and J_L gene segments are capable of recombining to form a rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence operably linked

to the immunoglobulin heavy chain constant region gene comprising the deletion or inactivating mutation in the nucleotide sequence encoding the C_H1 domain, and

wherein the heavy chain locus comprises non-human heavy chain variable region gene segments.

12. The genetically modified non-human animal of claim 1, wherein the animal does not comprise (i) a nucleic acid sequence comprising at least one unrearranged immunoglobulin light chain variable region (V_L) gene segment and at least one unrearranged immunoglobulin light chain joining (J_L) gene segment, wherein the unrearranged V_L and J_L gene segments are capable of recombining to form a rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence operably linked to the immunoglobulin heavy chain constant region gene comprising the deletion or inactivating mutation in the nucleotide sequence encoding the C_H1 domain, and wherein the heavy chain locus comprises human heavy chain variable region gene segments.

13. The genetically modified non-human animal of claim 12, wherein the human heavy chain variable region gene segments replace all or substantially all endogenous immunoglobulin heavy chain variable region gene segments.

14. The genetically modified non-human animal of claim 1, wherein the immunoglobulin light chain constant region gene sequence is a non-human immunoglobulin light chain constant region gene sequence.

15. The genetically modified non-human animal of claim 1, wherein the immunoglobulin light chain constant region gene sequence is a human immunoglobulin light chain constant region gene sequence.

16. The genetically modified non-human animal of claim 1, wherein the single rearranged light chain V_L/J_L gene sequence replaces all or substantially all endogenous immunoglobulin light chain variable region gene segments of the non-human animal.

17. The genetically modified non-human animal of claim 2, wherein the single rearranged human immunoglobulin light chain variable region V_L/J_L gene sequence comprises human germline V_L and human germline J_L gene segment sequences.
18. The genetically modified non-human animal of claim 1, wherein the endogenous immunoglobulin heavy chain locus further comprises an inactivated hinge region.
19. The genetically modified non-human animal of claim 1, wherein the deletion or inactivating mutation in a nucleotide sequence encoding a C_H1 domain is in an IgG1 sequence and wherein the endogenous immunoglobulin heavy chain locus further comprises a deletion or inactivating mutation of an immunoglobulin gene selected from the group consisting of IgD, IgG3, IgG2a, IgG2b, IgG2c, IgE, IgA, and a combination thereof.
20. The genetically modified non-human animal of claim 19, wherein the immunoglobulin heavy chain locus further comprises a deletion or inactivating mutation in the IgG2a and IgG2b immunoglobulin genes.
21. The genetically modified non-human animal of claim 19, wherein the immunoglobulin heavy chain locus further comprises a deletion or inactivating mutation in the IgG2b and IgG2c immunoglobulin genes.
22. The genetically modified non-human animal of claim 19, wherein the immunoglobulin heavy chain locus further comprises a deletion or inactivating mutation in the IgG3, IgD, IgA, and IgE immunoglobulin genes.
23. The genetically modified non-human animal of claim 1, wherein the non-human animal further comprises an Adam6 gene or portion thereof functional in a male non-human animal, wherein the Adam6 gene is an Adam6a gene, an Adam6b gene, or both.
24. The genetically modified non-human animal of claim 1, wherein the animal further comprises in its serum an antigen-specific single domain antigen binding protein that lacks a functional C_H1 domain.

