Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2689125 C 2015/10/20

(11)(21) **2 689 125**

(12) BREVET CANADIEN / CANADIAN PATENT

(13) **C**

(86) **Date de dépôt PCT/PCT Filing Date:** 2008/06/25
(87) **Date publication PCT/PCT Publication Date:** 2009/01/08
(45) **Date de délivrance/Issue Date:** 2015/10/20
(85) **Entrée phase nationale/National Entry:** 2009/12/01
(86) **N° demande PCT/PCT Application No.:** EP 2008/005136
(87) **N° publication PCT/PCT Publication No.:** 2009/003623
(30) **Priorité/Priority:** 2007/06/29 (EP07012774.1)

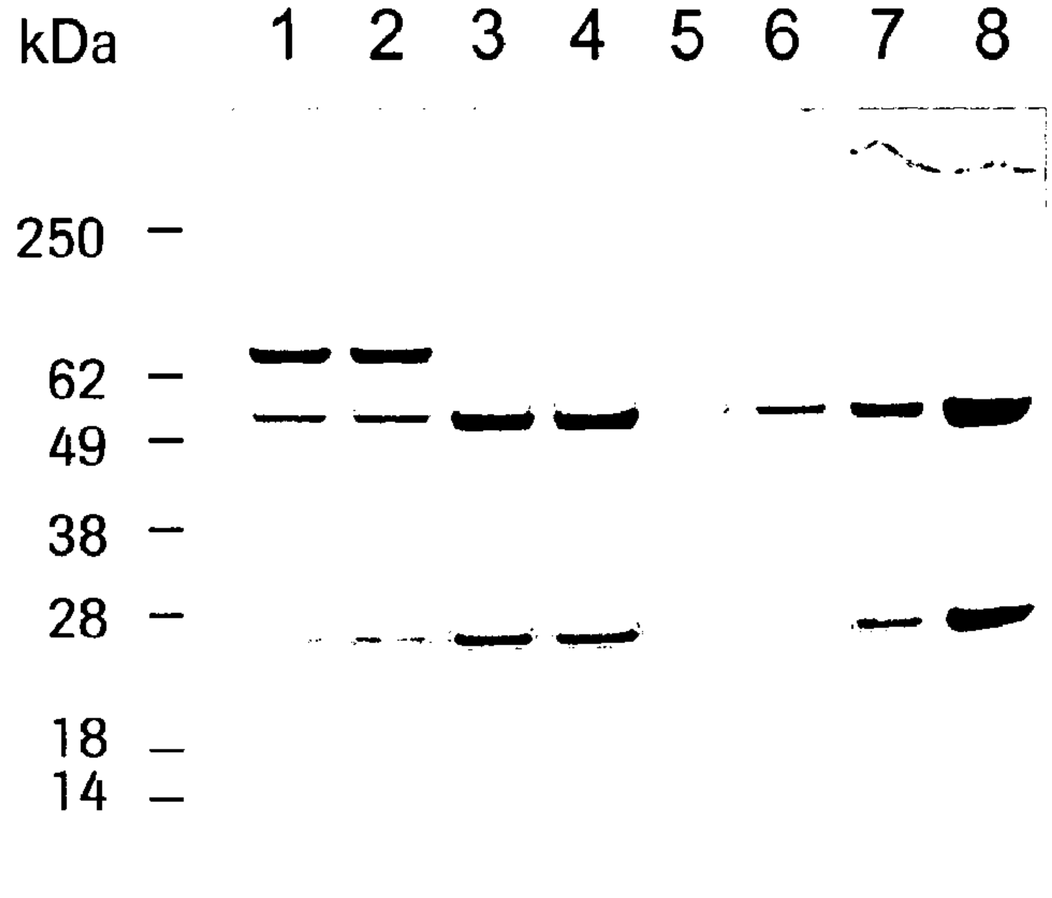
(51) **Cl.Int./Int.Cl.** ***C12N 15/13*** (2006.01), ***C07K 16/00*** (2006.01), ***C12N 15/67*** (2006.01), ***C12N 15/85*** (2006.01), ***C12P 21/08*** (2006.01)

(72) **Inventeurs/Inventors:**
GOEPFERT, ULRICH, DE;
HANSEN, SILKE, DE;
KNOETGEN, HENDRIK, DE;
KOPETZKI, ERHARD, DE;
PLOETTNER, OLIVER, DE

(73) **Propriétaire/Owner:**
F. HOFFMANN-LA ROCHE AG, CH

(74) **Agent:** GOWLING LAFLEUR HENDERSON LLP

(54) **Titre : MUTANT A CHAINE LOURDE CONDUISANT A UNE PRODUCTION AMELIOREE D'IMMUNOGLOBULINE**
(54) **Title: HEAVY CHAIN MUTANT LEADING TO IMPROVED IMMUNOGLOBULIN PRODUCTION**



(57) **Abrégé/Abstract:**
The current invention comprises a nucleic acid encoding the amino acid sequence of the C-terminal part of the C_H3-domain of an immunoglobulin of the class IgA or IgG, or the C-terminal part of the C_H4-domain of an immunoglobulin of the class IgE or IgM, wherein the glycine-lysine-dipeptide comprised in the amino acid sequence of the C-terminal part of the C_H3- or C_H4-domain is encoded by the nucleic acid ggaaaa, or the nucleic acid ggcaaa, or the nucleic acid gggaaa, or the nucleic acid gggaag, or the nucleic acid ggcaag, or the nucleic acid ggaaag.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 January 2009 (08.01.2009)

PCT

(10) International Publication Number
WO 2009/003623 A1

(51) International Patent Classification:
C07K 16/00 (2006.01) *C12N 15/67* (2006.01)
C12N 15/09 (2006.01) *C12N 15/00* (2006.01)

(21) International Application Number:
PCT/EP2008/005136

(22) International Filing Date: 25 June 2008 (25.06.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
07012774.1 29 June 2007 (29.06.2007) EP

(71) Applicant *(for all designated States except US)*: **F. HOFFMANN-LA ROCHE AG** [CH/CH]; Grenzacherstrasse 124, CH-4002 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **GOEPFERT, Ulrich** [DE/DE]; Elisabethstrasse 33, 80796 Muenchen (DE). **HANSEN, Silke** [DE/DE]; Egerlaender Strasse 3, 82393 Iffeldorf (DE). **KNOETGEN, Hendrik** [DE/DE]; Agathe-Fleissner-Weg 16, 82377 Penzberg (DE). **KOPETZKI, Erhard** [DE/DE]; Kastnerhofstrasse 21, 82377 Penzberg (DE). **PLOETTNER, Oliver** [DE/DE]; Krottenkopfstrasse 9, 81377 München (DE).

(74) Agent: **BURGER, Alexander**; Roche Diagnostics GmbH, Patent Department (TR-E), Postfach 11 52, 82372 Penzberg (DE).

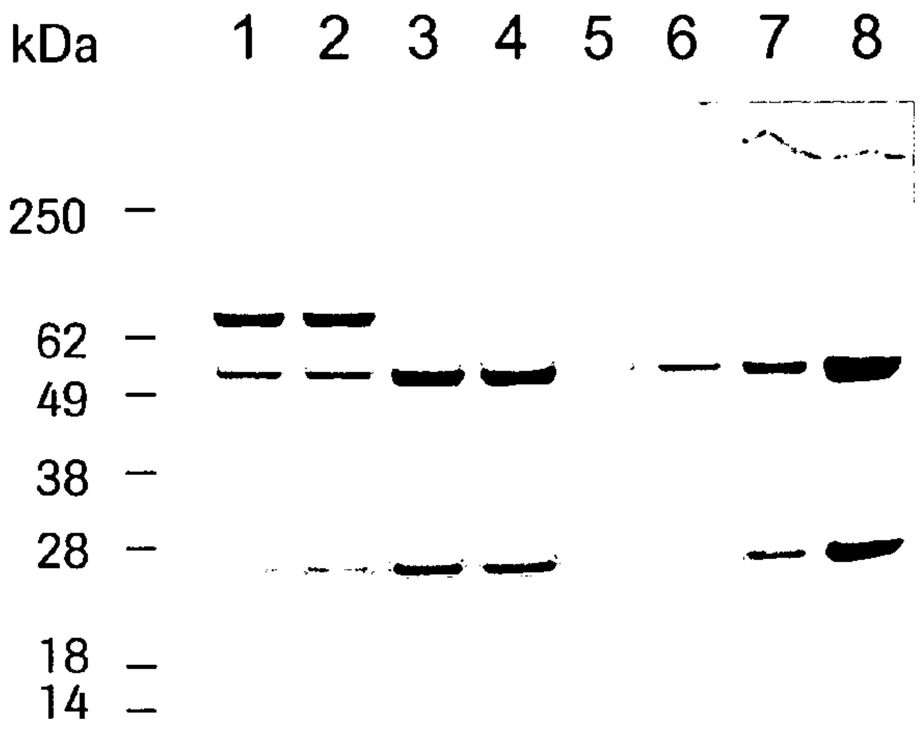
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, 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, 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, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: HEAVY CHAIN MUTANT LEADING TO IMPROVED IMMUNOGLOBULIN PRODUCTION

Figure 6



(57) Abstract: The current invention comprises a nucleic acid encoding the amino acid sequence of the C-terminal part of the C_H3-domain of an immunoglobulin of the class IgA or IgG, or the C-terminal part of the C_H4-domain of an immunoglobulin of the class IgE or IgM, wherein the glycine-lysine-dipeptide comprised in the amino acid sequence of the C-terminal part of the C_H3- or C_H4-domain is encoded by the nucleic acid ggaaaa, or the nucleic acid ggcaaa, or the nucleic acid gggaaa, or the nucleic acid gggaag, or the nucleic acid ggcaag, or the nucleic acid ggaaag.

WO 2009/003623 A1

WO 2009/003623 A1

Declaration under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report*

— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

Heavy Chain Mutant Leading To Improved Immunoglobulin Production

The current application describes methods and nucleic acids useful in the production of immunoglobulins in mammalian cells.

Background of the Invention

Expression systems for the production of recombinant polypeptides are well-known and reported in the state of the art literature. For the production of polypeptides used in pharmaceutical applications mammalian cells such as CHO cells, BHK cells, NS0 cells, Sp2/0 cells, COS cells, HEK cells, PER.C6® cells, and the like are employed. The nucleic acid encoding the polypeptide is introduced into the cell e.g. in a plasmid, such as, for example, an expression plasmid. The essential elements of an expression plasmid are a prokaryotic plasmid propagation unit, e.g. for Escherichia coli comprising an origin of replication and a selection marker, a eukaryotic selection marker, and one or more expression cassettes for the expression of the nucleic acid(s) of interest each of them comprising a promoter, a structural gene, and a transcription terminator including a polyadenylation signal. For transient expression in mammalian cells a mammalian origin of replication, such as the SV40 Ori or OriP, may be included. As a promoter a constitutive or inducible promoter can be selected. For optimized transcription a Kozak sequence may be included in the 5' untranslated region. For mRNA processing a polyadenylation signal may be included as well.

Proteins and especially immunoglobulins play an important role in today's medical portfolio. For human application every pharmaceutical substance has to meet distinct criteria. To ensure the safety of biopharmaceutical agents to humans substances, which would cause severe harm, have to be removed especially.

The splicing of mRNA is regulated by the occurrence of a splice donor site in combination with a splice acceptor site, which are located at the 5' end and 3' end of an intron, respectively. According to Watson et al. (Watson et al. (Eds), Recombinant DNA: A Short course, Scientific American Books, distributed by W.H. Freeman and Company, New York, New York, USA (1983)) are the consensus sequence of the 5' splice donor site ag|gtragt (exon|intron) and of the 3' splice acceptor site $(y)_nNcag|g$ (intron|exon) (r=purine base; y=pyrimidine base; n=integer; N=any natural base).

In 1980 first articles dealing with the origin of secreted and membrane bound forms of immunoglobulins have been published. The formation of the secreted (sIg) and the membrane bound (mIg) isoform results from alternative splicing of the heavy chain pre-mRNA. In the mIg isoform a splice donor site in the exon encoding the C-terminal domain of the secreted form (i.e. the C_H3 or C_H4 domain, respectively) and a splice acceptor site located at a distance downstream thereof are used to link the constant region with the downstream exons encoding the transmembrane domain.

A method to prepare synthetic nucleic acid molecules having reduced inappropriate or unintended transcriptional characteristics when expressed in a particular host cell is reported in WO 2002/016944. In WO 2006/042158 are reported nucleic acid molecules modified to enhance recombinant protein expression and/or reduce or eliminate mis-spliced and/or intron read through by products.

Therefore there exists a need for a recombinant production method for immunoglobulins with reduced by products.

Summary of the Invention

The current invention comprises in a first aspect a nucleic acid encoding the amino acid sequence of the C-terminal part of the C_H3-domain of an immunoglobulin of the class IgA or IgG, or the C-terminal part of the C_H4-domain of an immunoglobulin of the class IgE or IgM, wherein the glycine-lysine-dipeptide comprised in the primary amino acid sequence of the C-terminal part of the C_H3- or C_H4-domain is encoded by the nucleic acid ggaaaa, or the nucleic acid ggcaaa, or the nucleic acid gggaaa, or the nucleic acid ggaaag, or the nucleic acid ggcaag, or the nucleic acid gggaag.

In one embodiment the nucleic acid according to the invention encodes an amino acid sequence selected from the amino acid sequences SEQ ID NO: 1, 3, 4, 5, 6, 7, or 8. In another embodiment is the nucleic acid encoding the glycine-lysine-dipeptide preceded by the nucleotide g or a. In another embodiment is the nucleic acid encoding the glycine-lysine-dipeptide the nucleic acid ggaaaa, or the nucleic acid ggcaaa, or the nucleic acid gggaaa.

The second aspect of the current invention is a plasmid comprising the nucleic acid according to the invention, and the third aspect of the invention is a cell comprising the nucleic acid according to the invention.

A further aspect of the invention is a nucleic acid with the nucleotide sequence of SEQ ID NO: 17, 18, 19, 20, 21, 22, 23, 30, or 31.

