SUPER HUMANIZED ANTIBODIES

CROSS-REFERENCE TO RELATED APPLICATION

This application claim priority to U.S. provisional application number 60/305,111 filed July 12, 2001.

STATEMENT OF GOVERNMENT INTEREST

Development of this invention was supported, in part, by a grant from the U.S., National Institute of Health, grant number CA-18029.

TECHNICAL FIELD

The invention relates to methods of humanizing antibodies, particularly to humanizing antibodies by making chimeric antibodies containing CDR sequences from a non-human antibody and framework sequences of human antibodies, more particularly to methods of selecting appropriate human antibody framework sequences for performing the humanization, and still more particularly to using canonical CDR structure types of the non-human antibody in comparison to germline canonical CDR structure types of human antibodies as the basis for selecting the appropriate human framework sequences for a humanized antibody.

BACKGROUND OF THE INVENTION

Antibodies are natural proteins that the vertebrate immune system forms in response to foreign substances (antigens), primarily for defense against infection. For over a century, antibodies have been induced in animals under artificial conditions and harvested for use in therapy or diagnosis of disease conditions, or for biological research. Each individual antibody producing cell produces a single type of antibody with a chemically defined composition, however, antibodies obtained directly from animal serum in response

to antigen inoculation actually comprise an ensemble of non-identical molecules (i.e, polyclonal antibodies) made from an ensemble of individual antibody producing cells.

Hybridoma technology provided a method to propagate a single antibody-producing cell for an indefinite number of generations with a screening method to identify clones of cells producing an antibody that would react with a particular antigen. Development of this technology allowed production in unlimited quantities of structurally identical antibodies with essentially any desired antigenic specificity. Such antibodies are commonly called monoclonal antibodies, and most originate from rodents. Sequencing of monoclonal antibody genes allowed the primary amino acid structure of the antibody be defined.

The advent of recombinant DNA methodology enabled structural engineering of antibody genes and production of modified antibody molecules with properties not obtainable by hybridoma technology. In the therapeutic arena, one aim of this methodology has been to reduce the immunogenicity in humans of rodent monoclonal antibodies by modifying their primary amino acid structure. Reduction of the immunogenicity of therapeutic antibodies is desirable because induction of an immune response can cause a spectrum of adverse effects in a patient, ranging from accelerated elimination of the therapeutic antibody, with consequent loss of efficacy, to fatal anaphylaxis at the most extreme.

One strategy to reduce immunogenicity of foreign monoclonal antibodies has been to replace the light and heavy chain constant domains of the monoclonal antibody with analogous domains of human origin leaving the variable region domains of the foreign antibody intact. The variable region domains of the light and heavy chains are responsible for the interaction between the antibody and the antigen. The joining domains connecting variable domains to constant domains are situated in a region remote from the site of antigen binding, therefore, the joining domains between variable and constant domains generally do not interfere with antigen binding. Chimeric antibody molecules having mouse variable domains joined to human constant domains usually bind antigen with the same affinity constant as the mouse antibody from which the chimeric was derived. Such

chimeric antibodies are less immunogenic in humans than their fully murine counterparts. Nevertheless, antibodies that preserve entire murine variable domains tend to provoke immune responses in a substantial fraction of patients. For example, INFLIXIMAB™, a widely prescribed chimeric antibody that is considered safe, induced a human anti-chimeric antibody response in 7 out of 47 Crohns Disease patients. (Rutgeerts, P., et al (1999) *Efficacy and safety of retreatment with anti-tumor necrosis factor antibody (INFLIXIMAB) to maintain remission in Crohn's disease.* Gastroenterology 117, 761-769).

That humans would mount an immune response to whole murine variable domains was predictable, thus, efforts to obtain variable domains with more human character had begun even before clinical trials of such standard chimeric antibodies had been reported. One category of methods frequently referred to as "humanizing," aims to convert the variable domains of murine monoclonal antibodies to a more human form by recombinantly constructing an antibody variable domain having both mouse and human character. Humanizing strategies are based on several consensual understandings of antibody structure data. First, variable domains contain contiguous tracts of peptide sequence that are conserved within a species, but which differ between evolutionarily remote species, such as mice and humans. Second, other contiguous tracts are not conserved within a species, but even differ even between antibody producing cells within the same individual. Third, contacts between antibody and antigen occur principally through the non-conserved regions of the variable domain. Fourth, the molecular architecture of antibody variable domains is sufficiently similar across species that correspondent amino acid residue positions between species may be identified based on position alone, without experimental data.

Humanized strategies share the premise that replacement of amino acid residues that are characteristic of murine sequences with residues found in the correspondent positions of human antibodies will reduce the immunogenicity in humans of the resulting antibody. However, replacement of sequences between species usually results in reduction of antibody binding to its antigen. The art of humanization therefore lies in balancing replacement of the original murine sequence to reduce immunogenicity with the

need for the humanized molecule to retain sufficient antigen binding to be therapeutically useful. This balance has been struck using two approaches.

In one approach, exemplified by US Pat. No. 5,869,619 to Studnicka and by *Padlan (1991) A possible procedure for reducing the immunogenicity of antibody variable domains while preserving their ligand binding properties,* Molecular Immunology 28:489-498, characteristically human residues are substituted for murine variable domain residues that are determined or predicted (i) to play no significant chemical role in the interaction with antigen, and (ii) to be positioned with side chains projecting into the solvent, Thus, exterior residues remote from the antigen binding site are humanized, while interior residues, antigen binding residues, and residues forming the interface between variable domains remain murine. One disadvantage of his approach is that rather extensive experimental data is required to determine whether a residue plays no significant chemical role in antigen binding or will be positioned in the solvent in a particular three dimensional antibody structure.

In another more general approach, exemplified by U.S. Pat. No., 5,225,539 to Winter and by Jones et al (1986) *Replacing the complementarity determining regions in a human antibody with those from a mouse*, Nature 321:522-525, contiguous tracts of murine variable domain peptide sequence considered conserved are replaced with the correspondent tracts from a human antibody. In this more general approach, all variable domain residues are humanized except for the non-conserved regions implicated in antigen binding. To determine appropriate contiguous tracks for replacement, Winter, and Jones et al (1986) utilized a classification of antibody variable domain sequences that had been developed previously by Wu and Kabat (1970).

Wu and Kabat pioneered the alignment of antibody peptide sequences, and their contributions in this regard were several-fold: First. through study of sequence similarities between variable domains, they identified correspondent residues that to a greater or lesser extent were homologous across all antibodies in all vertebrate species, inasmuch as they adopted similar three-dimensional structure, played similar functional roles, interacted similarly with neighboring residues, and existed in similar chemical

environments. Second, they devised a peptide sequence numbering system in which homologous immunoglobulin residues were assigned the same position number. One skilled in the art can unambiguously assign what is now commonly called Kabat numbering, to any variable domain sequence, without reliance on any experimental data beyond the sequence itself. Third, for each Kabat-numbered sequence position, Kabat and Wu calculated variability, by which is meant the finding of few or many possible amino acids when variable domain sequences are aligned. They identified three contiguous regions of high variability embedded within four less variable contiguous regions. Other workers had previously noted variability approximately in these regions (hypervariable regions) and posited that the highly variable regions represented amino acid residues used for antigen binding. Kabat and Wu formally demarcated residues constituting these variable tracts, and designated these "complementarity determining regions" (CDRs), referring to chemical complementarity between antibody and antigen. A role in three-dimensional folding of the variable domain, but not in antigen recognition, was ascribed to the remaining less-variable regions, which are now termed "framework regions". Fourth, Kabat and Wu established a public database of antibody peptide and nucleic acid sequences, which continues to be maintained and is well known to those skilled in the art.

The humanization method disclosed by Winter and Jones using the Kabat classification results in a chimeric antibody comprising CDRs from one antibody and framework regions from another antibody that differs in species origin, specificity, subclass, or other characteristics. However, no particular sequences or properties were ascribed to the framework regions, indeed, Winter taught that any set of frameworks could be combined with any set of CDRs. Framework sequences have since been recognized as being important for conferring the three dimensional structure of an antibody variable region necessary retain good antigen binding. Thus, the general humanizing methods described by Winter and Jones have the disadvantage of frequently leading to inactive antibodies because these references do not provide information needed to rationally select among the many possible human framework sequences, those most likely to support antigen binding required by a particular CDR region from a non-human antibody.

Subsequent developments in the field have been refinements within the scope of Winter to deal with loss of avidity for antigen observed with some humanized antibodies relative to the avidity of the corresponding mouse antibodies. (Avidity is a quantitative measure of partitioning of an antibody, in the presence of antigen under conditions approximating chemical equilibrium, between free and antigen-bound forms. For reactions in solution not subject to multivalent binding effects, avidity is the same as affinity, the biochemical equilibrium constant.).