25. The genetically modified non-human animal of claim 24, wherein the animal further comprises an IgM heavy chain comprising a functional C_H1 domain.
26. The genetically modified non-human animal of claim 24, wherein the single domain antigen binding protein is a V_L-single domain binding protein encoded by the rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence operably linked to the immunoglobulin heavy chain constant region gene comprising the deletion or inactivating mutation in the nucleotide sequence encoding the C_H1 domain.
27. The genetically modified non-human animal of claim 25, further comprising a genetically engineered universal light chain encoded by the single rearranged immunoglobulin light chain variable region gene sequence comprising V_L and J_L gene segment sequences operably linked to an immunoglobulin light chain constant region gene sequence.
28. The genetically modified non-human animal of claim 25, wherein the single domain antigen binding protein is a V_L-single domain antigen binding protein encoded by the rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence operably linked to the immunoglobulin heavy chain constant region gene comprising the deletion or inactivating mutation in the nucleotide sequence encoding the C_H1 domain and wherein the non-human animal further comprises a genetically engineered universal light chain encoded by the single rearranged immunoglobulin light chain variable region gene sequence comprising V_L and J_L gene segment sequences operably linked to an immunoglobulin light chain constant region gene sequence.
29. The genetically modified non-human animal of claim 24, wherein at least one heavy chain of the single domain antigen binding protein further lacks a functional hinge region.
30. The genetically modified non-human animal of claim 24, wherein the single domain antigen binding protein has an IgG isotype selected from the group consisting of IgG1, IgG2a, IgG2b, IgG2c and IgG3.

31. The genetically modified non-human animal of claim 24, wherein the single domain antigen binding protein comprises a human variable domain and a non-human constant domain.
32. The genetically modified non-human animal of claim 24, wherein the single domain antigen binding protein is monomeric.
33. The genetically modified non-human animal of claim 25, wherein IgM heavy chain is associated with a cognate light chain.
34. The genetically modified non-human animal of claim 33, wherein the cognate light chain is a universal light chain encoded by the single rearranged immunoglobulin light chain variable region gene sequence comprising V_L and J_L gene segment sequences operably linked to an immunoglobulin light chain constant region gene sequence.
35. The genetically modified non-human animal of claim 24, further comprising a high titer of the single domain antigen binding protein.
36. The genetically modified non-human animal of claim 35, wherein the titer of the antigen specific single domain antigen binding protein is at least 1×10^2 $\mu\text{g/mL}$, at least 1×10^3 $\mu\text{g/mL}$, at least 1×10^4 $\mu\text{g/mL}$, or at least 1×10^5 $\mu\text{g/mL}$.
37. The genetically modified non-human animal of claim 35, wherein the titer of the single domain antigen binding protein is at least 5-fold more, at least 10-fold more, at least 100-fold more than a corresponding control animal that does not express a genetically engineered universal light chain encoded by the single rearranged immunoglobulin light chain variable region gene sequence comprising V_L and J_L gene segment sequences operably linked to an immunoglobulin light chain constant region gene sequence.
38. The genetically modified non-human animal of claim 35, wherein the titer is determined by enzyme linked immunosorbent assay.

39. The genetically modified animal of claim 1, wherein the animal is a rodent.
40. The genetically modified animal of claim 39, wherein the rodent is a rat or a mouse.
41. The genetically modified non-human animal of claim 40, wherein the animal is a mouse.
42. A genetically modified mouse comprising
- (a) a replacement at a mouse heavy chain locus of all or substantially all endogenous immunoglobulin heavy chain V, D, and J gene segments with either
 - (i) one or more unrearranged human immunoglobulin heavy chain V_H gene segments, one or more unrearranged human immunoglobulin heavy chain D_H gene segments, and one or more unrearranged human immunoglobulin heavy chain J_H gene segments, wherein the one or more human unrearranged immunoglobulin heavy chain V_H, D_H, and J_H gene segments are operably linked to a mouse heavy chain constant region gene sequence, or
 - (ii) one or more unrearranged human light chain V_L gene segments and one or more human unrearranged light chain J_L gene segments, wherein the one or more unrearranged human light chain V_L, and J_L gene segments are operably linked to mouse heavy chain constant region gene sequence,wherein the mouse heavy chain constant region gene sequence comprises a full-length IgM gene and a deletion or an inactivating mutation in a nucleotide sequence encoding a C_H1 domain in an IgG gene selected from the group consisting of an IgG1, IgG2a, IgG2b, IgG2c, IgG3, and a combination thereof,
 - (b) a replacement of all or substantially all endogenous immunoglobulin light chain V and J gene segments with a single rearranged human variable V_κ/J_κ gene sequence, and wherein the mouse expresses a B cell receptor that comprises an IgM heavy chain associated with a cognate light chain.
43. A method of making a genetically modified non-human animal according to claim 1 comprising

