

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 12-081792- -00011-0000

Fecha: 2013-04-26 15:41:12 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 329 CTOINFORMACION Folios: 18



SUSTENTO DE OPOSICIÓN AL EXPEDIENTE:

12-081792

Prorroga

TÍTULO:

**“INMUNOGLOBULINAS CON DOMINIO VARIABLE DUAL Y USOS DE LAS
MISMAS”**

Gaceta 655 de 31 de octubre de 2012,

Nº de publicación 1686.

Prioridad US 61/255.686 28/10/2009

Apoderado ALICIA LLOREDA RICAURTE

Sociedad: ABBVIE, INC.

Pedro A. Díaz M., mayor de edad, identificado con la cedula de ciudadanía No. 19.085.808 de Bogotá, obrando en calidad de Representante Legal de la compañía VITALIS S.A.C.I dentro del trámite de referencia, deseo sustentar y complementar la oposición presentada.

Nos basamos en el art 42, inciso 2º de la Decisión 486 de 2000 de la Comunidad Andina de Naciones, con los siguientes argumentos:

RATIFICACIÓN DE LA PETICIÓN

Comedidamente me permito insistir en los argumentos planteados en la oposición referida, los cuales, como se dijo en su momento, demuestran que la solicitud incumple los requisitos de patentabilidad establecidos por los artículos 14 y 18 de la Decisión 486 de 2000, y que por ello no merece recibir el beneficio patentario.

En su momento se mencionó que la solicitud colombiana en trámite 12-081-792 que se refiere a proteínas de unión multivalentes y multiespecíficas, no reúne los requisitos necesarios para otorgársele el privilegio de patente de invención.

Los anticuerpos reclamados no se definen o precisan mediante una estructura determinada sino que se limitan mediante opciones de cambio en la secuencia de anticuerpos duales con parejas que, sin presentar claridad, pueden ser seleccionables a través de sus moléculas diana o receptores.

Adicionalmente cualquier clase de anticuerpos duales estaría cubierto por las reivindicaciones, sin que halla sido efectuado realmente, o se tenga evidencia de su efectividad y uso terapéutico. De esta forma, no existe claridad en la definición, suficiencia ni sustento en la descripción de dichos anticuerpos según el artículo 28 de la D. 486 de la Comunidad Andina de Naciones.

Sobre los requisitos de Novedad y Nivel Inventivo, se demostró que la solicitud no cumple con las especificaciones mínimas entendiendo que los cambios en parejas de dominios afines a diferentes parejas de receptores, son de uso rutinario no solo en la técnica sino en la misma naturaleza, que permite que cada organismo efectúe mutaciones, adiciones, deleciones y sustituciones en secuencias de proteínas como los anticuerpos mejorado sus propiedades de afinidad al antígeno.

Con relación a la reivindicación 1 se encontró un documento que anula la novedad de la solicitud definido con número de publicación WO 2009/0215992, título "*Dual variable domain immunoglobulin and uses thereof*" publicado por Wu et al., el 28 de febrero de 2008.

Este documento describe una proteína de unión que comprende una cadena de polipéptidos, en donde dicha cadena de polipéptidos comprende VD1-(X1)_n-VD2-C-(X2)_n, en donde;

VD1 es un primer dominio variable de la cadena pesada;

VD2 es un segundo dominio variable de la cadena pesada;

C es un dominio constante de la cadena pesada;

X1 es un conector con la condición que no sea CH1;

X2 es una región Fc; y

n es 0 o 1 (ver párrafo 0013);

en donde la proteína de unión tiene la capacidad de unirse a un par de antígenos seleccionados del grupo que consiste de TNF (seq.2) (párrafo 0038 y 0041) y NGF (párrafo 0041)

Con relación a la reivindicación 3 por ejemplo, el documento WO 2009/0215992 describe una proteína de unión que comprende una cadena de polipéptidos, en donde dicha cadena de polipéptidos comprende VD1-(X1)_n-VD2-C-(X2)_n, en donde;

VD1 es un primer dominio variable de la cadena ligera (párrafo 0013);

VD2 es un segundo dominio variable de la cadena pesada;

C es un dominio constante de la cadena ligera;

X1 es un conector con la condición que no sea CH1;

X2 es una región Fc; y

n es 0 o 1 (ver párrafo 0015);

en donde la proteína de unión tiene la capacidad de unirse a un par de antígenos seleccionados del grupo que consiste de TNF (seq.2) (párrafo 0038 y 0041) y NGF (párrafo 0041)

El documento además describe que la proteína de unión corresponde a una de las reivindicaciones 1 o 3 donde n es 0 (párrafo 0013) esto es igual a la modalidad particular de la solicitud en su reivindicación 3. El documento de anterioridad enseña que la proteína de unión comprende dos primeras cadenas polipeptídicas y dos segundas cadenas de polipéptido (ver párrafo 0017), con lo cual también se anula la novedad de la modalidad particular de la reivindicación 9 de la solicitud.

Nuevamente, y con relación a la reivindicación 10 de la solicitud, la anterioridad WO 2009/0215992 describe además que la región Fc es seleccionada del grupo que consiste de secuencias nativas de regiones Fc y una variante de la secuencia de la región FC (ver

párrafo 0044). Esta región Fc además es seleccionada del grupo que consiste de regiones Fc de IgG1, IgG2, IgG3, IgG4, IgA, IgM, IgE, e Ig.

De acuerdo a lo anterior, es claro observar que la formula reclamada ya ha sido conocida y reclamada por otra patente anterior dando como lugar a que la invención no pueda ser reivindicada nuevamente, hecho que debería ser suficiente para negar de plano la presente solicitud de patente.

En la reivindicación 2 por ejemplo se reclaman proteínas de unión tales como las descritas pero que comprenden una secuencia de aminoácidos seleccionada del grupo que consiste de las SEQ ID NO. 28, 30, 32, 34, 36, 38, 40, 42 Y 44.

En esta sola reivindicación se presentan 9 diferentes anticuerpos a evaluar, con lo cual, y como es frecuente, el primero de los cuales determinaría el primer grupo de dicha reivindicación.

Haciendo un simple análisis en la base de datos de NCBI puede observarse que existe una cantidad de proteínas con secuencias que cumplirían con las características reclamadas, ya que se observan moléculas de inmunoglobulina con una cadena pesada de región variable.

En este caso, por citar solo uno, se encuentra el reporte de NCBI titulado “Immunoglobulin heavy chain variable region [Homo sapiens]” publicado el 21 de diciembre de 2003 mencionando la secuencia de aminoácidos:

>gi|39938155|gb|AAR32496.1| immunoglobulin heavy chain variable region [Homo sapiens]

KPTQTLTLTCTFSGLPVGTSGRPVGWFRQPPGQALEWLAVIYSDDDDRYNPHLEN
RLTVTRDAANNQVVLMTNMDPADKATYYCARLHSTMTTGYFDYWGQGVLVT
VSPASTKGPSVFPLAPCSRSTSESTAALGCLVK DYFP

Haciendo la alineación correspondiente, se obtienen:

Range 2: 111 to 123 [GenPeptGraphics](#) [Next Match](#) [Previous Match](#) [First Match](#)

Alignment statistics for match #2

Score	Expect	Identities	Positives	Gaps
41.8 bits(91)	0.004	13/13(100%)	13/13(100%)	0/13(0%)

Query 1 ASTKGPSVFPLAP 13
ASTKGPSVFPLAP
Sbjct 111 ASTKGPSVFPLAP 123

Teniendo en cuenta que la secuencia se halla repetida en la solicitud, se observa que al definir los fragmentos de actividad de las inmunoglobulinas como en este caso, dichas secuencias pueden ser modificadas de acuerdo a como se ha mencionado en la anterioridad WO 2009/0215992 incluyendo la secuencia dobles de incluso los mismos anticuerpos por lo que sería evidente para cualquier técnico medio versado en la materia encontrar las secuencias activas en inmunoglobulinas tales como las que aparecen en el documento de NCBI e incluirlas en el anticuerpo dual de WO 2009/0215992 dando como resultado los anticuerpos duales de la solicitud.

En este orden de ideas, no existe efecto técnico ventajoso sobre el estado de la técnica al realizar las sustituciones de dominios determinados que se unen a los antígenos divulgados de esta forma, y por lo tanto lo reclamado obedece a una simple selección de sustituciones del estado de la técnica pero que no demuestran efectividad superior con los mismos.

COMPLEMENTO DE INFORMACIÓN

Existe además, una gran cantidad de literatura que menciona las ventajas de realizar anticuerpos duales tales como los reclamados por la solicitud en estudio sin necesidad de realizar un experimento o desarrollar una inventiva que permita llegar a resultados sorprendentes.

Caso específico de la publicación Estadounidense US 2009/0259026 de título: "Ligand", publicada el 15 de octubre de 2009, que provee ligandos duales específicos que comprenden un primer dominio variable de inmunoglobulina con una primera especificidad de unión y un dominio variable de inmunoglobulina complementario o no complementario con una segunda especificidad de unión.

En el caso de esta anterioridad se observa como la producción de dichos ligandos duales fueron evaluados sobre el receptor de TNF (ver figuras 10 - 12) entendiéndose que cualquiera de estos anticuerpos duales específicos a TNF son el objeto de la solicitud reclamada y que no implica que la actividad de unión sea sorprendente.

El documento adicionalmente menciona los antígenos sobre los cuales la molécula de liga y gracias a dicha unión previene el efecto de los mismos receptores, estos pueden ser TNF- α y NGF tal como se observa en el siguiente aparte del párrafo 0043:

Neurturin, Nerve growth factor, β -NGF, NT-3, NT-4, Oncostatin M, PDGF-AA, PDGF-AB, PDGF-BB, PF-4, RANTES, SDF1 α , SDF10, SCF, SCGF, stem cell factor (SCF), TARC, TGF- α , TGF- β , TGF- β 2, TGF- β 3, tumour necrosis factor (TNF), TNF- α , TNF- β , TNF receptor I, TNF receptor II,

Adicionalmente, la formula de dichos anticuerpos duales era conocida y su mecanismo para identificar los antígenos a los cuales se une también fue empleada incluso en este documento de anterioridad ya que menciona que el anticuerpo debe comprender un primer dominio de unión al epítipo con una especificidad a un primer epítipo y un segundo dominio de unión al epítipo con una especificidad donde el método comprende los pasos de:

- a) seleccionar el primer dominio de unión al epítipo por su habilidad de unirse al primer epítipo.
- b) Seleccionar el segundo dominio de unión al epítipo por su habilidad de unirse a un segundo epítipo.
- c) Combinar los dominios de unión a epítipo tal que los dominios estén en una conformacional cerrada y
- d) Seleccionar los ligando de conformación multiespecífica cerrada por su habilidad de unirse a dichos primer y segundo epítipo y dicho segundo epítipo.

Los ejemplos específicos de la reivindicación 2 por ejemplo, obedecen a proteínas de unión caracterizadas por comprender secuencias de aminoácidos determinadas. Sin embargo, los cambios puntuales en secuencias CDR de anticuerpos monoclonales en el momento de la presentación de la solicitud son fácilmente determinados por análisis bioinformático.

