



(51) International Patent Classification:
C07K 16/00 (2006.01) *C12N 15/13* (2006.01)

(21) International Application Number:
PCT/EP2010/055440

(22) International Filing Date:
23 April 2010 (23.04.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09005776.1 24 April 2009 (24.04.2009) EP

(71) Applicant (for all designated States except US):
BOEHRINGER INGELHEIM INTERNATIONAL GMBH [DE/DE]; Binger Str. 173, 55216 Ingelheim Am Rhein (DE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BUCHNER, Johannes** [DE/DE]; Auf der Platte 15, 93346 Ihrlerstein (DE). **FEIGE, Matthias** [DE/DE]; Muellerstr. 26, 80469 Muenchen (DE). **MARCINOWSKI, Moritz** [DE/DE]; Koeniginstr. 37, 80539 Muenchen (DE).

(74) Agents: **HAMMANN, Heinz** et al.; c/o Boehringer Ingelheim GmbH, Binger Straße 173, 55216 Ingelheim Am Rhein (DE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: AN IMPROVED ANTIBODY DOMAIN AND ANTIBODY FRAGMENTS AND ANTIBODIES BASED THEREON

(57) Abstract: The present invention relates to an improved antibody domain and to antibody fragments and antibodies based thereon. More specifically, it relates to a C_H1 -domain which has been made more stable by selected mutation of specific residues. The C_H1-domain thus produced is natively folded and stable and therefore contributes to the stability and folding of an antibody fragment or an entire antibody. The present invention also relates to a nucleic acid coding for such an improved antibody domain or antibody fragment or antibody. The present invention also relates to a method of stabilizing a C_H1-domain of an antibody and/or of an antibody fragment comprising a C_H1-domain and/or of an antibody comprising a C_H1 -domain.



An improved antibody domain and antibody fragments and antibodies based thereon

5 The present invention relates to an improved antibody domain and to antibody fragments and antibodies based thereon. More specifically, it relates to a C_H1-domain which has been made more stable by selected mutation of specific residues. The C_H1-domain thus produced is natively folded and stable and therefore contributes to the stability and folding of an antibody fragment or an entire antibody. The present invention also relates to a nucleic acid coding for
10 such an improved antibody domain or antibody fragment or antibody. The present invention also relates to a method of stabilizing a C_H1-domain of an antibody and/or of an antibody fragment comprising a C_H1-domain and/or of an antibody comprising a C_H1-domain.

Antibodies are biotechnologically important proteins. The range of applications extends from
15 fundamental research, diagnostics to therapy, for example in the treatment of malignant diseases. Antibodies are complex glycosylated protein molecules. Antibodies occur in different varieties which are also known as so-called isotypes or classes. In mammals, there are typically five antibody isotypes known as IgA, IgD, IgE, IgG and IgM. These isotypes differ in their biological properties, functional locations and their ability to deal with different
20 antigens. Yet, structurally, they are remarkably similar. As an example, immunoglobulin G (IgG) is the most abundant antibody in the blood. It is a heterotetrameric glycoprotein assembled from two light and two heavy chains which are comprised of two and four compact Ig-domains, respectively, which are structurally almost identical (see figure 1). Antigen recognition and binding occurs via two identical antigen binding sites, also referred to as
25 paratopes (see figure 1). The antigen which is the target structure of the antibody is not only recognized in a highly specific manner, but its binding is moreover coupled to a plurality of so-called effector functions mediated by the Fc-fragment (see figure 1). Upon antigen binding, macrophages of the organism are stimulated to incorporate the antigen, and a target cell, once recognized as antigen, may be destroyed. Moreover, there also exist methods in which Fab-
30 fragments (see figure 1) of antibodies are used.

The basic structure of an intact antibody (immunoglobulin) is a Y-shaped molecule composed of two heavy (H) and two light (L) chains; in immunoglobulins A and M these molecules are assembled further into larger complexes. The L-chain consists of one variable (V_L) and one

- 2 -

constant (C_L) domain, and the IgG H-chains consist of one variable (V_H), and three constant domains (C_{H1} , C_{H2} , and C_{H3}). All these domains consist of about 100 residues, are homologous in their primary structure, and are believed to be independent structural units. All immunoglobulin domains appear to have essentially the same conformation, designated also
5 as the immunoglobulin fold, consisting of two layers of antiparallel β -sheet that are usually linked by a disulfide bond. The topology in this immunoglobulin fold typically has a specific connectivity of the individual β -strands which is also sometimes referred to as the Greek key topology which is named after a pattern found on Greek pottery.

10 With respect to the entire antibody structure, i. e. the arrangement of the various domains in the Y-shaped molecule, the polypeptide chains between the domains are susceptible to proteases to a different extent. Most susceptible is the hinge region linking the two arms to the base of Y. After cleavage at this site, the two arms are released individually which are known as Fab-fragments; the base is known as the Fc-fragment. Each Fab-fragment contains the V_L ,
15 C_L , V_H and C_{H1} -domains; the Fc-fragment has two copies of each of the C_{H2} - and C_{H3} -domains. Fv-fragments consist of only the V_H and V_L -domains. The connecting segments between the domains have varying degrees of flexibility and the individual parts of the antibody molecule may undergo considerable motion relative to each other.

20 Despite the development of advanced technologies for producing antibodies, their manufacture still remains a formidable task, also because of their complex structure and size which requires a complex folding and assembly process. Hence, the production of antibodies even now is still requiring attention and improvement. Antibodies are difficult to produce, and they may be relatively unstable.

25

Accordingly, it was an object of the present invention to provide for improved polypeptides for use in antibodies. It was also an object of the present invention to provide for improved antibodies. It was also an object of the present invention to provide for antibodies which are easier to produce and/or more stable.

30

The objects of the present invention are solved by a polypeptide having an amino acid sequence derived from a wildtype sequence of a C_{H1} -domain of an antibody, wherein in said

- 3 -

wildtype sequence, P167 and/or P189 of the wildtype sequence, or its (their) corresponding counterpart(s), is (are) replaced by

an amino acid residue independently at each occurrence selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, F, Y, W, preferably selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, more preferably A, V, T, and most preferably A, wherein the numbering of residues is with reference to the murine IgG1 wildtype heavy chain sequence using the Kabat numbering scheme.

In one embodiment, only P167 or its corresponding counterpart is replaced.

In one embodiment only P189 or its corresponding counterpart is replaced.

In another embodiment P167 and P189, or their corresponding counterparts, respectively, are replaced.

In one embodiment said wildtype sequence of said C_H1-domain is a murine sequence, preferably a murine IgG1 sequence, more preferably a sequence selected from SEQ ID NO: 1-8, and 18 even more preferably SEQ ID NO: 4.

In another embodiment, said wildtype sequence of said C_H1-domain is a human sequence, preferably a human IgG1 sequence, more preferably a sequence selected from SEQ ID NO: 9-17, more preferably selected from SEQ ID NO: 13-16, and even more preferably SEQ ID NO: 13.

In another embodiment, said wildtype sequence of said C_H1-domain is a goat sequence, chicken sequence, rat sequence or rabbit sequence.

In one embodiment said wildtype sequence of said C_H1-domain is of an Ig-isotype selected from IgA, IgA1, IgA2, IgD, IgE, IgG1, IgG2, IgG2a, IgG2b, IgG3, IgG4, and IgGM, preferably a sequence selected from SEQ ID NO: 1-17.

In another embodiment said wildtype sequence of said C_H1-domain, prior to any replacement of P167, P189 or its (their) counterpart(s) is a sequence selected from SEQ ID NO: 1-18,

- 4 -

preferably selected from SEQ ID NO: 4, 13-16 and 18, and more preferably selected from SEQ ID NO: 4, SEQ ID NO:18 and SEQ ID NO: 13.

5 In one embodiment the numbering of residues is with reference to the murine IgG1 C_H1-domain wildtype sequence, preferably SEQ ID NO:4 or SEQ ID NO:18, using the Kabat numbering scheme, and wherein said counterparts are residues in Ig C_H1-domain wildtype sequences other than murine IgG1 C_H1-domain wildtype sequence, such as C_H1-domain wildtype sequences of murine IgA, IgD, IgE, IgG2a, IgG2b, IgG3, IgM; human IgA1, IgA2, IgG, IgE, IgG1, IgG2, IgG3, IgG4, and IgM, or the same isotypes from other species, such as
10 goat, chicken, rat or rabbit, and wherein said counterparts are determined by sequence homology alignment or structure homology alignment of said murine IgG1 C_H1-domain wildtype sequence with another C_H1-domain wildtype sequence. In one embodiment of said sequence homology alignment, said counterparts are determined by being aligned with said P167 or P189 in an optimum alignment of sequences, i.e. with the best possible match of
15 residues (highest number of identical and/or homologous residues in aligned positions) between the aligned sequences.

In one embodiment said murine IgG1 C_H1-domain wildtype sequence and said another C_H1-domain wildtype sequence have a number of identical and/or homologous residues and their
20 respective positions in common, referred to as “conserved residues”, and wherein said sequence homology alignment occurs such that at least 20%, preferably at least 30%, preferably at least 40 %, preferably at least 50 %, preferably at least 60 %, preferably at least 70 %, preferably at least 80 %, preferably at least 90 %, preferably at least 95 %, more preferably at least 96 %, even more preferably at least 98 %, even more preferably at least
25 99 % of residues are conserved residues; and wherein said structure homology alignment occurs by superpositioning the tertiary structure of said murine IgG1 C_H1-domain wildtype sequence with the tertiary structure of another C_H1-domain wildtype sequence and obtaining maximum overlap of atomic coordinates of non-hydrogen protein atoms of said two wildtype sequences.