The fifth aspect of the invention is a method for the production of an immunoglobulin in a mammalian cell comprising the following steps:

a) transfecting a mammalian cell with a nucleic acid encoding an immunoglobulin heavy chain comprising a nucleic acid of SEQ ID NO: 17, 18, 19, 20, 21, 22, 23, 30, or 31, which encodes the C-terminal part of the immunoglobulin heavy chain,

b) cultivating the transfected mammalian cell under conditions suitable for the expression of the immunoglobulin,

c) recovering the immunoglobulin from the culture or the cell.

In one embodiment is the mammalian cell a CHO cell, a BHK cell, a NS0 cell, a Sp2/0 cell, a COS cell, a HEK cell, or a PER.C6® cell. Preferably the mammalian cell is a CHO cell, or a BHK cell, or a PER.C6® cell. In another embodiment is the mammalian cell transfected with two nucleic acids, wherein the first nucleic acid comprises an expression cassette encoding an immunoglobulin light chain, and wherein the second nucleic acid comprises an expression cassette encoding an immunoglobulin heavy chain comprising a nucleic acid of SEQ ID NO: 17, 18, 19, 20, 21, 22, 23, 30, or 31 encoding the C-terminal part of the immunoglobulin heavy chain.

The final aspect of the invention is a method for improving the expression of an immunoglobulin in a mammalian cell, wherein the nucleic acid encoding the immunoglobulin heavy chain comprises the nucleic acid ggaaaa, or the nucleic acid ggcaaa, or the nucleic acid gggaaa, or the nucleic acid ggaaag, or the nucleic acid ggcaag, or the nucleic acid gggaag encoding the glycine-lysine-dipeptide contained in the C_H3- or C_H4-domain.

Detailed Description of the Invention

The current invention comprises a nucleic acid encoding the amino acid sequence of the C-terminal part of the C_H3-domain of an immunoglobulin of the class IgA or IgG, or the C-terminal part of the C_H4-domain of an immunoglobulin of the class IgE or IgM, wherein the glycine-lysine-dipeptide comprised in the amino acid

sequence of the C-terminal part of the C_H3- or C_H4-domain is encoded by the nucleic acid ggaaaa, or the nucleic acid ggcaaa, or the nucleic acid gggaaa, or the nucleic acid ggaaag, or the nucleic acid ggcaag, or the nucleic acid gggaag, and wherein the nucleic acid encoding the glycine-lysine-dipeptide is optionally preceded either by the nucleotide g or the nucleotide a.

Methods and techniques useful for carrying out the current invention are known to a person skilled in the art and are described e.g. in Ausubel, F.M., ed., Current Protocols in Molecular Biology, Volumes I to III (1997), and Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). As known to a person skilled in the art enables the use of recombinant DNA technology the production of numerous derivatives of a nucleic acid and/or polypeptide. Such derivatives can, for example, be modified in one individual or several positions by substitution, alteration, exchange, deletion, or insertion. The modification or derivatisation can, for example, be carried out by means of site directed mutagenesis. Such modifications can easily be carried out by a person skilled in the art (see e.g. Sambrook, J., et al., Molecular Cloning: A laboratory manual (1999) Cold Spring Harbor Laboratory Press, New York, USA). The use of recombinant technology enables a person skilled in the art to transform various host cells with heterologous nucleic acid(s). Although the transcription and translation, i.e. expression, machinery of different cells use the same elements, cells belonging to different species may have among other things a different so-called codon usage. Thereby identical polypeptides (with respect to amino acid sequence) may be encoded by different nucleic acid(s). Also, due to the degeneracy of the genetic code, different nucleic acids may encode the same polypeptide.

A "nucleic acid" as used herein, refers to a polymeric molecule consisting of the individual nucleotides (also called bases) a, c, g, and t (or u in RNA), for example to DNA, RNA, or modifications thereof. This polynucleotide molecule can be a naturally occurring polynucleotide molecule or a synthetic polynucleotide molecule or a combination of one or more naturally occurring polynucleotide molecules with one or more synthetic polynucleotide molecules. Also encompassed by this definition are naturally occurring polynucleotide molecules in which one or more nucleotides are changed (e.g. by mutagenesis), deleted, or added. A nucleic acid can either be isolated, or integrated in another nucleic acid, e.g. in an expression

cassette, a plasmid, or the chromosome of a host cell. A nucleic acid is likewise characterized by its nucleic acid sequence consisting of individual nucleotides.

To a person skilled in the art procedures and methods are well known to convert an amino acid sequence, e.g. of a polypeptide, into a corresponding nucleic acid sequence encoding this amino acid sequence. Therefore, a nucleic acid is characterized by its nucleic acid sequence consisting of individual nucleotides and likewise by the amino acid sequence of a polypeptide encoded thereby.

The term “plasmid” includes e.g. shuttle and expression plasmids as well as transfection plasmids. The term “vector” is used synonymously for „plasmids“ within this application. Typically, a “plasmid” will also comprise an origin of replication (e.g. the ColE1 or oriP origin of replication) and a selection marker (e.g. an ampicillin, kanamycin, tetracycline, or chloramphenicol selection marker), for replication and selection, respectively, of the plasmid in bacteria.

An "expression cassette" refers to a construct that contains the necessary regulatory elements, such as promoter and polyadenylation site, for expression of at least the contained nucleic acid in a cell.

A “selection marker” is a nucleic acid that allows cells carrying the selection marker to be specifically selected for or against, in the presence of a corresponding selection agent. Typically, a selection marker will confer resistance to a drug or compensate for a metabolic or catabolic defect in the host cell. A selection marker can be positive, negative, or bifunctional. A useful positive selection marker is an antibiotic resistance gene. This selection marker allows the host cell transformed therewith to be positively selected for in the presence of the corresponding selection agent, e.g. the antibiotic. A non-transformed host cell is not capable to grow or survive under the selective culture conditions, i.e. in the presence of the selection agent, in the culture. Positive selection markers allow selection for cells carrying the marker, whereas negative selection markers allow cells carrying the marker to be selectively eliminated. Selection markers used with eukaryotic cells include, e.g., the genes for aminoglycoside phosphotransferase (APH), such as e.g. the hygromycin (hyg), neomycin (neo), and G418 selection markers, dihydrofolate reductase (DHFR), thymidine kinase (tk), glutamine synthetase (GS), asparagine synthetase, tryptophan synthetase (selection agent indole), histidinol dehydrogenase (selection agent histidinol D), and nucleic acids conferring resistance to puromycin,

bleomycin, phleomycin, chloramphenicol, Zeocin, and mycophenolic acid. Further marker genes are reported e.g. in WO 92/08796 and WO 94/28143.

The term "expression" as used herein refers to transcription and/or translation processes occurring within a cell. The level of transcription of a nucleic acid sequence of interest in a cell can be determined on the basis of the amount of corresponding mRNA that is present in the cell. For example, mRNA transcribed from a sequence of interest can be quantitated by RT-PCR or by Northern hybridization (see Sambrook et al., 1989, supra). Polypeptides encoded by a nucleic acid of interest can be quantitated by various methods, e.g. by ELISA, by assaying for the biological activity of the polypeptide, or by employing assays that are independent of such activity, such as Western blotting or radioimmunoassay, using immunoglobulins that recognize and bind to the polypeptide (see Sambrook et al., 1989, supra).

The term “cell” or "host cell" refers to a cell into which a nucleic acid, e.g. encoding a heterologous polypeptide, can be or is introduced / transfected. The term „cell“ includes both prokaryotic cells, which are used for propagation of plasmids, and eukaryotic cells, which are used for the expression of a nucleic acid. Preferably, the eukaryotic cells are mammalian cells. Preferably the mammalian cell is selected from the group of mammalian cells comprising CHO cells (e.g. CHO K1, CHO DG44), BHK cells, NS0 cells, SP2/0 cells, HEK 293 cells, HEK 293 EBNA cells, PER.C6® cells, and COS cells. As used herein, the expression "cell" includes the subject cell and its progeny. Thus, the words "transformant" and "transformed cell" include the primary subject cell and cells in cultures derived there from without regard for the number of transfers. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Variant progeny that have the same function or biological activity as screened for in the originally transformed cell are included.

A "polypeptide" is a polymer consisting of amino acids joined by peptide bonds, whether produced naturally or synthetically. Polypeptides of less than about 20 amino acid residues may be referred to as "peptides", whereas molecules consisting of two or more polypeptides or comprising one polypeptide of more than 100 amino acid residues may be referred to as “proteins”. A polypeptide may also comprise non-amino acid components, such as carbohydrate groups, metal ions, or carboxylic acid esters. The non-amino acid components may be added by the cell, in which the polypeptide is expressed, and may vary with the type of cell.

Polypeptides are defined herein in terms of their amino acid backbone structure or the nucleic acid encoding the same. Additions such as carbohydrate groups are generally not specified, but may be present nonetheless.

The term "amino acid" as used within this application denotes a group of carboxy α-amino acids, which directly or in form of a precursor can be encoded by a nucleic acid. The individual amino acids are encoded by nucleic acids each consisting of three nucleotides, so called codons or base-triplets. Each amino acid is encoded by at least one codon. For example the amino acid glycine can be encoded by each of the four nucleic acids (codons) gga, ggc, ggg, and ggt, whereas the amino acid lysine can only be encoded by the two nucleic acids aaa, and aag. This phenomenon is known as "degeneration of the genetic code". The group of amino acids comprises alanine (three letter code: ala, one letter code: A), arginine (arg, R), asparagine (asn, N), aspartic acid (asp, D), cysteine (cys, C), glutamine (gln, Q), glutamic acid (glu, E), glycine (gly, G), histidine (his, H), isoleucine (ile, I), leucine (leu, L), lysine (lys, K), methionine (met, M), phenylalanine (phe, F), proline (pro, P), serine (ser, S), threonine (thr, T), tryptophan (trp, W), tyrosine (tyr, Y), and valine (val, V).

As used herein, the term "immunoglobulin" denotes a protein consisting of one or more polypeptides substantially encoded by immunoglobulin genes. This definition includes variants such as mutated forms, i.e. forms with substitutions, deletions, and insertions of one or more amino acids, N-terminally truncated forms, fused forms, chimeric forms, as well as humanized forms. The recognized immunoglobulin genes include the different constant region genes as well as the myriad immunoglobulin variable region genes from, e.g., primates and rodents. Monoclonal immunoglobulins are preferred. Each of the heavy and light polypeptide chains of an immunoglobulin may comprise a constant region (generally the carboxyl terminal portion).

The term "monoclonal immunoglobulin" as used herein refers to an immunoglobulin obtained from a population of substantially homogeneous immunoglobulins, i.e. the individual immunoglobulins comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal immunoglobulins are highly specific, being directed against a single antigenic site. Furthermore, in contrast to polyclonal immunoglobulin preparations, which include different immunoglobulins directed against different antigenic sites (determinants or epitopes), each monoclonal immunoglobulin is directed against a single antigenic site on the antigen. In

addition to their specificity, the monoclonal immunoglobulins are advantageous in that they may be synthesized uncontaminated by other immunoglobulins. The modifier "monoclonal" indicates the character of the immunoglobulin as being obtained from a substantially homogeneous population of immunoglobulins and is not to be construed as requiring production of the immunoglobulin by any particular method.

"Humanized" forms of non-human (e.g. rodent) immunoglobulins are chimeric immunoglobulins that contain partial sequences derived from non-human immunoglobulin and from human immunoglobulin. For the most part, humanized immunoglobulins are derived from a human immunoglobulin (recipient immunoglobulin), in which residues from a hypervariable region are replaced by residues from a hypervariable region of a non-human species (donor immunoglobulin), such as mouse, rat, rabbit, or non-human primate, having the desired specificity and affinity (see e.g. Morrison, S.L., et al., Proc. Natal. Acad. Sci. USA 81 (1984) 6851-6855; US 5,202,238; US 5,204,244). In some instances, framework region (FR) residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized immunoglobulins may comprise further modifications, e.g. amino acid residues that are not found in the recipient immunoglobulin or in the donor immunoglobulin. Such modifications result in variants of such recipient or donor immunoglobulin, which are homologous but not identical to the corresponding parent sequence. These modifications are made to further refine immunoglobulin performance. In general, the humanized immunoglobulin will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable loops correspond to those of a non-human donor immunoglobulin and all or substantially all of the FRs are those of a human recipient immunoglobulin. The humanized immunoglobulin optionally will also comprise at least a portion of an immunoglobulin constant region, typically that of a human immunoglobulin. Methods for humanizing non-human immunoglobulin have been described in the art. Preferably, a humanized immunoglobulin has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers by substituting hypervariable region sequences for the corresponding sequences of a non-human immunoglobulin. Accordingly, such "humanized" immunoglobulins are chimeric

immunoglobulins, wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized immunoglobulins are typically human immunoglobulins in which some hypervariable region residues and possibly some framework region residues are substituted by residues from analogous sites in rodent or non-human primate immunoglobulins.

Depending on the amino acid sequence of the constant region of their heavy chains, immunoglobulins are divided in the classes: IgA, IgD, IgE, IgG, and IgM. Some of these may be further divided into subclasses (isotypes), e.g. IgG1, IgG2, IgG3, and IgG4, IgA1 and IgA2. According to the heavy chain constant regions the different classes of immunoglobulins heavy chains are denoted as α-, δ-, ε-, γ-, and μ-chain respectively. The constant region of a heavy chain of a full length human immunoglobulin of the class IgA, IgD, and IgG is constituted of a constant domain 1 (herein denoted as C_H1), a hinge region, a constant domain 2 (C_H2), and a constant domain 3 (C_H3). Human immunoglobulins of the class IgE and IgM comprise an additional fourth constant domain (C_H4). Furthermore are immunoglobulins of the class IgM polymers comprising multiple immunoglobulins, e.g. five or six, covalently linked by disulfide bonds. The C-terminal constant domain amino acid sequence of these different human immunoglobulin classes are listed in Table 1. The first amino acid of SEQ ID NO: 01 to 08 may be present or not, because this amino acid is encoded by two exons, the exon encoding the C-terminal constant domain and the exon encoding the preceding domain.