(a) modifying at least one non-human heavy chain constant region at an endogenous immunoglobulin heavy chain locus of the non-human animal such that the heavy chain constant region comprises a deletion or an inactivating mutation of a nucleotide sequence encoding a C_H1 domain of IgG, IgA, IgE, IgD, or a combination thereof, and

(b) either or both

(i) inserting a nucleic acid sequence into the endogenous immunoglobulin heavy chain locus, wherein the nucleic acid sequence comprises at least one unrearranged immunoglobulin light chain variable region (V_L) gene segment and at least one unrearranged immunoglobulin light chain joining (J_L) gene segment, wherein the unrearranged V_L and J_L gene segments are capable of recombining to form a rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence operably linked to the immunoglobulin heavy chain constant region gene comprising the deletion or inactivating mutation in the nucleotide sequence encoding the C_H1 domain, and/or

(ii) introducing an immunoglobulin light chain locus that comprises a single rearranged immunoglobulin light chain variable region V_L/J_L gene sequence comprising V_L and J_L gene segment sequences, wherein the single rearranged immunoglobulin light chain variable region gene sequence is operably linked to an immunoglobulin light chain constant region gene sequence.

44. The method of claim 43, wherein step (a) further comprises deleting or inactivating a hinge region of the IgG, IgA, IgE, IgD, or combination thereof, comprising the modified C_H1 domain.

45. The method of claim 43, wherein step (a) further comprises replacing one or more endogenous immunoglobulin heavy chain variable region gene segments of the heavy chain immunoglobulin locus with human heavy chain variable region gene segments such that expression of the heavy chain immunoglobulin locus results in a heavy chain variable domain that comprises human idiotypes, and wherein the method does not comprise (b)(i) inserting a nucleic acid sequence into the endogenous immunoglobulin heavy chain locus, wherein the nucleic acid sequence comprises at least one unrearranged immunoglobulin light chain variable region (V_L) gene segment and at least one unrearranged immunoglobulin light chain joining (J_L) gene segment, wherein the unrearranged V_L and J_L gene segments are capable of recombining to form

a rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence operably linked to the immunoglobulin heavy chain constant region gene comprising the deletion or inactivating mutation in the nucleotide sequence encoding the C_H1 domain.

46. The method of claim 43, wherein the step of inserting comprises replacing one or more endogenous immunoglobulin heavy chain variable region gene segments of the heavy chain immunoglobulin locus with the unrearranged light chain V_L and J_L gene segments such that expression of the heavy chain immunoglobulin locus results in a V_L —single domain binding protein comprising an immunoglobulin light chain variable domain and a immunoglobulin heavy chain constant domain that lacks a C_H1 domain.

47. The method of claim 46, wherein the immunoglobulin constant region gene is non-human.

48. The method of claim 47, wherein the unrearranged light chain V_L and J_L gene segments are human segments.

49. The method of claim 48, wherein the unrearranged light chain V_L and J_L gene segments are human kappa segments.

50. The method of claim 48, wherein the unrearranged light chain V_L and J_L gene segments are human lambda segments.

51. The method of claim 46, wherein the unrearranged light chain V_L and J_L gene segments replace all or substantially all endogenous non-human immunoglobulin heavy chain variable region gene segments at the endogenous non-human immunoglobulin heavy chain locus.

52. The method of claim 43, wherein the deletion or inactivating mutation of a nucleotide sequence encoding a C_H1 domain is in an IgG gene.

53. The method of claims 52, wherein the deletion or inactivating mutation of a nucleotide sequence encoding a C_H1 domain is in an IgG1 gene.