Por ejemplo, se puede observar la sección 1.2 del artículo científico de Ewert et al.,¹ que se emplean las técnicas observadas en la naturaleza de realizar cambios en los CDRs de los anticuerpos con el fin de obtener propiedades deseadas (ver último párrafo), mencionando que una de dichas propiedades es la estabilidad en la unión del anticuerpo a su molécula receptora.

Se menciona que las interacciones de unión en el complejo anticuerpo-ligando están determinadas (ver sección 1.4 de Ewert et al.,), entre otras, por:

- Las secuencias CDRs
- Recombinación genética de los segmentos V,D, y J
- Residuos de *loop* externo de las regiones variables
- Residuos cercanos a la interfase le heterodimero V_H y V_L.

En este sentido, el documento menciona la comparación de las secuencias con librerías de anticuerpos HuCAL, mencionando la formación de secuencias consenso que dan como resultado cambios en las secuencias que influyen en la actividad del anticuerpo.

El documento de Ewert et al., describe entonces métodos para diseñar y mejorar las propiedades biofísicas de cualquier anticuerpo o fragmento del mismo para hacerlo disponible en uso in vivo; dentro de estos métodos se tienen entonces se tiene la transferencia de una región de una secuencia CDR a un anticuerpo diferente con propiedades favorables. El otro método emplea una gran cantidad de datos disponibles para diseñar la secuencia de los anticuerpos.

Así, entendiendo que lo reivindicado exclusivamente para la reivindicación 2, obedece a diferentes secuencias de anticuerpo que únicamente obedecen a diferencias posibles mediante un análisis bioinformático, en CDRs u otro segmento de la secuencia, cualquier técnico medio versado en la materia identificaría la forma para diseñar nuevos anticuerpos basados en la información científica disponible, tal como la de Ewert et al.

¹ Stefan E., Annemarie H, y Plückthun A. "Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering" En: Methods 34 (2004) p. 184-199.

Todo lo anterior bajo la premisa de que existiera un desarrollo claro de la solicitud en estudio, pero como se ha visto, y teniendo en cuenta los documentos mencionados en la oposición presentada, no se ha demostrado de igual manera la novedad para la materia reclamada.

Todas las modificaciones observadas en los documentos, únicamente señalan la obviedad del diseño de anticuerpos como los presentados en la solicitud, y el hecho de que se trate de anticuerpos que se unen a TNF-alfa, se remiten a anticuerpos de referencia como el Adalimumab o infliximab los cuales podrían hacer parte de los anticuerpos biespecificos no novedosos ni inventivos de la solicitud, ya que dichos desarrollos han sido empleados comúnmente en el estado de la técnica.

PRUEBAS

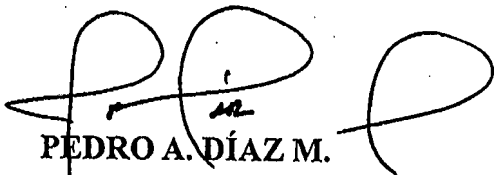
Ruego que se tengan como tales, sin perjuicio de que el Despacho considere oficiosamente, en particular de las allegadas por el suscrito a los expedientes mencionados en este escrito, las siguientes:

- ✓ Los datos incluidos en el presente documento resultado de bases bibliográficas.
- ✓ El documento WO 2009/0215992 A1 "*Dual variable domain immunoglobulin and uses thereof*" publicado el 27 de octubre de 2009.
- ✓ El documento US 2009/0259026 de título: "*Ligand*", publicada el 15 de octubre de 2009
- ✓ El artículo científico de Stefan E., Annemarie H, y Plückthun A. "*Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering*" En: Methods 34 (2004) p. 184-199.

NOTIFICACIONES

Recibiré las notificaciones a que haya lugar en la secretaría de su Despacho o en mi oficina localizada en la Carrera 7a No. 156-80, Oficina 1104, Tel. 6734340 Fax 6734407 de esta ciudad.

Señor Superintendente,

A handwritten signature in black ink, consisting of stylized, overlapping loops and curves, positioned above the printed name.

PEDRO A. DÍAZ M.

C.C. 19085808 de Bogotá

Representante Legal de VITALIS S.A.C.I.



US 20090259026A1

(19) **United States**

(12) **Patent Application Publication**
Tomlinson et al.

(10) Pub. No.: **US 2009/0259026 A1**

(43) Pub. Date: **Oct. 15, 2009**

(54) **LIGAND**

(76) Inventors: **Ian Tomlinson, Cambridge (GB);
Laurent Jaspers, Cambridge (GB);
Jasper Clube, Cambridge (GB);
Lucy J. Holt, Cambridge (GB);
Oliver Schon, Cambridge (GB)**

Correspondence Address:
**SMITHKLINE BEECHAM CORPORATION
CORPORATE INTELLECTUAL PROPERTY-
US, UW2220
P. O. BOX 1539
KING OF PRUSSIA, PA 19406-0939 (US)**

(21) Appl. No.: **11/704,832**

(22) Filed: **Feb. 8, 2007**

Related U.S. Application Data

(63) Continuation-in-part of application No. 11/023,959,
filed on Dec. 28, 2004, which is a continuation of
application No. PCT/GB03/02804, filed on Jun. 30,
2003.

(30) **Foreign Application Priority Data**

Dec. 27, 2002 (GB) 0230202.4
Jun. 30, 2003 (GB) PCT/GB02/03014
May 31, 2005 (WO) WO2005118642

Publication Classification

(51) **Int. Cl.**
C12P 21/08 (2006.01)
C07K 16/18 (2006.01)

(52) **U.S. Cl.** 530/387.3; 530/387.1

(57) **ABSTRACT**

The invention provides a dual-specific ligand comprising a first immunoglobulin variable domain having a first binding specificity and a complementary or non-complementary immunoglobulin variable domain having a second binding specificity.

Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering

Stefan Ewert,¹ Annemarie Honegger,* and Andreas Plückthun*

Biochemisches Institut, Universität Zürich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland

Available online 15 July 2004

Abstract

By combining the knowledge gained from an analysis of the biophysical properties of natural antibody variable domains, the effects of mutations obtained in directed evolution experiments, and the detailed structural comparison of antibodies, it has now become possible to engineer antibodies for higher thermodynamic stability and more efficient folding. This is particularly important when antibodies are to be used under conditions where the disulfide bonds cannot form, i.e., in intracellular applications (as “intrabodies”). We describe in detail two methods for the knowledge-based improvement of antibody stability and folding efficiency. While CDR grafting from a non-human to the most closely related human antibody framework is an established technique to reduce the immunogenicity of a therapeutic antibody, CDR grafting for stabilization implies the use of a more distantly related acceptor framework with superior biophysical characteristics. The use of such dissimilar frameworks requires particular attention to antigen contact residues outside the classical CDR definition and to residues capable of indirectly affecting the conformation of the antigen binding site. As a second alternative, the stability of a suboptimal framework can be improved by the introduction of point mutations designed to optimize key residue interactions. We describe the analysis methods used to identify such point mutations, which can be introduced all at once, while maintaining the framework features necessary for antigen binding. These rational approaches render the continued “rediscovery” of certain mutations by directed evolution unnecessary, but they can also be used in conjunction with such methods to discover even better molecules.

© 2004 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Intrabodies and scFv format

Antibodies are an important component of the immune system, which defends higher vertebrates against any kind of intrusion by microorganisms. To fulfill this task, a great variety of immunoglobulin variable domains have evolved, which can be divided by their sequence homology into different families. Antibodies are secreted by plasma cells and are designed by nature to act in extracellular fluids, where they can exploit the

stabilizing effect of disulfide bonds. The demands on stability have kept a selection pressure on immunoglobulin domains to retain an intradomain disulfide bond in all germline genes. Nevertheless, there is great interest in the use of antibodies for intracellular applications, be it for the modulation of cellular effects [1,2] or the use of intracellular selection systems [3–5]. Yet, the disulfide bonds cannot form in the reducing milieu of most intracellular compartments. As a consequence, the stability of the antibody in the absence of the disulfide is a crucial variable for these applications. Indeed, a few antibodies in the scFv format (in which the two variable domains are connected by a linker to give a single polypeptide chain [6,7]) have been identified, engineered or evolved, which have the ability to fold and be of sufficient stability in the absence of the disulfide bonds [8–10]. The purpose of this article was to review how to make other antibodies stable enough for this purpose.

* Corresponding authors: Fax: +41-1-635-5712.

E-mail addresses: honegger@bioc.unizh.ch (A. Honegger), plueckthun@bioc.unizh.ch (A. Plückthun).

¹ Present address: ESBATech AG, Wagistr. 21, CH-8952 Zürich-Schlieren, Switzerland.

1.2. Need for stable intrabodies

The intradomain disulfide bond contributes about 14 kJ/mol to the stability of antibody domains [10–12]. Therefore, antibody fragments expressed in a reducing environment, e.g., in the cytoplasm, are destabilized and a smaller fraction of these fragments, if any, is likely to fold to the correct native structure. This fact is believed to be responsible for the frequently observed reduced functional expression level of cytoplasmically expressed antibody fragments, as well as for their high tendency to form aggregates [13–15]. Nevertheless, a number of cytoplasmically expressed antibody fragments were reported to show specific biological effects (see [14,16] for representative examples). For many applications, the observed effects are insufficient, and such intrabodies would require further optimization by protein engineering. Furthermore, the general applicability of “intra-body” technology is strongly limited, as long as only a small subset of suitable specific antibodies has a biological effect.

A crucial assumption for the approach presented here is that there is actually a correlation between stability and biological activity. This was most directly demonstrated by Wörn et al. [17], who showed that a framework-engineered stabilized version of an intrabody showed significantly improved activity, while a destabilized point mutant of the same intrabody showed decreased effects *in vivo*. Similar results have subsequently been obtained in other systems [18]. These results indicate that stability engineering can result in improved performance of scFv fragments as intrabodies.

1.3. Principal strategies to arrive at stable intrabodies

As the natural antibody repertoire comprises frameworks of very different stability and aggregation properties (see below), any use of natural antibody domains from either monoclonal antibodies or natural libraries will, except in the rare case where a sufficiently stable antibody was obtained by sheer luck, require further steps of improvement. To achieve this goal, we can distinguish approaches involving selection and evolution from those involving engineering.

All genuine evolutionary approaches use random mutagenesis followed by selection in an iterative manner, with an attempt to reward the molecules that show improvement of the desired property. The crucial point in the selection for improved biophysical properties is to find a way to link antibody stability to binding activity, since most selection systems only exploit specific binding to the antigen. Stability must therefore be reflected in the number of native and active molecules of a given molecular species under a given set of conditions (reviewed in [19]). The selection pressure for active molecules during

phage display (reviewed in [20]) can be increased, e.g., by high temperature, addition of denaturants such as guanidine-hydrochloride (GdnHCl) or proteases [21,22]. Ribosome display [23] can also be used to select for increasing stability, if the *in vitro* translation steps between successive binding and amplification rounds are performed in the presence of DTT, which keeps the cysteines in the reduced state and thus increases the selection pressure towards better folding and more stable scFv fragments [24]. Intracellular selection systems for scFv fragments such as the yeast two-hybrid system [4] or the protein complementation assay (PCA) [5] act implicitly in a similar way, as the selected molecules must function in the reducing environment of the cytosol. In all these systems, care must be taken that specific binding (which should require correct folding) is the selection criterium, avoiding the enrichment of unspecifically associating sticky molecules.