30

In one embodiment the C_H1-domain wildtype sequence is murine IgA (SEQ ID NO: 1), and said corresponding counterparts to P167 and P189 are **P167 and P190**, respectively; or the C_H1-domain wildtype sequence is murine IgD (SEQ ID NO: 2), and said corresponding

- 5 -

counterparts to P167 and P189 are **P164 and no second residue**, respectively; or the C_H1-domain wildtype sequence is murine IgE (SEQ ID NO: 3), and said corresponding counterparts to P167 and P189 are **P163 and G180**, respectively; or the C_H1-domain wildtype sequence is murine IgG2a (SEQ ID NO: 5), and said corresponding counterparts to P167 and P189 are **P167 and P189**, respectively; or the C_H1-domain wildtype sequence is murine IgG2b (SEQ ID NO: 6), and said corresponding counterparts to P167 and P189 are **P166 and P188**, respectively; or the C_H1-domain wildtype sequence is murine IgG3 (SEQ ID NO: 7), and said corresponding counterparts to P167 and P189 are **no residue and P188**, respectively or wherein the C_H1-domain wildtype sequence is murine IgM (SEQ ID NO: 8), and said corresponding counterparts to P167 and P189 are **P170 and G194**, respectively; or the C_H1-domain wildtype sequence is human IgA1 (SEQ ID NO: 9), and said corresponding counterparts to P167 and P189 are **P167 and G193**, respectively; or the C_H1-domain wildtype sequence is human IgA2 (SEQ ID NO: 10), and said corresponding counterparts to P167 and P189 are **P167 and G193**, respectively; or the C_H1-domain wildtype sequence is human IgD (SEQ ID NO: 11), and said corresponding counterparts to P167 and P189 are **P168 and G193**, respectively; or the C_H1-domain wildtype sequence is human IgE (SEQ ID NO: 12), and said corresponding counterparts to P167 and P189 are **P169 and no second residue**, respectively; or the C_H1-domain wildtype sequence is human IgG1 (SEQ ID NO: 13), and said corresponding counterparts to P167 and P189 are **P167 and G190**, respectively; or the C_H1-domain wildtype sequence is human IgG2 (SEQ ID NO: 14), and said corresponding counterparts to P167 and P189 are **P167 and G190**, respectively; or the C_H1-domain wildtype sequence is human IgG3 (SEQ ID NO: 15), and said corresponding counterparts to P167 and P189 are **P167 and G190**, respectively; or the C_H1-domain wildtype sequence is human IgG4 (SEQ ID NO: 16), and said corresponding counterparts to P167 and P189 are **P167 and G190**, respectively; or the C_H1-domain wildtype sequence is human IgM (SEQ ID NO: 17), and said corresponding counterparts to P167 and P189 are **P170 and G194**, respectively.

As used herein, the terms “no residue” or “no second residue”, when used to denote the “corresponding counterparts”, are meant to indicate that, in such specific case, no such corresponding counterpart residue exists. For example, if, with reference to the C_H1-domain wildtype sequence of human IgE (SEQ ID NO:12), the statement is made “and said corresponding counterparts to P167 and P189 are P169 and no second residue”, this is meant

- 6 -

to mean that in the C_H1-domain wildtype sequence of human IgE (SEQ ID NO:12) there is no corresponding counterpart residue to P189 (of the murine IgG1 C_H1-domain wildtype sequence) and hence no mutation is to be introduced at such position, for the C_H1-domain wildtype sequence of human IgE (SEQ ID NO:12).

5

In one embodiment, the C_H1-domain wildtype sequence is murine IgG1 as shown in SEQ ID NO:18, and said corresponding counterparts to P167 and P189 are P167 and P189, respectively.

- 10 In one embodiment said corresponding counterparts are replaced by an amino acid residue independently at each occurrence selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, F, Y, W, preferably A, V, T, more preferably A.

- The objects of the present invention are also solved by an antibody fragment, preferably an Fab-fragment, or an antibody comprising the polypeptide according the present invention.
- 15

The objects of the present invention are also solved by a nucleic acid coding for a polypeptide according to the present invention or for an antibody fragment or antibody according to the present invention.

20

The objects of the present invention are also solved by a method of stabilizing a C_H1-domain or an antibody fragment, preferably an Fab-fragment, or an antibody, said antibody fragment or said antibody comprising a C_H1-domain, said method comprising the step:

- 25 replacing in a wildtype sequence of a C_H1-domain one or two amino acid residue(s) selected from P167 and P189, or its (their) corresponding counterpart(s), by an amino acid residue, independently at each occurrence selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, F, Y, W, preferably selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, more preferably A, V, T, and most preferably A, wherein the numbering of residues is with reference to the murine IgG1 wildtype sequence using the Kabat numbering scheme, and wherein, preferably, said step of replacing occurs by side-directed mutagenesis, and wherein said wildtype sequence of a C_H1-domain is an isolated polypeptide or is part of an antibody fragment or an antibody comprising said C_H1-domain wildtype sequence.
- 30

- 7 -

In one embodiment, the polypeptide in accordance with the present invention is no antibody or Fab-fragment, but consists of only a sequence of a C_H1-domain.

- 5 In one embodiment, the polypeptide in accordance with the present invention is no complete heavy chain, but only consists of a sequence of a C_H1-domain.

The term “a C_H1-domain”, as used herein, is meant to refer to an amino acid sequence which makes up the first constant domain in the heavy chain, after the variable domain of the heavy
10 chain, when viewed in the direction from the N-terminus to the C-terminus of the heavy chain. The term is meant to encompass the wildtype sequences of such C_H1-domains, as exemplified by the various C_H1-domains in Figures 8 and 9; the term is meant to also encompass variants of such wildtype sequences, wherein P167 and/or P189 or its (their) corresponding counterparts have been replaced by another amino acid or other amino acids. Such variants
15 are herein also sometimes referred to as being “derived from a wildtype sequence of C_H1-domain”. It should be noted that the numbering used herein refers to the numbering of murine IgG1 heavy chain using the Kabat scheme (see below), and the “corresponding counterparts” are amino acid residues in other C_H1-domains which residues have been identified through sequence homology alignments or structure homology alignments.

20

A person skilled in the art will be able to determine whether or not a sequence would qualify as a sequence of a C_H1-domain, based on sequence alignments made with known C_H1-domain sequences, as well as structural determinations of the corresponding sequences using structure
25 determination techniques, such as x-ray crystallography, and nuclear magnetic resonance with and without isotope labelling.

In the context of the present invention, the Kabat numbering is used. In this context, the amino acid numbering is based on the murine sequence of the heavy chain of IgG1. Accordingly, the terms “proline residue No. 167”, and “proline residue No. 189”, are based on
30 the murine sequence numbering of IgG1 but is meant to include and designate also the “corresponding counterpart” residues in other sequences, such as human sequences or sequences of other species or sequences of other Ig isotypes. It is therefore envisaged that, although the polypeptides/antibodies of the present invention are constructed and mutated in

- 8 -

the context of the numbering of the murine sequence, the replacements/mutations may also be “transferred” to another sequence. This can be done by determining the “equivalent” or “corresponding counterpart” residues between the two sequences, typically based on sequence or structural homology between the sequences of the two C_H1-domains. In order to establish

5 homology, the amino acid sequence of the first C_H1-domain is directly compared to the sequence of a second C_H1-domain. After aligning the sequences, using one or more of the homology alignment programs well known in the art, such as CLUSTALW; (for example using conserved residues between species), allowing for necessary insertions and deletions in order to maintain alignment (i. e. avoiding the elimination of conserved residues through

10 arbitrary deletion and insertion), the residues “equivalent” or “corresponding” to particular amino acid residues in the primary sequence of the first C_H1-domain are defined. Alignment of conserved residues preferably should conserve at least 20%, preferably at least 30%, preferably at least 40%, preferably at least 50%, preferably at least 60%, preferably at least 70%, preferably at least 80%, preferably at least 90%, preferably at least 95%, more

15 preferably at least 96%, even more preferably at least 98%, even more preferably at least 99% and most preferably 100% of residues. “Equivalent” or “corresponding counterpart” residues may also be defined by determining structural homology between a first and second C_H1-domain, that is at the level of tertiary structure for C_H1-domains whose structures have been determined. In this case, “equivalent” or “corresponding” residues are defined as those, for

20 which the atomic coordinates of two or more of the mainchain atoms of a particular amino acid residue of the first C_H1-domain or precursor (N on N, CA on CA, C on C and O on O) are within 0.13 nm and preferably 0.1 nm after alignment. Alignment is achieved after the best model has been oriented and positioned to give the maximum overlap of atomic coordinates of non-hydrogen protein atoms of the proteins. Regardless of how “equivalent” or

25 “corresponding” residues are determined, and regardless of the identity of the first C_H1-domain in which the replacement mutations are introduced, what is meant to be conveyed is that the C_H1-domain polypeptides according to the present invention may be constructed into any second C_H1-domain which has a significant sequence or structural homology with the first C_H1-domain.

30

It should be noted that the numbering adopted in the claims and throughout the description, with the exception of the sequence listing (see below), is the numbering according to the Kabat

- 9 -

scheme (Martin; 1996; PROTEINS: Structure, Function and Genetics, 25, 130-133), with reference to the complete heavy chain.

Accordingly, the numbering of residues for each of the sequences referred to herein is with
 5 reference to the entire heavy chain starting with residue number 1 at the N-terminus of the heavy chain and ending with residue No. X at the C-terminus of the heavy chain. Since the C_H1 -domain is situated within the heavy chain, its residue numbers, as shown in figures 8 and 9, start with a number > 1 , usually > 100 , given that the C_H1 -domain is preceded by the variable domain of the heavy chain (V_H -domain). Accordingly, the " C_H1 -domain" of murine
 10 IgG1, as used herein, comprises residues 114-214 of the entire heavy chain, using the Kabat scheme. Likewise, the " C_H1 -domain" of human IgG1, as used herein, comprises residues 114-215 of the entire heavy chain. The corresponding starting residue and the ending residue of the various C_H1 -domain sequences shown in Figures 8 and 9 are as follows:

mIgA: 114-212
 15 mIgD: 114-239
 mIgE: 114-219
 mIgG1: 114-214
 mIgG2a: 114-218
 mIgG2b 114-219
 20 mIgG3: 114-222
 mIgM: 114-221
 hIgA1: 114-213
 hIgA2: 114-213
 hIgD: 114-223
 25 hIgE: 114-216
 hIgG1: 114-215
 hIgG2: 114-214
 hIgG3: 114-215
 hIgG4: 114-214
 30 hIgM: 114-223

It should be noted at the same time that within the attached sequence listing and the sequences SEQ ID NO:1-18, which only provide sequences of various C_H1 -domains, numbering starts

- 10 -

with no. 1, despite the fact that the residue denoted thereby is not the first residue to appear in the complete heavy chain. For example residue no. 1 of the C_H1-domain of murine IgG1 (i.e. residue no. 1 in SEQ ID NO: 4), really is residue no. 114 with respect to the numbering of the entire heavy chain using the Kabat scheme. The same also applies to the residues of the wild type sequence, which, in accordance with the present invention, are replaced by another amino acid residue. For example, with respect to P167 and P189 of the C_H1-domain of murine IgG1, the numbers 167 and 189 refer to the respective position of these residues within the entire heavy chain, again using the Kabat scheme whereas in the isolated sequence of the C_H1-domain (SEQ ID NO:4), these residues would be numbered as P54 and P76, respectively. Similar considerations apply to all the other C_H1-domains shown herein.