54. The method of claim 43, wherein the heavy chain immunoglobulin locus comprises variable region gene segments encoding a human variable domain and a constant region gene encoding a non-human constant domain.
55. The method of claim 43, wherein the single rearranged immunoglobulin light chain variable region gene sequence operably linked to an immunoglobulin light chain constant region gene sequence encodes a universal light chain.
56. The method of claim 55, wherein the single rearranged immunoglobulin light chain variable region V_L/J_L gene sequence is a human V_L/J_L gene sequence.
57. The method of claim 56, wherein the human V_L/J_L gene sequence is a human $V_{\kappa 1-39}/J$ gene sequence, or a human $V_{\kappa 3-20}/J$ gene sequence.
58. The method of claim 57, wherein the human $V_{\kappa 1-39}/J$ gene sequence comprises a human $V_{\kappa 1-39}$ gene segment rearranged with a human $J_{\kappa 5}$ gene segment.
59. The method of claim 57, wherein the human $V_{\kappa 3-20}/J$ gene sequence comprises a human $V_{\kappa 3-20}$ gene segment rearranged with a human $J_{\kappa 1}$ gene segment.
60. The method of claim 43, wherein the single rearranged immunoglobulin light chain variable region gene sequence is operably linked to a non-human immunoglobulin light chain constant region gene sequence.
61. The method of claim 43, wherein the single rearranged immunoglobulin light chain variable region gene sequence is operably linked to a human immunoglobulin light chain constant region gene sequence.
62. The method of claim 43, wherein the single rearranged immunoglobulin light chain variable region replaces all or substantially all endogenous immunoglobulin light chain V_L and J_L gene segments of the non-human animal.

63. The method of claim 43, wherein the nucleic acid sequence and/or single rearranged immunoglobulin light chain variable region sequence is in the germline of the non-human animal.

64. The method of claim 43, wherein the C_H1 deletion or inactivating mutation is in IgG1 and the method further comprises (c) deleting or inactivating an immunoglobulin gene selected from the group consisting of IgD, IgG3, IgG2a, IgG2b, IgG2c, IgE, IgA, and a combination thereof.

65. The method of claim 64, wherein the IgG2a and IgG2b immunoglobulin genes are deleted or inactivated.

66. The method of claim 64, wherein the IgG2b and IgG2c immunoglobulin genes are deleted or inactivated.

67. The method of claim 64, wherein IgG3, IgD, IgA, and IgE immunoglobulin genes are deleted or inactivated.

68. The method of claim 43, wherein the non-human animal is a rodent.

69. The method of claim 68, wherein the rodent is a rat or a mouse.

70. The method of claim 69, wherein the rodent is a mouse.

71. A method of producing antigen-specific IgG single domain binding proteins lacking a C_H1 domain, in whole or in part, the method comprising the step(s) of:

(a) immunizing a genetically modified non-human animal with the antigen, wherein the genetically modified non-human animal comprises

(i) a deletion or inactivating mutation in a nucleotide sequence encoding a C_H1 domain of at least one endogenous immunoglobulin heavy chain constant region gene at an endogenous immunoglobulin heavy chain locus, wherein the at least one endogenous immunoglobulin heavy chain constant region gene is IgG, IgA, IgE, IgD, or a combination thereof, and

(ii) either or both

1. a nucleic acid sequence comprising at least one unrearranged immunoglobulin light chain variable region (V_L) gene segment and at least one unrearranged immunoglobulin light chain joining (J_L) gene segment, wherein the unrearranged V_L and J_L gene segments are capable of recombining to form a rearranged immunoglobulin light chain variable region (V_L/J_L) nucleotide sequence operably linked to the immunoglobulin heavy chain constant region gene comprising the deletion or inactivating mutation in the nucleotide sequence encoding the C_H1 domain, and/or

2. an immunoglobulin light chain locus that comprises a single rearranged immunoglobulin light chain variable region V_L/J_L gene sequence comprising V_L and J_L gene segment sequences, wherein the single rearranged immunoglobulin light chain variable region gene sequence is operably linked to an immunoglobulin light chain constant region gene sequence, and

(b) maintaining the mouse so that the mouse expresses:

(i) an IgM antibody comprising two IgM heavy chains associated with two cognate light chains; and

(ii) a single domain antigen binding protein comprising at least one heavy chain that lacks a functional C_H1 domain.