U.S. Pat. No. 5,693,761 to Queen et al, discloses one refinement on Winter for humanizing antibodies, and is based on the premise that ascribes avidity loss to problems in the structural motifs in the humanized framework which, because of steric or other chemical incompatibility, interfere with the folding of the CDRs into the binding-capable conformation found in the mouse antibody. To address this problem, Queen teaches using human framework sequences closely homologous in linear peptide sequence to framework sequences of the mouse antibody to be humanized. Accordingly, the methods of Queen focus on comparing framework sequences between species. Typically, all available human variable domain sequences are compared to a particular mouse sequence and the percentage identity between correspondent framework residues is calculated. The human variable domain with the highest percentage is selected to provide the framework sequences for the humanizing project. Queen also teaches that it is important to retain in the humanized framework, certain amino acid residues from the mouse framework critical for supporting the CDRs in a binding-capable conformation. Potential criticality is assessed from molecular models. Candidate residues for retention are typically those adjacent in linear sequence to a CDR or physically within 6 Å of any CDR residue.

In other approaches, criticality of particular framework amino acid residues is determined experimentally once a low-avidity humanized construct is obtained, by reversion of single residues to the mouse sequence and assaying antigen binding as described by Riechmann et al, (1988). Another example approach for identifying criticality of amino acids in framework sequences is disclosed by U.S. Pat. No. 5,821,337 to Carter et

al, and by U.S. Pat. No. 5,859,205 to Adair et al,. These references disclose specific Kabat residue positions in the framework, which, in a humanized antibody may require substitution with the correspondent mouse amino acid to preserve avidity. One of the disadvantages of the refinements by Queen, and the approaches of Ricechmann, Carter and Adair, is that a very large number of human framework sequences are required for comparison, and/or the guidelines for preserving critical amino acid residues are not completely sufficient to predict functionality. Accordingly, the resulting frameworks constructed, which are part human and part mouse, still frequently exhibit human immunogenicity or lowered antigen binding, thereby requiring numerous iterations in framework construction to obtain a suitable framework for therapeutic uses.

A second type of refinement to Winter is exemplified by Padlan et al (1995) *Identification of specificity-determining residues in antibodies*, FASEB J. 9:133-139; and Tamura et al (2000) *Structural correlates of an anti-carcinoma antibody: identification of specificity-determining residues (SDRs) and development of a minimally immunogenic antibody variant by retention of SDRs only*. J. Immunol. 164:1432-1441. These references share the premise that increasing the proportion of characteristically human sequence in a humanized antibody will reduce that antibody's immunogenicity, and accordingly disclose methods for grafting partial CDR sequences. Determination of the three-dimensional structure of antibody-antigen complexes showed that many residue positions assigned to the CDRs defined by Kabat and Wu rarely were directly involved in antigen binding. These references showed that grafting a subset of CDR residues would adequately transfer antigen binding in a humanized antibody. However, humanized framework sequences are still required, and these references do not teach methods for selecting adequate human framework sequences for use with a given set of mouse CDRs.

There is therefore, a need in the art for methods of humanizing antibodies that reliably identify suitable human framework sequences to support non-human CDR regions and to provide humanized antibodies that retain high antigen binding with low immunogenicity in humans, without the need for direct comparison of framework sequences, without need for determining critically important amino acid residues in the

framework, and without need for multiple iteration n construction to obtain humanized antibodies with suitable therapeutic properties.

.SUMMARY OF THE INVENTION

The present invention meets this need by providing methods for making humanized antibody of high affinity and low immunogenicity without need for comparing framework sequences between non-human and human antibodies and also provides humanized antibodies made thereby. Rather than relying on human framework sequences as the point of analysis, the methods provided herein rely on comparing canonical CDR structure types of the non-human antibody to CDR structure types of human antibodies, particularly as encoded by human germline sequences, to identify candidate human antibody sequences from which to obtain appropriate human frameworks.

More particularly, there is provided a method of making a humanized antibody that includes the acts of obtaining a peptide sequence for a subject variable region encoded by a non-human mature antibody gene and identifying a first set of canonical CDR structure types for at least two CDRs within the non-human antibody variable region. Then A library of peptide sequences for human antibody variable regions for human antibodies is also obtained. In a preferred embodiment, the library contains sequences for human germline variable regions as encoded by germline nucleic acid segments. In other embodiments, however, the library may include mature human antibody sequences. In either case, the method includes identifying canonical CDR structure types (*i.e.*, a second set of canonical CDR structure types) for at least two CDRs for each sequence within the library of human variable region sequences. From this library there is selected a subset of candidate sequences by comparing the first set of canonical CDR structure types to the second set of canonical CDR structure types (*i.e.*, comparing the mouse canonical CDR structure types to the human canonical CDR structure types at corresponding locations within the variable region) and selecting those human sequences where the second set of canonical CDR structure is the same as the first set of canonical CDR structure types for the CDR sequences at corresponding locations within the non-human and human variable

regions, respectively. The method uses these candidate human variable region sequences as a basis for constructing a chimeric molecule that includes at least two of the CDR sequences from the non-human variable region (*e.g.*, of the mouse CDRs) combined with the framework regions from candidate human variable region sequences. The result of the construction is that the chimeric antibody contains each of the non-human CDR sequences substituted for each of the human CDR sequences at corresponding locations in the variable regions so that the framework sequences in the chimeric antibody differs from the candidate human framework sequences by no more than 10 amino acid residues. In certain embodiments, the framework sequences of the chimeric antibody differ from the human framework sequences by no more than 5 amino acid residues. In other embodiments, the framework sequences of the chimeric antibody differs from the human framework sequences by no more than 2 amino acid residues. In most embodiments, the act of constructing the chimeric antibody molecule includes constructing a nucleic acid sequence that encodes the chimeric antibody sequences.

In typical embodiments, the method further includes ranking the members of the subset of candidated human sequences by comparing position by position similarity of amino acid residues of the non-human CDR sequences to the corresponding human CDR sequences according to a ranking criterion. In certain practices, the candidate of human sequences includes only the top 25% of the ranked members. In some embodiments, the ranking criterion includes a score of amino acid identity between the non-human and human CDR sequences at corresponding residue positions of at least one CDR, or at least two CDRs, or most typically each corresponding CDR. In other embodiments, the ranking criterion includes a score of amino acid homology between the non-human and human CDRs. at corresponding residue positions of at least one, at least two, or each of the corresponding CDRs. In still other embodiments , the ranking criterion includes both a score of amino acid identity as well as a score of amino acid homology for at least one, at least two or each of the corresponding CDRs. The method may be practiced using CDRs as defined by differing systems. For example, In certain embodiments, the CDRs are Kabat defined CDRs, in other embodiments, the CDRs are Chothia defined CDR loops.

The method is not limited to strictly using the exact CDR sequences of the non-human source or exact sequences of the human frameworks from the member sets. In certain embodiments, the method may also include substituting at least one amino acid residue of the non-human CDR sequences with a different amino acid, provided however, that no more than 4 residues are substituted in any of non-human light chain CDR1, light chain CDR2, light chain CDR 3, heavy chain CDR1, or heavy chain CDR3 and no more than 10 amino acids are substituted in non-human heavy chain CDR2. In other embodiments, the method may also include substituting at least one but no more than 10 amino acid residues of the human framework sequence with a different amino acid residue.

The method also recognizes that on certain occasions the non-human variable region may include a CDR sequence having a canonical type absent from human variable regions. In cases where each of three non-human CDRs is a light chain CDR, if one of three non-human CDR sequences is of a canonical structure type absent from the library of human variable region sequences, then the act of selecting the human sequences includes selecting a human variable region sequence with a CDR of a different canonical type than the absent non-human CDR type at the corresponding location, providing only that the different canonical human CDR type has a length no more than two amino acid residues smaller or larger than the length of the absent canonical CDR structure type of the non-human CDR. Typically, if the absent CDR sequences is of canonical type 1, then the act includes selecting a human sequence with a canonical type 2 CDR at the corresponding location, or if the non-human CDR sequences is of canonical type 5 then the act includes selecting a human sequence with a canonical type 4 or 3 CDR at the corresponding location.

In most embodiments, the non-human variable region is a mouse variable region. Similarly, in most embodiments the library of human variable region sequences includes a human V_k, V_λ, V_H, J_H, J_k or J_λ sequence as the source of the human frameworks. In most embodiments, the method includes assembling a chimeric antibody having both a chimeric variable light chain and a chimeric variable heavy chain, typically with human frameworks from V_k and V_H sequences. In typical embodiments, the chimeric variable

light chains and chimeric variable heavy chains are formed into an Fab fragment, or a $(Fab)'_2$ molecule, or a single chain Fv molecule, or the chimeric variable light chains and chimeric heavy chains are assembled with a human antibody constant region to form a complete antibody.