In a preferred embodiment, the polypeptides in accordance with the present invention only consist of the aforementioned amino acid sequences of the respective C_H1-domain, with the one or more replacements of amino acid residues, as outlined further above, and no other residues outside of the C_H1-domain, such as would for example occur in the regions connecting to and from the C_H1-domain within the heavy chain, or in regions of other domains, such as the V_H-domain or the C_H2-domain or the C_H3-domain. In another embodiment, the C_H1-domain may be part of an antibody fragment or an entire antibody.

In one preferred embodiment, said antibody fragment or antibody is a human or murine antibody fragment or antibody. In another embodiment, said antibody fragment or antibody is a goat, chicken, rat or rabbit antibody fragment or antibody.

The term “antibody”, as used herein, is meant to refer to a complete antibody or complete immunoglobulin, as opposed to a mere fragment thereof, which is herein also sometimes denoted as an “antibody fragment”.

The present inventors have surprisingly found that contrary to previous assumptions, the C_H1-domain, when studied and looked at on its own, i. e. without the other domains of the heavy or light chain present or covalently attached, in fact, is unfolded under native conditions. The term “native conditions”, as used herein is meant to refer to PBS buffer, pH 7.4, at 25°C. The term “unfolded” is meant to refer to a state of the polypeptide chain, wherein the chain does not show any defined structure, as measured by biophysical techniques such as CD (circular

- 11 -

dichroism) and nuclear magnetic resonance spectroscopy (NMR). The term “unfolded” is also sometimes referred to as “random coil”. A “random coil” protein is characterized by the absence of any secondary structure elements such as alpha-helix, β -strand, β -sheet, β -turn 3_{10} -helix, let alone the presence of any long range tertiary interactions. The terms “secondary”
5 and “tertiary” structural elements or structures or interactions are well known by someone skilled in the art and are used herein in the same sense, as would be understood by someone skilled in the art.

The fact that the C_{H1} -domain is unfolded under native conditions is surprising, given that
10 scientific evidence available so far (as an example Goto and Hamaguchi, 1982, JMB, 156:891-910; Feige et al., 2008, PNAS USA, 105:13373-13378) demonstrate that the adoption of structure of the Ig-proteins in their folding pathway starts around several hydrophobic amino acids at the core of the β -sheet structure of each domain and is completed in the reshuffling of the structure around this folding core. The evidence so far available
15 shows that all isolated domains such as C_{H1} , C_{H2} , C_{H3} , V_H etc. are autonomous structures which are folded and adopt a well defined structure under native conditions, i. e. they show a β -sheet sandwich structure, consisting of a typical fold of two layers of antiparallel β -sheets folded on top of each other to form a sandwich like structure.

Surprisingly, the present inventors could detect unexpected structural properties of the C_{H1} -
20 domain which also would have been expected to show such immunoglobulin fold, even if the C_{H1} -domain was looked at in isolation, i. e. without any of the other corresponding domains of the heavy chain or light chain, such as C_L , present or attached. The inventors have surprisingly found that the isolated C_{H1} -domain, i. e. without any of the other domains of the
25 heavy chain or light chain present or attached, is unfolded. Contrary to what one would have expected from previous results, the isolated C_{H1} -domain does not have any defined structure. Accordingly, this finding led to the assumption that a selected mutation at specific sites within the C_{H1} -domain will lead to a stabilisation of such C_{H1} -domain on its own and will therefore stabilize both the C_{H1} -domain on its own as well as an antibody comprising such C_{H1} -domain
30 within the heavy chain, as well as an antibody fragment, e. g. Fab-fragment containing such C_{H1} -domain. Accordingly, the present invention encompasses variants of the C_{H1} -domain which have improved properties with respect to assembly of the entire antibody and/or secretion of the C_{H1} -domain in cell culture experiments. Moreover, the present invention also

- 12 -

encompasses the stabilized, autonomously folding C_H1-domain, as well as larger proteins comprising such stabilized C_H1-domains. Such larger proteins could, for example be a Fab-fragment, the entire heavy chain or even an entire antibody.

- 5 The inventors have found that in order to cause the unmutated, i. e. wild type C_H1-domain to adopt a native structure, i. e. the β -sheet sandwich, as outlined above, the presence of the C_L-domain is necessary (see also figures 3 and 4 which show that C_H1 on its own adopts a random coil structure as shown by the CD-spectrum of figure 3 (trace "C_H1"). The C_L-domain on its own (trace "C_L") shows the typical β -sheet structure thus confirming that the C_L-
10 domain is an autonomous folding unit. The same applies to the complex, i. e. the two domains mixed together. The fourth trace in figure 3 is calculated trace wherein the C_H1-trace is subtracted from the combined trace "C_H1 + C_L". This calculated trace "C_H1 in complex" clearly shows that C_H1 in the presence of C_L also adopts a β -sheet sandwich structure.
- 15 The same picture emerges from NMR-spectra, wherein the C_H1-domain on its own shows the typical non-dispersed ¹H-¹⁵N HSQC spectrum in the amide region, whereas in the presence of an unlabeled C_L-domain, the amide signals are dispersed such as would be expected from a folded protein showing a well-defined structure (see figure 4). If one looks at the kinetics of the C_H1-folding process, the inventors have found that this is a rather slow process taking
20 several hours at room temperature, if C_H1 is folded in the presence of C_L (see figure 5).

Based on the aforementioned findings, the present inventors identified two residues within the C_H1-domain, which appear to be pivotal in the autonomy of the C_H1-folding. These residues are, with reference to the murine C_H1-domain of IgG1 P167 and P189. In other C_H1-domains,
25 the "corresponding counterpart" residues can be easily determined by optimum sequence homology alignment or structure homology alignment, whereby the maximum degree of identity between the two sequences aligned or between the two structures aligned is achieved. As outlined further above, the "corresponding counterpart" residues may also be determined by structural homology, rather than sequence alignments.

30

The C_H1-domain in accordance with the present invention, i. e. having one or two of the aforementioned residues replaced has improved properties in that it is secreted in higher yields in cell culture media and contributes to a better antibody assembly; the double mutant,

- 13 -

i.e. the variant in which both residues are replaced is autonomously folding and adopts a folded structure, when looked at in isolated form, i. e. in the absence of any of the other domains of the heavy or light chain, such as C_L, V_L, C_{H2}, C_{H3} etc.

- 5 Mutation of any of the residues may occur by means known to someone skilled in the art. A preferred way of producing such mutated C_{H1}-domains on their own or within an antibody is site-directed mutagenesis.

10 The C_{H1}-domains in accordance with the present invention have improved qualities in that they either show improved properties with respect to secretion in cell culture media or with respect to antibody assembly, or they have improved qualities in that they even adopt a folded structure on their own and thus are stable and will therefore aid in the structuring and assembly of an antibody fragment, e. g. Fab-fragment, or an entire antibody. In particular an autonomously folding and stabilized C_{H1}-domain will reduce the requirements of the cellular
15 folding machinery, exemplified by chaperone proteins.

In the following, reference is made to the figures wherein

20 Figure 1 shows an antibody of the IgG-isotype. The two light chains are lightly coloured, the heavy chains are represented in dark. The regions responsible for the antigen binding, for glycosylation, as well as those regions mediating effector functions (Fc-fragment) are shown.

25 Figure 2 shows the location of the C_{H1}-domain within the IgG. The heavy chain of the IgG is shown in dark, the light chain is shown lightly coloured. The C_{H1}-domain within the heavy chain is encircled.

Figure 3 shows the structural characterization of the C_{H1}-domain using circular dichroism (CD) spectroscopy. The isolated C_{H1}-domain ("C_{H1}") shows the CD-spectrum of an unfolded protein, whereas the isolated C_L-domain ("C_L") shows the typical spectrum of a β-sheet
30 protein. If both proteins are co-incubated (C_{H1} + C_L), the spectrum shows the typical trace of a β-sheet protein, and the calculated curve for C_{H1} within the co-incubation also shows such β-sheet protein trace (C_{H1} in complex, calculated).

- 14 -

Figure 4 shows the structuring of the C_H1-domain in the presence of the C_L-domain. The ¹H-¹⁵N HSQC NMR-spectrum in PBS at 25°C of the isolated C_H1-domain (shown in light colour) shows no dispersion of the amide region and is thus characteristic for an unfolded protein. Upon addition of the C_L-domain, the amide signals shift (spectrum shown in dark colour) and get more dispersed which is indicative of a structuring.

Figure 5 shows the folding kinetics of the C_H1-domain in the presence of the C_L-domain, as followed by the CD-signal at 205 nm. The folding takes several hours to be completed at 25°C in PBS.

Figure 6 shows a model of an exemplary C_H1-domain (murine from IgG1), with two proline residues marked which may be replaced in accordance with the present invention. A preferred replacement is by alanine, but other residues are also feasible. Such replacement leads to a folding of the C_H1-domain. The coordinates of this model are taken from pdb-file 12E8.

Figure 7 shows results of a mutant C_H1-domain in accordance with the present invention wherein both proline 167 and proline 189 have been replaced by alanine. The spectrum on the left show the CD-spectrum of a β-sheet (panel A), and the protein is stable, having a melting point of 50°C (panel B).