Table 1: C-terminal amino acid sequence of the different immunoglobulin classes. The glycine-lysine-dipeptides are underlined.

immunoglobulin class (C-terminal constant domain)	amino acid sequence of the C-terminal part (in N-terminal to C-terminal direction)	SEQ ID NO:
IgA1 (C_H3), IgA2 (C_H3)	GNTFRPEVHL LPPPSEELAL NELVTLTCLA RGFSPKDVLV RWLQGSQELP REKYL TWASR QEPSQGTTTF AVTSILRVAA EDWKKGDTFS CMVGHEALPL AFTQK TIDRL A<u>GK</u>PTHVNVS VVMAEVDGTC Y	01

immunoglobulin class (C-terminal constant domain)	amino acid sequence of the C-terminal part (in N-terminal to C-terminal direction)	SEQ ID NO:
IgD (C_H3)	AAQAPVKLSLN LLASSDPPEA ASWLLCEVSG FSPPNILLMW LEDQREVNTS GFAPARPPPQ PGSTTFWAWS VLRVPAPPSP QPATYTCVVS HEDSRTLLNA SRSLEVSY	02
IgE (C_H4)	GPRAAPEVYA FATPEWPGSR DKRTLACLIQ NFMPEDISVQ WLHNEVQLPD ARHSTTQPRK TKGSGFFVFS RLEVTRAEWE QKDEFICRAV HEAASPSQTV QRAVSVNPGK	03
IgM (C_H4)	GVALHRPDVY LLPPAREQLN LRESATITCL VTGFSPADVF VQWMQRGQPL SPEKYVTSAP MPEPQAPGRY FAHSILTVSE EEWNTGETYT CVVAHEALPN RVT ERTVDK STGKPTLYNV SLVMSDTAGT CY	04
IgG1 (C_H3)	GQPREPQVYT LPPSRDELTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTPPVLDS DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNHYTQKS LSLSPGK	05
IgG2 (C_H3)	GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTPPMLDS DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNHYTQKS LSLSPGK	06
IgG3 (C_H3)	GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESSGQPENN YNTTPPMLDS DGSFFLYSKL TVDKSRWQQG NIFSCSVMHE ALHNRFTQKS LSLSPGK	07

immunoglobulin class (C-terminal constant domain)	amino acid sequence of the C-terminal part (in N-terminal to C-terminal direction)	SEQ ID NO:
IgG4 (C_H3)	GQPREPQVYT LPPSQEEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTPPVLDS DGSFFLYSRL TVDKSRWQEG NVFSCSVMHE ALHNHYTQKS LSLSLGK	08

The term "C-terminal part of the C_H3-domain of an immunoglobulin of the class IgA or IgG, or the C-terminal part of the C_H4-domain of an immunoglobulin of the class IgE or IgM" denotes the amino acid sequences of an immunoglobulin heavy chain, which is located at the C-terminal end of a full-length or naturally occurring immunoglobulin heavy chain, whereby the C-terminus of said C-terminal part is identical to the C-terminus of the primary amino acid sequence of the immunoglobulin heavy chain. The term "primary amino acid sequence" denotes the amino acid sequence of an immunoglobulin heavy chain after the translation of the corresponding mRNA. This primary amino acid sequence may further be modified in the expressing cell after the mRNA translation e.g. by peptidases cleaving one or more C-terminal amino acids from the primary amino acid sequence. Therefore the primary amino acid sequence and the secreted amino acid sequence may not be identical but may differ by some amino acids at the C-terminus. In one embodiment the C-terminal part comprises at least the 100 C-terminal amino acids of an immunoglobulin heavy chain primary amino acid sequence, or preferably at least the 50 C-terminal amino acids of an immunoglobulin heavy chain primary amino acid sequences, or preferably the at least 20 C-terminal amino acids of an immunoglobulin heavy chain primary amino acid sequence. In one embodiment the nucleic acid according to the invention is encoding the amino acid sequence of the C-terminal part of the C_H3-domain of an immunoglobulin of the class IgG, or the C-terminal part of the C_H4-domain of an immunoglobulin of the class IgE. In a further emobiment the nucleic acid according to the invention is encoding the amino acid sequence of the C-terminal part of the C_H3 domain of an immunoglobulin of the class IgG.

The C-terminal constant domain amino acid sequences of the different human immunoglobulins are encoded by corresponding DNA sequences. In the genome

these DNA sequences contain coding (exonic) and non-coding (intronic) sequences. After transcription of the DNA to the pre-mRNA the pre-mRNA also contains these intronic and exonic sequences. Prior to translation the non-coding intronic sequences are removed during mRNA processing by splicing them out of the primary mRNA transcript to generate the mature mRNA. The splicing of the primary mRNA is controlled by a splice donor site in combination with a properly spaced apart splice acceptor site. The splice donor site is located at the 5' end and the splice acceptor site is located at the 3' end of an intronic sequence and both are partly removed during the pre-mRNA splicing process.

The term "properly spaced apart" denotes that a splice donor site and a splice acceptor site in a nucleic acid are arranged in such a way that all required elements for the splicing process are available and are in an appropriate position to allow the splicing process to take place.

The current invention comprises a nucleic acid encoding the amino acid sequence of the C-terminal part of the C_H3-domain of an immunoglobulin of the class IgA or IgG, or the C-terminal part of the C_H4-domain of an immunoglobulin of the class IgE or IgM, wherein the glycine-lysine-dipeptide comprised in said amino acid sequence of the C-terminal part of the C_H3- or C_H4-domain is encoded by the nucleic acid ggaaaa, or the nucleic acid ggcaaa, or the nucleic acid gggaaa.

It has been surprisingly found that with the nucleic acid according to the invention unwanted by product formation by unwanted splicing events can be reduced.

The term "glycine-lysine-dipeptide" as used within the current application denotes the peptide comprising in N-terminal to C-terminal direction the two amino acids glycine and lysine linked by a peptide bond. The term "glycine-lysine-dipeptide" as used within this application denotes a dipeptide fraction of a larger polypeptide or protein, which can be found at the beginning, within, or at the end of the larger polypeptide. The amino acid lysine can be encoded by the nucleic acids aaa and aag. Therefore it is another embodiment of the current invention that the glycine-lysine-dipeptide comprised in the amino acid sequence of the C-terminal part of the C_H3- or C_H4-domain of an immunoglobulin heavy chain is encoded by the nucleic acid ggaaag, or the nucleic acid ggcaag, or the nucleic acid gggaag. In another embodiment the nucleic acid encoding the glycine-lysine-dipeptide is preceded by the nucleotide a or g.

Nucleic acid sequences encoding the C-terminal constant domain of different human immunoglobulin classes and subclasses are listed in Table 2. For the C_H3 domain of human IgG1 and IgG2 two variant forms are known. In one embodiment encodes the nucleic acid a part of the C-terminal constant domain of an immunoglobulin heavy chain.

Table 2: Nucleic acid sequences encoding the C-terminal part of the heavy chain of the different human immunoglobulin classes.

immunoglobulin class (C-terminal constant domain)	nucleic acid sequence encoding the C-terminal part of the amino acid sequence of an immunoglobulin heavy chain	SEQ ID NO:
IgA1 (C_H3), IgA2 (C_H3)	ggcttcagcc ccaaggacgt gctggttcgc tggctgcagg ggtcacagga gctgccccgc gagaagtacc tgacttgggc atcccggcag gagcccagcc agggcaccac caccttcgct gtgaccagca tactgcgcgt ggcagccgag gactggaaga agggggacac cttctcctgc atggtgggcc acgaggccct gccgctggcc ttcacacaga agaccatcga ccgcttggcg ggtaaaccca cccatgtcaa tgtgtctgtt gtcatggcgg aggtggacgg cacctgctac	09
IgD (C_H3)	taccacccaa cgtccgtgac tgtcacctgg tacatgggga cacagagcca gccccagaga accttccctg agatacaaag acgggacagc tactacatga caagcagcca gctctccacc cccctccagc agtggcgcca aggcgagtac aaatgcgtgg tccagcacac cgccagcaag agtaagaagg agatcttccg ctggccaggt aggtcgcacc ggagatcacc cagaagggcc ccccaggacc cccagcacct tccactcagg gcctgaccac aaagacagaa gcaagggctg	10
IgE (C_H4)	tttgcgacgc cggagtggcc ggggagccgg gacaagcgca ccctcgcctg cctgatccag aacttcatgc ctgaggacat ctcggtgcag tggctgcaca acgaggtgca gctcccggac gcccggcaca gcacgacgca gccccgcaag accaagggct ccggcttctt cgtcttcagc cgcctggagg tgaccagggc cgaatgggag cagaaagatg agttcatctg ccgtgcagtc catgaggcag cgagcccctc acagaccgtc cagcgagcgg tgtctgtaaa tcccggtaaa	11
IgM (C_H4)	acgggcttct ctcccgcgga cgtcttcgtg cagtggatgc agaggggggca gcccttgtcc ccggagaagt atgtgaccag cgccccaatg cctgagcccc aggccccagg ccggtacttc gcccacagca tcctgaccgt gtccgaagag gaatggaaca cgggggagac ctacacctgc gtggtggccc atgaggccct gcccaacagg gtcaccgaga ggaccgtgga caagtccacc ggtaaaccca ccctgtacaa cgtgtccctg gtcatgtccg acacagctgg cacctgctac	12

immunoglobulin class (C-terminal constant domain)	nucleic acid sequence encoding the C-terminal part of the amino acid sequence of an immunoglobulin heavy chain	SEQ ID NO:
IgG1 (C_H3) variant 1	gtgtacaccc tgccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaaa	13
IgG1 (C_H3) variant 2	gggcagcccc gagaaccaca ggtgtacacc ctgccccat cccgggatga gctgaccaag aaccaggtca gcctgacctg cctggtcaaa ggcttctatc ccagcgacat cgccgtggag tgggagagca atgggcagcc ggagaacaac tacaagacca cgcctcccgt gctggactcc gacggctcct tcttcctcta cagcaagctc accgtggaca agagcaggtg gcagcagggg aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac acagaagagc ctctccctgt ctccgggtaa a	28
IgG2 (C_H3) variant 1	gtgtacaccc tgccccatc ccgggaggag atgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac acctcccatg ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaaa	14
IgG2 (C_H3) variant 2	gggcagcccc gagaaccaca ggtgtacacc ctgccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctggtcaaa ggcttctacc ccagcgacat ctccgtggag tgggagagca atgggcagcc ggagaacaac tacaagacca cacctcccat gctggactcc gacggctcct tcttcctcta cagcaagctc accgtggaca agagcaggtg gcagcagggg aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac gcagaagagc ctctccctgt ctccgggtaa a	29
IgG3 (C_H3)	gtgtacaccc tgccccatc ccgggaggag atgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcag cgggcagccg gagaacaact acaacaccac gcctcccatg ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg cagcagggga acatcttctc atgctccgtg atgcatgagg ctctgcacaa ccgcttcacg cagaagagcc tctccctgtc tccgggtaaa	15

immunoglobulin class (C-terminal constant domain)	nucleic acid sequence encoding the C-terminal part of the amino acid sequence of an immunoglobulin heavy chain	SEQ ID NO:
IgG4 (C_H3)	gtgtacaccc tgccccatc ccaggaggag atgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac agcaggctaa ccgtggacaa gagcaggtgg caggagggga atgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tctgggtaaa	16

In one embodiment the nucleic acid according to the invention encodes an amino acid sequence selected from the amino acid sequences of SEQ ID NO: 1, 3, 4, 5, 6, 7, or 8.

In one embodiment the nucleic acid encodes a part of the C-terminal constant domain of an immunoglobulin heavy chain of the class IgA, IgE, IgM, or IgG and is selected from the nucleic acids of SEQ ID NO: 17, 18, 19, 20, 21, 22, or 23, or 30, or 31. In another embodiment encodes the nucleic acid a part of the C-terminal constant domain of an immunoglobulin heavy chain of the class IgA, IgE, IgM, or IgG and is selected from the nucleic acids of SEQ ID NO: 17, 18, 19, 22, 23, 30, or 31. In a further embodiment encodes the nucleic acid a part of the C-terminal constant domain of an immunoglobulin heavy chain of the class IgA, IgE, IgM, or IgG and is selected from the nucleic acids of SEQ ID NO: 17, 18, 19, 22, or 23.

The nucleic acids with the nucleotide sequence of SEQ ID NO: 17 to 23 and 30 to 31 are also an aspect of the current invention. In one embodiment the nucleotide sequences of SEQ ID NO: 17, 18, 19, 22, 23, 30, or 31 are an aspect of the current invention. In another embodiment the nucleotide sequences of SEQ ID NO: 17, 18, 19, 22, or 23 are an aspect of the current invention.

In one embodiment encodes the nucleic acid the C-terminal constant domain of an immunoglobulin of the class IgA, IgE, IgM, or IgG, and comprises a nucleic acid selected from the nucleic acids of SEQ ID NO: 17, 18, 19, 20, 21, 22, or 23, or 30, or 31, which encodes a part of the C-terminal domain of the immunoglobulin heavy chain. In another embodiment encodes the nucleic acid the C-terminal constant domain of an immunoglobulin of the class IgA, IgE, IgM, or IgG, and comprises a

nucleic acid selected from the nucleic acids of SEQ ID NO: 17, 18, 19, 22, 23, 30, or 31, which encodes a part of the C-terminal domain of the immunoglobulin heavy chain. In a further embodiment encodes the nucleic acid the C-terminal constant domain of an immunoglobulin of the class IgA, IgE, IgM, or IgG, and comprises a nucleic acid selected from the nucleic acids of SEQ ID NO: 17, 18, 19, 22, or 23, which encodes a part of the C-terminal domain of the immunoglobulin heavy chain.