Yet another facet of this approach is to first select for binding or activity in the presence of disulfide bonds, e.g., by standard phage display, and then test the candidates for activity inside the cell [4,25]. This requires the screening of a large number of candidate molecules and appears to be a rather laborious approach.

This review aims to outline what can be achieved by rational engineering, based on our knowledge gained from the investigation of the properties of natural antibody variable domains, the results of many directed evolution experiments, and the evaluation of designed point mutants [17]. While this review will largely focus on the improvement of one given antibody, the same approaches can be and have been applied to library design [26], to create ensembles of stable molecules. Clearly, this is an iterative and ongoing process.

1.4. Biophysical properties of families are known

The germline contains a large collection of different V genes, and this variety is needed, since binding interactions are not limited to the CDR-3s, which derive their diversity from the genetic recombination of V, D, and J gene segments, but also involve germline-encoded CDR-1s and -2s as well as residues in the outer loop (for proteins) and residues close to the pseudo 2-fold axis relating V_H and V_L in the heterodimer interface (Fig. 1). Recently, we have analyzed the biophysical properties of human germline family specific consensus domains [27] derived from the Human Combinatorial Antibody Library (HuCAL[®]) [26]. In Fig. 2, an alignment of these consensus sequences representing the different human germline families is shown, with a header summarizing the location, surface exposure, variability, and likeliness of involvement in antigen binding and other interfaces. In the case of the V_H domains, we found that the hV_H3 germline family consensus domain was the most stable

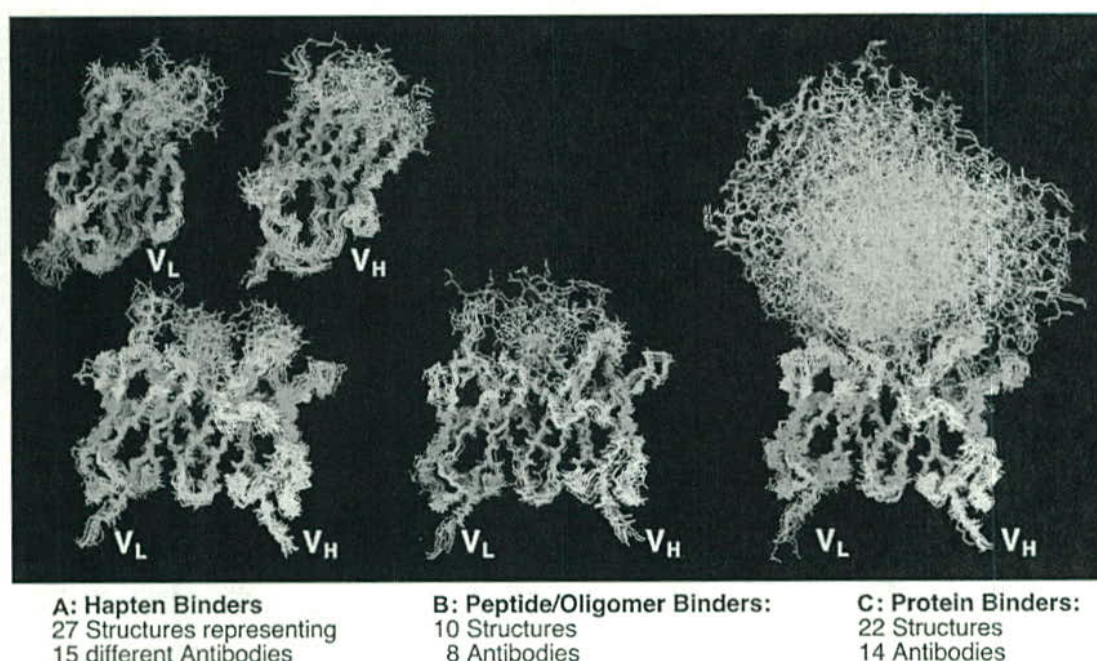


Fig. 1. Antibodies binding haptens, oligopeptides and oligosaccharides or proteins. Superposed crystal structures of antibody-antigen complexes were sorted into three classes according to the type of antigen. The antigens are colored pink. Structurally variable residues within the CDRs of the antibodies are shown in green, those at the N-terminus, to the N-terminal side of CDR-I and within the outer loop in cyan. The structurally least variable residues whose C α -positions were used for the least-squares superposition of the antibody fragments are shaded gray. Residues within the inner (dimer interface) β -sheet of V_L and V_H whose side chains contribute both to the dimer interface and, depending on the antigen, to antigen binding, are shown in yellow, orange, and red, depending on depth within the hapten binding cavity. It can be seen that there is extensive binding to residues which formally belong to the framework, either in the dimer interface region for hapten binders or binders of peptides (which frequently use a side-chain in a hapten-like binding mode), and to the outer loop in protein binders. Hapten binders commonly form a deep, funnel-shaped binding pocket enlarged by a long CDR-L1 and open CDR-H3 conformation, while protein binders preferentially utilize a relatively flat antigen binding surface characterized by a short CDR-L1 and a closed CDR-H3 conformation. This is one of the main reasons why the immune system uses different frameworks, rather than different CDRs all on the same framework. It also highlights the points to consider in loop grafting.

Fig. 2. Alignment of the human germline family consensus sequences, represented by the sequences of the Human Combinatorial Antibody Library (HuCAL) [26]. Homology models of the 14 domains have been deposited in the PDB (<http://www.rcsb.org/pdb>), models of the 49 Fv fragments which can be produced by the combination of the 7 V_L with the 7 V_H domains are available on the AAAAA website (<http://www.biochem.unizh.ch/antibody>). The sequences are color coded according to residue type (see legend). AHO numbering and alignment [30] is shown and used throughout this article. Kabat numbering [34] is shown for comparison. AHO numbering is identical for all types of Ig variable domains and places gaps and insertions centered on the turn positions highlighted in yellow, ensuring that residues carrying the same residue number in different variable domains are structurally equivalent. Dark gray shading indicates the structurally least variable positions used for least-squares superposition of structures. (A) Structural variability: average r.m.s. deviation from mean C α position (average of 185 V_K and 206 V_H structures representing >100 non-identical sequences, all the experimental Fv and Fab structures with a resolution better than 3.0 Å available in the PDB database at the time of the analysis) for structures aligned according to the AHO nomenclature [37]. Individual domains were excised from the corresponding PDB files and aligned by a least-squares fit of the C α -positions of the structurally least variable residues (3–7, 20–24, 41–47, 51–57, 78–82, 89–93, 102–108, and 138–144) to the corresponding C α -positions of a reference structure 1YEH [57] for V_K and 1MFD [58] for V_H. The mean C α positions for each residue were calculated and the average deviation for each residue position in the alignment is indicated by a color code (white: r.m.s. deviation <0.5 Å, yellow: 0.5–1 Å, yellow-orange: 1–1.5 Å, orange: 1.5–2 Å, orange-red: 2–4 Å, and red: >4 Å). (B) Average relative side chain accessibility: the side chain solvent accessible surface of each residue was calculated as percentage of the solvent accessible surface the same residue would have in the context of a poly-Ala peptide in extended conformation using the program NACCESS (<http://wolf.bms.umist.ac.uk/naccess/>) and converted to a color code (yellow: 0–10% relative solvent exposure; yellow-green: 10–25%; green: 25–50%; green-blue: 50–75%; blue: 75–100% and dark blue: >100%, fully solvent exposed). (C) Side chains contributing to the heterodimer interface between V_L and V_H: average reduction of the side chain accessible surface of each residue in the Fv fragment expressed as % of its accessible surface in the isolated V_L or V_H domain (white: 0% reduction of the solvent accessible surface area, yellow: 0–20%, yellow-orange: 20–40%, orange: 40–60%, red-orange: 60–80%, and red: 80–100%). (D) Side chains contributing to the interface between V_L and C_L or between V_H and C_H: average relative reduction of the side chain accessible surface of each residue in the Fab fragment compared to its accessible surface in the Fv fragment (white: 0% reduction, yellow: 0–20%, yellow-orange: 20–40%, orange: 40–60%, red-orange: 60–80%, and red: 80–100%). (E–G) Reduction of the side chain accessible surface upon formation of the complex of the Fv fragment with an antigen: side chains contributing to antigen binding in hapten binding antibodies (E), oligomer binders (F), and protein binders (G) (white: 0% reduction, yellow: 0–20%, yellow-orange: 20–40%, orange: 40–60%, red-orange: 60–80%, and red: 80–100%).

V_H domain, followed by the hV_{H1a} , hV_{H1b} , and hV_{H5} consensus domains with intermediate stabilities and little or no tendency to aggregate. hV_{H2} , hV_{H4} , and hV_{H6} domains, on the other hand, exhibited low cooperativity during denaturant-induced unfolding, lower production yields, and higher aggregation tendencies. The detailed analysis of hydrophobic core packing, hydrophobic contrast between surface and core, secondary structure propensity, and formation of salt bridges revealed that the V_{H3} domain had always found a solution which appeared better by all structural criteria, while the other V_H domains had some shortcomings, explaining the higher thermodynamic stability of hV_{H3} . Stable V_H domains derived from other species such as the camel V_{HH} domains [28,29], and particularly stable murine V_H domains share the main structural features which distinguish the human V_{H3} family from other human V_H clades. With the help of a sequence alignment grouped into V_H domains with favorable properties (families 1, 3, and 5) and unfavorable properties (families 2, 4, and 6), residues of the even-numbered V_H domains were identified and structurally analyzed which potentially decrease the folding efficiency and stability, leading to the observed unfavorable properties [27].

The biophysical properties of V_L domains differ to a smaller extent than those of V_H domains. In general, isolated V_L domains showed a higher thermodynamic stability and a higher yield of protein expressed in soluble form than isolated V_L domains. Interestingly, scFv with V_L domains (unstable as isolated domains) had very high thermodynamic stabilities as the assembled molecule. This was strongly dependent on the presence of a V_L -typical CDR-L3, which indicates that this high stability is caused by favorable V_H - V_L interface interactions in the scFv fragment.

1.5. Focus of this review

In this article, we describe and discuss two rational engineering methods to improve the biophysical properties of a given antibody fragment to enable its *in vivo* use as an intrabody. In the first method, known as CDR or loop grafting, the antigen-binding loops of a donor

framework are transferred to an acceptor framework with known favorable properties. The second method makes use of the large amount of data available concerning antibody domains to rationally engineer the antibody sequence. The approach takes into account sequence comparisons to the respective family-specific consensus sequences as well as the knowledge of the biophysical properties of the family consensus domains. These methods can be used alone, in combination, or in conjunction with evolutionary approaches for stability and folding efficiency, which have been reviewed elsewhere [19,20].