The methods are applicable to converting a subject antibody sequence of any subject species to a less immunogenic form suitable for use in an object species by making chimeric antibodies containing framework sequences of the object species in combination with CDRs from the subject species. In such cases, the foregoing methods are the same in the acts performed, where the variable region may be from any subject species and the object variable region may be from any object species for which the antibody will be used. Thus, for example, in various embodiments, a subject antibody may be chimierzied with framework sequences from bovine, porcine, murine or equine sources to form bovinized, porcinized, murinized, or equinized, respectively.

In another aspect, the invention provides compositions that include the chimeric antibody molecules made according to the disclosed methods. Because, the methods utilize a novel way of identifying the appropriate object framework sequence to combine with subject CDR sequences, the resulting chimeric antibodies made are also novel. Accordingly, there is herein provided, a humanized antibody that includes a chimeric antibody variable region containing at least two non-human CDR sequences fused adjacent to human variable region framework sequence. The human framework sequences are selected from a subset of framework sequences characterized by having no more than 10 amino acid residues that differ from a framework sequences in a human antibody variable region having at least two human CDR sequences with the same canonical structure type as the non-human CDR sequences for at least two corresponding CDR positions between the variable region of the chimeric antibody and the human antibody.

The non-human variable region CDRs are typically from a mouse. The human variable region sequence is typically a V_k ,V_λ, V_H, J_H, J_k or J_λ sequence. Most typically the chimeric antibody includes chimeric antibody sequences for each of a variable light chain and a variable heavy chain. In typical embodiments, the chimeric variable light

chains and chimeric variable heavy chains are formed into an Fab fragment. or a $(Fab)'_2$ molecule, or a single chain Fv molecule, or the chimeric variable light chains and chimeric heavy chains are assembled with a human antibody constant region in the form of a complete antibody. Most typically, the human variable region sequence is a sequence from a human germline variable region fragment. In other embodiments, the human variable regions sequence is a sequence from a human mature antibody.

In preferred embodiments, the humanized antibody has a dissociation constant for its antigen of at least 10^6 M^{-1}, preferably at least 10^7 M^{-1} and more preferably at least 10^8 M^{-1}. Typically the humanized antibody of does not elicit an immune response when administered to a human. Particular embodiments exemplifying the invention included humanized antibodies that bind a scorpion venom antigen, that bind a human CD28 receptor, that bind human lysozyme, or that bind a human glutamic acid decarboxylase (GAD65).

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts a library of human germline V_H gene segments.

Figure 2 depicts a library of human germline V_k gene segments.

Figure 3 depicts a portion of mouse D1.3 (anti-chicken lysozyme) antibody variable light chain sequence and a selected subset of human germline V_k variable region sequences having canonical CDRs of the same type as the mouse DL.3 light chain sequence at corresponding locations. The subset is ranked by similarity of amino acid sequences between the DL.3 CDRs and the human CDRs, with the highest ranked sequence depicted first.

Figure 4 depicts a portion of the mouse D1.3 antibody variable heavy chain sequence and a selected subset of human germline V_H variable region sequences having canonical CDRs of the same type as the DL.3. The subset is ranked by similarity of amino acid sequences of the corresponding CDRs analogously to Figure 3.

Figure 5 depicts amino acid sequences for a chimeric V_k variable region and V_H variable region for a humanized D1.3 antibody, illustrating one aspect of the invention.

Figure 6 depicts a nucleic acid sequence for a DNA construct that encodes (and expresses) the humanized chimeric D1.3 antibody of Figure 5, illustrating another aspect of the invention.

Figure 7 is a graph that illustrates antigen binding by the humanized D1.3 antibody, which has an affinity constant of greater than 10^8 M^{-1}, illustrating one embodiment of the invention.

Figure 8 depicts a portion of a mouse variable light chain sequence of an anti-human CD28 antibody designated 9.3 and a selected subset of human germline V_k variable region sequences having canonical CDRs of the same type as the mouse 9.3 variable light chain sequence at corresponding locations, which are ranked by similarity of amino acid sequences analagously to Figure 3.

Figure 9 depicts a portion of the mouse variable heavy chain sequence for the 9.3 antibody and a selected subset of human germline V_H variable region sequences having canonical CDRs of the same type as the mouse variable heavy chain sequence at corresponding locations also ranked by similarity of amino acid sequences.

Figure 10 depicts a humanized anti-human CD28 (Hu.9.3) Fab fragment with chimeric variable heavy and variable light chains, illustrating another embodiment of the invention.

Figure 11 is a graph that illustrates antigen binding by Hu9.3 Fab fragment, which has an affinity constant of greater than 10^6 M^{-1}, illustrating one embodiment of the invention

Figure 12. depicts a humanized anti-scorpion toxin Fab fragment with chimeric variable heavy and variable light chains, illustrating another embodiment of the invention.

Figure 13. depicts a humanized anti-human glutamic acid decarboxylase (GAD65) Fab fragment with chimeric variable heavy and variable light chains, illustrating another embodiment of the invention.

Figure 14 is a graph that illustrates antigen binding by the humanized anti GAD65 Fab fragment, which has an affinity constant of greater than 10^{11} M^{-1}, illustrating one embodiment of the invention

DETAILED DESCRIPTION OF THE INVENTION

In the description that follows, citation is made to various references that may assist one of ordinary skill in the art in understanding and practicing the invention to its fullest extent. Therefore, each reference cited in the description that follows is incorporated herein by reference in its entirety. To better aid in understanding various embodiments of the invention it may be helpful to explain the meanings of certain terms used herein.

A "mature antibody gene" is a genetic sequence encoding an immunoglobulin that is expressed, for example, in a lymphocyte such as a B cell, in a hybridoma or in any antibody producing cell that has undergone a maturation process so that the particular immunoglobulin is expressed. The term includes mature genomic, cDNA or other nucleic acid sequence that encodes such mature genes, which have been isolated and/or recombinantly engineered for expression in other cell types. Mature antibody genes have undergone various mutations and rearrangements that structurally distinguish them from antibody genes encoded in all cells other than lymphocytes. Mature antibody genes in humans, rodents, and many other mammals are formed by fusion of V and J gene segments in the case of antibody light chains and fusion of V, D, and J gene segments in the case of antibody heavy chains. Many mature antibody genes acquire point mutations subsequent to fusion, some of which increase the affinity of the antibody protein for a specific antigen

"Germline antibody genes" or gene fragments are immunoglobulin sequences encoded by non-lymphoid cells that have not undergone the maturation process that leads to genetic rearrangement and mutation for expression of a particular immunoglobulin. One of the advantages provided by various embodiments of the present invention stems from the recognition that germline antibody genes are more likely than

mature antibody genes to conserve essential amino acid sequence structures characteristic of individuals in the animal species, hence less likely to be recognized as from a foreign source when used therapeutically in that species. Figure 1 and Figure 2 show peptide sequences for human germline antibody genes encoding human variable heavy region (V_H) and variable light region (V_k) antibodies (*i.e.*,. immunoglobulins). Each of these list of sequences exemplify a library of human antibody genes, particularly a library of human germline antibody genes.

"A CDR" is the complement determining region within antibody variable sequences. There are three CDRs in each of the variable heavy and variable light sequences designated CDR1, CDR2 and CDR3, for each of the variable regions. The exact boundaries of these CDRs have been defined differently according to different systems, however, all have overlapping residues in what constitute the so called "hypervariable regions" within the variable sequences. The system described by Kabat (CITE) not only provides an unambiguous residue numbering system applicable to any variable region of an antibody, but also provides precise residue boundaries defining the three CDRs. These CDRs may be referred to as Kabat CDRs. Chothia and coworkers (CITE) found that certain sub portions within Kabat CDRs adopt nearly identical peptide backbone conformations, despite having great diversity at the level of amino acid sequence. These sub portions were designated as L1, L2 and L3 or H1, H2 and H3 where the "L" and the "H" designates the light chain and the heavy chains regions, respectively. These regions may be referred to as Chothia CDRs, which have boundaries that overlap with Kabat CDRs. Table I illustrates the overlap of Chotia and Kabat CDRs according to the residue numbering system of Kabat.

TABLE I

Chain	CDR	Kabat	Chothia
Light	CDR1	24 – 34	26 – 32

"	CDR2	50 – 56	50 – 52
"	CDR3	89 – 96	91 – 96
Heavy	CDR1	31 – 35	26 – 32
"	CDR2	50 – 65	52 – 56
"	CDR3	95 – 102	not uniquely defined

Other boundaries defining CDRs overlapping with the Kabat CDRs have been described by Padlan (1995) or MacCallum (1996). Still other CDR boundary definitions may not strictly follow one of the above systems, but will nonetheless overlap with the Kabat CDRs, although they may be shortened or lengthened in light of prediction or experimental findings that particular residues or groups of residues or even entire CDRs do not significantly impact antigen binding. The methods used herein may utilize CDRs defined according to any of these systems, although preferred embodiments use Kabat or Clothia defined CDRs.