Table 3: Nucleic acid sequences according to the invention encoding a part of the C-terminal constant domain amino acid sequence of an immunoglobulin heavy chain of different classes.

immunoglobulin class	nucleic acid sequence encoding the C-terminal amino acid sequence	SEQ ID NO:
IgA	ggcaaaccca cccatgtcaa tgtgtctgtt gtcatggcgg aggtggacgg cacctgctac	17
IgE	catgaggcag cgagcccctc acagaccgtc cagcgagcgg tgtctgtaaa tcccggcaaa	18
IgM	ggcaaaccca ccctgtacaa cgtgtccctg gtcatgtccg acacagctgg cacctgctac	19
IgG1 variant 1	atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggcaaa	20
IgG1 variant 2	atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tccgggcaaa	30
IgG2 variant 1	atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggcaaa	21
IgG2 variant 2	atctccgtgg agtgggagag caatgggcag ccggagaaca actacaagac cacacctccc atgctggact ccgacggctc cttcttcctc tacagcaagc tcaccgtgga caagagcagg tggcagcagg ggaacgtctt ctcatgctcc gtgatgcatg aggctctgca caaccactac acgcagaaga gcctctccct gtctccgggc aaa	31
IgG3	atgcatgagg ctctgcacaa ccgcttcacg cagaagagcc tctccctgtc tccgggcaaa	22
IgG4	atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tctgggcaaa	23

The nucleic acid according to the invention encodes at least a part of the C-terminal constant domain of an immunoglobulin of the class IgA, IgE, IgM, or IgG. The

term "C-terminal constant domain" denotes either the C_H3-domain of an immunoglobulin heavy chain of the class IgA or IgG, or the C_H4-domain of an immunoglobulin heavy chain of the class IgE or IgM. The expression "a part of" denotes a C-terminal fraction of the C-terminal constant domain of an immunoglobulin heavy chain of the class IgA, IgE, IgM, or IgG, of at least 20 consecutive amino acids, or of at least 50 consecutive amino acids, or of at least 100 consecutive amino acids of the primary amino acid sequence counted from the C-terminus in direction to the N-terminus of the immunoglobulin heavy chain.

Recombinant production of immunoglobulins is well-known in the state of the art and described, for example, in the review articles of Makrides, S.C., Protein Expr. Purif. 17 (1999) 183-202; Geisse, S., et al., Protein Expr. Purif. 8 (1996) 271-282; Kaufman, R.J., Mol. Biotechnol. 16 (2000) 151-160 Werner, R.G., Arzneimittelforschung - Drug Research 48 (1998) 870-880.

For the production of an immunoglobulin comprising an amino acid sequence encoded by a nucleic acid according to the invention the invention further comprises a method for the production of an immunoglobulin in a mammalian cell comprising the following steps:

a) transfecting the mammalian cell with a nucleic acid encoding an immunoglobulin heavy chain, wherein a nucleic acid selected from the nucleic acids of SEQ ID NO: 17, 18, 19, 20, 21, 22, or 23, or 30, or 31 encodes a part of the C-terminal domain of the immunoglobulin heavy chain,

b) cultivating the transfected mammalian cell under conditions suitable for the expression of the immunoglobulin,

c) recovering the immunoglobulin from the culture or the cell.

The term "under conditions suitable for the expression of the immunoglobulin" denotes conditions, which are used for the cultivation of a mammalian cell expressing an immunoglobulin and which are known to or can easily be determined by a person skilled in the art. It is also known to a person skilled in the art that these conditions may vary depending on the type of mammalian cell cultivated and type of immunoglobulin expressed. In general the mammalian cell is cultivated at a temperature, e.g. between 20°C and 40°C, and for a period of time sufficient to allow effective protein production of the immunoglobulin, e.g. for 4 to 28 days.

In one embodiment the transfecting the mammalian cell is with a nucleic acid encoding an immunoglobulin heavy chain, wherein a nucleic acid selected from the nucleic acids of SEQ ID NO: 17, 18, 19, 22, 23, 30, or 31 encodes a part of the C-terminal domain of the immunoglobulin heavy chain. In a further embodiment the transfecting the mammalian cell is with a nucleic acid encoding an immunoglobulin heavy chain, wherein a nucleic acid selected from the nucleic acids of SEQ ID NO: 17, 18, 19, 22, or 23 encodes a part of the C-terminal domain of the immunoglobulin heavy chain. In one embodiment the mammalian cell is transfected with a, i.e. one, nucleic acid comprising an expression cassette encoding an immunoglobulin heavy chain and an expression cassette encoding an immunoglobulin light chain. In another embodiment is the mammalian cell transfected with two nucleic acids, one comprising an expression cassette encoding an immunoglobulin light chain, and one comprising an expression cassette encoding an immunoglobulin heavy chain.

The immunoglobulin produced with the method according to the invention is preferably a heterologous immunoglobulin. The term „heterologous immunoglobulin“ denotes an immunoglobulin which is not naturally produced by said mammalian cell. The immunoglobulin produced according to the method of the invention is produced by recombinant means. Such methods are widely known in the state of the art and comprise protein expression in prokaryotic and eukaryotic cells with subsequent recovery and isolation of the heterologous immunoglobulin, and usually purification to a pharmaceutically acceptable purity. For the production, i.e. expression, of an immunoglobulin a nucleic acid encoding the light chain and a nucleic acid encoding the heavy chain, which is comprising the nucleic acid according to the invention, are inserted each into an expression cassette by standard methods. Nucleic acids encoding immunoglobulins are readily isolated and sequenced using conventional procedures. Hybridoma cells can serve as a source of such nucleic acids. The expression cassettes may be inserted into an expression plasmid(s), which is (are) then transfected into host cells, which do not otherwise produce immunoglobulins. Expression is performed in appropriate prokaryotic or eukaryotic host cells and the immunoglobulin is recovered from the cells after lysis or from the culture supernatant.

Different methods are well established and widespread used for protein recovery and purification, such as affinity chromatography with microbial proteins (e.g. protein A or protein G affinity chromatography), ion exchange chromatography

(e.g. cation exchange (carboxymethyl resins), anion exchange (amino ethyl resins) and mixed-mode exchange), thiophilic adsorption (e.g. with beta-mercaptoethanol and other SH ligands), hydrophobic interaction or aromatic adsorption chromatography (e.g. with phenyl-sepharose, aza-arenophilic resins, or m-aminophenylboronic acid), metal chelate affinity chromatography (e.g. with Ni(II)- and Cu(II)-affinity material), size exclusion chromatography, and electrophoretical methods (such as gel electrophoresis, capillary electrophoresis) (Vijayalakshmi, M.A. Appl. Biochem. Biotech. 75 (1998) 93-102).

The current invention further comprises a nucleic acid encoding an immunoglobulin heavy chain comprising the nucleic acid according to the invention. Furthermore comprises the invention a plasmid comprising the nucleic acid according to the invention and a cell comprising this plasmid.

Another aspect of the current invention is a method for improving the expression of an immunoglobulin in a mammalian cell, wherein the nucleic acid encoding the immunoglobulin heavy chain comprises the nucleic acid ggaaaa, or the nucleic acid ggcaaa, or the nucleic acid gggaaa, or the nucleic acid ggaaag, or the nucleic acid ggcaag, or the nucleic acid gggaag, encoding the glycine-lysine-dipeptide contained in the C_H3- or C_H4-domain. With this nucleic acid unwanted splicing events can be reduced or suppressed.

In one embodiment of the invention the immunoglobulin heavy chain is either an immunoglobulin heavy chain of a human antibody of the subclass IgG4 or an immunoglobulin heavy chain of a human antibody of the subclass IgG1, IgG2, or IgG3. In one embodiment the immunoglobulin heavy chain is a human immunoglobulin heavy chain and preferably either from human IgG4 subclass or a mutated immunoglobulin heavy chain from human IgG1 subclass. In another embodiment the immunoglobulin heavy chain is from human IgG1 subclass with mutations L234A and L235A. In a further embodiment the immunoglobulin heavy chain is a human IgG4 immunoglobulin heavy chain with the mutation S228P. In on embodiment the immunoglobulin heavy chain is of IgG4 subclass or of IgG1 or IgG2 subclass, with a mutation in L234, L235, and/or D265, and/or contains the PVA236 mutation. In other embodiments the mutations are S228P, L234A, L235A, L235E, and/or PVA236 (PVA236 means that the amino acid sequence ELLG (given in one letter amino acid code) from amino acid position 233 to 236 of IgG1 or EFLG of IgG4 is replaced by PVA). Preferred are the mutations S228P of IgG4, and L234A and L235A of IgG1 (numbering according to EU index of Kabat).

"Regulatory elements" as used herein, refer to nucleotide sequences present in cis, necessary for transcription and/or translation of the nucleic acid sequence encoding a polypeptide of interest. The transcriptional regulatory elements normally comprise a promoter upstream of the nucleic acid sequence to be expressed, transcriptional initiation and termination sites, and a polyadenylation signal sequence. The term "transcriptional initiation site" refers to the nucleic acid base in the nucleic acid corresponding to the first nucleic acid incorporated into the primary transcript, i.e. the mRNA precursor; the transcriptional initiation site may overlap with the promoter sequence. The term "transcriptional termination site" refers to a nucleotide sequence normally represented at the 3' end of a gene of interest to be transcribed, that causes RNA polymerase to terminate transcription. The polyadenylation signal sequence, or poly-A addition signal provides the signal for the cleavage at a specific site at the 3' end of eukaryotic mRNA and the post-transcriptional addition in the nucleus of a sequence of about 100-200 adenine nucleotides (polyA tail) to the cleaved 3' end. The polyadenylation signal sequence may include the consensus sequence AATAAA located at about 10-30 nucleotides upstream from the site of cleavage.

To produce a secreted polypeptide, the nucleic acid of interest includes a DNA segment that encodes a signal sequence/leader peptide. The signal sequence directs the newly synthesized polypeptide to and through the ER membrane where the polypeptide can be routed for secretion. The signal sequence is cleaved off by a signal peptidases during the protein crosses the ER membrane. As for the function of the signal sequence the recognition by the host cell's secretion machinery is essential. Therefore the used signal sequence has to be recognized by the host cell's proteins and enzymes of the secretion machinery.

Translational regulatory elements include a translational initiation (AUG) and stop codon (TAA, TAG or TGA). An internal ribosome entry site (IRES) can be included in some constructs.

A "promoter" refers to a polynucleotide sequence that controls transcription of a gene/structural gene or nucleic acid sequence to which it is operably linked. A promoter includes signals for RNA polymerase binding and transcription initiation. The promoters used will be functional in the cell type of the host cell in which expression of the selected sequence is contemplated. A large number of promoters including constitutive, inducible and repressible promoters from a variety of different sources, are well known in the art (and identified in databases such as

GenBank) and are available as or within cloned polynucleotides (from, e.g., depositories such as ATCC as well as other commercial or individual sources). A "promoter" comprises a nucleotide sequence that directs the transcription of a structural gene. Typically, a promoter is located in the 5' non-coding or untranslated region of a gene, proximal to the transcriptional start site of a structural gene. Sequence elements within promoters that function in the initiation of transcription are often characterized by consensus nucleotide sequences. These promoter elements include RNA polymerase binding sites, TATA sequences, CAAT sequences, differentiation-specific elements (DSEs; McGehee, R.E., et al., Mol. Endocrinol. 7 (1993) 551-560), cyclic AMP response elements (CREs), serum response elements (SREs; Treisman, R., Seminars in Cancer Biol. 1 (1990) 47-58), glucocorticoid response elements (GREs), and binding sites for other transcription factors, such as CRE/ATF (Treisman, R., Seminars in Cancer Biol. 1 (1990) 47-58; O'Reilly, M.A., et al., J. Biol. Chem. 267 (1992) 19938), AP2 (Ye, J., et al., J. Biol. Chem. 269 (1994) 25728), SP1, cAMP response element binding protein (CREB; Loeken, M.R., Gene Expr. 3 (1993) 253) and octamer factors (see, in general, Watson et al., eds., Molecular Biology of the Gene, 4th ed. (The Benjamin/Cummings Publishing Company, Inc. 1987), and Lemaigre, F.P. and Rousseau, G.G., Biochem. J. 303 (1994) 1-14). If a promoter is an inducible promoter, then the rate of transcription increases in response to an inducing agent. In contrast, the rate of transcription is not regulated by an inducing agent if the promoter is a constitutive promoter. Repressible promoters are also known. For example, the c-fos promoter is specifically activated upon binding of growth hormone to its receptor on the cell surface. Tetracycline (tet) regulated expression can be achieved by artificial hybrid promoters that consist e.g. of a CMV promoter followed by two Tet-operator sites. The Tet-repressor binds to the two Tet-operator sites and blocks transcription. Upon addition of the inducer tetracycline, Tet-repressor is released from the Tet-operator sites and transcription proceeds (Gossen, M. and Bujard, H. PNAS 89 (1992) 5547-5551). For other inducible promoters including metallothionein and heat shock promoters, see, e.g., Sambrook et al. (supra) and Gossen, M., Curr. Opin. Biotech. 5 (1994) 516-520. Among the eukaryotic promoters that have been identified as strong promoters for high-level expression are the SV40 early promoter, adenovirus major late promoter, mouse metallothionein-I promoter, Rous sarcoma virus long terminal repeat, Chinese hamster elongation factor 1 alpha (CHEF-1, see e.g. US 5,888,809), human EF-1 alpha, ubiquitin, and human cytomegalovirus immediate early promoter (CMV IE).

The "promoter" can be constitutive or inducible. An enhancer (i.e., a cis-acting DNA element that acts on a promoter to increase transcription) may be necessary to function in conjunction with the promoter to increase the level of expression obtained with a promoter alone, and may be included as a transcriptional regulatory element. Often, the polynucleotide segment containing the promoter will include enhancer sequences as well (e.g., CMV or SV40).

An "enhancer", as used herein, refers to a polynucleotide sequence that enhances transcription of a gene or coding sequence to which it is operably linked. Unlike promoters, enhancers are relatively orientation and position independent and have been found 5' or 3' (Lusky, M., et al., Mol. Cell Bio., 3 (1983) 1108-1122) to the transcription unit, within an intron (Banerji, J., et al., Cell, 33 (1983) 729-740) as well as within the coding sequence itself (Osborne, T.F., et al., Mol. Cell Bio., 4 (1984) 1293-1305). Therefore, enhancers may be placed upstream or downstream from the transcription initiation site or at considerable distances from the promoter, although in practice enhancers may overlap physically and functionally with promoters. A large number of enhancers, from a variety of different sources are well known in the art (and identified in databases such as GenBank) and available as or within cloned polynucleotide sequences (from, e.g., depositories such as the ATCC as well as other commercial or individual sources). A number of polynucleotides comprising promoter sequences (such as the commonly-used CMV promoter) also comprise enhancer sequences. For example, all of the strong promoters listed above may also contain strong enhancers (see e.g. Bendig, M.M., Genetic Engineering, 7 (Academic Press, 1988) 91-127).