2. Description of methods

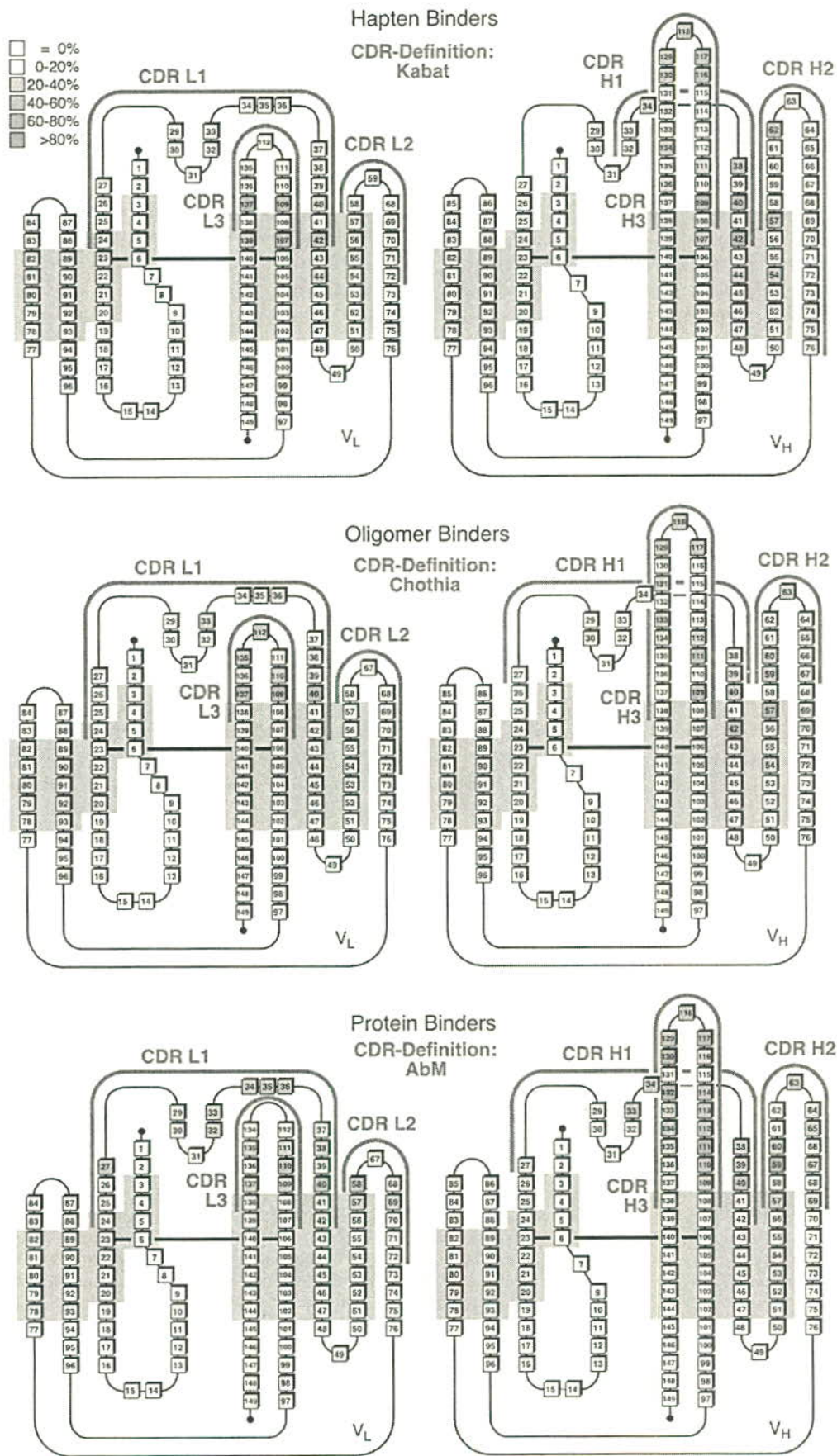
2.1. Loop grafting

2.1.1. Introduction

We will use the AHO nomenclature and numbering scheme [30] throughout, as it facilitates a structure-based discussion and, in particular, permits comparisons between the light and heavy chains, where structurally identical residues have the same number. The correspondence to the Kabat nomenclature can be found in [30] and Fig. 2.

We will first discuss the transfer of the antigen binding specificity from its original, poorly behaved framework to an antibody framework with superior properties (CDR graft, [31]). CDR grafts have primarily been introduced to reduce the immunogenicity of murine antibodies used in human *in vivo* applications. Most commonly, the human framework used as graft acceptor is the one whose sequence most closely resembles the original mouse sequence. However, the different antibody frameworks within a species show widely different intrinsic stabilities, as shown by the results of the detailed characterization of the human germline family consensus frameworks [27]. While a CDR graft to a closely related framework may limit the problems posed by non-CDR residues directly or indirectly affecting the antigen affinity, poor performance of the grafted molecule can result from an insufficiently stable acceptor framework,

Fig. 3. Average contribution of individual residue positions to the antibody–antigen interface. To determine the contribution of each residue to the antigen binding site, the accessibility of each residue was calculated for the Fv–antigen complex and for the Fv fragment without the antigen. The relative change in accessibility was calculated as $100 \times (\text{accessible surface of the side chain in the Fv fragment alone} - \text{accessible surface in the complex}) / \text{accessible surface in the Fv fragment alone}$. This value was converted to the color code indicated in the top right of the figure. This calculation of the relative change of side chain accessibility upon ligand interaction gives a better idea of the sensitivity of each position to mutations than the absolute contact area contributed by each residue that would weigh a large aromatic residue much more strongly than a glycine residue. While the contact analysis shown in this figure was performed on liganded/unliganded pairs created by simply deleting the antigen coordinates from the liganded structure, for the analysis of antigen-induced conformational changes, pairs of liganded and unliganded structures were identified by sequence comparison of the experimental structures. For comparison, the residues included in the CDR definitions of Kabat et al. [34], Chothia and Lesk [35], and those used by the AbM antibody modeling program (<http://antibody.bath.ac.uk/>) are indicated by blue lines. Note that the CDR definitions are totally independent of the type of antigen bound, and of course each CDR definition applies to all three types of antigens. Each type of antigen complex is shown with only one of the CDR definitions to save space.



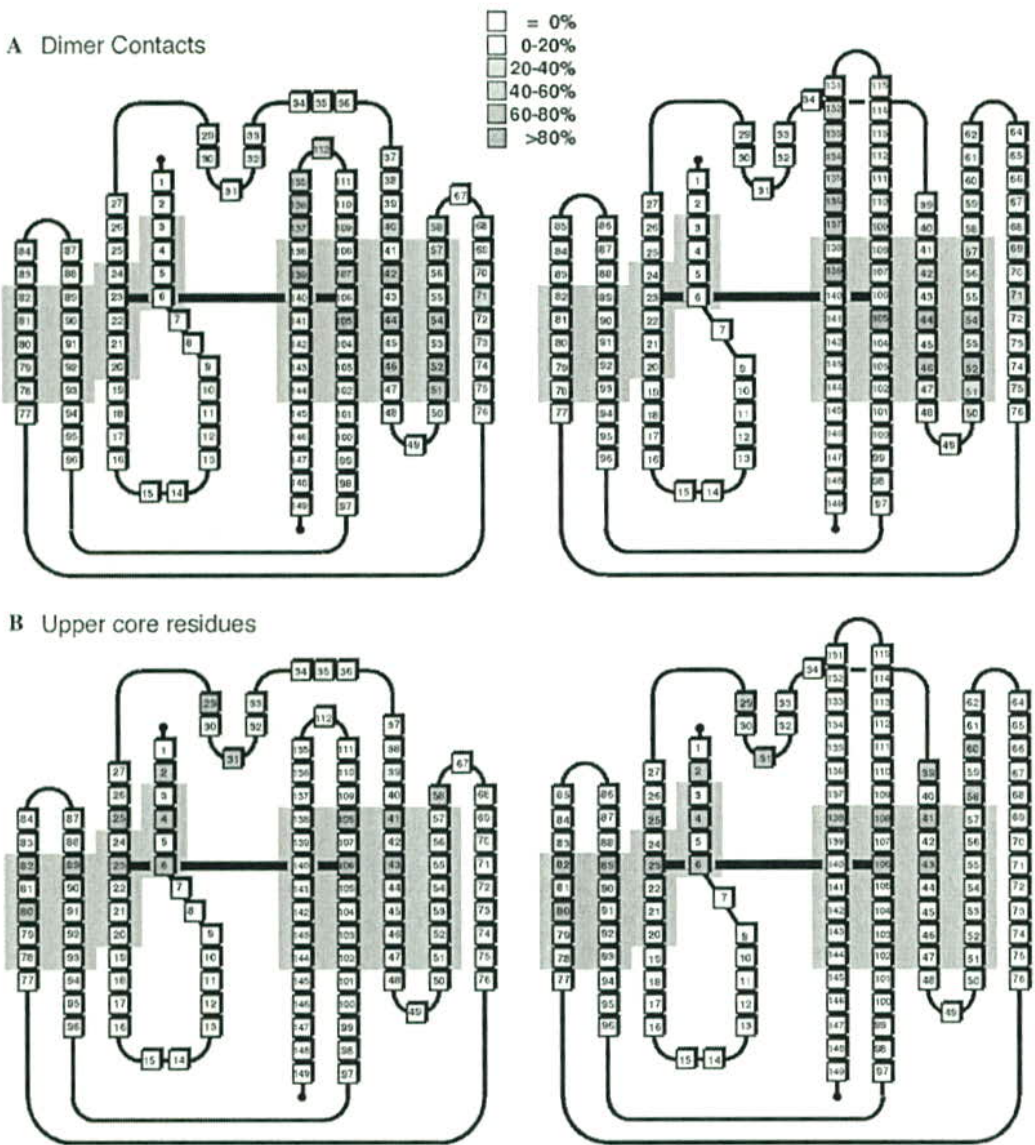


Fig. 4. Additional residues which need to be considered in a CDR graft. (A) Dimer interface residues: While the major dimer contact residues (indicated in red and red-orange) outside the CDR ranges normally are highly conserved, mutations in these residues can have a strong effect on antigen binding even if the residues make no contacts to the antigen. Gray areas underlie the residues which were used for least-squares superposition of the structures (3–7, 20–24, 41–47, 51–57, 78–82, 89–93, 102–108, and 138–144). (B) Packing of the core residues of the upper core (magenta) of the domains indirectly affects CDR conformation. The highly conserved central core residues (blue) shield the upper core from being affected by changes in the lower core packing and therefore make loop grafts to frameworks of a different structural subtype possible.

which is more likely to deform under the strain introduced by the grafted sequences.

The same technique of CDR grafting can be used to improve the biophysical properties of antibodies by grafting their antigen specificities to a framework with better biophysical properties [21,32,33]. In this case, the target framework often is not one closely related to that of the CDR donor, since the original framework showed, after all, inadequate biophysical properties. Therefore, particular care has to be taken to transfer not only contact residues, but all residues likely to directly or indirectly influence antigen binding. In general, simply combining the CDR sequences from one antibody with

the framework sequence of a second one is not sufficient to retain antigen recognition.

The schematic representations in Figs. 3 and 4 highlight the residues which have to be considered in the planning of a loop graft: not only the residues falling within the classical CDR definition, but also potential contact residues outside the CDRs (Fig. 3), dimer interface residues, whose substitution may affect the relative orientation of the V_L and V_H domain (Fig. 4A), and residues buried in the upper core of the domain may affect the CDR conformations (Fig. 4B). The key to success is to identify the residues to retain from the loop donor.

