antibodies is reported by Zhang, Z., et al., in Mass. Spectrom. Rev. 28 (2009) 147-176. Yu, X.C., et al. (Anal. Chem. 81 (2009) 9282-9290) report the identification of codon-specific serine to asparagine mistranslations in recombinant monoclonal antibodies by high-resolution mass spectrometry. The determination of protein oxidation by mass spectrometry and method transfer to quality control is reported by Houde, D., et al., J. Chromatogr. 1123 (2006) 189-198.

Summary ~~of the Invention~~

Herein is provided a method for the determination of amino acid sequence variants, i.e. mutations in the amino acid sequence, of a polypeptide. Indirectly therewith also nucleic acid sequence variants can be determined.

In more detail herein is reported a method for determining amino acid sequence mutations in a (produced) polypeptide, comprising the following steps: a) providing a sample of a (produced) polypeptide, b) incubating the polypeptide in the sample with a protease, c) performing a two dimensional analysis using reversed phase chromatography coupled with high resolution mass spectroscopy (FT-ICR / FT-orbitrap) and MS/MS analysis of the amino acid sequence fragments of the peptides, d) data evaluation by comparing the LC-MS data sets obtained for the samples side by side with the data set of a reference sample, by searching for differences in the signal intensities at given retention times and by evaluation of differential signals with respect to amino acid sequence mutations. The reference sample for data evaluation (step d) can be either a well characterized standard or one of the samples to be analyzed.

~~A first aspect as~~ There is reported herein ~~is~~ a method for determining an amino acid sequence variant, especially an amino acid mutation in the amino acid sequence, of a polypeptide, characterized in comprising the following steps:

a) providing at least two samples of the polypeptide,
b) incubating the samples each with a protease,
c) analyzing the incubated samples by a two dimensional data analysis comprising the combination of reversed phase chromatography separation with mass spectrometry analysis and/or MS/MS analysis,
d) defining the data-set obtained with one sample in step c) as reference sample and comparing the data-sets obtained with the other samples in step c) with the data-set of the reference sample, whereby every

difference determined is an amino acid sequence variation (mutation) of the polypeptide and thereby determining an amino acid sequence variant of the polypeptide.

There is ~~Another aspect as~~ reported herein ~~is~~ a method for determining an amino acid sequence mutation in the amino acid sequence of a polypeptide, characterized in comprising the following steps:

a) providing at least two samples of the polypeptide,
b) incubating the samples each with a protease,
c) analyzing the incubated samples by a two dimensional data analysis comprising the combination of reversed phase chromatography separation with mass spectrometry analysis and/or MS/MS analysis,
d) defining the data-set obtained with one sample in step c) as reference sample and comparing the data-sets obtained with the other samples in step c) with the data-set of the reference sample, whereby every difference determined is an amino acid sequence mutation of the polypeptide and thereby determining an amino acid sequence mutation of the polypeptide,

or the following steps:

a) providing at least one sample of the polypeptide,
b) incubating the sample with a protease,
c) analyzing the incubated sample by a two dimensional data analysis comprising the combination of reversed phase chromatography separation with mass spectrometry analysis and/or MS/MS analysis,
d) comparing the data-set obtained with the sample in step c) with the data-set of a reference sample determined with a method as reported in steps b) and c), whereby every difference determined is an amino acid sequence mutation of the polypeptide and thereby determining an amino acid sequence mutation of the polypeptide.

The provided samples may originate from different clones obtained by transient transfection of a nucleic acid encoding the polypeptide, or from stable transfected cell lines, or from cell lines of different age, or different fermentation scales, or fermentation runs under different conditions. In one embodiment of from 2 to 10,000 samples are provided, in another embodiment of from 2 to 1,000 samples are provided, and in a further embodiment of from 2 to 348 samples are provided.

embodiment the polypeptide in the samples is reduced and the free sulfhydryl residues are carboxymethylated prior to the incubation with the protease.

The samples are all treated separately, i.e. individually and not as mixture, in the steps according to the methods as reported herein. In one embodiment the samples are treated individually in the reducing step, the incubating step, and the analyzing step. In one embodiment the method is a method for high throughput determination.

In one embodiment the samples are incubated for 16 hours to 18 hours with the protease and immediately thereafter formic acid or trifluoro acetic acid are added. In a further embodiment the samples are incubated for 4 hours with the protease and immediately thereafter formic acid or trifluoro acetic acid are added.

