

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 donor splice site in the exon encoding the C-terminal domain of the secreted form (i.e. the C_H3 or C_H4 domain, respectively) and an acceptor splice 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. Magistrelli, G., et al. reported optimizing assembly and production of native bispecific antibodies by codon de-optimization (MABS 9 (2016) 231-239).

WO 2015/128509 reported expression constructs and methods for selecting host cells expressing polypeptides.

WO 2009/003623 reported a heavy chain mutant leading to improved immunoglobulin production.

Summary of the Invention

For the production of a therapeutic or diagnostic antibody high expression yields are the aim. One general option to achieve good expression rates is to optimize the codon usage of the encoding nucleic acids at first by adjusting it to the codon usage of the cell intended to express the exogenous nucleic acid. Such a codon adaptation or optimization can be done based on different established protocols.

But during such a codon adaptation and optimization process, e.g. non-paired splices sites, especially non-paired donor splice sites, can be generated unintentionally de novo. That is, e.g., during codon optimization a new donor splice site sequence is generated in the codon usage optimized nucleic acid by inadvertently generating a sequence motif inside the codon-optimized nucleic acid that follows a donor splice site consensus sequence. Such an event can take place independently on the organization of the codon usage optimize nucleic acid, i.e. it is possible for both cDNA as well as genomically organized nucleic acids. It is in fact an unintended side-result of the codon usage optimization process. Such a new donor splice site is

an additional artificial donor splice site and it does not have an associated target acceptor splice site. Thus, such non-paired donor splice sites can give rise to splicing events with a random, i.e. non-defined, acceptor splice site present somewhere in the transcribed mRNA. Thereby the expression yield is reduced due to the formation of by-products.

The ~~invention~~disclosure is based, at least in part, on the unexpected finding that the removal of non-paired donor splice sites in a codon usage optimized antibody heavy chain encoding nucleic acid needs only to be performed in the part of the nucleic acid encoding the variable domain of the heavy chain but not in the part encoding the constant region, i.e. for the constant region, e.g., the germline or wild-type human nucleic acid sequence can be used. Thereby the expression yield of the antibody heavy chain with correct length can be increased or becomes possible at all.

The current ~~invention~~disclosure is based, at least in part, on the finding that the introduction of amino acid sequence silent nucleotide changes (mutations) in the non-paired donor splice site consensus sequence NGGTA(G)AG (SEQ ID NO: 01) only in the codon optimized nucleic acid encoding the variable domain of an antibody heavy chain is sufficient to improve the expression yield.

~~One aspect of the current invention is~~The present disclosure provides a method for producing an antibody by cultivating a mammalian cell comprising/transfected with one or more (exogenous) nucleic acids encoding the antibody heavy chain and the antibody light chain (and expressing the antibody),

wherein the one or more (exogenous) nucleic acids are codon usage optimized for the codon usage of human cells and/or for the codon usage of the mammalian cell,

wherein in the nucleic acid encoding the heavy chain variable domain at least one (artificial) non-paired donor splice site is removed and, optionally, in the (human wild-type or human or hamster codon usage optimized) nucleic acid sequence encoding the heavy chain constant region (artificial) non-paired donor splice sites are not removed.

In one aspect, the present disclosure provides a method for producing an antibody of the human IgG1 subclass by cultivating a CHO cell that has been transfected with one or more expression cassettes comprising the nucleic acids encoding the heavy and the light chains of the antibody,

5 wherein the nucleic acids encoding the antibody heavy and light chains are codon usage optimized with the codon usage of human cells and/or the codon usage of CHO cells,

wherein in the part of the nucleic acid encoding the heavy chain variable domain at least one non-paired donor splice site is removed.

10 In another aspect, the present disclosure provides a method for producing an antibody of the human IgG1 subclass by cultivating a CHO cell comprising one or more nucleic acids encoding the heavy and the light chains of the antibody,

15 wherein in the nucleic acid encoding the heavy chain variable domain non-paired donor splice sites are removed by introducing amino acid sequence silent nucleotide changes in the non-paired donor splice site consensus sequence NGGTA(G)AG (SEQ ID NO: 01) in the codon NGG or the codon GGT or the codon GTA(G),

wherein non-paired donor splice sites according to the sequence of SEQ ID NO: 01 are not removed in the nucleic acid encoding the heavy chain constant region.