Figure 8 shows a comparison by multiple sequence alignment of C_H1-domain by multiple sequence alignment of C_H1-domains of various murine Ig-isotypes. The numbering of the C_H1-domains is in accordance with the Kabat numbering scheme. There is a conserved proline residue (underlined) which is rate determining for the folding of the C_H1-domain. The two proline residues or their corresponding counterparts (in some cases glycine like for IgM) which have been identified as being detrimental for the proper folding of the C_H1-domain are boxed. Because the two proline residues or their corresponding counterparts are conserved over antibody isotypes as well as over species (see also figure 9), such replacements in other C_H1-domains (other than murine IgG1) will have a similar effect of providing improved properties. The present invention envisages replacement of one or several of the boxed residues by amino acid residues, independently selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, F, Y, W, preferably selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, more preferably selected from A, V, T, and most preferably selected from A. It should be

- 15 -

noted that the numbering used in the sequences in figures 8 and 9 is with respect to the entire heavy chain, despite the fact that only C_H1-domain sequences are shown. Hence, for each of the aligned C_H1-domains, the respective first residue is denoted herein in the alignments of figures 8 and 9 with a residue number greater than 1 because each of these C_H1-domains is preceded by a variable domain, with reference to the entire heavy chain. The “offset” which is thus established is in accordance with the numbering scheme of Kabat as outlined above.

It should be noted that the murine wildtype sequence of the C_H1-domain as used as starting point in the present set of experiments has a tryptophan (W) in position 188 (as taken from the MAK33 antibody), whereas other murine IgG1 C_H1-domain wildtype sequences having Arginine (R) in position 188 are also contemplated (SEQ ID NO:18). Such wildtype sequence can, for example, be found in the SwissProt database as entry IGHG1_MOUSE (status: April 15, 2009) and is also contemplated as wildtype sequence “starting point” into which the respective mutation(s) is (are) introduced, as outlined further above. In this case, the “corresponding counterparts” to P167 and P189 are P167 and P189, respectively.

Figure 9 shows a comparison by multiple sequence alignment of the murine IgG1 C_H1-domain (mIgG1) (from the MAK33 antibody) with C_H1-domains of various human IgG-isotypes (hIgA1, hIgA2, etc.). Numbering is in accordance with the Kabat numbering scheme. The proline residues which are found to be detrimental for the folding of the C_H1-domain, or their corresponding counter parts, such as glycine at the equivalent positions, are shown boxed. In accordance with the present invention one or several of these proline and/or glycine residues are replaced by amino acid residues, independently selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, F, Y, W, preferably selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, more preferably selected from A, V, T, and most preferably selected from A; “m” as prefix denotes murine; “h” as prefix denotes human.

Figure 10 shows cell culture experiments to assess the influence of the P189A mutation on the assembly and secretion of complete antibody molecules. A significant increase in assembly and secretion is observed for the single point mutant P189A in comparison to the wt protein. H respectively L denote the heavy or the light chain. The completely assembled antibody corresponds to H₂L₂. BiP denotes the Hsp70 molecular chaperone BiP which binds to incompletely assembled antibodies and is thereby co-immunoprecipitated in cell lysates.

- 16 -

Lanes with no subscript indicate cell lysates, the subscript m stands for the medium, i.e. secreted protein. “ γ ” denotes antibodies used for immunoprecipitation which were directed against the IgG heavy chain, HA denotes antibodies used for immunoprecipitation were directed against the HA-tag.

5

In the enclosed sequence listing, the sequences are numbered in the order as they appear in figures 8 and 9 (read from the top to the bottom); hence SEQ ID NO: 1-8 are murine (“m”) sequences of the indicated isotypes; e.g. murine IgG1 C_H1-domain from the MAK33 antibody is SEQ ID NO:4.

10

SEQ ID NO:9-17 are human (“h”) sequences of the indicated isotypes. SEQ ID NO:9 is the sequence of the C_H1-domain of human IgA1, SEQ ID NO:10 is the sequence of the C_H1-domain of human IgA2 etc.

15 SEQ ID NO:18 is murine IgG1 C_H1-domain as found in SwissProt-entry IGHG1_MOUSE (status: April 15, 2009).

Moreover, reference is made to the following examples, which are given to illustrate, not to limit the present invention.

20

Examples

Example 1

Protein production. The C_H1 domain (Thr123-Arg215 of the pdb file 1FH5) was amplified from the murine IgG1 MAK33 cDNA and inserted into the pET28a vector (Novagen, Gibbstown, NJ, USA) without a tag. The protein was expressed as inclusion bodies as published (Feige et al., 2004). Isotope labeled C_H1 for NMR experiments was expressed in M9 minimal medium. The wild type C_L-domain was essentially purified as published (Feige et al., 2007). Inclusion bodies of the C_H1-domain were solubilized in 50 mM Tris/HCl, pH 7.5, 10 mM β -mercaptoethanol, 10 mM EDTA, 8 M urea and subsequently applied to a Q-Sepharose column equilibrated in 50 mM Tris/HCl, pH 7.5, 10 mM EDTA, 5 M urea. The protein did not bind to the column under these conditions. Refolding was carried out as published (Feige et al., 2007). After refolding, the protein was applied to a Superdex 75pg

30

- 17 -

(26/60) gel filtration column (GE Healthcare, München, Germany) equilibrated in PBS. All vectors were sequenced and protein masses were verified by mass spectrometry.

Optical spectroscopy. A Jasco J-720 spectropolarimeter was used for all CD measurements (Jasco, Gross-Umstadt, Germany). Far-UV CD spectra were recorded in a 0.2 mm quartz cuvette, far-UV kinetics in a 1 mm quartz cuvette. Spectra of the isolated domains were recorded at 45 μ M protein concentration, for the spectrum of the complex 15 μ M C_{H1} in the presence of 45 μ M C_L was used. Far-UV CD kinetics were recorded at 10 μ M protein concentration of each domain and followed at 205 nm. Spectra of the C_{H1} domain in the complex were calculated by subtraction of the spectrum of the isolated C_L domain from the spectrum of the complex, measured after a 4 h equilibration step at 25°C. All spectra were averaged 16 times and buffer corrected. Temperature melts were followed by the change in the far-UV CD signal at 216 nm with heating rate of 10°C/h.

NMR spectroscopy. Spectra of the C_{H1} domain in complex with the C_L domain were recorded at 25°C on Bruker DMX600 and DMX750 spectrometers (Bruker, Rheinstetten, Germany), whereas spectra for the assignment of the unfolded C_{H1} domain were measured at 12.5°C on a Bruker AVANCE900 spectrometer (Bruker, Rheinstetten, Germany). Backbone sequential assignment of the isolated C_{H1} domain was obtained by standard triple resonance experiments implemented with selective proton flip-back techniques for fast pulsing (Diercks et al., 2005). For all measurements of the folded C_{H1} domain in association with C_L, a twofold excess of unlabeled C_L was added to ¹⁵N or ¹⁵N, ¹³C labeled C_{H1}. Prior to steady state measurements, samples were incubated for at least 6 h at room temperature to ensure complete folding of the C_{H1} domain. Backbone sequential assignment of the assembled C_{H1} domain was achieved with standard triple resonance experiments with selective proton flip-back techniques for fast pulsing. The assignment of the carbon chemical shifts was limited to the C' and C ^{α} chemical shifts due to the relaxation properties of the whole protein complex. To verify the backbone resonance assignment NH residual dipolar couplings (RDCs) were determined. The sample was prepared as described above and aligned with nonionic liquid crystalline media (Ruckert and Otting, 2000). NH RDC values were extracted from IPAP-HSQC spectra using Bruker pulse sequences. The sequential information based on the C' and C ^{α} chemical shifts as well as the NH RDC values and the crystal structure of the folded C_{H1}

- 18 -

domain (pdb code: 1ORS) served as input for the software MARS (Jung and Zweckstetter, 2004b; Jung and Zweckstetter, 2004a).

Cell culture experiments. The murine IgG1 MAK33 light chain (LC_{WT}) and heavy chain (HC_{WT}) cDNAs were obtained with an intact signal sequence for expression in mammalian ER. An HA-epitope tag was engineered at the C-terminus of the wild type light chain for immunoprecipitation purposes. The heavy chain proline exchange mutants was generated by site-directed mutagenesis. The recombinant plasmids, along with a pMT vector encoding hamster BiP (Lee et al., 1999) were introduced into COS-1 cells (Gluzman, 1981) that were cultured as described (Lee et al., 1999) using FuGENE 6 transfection reagent (Roche, Indianapolis, USA) following the manufacturer's protocol. Metabolic labeling, cell lysis, immunoprecipitation, and visualization of the proteins were performed as described previously (Lee et al., 1999). Anti-rodent BiP antiserum (Hendershot et al., 1995), a monoclonal anti-HA (12CA5) antibody (kindly provided by Dr. Al Reynolds, Vanderbilt University, USA), and a goat anti-mouse Ig κ and goat anti-mouse γ antibodies (Southern Biotech, Birmingham, AL, USA) and Protein A Sepharose beads were used for immunoprecipitations. For metabolic labeling experiments, cells were cultured in DMEM lacking methionine and cysteine and labeled with ³⁵S Translabel (MP Biomedicals, Irvine, CA, USA) for the indicated times.

Example 2

Antibodies are modular structures composed of a series of structurally highly homologous domains. These domains can usually be produced and studied separately and represent independent structural units (Goto and Hamaguchi, 1982; Lilie et al., 1995). Analysis of the murine IgG1 C_H1 domain revealed that, in marked contrast to all antibody domains studied thus far (Goto and Hamaguchi, 1982; Thies et al., 1999; Feige et al., 2004; Rothlisberger et al., 2005), the isolated C_H1 domain is an unfolded protein, irrespective of whether its internal disulfide bridge is formed or not (far-UV CD spectrum in figure 3, data for the reduced protein not shown). To further characterize the unfolded state of the C_H1 domain under physiological conditions, iodide fluorescence quenching experiments were carried out. The

- 19 -

experiments indicate no significant differences in the burial of tryptophan residues between C_{H1} in PBS and in 3 M GdmCl (data not shown). Additionally, NMR experiments were recorded. The ¹H-¹⁵N spectra of the isolated C_{H1} domain are characteristic for an unfolded protein (figure 4). Taken together, these data argue against the presence of a significant amount of stable structure in the isolated C_{H1} domain. However, the pattern changed completely when the C_L domain, the cognate association partner of C_{H1} in the antibody, was added. Only then did we observe folding of the C_{H1} domain to a well defined β-sheet structure (figure 3-5). Thus, the C_L domain is necessary and sufficient to induce structure formation in C_{H1}. This folding process was observed only if the internal disulfide bridge in the C_{H1} domain was present (data not shown) and had an observable time constant of 50 min in kinetics followed by far-UV CD spectroscopy at 25°C (figure 5).