An "internal ribosome entry site" or "IRES" describes a sequence which functionally promotes translation initiation independent from the gene 5' of the IRES and allows two cistrons (open reading frames) to be translated from a single transcript in an animal cell. The IRES provides an independent ribosome entry site for translation of the open reading frame immediately downstream (downstream is used interchangeably herein with 3') of it. Unlike bacterial mRNA which can be polycistronic, i.e. encode several different polypeptides that are translated sequentially from the mRNAs, most mRNAs of animal cells are monocistronic and code for the synthesis of only one protein. With a monocistronic transcript in a eukaryotic cell, translation would initiate from the 5' most translation initiation site, terminate at the first stop codon, and the transcript would be released from the ribosome, resulting in the translation of only the first encoded polypeptide in the

mRNA. In a eukaryotic cell, a polycistronic transcript having an IRES operably linked to the second or subsequent open reading frame in the transcript allows the sequential translation of that downstream open reading frame to produce the two or more polypeptides encoded by the same transcript. The use of IRES elements in vector construction has been previously described, see, e.g., Pelletier, J., et al., Nature 334 (1988) 320-325; Jang, S.K., et al., J. Virol. 63 (1989) 1651- 1660; Davies, M.V., et al., J. Virol. 66 (1992) 1924-1932; Adam, M.A., et al. J. Virol. 65 (1991) 4985- 4990; Morgan, R.A., et al. Nucl. Acids Res. 20 (1992) 1293-1299; Sugimoto, Y, et al. Biotechnology 12 (1994) 694-698; Ramesh, N., et al. Nucl. Acids Res. 24 (1996) 2697-2700; and Mosser, D.D. et al, BioTechniques 22 (1997) 150-161).

"Operably linked" refers to a juxtaposition of two or more components, wherein the components so described are in a relationship permitting them to function in their intended manner. For example, a promoter and/or enhancer are operably linked to a coding sequence, if it acts in cis to control or modulate the transcription of the linked sequence. Generally, but not necessarily, the DNA sequences that are "operably linked" are contiguous and, where necessary to join two protein encoding regions such as a secretory leader and a polypeptide, contiguous and in (reading) frame. However, although an operably linked promoter is generally located upstream of the coding sequence, it is not necessarily contiguous with it. Enhancers do not have to be contiguous. An enhancer is operably linked to a coding sequence if the enhancer increases transcription of the coding sequence. Operably linked enhancers can be located upstream, within or downstream of coding sequences and at considerable distance from the promoter. A polyadenylation site is operably linked to a coding sequence if it is located at the downstream end of the coding sequence such that transcription proceeds through the coding sequence into the polyadenylation sequence. A translation stop codon is operably linked to an exonic nucleic acid sequence if it is located at the downstream end (3' end) of the coding sequence such that translation proceeds through the coding sequence to the stop codon and is terminated there. Linking is accomplished by recombinant methods known in the art, e.g., using PCR methodology and/or by ligation at convenient restriction sites. If convenient restriction sites do not exist, then synthetic oligonucleotide adaptors or linkers are used in accord with conventional practice.

"Heterologous DNA" or „heterologous polypeptide" refers to a DNA molecule or a polypeptide, or a population of DNA molecules or a population of polypeptides, that do not exist naturally within a given host cell. DNA molecules heterologous to

a particular host cell may contain DNA derived from the host cell species (i.e. endogenous DNA) so long as that host DNA is combined with non-host DNA (i.e. exogenous DNA). For example, a DNA molecule containing a non-host DNA segment encoding a polypeptide operably linked to a host DNA segment comprising a promoter is considered to be a heterologous DNA molecule. Conversely, a heterologous DNA molecule can comprise an endogenous structural gene operably linked with an exogenous promoter.

A peptide or polypeptide encoded by a non-host DNA molecule is a "heterologous" peptide or polypeptide.

An "expression plasmid" is a nucleic acid molecule encoding a protein to be expressed in a host cell. Typically, an expression plasmid comprises a prokaryotic plasmid propagation unit, e.g. for E.coli, comprising an origin of replication, and a selection marker, an eukaryotic selection marker, and one or more expression cassettes for the expression of the structural gene(s) of interest each comprising a promoter, a structural gene, and a transcription terminator including a polyadenylation signal. Gene expression is usually placed under the control of a promoter, and such a structural gene is said to be "operably linked to" the promoter. Similarly, a regulatory element and a core promoter are operably linked if the regulatory element modulates the activity of the core promoter.

An "isolated polypeptide" is a polypeptide that is essentially free from contaminating cellular components, such as carbohydrate, lipid, or other proteinaceous impurities associated with the polypeptide in nature. Typically, a preparation of isolated polypeptide contains the polypeptide in a highly purified form, i.e. at least about 80% pure, at least about 90% pure, at least about 95% pure, greater than 95% pure, or greater than 99% pure. One way to show that a particular protein preparation contains an isolated polypeptide is by the appearance of a single band following sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis of the protein preparation and Coomassie Brilliant Blue staining of the gel. However, the term "isolated" does not exclude the presence of the same polypeptide in alternative physical forms, such as dimers or alternatively glycosylated or derivatized forms.

„Transcription terminator“ as denoted within this application is a DNA sequence of 50-750 base pairs in length which gives the RNA polymerase the signal for termination of the mRNA synthesis. Very efficient (strong) terminators at the

3' end of an expression cassette are advisable to prevent the RNA polymerase from reading through particularly when using strong promoters. Inefficient transcription terminators can lead to the formation of an operon-like mRNA which can be the reason for an undesired, e.g. plasmid-coded, gene expression.

The following examples, sequence listing and figures are provided to aid the understanding of the present invention, the true scope of which is set forth in the appended claims. The scope of the claims should not be limited to the illustrative embodiments, but should be given the broadest interpretation consistent with the description as a whole.

Description of the Sequences

SEQ ID NO: 01	C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgA class immunoglobulin
SEQ ID NO: 02	C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgD class immunoglobulin
SEQ ID NO: 03	C-terminal part of the heavy chain constant domain (C_H4) amino acid sequence of a human IgE class immunoglobulin
SEQ ID NO: 04	C-terminal part of the heavy chain constant domain (C_H4) amino acid sequence of human IgM class immunoglobulin
SEQ ID NO: 05	C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgG1 class immunoglobulin
SEQ ID NO: 06	C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgG2 class immunoglobulin
SEQ ID NO: 07	C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgG3 class immunoglobulin
SEQ ID NO: 08	C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgG4 class immunoglobulin
SEQ ID NO: 09	Nucleic acid sequence encoding the C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgA class immunoglobulin

SEQ ID NO: 10	Nucleic acid sequence encoding the C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgD class immunoglobulin
SEQ ID NO: 11	Nucleic acid sequence encoding the C-terminal part of the heavy chain constant domain (C_H4) amino acid sequence of a human IgE class immunoglobulin
SEQ ID NO: 12	Nucleic acid sequence encoding the C-terminal part of the heavy chain constant domain (C_H4) amino acid sequence of a human IgM class immunoglobulin
SEQ ID NO: 13, 28	Nucleic acid sequence encoding the C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgG1 class immunoglobulin (variant 1 and 2)
SEQ ID NO: 14, 29	Nucleic acid sequence encoding the C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgG2 class immunoglobulin (variant 1 and 2)
SEQ ID NO: 15	Nucleic acid sequence encoding the C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgG3 class immunoglobulin
SEQ ID NO: 16	Nucleic acid sequence encoding the C-terminal part of the heavy chain constant domain (C_H3) amino acid sequence of a human IgG4 class immunoglobulin
SEQ ID NO: 17	Nucleic acid sequence according to the invention encoding a part of the C-terminal constant domain (C_H3) amino acid sequences of an IgA class immunoglobulin
SEQ ID NO: 18	Nucleic acid sequence according to the invention encoding a part of the C-terminal constant domain (C_H4) amino acid sequences of an IgE class immunoglobulin
SEQ ID NO: 19	Nucleic acid sequence according to the invention encoding a part of the C-terminal constant domain (C_H4) amino acid sequences of an IgM class immunoglobulin

SEQ ID NO: 20, 30 Nucleic acid sequence according to the invention encoding a part of the C-terminal constant domain (C_H3) amino acid sequences of an IgG1 class immunoglobulin (variant 1 and 2)

SEQ ID NO: 21, 31 Nucleic acid sequence according to the invention encoding a part of the C-terminal constant domain (C_H3) amino acid sequences of an IgG2 class immunoglobulin (variant 1 and 2)

SEQ ID NO: 22 Nucleic acid sequence according to the invention encoding a part of the C-terminal constant domain (C_H3) amino acid sequences of an IgG3 class immunoglobulin

SEQ ID NO: 23 Nucleic acid sequence according to the invention encoding a part of the C-terminal constant domain (C_H3) amino acid sequences of an IgG4 class immunoglobulin

SEQ ID NO: 24 to 27 Nucleic acid primers used in the Examples.

Description of the Figures

Figure 1 Annotated plasmid map of p4831.

Figure 2 SDS-PAGE and Western blot analysis of antibodies secreted by clone #23; a) Coomassie staining, b) Western blot analysis with peroxidase coupled anti-human immunoglobulin gammy chain-antibody, c) Western blot analysis with peroxidase coupled anti-human immunoglobulin kappa light chain-antibody. Lane 1: human anti-IGF-1R-reference antibody; lane 2: culture supernatant of CHO-DG44 clone #23 comprising the antibody of clone #23.

Figure 3 Annotated plasmid map of p4817.

Figure 4 Western blot analyses of immunoglobulins secreted by CHO-DXB11 transfected with p4831 or 4855. Lane 1: CHO cells transfected with p4831, lane 2: CHO cells transfected with p4855.

Figure 5 Annotated plasmid map of p5031.

Figure 6 SDS-PAGE analysis of antibodies secreted by HEK-293-EBNA cells after transfection with plasmid p5031 or p5032; Lanes 1 + 2: p5031, lanes 3 + 4: p5032, lanes 5 - 8: human anti-IGF-1R-antibody as reference antibody; 5: 0.2 µg, 6: 0.7 µg, 7: 2 µg, 8: 6 µg.

Figure 7 CHO-DG44 cells transfected with p5032 and selected for stable integration of the plasmid with Methotrexate (MTX). Antibodies from six clones were purified and analyzed by SDS-PAGE and Coomassie staining.

Materials and Methods

Recombinant DNA techniques

Standard methods were used to manipulate DNA as described in Sambrook et al., 1989 (supra). All molecular biological reagents were commercially available (if not indicated otherwise) and were used according to the manufacturer's instructions.

Nucleic acid sequence determination

DNA sequences were determined by double strand sequencing performed at MediGenomix GmbH (Martinsried, Germany).

DNA and protein sequence analysis and sequence data management

The GCG's (Genetics Computer Group, Madison, Wisconsin, USA) software package version 10.2 and Infomax's Vector NTI Advance suite version 8.0 was used for sequence creation, mapping, analysis, annotation, and illustration.

Cell culture techniques

CHO-DXB11 cells were grown in MEM alpha medium (Invitrogen Corp., Gibco®, Cat. No.: 22571) with 10 % FCS (fetal calf serum obtained from Hyclone, Thermo Fisher Scientific Inc., Cat. No.: SH3007.03).

HEK-293-EBNA cells (ATCC # CRL-10852) were cultivated in DMEM, supplemented with 2 mM glutamine (Gibco®, Cat. No.: 25030), 1 % (v/v) MEM non essential amino acids (Gibco®, Cat. No.: 11140), 10 % (v/v) ultra-low IgG FCS (Gibco®, Cat. No.: 16250), and 250 µg/ml G418 (Roche Applied Sciences, Roche

Diagnostics GmbH, Germany, Cat. No.: 1464981).

The medium for the cultivation of CHO-DG44 cells was MEM alpha medium (Gibco®, Cat. No.: 22561) supplemented with 10 % (v/v) dialyzed FCS (Gibco®, Cat. No.: 26400) and 2 % (v/v) HT supplement (Gibco®, Cat. No.: 41065). For the selection of stably transfected CHO DG44 cell lines the HT supplement was omitted and 20 to 500 nM Methotrexate (MTX) was added either alone or in combination with 400 µg/ml Hygromycin B (Roche Diagnostics GmbH, Roche Applied Sciences, Germany, Cat. No.: 10843555001).

All cell lines were maintained in humidified incubators at 37 °C with 5 % CO_2. Transfection of cells was either performed by nucleofection (Amaxa GmbH, Germany) or by lipofection using FuGENE 6 (Roche Diagnostics GmbH, Roche Applied Sciences, Germany, Cat. No.: 1815075).

Furthermore standard cell culture techniques were applied as described e.g. in Bonifacino, J. S., et al. (Eds) (2000) Current Protocols in Cell Biology, John Wiley and Sons, Inc.

Protein A-precipitation, SDS-PAGE and Western Blot

Immunoglobulins from cell culture supernatants were precipitated with Protein A-Sepharose beads and then analyzed by SDS/polyacrylamide gel electrophoresis (sodium dodecyl sulfate, SDS-PAGE) and Western-blotting.

For precipitation of immunoglobulins cell culture supernatants containing up to 7 µg immunoglobulin were diluted with TBS buffer (50 mM TRIS/HCl, pH 7.5, supplemented with 150 mM NaCl), 1 % (v/v) Nonidet-P40 (Roche Diagnostics GmbH, Roche Applied Sciences, Cat. No.: 1754599) to a final volume of 1 ml, and afterwards incubated for one hour with 15 µl wet volume Protein A-Sepharose beads. The beads were recovered by centrifugation and washed with TBS with 1 % (v/v) Nonidet P-40, thereafter with 2-fold concentrated PBS (phosphate buffered saline), and finally with 100 mM sodium citrate buffer, pH 5. After the final wash step the supernatant was removed completely from the beads. Bound immunoglobulins were eluted with 20 µl 2-fold concentrated LDS (lithium dodecyl sulfate) sample buffer (Invitrogen Corp.) containing 50 mM DTT (dithiothreitol). After 5 minutes incubation at 95 °C the suspension was centrifuged and the supernatant was recovered for further analysis.