111

2.1.2. Method

2.1.2.1. Potential antigen contact residues. Analysis of the contact residues involved in the many antibody–antigen complex structures currently available in the PDB indicates that residues outside the classical CDR definitions [34,35] can be involved in direct antigen contacts (see, e.g., Figs. 1 and 3). Depending on the type of antigen (hapten, linear oligomers such as peptides, oligosaccharides and oligonucleotides or folded protein), different binding modes apply and different, though overlapping, sets of residues are used for contact. It is preferable to determine the residues to be retained from the loop donor according to their potential contribution to antigen binding, depending on the type of antigen, rather than according to some fixed CDR definition. For protein binders, residues at the N-termini of the domains, the outer half of the CDR-1 loop, and the outer loops residues 82–89 may be directly involved in antigen contact, while for hapten and peptide binders, residues usually deeply buried in the V_L/V_H dimer interface and inaccessible to protein antigens (e.g. residues 44, 54, 107) are accessible to antigen contact (Figs. 1 and 3).

2.1.2.2. Residues contributing to the upper core. A layer of invariant residues (Cys 23, Cys 106, Trp 43, and Gln/Glu 6) divides the core of the immunoglobulin variable domain into an upper core, consisting of residues buried directly underneath the CDRs, whose packing can strongly affect antigen affinity, and a lower core, whose packing correlates with the framework subtype, but has little or no influence on antigen binding (Figs. 4B and 5). Upper core residue 31, which intercalates between the two β sheets of the immunoglobulin domain and divides the CDR-1 loop into an outer and an inner loop, is probably the prime mediator translating changes of upper core packing into changes of CDR-1 conformation (Fig. 4B [30]).

In V_κ , residues L2 and L4 of the N-terminus, residues L25, L29, L31, and L41 of CDR-L1, residues L58 of CDR-L2, residues L80, L82, and L89 of the outer loop, and residues L108 of CDR-L3 pack together to form the upper core of the domain. Some of the positions buried in V_κ are exposed in V_λ , due to the less ordered N-terminus and the different CDR-L1 conformations in lambda light chains. The upper core of the lambda domains is formed by residues L4 of the N-term, residues L25, L31, and L41 of CDR-L1, residue L58 of CDR-L2, residues L80, L82, and L89 of the outer loop and residues L108 and L138 of CDR-L3. In V_H , residues H2 and H4 of the N-term, residues H25, H29, H31, H39, and H41 of CDR-H1, residues H58 and H60 of CDR-H2, residues H80, H82, and H89 of the outer loop, and residues H108 and H138 of CDR-H3 pack together to form the upper core of the domain. While packing interactions in the lower core of the domain could conceivably affect CDR-H2

orientation, and while there exists a correlation between germline family, structural subtype, identity of these lower core residues, and CDR-H2 length and conformation in natural antibody domains, deliberate loop grafts to distantly related frameworks [33] showed no loss of binding affinity due to lower core mismatch.

2.1.2.3. N-terminal residues. Since the residues of the upper core interact, any substitution of one of these residues may affect the position and conformation of the others. Although in oligomer and protein binders, N-terminal residues 1 and 2 occasionally contribute to antigen contacts, a variety of N-terminal extensions, such as affinity tags, the scFv linker and sequence changes towards an overall consensus introduced to accommodate restriction sites for cloning only rarely affect antigen binding. Due to the flexibility of the N-terminal strand, conservative substitution of the core residues 2 and 4 is relatively well tolerated while non-conservative substitutions of residues 2 and 4 should be avoided. The molecule is more likely to adapt to a substitution which would cause steric problems or unfavorable interactions by adapting the main-chain conformation of the N-terminal residues than by adapting CDR-1 conformation.

In V_H domains, substitutions of residue H6 can have a drastic effect on stability and potentially on antigen affinity as it determines the conformation of framework 1 [36,37]. Residue 6 is an extremely conserved buried glutamine in light chains and can be either glutamine or glutamate in heavy chains. Since it is the focus of a network of buried hydrogen bonds linking the outer and inner β -sheets, involving the β -bulge in framework 4 and conserved Thr 143, its substitution by any other amino acids leaves several broken hydrogen bonds. Various authors have linked problems with the functionality of monoclonal antibody-derived scFv fragments to primer-induced Gln to Glu exchanges in position H6 (reviewed in [36,37]). Out-of-context mutations leading to an unusual combination of residues in positions H6–H10 can create more of problems than just the concerted shift to a different canonical framework type, since these can completely unravel strand A and result in unusual, strand-swapped dimers [37], while at least in the examples we studied, deliberate grafts to different well-folding framework subtypes did not lead to excessive loss of antigen binding affinity [33]. While out-of context Glu H6 to Gln substitutions seem to be less destabilizing than Gln to Glu mutations, the substitution still can lead to an unusual kink conformation, as the X-ray structure of the anti-peptide antibody anti-GCN4 demonstrates [38]; in this example, the out-of-context mutation was acquired during ribosome-display affinity maturation and improved antigen binding affinity.

2.1.2.4. Outer loop. Upper core residues L80, L82, and L89 of the outer loop of V_L domains and residues and

the corresponding residues H80, H82, and H89 in V_H are the most frequent culprits of failed CDR grafts (Fig. 4B). These residues not only affect the conformation of the outer loop, but also pack against core residue 31 in CDR-1, which intercalates between the two β -sheets of the immunoglobulin domain and divides CDR-1 into two loops. Pushing against these residues, the side chains of residues 80, 82, and 89 can affect CDR-1 conformation. While the outer loop residues are highly conserved in V_k domains, forming a “GSGSGT” motif, the outer loop sequences and conformations of V_L and V_H are more variable. The second Gly (L82) in the V_k outer loop motif has a positive Φ angle, and the loop responds to mutation to a non-Gly residue with a marked conformational change of the loop, forcing the outer loop to jut away from the domain core.

Upper core residues L108 and L138 are within CDR-L3 according to any of the CDR definitions, therefore included in classical loop grafts. Although H108 is just outside CDR-H3, the presence of a salt bridge between Arg H108 and Asp H137, present in a majority of the V_H domains, predisposes CDR-H3 to a “closed” conformation, packing on top of the hapten binding pocket, while a missing salt bridge facilitates the “open” conformation. Mutation of these two residues can strongly affect conformation and flexibility of the CDR-H3 loop.

2.1.2.5. Dimer interface residues. While most of the framework residues that strongly contribute to the dimer interface are highly conserved, unusual dimer interface residues in the loop donor may have to be retained to achieve full activity in a CDR graft. Differences of the relative orientation of V_L and V_H by up to 15° in different antibodies have been reported, and considerable differences in the domain orientation between free and liganded structures of the same antibody demonstrate the flexibility of this interface. While most of the key dimer interface residues are highly conserved (Fig. 4A), mutations of these residues can have a significant impact on antigen binding. This is of special importance for grafts involving murine V_L chains. These show several nonconservative substitutions in the dimer interface compared to even the closest human homologue (L44 is predominantly Tyr in human V_L , but is changed to Val in m V_L , L46: Gln to Glu, L49: Gly to Asp, L51: predominantly Ala to Leu, L52: Pro to Phe and L54: predominantly large hydrophobic residues to Gly). These exchanges between human and murine V_L can only be accommodated if the relative orientation of the domains adapts.

2.1.3. Method

1. Align the V_L and V_H sequences of the specificity donor and of the framework donor to each other and to the header of the alignment shown in Fig. 2, making sure to center the gaps as specified in [30]. A Microsoft

EXCEL worksheet containing this header is available for download on the AAAAA website (<http://www.biochem.unizh.ch/antibody>), together with a number of EXCEL visual basic macros facilitating the import and analysis of sequences in EXCEL.

2. Compare the sequence positions color coded as antigen contact residues for the type of antigen recognized by your antibody (header lines E: hapten, F: oligomer or G: protein in Fig. 2; Fig. 3). CDR length and sequence as well as any additional positions marked as potential antigen contacts should reflect the sequence of the specificity donor.
3. Check adjacent positions. If they contain structurally critical substitutions (Gly, Pro) in either loop donor or framework donor, also change them to the sequence of the specificity donor.
4. Change all residues of the upper core to the sequence of the specificity donor.
5. Check V_L/V_H heterodimer interface residues. While conservative substitutions in the peripheral contact residues indicated in yellow (Figs. 2 and 4A) should not affect antigen binding affinity, major dimer interface residues indicated in orange and red and non-conservative substitutions of peripheral contact residues should be changed to the sequence of the specificity donor.
6. All other positions should reflect the sequence of the framework donor.

2.1.4. Examples

The work of Jung et al. [21,32] presents an example of the improvement of stability and folding efficiency achieved by a loop graft from a murine framework to a human framework of the same structural subtype (m V_k 1 to h V_k 1 and m V_H 6 to h V_H 3), and describes the results of experiments to further improve the stability of the grafted scFv by directed in vitro evolution using phage display.

In another example, an anti-EpCAM antibody, used in tumor targeting, needed to be humanized and stabilized at the same time [33]. In this work, the CDRs of the murine scFv Moc31 were grafted to a framework of a different V_H subtype (m V_H 9 to h V_H 3), while the light chains were more similar (m V_k 2 to h V_k 1). While a classical loop graft indeed increased the stability of the scFv, an even more stable scFv was obtained by combining the hydrophobic lower core and the subtype-determining residues H6, H7, and H10 of the original murine V_H domain with the framework residues of h V_H 3. In this approach, the murine V_H domain was only resurfaced, preserving the original structural subtype.

Finally, an anti-GCN4 antibody to be used as an intrabody was stabilized by loop grafting [17] to an extremely stable, but highly divergent, framework. This work demonstrated that loop grafts to even the most highly divergent antibody frameworks are feasible, even

loop grafts from a murine lambda to a human kappa framework (mV_L1 to hV_K1) and between V_H domains representing different structural subtypes (mV_H2: Type I to mV_H4: Type II). At the same time the example demonstrates the importance of preserving dimer interface residues not involved in antigen contact in order to preserve the antigen binding affinity of the scFv.

2.2. Optimization of key residue interactions

2.2.1. Modified consensus approach

In vivo, in the process of affinity maturation the whole of the domain encoding sequence is subjected to random, independent mutations, and B-cells producing improved antigen receptors are selected and propagated by the immune system [39]. Even though selection for antigen binding plays a dominant role for any specific sequence, other factors including domain stability, assembly, and interaction between heavy and light chains or variable and constant domains, protease resistance, and the ability for export and secretion must have an influence on each position in the sequence of antibody domains.

A comparatively simple concept to improve the biophysical properties of an antibody is the consensus approach [40]. The underlying hypothesis is that destabilizing mutations are highly probable but are selectively neutral as long as the overall domain stability does not fall below a certain threshold. Conversely, random mutations resulting in increased thermodynamic stability are highly improbable in the absence of a positive selection. Consequently, the most frequent amino acid at any position in an alignment of homologous immunoglobulin variable domains is assumed to contribute most to the stability of the protein domain, and a mutation towards the consensus would be expected to be stabilizing. This consensus method has been applied to a V_K domain and 10 mutations were proposed of which six increased the stability [40], and in this case the consensus sequence was built from all light chains.