Example 3

If one of the suggested mutations in the C_{H1} is performed by site directed mutagenesis,, e.g. Pro 189 against alanine, the protein expressed and purified as described above (example 1), the murine C_{H1} domain still remains unfolded (data not shown). It shows only minor differences in the far-UV CD spectrum compared to the wild type protein. Nevertheless, already the single mutation has a significant impact on the overall folding and assembly efficiency of complete antibody molecules. *In vivo* experiments show that the assembly and secretion efficiency of this single point mutant, P189A, is significantly increased (figure 10) highlighting the critical role of the performed mutations.

Example 4

The C_{H1}-domain with proline 167 and proline 189 mutated to alanine was generated by site directed mutagenesis, recombinantly expressed and purified as the wild type C_{H1} domain (example 1). After size exclusion chromatography the identity of the protein was confirmed by MALDI-MS on a Bruker Ultraflex 2.

Unlike the wild type C_{H1} domain or the single proline to alanine mutants P167 and P189, respectively (example 3), the double mutant with proline 167 and proline 189 mutated to alanine is folded even in the absence of its cognate partner, the C_L-domain. The secondary structure of the mutant C_{H1} domain was analyzed at the physiological temperature of 37°C by CD spectroscopy. The far-UV CD spectrum depicted in figure 7 is typical for a protein consisting of β-sheets. The temperature transition as followed by CD spectroscopy at the

- 20 -

wavelength $\lambda = 216$ nm reveals that the simultaneous mutation of proline residues 167 and proline 189 transforms the natively unfolded C_H1 domain into a stably folded Ig domain with a transition midpoint at 50°C which is far above the physiological temperature.

5 Additionally, the C_H1 domain with proline 167 and proline 189 mutated to alanine is still able to interact with its cognate partner, the C_L domain of the light chain as observed by the change in intrinsic fluorescence (data not shown).

The features of the present invention disclosed in the specification, the claims and/or in the drawings may, both separately and in any combination thereof, be material for realizing the
10 invention in various forms thereof.

Reference List

- 15 Diercks,T., Daniels,M., and Kaptein,R. (2005). Extended flip-back schemes for sensitivity enhancement in multidimensional HSQC-type out-and-back experiments. *Journal of Biomolecular Nmr* 33, 243-259.
- Feige,M.J., Hagn,F., Esser,J., Kessler,H., and Buchner,J. (2007). Influence of the internal disulfide bridge on the folding pathway of the C-L antibody domain. *Journal of Molecular*
20 *Biology* 365, 1232-1244.
- Feige,M.J., Walter,S., and Buchner,J. (2004). Folding mechanism of the C(H)2 antibody domain. *Journal of Molecular Biology* 344, 107-118.
- Gluzman,Y. (1981). SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* 23, 175-182.
- 25 Goto,Y. and Hamaguchi,K. (1982). Unfolding and Refolding of the Constant Fragment of the Immunoglobulin Light Chain. *Journal of Molecular Biology* 156, 891-910.
- Hendershot,L.M., Wei,J.Y., Gaut,J.R., Lawson,B., Freiden,P.J., and Murti,K.G. (1995). In vivo expression of mammalian BiP ATPase mutants causes disruption of the endoplasmic reticulum. *Mol. Biol. Cell* 6, 283-296.
- 30 Jung,Y.S. and Zweckstetter,M. (2004a). Backbone assignment of proteins with known structure using residual dipolar couplings. *Journal of Biomolecular Nmr* 30, 25-35.
- Jung,Y.S. and Zweckstetter,M. (2004b). Mars - robust automatic backbone assignment of proteins. *Journal of Biomolecular Nmr* 30, 11-23.

- 21 -

Lee,Y.K., Brewer,J.W., Hellman,R., and Hendershot,L.M. (1999). BiP and immunoglobulin light chain cooperate to control the folding of heavy chain and ensure the fidelity of immunoglobulin assembly. *Molecular Biology of the Cell* 10, 2209-2219.

- 5 Lilie,H., Rudolph,R., and Buchner,J. (1995). Association of antibody chains at different stages of folding: prolyl isomerization occurs after formation of quaternary structure. *J. Mol. Biol.* 248, 190-201.

Rothlisberger,D., Honegger,A., and Pluckthun,A. (2005). Domain interactions in the Fab fragment: A comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. *Journal of Molecular Biology* 347, 773-789.

- 10 Ruckert,M. and Otting,G. (2000). Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *Journal of the American Chemical Society* 122, 7793-7797.

- 15 Thies,M.J.W., Mayer,J., Augustine,J.G., Frederick,C.A., Lilie,H., and Buchner,J. (1999). Folding and association of the antibody domain C(H)3: Prolyl isomerization precedes dimerization. *Journal of Molecular Biology* 293, 67-79.

- 22 -

Claims

1. A polypeptide having an amino acid sequence derived from a wildtype sequence of a
5 C_H1-domain of an antibody, wherein in said wildtype sequence, P167 and/or P189 of
the wildtype sequence, or its (their) corresponding counterpart(s), is (are) replaced by
an amino acid residue independently at each occurrence selected from A, V, T, L, I, S,
C, M, D, N, E, Q, K, R, H, F, Y, W, preferably selected from A, V, T, L, I, S, C, M, D,
N, E, Q, K, R, H, more preferably A, V, T, and most preferably A,
10 wherein the numbering of residues is with reference to the murine IgG1 wildtype heavy
chain sequence using the Kabat numbering scheme.
2. The polypeptide according to claim 1, wherein only P167 or its corresponding
counterpart is replaced.
15
3. The polypeptide according to claim 1, wherein only P189 or its corresponding
counterpart is replaced.
4. The polypeptide according to claim 1, wherein P167 and P189, or their corresponding
20 counterparts, respectively, are replaced.
5. The polypeptide according to any of claims 1-4, wherein said wildtype sequence of said
C_H1-domain is a murine sequence, preferably a murine IgG1 sequence, more preferably
a sequence selected from SEQ ID NO: 1-8 and 18, even more preferably SEQ ID NO: 4.
25
6. The polypeptide according to any of claims 1-4, wherein said wildtype sequence of said
C_H1-domain is a human sequence, preferably a human IgG1 sequence, more preferably
a sequence selected from SEQ ID NO: 9-17, more preferably selected from SEQ ID
NO: 13-16, and even more preferably SEQ ID NO: 13.
30
7. The polypeptide according to any of claims 1-6, wherein said wildtype sequence of said
C_H1-domain is of an Ig-isotype selected from IgA, IgA1, IgA2, IgD, IgE, IgG1, IgG2,

- 23 -

IgG2a, IgG2b, IgG3, IgG4, and IgGm, preferably a sequence selected from SEQ ID NO: 1-17.

- 5 8. The polypeptide according to any of claims 1-6, wherein said wildtype sequence of said C_H1-domain, prior to any replacement of P167, P189 or its (their) counterpart(s) is a sequence selected from SEQ ID NO: 1-18, preferably selected from SEQ ID NO: 4, 13-16 and 18, and more preferably selected from SEQ ID NO: 4, SEQ ID NO:18 and SEQ ID NO: 13.
- 10 9. The polypeptide according to any of claims 1-8, wherein the numbering of residues is with reference to the murine IgG1 C_H1-domain wildtype sequence using the Kabat numbering scheme, and wherein said counterparts are residues in Ig C_H1-domain wildtype sequences other than murine IgG1 C_H1-domain wildtype sequence, such as C_H1-domain wildtype sequences of murine IgA, IgD, IgE, IgG2a, IgG2b, IgG3, IgM; 15 human IgA1, IgA2, IgG, IgE, IgG1, IgG2, IgG3, IgG4, and IgM, and wherein said counterparts are determined by sequence homology alignment or structure homology alignment of said murine IgG1 C_H1-domain wildtype sequence with another C_H1-domain wildtype sequence.
- 20 10. The polypeptide according to claim 9, wherein said murine IgG1 C_H1-domain wildtype sequence and said another C_H1-domain wildtype sequence have a number of identical and/or homologous residues and their respective positions in common, referred to as “conserved residues”, and wherein said sequence homology alignment occurs such that at least 20%, preferably at least 30%, preferably at least 40 %, preferably at least 50 %, 25 preferably at least 60 %, preferably at least 70 %, preferably at least 80 %, preferably at least 90 %, preferably at least 95 %, more preferably at least 96 %, even more preferably at least 98 %, even more preferably at least 99 % of residues are conserved residues; and wherein said structure homology alignment occurs by superpositioning the tertiary structure of said murine IgG1 C_H1-domain wildtype sequence with the tertiary 30 structure of another C_H1-domain wildtype sequence and obtaining maximum overlap of atomic coordinates of non-hydrogen protein atoms of said two wildtype sequences.