20 In one ~~embodiment~~example the antibody is an antibody of the human IgG1 subclass. In one ~~embodiment~~example the antibody is a humanized antibody of the human IgG1 subclass. In one ~~embodiment~~example the constant region of the antibody comprises mutations suitable to induce heterodimerization or modify Fc-receptor binding.

In one embodimentexample the mammalian cell is a CHO cell.

In one embodimentexample the transfection is a transient transfection.

In one embodimentexample the one or more (exogenous) nucleic acids encoding the antibody are all cDNA.

5 In one embodimentexample the one or more (exogenous) nucleic acids encoding the antibody heavy chain and/or the antibody light chain are genomically organized DNA, i.e. have an intron-exon-organization.

10 In one embodimentexample the removal of the non-paired donor splice sites is by introducing an amino acid silent change (mutation) in the amino acid sequence NGGTA(G)AG (SEQ ID NO: 01). In one embodimentexample the amino acid sequence silent nucleotide change is introduced in the codon NGG or the codon GGT or the codon GTA(G) of SEQ ID NO: 01.

In one embodimentexample the codon usage optimization is done based on the human codon usage or based on the chinese hamster codon usage.

15 In one embodimentexample the nucleic acid encoding the antibody light chain is codon usage optimized, i.e. the variable domain and the constant region is codon usage optimized.

In one embodimentexample in the complete light chain encoding nucleic acid non-paired donor splice sites are removed.

20 In one embodimentexample the transfection is a stable transfection.

In one embodimentexample the method comprises the following steps:

a) cultivating the mammalian cell, and

b) recovering the antibody from the cell or the cultivation medium.

~~One aspect of the current invention is~~The present disclosure provides the use of the removal of non-paired donor splice sites only in a part of a human or hamster codon-usage optimized nucleic acid sequence encoding an antibody for reducing mis-splicing or/and increasing antibody expression yield when said nucleic acid is used to produce the antibody in CHO cells, whereby the part is the part that encodes the heavy chain variable domain. In another aspect, the present disclosure provides the use of the removal of non-paired donor splice sites in a part of a human or hamster codon-usage optimized nucleic acid sequence encoding an antibody, whereby the part is the part that encodes the heavy chain variable domain, for reducing mis-splicing or/and increasing antibody expression yield when said nucleic acid is used to produce the antibody in CHO cells.

In one ~~embodiment~~example non-paired donor splice sites are not removed in the part of the nucleic acid encoding the heavy chain constant region.

In one ~~embodiment~~example further in the part of the nucleic acid encoding the light chain non-paired donor splice sites are removed.

5 In one ~~embodiment~~example the removal of the non-paired splice sites is by introducing an amino acid silent mutation in the nucleotide sequence NGGTA(G)AG (SEQ ID NO: 01).

10 In one ~~embodiment~~example the removal of the non-paired splice sites is by introducing an amino acid silent mutation in the nucleotide sequence NGGTA(G)AG (SEQ ID NO: 01) in the codon NGG or the codon GGT or the codon GTA(G).

In one ~~embodiment~~example of ~~all aspects and embodiments~~the present disclosure the non-paired (donor) splice site is an artificial non-paired (donor) splice site.

15 In one ~~embodiment~~example of ~~all aspects and embodiments~~the present disclosure the non-paired (donor) splice site is an artificial non-paired (donor) splice site and has been generated during codon optimization.

Detailed Description of Embodiments of the Invention

20 For the production of a therapeutic or diagnostic antibody high expression yields are the aim. One option to achieve good expression rates is to optimize the codon usage of the encoding nucleic acids at first and adjust it to the codon usage of the cell intended to express the exogenous nucleic acid. This codon adaptation or optimization can be done based on different established protocols.

25 But during such a codon adaptation and optimization process, e.g. non-paired splices sites can be generated de novo. That is, during codon optimization a new donor splice site sequence is generated in the codon usage optimized nucleic acid. This is independent on the organization of the codon usage optimize nucleic acid, i.e. it is possible for both cDNA or genomically organized nucleic acids. It is in fact an unintended side-result of the codon usage optimization process. As such a new donor splice site is an additional artificial donor splice site it does not have an associated target acceptor splice site. Thus, such non-paired donor splice sites can give rise to
30 splicing events with a random, i.e. non-defined, acceptor splice site present somewhere in the transcribed mRNA. Thereby the expression yield is reduced.