However, this approach is based on the assumption that the antibody variable domain sequence diversity is based on independent random fluctuations around a single optimum. All frameworks of a given type (V_L, V_K or V_H) were averaged to a single “ideal” sequence. Since the antibody variable domains are a protein family with significant diversity, rather than point mutants of a single sequence, the key question to be discussed is from which part of the antibody sequence space the consensus should be built. The various germline families and structural subtypes differ from each other by groups of concerted sequence changes, presumably representing discrete optima which play an important role in antibody diversity. Different CDR-L lengths are derived from different germline families and require concerted adaptations of upper core packing and placement of turn-forming residue, e.g., a highly conserved Gly in L37

of CDR-L1 loops of length 17, not present in CDR-L1 loops of length 18, or a hydrophobic residue in L32 of long CDR-L3 replaced by predominantly hydrophilic residues in shorter CDR-L1s. Different structural subtypes of V_H domains carry mutually incompatible framework residues, which cannot simply be exchanged with those of other frameworks. It follows that family specific solutions are needed to create a variety of different frameworks with superior properties.

The new concept of the following method is to compare the sequence of a given antibody not only to the global consensus, but also to the specific subset of immunoglobulin sequences representing the structural specialization of different clades. The degree of conservation of any sequence position in the different clades can provide important clues towards the identification of important structural motives and residue interactions. It is useful to understand the structural rationale of the mutations in order to decide whether to ignore a specific deviation from the consensus or to mutate to either the general or to the context-dependent consensus. Even though a comprehensive account, based on the evaluation of a large number of mutations, is beyond the scope of the present review and will be presented elsewhere (Honegger et al., in preparation), a summary of the types of exchanges found will be of considerable help in understanding these approaches.

2.2.1.1. Core packing. Globular proteins are characterized by a well-packed hydrophobic core which plays an important role in protein folding and stability [41,42]. In antibody variable domains, a layer of invariant residues (Gln/Glu 6, Cys 23 and 106, Trp 43, and Thr 143) divide the hydrophobic core of the immunoglobulin variable domains into an upper core, whose packing affects the CDR conformation, and a lower core (see CDR grafting, above) (Fig. 5). Comparing the core packing in different antibody germline families, it is found that while a few of the core residues are highly conserved, in V_H domains most of the core packing varies in germline family-dependent and structural subtype-correlated manner (Fig. 6). This variation in core packing is characterized by complementary exchanges, a mutation to a smaller side chain in one position being compensated by a mutation to a larger residue in a second position, either in direct contact with the first or mediated through side chain conformational changes of intervening residues [36,43,44]. Therefore, individual residues can often not simply be transplanted from one framework to another. This is one of the main reasons why it is more reasonable to derive a subgroup-specific consensus sequence, rather than averaging over all families.

2.2.1.2. Hydrophobic surface residues. In principle, a fully solvent exposed hydrophobic residue should not affect protein stability either way, as it would be equally

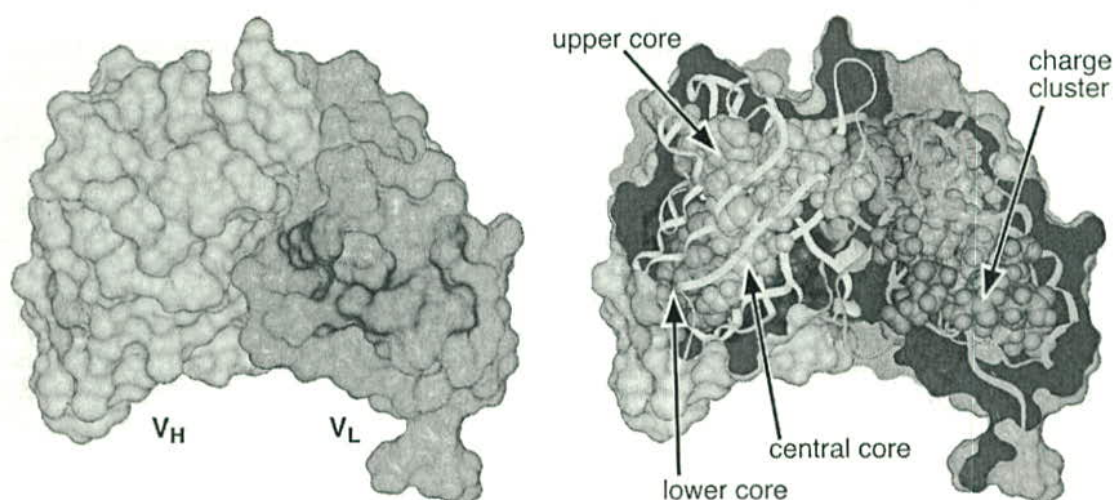


Fig. 5. Core residues of antibody variable domains. A layer of invariant residues (Gln/Glu 6, Cys 23 and 106, Trp 43, and Thr 143, shown in yellow) divides the hydrophobic core of the immunoglobulin variable domains into an upper core (green), whose packing affects the CDR conformation, and a lower core (orange). Buried and semiburied charged and polar residues surrounding the buried salt bridge between Arg/Lys 77 and Asp 100 form a charge cluster (red) that also contributes to the lower core of the domains. Surface residues are shown in blue (V_H) and magenta (V_L).

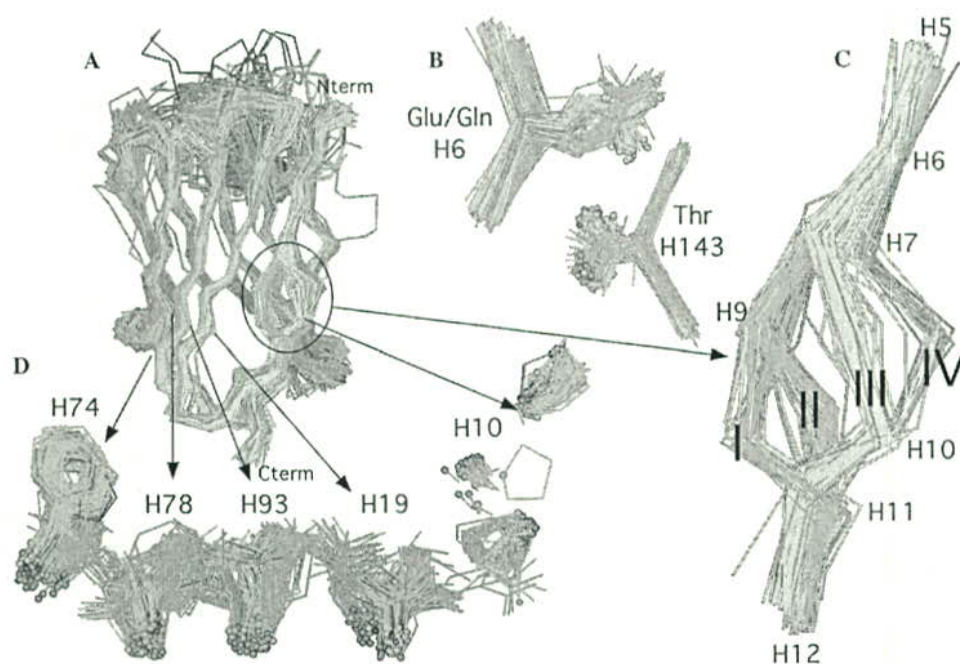


Fig. 6. Residues H6, H7, and H10 determine the framework 1 conformation of V_H domains. Correlation between the predicted structural subtype and the observed structure: 205 V_H domain structures were extracted from antibody Fab and Fv X-ray structures taken from the PDB (<http://www.rcsb.org/pdb/>). They represent 118 non-redundant structures (≥ 5 amino acids sequence difference between any pair of sequences). The domain structures were aligned by least squares superposition of the structurally least variable C α -positions in the V_H domain (H3–H6, H20–H24, H41–H47, H51–H57, H78–H82, H89–H93, H102–H108, and H138–H144). The molecules were color-coded according to the structural subtype predicted from the identities of the amino acids in positions H6, H7, and H10: Type I (magenta): H6 = Glu, H7 $\neq$ Pro, and H10 = Pro; Type II (pink): H6 = Glu, H7 $\neq$ Pro, and H10 = Gly; Type III (cyan): H6 = Gln, H7 $\neq$ Pro, and H10 = any amino acid; and Type IV (blue): H6 = Gln, H7 = Pro, and H10 = any amino acid. The figure was generated with InsightII (MSI/Biosym, San Diego, USA). In antibody fragments crystallized under mildly acidic conditions, Glu H6 can be protonated and show Gln-like hydrogen bonding [37]. (A) C α trace of entire V_H domain (B) Glu/Gln H6 and Thr H143 side chain conformation. (C) Framework 1 main-chain conformation. (D) Correlated structure and sequence changes across the V_H domain, first described by Saul and Poljak [43,44] who noticed the correlation between framework 1 structure and the types of amino acid found in position H10 (Kabat H9): Gly (as in Type II structures) correlated with germline clan hV_H3 and Ala/Ser (as in Types III and IV structures), correlated with germline class hV_H1. This in turn correlates with the nature of residues H74 and H78, and with the side chain orientation of the intervening residues H19 and H93. Type I structures, represented in human germline family hV_H2 and hV_H4, form yet a third class. Type III structures represented in hV_H1, hV_H5, and hV_H6 and Type IV structures behave the same with respect to this correlation. Type IV V_H domains, requiring a Pro in position H7, are not represented by any human V_H germline sequence, but make up about half of the murine V_H 1 germline sequences.

exposed in the native and the denatured state. In reality, however, hydrophobic surface residues frequently have a marked negative effect on the folding efficiency and aggregation of a protein during the folding process, i.e., the yield [45], since they provide opportunities for the formation of misfolded conformations or aggregates, there those residues can assume a buried position. The effects of such mutations are highly position dependent, possibly due to the destabilization of misfolded structures or aggregation interfaces, rather than on any effect on the equilibrium stability of the native or the unfolded state. This makes it very difficult to predict the effects of these mutations, since they predominantly affect structural states which cannot be observed directly. Since semi-exposed hydrophobic residues can contribute to domain stability by their interactions with neighboring amino acids, a trade-off between stability and folding efficiency can often be observed.

2.2.1.3. Buried charged and hydrophilic side chains. Buried charges are considered energetically unfavorable due to the high energetic cost of desolvating the charged groups. The energetic contribution from favorable electrostatic interactions of most buried salt bridges is not sufficient to compensate for the energetic cost of desolvating the charged groups [46–48]. While the buried hydrophilic side chains may not improve the equilibrium stability of the folded state compared to a well-packed hydrophobic core, they may have a decisive influence on the probability of the molecule eventually reaching that state. They destabilize alternative folded states, which otherwise could lead to a situation like that observed in a molten globule: large populations of somewhat differently packed folded states with very similar energies. Again, a subtle trade-off between stability and folding efficiency must be found.

2.2.1.4. Glu/GlnH6 and the determination of framework I class. Position 6 of the sequence of immunoglobulin variable domains is almost exclusively either a Gln or a Glu, which is fully buried (Fig. 6). The side chain carbonyl group of Glu/Gln in position 6 is within hydrogen bonding distance of three different main-chain amide hydrogens, the NH groups of Cys 106, of residue 141 and of Gly 142, and the presence of either Gln or Glu in heavy chains, as well as the conformational preferences of some of the neighboring residues have important consequences for the structure of variable domains [36,37].