- 24 -

11. The polypeptide according to any of the foregoing claims, wherein the C_H1-domain wildtype sequence is murine IgA (SEQ ID NO: 1), and said corresponding counterparts to P167 and P189 are P167 and P190, respectively; or wherein the C_H1-domain wildtype sequence is murine IgD (SEQ ID NO: 2), and said corresponding counterparts to P167 and P189 are P164 and no second residue, respectively; or wherein the C_H1-domain wildtype sequence is murine IgE (SEQ ID NO: 3), and said corresponding counterparts to P167 and P189 are P163 and G180, respectively; or wherein the C_H1-domain wildtype sequence is murine IgG2a (SEQ ID NO: 5), and said corresponding counterparts to P167 and P189 are PP167 and P189, respectively; or wherein the C_H1-domain wildtype sequence is murine IgG2b (SEQ ID NO: 6), and said corresponding counterparts to P167 and P189 are no residue and P188, respectively; or wherein the C_H1-domain wildtype sequence is murine IgG3 (SEQ ID NO: 7), and said corresponding counterparts to P167 and P189 are no residue and P188, respectively; or wherein the C_H1-domain wildtype sequence is murine IgM (SEQ ID NO: 8), and said corresponding counterparts to P167 and P189 are P170 and G194, respectively; or wherein the C_H1-domain wildtype sequence is human IgA1 (SEQ ID NO: 9), and said corresponding counterparts to P167 and P189 are P167 and G193, respectively; or wherein the C_H1-domain wildtype sequence is human IgA2 (SEQ ID NO: 10), and said corresponding counterparts to P167 and P189 are P167 and G193, respectively; or wherein the C_H1-domain wildtype sequence is human IgD (SEQ ID NO: 11), and said corresponding counterparts to P167 and P189 are P168 and G193, respectively; or wherein the C_H1-domain wildtype sequence is human IgE (SEQ ID NO: 12), and said corresponding counterparts to P167 and P189 are P169 and no second residue, respectively; or wherein the C_H1-domain wildtype sequence is human IgG1 (SEQ ID NO: 13), and said corresponding counterparts to P167 and P189 are P167 and G190, respectively; or wherein the C_H1-domain wildtype sequence is human IgG2 (SEQ ID NO: 14), and said corresponding counterparts to P167 and P189 are P167 and G190, respectively; or wherein the C_H1-domain wildtype sequence is human IgG3 (SEQ ID NO: 15), and said corresponding counterparts to P167 and P189 are P167 and G190, respectively; or wherein the C_H1-domain wildtype sequence is human IgG4 (SEQ ID NO: 16), and said corresponding counterparts to P167 and P189 are P167 and G190, respectively; or wherein the C_H1-domain wildtype sequence is human IgM (SEQ ID NO: 17), and said corresponding counterparts to P167 and P189 are P170 and G194,

- 25 -

respectively, or wherein the C_H1-domain wildtype sequence is murine IgG1 (SEQ ID NO:18), and said corresponding counterparts to P167 and P189 are P167 and P189, respectively.

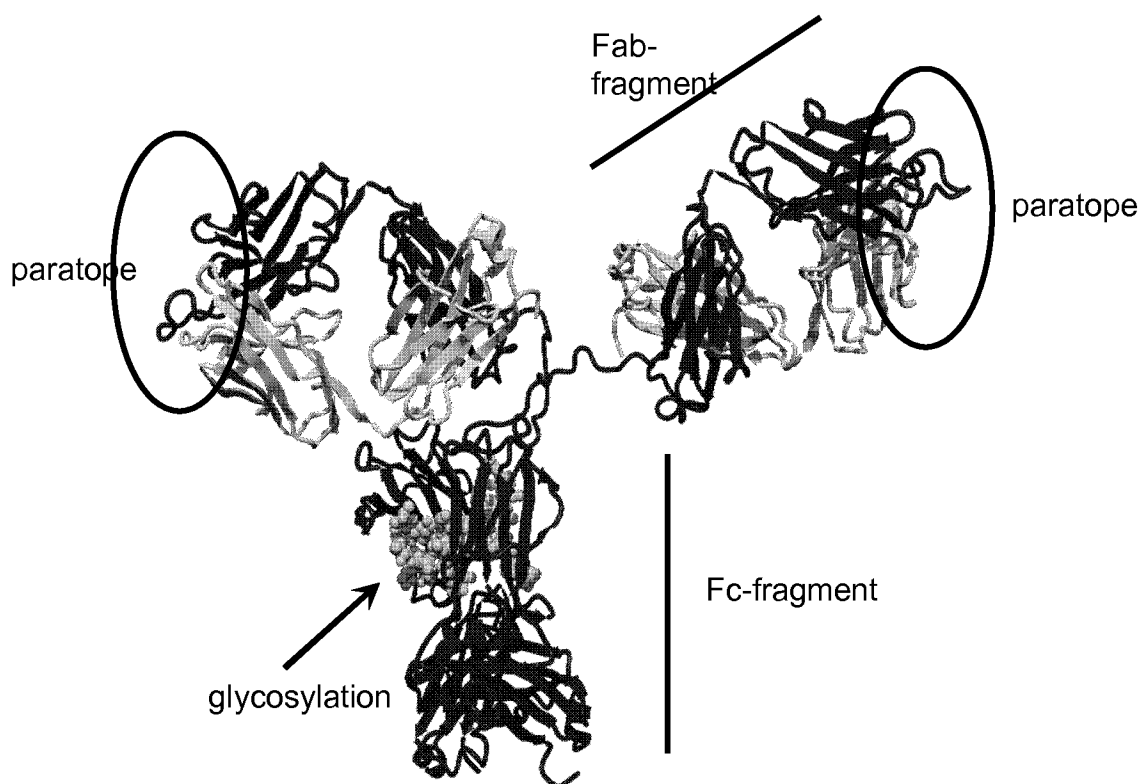
12. The polypeptide according to claim 11, wherein said corresponding counterparts are replaced by an amino acid residue independently at each occurrence selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, F, Y, W, preferably A, V, T, more preferably A.

13. An antibody fragment, preferably an Fab-fragment, or an antibody comprising the polypeptide according to any of claims 1-12.

14. A nucleic acid coding for a polypeptide according to any of claims 1-12 or for an antibody fragment or antibody according to claim 13.

15. A method of stabilizing a C_H1-domain or an antibody fragment, preferably an Fab-fragment, or an antibody, said antibody fragment or said antibody comprising a C_H1-domain, said method comprising the step:

replacing in a wildtype sequence of a C_H1-domain one or two amino acid residue(s) selected from P167 and P189, or its (their) corresponding counterpart(s), by an amino acid residue, independently at each occurrence selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, F, Y, W, preferably selected from A, V, T, L, I, S, C, M, D, N, E, Q, K, R, H, more preferably A, V, T, and most preferably A, wherein the numbering of residues is with reference to the murine IgG1 wildtype sequence using the Kabat numbering scheme, and wherein, preferably, said step of replacing occurs by site-directed mutagenesis, and wherein said wildtype sequence of a C_H1-domain is an isolated polypeptide or is part of an antibody fragment or an antibody comprising said C_H1-domain wildtype sequence.