SDS-PAGE: SDS-PAGE was performed using the NuPAGE® gel system (Invitrogen Corp.) according to the manufacturer's recommendations. Samples were loaded onto 10 % NuPAGE® Novex Bis/TRIS gels (Invitrogen Corp., Cat. No.: NP0301) and proteins were separated in reducing NuPAGE® MES SDS (4-morpholinoethanesulfonic acid/sodium dodecyl sulfate) running buffer. Typically 2 to 3 µg immunoglobulin per lane was loaded for Coomassie staining with Simply Blue Safe Stain® (Invitrogen Corp.) and 0.4 to 0.6 µg for Western-blotting.

Western Blot: For electro-transfer of proteins from SDS/polyacrylamide gels standard PVDF (polyvinylidene difluoride) or nitrocellulose membranes were used. After electro-transfer membranes were washed in TBS (tries buffered saline, 50 mM TRIS/HCl, pH 7.5, 150 mM NaCl). Nonspecific binding sites were blocked by incubation in TBS with 1 % (w/v) Western Blocking Reagent (Roche Diagnostics GmbH, Roche Applied Sciences, Cat. No.: 11921673001). Human immunoglobulin gamma heavy chains and kappa light chains were detected with peroxidase coupled polyclonal detection antibodies (see following paragraph for further details) diluted in TBS with 0.5 % (w/v) Western Blocking Reagent. After three wash steps with TBS with 0.05 % (v/v) Tween® 20, and one wash step with TBS, bound peroxidase coupled detection antibodies were detected by chemoluminescence using LumiLightPlus substrate solution (Roche Diagnostics GmbH, Roche Applied Sciences, Cat. No.: 12015196001) and LUMI-Imager F1 analyzer (Roche Diagnostics GmbH, Roche Applied Sciences).

Human immunoglobulin gamma heavy chains (H) and kappa light chains (L) were detected either simultaneously or separately. For simultaneous detection peroxidase coupled goat anti-human IgG (H+L)-antibody (Jackson ImmunoResearch Laboratories Inc., Cat. No.: 109-035-088) was used at a dilution of 1:2500 (v/v). For consecutive detection, membranes were first probed with peroxidase coupled rabbit anti-human Ig gamma-antibody (DAKO GmbH, Germany, code no. P0214) at a dilution of 1:1000 (v/v) or with peroxidase coupled $F(ab')_2$ goat anti-human Fc gamma-antibody (Jackson ImmunoResearch Laboratories, Cat. No.: 109-036-008) at a dilution of 1:7500 (v/v). After detection of the immunoglobulin gamma heavy chain, the membranes were stripped for 30 minutes in 62.5 mM TRIS/HCl, pH 6.7, supplemented with 2 % (w/v) SDS and 100 mM β-mercaptoethanol, at 50 °C. For the second detection the membranes were re-probed with peroxidase coupled rabbit anti-human Ig kappa-antibody (DAKO GmbH, code no. P0129) at a dilution of 1:1000 (v/v).

Example 1

Preparation of an expression plasmid for an immunoglobulin of class IgG1

Plasmid 4831 (denoted as p4831 in the following) is the expression plasmid for the expression of an anti-IGF-1R-antibody (genomically organized expression cassette with retained exon-intron organization) in eukaryotic cells (for sequences see e.g. US 2005/0008642, or EP 1 646 720). It comprises the following functional elements:

- an origin of replication derived from the vector pUC18 (pUC origin),

- a ß(beta)-lactamase gene conferring ampicillin resistance in E. coli (Amp),

- an expression cassette for the expression of a gamma 1-heavy chain comprising the following elements

  - the major immediate-early promoter and enhancer from the human cytomegalovirus (hCMV IE1 promoter),
  - a synthetic 5'UTR including a Kozak sequence,
  - a murine immunoglobulin heavy chain signal sequence including the signal sequence intron (L1_Intron_L2),
  - the cDNA for a heavy chain variable region (VH) arranged with a splice donor site at the 3' end,
  - the mouse immunoglobulin µ-enhancer region,
  - a human immunoglobulin heavy chain gamma 1-gene (IGHG1) including exons CH1, Hinge, CH2 and CH3, intervening introns and the 3'UTR bearing the polyadenylation signal sequence and optionally containing mutations,

- an expression cassette for the expression of a kappa-light chain comprising the following elements

  - the major immediate-early promoter and enhancer from the human cytomegalovirus (hCMV IE1 promoter),
  - a synthetic 5'UTR including a Kozak sequence,
  - a murine immunoglobulin heavy chain signal sequence including the signal sequence intron (L1_Intron_L2),
  - the cDNA for a light chain variable region arranged with a splice donor site at the 3' end (VL),
  - the intronic mouse Ig-kappa enhancer region,

- the human immunoglobulin kappa gene (IGK) including the IGKC exon and the IGK 3'UTR bearing the polyadenylation signal sequence.

- a Hygromycin B phosphotransferase transcription unit suitable for selection in eukaryotic cells including

- the SV40 early promoter and origin,
- the hygromycin B phosphotransferase coding sequence (HygB),
- the SV40 early polyadenylation signal

- an expression cassette for the expression of murine dihydrofolate reductase (DHFR) suitable for auxotrophic selection in eukaryotic cells including

- a shortened version of the SV40 early promoter and origin,
- the coding sequence for murine DHFR,
- the SV40 early polyadenylation signal.

An annotated plasmid map of p4831 is shown in Figure 1.

P4831 was transfected into CHO-DG44 cells and stable cell lines were isolated after selection with Hygromycin B and Methotrexate (MTX). Antibodies secreted by selected clone #23 were precipitated with Protein A-Sepharose beads and analyzed by SDS-PAGE and Coomassie staining (Figure 2 a)). In addition to the expected 50 kDa immunoglobulin gamma-1 heavy chain and the 25 kDa immunoglobulin kappa light chain, considerable amounts of an 80 kDa by-product protein were detected. This protein was recognized by anti-human immunoglobulin gamma chain antibodies (Figure 2 b)) as well as by anti-human immunoglobulin kappa chain antibodies (Figure 2 c)).

<u>Example 2</u>

Preparation of an expression plasmid for an immunoglobulin of class IgG1 with a modified CH3-domain

In order to prevent the generation of by-products resulting from aberrantly spliced gamma 1 pre-mRNA, the internal splice site of the CH3 exon of p4831 was destroyed by mutating the T in position 4573 to C. At the same time T4567 was replaced by C for removal of a BsmA I restriction site.

P4855 was constructed as follows. p4817, an ancestor plasmid of p4831 with the same gamma 1 heavy chain transcription unit, is composed of the following elements:

- an origin of replication from the vector pUC18 which allows replication of this plasmid in E. coli (pUC Ori)
- a beta-lactamase gene which confers ampicillin resistance in E. coli (Amp)
- an expression cassette for the expression of a gamma1-heavy chain comprising the following elements

  - the major immediate-early promoter and enhancer from the human cytomegalovirus (hCMV IE1 promoter),
  - a synthetic 5'UTR including a Kozak sequence,
  - a murine immunoglobulin heavy chain signal sequence including the signal sequence intron (L1_Intron_L2),
  - the cDNA for a heavy chain variable region (VH) arranged with a splice donor site at the 3' end,
  - the mouse immunoglobulin μ-enhancer region,
  - a human immunoglobulin heavy chain gamma 1-gene (IGHG1) including exons CH1, Hinge, CH2 and CH3, intervening introns and the 3'UTR bearing the polyadenylation signal sequence and optionally containing mutations,

The plasmid map of p4817 is shown in Figure 3.

P4817 was manipulated by site-directed mutagenesis using the QuickChange Site-Directed Mutagenesis Kit (Stratagene, Cat. No.: 200518) and the sequence specific oligonucleotides 1

agcctctccc tgtccccggg caaatgagtg cgacggccg SEQ ID NO: 24

and 2

cggccgtcgc actcatttgc ccggggacag ggagaggct SEQ ID NO: 25.

The 839 bp SfiI/SgrAI fragment of mutated p4817 was excised and ligated with the 13133 bp SgrAI/SfiI fragment of p4831 to form p4855.

Example 3

Expression of nucleic acids according to Example 1 and 2, Isolation of the produced immunoglobulin, and Analysis of the produced immunoglobulin

Plasmids p4831 and p4855 were transiently transfected into CHO-DXB11 cells. The cells were cultivated under non-selective conditions. After three days of cultivation the cell culture supernatants were harvested and the secreted immunoglobulins were purified with Protein A-Sepharose* beads. Western blot analysis of the immunoglobulins with anti-human IgG-antibodies (H+L) showed that the 80 kDa by-product had been expressed by the cells transfected with p4831 but not by the cells transfected with p4855 (Figure 4).

Example 4

Preparation of an expression plasmid for an immunoglobulin of class IgG4

Plasmid 5031 was designed for transient and stable expression of a human anti-P-selectin-antibody in eukaryotic tissue culture cell. For exemplary anti-P-selectin-antibodies see e.g. EP 1 737 891 or US 2005/0226876. P5031 is composed of the following elements

- an origin of replication from the vector pUC18 which allows replication of this plasmid in E. coli (pUC origin),

- a ß-lactamase gene which confers ampicillin resistance in E. coli (Amp),

- an expression cassette for the expression of a human gamma 4-heavy chain comprising the following elements:

  - the major immediate-early promoter and enhancer from the human cytomegalovirus (hCMV IE1 promoter),
  - a synthetic 5'UTR including a Kozak sequence,
  - a murine immunoglobulin heavy chain signal sequence including the signal sequence intron (L1_Intron_L2),
  - the cDNA for a heavy chain variable region (VH) arranged with a splice donor site at the 3' end,
  - the mouse Ig µ-enhancer region,
  - a human immunoglobulin heavy chain gamma 4-gene (IGHG4) including exons CH1, Hinge, CH2 and CH3, intervening introns and the

*Trademark

3'UTR bearing the polyadenylation signal sequence and optionally containing mutations,

- an expression cassette for the expression of the human kappa-light chain comprising the following elements

  - the major immediate-early promoter and enhancer from the human cytomegalovirus (hCMV IE1 promoter),
  - a synthetic 5'UTR including a Kozak sequence,
  - a murine immunoglobulin heavy chain signal sequence including the signal sequence intron (L1_Intron_L2),
  - the cDNA for a light chain variable region arranged with a splice donor site at the 3' end (VL),
  - the intronic mouse Ig-kappa enhancer region,
  - a human immunoglobulin kappa gene (IGK) including the IGKC exon and the IGK 3'UTR bearing the polyadenylation signal sequence,

- a Hygromycin B phosphotransferase transcription unit suitable for selection in eukaryotic cells including

  - the SV40 early promoter and origin,
  - the hygromycin B phosphotransferase coding sequence (HygB),
  - the SV40 early polyadenylation signal,

- an expression cassette for the expression of murine dihydrofolate reductase (DHFR) suitable for auxotrophic selection in eukaryotic cells including

  - a shortened version of the SV40 early promoter and origin,
  - the coding sequence for murine DHFR,
  - the SV40 early polyadenylation signal.

A plasmid map of p5031 is shown in Figure 5.

When HEK-293-EBNA cells were transfected with plasmid p5031, the cells produced immunoglobulins comprising an 80 kDa by-product (Figure 6, lane 1 and 2). This protein was bound by anti-human Ig gamma-antibody as well as by anti-human Ig kappa-antibody in Western-Blot analysis.

Example 5

Preparation of an expression plasmid for an immunoglobulin of class IgG4 with a modified CH3-domain

The modification was introduced according to Example 3. In brief, T4565 was mutated to C together with a second nucleotide, T4559, which was also exchanged to C for removal of a BsmAI restriction site. Oligonucleotide 3

gcctctccct gtccctgggc aaatgagtgc cagg SEQ ID NO: 26

and oligonucleotide 4

cctggcactc atttgcccag ggacagggag aggc SEQ ID NO: 27

were used for the site-directed mutagenesis. The obtained plasmid was named p5032.

Example 6

Expression of nucleic acids according to Example 5, Isolation of the produced immunoglobulin, and Analysis of the produced immunoglobulin

Plasmids p5031 and p5032 were transiently transfected into HEK-293-EBNA cells. The cells were cultivated under non-selective conditions. After three days of cultivation the cell culture supernatants were harvested and the secreted immunoglobulins were purified with Protein A-Sepharose beads. Western blot analysis of the immunoglobulins with anti-human IgG-antibodies (H+L) showed that cells transfected with p5031 expressed a 80 kDa by-product(Figure 6, lanes 1+2), whereas no by-product had been expressed by the cells transfected with p5032 (Figure 6, lanes 3+4).

Plasmid p5032 was transfected into HEK-293-EBNA cells, the 80 kDa protein was not expressed (Figure 6, lane 3 and 4). This clearly demonstrates that the 80 kDa fusion protein is a result of aberrant pre-mRNA splicing and that the production of such unwanted protein during transient expression can efficiently suppressed by the mutation of the internal CH3 splice site of the immunoglobulin heavy chain gamma 4 gene (IGHG4).

Mutation of the internal C_H3 splice site of IGHG4 prevented the expression of the 80 kDa also in stable cell lines. CHO-DG44 cells were transfected with p5032 and selected for stable integration of the plasmid with Methotrexate (MTX). Antibodies from 6 clones were purified and analyzed by SDS-PAGE and Coomassie staining (Figure 7). None of the antibodies contained the 80 kDa subunit.

Claims

1. A nucleic acid encoding the amino acid sequence of the C-terminal part of the C_H3-domain of an immunoglobulin of the class IgG, characterized in that the glycine-lysine-dipeptide comprised in said amino acid sequence of the C-terminal part of the C_H3- or C_H4-domain is encoded by the nucleic nucleic acid ggcaaa.

2. The nucleic acid according to claim 1, characterized in that said nucleic acid encodes an amino acid sequence selected from the group of amino acid sequences of SEQ ID NO: 5, 6, 7 and 8.

3. The nucleic acid according to claim 1 or 2, characterized in that the nucleic acid encoding said glycine-lysine-dipeptide is preceded by the nucleotide g or a.

4. The nucleic acid according to claim 1, characterized in that said C-terminal part comprises at least the 20 C-terminal amino acids of the immunoglobulin heavy chain primary amino acid sequence.