One aspect as reported herein is a method for determining a polypeptide with a mutant amino acid sequence, characterized in comprising the following steps:

a) providing at least two samples of the polypeptide,

b) incubating the samples each with the same protease,

c) analyzing the incubated samples by a two dimensional data analysis using a combination of a reversed phase liquid chromatography separation and a mass spectrometry analysis and/or MS/MS analysis,

d) defining the data set obtained with one sample in step c) as reference sample and comparing the data sets obtained with the other samples in step c) with the data set of the reference sample, whereby every amino acid sequence difference determined is an amino acid sequence mutation of the polypeptide and thereby determining a polypeptide with a mutant amino acid sequence.

Another aspect as reported herein is a method for producing a polypeptide comprising the following step: selecting a cell producing a polypeptide, whereby the polypeptide comprises the smallest number of amino acid sequence mutations of all processed samples or with respect to a reference sample or predetermined amino acid sequence determined with a method as defined herein.

There is ~~Another aspect as~~ reported herein ~~is~~ a method for producing a polypeptide comprising the step of selecting a cell producing a polypeptide, whereby the polypeptide comprises the smallest number and the smallest ratio, respectively, of amino acid sequence variations, i.e. mutations in the amino acid sequence, of all processed samples. There is ~~Another aspect as~~ reported herein ~~is~~ a method for producing a polypeptide comprising the step of selecting a cell producing a polypeptide, whereby the polypeptide comprises no detectable amino acid sequence mutation (variations). In one embodiment the smallest number and the smallest ratio, respectively, of amino acid sequence variations (mutations in the amino acid sequence) is determined with respect to a reference sample or a predetermined amino acid sequence.

In one embodiment the method comprises the steps of:

a) providing at least two cells comprising a nucleic acid encoding the polypeptide,
b) single depositing and cultivating the cells,
c) performing with the single deposited cells a method as reported herein,
d) selecting a cell producing a polypeptide, whereby the polypeptide comprises the smallest number, and/or smallest ratio of amino acid sequence variations (amino acid sequence mutations),
e) cultivating the selected cell,
f) producing a polypeptide by recovering the polypeptide from the cell or the cultivation medium.

In one embodiment the provided cell is selected from a transiently transfected cell, or a stable transfected cell, or an immune cell obtained from an animal after immunization.

Another aspect as reported herein is a method for producing an immunoglobulin comprising the steps of:

a) providing at least two cells comprising a nucleic acid encoding the immunoglobulin,
b) single depositing and cultivating the cells,
c) performing a method according to any one of claims 1 to 14,
d) selecting a cell producing an immunoglobulin, whereby the immunoglobulin comprises the smallest number of amino acid sequence mutations with respect to a reference sample,
e) cultivating the cell,
f) producing the polypeptide by recovering the polypeptide from the cell or the cultivation medium.

Detailed Description of the Invention

Therapeutic proteins, e.g. produced by recombinant methods, may be a mixture of molecules with slightly different amino acid sequences, whereby the difference is not only at the C- or N-terminus of the amino acid sequence but also within the amino acid sequence (e.g. amino acid mutations, exchanges, oxidation or thioether formation). Amino acid sequence mutations (variations) can be detected with a method as reported herein by peptide mapping methodology. These amino acid sequence mutations (variations) may emerge, for example, i) during nucleic acid replication by errors of the nucleic acid polymerase, or ii) during nucleic acid transcription by errors of the RNA polymerase or the splicing apparatus, or iii) during protein translation by acquisition of the wrong t-RNA or by t-RNAs loaded with the wrong amino acid, or iv) as unintended post-translational modification or incomplete removal of the signal peptide. Therefore, an amino acid sequence mutation (variation) can result from the integration/amplification process of the DNA encoding the respective polypeptide, or from the replication of the coding nucleic acid, or the transcription including post transcriptional modification (e.g. during mRNA splicing or RNA editing) or during translation. Amino acid variations resulting from a modification of individual amino acid residues after release of the molecule from the ribosome are defined as not necessarily unintended post translational modification and are not a mutation (variant) as termed herein.

From literature it is known that eukaryotic (as well as prokaryotic) enzymes have an average error rate of:

replication:	- 1 per 10^8 to 10^{12} replicated nucleic acid bases by normal polymerases,
	- 1 per 10^1 to 10^3 replicated nucleic acid bases by error-prone polymerases;
transcription:	1 per 10^4 to 10^5 transcribed nucleic acid base pairs;
translation:	1 per 10^3 to 10^4 incorporated amino acid residues.