**Figure 1**

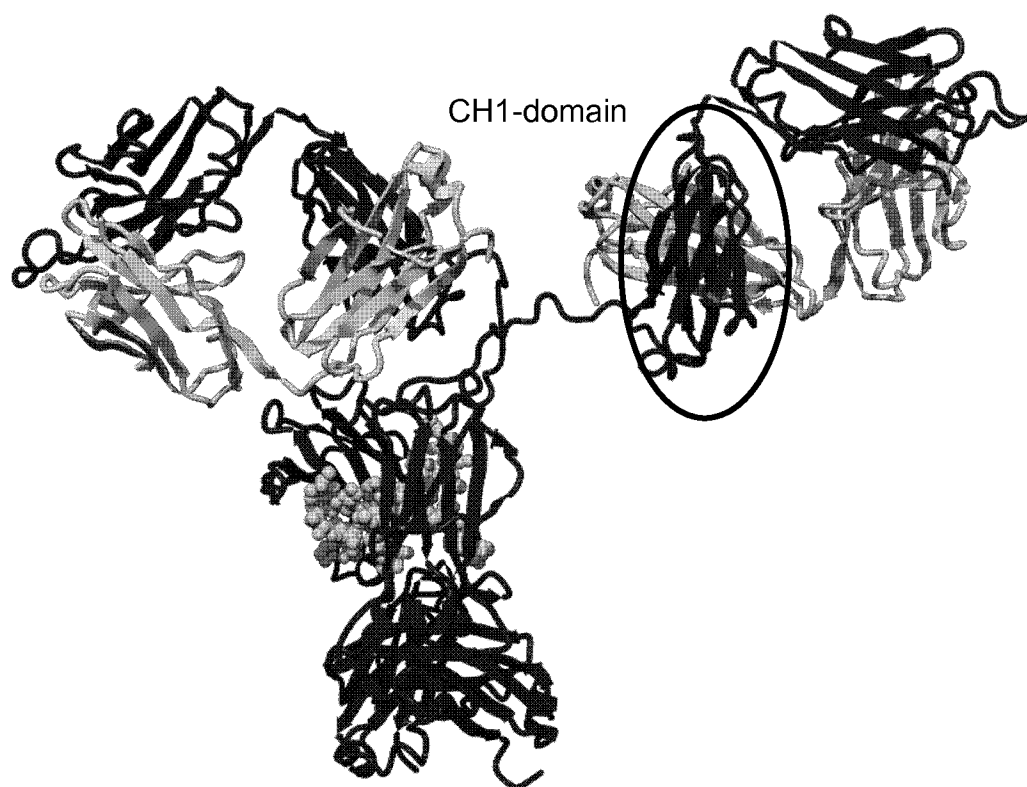


Figure 2

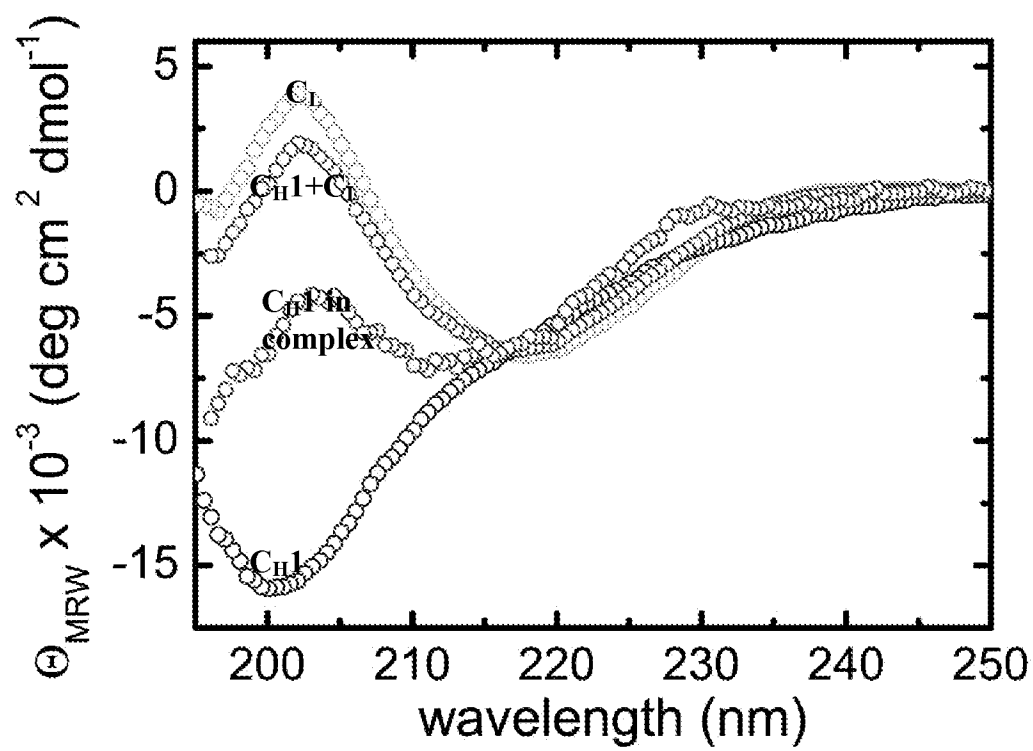


Figure 3

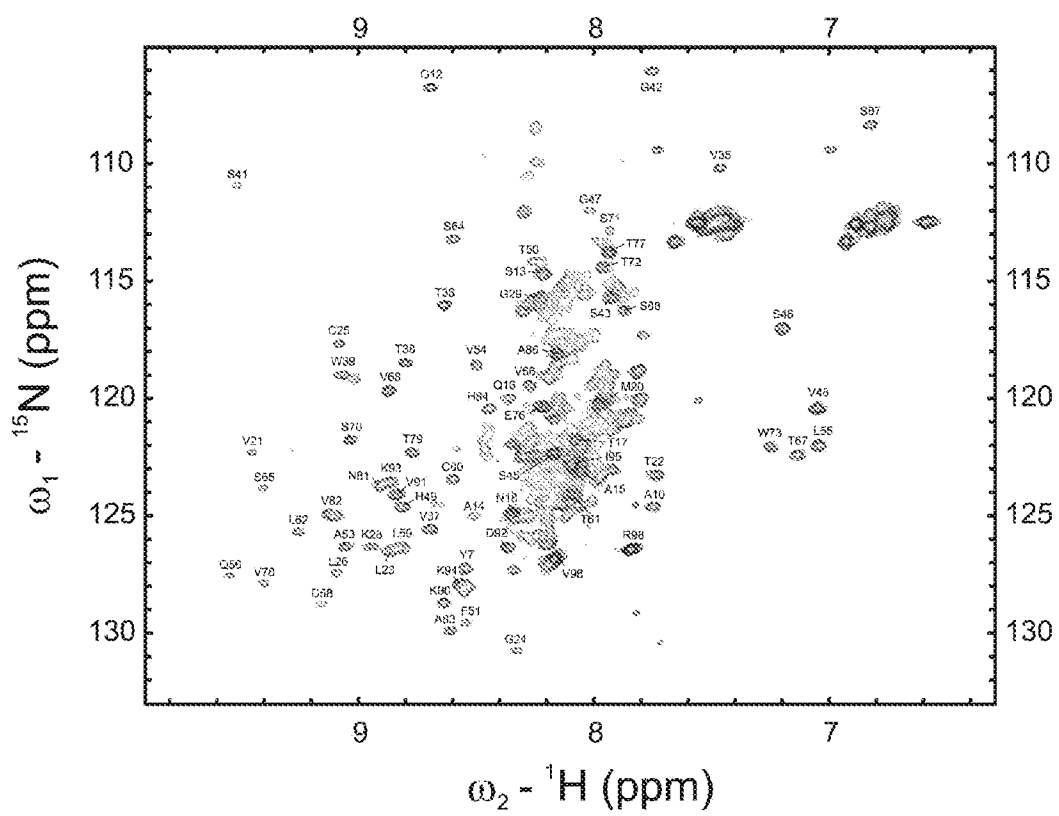
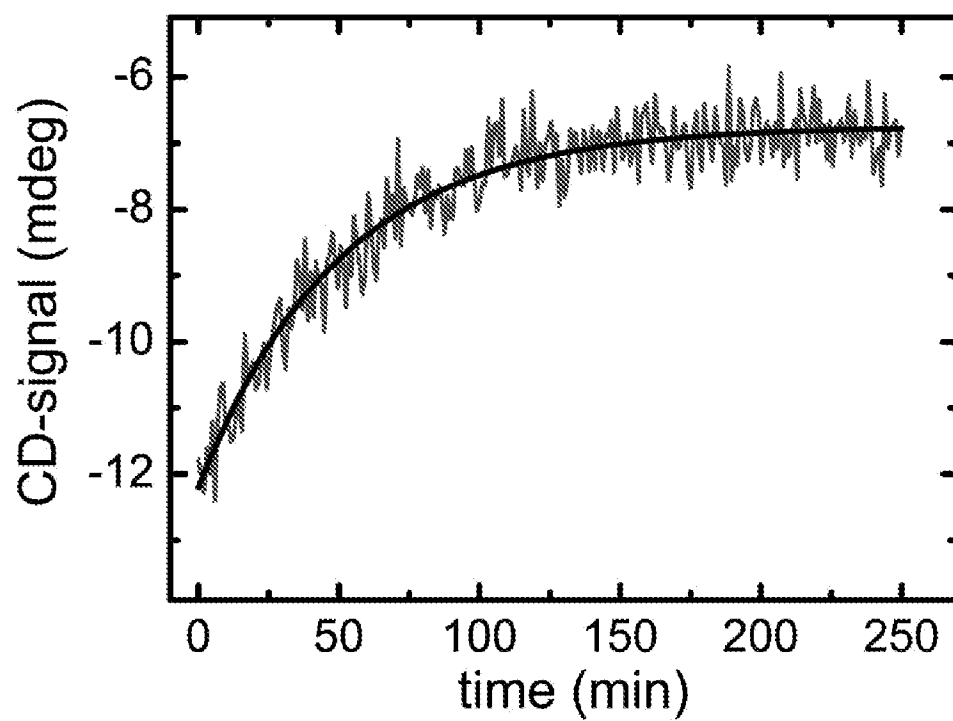


Figure 4

**Figure 5**

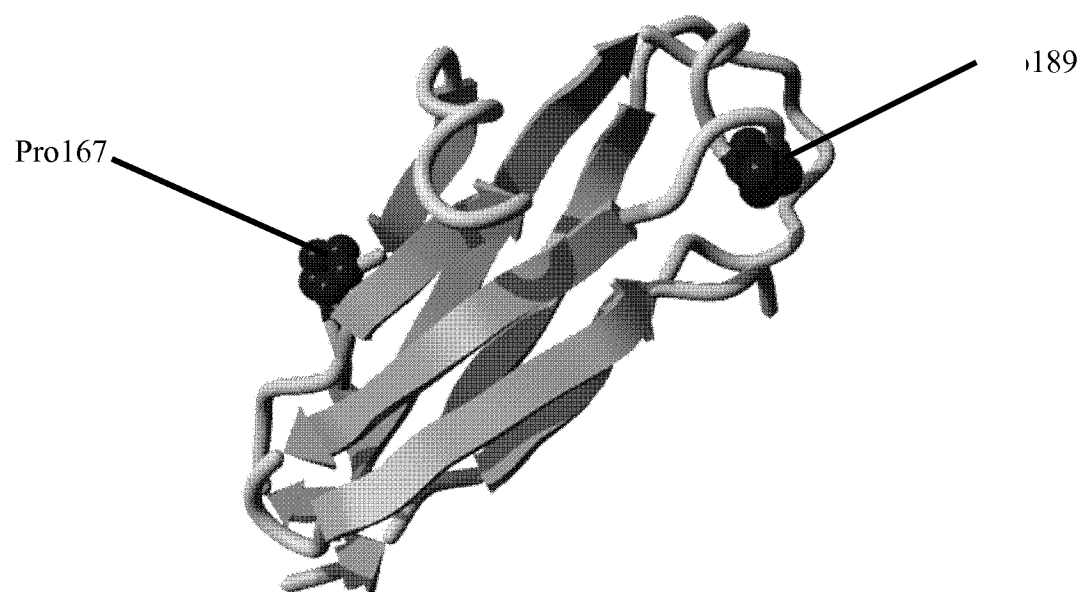
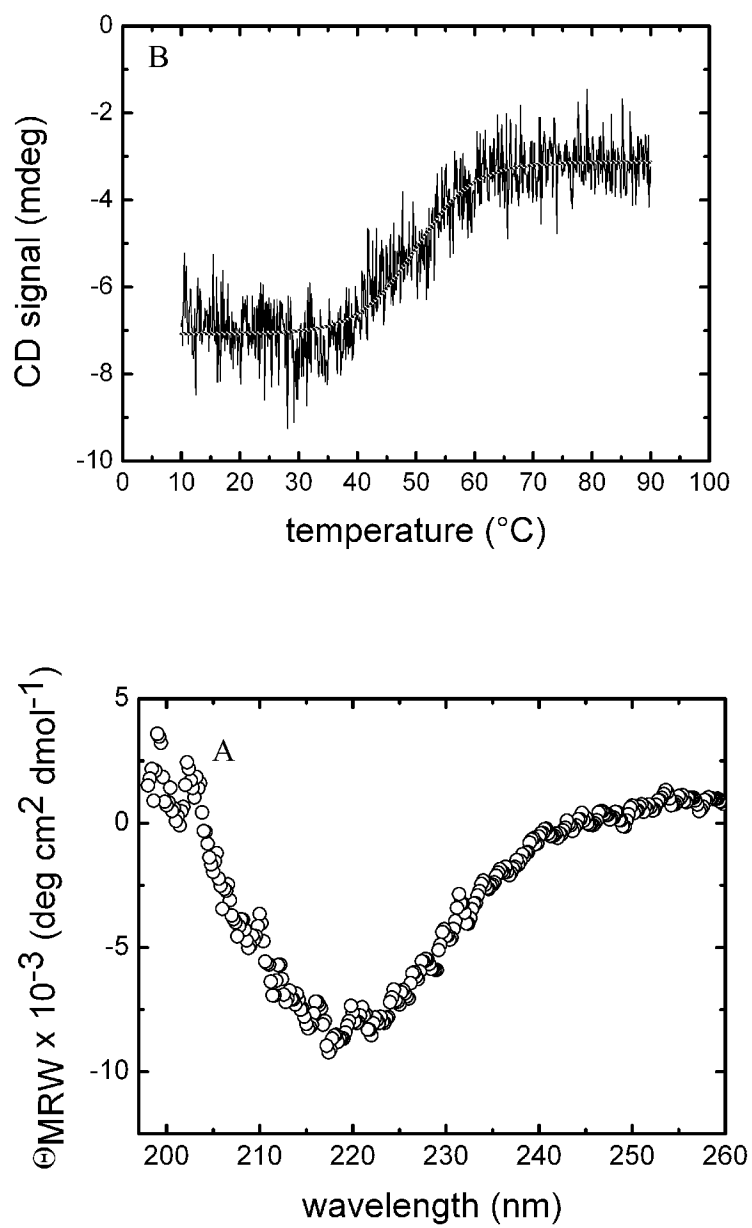