"Framework" or "framework sequence" are the remaining sequences of a variable region minus the CDRs. Because the exact definition of a CDR sequence can be determined by different systems, the meaning of a framework sequences is subject to correspondingly different interpretations. To clarify the meaning used herein, a framework sequence means those sequences within the variable region of an antibody other than those defined to be CDR sequences, so that the exact sequence of a framework depends only on how the CDR is defined. For example, the CDRs used in the methods provided herein are usually a subset of what is considered a Kabat CDR, but in the case of CDR1 of heavy chains for example, also includes residues that are classified as framework residues in the Kabat system.

"Canonical CDR structure types" are the structure types designated by Chothia (CITE). Chothia and coworkers found that critical portions of the CDRs of many antibodies adopt nearly identical peptide backbone conformations, despite great diversity at the level of amino acid sequence. Accordingly, Chothia defined for each CDR in each

chain one or a few "canonical structures". Each canonical structure specifies primarily a set of peptide backbone torsion angles for a contiguous segment of amino acid residues forming a loop. The canonical CDR structure types defined by Chothia are listed in Table II.

TABLE II

Chain	CDR	Canonical structure types
Kappa	CDR1	1 – 6
"	CDR2	1
"	CDR3	1 – 6
Heavy	CDR1	1 – 3
"	CDR2	1 – 4
Lambda	CDR1	1 – 4
"	CDR2	1
"	CDR3	1 – 2

"Corresponding CDRs" refer relatively to the CDRs between two different variable sequences that correspond in position within the two different variable sequences. Thus, for example, a mouse light chain CDR1 corresponds to a human light chain CDR1, and vice a versa, because each maps to a defined position in a Kabat numbering system, whether or not the actual boundary of the CDR is defined by Kabat, Chothia or some other system. Similarly, "corresponding" residues, sequences or amino acids refer relatively to the residue positions between two different peptide sequences mapped by the Kabat numbering system.

The objective of the methods provided herein, which may be called CDR grafting, method is to provide a prescription for arriving at appropriate human framework sequence for humanizing a subject non-human antibody. In all previous CDR grafting

methods, the choice of the humanized framework sequence was based on comparing the human framework to the subject (murine) frameworks . In contrast, the basis of the methods herein described are to chose the human antibody to provide the humanized framework based on similarity of its CDRs to those of the subject antibody, without regard to comparing the framework sequences between the two antibodies.

The similarity to the subject CDRs of candidate human antibody sequences is assessed for each domain at two levels. Primarily, identical three-dimensional conformations of CDR peptide backbones are sought. Experimentally determined atomic coordinates of the subject CDRs are seldom available, hence three-dimensional similarity is approximated by determining Chothia canonical structure types of the subject CDRs and excluding from further consideration candidates possessing different canonical structures. Secondarily, residue-to-residue homology between subject CDRs and the remaining human candidate CDRs is considered, and the candidate with the highest homology is chosen.

Choosing highest homology is based on various criterion used to rank candidate human variable regions having the same canonical structure as the subject the non-human variable regions. The criterion for ranking members of the selected set may be by amino acid sequence identity or amino acid homology or both. Amino acid identity is simple a score of position by position matches of amino acid residues. Similarity by amino acid homology is position by position similarity in residue structure of character. Homology may be scored, for example, according to the tables and procedures described by Henikoff and Henikoff, (1992) *Amino acid substitution matrices from protein blocks*, Proc. Natl. Acad. Sci 89: 10915-10919. or by the BLOSUM series described by Henikoff and Henikoff, (1996) .

The steps of the methods are as follow:

Determine the peptide sequences of the heavy and light chain variable domains of the subject antibody. These can be determined by any of several methods, such as DNA sequencing of the respective genes after conventional cDNA cloning; DNA sequencing of cloning products that have been amplified by the polymerase chain reaction

from reverse transcripts or DNA of the subject hybridoma line; or peptide sequencing of a purified antibody protein.

Apply the Kabat numbering system (Kabat et al, 1991) to the heavy and light chain sequences of the subject non-human antibody.

Determine canonical structure types for each of the CDRs of the subject non-human antibody. This determination is made from examination of the peptide sequence in light of the guidelines discussed in Chothia and Lesk (1987), Chothia et al (1992), Tomlinson et al (1995), Martin and Thornton (1996), and Al-Lazikani et al (1997). The salient features of canonical structure determination for each of the CDRs are as follows.

For heavy chain CDR1, three canonical structure types are currently known. Assignment of a new sequence is straightforward because each canonical structure type has a different number of residues. As described in Al-Lazikani *et. al* (1997), when Kabat numbering is assigned to the sequence, the numbering for residues 31 – 35 will be as follows for the respective canonical structures.

Canonical structure type 1: 31, 32, 33, 34, 35.

Canonical structure type 2: 31, 32, 33, 34, 35, 35a.

Canonical structure type 3: 31, 32, 33, 34, 35, 35a, 35b.

For heavy chain CDR2, four canonical structure types are currently known. Several have unique numbers of residues, and are easily distinguished from their unique Kabat numbering of positions 52 – 56, viz.:

Canonical structure type 1: 52, 53, 54, 55, 56.

Canonical structure type 4: 52, 52a, 52b, 52c, 53, 54, 55, 56.

Canonical structure types 2 and 3 for heavy chain CDR2 have equal numbers of residues, hence must be distinguished by clues within their sequence, as discussed by Chothia et al (1992). The Kabat numbering of the segment containing these clues is: 52, 52a, 53, 54, 55. Canonical structure type 2 has Pro or Ser at position 52a and Gly or Ser at position 55, with no restriction at the other positions. Canonical structure type 3 has Gly, Ser, Asn, or Asp at position 54, with no restriction at the other positions. These

criteria are sufficient to resolve the correct assignment in most cases. Additionally framework residue 71 is commonly Ala, Val, Leu, Ile, or Thr for canonical structure type 2 and commonly Arg for canonical structure type 3.

Heavy chain CDR3 is the most diverse of all the CDRs. It is generated by genetic processes, some of a random nature, unique to lymphocytes. Consequently, canonical structures for CDR3 have been difficult to predict. In any case, human germline V gene segments do not encode any part of CDR3; because the V gene segments end at Kabat position 94, whereas positions 95 to 102 encode CDR3. For these reasons, canonical structures of CDR3 are not considered for choosing candidate human sequences.

For light chain CDR1, six canonical structure types are currently known for CDR1 in kappa chains. Each canonical structure type has a different number of residues, hence assignment of a canonical structure type to a new sequence is apparent from the Kabat numbering of residue positions 27 – 31.

Canonical structure type 1: 27, 29, 30, 31.

Canonical structure type 2: 27, 28, 29, 30, 31.

Canonical structure type 3: 27, 27a, 27b, 27c, 27d, 27e, 27f, 28, 29, 30, 31.

Canonical structure type 4: 27, 27a, 27b, 27c, 27d, 27e, 28, 29, 30, 31.

Canonical structure type 5: 27, 27a, 27b, 27c, 27d, 28, 29, 30, 31.

Canonical structure type 6: 27, 27a, 28, 29, 30, 31.

For light chain CDR2, only a single canonical structure type is known for CDR2 in kappa chains, hence, barring exceptional subject antibody sequences, assignment is automatic.

For light chain CDR3, up to six canonical structure types have been described for CDR3 in kappa chains, but three of these are rare. The three common ones can be distinguished by their length, reflected in Kabat numbering of residue positions 91 – 97:

Canonical structure type 1: 91, 92, 93, 94, 95, 96, 97 (also with an obligatory Pro at position 95 and Gln, Asn, or His at position 90).

Canonical structure type 3: 91, 92, 93, 94, 95, 97.

Canonical structure type 5: 91, 92, 93, 94, 95, 96, 96a, 97.

After identifying the canonical CDR structure types of the subject non-human antibody, human genes of the same chain type (heavy or light) that have the same combination of canonical structure types as the subject antibody are identified to form a candidate set of human sequences. In preferred embodiments, only the peptide sequences of human germline immunoglobulin VH and Vk gene fragments are considered for comparison. Most of these gene fragments have been discovered and have already been assigned to a canonical structure type (Chothia et al, 1992, Tomlinson et al,1995). Additional V gene fragments not disclosed by tehse references are provide herein and appear among those sequneces listed in Figure 1 and Figure 2. For the heavy chain, conformity of CDR1 and CDR2 to the mouse canonical structure types is assessed, and genes that do not conform are excluded. For the light chain, conformity of CDR1 and CDR2 of each human sequence to the canonical structure types of the subject antibody is first assessed. The potential of residues 89 – 95 of a candidate Vk gene to form a CDR3 of the same canonical structure type as the subject antibody is assessed, by positing a fusion of the gene with a J region and a applying criteria for CDR3 canonical CDR structure type determination to the fused sequence, and non conforming sequences are excluded.