The side chain orientation of residue 6 has to adapt to a conformation which allows the residue to satisfy its hydrogen bonding requirements (Fig. 6B). As a consequence, the main-chain conformation adapts to subtle changes in its main-chain torsion angle. This is not very obvious for the main-chain conformation of H6 itself, but it affects the main-chain conformation of the residues following H6 (the so-called framework I region), which

form the kink between strands a and a' (Figs. 6 and 7). While the residues preceding H6 are locked up in strand a, belonging to the outer β -sheet of the domain, and cannot easily adapt their conformation, the shape of the kink is clearly correlated to the nature of the H6 residue [36] (Figs. 6A and C). If H6 is a glutamine, the H7 torsion angle lies in the range allowed for a proline, independent of the amino acid actually occupying this position. If H6 is a glutamate, H7 torsion angles are in a range disallowed for Pro. As a consequence, the effect of Gln H6 to Glu mutations is particularly severe for V_H domains with a Pro in position H7 [36,37]. In addition, a Glu in H6 requires either a Pro (Type I) or a Gly (Type II) in position H10, while Gln H6 is more permissive, tolerating any amino acid in that position.

Out-of-contact mutations, frequently introduced by the degenerate primers used to clone and amplify antibody variable domains and leading to an unusual combination of residues in H6, H7, and H10, can severely destabilize the V_H domain. If the domain cannot adapt without altering the conformation of the chain N-terminal to H6, in extreme cases completely unraveling strand a [37] and thus displacing upper core residues H2 and H4, these mutations also affect antigen binding. Due to the interaction of framework I with the domain core (Fig. 6D), even the concerted exchange of all three residues can destabilize the domain. An important consequence is that the framework I class should not be mixed or changed by mutations.

In V_L domains, different sequence lengths and the positioning of conserved proline residues determine the exact shape of the FR1 kink [30].

2.2.1.5. The Arg 77/Asp 100 charge cluster. Arg 77 and Asp 100 form a highly conserved buried salt bridge in the lower core of both the V_L and the V_H domains (Fig. 7). The absence of this interaction in light chains correlates with amyloid fibril formation [49]. In V_H domains, replacement of Arg H77 by Lys, as found in many murine V_H domains, already leads to a significant loss of stability [17,50]. Replacement by an uncharged residue, as in hV_H5, should have an even stronger effect. Surrounding polar and charged groups form a charge cluster around this central salt bridge and affect the degree of order and definition of the hydrogen bonds connecting the residues contributing to this charge cluster. This is present to the fullest extent only in V_H 3 domains [27].

2.2.1.6. Conserved proline and glycine residues. Gly, due to its exceptionally high flexibility, and Pro, with its very restricted torsional freedom, can have a strong effect both on folding efficiency and on stability. Highly conserved glycine positions (Fig. 8) often are positions with main-chain torsion angles disallowed for other amino acids, such as with positive Φ angles, frequently located in turns. Another possible reason for the conservation of a Gly

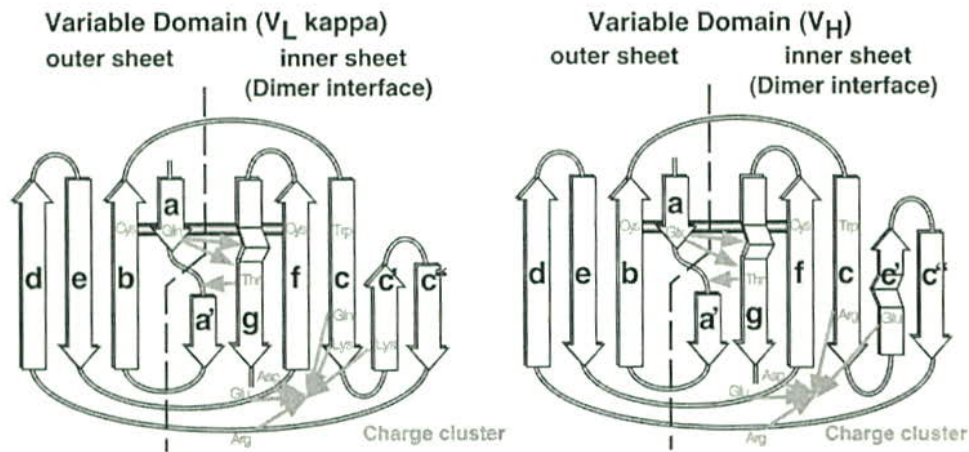


Fig. 7. Schematic representation of the secondary structures of immunoglobulin domains with the charge cluster indicated. The antigen binding site with the CDRs would be on the top, CDR-1 linking strands b to c, CDR-2 strands c' and c'' and CDR-3 strands f and g. The disulfide bond (res. 23 and 106) links strands b and f. Red arrows indicate conserved hydrogen bonding and charge interactions of buried hydrophilic side chains which have a major effect on stability and conformation: Gln/Glu in position 6, Thr 143, and the cluster of charged and hydrophilic residues interacting with the conserved salt bridge between positions 77 and 100, also involving residues 45, 47, 53, 97, 99, and 104.

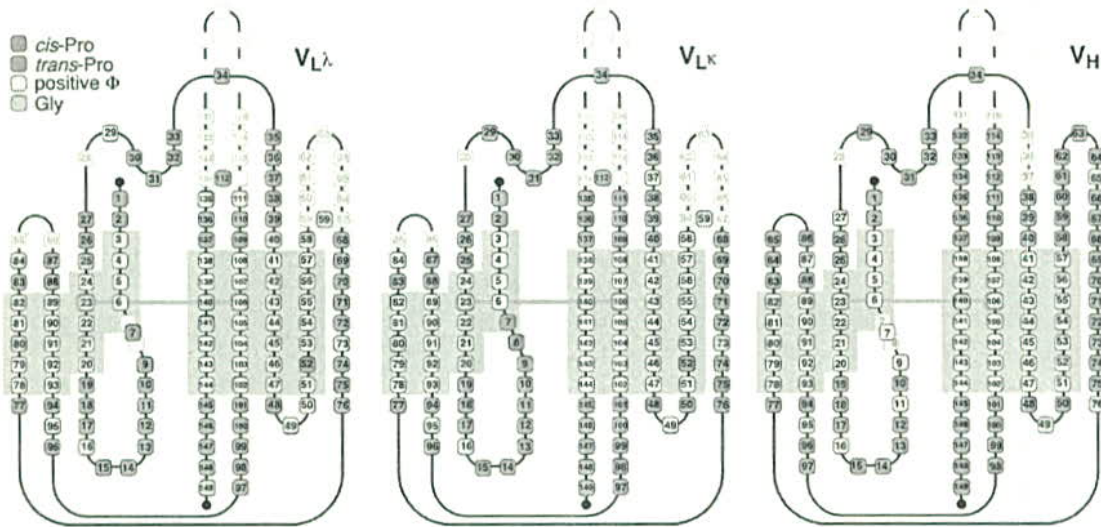


Fig. 8. Positions with conserved positive Φ angles, conserved Gly, *cis*- and *trans*-Pro Orange: conserved *trans*-Pro, red: *cis*-Pro, yellow: positions with positive Φ torsion angles, and green: additional conserved Gly positions.

residue can be the spatial constraints in buried positions, where larger side chain would lead to steric clashes.

Prolines are frequently located at position 2 of solvent exposed β -turns or in coil regions of loops, where they can have a significant stabilizing effect [51]. Proline residues are much more likely (5.7%) to assume a *cis*-peptide bond conformation than any other amino acid. Apart from one or two conserved *cis*-Pro in V_{κ} , there also are a fair number of conserved *trans*-prolines in the immunoglobulin variable domains, as well as some positions where Pro occurs in some germline sequences, but has a negative effect on folding. Again, the effects of mutating these residues are strongly context dependent.

2.2.1.7. Secondary structure propensities. Zhu and Blundell [52] analyzed the occurrence of amino acids in specific

positions in a β -sheet and gave propensities for each amino acid. These propensities reflect the ideal β -sheet composition in analogy to the consensus approach [40]. The exchange of a residue in a β -sheet with a low β -sheet propensity to a residue with high β -sheet propensity improves therefore the biophysical properties of an immunoglobulin.

2.2.2. Method

1. Identify the germline family association of your V_L and V_H sequence by comparing them to the different germline family consensus sequences.
2. Align the V_L and V_H sequence of your antibody to the global consensus, the germline family consensus, and the header of the alignment shown in Fig. 2, making sure to center the gaps as specified in [30]. An EXCEL

Table 1
Determination of framework-I class

Subtype	Subclass-defining residues ^a				Subclass-correlated core residues ^a			
	H6 ^b	H7	H9	H10	H19	H74	H78	H93
I	Glu	Not Pro	Gly	Pro	Leu	Leu	Ala/Val/Ile/Leu	Leu/Met
II	Glu	Not Pro	Gly	Gly	Leu	Val	Phe	Met
III	Gln	Not Pro	Gly	any	Leu/Val	Phe	Ala/Val	Leu
IV	Gln	Pro	Gly	any	Leu/Val	Phe	Ala/Val	Leu

^a According to [36].

^b Using the numbering scheme of Honegger and Plückthun [30].

worksheet containing this header and the human and murine germline family consensus sequences is available for download on the AAAAA website (<http://www.biochem.unizh.ch/antibody>), together with a number of EXCEL visual basic macros facilitating the import and analysis of sequences in EXCEL.

- Any position which should retain the amino acid type of the specificity donor in a loop graft (as described above) should also be retained in this method. These positions are to be excluded from the subsequent analysis.
- The subtype-determining residues at position H6, H7, and H10 [36] and the correlated core positions 19, 74, 78, and 93 should match those of the germline family consensus, except where this violates the rules outlined in Table 1.
- Line B of the header in Fig. 2 indicates the average solvent exposure of each position. Hydrophobic framework residues with a relative solvent exposure >75% (indicated in blue) should be replaced by hydrophilic ones, with the exception of Pro residues, which play an important structural role, and Val in positions L3 and H5, where the very high β -sheet propensity of Val outweighs the disadvantage of having a hydrophobic residue on the surface.
- Hydrophilic residues in fully buried positions (relative solvent exposure 0–25%, indicated by yellow and yellow-green in Fig. 2, line B) should be replaced by hydrophobic ones unless the global consensus indicates a highly conserved hydrophilic core residue. If the germline family consensus is hydrophobic, use this residue, otherwise use the hydrophilic residue closest in size to the one found in the germline family consensus. A Lys in position L13 is acceptable and does not need to be replaced, since its side chain amino group can reach the solvent.
- The key residues of the buried charge cluster, 77 and 100, should be Arg and Asp, independent of the consensus. The additional members of the charge cluster can be changed to the consensus of hV_K3, hV_L1, and hV_H3, respectively.
- Highly conserved Pro and Gly residues (global consensus, Fig. 8) should be conserved, even if not present in the individual sequence and the germline family consensus. Non-Gly residues in positions with conserved

positive Phi torsion angles (Fig. 8) should be replaced by Gly.

If possible, analyze the environment of the proposed mutations in a homology model of your antibody (if you are unable to build a homology model, have one built by submitting your sequence to the WAM antibody modeling website (<http://antibody.bath.ac.uk/index.html>), or use the structure of a closely related antibody (<http://www.rcsb.org/pdb>) or the human consensus Fv model representing the framework combination of your antibody (<http://www.biochem.unizh.ch/antibody>).

An important point is that it is not necessary to introduce these mutations one-by-one and investigate their contributions and their additivity. This has been done in the past in order to derive the above rules. However, for practical applications, we recommend to introduce a set of mutations all at once, making this into a very fast and practicable procedure.