Figure 6

**Figure 7**

CLUSTAL 2.0.10 multiple sequence alignment

```

mIgA      114 ESARNPTIYPLTLPPALSS--DPVIIGCLIHDFPSGTMNVTWGKSG---KDITTVNEP 167
mIgD      114 -DKKEPDMFLLSECKAPEEN-EKINLGCLCIGSQP---LKISWEPKK---SSIVEHVFP 164
mIgE      114 -SIRNPQLYPLKP-CKGT---ASMTLGCLVKDYFP-NPVTVTWYSDS---LNMSTVNEP 163
mIgG1     114 AKTTPPSVYPLAPGSAAQTN-SMVTLGCLVKGYFP-EPVTVTWNSGS---LSSGVHTEP 167
mIgG2a    114 AKTTAPSVYPLAPVCGDITG-SSVTLGCLVKGYFP-EPVTLTWNSGS---LSSGVHTEP 167
mIgG2b    114 -KTTPPSVYPLAPGCGDITG-SSVTLGCLVKGYFP-ESVTVTWNSGS---LSSSVHTEP 166
mIgG3     114 -TTTAPSVYPLVPGCSDTSG-SSVTLGCLVKGYFP-EPVTVKWNRYGA---LSSGVRTVS 166
mIgM      114 -SQSFPNVFPLVSCESPLSDKNLVAMGCLARDFLP-STISFTWNYQNNTEVIQGITFP 170
          *  ::  *                :  :***  .  *  :...*  ..

mIgA      168 PALASGGRYTMSSQLTLPAVECEP-GESEVKCSVQHD-----SNPVQEL 209
mIgD      165 SEMRNG-NYTMVLQVTVLASELN----LNHTCTINKPKRKEKPFKFPGSWDSQSSKRVT 219
mIgE      164 -ALGSE-LKVTTSQVTS----WGKS-AKNFTCHVTHPPSF-----NESRTILVR 205
mIgG1     168 AVLQSD-LYTLSSSVTVPSSTWP---SETVTCNVAHP-----ASSTKVVDK 208
mIgG2a    168 AVLQSD-LYTLSSSVTVTSSTWP---SQSITCNVAHP-----ASSTKVVDK 208
mIgG2b    167 ALLQSG-LYTMSSSVTVPSSTWP---SQTVTCNVAHP-----ASSTTVVDK 207
mIgG3     167 SVLQSG-FYSLSSLVTVPSSTWP---SQTVICNVAHP-----ASKTELIK 207
mIgM      171 TLRTGG-KYLATSQVLLSPKSILEGSDEYLVCKIHYG-----GKNKDLHV 214
          .      :                *  :                .

mIgA      210 DVN----- 212
mIgD      220 TLQAKNHSTEATKAITTKKD 239
mIgE      206 PVNITEPTLELLHS----- 219
mIgG1     209 KIVPRD----- 214
mIgG2a    209 KIEPRGPTIK----- 218
mIgG2b    208 KLEPSGPISTIN----- 219
mIgG3     208 RIEPRIPKPSTPPGS----- 222
mIgM      208 PIPAVAE----- 221
          :

```

Figure 8

CLUSTAL 2.0.10 multiple sequence alignment

```

mIgG1      114      AKTTPPSVYPLAPGSA--AQTNMVTLGCLVKGYFP-EPVTVTWNSGSLSSGVHTFPA-V 169
hIgA1      114      ASPTSPKVFPLSLCST---QPDGNVVIACLVQGFFPQEPLSVTWSESGQGVNTARNFPPSQ 170
hIgA2      114      ASPTSPKVFPLSLDST---PQDGNVVVACLVQGFFPQEPLSVTWSESGQNVNTARNFPPSQ 170
hIgD       114      APTKAPDVFPPIISGCR-HPKDNSPVVLACLIITGYHP-TSVTVTWYMGTSQSQPQRTFPE-I 170
hIgE       114      ASTQSPSVFPLTRCCKNIPSNATSVTLGCLATGYFP-EPVMVTWDTGSLNGTTMTLPATT 172
hIgG1      114      ASTKGPSVFPLAPSSK--STSGGTAALGCLVKDYFP-EPVTVSWNSGALTSGVHTFPA-V 169
hIgG2      114      ASTKGPSVFPLAPCSR--STSESTAALGCLVKDYFP-EPVTVSWNSGALTSGVHTFPA-V 169
hIgG3      114      ASTKGPSVFPLAPCSR--STSGGTAALGCLVKDYFP-EPVTVSWNSGALTSGVHTFPA-V 169
hIgG4      114      ASTKGPSVFPLAPCSR--STSESTAALGCLVKDYFP-EPVTVSWNSGALTSGVHTFPA-V 169
hIgM       114      GSASAPTLFPLVSCEN-SPSDTSSVAVGCLAQDFLP-DSITLSWKYKNNSDISSTRGFPS 171
          . . * ::*:          ....**  .: *  .: ::*          .

mIgG1      170      LQS-DLYTLSSSVTVPSSTWPS---ETVTCNVAHPASSTK-----VDKKIVPRD- 214
hIgA1      171      DASGDLYTTSSQLTLPATQCLAG--KSVTCHVKHYTNPSQ-----DVTVP---- 213
hIgA2      171      DASGDLYTTSSQLTLPATQCPDG--KSVTCHVKHYTNPSQ-----DVTVP---- 213
hIgD       171      QRRDSYYMTSSQLSTPLQQWRQ---GKYKCVVQHTASKSKKEIFRWPESPKAQASS- 223
hIgE       171      LTLSGHYATISLLTV-SGAWAK---QMFTCRVAHTPSSTDW-----VDNKTFE--- 216
hIgG1      173      LQSSGLYSLSSVTVPSSSISLT---QTYICNVNHHKPSNTK-----VDKKVEPKS- 215
hIgG2      170      LQSSGLYSLSSVTVPSSSNEGT---QTYTCNVNHHKPSNTK-----VDKTVERK-- 214
hIgG3      170      LQSSGLYSLSSVTVPSSSISLT---QTYTCNVNHHKPSNTK-----VDKRVELKT- 215
hIgG4      170      LQSSGLYSLSSVTVPSSSISLT---KTYTCNVNHHKPSNTK-----VDKRVESK-- 214
hIgM       172      VLRGGKYAATSQVLLPSKDVMOGTFDEHVVCVKVQHPNGNKEKN-----VPLPVIAELP 223
          . * * :          * * *  . . .

```

Figure 9

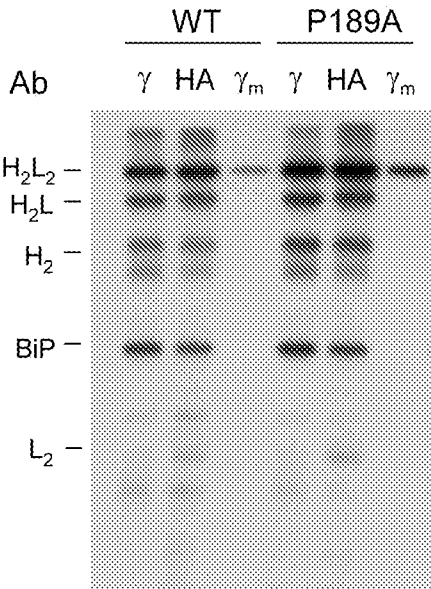


Figure 10

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2010/055440

A. CLASSIFICATION OF SUBJECT MATTER

INV. C07K16/00 C12N15/13

ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data, CHEM ABS Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 89/06690 A1 (GEN HOSPITAL CORP [US]) 27 July 1989 (1989-07-27) Registry number 127497-66-9 corresponds to a sequence of 729 amino acids length, encompasses SEQ ID N013 with an open position at amino acid P167 or its equivalent; page 23; example 2	1,2,6-8, 11-15
X	WO 95/15982 A2 (GENZYME CORP [US]) 15 June 1995 (1995-06-15) table 2; sequences 58,59 ----- -/--	1,3,5, 7-10,13, 14



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

G document member of the same patent family

Date of the actual completion of the international search

7 June 2010

Date of mailing of the international search report

12/07/2010

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040.
Fax: (+31-70) 340-3016

Authorized officer

Vadot, Pierre

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2010/055440

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE CA [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 8 December 2008 (2008-12-08), THIRUMENI, NAGARAJAN ET AL: "Fully human monoclonal antibodies against rabies virus" XP002548460 retrieved from STN Database accession no. 2008:1465924 abstract	1-15
A	& IN 2007 CH00 072 A1 (INDIAN INNUMOLOGICALS LIMITED, INDIA) 28 November 2008 (2008-11-28) -----	1-15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2010/055440

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 8906690	A1	27-07-1989	AU 611551 B2 13-06-1991
			AU 3281889 A 11-08-1989
			DE 68913771 D1 21-04-1994
			DE 68913771 T2 18-08-1994
			DK 174190 A 14-09-1990
			EP 0325262 A2 26-07-1989
			ES 2051897 T3 01-07-1994
			IE 63186 B1 22-03-1995
			JP 2980928 B2 22-11-1999
			JP 3502283 T 30-05-1991
			PT 89484 A 04-10-1989
			ZA 8900430 A 25-10-1989
WO 9515982	A2	15-06-1995	AU 696293 B2 03-09-1998
			AU 1432195 A 27-06-1995
			EP 0733070 A1 25-09-1996
			JP 9506262 T 24-06-1997