In another embodiment, appropriate when a variable domain of the subject antibody is of a canonical structure type not available in the human genome, human germline V genes that have three-dimensionally similar, but not identical, canonical structure types are considered for comparison. Such a circumstance often occurs with kappa chain CDR1 in murine antibodies, including two of the examples described below. All 6 possible canonical structure types have been observed at this CDR in murine antibodies, whereas the human genome encodes only canonical types 2, 3, 4 and 6. In these circumstances, a canonical CDR structure type having length of amino acid residues within two of the length of the amino acid residues of the subject non-human sequence may selected for the comparison. For example, where a type 1 canonical structure is found in the subject antibody, human Vk sequences with canonical structure type 2 should be used

for comparison. Where a type 5 canonical structure is found in the murine antibody, human Vk sequences with either canonical structure type 3 or 4 should be used for comparison.

In another embodiment, mature, rearranged human antibody sequences can be considered for the sequence comparison. Such consideration might be be warranted under a variety of circumstances, including but not limited to instances where the mature human sequence (1) is very close to germline; (2) is known not to be immunogenic in humans; or (3) contains a canonical structure type identical to that of the subject antibody, but not found in the human germline.

In preferred embodiments, for each of the candidate V genes with matching canonical structure types, residue to residue sequence identity and/or homology with the subject sequence is also evaluated to rank the candidate human sequences. In a specific embodiment, the residues evaluated are as follows.:

Chain	CDR	Residue positions
Kappa	1	26 – 32
"	2	50 – 52
"	3	91 – 96
Heavy	1	31 – 35
"	2	50 – 60

In preferred embodiments, residue-to-residue homology is first scored by the number of identical amino acid residues between the subject and the candidate human sequences. The human sequence used for subsequent construction of a converted antibody is chosen from among the 25 percent of candidates with the highest score. In other embodiments, appropriate when several candidate sequences have similar identity scores, similarity between non-identical amino acid residues may be additionally be considered. Aliphatic-with-aliphatic, aromatic-with-aromatic, or polar-with-polar matches between

subject and object residues are added to the scores. In another embodiment, quantitative evaluation of sequence homology may be performed using amino acid substitution matrices such as the BLOSUM62 matrix of Henikoff and Henikoff (1992).

Aa object sequence for the framework region C-terminal to CDR3 sequence is selected from the set of known human germline J segments. A preferred J peptide sequence is selected by evaluating residue to residue homology for each J segment for sequence positions for which CDR3 and J overlap, using the scoring criteria specified for the evaluation of candidate V genes as mentioned above. The J gene segment peptide sequence used for subsequent construction of a converted antibody is chosen from among the 25 percent of candidates with the highest score.

In one embodiment, the chimeric variable chain contains at least two CDRs from the subject non-human sequence, and framework sequences from the candidate human sequence. In other embodiments, a chimeric light chain contains three CDRs from the subject non-human sequence and framework sequences from the candidate human sequence. In other embodiments, a chimeric heavy chain contains at least two CDRs of the subject heavy chain, and framework sequence of the candidate human heavy chain. In another embodiment, a chimeric heavy chin contains each of the CDRs from the subject heavy chain and the framework sequences of the candidate human heavy chain. In still another embodiment, a chimeric antibody heavy chain contains CDRs 1 and 2 from the subject non-human sequence and residues 50-60 for CDR3 and residues 61-65 of a CDR from the candidate human heavy chain, along with the framework sequences of the candidate human sequence. In another embodiment, a chimeric heavy chain sequence contains each CDR from the subject non-human sequence, frameworks sequences 27-30 form the subject sequence, and the framework sequences from the candidated sequences. In all cases however, the chimeric antibody molecule contains no more than 10 amino acid residue in the framework sequence that differ from those in the framework sequence of the candidate human variable ration.

In another embodiment, appropriate when increased affinity of a humanized antibody is desired, residues within the CDRs of a converted antibody may be additionally

substituted with other amino acids. Typically, no more than four amino acid residues in a CDR are changed, and most typically no more than two residues in the CDR will be changed, except for heavy chain CDR 2, where as many as 10 residues may be changed. Similarly, in certain embodiments, some of the amino acids in the framework sequences may be changed. In all embodiments, no more than 10 amino acid residues are changed

The humanized antibody sequence is then physically assembled by methods of gene synthesis and recombinant protein expression known by those skilled in the art. The final form of the humanized sequences having the chimeric variable chains made by the methods disclosed herein may take many forms. Most typically, the chimeric antibodies will be made by construction a nucleic acid sequence encoding the chimeric variable chains, which are recombinantly expressed in a suitable cell type. One of most typical forms of the chimeric antibody will be an Fab antibody fragment. Other suitable forms of the chimeric antibody include $(Fab)'_2$ molecule, or a single chain Fv molecule. Still other forms may include further fusion to constant domains of a human antibody to form a complete antibody. In preferred embodiments, both light and heavy variable chains are humanized. However, in other embodiment the variable light and heavy chains may be of mixed, *i.e.*, with one fully mouse variable chain (either heavy or light) and the other being a humanized variable chain.

In most embodiments, the method will include screening candidate chimeric antibodies to select those having a dissociation constant for the antigen suitable for an intended use. In most embodiments the humanized antibody made according to these methods will have a dissociation constant of at least about 10^6 M^{-1}, at least about 10^7 M^{-1} or at least about 10^8 M^{-1} . A Kd of at least about 10^8 M^{-1} is preferred for most therapeutic uses.

The following Examples illustrate the present invention by showing specific embodiments for humanized antibodies that bind different types of antigens for purposes of illustration . It One of ordinary skill in the art will understand that many other specific embodiments may be created using the methods disclosed herein, and that the present invention is not limited by the specific examples.

EXAMPLE 1

HUMANIZED ANTI-CHICKEN LYSOZYME

The mouse antibody D1.3 binds to a chicken lysozyme antigen. The peptide sequence of the variable domains of D1.3 were obtained from the Protein Data Bank, accession number 1VFA. The light chain was numbered according to Kabat, and the mouse CDRs were assigned canonical structure types as follows:

Light chain CDR1, numbered according to Kabat, consists of the sequence:

24 25 26 27 28 29 30 31 32 33 34

R A S G N I H N Y L A

Because there are no insertions or deletions between residues 27 and 31, CDR1 has canonical structure type 2.

Light chain CDR2, numbered according to Kabat, consists of the sequence:

50 51 52 53 54 55 56

Y T T T L A D

This is not an exceptional sequence; its canonical structure type is type 1.

Light chain CDR3, numbered according to Kabat, consists of the sequence

89 90 91 92 93 94 95 96 97

Q H F W S T P R T

Because of the length and the Pro at position 95, this sequence is consistent with canonical structure type 1.

In the compilation in Figure 2 and in Tomlinson et al (1995), 21 non-redundant human germline Vk genes encode (1) CDR1 with canonical structure type 2, (2) CDR2 with canonical structure type 1, and (3) a sequence with the potential to form canonical structure type 1 at CDR3. These are listed in Figure 3 underneath the D1.3 Vk sequence. Their sequence at the residue positions comprising the Chothia canonical structure types is also given, and the human Vk genes in Figure 3 are stratified according to number of residue-to-residue identities in these sequences. L23 has 7 identities, whereas the next three entries on the list have 6. Furthermore, L23 has conserved residues positions

91 and 92, within CDR3, again superior to the next three candidates. L23 therefore is chosen for the humanizing construction.

Among the human Jk segments in Figure 3, none matches the Arg in D1.3 at position 96, and all are identical in the next three positions. Jk4, which replicates the GGG motif in D1.3 positions 99–101, is the best match for J segment, and is used for the humanizing construction.

The heavy chain variable domain of D1.3 was numbered according to Kabat, as shown in Figure 4. CDRs were assigned canonical structure types as follows.

The sequence in the region of heavy chain CDR1 is

27 28 29 30 31 32 33 34 35

F S L T G Y G V N

This sequence lacks any inserted residues, hence is assigned to canonical structure type 1.

The Kabat CDR2 of D1.3 has the sequence

50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65

M I W G D G N T D Y N S A L K S

Because there is no insertion between residues 52 and 56, CDR2 is assigned canonical structure type 1. Human germline VH genes predicted to have canonical structure type 1 at CDR1 and canonical structure type 1 at CDR2 were taken from Chothia et al (1992) and Figure 1, and are listed in Figure 4.