5. The nucleic acid according to claim 1, characterized in that said nucleic acid encodes a part of the C-terminal constant domain of an immunoglobulin heavy chain of the class IgG and is selected from the group of nucleic acids of SEQ ID NO: 20, 21, 22, 23, 30 and 31.

6. The nucleic acid according to any one of claims 1 to 5, characterized in that said nucleic acid encodes a part of the C-terminal constant domain of a human immunoglobulin heavy chain of the class IgG1 or IgG4.

7. A plasmid comprising the nucleic acid according to claim 1.

8. A cell transfected with the nucleic acid according to claim 1.

9. The cell according to claim 8, characterized in that said cell is a mammalian cell.

10. The cell according to claim 9, characterized in that said mammalian cell is selected from a CHO cell, a HEK cell, and a BHK cell.

11. A nucleic acid comprising the nucleotide sequence of SEQ ID NO: 20, 21, 22, 23, 30, or 31.

12. A method for the production of an immunoglobulin in a mammalian cell comprising the following steps:

a) transfecting said mammalian cell with a nucleic acid encoding an immunoglobulin heavy chain and an expression cassette encoding an immunoglobulin light chain, wherein a nucleic acid of SEQ ID NO: 20,21, 22, 23, 30, or 31 encodes the C-terminal part of the immunoglobulin heavy chain,

b) cultivating the transfected mammalian cell under conditions suitable for the expression of the immunoglobulin,

c) recovering the immunoglobulin from the culture or the cell.

13. The method according to claim 12, wherein said transfecting said mammalian cell is with a nucleic acid encoding an immunoglobulin heavy chain, wherein a nucleic acid of SEQ ID NO: 30 encodes the C-terminal part of the immunoglobulin heavy chain.

Figure 1

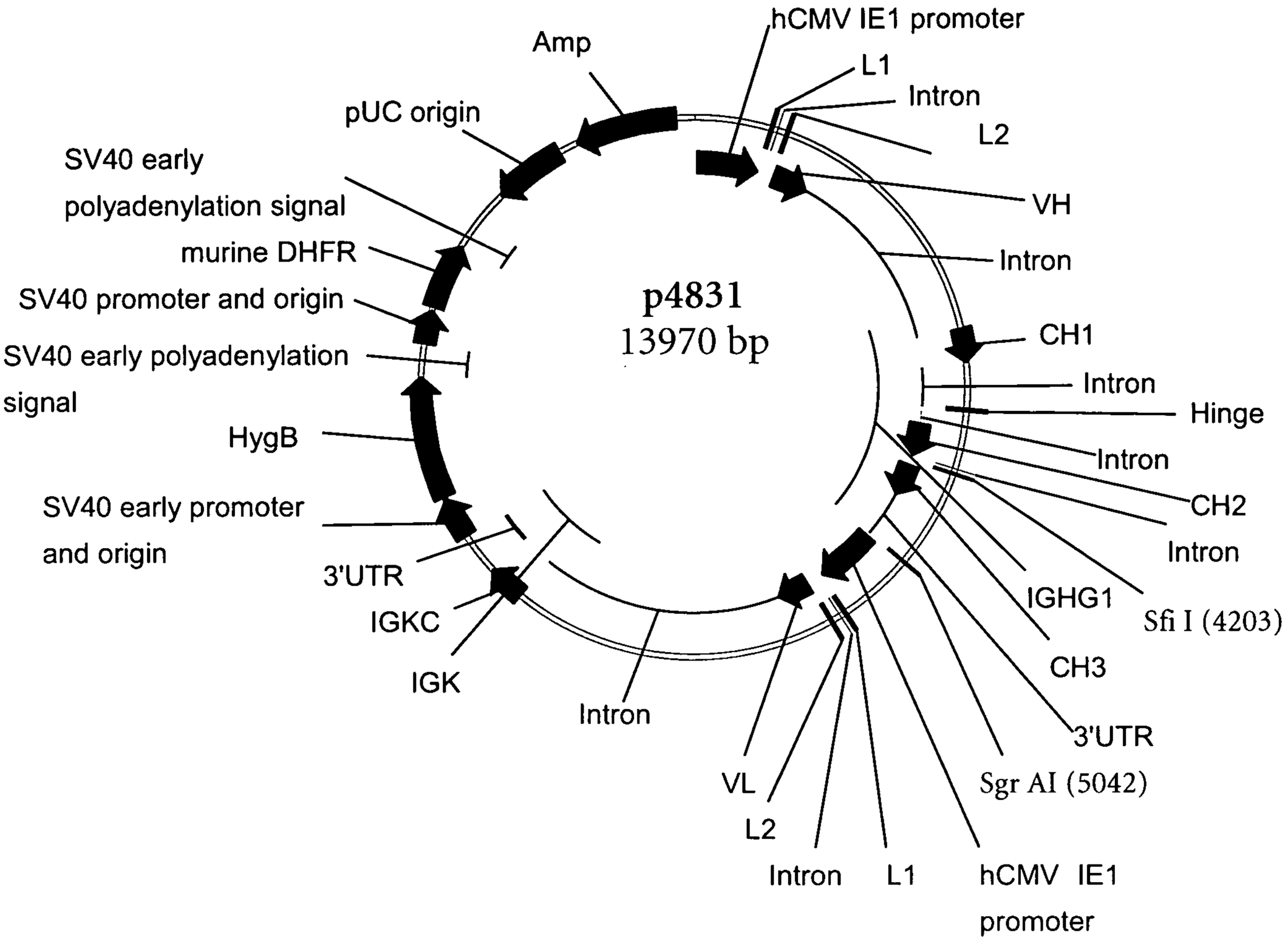

p4831
13970 bp
Amp
pUC origin
SV40 early
polyadenylation signal
murine DHFR
SV40 promoter and origin
SV40 early polyadenylation
signal
HygB
SV40 early promoter
and origin
3'UTR
IGKC
IGK
Intron
VL
L2
Intron
L1
hCMV IE1
promoter
Sgr AI (5042)
3'UTR
CH3
IGHG1
Sfi I (4203)
Intron
CH2
Intron
Hinge
Intron
CH1
Intron
VH
hCMV IE1 promoter
L1
Intron
L2

Figure 2

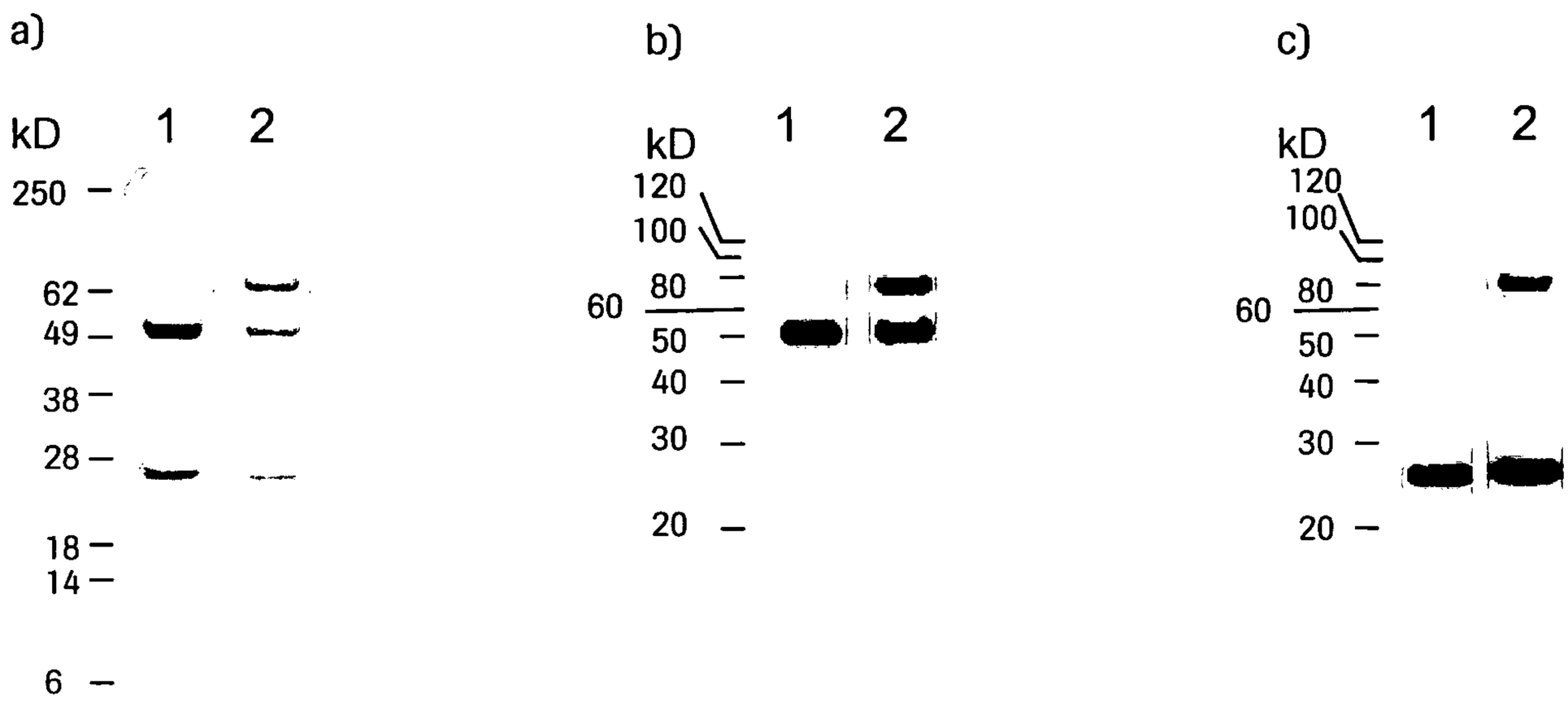

a)
kD
1
2
250
62
49
38
28
18
14
6
b)
kD
1
2
120
100
80
60
50
40
30
20
c)
kD
1
2
120
100
80
60
50
40
30
20

Figure 3

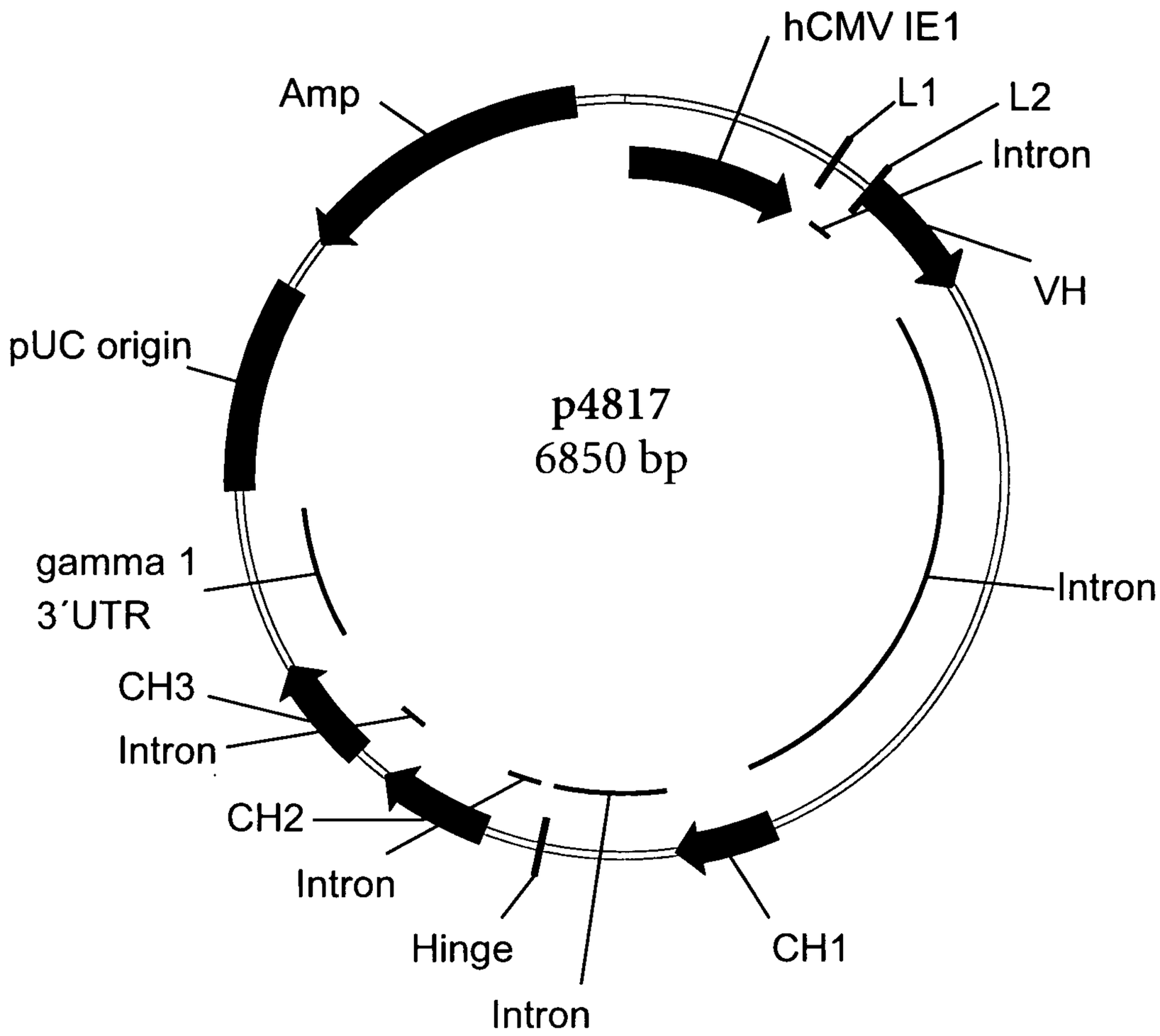

hCMV IE1
Amp
L1
L2
Intron
VH
pUC origin
p4817
6850 bp
gamma 1
3´UTR
Intron
CH3
Intron
CH2
Intron
Hinge
Intron
CH1