72. The method of claim 71, wherein the at least one heavy chain constant region gene is selected from the group consisting of IgG1, IgG2b, IgG2a, and a combination thereof.

73. The method of claim 72, wherein the heavy chain constant region gene is an IgG1 constant region gene.

74. The method of claim 71, wherein the single domain antigen binding protein lacks the C_H1 domain in whole.

75. The method of claim 71, wherein the single domain antigen binding protein comprises a human heavy chain variable region.

76. The method of claim 71, wherein the single domain antigen binding protein comprises a human light chain variable region.

77. The method of claim 71, wherein the cognate light chain is a universal light chain encoded by the single rearranged immunoglobulin light chain variable region sequence operably linked to an immunoglobulin light chain constant region gene sequence.

78. The method of claim 77, wherein the non-human animal does not express endogenous light chains.

79. The method of claim 77, wherein the single rearranged immunoglobulin light chain variable region sequence is a human V κ 1-39/J gene sequence, or a human V κ 3-20/J gene sequence.

80. The method of claim 79, wherein the human V κ 1-39/J gene sequence comprises a human V κ 1-39 gene segment rearranged with a human J κ 5 gene segment.

81. The method of claim 79, wherein the human V κ 3-20/J gene sequence comprises a human V κ 3-20 gene segment rearranged with a human J κ 1 gene segment.

82. The method of claim 71, further comprising
(c) isolating from the non-human animal a cell or protein that specifically binds the antigen, wherein the cell or protein comprises a somatically mutated single domain antigen binding protein.

83. The method of claim 82, further comprising
(d) isolating a first nucleic acid encoding the variable domain of the single domain antigen binding protein from the cell isolated in (c).

84. The method of claim 83, further comprising as last steps
(e) culturing a cell transfected with a vector in conditions sufficient for expression of the vector, wherein the vector comprises a second nucleic acid operably linked to a human heavy chain constant region gene, wherein the second nucleic acid is identical to or substantially identical to the first nucleic acid isolated in the step (d).

85. The method of claim 84, wherein the heavy chain constant region gene is a human IgG constant region gene is selected from the group consisting of IgG1, IgG2a, IgG2b, and IgG3.
86. The method of claim 82, further comprising the steps of
(d) collecting supernatant from a hybridoma culture, wherein the hybridoma is produced from the cell isolated in (c).
87. The method of claim 86, further comprising
(e) isolating a first nucleic acid encoding the variable domain of the single domain antigen binding protein from the hybridoma produced in (d).
88. The method of claim 87, further comprising
(f) culturing a cell transfected with a vector in conditions sufficient for expression of the vector; wherein the vector comprises a second nucleic acid operably linked to human heavy chain constant region gene, wherein the second nucleic acid is identical to or substantially identical to the first nucleic acid isolated in the step (e).
89. The method of claim 88, wherein the heavy chain constant region gene is a human IgG constant region gene selected from the group consisting of IgG1, IgG2a, IgG2b, and IgG3.
90. A hybridoma produced from the cell isolated according to the method of claim 82.
91. A nucleic acid isolated according to the method of claim 83.
92. A nucleic acid isolated according to the method of claim 87.
93. A cell comprising a nucleic acid identical to or substantially similar to the nucleic acid of claim 83.
94. A cell comprising a nucleic acid identical to or substantially similar to the nucleic acid of claim 87.

95. An isolated cell of the non-human animal of claim 1.
96. The isolated cell of claim 95, wherein the cell is an embryonic stem (ES) cell or a B cell.
97. A single domain antigen binding protein comprising a light chain variable domain operably linked to a heavy chain constant region that lacks a functional C_H1 domain.
98. A single domain antigen binding protein isolated from the non-human animal of claim 1.
99. A single domain antigen binding protein isolated from the mouse of claim 42.
100. A single domain antigen binding protein produced according to the method of claim 71.