Segments chosen for homology evaluation were 27 – 35, corresponding to Kabat CDR1 plus additional residues forming the Chothia canonical structure, and 50 – 60, corresponding to Kabat CDR2 less residues 61 – 65, which seldom participate directly in antigen binding. The first two entries have have 8 identities in these segments when compared to the mouse sequence, and the next five have 7 identities. The leading 25% of entries in the similarity ranking are thus the two genes with 8 identities and any of those with seven. Though any of these seven genes would be suitable candidates for a humanizing construction, several are preferred because of conservation in non-identical residues. Three that have Glu or Arg replacing Met at residue 50 are excluded because

burial of a charged side chain in the middle of a hydrophobic segment is likely to give an altered three-dimensional structure. V71-4 was thus chosen from the remaining four.

JH4 is clearly the best match to the C terminal end of CDR3.

A chimeric humanized antibody was designed by combining the Kabat CDRs of D1.3 with the Kabat frameworks encoded by V71-4, JH4, L23, and Jk4. The sequences of the heavy and light chain variable domains of this antibody are shown in Figure 5.

Synthetic variable domain genes encoding the humanized Vk and VH were prepared from synthetic oligonucleotides by the method of Ye et al (1992), incorporated herein by reference. These genes were then transferred to the Fab expression vector pAK19, described by Carter et al (1992), incorporated herein by reference. The DNA sequence of the synthetic genes and of the Fab expression cassette of pAK19 are shown in Figure 6. Recombinant Fab was expressed in E. coli, released from the periplasm by osmotic shock, and purified by chromatography on lysozyme-Sepharose.

The affinity of SHuD1.3 for lysozyme was determined by the fluorescence quench method described in by Foote and Winter (1992). This method relies on changes in the intrinsic tryptophan fluorescence of the antibody and antigen upon complex formation. In the experiment in Figure 7, 200 nM humanized D1.3 Fab was titrated with small aliquots of a concentrated lysozyme solution. Fluorescence data were fit by least squares to a titration equation to obtain a value and standard error for the dissociation constant, 23 ± 5 nM. By comparison, the Kd of D1.3 IgG is known to have a Kd of 4 nM (Foote and Winter, 1992). Thus the humanized antibody in example 1 has an identical antigenic specificity as the subject mouse antibody, and binds antigen with an affinity diminished by less than a factor of 6 relative to the subject antibody.

EXAMPLE 2

HUMANIZED ANTI HUMAN CD28

The mouse anti-human CD28 antibody designated 9.3 was used as the non-human subject antibody. The mouse 9.3 hybridoma line was isolated and is described by Hansen et al (1980).

The heavy and light chain variable region genes of 9.3 were cloned by reverse transcription and the polymerase chain reaction , starting with messenger RNA that had been isolated by a guanidinium isothiocyanate procedure (Chomczynski and Sacchi, 1987) followed by chromatography on oligo-dT columns. Amplification was primed using oligonucleotides complementary to the constant region and oligonucleotides corresponding to regions of the signal peptide or N-terminal framework sequence.

The light chain was numbered according to Kabat, and CDRs were assigned canonical structures types as follows, with reference to Figure 8.

Light chain CDR1, numbered according to Kabat, consists of the sequence

24 25 26 27 a b c d 28 29 30 31 32 33 34

R A S E S V E Y Y V T S L M Q

Because of the inserted residues between 27 and 31, CDR1 has canonical structure type 5.

Light chain CDR2, numbered according to Kabat, consists of the sequence

50 51 52 53 54 55 56

A A S N V E S

This is not an exceptional sequence; its canonical structure type is 1.

Light chain CDR3, numbered according to Kabat, consists of the sequence

89 90 91 92 93 94 95 96

Q Q S R K V P Y

Because of the length and the Pro at position 95, this sequence is consistent with canonical structure type 1.

Vk sequences with canonical structure type 5 at CDR1 are not represented in the human germline, but structures 3 and 4 resemble canonical structure type 5, and were considered further.

In the compilation in Figure 2, eight non-redundant human germline Vk genes encode (1) CDR1 with canonical structure type 3 or 4, (2) CDR2 with canonical

structure type 1, and (3) a sequence with the potential to form canonical structure type 1 at CDR3. These are listed in Figure 8 underneath the 9.3 Vk sequence. Their sequence at the Kabat CDR is also given. The human Vk genes in Figure 3 are ranked according to number of residue-to-residue identities in residue positions forming the Chothia canonical structure. The B3 gene has 7 identities in these position, whereas the next three on the list have 5, hence B3 was chosen for the humanizing construction. Had the scoring been based on Kabat CDR positions, rather than Chothia, B3 would still have been the leading candidate. The 5'-encoded Tyr residue of human Jk2 matched the corresponding position of 9.3 exactly, hence this germline fragment was used.

The heavy chain variable domain of 9.3 was numbered according to Kabat, as shown in Figure 9. CDRs were assigned canonical structure types as follows.

The sequence in the region of heavy chain CDR1 is

27 28 29 30 31 32 33 34 35

F S L S D Y G V H

This sequence lacks any inserted residues, hence is assigned to canonical structure type 1.

The Kabat CDR2 of 9.3 has the sequence

50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65

V I W A G G G T N Y N S A L M S

Because there is no insertion between residues 52 and 56, CDR2 is assigned canonical structure type 1.

Human germline VH genes predicted to have canonical structure type 1 at CDR1 and canonical structure type 1 at CDR2 were taken from Chothia et al (1992) and Figure 1, and are listed in Figure 9.

Segments chosen for homology evaluation were 27 – 35, corresponding to Kabat CDR1 plus additional residues forming the Chothia canonical structure, and 50 – 60, corresponding to Kabat CDR2 less residues 61 – 65, which seldom participate directly in antigen binding. Sequences were scored for number of identical residues when compared to 9.3, and are ranked by score in Figure 9. Gene DP-45 has the highest number of

identities, 10, in a residue-to-residue comparison with 9.3; the next 6 entries all have 9. DP-45 was chosen for the humanizing construction.

Of the human JH segments, JH4 had the closest homology to the C-terminal end of CDR3 in 9.3, hence was used in the construction.

Chimeric humanized antibody variable domains were designed by combining sequences as follows. The light chain variable domain consisted of Kabat CDR sequences of the 9.3 light chain, with the exception of Residue 34, which was thought to be not critical to antigen recognition, hence was made Ala, identical to the residue in B3 at that position; and framework sequences identical to B3 through residue 88 and identical to Jk2 from positions 98 – 108, with the exception of residues 70 and 72, which were left identical to 9.3 to preserve a glycosylation motif that these residues form in combination with residue 71. The heavy chain variable domain consisted of Kabat CDR sequences of the 9.3 heavy chain, with the exception of residues 60 - 65, which were thought to be not critical to antigen recognition and hence made identical to the sequence of DP-45 at those positions; and Kabat framework sequences identical to DP-45 through residue 94 and identical to JH4 from residue 103 - 113.

The sequences of the heavy and light chain variable domains of this antibody are shown in Figure 10. A recombinant Fab fragment with variable domains these sequences was prepared as described for Example 1, with the exception of using affinity chromatography on Protein-G Sepharose for purification. As a control, a Fab fragment comprised of mouse 9.3 variable domains and human constant domains was prepared by similar methods, as was a hybrid Fab fragment comprised of human constant domains, mouse 9.3 heavy chain variable domain, and humanized light chain variable domain.

The ability of the three Fabs to bind to CD28 was examined by ELISA. CD28Ig coated plates were incubated with Fab solutions at concentrations ranging from 1 pM to 10 mM. Binding was then assayed with an anti-human k immunoconjugate. The binding isotherms generated were processed to determine the equivalent concentration for half-maximal binding of the antibodies to CD28Ig (EC50) as described in Jin et al (1992), incorporated here by reference. This analysis, shown in Figure 11, indicated that the mouse

Fab had an EC50 of 20 nM, the EC50 of Hu9.3 was 630 nM, and the EC50 of the hybrid Fab was 30 nM. The similarity of the avidities of the hybrid and mouse Fabs showed that the reduction in binding by humanized 9.3 could be attributed to weakened interactions involving the heavy chain; thus humanization of the light chain alone caused minimal avidity loss.

EXAMPLE 3

HUMANIZED ANTI SCORPION TOXIN

The mouse anti-scorpion toxin antibody designated BCF2 was used as the subject non-human sequence for a humanized anti-scorpion toxin. The mouse BCF2 hybridoma line was described, and the efficacy of the BCF2 antibody in a mouse model demonstrated by Licea et al (1996). The sequence of the variable domains of BCF2 was disclosed by Selisko et al (1999), and is presented in Figure 12.

Canonical structure types of the light chain were determined as described before, and were type 5 for CDR1, type 1 for CDR2, and type 1 for CDR3. Canonical structure types of the heavy chain CDRs are type 1 for CDR1 and type 2 for CDR2. A humanized version of BCF2 was designed using the considerations discussed above for selection of human germline V and J gene sequences.