Figure 4

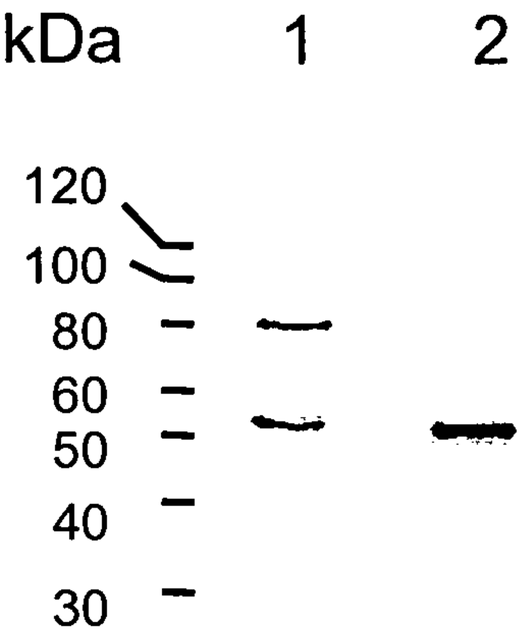


Figure 5

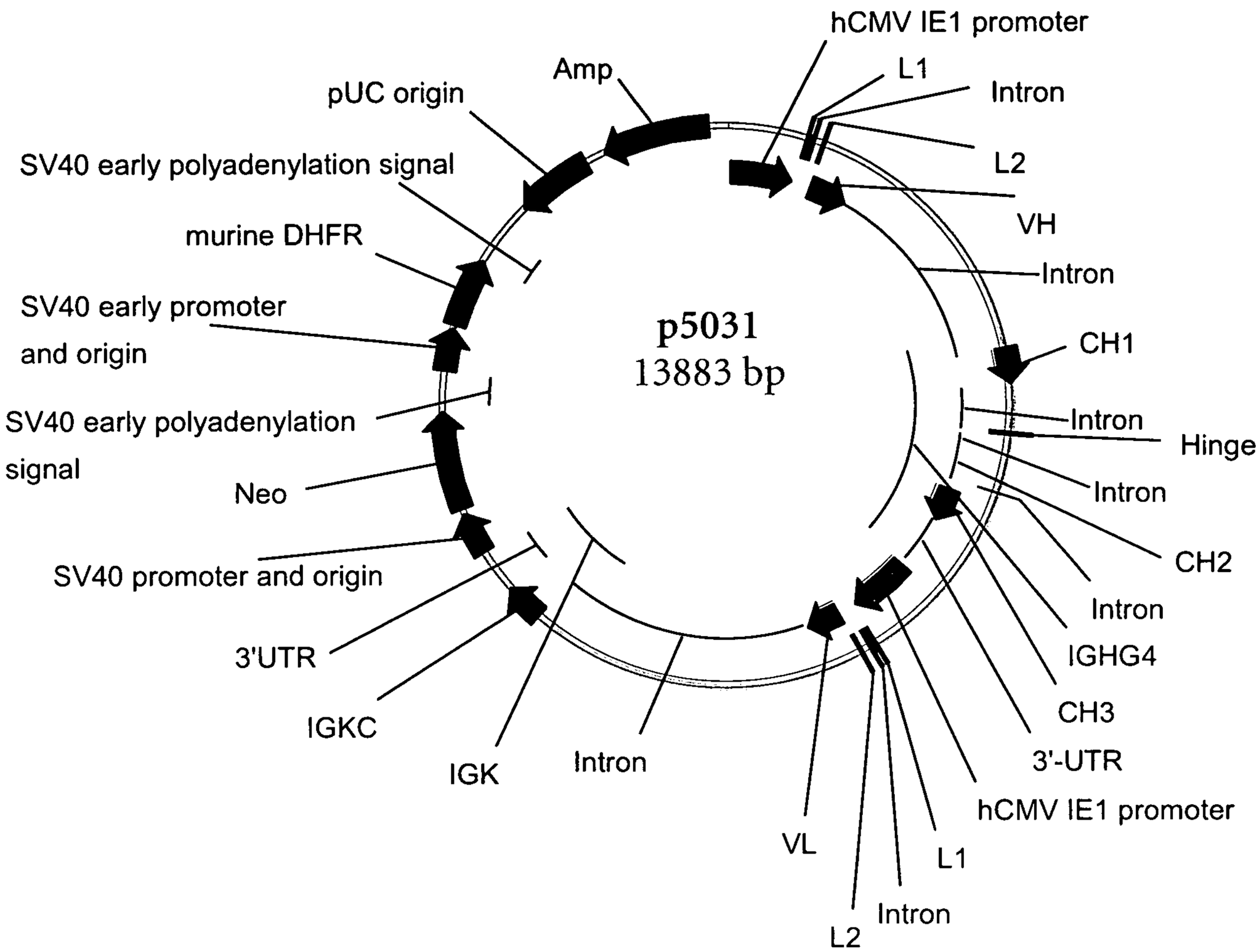

hCMV IE1 promoter
L1
Intron
L2
VH
Intron
CH1
Intron
Hinge
Intron
CH2
Intron
IGHG4
CH3
3'-UTR
hCMV IE1 promoter
L1
Intron
L2
VL
Intron
IGK
IGKC
3'UTR
SV40 promoter and origin
Neo
SV40 early polyadenylation signal
SV40 early promoter and origin
murine DHFR
SV40 early polyadenylation signal
pUC origin
Amp
p5031
13883 bp

Figure 6

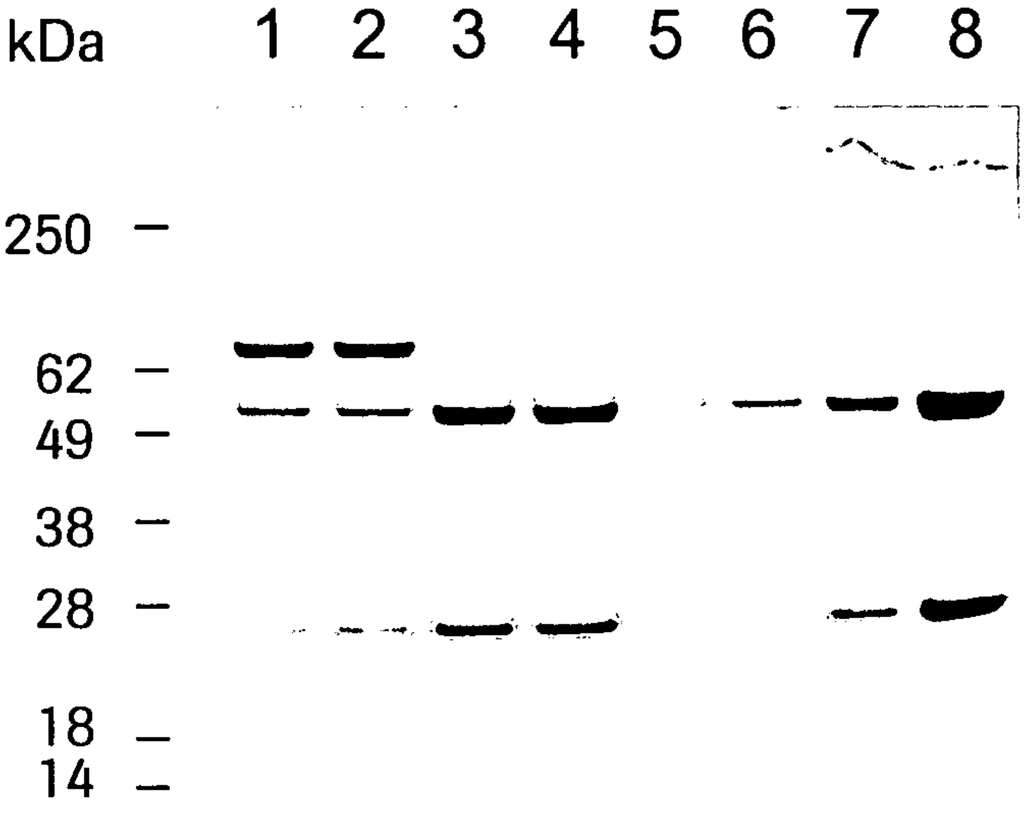
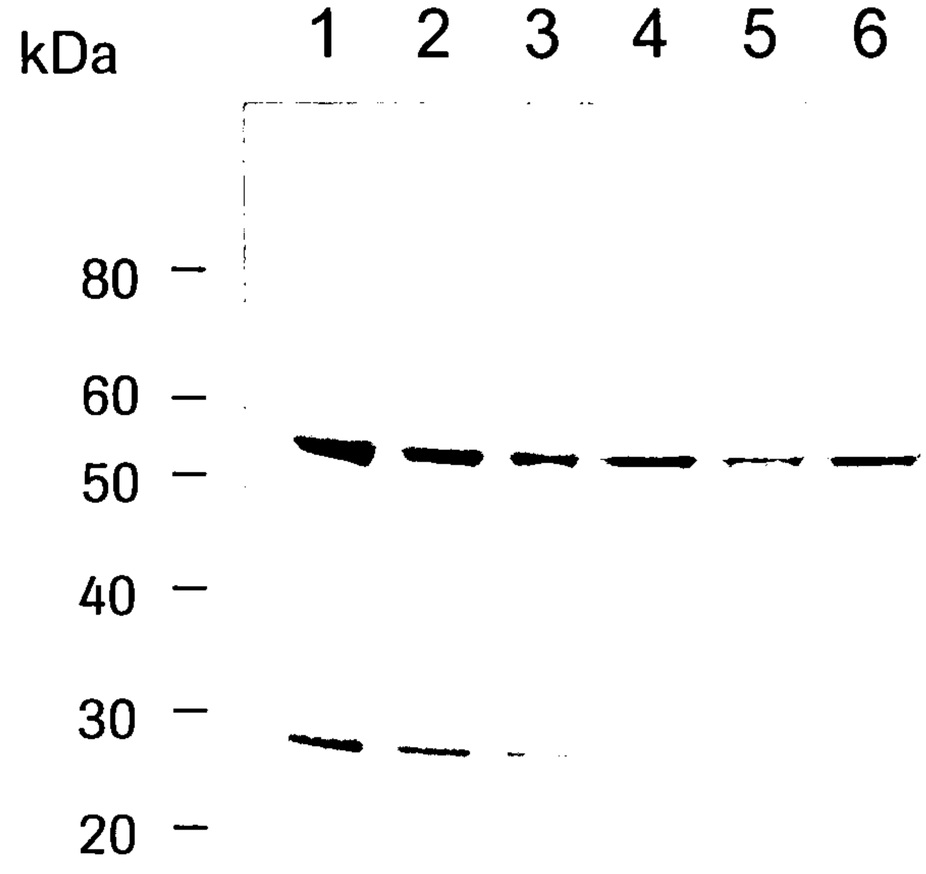
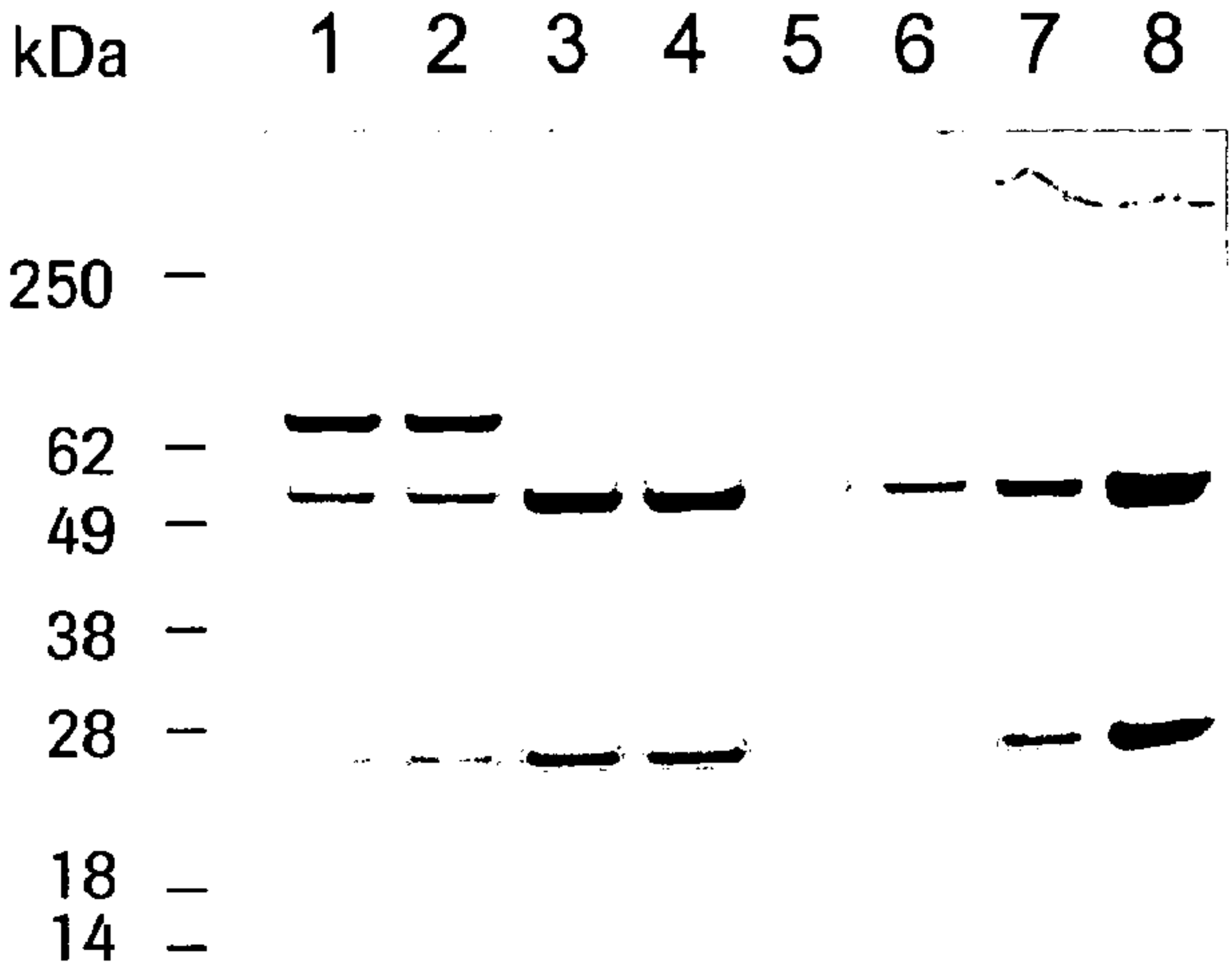


Figure 7




kDa
1 2 3 4 5 6 7 8
250
62
49
38
28
18
14