2.2.3. Examples

This method was tested successfully with two V_H6 containing antibodies [53]. All six proposed mutations to the family consensus of V_H3 improved either the expression yield of soluble protein, the thermodynamic stability or even both biophysical properties. Combining all 6 mutations increased the expression yield by a factor 4 to that of V_H3 containing antibodies and increased the thermodynamic stability, measured by denaturant-induced equilibrium unfolding, by 20.9 kJ/mol, just about compensating for the loss of stability caused by the absence of the disulfide bond in the reducing environment of the cytoplasm.

In another example, the expression of an anti-peptide antibody was dramatically improved by a similar series of mutations [54]. It needs to be stressed again that all mutations were introduced at once, with a rather minimal experimental effort.

3. Concluding remarks

3.1. Future aspects

Both in the process of designing loop grafting and in the process of introducing individual and groups of

mutations, the antibody sequence needs to be checked continuously for consistency with structural requirements. This process greatly profits from the availability of tables of preferred and allowed residues at all positions, which are now becoming available (Honegger et al., unpublished). Furthermore, once the rules for the identification of such mutations have been formulated with sufficient precision, the process lends itself to automation (Honegger et al., unpublished).

We therefore see, for the mid-term future, three strategies for antibody improvement. First, rule-based engineering, including CDR grafting, can be used, as outlined above, to “rescue” antibodies with particularly valuable biological effects or recognition properties. Second, evolutionary approaches (an iteration of randomization and selection) can be used to further refine any antibody, with or without prior rule-based engineering. Third, this knowledge can directly be used in the design of improved versions of future libraries.

While this issue is devoted to intrabodies, it should be pointed out that favorable biophysical properties such as high thermodynamic stability, high yield of natively folded protein, and the absence of aggregation-prone behavior are not only a prerequisite for intracellular but also extracellular applications of recombinant antibodies, particularly in medical applications, where the fact that antibodies are a natural library of human proteins of generally low immunogenicity can be of great importance. The very application of intracellular binders is now receiving some competition by non-antibody frameworks: recently libraries of synthetic ankyrin repeats have been introduced which do not contain any cysteines, are equal in affinity to antibodies, and greatly surpass them in expression yields and stability [55,56]. These libraries have been designed following the principles outlined here.

References

- [1] A. Cattaneo, S. Biocca, *Trends Biotechnol.* 17 (1999) 115–121.
- [2] W.A. Marasco, *Gene Ther.* 4 (1997) 11–15.
- [3] G. De Jaeger, E. Fiers, D. Eeckhout, A. Depicker, *FEBS Lett.* 467 (2000) 316–320.
- [4] M. Visintin, E. Tse, H. Axelsson, T.H. Rabbitts, A. Cattaneo, *Proc. Natl. Acad. Sci. USA* 96 (1999) 11723–11728.
- [5] E. Mössner, H. Koch, A. Plückthun, *J. Mol. Biol.* 308 (2001) 115–122.
- [6] R.E. Bird, K.D. Hardman, J.W. Jacobson, S. Johnson, B.M. Kaufman, S.M. Lee, T. Lee, S.H. Pope, G.S. Riordan, M. Whitlow, *Science* 242 (1988) 423–426.
- [7] J.S. Huston, D. Levinson, M. Mudgett-Hunter, M.S. Tai, J. Novotny, M.N. Margolies, R.J. Ridge, R.E. Brucoleri, E. Haber, R. Crea, et al., *Proc. Natl. Acad. Sci. USA* 85 (1988) 5879–5883.
- [8] A. Wörn, A. Plückthun, *FEBS Lett.* 427 (1998) 357–361.
- [9] K. Proba, A. Wörn, A. Honegger, A. Plückthun, *J. Mol. Biol.* 275 (1998) 245–253.
- [10] C. Frisch, H. Kolmar, A. Schmidt, G. Kleemann, A. Reinhardt, E. Pohl, I. Uson, T.R. Schneider, H.J. Fritz, *Fold. Des.* 1 (1996) 431–440.
- [11] Y. Goto, K. Hamaguchi, *J. Biochem.* 86 (1979) 1433–1441.
- [12] E. Ohage, B. Steipe, *J. Mol. Biol.* 291 (1999) 1119–1128.
- [13] S. Biocca, F. Ruberti, M. Tafani, P. Pierandrei-Amaldi, A. Cattaneo, *Biotechnology (NY)* 13 (1995) 1110–1115.
- [14] A.M. Mhashilkar, J. Bagley, S.Y. Chen, A.M. Szilvay, D.G. Helland, W.A. Marasco, *EMBO J.* 14 (1995) 1542–1551.
- [15] A. Cattaneo, S. Biocca, *Trends Biotechnol.* 17 (1999) 115–121.
- [16] L. Duan, O. Bagasra, M.A. Laughlin, J.W. Oakes, R.J. Pomerantz, *Proc. Natl. Acad. Sci. USA* 91 (1994) 5075–5079.
- [17] A. Wörn, A. Auf der Maur, D. Escher, A. Honegger, A. Barberis, A. Plückthun, *J. Biol. Chem.* 275 (2000) 2795–2803.
- [18] A. Rajpal, T.G. Turi, *J. Biol. Chem.* 276 (2001) 33139–33146.
- [19] A. Wörn, A. Plückthun, *J. Mol. Biol.* 305 (2001) 989–1010.
- [20] P. Forrer, S. Jung, A. Plückthun, *Curr. Opin. Struct. Biol.* 9 (1999) 514–520.
- [21] S. Jung, A. Honegger, A. Plückthun, *J. Mol. Biol.* 294 (1999) 163–180.
- [22] V. Sieber, A. Plückthun, F.X. Schmid, *Nat. Biotechnol.* 16 (1998) 955–960.
- [23] J. Hanes, A. Plückthun, *Proc. Natl. Acad. Sci. USA* 94 (1997) 4937–4942.
- [24] L. Jermutus, A. Honegger, F. Schwesinger, J. Hanes, A. Plückthun, *Proc. Natl. Acad. Sci. USA* 98 (2001) 75–80.
- [25] M. Visintin, G. Settanni, A. Maritan, S. Graziosi, J.D. Marks, A. Cattaneo, *J. Mol. Biol.* 317 (2002) 73–83.
- [26] A. Knappik, L. Ge, A. Honegger, P. Pack, M. Fischer, G. Wellenhofer, A. Hoess, J. Wölle, A. Plückthun, B. Virnekäs, *J. Mol. Biol.* 296 (2000) 57–86.
- [27] S. Ewert, T. Huber, A. Honegger, A. Plückthun, *J. Mol. Biol.* 325 (2003) 531–553.
- [28] S. Ewert, C. Cambillau, K. Conrath, A. Plückthun, *Biochemistry* 41 (2002) 3628–3636.
- [29] M. Dumoulin, K. Conrath, A. Van Meirhaeghe, F. Meersman, K. Heremans, L.G. Frenken, S. Muyldermans, L. Wyns, A. Matagne, *Protein Sci.* 11 (2002) 500–515.
- [30] A. Honegger, A. Plückthun, *J. Mol. Biol.* 309 (2001) 657–670.
- [31] P.T. Jones, P.H. Dear, J. Foote, M.S. Neuberger, G. Winter, *Nature* 321 (1986) 522–525.
- [32] S. Jung, A. Plückthun, *Protein Eng.* 10 (1997) 959–966.
- [33] J. Willuda, A. Honegger, R. Waibel, P.A. Schubiger, R. Stahel, U. Zangemeister-Wittke, A. Plückthun, *Cancer Res.* 59 (1999) 5758–5767.
- [34] E.A. Kabat, T.T. Wu, H.M. Perry, K.S. Gottesmann, C. Foeller, *Sequences of Proteins of Immunological Interest*, NIH Publication No. 91-3242, National Technical Information Service (NTIS), 1991.
- [35] C. Chothia, A.M. Lesk, *J. Mol. Biol.* 196 (1987) 901–917.
- [36] A. Honegger, A. Plückthun, *J. Mol. Biol.* 309 (2001) 687–699.
- [37] S. Jung, S. Spinelli, B. Schimmele, A. Honegger, L. Pugliese, C. Cambillau, A. Plückthun, *J. Mol. Biol.* 309 (2001) 701–716.
- [38] C. Zahnd, S. Spinelli, P. Amstutz, C. Cambillau, A. Plückthun, *J. Biol. Chem.* 279 (2004) 18870–18877.
- [39] C. Berek, M. Ziegner, *Immunol. Today* 14 (1993) 400–404.
- [40] B. Steipe, B. Schiller, A. Plückthun, S. Steinbacher, *J. Mol. Biol.* 240 (1994) 188–192.
- [41] J.D. Bernal, *Nature* 143 (1939) 663–667.
- [42] D.O. Alonso, K.A. Dill, *Biochemistry* 30 (1991) 5974–5985.
- [43] F.A. Saul, R.J. Poljak, *J. Mol. Biol.* 230 (1993) 15–20.
- [44] F.A. Saul, *Res. Immunol.* 145 (1994) 61–66.
- [45] L. Nieba, A. Honegger, C. Krebber, A. Plückthun, *Protein Eng.* 10 (1997) 435–444.
- [46] J.H. Hurley, W.A. Baase, B.W. Matthews, *J. Mol. Biol.* 224 (1992) 1143–1159.
- [47] C.D. Waldburger, J.F. Schildbach, R.T. Sauer, *Nat. Struct. Biol.* 2 (1995) 122–128.
- [48] C.V. Sindelar, Z.S. Hendsch, B. Tidor, *Protein Sci.* 7 (1998) 1898–1914.
- [49] L.R. Helms, R. Wetzel, *J. Mol. Biol.* 257 (1996) 77–86.
- [50] K. Proba, A. Honegger, A. Plückthun, *J. Mol. Biol.* 265 (1997) 161–172.

- [51] K. Watanabe, T. Masuda, H. Ohashi, H. Mihara, Y. Suzuki, *Eur. J. Biochem.* 226 (1994) 277–283.
- [52] Z.Y. Zhu, T.L. Blundell, *J. Mol. Biol.* 260 (1996) 261–276.
- [53] S. Ewert, A. Honegger, A. Plückthun, *Biochemistry* 42 (2003) 1517–1528.
- [54] M. Kaufmann, P. Lindner, A. Honegger, K. Blank, M. Tschopp, G. Capitani, A. Plückthun, M.G. Grütter, *J. Mol. Biol.* 318 (2002) 135–147.
- [55] P. Forrer, M.T. Stumpp, H.K. Binz, A. Plückthun, *FEBS Lett.* 539 (2003) 2–6.
- [56] A. Kohl, H.K. Binz, P. Forrer, M.T. Stumpp, A. Plückthun, M.G. Grütter, *Proc. Natl. Acad. Sci. USA* 100 (2003) 1700–1705.
- [57] B. Gigant, J.B. Charbonnier, Z. Eshhar, B.S. Green, M. Knossow, *Proc. Natl. Acad. Sci. USA* 94 (1997) 7857–7861.
- [58] D.R. Bundle, H. Baumann, J.R. Brisson, S.M. Gagne, A. Zdanov, M. Cygler, *Biochemistry* 33 (1994) 5183–5192.