The light chain variable domain consisted of Kabat CDR sequences of the BCF2 light chain; and framework sequences identical to the human gene A2/DPK12 through residue 88 and identical to Jk4 from positions 98 – 108. The heavy chain variable domain consisted of Kabat CDR sequences of the BCF2 heavy chain, with the exception of residues 62 - 65, which were thought to be not critical to antigen recognition and hence made identical to the sequence of 1-f/DP3 at those positions; and Kabat framework sequences identical to 1-f/DP3 through residue 94 and identical to JH6 from residue 103 - 113.

The sequences of the heavy and light chain variable domains of humanized BCF2 antibody are shown in Figure 12. A recombinant Fab fragment with variable domains having these sequences was prepared as described for Example 2. As a control, a

$(Fab)'_2$ fragment was prepared pepsin digestion of mouse BCF2 IgG obtained from hybridoma cells.

The ability of the two Fabs to bind to CD28 was examined using a BIAcore biosensor instrument, with toxin immobilized on the surface of the sensor chip and antibody in the supernatant. This method has been described by Jönsson et al (1991), incorporated herein by reference. Fab solutions at concentrations varying over at least a 10-fold range were then passed over the chip to observe the association phase. The sensorgram was continued with buffer alone in the fluid phase to observe dissociation. Affinity, as a dissociation equilibrium constant Kd, was determined from the ratio of the kinetic rate constants kon/koff. The respective affinities were 10 nM for the mouse $(Fab)'_2$ and 140 nM for the humanized version.

EXAMPLE 4

HUMANIZED ANTI-HUMAN GAD65

The mouse antibody to human glutamic acid decarboxylase 65 kilodalton isoform, NGAD65.

The mouse NGAD65 hybridoma line and sequences of its antibody variable domains were described by Hampe et al (2001) and the sequences are presented in Figure 13. The first two residues of the light chain are omitted because they were derived from the oligonucleotide used for cloning.

Canonical structure types of the light chain CDRs were determined to be type 4 for CDR1, type 1 for CDR2, and type 1 for CDR3. Canonical structure types of the heavy chain CDRs were determined to be type 1 for CDR1 and type 2 for CDR2.

A humanized version of NGAD65 was designed using the considerations discussed above for selection of human germline V and J gene sequences. The light chain variable domain consisted of Kabat CDR sequences of the NGAD65 light chain; and framework sequences identical to the human Vk gene A17/DPK18 through residue 88 and identical to Jk3 from positions 98 – 108. The heavy chain variable domain consisted of Kabat CDR sequences of the BCF2 heavy chain, with the exception of residues 61 - 65,

which were thought to be not critical to antigen recognition and hence made identical to the sequence of 1-v at those positions; and Kabat framework sequences identical to 1-f/DP3 through residue 94 and identical to JH4 from residue 103 - 113.

The sequences of the heavy and light chain variable domains of humanized NGAD65 antibody are shown in Figure 13. A recombinant Fab fragment with variable domains having these sequences was prepared as described for Example 2. As a control, a Fab fragment comprised of mouse NGAD65 variable and constant domains was prepared by similar methods.

The ability of the two Fabs to bind to antigen was examined by an immunoprecipitation assay. Radioactive human glutamic acid decarboxylase was prepared by in vitro translation with 35S-methionine. The labeled antigen was incubated overnight with various concentrations of either of the two Fab fragments. Protein G-Sepharose beads were then added to sequester Fab and any associated antigen. Radioactivity was determined by scintillation counting and the EC50 was determined visually from the midpoint of plots of bound radioactivity versus concentration of Fab fragment. Values of EC50 were obtained of 0.36 pM for the mouse Fab and 9 pM for the humanized Fab. Even given the 25-fold loss of affinity of the humanized antibody relative to the mouse antibody, the humanized will still bind antigen sufficiently to be used in human in therapy without need for further mutagenesis of the sequence to make up for the 25 fold loss in affinity.

The methods provided herein have been exemplified by the use of mouse mature antibody genes as a source of the first Chothia canonical CDR and human antibody genes as a source for the second Chothia canonical CDR. These examples are particularly suitable for the construction of humanized antibodies for use in human therapeutic applications. Such humanized antibodies contain sufficient mouse amino sequences to retain a three dimensional structure necessary for avid antigen binding but also contain sufficient human antibody sequences to prevent unwanted immunogenicity in humans. One of ordinary skill in the art will appreciate, however, that the methods disclosed herein are equally applicable to preparing converted antibodies that include chimeric hypervariable regions derived from any two different vertebrate species.

In a more general sense, the first antibody sequence, which is originally selected by virtue of its binding to an antigen, may be referred to as the "subject" antibody sequence. Typically the subject antibody sequence is of mouse or rat origin. The second antibody sequence, which is selected from antibody sequences of the target animal, may be referred to as the "object" antibody sequence. The object antibody sequence is typically from a human or a farm animal that is the object of therapeutic treatment. Antibody compositions containing the chimeric hypervariable regions according to the methods of this invention result in a third antibody sequence which may be generally designated a "converted" antibody sequence. The converted antibody sequence differs in certain defined structural features from each of the subject and the object antibody sequences and is identical in certain other defined structural features to each of the subject or object sequences.

REFERENCES

Carter, P., Kelley, R. F., Rodrigues, M. L., Snedecor, B., Covarrubias, M., Velligan, M. D., Wong, W. L. T., Rowland, A. M., Kotts, C. E., Carver, M. E., Yang, M., Bourell, J. H., Shepard, H. M. & Henner, D. (1992) High level Escherichia coli expression and production of a bivalent humanized antibody fragment. Bio/Technology 10, 163-167.

Chothia, C. & Lesk, A. M. (1987) Canonical structure types for the hypervariable regions of immunoglobulins. J. Mol. Biol. 96, 901-917.

Chothia, C., Lesk, A. M., Gherardi, E., Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. & Winter, G. (1992) Structural repertoire of the human VH segments. J. Mol. Biol. 227, 799–817.

Chomczynski, P. & Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156–159.

Foote, J. & Winter, G. (1992) Antibody residues affecting conformation of the hypervariable loops. J. Mol. Biol. 224, 487-499.

Hampe, C. S., Lundgren, P., Daniels, T. L., Hammerle, L. P., Marcovina, S. M. & Lernmark, A. (2001) A novel monoclonal antibody specific for the N-terminal end of GAD65. J Neuroimmunol. 113, 63-71.

Hansen, J. A., Martin, P. J. & Nowinski, R. C. (1980) Monoclonal antibodies identifying a novel T cell antigen and Ia antigens of human lymphocytes. Immunogenetics 10, 247-260.

Jin, L., Fendly, B. M. & Wells, J. A. (1992) High resolution functional analysis of antibody-antigen interactions. J. Mol. Biol. 226, 851–865.

Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S. & Winter, G. (1986) Replacing the complementarity-determining regions in a human antibody with those from a mouse. Nature 321, 522-525.

Jönsson, U., Fägerstam, L., Ivarsson, B., Lundh, K., Löfås, S., Persson, B., Roos, H., Rönnberg, I., Sjölander, S., Stenber, E., Ståhlberg, R., Urbaniczky, C., Östlin, H. & Malmqvist, M. (1991) Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. BioTechniques 11, 620-627.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Coeller, K. (1991) Sequences of proteins of immunological interest. 5th ed. 1991, Bethesda: US Dept. of Health and Human Services, PHS, NIH.

Licea, A. F., Becerril, B. & Possani, L. D. (1996) Fab fragments of the monoclonal antibody BCF2 are capable of neutralizing the whole soluble venom from the scorpion Centruroides noxius Hoffman. Toxicon 34, 843–847.

MacCallum, R. M., Martin, A. C. R. & Thornton, J. M. (1996) Antibody-antigen interactions: contact analysis and binding site topography. J. Mol. Biol. 262, 732-745.

Padlan. E. (1991) A possible procedure for reducing the immunogenicity of antibody variable domains while preserving their ligand binding properties, Molecular Immunology 28:489-498

Padlan, E. O., Abergel, C. & Tipper, J. P. (1995) Identification of specificity-determining residues in antibodies. FASEB J. 9, 133–139.

Riechmann, L., Clark, M., Waldmann, H. & Winter, G. (1988) Reshaping human antibodies for therapy. Nature 332, 323-327.

Rutgeerts, P., et al (1999) Efficacy and safety of retreatment with anti-tumor necrosis factor antibody (INFLIXIMAB) to maintain remission in Crohn's disease. Gastroenterology 117, 761-769

Selisko, B., Licea, A. F., Becerril, B., Zamudio, F., Possani, L. D. & Horjales, E. (1999) Antibody BCF2 against scorpion toxin Cn2 from Centruroides noxius Hoffman: primary structure and three-dimensional model as free Fv fragment and complexed with its antigen. Proteins 37, 130–143.

Tamura, M., Milenic, D., Iwahashi, M., E., P., Schlom, J. & Kashmiri, S. (2000) Structural Correlates of an Anticarcinoma Antibody: Identification of Specificity-Determining

Residues (SDRs) and Development of a Minimally Immunogenic Antibody Variant by Retention of SDRs Only. J. Immunol. 164, 1432–1441.

Tomlinson, I. M., Cox, J. P. L., Gherardi, E., Lesk, A. M. & Chothia, C. (1995) The structural repertoire of the human Vk domain. EMBO J. 14, 4628-4638.

Wu, T. T. & Kabat, E. A. (1970) An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. J. Exp. Med. 132, 211-250.

Ye, Q.-Z., Johnson, L. L. & Baragi, V. (1992) Gene synthesis and expression in E. coli for PUMP, a human matrix metalloproteinase. Biochem. Biophys. Res. Comm. 186, 143–149.Martin, A. C. R. & Thornton, J. M. (1996) Structural families in loops of homologous proteins: automatic classification, modelling and application to antibodies. J. Mol. Biol. 263, 800-815.

Henikoff, S. & Henikoff, J. G. (1992) Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89, 10915-10919.

Claims

1. A method of designing a molecule that binds an antigen, comprising,

determining a canonical CDR structure type for at least two CDRs in a variable region of a subject antibody that binds the antigen;

selecting an object variable region from an object antibody based on whether the object variable region has the same canonical CDR structure types at corresponding locations in the object variable region as the at least two canonical CDR structure types of the subject variable region; and

designing a molecule that binds the same antigen as the subject antibody comprising a chimeric variable region having the at least two CDRs from the subject variable region grafted into a framework sequence of the selected object variable region at the corresponding locations, where the framework sequence of the chimeric variable region differs from the framework sequence of the selected object variable region by no more than 10 amino acid residues.

2. The method of claim 1 wherein designing the chimeric variable region includes grafting each of three CDRs from the subject variable region into corresponding locations in the object framework sequence.

3. The method of claim 1 wherein the framework sequence of the chimeric variable region differs from the framework sequence of the object variable region by no more than 5 amino acid residues.

4. The method of claim 1 wherein the framework sequence of the chimeric variable region differs from the framework sequence of the object variable region by no more than 2 amino acid residues.

5. The method of claim 1 wherein selecting the object variable region includes obtaining a set of candidate object variable regions having the same canonical CDR structure types as the subject CDRs and ranking the set of candidate object variable regions by comparing position by position similarity of amino acid residues of the subject CDRs to the corresponding object CDRs according to a ranking criterion.

6. The method of claim 5 wherein the object variable region is selected from the top 25% of the ranked candidate variable regions.

7. The method of claim 5 wherein the ranking criterion includes a score of amino acid identity between the subject and object CDR sequences at corresponding residue positions of at least one CDR.

8. The method of claim 5 wherein the ranking criterion includes a score of amino acid identity between the subject and object CDR sequences at corresponding residue positions of at least two CDRs.

9. The method of claim 5 wherein the ranking criterion includes a score of amino acid identity between the subject and object CDR sequences at corresponding residue positions in each of three CDRs.

10. The method of claim 8 wherein the ranking criterion further includes a score of amino acid homology between the subject and object CDRs at corresponding residue positions of at least one CDR.

11. The method of claim 8 wherein the ranking criterion further includes a score of amino acid homology between the subject and object CDRs at corresponding residue positions of at least two CDRs.

12. The method of claim 8 wherein the ranking criterion further includes a score of amino acid homology between the subject and object CDRs at corresponding residue positions for each of three CDRs.

13. The method of claim 1 wherein the CDRs are Kabat defined CDRs.

14. The method of claim 1 wherein the CDRs are Chothia defined CDR loops.

15. The method of claim 1 wherein designing the chimeric variable region further includes substituting at least one amino acid residue of a subject CDR sequence with a different amino acid, with the proviso that no more than 4 residues are substituted in any of the subject's light

chain CDR1, light chain CDR2, light chain CDR 3, heavy chain CDR1, or heavy chain CDR3 and no more than 10 amino acids are substituted in the subject heavy chain CDR2.

16. The method of claim 1 wherein the act of designing the chimeric molecule further includes substituting at least one but no more than 10 amino acid residues of the selected framework sequence with a different amino acid residue.

17. The method of claim 1 wherein each of three subject CDRs is a light chain CDR, and if one of the three subject CDR sequences is of a canonical structure type absent from the object light chain variable region, then the act of selecting further includes selecting an object variable region with a CDR of a different canonical structure type than the subject CDR type that is absent, with the proviso that the different object canonical structure type has a length no more than two amino acid residues smaller or larger than the subject canonical structure type that is absent.

18. The method of claim 17 wherein if the absent CDR canonical structure type is of canonical type 1, then the act includes selecting an object variable region with a canonical type 2 CDR at the corresponding location, or if the subject CDR sequences is of canonical type 5 then the act includes selecting an object variable region with a canonical type 4 or 3 at the corresponding location.

19. The method of claim 1 wherein the subject variable region is a mouse variable region.

20. The method of claim 1 wherein the object variable region is a human variable region.

21. The method of claim 1 wherein the subject variable region is a mouse variable region and the object variable region is a human variable region.

22. The method of claim 1 wherein the object variable region sequence is selected from the group consisting of human V_k, V_λ, V_H, J_H, J_k and J_λ sequences.

23. The method of claim 1 wherein the act of designing the molecule that binds the antigen includes designing chimeric variable regions for each of the light chain and the heavy chain.

24. The method of claim 23 wherein the chimeric variable region includes the frameworks from human V_k, and V_H sequences.

25. The method of claim 23 wherein the chimeric variable light chains and chimeric heavy chains are designed to form a molecule selected from the group consisting of a Fab fragment, a $(Fab)'_2$ molecule, and a single chain Fv molecule.

26. The method of claim 23 wherein the chimeric variable light chains and chimeric variable heavy chains are further designed to include an antibody constant region domain for the object antibody to form a complete antibody.

27. The method of claim 1 wherein the selected object variable region is encoded by a human germline variable region gene.

28. The method of claim 1 wherein the selected object variable region is from a mature human antibody.

29. The method of claim 1 through 28 further comprising preparing a nucleic acid encoding said chimeric variable region.

30. A host capable of expressing a nucleic acid made according to claim 29.

31. The method of claim 1 further including the act of constructing a set of molecules with different object variable regions, determining a dissociation constant between the antigen and different members of the set of molecules, and selecting a molecule having a dissociation constant of at least 10^6 M^{-1}, at least 10^7 M^{-1} or at least 10^8 M^{-1}.

32. A molecule designed according to the method of claim 1-31 characterized in that said molecule binds the antigen, and has a chimeric variable region containing at the at least two CDRs from the subject variable region grafted into a framework sequence of the object variable region.

33. A humanized antibody variable region assembled into a molecule that binds an antigen, comprising

a chimeric antibody variable region containing at least two non-human CDR sequences grafted into a human variable region framework sequence, the chimeric variable region being characterized by having no more than 10 amino acid residues that differ from a framework sequence of a candidate human antibody variable region having at least two human CDR sequences with the same canonical structure type as the non-human CDR sequences for at least two corresponding CDR positions.

34. The humanized variable region of claim 33 wherein the non-human variable region is a mouse variable region.

35. The humanized variable region of claim 33 wherein the human variable region sequence is selected from the group consisting of V_k, V_λ, V_H, J_H, J_k and J_λ sequences.

36. The humanized variable region of claim 33 wherein the assembled molecule includes chimeric variable region sequences for each of a variable light chain and a variable heavy chain.

37. The humanized variable region of claim 36 wherein the chimeric variable light chains and chimeric variable heavy chains are assembled to form a molecule selected from the group consisting of a Fab fragment, a $(Fab)'_2$ molecule, and a single chain Fv molecule.

38. The humanized variable region of claim 36 wherein the chimeric molecule further includes a human antibody constant region domain to form a complete antibody.

39. The humanized variable region of claim 33 wherein the candidate human variable region is encoded by a human germline variable region gene.

40. The humanized variable region of claim 33 wherein the candidate human variable region is a sequence from a human mature antibody.

41. The humanized variable region of claim 33 wherein the assembled molecule has a dissociation constant for its antigen of at least at least 10^6 M^{-1}, at least 10^7 M^{-1} or at least 10^8 M^{-1}.

42. The humanized variable region of claim 33 wherein the assembled molecule does not elicit an immune response when administered to a human.

43. The humanized variable region of claim 33 wherein the assembled molecule binds a human CD28 antigen.